Московский государственный университет им. М. В. Ломоносова

Е. Е. Тыртышников

КУРС
ЛИНЕЙНОЙ АЛГЕБРЫ

Москва 2004–2005

Оглавление

Предисловие

Данная книга — это расширенный конспект лекций, прочитанных мною для студентов первого курса факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова в течение двух семестров 2004/2005 года.

В аудитории обсуждались прежде всего вопросы *основной части* каждой лекции. Кроме этого, почти каждая лекция содержит *дополнительную часть*, предназначенную для домашнего чтения и адресованную наиболее заинтересованным и успевающим студентам. В некоторых случаях дополнительная часть также разбита на две секции — одна из них набрана мелким шрифтом и должна считаться "более дополнительной".

Я стремился сохранить традиции в преподавании алгебры на факультете ВМиК МГУ — в частности, курс линейной алгебры естественным образом включает в себя и необходимые основы аналитической геометрии. В духе сложившихся традиций, безусловно, уже созданы хорошие учебники. Тем не менее, данная книга может оказаться полезной по ряду причин.

Во-первых, это именно лекции. Я избегал любых "длинных" ссылок и не думал о том, чтобы минимизировать неизбежные в лекциях напоминания и повторения. Вместе с этим изложение остается сжатым и лаконичным в той степени, которая свойственна ограниченным во времени лекциям.

Во-вторых, при обсуждении совершенно классических вопросов мне казалось важным показать, что изучаемая нами наука является живой, очень успешно развивающейся и прочно связанной с многими другими разделами математики. Как только появляется возможность сделать реальные шаги в направлении каких-либо особо впечатляющих достижений и современных проблем, я пытаюсь это делать. В каждом таком случае я считаю важным избегать чисто декларативного описания — если уж что-то обсуждается, то всегда с ясными формулировками и полными (почти всегда) доказательствами. Дополнительные части лекций служат именно этой цели.

В-третьих, в книге идет одновременное развитие нескольких тем — подобно тому, как это делается в полифоническом музыкальном произведении. Главная тема, конечно, — это все, что связано с концепцией линейной зависимости векторов. В качестве побочной (хотя и не менее значительной) темы в самом начале возникает понятие алгебраической операции и группы. Эта тема впоследствии приводит к важным понятиям кольца и поля, а затем и к своеобразной точке "контрапункта" (в той же музыкальной аналогии), когда свойства линейного пространства применяются к изучению расширений полей.

В дополнительных частях в сжатом и в то же время замкнутом виде можно найти весьма нетривиальные результаты, выходящие за рамки собственно линейной алгеб-

ры (например, вопросы о построении правильных n-угольников и разрешимости алгебраических уравнений). Общеизвестно, однако, что значение и сила линейной алгебры обусловлены прежде всего ее многочисленными приложениями.

Я согласен с тем, что линейной алгебре не следует учить "слишком абстрактно". Тем более, что есть возможность познакомиться с основными понятиями, работая с простыми для понимания объектами — матрицами, а не с абстрактными элементами линейных пространств. В то же время, мне кажется, что определенная доза абстракных понятий уместна и даже полезна на самой ранней стадии обучения. В самом деле, вряд ли можно считать чрезмерными усилия на освоение всего лишь определения группы и простейших ее свойств. Однако, если это сделать на раннем этапе обучения, то в дальнейшем находится много поводов для возвращения к этому понятию в связи с примерами групп, которые естественным образом возникают в разных местах курса.

Мне кажется, что упрощение формы изложения все же может сочетаться с более наполненным содержанием. По крайней мере, я стремился к этому. Линейная алгебра и ее приложения настолько фундаментальны и важны, что нет никаких оснований для реального сокращения объема обязательных базовых знаний в данной области.

В нашем курсе предмет линейной алгебры понимается в расширенном смысле, довольно часто мы оказываемся на территории смежных дисциплин — математического анализа, вычислительных методов и, конечно, общей алгебры. Границы являются условностью, как и в жизни. Особенно часто они пересекаются при разработке современных информационных и вычислительных технологий.

Например, одна из главных обязательных тем первого семестра — теория и методы исследования и решения систем линейных алгебраических уравнений. Материал вполне элементарный и, возможно, оставляющий впечатление абсолютной завершенности. Однако, практическая необходимость решения систем с миллионами уравнений и неизвестных и появление вычислительной техники с параллельным выполнением операций дали импульс к изучению новых, параллельных свойств алгоритмов. В данном случае успехи прямо связаны с ростом мощи компьютеров. В то же время - и мне особенно приятно сказать об этом - выход на радикально новый уровень возможностей был сделан благодаря новому математическому знанию, а не росту производительности компьютеров. Более того, для данной вполне классической задачи линейной алгебры потребовалось развитие фундаментальных вопросов из области математического анализа и теории приближений.

Отдельные места в книге содержат материал, который вообще нельзя найти в каких-либо учебниках и даже монографиях. В частности, это относится к теореме об обобщениях методов сопряженных градиентов. В еще большей степени — ко всему материалу заключительной лекции, посвященной многомерным массивам, тензорным рангам и полилинейным обобщениям сингулярного разложения матрицы.

В те времена, когда факультет ВМиК только появился, математики-вычислители часто сетовали на то, что в обязательных курсах мех-мата ничего не говорилось о возникших перед ними проблемах. В настоящее время можно уже говорить о том, что математикам-вычислителям часто не хватает знаний из традиционных именно для мех-мата разделов математики. Можно привести примеры рекордно эффективных вычислительных технологий, возникших на основе идей и аппарата казалось бы далеких от приложений областей — например, алгебраической топологии. В данной книге эти заявления все же останутся, к сожалению автора и читателей, не более чем декларациями.

Но ведь это лишь начало пути!

В любом деле очень важен начальный импульс. Для данной книги его генератором был В. А. Ильин, пригласивший меня прочитать лекции на ВМиК.

В Институте вычислительной математики Российской академии наук, где я имею честь работать, это предложение было горячо поддержано В. В. Воеводиным, В. П. Дымниковым и Г. И. Марчуком, попросившим меня в то же самое время помочь в организации на ВМиК новой кафедры — кафедры вычислительных технологий и моделирования, которой он стал заведовать.

Мне оставалось только согласиться и попытаться сделать то, о чем я, скорее всего, уже думал — попробовать рассказать студентам о линейной алгебре то, что я сам бы хотел услышать, когда был студентом. По крайней мере, самому мне это все пока нравится. Поэтому всем названным лицам выражаю искреннюю благодарность. Хочу поблагодарить также Н. Л. Замарашкина, Х. Д. Икрамова, Г. Д. Ким, В. С. Панферова и всех тех, кто уже сделал или еще сделает замечания по тексту лекций.

Лекция 1

ОСНОВНАЯ ЧАСТЬ

1.1 Линейные отображения и матрицы

В математике и других науках постоянно изучается зависимость одних величин от других. Обычно зависимость описывается различного типа функциями (отображениями, операторами). Простейший случай — линейные отображения. Строгие определения мы дадим позже. А пока предположим, что переменные $y_1, \ldots, y_m$ выражаются через $x_1, \ldots, x_n$ следующим образом:

$$\left\{\begin{array}{lcl} y_1 &=& a_{11}x_1 + \ldots + a_{1n}x_n, \\ & & \ldots \\ y_m &=& a_{m1}x_1 + \ldots + a_{mn}x_n, \end{array}\right. \qquad (*)$$

где коэффициенты считаются заданными постоянными величинами. Соберем все постоянные коэффициенты в прямоугольную таблицу и обозначим ее буквой A; составим также таблицы-столбцы из величин $x_1, \ldots, x_n$ и $y_1, \ldots, y_m$:

$$A = \left[\begin{array}{ccc} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \end{array}\right], \qquad x = \left[\begin{array}{c} x_1 \\ \ldots \\ x_n \end{array}\right], \qquad y = \left[\begin{array}{c} y_1 \\ \ldots \\ y_m \end{array}\right].$$

Такие таблицы и называются *матрицами*. Мы имеем целых три матрицы: размеров $m \times n$, $n \times 1$ и $m \times 1$. Соотношения $(*)$, описывающие зависимость y от x, запишем символически таким образом:

$$y = Ax. \qquad (**)$$

Возникает впечатление, что матрица A *умножается* на матрицу-столбец x, в результате чего появляется матрица-столбец y. Так оно и будет, если мы скажем, что соотношения $(*)$ суть *определение* операции $(**)$ умножения A на x.

Если $m = n$, то матрица называется *квадратной*. Квадратная матрица размеров $n \times n$ называется также *матрицей порядка* n.

1.2 Умножение матриц

Пусть $y_1, \ldots, y_m$ выражаются через $x_1, \ldots, x_n$ и при этом $x_1, \ldots, x_n$ выражаются через $z_1, \ldots, z_k$ следующим образом:

$$\left\{\begin{array}{lcl} y_1 &=& a_{11}x_1 + \ldots + a_{1n}x_n, \\ & & \ldots \\ y_m &=& a_{m1}x_1 + \ldots + a_{mn}x_n, \end{array}\right. \qquad \left\{\begin{array}{lcl} x_1 &=& b_{11}z_1 + \ldots + b_{1k}z_k, \\ & & \ldots \\ x_n &=& b_{n1}z_1 + \ldots + b_{nk}z_k. \end{array}\right.$$

Ясно, что $y_1, \ldots, y_m$ выражаются через $z_1, \ldots, z_k$ аналогичным образом. Матрицу из постоянных коэффициентов этой зависимости обозначим через C. Тогда

$$y = Ax, \quad x = Bz \ \text{ и } \ y = Cz.$$

Чтобы получить коэффициенты матрицы C, нужно подставить выражения для $x_1, \ldots, x_n$ через $z_1, \ldots, z_k$ в формулы, выражающие $y_1, \ldots, y_m$ через $x_1, \ldots, x_n$, и собрать коэффициенты при величинах $z_1, \ldots, z_k$. Получится вот что:

$$C = [c_{ij}], \ \text{ где } \ c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}. \qquad (*)$$

Определение. Матрица C вида $(*)$ называется *произведением* матриц A и B и обозначается $C = AB$.

Следствие. $y = A(Bz) = (AB)z$.

Часто говорят, что матрицы умножаются по правилу "строка на столбец". Число столбцов в первом сомножителе обязано, конечно, совпадать с числом строк во втором. Если мы пишем $C = AB$, то автоматически имеем в виду, что матрицы A и B не совсем уж произвольные.

1.3 Приятное свойство: ассоциативность

Теорема. $(AB)C = A(BC)$.

Доказательство. Пусть A — $m \times n$, B — $n \times k$, C — $k \times l$. Тогда

$$\begin{aligned} \{(AB)C\}_{ij} &= \sum_{p=1}^{k} \{AB\}_{ip} c_{pj} = \sum_{p=1}^{k} \left(\sum_{q=1}^{n} a_{iq} b_{qp} \right) c_{pj} \\ &= \sum_{q=1}^{n} a_{iq} \left(\sum_{p=1}^{k} b_{qp} c_{pj} \right) = \{A(BC)\}_{ij}. \end{aligned}$$

1.4 Не очень приятное свойство: некоммутативность

В общем случае $AB \neq BA$ — даже для квадратных матриц. Например,

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

1.5 Сложение матриц и умножение на число

Матрица $C = [c_{ij}]$ называется *суммой* матриц $A = [a_{ij}]$ и $B = [b_{ij}]$, если

$$c_{ij} = a_{ij} + b_{ij} \quad \text{для всех} \ \ i, j.$$

Матрицы A, B и $C = A + B$ – одинаковых размеров. Для операции сложения матриц выполняются сразу два приятных свойства:

$$A + (B + C) = (A + B) + C \quad \text{(ассоциативность)},$$

$$A + B = B + A \quad \text{(коммутативность)}.$$

Полезно ввести также операцию умножения матрицы на число. Если α — число, то матрица $C = \alpha A$ определяется как матрица тех же размеров с элементами $c_{ij} = \alpha a_{ij}$.

1.6 Умножение блочных матриц

Предположим, что матрицы A и B составлены из блоков A_{ij} и B_{ij}:

$$A = \begin{bmatrix} A_{11} & \dots & A_{1q} \\ \dots & \dots & \dots \\ A_{p1} & \dots & A_{pq} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{11} & \dots & B_{1r} \\ \dots & \dots & \dots \\ B_{q1} & \dots & B_{qr} \end{bmatrix},$$

где $A_{ij} \ - \ m_i \times n_j$, $B_{ij} \ - \ n_i \times k_j$. Тогда произведение $C = AB$ существует и его можно вычислять, используя операции умножения и сложения матриц-блоков:

$$C = \begin{bmatrix} C_{11} & \dots & C_{1r} \\ \dots & \dots & \dots \\ C_{p1} & \dots & C_{pr} \end{bmatrix}, \quad \text{где} \quad C_{ij} = \sum_{l=1}^{q} A_{il} B_{lj} \quad - \quad m_i \times k_j.$$

Докажите!

Можно сказать, что блочные матрицы умножаются по правилу “блочная строка на блочный столбец”. Мы очень скоро увидим, какую пользу может дать блочное умножение.

1.7 Вычислительный аспект умножения матриц

Пусть заданы $n \times n$-матрицы A и B и требуется вычислить их произведение $C = AB$. Вот классический алгоритм (программа на неком подобии алгоритмического языка Фортран):

```
DO i = 1, n
    DO j = 1, n
        DO k = 1, n
            c_ij = c_ij + a_ik b_kj
        END DO
    END DO
END DO.
```

Конечно, предварительно следует занулить элементы c_{ij}.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

1.8 Хороша ли программа?

Ответить на этот вопрос не очень просто. Прежде всего, нужен какой-то критерий — пусть это будет время исполнения программы. Но время зависит не только от типа компьютера. В строгом смысле, оно привязано к отдельно взятому компьютеру и зависит от его состояния на данный момент, от операционной системы и, конечно, от особенностей транслятора.

Чтобы что-то здесь понять, нужно отбросить очень много деталей и оставить нечто главное. Если все операции выполняются последовательно, то время работы можно считать пропорциональным числу операций. Мы пойдем дальше и будем подсчитывать лишь арифметические операции. Общее их число будем называть *арифметической сложностью* алгоритма.

Легко найти, что арифметическая сложность классического алгоритма умножения матриц равна $2n^3$ (n^3 умножений и n^3 сложений). Но хорошо ли это? Уверены ли мы в том, что это наилучший алгоритм?

1.9 Когда-то казалось, что да

Само понятие "наилучший" предполагает наличие некого множества возможных алгоритмов. Будем полагать, что алгоритм — это последовательность элементарных операций из конечного фиксированного набора элементарных операций. Для определенности, пусть это будут четыре арифметических действия.

Итак, математическая задача поставлена. Еще в недавнем прошлом многим казалось, что классический алгоритм самый лучший. Теперь уже ясно, что это не так.

1.10 Метод Винограда

Попробуйте-ка перемножить матрицы как-либо иначе — не по классическому алгоритму. Вероятно, впервые это сделал Виноград (в начале 60-х). Он догадался использовать следующее тождество:

$$\sum_{k=1}^{2m} a_{ik}b_{kj} = \sum_{k=1}^{m}(a_{i\,2k-1} + b_{2k\,j})(b_{2k-1\,j} + a_{i\,2k}) - \sum_{k=1}^{m} a_{i\,2k-1}a_{i\,2k} - \sum_{k=1}^{m} b_{2k\,j}b_{2k-1\,j}.$$

Пусть $n = 2m$. Ясно, что вторую и третью суммы для всех $1 \le i, j \le n$ можно найти, затратив $2nm = n^2$ умножений и $2nm = n^2$ сложений. Для первой суммы потребуется $n^2m = \frac{1}{2}n^3$ умножений и $3n^2m = \frac{3}{2}n^3$ сложений.

В итоге — по-прежнему, $2n^3$ операций (без учета порядка $n^2 \ll n^3$ операций), но теперь $\frac{1}{2}n^3$ умножений и $\frac{3}{2}n^3$ сложений! Поскольку умножение — операция более сложная, чем сложение, метод Винограда может представлять практический интерес.

1.11 Метод Штрассена

В 1965 году Штрассен нашел способ умножения 2×2-матриц с помощью всего лишь 7-ми умножений (в классическом методе 8 умножений). То, что придумал Штрассен, получается посредством вычисления тензорного ранга "многомерных матриц". Об этом мы поговорим в заключительной лекции курса. А пока давайте посмотрим на изобретение Штрассена "без комментариев": [1]

[1] См. задачу 5.4.21 из "Задачника по линейной алгебре" Х.Д.Икрамова.

$$
\begin{aligned}
\alpha_1 &= (a_{11}+a_{22})(b_{11}+b_{22}),\\
\alpha_2 &= (a_{21}+a_{22})b_{11},\\
\alpha_3 &= a_{11}(b_{12}-b_{22}),\\
\alpha_4 &= a_{22}(b_{21}-b_{11}),\\
\alpha_5 &= (a_{11}+a_{12})b_{22},\\
\alpha_6 &= (a_{21}-a_{11})(b_{11}+b_{12}),\\
\alpha_7 &= (a_{12}-a_{22})(b_{21}+b_{22}),
\end{aligned}
\qquad
\begin{aligned}
c_{11} &= \alpha_1+\alpha_4-\alpha_5+\alpha_7,\\
c_{12} &= \alpha_3+\alpha_5,\\
c_{21} &= \alpha_2+\alpha_4,\\
c_{22} &= \alpha_1+\alpha_3-\alpha_2+\alpha_6.
\end{aligned}
$$

Только очень ленивый человек не сможет проверить, что две матрицы порядка 2 умножаются правильно.

1.12 Рекурсия для $n \times n$-матриц

От метода умножения 2×2-матриц с 7-ю умножениями довольно легко перейти к методу умножения $n \times n$-матриц, требующему не более $7\,n^{\log_2 7}$ операций. Поскольку

$$\frac{7n^{\log_2 7}}{n^3} \to 0 \quad \text{при} \quad n \to \infty,$$

метод Штрассена асимптотически лучше классического метода.

Предположим, что $n = 2^L$ и будем рассматривать A и B как блочные 2×2-матрицы:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22}, \end{bmatrix}, \qquad A_{ij}, B_{ij} \quad - \quad \frac{n}{2} \times \frac{n}{2}.$$

Замечательно, что в штрассеновском методе умножения 2×2-матриц коммутативность не используется. Поэтому метод годится и для умножения блочных 2×2-матриц!

Итак, *задача размера* n сводится к 7-ми аналогичным *задачам размера* $\frac{n}{2}$. Для формирования этих 7-ми задач и для получения окончательного результата после решения этих 7-ми задач требуется 18 раз сложить блоки порядка $\frac{n}{2}$.

"Раскрутив"указанную рекурсию до конца, получим

$$7^{\log_2 n} = n^{\log_2 7}$$

умножений на последнем этапе. Общее число сложений на всех этапах составит

$$18 \sum_{k=1}^{L} 7^{k-1} \left(\frac{n}{2^k}\right)^2 = \frac{18}{4}\, n^2\, \frac{\left(\frac{7}{4}\right)^L - 1}{\frac{7}{4} - 1} \le 6 \cdot 7^L = 6\, n^{\log_2 7}$$

(нужно учесть, что $4^L = n^2$ и $7^L = n^{\log_2 7}$).

При практическом применении рекурсию не обязательно раскручивать до конца. Это вредно: $7\,n^{\log_2 7} > 2n^3$ даже при $n = 512$. Но при $n = 1\,024$ неравенство меняется в пользу Штрассена.

К настоящему времени придуманы и более быстрые (асимптотически) методы, чем метод Штрассена. Уже существуют методы с числом операций $\mathcal{O}(n^\alpha)$, где $\alpha < 2.42$. Никто не знает, каков минимальный показатель в таких оценках. Ясно лишь, что $\alpha \ge 2$.

1.13 Параллельная форма алгоритма

Арифметическая сложность алгоритма — вещь, конечно, важная в любом случае. Но с развитием компьютеров время становится "все менее пропорциональным" общему числу операций. Дело в том, что многие операции выполняются параллельно (одновременно).

Чтобы понять хоть что-то, нужно и теперь отбросить очень много деталей. Рассмотрим модель *бесконечного параллелизма*: имеется бесконечно много процессоров с неограниченной памятью, каждый может в любую единицу времени выполнить одну арифметическую операцию и мгновенно обменивается информацией с любым другим процессором.

Чтобы реализовать алгоритм на таком идеализированном компьютере, достаточно записать его в виде последовательности *ярусов* — наборов информационно несвязанных операций (их можно выполнять параллельно). Такое представление алгоритма называется его *параллельной формой*, число ярусов называется *высотой*, а максимальное число операций в одном ярусе — *шириной параллельной формы*.

Для любого алгоритма существует, очевидно, параллельная форма с минимальным числом ярусов. Это минимальное число ярусов называется *высотой алгоритма*. В модели бесконечного параллелизма минимальное время реализации алгоритма пропорционально его высоте.

1.14 Схема сдваивания и параллельное умножение матриц

Высота классического алгоритма умножения матриц имеет вид $\mathcal{O}(n)$. Докажите!

Легко получить и алгоритм высоты $\mathcal{O}(\log_2 n)$. Для этого достаточно построить алгоритм сложения n чисел, имеющий высоту $\mathcal{O}(\log_2 n)$. Такой алгоритм называется *схемой сдваивания*: нужно разбить числа на пары, найти суммы для каждой пары, затем разбить результаты на пары, найти суммы, и так далее.

1.15 Матрицы и рекуррентные вычисления

Рассмотрим последовательность величин $x_{-1}, x_0, x_1, \ldots$, в которой x_{-1}, x_0 заданы, а остальные величины вычисляются рекуррентно:

$$x_k = a_k x_{k-1} + b_k x_{k-2}, \quad k = 1, 2, \ldots, n. \qquad (*)$$

Коэффициенты a_k, b_k считаются заданными. Чтобы вычислить x_n, в силу $(*)$ требуется выполнить $O(n)$ арифметических операций. Число параллельных шагов – также $O(n)$. Возникает впечатление, что алгоритм с меньшей высотой параллельной формы получить нельзя.

Но это впечатление обманчиво. Запишем соотношения $(*)$ в матричной форме:

$$\begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix} = \begin{bmatrix} a_k & b_k \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ x_{k-2} \end{bmatrix},$$

или,

$$z_k = A_k z_{k-1},$$

$$z_k = \begin{bmatrix} x_k \\ x_{k-1} \end{bmatrix}, \quad z_{k-1} = \begin{bmatrix} x_{k-1} \\ x_{k-2} \end{bmatrix}, \quad A_k = \begin{bmatrix} a_k & b_k \\ 1 & 0 \end{bmatrix}.$$

Отсюда

$$z_n = Az_0, \quad A = A_n(A_{n-1}(\cdots(A_3(A_2A_1))\cdots).$$

Чтобы определить произведение матриц $A_nA_{n-1}\cdots A_1$, нужно свести его к вычислению произведений двух матриц. Это делается расстановкой скобок. Используя ассоциативность операции умножения матриц, можно доказать, что результат не будет зависить от порядка расстановки скобок; поэтому можно писать без скобок:

$$A = A_nA_{n-1}\cdots A_1.$$

Чтобы найти z_n (а значит, и x_n), сначала вычислим матрицу A. Для этого можно использовать ту же схему сдваивания: находим произведения

$$A_nA_{n-1},\ A_{n-2}A_{n-3},\ \ldots,\ A_2A_1,$$

затем попарные произведения полученных результатов, и так далее. Потребуется всего лишь $O(\log_2 n)$ параллельных шагов!

1.16 Модели и реальность

В модели бесконечного параллелизма мы отбрасываем, увы, слишком много деталей, которые следует учитывать. Я думаю, можно почувствовать проблемы параллельных вычислений, размышляя над следующей задачей-шуткой: “Один землекоп выкапывает яму глубиной 1 метр за 1 час. За какое время эту яму выкопают 100 землекопов? ”

Чтобы выполнять какую-то работу параллельно, необходимо такую работу иметь. В существующих алгоритмах работы для параллельного (одновременного) исполнения может быть недостаточно. Оперируя над общими данными, процессоры могут мешать друг другу. Как учесть все это в более адекватных и все же поддающихся анализу моделях – это трудный вопрос (которым мы, конечно, заниматься не собираемся).

Лекция 2

ОСНОВНАЯ ЧАСТЬ

2.1 Множества и элементы

Понятие *множества* вводится для обозначения совокупности *элементов*, объединенных каким-то общим признаком. Считается, что оно относится к первичным понятиям, которым не дается формального определения.

Запись $a \in M$ означает, что *элемент* a принадлежит *множеству* M. Запись $X \subset Y$ означает, что каждый элемент множества X принадлежит множеству Y. При этом X называется *подмножеством* для Y. Особо выделяется множество, в котором нет ни одного элемента. Оно называется *пустым* и обозначается символом $\emptyset$. По определению, $\emptyset \subset M \ \ \forall\, M$.

При описании множеств иногда возникают логические противоречия. Например, рассмотрим множество M, состоящее из одного числа, которое определяется как "наименьшее целое число, которое нельзя определить при помощи фразы, имеющей менее ста русских слов". Такое число должно существовать, поскольку число допустимых фраз, имеющих менее ста слов, конечно. В то же время оно определяется приведенной выше фразой, а в ней менее ста слов! [1]

В нашем курсе, к счастью, противоречий такого рода при задании множеств возникать не будет. Но даже при полной ясности с определением множества (например, множество корней уравнения) не всегда легко установить, сколько в нем элементов и будет ли оно вообще непустым.

Довольно часто множества будут задаваться перечислением своих элементов. Например, $M = \{1, 2, 3\}$ — множество, состоящее из трех чисел 1, 2, 3.

Кроме того, новые множества можно конструировать с помощью уже имеющихся множеств X и Y следующим образом:

- $A = X \cup Y \equiv \{a:\ a \in X \text{ или } \ a \in Y\}$ (объединение множеств);
- $B = X \cap Y \equiv \{b:\ b \in X \text{ и } \ b \in Y\}$ (пересечение множеств);
- $C = X \backslash Y \equiv \{c:\ c \in X,\ c \notin Y\}$ (разность множеств);
- $D = X \times Y \equiv \{d = (a, b):\ a \in X,\ b \in Y\}$ (декартово произведение множеств).

[1] Пример из учебника В. В. Воеводина "Линейная алгебра", Наука, 1980.

2.2 Отображения, функции, операторы

Все три слова из заголовка означают одно и то же — речь идет о правиле, по которому каждому элементу x множества X ставится в соответствие однозначно определенный элемент $y = f(x)$ множества Y. Задание правила равносильно выбору подмножества

$$\Gamma = \{(x, f(x)) : \ x \in X\} \ \subset \ X \times Y,$$

называемого *графиком* отображения (функции, оператора) f.

Элемент $y = f(x)$ называется *образом* элемента x, а x — *прообразом* элемента y при отображении f. Чтобы подчеркнуть, что f действует из X в Y, пишут так: $f : \ X \to Y$.

Множество $f(X) \equiv \{y : \ y = f(x) \text{ для некоторого } x \in X\}$ называется *образом* (множеством значений) отображения f.

Если $M \subset Y$, то множество $f^{-1}(M) \equiv \{x : f(x) \in M\}$ называется *полным прообразом* множества M. Если $M = \{y\}$, то пишут таким образом: $f^{-1}(y) = f^{-1}(M)$.

Отображение $f : \ X \to Y$ называется *обратимым*, если существует отображение $g : \ Y \to X$ такое, что $f(g(y)) = y \ \forall \ y \in Y$ и $g(f(x)) = x \ \forall \ x \in X$. При этом g называют *обратным отображением* для f и пишут $g = f^{-1}$.

Отображение f называется *взаимно-однозначным*, если для любого $y \in Y$ полный прообраз $f^{-1}(y)$ состоит ровно из одного элемента. Легко показать, что обратимость равносильна взаимной однозначности.

2.3 Алгебраические операции

Отображение $f : \ X \times X \to X$ называется *алгебраической операцией* на X. Пусть для обозначения такой операции используется символ $*$. Тогда запись $c = a * b$ означает, что $(a, b) \in X \times X$ и $c = f((a, b))$.

Если задано отображение $f : M \to X$ на непустом подмножестве $M \subset X \times X$, то f называется *частичной алгебраической операцией* на X. Таковой, в частности, является операция умножения матриц на множестве всех матриц.

Символ $*$ часто опускается, при этом пишут $ab = a * b$, называют операцию умножением, а элемент ab (если он существует) — произведением элементов a и b.

2.4 Ассоциативность и скобки

Частичная алгебраическая операция на X называется *ассоциативной*, если для любых $a, b, c \in X$ из существования произведений ab и bc вытекает существование произведений $a(bc)$, $(ab)c$ и равенство

$$a(bc) = (ab)c.$$

В этом случае естественно убрать скобки и писать $abc \equiv a(bc) = (ab)c$.

Теорема. *Пусть на X задана ассоциативная частичная алгебраическая операция и $x_1, \ldots, x_n$ — произвольные элементы из X, для которых существуют произведения x_1x_2, x_2x_3, …, $x_{n-1}x_n$. Тогда существует расстановка скобок, определяющая элемент*

$$x = x_1x_2 \ldots x_n,$$

при этом любая расстановка скобок дает один и тот же элемент x.

Доказательство. Проведем индукцию по n. Докажем сначала существование некоторой расстановки скобок, определяющей x. Согласно индуктивному предположению,

существует произведение $(x_1 \dots x_{n-2})x_{n-1}$. По условию теоремы существует также произведение $x_{n-1}x_n$. Таким образом, можно применить определение ассоциативности по отношению к элементам $a = x_1 \dots x_{n-2}$, $b = x_{n-1}$, $c = x_n$.

Пусть элементы a и b получаются при разных расстановках скобок. В любом случае имеем

$$a = (x_1 \dots x_k)(x_{k+1} \dots x_n), \qquad b = (x_1 \dots x_m)(x_{m+1} \dots x_n).$$

Пусть $k < m$. Тогда, в силу ассоциативности,

$$\begin{aligned} a &= (x_1 \dots x_k)((x_{k+1} \dots x_m)(x_{m+1} \dots x_n)) = \\ &((x_1 \dots x_k)(x_{k+1} \dots x_m))(x_{m+1} \dots x_n)) = (x_1 \dots x_m))(x_{m+1} \dots x_n) = b. \quad \square \end{aligned}$$

2.5 Ассоциативность при умножении матриц

Пусть нужно вычислить произведение трех прямоугольных матриц размеров $1 \times n$, $n \times 1$ и $1 \times n$:

$$A = BCD = [b_{11} \dots b_{1n}] \begin{bmatrix} c_{11} \\ \dots \\ c_{n1} \end{bmatrix} [d_{11} \dots d_{1n}].$$

В данном случае есть два варианта расстановки скобок:

$$A = B(CD) = [b_{11} \dots b_{1n}] \begin{bmatrix} c_{11}d_{11} & \dots & c_{11}d_{1n} \\ \dots & \dots & \dots \\ c_{n1}d_{11} & \dots & c_{n1}d_{1n} \end{bmatrix}, \tag{1}$$

$$A = (BC)D = [(b_{11}c_{11} + \dots + b_{1n}c_{n1})] [d_{11} \dots d_{1n}]. \tag{2}$$

Варианты (1) и (2) приводят к двум разным алгоритмам вычисления матрицы A. В силу ассоциативности результаты должны быть одинаковыми. Но арифметическая работа будет разная! Применяя правило "строка на столбец", получаем $2n^2$ умножений в случае (1) и всего $2n$ умножений в случае (2).

2.6 Группы

Непустое множество G с ассоциативной алгебраической операцией называется *группой*, если:

(1) существует элемент $e \in G$ такой, что $ae = ea = a$ для любого элемента $a \in G$;

(2) для любого элемента $a \in G$ существует элемент $b \in G$ такой, что $ab = ba = e$.

Элемент e определяется свойством (1) однозначно: если e_1 и e_2 — два таких элемента, то $e_1 = e_1e_2 = e_2$. Он называется *единичным*.

Элемент b из свойства (2) однозначно определяется по a: если b_1 и b_2 — два таких элемента, то $b_1 = b_1(ab_2) = (b_1a)b_2 = b_2$. Элемент b называется *обратным* для a. Обозначение: $b = a^{-1}$.

Для любых фиксированных $a, b \in G$ можно рассмотреть уравнения $ax = b$ (относительно x) и $ya = b$ (относительно y). Оба уравнения однозначно разрешимы: $x = a^{-1}b$ и $y = ba^{-1}$.

Группа называется *абелевой* (коммутативной), если $ab = ba$ для всех $a, b \in G$.

2.7 Примеры абелевых групп

1. $G = \mathbb{R}$ — множество вещественных чисел, операция — сложение чисел. Роль единичного элемента играет число 0.
2. $G = \mathbb{R}\backslash\{0\}$ — множество ненулевых вещественных чисел, операция — умножение чисел. Роль единичного элемента играет число 1.
3. $G = \mathbb{Q}$ — множество рациональных чисел, операция — сложение чисел. Роль единичного элемента играет число 0.
4. $G = \mathbb{Q}\backslash\{0\}$ — множество ненулевых рациональных чисел, операция — умножение чисел. Роль единичного элемента играет число 1.
5. G — множество ненулевых вещественных чисел вида $a+b\sqrt{2}$, где a, b — рациональные числа. Операция — умножение чисел.

Прежде всего, докажем, что произведение чисел из G принадлежит G:

$$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2},$$

из рациональности чисел a, b, c, d вытекает рациональность чисел $ac + 2bd$ и $ad + bc$. Далее, единичным элементом является число $1 = 1 + 0 \cdot \sqrt{2}$. Обратный элемент для $a + b\sqrt{2}$, как легко проверить, имеет вид

$$\left(\frac{a}{a^2 - 2b^2}\right) + \left(\frac{-b}{a^2 - 2b^2}\right)\sqrt{2}.$$

Задача. Пусть G — группа с единицей e. Докажите, что если $a^2 = e$ для любого $a \in G$, то группа G абелева.

2.8 Группа невырожденных диагональных матриц

Матрица $A = [a_{ij}]$ размеров $n \times n$ называется *диагональной*, если $a_{ij} = 0$ при $i \neq j$. Диагональная матрица A называется невырожденной, если $a_{ii} \neq 0$ при всех $1 \leq i \leq n$.

Множество невырожденных диагональных $n \times n$-матриц с вещественными элементами и операцией умножения матриц является абелевой группой. Роль единичного элемента играет матрица

$$I = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}.$$

Она называется *единичной матрицей*.

Задача. Матрица A порядка n коммутирует со всеми диагональными матрицами порядка n: $AB = BA$ для всех диагональных матриц B порядка n. Докажите, что A — диагональная матрица с равными элементами на диагонали.[2]

2.9 Группа невырожденных треугольных матриц

Матрица $A = [a_{ij}]$ размеров $n \times n$ назывется *нижней треугольной*, если $a_{ij} = 0$ при $i < j$, и *верхней треугольной*, если $a_{ij} = 0$ при $i > j$. Треугольная матрица называется невырожденной, если $a_{ii} \neq 0$ при всех $1 \leq i \leq n$.

Множество невырожденных нижних (верхних) треугольных матриц с вещественными элементами и операцией умножения матриц является группой (некоммутативной).

Доказательство состоит из трех этапов:

[2] Такие матрицы называются *скалярными*.

- проверить, что произведение невырожденных нижних (верхних) треугольных матриц является также нижней (верхней) треугольной матрицей;
- проверить, что роль единичного элемента играет единичная матрица I;
- проверить, что для невырожденной нижней (верхней) треугольной матрицы A разрешимы уравнения $AX = I$ и $YA = I$, при этом обе матрицы X и Y являются нижними (верхними) треугольными. После этого равенство $X = Y$ является уже очевидным.

2.10 Подгруппы

Подмножество $H \subset G$ называется *подгруппой* группы G, если оно является группой относительно операции, действующей в G. Для этого необходимо и достаточно, чтобы

- $ab \in H$ для любых элементов $a, b \in H$;
- $a^{-1} \in H$ для любого элемента $a \in H$.

Например, группа невырожденных диагональных матриц является подгруппой группы невырожденных нижних (верхних) треугольных $n \times n$-матриц.

2.11 Степени элемента

Зафиксируем произвольный элемент a в группе G и рассмотрим минимальную содержащую a подгруппу $H(a) \subset G$. Минимальность означает, что $H(a) \subset H$ для любой подгруппы H, содержащей a. Легко видеть, что

$$H(a) = \{a^k : \ k \text{ — целое число}\}.$$

По определению, $a^0 = e$, $a^k = a \dots a$ (a повторяется k раз) при целом положительном k, $a^{-k} = (a^{-1})^k$. Непосредственно из определения вытекает, что

$$a^{k+m} = a^k a^m \quad \text{для любых целых } k,\ m.$$

2.12 Циклические группы

Группа $H(a)$ называется *циклической* группой, порожденной элементом a. Минимальное целое $k > 0$ такое, что $a^k = e$, называется *порядком* элемента a. Если $a^k \neq e$ при всех $k > 0$, то a называется элементом *бесконечного порядка*.

Теорема. *Любая подгруппа циклической группы является циклической.*

Доказательство. Подгруппа $H \subset H(a)$ состоит из каких-то степеней элемента a:

$$H = \{a^{i_1}, a^{i_2}, \dots\}.$$

Пусть m — наименьшее целое положительное число среди $i_1, i_2, \dots$. Тогда ясно, что H содержит все элементы вида a^{mk}. Докажем, что в H не может быть других степеней элемента a. Пусть $a^n \in H$. Разделим n с остатком на m:

$$n = qm + r, \quad q, r \text{ — целые}, \quad 0 \leq r \leq m - 1.$$

Тогда $a^r = a^n a^{-qm} \in H$. В случае $r > 0$ получаем противоречие с минимальностью m. Поэтому $r = 0$. $\square$

Задача. Найти все подгруппы группы целых чисел $\mathbb{Z}$ относительно операции сложения чисел.

2.13 Конечные группы

Группа называется *конечной*, если в ней имеется конечное число элементов. В этом случае число элементов называется *порядком* группы.

Теорема Лагранжа. *В любой конечной группе порядок любой подгруппы является делителем порядка группы.*

Доказательство. Пусть $H = \{a_1, \ldots, a_m\}$ — подгруппа группы G порядка n. Возьмем элемент $a \in G \backslash H$ и рассмотрим множество

$$aH = \{aa_1, \ldots, aa_m\}.$$

Оно содержит m различных элементов: если $aa_i = aa_j$, то $a_i = a_j$. Кроме того, $aH \cap H = \emptyset$: если $a_i = aa_j$, то $a \in H$.

Если $H \cup aH = G$, то все доказано. Если нет, то существует $b \in G \backslash (H \cup aH)$. Множество

$$bH = \{ba_1, \ldots, ba_m\}$$

также содержит m различных элементов и при этом

$$H \cap aH = \emptyset, \quad H \cap bH = \emptyset, \quad aH \cap bH = \emptyset.$$

Если $H \cup aH \cup bH = G$, то все доказано. Если нет, действуем как и раньше. Поскольку число элементов в G конечно, на каком-то шаге мы получим разложение

$$H \cup aH \cup bH \cup \ldots \cup cH = G$$

с конечным числом попарно непересекающихся множеств H, aH, bH, $\ldots$, cH. $\square$

Задача. Докажите, что в любой бесконечной группе число различных подгрупп бесконечно.

Задача. В конечной группе G выбраны подгруппы H_1 и H_2 порядка n_1 и n_2, соответственно. Докажите, что число элементов в множестве $H_1H_2 = \{g \in G : g = h_1h_2, h_1 \in H_1, h_2 \in H_2\}$ равно n_1n_2/d, где d — число элементов в пересечении $H_1 \cap H_2$.

2.14 Смежные классы, нормальные делители, фактор-группы

Пусть H — подгруппа группы G и $a \in G$. Множества

$$aH = \{x : \ x = ah, \ \ h \in H\} \qquad \text{и} \qquad Ha = \{y : \ y = ha, \ \ h \in H\}$$

называются *левым смежным классом* и *правым смежным классом* группы G по подгруппе H.

Если $b \in aH$, то $bH = aH$ (докажите!) — отсюда вытекает, что левые (правые) смежные классы либо совпадают, либо не пересекаются (на этом факте и было основано доказательство теоремы Лагранжа).

Подгруппа H называется *нормальной подгруппой* или *нормальным делителем* группы G, если

$$aH = Ha \quad \forall\, a \in G. \qquad \Leftrightarrow \qquad aha^{-1} \in H \quad \forall\, a \in G \quad \forall\, h \in H.$$

Пусть K — множество различных смежных классов для нормального делителя $H \subset G$. Определим произведение смежных классов следующим образом:

$$(aH)(bH) \equiv (ab)H.$$

Прежде всего, нужно убедиться в том, что если $a_1 \in aH,\ b_1 \in bH$, то $(a_1b_1)H = (ab)H$ (то есть, определение корректно). Пусть $a_1 = ah_1,\ b_1 = bh_2,\ h_1, h_2 \in H$. Значит, если $h \in H$, то

$$(a_1b_1)h \ = \ ah_1bh_2h \ = \ (ab)(b^{-1}h_1b)(h_2h) \in (ab)H. \quad \square$$

Нетрудно проверить, что операция умножения смежных классов превращает множество K в группу. Эта группа называется *фактор-группой* группы G по нормальному делителю H. Обозначение: $K = G/H$.

Задача. Какие смежные классы являются подгруппами?

Задача. Докажите, что любая абелева группа порядка pq, где p и q — различные простые числа, является циклической.

2.15 Изоморфизмы групп

Рассмотрим группу H с операцией $*$ и группу G с операцией $\circ$. Обратимое отображение $f : \ H \to G$ называется *изоморфизмом*, если

$$f(a * b) = f(a) \circ f(b) \quad \forall \ a, b \in H. \qquad (\#)$$

Свойство $(\#)$ называется свойством *сохранения операций*. Легко видеть, что обратное отображение $f^{-1} : \ G \to H$ также является изоморфизмом. Группы H и G называются *изоморфными*. Обозначение: $H \simeq G$. Несмотря на формальные различия в определении элементов и операций, изоморфные группы можно считать одинаковыми с точки зрения свойств их операций.

Например, любые две конечные циклические группы одного порядка n будут изоморфными. Если $a^0, a^1, \ldots, a^{n-1}$ — все различные элементы группы H, то $a^n = a^0$ (докажите!). Пусть $b^0, b^1, \ldots, b^{n-1}$ — все различные элементы группы G. Тогда определим отображение f правилом $f(a^k) = b^k$. Оно является изоморфизмом, поскольку

$$f(a^{k+m}) \ = \ b^{k+m} \ = \ b^k\, b^m \ = \ f(a^k)\, f(a^m).$$

Задача. Докажите, что группа положительных рациональных чисел относительно умножения не изоморфна группе всех рациональных чисел с операцией сложения.

Задача. Найдите все группы, изоморфные любой своей неединичной подгруппе.

2.16 Гомоморфизмы групп

Отображение $f : \ H \to G$ называется *гомоморфизмом*, если выполняется свойство сохранения операций $(\#)$ (при этом обратимость отображения не требуется).

Обозначим через e_G единичный элемент группы G. Его полный прообраз $K = f^{-1}(e_G)$ называется *ядром гомоморфизма* f. Множество $f(H)$ называется *образом гомоморфизма* f.

Утверждение. *Ядро гомоморфизма $f : H \to G$ является нормальной подгруппой группы H. Образ гомоморфизма f является подгруппой группы G.*

Доказательство. Пусть e — единица группы H и K — ядро гомоморфизма f. Для любого $a \in H$ находим $f(ae) = f(a)f(e) = f(a) \ \Rightarrow \ f(e) = e_G$. Итак, $e \in K$.

Далее, если $a \in H$, то $f(e) = f(aa^{-1}) = f(a)f(a^{-1}) = e_G \;\Rightarrow\; f(a^{-1}) = (f(a))^{-1}$. Предположим, что $a \in K$. Тогда $f(a^{-1}) = e_G^{-1} = e_G \;\Rightarrow\; a^{-1} \in K$.

Если $f(a) = f(b) = e_G$, то $f(ab) = e_G e_G = e_G \;\Rightarrow\; ab \in K$.

Наконец, проверим нормальность подгруппы K. Пусть $a \in H$, $b \in K$. Тогда $f(aba^{-1}) = f(b) = e_G \;\Rightarrow\; aba^{-1} \in K$. $\square$

Теорема о гомоморфизме. *Пусть $f : H \to G$ — гомоморфизм группы H в группу G и пусть K — его ядро. Тогда $f(H) \simeq H/K$.*

Доказательство. Отображение $\Phi : H/K \to f(H)$ определим следующим образом:

$$\Phi(aK) = f(a), \quad a \in H.$$

Пусть $a_1 = ab_1$, $b_1 \in K$. Тогда $f(a_1) = f(a)$.

Обратно, если $f(a_1) = f(a)$, то $f(a_1 a^{-1}) = e_G \;\Rightarrow\; a_1 a^{-1} \in K$. Таким образом, отображение определено корректно (то есть, не зависит от выбора представителя a в смежном классе aK) и является взаимно-однозначным. Легко видеть, что оно сохраняет операции:

$$\Phi((aK)(bK)) = \Phi((ab)K) = f(ab) = f(a)f(b) = \Phi(aK)\Phi(bK). \qquad \square$$

Теорема показывает, что изучать образы группы при всевозможных гомоморфизмах можно “изнутри”: для полного описания соответствующих подгрупп группы G, в которой размещаются образы элементов, не требуется знание самой группы G — вопрос сводится к изучению фактор-групп по нормальным делителям заданной группы.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

2.17 Избыточность в определении группы

Пусть G — непустое множество с ассоциативной алгебраической операцией. Элемент $e \in G$ называется *правой единицей*, если $ae = a$ для всех $a \in G$. Элемент $b \in G$ называется *правым обратным* для $a \in G$ относительно правой единицы e, если $ab = e$.

Теорема. *Пусть G имеет правую единицу e, относительно которой для каждого элемента $a \in G$ существует правый обратный элемент. Тогда G является группой.*

Доказательство. Докажем, что правая единица e является единичным элементом. Возьмем произвольный элемент a и положим $c = ea$. Согласно условию теоремы, существуют $b, d \in G$ такие, что $ab = e$ и $bd = e$. Отсюда $a = ed$. Далее, $cb = e(ab) = e$, откуда $c = ed = a$.

Докажем теперь, что b является обратным элементом для a. Пусть $c = ba$. Тогда $cb = b(ab) = b$, и значит, $c = bd = e$. $\square$

Лекция 3

ОСНОВНАЯ ЧАСТЬ

3.1 Система линейных алгебраических уравнений

Система уравнений вида

$$\begin{cases} a_{11}x_1 & + \ldots + & a_{1k}x_k & = & b_1, \\ & & & \ldots & \\ a_{n1}x_1 & + \ldots + & a_{nk}x_k & = & b_n \end{cases} \tag{1}$$

относительно неизвестных величин $x_1, \ldots, x_k$ называется *системой линейных алгебраических уравнений*. Мы уже знаем, что с помощью матричных обозначений ее можно записать в виде

$$Ax = b, \qquad A = \begin{bmatrix} a_{11} & \ldots & a_{1k} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nk} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \ldots \\ x_k \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \ldots \\ b_n \end{bmatrix}.$$

Множество матриц размеров $n \times k$ с элементами $a_{ij} \in \mathbb{R}$, где $\mathbb{R}$ — множество вещественных чисел, обозначим $\mathbb{R}^{n\times k}$. В согласии с этим обозначением $\mathbb{R}^{n\times 1}$ и $\mathbb{R}^{k\times 1}$ — множества матриц-столбцов, имеющих, соответственно, n и k элементов. Для краткости будем писать $\mathbb{R}^n = \mathbb{R}^{n\times 1}$ и $\mathbb{R}^k = \mathbb{R}^{k\times 1}$ и называть матрицы-столбцы *векторами*.

Матрица $A \in \mathbb{R}^{n\times k}$ называется *матрицей коэффициентов*, вектор $b \in \mathbb{R}^n$ — *правой частью*, а вектор $x \in \mathbb{R}^k$ — *решением* системы (1).

3.2 Линейные комбинации

Для понимания сути дела исключительно полезна также следующая интерпретация системы (1). Согласно определению умножения матрицы на число, если $\alpha \in \mathbb{R}$, то

$$\alpha \begin{bmatrix} b_1 \\ \ldots \\ b_n \end{bmatrix} \equiv \begin{bmatrix} \alpha b_1 \\ \ldots \\ \alpha b_n \end{bmatrix}.$$

Пусть $a_1, \ldots, a_k$ — столбцы матрицы A:

$$A = [a_1, \ldots, a_k], \qquad a_1, \ldots, a_k \in \mathbb{R}^n.$$

Тогда соотношения (1) равносильны равенству между векторами

$$x_1 a_1 + \ \ldots \ + x_k a_k = b. \tag{2}$$

Выражение $x_1a_1 + \ldots + x_ka_k$ называется *линейной комбинацией* векторов $a_1, \ldots, a_k$, а числа $x_1, \ldots, x_k$ — *коэффициентами линейной комбинации*. Множество всевозможных линейных комбинаций векторов $a_1, \ldots, a_k$

$$L(a_1, \ldots, a_k) = \{\alpha_1 a_1 + \ldots + \alpha_k a_k : \alpha_1, \ldots, \alpha_k \in \mathbb{R}\}$$

называется *линейной оболочкой* векторов $a_1, \ldots, a_k$.

Таким образом, равенство (2) означает, что

$$b \in L(a_1, \ldots, a_k). \tag{3}$$

Другими словами, система (1) имеет решение (совместна) тогда и только тогда, когда правая часть b принадлежит линейной оболочке (является линейной комбинацией) столбцов матрицы коэффициентов.

3.3 Линейная зависимость

Векторы, все элементы которых равны нулю, называют *нулевыми векторами*, а иногда просто *нулями*. Любой нулевой вектор будем обозначать символом 0.

Линейная комбинация векторов называется *нетривиальной*, если хотя бы один из ее коэффициентов отличен от нуля. Система (другими словами, непустая упорядоченная совокупность конечного числа) векторов называется *линейно зависимой*, если для них существует нетривальная линейная комбинация, равная нулевому вектору.

Лемма 1. *Если $a_1, \ldots, a_k$ — линейно зависимая система $k > 1$ ненулевых векторов, то в ней существует вектор a_m, $m > 1$, являющийся линейной комбинацией предыдущих векторов:*

$$a_m \in L(a_1, \ldots, a_{m-1}).$$

Доказательство. Рассмотрим равную нулю нетривиальную линейную комбинацию

$$\alpha_1 a_1 + \ldots + \alpha_k a_k = 0,$$

и пусть m — наибольший номер такой, что $\alpha_m \neq 0$. Если $m = 1$, то $\alpha_1 a_1 = 0$ и, поскольку $\alpha_1 \neq 0$, получаем $a_1 = 0$, что противоречит условию леммы. Следовательно, $m > 1$. Тогда

$$\alpha_1 a_1 + \ldots + \alpha_m a_m = 0 \quad \Rightarrow \quad a_m = \left(-\frac{\alpha_1}{\alpha_m}\right) a_1 + \ldots + \left(-\frac{\alpha_{m-1}}{\alpha_m}\right) a_{m-1}. \quad \square$$

Задача. По заданным ненулевым числам $a_0, \ldots, a_{2n}$ составлены матрицы

$$A_k = \begin{bmatrix} a_0 & a_1 & a_2 & \ldots & a_k \\ a_1 & a_2 & a_3 & \ldots & a_{k+1} \\ a_2 & a_3 & a_4 & \ldots & a_{k+2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_k & a_{k+1} & a_{k+2} & \ldots & a_{2k} \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}, \qquad k = 1, \ldots, n,$$

при этом столбцы каждой из них линейно зависимы. Докажите, что существует число q такое, что $a_k = a_0 q^k$, $0 \leq k \leq 2n$.

3.4 Линейная независимость

Система векторов называется *линейно независимой*, если она не является линейно зависимой. Таким образом, если векторы $a_1, \ldots, a_k$ линейно независимы, то

$$\alpha_1 a_1 + \ldots + \alpha_k a_k = 0 \quad \Rightarrow \quad \alpha_1 = \ldots = \alpha_k = 0.$$

Лемма 2. *Любая подсистема линейно независимой системы является линейно независимой.*

Доказательство. Предположим, что подсистема линейно зависима. Значит, существует нетривиальная линейная комбинация векторов данной подсистемы, равная нулю. Тогда линейная комбинация векторов исходной системы с теми же коэффициентами при векторах из подсистемы и нулевыми коэффициентами при других векторах является нетривиальной линейной комбинацией, равной нулю. Получаем противоречие с линейной независимостью исходной системы. $\square$

Лемма 3. *Если вектор является линейной комбинацией линейно независимых векторов, то коэффициенты этой линейной комбинации определены единственным образом.*

Доказательство. Пусть векторы $a_1, \ldots a_k$ линейно независимы и

$$b \;=\; \alpha_1 a_1 \;+\; \ldots,\; +\alpha_k a_k \;=\; \beta_1 a_1 \;+\; \ldots,\; +\beta_k a_k.$$

Отсюда

$$(\alpha_1 - \beta_1) a_1 \;+\; \ldots \;+\; (\alpha_k - \beta_k) a_k \;=\; 0 \quad \Rightarrow \quad \alpha_1 - \beta_1 = \ldots = \alpha_k - \beta_k = 0. \qquad \square$$

Задача. Для каждого n найдите все значения параметра a, при которых столбцы трехдиагональной матрицы

$$A = \begin{bmatrix} a & 1 & & & \\ -1 & a & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & a & 1 \\ & & & -1 & a \end{bmatrix}$$

порядка n линейно независимы.

3.5 Транзитивность линейной зависимости

Важное (хотя и очевидное) свойство: если

$$L(c_1, \ldots, c_r) \subset L(b_1, \ldots, b_m) \quad \text{и} \quad L(b_1, \ldots, b_m) \subset L(a_1, \ldots, a_k),$$

то

$$L(c_1, \ldots, c_r) \subset L(a_1, \ldots, a_k).$$

3.6 Монотонность числа линейно независимых векторов

Лемма 4. *Пусть каждая из систем векторов $b_1, \ldots, b_m$ и $a_1, \ldots, a_k$ линейно независима, и предположим, что*

$$L(b_1, \ldots, b_m) \subset L(a_1, \ldots, a_k). \qquad (*)$$

Тогда $m \le k$.

Доказательство. Согласно $(*)$, система

$$b_1,\, a_1,\, \ldots,\, a_k$$

линейно зависима. В силу Леммы 1 существует вектор, являющийся линейной комбинацией предыдущих векторов, пусть это будет вектор

$$a_k \in L(b_1,\, a_1,\, \ldots,\, a_{k-1}).$$

Отсюда следует,что

$$L(a_1,\, \ldots,\, a_k) \subset L(b_1,\, a_1,\, \ldots,\, a_{k-1}).$$

В силу транзитивности линейной зависимости

$$L(b_1,\, \ldots,\, b_m) \subset L(b_1,\, a_1,\, \ldots,\, a_{k-1}),$$

поэтому система

$$b_2,\, b_1,\, a_1,\, \ldots,\, a_{k-1}$$

линейно зависима. В силу Леммы 1 и в этой системе существует вектор, линейно выражающийся через предыдущие, причем таковым не может быть вектор b_1 (векторы b_1, b_2 линейно независимы как подсистема линейно независимой системы (Лемма 2)). Не ограничивая общности, будем считать, что

$$a_{k-1} \in L(b_2,\, b_1,\, a_1,\, \ldots,\, a_{k-2}).$$

Предположим, что $m > k$. Тогда, продолжая предыдущие построения, на k-ом шаге получаем

$$L(a_1,\, \ldots,\, a_k) \subset L(b_k,\, b_{k-1},\, \ldots,\, b_1).$$

Следовательно, $b_{k+1} \in L(b_k,\, b_{k-1},\, \ldots,\, b_1)$, а это противоречит предположению о линейной независимости векторов $b_1,\, \ldots,\, b_m$. Полученное противоречие доказывает, что $m \le k$. $\square$

3.7 Базис и размерность

Линейно независимая система векторов $b_1,\, \ldots,\, b_m \in V = L(a_1,\, \ldots,\, a_k)$ называется *базисом* линейной оболочки V, если $L(b_1,\, \ldots,\, b_m) = V$.

Теорема о базисах. *Любые базисы линейной оболочки V содержат одно и то же число векторов.*

Доказательство. Пусть $b_1,\, \ldots,\, b_m$ и $c_1,\, \ldots,\, c_r$ — два базиса данной линейной оболочки. Ясно, что

$$L(b_1,\, \ldots,\, b_m) = L(c_1,\, \ldots,\, c_r).$$

Применяя Лемму 4 два раза, получаем два неравенства: $m \le r$ и $r \le m$. Отсюда $m = r$. $\square$

Определение. Число векторов в базисах линейной оболочки V называется ее *размерностью* и обозначается $\dim V$.

Теорема о размерности линейной оболочки: $\dim L(a_1,\, \ldots,\, a_k) \le k$.

Доказательство. Достаточно заметить, что в качестве базиса линейной оболочки заданной системы векторов можно выбрать их максимальную линейно независимую систему. $\square$

3.8 Дополнение до базиса

Лемма о дополнении до базиса. *Любая линейно независимая система векторов* $b_1, \ldots, b_m \in L(a_1, \ldots, a_k)$ *является подсистемой некоторого базиса данной линейной оболочки.*

Доказательство. Достаточно рассмотреть случай, когда векторы $a_1, \ldots, a_k$ линейно независимы. Система векторов $b_1, \ldots, b_m, a_1, \ldots, a_k$ линейно зависима. В силу Леммы 1 в ней существует вектор, линейно выражающийся через предыдущие. Уберем этот вектор и рассмотрим оставшуюся подсистему. Если она линейно независима, то и является базисом линейной оболочки $L(a_1, \ldots, a_k)$. Если нет, в ней имеется вектор, линейно выражающийся через предыдущие. Исключим и его из системы, рассмотрим оставшуюся подсистему, и так далее. В итоге система векторов $b_1, \ldots, b_m$ будет дополнена до базиса некоторыми из векторов $a_1, \ldots, a_k$. $\square$

3.9 Существование базиса

Для любой ли линейной оболочки существует базис? Согласно лемме о дополнении до базиса, базис существует, если в линейной оболочке существует линейно независимая подсистема векторов. Так будет, если существует хотя бы один ненулевой вектор.

Таким образом, базиса нет только в случае нулевой линейной оболочки, содержащей единственный вектор — нулевой. По определению, размерность нулевой линейной оболочки равна нулю.

3.10 Совместность системы линейных алгебраических уравнений

Теорема 1. *Система линейных алгебраических уравнений* $Ax = b, \quad A = [a_1, \ldots, a_k]$, *совместна тогда и только тогда, когда*

$$L(a_1, \ldots, a_k) = L(a_1, \ldots, a_k, b).$$

Доказательство. В любом случае имеем

$$L(a_1, \ldots, a_k) \subset L(a_1, \ldots, a_k, b). \qquad (*)$$

Если система совместна, то $b \in L(a_1, \ldots, a_k)$. Следовательно,

$$L(a_1, \ldots, a_k, b) \subset L(a_1, \ldots, a_k). \qquad (**)$$

Включения $(*)$ и $(**)$ доказывают равенство двух линейных оболочек. Если имеет место $(**)$, то очевидно, что $b \in L(a_1, \ldots, a_k)$, а это и означает совместность системы $Ax = b$. $\square$

Теорема 2. *Если* $n = k$, *то в случае линейной независимости векторов* $a_1, \ldots, a_n$ *система линейных алгебраических уравнений* $Ax = b$ *совместна и имеет единственное решение.*

Доказательство. Очевидно,

$$a_1, \ldots, a_n \in L(e_1, \ldots, e_n),$$

где $e_1, \ldots, e_n$ — столбцы единичной матрицы размеров $n \times n$ (на i-м месте в векторе e_i находится 1, а все остальные элементы равны 0). В силу теоремы о дополнении до базиса существует базис из $r \geq n$ векторов, содержащий векторы $a_1, \ldots, a_n$. В силу теоремы о размерности линейной оболочки $r \leq n$. По той же причине векторы $a_1, \ldots, a_n$ образуют базис в $L(a_1, \ldots, a_k, b)$. Поэтому $b \in L(a_1, \ldots, a_n)$, что и доказывает совместность системы. Единственность решения вытекает из Леммы 3. $\square$

Задача. Система линейных алгебраических уравнений вида

$$\begin{bmatrix} a_0 & a_1 & a_2 \\ a_1 & a_0 & a_1 \\ a_2 & a_1 & a_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

имеет решение, причем $x_1 \neq 0$. Докажите, что столбцы матрицы коэффициентов линейно независимы.

Лекция 4

ОСНОВНАЯ ЧАСТЬ

4.1 Индикатор линейной зависимости

Рассмотрим систему векторов $a_1, \ldots, a_n \in \mathbb{R}^n$ и попробуем сконструировать индикатор линейной зависимости — функцию $f(a_1, \ldots, a_n)$, которая равна нулю в случае линейной зависимости данной системы. При этом функция f должна иметь как можно более простой вид: пусть f будет линейна по каждому аргументу при фиксированных значениях остальных аргументов.

Дадим точную формулировку требований к функции f:

(A) для любого $1 \le i \le n$ функция *линейна* по i-му аргументу (функция должна иметь "простой вид"):

$$f(a_1, \ \ldots, a_{i-1}, \ \ \alpha a + \beta b, \ \ a_{i+1}, \ \ldots \ a_n) =$$

$$\alpha\, f(a_1, \ \ldots, a_{i-1}, \ \ a, \ \ a_{i+1}, \ \ldots, \ a_n) \ + \ \beta\, f(a_1, \ \ldots, a_{i-1}, \ \ b, \ \ a_{i+1}, \ \ldots, a_n)$$

для любых векторов $a, b \in \mathbb{R}^n$ и чисел $\alpha, \beta \in \mathbb{R}$;

(B) если система векторов $a_1, \ldots, a_n$ линейно зависима, то $f(a_1, \ldots, a_n) = 0$;

(C) функция принимает заданное ненулевое значение на заданной линейно независимой системе (условие нормировки):

$$f(e_1, \ldots, e_n) = 1,$$

где $e_1, \ldots, e_n$ — столбцы единичной матрицы размеров $n \times n$.

Функцию f с указанными свойствами будем называть индикатором линейной зависимости. Для ее построения нам понадобится понятие подстановки.

4.2 Подстановки и перестановки

Обратимое отображение $\sigma : \mathrm{N} \to \mathrm{N}$, где $\mathrm{N} = \{1, 2, \ \ldots, n\}$, называется *подстановкой* степени n. Для обозначения подстановки σ часто используется таблица

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix},$$

в которой числа $\sigma(1), \sigma(2), \ldots, \sigma(n)$ образуют *перестановку* чисел $1, 2, \ldots, n$ (это равносильно обратимости отображения σ).

Определим произведение подстановок a и b как отображение, получаемое последовательным выполнением (композицией) отображений b и a:

$$(ab)(i) = a(b(i)), \quad i \in \mathrm{N}.$$

Это алгебраическая операция на множестве всех подстановок степени n, относительно которой оно является группой. В самом деле, ассоциативность очевидна (этим свойством всегда обладает композиция отображений). Роль единичного элемента играет тождественное отображение

$$e = \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix},$$

а обратным элементом для σ является обратное отображение σ^{-1}.

Группа подстановок степени n называется *симметрической группой* степени n и обозначается S_n. Это один из важнейших примеров конечных групп (групп с конечным числом элементов; при этом число элементов называется *порядком* группы). Нетрудно проверить, что порядок группы S_n равен $n! = 1 \cdot 2 \cdot \ldots \cdot n$.

Название симметрической группы навеяно определением так называемых *симметрических функций*: так называется функция $F(x_1, \ldots, x_n)$, если она инвариантна относительно любых подстановок своих аргументов:

$$F(x_1, \ldots, x_n) = F(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \quad \forall\, \sigma \in S_n.$$

Пример симметрической функции (определяемой числовым параметром k):

$$F_k(x_1, \ldots, x_n) \;=\; \sum_{i=1}^{n} x_i^k.$$

4.3 Циклы и транспозиции

Подстановка $a \in S_n$ называется *циклом* длины k, если имеется k попарно различных номеров $i_1, \ldots, i_k \in \mathrm{N}$ таких, что

(1) $a(i_1) = i_2,\ a(i_2) = i_3, \ldots,\ a(i_{k-1}) = i_k,\ a(i_k) = i_1,$

(2) $a(i) = i \quad \forall\, i \in \mathrm{N} \setminus \{i_1, i_2, \ldots, i_k\}.$

Для обозначения цикла a удобно использовать запись

$$a = (i_1, \ldots, i_k).$$

Цикл длины 2 называется также *транспозицией*.

Циклы $a = (i_1, \ldots, i_k)$ и $b = (j_1, \ldots, j_m)$ называются *независимыми*, если

$$\{i_1, \ldots, i_k\} \;\cap\; \{j_1, \ldots, j_m\} \;=\; \emptyset.$$

(1) *Любые независимые циклы a и b коммутируют: $ab = ba$.*

(2) *Любая подстановка $\sigma \in S_n$ представима в виде произведения независимых циклов однозначно с точностью до порядка сомножителей.*

(3) *Любой цикл длины k представим в виде произведения $k-1$ транспозиций.*

(4) *Любая подстановка представима в виде произведения транспозиций.*

Утверждение **(1)** проверяется непосредственно: в случае независимых циклов $a = (i_1, \ldots, i_k)$ и $b = (j_1, \ldots, j_m)$ находим

$$(ab)(i) = (ba)(i) = a(i) \quad \text{при} \quad i \in \{i_1, \ldots, i_k\},$$

$$(ab)(i) = (ba)(i) = b(i) \quad \text{при} \quad i \in \{j_1, \ldots, j_m\},$$

$$(ab)(i) = (ba)(i) = i \quad \text{при} \quad i \notin \{i_1, \ldots, i_k\} \cup \{j_1, \ldots, j_m\}.$$

Чтобы доказать **(2)**, возьмем произвольный номер j и рассмотрим последовательность номеров $j, \sigma(j), \sigma^2(j), \ldots$. Имеется только n различных значений — поэтому для каких-то $k < l$ должно быть $\sigma^k(j) = \sigma^l(j)$, откуда получаем $\sigma^{l-k}(j) = j$. Пусть k — наименьший номер такой, что $\sigma^k(j) = j$. Тогда получаем цикл

$$a = (j, \sigma(j), \sigma^2(j), \ldots, \sigma^{k-1}(j)),$$

для которого

$$\sigma(i) = a(i) \quad \text{при} \quad i \in \{j, \sigma(j), \sigma^2(j), \ldots, \sigma^{k-1}(j)\}.$$

Ясно, что подстановка $\sigma_1 = \sigma a^{-1}$ оставляет на месте индексы

$$i \in \{j, \sigma(j), \sigma^2(j), \ldots, \sigma^{k-1}(j)\}.$$

Далее, возьмем $j_1 \notin \{j, \sigma(j), \sigma^2(j), \ldots, \sigma^{k-1}(j)\}$ и аналогичным образом построим цикл b, выполняющий преобразования вида

$$j_1 \ \to \ \sigma_1(j_1) \ \to \ \sigma_1^2(j_1) \ \to \ \ldots .$$

(Заметим, что $\sigma_1^l(j_1) = \sigma^l(j_1)$ для всех l.) Продолжая подобные построения, мы неизбежно придем к тождественной подстановке

$$\sigma\, a^{-1} b^{-1} \ \ldots \ c^{-1} = e,$$

откуда

$$\sigma = c \ \ldots \ ba.$$

По построению циклы $a,\ b,\ \ldots, c$ независимы.

Утверждение **(3)** доказывается проверкой, например, следующего равенства:

$$(i_1, \ldots, i_k) \ = \ (i_1, i_2)(i_2. i_3) \ \ldots \ (i_{k-1}, i_k).$$

Утверждение **(4)** очевидно вытекает из **(2)** и **(3)**.

Задача. Докажите, что все множество подстановок степени n можно упорядочить таким образом, что каждая следующая подстановка будет получаться из предыдущей путем умножения справа на некоторую транспозицию.

4.4 Четность подстановки

Подстановка может быть разложена в произведение транспозиций многими разными способами. Например,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 5 & 3 & 1 & 2 & 4 & 6 \end{pmatrix} =$$

$$(1,\,7)\,(7,\,6)\,(6,\,4)\,(2,\,5) \;=\; (1,\,7)\,(7,\,6)\,(6,\,4)\,(7,\,2)\,(7,\,5)\,(7,\,2).$$

Однако, число транспозиций в любом разложении одной и той же подстановки обладает следующим важным свойством.

Лемма о числе транспозиций. *Четность числа транспозиций не зависит от способа представления подстановки в виде произведения транспозиций.*

Доказательство. Для заданной подстановки $\sigma \in S_n$

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

назовем *инверсией* пару $(i,\,j)$, если $i < j$, но $\sigma(i) > \sigma(j)$. Пусть $\delta(\sigma)$ — общее число инверсий для σ. Докажем, что для любой транспозиции τ разность $\delta(\sigma\tau) - \delta(\sigma)$ будет нечетным числом. Пусть $\tau = (i,\,j),\;\; i < j$. Тогда

$$\sigma\tau = \begin{pmatrix} 1 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ \sigma(1) & \dots & \sigma(i-1) & \sigma(j) & \sigma(i+1) & \dots & \sigma(j-1) & \sigma(i) & \sigma(j+1) & \dots & \sigma(n) \end{pmatrix}.$$

Предположим, что подстановка σ имеет k инверсий среди пар вида

$$(i,\,l), \quad \text{где} \quad l \;\in\; \{i+1, i+2,\,\dots,\,j-1\}, \tag{$*$}$$

m инверсий среди пар вида

$$(l,\,j), \quad \text{где} \quad l \;\in\; \{i+1, i+2,\,\dots,\,j-1\}, \tag{$**$}$$

и еще s инверсий среди любых других пар. Тогда $\sigma\tau$ будет иметь $j-i-1-k$ инверсий среди пар вида $(*)$ и $j-i-1-m$ инверсий среди пар вида $(**)$. Кроме того, среди любых других пар подстановка $\sigma\tau$ будет иметь $s+1$ инверсию, если пара $(i,\,j)$ не была инверсией, и $s-1$ в противном случае. Таким образом,

$$\delta(\sigma) \;=\; k \;+\; m \;+\; s, \qquad \delta(\sigma\tau) \;=\; (i-j-1-k) \;+\; (i-j-1-m) \;+\; s \;\pm\; 1.$$

Отсюда

$$\delta(\sigma\tau) - \delta(\sigma) \;=\; 2(i-j-1-k-m) \;\pm\; 1. \qquad \square$$

Следствие. *Четность числа транспозиций в разложении подстановки совпадает с четностью ее числа инверсий.*

Определение. Подстановка называется *четной*, если она является произведением четного числа транспозиций, и *нечетной* в противном случае.

Замечание. Рассмотрим функцию

$$\Delta(x_1,\,\dots,\,x_n) = \prod_{i<j} (x_j - x_i).$$

Тогда для любой подстановки $\sigma \in S_n$ имеет место одно из двух:

$$\Delta(x_1, \ldots, x_n) \;=\; \Delta(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \qquad \text{либо} \qquad \Delta(x_1, \ldots, x_n) \;=\; -\Delta(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Четные подстановки и только они знак сохраняют (первый случай), нечетные и только они знак меняют (второй случай).

Множество всех четных подстановок в S_n образует подгруппу (докажите!), которая называется *знакопеременной группой* степени n и обозначается A_n.

Задача. Докажите, что любую четную подстановку степени $n \geq 3$ можно представить в виде произведения циклов длины 3.

4.5 Единственность индикатора линейной зависимости

Вернемся к построению индикатора линейной зависимости — функции $f(a_1, \ldots, a_n)$ от векторов

$$a_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{n1} \end{bmatrix}, \quad a_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \cdots \\ a_{n2} \end{bmatrix}, \quad \ldots\ , \quad a_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdots \\ a_{nn} \end{bmatrix},$$

удовлетворяющей требованиям (A), (B), (C). Легко видеть, что

$$a_1 \;=\; \sum_{i_1=1}^{n} a_{i_1 1}\, e_{i_1}, \quad a_2 \;=\; \sum_{i_2=1}^{n} a_{i_1 2}\, e_{i_2}, \quad \ldots\ , \quad a_n \;=\; \sum_{i_n=1}^{n} a_{i_n n}\, e_{i_n},$$

где $e_1, e_2, \ldots, e_n$ — столбцы единичной матрицы размеров $n \times n$.

Если искомая функция f существует, то свойство (A) линейности по каждому аргументу приводит к выражению

$$f(a_1, \ldots, a_n) \;=\; \sum_{i_1=1}^{n} \ldots \sum_{i_n=1}^{n} a_{i_1 1}\, a_{i_2 2}\, \ldots\ a_{i_n n}\ \ f(e_{i_1}, e_{i_2}, \ldots, e_{i_n}).$$

Согласно требованию (B), $f = 0$ на любой линейно зависимой системе векторов. Очевидно, система векторов $e_{i_1}, e_{i_2}, \ldots, e_{i_n}$ будет линейно зависимой в том и только том случае, когда среди этих векторов есть равные (если все эти векторы попарно различны, то они образуют перестановку столбцов единичной матрицы). Следовательно, исключая из суммирования заведомые нули, находим

$$f(a_1, \ldots, a_n) \;=\; \sum_{\sigma \in S_n} a_{\sigma(1)1}\, a_{\sigma(2)2}\ \ldots\ a_{\sigma(n)n}\ f(e_{\sigma(1)i}, e_{\sigma(2)}, \ldots, e_{\sigma(n)}).$$

Далее, из требований (A) и (B) вытекает, что f должна менять знак при перестановке любых двух аргументов. Докажем это, например, для первого и второго аргументов. Учтем, что $f = 0$ в случае равных аргументов и воспользуемся линейностью по каждому аргументу:

$$0 = f(a_1 + a_2,\ a_1 + a_2,\ a_3,\ \ldots,\ a_n) =$$

$$f(a_1,\ a_1,\ a_3,\ \ldots,\ a_n) + f(a_1,\ a_2,\ a_3,\ \ldots,\ a_n) + f(a_2,\ a_1,\ a_3,\ \ldots,\ a_n) + f(a_2,\ a_2,\ a_3,\ \ldots,\ a_n).$$

Первое и четвертое слагаемые имеют совпадающие векторы и поэтому равны нулю. Отсюда

$$f(a_1,\, a_2,\, a_3,\, \ldots,\, a_n) = -f(a_2,\, a_1,\, a_3,\, \ldots,\, a_n).$$

Следовательно, если подстановка σ является транспозицией, то

$$f(e_{\sigma(1)},\, e_{\sigma(2)},\, \ldots,\, e_{\sigma(n)}) = -f(e_1,\, e_2,\, \ldots,\, e_n) = -1.$$

В общем случае подстановку σ можно разложить в произведение транспозиций. Пусть $\delta(\sigma)$ есть число транспозиций в каком-либо из разложений. Тогда

$$f(e_{\sigma(1)},\, e_{\sigma(2)},\, \ldots,\, e_{\sigma(n)}) = (-1)^{\delta(\sigma)}.$$

По лемме о числе транспозиций, четность числа $\delta(\sigma)$ не зависит от конкретного разложения в произведение транспозиций, поэтому величина $(-1)^{\delta(\sigma)}$ зависит только от σ. Назовем ее *знаком подстановки* и обозначим через $\operatorname{sgn}(\sigma)$. Окончательно,

$$f(a_1,\, \ldots,\, a_n) \;=\; \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\;\, a_{\sigma(1)1}\, a_{\sigma(2)2}\;\, \ldots\;\, a_{\sigma(n)n}. \qquad (*)$$

Мы доказали важное

Утверждение. *Если функция — индикатор линейной зависимости существует, то она определяется формулой* $(*)$.

4.6 Определитель

Определение. Функция вида $(*)$ называется *определителем* (детерминантом) матрицы A со столбцами $a_1,\, a_2,\, \ldots,\, a_n$ и обозначается $\det A$ или $|A|$.

Таким образом, если $A = [a_{ij}]$ — матрица размеров $n \times n$, то

$$\det A \;=\; |A| \;=\; \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \;=\; \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\;\, a_{\sigma(1)1}\, a_{\sigma(2)2}\;\, \ldots\;\, a_{\sigma(n)n}. \qquad (\triangle)$$

Частные случаи:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \;=\; a_{11}\, a_{22} - a_{21}\, a_{12},$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \;=\; a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{13}\, a_{21}\, a_{32} - a_{13}\, a_{22}\, a_{31} - a_{12}\, a_{21}\, a_{33} - a_{11}\, a_{23}\, a_{32}.$$

В общем случае сумма $(\triangle)$ содержит $n!$ членов, в каждом из них перемножаются n элементов матрицы, причем никакие два элемента в одном произведении не принадлежат одной строке или одному столбцу.

Несмотря на то, что определитель вводится как функция от матрицы, исторически понятие определителя сформировалось в 18 веке (сначала в трудах Лейбница и Крамера, затем теория определителей была развита в работах Вандермонда, Лапласа, Коши и К.Якоби) — намного раньше понятия матрицы, введенного в алгебру Гамильтоном и Кэли в середине 19 века. Конечно, с самого начала определитель связывался с квадратной таблицей $n \times n$ чисел (поэтому говорили об определителе *порядка* n). Это были, в частности, таблицы коэффициентов "квадратной"системы линейных алгебраических

уравнений. Но такие таблицы стали называть матрицами позже — когда для них ввели операцию умножения.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

4.7 Знакопеременная группа

Название знакопеременной группы A_n (группы всех четных подстановок степени n) навеяно следующим построением. Рассмотрим отображение

$$\mathrm{sgn} : S_n \to K = \{1,\ -1\}, \qquad \mathrm{sgn}(\sigma) = \begin{cases} 1, & \sigma\text{— четная подстановка,} \\ -1, & \sigma\text{— нечетная подстановка.} \end{cases}$$

На множестве "знаков" K введем операцию умножения так же, как для целых чисел. Тогда K превращается в абелеву группу, а отображение sgn сохраняет операции:

$$\mathrm{sgn}(\sigma_1\sigma_2) = \mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2) \qquad \forall\ \ \sigma_1,\ \sigma_2 \in S_n.$$

Поэтому sgn является гомоморфизмом группы S_n на группу K.

Напомним, что *ядром гомоморфизма* называется множество всех элементов группы, которые переводятся данным гомоморфизмом в единичный элемент (вообще говоря, другой группы — содержащей образы элементов при данном отображении). Таким образом, ядром гомоморфизма sgn является в точности знакопеременная группа A_n.

Погруппа A_n является в S_n нормальным делителем, поскольку ядро любого гомоморфизма группы является ее нормальным делителем. Вот, впрочем, прямая проверка того, что A_n есть нормальный делитель группы S_n: если $\sigma \in S_n$ и $h \in A_n$, то, очевидно, $\sigma h \sigma^{-1} \in A_n \quad \Rightarrow \quad \sigma A_n = A_n \sigma$ (левые смежные классы совпадают с правыми).

В данном случае имеется всего два различных смежных класса группы S_n по нормальной подгруппе A_n: $eA_n = A_n$ и τA_n, где e — тождественная подстановка, а τ — произвольная нечетная подстановка (например, транспозиция). В самом деле, если σ_1 и σ_2 одной четности, то $h = \sigma_1^{-1}\sigma_2 \in A_n \quad \Rightarrow \quad \sigma_1 A_n = \sigma_2 A_n$. Таким образом, факторгруппа S_n/A_n состоит из двух смежных классов. Она изоморфна группе "знаков" K: изоморфизм осуществляется отображением $\sigma A_n \to \mathrm{sgn}(\sigma)$ (здесь мы имеем частный случай более общей теоремы о гомоморфизме из Лекции 2).

4.8 Подгруппы симметрической группы

Теорема. *Любая конечная группа порядка n изоморфна некоторой подгруппе симметрической группы S_n.*

Доказательство. Пусть группа G имеет элементы $g_1,\ \ldots,\ g_n$. Тогда для любого i элементы $g_i g_1,\ \ldots,\ g_i g_n$ представляют собой перестановку элементов $g_1,\ \ldots,\ g_n$. Обозначим соответствующую подстановку через σ_i и определим отображение $\Phi : G \to S_n$ правилом $\Phi(g_i) = \sigma_i$. Очевидно, $\Phi(g_i g_j) = \sigma_i \sigma_j$. Поэтому Φ является гомоморфизмом $\Rightarrow$ его образ $\Phi(G)$ является подгруппой в S_n.

Остается заметить, что $\Phi(g_i) = \Phi(g_j) \ \Leftrightarrow\ g_i = g_j$. $\square$

4.9 Четность без инверсий

То, что четность числа транспозиций в любом разложении подстановки одна и та же, можно доказать и без подсчета числа инверсий. Это сразу же вытекает из следующего наблюдения.

Утверждение. *В любом разложении тождественной подстановки в произведение транспозиций их число четно.*

Доказательство. Пусть тождественная подстановка $e \in S_n$ разложена в произведение транспозиций $e = (ij)\dots(kl)$, в котором среди индексов i, j, $\dots$, k, l имеется ровно s различных. Ясно, что $2 \leq s \leq n$ и в случае $s = 2$ утверждение очевидно. Проведем индукцию по s. Пусть $s \geq 3$. Не ограничивая общности, можно считать, что индексы равны 1, $\dots$, s. Легко проверить, что $(1l)(kl) = (1k)(1l)$ для любых $k, l \neq 1$ и $(1l)(ij) = (ij)(1l)$ при $\{i, j\} \neq \{1, l\}$. Поэтому можно передвинуть все транспозиции вида $(1l)$ вправо и получить другое разложение

$$e \;=\; (i_1j_1)\dots(i_kj_k)\;(1l_1)\dots(1l_m)$$

с тем же числом транспозиций. Далее, если $l_1 = l_2$, то $(1l_1)(1l_2) = e$ и в последнем разложении можно убрать пару транспозиций $(1l_1)$, $(1l_2)$. Если же $l_1 \neq l_2$, то, используя равенство $(1l_1)(1l_2) = (l_1l_2)(1l_1)$, можно получить разложение с тем же числом транспозиций и меньшим на 1 числом транспозиций, содержащих индекс 1:

$$e \;=\; (i_1j_1)\dots(i_kj_k)(l_1l_2)\;(1l_1)(1l_3)\dots(1l_m).$$

Продолжая таким же образом, придем к разложению с числом транспозиций, уменьшенным на четное число, и, возможно, всего лишь одной транспозицией вида $(1l)$:

$$e \;=\; (i_1j_1)\dots(i_pj_p)\;(1l).$$

Поскольку $i_1, j_1, \dots, i_p, j_p \neq 1$, подстановка e переводит l в 1, что невозможно, так как она является тождественной. Поэтому

$$e \;=\; (i_1j_1)\dots(i_pj_p),$$

где индексы $i_1, j_1, \dots, i_p, j_p$ принимают значения от 2 до s. По индуктивному предположению, число p четно. $\square$

Лекция 5

ОСНОВНАЯ ЧАСТЬ

5.1 Определитель транспонированной матрицы

Пусть имеется прямоугольная матрица размеров $m \times n$:

$$A = [a_{ij}], \qquad 1 \le i \le m, \quad 1 \le j \le n.$$

Если поменять местами строки и столбцы, то получается новая матрица — размеров $n \times m$. Она называется *транспонированной* по отношению к A и обозначается $A^\top$:

$$A^\top = [a_{ji}], \qquad 1 \le j \le n, \quad 1 \le i \le m.$$

Утверждение. *Для любой квадратной матрицы* $\det A^\top = \det A$.

Доказательство. Согласно определению транспонированной матрицы и формуле ($\triangle$) из Лекции 4 для определителя матрицы порядка n,

$$\det A^\top = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\ a_{1\,\sigma(1)} \ldots, a_{n\,\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\ a_{\sigma^{-1}(1)\,1} \ldots, a_{\sigma^{-1}(n)\,n}$$

$$= \sum_{\sigma^{-1} \in S_n} \operatorname{sgn}(\sigma^{-1})\ a_{\sigma^{-1}(1)\,1} \ldots, a_{\sigma^{-1}(n)\,n} = \det A.$$

В последнем равенстве было принято во внимание, что $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$. $\square$

5.2 Определитель как функция столбцов (строк) матрицы

(1) *Определитель как функция столбцов матрицы является линейной функцией относительно каждого столбца: если* $A = [a_1, \ldots, a_n]$ *и* $a_i = \alpha p + \beta q$ *— линейная комбинация столбцов* p *и* q*, то*

$$\det A \ = \ \alpha\ \det A_p \ + \ \beta\ \det A_q,$$

где матрицы A_p *и* A_q *получаются из* A *заменой столбца* a_i *на* p *и* q*, соответственно.*

Доказательство. В соответствии с определением,

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\ a_{\sigma(1)1}\, a_{\sigma(2)2}\ \ldots\ a_{\sigma(i)i}\ \ldots\ a_{\sigma(n)n}$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\ a_{\sigma(1)1}\, a_{\sigma(2)2}\ \ldots\ (\alpha p_{\sigma(i)i} + \beta q_{\sigma(i)i})\ \ldots\ a_{\sigma(n)n}$$

$$= \quad \alpha \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\ a_{\sigma(1)1}\, a_{\sigma(2)2}\ \ldots\ p_{\sigma(i)i}\ \ldots\ a_{\sigma(n)n}$$
$$+\ \beta \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma)\ a_{\sigma(1)1}\, a_{\sigma(2)2}\ \ldots\ q_{\sigma(i)i}\ \ldots\ a_{\sigma(n)n}\ =\ \alpha\ \det A_p\ +\ \beta\ \det A_q. \quad \square$$

(2) *Определитель меняет знак при перестановке двух столбцов.*

Доказательство. Пусть матрица $B = [b_{ij}]$ отличается от A перестановкой столбцов a_i и a_j. Тогда для любой подстановки $\sigma \in S_n$

$$a_{\sigma(1)1}\ \ldots\ a_{\sigma(n)n}\ =\ b_{(\sigma\tau)(1)1}\ \ldots\ b_{(\sigma\tau)(n)n},$$

где $\tau = (i,\, j)$, и поскольку транспозиция меняет знак подстановки,

$$\operatorname{sgn}(\sigma\tau)\ =\ -\operatorname{sgn}(\sigma).$$

Легко видеть, что отображение $\sigma \to \sigma\tau$ задает взаимно-однозначное соответствие между подстановками. Каждый член суммы вида $(\triangle)$ определяется одной и только одной подстановкой. Подстановки σ и $\sigma\tau$ в разложениях $\det A$ и $\det B$ определяют члены с произведением одних и тех же элементов (в разном порядке), но с противоположными знаками. Значит, $\det A = -\det B$. $\square$

(3) *Если столбцы матрицы линейно зависимы, то ее определитель равен нулю.*

Доказательство. Прежде всего заметим, что определитель с двумя равными столбцами равен нулю, поскольку в силу утверждения **(2)** он равен себе самому с противоположным знаком.

Если столбцы $a_1, a_2, \ldots, a_n$ линейно зависимы, то хотя бы один из них линейно выражается через остальные. Пусть

$$a_i\ =\ \sum_{k\neq i} \alpha_k\, a_k.$$

Обозначим через B матрицу, полученную из A заменой столбца a_i на

$$a_i - \sum_{k\neq i} \alpha_k\, a_k\ =\ 0.$$

Опираясь на уже установленное свойство **(1)**, находим

$$0 = \det B = \det A - \sum_{k\neq i} \alpha_k\ \det A_k,$$

где матрица A_k получается из A заменой i-го столбца на a_k. Ясно, что в A_k равны i-й и k-й столбцы, поэтому $\det A_k = 0$. Таким образом, $\det A = \det B = 0$. $\square$

(4) *Определитель как функция строк матрицы обладает свойствами, аналогичными* **(1), (2), (3)**.

Доказательство. Достаточно учесть, что $\det A = \det A^\top$, и рассмотреть $\det A$ как функцию столбцов матрицы $A^\top$. $\square$

Задача. Даны матрицы-столбцы $u_1, \ldots, u_k, v_1, \ldots, v_k \in \mathbb{R}^n$ и $A = u_1 v_1^\top + \ldots + u_k v_k^\top$. Доказать, что $\det A = 0$, если $k < n$.

5.3 Существование индикатора линейной зависимости

Теорема. *Индикатор линейной зависимости (функция, наделенная свойствами* **(A)**, **(B)**, **(C)** *из первого раздела Лекции 4) существует, единствен и является определителем.*

Свойства **(A)** и **(B)** индикатора линейной зависимости совпадают с установленными выше свойствами определителя **(1)** и **(3)**. Свойство **(C)** означает, что определитель единичной матрицы равен 1 и является следствием следующего более общего утверждения.

Утверждение. *Определитель диагональной матрицы равен произведению элементов ее диагонали:*

$$\det \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix} = a_{11} \ldots a_{nn}.$$

Доказательство. Для диагональной матрицы в сумме ($\triangle$) для ее определителя есть только одно ненулевое слагаемое, равное произведению элементов главной диагонали. $\square$

5.4 Подматрицы и миноры

Для заданной матрицы $A = [a_{ij}]$ можно выбрать какие-то из ее строк и столбцов и составить таблицу элементов, расположенных на пересечении выбранных строк и столбцов. Такая таблица называется *подматрицей* матрицы A.

Пусть A — квадратная матрица порядка n. Чтобы задать квадратную подматрицу порядка k, нужно указать номера содержащих ее строк $1 \le i_1 < \ldots < i_k \le n$ и столбцов $1 \le j_1 < \ldots < j_k \le n$. Обозначим через $\mathcal{N}_k$ множество всех систем номеров $(i_1, \ldots, i_k)$, упорядоченных по возрастанию $1 \le i_1 < \ldots < i_k \le n$. Тогда задание подматрицы равносильно выбору двух конкретных систем номеров

$$I = (i_1, \ldots, i_k) \in \mathcal{N}_k, \qquad J = (j_1, \ldots, j_k) \in \mathcal{N}_k.$$

Подматрица на строках с номерами из I и столбцах с номерами из J обозначается

$$A(I,\, J) \,=\, [a_{i_p\, j_q}], \qquad 1 \le p \le k, \quad 1 \le q \le k.$$

Пусть $I' = (i'_1, \ldots, i'_m)$ — еще одна система номеров, упорядоченных по возрастанию $1 \le i'_1 < \ldots < i'_m \le n$. Назовем систему I' *дополнительной* для $I = (i_1, \ldots, i_k)$, если

$$\{i_1,\, \ldots,\, i_k\} \cap \{i'_1,\, \ldots,\, i'_m\} = \emptyset, \quad \{i_1,\, \ldots,\, i_k\} \cup \{i'_1,\, \ldots,\, i'_m\} = \{1,\, \ldots,\, n\}.$$

Очевидно, в этом случае $k + m = n$.

Пусть заданы системы строчных и столбцовых номеров $I, J \in \mathcal{N}_k$ и пусть I' и J' — дополнительные системы, соответственно, для I и J. Подматрица $A(I',\, J')$ порядка $m = n - k$ называется дополнительной подматрицей по отношению к подматрице $A(I,\, J)$ порядка k.

Определитель подматрицы порядка k называется также *минором* порядка k, а определитель соответствующей дополнительной подматрицы — *дополнительным минором*.

5.5 Замечание о подстановках

Как мы знаем, подстановка σ степени n задается таблицей

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}.$$

Поскольку отображение полностью определяется указанием соответствий $i \to \sigma(i)$, порядок столбцов в этой таблице не имеет значения. Другими словами, для любой подстановки $\pi \in S_n$ таблица

$$\widetilde{\sigma} = \begin{pmatrix} \pi(1) & \pi(2) & \dots & \pi(n) \\ \sigma(\pi(1)) & \sigma(\pi(2)) & \dots & \sigma(\pi(n)) \end{pmatrix}$$

однозначно определяет ту же самую подстановку $\sigma = \widetilde{\sigma}$.

При этом очевидно, что четность числа инверсий для σ совпадает с четностью суммы числа инверсий для подстановок π и $\sigma\pi$ (поскольку четность числа инверсий для произведения $\sigma\pi$ совпадает с четностью суммы числа инверсий для σ и π). Отсюда ясно, что если подстановка задана таблицей вида

$$\begin{pmatrix} s(1) & s(2) & \dots & s(n) \\ t(1) & t(2) & \dots & t(n) \end{pmatrix}, \qquad s, t \in S_n,$$

то ее знак равен произведению знаков подстановок s и t.

5.6 Разбиение множества подстановок на подмножества

Пусть $J = (j_1, \dots, j_k)$ — фиксированная система номеров и $(j'_1, \dots, j'_m)$ — система дополнительных номеров. Таким образом, $m = n - k$. Возьмем любую систему номеров $I = (i_1, \dots, i_k) \in \mathcal{N}_k$ с дополнительной системой номеров $(i'_1, \dots, i'_m)$ и рассмотрим подстановки степени n вида

$$\sigma = \sigma_{I,J}(\pi, \tau) = \begin{pmatrix} j_1 & \dots & j_k & j'_1 & \dots & j'_m \\ i_{\pi(1)} & \dots & i_{\pi(k)} & i'_{\tau(1)} & \dots & i'_{\tau(m)} \end{pmatrix}, \qquad \pi \in S_k, \quad \tau \in S_m. \qquad (*)$$

Множество всех таких подстановок при фиксированных I, J обозначим

$$S_n(I, J) \;=\; \{\sigma_{I,J}(\pi, \tau) : \quad \pi \in S_k, \quad \tau \in S_m\}.$$

Любой системе номеров $I = (i_1, \dots, i_k) \in \mathcal{N}_k$ поставим в соответствие число

$$\nu(I) = i_1 + \dots + i_k.$$

Лемма. *При фиксированной системе J подмножества $S_n(I, J)$ не пересекаются при разных $I \in \mathcal{N}_k$ и их объединение дает множество всех подстановок степени n. Если $\sigma = \sigma_{I,J}(\pi, \tau) \in S_n(I, J)$, то*

$$\operatorname{sgn}(\sigma_{I,J}(\pi, \tau)) \;=\; \operatorname{sgn}(\pi)\, \operatorname{sgn}(\tau)\, (-1)^{\nu(I)+\nu(J)}.$$

Доказательство. Первое утверждение леммы о разбиении S_n на непересекающиеся подмножества вида $S_n(I, J)$ очевидно.

В силу сделанного выше замечания о подстановках, знак подстановки $\sigma_{I,J}(\pi,\,\tau)$, определяемой таблицей $(*)$, есть произведение знаков подстановок вида

$$\begin{pmatrix} 1 & \dots & k & k+1 & \dots & k+m \\ j_1 & \dots & j_k & j'_1 & \dots & j'_m \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 1 & \dots & k & k+1 & \dots & k+m \\ i_{\pi(1)} & \dots & i_{\pi(k)} & i'_{\tau(1)} & \dots & i'_{\tau(m)} \end{pmatrix}.$$

Подсчитаем число инверсий для первой подстановки. Принимая во внимание упорядоченность номеров в системах $(j_1,\dots,j_k)$, $(j'_1,\dots,j'_m)$ и их взаимную дополнительность, приходим к выводу о том, инверсию могут образовывать только пары вида

$$(p,q), \qquad \text{где} \qquad p\in\{1,\dots,k\}, \quad q\in\{k+1,\dots,k+m\}. \qquad (**)$$

При этом ясно, что j_1 порождает j_1-1 инверсий, j_2 порождает j_2-2 инверсий, и так далее: общее число инверсий, таким образом, равно

$$(j_1-1)+(j_2-2)+\dots+(j_k-k).$$

Число инверсий для второй подстановки включает три слагаемых:

(1) число инверсий среди пар вида вида $(**)$;

(2) число инверсий среди пар вида (p,q), где $p,q\in\{1,\dots,k\}$;

(3) число инверсий среди пар вида (p,q), где $p,q\in\{k+1,\dots,k+m\}$.

Первое число равно, по аналогии с рассмотренным выше случаем,

$$(i_1-1)+(i_2-2)+\dots+(i_k-k),$$

второе — числу инверсий $\delta(\pi)$ для подстановки $\pi\in S_k$, третье — числу инверсий $\delta(\tau)$ для подстановки $\tau\in S_m$. Таким образом, четность числа инверсий для подстановки $\sigma(\pi,\,\tau)$ совпадает с четностью числа

$$\delta(\pi)+\delta(\tau)+(i_1+\dots+i_k)+(j_1+\dots+j_k) \;=\; \delta(\pi)+\delta(\tau)+\nu(I)+\nu(J). \qquad \square$$

5.7 Теорема Лапласа

Теорема Лапласа. *Пусть A — квадратная матрица порядка n. Зафиксируем любую систему k столбцов, выбрав $J\in\mathcal{N}_k$. Тогда вычисление определителя матрицы A сводится к вычислению миноров на фиксированных k столбцах и их дополнительных миноров:*

$$\det A \;=\; \sum_{I\in\mathcal{N}_k} \det A(I,\,J)\;\det A(I',\,J')\;(-1)^{\nu(I)+\nu(J)}.$$

Доказательство. Опираясь на результат леммы предыдущего раздела, находим

$$\det A=\sum_{I\in\mathcal{N}_k}\left(\sum_{\pi\in S_k}\sum_{\tau\in S_m}(a_{i_{\pi(1)}j_1}\dots a_{i_{\pi(k)}j_k})\;(a_{i'_{\tau(1)}j'_1}\dots a_{i'_{\tau(m)}j'_m})\;\operatorname{sgn}(\sigma(\pi,\,\tau))\right)$$

$$= \sum_{I\in\mathcal{N}_k} \left(\sum_{\pi\in S_k} \sum_{\tau\in S_m} (a_{i_{\pi(1)}j_1} \dots a_{i_{\pi(k)}j_k})\ (a_{i'_{\tau(1)}j'_1} \dots a_{i'_{\tau(m)}j'_m})\ \mathrm{sgn}(\pi)\,\mathrm{sgn}(\tau)) \right)\ (-1)^{\nu(I)+\nu(J)}$$

$$= \sum_{I\in\mathcal{N}_k} \left(\sum_{\pi\in S_k} (a_{i_{\pi(1)}j_1} \dots a_{i_{\pi(k)}j_k})\ \ \mathrm{sgn}(\pi) \right)\ \left(\sum_{\tau\in S_m} (a_{i'_{\tau(1)}j'_1} \dots a_{i'_{\tau(m)}j'_m})\ \mathrm{sgn}(\tau)) \right)\ (-1)^{\nu(I)+\nu(J)}.$$

Остается заметить, что первая и вторая скобки дают, соответственно, $\det A(I,J)$ и $\det A(I',J')$. $\square$

Величину $\det A(I',J')(-1)^{\nu(I)+\nu(J)}$ называют *алгебраическим дополнением* минора $\det A(I,J)$. Таким образом, теорема Лапласа утверждает, что при выборе любой системы столбцов *определитель матрицы равен сумме всевозможных миноров, расположенных на заданных столбцах, умноженных на их алгебраические дополнения.*

Поскольку $\det A = \det A^\top$, имеет место и такой вариант теоремы Лапласа: *при выборе любой системы строк определитель матрицы равен сумме всевозможных миноров, расположенных на данных строках, умноженных на их алгебраические дополнения.*

Задача. Матрица B с определителем $b = \det B$ получена из A с определителем $a = \det A$ прибавлением числа $c \neq 0$ к каждому элементу. Найти суммы алгебраических дополнений всех элементов (подматриц порядка 1) для A и для B.

5.8 Определитель блочно-треугольной матрицы

Рассмотрим блочно-треугольную матрицу порядка n:

$$A = \begin{bmatrix} P & R \\ 0 & Q \end{bmatrix}, \qquad P \in \mathbb{R}^{k\times k}, \quad Q \in \mathbb{R}^{m\times m}, \quad k+m=n.$$

Применение теоремы Лапласа для системы первых k столбцов (или строк) сразу же дает полезную формулу

$$\det A \ = \ \det P\ \det Q.$$

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

5.9 Функциональное доказательство теоремы Лапласа

Рассмотрим выражение

$$f(A) \ = \ \sum_{I\in\mathcal{N}_k} \det A(I,\,J)\ \det A(I',\,J')\ (-1)^{\nu(I)+\nu(J)}$$

как функцию строк матрицы A и докажем, что она обладает тремя свойствами:

- $f(A)$ линейна по каждому аргументу;
- если строки матрицы A линейно зависимы, то $f(A)=0$;
- если A — единичная матрица, то $f(A)=1$.

Первое и третье свойства очевидны. Для того чтобы доказать второе свойство, достаточно установить, что $f(A)$ меняет знак при перестановке двух строк. Более того, достаточно рассмотреть перестановку двух соседних строк. Пусть это будут строки с номерами s и $s+1$. Матрицу с переставленными строками обозначим B.

Пусть I, $\widehat{I} \in \mathcal{N}_k$. Определим на $\mathcal{N}_k$ взаимно-однозначное отображение, при котором I переходит в $\widehat{I}$, следующим правилом. Если s и $s+1$ оба входят или оба не входят в систему номеров I, то пусть $\widehat{I} = I$. Если s принадлежит I, а $s+1$ нет, то пусть $\widehat{I}$ получается из I заменой номера s на $s+1$. Если $s+1$ принадлежит I, а s нет, то пусть $\widehat{I}$ получается из I заменой номера $s+1$ на s. Ясно, что

$$f(B) = \sum_{I \in \mathcal{N}_k} \det B(I,J) \det B(I',J')(-1)^{\nu(I)+\nu(J)} = \Sigma_1(B) + \Sigma_2(B),$$

где Σ_1 содержит члены, для которых $I = \widehat{I}$, а Σ_2 — члены, для которых $I \neq \widehat{I}$.

Нам нужно доказать, что $f(B) = -f(A)$. Рассмотрим члены, для которых $I = \widehat{I}$:

- $s,\ s+1 \in I \quad \Rightarrow \quad \det B(I,J) = -\det A(I,J), \quad \det B(I',J') = \det A(I',J').$
- $s,\ s+1 \notin I \quad \Rightarrow \quad \det B(I,J) = \det A(I,J), \quad \det B(I',J') = -\det A(I',J').$

При этом $\nu(I) = \nu(\widehat{I})$ (поскольку $I = \widehat{I}$). Отсюда $\Sigma_1(B) = -\Sigma_1(A)$.

Теперь рассмотрим члены, для которых $I \neq \widehat{I}$. Заметим, что если I переходит в $\widehat{I}$, то $\widehat{I}$ переходит в I. Следовательно, сумма Σ_2 разбивается на сумму пар членов, отвечающих I и $\widehat{I}$. При этом находим

$$\det B(I,J) = \det A(\widehat{I},J), \quad \det B(I',J') = \det A(\widehat{I}',J'),$$

$$\det B(\widehat{I},J) = \det A(I,J), \quad \det B(\widehat{I}',J') = \det A(I',J').$$

В то же время, $\nu(\widehat{I}) = \nu(I) \pm 1$. Поэтому $\Sigma_2(B) = -\Sigma_2(A)$. Таким образом, функция $f(A)$ является индикатором линейной зависимости, а в силу его единственности — определителем матрицы A. $\square$

5.10 Определители с нулевыми членами

Теорему Лапласа удобно применять, когда среди миноров на выбранных столбцах (или строках) оказывается много нулевых. Часто это связано с наличием большого числа нулей в матрице. Иногда нулей оказывается настолько много, что каждый член определителя содержит нулевой сомножитель и поэтому равен нулю. Очевидно, так обстоит дело, если матрица имеет нулевой столбец или нулевую строку. Следующее утверждение представляет собой нетривиальное обобщение этого наблюдения.

Теорема Холла. *Для того чтобы все члены определителя матрицы порядка n были равны нулю, необходимо и достаточно существование нулевой подматрицы размеров $p \times q$ с условием $p+q > n$.*

Доказательство достаточности является простым упражнением. А вот доказательство необходимости требует уже изрядной изобретательности.

Доказательство необходимости. Проведем индукцию по n. При $n = 1$ утверждение очевидно. Предположим, что оно доказано для любых матриц порядка $k \leq n$, и рассмотрим матрицу A, в которой каждый член определителя содержит нулевой элемент матрицы. Если все ее элементы равны нулю, то утверждение уже доказано. Пусть имеется хотя бы один ненулевой элемент. Пусть $a_{1n} \neq 0$. Тогда

$$A = \begin{bmatrix} a_{11} & \dots & a_{1\,n-1} & a_{1n} \\ & & & a_{2n} \\ & B & & \dots \\ & & & a_{nn} \end{bmatrix},$$

причем любой член определителя матрицы B обязан содержать нулевой множитель. По индуктивному предположению, в B имеется нулевая подматрица $0_{k\times l}$ размеров $k \times l$ с условием $k+l > n-1$. Если $k+l > n$, то эта подматрица является искомой.

Остается рассмотреть случай $k+l = n$. Не ограничивая общности, предположим, что A имеет вид

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0_{k\times l} & A_{22} \end{bmatrix}.$$

Подматрицы A_{11} и A_{22} квадратные — порядка l и k, соответственно. В силу исходного предполодения о матрице A, если хотя бы один член определителя A_{11} ненулевой, то все члены определителя A_{22}

равны нулю. По индуктивному предположению, в A_{22} имеется нулевая $r \times s$-подматрица с условием $r+s > k$. Не ограничивая общности, предполодим, что она находися на последних r строках и столбцах с номерами от $l+1$ до $l+s$. Рассмотрим подматрицу Z на пересечении последних $p = r$ строк и $q = l+s$ и столбцов. Легко видеть, что $Z = 0$, при этом $p + q = l + r + s > l + k = n$.

Если все члены определителя A_{11} равны нулю, то индуктивное предположение можно применить непосредственно к A_{11}. Искомая нулевая подматрица в A строится аналогичным образом. $\square$

Заметим, что теорема Холла появилась в 1935 году в связи с изучением комбинаторных задач (а именно, задачи о паросочетаниях).

Лекция 6

ОСНОВНАЯ ЧАСТЬ

6.1 Обратная матрица

Матрица A порядка n называется *обратимой*, если существует матрица X порядка n такая, что

$$AX = XA = I,$$

где I — единичная матрица порядка n; X называется *обратной матрицей* для A.

Может существовать только одна обратная матрица: если $AX = XA = I$ и $AY = YA = I$, то $X = X(AY) = (XA)Y = Y$. Обозначение для обратной матрицы: $X = A^{-1}$.

Задача. Найти все обратимые матрицы A порядка n, для которых все элементы A и A^{-1} неотрицательны. Доказать, что множество всех таких матриц образует группу относительно операции умножения матриц.

Задача. Пусть A, B — произвольные матрицы порядка n; I и 0 — единичная и нулевая матрицы порядка n. Доказать, что

$$\begin{bmatrix} I & A & 0 \\ 0 & I & B \\ 0 & 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -A & AB \\ 0 & I & -B \\ 0 & 0 & I \end{bmatrix}.$$

(Отсюда следует, что любой алгоритм вычисления обратной матрицы порядка n с числом операций $s(n)$ порождает алгоритм умножения матриц порядка n с числом операций $s(3n)$).

6.2 Критерий обратимости матрицы

Теорема. *Квадратная матрица обратима тогда и только тогда, когда ее столбцы образуют линейно независимую систему.*

Доказательство. Пусть матрица A порядка n имеет линейно независимые столбцы. Согласно результатам Лекции 3 о совместности систем линейных алгебраических уравнений, каждая из систем

$$Ax_1 = e_1, \;\; Ax_2 = e_2, \;\; \ldots, \;\; Ax_n = e_n,$$

где $e_1, e_2, \ldots, e_n$ — столбцы единичной матрицы, имеет единственное решение. Пусть $X = [x_1, x_2, \ldots, x_n]$. Тогда $AX = I$.

Столбцы матрицы X линейно независимы. В самом деле, пусть некоторая линейная комбинация этих столбцов равна нулю:

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0.$$

Это означает, что

$$X \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix} = 0.$$

Следовательно,

$$AX \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix} = 0.$$

Отсюда $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$. Таким образом, для X существует матрица Y такая, что $XY = I$. Докажем, что $Y = A$. В самом деле, $A = A(XY) = (AX)Y = Y$. [1]

Теперь предположим, что $A = [a_1, \ldots, a_n]$ — обратимая матрица, и рассмотрим равную нулю линейную комбинацию ее столбцов:

$$\alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_n a_n = 0.$$

Данное равенство запишем в следующем виде:

$$A \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix} = 0.$$

Умножая обе части слева на A^{-1}, находим $\alpha_1 = \ldots = \alpha_n = 0$. Следовательно, столбцы $a_1, \ldots a_n$ линейно независимы. $\square$

6.3 Обращение и транспонирование

Утверждение. $(AB)^\top = B^\top A^\top$.

Доказательство. $(AB)_{ij} = \sum\limits_k (A)_{ik}(B)_{kj} = \sum\limits_k (B)_{kj}(A)_{ik} = \sum\limits_k (B^\top)_{jk}(A^\top)_{ki} = (B^\top A^\top)_{ji}$. $\square$

Из равенства $AX = XA = I$ получаем $X^\top A^\top = A^\top X^\top = I$. Таким образом, *матрица обратима тогда и только тогда, когда обратима ее транпонированная матрица.* При этом

$$X^\top = (A^\top)^{-1} = (A^{-1})^\top.$$

Обозначение: $A^{-\top} \equiv X^\top$.

Как следствие, получаем "строчный" аналог критерия обратимости матрицы: *обратимость матрицы равносильна линейной независимости ее строк.*

[1] По существу, здесь воспроизводится часть доказательства более общего утверждения, связанного с избыточностью рассмотренного нами определения группы (см. раздел из дополнительной части Лекции 2).

6.4 Группа обратимых матриц

Множество обратимых $n \times n$-матриц относительно операции умножения образует группу. Для доказательства есть все, кроме факта обратимости произведения обратимых матриц. Но это проверяется непосредственно: если A и B обратимы, то

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}\,I\,B = I,$$

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A^{-1}\,I\,A = I.$$

Отсюда

$$(AB)^{-1} = B^{-1}\,A^{-1}.$$

Задача. Является ли группа невырожденных верхних треугольных матриц нормальным делителем группы всех невырожденных матриц данного порядка?

6.5 Обращение невырожденной матрицы

Квадратная матрица с отличным от нуля определителем называется *невырожденной.*

Теорема. *Если матрица A порядка n невырожденная, то она обратима и при этом*

$$A^{-1} \;=\; \frac{1}{\det A}\,\widetilde{A}^{\top},$$

где $\widetilde{A} = [A_{ij}]$ — матрица порядка n, в которой элемент A_{ij} есть алгебраическое дополнение к элементу a_{ij} в матрице A.

Доказательство. Заметим, что

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \begin{cases} \det A, & k = i, \\ 0, & k \neq i. \end{cases} \qquad (*)$$

При $k = i$ равенство $(*)$ получается применением теоремы Лапласа при разложении определителя матрицы A по k-й строке. При $k \neq i$ левая часть равенства $(*)$ представляет собой разложение по k-й строке определителя матрицы, полученной из A заменой k-й строки на i-ю. Такой определитель равен нулю — как определитель матрицы с двумя одинаковыми строками. Далее,

$$\sum_{j=1}^{n} a_{ij}A_{kj} = \sum_{j=1}^{n} a_{ij}(\widetilde{A}^{\top})_{jk}, \qquad 1 \leq i, k \leq n.$$

Поэтому, в силу соотношений $(*)$,

$$A\widetilde{A}^{\top} \;=\; \begin{bmatrix} \det A & & \\ & \ddots & \\ & & \det A \end{bmatrix} \;=\; \det A \;\cdot\; I.$$

Используя теорему Лапласа для разложения определителя по k-му столбцу, находим

$$\sum_{i=1}^{n} a_{ij}A_{ik} = \begin{cases} \det A, & k = j, \\ 0, & k \neq j. \end{cases} \qquad \Rightarrow$$

$$\widetilde{A}^{\top} A \;=\; \begin{bmatrix} \det A & & \\ & \ddots & \\ & & \det A \end{bmatrix} \;=\; \det A \;\cdot\; I. \qquad \square$$

Задача. Докажите, что любую невырожденную матрицу можно сделать вырожденной, изменив лишь один из ее элементов.

6.6 Правило Крамера

Теорема. *Пусть A — невырожденная матрица порядка n. Тогда система линейных алгебраических уравнений $Ax = b$ имеет и притом единственное решение x с компонентами*

$$x_i \;=\; \frac{\det A_i}{\det A}, \qquad 1 \le i \le n,$$

где A_i — матрица, получаемая из A заменой i-го столбца на b.

Доказательство. Согласно теореме об обращении невырожденной матрицы,

$$x \;=\; A^{-1}b \;=\; \frac{1}{\det A}\,\widetilde{A}^{\top} b \qquad \Rightarrow \qquad x_i \;=\; \frac{1}{\det A}\,\sum_{j=1}^{n} A_{ji} b_j, \quad 1 \le i \le n.$$

Остается заметить, что сумма $\sum\limits_{j=1}^{n} A_{ji} b_j$ есть разложение по i-му столбцу определителя матрицы A_i. $\square$

6.7 Определитель произведения матриц

Теорема. *Определитель произведения квадратных матриц равен произведению их определителей.*

Доказательство. Пусть $A = [a_1, \ldots, a_n]$ — матрица порядка n со столбцами $a_1, \ldots, a_n$ и $B = [b_{ij}]$. Тогда любой столбец матрицы AB есть линейная комбинация столбцов матрицы A с коэффициентами из соответствующего столбца матрицы B:

$$AB \;=\; \left[\sum_{i_1=1}^{n} b_{i_1\,1}\; a_{i_1},\; \ldots,\; \sum_{i_n=1}^{n} b_{i_n\,n}\; a_{i_n}\right].$$

Используя линейность определителя по каждому столбцу, получаем

$$\begin{aligned} \det(AB) \;&=\; \sum_{i_1=1}^{n} \ldots \sum_{i_n=1}^{n} b_{i_1 1} \ldots b_{i_n n}\; \det[a_{i_1},\, \ldots,\, a_{i_n}] \\ &=\; \sum_{\sigma \in S_n} b_{\sigma(1)1} \ldots b_{\sigma(n)n}\; \det[a_{\sigma(1)},\, \ldots,\, a_{\sigma(n)}] \\ &=\; \left(\sum_{\sigma \in S_n} b_{\sigma(1)1} \ldots b_{\sigma(n)n}\, \operatorname{sgn}(\sigma)\right) \det A \;=\; \det B \;\cdot\; \det A. \qquad \square \end{aligned}$$

6.8 Обратимость и невырожденность

Теорема. *Квадратная матрица обратима тогда и только тогда, когда она невырожденная.*

Доказательство. Пусть для матрицы A существует обратная матрица A^{-1}. Тогда $AA^{-1} = I$ и, в силу теоремы об определителе произведения матриц,

$$\det A \cdot \det A^{-1} = \det I = 1 \qquad \Rightarrow \qquad \det A \neq 0.$$

Если $\det A \neq 0$, то A обратима по теореме об обращении невырожденной матрицы. $\square$

Следствие. *Столбцы матрицы A линейно независимы тогда и только тогда, когда* $\det A \neq 0$.

Задача. Пусть I_n и I_m — единичные матрицы порядка n и m. Докажите, что для любых матриц A размеров $m \times n$ и B размеров $n \times m$ из обратимости $I_m - AB$ вытекает обратимость $I_n - BA$.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

6.9 Матрицы с диагональным преобладанием

Отметим полезное достаточное условие обратимости матрицы. Пусть для элементов матрицы $A = [a_{ij}]$ порядка n выполняются соотношения

$$|a_{ii}| \; > \sum_{\substack{1 \le j \le n \\ j \neq i}} |a_{ij}|, \qquad i = 1, 2 \ldots, n.$$

В таких случаях A называется матрицей с *диагональным преобладанием по строкам*. Если имеют место соотношения

$$|a_{jj}| \; > \sum_{\substack{1 \le i \le n \\ i \neq j}} |a_{ij}|, \qquad j = 1, 2 \ldots, n,$$

то A называется матрицей с *диагональным преобладанием по столбцам*.

Теорема. *Любая матрица с диагональным преобладанием по строкам или по столбцам является обратимой.*

Доказательство. Пусть A — матрица с диагональным преобладанием по строкам. Докажем, что ее столбцы линейно независимы. Для этого приравняем нулю их линейную комбинацию с коэффициентами $x_1, \ldots, x_n$:

$$A \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix} \; = \; 0.$$

Выберем строку с номером i таким, что $|x_i| \ge |x_j|$ для всех j. Тогда

$$0 \; = \; \left| a_{ii} x_i + \sum_{\substack{1 \le j \le n \\ j \neq i}} a_{ij} x_j \right| \; \ge \; \left(|a_{ii}| - \sum_{\substack{1 \le j \le n \\ j \neq i}} |a_{ij}| \right) \; |x_i|.$$

Поскольку величина в скобках положительная, получаем $x_i = 0 \;\; \Rightarrow \;\; x_j = 0 \;\; \forall\, j$. Обратимость матрицы с диагональным преобладанием по столбцам доказывается с помощью перехода к транспонированной матрице. $\square$

6.10 Определитель и возмущения

Можно доказать, что если определитель матрицы отличен от нуля, то при всех *достаточно малых* изменениях (в математике часто говорят — *возмущениях*) элементов матрицы определитель не станет нулем.

Задача. Докажите, что $\det(I+F) \neq 0$, если каждый элемент матрицы-возмущения F порядка n по модулю меньше $1/n$.

Однако, по величине определителя трудно судить, насколько малы должны быть соответствующие возмущения. Например, рассмотрим двухдиагональные матрицы порядка n с возмущением ε только одного элемента — в левом нижнем углу:

$$A(\varepsilon) = \begin{bmatrix} 1 & 2 & & & 0 \\ & 1 & 2 & & \\ & & \ddots & \ddots & \\ & 0 & & 1 & 2 \\ \varepsilon & & & & 1 \end{bmatrix}.$$

При $\varepsilon = 0$ имеем $\det A(0) = 1$. В общем случае, применяя теорему Лапласа для разложения определителя по первому столбцу, находим

$$\det A(\varepsilon) \;=\; 1 + \varepsilon \cdot (-1)^{n+1}\, 2^{n-1}.$$

При $\varepsilon = (-1)^n/2^{n-1}$ получаем $\det A(\varepsilon) = 0$. Пусть, например, $n = 100$. Как видим, невырожденная матрица с определителем 1 превращается в вырожденную при весьма малом возмущении!

Лекция 7

ОСНОВНАЯ ЧАСТЬ

7.1 Разделение переменных и матрицы

При изучении функций от двух переменных особую роль играют функции с *разделенными переменными* $f(x,y) = u(x)v(y)$ или суммы таких функций [1]

$$f(x,y) = u_1(x)v_1(x) + \ldots + u_r(x)v_r(y).$$

Пусть дана $m \times n$-матрица A. Ее элемент a_{ij} можно рассматривать как функцию от дискретных переменных $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$. В данном случае разделение переменных означает, что

$$a_{ij} = u_i v_j, \qquad 1 \le i \le m, \;\; 1 \le j \le n.$$

Отсюда легко вывести, что A есть произведение столбца и строки: [2]

$$A = uv^\top, \qquad u = \begin{bmatrix} u_1 \\ \cdots \\ u_m \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ \cdots \\ v_n \end{bmatrix}.$$

7.2 Скелетное разложение

Теперь предположим, что $a_{ij} = u_{i1}v_{j1} + \ldots + u_{ir}v_{jr}$, $1 \le i \le m$, $1 \le j \le n$. В этом случае A является суммой r матриц вида "столбец на строку"

$$A = \sum_{k=1}^{r} u_k v_k^\top, \qquad u_k = \begin{bmatrix} u_{1k} \\ \cdots \\ u_{mk} \end{bmatrix}, \quad v_k = \begin{bmatrix} v_{1k} \\ \cdots \\ v_{nk} \end{bmatrix}.$$

Это же равенство, записанное в виде произведения двух матриц

$$A = UV^\top = \begin{bmatrix} u_{11} & \ldots & u_{1r} \\ \cdots & \cdots & \cdots \\ u_{m1} & \ldots & u_{mr} \end{bmatrix} \begin{bmatrix} v_{11} & \ldots & v_{n1} \\ \cdots & \cdots & \cdots \\ v_{1r} & \ldots & u_{nr} \end{bmatrix}, \qquad U = [u_1, \ldots, u_r], \quad V = [v_1, \ldots, v_r], \qquad (*)$$

называется *скелетным разложением* матрицы A. Оно означает, что каждый столбец матрицы A есть линейная комбинация столбцов матрицы U, а каждая строка матрицы A есть линейная комбинация строк матрицы $V^\top$. Отсюда сразу же вытекает

Теорема. *Размерность линейной оболочки, натянутой на столбцы матрицы A, совпадает с размерностью линейной оболочки, натянутой на ее строки:*

$$\dim L(a_1, \ldots, a_n) = \dim L(\widehat{a}_1, \ldots, \widehat{a}_m), \qquad A = [a_1, \ldots, a_n] = \begin{bmatrix} \widehat{a}_1^\top \\ \cdots \\ \widehat{a}_m^\top \end{bmatrix}.$$

[1] Разделение переменных с большим успехом применяется для приближения функций общего вида.
[2] Матрица $uv^\top$ иногда называется *внешним произведением* векторов $u \in \mathbb{C}^m$ и $v \in \mathbb{C}^n$.

Доказательство. Пусть столбцы матрицы U образуют базис линейной оболочки $L(a_1, \ldots, a_n)$, а j-й столбец матрицы $V^\top$ состоит из коэффициентов их линейной комбинации, дающей столбец a_j. Тогда, очевидно, $A = UV^\top$. В силу предваряющего теорему замечания, размерность линейной оболочки строк матрицы A не выше числа строк матрицы $V^\top$, которое равно, по построению, размерности линейной оболочки столбцов матрицы A. Противоположное неравенство доказывается аналогично — роль столбцов и строк меняется транспонированием. $\square$

7.3 Ранг матрицы

Размерность линейной оболочки столбцов (строк) матрицы иногда называется ее столбцовым (строчным) рангом. Поскольку столбцовый и строчный ранги совпадают, их общее значение было бы естественно называть просто рангом матрицы.

Однако, обычно дается другое определение: *рангом* матрицы называется наивысший порядок ее отличных от нуля миноров. Соответствующие минор и подматрица называются *базисным минором* и *базисной подматрицей*. В силу уже установленной эквивалентности обратимости и невырожденности, ранг матрицы равен наивысшему порядку обратимых подматриц в данной матрице. Обозначение: rankA.

Два очевидных свойства ранга матрицы A размеров $m \times n$:

$$\mathrm{rank}A \le \min(m, n), \qquad \mathrm{rank}A \;=\; \mathrm{rank}A^\top.$$

Менее очевидно, что *наивысший порядок отличных от нуля миноров матрицы совпадает с ее столбцовым и строчным рангом.* Давайте это докажем.

7.4 Окаймление обратимой подматрицы

Начнем с полезного вспомогательного предложения. Пусть матрица Q порядка $k+1$ имеет блочный вид

$$Q = \begin{bmatrix} P & v \\ u^\top & c \end{bmatrix}, \qquad P \in \mathbb{R}^{k\times k}, \quad u, v \in \mathbb{R}^{k\times 1}.$$

В этом случае Q называется *окаймлением подматрицы* P.

Лемма о необратимом окаймлении. *Если подматрица P обратима, а ее окаймление Q является необратимой матрицей, то последний столбец матрицы Q есть линейная комбинация первых k столбцов.*

Доказательство. Используя правило умножения блочных матриц (см. Лекцию 1), легко проверить справедливость равенства

$$\begin{bmatrix} I & 0 \\ -u^\top P^{-1} & 1 \end{bmatrix} \begin{bmatrix} P & v \\ u^\top & c \end{bmatrix} \;=\; \begin{bmatrix} P & v \\ 0 & \gamma \end{bmatrix}, \qquad \gamma = c - u^\top P^{-1} v.$$

Обозначим матрицу в правой части через M. Как произведение обратимой и необратимой матриц, M не может быть обратимой матрицей. Но она имеет блочно-треугольный вид, и если бы блоки P и γ были оба обратимы, то M имела бы обратную матрицу вида

$$M^{-1} \;=\; \begin{bmatrix} P^{-1} & -P^{-1} v \gamma^{-1} \\ 0 & \gamma^{-1} \end{bmatrix}.$$

(Равенство $MM^{-1} = I$ проверяется непосредственно.) Поскольку M не является обратимой матрицей, непременно

$$\gamma = c - u^\top P^{-1} v = 0 \qquad \Rightarrow \qquad c = u^\top P^{-1} v.$$

Следовательно,

$$\begin{bmatrix} P \\ u^\top \end{bmatrix} (P^{-1} v) \;=\; \begin{bmatrix} v \\ c \end{bmatrix}. \qquad \square$$

7.5 Теорема о базисном миноре

Теорема. *Столбцы (строки), содержащие базисный минор, являются линейно независимыми, при этом любой столбец (любая строка) данной матрицы является их линейной комбинацией.*

Доказательство. Не ограничивая общности, предположим, что базисная подматрица P порядка k расположена в левом верхнем углу матрицы A размеров $m \times n$. Таким образом,

$$A = \begin{bmatrix} P & v_{k+1} & \dots & v_n \\ u_{k+1}^\top & a_{k+1\,k+1} & \dots & a_{k+1\,n} \\ \dots & \dots & \dots & \dots \\ u_m^\top & a_{m\,k+1} & \dots & a_{mn} \end{bmatrix}, \qquad u_{k+1}, \dots, u_m,\ v_{k+1}, \dots, v_n \in \mathbb{R}^{k \times 1}.$$

По условию теоремы, любая подматрица порядка $k+1$ вида

$$M \;=\; \begin{bmatrix} P & v_j \\ u_i^\top & a_{ij} \end{bmatrix}, \qquad i, j > k,$$

является необратимой. По лемме о необратимом окаймлении обратимой подматрицы, последний столбец в M есть линейная комбинация первых k столбцов. При этом коэффициенты данной линейной комбинации не зависят от i (поскольку определяются вектором $P^{-1} v_j$). Значит, j-й столбец матрицы A при $j > k$ есть линейная комбинация первых k столбцов. Линейная независимость первых k столбцов доказывается следующим образом: пусть их линейная комбинация с коэффициентами $\alpha_1, \dots, \alpha_k$ равна 0, тогда

$$P \begin{bmatrix} \alpha_1 \\ \dots \\ \alpha_k \end{bmatrix} = 0 \qquad \Rightarrow \qquad \begin{bmatrix} \alpha_1 \\ \dots \\ \alpha_k \end{bmatrix} = 0.$$

Утверждение теоремы относительно строк доказывается переходом к транспонированной матрице. $\square$

Следствие. *Ранг матрицы совпадает с ее строчным и столбцовым рангом.*

Замечание. Теорема о базисном миноре не использует доказанной ранее теоремы о равенстве столбцового и строчного рангов. По существу, она дает еще одно доказательство этой теоремы.

Задача. Пусть A — $n \times n$-матрица ранга k, а B — любая невырожденная подматрица порядка k. Обозначим через R подматрицу размеров $k \times n$, состоящую из строк матрицы A, содержащих подматрицу B, а через C — подматрицу размеров $n \times k$, состоящую из столбцов, содержащих B. Доказать, что

$$A = CB^{-1}R.$$

7.6 Ранги и матричные операции

Утверждение 1. *Ранг суммы матриц не превосходит суммы их рангов:*

$$\operatorname{rank}(A+B) \;\le\; \operatorname{rank} A \;+\; \operatorname{rank} B.$$

Доказательство. Очевидно, A и B должны иметь одинаковое число столбцов:

$$A = [a_1, \ldots, a_n], \qquad B = [b_1, \ldots, b_n].$$

Ясно, что $L(a_1 + b_1, \ldots, a_n + b_n) \;\subset\; L(a_1, \ldots, a_n,\; b_1, \ldots, b_n)$. В меньшей линейной оболочке выберем какую-либо систему векторов, образующую базис. Согласно лемме о дополнении до базиса, базис в большей линейной оболочке можно получить путем дополнения данной системы какими-то векторами из большей линейной оболочки. Поэтому

$$\operatorname{rank}(A+B) = \dim L(a_1 + b_1, \ldots, a_n + b_n) \;\le\; \dim L(a_1, \ldots, a_n,\; b_1, \ldots, b_n).$$

Пусть $p = \operatorname{rank} A$, $q = \operatorname{rank} B$, и предположим, не ограничивая общности, что базис в $L(a_1, \ldots, a_n)$ образуют первые p векторов, а базис в $L(b_1, \ldots, b_n)$ — первые q векторов. Тогда $L(a_1, \ldots, a_n,\; b_1, \ldots, b_n) \;=\; L(a_1, \ldots, a_p,\; b_1, \ldots, b_q) \quad \Rightarrow$

$$\dim L(a_1, \ldots, a_n,\; b_1, \ldots, b_n) \;\le\; p + q. \qquad \square$$

Утверждение 2. *Ранг произведения матриц не превосходит ранга каждого из сомножителей:*

$$\operatorname{rank}(AB) \;\le\; \min(\operatorname{rank} A,\; \operatorname{rank} B).$$

Доказательство. Достаточно заметить, что каждый из столбцов матрицы AB является линейной комбинацией столбцов матрицы A. Поэтому линейная оболочка столбцов матрицы AB содержится в линейной оболочке столбцов матрицы A. Следовательно, $\operatorname{rank}(AB) \le \operatorname{rank} A$. Далее,

$$\operatorname{rank}(AB) = \operatorname{rank}(AB)^\top = \operatorname{rank}(B^\top A^\top) \;\le\; \operatorname{rank} B^\top \;=\; \operatorname{rank} B. \qquad \square$$

Утверждение 3. *Ранг матрицы не изменяется при умножении ее слева или справа на обратимую матрицу.*

Доказательство. Пусть $B = PAQ$, где P и Q — обратимые матрицы. [3] В силу предыдущего утверждения, $\operatorname{rank} B \le \operatorname{rank} A$. В то же время, $A = P^{-1}BQ^{-1} \quad \Rightarrow$ $\operatorname{rank} A \le \operatorname{rank} B$. $\square$

Задача 1. Пусть A и B — матрицы ранга 1. Докажите, что если $AB = BA$, то ранг матрицы $A + B$ не больше 1.

Задача 2. Матрица A имеет r столбцов, а матрица B имеет r строк. Докажите, что

$$r \;\ge\; \operatorname{rank}(A) + \operatorname{rank}(B) - \operatorname{rank}(AB).$$

[3] Конечно, порядок P равен числу строк, а порядок Q — числу столбцов матрицы A.

7.7 Однородная система линейных алгебраических уравнений

Система линейных алгебраических уранений с нулевой правой частью

$$Ax = 0 \qquad (*)$$

называется *однородной*. Пусть в данной системе имеется m уравнений и n неизвестных. Тогда матрица коэффициентов A имеет размеры $m \times n$. Рассмотрим A как систему столбцов $A = [a_1, \ldots, a_n]$ и предположим, что ее ранг равен r. Не ограничивая общности, будем считать, что базисная подматрица в A расположена на первых r столбцах — будем называть их *базисными*. Отвечающие базисным столбцам компоненты решения $x_1, \ldots, x_r$ будем также называть *базисными*, а оставшиеся компоненты $x_{r+1}, \ldots, x_n$ — *свободными*. Таким образом, вектор-решение имеет вид

$$x = \begin{bmatrix} x_1 \\ \cdots \\ x_r \\ \\ x_{r+1} \\ \cdots \\ x_n \end{bmatrix}.$$

Система $(*)$ равносильна равенству

$$x_1 a_1 + \ldots + x_r a_r = -x_{r+1} a_{r+1} - \ldots - x_n a_n. \qquad (**)$$

По теореме о базисной подматрице, столбцы $a_{r+1}, \ldots, a_n$ принадлежат линейной оболочке столбцов $a_1, \ldots, a_r$. Поэтому при любом выборе значений свободных неизвестных значения базисных неизвестных, удовлетворяющих равенству $(**)$, существуют и определяются однозначно. Таким образом, существуют $n - r$ векторов вида

$$v_1 = \begin{bmatrix} x_{11} \\ \cdots \\ x_{r1} \\ \\ 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} x_{12} \\ \cdots \\ x_{r2} \\ \\ 0 \\ 1 \\ \cdots \\ 0 \end{bmatrix}, \quad \ldots \quad v_{n-r} = \begin{bmatrix} x_{1\,n-r} \\ \cdots \\ x_{r\,n-r} \\ \\ 0 \\ 0 \\ \cdots \\ 1 \end{bmatrix},$$

каждый из который является решением системы $(*)$:

$$Av_1 = 0, \ \ldots, \ Av_{n-r} = 0.$$

Векторы $v_1, \ldots, v_{n-r}$ линейно независимы:

$$\alpha_1 v_1 + \ldots + \alpha_{n-r} v_{n-r} = \begin{bmatrix} * \\ \cdots \\ * \\ \\ \alpha_1 \\ \alpha_2 \\ \cdots \\ \alpha_{n-r} \end{bmatrix} = 0 \qquad \Rightarrow \qquad \alpha_1 = \ldots = \alpha_{n-r} = 0.$$

Кроме того, любая линейная комбинация этих векторов является решением системы $(*)$ и, более того, если x есть произвольное решение системы $(*)$, то

$$x = x_{r+1} v_1 + \ldots + x_n v_{n-r}.$$

Таким образом, мы доказали следующее важное утверждение.

Теорема. *Множество решений однородной системы линейных алгебраических уравнений* $(*)$ *совпадает с линейной оболочкой* $L(v_1, \ldots, v_{n-r})$ *линейно независимых векторов* $v_1, \ldots, v_{n-r}$.

Следствие. $\dim L(v_1, \ldots, v_{n-r}) = n - r$.

Линейно независимую систему решений $w_1, \ldots, w_k$ системы $Ax = 0$ называют *фундаментальной системой*, если ее линейная оболочка $L(w_1, \ldots, w_k)$ совпадает с множеством всех решений однородной системы $Ax = 0$.

Следствие. *Число векторов в любой фундаментальной системе решений для* $Ax = 0$ *равно* $n - r$, *где* $r = \operatorname{rank} A$.

Для доказательства достаточно заметить, что линейная оболочка векторов фундаментальной системы решений имеет базис из построенных выше векторов $v_1, \ldots, v_{n-r}$.

7.8 Теорема Кронекера–Капелли

Рассмотрим систему линейных алгебраических уравнений $Ax = b$ с $m \times n$-матрицей $A = [a_1, \ldots, a_n]$. Матрица $[A, b] = [a_1, \ldots, a_n, b]$ называется *расширенной матрицей* данной системы.

Теорема. *Система* $Ax = b$ *совместна тогда и только тогда, когда ранг матрицы коэффициентов совпадает с рангом расширенной матрицы:*

$$\operatorname{rank} A \;=\; \operatorname{rank}[A, b].$$

Доказательство. Мы уже знаем (см. Лекцию 3), что совместность системы $Ax = b$ равносильна равенству линейных оболочек $L(a_1, \ldots, a_n) = L(a_1, \ldots, a_n, b)$. Остается заметить, что $\operatorname{rank} A = \dim L(a_1, \ldots, a_n)$ и $\operatorname{rank}[A, b] = \dim L(a_1, \ldots, a_n, b)$. $\square$

7.9 Общее решение системы линейных алгебраических уравнений

Если U и V — два множества векторов из $\mathbb{R}^n$, то суммой $U + V$ называется множество, составленное из всевозможных сумм векторов вида $u + v$, где $u \in U$, $v \in V$.

Теорема. *Предположим, что система* $Ax = b$ *совместна, и зафиксируем произвольное частное решение* u $(Au = 0)$. *Тогда множество всех решений системы* $Ax = b$ *имеет вид* $u + V$, *где* V — *множество всех решений соответствующей однородной системы* $Ax = 0$.

Доказательство. Пусть x — произвольное решение системы $Ax = b$. Тогда, очевидно, $A(x - u) = 0 \;\Rightarrow\; x - u \in V \;\Rightarrow\; x \in u + V$. Далее, возьмем произвольный вектор $x \in u + V \;\Rightarrow\; x = u + v,\ v \in V \;\Rightarrow\; A(u + v) = Au + Av = b + 0 = b$. $\square$

Следствие. *Общее решение совместной системы* $Ax = b$ *имеет вид*

$$x \;=\; u \,+\, c_1 v_1 \,+\, \ldots \,+\, c_{n-r} v_{n-r},$$

где u — произвольное частное решение данной системы, $v_1, \ldots, v_{n-r}$ — линейно независимые решения соответствующей однородной системы, $r = \operatorname{rank} A$, а коэффициенты $c_1, \ldots, c_{n-r}$ — произвольные числа.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

7.10 Неустойчивость ранга

Матрица называется матрицей *полного ранга*, если ее ранг совпадает с одним из ее размеров (то есть, имеет максимально возможное значение). В противном случае говорят о матрице *неполного ранга*.

Можно доказать, что если A есть $m \times n$-матрица полного ранга, то при всех достаточно малых $\varepsilon > 0$ матрица $A+F$, где все элементы матрицы-возмущения F по модулю меньше ε, будет также матрицей полного ранга.

В то же время, если A имеет неполный ранг, то для любого сколь угодно малого ε существует матрица-возмущение F с элементами по модулю не больше ε, для которой $A + F$ будет матрицей полного ранга. Например, матрица

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

имеет ранг 2, но для любого $\varepsilon \neq 0$ матрица

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \varepsilon & 0 \end{bmatrix}$$

имеет, очевидно, ранг 3.

Лекция 8

ОСНОВНАЯ ЧАСТЬ

8.1 Исключение неизвестных

Если задана система линейных алгебраических уравнений $Ax = b$ и требуется найти ее общее решение или установить несовместность, то это удобнее всего сделать путем последовательного *исключения неизвестных*: если в каком-то уравнении коэффициент при x_1 отличен от нуля, то можно исключить x_1 из всех других уравнений путем вычитания данного уравнения, предварительно умноженного на подходящим образом выбранные коэффициенты; если среди уравнений, уже не содержащих x_1, имеется уравнение с ненулевым коэффициентом при x_2, то x_2 можно аналогичным образом исключить из всех других уравнений, кроме данного и первого уравнения, содержащего x_1, и так далее.

На каждом шаге исключения получается новая система, которая, очевидно, равносильна исходной. Если возникло уравнение, в котором все коэффициенты при неизвестных равны нулю, а в правой части получилось отличное от нуля число, то система не имеет решений. В противном случае система совместна и описанный способ позволяет с легкостью выписать ее общее решение.

8.2 Элементарные матрицы

Каждый шаг описанного выше исключения неизвестных преобразует систему $Ax = b$ в равносильную систему вида $(PA)x = Pb$, где P — некоторая обратимая матрица. Если потребовалось k шагов, то в итоге возникает последовательность равносильных систем

$$Ax = b, \quad (P_1A)x = P_1b, \quad (P_2P_1A)x = P_2P_1b, \quad \dots \ , \quad (P_k \dots P_2P_1A)x = P_k \dots P_2P_1b.$$

Матрица коэффициентов последней системы имеет настолько простой вид, что решение соответствующей системы осуществляется уже очевидным образом.

Каждая из матриц P_l, $1 \leq l \leq k$, может быть представлена как произведение двух матриц:

$$P_l = Z_l\Pi_l,$$

где Z_l отличается от I (единичной матрицы) только в позициях ниже главной диагонали

какого-то одного (пусть j-го) столбца:

$$Z_l = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & (Z_l)_{j+1\,j} & & \\ & & \dots & \ddots & \\ & & (Z_l)_{nj} & & 1 \end{bmatrix},$$

а Π_l получается из I перестановкой строк. Матрицы Π_l и Z_l указанного специального вида будем для краткости называть *элементарными матрицами*; матрица Π_l называется *матрицей перестановки*, а Z_l — *матрицей модификации строк*. Их роль в процессе исключения объясняется следующими фактами:

- матрица $\Pi_l A$ отличается от A перестановкой строк;
- если Z_l отличается от единичной матрицы j-м столбцом, то матрица $Z_l A$ имеет первые j строк те же, что и в матрице A, а i-я строка при $i > j$ есть сумма i-й строки и взятой с некоторым коэффициентом j-й строки матрицы A.

Утверждение 1. *Любая матрица перестановки* Π *обратима и при этом*

$$\Pi^{-1} = \Pi^\top.$$

Утверждение 2. *Любая матрица модификации строк* $Z = Z_l$ *обратима и при этом обратная матрица получается из* Z *изменением знаков поддиагональных элементов:*

$$Z^{-1} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & -(Z_l)_{j+1\,j} & & \\ & & \dots & \ddots & \\ & & -(Z_l)_{nj} & & 1 \end{bmatrix},$$

Доказательство сводится к непосредственной проверке равенств $\Pi\Pi^{-1} = I,\ \ ZZ^{-1} = I$.

8.3 Ступенчатые матрицы

Будем говорить, что матрица $S = [s_{ij}]$ размеров $m \times n$ является *верхней ступенчатой* с числом ступеней k, если существуют номера $1 \le j_1 < \ldots < j_k \le m$, для которых:

- если $1 \le i \le k$, то $s_{ij} \ne 0$ при $j = j_i$ и $s_{ij} = 0$ при всех $1 \le j \le j_i - 1$;
- если $k + 1 \le i \le m$, то $s_{ij} = 0$ при всех $1 \le j \le n$.

Матрица S называется *нижней ступенчатой* с числом ступеней k, если $S^\top$ является верхней ступенчатой с числом ступеней k.

Утверждение. *Ранг ступенчатой матрицы с числом ступеней* k *равен* k.

Доказательство. Рассмотрим верхнюю ступенчатую матрицу S и докажем, что ее строчный ранг (размерность линейной оболочки строк) равен числу ступеней k. Ясно,

что S имеет ровно k ненулевых строк. Докажем, что эти строки линейно независимы. Приравняем нулю их линейную комбинацию с коэффициентами $\alpha_1, \ldots, \alpha_k$:

$$[\alpha_1, \ldots, \alpha_k]S = [0, \ldots, 0].$$

Отсюда

$$\alpha_1 \, s_{1i_1} = 0 \qquad \Rightarrow \qquad \alpha_1 = 0.$$

Далее,

$$0 \cdot s_{1\,i_2} \; + \; \alpha_2 s_{2\,i_2} = 0 \qquad \Rightarrow \qquad \alpha_2 = 0.$$

Продолжая подобным образом, находим $\alpha_1 = \ldots = \alpha_k = 0$. В случае нижней ступенчатой матрицы S ее столбцовый ранг, очевидно, совпадает со строчным рангом верхней ступенчатой матрицы $S^\top$. $\square$

8.4 Приведение к ступенчатой форме

Теорема 1. *Для любой $m \times n$-матрицы A ранга r существует обратимая матрица P, представимая в виде произведения конечного числа элементарных матриц и такая, что матрица $S = PA$ является верхней ступенчатой с числом ступеней r.*

Доказательство. Обозначим через j_1 номер первого столбца матрицы A, в котором есть хотя бы один ненулевой элемент. (Если таковой столбец отсутствует, то $A = 0$ и теорема уже доказана.) С помощью умножения слева на некоторую матрицу перестановки Π_1 ненулевой элемент можно переместить в позицию $(1, j_1)$. Далее с помощью умножения слева на некоторую матрицу модификации строк Z_1 можно получить матрицу с нулями в поддиагональных позициях j_1-го столбца и сохранением нулей в предыдущих столбцах. Очевидно, преобразованная матрица имеет блочный вид (через $0_{p\times q}$ мы обозначаем нулевой блок размеров $p \times q$)

$$Z_1\Pi_1 A \;=\; \begin{bmatrix} 0_{1\times(i_1-1)} & u^\top \\ 0_{(m-1)\times(i_1-1)} & B \end{bmatrix}, \qquad u \in \mathbb{R}^{(n-i_1+1)\times 1}.$$

Сделаем индуктивное предположение о существовании матрицы Q, являющейся произведением элементарных матриц порядка $m-1$ и такой, что QB имеет верхнюю ступенчатую форму. Рассмотрим матрицу

$$P = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} Z_1\Pi_1.$$

Легко видеть, что P есть произведение элементарных матриц и при этом $S = PA$ имеет верхнюю ступенчатую форму. Пусть число ступеней равно k. Значит, строчный ранг матрицы S равен $k \quad \Rightarrow \quad k = \operatorname{rank} S = \operatorname{rank} A = r$. $\square$

Теорема 2. *Для любой $m \times n$-матрицы A ранга r существует обратимая матрица Q, представимая в виде произведения конечного числа элементарных матриц и такая, что матрица $AQ^\top$ является нижней ступенчатой с числом ступеней r.*

Доказательство. Достаточно применить теорему 1 к матрице $A^\top$ и заметить, что если матрица $QA^\top$ — верхняя ступенчатая, то матрица $(QA^\top) = AQ^\top$ будет нижней ступенчатой.

8.5 Приведение к диагональной форме

Теорема. *Для любой $m \times n$-матрицы A ранга r существуют обратимые матрицы P и Q, представимые в виде произведения конечного числа элементарных матриц и такие, что матрица $B = PAQ^\top$ имеет ненулевые элементы $b_{11}, \ldots, b_{rr}$, а все остальные ее элементы равны нулю.*

Доказательство. Сначала приведем A к верхней ступенчатой форме $S = PA$, а затем заметим, что нижняя ступенчатая форма $SQ^\top$, получаемая согласно построениям теоремы 2, будет иметь требуемую диагональную форму. □

8.6 Эквивалентные матрицы

Матрицы A и B называются *эквивалентными*, если существуют обратимые матрицы P и Q такие, что $B = PAQ$.

Теорема. *Матрицы A и B эквивалентны тогда и только тогда, когда они имеют одинаковые размеры и одинаковые ранги.*

Доказательство. В силу теоремы о приведении к диагональной форме, каждая из матриц A и B эквивалентна прямоугольной диагональной матрице — обозначим их через D_A и D_B. При этом очевидно, что D_A и D_B эквивалентны тогда и только тогда, когда они имеют одинаковое число ненулевых диагональных элементов. Последнее означает, что $\operatorname{rank} D_A = \operatorname{rank} D_B$. Остается учесть, что $\operatorname{rank} A = \operatorname{rank} D_A$ и $\operatorname{rank} B = \operatorname{rank} D_B$. □

8.7 Метод Гаусса и LU-разложение

Рассмотренный выше метод исключения неизвестных обычно называют *методом Гаусса*. Пусть он применяется к системе $Ax = b$ с невырожденной матрицей A. В данном случае верхняя ступенчатая матрица, к которой приводится матрица A, оказывается верхней треугольной матрицей.

Метод исключения неизвестных можно трактовать как метод исключения элементов матрицы с целью приведения ее к более простому виду. Если можно обойтись без перестановки уравнений (строк матрицы), то метод Гаусса для матрицы порядка n состоит в последовательном исключении элементов в столбцах от 1-го до $n-1$-го и приводит к равносильной системе

$$(Z_{n-1} \ldots Z_2 Z_1 A)x = Z_{n-1} \ldots Z_2 Z_1 b, \qquad (*)$$

где Z_1, …, Z_{n-1} — матрицы модификации строк, причем Z_i отличается от I в точности i-м столбцом. Каждая из матриц $Z_1, \ldots, Z_{n-1}$ является нижней треугольной, поэтому их произведение

$$\widehat{L} = Z_{n-1} \ldots Z_1$$

является также нижней треугольной матрицей. Матрица коэффициентов системы $(*)$

$$U = Z_{n-1} \ldots Z_1 A = \widehat{L}\, A$$

является верхней треугольной. Матрица $L = \widehat{L}^{-1}$ является также нижней треугольной. Следовательно, метод Гаусса порождает разложение матрицы A в произведение нижней и верхней треугольной матриц

$$A = LU.$$

При этом L имеет на главной диагонали единицы, а U является невырожденной матрицей (в силу невырожденности A). Такое разложение называется *LU-разложением.*

Подсчитаем число арифметических операций при приведении A к верхней треугольной матрице U. На i-м шаге требуется получить $n-i$ нулей ниже диагонали в i-м столбце. При получении нуля на пересечении i-го столбца и l-й строки при $l > i$ из l-й строки вычитается i-я строка, предварительно умноженная на коэффициент, выбор которого и обеспечивает получение данного нуля. Поскольку в рассматриваемых строках может быть только $n-i$ ненулевых элементов, число умножений (и вычитаний) при получении одного нуля в i-м столбце равно $(n-i)^2$. Всего потребуется

$$(n-1)^2 + (n-2)^2 + \ldots + 1^2 \;=\; \frac{1}{3}n^3 + O(n^2)$$

умножений и столько же вычитаний; через $O(n^2)$ обозначен многочлен от n степени 2.

Чтобы найти решение системы $Ax = b$, требуется выполнить еще два действия:

- вычислить вектор $Z_{n-1} \ldots Z_1 b$;
- найти решение системы с верхней треугольной матрицей U.

Каждое из этих действий требует лишь $O(n^2)$ арифметических операций — на порядок меньше, чем приведение к верхнему треугольному виду.

Задача. Невырожденная матрица и обратная к ней разбиты на блоки одинаковых размеров:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}.$$

Доказать, что блок A_{11} невырожден тогда и только тогда, когда невырожден блок B_{22}.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

8.8 LU-разложение и строго регулярные матрицы

Допустим, что невырожденная матрица A имеет LU-разложение: $A = LU$. Обозначим через A_k, L_k, U_k подматрицы порядка k, расположенные в левом верхнем углу матриц A, L, U, соответственно, и рассмотрим равенство блочных матриц

$$A \equiv \begin{bmatrix} A_k & P \\ Q & \widetilde{A}_k \end{bmatrix} = \begin{bmatrix} L_k & 0 \\ V & \widetilde{L}_k \end{bmatrix} \begin{bmatrix} U_k & W \\ 0 & \widetilde{U}_k \end{bmatrix}.$$

Отсюда вытекает, что

$$A_k = L_k U_k, \qquad k = 1, \ldots, n.$$

Очевидно, что матрицы L_k и U_K невырожденные (как треугольные матрицы с ненулевой диагональю). Поэтому подматрица A_k должна быть невырожденной. Матрица A, в которой все подматрицы A_k невырожденные, называется *строго регулярной.*

Таким образом, для существования LU-разложения невырожденной матрицы A необходимо, чтобы она была строго регулярной.

Можно доказать, что это условие является также и достаточным. В самом деле, пусть уже построено LU-разложение для подматрицы $A_k = L_k U_k$. Тогда

$$\begin{bmatrix} L_k^{-1} & 0 \\ -QA_k^{-1} & I \end{bmatrix} \begin{bmatrix} A_k & P \\ Q & \widetilde{A}_k \end{bmatrix} = \begin{bmatrix} U_k & L_k^{-1}P \\ 0 & W \end{bmatrix}, \qquad W = \widetilde{A}_k - QA_k^{-1}P. \qquad (\#)$$

Блок W называется *дополнением по Шуру* блока A_k в матрице A. Из равенства (#) и строгой регулярности A можно вывести, что W является также строго регулярной матрицей. Предположим, что для W уже построено LU-разложение $W = \widetilde{L}_k\widetilde{U}_k$. Тогда положим

$$L = \begin{bmatrix} L_k & 0 \\ -QA_k^{-1}L_k & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & \widetilde{L}_k \end{bmatrix}, \qquad U = \begin{bmatrix} U_k & L_k^{-1}P \\ 0 & \widetilde{U}_k \end{bmatrix}.$$

Полученная таким образом матрица L — верхняя треугольная. Равенство $LU = A$ проверяется прямым вычислением.

8.9 Вычисление обратной матрицы

Обратную матрицу можно вычислить, используя конструкции того же метода Гаусса. Если получено разложение $A = LU$, то, поскольку $A^{-1} = U^{-1}L^{-1}$, достаточно научиться вычислять обратные к верхней и нижней треугольным матрицам. Общее число арифметических операций будет $O(n^3)$.

Однако, в 1965 году появилась работа Штрассена с заголовком "Метод Гаусса не оптимален", в которой впервые было показано, что существуют и более быстрые алгоритмы. Пусть имеется алгоритм умножения двух $n \times n$-матриц с числом операций $\leq cn^\gamma$ (например, в дополнительной части Лекции 1 обсуждается алгоритм Штрассена, для которого $\gamma = \log_2 7 < 3$). Тогда в случае строго регулярной матрицы A можно построить алгоритм вычисления A^{-1} с числом операций $O(n^\gamma)$.

Для простоты предположим, что $n = 2^p$. Разобьем A на блоки порядка $n/2$ и рассмотрим следующее равенство:

$$\begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & W \end{bmatrix}, \qquad W = A_{22} - A_{21}A_{11}^{-1}A_{12}.$$

Из невырожденности A и A_{11} вытекает невырожденность блока W. Более того, блоки A_{11} (что очевидно) и W (докажите!) наследуют строгую регулярность матрицы A. Нетрудно проверить, что

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & W \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}W^{-1} \\ 0 & W^{-1} \end{bmatrix}.$$

Таким образом,

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}W^{-1} \\ 0 & W^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}. \qquad (*)$$

Полученные из $(*)$ выражения для блоков матрицы A^{-1} иногда называют *формулами Фробениуса*.

Для нашей цели формула $(*)$ интересна тем, что показывает, как обращение матрицы порядка n сводится к двум аналогичным задачам для матриц A_{11} и W порядка $n/2$. Для реализации указанной редукции требуется выполнить несколько умножений матриц порядка $n/2$. Общие затраты на всех шагах редукции пропорциональны

$$\left(\frac{n}{2}\right)^\gamma + 2\left(\frac{n}{2^2}\right)^\gamma + 2^2\left(\frac{n}{2^3}\right)^\gamma + \ldots \quad \leq \quad \frac{n^\gamma}{2^\gamma}\,\frac{1}{1-1/2^\gamma} \;=\; \frac{n^\gamma}{2^\gamma - 1}.$$

Вслед за открытием Штрассена появилась работа других авторов под названием "Метод Штрассена не оптимален". Но лидерство нового алгоритма было не очень долгим. Соревнование по построению все более быстрых алгоритмов обращения матриц и решения систем продолжается до сих пор, а вопрос об оптимальном алгоритме с точки зрения числа операций остается открытым.

Еще менее ясным является вопрос об алгоритме с минимальным числом параллельных шагов (хотя бы в модели бесконечного параллелизма). Довольно давно был придуман алгоритм, в котором число параллельных шагов в случае матрицы общего вида есть $O(\log_2^2 n)$. Никто не знает, можно ли построить более быстрый параллельный алгоритм. Любопытно, что предъявленный алгоритм не имеет ничего общего ни с методом Гаусса, ни с методом Штрассена. Кроме того, даже для треугольной матрицы не известно алгоритма с меньшим числом параллельных шагов (по порядку зависимости от n).

Задача. Докажите, что определитель строго регулярной матрицы порядка n можно вычислить за $O(n^{\log_2 7})$ арифметических операций.

Лекция 9

ОСНОВНАЯ ЧАСТЬ

9.1 Метод координат

Нашим исследованиям линейной зависимости и линейных оболочек векторов (матриц-столбцов) можно дать наглядную геометрическую интерпретацию. Как скоро выяснится, определитель также имеет замечательный геометрический смысл. Особенно важно то, что "алгебраизация" геометрических понятий дает мощный алгебраический инструмент для решения задач геометрии.

К основным объектам геометрии относятся точки, прямые и плоскости в геометрическом пространстве. Если A и B — точки прямой, то пусть $[AB]$ обозначает отрезок прямой — множество точек данной прямой, расположенных между точками A и B; $|AB|$ — длина отрезка $[AB]$.

Будем опираться на то, что между вещественными числами и точками прямой существует взаимно-однозначное соответствие $x \leftrightarrow P(x)$, которое полностью определяется заданием двух точек $P(0)$, $P(1)$ и обладает следующими свойствами:

- если $x \neq 0$ и точки $P(x)$ и $P(1)$ находятся по одну сторону от точки $P(0)$, то $x > 0$; в противном случае $x < 0$;
- $|P(0)P(x)| \;=\; |x|\,|P(0)P(1)|$.

Прямую, для которой установлено указанное соответствие, будем называть *числовой осью*, а число x — *координатой* точки $P(x)$.

Заметим, что при выборе произвольного вещественного числа a соответствие $x \leftrightarrow P(x+a)$ будет также взаимно-однозначным. Это позволяет переносить точку $P(0)$ в любую заданную точку данной прямой.

Рассмотрим прямые l_x, l_y, l_z, проходящие через общую точку O и не лежащие в одной плоскости. Пусть каждая из этих прямых является числовой осью с соответствиями

$$x \leftrightarrow P_{l_x}(x), \quad y \leftrightarrow P_{l_y}(y), \quad z \leftrightarrow P_{l_z}(z),$$

дающими общую точку

$$P_{l_x}(0) = P_{l_y}(0) = P_{l_z}(0) = O.$$

Пусть (x, y, z) — система трех вещественных чисел, определяющих точки $X = P_{l_x}(x)$, $Y = P_{l_y}(y)$, $Z = P_{l_z}(z)$ на прямых l_x, l_y, l_z, соответственно. Рассмотрим три плоскости:

- π_x — плоскость, проходящая через точку X параллельно прямым l_y и l_z;
- π_y — плоскость, проходящая через точку Y параллельно прямым l_x и l_z;

- π_z — плоскость, проходящая через точку Z параллельно прямым l_x и l_y.

Легко видеть, что плоскости π_x, π_y, π_z пересекаются в одной точке $M = M(x, y, z)$. Таким образом устанавливается взаимно-однозначное соответствие

$$(x, y, z) \leftrightarrow M(x, y, z).$$

Точки X, Y, Z называются *проекциями* точки M на прямые l_x, l_y, l_z параллельно плоскостям, соответственно, π_x, π_y, π_z. Числа x, y, z называются *координатами* точки $M = M(x, y, z)$, а система числовых осей l_x, l_y, l_z — *аффинной системой координат*. Точка O называется *началом* (или *центром*) системы координат.

Эпитет "аффинная" по отношению к системе координат означает только то, что углы между осями могут не быть прямыми, а длины отрезков $[OP_{l_x}(1)]$, $[OP_{l_y}(1)]$, $[OP_{l_z}(1)]$ не обязательно равные. Если углы между осями прямые, а длины указанных отрезков равны 1, то система координат называется *декартовой*.

9.2 Направленные отрезки

Любую упорядоченную пару точек A, B будем называть *направленным отрезком* с *началом* в точке A и *концом* в точке B. Обозначение: $\overrightarrow{AB}$.

Если имеется система координат с началом в точке O, то направленный отрезок вида $\overrightarrow{OA}$ называется *радиус-вектором* точки A. Координаты точки A называются также координатами радиус-вектора $\overrightarrow{OA}$.

Точка $A \neq B$ разбивает прямую AB на два луча: луч $[AB)$, состоящий их точек данной прямой, лежащих вместе с B по одну сторону от A и дополнительный луч, состоящий из точек, лежащих по другую сторону (точка A для двух лучей является общей). Два луча на одной прямой называются одинаково направленными, если их пересечение является лучом (и противоположно направленными, если их пересечение является отрезком). Если прямые AB и CD не совпадают, то лучи $[AB)$ и $[CD)$ называются одинаково направленными, если эти прямые параллельны и точки B и D лежат по одну сторону от прямой AC.

Предположим, что $A \neq B$, рассмотрим отрезок $\overrightarrow{AB}$, и пусть C — произвольная точка. Проведем через C прямую, параллельную прямой AB или совпадающую с ней в случае $C \in AB$. На этой прямой можно найти ровно две точки D_1 и D_2 такие, что $|CD_1| = |CD_2| = |AB|$. Выберем из них такую точку $D \in \{D_1, D_2\}$, для которой лучи $[AB)$ и $[CD)$ одинаково направлены. Направленный отрезок $\overrightarrow{CD}$ будем считать *равным* $\overrightarrow{AB}$. (Часто говорят также, что $\overrightarrow{CD}$ получается из $\overrightarrow{AB}$ *параллельным переносом*.)

Данным построением не охвачен случай $A = B$. Направленные отрезки $\overrightarrow{AA}$ и $\overrightarrow{CC}$ будем считать равными *по определению* и называть их *нулевыми*.

Отметим формальную "несимметричность" в данном определении: $\overrightarrow{CD}$ равен $\overrightarrow{AB}$, но будет ли $\overrightarrow{AB}$ равен $\overrightarrow{CD}$? Ответ, к счастью, положительный — в силу того, что направленный отрезок $\overrightarrow{AB}$ получается из $\overrightarrow{CD}$ с помощью точно такой же конструкции.

Заметим, что все случаи при определении равенства направленных отрезков можно свести к одному случаю, если принять формально другое (и притом "симметричное") определение: назовем направленные отрезки $\overrightarrow{AB}$ и $\overrightarrow{CD}$ равными, если середины

отрезков $[AD]$ и $[BC]$ совпадают. Эквивалентность нового определения предыдущему вытекает из общеизвестных свойств параллелограмма.

Обратим внимание на то, что при фиксированной системе координат любой направленный отрезок равен некоторому и только одному радиус-вектору.

9.3 Отношение эквивалентности

Любое непустое подмножество $M \subset X \times X$ определяет на множестве X *бинарное отношение* между его элементами:

$$x \overset{M}{\sim} y \quad \Leftrightarrow \quad (x, y) \in M.$$

ПРИМЕР. X — множество всех матриц, M — множество таких пар матриц (A, B), для которых существует произведение AB. Ясно, что имеются пары матриц, не входящие в M. Кроме того, если $A \overset{M}{\sim} B$, то отсюда не следует, что $B \overset{M}{\sim} A$.

Бинарное отношение M на X называется *отношением эквивалентности*, если выполняются следующие три свойства:

- $x \overset{M}{\sim} x$ для всех элементов $x \in X$ (рефлексивность);
- если $x \overset{M}{\sim} y$, то $y \overset{M}{\sim} x$ (симметричность);
- если $x \overset{M}{\sim} y$ и $y \overset{M}{\sim} z$, то $x \overset{M}{\sim} z$ (транзитивность).

Если на X задано отношение эквивалентности M и $x \overset{M}{\sim} y$, то x и y называются *эквивалентными* элементами. Множество всех элементов из X, эквивалентных некоторому элементу $a \in X$, называется *классом эквивалентности, порожденным* элементом a.

Теорема. *Непустое множество X с отношением эквивалентности является объединением непересекающихся подмножеств, каждое из которых состоит из элементов, эквивалентных между собой и не эквивалентных ни одному из элементов других подмножеств.*

Доказательство. Пусть $X(a)$ обозначает класс эквивалентности, порожденный элементом $a \in X$. Выберем произвольный элемент x и рассмотрим его класс эквивалентности $X(a)$. Если $b, c \in X(a)$, то каждый из них эквивалентен a, а значит, в силу транзитивности, b и c эквивалентны между собой ($b \sim a,\ a \sim c \ \Rightarrow\ b \sim c$). Ясно также, что $X(b) = X(c) = X(a)$ (то есть, класс эквивалентности порождается любым своим представителем).

По определению, $X(a)$ содержит *абсолютно все* элементы, эквивалентные a. Поэтому если $b \notin X(a)$, то b не является эквивалентным a. Отсюда следует, что классы эквивалентности $X(a)$ и $X(b)$ не пересекаются: если бы имелся элемент $c \in X(a) \bigcap X(b)$, то это бы означало, что $b \in X(a) \quad \Rightarrow \quad X(a) = X(b)$.

Таким образом, для произвольных элементов a и b классы эквивалентности $X(a)$ и $X(b)$ либо не пересекаются, либо совпадают. Очевидно, $X = \bigcup\limits_{a \in X} X(a)$. Для завершения доказательства остается исключить из этого объединения совпадающие классы эквивалентности. $\square$

ПРИМЕР 1. Пусть G — произвольная (не обязательно абелева) группа. Элементы $a,\ b \in G$ называются *сопряженными*, если для некоторого $h \in G$ (зависящего от a и b)

выполняется равенство $a = hbh^{-1}$. Сопряженность элементов — это бинарное отношение на G, которое, как легко проверить, является отношением эквивалентности.

ПРИМЕР 2. Пусть $\mathbb{Z}$ — множество целых чисел, а p — некоторое натуральное (целое положительное) число. Целые числа x и y называются *сравнимыми по модулю* p, если они имеют одинаковые остатки при делении на p (это означает, что разность $x - y$ делится нацело на p, то есть $x - y = kp$ для некоторого целого k). Обозначение: $x = y\,(\mathrm{mod} p)$.

Пусть $x \sim y \;\Leftrightarrow\; x - y\,(\mathrm{mod})$. Это бинарное отношение на $\mathbb{Z}$ является отношением эквивалентности. В данном случае имеется ровно p различных классов эквивалентности

$$\mathbb{Z}(0),\ \mathbb{Z}(1),\ \ldots,\ \mathbb{Z}(p-1),$$

называемых обычно *вычетами по модулю* p.

9.4 Свободный вектор

Утверждение. *Отношение равенства направленных отрезков является отношением эквивалентности.*

Доказательство. Рассмотрим два бинарных отношения на множестве направленных отрезков:

- равенство длин: $\overrightarrow{AB} \overset{M_1}{\sim} \overrightarrow{CD} \;\Leftrightarrow\; |AB| = |CD|$;
- сонаправленность: $\overrightarrow{AB} \overset{M_2}{\sim} \overrightarrow{CD} \Leftrightarrow$ лучи $[AB)$ и $[CD)$ одинаково направлены.

Легко проверяется, что каждое из отношений M_1 и M_2 является отношением эквивалентности. Но $\overrightarrow{AB} = \overrightarrow{CD}$ тогда и только тогда, когда $\overrightarrow{AB} \overset{M_1}{\sim} \overrightarrow{CD}$ и одновременно $\overrightarrow{AB} \overset{M_2}{\sim} \overrightarrow{CD}$. $\square$

Определение. Любой класс эквивалентности направленных отрезков называется *свободным вектором*, или, короче, *вектором*.

Согласно определению, свободный вектор $\vec{a}$ содержит все эквивалентные между собой направленные отрезки. При этом для любой точки A существует единственная точка B такая, что $\overrightarrow{AB} \in \vec{a}$. В частности, при фиксированной системе координат всегда имеется один и только один радиус-вектор, принадлежащий $\vec{a}$.

Пусть V — множество всех точек геометрического пространства. Тогда вектор $\vec{a}$ задает следующее взаимно-однозначное отображение $V \to V$: точка $A \in V$ переходит в точку $B \in V$ такую, что $\overrightarrow{AB} \in \vec{a}$. Такое отображение называется *параллельным переносом*, или *сдвигом на вектор* $\vec{a}$.

Традиционно допускаемый элемент вольности в обозначениях: вместо $\overrightarrow{AB} \in \vec{a}$ принято писать $\vec{a} = \overrightarrow{AB}$ (вектор как класс эквивалентности отождествляется с любым его представителем).

9.5 Линейные операции над векторами

Сумма векторов: пусть $\overrightarrow{AB} \in \vec{a}$ и $\overrightarrow{BC} \in \vec{b}$, тогда $\vec{c} = \vec{a} + \vec{b}$ определяется как вектор, порожденный направленным отрезком $\overrightarrow{AC}$.

Важно, что получаемый таким образом вектор $\vec{c}$ не зависит от выбора точки A. В самом деле, пусть $\overrightarrow{PQ} \in \vec{a}$ и $\overrightarrow{QR} \in \vec{b}$. Тогда из равенства треугольников $\triangle ABC$ и $\triangle PQR$ вытекает равенство длин и сонаправленность направленных отрезков $\overrightarrow{AC}$ и $\overrightarrow{PR}$, а значит, и их равенство.

Множество свободных векторов относительно операции сложения векторов образует абелеву группу. Роль единичного элемента играет нулевой вектор $\vec{0} = \overrightarrow{AA}$. Для $\vec{a} = \overrightarrow{AB}$ обратным элементом является $\vec{b} = \overrightarrow{BA}$. В данном контексте вектор $\vec{b}$ назывется *противоположным вектором* для $\vec{a}$ и обозначается $\vec{b} = -\vec{a}$.

Умножение вектора на число: пусть $\overrightarrow{AB} \in \vec{a}$, тогда $\alpha\,\vec{a}$ определяется как вектор, порождаемый направленным отрезком $\overrightarrow{AC}$, который имеет длину $|AC| = |\alpha|\,|AB|$ и, если $\alpha \neq 0$, является одинаково направленным с $\overrightarrow{AB}$ при $\alpha > 0$ и противоположно направленным при $\alpha < 0$. Несложно убедиться в том, что вектор $\alpha\,\vec{a}$ не зависит от выбора точки A.

Несложно проверить, что $\alpha\,(\beta\,\vec{a}) = (\alpha\beta)\,\vec{a}$ для любых вещественных чисел α, β. Полезно также заметить, что $1 \cdot \vec{a} = \vec{a}$, а $(-1) \cdot \vec{a}$ есть вектор, противоположный к $\vec{a}$.

Кроме того, операции сложения векторов и умножения вектора на число связаны следующими законами дистрибутивности:

$$(\alpha + \beta)\,\vec{a} = (\alpha\,\vec{a}) + (\beta\,\vec{a}), \qquad \alpha\,(\vec{a} + \vec{b}) = (\alpha\,\vec{a}) + (\alpha\,\vec{b}).$$

9.6 Координаты вектора

Пусть фиксирована некоторая аффинная система координат. Как уже отмечалось, каждому свободному вектору соответствует один и только один радиус-вектор. Его координаты и будем называть координатами данного свободного вектора.

Пусть O — начало системы координат с числовыми осями l_x, l_y, l_z и точками $X \in l_x$, $Y \in l_y$, $Z \in l_z$, соответствующими числу 1 на данных осях. Система векторов

$$\vec{e}_x = \overrightarrow{OX}, \quad \vec{e}_y = \overrightarrow{OY}, \quad \vec{e}_z = \overrightarrow{OZ}$$

называется для данной системы координат *базисной* (иногда также *реперной*).

Непосредственно из определения координат точки и линейных операций над векторами вытекает следующее

Утверждение 1. *Пусть x, y, z — координаты вектора $\vec{a} = \overrightarrow{OA}$ в системе координат с базисными векторами $\vec{e}_x$, $\vec{e}_y$, $\vec{e}_z$. В этом и только в этом случае имеет место разложение*

$$\vec{a} \;=\; x\,\vec{e}_x \;+\; y\,\vec{e}_y \;+\; z\,\vec{e}_z.$$

Векторы $x\,\vec{e}_x$, $y\,\vec{e}_y$, $z\,\vec{e}_z$ называются *проекциями* вектора $\vec{a}$ на прямые l_x, l_y, l_z (они,

как легко видеть, не зависят от способа превращения прямых в числовые оси).

Утверждение 2. *Пусть x_a, y_a, z_a и x_b, y_b, z_b — координаты векторов $\vec{a} = \overrightarrow{OA}$ и $\vec{b} = \overrightarrow{OB}$, соответственно. Тогда вектор $\vec{c} = \vec{a} + \vec{b}$ имеет координаты*

$$x_c = x_a + x_b, \quad y_c = y_a + y_b, \quad z_c = z_a + z_b,$$

а вектор $\vec{d} = \alpha\,\vec{a}$ для любого вещественного числа α имеет координаты

$$x_d = \alpha\, x_a, \quad y_d = \alpha\, y_a, \quad z_d = \alpha\, z_a.$$

Для доказательства достаточно установить, что проекция суммы векторов для каждой оси есть сумма проекций данных векторов, а проекция вектора, умноженного на число, есть умноженная на это число проекция данного вектора.

9.7 Изоморфизм и линейная зависимость

Пусть V — множество всех свободных векторов. Каждый свободный вектор можно отождествить с соответствующим ему радиус-вектором, а каждый радиус-вектор вида $\overrightarrow{OA}$ — с точкой A геометрического пространства.

Утверждение 2 позволяет установить такое взаимно-однозначное соответствие между множеством свободных векторов V и множеством матриц-столбцов $\mathbb{R}^3$, при котором *сохраняются операции* сложения векторов и умножения векторов на числа: если $\vec{a} \leftrightarrow \mathbf{a} \in \mathbb{R}^3$ и $\vec{b} \leftrightarrow \mathbf{b} \in \mathbb{R}^3$, то

$$\vec{a} + \vec{b} \leftrightarrow \mathbf{a} + \mathbf{b}, \qquad \alpha\,\vec{a} \leftrightarrow \alpha\,\mathbf{a}.$$

Взаимно-однозначные отображения, сохраняющие операции, принято называть *изоморфизмами*, а множества, между которыми такое соответствие установлено, *изоморфными*. Таким образом, множество свободных векторов V изоморфно $\mathbb{R}^3$.

Понятия линейной зависимости и линейной независимости систем свободных векторов вводятся точно так же, как и для матриц-столбцов. То же относится к понятию линейных оболочек. Учитывая изоморфизм, в случае свободных векторов мы можем использовать результаты уже выполненного для матриц-столбцов исследования линейной зависимости и связанных с ней понятий базиса и размерности линейной оболочки. Легко видеть, что введенные выше базисные векторы $\vec{e}_x$, $\vec{e}_y$, $\vec{e}_z$ являются линейно независимыми, а все множество свободных векторов есть их линейная оболочка:

$$V = L(\vec{e}_x,\, \vec{e}_y,\, \vec{e}_z),$$

при этом

$$\dim V = 3.$$

9.8 Коллинеарные и компланарные векторы

Определение 1. Векторы называются *коллинеарными*, если среди порождающих их направленных отрезков имеются принадлежащие одной прямой.

Определение 2. Векторы называются *компланарными*, если среди порождающих их направленных отрезков имеются принадлежащие одной плоскости.

Линейная оболочка любой системы коллинеарных векторов, содержащей хотя бы один ненулевой вектор, имеет размерность 1. Верно и обратное: все векторы из линейной оболочки размерности 1 являются коллинеарными.

Линейная оболочка любой системы компланарных векторов, в которой имеется хотя бы одна пара неколлинеарных векторов, имеет размерность 2. Все векторы из линейной оболочки размерности 2 являются компланарными.

Будем отождествлять свободные векторы с порождающими их радиус-векторами. Тогда множество всех векторов, коллинеарных заданному вектору, представляет собой прямую, проходящую через начало координат. Множество всех векторов, компланарных заданной паре неколлинеарных векторов, представляет собой проходящую через начало кооординат плоскость.

Прямая $l = AB$, проходящая через точки A и B, представляет собой множество точек (радиус-векторов) следующего вида:

$$l = \{M : \overrightarrow{OM} = \overrightarrow{OA} + t\overrightarrow{AB},\ t \in \mathbb{R}\}. \tag{1}$$

Вектор $\overrightarrow{AB}$ (параллельный прямой l) называется *направляющим вектором* для l.

Плоскость π, проходящая через три не лежащие на одной прямой точки A, B, C, есть множество точек (радиус-векторов) вида

$$\pi = \{M : \overrightarrow{OM} = \overrightarrow{OA} + u\overrightarrow{AB} + v\overrightarrow{AC},\ u, v \in \mathbb{R}\}. \tag{2}$$

9.9 Прямая на плоскости

В качестве геометрического пространства часто рассматривается плоскость. В этом случае система координат состоит из двух осей и устанавливает взаимно-однозначное соответствие между системами двух вещественных чисел (x, y) и точками (радиус-векторами) плоскости.

Пусть A и B — вещественные числа, не равные нулю одновременно. Уравнение вида

$$Ax + By + C = 0 \tag{$*$}$$

называется *общим уравнением прямой на плоскости.*

Теорема. Пусть на плоскости фиксирована аффинная система координат. Множество точек с координатами x, y, удовлетворяющими уравнению $(*)$, представляет собой прямую, и при этом любая прямая может быть задана уравнением вида $(*)$.

Доказательство. Пусть l — прямая, проходящая через точки (x_0, y_0) и (x_1, y_1). Тогда, согласно (1), прямая l состоит из точек (x, y) таких, что

$$\begin{cases} x &= x_0 + tp_x, \\ y &= y_0 + tp_y, \end{cases} \qquad t \in \mathbb{R}, \qquad p_x = x_1 - x_0, \quad p_y = y_1 - y_0..$$

Отсюда (как определитель с линейно зависимыми столбцами)

$$\det \begin{bmatrix} x - x_0 & p_x \\ y - y_0 & p_y \end{bmatrix} = 0.$$

$$\Rightarrow \quad Ax + by + C = 0, \quad \text{где} \quad A = p_y,\ B = -p_x,\ C = -p_y x_0 + p_x y_0.$$

Теперь рассмотрим множество точек (x, y), удовлетворяющих уравнению $(*)$. Поскольку хотя бы одно из чисел A, B отлично от нуля, это множество заведомо не пусто. Пусть $Ax_0 + By_0 + C = 0$. Тогда $(*)$ равносильно уравнению

$$A(x - x_0) + B(y - y_0) = 0.$$

Предположим для определенности, что $A \neq 0$. Тогда $x - x_0 = (y - y_0)(-B/A)$. Положим $p_x = -B/A,\ p_y = 1$. Тогда при $t = y - y_0$ находим

$$\begin{cases} x &= x_0 + tp_x, \\ y &= y_0 + tp_y. \end{cases}$$

Если t — произвольное вещественное число, то определенные таким образом x, y удовлетворяют уравнению $(*)$. $\square$

9.10 Плоскость в пространстве

Пусть A, B, C — вещественные числа, не равные нулю одновременно. Уравнение вида

$$Ax + By + Cz + D = 0 \qquad (\#)$$

называется *общим уравнением плоскости*.

Теорема. Пусть в пространстве фиксирована аффинная система координат. Множество точек с координатами x, y, z, удовлетворяющими уравнению $(\#)$, представляет собой плоскость, и при этом любая плоскость может быть задана уравнением вида $(\#)$.

Доказательство. Пусть π — плоскость, проходящая через точки (x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2). Тогда, согласно (2), плоскость π состоит из точек (x, y, z) таких, что

$$\begin{cases} x &= x_0 + up_x + vq_x, \\ y &= y_0 + up_y + vq_y, \\ z &= z_0 + up_z + vq_z, \end{cases} \qquad u, v \in \mathbb{R},$$

$$(p_x, p_y, p_z) = (x_1 - x_0, y_1 - y_0, z_1 - z_0), \quad (q_x, q_y, q_z) = (x_2 - x_0, y_2 - y_0, z_2 - z_0).$$

Отсюда

$$\det \begin{bmatrix} x - x_0 & p_x & q_x \\ y - y_0 & p_y & q_y \\ z - z_0 & p_z & q_z \end{bmatrix} = 0.$$

Как уравнение относительно x, y, z, очевидно, это уравнение имеет вид $(\#)$.

Теперь рассмотрим множество (заведомо непустое) точек (x, y, z), удовлетворяющих уравнению $(\#)$. Пусть $Ax_0 + By_0 + Cz_0 + D = 0$. Тогда $(\#)$ равносильно уравнению

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0.$$

Предположим, что $A \neq 0$. Тогда $x - x_0 = (y - y_0)(-B/A) + (z - z_0)(-C/A)$. Положим

$$(p_x,\ p_y,\ p_z) = (-B/A,\ 1,\ 0), \qquad (q_x,\ q_y,\ q_z) = (-C/A,\ 0,\ 1).$$

Тогда при $u = y - y_0,\ v = z - z_0$ находим

$$\begin{cases} x &= x_0 + up_x + vq_x, \\ y &= y_0 + up_y + vq_y, \\ z &= z_0 + up_z + vq_z. \end{cases}$$

Если u, v — произвольные вещественные числа, отсюда получаем x, y, z, удовлетворяющие $(\#)$. $\square$

Лекция 10

ОСНОВНАЯ ЧАСТЬ

10.1 Скалярное произведение геометрических векторов

Длиной вектора $\overrightarrow{a}$ называется длина порождающего его направленного отрезка (направленные отрезки, порождающие один и тот же вектор, равны и поэтому имеют одинаковую длину). Обозначение для длины: $|\vec{a}|$. Углом $\phi(\vec{a},\,\vec{b})$ между ненулевыми векторами $\vec{a}=\overrightarrow{OA},\ \ \vec{b}=\overrightarrow{OB}$ называется угол между сторонами OA и OB в треугольнике OAB.

Скалярным произведением векторов $\vec{a}$ и $\vec{b}$ называется величина

$$(\vec{a},\,\vec{b}) \;=\; \begin{cases} |\vec{a}|\;|\vec{b}|\;\cos\phi(\vec{a},\,\vec{b}), & \vec{a}\neq\vec{0}\ \text{ и }\ \vec{b}\neq\vec{0},\\ 0, & \vec{a}=\vec{0}\ \text{ или }\ \vec{b}=\vec{0}. \end{cases}$$

В силу определения очевидно, что

$$(\vec{a},\,\vec{a})>0 \quad \text{при} \quad \vec{a}\neq\vec{0}; \qquad (\vec{a},\,\vec{a})=0 \;\;\Leftrightarrow\;\; \vec{a}=\vec{0}. \tag{1}$$

Так же очевидно, что

$$(\vec{a},\,\vec{b})=(\vec{b},\,\vec{a}) \qquad \forall\ \ \vec{a},\ \vec{b}. \tag{2}$$

Если векторы $\vec{a}=\overrightarrow{OA},\ \vec{b}=\overrightarrow{OB}$ неколлинеарны, то в плоскости, проходящей через точки O, A, B можно ввести декартову систему координат с началом в точке O и первой осью, совпадающей с прямой OB и дающей точке B положительную координату. Тогда величина $|\vec{a}|\,\cos\phi(\vec{a},\,\vec{b})$ будет в точности координатой точки A на данной оси. Отсюда сразу же вытекают важные свойства линейности скалярного произведения по первому аргументу:

$$(\vec{a}+\vec{b},\,\vec{c}) \;=\; (\vec{c},\,\vec{a}) \;+\; (\vec{c},\,\vec{b}) \qquad \forall\ \ \vec{a},\ \vec{b}, \tag{3}$$

$$(\alpha\,\vec{a},\,\vec{b}) \;=\; \alpha\;(\vec{a},\,\vec{b}) \qquad \forall\vec{a},\ \vec{b}, \quad \forall\ \ \alpha\in\mathbb{R}. \tag{4}$$

Свойство (2) сразу же дает аналогичные свойства линейности скалярного произведения и по второму аргументу.

Векторы называются *ортогональными*, если их скалярное произведение равно нулю.

10.2 Скалярное произведение и координаты

Пусть задана декартова система координат с базисными векторами $\vec{e}_1,\ \vec{e}_2,\ \vec{e}_3$. Тогда

$$(\vec{e}_i,\, \vec{e}_j) \ = \ \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \qquad (*)$$

Теорема. *Пусть в заданной декартовой системе координат вектор $\vec{a}$ имеет координатами a_1, a_2, a_3, а вектор $\vec{b}$ — координаты b_1, b_2, b_3. Тогда имеет место формула*

$$(\vec{a},\, \vec{b}) \ = \ a_1b_1 + a_2b_2 + a_3b_3. \qquad (\#)$$

Доказательство. Запишем

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3, \qquad \vec{b} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3.$$

Опираясь на свойства скалярного произведения $(2) - (4)$ и сооношения $(*)$, находим

$$\begin{aligned} (\vec{a},\, \vec{b}) \ &= \ (a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3,\ b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3) \\ &= \ \sum_{i=1}^{3}\sum_{j=1}^{3} a_ib_j(\vec{e}_i,\, \vec{e}_j) \ = \ a_1b_1 + a_2b_2 + a_3b_3. \quad \square \end{aligned}$$

Замечание 1. Если в некоторой системе координат скалярное произведение любых векторов $\vec{a}$ и $\vec{b}$ вычисляется по формуле $(\#)$, то данная система координат декартова.

Замечание 2. В случае декартовой системы координат для векторов на плоскости формула $(\#)$ приобретает вид

$$(\vec{a},\, \vec{b}) \ = \ a_1b_1 + a_2b_2.$$

10.3 Об обобщениях

Формула $(\#)$ и свойства $(1) - (4)$ дают основу для введения скалярного произведения в более общих случаях — для объектов, уже не являющихся векторами в геометрическом пространстве.

Например, если $a = [a_1, \ldots, a_n]^\top,\ b = [b_1, \ldots, b_n]^\top$ — матрицы-столбцы из $\mathbb{R}^n$, то можно определить их скалярное произведение по аналогии с формулой $(\#)$:

$$(a,\, b) \ = \ a_1b_1 + \ldots + a_nb_n. \qquad (*)$$

Есть и другая идея, имеющая более общий характер — взять за основу свойства $(1) - (4)$ и называть скалярным произведением любую функцию от матриц-столбцов a, b, удовлетворяющую аксиомам $(1) - (4)$.

Для геометрических векторов скалярное произведение определялось на основе таких понятий, как длина вектора и угол между векторами. В более общих случаях проще ввести каким-то образом скалярное произведение и уже с его помощью вводить понятия длины и угла:

$$|a| = \sqrt{(a,\, a)}, \qquad \cos\phi(a,\, b) = \frac{(a,\, b)}{|a|\ |b|}.$$

Например, опираясь на $(*)$, можно ввести таким образом длину и угол для векторов $a, b \in \mathbb{R}^n$. При этом важно, что

$$|(a,\, b)| \le \sqrt{(a,\, a)}\sqrt{(b,\, b)}.$$

Это неравенство (известное как неравенство Коши–Буняковского–Шварца) легко выводится из $(*)$, но в действительно оно справедливо для всех мыслимых способов задания скалярного произведения — подробный разговор на эту тему будет позже.

10.4 Ориентация системы векторов

Понятие ориентации для тройки (системы из трех) некомпланарных векторов вводится в буквальном смысле слова "на пальцах": тройка векторов называется *правой*, если их можно расположить как большой, несогнутый [1] указательный и средний пальцы правой руки; тройка векторов называется *левой*, если их можно расположить как большой, несогнутый указательный и средний пальцы левой руки.

Очевидно, может возникнуть желание освободиться от анатомической компоненты этого определения. Например, таким образом: тройка векторов называется правой, если кратчайший поворот от первого вектора ко второму происходит против часовой стрелки, если он наблюдается из конца третьего вектора.

Конечно, остается чувство неудовлетворения по поводу обоих определений. Но оно имеет неустранимый характер — в силу фундаментальных причин. Дело в том, что любые тройки некомпланарных векторов могут иметь ровно два типа ориентации, а фиксация одного из них, вообще говоря, произвольна. [2]

Можно выбрать произвольную декартову систему координат и объявить, что тройка ее базисных векторов имеет, скажем, "правильную ориентацию". Пусть вектор $\vec{a}$ имеет координаты $a_1,\ a_2,\ a_3$, вектор $\vec{b}$ — координаты $b_1,\ b_2,\ b_3$, вектор $\vec{c}$ — координаты $c_1,\ c_2,\ c_3$. Тройку векторов $\vec{a},\ \vec{b},\ \vec{c}$ можно назвать тройкой "правильной ориентации", если

$$\det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} > 0.$$

Если определитель меньше нуля, то это будет тройка "неправильной ориентации". Таким образом, определение ориентации зависит от объявления типа ориентации для исходной системы координат.

Аналогичным образом можно ввести понятие ориентации для пар векторов на плоскости и даже для систем n матриц-столбцов из $\mathbb{R}^n$.

10.5 Векторное и смешанное произведения

Векторным произведением неколлинеарных векторов $\vec{a}$ и $\vec{b}$ называется вектор $\vec{c}$ такой, что:

- вектор $\vec{c}$ ортогонален $\vec{a}$ и $\vec{b}$;
- тройка векторов $\vec{a},\ \vec{b},\ \vec{c}$ является правой;

[1] Если указательный палец согнуть, то получится совсем не то.

[2] Как утверждает М. М. Постников, в старые времена правыми называли как раз сегодняшние левые тройки.

- $|\vec{c}| = |\vec{a}|\,|\vec{b}|\,\sin\phi(\vec{a},\,\vec{b})$.

Если $\vec{a}$ и $\vec{b}$ коллинеарны, то $\vec{c} = \vec{0}$. Обозначение: $\vec{c} = [\vec{a},\,\vec{b}]$.

Если $\vec{a} = \overrightarrow{OA},\ \vec{b} = \overrightarrow{OB}$, то длина вектора $\vec{c}$, очевидно, равна площади параллелограмма со сторонами OA и OB.

Число, равное скалярному произведению векторов $[\vec{a},\,\vec{b}]$ и $\vec{c}$, называется *смешанным произведением* векторов $\vec{a},\ \vec{b},\ \vec{c}$. Обозначение:

$$(\vec{a},\,\vec{b},\,\vec{c}) \;=\; ([\vec{a},\,\vec{b}],\ \vec{c}).$$

Теорема. *Пусть векторы $\vec{a},\ \vec{b},\ \vec{c}$ некомпланарны и V — объем параллелепипеда, натянутого на приведенные к общему началу векторы $\vec{a},\ \vec{b},\ \vec{c}$. Тогда*

$$(\vec{a},\,\vec{b},\,\vec{c}) \;=\; \begin{cases} V, & \text{если тройка } \vec{a},\ \vec{b},\ \vec{c} \text{ правая;} \\ -V, & \text{в противном случае.} \end{cases}$$

Если векторы $\vec{a},\ \vec{b},\ \vec{c}$ компланарны, то

$$(\vec{a},\,\vec{b},\,\vec{c}) \;=\; 0.$$

Доказательство. Предположим, что $\vec{a} = \overrightarrow{OA},\ \vec{b} = \overrightarrow{OB},\ \vec{c} = \overrightarrow{OC}$, и пусть $\overrightarrow{OD} = [\vec{a},\,\vec{b}]$. Согласно определению смешанного произведения,

$$(\vec{a},\,\vec{b},\,\vec{c}) = |\overrightarrow{OD}|\,\gamma, \qquad \text{где} \qquad \gamma \;=\; \cos\phi(\overrightarrow{OD},\,\overrightarrow{OC}).$$

Ясно, что $|\gamma|$ есть длина перпендикуляра, опущенного из точки C на плоскость OAB (высота параллелепипеда). При этом $\gamma > 0$, если точки D и C находятся по одну сторону от плоскости OAB, и $\gamma < 0$, если эти точки оказались по разные стороны от данной плоскости. В первом случае тройка векторов $\overrightarrow{OA},\ \overrightarrow{OB},\ \overrightarrow{OC}$ правая, во втором — левая. □

Следствие 1.

$$(\vec{a},\,\vec{b},\,\vec{c}) = (\vec{b},\,\vec{c},\,\vec{a}) = (\vec{c},\,\vec{a},\,\vec{b}) = -(\vec{b},\,\vec{a},\,\vec{c}) = -(\vec{a},\,\vec{c},\,\vec{b}) = -(\vec{c},\,\vec{b},\,\vec{a}).$$

Доказательство. Достаточно заметить, что тройки векторов

$$\{\vec{a},\ \vec{b},\ \vec{c}\}, \qquad \{\vec{b},\ \vec{c},\ \vec{a}\}, \qquad \{\vec{c},\ \vec{a},\ \vec{b}\}$$

имеют одинаковую ориентацию, противоположную ориентации троек векторов

$$\{\vec{b},\ \vec{a},\ \vec{c}\}, \qquad \{\vec{a},\ \vec{c},\ \vec{b}\}, \qquad \{\vec{c},\ \vec{b},\ \vec{a}\}. \quad \square$$

Следствие 2. *Смешанное произведение $(\vec{a},\,\vec{b},\,\vec{c})$ линейно по каждому аргументу.*
Доказательство. Из свойств скалярного произведения сразу же вытекает линейность по третьему аргументу. Остается заметить, что тройки $\{\vec{a},\ \vec{b},\ \vec{c}\}$, $\{\vec{b},\ \vec{c},\ \vec{a}\}$, $\{\vec{c},\ \vec{a},\ \vec{b}\}$ имеют одинаковую ориентацию. Поэтому

$$(\vec{a},\ \vec{b},\ \vec{c}) \;=\; (\vec{b},\ \vec{c},\ \vec{a}) \;=\; (\vec{c},\ \vec{a},\ \vec{b}).$$

Следовательно, имеет место линейность по первому и второму аргументам. $\square$

Следствие 3. *Векторное произведение антисимметрично:*

$$[\vec{a},\,\vec{b}] = -[\vec{b},\,\vec{a}].$$

Доказательство. Пусть $\vec{d} = [\vec{a},\,\vec{b}] + [\vec{b},\,\vec{a}]$. Тогда

$$(\vec{d},\,\vec{d}) = (\vec{a},\,\vec{b},\,\vec{d}) + (\vec{b},\,\vec{a},\,\vec{d}) = (\vec{a},\,\vec{b},\,\vec{d}) - (\vec{a},\,\vec{b},\,\vec{d}) = 0 \quad \Rightarrow \quad \vec{d} = \vec{0}. \quad \square$$

Следствие 4. *Векторное произведение линейно по каждому аргументу.*
Доказательство. Докажем, что $[\vec{a}+\vec{b},\,\vec{c}] = [\vec{a},\,\vec{c}] + [\vec{b},\,\vec{c}]$. Для этого рассмотрим вектор

$$\vec{d} = [\vec{a}+\vec{b},\,\vec{c}] - [\vec{a},\,\vec{c}] - [\vec{b},\,\vec{c}].$$

Используя линейность смешанного произведения по каждому аргументу, находим

$$(\vec{d},\,\vec{d}) = (\vec{a}+\vec{b},\,\vec{c},\,\vec{d}) - (\vec{a},\,\vec{c},\,\vec{d}) - (\vec{b},\,\vec{c},\,\vec{d}) = 0 \quad \Rightarrow \quad \vec{d} = \vec{0}.$$

Аналогично, пусть $\vec{d} = [\alpha\,\vec{a},\,\vec{b}] - \alpha\,[\vec{a},\,\vec{b}]$. Тогда

$$(\vec{d},\,\vec{d}) = (\alpha\,\vec{a},\,\vec{b},\,\vec{d}) - \alpha\,(\vec{a},\,\vec{b},\vec{d}) = 0 \quad \Rightarrow \quad \vec{d} = \vec{0}.$$

Линейность по второму аргументу вытекает из утверждения 3. $\square$

Отметим также два простых, но полезных предложения.

Критерий коллинеарности. *Векторы $\vec{a},\ \vec{b}$ коллинеарны тогда и только тогда, когда $[\vec{a},\,\vec{b}] = 0$.*

Критерий компланарности. *Векторы $\vec{a},\ \vec{b},\ \vec{c}$ компланарны тогда и только тогда, когда $(\vec{a},\,\vec{b},\vec{c}) = 0$.*

10.6 Векторное произведение в декартовых координатах

Пусть $\vec{e}_1,\ \vec{e}_2,\ \vec{e}_3$ — базисные векторы декартовой системы координат. Легко проверить, что

$$[\vec{e}_1,\,\vec{e}_2] = \vec{e}_3, \qquad [\vec{e}_2,\,\vec{e}_3] = \vec{e}_1, \qquad [\vec{e}_3,\,\vec{e}_1] = \vec{e}_2.$$

Для векторов $\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3, \quad \vec{b} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3$ получаем

$$\begin{aligned}
[\vec{a},\,\vec{b}] \;&=\; a_1b_1[\vec{e}_1,\,\vec{e}_1] + a_1b_2[\vec{e}_1,\,\vec{e}_2] + a_1b_3[\vec{e}_1,\vec{e}_3] \\
&+\; a_2b_1[\vec{e}_2,\,\vec{e}_1] + a_2b_2[\vec{e}_2,\,\vec{e}_2] + a_2b_3[\vec{e}_2,\vec{e}_3] \\
&+\; a_3b_1[\vec{e}_3,\,\vec{e}_1] + a_3b_2[\vec{e}_3,\,\vec{e}_2] + a_3b_3[\vec{e}_3,\vec{e}_3] \\
&=\; (a_2b_3 - a_3b_2)\,\vec{e}_1 \;-\; (a_1b_3 - a_3b_1)\,\vec{e}_2 \;+\; (a_1b_2 - a_2b_1)\,\vec{e}_3.
\end{aligned}$$

Полученный результат легче всего запомнить, увидев в нем формальное применение теоремы Лапласа для разложения определителя по первой строке:

$$[\vec{a},\,\vec{b}] \;=\; \det\begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$

10.7 Смешанное произведение в декартовых координатах

После уже выполненного нами исследования индикатора линейной зависимости, приведшего нас к понятию определителя, ответ в данном случае требует одного-единственного вычисления:

$$(\vec{e}_1,\, \vec{e}_2,\, \vec{e}_3) = 1.$$

Таким образом, смешанное произведение $(\vec{a}, \vec{b}, \vec{c})$ как функция столбцов

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

составленных из координат векторов $\vec{a},\ \vec{b},\ \vec{c}$, является индикатором линейной зависимости, и, в силу его единственности, совпадает с определителем:

$$(\vec{a}, \vec{b}, \vec{c}) = \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}.$$

Следствие. *Определитель по абсолютной величине — это объем параллелепипеда, натянутого на векторы, определяемые его столбцами.*

10.8 Нормали к прямой и плоскости

Ненулевой вектор, ортогональный прямой, называется ее *нормалью*. Если прямая на плоскости с декартовой системой координат задана общим уравнением

$$Ax + By + C = 0,$$

то вектор $\vec{n} = (A,\, B)$ ортогонален любому вектору на данной прямой. В самом деле, любой вектор на прямой имеет вид $\vec{l} = (x_2 - x_1,\, y_2 - y_1)$, где $(x_1,\, y_1)$ и $(x_2,\, y_2)$ — две точки на данной прямой. Подставляя координаты точек в общее уравнений прямой на плоскости, находим

$$A(x_2 - x_1) + B(y_2 - y_1) = 0 \quad \Leftrightarrow \quad (\vec{n},\, \vec{l}) = 0.$$

Аналогично, ненулевой вектор, ортогональный плоскости, называется нормалью для данной плоскости. Если плоскость в пространстве с декартовой системой координат задана общим уравнением

$$Ax + By + Cz + D,$$

то вектор $\vec{n} = (A,\, B,\, C)$ — ее нормаль.

Используя векторное произведение, нормаль можно построить, имея пару неколлинеарных векторов $\vec{a}$ и $\vec{b}$, принадлежащих плоскости: вектор $[\vec{a},\, \vec{b}]$ ортогонален плоскости векторов $\vec{a}$ и $\vec{b}$. (Конечно, нормаль к плоскости определена однозначно с точностью до ненулевого коэффициента.)

10.9 Расстояние от точки до прямой на плоскости

Рассмотрим прямую $l:\ Ax+By+C=0$ на плоскости с декартовой системой координат и точку $M_0=(x_0,y_0)\notin l$. Для того чтобы найти расстояние $\rho(M_0,\,l)$ от точки M_0 до прямой l, нужно выполнить такие действия:

- провести через точку M_0 прямую l_0, ортогональную прямой l;
- найти точку $M_1=(x_1,y_1)$ пересечения прямых l_0 и l;
- вычислить длину отрезка M_0M_1.

Мы уже знаем, что вектор $\vec{n}=(A,\,B)$ ортогонален прямой l. Поэтому прямая l_0 есть множество точек вида

$$l_0=\{(x,y):\ x=x_0+At,\ y=y_0+Bt,\quad t\in\mathbb{R}\}.$$

Найдем значение параметра t, при котором $(x,y)\in l$:

$$A(x_0+At)+B((y_0+Bt)+C=0\quad\Rightarrow\quad t=-\frac{Ax_0+By_0+C}{A^2+B^2}.$$

Далее, $\overrightarrow{M_0M_1}=(At,\,Bt)\quad\Rightarrow$

$$\rho(M_0,\,l)=|\overrightarrow{M_0M_1}|=\frac{|Ax_0+By_0+C|}{\sqrt{A^2+B^2}}.$$

10.10 Расстояние от точки до плоскости

Рассмотрим плоскость $\pi:\ Ax+By+Cz+D=0$ в геометрическом пространстве с декартовой системой координат и точку $M_0=(x_0,y_0,z_)\notin\pi$. Расстояние $\rho(M_0,\,l)$ от точки M_0 до плоскости π вычисляется в полной аналогии со случаем точки и прямой на плоскости:

$$\rho(M_0,\,\pi)=\frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}.$$

10.11 Критерии параллельности вектора прямой и плоскости

Пусть на плоскости заданы прямая $l:Ax+By+c=0$ и вектор $\vec{v}=(v_1,v_2)$. Если система координат декартова, то вектор $\vec{n}=(A,B)$ является нормалью к прямой l. Поэтому вектор $\vec{v}$ параллелен прямой l тогда и только тогда, когда $(\vec{v},\vec{n})=0$. Учитывая вид скалярного произведения в декартовых координатах, получаем

$$\vec{v}\parallel l\quad\Leftrightarrow\quad Av_1+Bv_2=0. \tag{1}$$

Для плоскости $\pi:Ax+By+Cz+D=0$ и вектора $\vec{v}=(v_1,v_2,v_3)$ в пространстве с декартовой системой координат получаем аналогичный критерий параллельности:

$$\vec{v}\parallel\pi\quad\Leftrightarrow\quad Av_1+Bv_2+Cv_3=0. \tag{2}$$

Заметим, что критерии параллельности (1) и (2) остаются в силе и в случае произвольной аффинной системы координат (докажите!).

10.12 Полуплоскости и полупространства

Пусть на плоскости дана прямая $l : Ax + By + C = 0$. Тогда любая точка $P = (x, y)$ на плоскости принадлежит одному из трех множеств

$$l = \{(x, y) : Ax + By + C = 0\},$$

$$\pi^+ \;=\; \{(x, y) : Ax + By + C > 0\}, \quad \pi^- \;=\; \{(x, y) : Ax + By + C < 0\}.$$

Говорят, что прямая l делит плоскость на две *полуплоскости* π^+ и π^-.

Возьмем две точки $P = (x_1, y_1)$ и $Q = (x_2, y_2)$, тогда любая точка отрезка PQ имеет координаты

$$x = x_1 + t(x_2 - x_1) = (1 - t)x_2 + tx_1, \quad y = y_1 + t(y_2 - y_1) = (1 - t)y_2 + ty_1, \qquad 0 \le t \le 1.$$

Отсюда ясно, что если P и Q принадлежат обе одному из множеств π^+ или π^-, то все точки отрезка PQ принадлежат тому же множеству.

Множество точек, содержащее вместе с любыми двумя точками все точки соединяющего их отрезка, называется *выпуклым*. Таким образом, каждое из множеств l, π^+, π^- является выпуклым.

Теперь предположим, что $P \in \pi^+$, но $Q \in \pi^-$. Тогда уравнение

$$A(x_1 + (t(x_2 - x_1)) + B(y_1 + t(y_2 - y_1)) + C = 0$$

выполняется при

$$t \;=\; \frac{Ax_1 + By_1 + C}{(Ax_1 + By_1 + C) - (Ax_2 + By_2 + C)},$$

откуда видно, что $0 < t < 1$. Следовательно, некоторая точка отрезка PQ принадлежит прямой l.

Итак, *две точки принадлежат одной полуплоскости относительно заданной прямой l в том и только том случае, когда соединяющий их отрезок не имеет общих точек с прямой l.*

Аналогично, плоскость $\pi : Ax + By + Cz + D = 0$ делит пространство на два полупространства

$$\pi^+ = \{(x, y, z) : Ax + By + Cz + D > 0\}, \qquad \pi^- = \{(x, y, z) : Ax + By + Cz + D < 0\}.$$

При этом *две точки принадлежат одному полупространству относительно заданной плоскости π в том и только том случае, когда соединяющий их отрезок не пересекается с плоскостью π.*

Лекция 11

ОСНОВНАЯ ЧАСТЬ

11.1 Линейные пространства

При изучении линейной зависимости векторов, линейных оболочек, базисов, размерностей в предыдущих лекциях мы полагали, что векторы — это матрицы-столбцы с вещественными элементами. Впрочем, при изучении ранга матрицы речь уже заходила также о линейной зависимости и независимости строк матрицы. Конечно, с формальной точки зрения строки можно транспонировать и снова иметь дело с матрицами-столбцами. Однако, все перечисленные выше понятия и многие полученные факты без всяких изменений можно применять и в случае, когда под векторами понимаются матрицы каких-либо фиксированных размеров. Уже одно это заставляет подумать о введении более общего (и более абстрактного) понятия вектора.

Кроме того, изучая базисы и размерности, мы имели дело исключительно с линейными оболочками векторов, а это не всегда удобно: например, множество всех решений однородной системы линейных алгебраических уравнений $Ax = 0$ является, конечно, линейной оболочкой векторов фундаментальной системы решений, но было бы полезно иметь право обсуждать свойства этого множества без упоминания об образующей его системе векторов.

Давайте скажем, что векторы — это элементы некоторого непустого множества V, на котором определены две операции: сложение векторов (если $a, b \in V$, то $a + b \in V$) и умножение векторов на вещественные числа (если $a \in V$ и $\alpha \in \mathbb{R}$, то $\alpha a \in V$) . Потребуем, чтобы данные операции обладали следующими свойствами:

- $(a + b) + c = a + (b + c) \quad \forall \ \ a, b \in V \quad$ (ассоциативность сложения векторов);
- существует особый вектор 0, называемый *нулевым вектором*, такой что
$$a + 0 = 0 + a = a \quad \forall a \in V;$$
- для любого вектора $a \in V$ существует вектор $b \in V$ такой, что
$$a + b = b + a = 0;$$
- $a + b = b + a \quad \forall \ \ a, b \in V \quad$ (коммутативность сложения векторов);
- $\alpha(\beta\, a) = (\alpha\beta)\, a \quad \forall \ \ \alpha, \beta \in \mathbb{R}, \ \ \forall \ \ a \in V;$
- $(\alpha + \beta)\, a = (\alpha a) + (\beta a) \quad \forall \ \ \alpha, \beta \in \mathbb{R}, \ \ \forall a \in V \quad$ (дистрибутивность);

- $\alpha(a+b) = (\alpha a) + (\alpha b) \quad \forall \ \alpha \in \mathbb{R}, \quad \forall \ a, b \in V$ (дистрибутивность);
- $1 \cdot a = a \quad \forall \ a \in V$. [1]

В таких случаях множество V называется вещественным *линейным пространством*. Часто встречающийся термин-синоним — *векторное пространство*.

Заметим, что множество V относительно операции сложения векторов является абелевой группой. Роль единичного элемента играет нулевой вектор. Вектор b такой, что $a + b = b + a = 0$, называется *противоположным* к вектору a и обозначается $b \equiv -a$.

Некоторые привычные свойства данных операций, ранее свободно применявшихся к матрицам-столбцам, в рассмотренном более абстрактном случае *нуждаются в доказательствах*.

Утверждение 1. $0 \cdot a = 0 \quad \forall \ a \in V$.
Доказательство. В силу дистрибутивности, $0 \cdot a = (0+0) \cdot a = (0 \cdot a) + (0 \cdot a)$. Далее, пусть $b = -(0 \cdot a)$ (противоположный вектор к вектору $0 \cdot a$). Тогда $0 = b + (0 \cdot a) = (b + (0 \cdot a)) + (0 \cdot a) \quad \Rightarrow \quad 0 = 0 \cdot a$. $\square$

Утверждение 2. $(-1) \cdot a = -a \quad \forall \ a \in V$.
Доказательство. В силу утверждения 1 и дистрибутивности, $0 = 0 \cdot a = (1 + (-1)) \cdot a = (1 \cdot a) + ((-1) \cdot a) = a + ((-1) \cdot a)$. $\square$

Утверждение 3. *Если* $\alpha \cdot a = 0$, *то либо* $\alpha = 0$, *либо* $a = 0$.
Доказательство. [2] Пусть $\alpha \neq 0$. Тогда

$$a = 1 \cdot a = \left(\left(\frac{1}{\alpha}\right)\alpha\right) \cdot a = \left(\frac{1}{\alpha}\right)(\alpha \cdot a) = \left(\frac{1}{\alpha}\right) \cdot 0 = 0. \qquad \square$$

Как и раньше, для любых чисел $\alpha_1, \ldots, \alpha_n$ вектор w вида

$$w = \alpha_1 a_1 + \ldots + \alpha_n a_n$$

называется линейной комбинацией векторов $a_1, \ldots a_n$, а множество всех линейных комбинаций со всеми возможными значениями коэффициентов $\alpha_1, \ldots, \alpha_n$ называется линейной оболочкой векторов $a_1, \ldots, a_n$ и обозначается $L(a_1, \ldots, a_n)$.

11.2 Примеры бесконечномерных линейных пространств

(1) *Множество функций с вещественными значениями на отрезке* $[0, 1]$.

Сумма $f + g$ функций f и g определяется как функция со значениями $(f+g)(x) = f(x) + g(x)$. При умножении функции на число получается функиця αf, определяемая правилом $(\alpha f)(x) = \alpha f(x)$. Роль нулевого вектора выполняет функция, тождественно равная нулю.

[1] Данное свойство равносильно тому, что каждый вектор a можно представить в виде $a = \alpha b$ для некоторого вектора b и некоторого числа α. В самом деле, если это свойство выполнено, то можно взять $b = a$ и $\alpha = 1$. С другой стороны, пусть выполнение этого свойства не предполагается, но известно, что $a = \alpha b$. Тогда, используя аксиому $\alpha(\beta\, a) = (\alpha\beta)\, a$, получаем $1 \cdot (\alpha b) = (1 \cdot \alpha) \cdot b = \alpha b \quad \Rightarrow \quad 1 \cdot a = a$.

[2] Утверждение нельзя получить без аксиомы $1 \cdot a = a$. В самом деле, возьмем любую абелеву группу V с нулевым элементом 0 и определим умножение на число правилом $\alpha a = 0$ для всех чисел α и векторов $a \in V$. При этом будут выполнены все аксиомы линейного пространства, кроме данной.

(2) *Множество бесконечных последовательностей* $\{x_k\}_{k=1}^{\infty}$.

Сумма последовательностей $\{x_k\}$ и $\{y_k\}$ определяется как последовательность $\{z_k\}$ с членами $z_k = x_k + y_k$. Произведение последовательности $\{x_k\}$ на число α определяется как последовательность $\{z_k\}$ с членами $z_k = \alpha x_k$. Роль нулевого вектора выполняет последовательность, в которой все элементы равны нулю.

(3) *Множество сходящихся последовательностей* $\{x_k\}_{k=1}^{\infty}$.

Операции определяются так же, как и в случае произвольных бесконечных последовательностей. Необходимо лишь заметить, что сумма сходящихся последовательностей останется сходящейся последовательностью, а умножение сходящейся последовательности на число также дает сходящуюся последовательность.

Примеры **(1)-(3)** замечательны тем, что соответствующие линейные пространства не являются линейной оболочкой какого-то конечного числа своих векторов. Такие линейные пространства называются *бесконечномерными*.

Докажем, например, бесконечномерность пространства функций. Предположим от противного, что оно совпадает с линейной оболочкой каких-то функций $f_1, \ldots, f_n$. Тогда любая функция f имеет вид

$$f(x) = \alpha_1 f_1(x) + \ldots + \alpha_n f_n(x). \qquad (*)$$

Выберем n попарно различных точек $x_1, \ldots, x_n \in [0, 1]$ и для произвольно выбранной функции f рассмотрим систему уравнений

$$\begin{array}{ccccccc} \alpha_1(f)\, f_1(x_1) & + & \ldots & + & \alpha_n(f)\, f_n(x_1) & = & f(x_1), \\ \ldots & & \ldots & & \ldots & & \ldots \\ \alpha_1(f)\, f_1(x_n) & + & \ldots & + & \alpha_n(f)\, f_n(x_n) & = & f(x_n). \end{array}$$

Это есть система линейных алгебраических уравнений относительно $\alpha_1(f), \ldots, \alpha_n(f)$. Если матрица коэффициентов данной системы необратима, то заведомо решение существует не для любой правой части. Тогда равенство $(*)$ не выполняется хотя бы для одной функции f. Следовательно, матрица коэффициентов должна быть обратимой. Поэтому для заданной функции f коэффициенты $\alpha_1(f), \ldots, \alpha_n(f)$ определены однозначно.

Пусть теперь точка $x_* \in [0, 1]$ не совпадает ни с одной из точек $x_1, \ldots, x_n$. Заведомо существует функция g такая, что $g(x_i) = f(x_i)$ при $i = 1, \ldots, n$, но $g(x_*) \neq f(x_*)$. Ясно, что $\alpha_i(f) = \alpha_i(g)$ при $i = 1, \ldots, n$, откуда $f = g$, чего быть не может, поскольку $f(x_*) \neq g(x_*)$. $\square$

11.3 Примеры конечномерных линейных пространств

Линейные пространства, представляющие собой линейную оболочку некоторого конечного числа своих векторов, называются *конечномерными*.

(1) *Множество многочленов порядка* n.

Многочленом (полиномом) от x порядка n называется функция, представимая в виде

$$f(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \ldots + a_1 x + a_0.$$

Если $a_k \neq 0$ и $a_{k+1} = \ldots = a_{n-1} = 0$, то k называется *степенью* многочлена $f(x)$. Выражения вида ax^i называются *одночленами*, или *мономами*.

Сумма многочленов и умножение многочлена на число определяются так же, как в случае функций общего вида. При этом ясно, что результаты этих операций остаются многочленами. Очевидно, линейное пространство всех многочленов порядка n является линейной оболочкой одночленов вида

$$x^{n-1}, \quad x^{n-2}, \quad \ldots, \quad x^1, \quad x^0 \equiv 1.$$

(2) *Множество $m \times n$-матриц с фиксированными размерами m и n.*

В данном случае сложение векторов определяется как сложение матриц, а умножение вектора на число — как умножение матрицы на число.

Обозначим через $E^{kl} = [(E^{kl})_{ij}]$ матрицу размеров $m \times n$ с элементами вида

$$(E^{kl})_{ij} = \begin{cases} 1, & i = k, j = l, \\ 0, & \text{иначе.} \end{cases}$$

Таких матриц ровно mn и очевидно, что все пространство $m \times n$-матриц является их линейной оболочкой.

(3) *Множество всех решений однородной системы линейных алгебраических уравнений $Ax = b$.*

Если ранг $m \times n$-матрицы A равен r, то фундаментальная система решений данной однородной системы содержит $n - r$ векторов, а все множество решений совпадает с их линейной оболочкой.

Данное линейное пространство называется *нуль-пространством*, или *ядром* матрицы A. Обозначение: $\ker A$ (в некоторых книгах $\operatorname{null} A$).

(4) *Множество всех столбцов вида $y = Ax$ (для заданной матрицы A).*

Это хорошо знакомое нам множество, совпадающее с линейной оболочкой столбцов матрицы A. Оно называется *образом* матрицы A. Обозначение: $\operatorname{im} A$.

11.4 Базис и размерность

Пусть V — конечномерное пространство. По определению, оно является линейной оболочкой конечного числа своих векторов:

$$V = L(a_1, \ldots, a_n).$$

Понятия линейно зависимой и линейно независимой систем векторов в абстрактном случае ничем не отличаются от тех же понятий в случае матриц-столбцов. То же справедливо в отношении базиса и размерности:

- V можно представить как линейную оболочку некоторой линейно независимой подсистемы векторов $a_1, \ldots, a_n$;
- базис в пространстве V определяется как любая линейно независимая система векторов, для которой V является линейной оболочкой; любые два базиса в V содержат одинаковое число векторов; число векторов в базисе называется размерностью пространства V и обозначается $\dim V$;

- любую линейно независимую систему векторов из V можно достроить до базиса V; более того, это можно сделать с помощью части векторов $a_1, \ldots, a_n$.

Доказательства этих предложений повторяют доказательства из Лекции 3 для частного случая линейных пространств — когда под векторами подразумевались матрицы-столбцы.

11.5 Подпространства линейного пространства

Непустое множество $W \subset V$ называется *подпространством* линейного пространства V, если оно само является линейным пространством относительно операций, действующих в V. Ясно, что для того чтобы W было подпространством, необходимо и достаточно, чтобы для любых векторов $a, b \in W$ и любого числа α имели место включения $a+b \in W$ и $\alpha a \in W$.

Если векторы $a_1, \ldots, a_n$ принадлежат подпространству W, то $L(a_1, \ldots, a_n) \subset W$.

ПРИМЕР. Рассмотрим множество V всех свободных векторов на плоскости с системой координат с началом в точке O. Поскольку каждый свободный вектор порождается одним и только одним радиус-вектором, любое подмножество свободных векторов можно отождествлять с подмножеством радиус-векторов $\overrightarrow{OA}$ или их концов — точек A.

Множество V, очевидно, является линейным пространством. Любая прямая, проходящая через начало координат, является подпространством в V. В то же время, если l — прямая, не проходящая через начало координат, то она подпространством не является: пусть $A, B \in l$ и $\overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{OB}$; ясно, что $C \notin l$.

Задача. Докажите, что линейное пространство $\mathbb{R}^n$ нельзя представить в виде объединения конечного числа множеств, каждое из которых не совпадает с $\mathbb{R}^n$ и является его подпространством.

11.6 Сумма и пересечение подпространств

Пусть P и Q — подпространства линейного пространства V. Под суммой $P+Q$ понимается множество всех векторов вида $w = p + q$, где $p \in P$, $q \in Q$. Под пересечением $P \cap Q$ понимается обычное пересечение множеств.

Утверждение. *Множества $P+Q$ и $P \cap Q$ являются подпространствами в V.*

Доказательство.

(1) Рассмотрим произвольную линейную комбинацию векторов $w_1, w_2 \in P+Q$. По определению множества $P+Q$, $w_1 = p_1 + q_1$ и $w_2 = p_1 + q_2$, где $p_1, p_2 \in P$ и $q_1, q_2 \in Q$. Тогда

$$\alpha_1 w_1 + \alpha_2 w_2 = (\alpha_1 p_1 + \alpha_2 p_2) + (\alpha_1 q_1 + \alpha_2 q_2) \in P + Q,$$

поскольку вектор в первой скобке принадлежит P, а вектор второй скобки принадлежит Q (P и Q — подпространства, поэтому они содержат все линейные комбинации своих векторов).

(2) Аналогично, рассмотрим линейную комбинацию векторов $w_1, w_2 \in P \cap Q$:

$$\alpha w_1 + \alpha_2 w_2 \in P \quad \text{и одновременнно} \quad \alpha_1 w_1 + \alpha_2 w_2 \in Q$$

$$\Rightarrow \qquad \alpha w_1 + \alpha_2 w_2 \in P \cap Q. \quad \square$$

Заметим, что любые два подпространства имеют непустое пересечение: каждое из них содержит, по крайней мере, нулевой вектор.

Теорема Грассмана. *Пусть W_1 и W_2 — конечномерные подпространства линейного пространства V. Тогда*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2).$$

Доказательство. Рассмотрим базис $g_1, \ldots, g_r$ подпространства $W_1 \cap W_2$ и дополним его сначала до базиса W_1

$$g_1, \ldots, g_r,\ p_{r+1}, \ldots, p_{r+k}, \qquad r + k = \dim W_1,$$

а затем до базиса W_2

$$g_1, \ldots, g_r,\ q_{r+1}, \ldots, q_{r+m}, \qquad r + m = \dim W_2.$$

Очевидно,

$$W_1 + W_2 = L(g_1, \ldots, g_r,\ p_{r+1}, \ldots, p_{r+k},\ q_{r+1}, \ldots, q_{r+m}).$$

Поэтому остается доказать линейную независимость векторов, порождающих данную линейную оболочку. Пусть

$$\alpha_1 g_1 + \ldots + \alpha_r g_r \ + \ \beta_{r+1} p_{r+1} + \ldots + \beta_{r+k} p_{r+k} \ + \ \gamma_{r+1} q_{r+1} + \ldots + \gamma_{r+m} q_{r+m} \ = \ 0.$$

Отсюда

$$\alpha_1 g_1 + \ldots + \alpha_r g_r \ + \ \beta_{r+1} p_{r+1} + \ldots + \beta_{r+k} p_{r+k} \ = \ -(\gamma_{r+1} q_{r+1} + \ldots + \gamma_{r+m} q_{r+m}) \in W_1 \cap W_2.$$

Поскольку $W_1 \cap W_2 = L(g_1, \ldots, g_r)$, для каких-то коэффициентов $\delta_1, \ldots, \delta_r$ имеем

$$\delta_1 g_1 + \ldots + \delta_r g_r \ = \ -(\gamma_{r+1} q_{r+1} + \ldots + \gamma_{r+m} q_{r+m}).$$

Это равносильно равенству

$$\delta_1 g_1 + \ldots + \delta_r g_r \ + \ \gamma_{r+1} q_{r+1} + \ldots + \gamma_{r+m} q_{r+m} = 0$$

$$\Rightarrow \quad \delta_1 = \ldots = \delta_r = \gamma_{r+1} = \ldots = \gamma_{r+m} = 0 \ \Rightarrow \quad \alpha_1 = \ldots = \alpha_r = \beta_{r+1} = \ldots = \beta_{r+m} = 0. \quad \square$$

Лекция 12

ОСНОВНАЯ ЧАСТЬ

12.1 Разложение по базису

Пусть V — вещественное линейное пространство размерности n и $f_1, \ldots, f_n$ — некоторый его базис. Тогда любой вектор $v \in V$ имеет однозначное разложение по данному базису

$$v = x_1 f_1 + \ldots + x_n f_n.$$

Коэффициенты $x_1, \ldots, x_n$ называются координатами вектора v в данном базисе. Понятно, что между элементами линейного пространства V и множества столбцов $\mathbb{R}^n$ имеется взаимно-однозначное соответствие

$$v \quad \leftrightarrow \quad x = \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}.$$

При выборе другого базиса $g_1, \ldots, g_n$ возникает еще одно взаимно-однозначное соответствие между теми же множествами:

$$v = y_1 g_1 + \ldots + y_n g_n \quad \leftrightarrow \quad y = \begin{bmatrix} y_1 \\ \ldots \\ y_n \end{bmatrix}.$$

Рассмотрим разложения

$$\begin{array}{ccccccc} f_1 & = & p_{11} g_1 & + & \ldots & + & p_{n1} g_n, \\ \ldots & & \ldots & & \ldots & & \\ f_n & = & p_{1n} g_1 & + & \ldots & + & p_{nn} g_n \end{array} \qquad (*)$$

и введем $n \times n$-матрицу $P = [p_{ij}]$. Используя $(*)$, легко получить соотношение

$$y = Px,$$

позволяющее переходить от координат вектора в базисе $\{f_i\}$ к координатам того же вектора в базисе $\{g_i\}$. (Докажите, что P обратима — это очень легко.)

Матрицу P логично было бы называть матрицей перехода от базиса $\{f_i\}$ к базису $\{g_i\}$. Но она все же называется обычно матрицей перехода от базиса $\{g_i\}$ к базису $\{f_i\}$. Впрочем, дело не в названии — важно, конечно, правильно ею пользоваться при пересчете координат!

12.2 Изоморфизм линейных пространств

Два вещественных линейных пространства V и W называются *изоморфными*, если существует взаимно-однозначное отображение $\Phi : V \to W$, сохраняющее операции в следующем смысле:

$$\Phi(a+b) = \Phi(a) + \Phi(b), \quad \Phi(\alpha\, a) = \alpha\, \Phi(a) \qquad \forall\ \ a, b \in V, \ \ \forall\ \ \alpha \in \mathbb{R}.$$

Само отображение Φ называется при этом *изоморфизмом*.

Заметим, что операции сложения векторов и умножения на число в левой и правой частях данных равенств, вообще говоря, *разные*! Операции слева действуют в V, а операции справа — в W. Тем не менее, если установлено, что пространства изоморфны, то это означает их неразличимость с точки зрения *свойств операций*.

Утверждение. $\Phi(0) = 0, \qquad \Phi(-a) = -\Phi(a) \ \ \forall\ a \in V.$

Доказательство. $\Phi(0) = \Phi(0+0) = \Phi(0) + \Phi(0)$. Прибавим к обеим частям вектор $b = -\Phi(a)$ (вектор, противоположный к $\Phi(a)$):

$$0 = b + \Phi(0) = (b + \Phi(0)) + \Phi(0) = 0 + \Phi(0) = \Phi(0) \quad \Rightarrow \quad \Phi(0) = 0. \qquad \square$$

На множестве всех вещественных линейных пространств изоморфизм порождает, очевидно, отношение эквивалентности. Важно, что исследования, выполненные для одного пространства V, сразу же переносятся на любое изоморфное ему пространство. Например, векторы $a_1, \ldots, a_n \in V$ линейно зависимы тогда и только тогда, когда линейно зависимы векторы $\Phi(a_1), \ldots, \Phi(a_n)$.

Теорема. *Любое вещественное линейное пространство V размерности $n = \dim V$ изоморфно $\mathbb{R}^n$.*

Доказательство. Выберем какой-нибудь базис $a_, \ldots, a_n$ в пространстве V и определим отображение Φ следующим образом:

$$\Phi(v) = \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix},$$

где $x_1, \ldots, x_n$ — коэффициенты разложения вектора v по данному базису:

$$v \ = \ x_1 a_1 + \ \ldots \ + x_n a_n.$$

Сохранение операций проверяется очевидным образом. $\square$

Следствие. *Любые конечномерные вещественные пространства одинаковой размерности изоморфны.*

12.3 Пространство многочленов

Пусть $\mathcal{P}_n$ — линейное пространство многочленов порядка n с вещественными коэффициентами. Докажем, что $\mathcal{P}_n$ изоморфно $\mathbb{R}^n$.

Любой многочлен $p(x)$ порядка n имеет вид

$$p(x) = p_{n-1} x^{n-1} + \ldots + p_1 x + p_0. \qquad (*)$$

Поэтому кажется, что с определением изоморфизма Φ нет проблем:

$$\Phi(p(x)) = \begin{bmatrix} p_0 \\ \dots \\ p_{n-1} \end{bmatrix}.$$

Действительно, это отображение сохраняет операции. Но будет ли оно взаимно-однозначным?

Если под многочленом понимается *формальное выражение* вида $(*)$ и при этом равенство многочленов *определяется* как равенство всех коэффициентов при одинаковых степенях x, то взаимная однозначность очевидна.

Если же под многочленом понимается *функция от x вида* $(*)$, то равенство многочленов определяется как равенство функций. В этом случае требуется *доказать*, что коэффициенты в представлении $(*)$ определяются по функции $p(x)$ однозначно. Для этого достаточно установить линейную независимость одночленов

$$x^0,\ x^1,\ \dots,\ x^{n-1}$$

как функций от x.

Предположим от противного, что данные одночлены линейно зависимы. Поскольку это ненулевые функции, существует одна из них, линейно выражающаяся через предыдущие:

$$x^k = \alpha_0 + \alpha_1 x + \dots + \alpha_{k-1} x^{k-1}.$$

Понятно, что такого быть не может, если эти функции рассматриваются как функции на всей оси $(-\infty,\ \infty)$: поделим обе части на x^k и перейдем в обеих частях к пределу при $x \to \infty$; слева получится 1, а справа 0.

Как быть, если эти функции рассматриваются на конечном отрезке, например, на $[0,\ 1]$? Предположим, что

$$p_0 + p_1 x + \dots + p_{n-1} x^{n-1} = 0 \quad \forall\ \ x \in [0,\ 1]. \qquad (\#)$$

В этом случае можно поступить следующим образом. Выберем произвольные попарно различные числа $x_1 \dots, x_n \in [0,\ 1]$. Равенство $(\#)$ имеет место при всех $x \in [0,\ 1]$, поэтому мы имеем право рассмотреть его только для выбранных значений $x = x_1, \dots, x_n$:

$$\begin{cases} p_0 \cdot 1 + p_1 \cdot x_1 + \dots + p_{n-1} \cdot x_1^{n-1} &= 0, \\ \dots \\ p_0 \cdot 1 + p_1 \cdot x_n + \dots + p_{n-1} \cdot x_n^{n-1} &= 0. \end{cases}$$

Это однородная система линейных алгебраических уравнений с матрицей коэффициентов

$$A = \begin{bmatrix} 1 & x_1 & \dots & x_1^{n-1} \\ \dots & \dots & \dots & \dots \\ 1 & x_n & \dots & x_n^{n-1} \end{bmatrix}.$$

Матрица такого вида называется *транспонированной матрицей Вандермонда*, а матрица $A^\top$ — *матрицей Вандермонда* порядка n для чисел $x_1, \dots, x_n$. Обозначение: $A^\top = V(x_1, \dots, x_n)$.

Утверждение. *Определитель матрицы Вандермонда $V(x_1, \dots, x_n)$ в случае попарно различных чисел $x_1,\ \dots\ x_n$ равен*

$$\det V(x_1, \dots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Доказательство. Определитель

$$\det V(x_1,\ldots,x_n) = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ \ldots & \ldots & \ldots & \ldots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{bmatrix}$$

не изменится, если из i-й вычесть $i-1$-ую строку, умноженную на x_n. При этом в последнем столбце i-й элемент станет нулем. Указанные действия выполним последовательно для строк с номерами $i = n, n-1, \ldots, 2$. В результате находим

$$\det V(x_1,\ldots,x_n) = \det \begin{bmatrix} 1 & 1 & \ldots & 1 & 1 \\ x_1 - x_n & x_2 - x_n & \ldots & x_{n-1} - x_n & 0 \\ \ldots & \ldots & \ldots & & \\ x_1^{n-2}(x_1 - x_n) & x_2^{n-2}(x_2 - x_n) & \ldots & x_{n-1}^{n-2}(x_{n-1} - x_n) & 0 \end{bmatrix}.$$

Применим теорему Лапласа для разложения определителя по последнему столбцу:

$$\begin{aligned} \det V(x_1,\ldots,x_n) &= (-1)^{n+1}\,(x_1 - x_n)(x_2 - x_n)\ldots(x_{n-1} - x_n)\ \det V(x_1,\ldots,x_{n-1}) \\ &= (x_n - x_1)(x_n - x_1)\ldots(x_n - x_{n-1})\ \det V(x_1,\ldots,x_{n-1}). \end{aligned}$$

Доказательство завершается по индукции. □

Следствие 1. *Определитель Вандермонда в случае попарно различных чисел отличен от нуля.*

Следствие 2. *Если функция вида* $(*)$ *равна нулю для* n *попарно различных значений* x*, то*

$$p_0 = p_1 = \ldots = p_{n-1} = 0.$$

Отсюда вытекает линейная независимость одночленов как функций на любом фиксированном отрезке.

Задача. Даны попарно различные числа $x_1, \ldots, x_n, y_1, \ldots, y_n$ и известно, что для каких-то чисел $u_1, \ldots, u_n$ выполняются равенства

$$\prod_{k=1}^{n}(x - y_k)\sum_{j=1}\frac{u_j}{x - y_j} = 0 \quad \text{при} \quad x = x_1, \ldots, x_n.$$

Доказать, что $u_1 = \ldots = u_n = 0$. Вывести отсюда невырожденность $n \times n$-матрицы с элементами $1/(x_i - y_j)$.

12.4 Прямая сумма подпространств

В линейном пространстве наряду с разложениями векторов по базису часто представляют интерес также разложения векторов по некоторым системам подпространств.

Пусть $W_1, \ldots, W_m$ — подпространства в линейном пространстве V. Множество

$$W = W_1 + \ldots + W_m \equiv \{w = w_1 + \ldots + w_m :\ w_1 \in W_1, \ldots, w_m \in W_m\}$$

называется *суммой* подпространств $W_1, \ldots, W_m$. Конечно, W является подпространством в V (доказательство для суммы двух подпространств легко адаптируется и к случаю суммы m подпространств).

В случае, если для любого вектора $w \in W_1 + \ldots + W_m$ в разложении

$$w = w_1 + \ldots + w_m, \quad w_1 \in W_1, \ \ldots, \ w_m \in W_m,$$

векторы $w_1 \in W_1, \ \ldots, \ w_m \in W_m$ определяются однозначно, сумма данных подпространств называется *прямой суммой*. Подпространства, сумма которых является прямой, называются *линейно независимыми*.

Если $e_1, \ldots, e_n$ — любая линейно независимая система векторов, то сумма их линейных оболочек

$$W = L(e_1) + \ \ldots \ + \ L(e_n)$$

является прямой суммой. Это наблюдение обобщается следующим образом.

Теорема. *Пусть V — конечномерное пространство и $W_1, \ \ldots, \ W_m$ — его подпространства. Сумма*

$$W = W_1 + \ldots + W_n$$

является прямой суммой тогда и только тогда, когда объединение произвольно выбранных базисов для $W_1, \ \ldots, \ W_m$ дает базис подпространства W.

Доказательство. Пусть W является прямой суммой. Предположим, что $\dim W_i = n_i$, и рассмотрим $W_1, \ldots, W_m$ как линейные оболочки своих базисов:

$$W_1 = L(v_{11}, \ldots, v_{n_1 1}), \qquad \ldots \qquad W_m = L(v_{1m}, \ldots, v_{n_m m}).$$

Докажем, что объединение базисов образует базис в V. Ясно, что W есть линейная оболочка объединения базисов:

$$W = L(v_{11}, \ \ldots, \ v_{n_1 1}, \ \ \ldots, \ \ v_{1m}, \ \ldots, \ v_{n_m m}).$$

Поэтому остается лишь убедиться в линейной независимости векторов объединения базисов. Пусть w — произвольная линейная комбинация этих векторов. Запишем w виде $w = w_1 + \ldots + w_m$, где $w_i \in W_i, \ i = 1, \ldots, m$. Если $w = 0$, то в силу единственности векторов $w_1 \in W_1, \ldots, w_m \in W_m$ данного разложения получаем $w_1 = \ldots = w_m = 0$. Отсюда следует, что все коэффициенты в разложении $w = 0$ по объединенной системе $v_{11}, \ \ldots, \ v_{n_1 1}, \ \ \ldots, \ \ v_{1m}, \ \ldots, \ v_{n_m m}$ равны нулю.

Пусть теперь объединение базисов подпространств $W_1, \ldots, W_m$ дает базис W. Единственность разложения вектора по базису означает единственность векторов $w_1 \in W_1, \ldots, w_m \in W_m$ в разложении $w = w_1 + \ldots + w_m$. $\square$

12.5 Дополнительные пространства и проекции

Если линейное пространство V является прямой суммой своих подпространств

$$L + M = V,$$

то M называется *дополнительным пространством* для L. В силу симметрии суммы очевидно, что L является дополнительным для M.

В таких случаях для любого вектора $v \in V$ существует единственное разложение

$$v = x + y, \quad \text{где} \quad x \in L, \ \ y \in M.$$

Вектор x называется *проекцией* вектора v на подпространство L *параллельно* (вдоль подпространства) M, а y — проекцией вектора v на M параллельно L.

Утверждение. *Сумма двух подпространств $L+M$ является прямой тогда и только тогда, когда $L \cap M = \{0\}$.*

Доказательство. Пусть сумма прямая и $x \in L \cap M$. Тогда мы имеем два разложения $x = x + 0$ и $x = 0 + x$, в которых первый вектор из L, а второй из M. В силу единственности компонент разложения, $x = 0$.

Пусть теперь $L \cap M = \{0\}$, и пусть $x_1 + y_1 = x_2 + y_2,\ \ x_1, x_2 \in L,\ \ y_1, y_2 \in M$. Отсюда $x_1 - x_2 = -(y_1 - y_2) \in L \cap M \quad \Rightarrow \quad x_1 - x_2 = -(y_1 - y_2) = 0 \quad \Rightarrow \quad x_1 = x_2,\ y_1 = y_2$. $\square$

ПРИМЕР. В линейном пространстве V всех свободных векторов (радиус-векторов) рассмотрим плоскость L и прямую M, проходящие через начало координат. Если прямая лежит в плоскости, то сумма $L+M$ равна L и не является прямой. Во всех других случаях имеем прямую сумму $V = L + M$ и можем рассматривать проекции радиус-вектора $\overrightarrow{OA}$ (точки A) на данную плоскость параллельно прямой и на прямую параллельно плоскости.

12.6 Вычисление подпространства

Под вычислением конечномерного подпространства W в линейном пространстве V обычно понимается построение какого-либо его базиса — линейно независимой системы векторов $w_1, \ldots, w_k$ такой, что $W = L(w_1, \ldots, w_k),\ \ k = \dim W$.

ПРИМЕР. Пусть подпространства $W_1, W_2 \subset \mathbb{R}^5$ определяются следующим образом:

- W_1 — множество всех решений однородной системы

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 &= 0, \\ x_4 + x_5 &= 0; \end{cases}$$

- W_2 — множество всех решений однородной системы

$$\begin{cases} x_1 + x_2 + \ x_3 + x_4 + \ x_5 &= 0, \\ 2x_3 + x_4 + 2x_5 &= 0, \\ x_5 &= 0. \end{cases}$$

Требуется вычислить подпространства $W_1,\ W_2,\ W_1 + W_2$ и $W_1 \bigcap W_2$.

Обозначим матрицы коэффициентов данных систем через A и B. Тогда

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \qquad W_1 = \ker A, \quad W_2 = \ker B.$$

Матрицы A и B имеют верхнюю ступенчатую форму с числом ступененей 2 и 3, соответствввенно. Поэтому

$$\operatorname{rank} A = 2, \quad \operatorname{rank} B = 3 \quad \Rightarrow \quad \dim W_1 = 5 - 2 = 3, \quad \dim W_2 = 5 - 3 = 2.$$

(Размерности определяются теоремой о строении множества решений однородной системы линейных алгебраических уравнений.)

Далее, в матрице A базисными являются, например, столбцы с номерами $1, 4$. Поэтому в качестве базисных можно выбрать неизвестные x_1, x_4; остальные неизвестные x_2, x_3, x_5 будут свободными:

$$\begin{array}{lll} x_2 = 1,\ x_3 = 0,\ x_5 = 0 & \Rightarrow & x_1 = -1,\ x_4 = 0; \\ x_2 = 0,\ x_3 = 1,\ x_5 = 0 & \Rightarrow & x_1 = -1,\ x_4 = 0; \\ x_2 = 0,\ x_3 = 0,\ x_5 = 1 & \Rightarrow & x_1 = 0,\ \ x_4 = -1. \end{array}$$

Таким образом,

$$W_1 = L(p_1, p_2, p_3), \quad \text{где} \quad p_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad p_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad p_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

В матрице B базисными являются, например, столбцы с номерами $1, 3, 5$. Неизвестные x_1, x_3, x_5 — базисные, а неизвестные x_2, x_4 — свободные:

$$\begin{array}{lll} x_2 = 1,\ x_4 = 0 & \Rightarrow & x_1 = -1,\ x_3 = 0,\ x_5 = 0; \\ x_2 = 0,\ x_4 = 1 & \Rightarrow & x_1 = -1/2,\ x_3 = -1/2,\ x_5 = 0. \end{array}$$

Таким образом,

$$W_2 = L(q_1, q_2), \quad \text{где} \quad q_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad q_3 = \begin{bmatrix} -1/2 \\ 0 \\ -1/2 \\ 1 \\ 0 \end{bmatrix}.$$

Далее, $W_1 + W_2 = L(p_1,\ p_2,\ p_3,\ q_1,\ q_2) = \operatorname{im} C$, где

$$C = [p_1,\ p_2,\ p_3,\ q_1,\ q_2] = \begin{bmatrix} -1 & -1 & 0 & -1 & -1/2 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1/2 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Простое вычисление показывает, что $\operatorname{rank} C = 4$, а базисными являются, например, столбцы с номерами $1, 2, 3, 5$. Поэтому векторы $p_1,\ p_2,\ p_3,\ q_2$ линейно независимы и

$$W_1 + W_2 = L(p_1,\ p_2,\ p_3,\ q_2), \qquad \dim(W_1 + W_2) = 4.$$

Наконец, в силу теоремы Грассмана, $\dim W_1 \bigcap W_2 = 3 + 2 - 4 = 1$. В данном случае можно заметить, что $p_1 = q_1 \quad \Rightarrow \quad p_1 \in W_1 \bigcap W_2$. Следовательно,

$$W_1 \bigcap W_2 = L(p_1).$$

Конечно, для поиска пересечения в данном случае можно также заметить, что вектор x принадлежит $W_1 \bigcap W_2$ тогда и только тогда, когда $Ax = 0$ и $Bx = 0$. Таким образом,

$$W_1 \bigcap W_2 = \ker A \bigcap \ker B = \left\{ x : \begin{bmatrix} A \\ B \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}.$$

В общем случае вычисление пересечения подпространств

$$W = L(a_1, \ldots, a_k) \bigcap L(b_1, \ldots, b_m)$$

сводится к решению однородной системы линейных алгебраических уравнений

$$x_1 a_1 + \ldots + x_k a_k \quad + \quad y_1 b_1 + \ldots + y_m b_m \;=\; 0 \qquad (*)$$

с неизвестными $x_1, \ldots, x_k, \;\; y_1, \ldots, y_m$. Из равенства $(*)$ ясно, что

$$v = x_1 a_1 + \ldots + x_k a_k = -(y_1 b_1 + \ldots + y_m b_m) \in W.$$

Пусть $r = \operatorname{rank}[a_1, \ldots, a_k, \; b_1, \ldots, b_m]$. Тогда фундаментальная система векторов для $(*)$ состоит из $k + m - r$ векторов вида

$$\begin{bmatrix} s_{11} \\ \cdots \\ s_{k\,1} \\ \cdots \end{bmatrix}, \;\; \ldots \;, \;\; \begin{bmatrix} s_{1\,k+m-r} \\ \cdots \\ s_{k\,k+m-r} \\ \cdots \end{bmatrix},$$

где компоненты $s_{1j}, \ldots, s_{kj}$ соответствуют неизвестным $x_1, \ldots, x_k$. После того, как фундаментальная система построена, получаем

$$W = L(v_1, \; \ldots, \; v_{k+m-r}), \qquad v_j \;=\; s_{1j}\, a_1 \;+\; \ldots \;+\; s_{kj}\, a_k, \quad j = 1, \ldots, k + m - r.$$

Предположим, что

$$\dim L(a_1, \ldots, a_k) = k, \quad L(b_1, \ldots, b_m) = m.$$

Тогда в силу теоремы Грассмана $\dim W = k + m - r$, поэтому векторы $v_1, \; \ldots, \; v_{k+m-r}$ будут линейно независимы.

Лекция 13

ОСНОВНАЯ ЧАСТЬ

13.1 Линейные многообразия

Пусть W — подпространство в линейном пространстве V и x — некоторый вектор из V. Множество векторов вида

$$M = \{v = x + w : \ w \in W\}$$

называется *линейным многообразием* в V. Обозначение: $M = x + W$.

Подпространство W называется *направляющим пространством* для M и определяется по множеству M однозначно. В самом деле, пусть

$$M = x_1 + W_1 = x_2 + W_2.$$

Отсюда $x_1 - x_2 \in W_1 \bigcap W_2$. Пусть $y \in W_1$. Тогда $y + (x_1 - x_2) \in W_2 \quad \Rightarrow \quad y \in W_2$. Аналогично, если $y \in W_2$, то $y + (x_2 - x_1) \in W_1 \quad \Rightarrow \quad y \in W_1$. $\square$

Для $M = x + W$ вектор x называется *вектором сдвига*. В качестве вектора сдвига можно взять любой вектор из M:

$$M = x + W = y + W \qquad \forall \ y \in M.$$

Действительно, пусть $y = x + w_0$ для какого-то $w_0 \in W$. Тогда, если $z = x + w$ при некотором $w \in W$, то $z = y + (x - y) + w = y + (-w_0 + w)$. Значит, $x + W \subset y + W$. Обратное включение доказывается аналогично. $\square$

Если W — конечномерное пространство, то его размерность называется также размерностью линейного многообразия $M = x + W$.

ПРИМЕР 1. Множество решений системы $Ax = b$ с $m \times n$-матрицей ранга r представляет собой линейное многообразие $v + W$, где v — частное решение данной системы и $W = \ker A$.

ПРИМЕР 2. Прямая на плоскости или в трехмерном пространстве — это линейное многообразие размерности 1. Плоскость в трехмерном пространстве — это линейное многообразие размерности 2.

При изучении линейных многообразий элементы векторного пространства обычно называют *точками*. По аналогии с геометрическим пространством, можно думать о векторах как о радиус-векторах, отложенных от общей начальной точки, отождествляемой с нулевым вектором.

13.2 Аффинные множества

Пусть $x \neq y$ — две точки в линейном пространстве V. Множество точек вида

$$l = \{z = x + t(y - x), \quad t \in \mathbb{R}\}$$

называется *прямой*, проходящей через точки x и y. Множество $M \subset V$ называется *аффинным множеством*, если вместе с любыми двумя точками $x \neq y$ оно содержит все точки проходящей через них прямой.

Утверждение. *Линейные многообразия и только они являются аффинными множествами.*

Доказательство. Пусть $M = x_0 + L$ — линейное многообразие с направляющим пространством L и сдвига x_0. Пусть $x = x_0 + u$, $u \in L$ и $y = x_0 + v$, $v \in L$. Тогда $v - u \in L$ и поэтому $x + t(y - x) = x_0 + t(v - u) \in L$ для любых t.

Теперь предположим, что M — аффинное множество. Зафиксируем точку $x_0 \in M$ и рассмотрим множество

$$L = \{z \in V : z = x - x_0, \ x \in M\}.$$

Докажем, что L — линейное подпространство. Во-первых, если $z \in L$, то

$$\alpha z = (x_0 + \alpha(x - x_0)) - x_0 \in L.$$

Во-вторых, если $z_1 = x_1 - x_0 \in L$ и $z_2 = x_2 - x_0 \in L$, то $z_1 + z_2 = 2z$, где $z = \dfrac{x_1 + x_2}{2} - x_0$. Остается заметить, что

$$\frac{x_1 + x_2}{2} = x_1 + \frac{x_2 - x_1}{2} \in M \quad \Rightarrow \quad z \in L \quad \Rightarrow \quad 2z \in L \quad \Rightarrow \quad z_1 + z_2 \in L. \quad \square$$

Любое множество точек S содержится, конечно, в некотором аффинном множестве (например, в V). Пусть M — пересечение всех таких аффинных множеств. Ясно, что M будет тоже аффинным множеством, причем наименьшим аффинным множеством, содержащим S. Оно называется *аффинной оболочкой* множества S.

13.3 Гиперплоскости

Пусть V — вещественное линейное пространство размерности n. Любое линейное многообразие $M = v_0 + L \subset V$ размерности $n - 1$ называется *гиперплоскостью*.

Поскольку V изоморфно $\mathbb{R}^n$, давайте считать, что $V = \mathbb{R}^n$, и рассмотрим уравнение относительно вещественных неизвестных $x_1, \ldots, x_n$

$$c_1 x_1 + \ldots + c_n x_n = b. \qquad (*)$$

где хотя бы одно из чисел c_i отлично от нуля.

Утверждение 1. *Множество всех векторов из R^n с координатами $x_1, \ldots, x_n$, удовлетворяющими уравнению $(*)$, есть гиперплоскость. Кроме того, любая гиперплоскость может быть задана как множество решений некоторого уравнения вида $(*)$.*

Доказательство. Уравнение $(*)$ — это частный случай системы линейных алгебраических уравнений, состоящей их одного уравнения. Матрица коэффициентов имеет

размеры $1 \times n$, и, поскольку не все c_i равны нулю, ее ранг равен 1. Очевидно, система совместна. Обозначим через $v_1, \ldots, v_{n-1}$ векторы фундаментальной системы решений, и пусть v_0 — произвольное частное решение. Тогда множество решений системы $(*)$ имеет вид $v_0 + L(v_1, \ldots, v_{n-1})$ и поэтому является гиперплоскостью.

Пусть $M = v_0 + L(v_1, \ldots, v_{n-1})$ — произвольная гиперплоскость. Образуем $n \times (n-1)$-матрицу $B = [v_1, \ldots, v_{n-1}]$ и рассмотрим уравнение

$$B^\top \begin{bmatrix} c_1 \\ \ldots \\ c_n \end{bmatrix} = 0.$$

Ранг матрицы коэффициентов равен $n-1 \quad \Rightarrow \quad$ система имеет нетривиальное решение $c = [c_1, \ldots, c_n]^\top$. Очевидно,

$$M = \{x = v_0 + Bz, \quad z \in \mathbb{R}^{n-1}\}.$$

Умножив обе части равенства $x = v_0 + Bz$ слева на матрицу-строку $c^\top$, находим

$$c^\top x = x^\top v_0 + (c^\top B)z = c^\top v_0 \quad \Rightarrow \quad c_1 x_1 + \ldots + c_n x_n = b, \quad b = c^\top v_0.$$

Остается заметить, что v_0 есть частное решение полученной системы, а столбцы матрицы B образуют фундаментальную систему решений для соответствующей однородной системы. $\square$

Утверждение 2. *Любое линейное многообразие размерности k является пересечением $n-k$ гиперплоскостей.*

Доказательство. Пусть данное многообразие имеет вид $M = v_0 + L(v_1, \ldots, v_k)$. Тогда $x \in M$ есть вектор вида $x = v_0 + Bz$, где $B = [v_1, \ldots, v_k]$, $z \in \mathbb{R}^k$. Рассмотрим уравнение

$$B^\top y = 0.$$

Поскольку $\mathrm{rank} B = k$, фундаментальная система решений содержит $n-k$ векторов. Обозначим их через $a_1, \ldots, a_{n-1}$. Далее,

$$a_i^\top x = a_i^\top v_0 + (a_i^\top B)z = a_i^\top v_0.$$

Следовательно, x принадлежит пересечению гиперплоскостей

$$a_i^\top x = b_i, \quad b_i = a_i^\top v_0, \quad 1 \le i \le n-k.$$

В то же время, пересечение этих гиперплоскостей есть линейное многообразие той же размерности. $\square$

Заметим, что системы гиперплоскостей, дающие в пересечении M, можно выбрать многими способами. Из доказательства видно, что их столько, сколько имеется фундаментальных систем решений уравнения $B^\top y = 0$.

13.4 Полупространства

Любая гиперплоскость $\pi : c_1 x_1 + \ldots + c_n x_n = b \;\; (c^\top x = b)$ выделяет в $\mathbb{R}^n$ два подмножества:

$$\pi_- = \{x : c^\top x \le b\}, \quad \pi_+ = \{x : c^\top x \ge b\}, \qquad \pi_- \cap \pi_+ = \pi.$$

Эти подмножества называются (“отрицательным” и “положительным”) *полупространствами*. В случае плоскости в трехмерном пространстве они уже изучались в разделе 7.12.

Утверждение. *Точки $x, y \notin \pi$ принадлежат разным полупространствам тогда и только тогда, когда $x + t(y - x) \in \pi$ при некотором $0 < t < 1$.*

Доказательство. Пусть для определенности $x \in \pi_-$ и $y \in \pi_+$. Тогда уравнение

$$c^\top(x + t(y - x)) = b$$

имеет решение

$$t = \frac{b - c^\top x}{c^\top(y - x)} = \frac{b - c^\top x}{(b - c^\top x) - (b - c^\top y)},$$

причем с очевидностью $0 < t < 1$. Если же $x, y \in \pi_-$ (π_+), то при любом $0 \le t \le 1$ находим $x + t(y - c) \in \pi_-$ (π_+). $\square$

13.5 Выпуклые множества

Пусть V — линейное пространство и $x, y \in V$. Множество точек вида $x + t(y - x) = (1 - t)x + ty$, $0 \le t \le 1$, называется *отрезком*, соединяющим x и y. Множество $M \subset V$ называется *выпуклым*, если вместе с любыми двумя точками оно содержит все точки соединяющего их отрезка. Точки, получаемые при $0 < t < 1$, называются *внутренними точками* отрезка.

Любые полупространства в $\mathbb{R}^n$ — выпуклые множества. То же верно и для любого пересечения конечного числа полупространств. Это следствие более общего и очевидного факта: *пересечение любого конечного числа выпуклых множеств является выпуклым множеством.*

В определенном смысле двойственный подход к построению выпуклых множеств такой. Пусть $v_1, \ldots, v_k \in V$. Тогда вектор

$$v = t_1 v_1 + \ldots + t_k v_k, \quad t_i \ge 0, \quad t_1 + \ldots + t_k = 1,$$

называется *выпуклой комбинацией* векторов $v_1, \ldots, v_k$. Множество всех возможных выпуклых комбинаций заданных векторов называется их *выпуклой оболочкой.*

Утверждение 1. *Выпуклая оболочка векторов является выпуклым множеством.*

Доказательство. Пусть $x = \alpha_1 v_1 + \ldots + \alpha_k v_k$ и $y = \beta_1 v_1 + \ldots \beta_k v_k$. Тогда при $0 \le t \le 1$ получаем

$$(1 - t)x + ty = \sum_{i=1}^{k} ((1 - t)\alpha_i + t\beta_i)v_i.$$

Если $\sum \alpha_i = \sum \beta_i = 1$ и $\alpha_i,\ \beta_i \ge 0$, то, очевидно,

$$\sum_{i=1}^{k} ((1 - t)\alpha_i + t\beta_i) = 1, \quad (1 - t)\alpha_i + t\beta_i \ge 0 \qquad \square$$

Например, в трехмерном пространстве выпуклая оболочка трех точек, не лежащих

на одной прямой, представляет собой треугольник с вершинами в этих точках. Выпуклая оболочка четырех точек, не лежащих в одной плоскости, есть тетраэдр.

Утверждение 2. *Пусть M — выпуклое множество. Тогда вместе с любой системой точек M содержит целиком и их выпуклую оболочку.*

Доказательство. Если $t_1 > 0$, то

$$\sum_{i=1}^{k} t_i v_i = t_1 v_1 + (1 - t_1) \left(\sum_{i=2}^{k} \frac{t_i}{1 - t_1} v_i \right), \quad \sum_{i=2}^{k} \frac{t_i}{1 - t_1} = 1.$$

Далее проводим индукцию по числу точек k. $\square$

Любое (в том числе и бесконечное) множество точек S содержится в некотором выпуклом множестве (достаточно учесть, что любое аффинное множество является выпуклым). Пересечение всех таких множеств будет наименьшим выпуклым множеством, содержащим S. Оно называется *выпуклой оболочкой множества* S. Легко видеть, что если S — конечная система точек, то ее выпуклая оболочка совпадает с выпуклой оболочкой множества S.

13.6 Аффинная независимость

Точки v_0, v_1, …, v_k в n-мерном пространстве называются *аффинно независимыми*, если векторы $v_1 - v_0$, …, $v_k - v_0$ линейно независимы. Равносильное "симметричное" определение: векторы v_0, …, v_k аффинно независимы, если из равенств $\alpha_0 v_0 + \ldots + \alpha_k v_k = 0$, $\alpha_0 + \ldots + \alpha_k = 0$ вытекает, что $\alpha_0 = \ldots = \alpha_k = 0$. В самом деле, из этих равенств находим $\alpha_1(v_1 - v_0) + \ldots + \alpha_k(v_k - v_0) = 0$, при этом необходимость условий $\alpha_1 = \ldots = \alpha_k = 0$ равносильна линейной независимости векторов $v_1 - v_0$, …, $v_k - v_0$.

Задача. Докажите, что в любой аффинно независимой системе с числом векторов $k+1$ можно выбрать линейно независимую подсистему с числом векторов k.

Выпуклая оболочка аффинно независимых векторов v_0, v_1, …, v_k называется *симплексом размерности* k. Точки v_0, …, v_k называются *вершинами* симплекса. Согласно определению, размерность симплекса не зависит от размерности пространства V. Размерностью произвольного выпуклого множества называют максимальную размерность принадлежащих ему симплексов.

Среди точек в выпуклом множестве M особый интерес представляют его *угловые точки* — так называются точки из M, не являющиеся внутренней точкой ни для одного отрезка, лежащего в M. Например, круг на плоскости является выпуклым множеством, а его угловые точки — это точки граничной окружности.

Утверждение. *Угловыми точками симплекса являются его вершины и только они.*

Доказательство. Пусть v_0, …, v_k — вершины заданного симплекса M. Докажем, что v_j является угловой точкой. От противного: пусть $v_j = tx + (1-t)y$ при $0 < t < 1$ и $x \neq y$:

$$x = \sum_{i=0}^{k} \alpha_i v_i, \quad y = \sum_{i=0}^{k} \beta_i v_i, \qquad \sum_{i=0}^{k} \alpha_i = \sum_{i=0}^{k} \beta_i = 1, \quad \alpha_i,\ \beta_i \geq 0.$$

Отсюда

$$\sum_{\substack{i=1 \\ i \neq j}} (t\alpha_i + (1-t)\beta_i)(v_i - v_j) = 0 \ \Rightarrow \ t\alpha_i + (1-t)\beta_i = 0 \ \Rightarrow \ \alpha_i = \beta_i = 0.$$

Итак, $x = y$, а мы исходили из того, что $x \neq y$.

Пусть теперь $x \in M$ — произвольная точка симплекса, отличная от его вершин. Это значит, что

$$x = \sum_{i=0}^{k} t_i v_i$$

и $0 < t_j < 1$ хотя бы для одного j. Не ограничивая общности, предположим, что $0 < t_0 < 1$. Тогда

$$x = t_0 v_0 + (1 - t_0) w, \quad \text{где} \quad w = \sum_{i=1}^{k} (t_i/(1 - t_0)) v_i \in M. \qquad \square$$

В действительности для широкого класса выпуклых множеств имеет место элегантный и глубокий факт, к доказательству которого мы пока не готовы: любая точка в них является выпуклой комбинацией конечного числа угловых точек. [1]

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

13.7 Линейные неравенства и минимизация

Большое число прикладных задач (составление расписаний, управление производством, оптимизация диеты, портфеля инвестиций и т. п.) связано с минимизацией (максимизацией) вещественной функции $f(x)$ от $x = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$ вида

$$f(x) = c^\top x = c_1 x_1 + \ldots + c_n x_n, \quad c_i \in \mathbb{R}, \quad c = [c_1, \ldots, c_n]^\top \neq 0,$$

на множестве точек M, заданном линейными неравенствами

$$\begin{aligned} a_{11}\, x_1 + \ldots + a_{1n} x_n &\leq b_1, \\ &\ldots \\ a_{m1} x_1 + \ldots + a_{mn} x_n &\leq b_m. \end{aligned}$$

Ясно, что M есть пересечение конечного числа полупространств. Предположим дополнительно, что координаты точек из M ограничены. В таких случаях M называют выпуклым многогранником. Интуитивно понятно, что можно говорить о гранях — детали важны, но это предмет отдельного курса.

Уравнение $f(x) = b$ при любом фиксированном b определяет гиперплоскость. Очевидно, $f(x + tc) > f(x)$ при $t > 0$. Более того, $f(x + td) > f(x)$ при $t > 0$, если $c^\top d > 0$ (докажите!). Отсюда можно вывести, что минимум $f(x)$ должен достигаться в *угловых точках* множества M (возможно, не только в них). Простая геометрическая идея поиска минимума заключается в переборе всех угловых точек. Конечно, его можно организовать так, чтобы следующая угловая точка лежала в той же грани и уменьшала значение $f(x)$. Формализация данной идеи привела в свое время к так называемому *симплекс-методу*. До сих пор это один из основных методов решения задач с линейными ограничениями и линейной целевой функцией $f(x)$ — такие задачи относятся к задачам *линейного программирования*. Другой эффективный класс методов использует внутренние точки множества M и получил общее название *методов внутренней точки*. Конечно, весь этот круг вопросов составляет отдельную и обширную область с развитым математическим аппаратом и многочисленными приложениями.

[1] Достаточно потребовать, чтобы выпуклое множество было ограниченным и замкнутым — строгие определения будут в лекциях второго семестра.

Лекция 14

ОСНОВНАЯ ЧАСТЬ

14.1 Комплексные числа

Как известно, квадратное уравнение с вещественными коэффициентами может не иметь вещественных решений. Формально положение легко поправить, введя для обозначения несуществующих решений некие "абстрактные числа". Но одних обозначений, конечно, мало. Важно определить операции сложения и умножения для новых чисел таким образом, чтобы остались в силе привычные свойства этих операций над вещественными числами.

В качестве "абстрактных чисел" рассмотрим 2×2-матрицы специального вида

$$z = z(a,b) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \qquad a,\ b \in \mathbb{R}. \tag{$*$}$$

Обозначим через $\mathbb{C}$ множество всех таких матриц. Операции сложения и умножения "абстрактных чисел" определим как соответствующие операции над матрицами. Элементарно проверяются, что они обладают следующими свойствами.

(1) *Если* $u,\ v \in \mathbb{C}$, *то* $u+v \in \mathbb{C}$ и $uv \in \mathbb{C}$.

(2) *Любая ненулевая матрица* $z = z(a,b) \in \mathbb{C}$ *обратима, а соответствующая обратная матрица имеет вид*

$$z^{-1} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix}, \qquad c = \frac{a}{\sqrt{a^2+b^2}}, \quad d = \frac{-b}{\sqrt{a^2+b^2}}.$$

(3) *Множество* $\mathbb{C}$ *относительно операции сложения матриц является абелевой группой.*

(4) *Множество* $\mathbb{C}\backslash\{0\}$ *относительно операции умножения матриц является абелевой группой.*

(5) *Имеет место дистрибутивность:* $z(u+v) = zu + zv \quad \forall \quad u,v,z \in \mathbb{C}$.

Если в утверждениях (3)-(5) заменить $\mathbb{C}$ на $\mathbb{R}$, то получатся основные свойства операций над вещественными числами. Поэтому элементы множества $\mathbb{C}$ логично рассматривать как числа. Это и будут так называемые *комплексные числа.*

Вещественные числа a и b называются соответственно *вещественной* и *мнимой* частью комплексного числа $z = z(a,b)$ Рассмотрим две особые матрицы вида $(*)$:

$$\mathbf{e} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathbf{i} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Легко видеть, что

$$z = z(a,b) = a\mathbf{e} + b\mathbf{i}, \qquad a,b \in \mathbb{R}. \qquad (**)$$

Матрица $\mathbf{e}$ выполняет роль единичного элемента относительно операции умножения. Матрицу вида $a\mathbf{e}$ естественно отождествить с вещественным числом a. Тогда $\mathbf{e} = 1 \cdot \mathbf{e}$ отождествится с числом 1, а соотношение $(**)$ примет вид

$$z = a + b\mathbf{i},$$

и при этом, как легко проверить,

$$\mathbf{i}^2 = -1 \qquad (-1 \text{ отождествляется с матрицей } -\mathbf{e}).$$

Несложно проверить, что уравнение $z^2 = -1$ имеет на множестве $\mathbb{C}$ в точности два решения $z = \pm\mathbf{i}$. Отсюда можно вывести, что любое квадратное уравнение с вещественными коэффициентами имеет два (иногда совпадающих) решения из $\mathbb{C}$. Мы скоро увидим, что то же верно и для квадратных уравнений с комплексными коэффициентами.

Конечно, комплексные числа можно было бы ввести без использования матриц — сказав, что это пары (a,b) вещественных чисел, для которых операции определяются правилами

$$(a,b) + (c,d) = (a+c, b+d), \quad (a,b)(c,d) = (ac - db, ad + bc).$$

Придется изменить лишь некоторые детали доказательства свойств (3)-(5).

Наш интерес к использованию матриц вида $(*)$ объясняется тем, что они позволили нам получить искомые "абстрактные числа" как уже знакомые объекты с хорошо изученными свойствами.

14.2 Комплексная плоскость

Рассмотрим плоскость с декартовой системой координат. Пусть (a,b) — точка (радиус-вектор) с координатами a, b. Очевидно, $(a,b) \leftrightarrow z = a + b\mathbf{i}$ есть взаимно-однозначное соответствие между точками (радиус-векторами) плоскости и комплексными числами. Плоскость, точки (радиус-векторы) которой используются для изображения комплексных чисел, называется *комплексной плоскостью*.

Рассмотрим комплексное число $z = a + a\mathbf{i}$. Длина отвечающего ему радиус-вектора, равная $\sqrt{a^2 + b^2}$, называется *модулем* комплексного числа z и обозначается $|z|$. Угол ϕ между радиус-вектором для $z \neq 0$ и положительным направлением первой оси (оси абсцисс), называется *аргументом* комплексного числа z. Обозначение: $\phi = \arg z$. Конечно, аргумент определен с точностью до слагаемого, кратного 2π.

Числу $z = 0$ можно приписать любое значение аргумента.

Очевидно,

$$z = |z|\,(\cos\phi + \mathbf{i}\sin\phi), \qquad \phi = \arg z.$$

Такая форма представления комплексного числа называется его *тригонометрической формой*.

Заметим, что сумме комплексных чисел соответствует сумма соответствующих радиус-векторов. Отсюда получаем очень полезное неравенство (неравенство треугольника):

$$|u + v| \leq |u| + |v| \qquad \forall \ u, v \in \mathbb{C}.$$

При умножении z на комплексное число

$$w \;=\; |w|\,(\cos\psi \;+\; \mathbf{i}\,\sin\psi), \qquad \psi = \arg w,$$

получается

$$\begin{aligned} zw &= |z|\,|w|\,(\cos\phi \;+\; \mathbf{i}\,\sin\psi)\,(\cos\psi \;+\; \mathbf{i}\,\sin\psi) \\ &= |z|\,|w|\,((\cos\phi\,\cos\psi \;-\; \sin\phi\,\sin\psi) \;+\; \mathbf{i}\,(\cos\phi\,\sin\psi \;+\; \sin\phi\,\cos\psi)) \\ &= |z|\,|w|\,(\cos(\phi+\psi) \;+\; \mathbf{i}\,\sin(\phi+\psi)). \end{aligned}$$

Таким образом, *при умножении комплексных чисел модули перемножаются, а аргументы складываются.*

Комплексное число $a - b\mathbf{i}$ называется *сопряженным* к $z = a + b\mathbf{i}$. Обозначение: $\bar{z} = a - b\mathbf{i}$. На комплексной плоскости радиус-вектор для $\bar{z}$ получается из радиус-вектора для z симметричным отражением относительно первой оси. Заметим также, что $\bar{z}\,z = |z|^2$.

14.3 Преобразования плоскости

С помощью комплекных чисел можно задавать взимно-однозначные отображения плоскости на себя. Например, фиксируем $w \in \mathbb{C}$ и рассмотрим отображение $z \;\to\; z + w$. Это параллельный перенос (сдвиг) точек на вектор, заданный комплексным числом w.

Далее, рассмотрим отображение $z \;\to\; wz$ в предположении, что $|w| = 1$. В силу того, что $|w| = 1$, находим $|wz| = |z|$. При этом радиус-вектор для wz получается поворотом радиус-вектора для z на угол $\phi = \arg w$. Таким образом, умножение комплексных чисел на фиксированное комплексное число w с модулем 1 задает поворот на угол, равный аргументу числа w.

Умножение на вещественное число $\rho > 0$ задает *гомотетию* — каждый радиус-вектор умножается на ρ (растягивается в ρ раз).

Поскольку в случае $w \neq 0$ можно записать $w = |w|\,\widetilde{w}$, где $\widetilde{w} = w/|w|$ и, следовательно, $|\widetilde{w}| = 1$, умножение на произвольное комплексное число $w \neq 0$ сводится к композиции (последовательному выполнению) двух отображений: поворота и гомотетии.

Преобразование вида $z \;\to\; \bar{z}$ также является взаимно-однозначным. Это симметричное отражение относительно первой оси. Но оно уже не представимо в виде композиции поворотов, гомотетий и параллельных переносов. Сказанное означает, что ни для каких комплексных чисел a, b нельзя получить равенство $\bar{z} = a + bz$, верное для всех $z \in \mathbb{C}$. Докажите!

Утверждение. *Множество $\mathcal{T}$ отображений комплексной плоскости вида*

$$\Phi(z) = a + bz \quad \textit{или} \quad \bar{\Phi}(z) = a + b\bar{z}, \quad \textit{где} \quad a, b \in \mathbb{C}, \quad |b| = 1,$$

образует группу относительно композиции отображений.

Доказательство. Композиция отображений $\Phi\Psi$ определяется следующим правилом: $(\Phi\Psi)(z) = \Phi(\Psi(z))$. Пусть $\Phi(z) = a + bz$ и $\Psi(z) = c + dz$ принадлежат $\mathcal{T}$. Это означает, что $|b| = |d| = 1$. Тогда

$$\Phi(\Psi(z)) = a + b(c + dz) = (c + bc) + (bd)z.$$

Поскольку $|bd| = |b||d| = 1$, данное отображение также принадлежит $\mathcal{T}$. Роль единичного элемента выполняет тождественное отображение $z \;\to\; z$, которое, очевидно,

принадлежит $\mathcal{T}$. Далее, если $w = a+bz$, то $z = a-\bar{b}w$. Поскольку $|-\bar{b}| = 1$, отображение, обратное к Φ, также принадлежит $\mathcal{T}$.

Теперь заменим Φ на $\bar{\Phi}$ или Ψ на $\bar{\Psi}$. Композиция таких отображений и обратные к ним также принадлежат $\mathcal{T}$ — для проверки нужны выкладки, аналогичные предыдущим. $\square$

Взаимно-однозначное отображение плоскости $z \to \Phi(z)$ называется *движением*, если оно сохраняет расстояние между точками: $|\Phi(z_1) - \Phi(z_2)| = |z_1 - z_2| \quad \forall\ z_1, z_2 \in \mathbb{C}$.

Из наших предыдущих обсуждений понятно, что любое отображение из $\mathcal{T}$ является композицией параллельных переносов, поворотов и симметричных отражений. Каждое из данных отображений специального вида является движением. Поэтому любое отображение из $\mathcal{T}$ есть движение. Верно и обратное — это весьма примечательный факт, дающий полное описание всех мыслимых движений (и требующий более обстоятельного доказательства, на котором мы не будем останавливаться).

Пример более сложного отображения: $z \to 1/z$. Оно не определено при $z = 0$, но является взаимно-однозначным на множестве $\mathbb{C}\backslash\{0\}$. Часто к комплексной плоскости добавляется абстрактная *бесконечно удаленная* точка ∞, в результате чего появляется *расширенная* комплексная плоскость $\overline{\mathbb{C}} = \mathbb{C}\bigcup\{\infty\}$. Тогда отображение $z \to 1/z$ можно превратить во взаимно-однозначное отображение на $\overline{\mathbb{C}}$, приняв соглашение о том, что 0 переходит в ∞, а ∞ переходит в 0. Отображение $z \to 1/z$ представляет собой частный случай так называемых *дробно-линейных* отображений вида

$$z \to \Phi(z) = \frac{a + bz}{c + dz},$$

где a, b, c, d — фиксированные комплексные числа, причем предполагается, что $\Phi(z)$ не является тождественной константой: $ad - bc \neq 0$.

Если $d = 0$, то дробно-линейное отображение сводится к рассмотренному выше. Предположим, что $d \neq 0$. Тогда $\Phi(z)$ не определено при $z = -c/d$. Если условиться, что $\Phi(-c/d) = \infty$ и $\Phi(\infty) = -c/d$, то Φ будет взаимно-однозначным отображением на расширенной комплексной плоскости. Дробно-линейные отображения обладают рядом замечательных геометрических свойств (например, они переводят окружности и прямые в окружности или прямые) и играют важную роль в теории функций комплексного переменного.

14.4 Корни из единицы

Комплексное число z называется корнем из единицы степени n, если $z^n = 1$.

Формула Муавра. *Если* $z = |z|\,(\cos\phi\ +\ \mathbf{i}\sin\phi)$, *то*

$$z^n = |z|^n\,(\cos(n\phi)\ +\ \mathbf{i}\sin(n\phi)).$$

Доказательство. Достаточно учесть, что при умножении комплексных чисел модули перемножаются, а аргументы складываются. $\square$

Следствие. *Существует ровно* n *различных корней из единицы степени* n. *Это комплексные числа вида*

$$z_k = \cos\left(\frac{2\pi\,k}{n}\right)\ +\ \mathbf{i}\sin\left(\frac{2\pi\,k}{n}\right), \qquad k = 0,\,1,\,\ldots,\,n-1.$$

Доказательство. Пусть $z = |z|\,(\cos\phi \;+\; \mathbf{i}\sin\phi)$ есть корень из единицы степени n. Тогда, согласно формуле Муавра, $|z| = 1$ и $\cos(n\phi) = 1 \;\;(\Rightarrow \;\; \sin(n\phi) = 0)$. Отсюда

$$\phi = \frac{2\pi\,k}{n}, \quad k = 0,\, \pm 1,\, \pm 2,\, \ldots\,.$$

Следовательно, при любом целом k комплексное число вида

$$z_k = \cos\left(\frac{2\pi\,k}{n}\right) \;+\; \mathbf{i}\,\sin\left(\frac{2\pi\,k}{n}\right)$$

является корнем из единицы степени n. В силу периодичности синуса и косинуса очевидно, что $z_k = z_l$, если

$$l \;=\; k \,+\, m\,n, \quad m = 0,\, \pm 1,\, \pm 2,\, \ldots\,.$$

Если же $0 \le k, l \le n-1$, то равенство $z_k = z_l$ означает, что

$$\cos\left(\frac{2\pi\,k}{n}\right) = \cos\left(\frac{2\pi\,l}{n}\right) \quad \Rightarrow \quad k = l \;\text{ или }\; k = n - l.$$

Кроме этого, должно выполняться равенство

$$\sin\left(\frac{2\pi\,k}{n}\right) = \sin\left(\frac{2\pi\,l}{n}\right).$$

При $k = n - l$ получаем

$$\sin\left(\frac{2\pi\,k}{n}\right) = -\sin\left(\frac{2\pi\,k}{n}\right) \quad \Rightarrow \quad n \text{ четно и } k = n/2 \quad \Rightarrow \quad k = l.$$

Итак, в любом случае из равенства $z_k = z_l$ при $0 \le k, l \le n-1$ следует, что $k = l$. Таким образом, среди чисел $z_0,\, \ldots,\, z_{n-1}$ нет одинаковых. $\square$

Замечание. Корни из единицы степени n располагаются на единичной окружности $|z| = 1$ в вершинах правильного n-угольника.

14.5 Группа корней из единицы степени n

Введем обозначение $\mathcal{K}_n$ для множества корней из единицы степени n. Мы только что доказали, что $\mathcal{K}_n$ содержит ровно n комплексных чисел.

Множество $\mathcal{K}_n$ является, как легко видеть, группой относительно операции умножения комплексных чисел. Более того, $\mathcal{K}_n$ *является циклической группой*. В самом деле, $z_k = z_1^k$.

Корень z_m называется *первообразным корнем* из единицы степени n, если его степени дают все множество $\mathcal{K}_n$.

Предположим, что z_m — первообразный корень. Тогда равенство $z_m^p = 1$ в случае $0 < p \le n$ влечет за собой равенство $p = n$ (если $z_1^p = 1$, то степени числа z_1 не могут породить более, чем p чисел).

Утверждение 1. *Корень из единицы $z_m \in \mathcal{K}_n$ при $m \ge 1$ является первообразным тогда и только тогда, когда числа m и n взаимно просты (наибольший общий делитель этих чисел равен 1).*

Доказательство. Предположим, что z_m является первообразным корнем, но числа m и n все же имеют наибольший общий делитель $d > 1$: $n = dp$ и $m = dq$ при целых

$p,\ q,\ d > 1$. Тогда $z_m^p = z_1^{mp} = z_1^{d\,qp} = z_1^{dn} = 1$ при $0 < p < n \quad \Rightarrow$ степени числа z_m не могут породить более, чем $p < n$ чисел $\quad \Rightarrow$ корень z_m не может быть первообразным.

Пусть теперь m и n взаимно просты. Докажем, что z_m является первообразным корнем. Для этого достаточно установить, что если $z_m^k = z_m^l$ при $0 \le k, l \le n-1$, то $k = l$. В самом деле, $m(k-l)$ должно нацело делиться на n. Поскольку m и n взаимно просты, $k - l$ должно делиться на $n \quad \Rightarrow \quad k = l$. $\square$

Утверждение 2. *Сумма всех корней из единицы степени n равна нулю.*

Доказательство. Поскольку $z_k = z_1^k$, требуется найти сумму членов геометрической прогрессии:

$$\sum_{k=0}^{n-1} z_k \;=\; \sum_{k=0}^{n-1} z_1^k \;=\; \frac{z_1^n - 1}{z_1 - 1} \;=\; 0. \qquad \square$$

14.6 Матрицы с комплексными элементами

Множество матриц размеров $m \times n$ с комплексными элементами обозначается $\mathbb{C}^{m\times n}$. Если $A = [a_{ij}] \in \mathbb{C}^{m\times n}$, то матрица тех же размеров с заменой элементов на комплексно сопряженные к ним часто обозначается через $\bar{A} = [\bar{a}_{ij}]$.

Матрица $\bar{A}^\top$ называется *сопряженной* к A матрицей. Обозначение: $A^* = \bar{A}^\top$.

Отметим некоторые свойства сопряженных матриц:

- $(AB)^* = B^* A^*$;
- $\det A^* \;=\; \overline{\det A}$;
- матрица A обратима тогда и только тогда, когда обратима сопряженная матрица A^*, при этом $(A^*)^{-1} = (A^{-1})^*$.

14.7 Квадратные уравнения

Рассмотрим произвольное квадратное уравнение $z^2 + az + b = 0$ с комплексными коэффициентами a, b. После выполнения традиционных преобразований

$$z^2 + az + b \;=\; \left(z^2 \;+\; 2\,\frac{a}{2}\,z \;+\; \left(\frac{a}{2}\right)^2\right) \;+\; \left(b - \left(\frac{a}{2}\right)^2\right) \;=\; \left(z + \frac{a}{2}\right)^2 \;+\; \left(b - \frac{a^2}{4}\right)$$

получаем равносильное уравнение

$$\left(z \;+\; \frac{a}{2}\right)^2 \;=\; D, \qquad D \;\equiv\; \frac{a^2}{4} \;-\; b.$$

Если $D = 0$, то единственное решение имеет вид

$$z \;=\; \frac{a}{2}.$$

Если D вещественное и $D > 0$, то получаем пару вещественных решений

$$z_\pm \;=\; -\frac{a}{2} \;\pm\; \sqrt{D},$$

а при $D < 0$ получаем пару комплексных решений

$$z_\pm \;=\; -\frac{a}{2} \;\pm\; \mathbf{i}\sqrt{|D|}.$$

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

14.8 Кубические уравнения

Произвольное кубическое уравнение

$$z^3 + a_2 z^2 + a_1 z + a_0 = 0$$

с помощью замены $z = x - a_2/3$ приводится к виду

$$x^3 + px + q = 0. \qquad (*)$$

Будем искать x в виде $x = u + v$. Тогда

$$u^3 + 3u^2 + 3uv^2 + v^3 + p(u+v) + q \;=\; (u^3 + v^3 + q) + (3uv + p)(u+v) = 0.$$

Очевидно, $x = u + v$ будет решением уравнения $(*)$, если

$$\begin{cases} u^3 + v^3 &= -q, \\ uv &= -p/3. \end{cases} \qquad \Rightarrow \qquad \begin{cases} u^3 + v^3 &= -q, \\ u^3 v^3 &= -p^3/27. \end{cases}$$

Два комплексных числа u^3 и v^3 с заданной суммой и заданным произведением находятся как корни квадратного уравнения

$$w^2 + qw - \frac{p^3}{27} = 0 \qquad \Rightarrow \qquad \begin{aligned} w_1 &= u^3 = -q/2 + \sqrt{q^2/4 + p^3/27}, \\ w_2 &= v^3 = -q/2 - \sqrt{q^2/4 + p^3/27}. \end{aligned}$$

В результате получается следующая *формула Кардано*: [1]

$$x = \sqrt[3]{-q/2 + \sqrt{q^2/4 + p^3/27}} \;+\; \sqrt[3]{-q/2 - \sqrt{q^2/4 + p^3/27}}.$$

При применении формулы Кардано следует иметь в виду, что для каждого из кубических корней существуют три комплексных значения, которые нельзя выбирать независимо: их произведение uv должно быть равно $-p/3$. Даже в случае вещественных корней формула Кардано, как правило, дает их представление с использованием комплексных значений кубических корней.

[1] Это тот самый Кардано, который известен автомобилистам как изобретатель способа передачи вращения с одного вала на другой. Данная формула опубликована им в 16-м веке, но известно, что она была открыта другими итальянскими математиками. Ученики Кардано нашли также способ решения уравнений 4-й степени.

14.9 Уравнения четвертой степени

Общее уравнение четвертой степени

$$z^4 + a_3 z^3 + a_2 z^2 + a_1 z + a_0 = 0$$

с помощью замены $z = x - a_3/4$ приводится к виду

$$x^4 + px^2 + qx + r = 0. \qquad (*)$$

Данное уравнение может быть сведено к кубическому. Наиболее простой способ для этого был предложен итальянским математиком Феррари. Идея состоит в том, чтобы представить левую часть уравнения $(*)$ как разность двух квадратов:

$$x^4 + px^2 + qx + r \;=\; (x^2 + y/2)^2 - ((y-p)x^2 - qx + (y^2/4 - r)).$$

Квадратный трехчлен $ax^2 + bx + c$ является квадратом двучлена $\alpha x + \beta$ в том и только том случае, когда его дискриминант равен нулю. Поэтому потребуем, чтобы y был решением кубического уравнения

$$q^2 - 4(y-p)(y^2/4 - r) \;=\; 0.$$

Тогда для некоторых $\alpha,\ \beta$

$$x^4 + px^2 + qx + r \;=\; (x^2 + y/2)^2 - (\alpha x + \beta)^2 \;=\; (x^2 + y/2 + \alpha x + \beta)(x^2 + y/2 - \alpha x - \beta).$$

Таким образом, получение решений для $(*)$ сводится к решению одного кубического и нескольких квадратных уравнений.

В начале 19-го века Руффини и Абель независимо друг от друга доказали, что для общего алгебраического уравнения n-й степени при $n \geq 5$ формулы, выражающей корни через радикалы, не существует. В 1830 г. Эварист Галуа создал теорию, позволяющую выяснить разрешимость или неразрешимость в радикалах любого конкретного уравнения n-й степени (см. дополнительную часть Лекции 18).

Лекция 15

ОСНОВНАЯ ЧАСТЬ

15.1 Кольца и поля

В процессе развития математики постоянно находились причины для того, чтобы вводить более общие понятия числа. Общеизвестна, по крайней мере, такая цепочка расширений (на прошлой лекции мы как раз завершили построение ее последнего звена):

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C},$$

$\mathbb{N} = \{1, 2, \dots\}$ — натуральные числа; [1]

$\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ — целые числа,

$\mathbb{Q} = \{p/q,\ p \in \mathbb{Z},\ q \in \mathbf{N}\}$ — рациональные числа,

$\mathbb{R}$ — вещественные числа,

$\mathbb{C}$ — комплексные числа.

Множество целых чисел $\mathbb{Z}$ послужило прототипом для введения понятия *кольца*, а множества $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ — для понятия *поля*.

Пусть на непустом множестве K действуют две алгебраические операции: сложение (обозначаемое знаком $+$) и умножение (обозначаемое точкой или "пустым местом"), и пусть эти операции обладают следующими свойствами:

- множество K относительно операции сложения является абелевой группой;
- выполняются законы дистрибутивности:

$$a(b+c) = ab + ac, \quad (b+c)a = ba + ca \qquad \forall\ \ a, b, c \in K;$$

- операция умножения ассоциативна.

В таких случаях множество K называется (ассоциативным) *кольцом*. (В некоторых книгах по алгебре в определение кольца ассоциативность умножения не включается.)

Единичный элемент относительно операции сложения в кольце называется *нулевым* и обозначается символом 0. Элемент, обратный относительно сложения к элементу a, называется *противоположным* к a и обозначается $-a$.

Утверждение 1. $0 \cdot a = a \cdot 0 = 0 \quad \forall\ \ a \in K.$

[1] По словам Кронекера, "Бог создал натуральные числа, все остальное придумал человек".

Доказательство. Пусть $b = -(0 \cdot a)$ (элемент, противоположный к $0 \cdot a$). В силу дистрибутивности, $0 \cdot a = (0 + 0) \cdot a = (0 \cdot a) + (0 \cdot a)$. Прибавим b к обеим частям: $0 = b + (0 \cdot a) = (b + 0 \cdot a) + (0 \cdot a) = 0 + (0 \cdot a) = 0 \cdot a$. $\square$

Если умножение коммутативно, то K называется *коммутативным кольцом*. Если существует единичный элемент относительно операции умножения, то кольцо называется *кольцом с единицей*.

Пусть P — коммутативное кольцо с единицей, для которого множество $P\backslash\{0\}$ относительно операции умножения является абелевой группой. В таких случаях множество P называется *полем*.

Группа $P\backslash\{0\}$ по умножению называется *мультипликативной группой* поля P.

Единичный элемент кольца с единицей или поля относительно операции умножения обозначается обычно символом 1.

Утверждение 2. Если K — кольцо с единицей, то $(-1) \cdot a = a \cdot (-1) = -a \quad \forall\ a \in K$.

Доказательство. В силу дистрибутивности и утверждения 1, $0 = (1 + (-1) \cdot a = 1 \cdot a + (-1) \cdot a = a + (-1) \cdot a$. $\square$

Задача. Пусть a и b — элементы кольца с единицей e. Докажите, что из обратимости элемента $e - ab$ в данном кольце вытекает обратимость элемента $e - ba$.

15.2 Делители нуля

В некоторых кольцах существуют ненулевые элементы a, b такие, что $ab = 0$. Такие элементы a, b называются *делителями нуля*.

Утверждение 3. В поле не может быть делителей нуля: $ab = 0 \ \Rightarrow\ a = 0$ или $b = 0$.

Доказательство. Пусть $ab = 0$. Если $a = 0$, то утверждение доказано. Предположим, что $a \neq 0$. Тогда для a существует обратный элемент a^{-1} $(a^{-1}a = aa^{-1} = 1)$. В силу утверждения 1 и ассоциативности умножения, $0 = a^{-1} \cdot 0 = a^{-1}(ab) = (a^{-1}a)\, b = 1 \cdot b = b$. $\square$

ПРИМЕРЫ:

(1) K — множество четных целых чисел. Операции — сложение и умножение целых чисел. Это коммутативное кольцо без единицы. Кольцо не имеет делителей нуля.

(2) $K = \mathbb{R}^{n \times n}$ (множество всех $n \times n$-матриц). Операции — сложение и умножение матриц. Это некоммутативное кольцо с единицей. Кольцо имеет делители нуля. Например, в случае $n = 2$ находим

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = 0.$$

(3) K — множество всех чисел вида $a + b\sqrt{2}$, где $a, b \in \mathbb{Q}$. Операции — сложение и умножение вещественных чисел. Ясно, что сумма чисел такого вида и их произведения будут числами такого же вида. Поэтому очевидно, что K — коммутативное кольцо с единицей $1 = 1 + 0 \cdot \sqrt{2}$.

В данном случае K является полем: множество $K\backslash\{0\}$ относительно операции умножения чисел является абелевой группой (см. пример абелевой группы из Лекции 2).

15.3 Кольцо вычетов

Напомним, что вычеты — это специальные подмножества целых чисел (см. пример классов эквивалентности из Лекции 6).

Зафиксируем целое число $p > 1$. Для любого $a \in \mathbb{Z}$ обозначим через $\mathbb{Z}(a)$ множество всех целых чисел, имеющих при делении на p такой же остаток, как и число a (сравнимых с a по модулю p). Множества $\mathbb{Z}(a)$ и называются *вычетами по модулю* p.

Множество всех вычетов по модулю p обозначается $\mathbb{Z}_p$. Всего имеется ровно p различных вычетов по модулю p:

$$\mathbb{Z}_p \;=\; \{\mathbb{Z}(0),\ \mathbb{Z}(1),\ \ldots,\ \mathbb{Z}(p-1)\}.$$

Определения операций сложения и умножения вычетов:

$$\mathbb{Z}(a) + \mathbb{Z}(b) = \mathbb{Z}(a+b), \qquad \mathbb{Z}(a)\mathbb{Z}(b) = \mathbb{Z}(ab).$$

Данные определения корректны в силу следующего элементарного наблюдения:

$$\mathbb{Z}(c+d) = \mathbb{Z}(a+b), \qquad \mathbb{Z}(cd) = \mathbb{Z}(ab) \qquad \forall\ \ c \in \mathbb{Z}(a),\ \ \forall\ \ d \in \mathbb{Z}(b).$$

Столь же элементарно проверяется, что относительно операций сложения и умножения множество вычетов $\mathbb{Z}_p$ является коммутативным кольцом с единицей.

Теорема. *В случае простого p и только в этом случае кольцо вычетов по модулю p является полем.*

Доказательство. Пусть p не является простым числом $\;\Rightarrow\;$ $p = ab$ при $1 < a, b < p$ $\;\Rightarrow\;$ $\mathbb{Z}(a)\,\mathbb{Z}(b) = \mathbb{Z}(ab) = \mathbb{Z}(p) = \mathbb{Z}(0) = 0$. Значит, $\mathbb{Z}_p$ имеет делители нуля, и согласно утверждению 3, $\mathbb{Z}_p$ не может быть полем при составном p.

Теперь предположим, что p — простое число. Докажем, что для любого вычета $\mathbb{Z}(a)$ при $1 \le a \le p-1$ существует вычет $\mathbb{Z}(b)$ такой, что $\mathbb{Z}(a)\mathbb{Z}(b) = \mathbb{Z}(1) = 1$. Для этого рассмотрим числа вида ka и их остатки от деления на p:

$$1 \cdot a = pq_1 + r_1, \qquad 2 \cdot a = pq_2 + r_2, \qquad \ldots\ , \qquad (p-1) \cdot a = pq_{p-1} + r_{p-1}, \tag{1}$$

$$q_1, r_1,\ \ldots,\ q_{p-1}, r_{p-1} \in \mathbb{Z}, \qquad 0 \le r_1, \ldots, r_{p-1} \le p-1.$$

Ни один из остатков $r_1, \ldots, r_{p-1}$ не равен нулю, иначе a делилось бы на p. Кроме того, среди них нет совпадающих. Предположим, что $r_k = r_m$. Тогда $(k-m)a = p(q_k - q_m)$. Поскольку a и p взаимно простые, $k-m$ делится на p.

Однако, при $k, m = 1, 2, \ldots, p-1$ очевидно, что $|k-m| < p \;\Rightarrow\; k-m = 0$. Таким образом,

$$\{r_1,\ r_2,\ \ldots,\ r_{p-1}\} \;=\; \{1,\ 2,\ \ldots,\ p-1\}. \tag{2}$$

Значит, при некотором k непременно $r_k = 1 \;\Rightarrow\; \mathbb{Z}(a)\,\mathbb{Z}(r_k) = \mathbb{Z}(1) = 1$. $\square$

Замечание. В проведенных рассуждениях фактически содержится доказательство "малой" теоремы Ферма: *если p — простое число и a взаимно просто с p, то число $a^{p-1}-1$ делится на p.* В самом деле, перемножая равенства (1) и учитывая (2), получаем, что $(p-1)!\,(a^{p-1}-1)$ делится на p. Поскольку $(p-1)!$ и p взаимно просты, на p обязано делиться число $a^{p-1}-1$.

Как видим, кольца $\mathbb{Z}_p$ дают примеры *конечных колец*, а при простом p — также примеры *конечных полей* (то есть, колец и полей с конечным числом элементов).

Конечные поля играют важную роль в прикладных вопросах математики — например, в теории кодирования, обнаружения и исправления ошибок при передаче информации по различным каналам связи.

15.4 Вложения и изоморфизмы

Пусть M — непустое подмножество в K. Если K — кольцо, то M называется его *подкольцом*, если оно является кольцом относительно операций, действующих в K. Если K — поле, то M называется его *подполем*, если оно является полем относительно тех же операций, котоорые действуют в K. В таких случаях говорят, что M *вложено* в K, или K является *расширением* кольца (поля) M.

В различных построениях могут возникать кольца или поля, неразличимые с точки зрения свойств действующих в них операций. Одинаковость свойств операций в L и M означает существование взаимно-однозначного отображения $\Phi : L \to M$, сохраняющего операции:

$$\Phi(a+b) = \Phi(a) + \Phi)b), \qquad \Phi(ab) = \Phi(a)\,\Phi(b) \qquad \forall\ \ a, b \in L.$$

Такое отображение Φ называется *изоморфизмом*, а L и M — *изоморфными*.

Обычно K называют расширением кольца (поля) L и в тех случаях, когда L изоморфно некоторому его подкольцу (подполю) M.

Пусть $\mathbf{1}$ — единичный элемент поля P. Рассмотрим суммы, состоящие из p слагаемых вида

$$p \cdot \mathbf{1} \ = \ \mathbf{1} + \ldots + \mathbf{1}.$$

Подчеркнем, что правая часть есть *определение* выражения $p \cdot \mathbf{1}$ (p не является элементом нашего поля и, стало быть, речь не идет об умножении двух элементов поля). Минимальное p такое, что $p \cdot \mathbf{1} = 0$, называется *характеристикой* поля P.

Утверждение 1. *Если поле имеет характеристику $p \geq 1$, то число p простое.*

Доказательство. Предположим от противного, что $p = mk$. Тогда $0 = (mk) \cdot \mathbf{1} = (m \cdot \mathbf{1})(k \cdot \mathbf{1})$. Это невозможно, так как в поле не бывает делителей нуля. □

Утверждение 2. *Любое поле характеристики $p \geq 1$ может рассматриваться как расширение поля вычетов $\mathbb{Z}_p$.*

Доказательство. В поле характеристики p имеется, по крайней мере, p различных элементов вида $k \cdot \mathbf{1}$, $k = 1, \ldots, p$. Легко проверяется, что составленное из них множество является подполем. Изоморфизм данного подполя с $\mathbb{Z}_p$ устанавливается отображением $\Phi(k \cdot \mathbf{1}) = \mathbb{Z}(k)$. □

Следствие. *Любое конечное поле может рассматриваться как расширение некоторого поля вычетов.*

Задача 1. Пусть P — числовое поле и при этом $\mathbb{R} \subset P \subset \mathbb{C}$. Докажите, что $P = \mathbb{R}$ либо $P = \mathbb{C}$.

Задача 2. Найдите все поля, вложенные в поле $\mathbb{Q}$.

15.5 Число элементов в конечном поле

Утверждение 3. *В конечном поле число элементов обязательно имеет вид $n = p^m$, где p — простое, m — натуральное число.*

Доказательство. Если p — характеристика конечного поля F, то, согласно утверждению 2, F является расширением поля вычетов по простому модулю p: $\mathbb{Z}_p \subset F$.

По аналогии с нашими исследованиями в случае вещественного линейного пространства, элементы $a_1, \ldots, a_m \in F$ назовем линейно независимыми над $\mathbb{Z}_p$, если из равенства $\alpha_1 a_1 + \ldots + \alpha_m a_m = 0$ с коэффициентами $\alpha_1, \ldots, \alpha_m \in \mathbb{Z}_p$ вытекает, что $\alpha_1 = \ldots = \alpha_m = 0$. Пусть m — максимально возможное число элементов, линейно независимых над $\mathbb{Z}_p$. Тогда любой элемент $v \in F$ имеет вид

$$v = \alpha_1 a_1 \ + \ \ldots \ + \ \alpha_m a_m, \qquad \alpha_1, \ldots, \alpha_m \in \mathbb{Z}_p.$$

Для каждого из коэффициентов α_i возможно p различных значений $\Rightarrow \quad n = p^m$. $\square$

Конечные поля принято называть также *полями Галуа*. Мы доказали, что для существования поля Галуа необходимо, чтобы его число элементов имело вид $n = p^m$. Но существуют ли поля Галуа для произвольного n такого вида? Ответ положительный, но на конструировании соответствующих полей мы останавливаться не будем.

Задача. Докажите существование поля из четырех элементов.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

15.6 Поле частных

Теорема. *Любое коммутативное кольцо без делителей нуля может быть вложено в поле.*

Доказательство. Пусть K — коммутативное кольцо без делителей нуля. Чтобы расширить его до поля, рассмотрим *формальные частные* вида $\frac{a}{b}$, где $a, b \in K$ и $b \neq 0$. Назовем формальные частные $\frac{a}{b}$ и $\frac{c}{d}$ равными, если $ad = bc$. Данное отношение равенства является, очевидно, рефлексивным ($\frac{a}{b} = \frac{a}{b}$) и симметричным. Но оно также транзитивно. В самом деле,

$$\frac{a}{b} = \frac{c}{d} \quad \Leftrightarrow \quad ad = bc, \qquad \frac{c}{d} = \frac{p}{q} \quad \Leftrightarrow \quad cq = dp.$$

Отсюда $(aq - bp)(cd) = 0$ и, в силу отсутствия делителей нуля, $aq - bp = 0 \quad \Rightarrow$

$$aq = bp \quad \Leftrightarrow \quad \frac{a}{b} = \frac{p}{q}.$$

Следовательно, отношение равенства является на множестве всевозможных формальных частных отношением эквивалентности. Поэтому все множество формальных частных разбивается на непересекающиеся классы эквивалентности.

Пусть $K\left(\frac{a}{b}\right)$ обозначает класс эквивалентности, порождаемый формальным частным $\frac{a}{b}$. Как мы уже знаем, класс эквивалентности однозначно определяется любым своим представителем: если $\frac{c}{d} \in K\left(\frac{a}{b}\right)$, то $K\left(\frac{c}{d}\right) = K\left(\frac{a}{b}\right)$; поэтому традиционно он отождествляется с любым своим представителем.

Операции сложения и умножения классов эквивалентности формальных частных определим по аналогии с заданием операций для рациональных чисел:

$$K\left(\frac{a}{b}\right) + K\left(\frac{c}{d}\right) = K\left(\frac{ad + bc}{bd}\right), \qquad K\left(\frac{a}{b}\right) K\left(\frac{c}{d}\right) = K\left(\frac{ac}{bd}\right).$$

Проверка того, что результаты этих операций не зависят от выбора представителей в классах эквивалентности $K\left(\frac{a}{b}\right)$ и $K\left(\frac{c}{d}\right)$, осуществляется вполне рутинным образом.

Легко видеть, что множество формальных частных есть коммутативное кольцо с единицей $1 = K\left(\frac{a}{a}\right)$. При этом $0 = K\left(\frac{0}{a}\right)$. Любой ненулевой элемент имеет вид $K\left(\frac{a}{b}\right)$, где $a \neq 0$. Очевидно, элемент $K\left(\frac{b}{a}\right)$ будет к нему обратным.

Итак, множество классов $K\left(\frac{a}{b}\right)$ есть поле. Почему оно может считаться расширением кольца K? Для этого рассмотрим взаимно-однозначное соответствие $a \ \leftrightarrow \ K\left(\frac{ac}{c}\right)$ и заметим, что оно сохраняет операции:

$$a + b \ \leftrightarrow \ K\left(\frac{ac}{c}\right) + K\left(\frac{bc}{c}\right), \qquad ab \ \leftrightarrow \ K\left(\frac{ac}{c}\right) K\left(\frac{bc}{c}\right).$$

Остается договориться об отождествлении элемента $a \in K$ с классом эквивалентности $K\left(\frac{ac}{c}\right)$ (конечно, не зависящим от выбора $c \neq 0$). $\square$

Построенное поле формальных частных является *минимальным полем*, содержащим K — в том смысле, что любое поле, содержащее K, должно содержать и данное поле частных (это очевидно — вместе с любыми двумя элементами поле содержит также их частное).

15.7 Мультипликативная группа поля вычетов

Теорема. *Мультипликативная группа поля вычетов является циклической.*

Доказательство. Пусть G — мультипликативная группя поля $\mathbb{Z}_p$; обозначим ее элементы через $g_1 = \mathbb{Z}(1),\ \ldots,\ g_{p-1} = \mathbb{Z}(p-1)$. Требуется доказать, что эти элементы суть степени какого-то одного из них.

Для любого $g \in G$ рассмотрим степени g^k при целых $k \geq 1$. Обязательно найдется k такое, что $g^k = 1$. Наименьшее k с таким свойством называется *порядком* элемента g. Пусть $d(g)$ — порядок элемента g. Заметим, что все целые степени элемента g образуют в G подгруппу с числом элементов, равным $d(g)$. В силу теоремы Лагранжа (см. дополнительную часть Лекции 2), $d(g)$ является делителем числа $p-1$.

Обозначим через $\mathcal{D}_n$ множество всех натуральных делителей натурального числа n, а через $\phi(n)$ — общее количество взаимно простых с n целых чисел в промежутке от 1 до n. В теории чисел функция $\phi(n)$ называется *функцией Эйлера*. Можно доказать, что имеет место следующая *формула Гаусса*:

$$\sum_{d \in \mathcal{D}_n} \phi(d) \ = \ n. \qquad (*)$$

Действительно, пусть $d \in \mathcal{D}_n$ и $M(d)$ — множество чисел вида $k(n/d)$, где k пробегает множество взаимно простых с d чисел в промежутке от 1 до d. При $d_1 \neq d_2$ множества $M(d_1)$ и $M(d_2)$ не пересекаются. Если бы они имели общий элемент $k_1(n/d_1) = k_2(n/d_2)$, где k_1 взаимно просто с d_1, а k_2 взаимно просто с d_2, то это означало бы, что $k_1 d_2 = k_2 d_1$. Отсюда бы вытекало, что d_2 делится на d_1 и одновременно d_1 делится на $d_2 \quad \Rightarrow \quad d_1 = d_2$. Число элементов в $M(d)$ равно $\phi(d)$. В то же время, любое число из промежутка от 1 до n попадает в какое-то из множеств $M(d)$ — формула Гаусса доказана.

Далее, пусть $\psi(d)$ обозначает количество элементов группы G, имеющих порядок d. Если $g \in G$ имеет порядок d, то каждый из элементов g^k, где k — взаимно простое с d число из промежутка от 1 до d, также имеет порядок d. В самом деле, из равенства $(g^k)^m = g^{km} = g^0$ при $1 \leq m \leq d$ следует, что km делится на d, и значит, если k и d взаимно просты, то m делится на $d \quad \Rightarrow \quad m = d$. Отсюда

$$\psi(d) \ \geq \ \phi(d) \qquad \forall \ \ d \in \mathcal{D}_{p-1}. \qquad (\#)$$

Теперь примем во внимание очевидное равенство (каждый элемент группы G имеет какой-то порядок d, являющийся делителем порядка группы G, равного $p-1$)

$$\sum_{d \in \mathcal{D}_{p-1}} \psi(d) = p-1$$

и вычтем из него равенство $(*)$ при $n = p-1$:

$$\sum_{d \in \mathcal{D}_{p-1}} (\psi(d) - \phi(d)) \ = \ 0.$$

В силу $(\#)$, имеем равную нулю сумму неотрицательных чисел, поэтому каждое из них равно нулю $\Rightarrow$

$$\psi(d) = \phi(d) \qquad \forall \ \ d \in \mathcal{D}_{p-1}.$$

В частности, $\psi(p-1) = \phi(p-1) \quad \Rightarrow \quad$ в группе G число элементов порядка $p-1$ равно $\phi(p-1) \geq 1$. $\square$

Лекция 16

ОСНОВНАЯ ЧАСТЬ

16.1 Линейные пространства над полем

Пусть P — произвольное поле, элементы которого называются *числами*, и V — непустое множество, элементы которого называются *векторами*.

Предположим, что на V определены две операции: сложение векторов и умножение векторов на числа (элементы из поля P), и пусть эти операции удовлетворяют тем же требованиям (аксиомам), которые были сформулированы при определении вещественного линейного пространства — с тем только отличием, что всюду под числом подразумевается элемент из поля P. В таких случаях V называется *линейным пространством над полем P*, или *векторным пространством над полем P*.

Понятия линейной зависимости и независимости векторов линейного пространства над полем P вводятся так же, как и в случае вещественного линейного пространства. Точно так же вводятся понятия линейной оболочки, базиса, размерности, подпространства (суммы подпространств, пересечения и т. д.). Сохраняются все факты, полученные ранее при исследовании этих понятий.

Заметим, что иногда одно и то же множество векторов V можно рассматривать как линейное пространство над разными полями. Соответствующие линейные пространства должны считаться разными.

ПРИМЕРЫ:

(1) V — множество комплексных чисел (в роли векторов), $P = \mathbb{R}$ — поле вещественных чисел. Сложение векторов определяется как сложение комплексных чисел. Операция умножения векторов на числа из поля $\mathbb{R}$ так же определяется как умножение двух чисел — комплексного и вещественного. Это конечномерное линейное пространство над полем $\mathbb{R}$. Как легко видеть, $\dim V = 2$.

(2) V — множество комплексных чисел, $P = \mathbb{Q}$ — поле рациональных чисел. Сложение векторов определяется как сложение комплексных чисел. Операция умножения векторов на числа из поля $\mathbb{Q}$ так же определяется как умножение двух чисел — комплексного и рационального. Данное линейное пространство является бесконечномерным. [1]

(3) V — множество $m \times n$ матриц с элементами из произвольного поля P. Обозначение: $V = P^{m \times n}$.

[1] Попробуйте доказать, что имеются, по крайней мере, три линейно независимых вектора: 1, $\sqrt[2]{2}$, $\sqrt[3]{2}$. Доказательство бесконечномерности можно извлечь из неразложимости над $\mathbb{Q}$ круговых многочленов простого порядка (см. дополнительную часть к данной лекции).

Сложение векторов определяется как сумма матриц: $[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$. Умножение вектора на число $\alpha \in P$ определяется как умножение матрицы на число: $\alpha[a_{ij}] = [\alpha\, a_{ij}]$.

В данном случае V — конечномерное линейное пространство над полем P; $\dim V = mn$.

16.2 Многочлены над полем

Многочлены от x над полем P — это формальные выражения вида

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n, \qquad a_0,\ a_1,\ \ldots,\ a_n \in P. \tag{$*$}$$

В данном случае x всего лишь символ. Если $a_n \neq 0$, то говорят, что $p(x)$ — *многочлен степени* n. Обозначение: $\deg p(x) = n$. Многочлены нулевой степени называются *константами* и обычно отождествляются с элементами поля P. Многочлен, все коэффициенты которого равны 0, называется *нулевым*. Для нулевого многочлена степень не определена.

Конечно, можно было бы, как и в случае вещественных коэффициентов, рассматривать $p(x)$ как функцию от $x \in P$. Мы не делаем это по следующей причине. Пусть, например, $P = \mathbb{Z}_2 = \{0,\ 1\}$. Тогда $x = x^2 \quad \forall \quad x \in \mathbb{Z}_2$. Как видим, многочлены с разными коэффициентами могут оказаться равными как функции, а нам все же кажется полезным иметь такое определение, при котором они будут различными.

Итак, в случае произвольного поля P мы рассматриваем многочлены именно как формальные выражения от какой-то буквы. При использовании буквы x множество всех многочленов любых степеней обозначается через $P[x]$.

Определение. Будем говорить, что многочлен $p(x)$ вида $(*)$ имеет коэффициент a_i при степени x^i для всех i от 0 до n и коэффициент 0 при любой степени x^i, где $i \geq n+1$. Многочлены от x над полем P называются равными, если они имеют одинаковые коэффициенты при одинаковых степенях буквы x.

Таким образом, многочлены x и x^2 над полем $\mathbb{Z}_2$ считаются различными (хотя и совпадают как функции от $x \in \mathbb{Z}_2$).

Рассмотрим два многочлена из множества $P[x]$:

$$p(x) = a_0 + a_1 x + \ldots + a_{n_p} x^{n_p}, \qquad a_i = 0 \quad \text{при} \quad i \geq n_p + 1,$$

$$q(x) = b_0 + b_1 x + \ldots + b_{n_q} x^{n_q}, \qquad b_i = 0 \quad \text{при} \quad i \geq n_q + 1.$$

Суммой многочленов называется многочлен $p(x) + q(x) = s_0 + s_1 x + \ldots$, в котором коэффициент при x^i равен

$$s_i \ = \ a_i + b_i, \qquad i \geq 0.$$

Произведением многочленов называется многочлен $p(x)\, q(x) = t_0 + t_1 x + \ldots$, в котором коэффициент при x^i равен

$$t_i \ = \ \sum_{k+l=i}^{i} a_k b_l, \qquad i \geq 0.$$

Именно такой многочлен получится, если привычным способом раскрыть скобки и привести подобные члены в выражении

$$(a_0 + a_1 x + \ldots + a_{n_p} x^{n_p})(b_0 + b_1 x + \ldots + b_{n_q} x^{n_q}) \ =$$

$$(a_0b_0) \ + \ (a_1b_0 + a_0b_1)x \ + \ (a_2b_0 + a_1b_1 + a_0b_2)x^2 \ + \ \ldots \ + \ (a_{n_p}b_{n_q})x^{n_p+n_q}.$$

Важное (хотя и очевидное) наблюдение:

$$\deg(p(x)\,q(x)) \ = \ \deg p(x) \ + \ \deg q(x). \qquad (\#)$$

16.3 Кольцо многочленов

Утверждение. *Множество многочленов $P[x]$ относительно операций сложения и умножения многочленов является коммутативным кольцом с единицей. Делителей нуля в $P[x]$ нет.*

Доказательство. Ввиду очевидности того, что сложение превращает $P[x]$ в абелеву группу, перейдем сразу к изучению свойств операции умножения. Наряду с $p(x)$ и $q(x)$, рассмотрим еще один многочлен

$$r(x) = c_0 + c_1x + \ldots + c_{n_r}x^{n_r}, \qquad c_i = 0 \quad \text{при} \quad i \geq n_r + 1.$$

Пусть $(p(x)q(x))r(x) = u_0 + u_1x + \ldots$; $p(x)(q(x)r(x)) = v_0 + v_1x + \ldots$. Тогда, согласно определению операции умножения,

$$u_i \ = \ \sum_{j+m=i} \left(\sum_{k+l=j} a_kb_l \right) c_m \ = \ \sum_{k+l+m=i} a_kb_lc_m \ = \ \sum_{k+j=i} a_k \left(\sum_{l+m=j} b_lc_m \right) \ = \ v_i.$$

Таким образом, умножение многочленов ассоциативно. Дистрибутивность проверяется очевидным образом. Коммутативность умножения также очевидна. Роль единицы выполняет многочлен 1. Отсутствие делителей нуля вытекает из свойства (#). $\square$

Заметим, что $P[x]$ можно рассматривать и как линейное пространство над полем P (сложение векторов определяется как сложение многочленов, умножение векторов на элементы поля P — как умножение многочленов на нулевой многочлен и многочлены нулевой степени, отождествляемые с элементами поля P).

Линейное пространство $P[x]$ бесконечномерно (при определении многочлена как формальной суммы одночленов линейная независимость любой системы одночленов с разными степенями очевидна). Множество многочленов $P_n[x]$ степени n или ниже является подпространством размерности $n + 1$.

16.4 Деление с остатком

Утверждение. *Для любой пары многочленов $f(x), g(x) \in P[x]$ в случае $g(x) \neq 0$ существуют и единственны многочлены $q(x), r(x) \in P[x]$ такие, что*

$$f(x) = g(x)q(x) + r(x), \qquad \deg r(x) < \deg g(x) \quad \textit{либо} \quad r(x) = 0. \qquad (*)$$

Доказательство. Пусть $f(x) = a_nx^n + \ldots + a_0$, $g(x) = b_mx^m + \ldots + b_0$, причем $b_m \neq 0$. Если $\deg f(x) < \deg g(x)$, то существование доказано: $q(x) = 0$ и $r(x) = f(x)$. Если $\deg f(x) \geq \deg g(x)$, то положим

$$f_1(x) = f(x) - \left(\frac{a_n}{b_m}\, x^{n-m} \right) g(x) \quad \Rightarrow \quad \deg f_1(x) < \deg f(x) \quad \text{либо} \quad f_1(x) = 0.$$

Воспользуемся индукцией по степени $f(x)$. Если уже найдено представление

$$f_1(x) = g(x)q_1(x) + r_1(x), \qquad \deg r_1(x) < \deg g(x) \quad \text{либо} \quad r_1(x) = 0,$$

то $(*)$ получается при выборе

$$q(x) = \left(\frac{a_n}{b_m}\, x^{n-m}\right) + q_1(x), \qquad r(x) = r_1(x).$$

Докажем единственность. Пусть имеется еще одна пара многочленов $\widetilde{q}(x)$ и $\widetilde{r}(x)$, удовлетворяющих соотношению $(*)$. Тогда

$$g(x)(q(x) - \widetilde{q}(x)) \;=\; r(x) - \widetilde{r}(x).$$

Если $q(x) - \widetilde{q}(x) \neq 0$, то степень многочлена в левой части не меньше степени $g(x) \;\Rightarrow$ $\deg(r(x) - \widetilde{r}(x)) \geq \deg g(x)$. Это невозможно, потому что при вычитании многочленов степень результата не выше степени каждого из них $\;\Rightarrow\;$ $q(x) = \widetilde{q}(x) \;\Rightarrow\; r(x) = \widetilde{r}(x)$. $\square$

Многочлен $r(x)$ из равенства $(*)$ называется *остатком*, а $q(x)$ — *неполным частным* при делении многочлена $f(x)$ на $g(x) \neq 0$. Если $r(x) = 0$, то говорят, что $f(x)$ *делится на* $g(x)$, или $g(x)$ *является делителем* многочлена $f(x)$.

16.5 Наибольший общий делитель

Пусть многочлен $d(x) \in P[x]$ является общим делителем многочленов $f(x)$ и $g(x)$ из $P[x]$. Он называется *наибольшим общим делителем*, если любой общий делитель этих многочленов является также и его делителем. Обозначение: $d(x) = (f(x), g(x))$. Многочлены называются *взаимно простыми* над полем P, если их наибольший общий делитель имеет нулевую степень.

Из определения ясно, что наибольший общий делитель многочленов определен однозначно с точностью до ненулевого множителя (многочлена нулевой степени), принадлежащего полю P. В случае взаимно простых многочленов $f(x)$ и $g(x)$ всегда можно считать, что $(f(x), g(x)) = 1$.

Предположим, что $\deg f(x) \geq \deg g(x)$. Наибольший общий делитель многочленов $f(x)$ и $g(x)$ можно найти с помощью *алгоритма Евклида*, представляющего собой цепочку делений с остатком следующего вида:

$$\begin{array}{rcrcll}
f(x) & = & g(x)q_1(x) & + & r_1(x), & \deg r_1(x) < \deg g(x), \\
g(x) & = & r_1(x)q_2(x) & + & r_2(x), & \deg r_2(x) < \deg r_1(x), \\
r_1(x) & = & r_2(x)q_3(x) & + & r_3(x), & \deg r_3(x) < \deg r_2(x), \\
\dots & \dots & \dots & & & \\
r_{k-2}(x) & = & r_{k-1}(x)q_k(x) & + & r_k(x), & \deg r_k(x) < \deg r_{k-1}(x), \\
r_{k-1}(x) & = & r_k(x)q_{k+1}(x). & & &
\end{array}$$

При последовательном делении с остатком степень остатка понижается на каждом шаге. В данной цепочке $r_k(x)$ — последний ненулевой остаток.

Утверждение. $r_k(x) = (f(x), g(x))$.

Доказательство. Просматривая данные равенства снизу вверх, легко убедиться в том, что $r_k(x)$ является общим делителем многочленов $f(x)$ и $g(x)$. Пусть $\widetilde{d}(x)$ — любой

общий делитель для $f(x)$ и $g(x)$. Просматривая те же равенства сверху вниз, получаем, что $\widetilde{d}(x)$ является делителем для $r_k(x)$. Следовательно, $r_k(x) = (f(x), g(x))$. $\square$

Теорема о наибольшем общем делителе. *Для любых многочленов $f(x), g(x) \in P[x]$ существуют многочлены $\phi(x), \psi(x) \in P[x]$ такие, что*

$$f(x)\phi(x) + g(x)\psi(x) \;=\; d(x), \qquad d(x) = (f(x), g(x)).$$

Доказательство. Искомые многочлены конструктивно получаются на основе алгоритма Евклида. Если уже получены равенства

$$r_{i-2}(x) = f(x)\,\phi_{i-2}(x) + g(x)\,\psi_{i-2}(x), \quad r_{i-1}(x) = f(x)\,\phi_{i-1}(x) + g(x)\,\psi_{i-1}(x),$$

то из них нетрудно вывести, что $r_i(x) = f(x)\,\phi_i(x) + g(x)\,\psi_i(x)$, где

$$\phi_i(x) = \phi_{i-2}(x) - \psi_{i-1}(x)\,q_i(x), \quad \psi_i(x) = \psi_{i-2}(x) - \psi_{i-1}(x)\,q_i(x).$$

Требуемое равенство получается при $i = k$. $\square$

Следствие. *Для взаимно простых многочленов $f(x), g(x) \in P[x]$ существуют многочлены $\phi(x), \psi(x) \in P[x]$ такие, что $f(x)\,\phi(x) + g(x)\,\psi(x) = 1$.*

Замечание. Любой многочлен вида $f(x)\phi(x) + g(x)\psi(x)$ делится на $d(x) = (f(x), g(x))$ (поэтому, в частности, его степень не меньше степени $d(x)$).

16.6 Значения многочлена и корни

Пусть $f(x) = a_0 + a_1x + \ldots + a_nx^n \in P[x]$ и $\theta \in P$. Определим $f(\theta)$ естественным образом: $f(\theta) = a_0 + a_1\theta + \ldots + a_n\theta^n$. Ясно, что $f(\theta) \in P$. Оно и называется *значением* многочлена $f(x)$ при $x = \theta$. Элемент θ называется *корнем* многочлена $f(x)$, если $f(\theta) = 0$.

Теорема Безу. *Если $f(x) \in P[x]$ и $f(\theta) = 0$ для некоторого $\theta \in P$, то $f(x)$ делится на $x - \theta$.* [2]

Доказательство. Выполнив деление с остатком, находим $f(x) = (x - \theta)q(x) + r(x)$, где $r(x) = 0$ либо $\deg r(x) = 0$. Если $r(x) = 0$, то все доказано. Случай $\deg r(x) = 0$ ведет к противоречию: $0 = f(\theta) = (\theta - \theta)\,q(\theta) + r(\theta) = r(\theta) \quad \Rightarrow \quad r(x) = 0$. В то же время, согласно нашим определениям, многочлен нулевой степени не может быть равен нулевому многочлену. $\square$

Ненулевой многочлен $f(x) \in P[x]$ называется разложимым над P, если существуют многочлены ненулевой степени $p(x), q(x) \in P[x]$ такие, что $f(x) = p(x)q(x)$. В противном случае многочлен $f(x)$ называется *неразложимым*, или *неприводимым* над P.

Из теоремы Безу вытекает, что неразложимый над P многочлен не может иметь корней из P, а произвольный многочлен степени n над P не может иметь более n корней.

16.7 Присоединение корня

Нередко приходится рассматривать многочлены над полем P, не имеющие корней из P. Такие многочлены могут, тем не менее, иметь корень в каком-либо расширении F

[2] Данное предложение обычно приводится в качестве главного следствия из теоремы Безу — утверждения о том, что $r(\theta) = f(\theta)$ для остатка $r(x)$ от деления $f(x)$ на $x - \theta$.

поля P. Элемент $\theta \in F$ называется *алгебраическим над полем* P, если он является корнем многочлена над P. Многочлен над P минимальной степени с корнем θ называется *минимальным многочленом* для θ над полем P.

Будем рассматривать только такие расширения поля P, которые вложены в F. Пусть $\theta \in F$. Поле называется *минимальным θ-расширением* поля P, если оно содержит θ и вложено в любое поле, содержащее P и θ. Обозначение: $P(\theta)$.

В более общем случае, если $\theta_1, \ldots, \theta_k \in F$, то через $P(\theta_1, \ldots, \theta_k)$ обозначается минимальное поле, содержащее P и элементы $\theta_1, \ldots, \theta_k$. Минимальность означает, что данное поле вложено в любое поле, содержащее P и $\theta_1, \ldots, \theta_k$.

Если $\theta \notin P$, то говорят, что поле $P(\theta)$ получено из P присоединением элемента θ. Расширение такого типа называется *простым алгебраическим*, если θ является корнем некоторого многочлена из $P[x]$.

Теорема о присоединении корня. *Минимальный многочлен для θ определяется однозначно с точностью до ненулевого множителя. Если n — его степень, то минимальное θ-расширение поля P имеет вид*

$$P(\theta) = \{s \in F : \ s = a_0 + a_1\theta + \ldots + a_{n-1}\theta^{n-1}, \quad a_0,\ a_1,\ \ldots,\ a_{n-1} \in P\}. \qquad (*)$$

Доказательство. Предположим, что $f(x)$ и $g(x)$ — два минимальных многочлена для θ (оба степени n). Тогда их наибольший общий делитель $d(x) \in P[x]$ имеет вид

$$d(x) = f(x)\phi(x) + g(x)\psi(x), \quad \text{где} \quad \phi(x), \psi(x) \in P[x].$$

Отсюда $d(\theta) = 0$. Поэтому $\deg d(x) = n \quad \Rightarrow \quad$ каждый из многочленов $f(x)$ и $g(x)$ отличается от $d(x)$ лишь ненулевым множителем.

Обозначим через M множество, определенное правой частью $(*)$. Очевидно, что $M \subset P(\theta)$. Поэтому остается только доказать, что M — подполе.

Возьмем произвольный многочлен $p(x)$ над полем P и заметим, что $p(\theta) \in M$. Для доказательства разделим $p(x)$ с остатком на минимальный многочлен $f(x)$:

$$p(x) = f(x)q(x) + r(x) \qquad \Rightarrow \qquad p(\theta) = r(\theta).$$

Ясно, что $r(\theta)$ есть сумма элементов $1,\ \theta,\ \ldots,\ \theta^{n-1}$ с коэффициентами из поля P. Поэтому $r(\theta) \in M$.

Произведение двух элементов из M является, очевидно, значением некоторого многочлена $p(x) \in P[x]$ при $x = \theta$. Поэтому оно принадлежит M. Далее, любой элемент из M имеет вид $p(\theta)$, где многочлен $p(x) \in P[x]$ имеет степень не выше $n - 1$. Многочлен $f(x)$, очевидно, неразложим, поэтому многочлены $p(x)$ и $f(x)$ взаимно просты. По следствию из теоремы о наибольшем общем делителе, существуют многочлены $\phi(x), \psi(x) \in P[x]$ такие, что

$$f(x)\phi(x) + p(x)\psi(x) = 1 \qquad \Rightarrow \qquad p(\theta)\psi(\theta) = 1. \qquad \square$$

Задача. Поле P — минимальное числовое поле, содержащее поле рациональных чисел $\mathbb{Q}$ и $\sqrt[5]{2}$. Докажите, что поле P есть линейное пространство над полем $\mathbb{Q}$ и найдите его размерность.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

16.8 Построения циркулем и линейкой

Наши исследования линейной зависимости, полей и многочленов уже сейчас позволяют разобраться с многими не очень простыми вопросами. Давайте посмотрим, как они применяются к анализу построений, выполняемых с помощью лишь циркуля и линейки. Вот знаменитые примеры таких задач:

- построить ребро куба, объем которого в два раза больше объема заданного куба (задача об удвоении куба);
- построить правильный n-угольник, вписанный в заданную окружность.

Вопрос о том, что можно и что нельзя построить с помощью циркуля и линейки, оказался трудным и не поддавался решению на протяжении многих веков.

Используя метод координат, мы можем свести вопрос о геометрических построениях к нахождению некоторой специальной цепочки расширений полей, начинающейся с поля рациональных чисел. Все поля вложены, конечно, в поле вещественных чисел.

Не ограничивая общности, можно считать, что ребро заданного куба и радиус заданной окружности равны 1. Опираясь на теорему Фалеса, мы можем построить с помощью циркуля и линейки любой отрезок рациональной длины.

Пусть алгоритм построения представляет собой последовательность из m шагов. На начальном (нулевом) шаге мы имеем любые точки с координатами из поля $\mathbb{Q}_0 = \mathbb{Q}$. Далее предположим, что к началу i-го шага мы имеем любые точки с координатами из некоторого поля $\mathbb{Q}_{i-1}$. Тогда на i-м шаге выполняется одно из трех допустимых построений:

(a) пересечение двух прямых, проходящих через точки с координами из $\mathbb{Q}_{i-1}$;

(b) пересечение прямой и окружности — в предположении, что прямая проходит через пару точек с координатами из $\mathbb{Q}_{i-1}$, центр окружности есть точка с координатами из $\mathbb{Q}_{i-1}$, а ее радиус — число из $\mathbb{Q}_{i-1}$;

(c) пересечение двух окружностей — с тем же предположением относительно центра и радиуса.

Не очень трудно убедиться в том, что каждое из допустимых построений дает точки, координаты которых принадлежат полю $\mathbb{Q}_{i-1}$ либо некоторому его расширению

$$\mathbb{Q}_i = \mathbb{Q}_{i-1}(\theta_i), \quad \text{где} \quad \theta_i \notin \mathbb{Q}_{i-1}, \quad \text{но} \quad D_i \equiv \theta_i^2 \in \mathbb{Q}_{i-1}.$$

Перенумеруем подряд только те поля, которые не совпадают с предыдущим полем. После этого получаем цепочку из $k \le m$ расширений вида

$$\mathbb{Q} = \mathbb{Q}_0 \ \subset \ \mathbb{Q}_1 \ \subset \ \ldots \mathbb{Q}_{k-1} \ \subset \ \mathbb{Q}_k, \tag{1}$$

$$\mathbb{Q}_i = \mathbb{Q}_{i-1}(\theta_i), \qquad \theta_i \notin \mathbb{Q}_{i-1}, \qquad D_i = \theta_i^2 \in \mathbb{Q}_i, \quad i = 1, \ldots, k. \tag{2}$$

Теперь мы в состоянии доказать, например, следующий результат.

Теорема. *Задача об удвоении куба неразрешима с помощью циркуля и линейки.*

Доказательство. В данном случае цель построений — отрезок длины $2^{1/3}$. Если построение возможно, то существует такая цепочка расширений, в которой $2^{1/3} \in \mathbb{Q}_k$, но $2^{1/3} \notin \mathbb{Q}_{k-1}$. Следовательно,

$$2^{1/3} = a + b\theta_k, \quad a, b \in \mathbb{Q}_{k-1}, \quad b \neq 0.$$

Возводя в куб, находим

$$2 = a^3 + 3a^2\theta_k + 3ab^2\, D_k + b^3 D_k\, \theta_k \quad \Rightarrow \quad 2 - a^3 - 3ab^2\, D_k = (3a^2 + b^2 D_k) b\, \theta_k.$$

Учитывая, что $b \neq 0$ и $3a^2 + b^2 D_k > 0$, получаем

$$\theta_k = \frac{2 - a^3 - 3ab^2 D_k}{(3a^2 + b^2 D_k) b} \in \mathbb{Q}_{k-1},$$

что противоречит нашим предположениям. □

Исследование вопроса о построении правильных n-угольников менее элементарно. Тем не менее, мы находимся буквально в двух шагах, например, от доказательства невозможности построения правильного 7-угольника. Один из этих шагов связан с изучением расширений полей как линейных пространств и включает легко доказываемую теорему о размерностях этих пространств. Другой шаг эквивалентен доказательству неразложимости над полем рациональных чисел многочлена $f(x) = 1 + x + \ldots + x^{n-1}$ при простом n.

16.9 Конечные расширения полей

Предположим, что поле P вложено в поле F. Тогда элементы из F можно рассматривать как векторы. Суммой векторов можно назвать их сумму как элементов поля F. Умножение векторов (элементов F) на числа (элементы P) можно определить естественным образом как умножение двух элементов: один (вектор) из поля F, другой (число) — из поля P. Все аксиомы линейного пространства, как легко проверить, выполнены. Поэтому F можно рассматривать как линейное пространство над полем P.

Поле F называется *конечным расширением* поля P, если оно является конечномерным как линейное пространство над полем P. Размерность данного линейного пространства называется *степенью расширения* и обозначается $(F : P)$.

Предположим, что поле P вложено в поле F, а F вложено в поле H: $P \subset F \subset H$. Тогда можно рассматривать следующие три расширения:

$$P \subset F, \qquad F \subset H, \qquad P \subset H. \tag{$*$}$$

Теорема. *Из конечности первых двух расширений вида* $(*)$ *вытекает конечность третьего расширения, а из конечности третьего — конечность первых двух расширений. При этом степени расширений связаны соотношением*

$$(H : P) \;=\; (H : F)\,(F : P).$$

Доказательство. Предположим конечность расширений $P \subset F$ и $F \subset H$. Пусть $a_1, \ldots, a_m$ — элементы поля F, образующие базис линейного пространства F над полем P. Аналогично, пусть $b_1, \ldots, b_n$ — элементы поля H, образующие базис линейного пространства H над полем F. Очевидно, любой элемент $h \in H$ можно представить в виде

$$h = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} s_{ij} a_i \right) b_j = \sum_{i=1}^{m} \sum_{j=1}^{n} s_{ij} (a_i b_j), \qquad s_{ij} \in P.$$

Таким образом, любой элемент $h \in H$ представим в виде линейной комбинации mn элементов поля $H \quad \Rightarrow \quad$ линейное пространство H над полем P конечномерно и его размерность не выше mn.

Остается доказать линейную независимость элементов (векторов)

$$a_i b_j, \qquad 1 \le i \le m, \quad 1 \le j \le n.$$

Пусть $h = 0$. Тогда, поскольку $b_1, \ldots, b_n$ есть базис линейного пространства H над полем F, находим

$$\sum_{i=1}^{m} s_{ij} a_i = 0, \qquad 1 \le j \le m.$$

Поскольку элементы (векторы) $a_1, \ldots, a_m$ образуют базис в линейном пространстве F над полем P, отсюда получаем $s_{ij} = 0$ для всех i, j. Следовательно, размерность линейного пространства H над полем P в точности равна mn.

Теперь предположим, что расширение $P \subset H$ конечно. В данном случае конечность расширения $P \subset F$ очевидна. Если бы имелось n элементов (векторов) линейного пространства H над полем F, то, повторяя предыдущее рассуждение, мы бы установили линейную независимость векторов $a_i b_j$ — как элементов линейного пространства H над полем P. Значит, $mn \le (H : P)$, то есть, расширение $F \subset H$ конечно. $\square$

Следствие. *Степень расширения $\mathbb{Q} \subset \mathbb{Q}_k$, получаемого в цепочке расширений (1), (2), равна 2^k.*

Доказательство. Согласно теореме о минимальном θ-расширении, каждое из расширений $\mathbb{Q}_{i-1} \subset \mathbb{Q}_i$ в цепочке (1), (2) имеет степень 2. $\square$

16.10 Круговые многочлены простой степени

Речь идет о многочленах $f(x) = 1 + x + \ldots + x^{n-1} = \frac{x^n - 1}{x - 1}$ при простом n.

Теорема. *Многочлен $f(x)$ при простом n неразложим над полем рациональных чисел.*

Доказательство. Легко доказывается, что разложимость $f(x)$ над $\mathbb{Q}$ равносильна возможности его представления в виде $f(x) = g(x)h(x)$, где ненулевые многочлены $g(x)$ и $h(x)$ имеют целочисленные коэффициенты. [3]

[3] Вообще можно доказать, что многочлен с целочисленными коэффициентами разложим над $\mathbb{Q}$ тогда и только тогда, когда он разложим в произведение двух многочленов с целочисленными коэффициентами. Это можно вывести из следующей леммы.

Лемма Гаусса. *Для любых целочисленных многочленов $f(x) = a_0 + \ldots + a_m x^m$ и $g(x) = b_0 + \ldots + b_n x^n$ наибольший общий делитель C всех коэффициентов произведения $f(x)g(x) = c_0 + \ldots + c_{m+n} x^{m+n}$ равен произведению наибольшего общего делителя A всех коэффициентов $f(x)$ и наибольшего общего*

Заменив каждый из коэффициентов на порождаемый им вычет по простому модулю n, получим многочлены $f_n(x),\ g_n(x),\ h_n(x)$ над полем $\mathbb{Z}_n$ и равенство $f_n(x) = g_n(x)h_n(x)$. Используя разложение для бинома Ньютона, несложно получить следующее равенство многочленов над $\mathbb{Z}_n$: $x^n - 1 = (x-1)^n$. Поэтому в поле $\mathbb{Z}_n$ справедливы разложения

$$f_n(x) = (x-1)^{n-1}, \quad g_n(x) = (x-1)^{m_1}, \quad h_n(x) = (x-1)^{m_2}, \quad m_1 + m_2 = n-1.$$

Следовательно, каждое из целых чисел $g(1)$ и $h(1)$ делится на $n \ \Rightarrow \ f(1) = g(1)h(1)$ делится на n^2. Но это невозможно, так как $f(1) = n$. $\square$

Еще один (пожалуй, даже более простой) подход: доказать неразложимость многочлена $f(x+1)$.

Признак Эйзенштейна. *Пусть дан многочлен $F(x) = a_0 + \ldots + a_n x^n$ с целыми коэффициентами, в котором $a_0, \ldots, a_{n-1}$ делятся на некоторое простое число $p > 1$ и при этом a_0 не делится на p^2. Если a_n не делится на p, то $F(x)$ нельзя представить в виде произведения мсногочленов с целыми коэффициентами.*

Доказательство. Запишем $F(x) = (b_0 + \ldots + b_k x^k)(c_0 + \ldots + c_m x^m)$. Тогда $b_0 c_0 = a_0$ делится на p, но не на p^2. Поэтому одно и только одно из чисел $b_0,\ c_0$ делится на p. Пусть это будет c_0. Среди коэффициентов $c_0, \ldots, c_m$ должен быть не делящийся на p (иначе a_n делится на p). Пусть c_i — первый такой коэффициент. Тогда $a_i = b_0 c_i + (b_1 c_{i-1} + \ldots b_i c_0)$ не делится на p (число в скобках делится на p, а произведение $b_0 c_i$ не делится на p). Отсюда $i = n \le m \ \Rightarrow \ m = n$. $\square$

Остается заметить, что в случае $f(x) = 1 + x + \ldots + x^{n-1}$ при простом n многочлен $F(x) = f(x+1)$ имеет старший коэффициент 1, а все остальные коэффициенты делятся на n.

16.11 Правильные n-угольники

Теперь мы готовы к тому, чтобы доказать, например, что правильный 7-угольник с помощью циркуля и линейки построить нельзя. Более того, для возможности построения правильного n-угольника мы выведем некоторое необходимое условие. (Оно же является и достаточным, но мы докажем только необходимость.)

Будем исходить из того, что вершины вписанного в единичную окружность правильного n-угольника располагаются на корнях из единицы степени n. Предположим, что n — простое число. Пусть существует цепочка вида (1), (2), в которой поле $\mathbb{Q}_k$ содержит координаты всех корней из единицы степени n. Ясно, что минимальная цепочка должна приводить к равенству

$$\mathbb{Q}_k \ = \ \mathbb{Q}(\theta), \qquad \theta = \varepsilon + \varepsilon^{-1} = 2\cos\left(\frac{2\pi}{n}\right), \qquad \varepsilon = \cos\left(\frac{2\pi}{n}\right) \ + \ \mathbf{i}\cos\left(\frac{2\pi}{n}\right).$$

Далее, рассмотрим расширение $\mathbb{Q}(\theta) \subset \mathbb{Q}(\theta)(\varepsilon) = \mathbb{Q}(\varepsilon)$. Поскольку ε является корнем квадратного уравнения

$$x^2 - \theta x + 1 = 0$$

с коэффициентами из поля $\mathbb{Q}(\theta)$, степень расширения $\mathbb{Q}_k \subset \mathbb{Q}(\varepsilon)$ равна 2. Как мы уже знаем, степень расширения $\mathbb{Q} \subset \mathbb{Q}_k$ равна 2^k. Поэтому степень расширения $\mathbb{Q} \subset \mathbb{Q}(\varepsilon)$

делителя B всех коэффициентов $g(x)$.

Доказательство. Ясно, что C делится на AB. Поэтому, не ограничивая общности, можно считать, что $A = B = 1$. Пусть C делится на простое число $p > 1$. Хотя бы один из коэффициентов $a_0, \ldots, a_m$ и хотя бы один из коэффициентов $b_0, \ldots, b_n$ не делится на p. Обозначим через a_r и b_s первые из коэффициентов, не делящиеся на p. Тогда $c_{r+s} = a_r b_s + (a_{r-1}b_{s+1} + \ldots + a_{r+1}b_{s-1} + \ldots)$. Число в скобках делится на p. Поэтому c_{r+s} не может делиться на p. $\square$

равна

$$(\mathbb{Q}(\varepsilon) : \mathbb{Q}) \;=\; (\mathbb{Q}(\varepsilon) : \mathbb{Q}_k)\,(\mathbb{Q}_k : \mathbb{Q}) \;=\; 2^{k+1}.$$

В то же время, в случае простого n расширение $\mathbb{Q} \subset \mathbb{Q}(\varepsilon)$ имеет степень $n-1$ (равную степени минимального многочлена для ε над полем $\mathbb{Q}$). Таким образом, мы доказали следующее утверждение.

Лемма. *Для возможности построения правильного n-угольника в случае простого n необходимо, чтобы n имело вид $n = 2^L + 1$.*

Из нашего рассуждения вытекает, что $L = k+1$. Заметим, что если число $n = 2^L+1$ простое, то L должно иметь вид $L = 2^m$ (если $L = MN$ при нечетном M, то число $(2^N)^M - 1$ делится на $2^N - 1$ и поэтому не может быть простым).

Следствие. *Построение правильного 7-угольника с помощью циркуля и линейки невозможно.*

Доказательство. $7 \neq 2^L + 1$. $\square$

Теорема. *Для возможности построения правильного n-угольника необходимо, чтобы любой нечетный простой сомножитель числа n имел вид $2^L + 1$.*

Доказательство. Достаточно заметить, что если n-угольник строится с помощью циркуля и линейки, то строится также любой правильный многоугольник с числом сторон, равным любому делителю числа n. Случай простых нечетных делителей сводится к применению доказанной выше леммы. $\square$

Исследование вопроса о построении правильных n-угольников — одно из самых ранних достижений Гаусса. Он доказал достаточность полученного выше условия. В частности, Гаусс описал конкретный алгоритм построения правильного 17-угольника (заметим, что $17 = 2^4 + 1$) — теперь мы понимаем, что для этого достаточно найти конкретную цепочку расширений вида (1), (2). Гаусс писал также о том, что данное условие является необходимым. [4]

16.12 Эндоморфизмы и автоморфизмы

Рассмотрим еще одно доказательство неразложимости многочлена

$$f(x) = 1 + x + \ldots + x^{n-1}$$

над $\mathbb{Q}$ при простом n. Оно является более длинным, но приоткрывает связи с некоторыми очень плодотворными идеями и понятиями алгебры (в частности, с автоморфизмами полей — их детальное изучение составляет предмет теории Галуа и выходит за рамки нашего курса).

Пусть F — поле и $\Phi : F \to F$ — отображение, сохраняющее операции:

$$\Phi(a+b) = \Phi(a) + \Phi(b), \qquad \Phi(ab) = \Phi(a)\,\Phi(b) \qquad \forall\;\; a, b \in F.$$

Взаимная однозначность не предполагается. В таких случаях Φ называется *эндоморфизмом* поля F. [5] Если F является расширением поля P, то особый интерес представляют эндоморфизмы, оставляющие

[4] Однако, специалисты по истории вопроса говорят, что доказательство необходимости в рукописях Гаусса не было обнаружено.

[5] В более общем случае, когда $\Phi(F)$ принадлежит другому полю, отображение Φ со свойством сохранения операций называется *гомоморфизмом*.

на месте элементы поля P — они называются *эндоморфизмами F над P*. Пусть $\mathcal{E}(F,P)$ обозначает множество всех эндоморфизмов поля F над полем P.

Утверждение 1. *Пусть $f(x)$ — произвольный многочлен над полем P и $\theta \in F$ — его корень: $f(\theta)=0$. Тогда для любого эндоморфизма $\Phi \in \mathcal{E}(F,P)$ элемент $\Phi(\theta)$ является корнем того же многочлена: $f(\Phi(\theta))=0$.*

Доказательство. Пусть $f(x) = a_0 + a_x + \ldots + a_n x^n$, где $a_i \in P$. Тогда $0 = \Phi(0) = \Phi(f(\theta)) = a_0 + a_1\Phi(\theta) + \ldots + a_n(\Phi(\theta))^n = f(\Phi(\theta))$. □

Изучим подробнее эндоморфизмы для поля, получаемого из поля рациональных чисел $\mathbb{Q}$ присоединением всех корней из единицы степени n (достаточно присоединить лишь один корень — такой, степени которого порождают все множество корней):

$$P \;=\; \mathbb{Q}, \qquad F \;=\; \mathbb{Q}(\varepsilon), \qquad \varepsilon = \cos\left(\frac{2\pi}{n}\right) \;+\; \mathbf{i}\cos\left(\frac{2\pi}{n}\right).$$

Утверждение 2. *Множество $\mathcal{E}(\mathbb{Q}(\varepsilon),\mathbb{Q})$ состоит ровно из n эндоморфизмов Φ_i, однозначно определяемых образом элемента ε: $\Phi_i(\varepsilon)=\varepsilon^i$, $i=1,\ldots,n-1$.*

Доказательство. Пусть $\Phi \in \mathcal{E}(\mathbb{Q}(\varepsilon),\mathbb{Q})$. Тогда, в силу утверждения 1, $\Phi(\varepsilon)=\varepsilon^i$ для некоторого i от 0 до $n-1$.

Теперь докажем, что для любого i существует эндоморфизм $\Phi_i \in \mathcal{E}(\mathbb{Q}(\varepsilon),\mathbb{Q})$ такой, что $\Phi_i(\varepsilon)=\varepsilon^i$.

Пусть $f(x)$ — минимальный многочлен для ε над полем $\mathbb{Q}$. Заметим, что ε есть корень уравнения

$$\frac{x^n-1}{x-1} \;=\; 1 + x + \ldots + x^{n-1} \;=\; 0 \qquad \Rightarrow \quad m \equiv \deg f(x) \le n-1.$$

В силу теоремы о присоединении корня, любой элемент $z \in \mathbb{Q}(\varepsilon)$ однозначно представим в виде

$$z \;=\; \sum_{k=0}^{m} a_k\,\varepsilon^k, \qquad a_k \in \mathbb{Q} \quad \forall \;\; k.$$

Определим отображение $\Phi_i : \mathbb{Q}(\varepsilon) \to \mathbb{Q}(\varepsilon)$ формулой

$$\Phi_i\Big(\sum_{k=0}^{m} a_k\,\varepsilon^k\Big) \;=\; \sum_{k=0}^{m} a_k \varepsilon^{ik}.$$

Легко проверяется, что оно является эндоморфизмом поля $\mathbb{Q}(\varepsilon)$ и оставляет на месте числа из $\mathbb{Q}$. □

Утверждение 3. *В случае простого n любой из эндоморфизмов Φ_i утверждения 2 при $1 \le i \le n-1$ задает взаимно-однозначное отображение множества $\{\varepsilon,\ \varepsilon^2,\ \ldots,\ \varepsilon^{n-1}\}$ на себя, причем каждое такое отображение является циклическим:*

$$\varepsilon^{i_1} \to \varepsilon^{i_2} \to \ldots \to \varepsilon^{i_{n-2}} \to \varepsilon^{i_{n-1}} \to \varepsilon^{i_1},$$

где $i_1,\ \ldots,\ i_{n-1}$ — некоторая перестановка номеров $1,\ 2,\ \ldots,\ n-1$.

Доказательство. Мы знаем, что мультипликативная группа поля вычетов по простому модулю является циклической (см. дополнительную часть Лекции 14). Поэтому существует m в промежутке от 2 до $n-1$ такое, что остатки при делении на n чисел $m,\ m^2,\ m^3,\ \ldots,\ m^{n-1}$ образуют перестановку чисел $\{1,\ 2,\ 3,\ \ldots,\ n-1\}$. Рассмотрим эндоморфизм $\Phi \in \mathcal{E}(\mathbb{Q}(\varepsilon),\mathbb{Q})$ такой, что $\Phi(\varepsilon)=\varepsilon^m$. Очевидно, он действует таким образом:

$$\varepsilon \to \varepsilon^m \to \varepsilon^{m^2} \to \varepsilon^{m^3} \to \ldots \varepsilon^{m^{n-2}} \to \varepsilon^{m^{n-1}} = \varepsilon.$$

Эндоморфизмы $\Phi,\ \Phi^2,\ \ldots,\ \Phi^{n-1}$ являются, очевидно, различными и ни один из них не совпадает с $\Phi_0 \;\Rightarrow\; \{\Phi,\ \Phi^2,\ \ldots,\ \Phi^{n-1}\} = \{\Phi_1,\ \ldots,\ \Phi_{n-1}\}$. Остается заметить, что при любом k отображение Φ^k реализует циклическую подстановку на множестве корней $\{\varepsilon,\ \varepsilon^2,\ \ldots,\ \varepsilon^{n-1}\}$. □

Утверждение 4. *При простом n минимальный многочлен для ε над полем $\mathbb{Q}$ равен $f(x) = 1 + x + \ldots + x^{n-1}$.*

Доказательство. Достаточно убедиться в неразложимости многочлена $f(x)$ над полем $\mathbb{Q}$. Предположим, что $f(x) = u(x)v(x)$, где $u(x), v(x) \in \mathbb{Q}[x]$. Выберем любое k от 1 до $n-1$ и рассмотрим

эндоморфизм $\Phi = \Phi_k$. Пусть степень многочлена $u(x)$ равна m. Тогда он имеет m различных корней $z_1, \ldots, z_m \subset \{\varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}\}$ (следствие из теоремы Безу). Согласно утверждению 1, все числа $z_1, \Phi(z_1), \Phi^2(z_1), \ldots, \Phi^{n-2}(z_1)$ являются корнями $u(x)$. В силу утверждения 3 эти числа попарно различны $\Rightarrow$ $m = n - 1$. $\square$

Эндоморфизмы поля, являющиеся взаимно-однозначными отображениями, называются *автоморфизмами*.

Утверждение 5. *При простом n эндоморфизмы $\Phi_1, \ldots, \Phi_{n-1}$ утверждения 2 являются автоморфизмами поля $\mathbb{Q}(\varepsilon)$, оставляющими на месте элементы поля $\mathbb{Q}$, и исчерпывают все множество автоморфизмов такого типа.*

Доказательство. Данные отображения взаимно-однозначны в силу теоремы о присоединении корня. В то же время, любой автоморфизм Φ, оставляющий на месте элементы из $\mathbb{Q}$, переводит ε в ε^i для какого-то i от 1 до $n-1$ (при автоморфизме ε не может перейти в 1) $\Rightarrow$ Φ совпадает с одним из эндоморфизмов Φ_i. $\square$

16.13 Алгебраические числа

Комплексное число называется *алгебраическим*, если оно является корнем многочлена над полем рациональных чисел. В противном случае оно называется *трансцендентным*. Изученные нами свойства конечных расширений полей делают почти очевидным следующее утверждение.

Теорема. *Множество всех алгебраических чисел относительно операций сложения и умножения комплексных чисел является полем.*

Доказательство. Пусть α и β являются корнями каких-то многочленов над $\mathbb{Q}$. Рассмотрим поле $\mathbb{Q}(\alpha)$, полученное присоединением к $\mathbb{Q}$ элемента α, и поле $\mathbb{Q}(\alpha)(\beta)$, полученное из $\mathbb{Q}(\alpha)$ присоединением элемента β — корня многочлена из кольца $\mathbb{Q}(\alpha)[x]$ (ясно, что $\mathbb{Q}[x] \subset \mathbb{Q}(a)[x]$). Тогда расширение $\mathbb{Q} \subset \mathbb{Q}(\alpha)(\beta)$ является конечным расширением. Очевидно, что любой элемент γ конечного расширения поля $\mathbb{Q}$ является корнем некоторого многочлена над $\mathbb{Q}$, иначе элементы $1, \gamma, \gamma^2, \ldots, \gamma^n$ были бы линейно независимы над $\mathbb{Q}$ при любом n. $\square$

Лекция 17

ОСНОВНАЯ ЧАСТЬ

17.1 Комплексные многочлены

Замечательно, что в наиболее интересных случаях — а именно, для *комплексных многочленов* (многочленов с комплексными коэффициентами) — можно получить точное утверждение о существовании корней: любой многочлен степени $n > 1$ имеет корень, являющийся комплексным числом. Данное утверждение традиционно называется *основной теоремой алгебры*.

Оно занимает действительно особое место в ряде разделов математики — многие из них имеют для нее свои собственные доказательства. Все известные доказательства в той или иной мере используют понятие *непрерывности*. Мы изложим доказательство, основанное на методе Даламбера [1] и требующее от нас наименьшей подготовительной работы.

Мы будем рассматривать многочлен $f(z) \in \mathbb{C}[z]$ как функцию от $z \in \mathbb{C}$. При этом равенство многочленов как функций влечет за собой также их равенство как формальных выражений от степеней буквы z. Для доказательства можно практически повторить рассуждение, проведенное в случае вещественных многочленов. А можно это сделать и так: из теоремы Безу ясно, что многочлен степени n не может иметь более, чем n корней; если $f(z) = g(z)$ для всех z, то многочлен $f(z) - g(z)$ имеет бесконечно много корней, поэтому он обязан быть нулевым многочленом.

17.2 Последовательности комплексных чисел

Пусть задана последовательность комплексных чисел $z_k, \quad k = 1,\ 2,\ \ldots$. Она называется сходящейся к точке z_0, если для любого $\varepsilon > 0$ существует номер $N = N(\varepsilon)$ такой, что для всех $k \geq N$ выполняется неравенство $|z_k - z_0| \leq \varepsilon$. (Согласно определению, понятие сходимости для комплекных последовательностей сводится к сходимости к нулю вещественной последовательности $|z_k - z|$.) Обозначение: $\lim\limits_{k\to\infty} z_k = z_0$ или $z_k \to z_0$.

Теорема Больцано-Вейерштрасса. *Для произвольной последовательности z_k точек прямоугольника $\Pi = [A, B] \times [C, D]$ существует подпоследовательность z_{k_i}, сходящаяся к некоторой точке $z_0 \in \Pi$.*

Доказательство. Запишем $z_k = x_k + \mathbf{i}\, y_k, \quad x_k, y_k \in \mathbb{R}$. Очевидно, $x_k \in [A, B]$ и $y_k \in [C, D]$. В силу теоремы Больцано-Вейерштрасса для вещественных последовательностей на отрезке, существует подпоследовательность x_{k_i}, сходящаяся к вещест-

[1] Заметим, что Даламбер не мог дать полного доказательства — в его время не было строгого понятия непрерывной функции.

венному числу $x_0 \in [A, B]$. Рассмотрим соответствующую подпоследовательность точек $z_{k_i} = x_{k_i} + \mathbf{i}\, y_{k_i}$. Поскольку $y_{k_i} \in [C, D]$, по той же причине найдется подпоследовательность $y_{k_{i_j}}$, сходящаяся к вещественному числу $y_0 \in [C, D]$. При этом $x_{k_{i_j}} \to x_0$ (как подпоследовательность сходящейся последовательности). Пусть $z_0 = x_0 + \mathbf{i}\, y_0$. Тогда

$$|z_{k_{i_j}} - z_0| \le |x_{k_{i_j}} - x_0| + |y_{k_{i_j}} - y_0| \to 0. \qquad \square$$

17.3 Непрерывные функции на комплексной плоскости

Рассмотрим функцию $\Phi(z)$, определенную при всех $z \in \mathbb{C}$ и принимающую вещественные значения. Функция $\Phi(z)$ называется *непрерывной в точке* z_0, если для для любой последовательности z_k, сходящейся к z_0, последовательность значений $\Phi(z_k)$ сходится к $\Phi(z_0)$.

Теорема Вейерштрасса. *Пусть функция $\Phi(z)$ непрерывна во всех точках прямоугольника $\Pi = [A, B] \times [C, D]$. Тогда существуют точки $z_*,\ z^* \in \Pi$ такие, что*

$$\Phi(z_*) \ \le \ \Phi(z) \ \le \ \Phi(z^*) \qquad \forall z \in \Pi.$$

Доказательство. Докажем существование точки z^*. Прежде всего, убедимся в том, что функция $\Phi(z)$ ограничена сверху. Если это не так, то существует последовательность z_k со свойством $\Phi(z_k) > k$. По теореме Больцано-Вейерштрасса, она обладает сходящейся подпоследовательностью $z_{k_i} \to z_0 \in \Pi$. В силу непрерывности, $\Phi(z_{k_i}) \to \Phi(z_0)$, а это противоречит неравенствам $\Phi(z_{k_i}) > k_i$, выполняющимся при всех k_i. Поэтому существует вещественное число M такое, что $\Phi(z) \le M$ для всех $z \in \Pi$. Число M называется *верхней гранью* для $\Phi(z)$.

Рассмотрим множество вещественных чисел $\Phi(\Pi) = \{x :\ x = \Phi(z),\ \ z \in \Pi\}$. Поскольку оно ограничено сверху, то для него существует *точная верхняя грань* M^* — такая верхняя грань, которая либо принадлежит множеству, либо к ней сходится некоторая последовательность отличных от нее чисел из данного множества. [2] Итак, пусть $z_k \in \Pi$ и $\Phi(z_k) \to M^*$. По теореме Больцано-Вейерштрасса, имеется подпоследовательность z_{k_i}, сходящаяся к некоторой точке $z^* \in \Pi$. В силу непрерывности,

$$M^* \ = \ \lim_{k \to \infty} \Phi(z_{k_i}) \ = \ \Phi(z^*).$$

Очевидно, что функция $\Psi(z) = -\Phi(z)$ ограничена сверху тогда и только тогда, когда $\Phi(z)$ ограничена снизу. Значит, нами доказано также существование *нижней грани* для $\Phi(z)$. Опираясь на уже доказанное утверждение, заключаем, что для $\Psi(z)$ существует точка $z_* \in \Pi$ такая, что $\Psi(z) \le \Psi(z_*)$ для всех $z \in \Pi$. Отсюда $\Phi(z_*) \le \Phi(z)$ для всех $z \in \Pi$. $\square$

17.4 Свойства модуля многочлена

Рассмотрим произвольный многочлен

$$f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n \qquad (\#)$$

[2] Данный факт доказывается в курсе математического анализа.

с комплекными коэффициентами и старшим коэффициентом $a_n = 1,\ n \geq 1$.

Лемма о непрерывности модуля многочлена. *Функция* $\Phi(z) = |f(z)|$ *непрерывна при всех* $z \in \mathbb{C}$.

Доказательство. Для доказательства непрерывности $\Phi(z)$ в точке $z = z_0$ достаточно установить непрерывность функции $\Phi(z_0+h)$ от $h \in \mathbb{C}$ в точке $h = 0$. Ясно, что $f(z_0+h)$ есть многочлен от h:

$$f(z_0 + h) = b_0 + b_1 h + \ldots + b_{n-1}h^{n-1} + h^n, \qquad \text{где} \quad b_0 = f(z_0).$$

Отсюда находим

$$\begin{aligned} |\Phi(z_0 + h) - \Phi(z_0)| &= ||f(z_0 + h)| - |f(z_0)|| \leq |f(z_0 + h) - f(z_0)| \\ &\leq |b_1 h + \ldots + b_{n-1}h^{n-1} + h^n| \\ &\leq |b_1||h| + \ldots + |b_{n-1}||h|^{n-1} + |h|^n. \qquad \square \end{aligned}$$

Лемма о росте модуля многочлена. *Для любого числа* $M > 0$ *существует* $R > 0$ *такое, что из неравенства* $|z| \geq R$ *вытекает, что* $|f(z)| \geq M$.

Доказательство. Учитывая, что $|z^i| = |z|^i$, получаем

$$|f(z)| \geq |z^n| - |a_0 + a_1 z + \ldots + a_{n-1}z^{n-1}| \geq |z|^n - |a_0| - |a_1||z| - \ldots - |a_{n-1}||z|^{n-1}.$$

Обозначим через A максимальное из чисел $|a_0|, \ldots, |a_{n-1}|$. Тогда при $|z| \geq 1$ находим

$$|f(z)| \geq |z|^n \left(1 - \frac{nA}{|z|}\right).$$

Для любого заданного $M > 0$ положим

$$R = \max\{1,\ 2nA,\ \sqrt[n]{2M}\}.$$

Легко видеть, что если $|z| \geq R$, то

$$|f(z)| \geq R^n \left(1 - \frac{nA}{2nA}\right) = R^n/2 \geq \frac{2M}{2} = M. \qquad \square$$

17.5 Основная теорема алгебры

Пусть $f(z)$ — произвольный многочлен вида (#).

Лемма Даламбера. *Если в некоторой точке* $z \in \mathbb{C}$ *выполняется неравенство* $|f(z)| > 0$, *то найдется* $h \in \mathbb{C}$ *такое, что* $|f(z + h)| < |f(z)|$.

Доказательство. Утверждение очевидно в случае $n = 1$. Поэтому предположим, что $n \geq 2$. Фиксируем $z \in \mathbb{C}$ и рассмотрим $f(z + h)$ как многочлен от h:

$$f(z + h) = f(z) + b_1 h + \ldots + b_{n-1}h^{n-1} + h^n.$$

Пусть b_m — первый ненулевой коэффициент ($\Rightarrow\ b_1 = \ldots = b_{m-1} = 0$). Тогда

$$f(z + h) = f(z) + b_m h^m + g(h)h^{m+1}, \qquad g(h) = b_{m+1} + \ldots + b_{n-1}h^{n-m-2} + h^{n-m-1}.$$

Определим комплексное число ζ равенством $\zeta^m = -f(z)/b_m$ и будем искать h в виде

$$h = \zeta t, \quad t > 0.$$

Ясно, что

$$|f(z) + b_m h^m| \;=\; |f(z)(1 - t^m)| \;=\; |f(z)|(1 - t^m) \;<\; |f(z)| \quad \text{при} \quad t > 0.$$

При этом на отрезке $0 \le t \le 1$ для некоторого $B > 0$ имеем

$$|g(\zeta t)\,(\zeta t)^{m+1}| \;\le\; Bt^{m+1}.$$

Следовательно, если $0 < t \le 1$, то

$$|f(z + \zeta t h)| \;<\; |f(z)|(1 - t^m) + Bt^{m+1} \;=\; |f(z)| \;+\; (Bt - |f(z)|)\,t^m.$$

При $0 < t \le \min(1,\ |f(z)|/B)$ получаем $|f(z + \zeta t)| \;<\; |f(z)|$. $\square$

Основная теорема алгебры. *Любой многочлен с комплексными коэффициентами степени выше нулевой имеет хотя бы один комплексный корень.*

Доказательство. Пусть $M = |f(0)|$. Если $M = 0$, то все доказано. Предположим, что $M > 0$. Согласно лемме о росте модуля многочлена, при всех $|z| \ge R$ имеем $|f(z)| \ge M$. Рассмотрим квадрат $\Pi = [-R, R] \times [-R, R]$. Функция $|f(z)|$ непрерывна при всех $z \in \mathbb{C}$ и, в частности, при всех $z \in \Pi$. По теореме Вейерштрасса, существует $z_* \in \Pi$ такое, что $|f(z_*)| \le |f(z)|$ при всех $z \in \Pi$. Очевидно, что $|f(z_*)| \le M$ и, кроме того, $M \le |f(z)|$ для любых точек $z \notin \Pi \;\Rightarrow$

$$|f(z_*)| \le |f(z)| \qquad \forall \;\; z \in \mathbb{C}. \tag{$*$}$$

Если $|f(z_*)| > 0$, то, по лемме Даламбера, при некотором $h \in \mathbb{C}$ получаем $|f(z_* + h)| < |f(z)|$, что противоречит неравенствам $(*)$. Таким образом, $|f(z_*)| = 0 \;\Rightarrow\; z_*$ является искомым корнем: $f(z_*) = 0$. $\square$

17.6 Разложение комплексных многочленов

Многочлены первой степени называют также *линейными многочленами.*

Теорема. *Любой комплексный многочлен $f(z)$ степени $n > 0$ разлагается в $\mathbb{C}[z]$ на n линейных множителей:*

$$f(z) \;=\; a\,(z - z_1)\,\ldots\,(z - z_n), \qquad a,\ z_1, \ldots,\ z_n \in \mathbb{C}. \tag{$*$}$$

Данное разложение единственно с точностью до порядка сомножителей.

Доказательство. По основной теореме алгебры, $f(z)$ имеет хотя бы один комплексный корень — пусть это будет z_1. Согласно теореме Безу, многочлен $f(z)$ делится на линейный многочлен $z - z_1$: $f(z) = (z - z_1)f_1(z)$. Если $\deg f_1(z) = 0$, то искомое разложение уже получено. Если $\deg f_1(z) > 0$, то и этот многочлен имеет хотя бы один корень — пусть это будет z_2. Таким образом, $f(z) = (z - z_1)(z - z_2)f_2(z)$. Если $\deg f_2(z) = 0$, то разложение получено. Если нет, то $f_2(z)$ также имеет комплексный корень, и так далее. Ясно, что число a равно старшему коэффициенту многочлена $f(z)$.

Теперь предположим, что имеются два разложения:

$$f(z) \;=\; a\,(z - z_1)\,\ldots\,(z - z_n) \;=\; \widetilde{a}\,(z - \widetilde{z}_1)\,\ldots\,(z - \widetilde{z}_m).$$

Степень многочлена в правой части, очевидно, равна $m \;\Rightarrow\; m = n$. Кроме того, $a = \widetilde{a}$ (это старший коэффициент многочлена $f(z)$). Далее, $(z_1 - \widetilde{z}_1)\ldots(z_1 - \widetilde{z}_n) = 0 \;\Rightarrow$ хотя бы одна из скобок равна нулю $\;\Rightarrow\;$ z_1 совпадает с каким-то из чисел $\widetilde{z}_i$. После перенумерации всегда можно считать, что $z_1 = \widetilde{z}_1$. Итак,

$$(z - z_1)\ \big((z - z_2)\ldots(z - z_n)\ -\ (z - \widetilde{z}_2)\ldots(z - \widetilde{z}_n)\big)\ =\ 0.$$

Отсутствие в $\mathbb{C}[z]$ делителей нуля означает, что

$$(z - z_2)\ldots(z - z_n)\ =\ (z - \widetilde{z}_2)\ldots(z - \widetilde{z}_n).$$

Рассуждая аналогичным образом, приходим (после перенумерации корней) к равенству $z_2 = \widetilde{z}_2$, и так далее. $\square$

Следствие. *Любой комплексный многочлен $f(x)$ степени $n > 0$ имеет единственное разложение вида*

$$f(z) = a\,(z - \zeta_1)^{k_1}\ \ldots,\ (z - \zeta_m)^{k_m}, \qquad k_1,\ \ldots,\ k_m > 0, \qquad k_1 + \ldots + k_m = n, \qquad (**)$$

$$\zeta_i \neq \zeta_j \quad \textit{при} \quad i \neq j, \qquad a,\ \zeta_1,\ \ldots,\ \zeta_m \in \mathbb{C}.$$

Разложение вида $(**)$ иногда называется *комплексным каноническим* разложением многочлена $f(z)$. Число k_i называется *кратностью* корня ζ_i. Корень ζ_i называется *кратным*, если $k_i > 1$, и *простым*, если $k_i = 1$.

Согласно $(**)$, многочлен $f(z)$ имеет m попарно различных корней. В разложении $(*)$ некоторые из чисел $z_1, \ldots, z_m$ могут совпадать: если $z_i = \zeta_j$, то имеется ровно k_j чисел, равных ζ_j. Нередко полученную выше теорему формулируют таким образом: *любой комплексный многочлен степени $n > 0$ имеет ровно n комплексных корней с учетом кратностей.*

17.7 Разложение вещественных многочленов

Рассмотрим *вещественный многочлен* (многочлен с вещественными коэффициентами) $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ и предположим, что число $z \in \mathbb{C}$ является его корнем. Тогда комплексно сопряженное число $\overline{z}$ также является корнем (в силу вещественности коэффициентов $\overline{a_i} = a_i$ для всех i):

$$f(\overline{z})\ =\ a_0 + a_1\overline{z} + \ldots + a_n\overline{z}^n\ =\ \overline{a_0} + \overline{a_1 z} + \ldots + \overline{a_n z^n}\ =\ \overline{f(z)}\ =\ 0.$$

Если $\overline{z} \neq z$, то *квадратичный многочлен* (многочлен степени 2)

$$\phi(x)\ =\ (x - z)(x - \overline{z})\ =\ x^2 - (z + \overline{z})\,x + |z|^2$$

имеет, очевидно, вещественные коэффициенты и является неразложимым в $\mathbb{R}[x]$.

Теорема. *Любой вещественный многочлен $f(x)$ степени $n > 0$ разлагается в $\mathbb{R}[x]$ на линейные и неразложимые квадратичные множители:*

$$f(x) = a(x - x_1)\ldots(x - x_M)\ \phi_1(x)\ldots\phi_N(x), \qquad M + 2N = n,$$

$$a,\ x_1,\ \ldots,\ x_M \in \mathbb{R},$$

$$\phi_i(x) = x^2 + s_i x + t_i, \quad s_i,\, t_i \in \mathbb{R}, \quad i = 1,\, \ldots,\, N.$$

Данное разложение единственно с точностью до порядка сомножителей.

Доказательство. Многочлен $f(x)$ имеет n комплексных корней $z_1,\ \ldots,\ z_n$ с учетом кратностей. Пусть ровно M из них являются вещественными. Тогда остальные $n - M$ корней разбиваются на пары комплексно сопряженных чисел ($\Rightarrow$ число $n - M$ должно быть четным: $n - M = 2N$). Вещественные корни дают M линейных множителей, а пары комплексно сопряженных чисел дают N неразложимых квадратичных множителей. Тем самым существование искомого разложения доказано. Допустим, что имеются два разложения такого вида:

$$f(x) = a(x - x_1)\ldots(x - x_M)\ \phi_1(x)\ldots\phi_N(x) = \widetilde{a}(x - \widetilde{x}_1)\ldots(x - \widetilde{x}_{M'})\ \widetilde{\phi}_1(x)\ldots\widetilde{\phi}_{N'}(x).$$

Ясно, что $a = \widetilde{a}$ (это старший коэффициент $f(x)$). Далее, полный набор комплексных корней с учетом кратностей определен однозначно $\Rightarrow$ вещественные корни с учетом кратностей определены однозначно $\Rightarrow$

$$a(x - x_1)\ldots(x - x_M) = \widetilde{a}(x - \widetilde{x}_1)\ldots(x - \widetilde{x}_{M'}), \quad M = M' \quad \Rightarrow \quad N = N'.$$

Поскольку в $\mathbb{R}[x]$ делителей нуля нет, получаем

$$\phi_1(x)\ldots\phi_N(x) \ = \ \widetilde{\phi}_1(x)\ldots\widetilde{\phi}_N(x).$$

Пусть $\phi_1(z) = 0 \quad \Rightarrow \quad \phi_1(x) = (x - z)(x - \overline{z})$. Далее, $\widetilde{\phi}_1(z)\ldots\widetilde{\phi}_N(z) = 0 \quad \Rightarrow$ хотя бы один из множителей равен нулю. Пусть, например, $\widetilde{\phi}_1(z) = 0 \ \Rightarrow \ \widetilde{\phi}_1(x) = (x - z)(x - \overline{z})$. Таким образом,

$$\phi_1(x) = \widetilde{\phi}_1(x) \quad \Rightarrow \quad \phi_2(x)\ldots\phi_N(x) \ = \ \widetilde{\phi}_2(x)\ldots\widetilde{\phi}_N(x).$$

Далее по индукции. $\square$

Следствие. *Любой вещественный многочлен нечетной степени имеет хотя бы один вещественный корень.*

Замечание. Последнее утверждение можно было бы доказать и непосредственно — без использования основной теоремы алгебры. Достаточно доказать, что $f(x)$ (как функция от $x \in \mathbb{R}$) имеет положительный знак при достаточно больших положительных x и отрицательный знак при достаточно больших отрицательных x. После этого использовать непрерывность $f(x)$ и теорему Ролля.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

17.8 Кратные корни и производные

Производной многочлена $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ называется многочлен

$$f'(x) \ = \ a_1 \ + \ 2a_2 x \ + \ \ldots \ + \ n a_n x^{n-1}.$$

Рассматривая $f(x)$ как функцию от x (например, в случае вещественных коэффициентов) и вычисляя производную по правилам математического анализа, мы получим,

очевидно, функцию, совпадающую с $f'(x)$.

Утверждение. *Многочлен $f(x)$ над числовым полем $K \subset \mathbb{C}$ имеет только простые корни тогда и только тогда, когда многочлены $f(x)$ и $f'(x)$ взаимно просты.*

Доказательство. Пусть $f(x)$ имеет корень θ кратности k. Тогда

$$f(x) = (x-\theta)^k g(x), \quad g(\theta) \neq 0. \qquad \Rightarrow \qquad f'(x) = k(x-\theta)^{k-1} g(x) + (x-\theta)^k g'(x).$$

При $k \geq 2$ находим $f'(\theta) = 0$. Поэтому θ является общим корнем многочленов $f(x)$ и $f'(x) \quad \Rightarrow \quad$ их наибольший общий делитель имеет степень ≥ 1. $\square$

Важное наблюдение: если $f(x) \in K[x]$, то $f'(x) \in K[x]$. Поэтому все коэффициенты их наибольшего общего делителя принадлежат тому же полю K. Отсюда получам полезное

Следствие. *Минимальный многочлен над полем $K \subset \mathbb{C}$ для любого числа $\theta \in \mathbb{C}$ имеет только простые корни.*

17.9 Поле разложения

Рассмотрим многочлен

$$f(x) = a_0 + \ldots + a_{n-1}x^{n-1} + x^n = \prod_{i=1}^{n}(x - x_i) \in K[x], \qquad K \subset \mathbb{C}.$$

Поле $L = K(x_1, \ldots, x_n)$ называется *полем разложения* многочлена $f(x)$. Конечно, L может быть получено из K путем последовательного присоединения корней $x_1, \ldots, x_n$:

$$K \subset K(x_1) \subset K(x_1)(x_2) \subset \ldots \subset K(x_1)(x_2)\ldots(x_n) = L.$$

В действительности поле L можно получить из K присоединением всего лишь какого-то одного числа $\theta \in L$ (вообще говоря, θ отлично от корней $f(x)$). Данный результат получается с помощью последовательного применения следующей леммы.

Лемма. *Пусть α и β являются корнями многочленов над полем $K \subset \mathbb{C}$. Тогда*

$$K(\alpha)(\beta) \;=\; K(\theta)$$

для какого-то числа $\theta \in K(\alpha)(\beta)$.

Доказательство. Пусть $F(x)$ — минимальный многочлен над K для α, имеющий (как мы знаем, простые) корни $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_k$, а $G(x)$ — минимальный многочлен над K для β, имеющий корни $\beta_1 = \beta, \beta_2, \ldots, \beta_m$. Число θ попытаемся найти в виде

$$\theta = \alpha_1 + c\beta_1, \qquad c \neq 0, \quad c \in K,$$

причем выберем c так, чтобы

$$(\theta - \alpha_i)/c \neq \beta_j \quad \text{при всех } i, j, \text{ кроме } i = j = 1.$$

Тогда многочлен $\Phi(x) = G((\theta - x)/c)$ имеет своим корнем α_1, но не $\alpha_2, \ldots, \alpha_k$. Значит, $\Phi(x)$ и $F(x)$ имеют в точности один общий корень α_1. Поэтому их наибольший общий делитель равен $x - \alpha_1$. Но он является многочленом над $K(\theta)$ (поскольку таковы $\Phi(x)$ и $F(x)$). Отсюда $\alpha_1 \in K(\theta)$. $\square$

17.10 Корни многочленов над произвольным полем

Пусть задан многочлен над абстрактным полем P. Он может не иметь корней в P, но получить их в более широком поле F. Всегда ли найдется поле F с таким свойством?

Мы уже знаем, что для комплексных многочленов ответ положительный. Это можно доказать и для произвольного поля, причем легче, чем основную теорему алгебры (потому что в последней F является заранее предписанным полем).

Теорема о существовании корня. *Для произвольного многочлена над полем P, имеющего степень выше нулевой, существует расширение поля P, в котором он имеет корень.*

Доказательство. Рассмотрим многочлен $f(x) \in P[x]$ степени $n \geq 1$ и введем следующее бинарное отношение на множестве $P[x]$: $u(x) \sim v(x)$, если $u(x)$ и $v(x)$ имеют одинаковые остатки от деления на $f(x)$. Легко проверить, что это есть отношение эквивалентности. Поэтому все множество многочленов над P разбивается на непересекающиеся классы эквивалентности. Класс многочленов, эквивалентных $u(x)$, обозначим через $[u(x)]$, а все множество классов эквивалентности — через F.

Данная конструкция напоминает вычеты по модулю n, поэтому каждый класс эквивалентных многочленов будем также называть *вычетом* относительно многочлена $f(x)$. Вычетов ровно столько, сколько имеется разных остатков от деления на $f(x)$ — не меньше, чем элементов в поле P (разные многочлены нулевой степени принадлежат, очевидно, разным вычетам).

Определим операции сложения и умножения элементов из F:

$$[u(x)] + [v(x)] = [u(x) + v(x)], \qquad [u(x)][v(x)] = [u(x)\,v(x)], \qquad u(x), v(x) \in P[x].$$

Легко проверяется, что их результаты не зависят от выбора конкретных представителей в классах $[u(x)]$ и $[v(x)]$ и что данные операции превращают F в кольцо.

Не ограничивая общности, предположим, что $f(x)$ является неразложимым над полем P. Тогда, опять-таки по аналогии с вычетами по простому модулю n, множество F оказывается полем. В самом деле, роль единичного элемента, очевидно, выполняет вычет [1], порождаемый константой (многочленом нулевой степени) 1. Рассмотрим ненулевой вычет $[u(x)] \in F$. Многочлены $u(x)$ и $f(x)$ взаимно просты в силу неразложимости $f(x)$. По теореме о наибольшем общем делителе, существуют многочлены $\phi(x), \psi(x) \in P[x]$ такие, что

$$u(x)\phi(x) + f(x)\psi(x) = 1 \quad \Rightarrow \quad [u(x)][\phi(x)] = [1].$$

Вычет [a], порожденный многочленом нулевой степени (константой) $a \in P$, будем отождествлять с a. Таким образом, $P \subset F$, а многочлен $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_0$ можно рассматривать также как многочлен над полем F:

$$f(x) = x^n + [a_{n-1}]x^{n-1} + \ldots + [a_0] \ \in \ F[x].$$

Тогда

$$0 = [0] = [f(x)] = f([x]).$$

Это означает, что вычет $[x] \in F$ является корнем многочлена $f(x)$. $\square$

Следствие. *Для любого многочлена $f(x) \in P[x]$ степени $n > 0$ существует расширение F поля P, в котором $f(x)$ разлагается на n линейных множителей:*

$$f(x) = a(x - z_1)\ldots(x - z_n), \qquad a \in P, \quad z_1, \ldots, z_n \in F.$$

Лекция 18

ОСНОВНАЯ ЧАСТЬ

18.1 Формулы Виета

Рассмотрим комплексный многочлен $f(x)$ степени n со старшим коэффициентом 1 и его разложение на линейные множители:

$$f(x) \;=\; a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n \;\;=\;\; (x - x_1)\ldots(x - x_n).$$

Раскрывая скобки в правой части и приравнивая коэффициенты при одинаковых степенях x, получаем *формулы Виета*:

$$\begin{array}{rcl}
a_{n-1} &=& -(x_1 + x_2 + \ldots + x_n),\\
a_{n-2} &=& \;\;(x_1x_2 + x_1x_3 + \ldots + x_{n-1}x_n),\\
a_{n-3} &=& -(x_1x_2x_3 + x_1x_2x_4 + \ldots + x_{n-2}x_{n-1}x_n),\\
\ldots & \ldots & \ldots\\
a_{n-k} &=& (-1)^k \sum\limits_{1\le i_1 < \ldots < i_k \le n} x_{i_1}\,\ldots\,x_{i_k},\\
\ldots & \ldots & \ldots\\
a_0 &=& (-1)^n\;x_1\ldots x_n.
\end{array}$$

Выражения вида

$$\sigma_k = \sigma_k(x_1, \ldots, x_n) = \sum_{1\le i_1 < \ldots < i_k \le n} x_{i_1}\,\ldots\,x_{i_k}, \qquad k = 1, \ldots, n, \qquad (*)$$

называются *элементарными симметрическими многочленами* от $x_1, \ldots, x_n$. Таким образом, коэффициенты многочлена $f(x)$ выражаются через элементарные симметрические многочлены от его корней $x_1, \ldots, x_n$:

$$a_{n-k} \;=\; (-1)^k\;\sigma_k, \qquad k = 1, \ldots, n.$$

18.2 Многочлены от n переменных

Формальное выражение $x_1^{\alpha_1}\,\ldots\,x_n^{\alpha_n}$, где $\alpha_1, \ldots, \alpha_n$ — неотрицательные целые степени, называется *одночленом степени* $\alpha_1 + \ldots + \alpha_n$ *от переменных* $x_1, \ldots, x_n$. Равенство $\alpha_i = 0$ допускается (в этом случае одночлен не содержит x_i).

Многочленом от переменных $x_1, \ldots, x_n$ *над полем* P называется формальная сумма одночленов от $x_1, \ldots, x_n$ с коэффициентами из поля P. *Степенью* многочлена называется наивысшая степень входящих в него одночленов. Например, многочлен

$$f(x_1, x_2, x_3) \;=\; x_1^3x_2^2x_3 + x_1^2x_2^3x_3 + x_1x_2x_3^4 + x_1x_2 + x_3$$

имеет степень 6. Как видим, в состав $f(x_1, x_2, x_3)$ входят 3 одночлена наивысшей степени.

Полагаем $x_1^{\alpha_1} \dots x_n^{\alpha_n} = x_1^{\beta_1} \dots x_n^{\beta_n}$, если $\alpha_i = \beta_i$ для всех i. Один и тот же многочлен допускает много формально различных представлений в виде суммы одночленов с коэффициентами из поля P. Однако мы всегда можем перейти к *стандартному представлению*, в котором каждый одночлен встречается только один раз — процедура перехода называется *приведением подобных членов* и заключается в замене всех одинаковых одночленов одним одночленом с коэффициентом, равным сумме соответствующих коэффициентов. Многочлены f и g называются равными, если они имеют равные коэффициенты для равных одночленов в своих стандартных представлениях.

Суммой многочленов $f+g$ называется многочлен с коэффициентами, равными сумме коэффициентов для соответствующих одночленов, входящих в f и g. Произведением многочленов $f = \sum a_{\alpha_1,\dots,\alpha_n} x_1^{\alpha_1} \dots x_n^{\alpha_n}$ и $g = \sum b_{\beta_1,\dots,\beta_n} x_1^{\beta_1} \dots x_n^{\beta_n}$ называется многочлен fg, состоящий из всех членов вида

$$(a_{\alpha_1,\dots,\alpha_n}\, b_{\beta_1,\dots,\beta_n})\; x_1^{\alpha_1+\beta_1} \dots x_n^{\alpha_n+\beta_n}.$$

Таким образом, умножение многочленов выполняется по привычным правилам раскрытия скобок и приведения подобных членов.

Множество всех многочленов от $x_1, \dots, x_n$ над полем P обозначается через $P[x_1, \dots, x_n]$. Относительно операций сложения и умножения многочленов оно является коммутативным кольцом с единицей и без делителей нуля.

18.3 Лексикографическое упорядочение

При изучении многочленов от $x_1, \dots, x_n$ часто используется *лексикографическое* (*словарное*) упорядочение входящих в них одночленов:

$$x_1^{\alpha_1} \dots x_n^{\alpha_n} \quad \text{старше (выше)} \quad x_1^{\beta_1} \dots x_n^{\beta_n},$$

если для некоторого $1 \le k \le n$ выполняются соотношения

$$\alpha_1 = \beta_1, \quad \dots, \quad \alpha_{k-1} = \beta_{k-1}, \quad \alpha_k > \beta_k.$$

В дальнейшем под *старшим членом* многочлена будет пониматься взятый с соответствующим ненулевым коэффициентом одночлен, являющийся наивысшим при лексикографическом упорядочении одночленов стандартного представления данного многочлена. Очевидно, старший член определен однозначно.

Легко проверяется, что старший член произведения двух многочленов равен произведению их старших членов.

18.4 Симметрические многочлены

Многочлен $f(x_1, \dots, x_n)$ называется *симметрическим*, если для любой подстановки σ степени n

$$f(x_1, \dots, x_n) \;=\; f(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Важными примерами симметрических многочленов являются элементарные симметрические многочлены σ_k, присутствующие в формулах Виета.

Теорема о симметрических многочленах. *Для любого симметрического многочлена $f(x_1, \dots, x_n)$ существует и единствен многочлен g от n переменных такой,*

что

$$f(x_1, \ldots, x_n) \;=\; g(\sigma_1, \ldots, \sigma_n),$$

где $\sigma_k = \sigma_k(x_1, \ldots, x_n)$ *— элементарные симметрические многочлены вида* $(*)$.

Доказательство. Пусть $a\, x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ — старший член многочлена $f(x_1, \ldots, x_n)$. Тогда в случае симметрического многочлена обязательно выполняются неравенства $\alpha_1 \geq \ldots \geq \alpha_n$. Если бы это было не так, то данный член не был бы старшим: в симметрическом многочлене вместе с одночленом $x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ должны присутствовать все одночлены вида $x_{\sigma(1)}^{\alpha_1} \ldots x_{\sigma(n)}^{\alpha_n}$ для любой подстановки σ. Рассмотрим многочлен

$$\phi(\sigma_1, \ldots, \sigma_n) \;=\; a\, \sigma_1^{\alpha_1 - \alpha_2} \ldots \sigma_{n-1}^{\alpha_{n-1} - \alpha_n} \sigma_n^{\alpha_n}. \qquad (1)$$

Его можно рассматривать также как многочлен от $x_1, \ldots, x_n$, для которого старший член будет, очевидно, равен

$$a\, x_1^{\alpha_1 - \alpha_2} (x_1 x_2)^{\alpha_2 - \alpha_3} \ldots (x_1 \ldots x_{n-1})^{\alpha_{n-1} - \alpha_n} (x_1 \ldots x_{n-1} x_n)^{\alpha_n} \;=\; a\, x_1^{\alpha_1} \ldots x_n^{\alpha_n}. \qquad (2)$$

Поэтому старший член многочлена

$$f_1(x_1, \ldots, x_n) \;=\; f(x_1, \ldots, x_n) - \phi(\sigma_1, \ldots, \sigma_n)$$

будет младше старшего члена для $f(x_1, \ldots, x_n)$. Аналогичным образом от f_1 можно перейти к многочлену f_2 с меньшим старшим членом, и так далее. В силу конечности общего числа членов данная процедура должна на каком-то шаге дать нулевой многочлен.

Для доказательства единственности многочлена g достаточно показать, что если $g(\sigma_1, \ldots, \sigma_n) \neq 0$, то и $f(x_1, \ldots, x_n) \neq 0$. Другими словами, нужно проверить, что после замены σ_k на соответствующие многочлены от $x_1, \ldots, x_n$ и приведения подобных членов останется хотя бы один ненулевой член. Любой член многочлена g можно записать в виде (1) с $\alpha_1 \geq \ldots \geq \alpha_n$. Как многочлен от $x_1, \ldots, x_n$, многочлен ϕ имеет своим старшим членом (2). Старшим членом для g как многочлена от $x_1, \ldots, x_n$ будет наивысший из членов такого вида. Он определен однозначно и поэтому не может сократиться при приведении подобных членов. $\square$

Следствие. *Значение любого симметрического многочлена* $\phi(x_1, \ldots, x_n)$ *при замене переменных на корни многочлена* $f(x) = a_0 + a_1 x + \ldots + a_{n-1} x_{n-1} + x^n$ *над полем* P *является элементом поля* P.

Доказательство. Симметрический многочлен является многочленом от элементарных симметрических многочленов. Если считать переменные корнями для $f(x)$, то, в силу формул Виета, он будет многочленом над тем же полем P от коэффициентов $a_0, \ldots, a_{n-1}$, которые являются элементами поля P. $\square$

18.5 Ньютоновы суммы

Пусть задан многочлен $f(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + x^n$, и пусть $x_1, \ldots, x_n$ — все его корни с учетом кратностей. Выражения

$$s_k \;=\; x_1^k + x_2^k + \ldots + x_n^k, \qquad k = 1,\ 2,\ \ldots,$$

называются *ньютоновыми суммами* для $f(x)$.

Ясно, что s_k — симметрический многочлен от корней $x_1, \ldots, x_n$. Поэтому s_k есть значение многочлена от элементарных симметрических многочленов и, следовательно, от коэффициентов $a_0, \ldots, a_{n-1}$. Таким образом, ньютоновы суммы конструктивно выражаются через коэффициенты многочлена $f(x)$ — их можно найти, не зная корни.

На вычислении ньютоновых сумм легко построить также некоторый метод приближенного вычисления корней многочлена $f(x)$. [1] Предположим, что

$$|x_1| > |x_2| \geq \ldots \geq |x_n|.$$

Тогда

$$\frac{s_{k+1}}{s_k} \;=\; x_1\,\frac{1+\left(\frac{x_2}{x_1}\right)^{k+1}+\ldots+\left(\frac{x_n}{x_1}\right)^{k+1}}{1+\left(\frac{x_2}{x_1}\right)^{k}+\ldots+\left(\frac{x_n}{x_1}\right)^{k}} \;\to\; x_1 \qquad \text{при} \qquad k\to\infty.$$

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

18.6 Еще одно доказательство основной теоремы алгебры

Доказательство на основе симметрических многочленов и формул Виета сложнее того, что уже обсуждалось. Но оно использует понятие непрерывности "минимальным" образом.

(1) Пусть $f(x)$ — многочлен степени $n>0$ с вещественными коэффициентами. Мы знаем, что в некотором поле F он разлагается на линейные множители и, следовательно, имеет n корней $x_1, \ldots, x_n$ с учетом кратностей. Наша цель — доказать, что хотя бы один из этих корней является комплексным числом.

(2) Если n нечетно, что данный факт получается очень легко — это единственное место, где используется непрерывность. Легко видеть, что $f(x)$ — непрерывная функция от x. Поскольку n нечетно, многочлен $f(x)>0$ при $x\geq b$ для некоторого $b>0$ и $f(x)<0$ при $x\leq a$ для некоторого $a<0$. По теореме Ролля из математического анализа, существует число $c\in[a,b]$ такое, что $f(c)=0$.

(3) Предположим, что $n=2^k p$, где p нечетно, и будем вести индукцию по k. При $k=0$ существование комплексного (даже вещественного) корня уже доказано. Пусть $k>0$. Тогда возьмем произвольное вещественное число c и рассмотрим многочлен

$$\mathcal{F}_c(x) \;=\; \prod_{1\leq i<j\leq n}(x-x_{ij}^c), \qquad x_{ij}^c \;=\; c\,x_ix_j+x_i+x_j.$$

В силу формул Виета и определения x_{ij}, коэффициенты $\mathcal{F}_c(x)$ — симметрические функции от корней вещественного многочлена $f(x)$ $\;\Rightarrow\;$ они вещественны. Степень $\mathcal{F}_c(x)$ равна $(n^2-n)/2=2^{k-1}q$, где $q=(2^kp-1)p$ — нечетное число. Поэтому, согласно предположению индукции, многочлен $\mathcal{F}_c(x)$ имеет хотя бы один комплексный корень — пусть он получается при $i=i(c),\;\; j=j(c)$.

(4) Вещественных чисел c бесконечно много, а индексы $i(c), j(c)$ могут принимать лишь конечное число значений $\;\Rightarrow\;$ для некоторых вещественных чисел $c_1\neq c_2$ имеют место равенства $i=i(c_1)=i(c_2),\;\; j=j(c_1)=j(c_2)$. $\;\Rightarrow$

$$\left\{\begin{array}{rcl} c_1\,x_ix_j+x_i+x_j &=& z_1\in\mathbb{C}\\ c_2\,x_ix_j+x_i+x_j &=& z_2\in\mathbb{C}\end{array}\right. \quad\Rightarrow\quad x_ix_j=\frac{z_1-z_2}{c_1-c_2}\in\mathbb{C} \quad\Rightarrow\quad x_i+x_j\in\mathbb{C}.$$

[1] На практике для этой цели все же используются другие методы — с более быстрой сходимостью.

Следовательно, x_ix_j и x_i+x_j являются корнями квадратного уравнения с комплексными коэффициентами $\Rightarrow \quad x_i, x_j \in \mathbb{C}$.

(5) Итак, доказано, что любой вещественный многочлен степени $n > 0$ имеет хотя бы один комплексный корень. Пусть

$$f(x) = a_0 + a_1x + \ldots a_{n-1}x^{n-1} + x^n$$

имеет комплексные коэффициенты. Тогда введем "сопряженный" многочлен

$$\bar{f}(x) = \bar{a_0} + \bar{a}_1x + \ldots + \bar{a}_{n-1}x^{n-1} + x^n$$

и рассмотрим многочлен $g(x) = f(x)\bar{f}(x)$. Легко проверить, что $g(x)$ имеет вещественные коэффициенты. По доказанному выше, он имеет комплексный корень z. Таким образом,

$$g(z) = f(z)\bar{f}(z) = f(z)\overline{f(\bar{z})} = 0 \qquad \Rightarrow \qquad f(z) = 0 \quad \text{или} \quad f(\bar{z}) = 0. \qquad \square$$

18.7 Нормальные поля и поля разложения

Формулы Виета и теорема о симметрических многочленах с большой пользой применяются при изучении расширений полей, содержащих корни тех или иных многочленов.

Фиксируем числовое поле $K \subset \mathbb{C}$ и будем рассматривать его конечные расширения $K \subset L$. Последнее означает, что поле L можно рассматривать как конечномерное линейное пространство над полем K. Отсюда вытекает, что в L любой элемент является корнем некоторого неразложимого многочлена над K.

В теории Галуа особый интерес вызывают *нормальные расширения*. Это конечные расширения $K \subset L$ с особым свойством: если хотя бы один корень неразложимого над K многочлена степени n принадлежит L, то все его n комплексных корней принадлежат L. В таких случаях говорят также, что L является *нормальным полем над* K или *нормально над* K.

Пусть $L = K(\theta_1, \ \ldots, \ \theta_n)$ — поле разложения некоторого (возможно, разложимого) многочлена $f(x) \in K[x]$ степени n.

Теорема. *Поле разложения L любого многочлена над K является нормальным над K, а любое нормальное над K поле является полем разложения некоторого многочлена над K.*

Доказательство. Пусть $L = K(\theta_1, \ \ldots, \ \theta_n)$ — поле разложения многочлена

$$f(x) = (x - \theta_1)\ldots(x - \theta_n) \in K[x]$$

(корни в поле L, а коэффициенты принадлежат меньшему полю K). Ясно, что поле L можно получить последовательным присоединением отдельных корней. Из теоремы о присоединении корня (из Лекции 15) легко вывести, что любой элемент $\alpha \in L$ имеет вид $\alpha = g(\theta_1, \ \ldots, \ \theta_n)$, где $g(x_1, \ \ldots, \ x_n)$ — многочлен от n переменных с коэффициентами из поля K. Рассмотрим следующий многочлен:

$$\Psi(x) = \prod_{\sigma \in S_n} (x - g(\theta_{\sigma(1)}, \ \ldots, \ \theta_{\sigma(n)})).$$

В силу формул Виета и теоремы о симметрических многочленах, его коэффициенты принадлежат полю K.

Докажем нормальность поля L. Пусть $\alpha \in Ł$ — корень неразложимого многочлена $\phi(x) \in K[x]$ и β — любой другой корень $\phi(x)$. Поскольку $\phi(x)$ и $\Psi(x)$ имеют общий корень α, он является также корнем их наибольшего общего делителя. В силу алгоритма Евклида, коэффициенты наибольшего общего делителя принадлежат полю K. Поэтому он лишь ненулевым множителем может отличаться от $\phi(x)$ (в силу неразложимости $\phi(x)$). Значит, $\Psi(x)$ делится на $\phi(x) \quad \Rightarrow \quad \beta$ имеет вид $\beta = g(\theta_{\sigma(1)}, \ \ldots, \ \theta_{\sigma(n)})$ для какой-то подстановки $\sigma \in S_n \quad \Rightarrow \quad \beta \in L$.

Вторая часть утверждения доказывается очевидным образом. $\square$

18.8 Радикальные расширения

Рассмотрим алгебраическое уравнение $f(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} + x^n = 0$ с коэффициентами из числового поля $K \subset \mathbb{C}$. Пусть L — поле разложения $f(x)$.

Вопрос об формуле, выражающей корни $f(x)$ через коэффициенты с помощью арифметических операций и операций извлечения корня любой предписанной степени (короче, *в радикалах*), сводится к вопросу о существовании конечной цепочки так называемых *радикальных* расширений

$$K = K_0 \subset K_1 \subset K_2 \subset \ldots \subset K_m = \widetilde{L}, \qquad K_i = K_{i-1}(\theta_i), \quad \theta_i^{n_i} = D_i \in K_{i-1}, \qquad (*)$$

где $\widetilde{L}$ — расширение поля L, являющееся нормальным над K. [2] В теории Галуа данный вопрос сводится к изучению группы $\mathrm{Aut}(L, K)$ *автоморфизмов L над K* (взаимно-однозначных отображений L на себя, сохраняющих операции и оставляющих на месте все элементы поля K) и ее подгрупп.

Простейший пример: $f(x) = x^n - a = 0$, $a \in K$. В данном случае очевидно, что $L = K(\varepsilon, \zeta)$, где ε — первообразный корень из единицы степени n, а ζ — любой число такое, что $\zeta^n = a$. Пусть Φ — произвольный автоморфизм L над K. Тогда $\Phi(\varepsilon)$ также является корнем из единицы степени n.

Предположим, что n — простое число. В этом случае мы уже имеем описание всех автоморфизмов поля $\mathbb{Q}(\varepsilon)$ над $\mathbb{Q}$ (см. Лекцию 15): любой автоморфизм $\Psi \in \mathrm{Aut}(\mathbb{Q}(\varepsilon), \mathbb{Q})$ однозначно определяется заданием образа ε^k для ε, а группа $\mathrm{Aut}(\mathbb{Q}(\varepsilon), \mathbb{Q})$ является циклической группой порядка $n-1$.

Возьмем автоморфизм $\Phi \in \mathrm{Aut}(L, K)$ и поставим ему в соответствие автоморфизм $\Psi \in \mathrm{Aut}(\mathbb{Q}(\varepsilon), \mathbb{Q})$ такой, что $\Psi(\varepsilon) = \Phi(\varepsilon)$. Данное соответствие, как несложно проверить, является гомоморфизмом группы $G = \mathrm{Aut}(L, K)$ в группу $\mathrm{Aut}(\mathbb{Q}(\varepsilon), \mathbb{Q})$. Пусть H — ядро этого гомоморфизма. Тогда, в силу теоремы о гомоморфизме, фактор-группа G/H изоморфна некоторой подгруппе группы $\mathrm{Aut}(\mathbb{Q}(\varepsilon), \mathbb{Q})$. Будучи подгруппой циклической группы, данная подгруппа является циклической.

Таким образом, *если L — поле разложения для $x^n - a$, $a \in K$, то при простом n группа $G = Aut(L, K)$ имеет нормальную подгруппу H с циклической фактор-группой G/H.*

18.9 Автоморфизмы и расширения

Утверждение. *Для любого конечного расширения $K \subset L$ число автоморфизмов L над K не превышает степени расширения.*

Доказательство. L может быть получено из K присоединением какого-то одного числа $\theta \in L$. По теореме о присоединении корня, степень минимального над K многочлена $R(x)$ для θ равна степени расширения $K \subset L$. При любом автоморфизме $g \in \mathrm{Aut}(L, K)$ имеем $R(g(\theta)) = 0$. Поэтому число автоморфизмов не больше числа корней многочлена $R(x)$, принадлежащих полю L. $\square$

Число автоморфизмов L над K модет оказаться меньше степени расширения. Пусть, например, $K = \mathbb{Q}$ и $L = \mathbb{Q}(\sqrt[3]{2})$. Тогда единственным автоморфизмом L над K является тождественное отображение (докажите!), а степень расширения равна 3.

18.10 Расширения Галуа

Если число автоморфизмов равно степени расширения, то соответствующее расширение называется *расширением Галуа*, а группа автоморфизмов — *группой Галуа*.

Утверждение. *Для того чтобы расширение $K \subset L$ было расширением Галуа, необходимо и достаточно, чтобы оно было нормальным.*

Доказательство. Пусть $L = K(\theta)$ и $R(x)$ — минимальный многочлен для θ. Пусть его степень равна m. По теореме о присоединении корня, любой элемент из L имеет вид $\alpha_0 + \alpha_1\theta + \ldots + \alpha_{m-1}\theta^{m-1}$, $\alpha_i \in K$. Заметим, что все корни $R(x)$ простые. Если поле L нормально над K, то все они принадлежат L. Пусть ζ — любой корень $R(x)$. Тогда отображение $\alpha_0 + \alpha_1\theta + \ldots + \alpha_{m-1}\theta^{m-1} \to \alpha_0 + \alpha_1\zeta + \ldots + \alpha_{m-1}\zeta^{m-1}$ является автоморфизмом L над K. Значит, число автоморфизмов равно степени расширения.

[2] Можно доказать, что от цепочки радикальных расширений вида $(*)$, дающей некоторое поле $\widetilde{L}$, всегда можно перейти к цепочке таких радикальных расширений, которая дает в итоге поле, содержащее $\widetilde{L}$ и нормальное над K. Для первого знакомства с данным кругом идей можно полагать для простоты, что изучается случай $\widetilde{L} = L$.

В любом случае число автоморфизмов не больше числа различных корней $R(x)$, принадлежащих полю L. Если все m корней $R(x)$ принадлежат L, то L — поле разложения для $R(x)$ и поэтому является нормальным над K. $\square$

18.11 Промежуточные поля и подгруппы

Пусть H — подгруппа группы $G = \mathrm{Aut}(L, K)$. Обозначим через L^H множество всех элементов $a \in L$ таких, что $h(a) = a \ \ \forall \ \ h \in H$. Легко доказать, что множество L^H является *промежуточным* полем — то есть, полем в цепочке $K \subset L^H \subset L$.

(1) Пусть P — промежуточное поле в цепочке $K \subset P \subset L$. Легко доказать, что из нормальности L над K вытекает его нормальность также над полем P.

(2) Итак, $K \subset L^H \subset L$. Если $K \subset L$ — расширение Галуа, то $L^H \subset L$ — также расширение Галуа $\Rightarrow$ порядок группы H равен степени расширения $(L : L^H)$. Если $H = G$, то $(L : L^H) = (L : K) \ \Rightarrow$ $L^H = K$. Верно и обратное: *если $L^H = K$, то расширение $K \subset L$ есть расширение Галуа и $H = G$.*

Пусть $L = K(\theta)$. Тогда легко убедиться в том, что минимальный для θ многочлен над K имеет вид

$$\Phi(x) = \prod_{g \in G} (x - g(\theta)).$$

Отсюда следует, что L — поле разложения для $\Phi(x)$. Поэтому $K \subset L$ — нормальное расширение. Далее, многочлен

$$\phi(x) = \prod_{h \in H} (x - h(\theta))$$

является многочленом над L^H (в силу все тех же формул Виета и теоремы о симметрических многочленах) и, следовательно, над $K = L^H$. В силу неразложимости минимального многочлена $\Phi(x) = \phi(x)$ $\Rightarrow$ число автоморфизмов в подгруппе H равно числу автоморфизмов в группе $G \ \Rightarrow \ H = G$. $\square$

(3) Для любого промежуточного поля P группа $H = \mathrm{Aut}(L, P)$ является подгруппой группы $G = \mathrm{Aut}(L, K)$. [3]

Теорема. *Если L — нормальное расширение поля K, то $P = L^H$ — нормальное поле над K тогда и только тогда, когда H — нормальный делитель группы $G = Aut(L, K)$; при этом группа $Aut(P, K)$ изоморфна фактор-группе G/H.*

Доказательство. Пусть $P = K(\zeta)$. Тогда любой элемент $a \in P$ имеет вид $a = \sum \alpha_i \zeta^i, \ \ \alpha_i \in K$. Ясно, что ζ — корень своего минимального многочлена и $g(\zeta)$ будет его же корнем для любого автоморфизма $g \in G$.

Если поле P нормально над K, то все корни данного многочлена принадлежат $P \ \Rightarrow \ g(\zeta) \in P$. Значит, $(g^{-1}hg)(\zeta) = (g^{-1}g)(\zeta) = \zeta \ \ \forall \ h \in H \Rightarrow \ (g^{-1}hg)(\sum \alpha_i \zeta^i) = \sum \alpha_i \zeta^i$. Таким образом, $g^{-1}hg \in H \ \forall h \in H, \ \forall g \in G \ \Rightarrow \ H$ является нормальной подгруппой группы G.

Если H — нормальный делитель группы G, то $(g^{-1}hg)(\zeta) = \zeta \ \ \forall h \in H, \ \forall g \in G$. Отсюда $h(g(\zeta)) = g(\zeta) \ \ \forall \ h \in H \ \Rightarrow \ g(\zeta) \in P$. Таким образом, каждый автоморфизм $g \in G$ при действии на числа из P переводит их в числа из P, порождая тем самым автоморфизм поля P над K. При этом все автоморфизмы вида hg, где $h \in H$, порождают один и тот же автоморфизм поля P над K. Автоморфизмы $g_1, g_2 \in G$ оставляют разные "следы" на P тогда и только тогда, когда $g_1 g_2^{-1} \notin H$. Следовательно, число автоморфизмов P над K равно числу различных смежных классов группы G по нормальному делителю $H \ \Rightarrow \ $ оно равно степени расширения $(P : K) = (L : K)/(L : P) \ \Rightarrow$ поле P нормально над K. Итак, каждому смежному классу ставится в соответствие порождаемый любым его представителем автоморфизм P над K — это и есть изоморфизм между G/H и $\mathrm{Aut}(P, K)$. $\square$

18.12 Разрешимость алгебраических уравнений

Пусть $f(x)$ — многочлен степени n над полем $K \subset \mathbb{C}$, и предположим, что $f(x)$ имеет n простых корней $\theta_1, \ldots, \theta_n$ и $g \in G = \mathrm{Aut}(L, K), \ \ $ где L — поле разложения $f(x)$. Легко видеть, что $g(\theta_i) = \theta_{\sigma(i)}$ для

[3] Отсюда, в частности, вытекает конечность числа промежуточных полей.

некоторой подстановки $\sigma \in S_n$. Несложно придти к выводу о том, что группа G изоморфна подгруппе симметрической группы S_n (поэтому о ней обычно говорят просто как о подгруппе в S_n).

Изучение цепочек радикальных расширений вида $(*)$ можно свести к изучению специальных подгрупп группы Галуа — нормальных делителей с абелевой (более того, даже с циклической) фактор-группой. В самом деле, можно ограничиться рассмотрением таких цепочек, в которых каждое звено дает поле разложения некоторого многочлена $x^p - a$ при простом p. Мы уже знаем, что группа Галуа такого расширения имеет нормальный делитель с циклической (а значит, и абелевой) фактор-группой.

В группе S_n при $n \geq 5$ нормальных делителей с абелевой фактор-группой слишком мало — одна лишь знакопеременная группа (см. доказательство в разделе 18.13).

(4) В конечном счете отсюда получаются примеры неразрешимых в радикалах алгебраических уравнений степени $n \geq 5$. Неразрешимым будет любое уравнение степени $n \geq 5$, для которого группа Галуа совпадает с S_n.

(5) Подгруппа G группы S_n называется *транзитивной*, если для любых номеров i, j от 1 до n существует подстановка $\sigma \in G$ такая, что $\sigma(i) = j$.

Утверждение. *Если $f(x)$ — неразложимый многочлен над полем $K \subset \mathbb{C}$, то группа $G = Aut(L, K)$ изоморфна некоторой транзитивной группе подстановок.*

Доказательство. Мы знаем, что неразложимый многочлен над $K \subset \mathbb{C}$ имеет только простые корни. Пусть α и β — различные корни $f(x)$. Рассмотрим многочлен

$$\Psi(x) \;=\; \prod_{g \in G} (x - g(\alpha)).$$

В силу формул Виета, коэффициенты $\Psi(x)$ остаются на месте при всех автоморфизмах из G. Поэтому, опираясь на предложение **(2)**, заключаем, что они принадлежат полю K. Поскольку $f(\alpha) = \Psi(\alpha) = 0$, многочлены $f(x)$ и $\Psi(x)$ имеют общий корень $\Rightarrow$ их наибольший общий делитель над K имеет степень $\geq 1 \;\Rightarrow\; f(x)$ является делителем для $\Psi(x)$. Следовательно, β содержится среди элементов вида $g(\alpha)$. □

(6) Справедливо следующее утверждение: *любая транзитивная подгруппа G группы S_n, содержащая хотя бы одну транспозицию, при простом n совпадает с S_n.*

Вот схема доказательства. Введем отношение эквивалентности: $i \sim j \;\Leftrightarrow\; (ij) \in G$. Транзитивность данного отношения следует из равенства $(ij)(jk)(ij) = (ik)$. Транзитивность группы G позволяет доказать, что классы эквивалентности содержат одно и то же число номеров. Поэтому при простом n имеется ровно один класс эквивалентности. Следовательно, G содержит все транспозиции.

(7) Пусть $K = \mathbb{Q}$. Многочлен $f(x) = x^5 - 4x - 2$ является наразложимым над $\mathbb{Q}$ и имеет три различных вещественных θ_1, θ_2, θ_3 корня и два комплексно сопряженных корня $\zeta, \overline{\zeta}$ (докажите!). В данном случае группа Галуа транзитивна и содержит транспозицию (автоморфизм, переводящий ζ в $\overline{\zeta}$ и оставляющий на месте θ_1, θ_2, θ_3). Таким образом, для данного многочлена группа Галуа совпадает с S_5.

Наше обсуждение является, конечно, лишь беглым очерком некоторых идей, развиваемых в данном разделе алгебры.

18.13 Нормальные делители симметрической группы

При построении радикальных расширений ключевую роль играют нормальные подгруппы с абелевой фактор-группой. Связанное с ними свойство подгрупп симметрической группы доказывается легко.

Утверждение 1. *Если H — нормальный делитель группы G с абелевой фактор-группой G/H, то H содержит все элементы вида $aba^{-1}b^{-1}$, где $a, b \in G$.*

Доказательство. $H(ab) = H(ba) \;\Rightarrow\; aba^{-1}b^{-1} \in H$. □

Утверждение 2. *Пусть $H \neq S_n$ — нормальный делитель группы S_n с абелевой фактор-группой*

S_n/H, и предположим, что $n \geq 5$. Тогда H совпадает со знакопеременной группой.

Доказательство. Возьмем два тройных цикла (цикла длины 3) $a = (ijk)$, $b = (ijm)$. Тогда

$$aba^{-1}b^{-1} \;=\; (ijk)(ijm)(kji)(mji) \;=\; (ij)(km).$$

Значит, H содержит все произведения пар независимых транспозиций. При $n \geq 5$ пары независимых транспозиций порождают все тройные циклы:

$$(ij)(kl)\,(ik)(jm)\,(il)(km) \;=\; (ikj).$$

Тройные циклы и произведения пар независимых транспозиций порождают все четные подстановки. □

Отметим также (без доказательства), что при $n \geq 5$ знакопеременная группа вообще не обладает нормальными делителями, отличными от нее самой или подгруппы, состоящей из одной лишь тождественной подстановки. Такие группы называются *простыми*. Классификация простых конечных групп была завершена лишь в 1980-х годах.

18.14 Группы при построении правильных многоугольников

Мы уже изучали вопрос о построении правильного n-угольника с помощью циркуля и линейки (см. раздел 16.11) — напомним, что он сводится к построению специальной цепочки расширений поля рациональных чисел, в которой каждое промежуточное поле имеет степень 2 над предыдущим полем.

Цепочка завершается построением поля, содержащего нужное нам число — длину стороны правильного n-угольника. Степень данного поля над $\mathbb{Q}$ с необходимостью равна 2^k. В свете теории Галуа это означает, что доказанное нами ранее необходимое условие на число сторон ($n = 2^k + 1$) вызвано тем, что группа Галуа для кругового многочлена простой степени содержит 2^k элементов.

Чтобы доказать достаточность этого условия, нужно доказать существование упомянутой выше специальной цепочки расширений поля $\mathbb{Q}$. Теория Галуа позволяет свести вопрос к доказательству существования специальной цепочки нормальных подгрупп группы порядка 2^k. Путь к доказательству достаточности условия на число сторон n открывается следующим наблюдением: *если группа G имеет порядок 2^k, то она обладает нормальной подгруппой порядка* 2. В действительности имеет место более общая

Теорема. *Пусть группа G имеет порядок p^k, где $p > 1$ — простое число.* [4] *Тогда G обладает нормальной подгруппой порядка p.*

Доказательство требует некоторой подготовки. Элементы $a,\ b \in G$ называются *сопряженными*, если $a = hbh^{-1}$ для некоторого $h \in G$. Нетрудно проверить, что сопряженность элементов — это отношение эквивалентности на G. Поэтому конечная группа G является объединением конечного числа (скажем, m) непересекающихся классов эквивалентности

$$G = K_1 \cup \ldots \cup K_m. \qquad (*)$$

Лемма 1. *В произвольной конечной группе G число элементов, сопряженных с заданным элементом a, является делителем порядка группы.*

Доказательство. Пусть $G(a) = \{h_1 a h_1^{-1}, \ldots, h_s a h_s^{-1}\}$ — множество всех элементов, сопряженных с a. Заметим, что

$$h_i a h_i^{-1} = h_j a h_j^{-1} \quad\Leftrightarrow\quad (h_j^{-1}h_i)a = a(h_j^{-1}h_i).$$

Обозначим через $H(a)$ множество всех элементов из G, коммутирующих с a. Элементарно проверяется, что $H(a)$ является подгруппой в G (подгруппа $H(a)$ называется *централизатором* элемента a). Таким образом,

$$h_i a h_i^{-1} = h_j a h_j^{-1} \quad\Leftrightarrow\quad h_j^{-1}h_i \in H(a) \quad\Leftrightarrow\quad h_i H(a) = h_j H(a).$$

Следовательно, число сопряженных с a элементов равно числу смежных классов группы G по подгруппе $H(a)$. □

Лемма 2. *В произвольной группе G порядка p^k существует элемент $a \neq e$ (отличный от единицы), коммутирующий со всеми элементами из G.*

Доказательство. Рассмотрим разложение $(*)$ группы G на непересекающиеся классы сопряженных элементов. Согласно лемме 1, порядок K_i имеет вид p^{k_i} (делитель числа p^k). Отсюда ясно, что число

[4] Такие группы называются *примарными*.

классов K_i, состоящих из одного элемента, должно делиться на $p \;\Rightarrow\;$ существует элемент $a \neq e$ такой, что $a = hah^{-1} \quad \forall \;\; h \in G \quad \Rightarrow \quad ah = ha \quad \forall \;\; h \in G$. $\square$

Доказательство теоремы. Согласно лемме 2, имеется элемент $a \neq e$, коммутирующий со всеми элементами из G. Пусть его порядок равен p^l. Тогда элемент

$$b = a^{p^{l-1}}$$

имеет порядок p. Циклическая группа, порожденная элементом b, является нормальным делителем, так как степени элемента b коммутируют со всеми элементами из G. $\square$

Лекция 19

ОСНОВНАЯ ЧАСТЬ

19.1 Алгебраические многообразия

Пусть $f(x_1, \ldots, x_n)$ — многочлен степени k от переменных $x_1, \ldots, x_n$. Множество

$$M = \{x = [x_1, \ldots, x_n]^\top : f(x_1, \ldots, x_n) = 0\}$$

называется *алгебраическим многообразием* [1] порядка k. Очевидно, это понятие обобщает понятие линейного многообразия в n-мерном пространстве.

В общем случае строение множества M весьма сложно. Однако, при его изучении часто помогает очень простая идея — давайте попытаемся упростить вид уравнения $f = 0$ с помощью замены переменных $x = Py$, где P — невырожденная матрица порядка n. Замена переменных связана с переходом к другому базису в том же n-мерном пространстве.

Утверждение. *Пусть P — произвольная невырожденная матриц порядка n и $g(y_1, \ldots, y_n) = f(x_1, \ldots, x_n)$, где $[x_1, \ldots, x_n]^\top = P[y_1, \ldots, y_n]^\top$. Тогда степень многочлена g равна степени многочлена f.*

Доказательство. Пусть $P = [p_{ij}]$. Тогда

$$x_1^{k_1} \ldots x_n^{k_n} = \left(\sum_{j=1}^{n} p_{1j} y_j\right)^{k_1} \ldots \left(\sum_{j=1}^{n} p_{nj} y_j\right)^{k_n}.$$

Отсюда ясно, что степень g не выше степени f. Противоположное неравенство доказывается с помощью замены $y = P^{-1}x$. $\square$

19.2 Квадратичные многочлены от двух переменных

Рассмотрим квадратичный многочлен с вещественными коэффициентами

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33}$$

как функцию от декартовых координат x, y на плоскости и исследуем строение множества точек (x, y), удовлетворяющих уравнению $f(x, y) = 0$.

Многочлен $f(x, y)$ имеет три типа слагаемых:

$$f(x, y) = f_2(x, y) + f_1(x, y) + f_0,$$

[1] Подробным изучением алгебраических многообразий занимается алгебраическая геометрия.

$f_2(x,y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$ — *квадратичная часть*, $f_1(x,y) = a_{13}x + a_{23}y$ — *линейная часть*, $f_0 = a_{33}$ — *свободный член*. Квадратичная и линейная части записываются с помощью матричных операций таким образом:

$$f_2(x,y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad f_1(x,y) = 2 \begin{bmatrix} a_{13} & a_{23} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Кроме того, легко проверяется, что

$$f(x,y) \;=\; \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}.$$

Попробуем найти такую декартову систему, в которой уравнение $f(x,y) = 0$ получит более простой вид. Множество его решений принято называть *линией (кривой) второго порядка*.

19.3 Поворот декартовой системы координат

Исходную декартову систему координат повернем против часовой стрелки на угол ϕ. Тогда базисные векторы $\mathbf{e}_1, \mathbf{e}_2$ перейдут в новые базисные векторы, соответственно,

$$\widetilde{\mathbf{e}}_1 = \cos\phi\,\mathbf{e}_1 + \sin\phi\,\mathbf{e}_2, \quad \widetilde{\mathbf{e}}_1 = -\sin\phi\,\mathbf{e}_1 + \cos\phi\,\mathbf{e}_2.$$

Старые координаты x, y будут выражаться через новые координаты $\widetilde{x}, \widetilde{y}$ следующим образом:

$$\begin{bmatrix} x \\ y \end{bmatrix} \;=\; \widetilde{x}\,\widetilde{\mathbf{e}}_1 + \widetilde{y}\,\widetilde{\mathbf{e}}_2 \;=\; \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix}.$$

Легко проверяется, что

$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}^{-1} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

В новых координатах квадратичная часть $f_2(x,y)$ принимает вид

$$f_2 \;=\; \begin{bmatrix} \widetilde{x} & \widetilde{y} \end{bmatrix} \left(\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \right) \begin{bmatrix} \widetilde{x} \\ \widetilde{y} \end{bmatrix}.$$

Матрица в скобках есть произведение трех матриц вида $\widetilde{A} = Q^\top A Q$, причем A — симметричная матрица: $A^\top = A$. Отсюда

$$\widetilde{A}^\top = (Q^\top A Q)^\top = Q^\top A^\top (Q^\top)^\top = Q^\top A Q = \widetilde{A}.$$

Значит, $\widetilde{A} = [\widetilde{a}_{ij}]$, $1 \le i, j \le 2$, остается симметричной матрицей.

Попытаемся выбрать угол ϕ так, чтобы матрица $\widetilde{A}$ приобрела диагональный вид:

$$\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \qquad (*)$$

Таким образом, требуется занулить элемент

$$\begin{aligned} \widetilde{a}_{12} = \widetilde{a}_{21} \;&=\; (\cos^2\phi - \sin^2\phi)\; a_{12} \;-\; \sin\phi\cos\phi\;(a_{11} - a_{22}) \\ &=\; \cos(2\phi)\; a_{12} \;-\; \sin(2\phi)\;\frac{a_{11} - a_{22}}{2} \;=\; 0. \end{aligned}$$

Если $a_{12} = 0$, то можно взять $\phi = 0$. Если $a_{12} \neq 0$, то надо решить уравнение

$$\operatorname{ctg}(2\phi) \;=\; \frac{a_{11} - a_{22}}{2a_{12}}.$$

Очевидно, решение существует. Поэтому всегда найдется ϕ такое, что имеет место равенство $(*)$. Доказано следующее

Утверждение. *С помощью поворота исходной декартовой системы координат на некоторый угол ϕ уравнение $f(x, y) = 0$ преобразуется в новых координатах к виду*

$$f_2 \;=\; \lambda_1 \widetilde{x}^2 \;+\; \lambda_2 \widetilde{y}^2 \;+\; 2b_{13}\widetilde{x} \;+\; 2b_{23}\widetilde{y} \;+\; b_{33} \;=\; 0,$$

где

$$\begin{bmatrix} b_{13} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{13} & a_{23} \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}, \quad b_{33} = a_{33}.$$

19.4 Сдвиг декартовой системы координат

Естественно предположить, что квадратичная часть f_2 не является тождественным нулем. Значит, λ_1 и λ_2 не равны нулю одновременно.

Случай 1: $\boldsymbol{\lambda_1 \neq 0,\ \lambda_2 \neq 0}$. Выделим в квадратичной части полные квадраты:

$$\lambda_1 \widetilde{x}^2 + 2b_{13}\widetilde{x} \;=\; \lambda_1 \left(\widetilde{x}^2 + 2\frac{b_{13}}{\lambda_1}\widetilde{x} + \frac{b_{13}^2}{\lambda_1^2} \right) \;-\; \frac{b_{13}^2}{\lambda_1^2} \;=\; \lambda_1 \left(\widetilde{x} + \frac{b_{13}}{\lambda_1} \right)^2 \;-\; \frac{b_{13}^2}{\lambda_1^2},$$

$$\lambda_2 \widetilde{y}^2 + 2b_{23}\widetilde{y} \;=\; \lambda_2 \left(\widetilde{y}^2 + 2\frac{b_{23}}{\lambda_2}\widetilde{y} + \frac{b_{23}^2}{\lambda_2^2} \right) \;-\; \frac{b_{23}^2}{\lambda_2^2} \;=\; \lambda_2 \left(\widetilde{y} + \frac{b_{23}}{\lambda_2} \right)^2 \;-\; \frac{b_{23}^2}{\lambda_2^2}.$$

Осуществим сдвиг декартовой системы координат $\widetilde{x}$ и $\widetilde{y}$, поместив ее начало в точку

$$O' \;=\; \left(-\frac{b_{13}}{\lambda_1},\; -\frac{b_{23}}{\lambda_2} \right).$$

Новые координаты x' и y' выражаются через $\widetilde{x}$ и $\widetilde{y}$ следующим образом:

$$x' \;=\; \widetilde{x} \;+\; \frac{b_{13}}{\lambda_1}, \quad y' \;=\; \widetilde{y} \;+\; \frac{b_{23}}{\lambda_2}.$$

В новых координатах уравнение $f(x, y) = 0$ теряет линейную часть и принимает вид

$$\lambda_1 (x')^2 \;+\; \lambda_2 (y')^2 \;+\; c \;=\; 0, \tag{1}$$

$$c \;=\; b_{33} \;-\; \frac{b_{13}^2}{\lambda_1^2} \;-\; \frac{b_{23}^2}{\lambda_2^2}.$$

Случай 2: $\boldsymbol{\lambda_1 = 0,\ \lambda_2 \neq 0}$. Переносим начало координат в точку

$$O' \;=\; \left(0,\; -\frac{b_{23}}{\lambda_2} \right).$$

В новых координатах

$$\widehat{x} \;=\; \widetilde{x}, \quad \widehat{y} \;=\; \widetilde{y} \;+\; \frac{b_{23}}{\lambda_2}$$

уравнение $f(x,y)=0$ получает вид

$$\lambda_2\,\widehat{y}^2 \;+\; 2b\,\widehat{x} \;+\; c \;=\; 0, \qquad b \;=\; b_{23}, \qquad c \;=\; b_{33} - \frac{b_{23}^2}{\lambda_2^2}.$$

Если $b \neq 0$, выполним еще один перенос начала системы координат — в точку $(-c/b,\ 0)$. В новых координатах

$$x' \;=\; \widehat{x} \;+\; \frac{c}{2b}, \quad y' \;=\; \widehat{y}$$

уравнение $f(x,y)=0$ приобретает форму

$$\lambda_2\,(y')^2 \;+\; 2b\,x' \;=\; 0. \tag{2}$$

Если $b=0$, получам уравнение (положим для унификации $x'=\widehat{x},\ y'=\widehat{y}$)

$$\lambda_2\,(y')^2 \;+\; c \;=\; 0. \tag{3}$$

Случай $\lambda_1 \neq 0,\ \lambda_2 = 0$ сводится к случаю 2 дополнительным поворотом системы координат на угол $\pi/2$. Доказано следующее

Утверждение. *С помощью поворота и сдвига исходной системы координат уравнение* $f(x,y)=0$ *приводится в новых координатах к виду* (1), (2) *или* (3).

19.5 Эллипс

Пусть в некоторой декартовой системе координат уравнение $f(x,y)=0$ имеет вид (1), где λ_1 и λ_2 — ненулевые числа одинакового знака. Уберем штрихи и рассмотрим новую систему в качестве исходной.

Не ограничивая общности, можно считать, что $0 < \lambda_1 \leq \lambda_2$ (если оба числа отрицательны, то можно поменять знак в обеих частях уравнения; если $\lambda_1 > \lambda_2$, то можно поменять их местами с помощью поворота на угол $\pi/2$). Если при этом $c>0$, то изучаемое множество пусто. Если $c=0$, в нем только одна точка $(0,\ 0)$. Предположим, что $c<0$. Тогда, положив $a=-c/\lambda_1,\ b=-c/\lambda_2$, уравнение (1) можно записать в виде

$$\frac{x^2}{a^2} \;+\; \frac{y^2}{b^2} \;=\; 1, \quad a \geq b > 0. \tag{1'}$$

Определение 1. Множество точек (x,y), удовлетворяющих уравнению (1′), называется *эллипсом* с полуосями a и b.

Точки $F_- = (-c,\ 0)$ и $F_+ = (c,\ 0)$, где $c = \sqrt{a^2-b^2} \geq 0$, называются, соответственно, отрицательным и положительным *фокусами* эллипса. Число $e = c/a$ называется *эксцентриситетом* эллипса. Заметим, что $0 \leq e < 1$. Прямые $l_- : x = -a/e$ и $l_+ : x = a/e$ называются, сответственно, отрицательной и положительной *директрисами* эллипса.

Пусть точка $M = (x,\ y)$ удовлетворяет уравнению (1′). Найдем сумму расстояний от нее до фокусов:

$$\begin{aligned} |MF_-| + |MF_+| \;&=\; \sqrt{(x+c)^2+y^2} \;+\; \sqrt{(x-c)^2+y^2} \\ &=\; \sqrt{x^2(1-b^2/a^2)+2xc+b^2+c^2} \;+\; \sqrt{x^2(1-b^2/a^2)-2xc+b^2+c^2}. \end{aligned}$$

Заметим, что $1-b^2/a^2 = e^2$ и $b^2+c^2 = (c/e)^2 = a^2$. Кроме того, $|ex| \leq a$. Поэтому

$$\begin{aligned} |MF_-| + |MF_+| \;&=\; \sqrt{(ex)^2+2(ex)(c/e)+(c/e)^2} \;+\; \sqrt{(ex)^2-2(ex)(c/e)+(c/e)^2} \\ &=\; |a+ex| \;+\; |a-ex| \;=\; (a+ex) \;+\; (a-ex) \;=\; 2a. \end{aligned}$$

Таким образом, *сумма расстояний от любой точки эллипса* $(1')$ *до его фокусов постоянна и равна* $2a$.

Определение 2. Множество тех и только тех точек плоскости, для которых сумма расстояний до заданных точек постоянна, называется эллипсом.

Мы уже выяснили, что все точки эллипса как множества из определения 1 принадлежат множеству из определения 2. Рассмотрим теперь эллипс как множество, данное определением 2. Выберем декартову систему, в которой заданные точки F_- и F_+ получают координаты $(-c,\ 0)$ и $(c,\ 0)$. Постоянную сумму расстояний будем считать равной $2a$. Тогда

$$\left(\sqrt{(x-c)^2+y^2}\right)^2 = \left(2a - \sqrt{(x+c)^2+y^2}\right)^2 \quad \Rightarrow \quad a\sqrt{(x+c)^2+y^2} = a^2 + xc.$$

Еще одно возведение в квадрат дает $b^2x^2 + a^2y^2 = a^2b^2 \quad \Rightarrow \quad (1')$. Следовательно, определения 1 и 2 эквивалентны.

В случае $e = 0$ эллипс есть окружность радиуса $a = b$. Пусть $e > 0$. Выше мы получили равенства

$$|MF_-| = |a + ex| = e|x + (a/e| \quad \Rightarrow \quad \frac{|MF_-|}{|x + (a/e)|} = e,$$

$$|MF_+| = |a - ex| = e|x - (a/e)| \quad \Rightarrow \quad \frac{|MF_+|}{|x - (a/e)|} = e.$$

Возводя каждое из последних равенств в квадрат, получаем $(1')$. Таким образом, доказано следующее

Утверждение. *Множество тех и только тех точек плоскости, для которых отношение расстояний до заданной точки и заданной прямой постоянно и равно* $0 < e < 1$, *является эллипсом.*

Ясно, что для выбора точки (фокуса) и соответствующей прямой (директрисы), определяющих один и тот же эллипс, имеются в точности две возможности.

19.6 Гипербола

По-прежнему, пусть в декартовой системе координат уравнение $f(x, y) = 0$ получает вид (1), но λ_1 и λ_2 имеют разные знаки. Если при этом свободный член оказался равен нулю, то получаем пару прямых, проходящих через начало координат. Предположим, что свободный член отличен от нуля. Ясно, что в этом случае уравнение можно записать в виде

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \qquad (1'')$$

где $a,\ b$ — некоторые положительные числа (возможно, для того потребуется дополнительно повернуть систему координат на угол $\pi/2$).

Определение 1. Множество точек $(x,\ y)$, удовлетворяющих уравнению $(1'')$, называется *гиперболой* с полуосями a и b.

Легко видеть, что точки $(x,\ y)$ гиперболы $(1'')$ находятся в объединении двух непересекающихся областей плоскости (как говорят, распадаются на *две ветви*)

$$\mathcal{D}_+ = \{(x,y):\ x \ge a, \quad |y| \le (b/a)|x|\}, \quad \mathcal{D}_- = \{(x,y):\ x \le -a, \quad |y| \le (b/a)|x|\}.$$

Прямые $h_+ : y = (b/a)x$ и $h_- : y = -(b/a)x$ называются *асимптотами* гиперболы.

Пусть $x > 0$ и $y = y(x)$ — единственное значение для y такое, что $y > 0$ и точка (x, y) удовлетворяет уравнению $(1'')$. Очевидно, расстояние от точки $(x,\ y((x))$ до асимптоты h_+ не превышает

$$|(b/a)x - y(x)| = \frac{b^2}{|(b/a)x + y(x)|} \ \to\ 0 \quad \text{при} \quad x \ \to\ +\infty.$$

При $x > 0$ и $y < 0$ соответствующие точки $(x,\ y(x))$ гиперболы приближаются к асимптоте l_-. Аналогичные наблюдения справедливы также для точек гиперболы при $x < 0$.

Точки $F_- = (-c,\ 0)$ и $F_+ = (c,\ 0)$, где $c = \sqrt{a^2 + b^2} > 0$, называются, соответственно, отрицательным и положительным *фокусами* гиперболы. Число $e = c/a$ называется *эксцентриситетом* гиперболы. Заметим, что в случае гиперболы $e > 1$. Прямые $l_{-1} : x = -a/e$ и $l_+ : x = a/e$ называются, соответственно, отрицательной и положительной *директрисами* гиперболы.

Найдем расстояния от произвольной точки $M = (x, y)$, удовлетворяющей $(1'')$, до фокусов (выкладки проводятся в полной аналогии со случаем эллипса):

$$|MF_-| = \sqrt{(x + c)^2 + y^2} = |a + ex|, \tag{A}$$

$$|MF_+| = \sqrt{(x - c)^2 + y^2} = |a - ex|. \tag{B}$$

Поскольку $|x| \geq a$ и $e > 1$, получаем

$$|a + ex| - |a - ex| = \begin{cases} (ex + a) - (ex - a) \ = \ 2a, & x > 0, \\ -(ex + a) + (ex - a) \ = \ -2a, & x < 0. \end{cases}$$

Таким образом, *абсолютная величина разности расстояний от любой точки гиперболы* $(1'')$ *до ее фокусов постоянна и равна* $2a$.

Определение 2. Множество тех и только тех точек плоскости, для которых абсолютная величина разности расстояний до двух заданных точек постоянна, называется гиперболой.

Пусть точка (x, y) принадлежит множеству из определения 2. Введем декартовы координаты таким образом, что заданные точки получают координаты $(-c,\ 0)$ и $(c,\ 0)$. Постоянную абсолютную величину разности расстояний обозначим через $2a$. Тогда

$$\left|\sqrt{(x + c)^2 + y^2} - \sqrt{(x - c)^2 + y^2}\right| = 2a \quad \Rightarrow$$

$$(x + c)^2 + y^2 \ = \ \left(2a - \sqrt{(x - c)^2 + y^2}\right)^2 \quad \Rightarrow \quad (1'').$$

Таким образом, определения 1 и 2 эквивалентны.

Формулы (A) и (B) делают очевидным также следующее

Утверждение. *Множество тех и только тех точек плоскости, для которых отношение расстояний до заданной точки и заданной прямой постоянно и равно* $e > 1$, *является гиперболой.*

Из наших построений следует, что имеются ровно две возможности для выбора точки (фокуса) и соответствующей прямой (директрисы), определяющих одну и ту же гиперболу.

19.7 Парабола

Пусть уравнение $f(x,y)$ имеет вид (2). Можно считать, что $\lambda_2 > 0$ и $b < 0$ (этого всегда можно добиться умножением уравнения на (-1) и дополнительным поворотом системы координат на угол π). Уберем штрихи, рассматривая новую систему в качестве исходной. Положив $p = -b/\lambda_2$, получаем уравнение

$$y^2 \;=\; 2px, \quad p > 0. \tag{2'}$$

Определение 1. Множество точек (x,y), удовлетворяющих уравнению (2′), называется *параболой с фокальным параметром* p.

Точка $F = (p/2,\ 0)$ называется *фокусом* параболы (2′). Прямая $l : x = -p/2$ называется *директрисой* параболы (2′).

Пусть $M = (x,y)$ — произвольная точка параболы. Расстояние от нее до фокуса имеет вид

$$\begin{aligned} |MF| \;=\; \sqrt{(x-p/2)^2+y^2} \;&=\; \sqrt{x^2-px+(p/2)^2+2px} \\ &=\; \sqrt{x^2+2x(p/2)+(p/2)^2} \;=\; |x+p/2|. \end{aligned}$$

Итак, *расстояние от любой точки параболы до фокуса* $|F|$ *равно расстоянию от нее до директрисы* l.

Определение 2. Множество тех и только тех точек, для которых расстояние до заданной точки F равно расстоянию до заданной прямой l, называется параболой.

Путь расстояние от заданной точки до заданной прямой равно p. Выберем систему координат таким образом, что $F = (p/2,\ 0)$ и $l : x = -p/2$. Если $|MF| = |x + p/2|$, то, возводя это равенство в квадрат, получаем (2′). Таким образом, определения 1 и 2 действительно эквивалентны.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

19.8 Классификация линий второго порядка

Мы уже доказали, что любая линия второго порядка в некоторой декартовой системе координат удовлетворяет одному из уравнений (1), (2) или (3). Для описания всех возможных случаев иногда предлагается следующая классификация:

(1)

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \;\text{ (эллипс)};\quad \frac{x^2}{a^2}+\frac{y^2}{b^2}=-1 \;\text{ (мнимый эллипс)};$$

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=0 \;\text{ (пара мнимых пересекающихся прямых)};$$

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1 \;\text{ (гипербола)};\quad \frac{x^2}{a^2}-\frac{y^2}{b^2}=0 \;\text{ (пара пересекающихся прямых)};$$

(2)

$$y^2 = 2px \;\text{ (парабола)};$$

(3)

$$y^2 = a^2 \;\text{ (пара параллельных прямых)};\quad y^2 = 0 \;\text{ (пара совпадающих прямых)};$$

$$y^2 = -a^2 \;\text{ (пара мнимых параллельных прямых)}.$$

19.9 Инварианты линии второго порядка

Рассмотрим общее уравнение $f(x,y) = a_{11}x^2 + 2a_{12}xy + a_{33}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0$ в заданной декартовой системе координат и определители

$$I_2 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}, \quad I_3 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}.$$

Теорема об инвариантах. *Значения определителей I_2 и I_3 не изменяются при переходе от заданной декартовой к любой декартовой системе координат.*

Доказательство. Пусть переход к новой декартовой системе координат задается формулами

$$\begin{cases} x &= p_{11}x' + p_{12}y' + c_1, \\ y &= p_{21}x' + p_{22}y' + c_2. \end{cases} \quad \Rightarrow \quad \begin{bmatrix} x \\ y \end{bmatrix} = P \begin{bmatrix} x' \\ y' \end{bmatrix} + c, \quad P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}, \quad c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Важное наблюдение: $P^\top P = I$ (в силу ортогональности базисных векторов декартовых систем).

Обозначим через A_2 и A_3 матрицы в определителях I_2 и I_3. Пусть $\widetilde{A}_2$ и $\widetilde{A}_3$ — аналогичные матрицы в новой системе координат. Тогда $\widetilde{A}_2 = P^\top A_2 P$, $\widetilde{A}_3 = Q^\top A_3 Q$, где

$$Q = \begin{bmatrix} P & c \\ 0 & 1 \end{bmatrix}, \quad Q^\top = \begin{bmatrix} P^\top & 0 \\ c^\top & 1 \end{bmatrix} \quad \Rightarrow \quad \det Q = \det P, \quad \det Q^\top = \det P^\top.$$

Определитель произведения матриц равен произведению определителей $\Rightarrow$

$$\det \widetilde{A}_2 \;=\; \det P^\top \;\det A_2 \;\det P \;=\; \det (P^\top P) \;\det A_2 \;=\; \det A_2,$$

$$\det \widetilde{A}_3 \;=\; \det P^\top \;\det A_3 \;\det P \;=\; \det (P^\top P) \;\det A_3 \;=\; \det A_3. \quad \square$$

Определение. Определители I_2 и I_3 называются *инвариантами* линии второго порядка.

19.10 Определение типа линии

Если в какой-либо декартовой системе координат получается уравнение вида (1), то, очевидно, $I_2 = \lambda_1\lambda_2 \neq 0$. Для того чтобы линия была эллипсом, необходимо, чтобы $I_2 > 0$. Для гиперболы необходимо, чтобы $I_2 < 0$. Если же $I_2 = 0$, то соответствующая линия относится к случаю (2) или (3).

Если получается уравнение вида (2), то

$$I_3 \;=\; \det \begin{bmatrix} 0 & 0 & b \\ 0 & \lambda_2 & 0 \\ b & 0 & 0 \end{bmatrix} \;=\; -\lambda_2 b^2 \;\neq\; 0.$$

Для уравнения вида (3) находим

$$I_3 \;=\; \det \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & c \end{bmatrix} \;=\; 0.$$

Инварианты полезны для определения типа линии и в том случае, когда общее уравнение задано в произвольной аффинной системе координат.

Теорема о знаках инвариантов. *Знаки определителей I_2 и I_3 не изменяются при переходе к любой аффинной системе координат.*

Доказательство. В случае аффинных систем нельзя утверждать, что $P^\top P = I$. Однако, $\det \widetilde{A}_2 = \det A_2 \; (\det P)^2$, $\det \widetilde{A}_3 = \det A_3 \; (\det Q)^2$. $\square$

Лекция 20

ОСНОВНАЯ ЧАСТЬ

20.1 Квадратичные многочлены от трех переменных

Рассмотрим вещественный квадратичный многочлен

$$f(x,y,z) = a_{11}x^2 + 2a_{12}xy + 2a_{13}xz + 2a_{22}y^2 + a_{23}yz + a_{33}z^2 + 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44}$$

от декартовых координат x, y, z в геометрическом пространстве и исследуем множество решений уравнения $f(x,y,z) = 0$ — его принято называть *поверхностью второго порядка*.

Легко проверить, что

$$f(x,y,z) = \begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix},$$

а квадратичная часть многочлена $f(x,y,z)$ имеет вид

$$f_2(x,y,z) = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Как и в случае двух переменных, попробуем перейти к более удобной декартовой системе координат.

20.2 Декартовы системы и ортогональные матрицы

Пусть e_1, e_2, e_3 и $\widetilde{e}_1$, $\widetilde{e}_2$, $\widetilde{e}_3$ — базисные векторы двух декартовых систем координат с общим началом. Выразим векторы второй системы в виде линейных комбинаций векторов первой системы

$$\begin{aligned} \widetilde{e}_1 &= p_{11}e_1 + p_{21}e_2 + p_{31}e_3, \\ \widetilde{e}_2 &= p_{12}e_1 + p_{22}e_2 + p_{32}e_3, \\ \widetilde{e}_3 &= p_{13}e_1 + p_{23}e_2 + p_{33}e_3 \end{aligned}$$

и заметим, что

$$\begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} (\widetilde{e}_1,\widetilde{e}_1) & (\widetilde{e}_1,\widetilde{e}_2) & (\widetilde{e}_1,\widetilde{e}_3) \\ (\widetilde{e}_2,\widetilde{e}_1) & (\widetilde{e}_2,\widetilde{e}_2) & (\widetilde{e}_2,\widetilde{e}_3) \\ (\widetilde{e}_3,\widetilde{e}_1) & (\widetilde{e}_3,\widetilde{e}_2) & (\widetilde{e}_3,\widetilde{e}_3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Таким образом, матрица перехода P для базисов двух декартовых систем координат удовлетворяет матричному равенству

$$P^\top P = I. \qquad (*)$$

Ясно и то, что если матрица перехода обладает свойством $(*)$, то декартова система переходит в декартову.

Определение. Квадратная вещественная матрица P, удовлетворяющая равенству $(*)$, называется *ортогональной*.

Ортогональные матрицы порядка 2 осуществляют переход между базисами декартовых систем на плоскости. Таковы, в частности, матрицы перехода, реализующие поворот (см. лекцию 18).

Данное нами определение применимо и для матриц, порядок которых больше 3. Согласно $(*)$, для ортогональных матриц произвольного порядка обращение сводится к транспонированию:

$$P^{-1} = P^\top.$$

Кроме того, произведение двух ортогональных матриц остается ортогональной матрицей: если $P^\top P = Q^\top Q = I$, то $(PQ)^\top(PQ) = Q^\top(P^\top P)Q = Q^\top\ I\ Q = Q^\top Q = I$. Очевидно, единичная матрица I является ортогональной матрицей.

Следовательно, *множество всех ортогональных матриц фиксированного порядка относительно операции умножения матриц образует группу.*

Утверждение. *Пусть $y = Px$, где P — произвольная ортогональная матрица порядка n и $x \in \mathbb{R}^{n\times 1}$. Тогда сумма квадратов элементов матрицы-столбца y равна сумме квадратов элементов матрицы-столбца x.*

Доказательство.

$$y_1^2 + \ldots + y_n^2 \;=\; y^\top y \;=\; (Px)^\top(Px) \;=\; x^\top(P^\top P)\,x \;=\; x^\top x \;=\; x_1^2 + \ldots + x_n^2. \qquad \square$$

Следствие. *Пусть $B = PAQ$, где P, Q — произвольные ортогональные матрицы порядка n и A — произвольная вещественная матрица порядка n. Тогда сумма квадратов элементов матрицы B равна сумме квадратов элементов матрицы A.*

20.3 Метод вращений

Попробум упростить квадратичную часть $f_2(x, y)$, используя ту же идею поворота системы координат, как и в случае плоскости. Теперь, однако, у нас есть три координатных плоскости, порождаемые тремя парами координатных осей. Цель вращения — получить нуль вместо какой-нибудь одной пары элементов $a_{ij} = a_{ji}$ при $i \neq j$. Рассмотрим три возможности:

$$\begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \widetilde{a}_{11} & 0 & \widetilde{a}_{13} \\ 0 & \widetilde{a}_{22} & \widetilde{a}_{23} \\ \widetilde{a}_{13} & \widetilde{a}_{23} & \widetilde{a}_{33} \end{bmatrix}, \quad (1)$$

$$\begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix} = \begin{bmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} & 0 \\ \widetilde{a}_{12} & \widetilde{a}_{22} & \widetilde{a}_{23} \\ 0 & \widetilde{a}_{23} & \widetilde{a}_{33} \end{bmatrix}, \quad (2)$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} & \widetilde{a}_{13} \\ \widetilde{a}_{12} & \widetilde{a}_{22} & 0 \\ \widetilde{a}_{13} & 0 & \widetilde{a}_{33} \end{bmatrix}. \quad (3)$$

Обозначим через d_0, h_0 и d_1, h_1 суммы квадратов диагональных и внедиагональных элементов исходной и новой матриц в каждом из трех случаев:

$$d_0 = a_{11}^2 + a_{22}^2 + a_{33}^2, \quad h_0 = 2a_{12}^2 + 2a_{13}^2 + 2a_{23}^2,$$
$$d_1 = \widetilde{a}_{11}^2 + \widetilde{a}_{22}^2 + \widetilde{a}_{33}^2, \quad h_1 = 2\widetilde{a}_{12}^2 + 2\widetilde{a}_{13}^2 + 2\widetilde{a}_{23}^2.$$

Согласно отмеченным выше свойствам ортогональных матриц,

$$d_1 + h_1 = d_0 + h_0 \quad \Rightarrow \quad h_1 = h_0 - (d_1 - d_0).$$

По той же причине в случае (1) имеем $a_{11}^2 + a_{22}^2 + 2a_{12}^2 = \widetilde{a}_{11}^2 + \widetilde{a}_{22}^2$ и, поскольку $a_{33} = \widetilde{a}_{33}$, $d_1 - d_0 = 2a_{12}^2$. В случае (2) $d_1 - d_0 = 2a_{13}^2$, а в случае (3) $d_1 - d_0 = 2a_{23}^2$.

Пусть индексы i, j определяют координатную плоскость, в которой проводится вращение (и указывают на то, какое из соотношений (1), (2) или (3) имеет место). Выберем их таким образом, чтобы исключаемый элемент a_{ij} был максимальным по модулю. Тогда, очевидно,

$$d_1 - d_0 \;=\; 2a_{ij}^2 \;\geq\; 2(h_0/6) \;=\; h_0/3 \qquad \Rightarrow \qquad h_1 \;\leq\; \frac{2}{3}\, h_0.$$

Пусть $A_0 = [a_{ij}]$ и $A_1 = [\widetilde{a}_{ij}]$. Рассматривая A_1 в качестве новой исходной матрицы, выберем в ней максимальный по модулю внедиагональный элемент и, занулив его с помощью вращения, получим матрицу A_2. Продолжая действовать таким же образом и далее, построим последовательность матриц $A_k = [a_{ij}^{(k)}]$, $k = 0, 1, \ldots$.

Пусть h_k обозначает сумму квадратов внедиагональных элементов матрицы A_k. Тогда

$$h_k \;\leq\; \left(\frac{2}{3}\right)^k h_0 \to\; 0 \quad \text{при} \quad k \to \infty.$$

Следовательно, при любых фиксированных $i \neq j$ последовательность внедиагональных элементов $a_{ij}^{(k)}$ сходится к нулю при $k \to \infty$.

20.4 Вложенные подпоследовательности

Лемма об ограниченных последовательностях. *Пусть имеется конечное число ограниченных последовательностей* $\{s_1^{(k)}\}, \ldots, \{s_m^{(k)}\}$, $k = 1, 2, \ldots$. *Тогда можно выбрать последовательность номеров* $k_1 < k_2 < \ldots$ *таким образом, что каждая из подпоследовательностей* $\{s_1^{(k_l)}\}, \ldots, \{s_m^{(k_l)}\}$, $l = 1, 2, \ldots$, *будет сходящейся.*

Доказательство. Из ограниченной последовательности $\{s_k\}$ выбираем сходящуюся подпоследовательность s_{k_l} и вместо исходных последовательностей рассматриваем подпоследовательности $\{s_1^{(k_l)}\}, \ldots, \{s_m^{(k_l)}\}$, $l = 1, 2, \ldots$. Они остаются, конечно, ограниченными и при этом первая из них будет сходящейся. Теперь уже из ограниченной последовательности $\{s_2^{k_l}\}$ выберем сходящуюся подпоследовательность (подпоследовательность подпоследовательности — по отношению к исходной последовательности) и переходим к подпоследовательностям $\{s_1^{(k_{l_i})}\}, \ldots, \{s_m^{(k_{l_i})}\}$, $i = 1, 2, \ldots$. Полученные вложенные подпоследовательности будут, по-прежнему, ограниченными, но сходящимися являются уже первые две. И так далее. $\square$

20.5 Диагонализация в пределе

Вернемся к методу вращений. Будут ли сходиться к конечным пределам последовательности диагональных элементов $a_{ii}^{(k)}$ — для нашей ближайшей цели не очень важно. Каждая из них является ограниченной и поэтому обладает сходящейся подпоследовательностью. Более того, по лемме об ограниченных последовательностях, имеется подпоследовательность матриц A_k, в которой каждая из последовательностей диагональных элементов сходится к какому-то пределу.

Чтобы не загромождать обозначения, будем считать, что A_k и есть та самая подпоследовательность, для которой все последовательности $a_{ij}^{(k)}$ являются сходящимися (как мы знаем, при $i \neq j$ к нулю). Пусть

$$\lim_{k\to\infty} a_{ii}^{(k)} \;=\; \lambda_i, \qquad i = 1,\ 2,\ 3.$$

Понятно, что

$$A_k = P_k^\top A_0 P_k, \quad k = 1,\ 2,\ \dots, \qquad (\#)$$

где матрицы $P_k = [p_{ij}^{(k)}]$ являются произведениями использованных матриц вращения (из соотношений (1), (2) или (3)). Поэтому при любом k матрица P_k является ортогональной (как произведение ортогональных матриц). Следовательно, сумма квадратов всех элементов матрицы P_k при любом k одинакова (и равна 3). Значит, каждая последовательность $p_{ij}^{(k)}$ является ограниченной при $k \to \infty$ и поэтому обладает сходящейся подпоследовательностью.

По лемме об ограниченных последовательностях, существует подпоследовательность матриц P_k, в которой каждая последовательность $p_{ij}^{(k)}$ будет сходящейся. Для упрощения обозначений будем считать, что P_k и есть именно такая подпоследовательность. Пусть

$$\lim_{k\to\infty} p_{ij}^{(k)} \;=\; p_{ij}, \quad i,j = 1,\ 2,\ 3.$$

При каждом k выполняется равенство (#). Переходя к пределу в соответствующих поэлементных равенствах, получаем

$$\Lambda \;\equiv\; \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \;=\; P^\top A_0 P.$$

Кроме того, для каждого k имеем $P_k^\top P_k = I \;\Rightarrow\; P^\top P = I$. В итоге доказана следующая важная

Теорема. *Для любой вещественной симметричной матрицы A порядка 3 существуют ортогональная матрица P и диагональная матрица Λ такие, что*

$$\Lambda \;=\; P^\top A P.$$

Следствие. *Существует декартова система координат, в которой уравнение поверхности второго порядка имеет вид*

$$f(\widetilde{x},\widetilde{y},\widetilde{z}) \;=\; \lambda_1\widetilde{x}^2 + \lambda_2\widetilde{y}^2 + \lambda_3\widetilde{z}^2 + 2b_1\widetilde{x} + 2b_2\widetilde{y} + 2b_3\widetilde{z} + a \;=\; 0.$$

20.6 Диагонализация вещественных симметричных матриц

В действительности тот же метод вращений позволяет получить более общую теорему.

Теорема о диагонализации вещественных симметричных матриц. *Вещественная симметричная матрица A произвольного порядка n приводится к диагональной матрице Λ с помощью некоторой ортогональной матрицы P:*

$$\Lambda \;=\; P^{\top} A P.$$

Доказательство. Начиная с $A_0 = A$, построим последовательность матриц $A_k = [a_{ij}^{(k)}]$, $k = 0,\ 1,\ \ldots$, в которой A_k получается из A_{k-1} путем умножения слева и справа на матрицы вращения:

$$A_k \;=\; R_k^{\top} A_{k-1} R_k, \qquad (*)$$

где R_k отличается от единичной матрицы I лишь четырьмя элементами 2×2-подматрицы, раположенной на пересечении строк и столбцов с номерами $i < j$ и равной

$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

Матрица R_k осуществляет поворот на угол ϕ в координатной плоскости, определяемой номерами $i \neq j$. Любая матрица вращения такого вида является, очевидно, ортогональной.

Ясно, что симметричность матрицы $A_0 = A$ наследуется всеми матрицами A_k. Из предыдущих исследований мы уже знаем, что ϕ можно выбрать таким образом, что $a_{ij}^{(k)} = a_{ji}^{(k)} = 0$. Обозначим через d_k и h_k суммы квадратов диагональных и внедиагональных элементов матрицы A_k. Тогда

$$\left(a_{ii}^{(k)}\right)^2 + \left(a_{jj}^{(k)}\right)^2 = \left(a_{ii}^{(k-1)}\right)^2 + \left(a_{jj}^{(k-1)}\right)^2 + 2\left(a_{ij}^{(k-1)}\right)^2 \quad \Rightarrow \quad d_k - d_{k-1} = 2\left(a_{ij}^{(k-1)}\right)^2.$$

Для каждого k будем выбирать плоскость вращения (номера $i < j$) такми образом, чтобы исключаемый элемент $a_{ij}^{(k-1)}$ был максимальным по модулю среди всех внедиагональных элементов матрицы A_{k-1}. Общее число внедиагональных элементов равно $n^2 - n$. Поэтому

$$\left(a_{ij}^{(k-1)}\right)^2 \;\geq\; \frac{h_{k-1}}{n^2 - n}.$$

Отсюда, учитывая равенство $d_k + h_k = d_{k-1} + h_{k-1}$, получаем

$$h_k \;\leq\; h_{k-1} \;-\; \frac{2}{n^2 - n}\, h_{k-1} \;\leq\; \left(1 - \frac{2}{n^2 - n}\right)^k h_0 \;\;\to\;\; 0 \quad \text{при} \quad k \to \infty.$$

Из соотношений $(*)$ вытекает, что

$$A_k \;=\; P_k^{\top} A P_k, \quad k = 1,\ 2,\ \ldots,$$

где для всех k матрицы $P_k = [p_{ij}^{(k)}]$ являются ортогональными (как произведения ортогональных матриц).

Для любых фиксированных i, j последовательности $a_{ij}^{(k)}$, $p_{ij}^{(k)}$ являются ограниченными. По лемме об ограниченных последовательностях, существует последовательность

номеров $k_1 < k_2 < \ldots$ такая, что каждая из подпоследовательностей $a_{ij}^{(k_l)}$, $p_{ij}^{(k_l)}$ будет сходящейся. Заметим, что $a_{ij}^{(k_l)} \to 0$ при $i \neq j$. Пусть

$$\lim_{l\to\infty} a_{ii}^{(k_l)} = \lambda_i, \quad i = 1, \ldots, n, \qquad \lim_{l\to\infty} p_{ij}^{(k_l)} = p_{ij}, \quad i, j = 1, \ldots, n.$$

Введем матрицы

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \qquad P = [p_{ij}].$$

Для всех $l = 1, 2, \ldots$ имеем $A_{k_l} = P_{k_l}^\top A P_{k_l}$. Переходя к пределу в поэлементных равенствах, получаем

$$\Lambda = P^\top A P.$$

Из условий ортогональности $P_{k_l}^\top P_{k_l} = I$ вытекает, что в пределе $P^\top P = I$. Значит, матрица P является ортогональной. $\square$

Мы только что получили один из важнейших результатов — как для самой теории матриц, так и для ее многочисленных приложений. В нашем курсе мы еще вернемся к его обсуждению в связи с рядом фундаментальных понятий линейной алгебры. Наше доказательство замечательно своей конструктивностью: оно дает одновременно и метод приближенного вычисления матриц Λ и P. Это один из ранних практических методов вычислительной алгебры, предложенный К. Якоби в 1846 году. [1]

Задача. Дана симметричная матрица $A \in \mathbb{R}^{n\times n}$ с ненулевой суммой элементов главной диагонали. Доказать существование ортогональной матрицы $Q \in \mathbb{R}^{n\times n}$ такой, что в матрице $Q^\top A Q$ все элементы главной диагонали одинаковы.

[1] Последние результаты по изучению метода вращений принадлежат совсем недавнему прошлому: в 1990-х годах были обнаружены его особые возможности, связанные с высокоточным вычислением малых по модулю элементов матрицы Λ.

Лекция 21

ОСНОВНАЯ ЧАСТЬ

21.1 Приведенные уравнения поверхности второго порядка

При изучении линий второго порядка мы установили, что любая из них в какой-либо декартовой системе координат описывается одним из основных (как иногда говорят, *приведенных*) уравнений

$$(1)\quad \lambda_1 x^2 + \lambda_2 y^2 + c = 0, \qquad (2)\quad \lambda_2 y^2 + 2bx = 0, \qquad (3)\quad \lambda_2 y^2 + c = 0,$$

в которых все коэффициенты ненулевые, за исключением, быть может, c. В случае поверхности второго порядка исходной точкой для вывода приведенных уравнений является возникающее в некоторой декартовой системе уравнение вида

$$\lambda_1 x^2 \ + \ \lambda_2 y^2 \ + \ \lambda_3 z^2 \ + \ 2b_1 x \ + \ 2b_2 y \ + \ 2b_3 z \ + \ a \ = \ 0.$$

Если $\lambda_1, \lambda_2, \lambda_3 \neq 0$, то с помощью переноса начала координат (сдвига) можно получить уравнение вида

$$\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + c = 0.$$

Пусть $\lambda_3 \neq 0$. Тогда в линейной части с помощью сдвига можно убрать члены, содержащие x и y. В результате появится уравнение вида $\lambda_1 x^2 + \lambda_2 y^2 + 2bz + c = 0$. Если $b \neq 0$, то сдвиг позволяет перейти более простому уравнению $\lambda_1 x^2 + \lambda_2 y^2 + 2bz = 0$. Если же $b = 0$, то получается уравнение вида $\lambda_1 x^2 + \lambda_2 y^2 + c = 0$.

Теперь предположим, что $\lambda_2 = \lambda_3 = 0$. После исключения члена с x в линейной части (путем сдвига) получим уравнение $\lambda_1 x^2 + 2b_2 y + 2b_3 z + c = 0$. Далее, с помощью поворота в плоскости координат y и z в линейной части можно избавиться от члена, содержащего z:

$$\begin{bmatrix} b_2 & b_3 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} b & 0 \end{bmatrix}.$$

В самом деле, выберем ϕ так, чтобы $-b_2 \sin\phi + b_3 \cos\phi = 0$. Таким образом, имеется декартова система координат, в которой заданная поверхность описывается уравнением $\lambda_1 x^2 + 2by + c = 0$. Если $b \neq 0$, то с помощью сдвига легко перейти к уравнению $\lambda_1 x^2 + 2by = 0$. Если $b = 0$, то получается уравнение $\lambda_1 x^2 + c = 0$. В итоге доказано следующее

Утверждение. *Для произвольной поверхности второго порядка в некоторой декартовой системе координат получается приведенное уравнение одного из пяти типов:*

$$(1)\quad \lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + c = 0, \qquad (2)\quad \lambda_1 x^2 + \lambda_2 y^2 + 2bz = 0, \qquad (3)\quad \lambda_1 x^2 + \lambda_2 y^2 + c = 0,$$

$$(4)\quad \lambda_1 x^2 + 2by = 0, \qquad (5)\quad \lambda_1 x^2 + c = 0.$$

Все коэффициенты ненулевые, кроме, возможно, свободного члена c.

21.2 Эллипсоид

Пусть в приведенном уравнении типа (1) коэффициенты $\lambda_1,\ \lambda_2,\ \lambda_3$ имеют одинаковый знак, противоположный знаку свободного члена c. Тогда уравнение приводится к виду

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \qquad a,\ b,\ c > 0.$$

Множество точек (x, y, z), удовлетворяющих этому уравнению, называется *эллипсоидом* с полуосями $a,\ b,\ c$.

Заметим, что эллипсоид целиком содержится в параллелепипеде

$$|x| \le a, \quad |y| \le b, \quad |z| \le c.$$

Ясно, что в сечении поверхности второго порядка плоскостью получается некоторая линия второго порядка. Легко проверяется, что для эллипсоида в любом сечении плоскостью возникает эллипс (вырождающийся в точку, когда плоскость касается эллипсоида).

21.3 Однополостный гиперболоид

Пусть приведенное уравнение имеет тип (1) с отличным от нуля свободным членом. Предположим, что знак одного из коэффициентов при квадратах равен знаку свободного члена и противоположен знаку двух других коэффициентов. Тогда в некоторой декартовой системе координат получается уравнение вида:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \qquad a,\ b,\ c > 0.$$

Множество удовлетворяющих ему точек (x, y, z) называется *однополостным гиперболоидом.*

В любом сечении однополостного гиперболоида плоскостью $x + D = 0$ или $y + D = 0$ возникает гипербола.

По отношению к однополостному гиперболоиду множество всех точек пространства разбивается на три части:

$$\begin{array}{lll} \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 & \text{(точки поверхности)}, \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} < 1 & \text{(внутренние точки)}, \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} > 1 & \text{(внешние точки)}. \end{array}$$

Множество внутренних точек является *связным*: вместе с любыми двумя точками оно целиком содержит все точки некоторой соединяющей их ломаной (состоящей из конечного числа отрезков) линии. Отсюда и название — “однополостный”.

21.4 Линейчатая поверхность

Интересно отметить, что однополостный гиперболоид является примером *линейчатой* поверхности. Так называются поверхности, состоящие из всех точек некоторого бесконечного множества прямых.

Утверждение. *Через каждую точку однополостного гиперболоида S проходят в точности две различные прямые, все точки которых принадлежат S.*

Доказательство. Изменив масштаб, перейдем к аффинной системе координат, в которой уравнение поверхности S будет иметь вид $x^2 + y^2 - z^2 = 1$. Пусть прямая l описывается параметрическими уравнениями

$$x = x_0 + p_1 t, \quad y = y_0 + p_2 t, \quad z = z_0 + p_3 t,$$

а направляющий вектор (p_1, p_2, p_3) выбирается так, чтобы все ее точки принадлежали поверхности S:

$$(x_0+p_1t)^2+(y_0+p_2t)^2-(z_0+p_3t)^2 = 1 \ \ \forall\ t \in \mathbb{R} \quad \Leftrightarrow \quad \left\{ \begin{array}{ccccccc} p_1^2 & + & p_2^2 & - & p_3^2 & = & 0, \\ p_1x_0 & + & p_2y_0 & - & p_3z_0 & = & 0, \\ x_0^2 & + & y_0^2 & - & z_0^2 & = & 1. \end{array} \right.$$

Легко видеть, что $p_3 \neq 0$. Поэтому направляющий вектор можно нормировать, взяв $p_3 = 1$. Тогда $p_1^2 + p_2^2 = 1$, $p_1x_0 + p_2y_0 = z_0$. Предположим, что $y_0 \neq 0 \ \Rightarrow$ $p_2 = (z_0 - p_1x_0)/y_0 \ \Rightarrow \ p_1^2 + (z_0 - p_1x_0)^2/y_0^2 = 1$. Таким образом,

$$(x_0^2 + y_0^2)p_1^2 \ - \ 2(x_0z_0)p_1 \ + \ (z_0^2 - y_0^2) \ = \ 0.$$

Вычисляем дискриминант: $D = x_0^2z_0^2-(x_0^2+y_0^2)(z_0^2-y_0^2) = y_0^2(x_0^2+y_0^2-z_0^2) = y_0^2$. Поскольку $y_0 \neq 0$, для p_1 получаем в точности два различных значения. Поскольку $p_3 = 1$, соответствующие направляющие векторы, очевидно, неколлинеарны. Они дают две различные прямые, целиком принадлежащие S и проходящие через точку (x_0, y_0, z_0). Случай $x_0 \neq 0$ разбирается аналогично. $\square$

Замечание. Для поиска тех же самых прямых на поверхности S можно записать ее уравнение в виде

$$\left(\frac{x}{a} - \frac{z}{c}\right)\left(\frac{x}{a} + \frac{z}{c}\right) = \left(1 - \frac{y}{b}\right)\left(1 + \frac{y}{b}\right)$$

и рассмотреть два семейства пар плоскостей

$$\alpha\left(\frac{x}{a} - \frac{z}{c}\right) = \beta\left(1 - \frac{y}{b}\right), \quad \beta\left(\frac{x}{a} + \frac{z}{c}\right) = \alpha\left(1 + \frac{y}{b}\right),$$
$$\gamma\left(\frac{x}{a} - \frac{z}{c}\right) = \delta\left(1 + \frac{y}{b}\right), \quad \delta\left(\frac{x}{a} + \frac{z}{c}\right) = \gamma\left(1 - \frac{y}{b}\right),$$

определяемых парами не равных одновременно нулю параметров $\alpha,\ \beta$ и $\gamma,\ \delta$. Можно доказать, что для каждой пары плоскостей в пересечении получается прямая, целиком принадлежащая S.

21.5 Двуполостный гиперболоид

Пусть в приведенном уравнении типа (1) знак одного из коэффициентов при квадратах противоположен знаку свободного члена и знаку двух других коэффициентов. Тогда оно приводится к виду

$$\frac{x^2}{a^2} \ + \ \frac{y^2}{b^2} \ - \ \frac{z^2}{c^2} \ = \ -1, \qquad a,\ b,\ c > 0.$$

Множество точек (x, y, z), удовлетворяющих данному уравнению, называется *двуполостным гиперболоидом.*

Легко видеть,что двуполостный гиперболоид не имеет точек в полосе $|z| < c$. Множество его внутренних точек, определяемое неравенством $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} < -1$, разбивается на два связных множества. Отсюда и название — "двуполостный".

21.6 Эллиптический конус

Если в приведенном уравнении типа (1) знак одного из коэффициентов при квадратах противоположен знаку двух других коэффициентов, а свободный член равен нулю, то уравнение можно записать в виде

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

Множество удовлетворяющих ему точек (x, y, z) называется *эллиптическим конусом.*

21.7 Эллиптический параболоид

Теперь рассмотрим приведенное уравнение типа (2). Предположим, что λ_1 и λ_2 имеют одинаковые знаки. Тогда в некоторой декартовой системе данная поверхность описывается уравнением

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z, \qquad a,\ b > 0.$$

Множество удовлетворяющих ему точек называется *эллиптическим параболоидом.*

Название навеяно рассмотрением сечений в плоскостях $z + D = 0$ (эллипсы) и в плоскостях $x + D = 0$ или $y + D = 0$ (параболы).

21.8 Гиперболический параболоид

Если в приведенном уравнении типа (2) коэффициенты при квадратах имеют разные знаки, то получается уравнение

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z, \qquad a,\ b > 0,$$

которое определяет *гиперболический параболоид.*

Название объясняется видом кривых, получаемых в сечениях плоскостями $z+D = 0$ (гиперболы) и плоскостями $x + D = 0$ или $y + D = 0$ (параболы).

Это еще один пример линейчатой поверхности: *каждая точка гиперболического параболоида принадлежит двум различным прямым, целиком принадлежащим данной поверхности.* Доказательство проводится по аналогии со случаем однополостного гиперболоида.

21.9 Цилиндрические поверхности

Приведенные уравнения типов (3)−(5) не зависят от z. Поэтому кривые второго порядка в сечениях любой плоскостью вида $z+D = 0$ одинаковы Соответствующие поверхности называются *цилиндрическими.*

Лекция 22

ОСНОВНАЯ ЧАСТЬ

22.1 Нормированное пространство

В дальнейшем любые линейные пространства будут предполагаться вещественными или комплексными. Наша ближайшая цель — ввести важное обобщение понятия длины геометрического вектора и модуля комплексного числа.

Пусть V — линейное пространство над полем P, где $P = \mathbb{R}$ или $P = \mathbb{C}$. Каждому вектору $x \in V$ припишем вещественное число $||x||$ так, чтобы выполнялись следущие свойства:

(1) $||x|| \geq 0 \quad \forall \ x \in V, \qquad ||x|| = 0 \quad \Leftrightarrow \quad x = 0;$

(2) $||\alpha x|| = |\alpha| \, ||x|| \quad \forall \ x \in V, \ \forall \ \alpha \in P$ (положительная однородность);

(3) $||x + y|| \leq ||x|| + ||y|| \quad \forall \ x, y \in V$ (неравенство треугольника).

Число $||x||$ называется *нормой* вектора x. Линейное пространство V, снабженное нормой, называется *нормированным пространством*.

В одном и том же линейном пространстве норму можно ввести очень многими способами. Например, пусть $V = \mathbb{C}^n$ и $\lambda_1, \ldots, \lambda_n$ — произвольные положительные числа. Если $x = [x_1, \ldots, x_n]^\top$, то пусть

$$||x|| \equiv \sum_{i=1}^{n} \lambda_i |x_i|.$$

Легко проверить, что соответствие $x \mapsto ||x||$ обладает свойствами (1), (2), (3).

Чтобы построить другие, наиболее популярные примеры норм в $\mathbb{C}^n$, нам понадобятся некоторые неравенства, опирающиеся на свойства выпуклых функций.

22.2 Выпуклые функции и неравенства

Вещественная функция $f(x)$ называется *выпуклой* на интервале $I = (a, b)$, если для любых $x, y \in I$ и любого числа $0 \leq t \leq 1$ выполняется неравенство

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y). \qquad (*)$$

Функция $g(x)$ называется *вогнутой* на I, если $f(x) \equiv -g(x)$ выпукла на I.

Теорема. *Пусть функция $f(x)$ дважды дифференцируема на I и $f''(x)$ — ее вторая производная. Если $f''(x) \geq 0$ при всех $x \in I$, то $f(x)$ выпукла на I.*

Доказательство. При $x = y$ неравенство $(*)$ превращается в равенство. При $t = 0$ или $t = 1$ равенство получается при любых x, y. Поэтому предположим, что $a < x < y < b$

и $0 < t < 1$. Тогда для $z = tx + (1-t)y$ имеем $x < z < y$. По теореме Лагранжа из математического анализа, существуют точки ξ и η такие, что

$$\frac{f(z)-f(x)}{z-x} = f'(\xi), \quad x < \xi < z, \qquad \frac{f(y)-f(z)}{y-z} = f'(\eta), \quad z < \eta < y.$$

По той же теореме, для некоторой точки ζ получаем

$$\frac{f(y)-f(z)}{y-z} - \frac{f(z)-f(x)}{z-x} = f''(\zeta)\,(\eta-\xi) \geq 0, \quad \xi < \zeta < \eta.$$

Остается учесть, что $t = (z-x)/(y-x)$ и заметить, что левая часть имеет вид

$$\frac{f(x)(z-x)+f(y)(z-x)-f(z)(y-x)}{(y-z)(z-x)} = \frac{tf(x)+(1-t)f(y)-f(z)}{(y-z)(z-x)}\,(y-x). \qquad \square$$

Следствие. *Функция* $\ln x$ *является вогнутой.*
Доказательство. $(\ln x)'' = -1/x^2 < 0$. $\square$

Отсюда, например, можно сразу же вывести *неравенство между средним арифметическим и средним геометрическим чисел* $x_1, \ldots, x_n > 0$:

$$\sqrt[n]{x_1 \ldots x_n} \leq \frac{x_1 + \ldots + x_n}{n}.$$

В самом деле, используя вогнутость логарифма, находим

$$\ln\left(\frac{x_1+\ldots+x_n}{n}\right) \geq \frac{\ln x_1 + \ldots + \ln x_n}{n} \geq \ln \sqrt[n]{x_1 \ldots x_n}. \qquad \square$$

22.3 Неравенства Гельдера и Минковского

Лемма. *Пусть положительные числа* p, q *таковы, что* $\dfrac{1}{p} + \dfrac{1}{q} = 1$. *Тогда*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \qquad \forall \;\; a, b \geq 0.$$

Доказательство. В силу вогнутости логарифма,

$$\ln(ab) = \frac{\ln a^p}{p} + \frac{\ln b^q}{q} \leq \ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right). \quad \square$$

Неравенство Гельдера. *В условиях леммы для любых комплексных чисел* $x_1, \ldots, x_n$ *и* $y_1, \ldots, y_n$ *справедливо неравенство*

$$\left|\sum_{i=1}^n x_i y_i\right| \leq \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \left(\sum_{i=1}^n |y_i|^q\right)^{1/q}.$$

Доказательство. Пусть

$$a = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}, \quad b = \left(\sum_{i=1}^n |y_i|^q\right)^{1/q}.$$

В случае $a = 0$ или $b = 0$ неравенство $(*)$ очевидно. Если $a \neq 0$ и $b \neq 0$, то, используя лемму для чисел $|x_i|/a$ и $|y_i|/b$, находим

$$(|x_i|/a)\,(|y_i|/b) \;\leq\; \frac{|x_i|^p/a^p}{p} + \frac{|y_i|^q/b^q}{q}, \quad i = 1, \ldots, n.$$

Складывая эти неравенства, получаем

$$\left(\sum_{i=1}^{n} |x_i y_i|\right)/(ab) \;\leq\; \frac{1}{p} \;+\; \frac{1}{q} \;=\; 1. \qquad \square$$

Неравенство Минковского. *Пусть $p \geq 1$, $x_1, \ldots, x_n$ и $y_1, \ldots, y_n$ — произвольные комплексные числа. Тогда*

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{1/p} \;\leq\; \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \;+\; \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}.$$

Доказательство. При $p = 1$ неравенство проверяется очевидным образом. В случае $p > 1$ имеем, очевидно,

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \sum_{i=1}^{n} |x_i + y_i|^{p-1}\,|x_i + y_i| \leq \sum_{i=1}^{n} |x_i|\;|x_i + y_i|^{p-1} \;+\; \sum_{i=1}^{n} |y_i|\;|x_i + y_i|^{p-1}.$$

Для каждой из сумм справа применим неравенство Гельдера, взяв $q = p/(p-1) \quad \Rightarrow \quad \dfrac{1}{p} + \dfrac{1}{q} = 1$. Получаем

$$\sum_{i=1}^{n} |x_i + y_i|^p \leq \left(\left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{1/p}\right)\;\left(\sum_{i=1}^{n} |x_i + y_i|^{(p-1)q}\right)^{1/q}$$

Остается заметить, что $(p-1)q = p$ и $1 - 1/q = 1/p$. $\square$

22.4 Нормы Гельдера

Пусть $x = [x_1, \ldots, x_n]^\top \in \mathbb{C}^n$. При $p \geq 1$ положим

$$||x||_p \;=\; \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

Заметим также, что при фиксированном x величина $||x||_p$ при $p \to \infty$ имеет предел, равный $\max\limits_{1 \leq i \leq n} |x_i|$. Поэтому разумно принять обозначение

$$||x||_\infty \;=\; \max_{1 \leq i \leq n} |x_i|.$$

Величины $||x||_p$ называются *p-нормами* или *нормами Гельдера*.

Неравенства Гельдера и Минковского сохраняют силу при $p = \infty$ (в этом случае $q = 1$). Для доказательства достаточно перейти к пределу при $p \to \infty$.

Теорема. *При любом $p \geq 1$, включая $p = \infty$, величина $||x||_p$ является нормой на $\mathbb{C}^n$.*

Доказательство. Свойства (1) и (2) нормы очевидны. Неравенство треугольника есть не что иное, как неравенство Минковского. $\square$

22.5 Зачем нужны нормы?

Прежде всего, это удобный инструмент для изучения пределов в линейном пространстве.

Последовательность векторов $x^k \in V$ называется *сходящейся к вектору* $x \in V$, если числовая последовательность $||x^k - x||$ сходится к нулю при $k \to \infty$. Вектор x называется *пределом* последовательности x^k. Обозначения: $x = \lim\limits_{k\to\infty} x^k$ или $x^k \to x$ при $x \to \infty$.

Последовательность, сходящаяся к какому-нибудь вектору, называется просто *сходящейся*. Это оправдано, поскольку двух различных пределов быть не может. Если $x^k \to x$ и $x^k \to y$, то

$$||x - y|| = ||(x - x^k) - (y - x^k|| \le ||x - x^k|| + ||y - x^k|| \to 0 \quad \Rightarrow \quad x = y. \quad \square$$

В конечномерном пространстве V при изучении сходимости можно, в принципе, обойтись и без норм. Фиксировав какой-нибудь базис $e_1, \ldots, e_n \in V$, мы могли бы рассмотреть разложения

$$x^k \;=\; \sum_{i=1}^{n} x_i^k \, e_i$$

и называть последовательность векторов x^k сходящейся, если сходятся координатные последовательности x_i^k при всех i. Такое понятие сходимости не будет зависеть от выбора базиса (докажите!). Легко видеть также, что *из покоординатной сходимости в конечномерном пространстве вытекает сходимость по любой норме.* Действительно, пусть $x_i^k \to x_i$. Тогда, взяв $x = \sum\limits_i x_i \, e_i$, получаем

$$||x^k - x|| \;\le\; \sum_{i=1}^{n} |x_i^k - x_i| \, ||e_i||. \qquad \square$$

Более того, имеет место и менее очевидный факт: в конечномерном пространстве из сходимости по любой норме вытекает покоординатная сходимость. Мы скоро это докажем.

Тем не менее, даже в конечномерном пространстве исследовать сходимость с помощью норм очень удобно: все сводится к изучению лишь одной числовой последовательности $||x^k - x||$. Это тем более важно, когда пространство бесконечномерно!

22.6 Нормы в бесконечномерном пространстве

ПРИМЕР 1. Пусть $C[a,b]$ — линейное пространство функций, непрерывных на отрезке $[a,b]$. Для функции $f \in C[a,b]$ наиболее часто используется норма

$$||f||_C \;=\; \max_{a\le x\le b} |f(x)|,$$

называемая *C-нормой* (иногда также *равномерной* или *чебышевской* [1]).

ПРИМЕР 2. Пусть $C^1[a,b]$ — линейное пространство функций, непрерывных на отрезке $[a,b]$ вместе с первой производной [2]. В данном случае норму можно ввести, например, так:

$$||f||_{C^1} \;=\; \max_{a\le x\le b} (|f(x)| + |f'(x)|).$$

[1] В честь знаменитого русского математика Пафнутия Львовича Чебышева.

[2] Чтобы рассматривать $f'(x)$ в точках a и b, можно считать функцию $f(x)$ определенной и дифференцируемой на более широком интервале, накрывающем $[a,b]$.

Заметим, что сходимость последовательности функций из $C^1[a,b]$ по норме C^1 влечет за собой сходимость по норме C. Обратное, однако, не верно: последовательность функций

$$f^k(x) = \frac{\sin kx}{k}$$

принадлежит $C^1[a,b]$ и сходится по норме C к нулю, но не является сходящейся по норме C^1 (докажите!).

Таким образом, в бесконечномерных пространствах разные нормы определяют, вообще говоря, разные типы сходимости. В этом отношении конечномерные пространства отличаются принципиально: в них сходимость по какой-либо норме равносильна сходимости по любой другой норме — это фундаментальный факт, который скоро будет доказан. Он вроде бы означает, что в конечномерных пространствах можно ограничиться изучением какой-нибудь одной нормы. Тем не менее, это не так! В огромном числе вопросов конечномерные пространства возникают как подпространства бесконечномерного нормированного пространства. Поэтому нормы в них должны порождаться нормой соответствующего бесконечномерного пространства. А мы только что выяснили, что для бесконечномерных пространств разные нормы могут отличаться существенным образом.

22.7 Метрическое пространство

В понятии предела аксиомы линейного пространства используются, на самом деле, не очень существенным образом — норма разности двух векторов легко заменяется более общим понятием расстояния между двумя векторами.

Пусть M — непустое множество и $\rho(x,y)$ — вещественная функция от элементов $x, y \in M$, обладающая следующими свойствами:

(1) $\rho(x,x) \geq 0 \quad \forall \ \ x \in M, \qquad \rho(x,x) = 0 \quad \Leftrightarrow \quad x = 0;$

(2) $\rho(x,y) = \rho(y,x) \quad \forall \ \ x, y \in M;$

(3) $\rho(x,y) \ \leq \ \rho(x,z) \ + \ \rho(z,y) \quad \forall \ \ x, y, z \in M.$

В таких случаях M называется *метрическим пространством*, а $\rho(x,y)$ — *расстоянием* между элементами x и y.

Любое нормированное пространство является метрическим пространством с расстоянием

$$\rho(x,y) \ = \ ||x - y||.$$

Однако, метрическое пространство в общем случае не предполагает наличия каких-либо операций над его элементами. Например, произвольное непустое множество M будет метрическим пространством, если $\rho(x,y) = 0$ при $x = y$ и $\rho(x,y) = 1$ при $x \neq y$.

22.8 Пределы и полнота

Пусть M — метрическое пространство. Последовательность элементов $x^k \in M$ называется *сходящейся* в M, если существует элемент $x \in M$ такой, что числовая последовательность $\rho(x^k, x)$ сходится к нулю при $k \to \infty$. Как и в нормированном пространстве, двух разных пределов быть не может: если $x^k \to x$ и $x^k \to y$, то

$$\rho(x,y) \ \leq \ \rho(x,x^k) \ + \ \rho(x^k,y) \ \to \ 0 \quad \Rightarrow \quad x = y.$$

Последовательность $x^k \in M$ называется *фундаментальной* или *последовательностью Коши*, [3] если для любого $\varepsilon > 0$ существует номер $N = N(\varepsilon)$ такой, что при $k, l > N$ выполняется неравенство $\rho(x^k, x^l) < \varepsilon$.

Из неравенства $\rho(x^k, x^l) \le \rho(x^k, x) + \rho(x, x^l)$ очевидно, что *любая сходящаяся последовательность является последовательностью Коши*. Обратное в общем случае не верно. Например, любой интервал $M = (a, b)$ вещественной оси можно рассматривать как метрическое пространство с расстоянием $\rho(x, y) = |x - y|$. Последовательность $x^k = a + (b - a)/k$ является фундаментальной, но не может сходиться ни к какому элементу из M (ее пределом должно бы быть число a, но $a \notin M$).

Метрическое пространство называется *полным*, если в нем любая фундаментальная последовательность является сходящейся.

В начальных курсах математического анализа обычно доказывается, что фундаментальные последовательности чисел из $\mathbb{R}$ являются сходящимися в $\mathbb{R}$ — таким образом, метрическое пространство $\mathbb{R}$ с расстоянием $\rho(x, y) = |x - y|$ является полным.

Все понятия и факты, полученные для метрических пространств, переносятся на произвольные нормированные пространства. При этом всегда предполагается, что расстояние в них вводится с помощью нормы: $\rho(x, y) = ||x - y||$. Полное нормированное пространство называется также *банаховым*. [4]

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

22.9 Пополнение пространства

Пример интервала (a, b) (неполного метрического пространства с расстоянием $\rho(x, y) = |x - y|$) наводит на мысль о том, что если в неполном пространстве не хватает точек для того, чтобы представлять пределы всех возможных фундаментальных последовательностей, то его с легкостью можно расширить до полного метрического пространства. Эта идея реализуется с помощью простой алгебраической конструкции пополнения.

Пусть M — произвольное метрическое пространство. Рассмотрим множество всех фундаментальных последовательностей точек из M, введем на нем отношение эквивалентности

$$\{x^k\} \sim \{y^k\} \quad \Leftrightarrow \quad \rho(x^k, y^k) \to 0 \ \text{ при } \ k \to \infty,$$

и обозначим через M' множество всех классов эквивалентности. Расстояние на M' определим таким образом: если $[\{x^k\}]$ и $[\{y^k\}]$ — классы эквивалентности, порождаемые фундаментальными последовательностями $\{x^k\}$ и $\{y^k\}$, то пусть

$$\rho'([\{x^k\}],\ [\{y^k\}]) \ = \ \lim_{k\to\infty} \rho(x^k, y^k).$$

Конечный предел существует потому, что числовая последовательность $\rho(x^k, y^k)$ фундаментальна — это легко получается из неравенства

$$|\rho(x^k, y^k) - \rho(x^l, y^l)| \ \le \ \rho(x^k, x^l) + \rho(y^k, y^l).$$

Важно также, что предел не зависит от выбора конкретных последовательностей в классах эквивалентности: если $\{x^k\} \sim \{u^k\}$ и $\{y^k\} \sim \{v^k\}$, то

$$|\rho(u^k, v^k) - \rho(x^k, y^k)| \ \le \ \rho(u^k, x^k) + \rho(y^k, v^k) \ \to \ 0.$$

Аксиомы метрического пространства для M' с расстоянием ρ' проверяются без каких-либо затруднений.

Элемент $a \in M$ будем отождествлять с классом эквивалентности для последовательности, все члены которой одинаковы и равны a:

$$a \quad \leftrightarrow \quad A = [\{a^k\}], \quad \text{где} \quad a^k = a \quad \forall \ k.$$

[3] Еще одно (красивое, но редко используемое) название — *сходящаяся в себе*.

[4] В честь польского математика, профессора Львовского университета Стефана Банаха.

Путь $a, b \in M$ и $A, B \in M'$. Тогда если $a \leftrightarrow A$ и $b \leftrightarrow B$, то $\rho(a,b) = \rho'(A,B)$.

Таким образом, можно считать, что M вложено в M'; построенное нами метрическое пространство M' называется *пополнением* метрического пространства M.

Утверждение. *M' является полным метрическим пространством.*

Доказательство. Пусть классы эквивалентности $X_1 = [\{x_1^k\}]$, $X_2 = [\{x_2^k\}]$, ... образуют в M' фундаментальную последовательность.

При любом фиксированном l последовательность x_l^k, $k = 1, 2, \ldots$, является фундаментальной. Поэтому существует номер $n_l \geq l$ такой, что $\rho(x_l^{n_l}, x_l^k) < 1/l$ при всех $k \geq n_l$. Определим последовательность $\{y^l\}$ равенствами $y^l = x_l^{n_l}$, $l = 1, 2 \ldots$, и докажем, что она является фундаментальной. Фиксируем произвольное $\varepsilon > 0$. В силу фундаментальности $\{X_l\}$ существует номер $N = N(\varepsilon)$ такой, что при $l, m > N$ имеем $\rho'(X_l, X_m) < \varepsilon$, то есть,

$$\exists\, N' = N'(l,m,N): \quad l,m > N,\ k > N' \quad \Rightarrow \quad \rho(x_l^k, x_m^k) < \varepsilon.$$

Для любых $l, m > \max\{N, \varepsilon^{-1}\}$ и $k > \max\{n(l,m,\varepsilon),\ n_l,\ n_m\}$ находим

$$\rho(y^l, y^m) \ = \ \rho(x_l^{n_l}, x_m^{n_m}) \ \leq \ \rho(x_l^{n_l}, x_l^k) \ + \ \rho(x_l^k, x_m^k) \ + \ \rho(x_m^k, x_m^{n_m}) \ < \ 3\varepsilon.$$

Остается ввести класс эквивалентности $Y = [\{y^k\}] \in M'$ и доказать, что $X^l \to Y$. Это вытекает из неравенства

$$\rho(x_l^k, y^k) \ \leq \ \rho(x_l^k, x_l^{n_l}) + \rho(x_l^{n_l}, x_k^{n_k}). \qquad \square$$

Заметим, что в M' нет "лишних" элементов: каждый элемент $Y \in M'$ является пределом последовательности элементов из M (докажите!).

Ту же технику можно применить для пополнения нормированного пространства M с нормой $||\cdot||_M$. В данном случае следует ввести на M' операции сложения классов эквивалентности и умножения их на числа как операции над порождающими эти классы последовательностями. Эти операции не выводят из множества M', так как сумма фундаментальных последовательностей, умноженных на любые числа, остается фундаментальной последовательностью.

Таким образом, M' становится линейным пространством, а M можно рассматривать как его подпространство. Норма в M' для $[\{x^k\}]$ определяется следующим образом:

$$||\,[\{x^k\}]\,||_{M'} \ = \ \lim_{k\to\infty} ||x^k||_M.$$

Существование предела следует из неравенства $|\,||x^k||_M - ||x^l||_M\,| \leq ||x^k - x^l||_M$.

Лекция 23

ОСНОВНАЯ ЧАСТЬ

23.1 Множества в метрическом пространстве

Пусть M — метрическое пространство, $a \in M$ и $r > 0$. Множества

$$M(a,r) \;=\; \{x \in M : \rho(a,x) < r\}, \qquad \overline{M}(a,r) \;=\; \{x \in M : \rho(a,x) \le r\}.$$

называются *открытым шаром* и *замкнутым шаром* радиуса r с центром в точке a.

Пусть S — какое-то множество точек в метрическом пространстве M. Множество S называется *ограниченным*, если оно целиком содержится в некотором шаре.

Точка $a \in S$ называется *внутренней* для S, если она содержится в S вместе с некоторым открытым шаром. Множество S называется *открытым* в M, если любая его точка является внутренней. Пустое множество по определению считается открытым.

Пусть $x \in M$ и существует последовательность точек $x^k \in S$, сходящаяся к x. В этом случае x называется *точкой прикосновения* для S. Если $x^k \ne x$ для всех k, то x называется *предельной точкой* для S. Очевидно, любая точка прикосновения, не принадлежащая множеству S, является для него предельной.

Замыканием множества S называется оно само плюс все его предельные точки. Обозначение: $[S]$. Множество S называется замкнутым, если оно содержит все свои предельные точки: $[S] = S$. Несложно проверить, что S замкнуто в том и только в том случае, когда дополнительное в M множество $O = M \backslash S$ является открытым.

Задача. Всегда ли замыкание открытого шара совпадает с замкнутым шаром с тем же центром и радиусом?

Множество S называется *компактным*, если из любой последовательности точек $x^k \in S$ можно выделить подпоследовательность, сходящуюся к некоторой точке $x \in S$.

Ясно, что компактное множество обязано быть замкнутым. Обратное не верно: например, $S = M$ всегда является замкнутым множеством, но может и не быть компактным. Заметим также, что любое компактное множество S является ограниченным (подпоследовательность неограниченной последовательности не может быть сходящейся, так как не может быть ограниченной).

В начальных курсах анализа рассматривается метрическое пространство $\mathbb{R}$ с расстоянием $\rho(x,y) = |x-y|$, а компактным принято называть любое замкнутое и ограниченное множество точек из $\mathbb{R}$. В данном случае это определение равносильно нашему определению компактности. Более того, мы скоро докажем, что эти два определения

равносильны и в случае произвольных конечномерных нормированных пространств. Однако, в бесконечномерных пространствах замкнутость и ограниченность недостаточны для выделения сходящейся подпоследовательности.

23.2 Компактность и непрерывность

Вещественная функция $f(x)$, определенная для точек x метрического пространства M, называется *непрерывной* в точке $x \in M$, если для любой последовательности x^k, сходящейся к x, последовательность значений $f(x^k)$ сходится к $f(x)$.

Теорема Вейерштрасса. *Для любой вещественной функции $f(x)$, непрерывной во всех точках компактного множества S, существуют точки $x_{\min}, x_{\max} \in S$ такие, что $f(x_{\min}) \le f(x) \le f(x_{\min})$ для всех $x \in S$.*

Доказательство. Если предположить, что $f(x^k) > k$ для некоторой последовательности точек $x^k \in S$, то возникает противоречие с возможностью выделения сходящейся подпоследовательности: если $x^{k_i} \to x$, то $f(x^{k_i}) \to f(x)$, но $f(x^{k_i})$ не может сходиться, так как не является ограниченной. Поэтому $f(x)$ ограничена сверху. Пусть $c_{\max}$ — точная верхняя грань множества значений $\{f(x), x \in S\}$. Тогда для каждого k найдется точка $x^k \in S$ такая, что $c_{\max} - 1/k \le f(x^k) \le c_{\max}$. Выберем сходящуюся подпоследовательность $x^{k_i} \to x$ и перейдем в последних неравенствах к пределу $\Rightarrow \quad f(x) = c_{\max}$.

Ограниченность снизу и существование точки минимума доказывается переходом к $g(x) = -f(x)$. $\square$

23.3 Компактность единичной сферы

Рассмотрим единичную сферу в пространстве $\mathbb{C}^n$ относительно 2-нормы:

$$S_2 \;=\; \{x \in \mathbb{C}^n : ||x||_2 = 1\} \;=\; \{x = [x_1, \ldots, x_n]^\top : \sum_{i=1}^{n} |x|_i^2 = 1\}.$$

Лемма 1. *Единичная сфера S_2 в пространстве $\mathbb{C}^n$ компактна относительно 2-нормы.*

Доказательство. Рассмотрим произвольную последовательность векторов

$$x^k = [x_1^k, \ldots, x_n^k]^\top \in S_2.$$

Соответствующие координатные последовательности удовлетворяют неравенствам

$$|x_1^k| \le 1, \;\; |x_2^k| \le 1, \;\ldots,\; |x_n^k| \le 1.$$

Согласно лемме об ограниченных последовательностях (см. Лекцию 19), существует подпоследовательность номеров $k_1 < k_2 < \ldots$ такая, что каждая из координатных последовательностей $x_i^{k_l}$ будет сходиться и удовлетворять равенству

$$\sum_{i=1}^{n} |x_i^{k_l}|^2 \;=\; 1. \qquad (*)$$

Пусть $x_i = \lim\limits_{l\to\infty} x_i^{k_l}$ и $x = [x_1, \ldots, x_n]^\top$. Тогда

$$||x^{k_l} - x||_2 \;=\; \left(\sum_{i=1}^{n} |x_i^{k_l} - x_i|^2\right)^{1/2} \to 0.$$

Переходя в $(*)$ пределу, получаем $x \in S_2$. $\square$

Лемма 2. *Для произвольной нормы $||\cdot||$ в пространстве $\mathbb{C}^n$ функция $f(x) = ||x||$ является непрерывной относительно 2-нормы.*

Доказательство. Пусть $x^k = [x_1^k, \ldots, x_n^k]^\top \to x = [x_1, \ldots, x_n]^\top$. Тогда, используя неравенство треугольника для норм, находим

$$|f(x^k) - f(x)| = |\,||x^k|| - ||x||\,| \;\le\; ||x^k - x|| \;\le\; \sum_{1\le i\le n} |x_i^k - x_i|\,||e_i||,$$

где $e_i = [0, \ldots, 1, \ldots 0]^\top$ — вектор из нулей, кроме i-й компоненты, равной 1. Правая часть стремится к нулю при

$$||x^k - x||_2 \;=\; \left(\sum_{i=1}^n |x_i^k - x_i|^2\right)^{1/2} \;\to\; 0. \qquad \square$$

Лемма 3. *Для любой нормы $||\cdot||$ на $\mathbb{C}^n$ существуют константы $c_1, c_2 > 0$ такие, что*

$$c_1 \;\le\; ||x|| \;\le\; c_2 \qquad \forall\;\; x \in S_2.$$

При этом $c_1 = ||x^1||$, $c_2 = ||x^2||$ для некоторых векторов $x^1,\ x^2 \in S_2$.

Доказательство. Достаточно заметить, что функция $f(x) = ||x||$ непрерывна относительно 2-нормы на множестве S_2, компактном относительно 2-нормы. $\square$

23.4 Эквивалентные нормы

Две нормы $||\cdot||_{(a)}$ и $||\cdot||_{(b)}$ на одном и том же линейном пространстве V называются *эквивалентными*, если существуют константы $c_1,\ c_2 > 0$ такие, что

$$c_1\,||x||_{(a)} \;\le\; ||x||_{(b)} \;\le\; c_2\,||x||_{(a)} \qquad \forall\;\; x \in V.$$

Теорема. *Если V конечномерно, то любые нормы на нем эквивалентны.*

Доказательство. Прежде всего, заметим, что любая норма $||\cdot||$ на $\mathbb{C}^n$ эквивалентна $||\cdot||_2$. Пусть $x \in \mathbb{C}^n \;\;\Rightarrow\;\; x/||x||_2 \in S_2$. По лемме 3, $c_1 \;\le\; \|x/||x||_2\| \;\le\; c_2 \;\;\Rightarrow$

$$c_1\,||x||_2 \;\le\; ||x|| \;\le\; c_2\,||x||_2 \qquad \forall\;\; x \in \mathbb{C}^n.$$

Отсюда легко вывести эквивалентность любых двух норм на $\mathbb{C}^n$.

В случае произвольного конечномерного пространства V с нормой $||\cdot||_V$ фиксируем в нем произвольный базис e_1, …, e_n и рассмотрим взаимно-однозначное соответствие

$$v \;\leftrightarrow\; [x_1, \ldots, x_n]^\top, \qquad v = \sum_{i=1}^n x_i e_i.$$

Используя его, введем норму на $\mathbb{C}^n$ следующим образом:

$$||[x_1, \ldots, x_n]^\top||_V \;\equiv\; \left\| \sum_{i=1}^n x_i e_i \right\|_V.$$

Свойства нормы проверяются непосредственно. Введем также еще одну норму на V:

$$\left\| \sum_{i=1}^{n} x_i e_i \right\|_2 \equiv ||[x_1, \ldots, x_n]^\top||_2.$$

Уже установленная эквивалентность любых двух норм на $\mathbb{C}^n$ доказывает, очевидно, эквиваленость данных (а значит, и любых) норм в пространстве V. $\square$

Следствие. *Сходимость по любой норме в конечномерном пространстве равносильна поокоординатной сходимости.*

Заметим, что нам уже встречались нормы, которые не могут быть эквивалентными: это C-норма и C^1-норма в пространстве $C^1[a,b]$ функций, непрерывных на отрезке $[a,b]$ вместе с первой производной:

$$||f||_C = \max_{a\le x\le b} |f(x)|, \qquad ||f||_{C^1} = \max_{a\le x\le b} (|f(x)| + |f'(x)|).$$

В самом деле, последовательность функций $f^k(x) = \sin kx/k$ является сходящейся в норме C, но расходится в норме C^1 (поскольку не является последовательностью Коши в данной норме). Отсюда, кстати, получаем (не очень прямое!) доказательство бесконечномерности линейного пространства $C^1[a,b]$.

23.5 Компактность замкнутых ограниченных множеств

Теорема. *В конечномерном нормированном пространстве множество является компактным тогда и только тогда, когда оно замкнуто и ограничено.*

Доказательство. Мы уже знаем, что компактное множество в метрическом пространстве всегда является замкнутым и ограниченным. Пусть множество S замкнуто и ограничено относительно какой-то нормы в $\mathbb{C}^n$. В силу эквивалентности норм в конечномерном пространстве, S также замкнуто и ограничено относительно 2-нормы. Поэтому любая последовательность векторов из S имеет ограниченные координатные последовательности. По лемме об ограниченных последовательностях, мы можем выбрать подпоследовательность, сходящуюся в 2-норме к какому-то вектору $x \in S$. Эта же подпоследовательность будет сходиться и относительно любой другой нормы. $\square$

Отсюда вытекает, например, компактность единичной сферы и компактность замкнутого шара в любом конечномерном пространстве относительно любой нормы.

23.6 Наилучшие приближения

Пусть $x \in V$ и L — непустое множество векторов из V. Величину

$$\gamma = \inf_{z\in L} ||x - z||$$

называют *расстоянием* между x и L. Вектор $z_0 \in L$ называется *элементом наилучшего приближения* для x на L, если $\gamma = ||x - z_0||$.

Лемма о наилучшем приближении. *Пусть L — конечномерное подпространство в нормированном пространстве V. Тогда для любого $x \in V$ существует вектор $z_0 \in L$ такой, что $||x - z_0|| \le ||x - z|| \quad \forall \quad z \in L$.*

Доказательство. Фиксируем $\varepsilon > 0$ и рассмотрим любой вектор z такой, что $||x-z|| \le \gamma + \varepsilon$. Отсюда $||z|| \le R \equiv \gamma + \varepsilon + ||x||$. Поэтому очевидно, что

$$\gamma = \inf_{z\in L,\ ||z||\le R} ||x - z||.$$

Функция $f(z) = ||x - z||$ непрерывна на замкнутом шаре $||z|| \le R$ конечномерного пространства L. По теореме Вейерштрасса, $\gamma = ||x - z_0||$ для некоторого $z_0 \in L$. $\square$

Заметим, что существование элемента наилучшего приближения очевидно также для компактных множеств L.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

23.7 Подпространства и замкнутость

Если V — нормированное пространство, то любое его подпространство $L \subset V$ конечной размерности будет замкнутым множеством.

В самом деле, если $L = L(e_1, \ldots, e_m)$ и $x^k = \sum\limits_{i=1}^{m} x_i^k e_i \to x \in V$, то последовательность x^k ограничена по норме пространства V. Следовательно, она принадлежит какому-то замкнутому ограниченному шару $Z \subset L$ в конечномерном пространстве L. В силу компактности Z существует подпоследовательность x^{k_l}, сходящаяся к вектору из $Z \;\Rightarrow$ $x \in Z \subset L$.

Если подпространство L бесконечномерно, то оно может и не быть замкнутым. [1]

Задача. Дана матрица $A \in \mathbb{R}^{m \times n}$. Доказать замкнутость множества

$$\{y = Ax, \;\; x = [x_1, \ldots, x_n]^\top, \;\; x_1, \ldots, x_n \ge 0\}.$$

23.8 Единичная сфера в бесконечномерном пространстве

Пусть V — нормированное пространство с нормой $||\cdot||$ и $S = \{x \in V : ||x|| = 1\}$ — единичная сфера.

Теорема. *Единичная сфера S компактна в нормированном пространстве V тогда и только тогда, когда V конечномерно.*

Доказательство. По существу, нужно доказать лишь то, что в бесконечномерном пространстве V сфера S не является компактным множеством. Предположим, что каким-то образом найдены векторы $x_1, \ldots, x_k$ такие, что

$$||x_1|| \;=\; \ldots \;=\; ||x_k|| \;=\; 1, \qquad ||x_i - x_j|| \ge 1 \;\text{ при }\; i \ne j. \qquad (**)$$

Построим вектор x_{k+1} такой, что $||x_{k+1}|| = 1$ и $||x_i - x_{k+1}|| \ge 1$ при $1 \le i \le k$.

Будучи бесконечномерным, V содержит $y \notin L_k = L(x_1, \ldots, x_k)$. По лемме о наилучшем приближении, для некоторого $z_0 \in L_k$

$$\gamma \;=\; \inf_{z \in L_k} ||y - z|| \;=\; ||y - z_0||.$$

Положим

$$x_{k+1} \;=\; (y - z_0)/\gamma.$$

[1] В некоторых книгах под подпространствами в бесконечномерном случае понимаются только замкнутые подпространства, а подпространства в традиционном для нас смысле называются линейными многообразиями (и могут не быть замкнутыми). Напомним, что в нашем курсе линейным многообразием называется множество вида $x + L = \{x + h : h \in L\}$, где x — заданный вектор сдвига, а L — заданное направляющее подпространство.

Тогда $||x_{k+1}|| = 1$ и, кроме того,

$$\min_{z\in L_k} ||x_{k+1} - z|| \;=\; \min_{z\in L_k} ||(y - z_0)/\gamma - z/\gamma|| \;=\; \frac{1}{\gamma}\min_{z\in L_k} ||y - z|| \;=\; 1.$$

Поскольку $x_i \in L_k$ при $1 \le i \le k$, находим $||x_{k+1} - x_i|| \ge \inf\limits_{z\in L_k} ||x_{k+1} - z|| = 1$.

Таким образом, к системе векторов $x_1, \ldots, x_k$ можно добавить вектор x_{k+1} с сохранением соотношений вида $(**)$. Любая подпоследовательность последовательности векторов x_k таких, что $||x_i - x_j|| \ge 1$ при $i \ne j$, обладает тем же свойством и поэтому не может быть фундаментальной. $\square$

23.9 Геометрические свойства единичных шаров

Пусть дана произвольная норма $||\cdot||$ на $\mathbb{C}^n$, а замкнутый единичный шар

$$Z = \{x \in \mathbb{C}^n : ||x|| \le 1\}$$

рассматривается как некоторое множество в пространстве $\mathbb{C}^n$ с 2-нормой. Легко показать, что имеют место такие свойства:

(1) Z является замкнутым и ограниченным.

(2) Z содержит нулевой вектор в качестве внутренней точки.

(3) Если $x \in Z$, то $tx \in Z$ для всех $|t| \le 1$.

(4) Если $x,\ y \in Z$, то $tx + (1-t)y \in Z$ для всех $0 \le t \le 1$ (множества с таким свойством называются *выпуклыми*).

Теорема. *Для того чтобы множество $Z \subset \mathbb{C}^n$ было замкнутым единичным шаром для какой-нибудь нормы на $\mathbb{C}^n$, необходимо и достаточно выполнение свойств* **(1)**–**(4)**.

Доказательство. Рассмотрим множество Z, обладающее указанными свойствами, и попытаемся ввести норму таким образом: [2]

$$f(x) = \inf\{t > 0 : x/t \in Z\}, \quad x \in \mathbb{C}^n. \qquad (\#)$$

Прежде всего, заметим, что $f(x)$ принимает конечные значения для всех x. Согласно условию **(2)**, в Z содержится окрестность нуля вида $O = \{||x||_2 < \varepsilon\}$, где $\varepsilon > 0$. Поэтому для любого $x \ne 0$ имеем $x/t \in O \subset Z$ при $t > ||x||_2/\varepsilon \;\Rightarrow\; f(x) \le ||x||_2/\varepsilon$. Ясно также, что $f(0) = 0$ и $f(x) > 0$ при $x \ne 0$ (первое свойство нормы).

Второе свойство (положительная однородность) доказывается так. Пусть $t_k \to f(x)$ и $x/t_k \in Z$. Предположим, что $\alpha \ne 0$. Поскольку $x/t_k \in Z$, то, в силу свойства **(3)**,

$$(\alpha/|\alpha|)(x/t_k) \in Z \;\Rightarrow\; (\alpha x)/(|\alpha_k|\, t_k) \in Z \;\Rightarrow\; f(\alpha x) \le |\alpha|\, t_k \;\to\; |\alpha|\, f(x).$$

Следовательно, $f(\alpha x) \le |\alpha|\, f(x)$. Противоположное неравенство доказывается аналогично — с выбором последовательности $t_k \to f(\alpha x),\ (\alpha x)/t_k \in Z$.

Докажем неравенство треугольника. Пусть

$$\alpha_k \to f(x),\ \ x/\alpha_k \in Z, \qquad \beta_k \to f(y),\ \ y/\beta_k \in Z.$$

Согласно выпуклости Z, находим

$$\frac{\alpha_k}{\alpha_k + \beta_k}(x/\alpha_k) + \frac{\beta_k}{\alpha_k + \beta_k}(y/\beta_k) \;=\; (x+y)/(\alpha_k + \beta_k) \in Z.$$

Отсюда $f(x+y) \le \alpha_k + \beta_k \;\to\; f(x) + f(y)$. $\square$

Заметим, что Минковский определял нормы именно с помощью функции вида $(\#)$ и множеств, обладающих свойствами **(1)**–**(4)**. Аксиоматический подход к определению нормы был предложен несколько позже (в 1922 году) независимо Банахом и Винером.

Доказанная нами теорема легко обобщается на случай бесконечномерных пространств. Все остается без изменений, если вместо 2-нормы выбрать и зафиксировать любую норму, относительно которой будут затем определяться понятия сходимости, окрестности, замкнутости и ограниченности.

[2] Функция такого вида называется *функционалом Минковского*.

23.10 Топологические пространства

В действительности при изучении сходимости понятие расстояния нужно лишь для того, чтобы определять, какие точки считаются "близкими". В метрическом пространстве M можно объявить, что "близкие" точки — это точки, входящие в одно и то же открытое множество. Обычно любое открытое множество, содержащее заданную точку, называется также ее *окрестностью*. Последовательность точек $x^k \in M$ сходится к точке $x \in M$, если в любой ее окрестности содержатся все точки x^k, начиная с некоторой. Это предложение не опирается явным образом на понятие расстояния и часто принимается в качестве определения сходящейся последовательности.

Обозначим через $\mathcal{T}$ систему всех открытых множеств точек из M. Несложно проверить, что система $\mathcal{T}$ обладает следующими свойствами:

(1) $\mathcal{T}$ содержит M и пустое множество $\emptyset$;

(2) объединение любого (конечного или бесконечного) числа множеств из $\mathcal{T}$ принадлежит $\mathcal{T}$;

(3) пересечение любого конечного числа множеств из $\mathcal{T}$ принадлежит $\mathcal{T}$. [3]

Пусть теперь M — произвольное непустое множество, а $\mathcal{T}$ — произвольная система его подмножеств, обладающая свойствами (1) – (3). Тогда $\mathcal{T}$ называется *топологией* на M, сами множества, входящие в $\mathcal{T}$, объявляются *открытыми*, а множество M, снабженное топологией, называется *топологическим пространством*.

В топологическом пространстве сходимость определяется отмеченным выше образом. Понятие предельной точки, замыкания и замкнутого множества опираются исключительно на понятие сходящейся последовательности и вводятся так же, как в метрическом пространстве.

23.11 Компактные множества в топологическом пространстве

Открытым покрытием множества $S \subset M$ в топологическом пространстве M называется любая совокупность открытых множеств, объединение которых содержит S. Покрытие, состоящее из части данных множеств, называется *подпокрытием*, а если оно состоит их конечного числа открытых множеств, то — *конечным подпокрытием*.

Множество S называется *компактным в топологическом пространстве* M, если из любого его открытого покрытия можно выделить конечное подпокрытие.

Утверждение. *Любое компактное в топологическом пространстве множество замкнуто и таково, что из любой принадлежащей ему последовательности точек можно выделить сходящуюся подпоследовательность.*

Доказательство. Если множество $\{x^k\}$ имеет предельную точку, принадлежащую заданному компактному множеству S, то все доказано. Если это не так, то для каждой точки $x \in S$ существует открытое множество O_x, содержащее лишь конечное число точек последовательности x^k. Очевидно, множества O_x образуют открытое покрытие множества S $\Rightarrow$ существует конечное подпокрытие $\Rightarrow$ в множестве S имеется лишь конечное число точек множества $\{x^k\}$ $\Rightarrow$ последовательность x^k имеет бесконечное число одинаковых точек. $\square$

Теорема. *Для того чтобы множество в метрическом пространстве было компактным в соответствующем топологическом пространстве, необходимо и достаточно, чтобы оно было замкнутым и таким, что в любой принадлежащей ему последовательности выделяется сходящаяся подпоследовательность.*

Доказательство достаточности. Пусть речь идет о множестве S. Прежде всего, заметим, что для любого $\varepsilon > 0$ оно покрывается конечной системой открытых шаров радиуса не больше ε. [4] Если это не так для какого-то ε, то существует точка $a_1 \in S$ такая, что S не покрывается шаром $M(a_1, \varepsilon)$ $\Rightarrow$ $\rho(a_1, a_2) \geq \varepsilon$ для некоторой точки $a_2 \in S$ и при этом S не покрывается системой двух шаров $M(a_1, \varepsilon)$ и $M(a_2, \varepsilon)$, и так далее. В итоге получается последовательность точек $a_k \in S$ таких, что $\rho(a_i, a_j) \geq \varepsilon$ при $i \neq j$ $\Rightarrow$ из последовательности a_k нельзя выделить сходящуюся подпоследовательность.

Рассмотрим конечные покрытия шарами последовательно для $\varepsilon = 1, 1/2, 1/3, \ldots$ и обозначим через $\mathcal{B}$ множество всех этих шаров. Пусть имеется произвольное открытое покрытие множества S.

[3] Пересечение бесконечного числа открытых множеств может и не быть открытым (например, пересечение всех открытых множеств, содержащих данную точку).

[4] Такое покрытие называется ε-сетью.

Любая точка любого открытого множества принадлежит некоторому шару из $\mathcal{B}$. Поэтому существует открытое покрытие S некоторой последовательностью шаров из $\mathcal{B}$. Следовательно, из заданного открытого покрытия множества S можно выбрать счетное подпокрытие — другими словами, S покрывается некоторой последовательностью открытых множеств O_k.

Если из системы множеств O_k нельзя выбрать какое-либо конечное подпокрытие множества S, то каждое из замкнутых множеств $Z_k = S\setminus\left(\bigcup\limits_{1\le i\le k} O_i\right)$ непустое. При этом $Z_1 \supset Z_2 \supset Z_3 \supset \ldots$. Пусть $x_k \in Z_k \subset S$. Выделим из последовательности x_k сходящуюся подпоследовательность и обозначим ее предел через x. В силу замкнутости S, $x \in S$. Для какого-то номера i имеем $x \in O_i$. Но O_i не имеет общих точек с с любым из множеств Z_k, начиная с некоторого номера. Поэтому последовательность x_k не может сходиться к x. Полученное противоречие означает, что из покрытия S множествами O_k можно выделить конечное подпокрытие. $\square$

Лекция 24

ОСНОВНАЯ ЧАСТЬ

24.1 Евклидово пространство

Пусть V — вещественное линейное пространство, на котором каждой упорядоченной паре векторов $x, y \in V$ поставлено в соответствие вещественное число (x, y) таким образом, что:

(1) $(x, x) \geq 0 \ \ \forall \ \ x \in V; \ \ (x, x) = 0 \ \ \Leftrightarrow \ \ x = 0;$

(2) $(x, y) = (y, x) \ \ \forall \ \ x, y \in V;$

(3) $(x + y, z) = (x, z) + (y, z) \ \ \forall \ \ x, y, z \in V;$

(4) $(\alpha x, y) = \alpha(x, y) \ \ \forall \ \ \alpha \in \mathbb{R}, \ \ \forall \ \ x \in V.$

Число (x, y) называется *скалярным произведением* векторов x и y. Вещественное линейное пространство со скалярным произведением называется *евклидовым*.

В $\mathbb{R}^n$ скалярное произведение векторов $x = [x_1, \ldots, x_n]^\top$, $y = [y_1, \ldots, y_n]^\top$ часто вводится как сумма парных произведений координат:

$$(x, y) \ = \ \sum_{i=1}^{n} x_i y_i \ = \ y^\top x. \qquad (*)$$

Оно называется *естественным* скалярным произведением на $\mathbb{R}^n$. Но на $\mathbb{R}^n$ скалярное произведение можно ввести и многими другими способами: например, если фиксировать числа $\lambda_1, \ldots, \lambda_n > 0$, то выражение

$$(x, y) \ = \ \sum_{i=1}^{n} \lambda_i x_i y_i \ = \ y^\top \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} x$$

обладает свойствами (1)–(4) и, следовательно, задает скалярное произведение.

24.2 Унитарное пространство

В $\mathbb{C}^n$ выражение $(*)$, очевидно, уже не является скалярным произведением: пусть $n = 2$ и $x = [1, \mathbf{i}]^\top$, тогда $(x, x) = 1^2 + \mathbf{i}^2 = 0$. Вообще в ненулевом комплексном пространстве аксиомы (1), (4) не совместимы с аксиомой (2): $(\mathbf{i}x, \mathbf{i}x) = -(x, x) \ \ \Rightarrow \ \ $ если $(x, x) > 0$, то $(\mathbf{i}x, \mathbf{i}x) < 0$.

Пусть V — комплексное линейное пространство. Теперь при определении скалярного произведения (x, y) предполагается, что число (x, y) в общем случае комплексное, а набор аксиом модифицируется таким образом:

$(1')$ $(x,x) \geq 0 \ \ \forall \ \ x \in V; \ \ (x,x) = 0 \ \ \Leftrightarrow \ \ x = 0;$

$(2')$ $(x,y) = \overline{(y,x)} \ \ \forall \ \ x,y \in V$ (черта означает комплексное сопряжение);

$(3')$ $(x+y,z) = (x,z) + (y,z) \ \ \forall \ \ x,y,z \in V;$

$(4')$ $(\alpha x, y) = \alpha(x,y) \ \ \forall \ \ \alpha \in \mathbb{C}, \ \ \forall \ \ x \in V.$

Комплексное линейное пространство со скалярным произведением называется *унитарным*.

Аксиомы евклидова и унитарного пространств отличаются лишь комплекным сопряжением во второй аксиоме и, конечно, тем, что в вещественном пространстве все числа и само скалярное произведение вещественны. Заметим, что то в любом случае *скалярный квадрат* (x,x) обязан быть неотрицательным вещественным числом.

В отличие от $(*)$, в $\mathbb{C}^n$ *естественное* скалярное произведение векторов $x = [x_1, \ldots, x_n]^\top$, $y = [y_1, \ldots, y_n]^\top$ вводится так:

$$(x,y) \ = \ \sum_{i=1}^{n} x_i \overline{y_i} \ = \ y^* x.$$

24.3 Билинейные и полуторалинейные формы

В аксиомах скалярного произведения свойства (3), (4) отражают линейность функции (x,y) от векторов x и y по первому аргументу. В евклидовом пространстве аксиома (2) дает нам линейность и по второму аргументу.

Функция $f(x,y)$ называется *билинейной формой*, если она линейна по каждому из аргументов:

$$(\alpha x + \beta y, z) = \alpha(x,z) + \beta(y,z), \quad (z, \alpha x + \beta y) = \alpha(z,x) + \beta(z,y) \ \ \forall \ \ x,y,z \in V, \ \ \forall \ \ \alpha, \beta.$$

Таким образом, скалярное произведение в евклидовом пространстве является билинейной формой с дополнительными условиями (1) и (2).

Функция $f(x,y)$ называется *полуторалинейной формой*, если

$$(\alpha x + \beta y, z) = \alpha(x,z) + \beta(y,z), \quad (z, \alpha x + \beta y) = \overline{\alpha}(z,x) + \overline{\beta}(z,y) \ \ \forall \ \ x,y,z \in V, \ \ \forall \ \ \alpha, \beta \in \mathbb{C}.$$

Очевидно, скалярное произведение в унитарном пространстве является полуторалинейной формой с дополнительными условиями $(1')$ и $(2')$.

24.4 Длина вектора

Пусть V — произвольное пространство со скалярным произведением. Величина

$$|x| = \sqrt{(x,x)}$$

называется *длиной* вектора $x \in V$.

Неравенство Коши–Буняковского–Шварца. *Для любых векторов* $x, y \in V$

$$|(x,y)| \ \leq \ |x| \, |y|, \qquad (*)$$

причем равенство достигается в том и только том случае, когда x и y линейно зависимы.

Доказательство. Комплексное число (x, y) запишем в тригонометрической форме

$$(x, y) = |(x, y)|\,\xi, \quad \xi \;=\; \cos\phi + \mathbf{i}\sin\phi.$$

Если $y = 0$, то в $(*)$ имеет место равенство. Пусть $y \neq 0$. Для произвольного $t \in \mathbb{R}$ рассмотрим выражение

$$(x + t\xi y,\, x + t\xi y) \;=\; (x, x) + t\xi\overline{(x, y)} + t\overline{\xi}(x, y) + \xi\overline{\xi}(y, y) \;=\; t^2|y|^2 + 2t|(x, y)| + |x|^2 \;\geq\; 0.$$

Неотрицательность квадратного трехчлена от переменной t означает неположительность его дискриминанта:

$$D = |(x, y)|^2 - |x|^2|y|^2 \;\leq\; 0 \quad\Rightarrow\quad |(x, y)| \;\leq\; |x|\,|y|.$$

Предположим, что при $y \neq 0$ в $(*)$ имеет место равенство $\Rightarrow$ $D = 0$ $\Rightarrow$ для некоторого вещественного t получаем

$$(x + t\xi y,\, x + t\xi y) \;=\; 0 \quad\Rightarrow\quad x + t\xi y = 0.$$

Очевидно также, что если $y = 0$ или $x = \alpha y$, то $(*)$ обращается в равенство. $\square$

Следствие. *Длина является векторной нормой на V.*

Доказательство. Первые два свойства нормы очевидны, а неравенство треугольника вытекает из неравенства Коши-Буняковского-Шварца:

$$|x + y|^2 \;=\; |x|^2 + |y|^2 + (x, y) + (y, x) \;\leq\; |x|^2 + |y|^2 + 2|x|\,|y| \;=\; (|x| + |y|)^2. \qquad \square$$

Пространство со скалярным произведением, полное относительно нормы $||\cdot|| = |\cdot|$, обычно называется *гильбертовым*.

24.5 Тождество параллелограмма

Итак, любое пространство со скалярным произведением обладает специальной нормой, порожденной скалярным произведением. Зададим вопрос: какие нормы на V могут порождаться каким-нибудь скалярным произведением?

Ответ связан со следующим *тождеством параллелограмма*:

$$||x + y||^2 + ||x - y||^2 \;=\; 2(||x||^2 + ||y||^2) \quad \forall\;\; x, y \in V.$$

Легко проверить, что длина вектора $||x|| = |x|$ (то есть, норма, порожденная скалярным произведением) удовлетворяет данному тождеству. Но верно и обратное.

Теорема. *Норма $||\cdot||$ порождается каким-то скалярным произведением в том и только том случае, когда для нее выполняется тождество параллограмма.*

Доказательство. Пусть V — пространство со скалярным произведением. Запишем $(x, y) = a + \mathbf{i}b$, где $a, b \in \mathbb{R}$. Тогда если $||x|| = |x|$, то

$$\begin{array}{rcccl} ||x + y||^2 & = & (x, x) + (x, y) + (y, z) + (y, y) & = & ||x||^2 + ||y||^2 + 2a, \\ ||x + \mathbf{i}y||^2 & = & (x, x) + \mathbf{i}(y, x) - \mathbf{i}(x, y) + (y, y) & = & ||x||^2 + ||\mathbf{i}y||^2 + 2b. \end{array}$$

Отсюда $a = f(x,y)$ и $b = g(x,y)$, где

$$f(x,y) = \frac{1}{2}(||x+y||^2 - ||x||^2 - ||y||^2), \qquad g(x,y) = \frac{1}{2}(||x+\mathbf{i}y||^2 - ||x||^2 - ||\mathbf{i}y||^2).$$

Теперь предположим, что V — нормированное пространство. Если норма порождается скалярным произведением, то последнее обязано иметь вид

$$(x,y) \;=\; f(x,y) + \mathbf{i}g(x,y). \qquad (*)$$

Рассмотрим $(*)$ как определение функции (x,y) и докажем, что она обладает всеми свойствами скалярного произведения.

Легко видеть, что $(x,x) = ||x||^2$. Поэтому первая аксиома очевидна. Так же легко проверяется, что $(x,y) = \overline{(y,x)}$.

Теперь предположим, что норма удовлетворяет тождеству параллелограмма и докажем, что функция (x,y) линейна по первому аргументу (третья и четвертая аксиомы). Для этого достаточно доказать линейность по первому аргументу функции $f(x,y)$ (линейность $g(x,y)$ по первому аргументу будет очевидным следствием).

Докажем сначала, что $f(x+y,z) = f(x,z) + f(y,z)$. Из определения f и тождества параллелограмма видно, что

$$f(x,z) = \frac{1}{4}(||x+z||^2 - ||x-z||^2), \qquad f(y,z) = \frac{1}{4}(||y+z||^2 - ||y-z||^2).$$

Запишем $x+z = u+v$, $y+z = u-v \;\Rightarrow\; u = \frac{1}{2}(x+y+2z)$, $v = \frac{1}{2}(x-y)$. В силу тождества параллелограмма для векторов u и v,

$$||x+z||^2 + ||y+z||^2 \;=\; \frac{1}{2}(||(x+y+z)+z||^2 + \frac{1}{2}||x-y||^2.$$

Аналогично,

$$||x-z||^2 + ||y-z||^2 \;=\; \frac{1}{2}(||(x+y-z)-z||^2 + \frac{1}{2}||x-y||^2.$$

По тому же тождеству параллелограмма для $x+y+z$ и z,

$$\frac{1}{2}(||(x+y+z)+z||^2 \;=\; ||x+y+z||^2 + ||z||^2 - \frac{1}{2}||x+y||^2,$$

$$\frac{1}{2}(||(x+y-z)-z||^2 \;=\; ||x+y-z||^2 + ||z||^2 - \frac{1}{2}||x+y||^2.$$

Отсюда

$$f(x,z) + f(y,z) = \frac{1}{4}(||x+y+z||^2 - ||x+y-z||^2) = f(x+y,z).$$

Теперь докажем, что $f(\alpha x, y) = \alpha\, f(x,y)$ для любого $\alpha \in \mathbb{R}$. Пусть $\alpha = \frac{m}{n}$ — рациональное число. Тогда, пользуясь уже доказанным свойством, находим

$$_n f(\tfrac{1}{n}x,\, y) \;=\; f(_n(\tfrac{1}{n}x),\, y) \;=\; f(x,y) \;\;\Rightarrow\;\; f(\tfrac{1}{n}x,\, y) \;=\; \tfrac{1}{n}f(x,y) \;\;\Rightarrow$$

$$f(\tfrac{m}{n}x,\, y) = f(_m(\tfrac{1}{n}x,\, y) = {}_m f(\tfrac{1}{n}x,\, y) = \tfrac{m}{n}f(x,y).$$

Произвольное вещественное α представим как предел последовательности рациональных $\alpha_k \to \alpha$. Несложно убедиться в том, что функция $f(x,y)$ непрерывна по x. Поэтому в равенствах $f(\alpha_k x, y) = \alpha_k f(x,y)$ можно перейти к пределу при $k \to \infty$.

Таким образом, мы доказали равенство $(\alpha x, y) = (\alpha x, y)$ пока только для вещественных α. Оно будет верно для любых комплексных α, если мы установим, что $(\mathbf{i}x, y) = \mathbf{i}(x,y)$. Это вытекает непосредственно из определения $(*)$ и вида функций $f(x,y)$ и $g(x,y)$. $\square$

24.6 Ортогональность векторов

Скалярное произведение позволяет ввести общее понятие ортогональности векторов: x и y называются *ортогональными*, если $(x, y) = 0$. Обозначение: $x \perp y$.

Заметим, что в одном и том же пространстве скалярное произведение можно ввести многими разными способами, и векторы, ортогональные для какого-то скалярного произведения, могут быть не ортогональными по отношению к другому скалярному произведению.

В евклидовом пространстве можно ввести также общее понятие *угла* $\phi = \phi(x, y)$ между векторами x и y. По определению,

$$\cos\phi \;=\; \frac{(x,y)}{|x|\,|y|}.$$

Нужно заметить, что правая часть по модулю не больше 1 (в силу неравенства Коши-Буняковского-Шварца). Для ортогональных векторов $\phi = \pi/2$. По понятной причине данное определение угла *не переносится* на случай унитарных пространств. Но понятие ортогональности работает, конечно, и там.

Теорема Пифагора. *Если* $x \perp y$, *то* $|x + y|^2 = |x|^2 + |y|^2$.

Доказательство. Пусть $(x, y) = 0$. Тогда $(x+y, x+y) = (x, x)+(x, y)+(y, x)+(y, y) = (x, x) + (y, y)$. $\square$

Замечание. В евклидовом (но не в унитарном!) пространстве теорему Пифагора можно обратить: если $|x + y|^2 = |x|^2 + |y|^2$, то $(x, y) = 0$ — это очевидно, поскольку $(x, y) + (y, x) = 2(x, y)$. Однако, последнее не верно для произвольных векторов в унитарном пространстве.

24.7 Ортогональность множеств

Пусть V — пространство со скалярным произведением и $L \subset V$ — произвольное непустое подмножество векторов. Вектор $x \in V$ называется *ортогональным* множеству L, если $(x, y) = 0$ для всех $y \in L$. Обозначение: $x \perp L$. По определению, множества L и M *ортогональны*, если $(x, y) = 0$ для любых $x \in L$ и $y \in M$. Обозначение: $L \perp M$.

Множество M всех векторов из V, каждый из которых ортогонален заданному множеству L, называется его *ортогональным дополнением* в пространстве V. Обозначение: $M = L^\perp$.

Утверждение. *Для любого множества* L *его ортогональное дополнение* $L^\perp$ *является подпространством. При этом* $L \subset (L^\perp)^\perp$.

Доказательство. Пусть $x, y \in L^\perp$. Тогда $(x, z) = (y, z) = 0 \;\forall\; z \in L \;\Rightarrow\; (\alpha x+\beta y, z) = \alpha(x, y) + \beta(y, z) = 0 \;\forall\; z \in L \;\Rightarrow\; \alpha x + \beta y \in L^\top$.

По определению, множество $(L^\perp)^\perp$ содержит все векторы, ортогональные $L^\top$, а значит, и все векторы из множества L. $\square$

24.8 Ортогональная сумма подпространств

Напомним, что *суммой подпространств* $L_1, L_2, \ldots, L_m$ называется множество L всех векторов вида $x = x_1 + x_2 + \ldots + x_m$, где $x_i \in L_i$ для всех i. Элементарно проверяется, что L — подпространство. Обозначение: $L = L_1 + \ldots + L_m$.

Напомним также, что L называется *прямой суммой*, если подпространства L_i ненулевые и каждый вектор $x \in L$ имеет единственное разложение вида $x = x_1 + \ldots + x_m$, где $x_i \in L_i$ (если $x = x'_1 + \ldots + x'_m$ и $x'_i \in L_i \ \ \forall\, i$, то непременно $x'_i = x_i \ \ \forall\, i$).

Сумма $L = L_1 + \ldots + L_m$ ненулевых подпространств называется *ортогональной суммой*, если $L_i \perp L_j$ при $i \neq j$. Обозначение: $L = L_1 \oplus \ldots \oplus L_m$.

Утверждение. *Ортогональная сумма подпространств $L = L_1 \oplus \ldots \oplus L_m$ является прямой суммой. Кроме того, если $x_i \in L_i$, то*

$$|x_1 + \ldots + x_m|^2 \ = \ |x_1|^2 + \ldots + |x_m|^2. \qquad (*)$$

Доказательство. Докажем сначала $(*)$. Учитывая, что $(x_i, x_j) = 0$ при $i \neq j$, находим

$$|x_1 + \ldots + x_m|^2 \ = \ \sum_{i=1}^{m}\sum_{j=1}^{m}(x_i, x_j) \ = \ \sum_{i=1}^{m}(x_i, x_i) = \sum_{i=1}^{m}|x_i|^2.$$

Далее, пусть $x_1 + \ldots + x_m = x'_1 + \ldots + x'_m$, где $x_i, x'_i \in L_i \ \ \forall\, i$. Тогда

$$0 = |(x_1 - x'_1) + \ldots + (x_m - x'_m)|^2 = |x_1 - x'_1|^2 + \ldots + |x_m - x'_m|^2 \ \ \Rightarrow \ \ x_i = x'_i \ \ \forall \ \ i. \qquad \square$$

Следствие 1. *Конечная система ненулевых попарно ортогональных векторов является линейно независимой.*

Доказательство. Пусть векторы $x_1, \ldots, x_m$ попарно ортогональны и отличны от нуля. Тогда сумма линейных оболочек $L(x_1), \ldots, L(x_m)$ является ортогональной суммой, и если $\alpha_1 x_1 + \ldots + \alpha_m x_m = 0$, то, согласно $(*)$,

$$0 = |\alpha_1 x_1 + \ldots + \alpha_m x_m|^2 \ = \ |\alpha_1|^2|x_1|^2 + \ldots + |\alpha_m|^2|x_m|^2 \ \ \Rightarrow \ \ \alpha_1 = \ldots = \alpha_m = 0. \qquad \square$$

Следствие 2. *Если ненулевые подпространства $L_1, \ldots, L_m$ конечномерны и попарно ортогональны, то*

$$\dim(L_1 \oplus \ldots \oplus L_m) = \dim L_1 + \ldots + \dim L_m.$$

Достаточно вспомнить, что для прямой суммы конечномерных подпространств L_i базис получается объединением базисов в подпространствах L_i (см. Лекцию 12).

Лекция 25

ОСНОВНАЯ ЧАСТЬ

25.1 Матрица Грама

Пусть дана система векторов $v_1, \ldots, v_n$ и пусть

$$x = \alpha_1 v_1 + \ldots + \alpha_n v_n, \quad y = \beta_1 v_1 + \ldots + \beta_n v_n.$$

Тогда прямое вычисление дает

$$(x, y) = \sum_{j=1}^{n} \overline{\beta}_j \left(\sum_{i=1}^{n} (v_j, v_i)\alpha_j \right) = b^* G a, \qquad (*)$$

где

$$G = G(v_1, \ldots, v_n) = \begin{bmatrix} (v_1, v_1) & \ldots & (v_n, v_1) \\ \ldots & \ldots & \ldots \\ (v_1, v_n) & \ldots & (v_n, v_n) \end{bmatrix}, \quad a = \begin{bmatrix} \alpha_1 \\ \ldots \\ \alpha_n \end{bmatrix}, \quad b = \begin{bmatrix} \beta_1 \\ \ldots \\ \beta_n \end{bmatrix}.$$

Матрица G из скалярных произведений системы векторов называется ее *матрицей Грама.* [1]

Теорема о матрице Грама. *Система векторов $v_1, \ldots, v_n$ линейно зависима тогда и только тогда, когда ее матрица Грама вырожденная.*

Доказательство. Пусть $x = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Используя $(*)$ при $x = y$, находим

$$(x, x) = a^* G a, \quad a = [\alpha_1, \ldots, \alpha_n]^\top. \qquad (\#)$$

Если G — вырожденная матрица, то существует столбец $a \neq 0$ такой, что $Ga = 0 \quad \Rightarrow$ $x = 0 \quad \Rightarrow$ система векторов $v_1, \ldots, v_n$ линейно зависима.

Обратно, если эта система линейно зависима, то $x = 0$ при некотором $a \neq 0$. Легко видеть, что $Ga = [(x, v_1), \ldots, (x, v_n)]^\top = 0$ есть равная нулю нетривиальная линейная комбинация столбцов матрицы $G \quad \Rightarrow$ столбцы G линейно зависимы $\quad \Rightarrow \quad G$ вырожденная. $\square$

25.2 Скалярное произведение в конечномерном пространстве

Пусть $v_1, \ldots, v_n$ — базис в n-мерном пространстве V. Тогда скалярное произведение векторов $x = \alpha_1 v_1 + \ldots + \alpha_n v_n$ и $y = \beta_1 v_1 + \ldots + \beta_n v_n$ имеет вид $(*)$, где G — матрица Грама, $a = [\alpha_1, \ldots, \alpha_n]^\top$, $b = [\beta_1, \ldots, \beta_n]^\top$.

[1] Обратим внимание на то, что элемент в позиции i, j имеет вид (v_j, v_i). Часто матрицей Грама называют $G^\top$ (в вещественном случае, конечно, $G^\top = G$).

Какими свойствами должна обладать матрица, чтобы являться матрицей Грама для линейно независимой системы?

Во-первых, любая матрица Грама обладает свойством $G^* = G$. Матрицы с таким свойством называются *самосопряженными* или *эрмитовыми*. [2] В вещественном случае $G^* = G^\top$, а матрицы со свойством $G^\top = G$ называются *симметричными*.

Во-вторых, согласно (#), $a^*Ga > 0$ для всех $a \neq 0$, причем если V — вещественное пространство, то $a \in \mathbb{R}^n$, а если комплексное, то $a \in \mathbb{C}^n$. Любая матрица с таким свойством в случае $a \in \mathbb{C}^n$ называется *положительно определенной*. Вещественная матрица с тем же свойством, когда $a \in \mathbb{R}^n$, называется *вещественной положительно определенной*.

Итак, любая матрица Грама в случае унитарного пространства является эрмитовой положительно определенной. Но верно и обратное. Пусть G — произвольная эрмитова положительно определенная матрица. Тогда легко проверяется, что функция

$$f(a,b) \;=\; b^*Ga, \quad a,b \in \mathbb{C}^n, \tag{!}$$

задает скалярное произведение на $\mathbb{C}^n$ и G является матрицей Грама системы стандартных базисных векторов $e_1, \ldots, e_n$ (e_i имеет 1 на i-м месте и 0 в остальных позициях). Таким образом, формула (!) определяет *общий вид скалярного произведения* в пространстве $\mathbb{C}^n$.

Соответствия $a \leftrightarrow x$, $b \leftrightarrow y$ (задающие изоморфизм V и $\mathbb{C}^n$) позволяют с помощью $f(a,b)$ ввести скалярное произведение и на V.

25.3 Перпендикуляр и проекция

Пусть V — пространство со скалярным произведением и L — его подпространство размерности m. Мы уже знаем, что для любого $x \in V$ существует элемент наилучшего приближения $z_0 \in L$ — такой, что $|x - z_0| \leq |x - z|$ для всех $z \in L$. В данном специальном случае — для нормы, порожденной скалярным произведением — имеет место единственность z_0 и есть очень простой способ его получения.

Исходим из того, что в L задан базис $z_1, \ldots, z_m$. Тогда

$$z_0 \;=\; \alpha_1 z_1 + \ldots + \alpha_m z_m.$$

Найдем коэффициенты $\alpha_1, \ldots, \alpha_m$ из условия

$$x - z_0 \perp L \;\;\Leftrightarrow\;\; (x - z_0,\; z_1) = 0,\; \ldots,\; (x - z_0,\; z_m) = 0 \;\;\Leftrightarrow$$

$$\left\{\begin{array}{ccccccc} \alpha_1(z_1,z_1) & + & \ldots & + & \alpha_m(z_m,z_1) & = & (x,z_1), \\ \alpha_1(z_1,z_2) & + & \ldots & + & \alpha_m(z_m,z_2) & = & (x,z_2), \\ \quad\ldots & & & & & & \\ \alpha_1(z_1,z_m) & + & \ldots & + & \alpha_m(z_m,z_m) & = & (x,z_m). \end{array}\right.$$

Очевидно, имеем систему линейных алгебраических уравнений, для которой матрица коэффициентов совпадает с матрицей Грама $G = G(z_1, \ldots, z_m)$ системы векторов $z_1, \ldots, z_m$. По теореме о матрице Грама, для линейно независимой системы она невырожденная $\Rightarrow$ система относительно $\alpha_1, \ldots \alpha_m$ имеет и притом единственное решение

[2] В честь французского математика Шарля Эрмита (1822–1901).

$\Rightarrow$ вектор z_0, подчиненный условию $x - z_0 \perp L$, существует и единствен.

Вектор $h \equiv x - z_0$ в случае $h \perp L$, $z_0 \in L$ называется *перпендикуляром*, опущенным из x на L, а z_0 — *ортогональной проекцией* вектора x на L.

Теорема о перпендикуляре. *Для любого вектора x и конечномерного подпространства L существуют и единственны перпендикуляр $h \perp L$ и проекция $z_0 \in L$ такие, что $x = z_0 + h$. При этом*

$$|h| \;=\; |x - z_0| \;<\; |x - z| \quad \forall \;\; z \in L,\; z \neq z_0.$$

Доказательство. Остается доказать лишь то, что z_0 — однозначно определенный элемент наилучшего приближения на L для вектора x. Пусть z — произвольный вектор из L. Тогда $x - z = (x - z_0) + (z_0 - z)$, где $x - z_0 \perp L$ и $z_0 - z \in L$. Отсюда вытекает, что $x - z_0$ и $z_0 - z$ суть перпендикуляр и ортогональная проекция на L для вектора $x - z$. По теореме Пифагора,

$$|x - z|^2 \;=\; |x - z_0|^2 + |z_0 - z|^2 \quad \Rightarrow \quad |x - z_0| \;<\; |x - z| \;\; \forall\, z \neq z_0. \qquad \square$$

Следствие. *Если L — конечномерное подпространство, то $L \;=\; (L^\perp)^\perp$.*

Доказательство. Мы уже знаем, что $L \subset (L^\perp)^\perp$. Возьмем $x \in (L^\perp)^\perp$ и опустим из него перпендикуляр h на L. Согласно определению ортогонального дополнения, $h \in L^\perp$. В то же время, $h \perp L^\perp \;\; \Rightarrow \;\; (h, h) = 0 \;\; \Rightarrow \;\; h = 0$. Значит, $x \in L \;\; \Rightarrow \;\; (L^\perp)^\perp \subset L$. $\square$

25.4 Ортогональные системы

Система ненулевых векторов $x_1, \ldots, x_n$ называется *ортогональной*, если

$$(x_i, x_j) = 0, \quad i \neq j, \qquad (*)$$

и *ортонормированной*, если, дополнительно, $|x_1| = \ldots = |x_n| = 1$. Таким образом, матрица Грама для ортогональной системы является диагональной с положительными диагональными элементами, а для ортонормированной — единичной матрицей.

Рассмотрим пространство $\mathbb{C}^n$ с естественным скалярным произведением и ортонормированную систему вектор-столбцов

$$x_1 = \begin{bmatrix} x_{11} \\ \cdots \\ x_{n1} \end{bmatrix}, \;\ldots,\; x_n = \begin{bmatrix} x_{1n} \\ \cdots \\ x_{nn} \end{bmatrix} \;\in\; \mathbb{C}^n.$$

Составим из них $n \times n$-матрицу

$$X = [x_1, \ldots, x_n] = \begin{bmatrix} x_{11} & \ldots & x_{1n} \\ \cdots & \cdots & \cdots \\ x_{n1} & \ldots & x_{nn} \end{bmatrix}$$

и заметим, что соотношения $(*)$ равносильны матричному равенству

$$X^* X \;=\; \begin{bmatrix} (x_1, x_1) & \ldots & (x_n, x_1) \\ \cdots & \cdots & \cdots \\ (x_1, x_n) & \ldots & (x_n, x_n) \end{bmatrix} \;=\; I.$$

Матрица $X \in \mathbb{C}^{n \times n}$ со свойством $X^* X = I$ называется *унитарной*. Таким образом, любая квадратная матрица с ортонормированной системой столбцов является унитарной, а любая унитарная матрица имеет ортонормированную систему столбцов. Ясно

также, что для унитарности матрицы необходимо и достаточно, чтобы она имела ортонормированную систему строк (докажите!).

Вещественная унитарная матрица называется *ортогональной*. Ранее мы уже отмечали, что множество всех ортогональных матриц порядка n является группой относительно операции умножения матриц. То же справедливо и по отношению к множеству всех унитарных матриц порядка n.

25.5 Процесс ортогонализации

Из теоремы о перпендикуляре сразу же вытекает, что в любом конечномерном пространстве V существует ортонормированный базис.

В самом деле, возьмем в V произвольный базис $v_1, \ldots, v_n$ и предположим, что в линейной оболочке $L_{n-1} = L(v_1, \ldots, v_{n-1})$ уже построен ортонормированный базис из векторов $q_1, \ldots, q_{n-1}$ (конечно, $L_{n-1} = L(q_1, \ldots, q_{n-1})$). Пусть h_n — перпендикуляр, опущенный из вектора v_n на L_{n-1}. Ясно, что $h_n \neq 0$ (иначе $v_n \in L_{n-1} \;\Rightarrow\;$ система $v_1, \ldots, v_n$ линейно зависима). Положим $q_n = h_n/|h_n|$. Тогда система $q_1, \ldots, q_n$ и будет искомым ортонормированным базисов в V.

Заметим, что в построенном базисе для любого $k = 1, \ldots, n$ первые k векторов $q_1, \ldots, q_k$ образуют ортонормированный базис в линейной оболочке $L_k = L(v_1, \ldots, v_k)$. Таким образом,

$$L(q_1, \ldots, q_k) \;=\; L(v_1, \ldots, v_k), \quad k = 1, \ldots, n.$$

Реальные вычисления начинаются с получения вектора $q_1 = v_1/|v_1|$. Затем из вектора v_2 опускается на L_1 перпендикуляр h_2 и нормируется: $q_2 = h_2/|h_2|$. И так далее. Опуская перпендикуляр на L_k, разумно искать разложение соответствующей проекции не по исходной системе $v_1, \ldots, v_k$, а по уже построенной ортонормированной системе $q_1, \ldots, q_k$. Выгода очевидна: матрица Грама для $q_1, \ldots, q_k$ является единичной!

Данный алгоритм называется *процессом ортогонализации Грама–Шмидта*. Вот его формальное описание:

$$h_k \;=\; v_k - \sum_{i=1}^{k-1} (v_k, q_i)\, q_i, \quad q_k = h_k/|h_k|, \qquad k = 1, \ldots, n.$$

25.6 Дополнение до ортогонального базиса

Пусть V — пространство размерности n со скалярным произведением.

Лемма о дополнении до ортогонального базиса. *Любая ортогональная (ортонормированная) система векторов $v_1, \ldots, v_k \in V$ может быть достроена какими-то векторами из V до ортогонального (ортонормированного) базиса в V.*

Доказательство. Дополним $v_1, \ldots, v_k$ какими-нибудь векторами до базиса в V, а затем к полученному базису применим процесс ортогонализации. $\square$

Следствие. *Если L_k — подпространство размерности k, то $\dim L_k^\perp = n - k$. При этом*

$$V \;=\; L_k \oplus L_k^\perp.$$

Доказательство. В V существует ортонормированный базис $q_1, \ldots, q_n$ такой, что $L_k = L(q_1, \ldots, q_k)$. При этом очевидно, что любой вектор, ортогональный L_k, есть линейная комбинация векторов $q_{k+1}, \ldots, q_n$. $\square$

25.7 Биортогональные системы

Пусть V — линейное пространство со скалярным произведением $(\cdot\,,\cdot)$ и $L \subset V$ — подпространство размерности m.

Системы векторов $u_1, \ldots, u_m \in L$ и $v_1, \ldots, v_m \in L$ называются *биортогональными*, если

$$(u_i, v_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Говорят также, что каждая из систем является биортогональной для другой системы.

Утверждение 1. *В случае биортогональности каждая из систем $u_1, \ldots, u_m$ и $v_1, \ldots, v_m$ является линейно независимой.*

Доказательство. Пусть $z \equiv \alpha_1 u_1 + \ldots + \alpha_m u_m = 0$. Используя биортогональность, находим $(z, v_i) = \alpha_i = 0$. $\square$

Утверждение 2. *Для любой линейно независимой системы $u_1, \ldots, u_m \in L$, $\dim L = m$, существует единственная биортогональная система $v_1, \ldots, v_m \in L$.*

Доказательство. Построим в L ортонормированный базис $p_1, \ldots, p_m$, рассмотрим разложения

$$u_i = \sum_{k=1}^{m} a_{ik} p_k, \quad i = 1, \ldots, m,$$

и образуем $m \times m$-матрицу $A = [a_{ij}]$. Столбцы A линейно независимы $\Rightarrow$ матрица A обратима. Будем искать $v_1, \ldots, v_m$ в виде

$$v_j = \sum_{k=1}^{m} b_{kj} p_k, \quad j = 1, \ldots, m. \qquad (*)$$

В случае биортогональности получаем

$$(u_i, v_j) = \sum_{k=1}^{m} a_{ik} \bar{b}_{kj} \quad \Rightarrow \quad B \equiv [\bar{b}_{kj}] = A^{-1},$$

что доказывает единственность системы $v_1, \ldots, v_m$. Чтобы доказать существование, положим $b_{kj} = \overline{(A^{-1})}_{kj}$ и определим v_j формулой $(*)$. $\square$

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

25.8 QR-разложение матрицы

Пусть $A \in \mathbb{C}^{n \times m}$ имеет линейно независимые столбцы $a_1, \ldots, a_m \in \mathbb{C}^n$ и к ним применяется процесс ортогонализации Грама–Шмидта с использованием естественного скалярного произведения.. Пусть в результате получаются ортонормированные векторы $q_1, \ldots, q_m \in \mathbb{C}^m$.

Соотношения $a_k \in L(q_1, \ldots, q_k)$ выполняются при $k = 1, \ldots, m$ и означают, что для каких-то чисел r_{ik} имеют место равенства

$$a_k \;=\; \sum_{i=1}^{k} r_{ik}\, q_i, \quad k = 1, \ldots, m,$$

или, в матричном виде,

$$A \;=\; QR, \quad Q = [q_1, \ldots, q_m], \quad R = \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1m} \\ & r_{22} & \ldots & r_{2m} \\ & & \ddots & \ldots \\ & & & r_{mm} \end{bmatrix}.$$

Определение. Разложение $A = QR$, где Q имеет ортонормированные столбцы, а R — верхняя треугольная матрица, называется *QR-разложением* матрицы A.

Таким образом, мы только что доказали, что *для любой прямоугольной матрицы с линейно независимыми столбцами существует QR-разложение.* В частности, оно существует для любой невырожденной матрицы. В действительности справедлива более общая

Теорема. *Любая прямоугольная матрица, в которой число строк не меньше числа столбцов, обладает QR-разложением с верхней ступенчатой матрицей R.*

Доказательство. Пусть a_{i_1} — первый ненулевой столбец матрицы A, a_{i_2} — первый столбец такой, что $a_{i_2} \notin L(a_{i_1})$, a_{i_3} — первый столбец такой, что $a_{i_3} \notin L(a_{i_1}, a_{i_2})$, и так далее. В итоге получаем в A базисную систему столбцов

$$a_{i_1}, \ldots, a_{i_r}, \quad i_1 < i_2 < \ldots < i_r,$$

обладающую такими свойствами:

$a_j = 0$ при $j < i_1$;

$a_j \in L(a_{i_1}, \ldots, a_{i_l})$ при $i_l < j < i_{l+1}, \quad l = 1, \ldots, r-1$;

$a_j \in L(a_{i_1}, \ldots, a_{i_r})$ при $i_r < j$.

Найдем QR-разложение

$$[a_{i_1}, \ldots, a_{i_r}] = [q_{i_1}, \ldots, q_{i_r}]R_r.$$

Систему столбцов $q_{i_1}, \ldots, q_{i_r}$ дополним до ортонормированного базиса в n-мерном пространстве столбцов и из полученных столбцов составим матрицу Q, сохранив первоначальные столбцы в позициях $i_1, \ldots, i_r$.

Записав $A = QR$, видим, что в матрице R первые r элементов i_l-го столбца те же, что в l-м столбце матрицы R_r. В то же время, j-й столбец при $i_l < j < i_{l+1}$ имеет нули в позициях ниже i_l-й. $\square$

Задача. Пусть $A \in \mathbb{C}^{n\times n}$ имеет столбцы $a_1, \ldots, a_n \in C^n$. Докажите неравенство

$$|\det A| \;\le\; \prod_{j=1}^{n} ||a_j||_2.$$

25.9 Потеря ортогональности при вычислениях

Попробуйте реализовать процесс ортогонализации Грама–Шмидта на компьютере. По завершении вычислений законно желание проверить, "насколько ортогональными" будут *вычисленные векторы* $\widetilde{q}_1, \ldots, \widetilde{q}_n \in \mathbb{C}^n$.

В силу ошибок округления они, конечно, отличаются от точных ортонормированных векторов $q_1, \ldots, q_n$. Однако, проверка может Вас и удивить: в большом числе случаев скалярные произведения вычисленных векторов $(\widetilde{q}_i, \widetilde{q}_j)$ при $i \neq j$ совсем не похожи на нули.

Причину понять нетрудно. Допустим, что все хорошо на первых k шагах:

$$(\widetilde{q}_i, \widetilde{q}_i) \approx 1, \quad (\widetilde{q}_i, \widetilde{q}_j) \approx 0, \quad i \neq j, \qquad 1 \leq i, j \leq k.$$

Далее, пусть вычисленный перпендикуляр $\widetilde{h}_{k+1}$ таков, что

$$(\widetilde{h}_{k+1}, \widetilde{q}_i) \;=\; \varepsilon, \qquad \varepsilon \approx 0. \tag{$*$}$$

После нормировки, тем не менее,

$$(\widetilde{q}_{k+1}, \widetilde{q}_i) \;=\; (\widetilde{h}_{k+1}, \widetilde{q}_i)/|\widetilde{h}_{k+1}| \;=\; \varepsilon/|\widetilde{h}_{k+1}|.$$

Отсюда видно, что ортогональность утрачивается при достаточно малой длине приближенного перпендикуляра $\widetilde{h}_{k+1}$. Последнее означает, что вектор a_{k+1} близок к линейной комбинации векторов $a_1, \ldots, a_k$.

Что же делать? Хороший рецепт — "задержаться" на $k+1$-м шаге и повторить p раз вычисления $k+1$-го шага с заменой a_{k+1} на $\widetilde{h}_{k+1}$. Это так называемая процедура p-кратной *реортогонализации.*

В результате величина ε в соотношениях типа $(*)$ уменьшается и может быть сделана сколь угодно малой. Действительно, пусть $Q_k = [q_1, \ldots, q_k]$, $\widetilde{Q}_k = [\widetilde{q}_1, \ldots, \widetilde{q}_k]$. Тогда

$$h_{k+1} = a_{k+1} - \sum_{i=1}^{k} q_i\,(q_i^* a_{k+1}) = a_{k+1} - \sum_{i=1}^{k} (q_i\,q_i^*) a_{k+1} = (I - Q_k Q_k^*) a_{k+1} \quad \Rightarrow$$

$$\begin{bmatrix} (h_{k+1}, q_1) \\ (h_{k+1}, q_2) \\ \cdots \\ (h_{k+1}, q_k) \end{bmatrix} = Q_k^* h_{k+1} = Q_k^* (I - Q_k Q_k^*) a_{k+1} = (I_k - Q_k^* Q_k)\, Q_k^* a_{k+1}.$$

Здесь I — единичная матрица порядка n, а I_k — единичная матрица порядка k. Если погрешности имели место только при вычислении первых k векторов, а на $k+1$-м шаге их не было, то для вычисленных векторов получаем

$$\widetilde{Q}_k^* \widetilde{h}_{k+1} = (I_k - \widetilde{Q}_k^* \widetilde{Q})\, \widetilde{Q}_k^* a_{k+1}.$$

Пусть реортогонализация повторяется p раз без нормировки перпендикуляра и в результате получается вектор $h^{(p)}$. Тогда

$$\widetilde{Q}_k^* h^{(p)} = (I_k - \widetilde{Q}_k^* \widetilde{Q})^p\, \widetilde{Q}_k^* a_{k+1}.$$

Легко проверить, что в случае достаточно малых элементов матрицы $I_k - \widetilde{Q}_k^* \widetilde{Q}_k$ (то есть, в случае "приемлемой" ортогональности первых k векторов)

$$(I_k - \widetilde{Q}_k^* \widetilde{Q}_k)^p \to 0 \;\text{ при }\; p \to \infty.$$

Поэтому $\widetilde{Q}_k^* h^{(p)} \to 0$ при $p \to \infty$.

Замечательно то, что метод реортогонализации позволяет добиться "хорошей" ортогональности вектора $\widetilde{q}_{k+1}$ к векторам $\widetilde{q}_1, \ldots, \widetilde{q}_k$ даже том случае, когда они сами ортогональны с существенно меньшей точностью.

25.10 Обобщение теоремы о перпендикуляре

Теорему о перпендикуляре можно доказать с помощью совершенно другой техники — менее конструктивной, но работающей также в случае бесконечномерного подпространства L.

Теорема. *Пусть V — гильбертово пространство, а L — его замкнутое подпространство. Тогда для любого вектора x существуют и единственны перпендикуляр $h \perp L$ и проекция $z_0 \in L$ такие, что $x = z_0 + h$. При этом*

$$|h| \;=\; |x - z_0| \;<\; |x - z| \quad \forall \;\; z \in L,\; z \neq z_0.$$

Доказательство. Пусть $x \neq L$ и $\gamma = \inf\limits_{z \in L} |x - z|$ — расстояние между x и L. Рассмотрим последовательность $z_n \in L$ со свойством $\gamma^2 \leq |x - z_n|^2 \leq \gamma^2 + 1/n$. В силу тождества параллелограмма,

$$|x - z_n|^2 + |x - z_m|^2 = 2|\frac{1}{2}(z_m - z_n)|^2 + 2|x - \frac{1}{2}(z_n + z_m)|^2 \quad \Rightarrow \quad |z_n - z_m|^2 \leq 2(1/n + 1/m).$$

Таким образом, последовательность z_n является фундаментальной и, в силу полноты гильбертова пространства, сходится к какому-то вектору $z_0 \in V$. Из замкнутости L вытекает, что $z_0 \in L$.

Докажем теперь, что $y = x - z_0 \perp L$.

Возьмем любой вектор $z \in L$ и запишем $(y, z) = a + ib,\; a, b \in R$. Если $a \neq 0$, то пусть $\tau = a/|a|$. Для любого $\varepsilon > 0$ находим

$$\gamma^2 \leq |x - z_0 - \varepsilon\tau z|^2 = |y - \varepsilon\tau z|^2 \leq |y|^2 - 2\varepsilon|a| + \varepsilon^2|z|^2 = \gamma^2 - 2\varepsilon|a| + \varepsilon^2|z|^2 \quad \Rightarrow \quad |a| \leq \varepsilon|z|^2/2.$$

В силу произвольности ε должно быть $a = 0$. Аналогичная выкладка (с заменой z на $\mathbf{i}z$) позволяет доказать, что и $b = 0$.

То, что z_0 является для x единственным элементом наилучшего приближения на L, доказывается так же, как в конечномерном случае (с помощью теоремы Пифагора). $\square$

Замечание. Доказано, по существу, что элемент наилучшего приближения z_0 существует для произвольного замкнутого выпуклого множества L. Знание о том, что L — подпространство, требуется лишь для доказательства ортогональности $x - z_0 \perp L$ и единственности z_0.

Лекция 26

ОСНОВНАЯ ЧАСТЬ

26.1 Линейные функционалы

Пусть V — линейное пространство над числовым полем P и $f(x)$ — функция от вектора $x \in V$ с числовыми значениями. Такие функции принято называть *функционалами*. Если выполняется *свойство линейности*

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \quad \forall \ \alpha, \beta \in P, \ \ \forall \ \ x, y \in V,$$

то функция f называется *линейным функционалом*.

Пусть теперь V — нормированное пространство[1]. Линейный функционал называется *ограниченным*, если для некоторой константы $c > 0$

$$|f(x)| \ \leq \ c||x||_V \quad \forall \ \ x \in V. \qquad (*)$$

Утверждение 1. *Для ограниченности линейного функционала необходима и достаточна его непрерывность.*

Доказательство. Если выполняется $(*)$, то из сходимости $||x_k - x||_V \to 0$ при $k \to \infty$ следует, что $|f(x_k) - f(x)| = |f(x_k - x)| \leq c||x_k - x||_V \to 0$.

Если линейный функционал $f(x)$ непрерывен, то покажем, что он ограничен на единичной сфере $S = \{x : ||x||_V = 1\}$. Если это не так, то для какой-то последовательности векторов $x_k \in S$ имеем $|f(x_k)| \to \infty$.

Отсюда $||x_k/|f(x_k)|\,||_V \ \to \ 0 \quad \Rightarrow \quad x_k/|f(x_k)| \ \to \ 0$. В силу непрерывности, $f(x_k/|f(x_k)|) \to f(0) = 0$, что невозможно, так как $|f(x_k/|f(x_k)|)| = |f(x_k)|/|f(x_k)| = 1$.

Итак, $|f(x)| \leq c$ для всех x таких, что $||x||_V = 1$. Следовательно,

$$|f(x/||x||_V)| \leq c \quad \Rightarrow \quad |f(x)| \leq c||x||_V \ \ \forall \ \ x \in V. \qquad \square$$

Замечание. Для линейного функционала непрерывность в какой-то одной точке равносильна непрерывности во всех точках пространства.

Утверждение 2. *Если V конечномерно, то любой линейный функционал на V является ограниченным.*

Доказательство. Пусть $v_1, \ldots, v_n$ — базис в V. Если $x = x_1 v_1 + \ldots + x_n v_n$, то

$$|f(x)| \leq \sum_{i=1}^{n} |x_i| \, |f(v_i)| \leq c \sum_{i=1}^{n} |x_i|, \quad c \equiv \max_{1 \leq i \leq n} |f(v_i)|.$$

[1] Значит, $P = \mathbb{C}$ или $P = \mathbb{R}$.

В конечномерном пространстве из сходимости по норме вытекает покоординатная сходимость. Поэтому если $x^k \to 0$ при $k \to \infty$, то $x_i^k \to 0$. Отсюда $|f(x^k)| \to 0$. Значит, функционал непрерывен при $x = 0$. $\square$

26.2 Сопряженное пространство

Операции сложения и умножения на число для линейных функционалов определяются естественным образом.

Пусть $f(x)$ и $g(x)$ — линейные функционалы на V. Тогда их суммой называется функция $h = f + g : V \to \mathbb{C}$, определенная правилом $h(x) \equiv f(x) + g(x)$. Для $\alpha \in \mathbb{C}$ функция $h = \alpha f : V \to \mathbb{C}$ определяется правилом $h(x) \equiv \alpha f(x)$.

Элементарно проверяется, что $f + g$ и αf остаются линейными функционалами. Таким образом, множество всех всех линейных функционалов на V превращается в линейное пространство.

Особый интерес представляет множество всех ограниченных линейных функционалов. Оно тоже является линейным пространством, поскольку сложение и умножение на число непрерывных функций сохраняют свойство непрерывности.

Линейное пространство всех ограниченных линейных функционалов на V называется *сопряженным пространством* для V. Обозначение: V^*.

Нормой функционала $f \in V^*$ называется величина

$$||f|| \;=\; \sup_{||x||_V=1} |f(x)|.$$

Конечность $||f||$ вытекает из ограниченности f. Аксиомы векторной нормы проверяются очевидным образом.

26.3 Примеры линейных функционалов

(1) Пусть $\mathcal{P}$ — линейное пространство всех вещественных многочленов на отрезке $[-1, 1]$ с C-нормой $||p||_C = \sup\limits_{-1\le x\le 1} |p(x)|$. Пусть $p'(x)$ обозначает производную многочлена $p(x)$ (ясно, что $p' \in \mathcal{P}$). Функционал $f : \mathcal{P} \to \mathbb{R}$, заданный правилом

$$f(p) \equiv p'(1), \quad p \in \mathcal{P},$$

является, очевидно, линейным, но не ограниченным (докажите!).

Задача. Докажите, что функционал $f(p) = p'(0)$ также не будет ограниченным.

(2) В том же пространстве $\mathcal{P}$ функционал $f(p) = p(0)$ является ограниченным линейным функционалом.

(3) Функционал $f(p) = \int\limits_{-1}^{1} p(x)dx$ является линейным и ограниченным на $\mathcal{P}$. [2]

(4) Рассмотрим пространство $\mathbb{C}^n$ с любой нормой, и пусть даны числа $c_1, \ldots, c_n$. Пусть $x = [x_1, \ldots, x_n]^\top \in \mathbb{C}^n$ и $f(x) = c_1x_1 + \ldots + c_nx_n$. Это ограниченный линейный функционал на $\mathbb{C}^n$.

[2] В данном случае достаточно уметь интегрировать лишь многочлены.

26.4 Размерность дополнительного пространства

Множество $L = \{x \in V : f(x) = 0\}$ называется *ядром* или *нуль-пространством* линейного функционала $f : V \to \mathbb{C}$. Обозначение: $L = \ker f$. Легко видеть, что L — подпространство. Если $\dim V = n$ и функционал не равен нулю тождественно, то $\dim L = n - 1$ (докажите!). Мы собираемся доказать, что в бесконечномерном случае конечной (и равной 1) оказывается размерность так называемого дополнительного подпространства.

Подпространство L' в пространстве V называется *дополнительным* для подпространства L, если разложение $V = L + L'$ является прямой суммой. Размерность дополнительного пространства называется *коразмерностью* подпространства L.

Если V конечномерно, то его базис можно получить объединением базисов в L и L'. Поэтому $\dim L' = \dim V - \dim L \quad \Rightarrow$ коразмерность одна и та же для любого дополнительного пространства. То же верно и в бесконечномерном случае.

Скажем, что $a \sim b$, если $a - b \in L$. Это отношение эквивалентности на V. Поэтому V разбивается на множество непересекающихся классов эквивалентности.

Пусть классы $[a]$ и $[b]$ порождены векторами a и b. Естественные определения операций сложения и умножения на число

$$[a] + [b] = [a + b], \quad \alpha[a] = [\alpha a]$$

корректны, так как их результаты не зависят от выбора представителей в классах эквивалентности. Таким образом, множество классов эквивалентности превращается в линейное пространство над тем же полем, что и пространство V. Оно называется *факторпространством*. Обозначение: V/L.

Утверждение. *Любое дополнительное для L подпространство изоморфно факторпространству V/L.*

Доказательство. Для $a \in L'$ пусть $\Phi(a) = [a]$. Очевидно, отображение $\Phi : L' \to V/L$ сохраняет операции и $\Phi(L) = V/L$. Кроме того, если $\Phi(a) = \Phi(b)$, то $a \sim b \quad \Rightarrow$ $a - b \in L$ и одновременно $a - b \in L' \quad \Rightarrow \quad a - b = 0$. Значит, Φ — сохраняющее операции взаимно-однозначное отображение L' на V/L — другими словами, изоморфимзм. $\square$

Следствие. *Для любых двух разложений в прямую сумму $V = L + L' = L + L''$ размерности дополнительных пространств L' и L'' одинаковы.*

26.5 Линейные функционалы и гиперплоскости

Пусть $L = \ker f$. Если $L = V$, то функционал тождественно равен нулю (и поэтому называется нулевым или тривиальным).

Пусть $L \neq V$. Тогда существует вектор x_0, для которого $f(x_0) \neq 0$. Для произвольного вектора $x \in V$ находим

$$f(x - \alpha x_0) = 0 \ \text{ при } \ \alpha = f(x)/f(x_0) \quad \Rightarrow \quad x = z + \alpha x_0, \ \ z \in L.$$

Очевидно, α однозначно определяется условием $z \in L$. Поэтому V есть прямая сумма подпространств L и $L(x_0)$. Таким образом, *ядро нетривиального линейного функционала имеет коразмерность, равную 1.*

Теперь рассмотрим множество $M_c = \{x \in V : f(x) = c\}$. Если $f(x_0) = c$, то, очевидно, $M_c = x_0 + L$. Таким образом, M_c есть линейное многообразие с направляющим пространством L коразмерности 1. В таких случаях линейное многообразие называется

гиперплоскостью. Легко видеть, что отображение $f(x) \mapsto M(f) = \{x \in V : f(x) = 1\}$ является взаимно-однозначным соответствием между линейными функционалами и гиперплоскостями.

Пусть $\dim V = n$ и $e_1, \ldots, e_n$ — базис в V. В данном случае ясно, что любой линейный функционал имеет вид $f(x_1e_1 + \ldots + x_ne_n) = c_1x_1 + \ldots + c_nx_n$, где $c_i = f(e_i)$. Таким образом, любая гиперплоскость в n-мерном пространстве имеет вид

$$c_1x_1 + \ldots + c_nx_n = c, \qquad (*)$$

где $x_1, \ldots, x_n$ — координаты разложения вектора по выбранному базису.

26.6 Опорные гиперплоскости

Уравнение гиперплоскости $(*)$ в $\mathbb{R}^n$ удобно записывать в виде

$$(x, h) = c, \quad \text{где} \quad h = [c_1, \ldots, c_n]^\top.$$

Гиперплоскость, проходящая через точку x_0, задается уравнением $(x, h) = (x_0, h)$. Под скалярным произведением здесь понимается естественное скалярное произведение в $\mathbb{R}^n$.

Пусть $M \subset \mathbb{R}^n$ — некоторое множество. Точка $x_0 \in M$ называется *граничной* для M, если в любой ее окрестности имеются точки $u \in M$ и $v \notin M$. Гиперплоскость $\pi : (x, h) = (x_0, h)$, проходящая через граничную точку $x_0 \in M$, называется *опорной гиперплоскостью* для M, если $(x, h) \le (x_0, h) \ \ \forall \ \ x \in M$.

Лемма о наилучшем приближении на выпуклом множестве. *Пусть $M \subset \mathbb{R}^n$ — замкнутое выпуклое множество. Тогда для любой точки $x \notin M$ существует единственная точка $z_0 \in M$ такая, что*

$$|x - z_0| \ = \ \rho \equiv \inf_{z \in M} |x - z|.$$

При этом $(x - z_0,\, z - z_0) \le 0 \ \ \forall \ \ z \in M$.

Доказательство. Пусть $|x - z_k| \to \rho, \quad z_k \in M$. В силу ограниченности длин $|z_k|$, найдется подпоследовательность $z_{k_l} \to z_0 \in M$. Положим $h = x - z_0$. С помощью предельного перехода получаем $|h| = \rho$. Далее, если $z \in M$ и $v \equiv z - z_0$, то, в силу выпуклости M, $z_0 + \varepsilon v \in M$ для всех $0 \le \varepsilon \le 1$. Следовательно,

$$\rho^2 \le |x - (z_0 + \varepsilon v)|^2 = (h - \varepsilon v, h - \varepsilon v) = \rho^2 - 2\varepsilon(h, v) + \varepsilon^2|v|^2 \quad \Rightarrow$$

$$(h, v) \le \varepsilon|v|^2/2 \quad \forall \ \ 0 < \varepsilon \le 1 \quad \Rightarrow \quad (h, v) \le 0.$$

Если $|x - (z_0 + v)| = \rho$, то $|v|^2 = 2(h, v) \le 0 \quad \Rightarrow \quad v = 0$. $\square$

Теорема. *Через любую граничную точку замкнутого выпуклого множества $M \subset \mathbb{R}^n$ проходит хотя бы одна опорная гиперплоскость для M.*

Доказательство. Любая граничная точка $x_0 \in M$ есть предел некоторой последовательности внешних для M точек: $x_k \to x_0, \ \ x_k \notin M$. В силу леммы, для каждой точки x_k существует элемент наилучшего приближения $z_k \in M$: $|x_k - z_k| \le |x_k - z| \ \ \forall \ \ z \in M$, при этом для любой точки $z \in M$ имеет место неравенство $(p_k, z) \le (p_k, z_0)$, где $p_k = h_k/|h_k|$,

$h_k = x_k - z_k$. Из последовательности векторов p_k выберем подпоследовательность, сходящуюся к некоторому вектору p; очевидно, $|p| = 1$. Тогда для любой точки $z \in M$ выполняется неравенство $(z, p) \le (x_0, p)$. $\square$

Следствие. *Если $x_0 \notin M$, то существует гиперплоскость $(x, h) = (x_0, h)$ такая, что $(x, h) < (x_0, h) \ \ \forall \ \ x \in M$.*

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

26.7 Строение выпуклых множеств

Существование опорных гиперплоскостей позволяет утверждать, что в $\mathbb{R}^n$ любое замкнутое выпуклое множество является пересечением (возможно, бесконечного) числа замкнутых полупространств. Примечателен также следующий факт.

Теорема. *Для любой точки замкнутого ограниченного выпуклого множества $M \subset \mathbb{R}^n$ можно найти конечное число угловых точек множества M и их выпуклую комбинацию, равную данной точке.*

Доказательство. Пусть заданное выпуклое множество содержится в линейном многообразии размерности n. Утверждение очевидно, если $n = 1$. Проведем индукцию по n. Начнем с произвольной граничной точки $x_0 \in M$. Рассмотрим проходящую через нее опорную гиперплоскость $\pi : (x, h) = (x_0, h)$. Пересечение $N = M \cap \pi$ есть замкнутое ограниченное выпуклое множество, принадлежащее линейному многообразию размерности $n - 1$. По индуктивному предположению, любая точка N' будет выпуклой комбинацией его угловых точек.

Можно проверить, что угловые точки N являются также угловыми точками множества M. В самом деле, пусть точка $x \in N$ является внутренней точкой отрезка, соединяющего $a, b \in M, \ \ a \ne b$. Очевидно, a и b должны принадлежать опорной гиперплоскости π. А это означает, что x не является угловой точкой для N.

Далее, пусть x_0 — внутренняя точка множества M. Проведем через нее прямую, пересекающуюся с границей множества M в точках x_1 и x_2. Очевидно, x_0 является выпуклой комбинацией точек x_1 и x_2, а они, в свою очередь, являются выпуклыми комбинациями угловых точек для пересечений M с проходящими через них опорными гиперплоскостями. $\square$

Следствие. *Минимальное значение линейной функции $f(x) = c^\top x = c_1 x_1 + \ldots + c_n x_n$, $c, x \in \mathbb{R}^n$, на замкнутом ограниченном выпуклом множестве $M \subset \mathbb{R}^n$ достигается в некоторой угловой точке.*

Доказательство. Пусть минимальное значение $f(x)$ достигается в точке $x_0 \in M$. Как и любая точка M, x_0 является выпуклой комбинацией конечного числа угловых точек: $x_0 = s_1 x_1 + \ldots_m x_m$, $s_i \ge 0$, $s_1 + \ldots + s_m = 1$. Отсюда

$$f(x_0) = s_1 f(x_1) + \ldots + s_m f(x_m) \ge (s_1 + \ldots + s_m) \min_{1 \le i \le m} f(x_i) \ = \ \min_{1 \le i \le m} f(x_i). \quad \square$$

26.8 Линейные неравенства

Вопросы о системах линейных неравенств являются, по существу, вопросами о свойствах пересечений полупространств $(x, a_k) \le \gamma_k, \ \ 1 \le k \le m$. При этом важно, конечно,

знать, в каких случаях какие-то неравенства являются следствием других неравенств. Основой для ответа на данный вопрос является следующий результат.

Теорема Фаркаша. *Пусть $a, a_1, \ldots, a_m \in \mathbb{R}^n$, и предположим, что неравенство $(x, a) \le 0$ является следствием системы неравенств $(x, a_k) \le 0$, $1 \le k \le m$. Это возможно в том и только том случае, когда $a = s_1 a_1 + \ldots + s_m a_m$ для некоторых неотрицательных чисел $s_1, \ldots, s_m$.*

Доказательство. Если $a = s_1 a_1 + \ldots + s_m a_m$ при $s_i \ge 0$, то неравенство $(x, a) \le 0$ следует из неравенств $(x, a_k) \le 0$ очевидным образом. Рассмотрим множество

$$M = \{v \in \mathbb{R}^n : v = \sum_{k=1}^{m} s_k a_k, \quad s_k \ge 0,\ 1 \le k \le m\}.$$

Это выпуклое и замкнутое множество (докажите!). Поэтому если $a \notin M$, то существует элемент наилучшего приближения $z_0 \in M$: $|a - z_0| \le |a - z| \quad \forall \quad z \in M$. Положим $x_0 = a - z_0$. Тогда $(x_0,\, z - z_0) \le 0 \quad \forall \quad z \in M \quad \Rightarrow \quad (x_0,\, a_k) \le 0,\ 1 \le k \le m$. Кроме того, $(x_0, z) < (x_0, a) \quad \forall \quad z \in M$. Поскольку $0 \in M$, находим $0 < (x_0, a)$. Таким образом, неравенство $(x, a) \le 0$ нарушается для вектора $x = x_0$, который удовлетворяет системе неравенств $(x_0,\, a_k) \le 0, \quad 1 \le k \le m$. $\square$

26.9 Поиск точки в пересечении гиперплоскостей

Гиперплоскость в $\mathbb{C}^n$ — это линейное многообразие размерности $n-1$. Пусть задано m гиперплоскостей

$$\begin{array}{rcl} a_{11}x_1 + \ldots + a_{1n}x_n & = & b_1, \\ & \ldots & \\ a_{m1}x_1 + \ldots + a_{mn}x_n & = & b_m. \end{array} \qquad (*)$$

Обозначим i-ю гиперплоскость через M_i. Очевидно, их пересечение $M = M_1 \cap \ldots \cap M_m$ состоит из векторов $[x_1,\, \ldots,\, x_n]^\top$, удовлетворяющих системе линейных алгебраических уравнений $(*)$. Следовательно, если пересечение m гиперплоскостей не пусто, то оно является линейным многообразием размерности $n - r$, где r — ранг матрицы коэффициентов системы $(*)$.

Пусть в $\mathbb{C}^n$ введено естественное скалярное произведение. Тогда M_i можно задать уравнением $(x, a_i) = b_i$, где $a_i = [\overline{a}_{i1},\, \ldots,\, \overline{a}_{in}]^\top$, а систему $(*)$ записать в виде

$$(x, a_1) = b_1, \quad \ldots, \quad (x, a_m) = b_m.$$

Направляющее подпространство для M_i имеет вид $L_i = \{x : (x, a_i) = 0\} \quad \Rightarrow \quad a_i \perp L_i$.

Предположим, что пересечение гиперплоскостей M не пусто. Ясно, что $M = \widetilde{x} + L$, где $\widetilde{x} \in M$ — частное решение системы $(*)$, а $L = L_1 \cap \ldots \cap L_m$ — линейное подпространство всех решений соответствующей однородной системы.

Для приближенного вычисления частного решения системы $(*)$ попробуем использовать простую геометрическую идею. Возьмем произвольный вектор $x^0 \in \mathbb{C}^n$, найдем ближайший к нему вектор $x^1 \in M_1$, затем ближайший к x^1 вектор $x^2 \in M_2$, и так далее. Получив $x^m \in M_m$, будем повторять те же действия циклически: найдем ближайший к x^m вектор $x^{m+1} \in M_1$, и так далее. [3]

Обозначим через $\hat{x}$ ближайший к x^0 вектор из L, и пусть $z^k \equiv x^k - \hat{x}$. Очевидно,

$$z^k \in L_k, \qquad z_k - z_{k-1} = x^k - x^{k-1} \perp L_k.$$

[3] Метод описан в работе польского математика Качмажа (1937 г.).

Таким образом, $x^k = x^{k-1} + ta_k$, где t определяется условием $(x^k, a_k) = b_k \quad \Rightarrow$

$$x^k \;=\; x^{k-1} + \frac{b_k - (x^{k-1}, a_k)}{(a_k, a_k)} a_k.$$

Будем считать, что

$$M_j = M_{km+j}, \quad L_j = L_{km+j}, \quad a_j = a_{km+j}, \qquad 1 \le j \le m, \quad k = 1, 2 \ldots .$$

Утверждение. *$x^k \to \hat{x}$ при $k \to \infty$.*

Доказательство. По теореме Пифагора, $|z^0|^2 \;=\; |z^k|^2 + \sum\limits_{j=0}^{k-1} |z^{j+1} - z^j|^2$. Отсюда

$$\rho(z^k, L_i) \equiv \inf_{y \in L_i} |z^k - y| \le \sum_{j=1}^{m} |z^{k+j+1} - z^{k+j}| \to 0, \quad k \to \infty.$$

Функция $\rho(v, L_i) = \inf\limits_{y \in l_i} |v - y|$ непрерывна по v: пусть $p_1, p_2 \in L_i$, $v - p_1 \perp L_i$, $w - p_2 \perp L_i$; тогда

$$|\rho(v, L_i) - \rho(w, L_i)| = |\;|v - p_1| - |w - p_2|\;| \le |(v - w) - (p_1 - P_2)| \le |v - w|.$$

Поэтому если $z^k \to z$, то $\rho(z, L_i) = 0$ при $1 \le i \le m$. Следовательно, $z \in L$.

Кроме того, $z^1 - z^0 \perp L_1 \;\Rightarrow\; z^1 - z^0 \perp L$. Поскольку $z^0 \perp L$, находим $z^1 = z^0 + (z^1 - z^0) \perp L$. Аналогично, $z^k \perp L$ для всех k. Значит, $z \perp L$. Таким образом, $z = 0$.

Ясно, что числовая последовательность $|z^k|$ монотонно убывает и поэтому сходится. Последовательнось векторов z^k ограничена и поэтому, в силу конечной размерности пространства, имеет сходящуюся подпоследовательность. Если z — ее предел, то, согласно предыдущему, $z = 0$. Значит, $|z^k| \to 0 \quad \Rightarrow \quad z^k \to 0$. $\square$

26.10 Линейные функционалы и скалярные произведения

Пусть V — произвольное пространство со скалярным произведением. Фиксируем любой вектор $z \in V$ и рассмотрим функцию $f(x) = (x, z)$. Из неравенства Коши–Буняковского–Шварца вытекает, что $f(x)$ — ограниченный линейный функционал на V. Замечательно, что данный пример имеет обший характер.

Теорема Рисса. *Если V — гильбертово пространство, то для любого ограниченного линейного функционала f существует вектор $h \in V$ такой, что*

$$f(x) = (x, h) \quad \forall \;\; x \in V \quad \textit{и} \quad ||f|| = |h|.$$

Доказательство. Рассмотрим в V линейное подпространство

$$L = \{x \in V : f(x) = 0\}$$

и его ортогональное дополнение $M = L^\top$.

Предположим сначала, что в M имеется ненулевой вектор h_0. Тогда $f(h_0) \ne 0$ (иначе $h_0 \in L$ и $h_0 \perp L \;\Rightarrow\; h_0 = 0$). Если h — произвольный вектор из M и $\alpha = f(h)/f(h_0)$, то $z \equiv h - \alpha h_0 \in L$ и одновременно $z \perp L \;\Rightarrow\; z = 0$.

Следовательно, $\dim M = 1$ и любой вектор $x \in V$ допускает единственное разложение $x = \alpha h_0 + z$, где $z \in L$. Положим $\alpha = f(h_0)/|h_0|^2$ и $h = \alpha h_0$. Тогда $f(x) = \alpha f(h_0) = (x, h)$. Кроме того, $|f(x)| \le |h||x|$ и $f(h/|h|) = |h| \quad \Rightarrow \quad ||f|| = |h|$.

Заметим, что полнота гильбертова пространства пока еще не использовалась. Она нужна лишь для того, чтобы рассмотреть особый случай, когда подпространство M нулевое, и доказать, что в этом случае $L = V \quad \Rightarrow \quad f(x) = 0 = (x, 0) \;\; \forall \;\; x \in V$. Для этого достаточно заметить, что L — замкнутое подпространство и воспользоваться обобщением теоремы о перпендикуляре. Таким образом, если $L \ne V$, то $L^\top$ содержит ненулевой вектор. $\square$

26.11 Дуальные нормы

В $\mathbb{C}^n$ любой линейный функционал имеет вид $f(x) = z^*x$, где z — некоторый фиксированный вектор из $\mathbb{C}^n$. В силу взаимно-однозначного соответствия $f \leftrightarrow z$ сопряженное пространство в данном случае естественным образом отождествляется с $\mathbb{C}^n$. Пусть в $\mathbb{C}^n$ задана какая-то векторная норма $||\cdot||$. Тогда норма линейного функционала f может рассматриваться как норма вектора z и, таким образом, является векторной нормой на том же пространстве $\mathbb{C}^n$:

$$||z||' \;=\; \sup_{||x||=1} |z^*x|.$$

Норма $||\cdot||'$ называется дуальной для нормы $||\cdot||$.

Утверждение. *В пространстве $\mathbb{C}^n$ с гельдеровской нормой $||\cdot||_p$, $p \geq 1$, дуальная норма есть $||\cdot||_q$, где $1/p + 1/q = 1$.*

Доказательство. Из неравенства Гельдера $|z^*x| \leq ||z||_q||x||_p$ следует, что $||z||' \leq ||z||_q$. В то же время, равенство легко получается при выборе x вида $x = \alpha z$. $\square$

Интересно отметить, что *норма $||\cdot||''$, дуальная к дуальной норме $||\cdot||'$, совпадает с исходной нормой $||\cdot||$*. По определению,

$$||x||'' = \sup_{||y||'=1} |y^*x| \leq ||x||.$$

Если $n = 1$, то $||x|| = c|x|$ для какого-то $c > 0$. Поэтому

$$||y||' = \sup_{x\neq 0} \frac{|y||x|}{||x||} = \frac{|y|}{c} \quad \Rightarrow \quad ||x||'' = \sup_{y\neq 0} \frac{|y||x|}{||y||'} = c|x| = ||x||.$$

При $n > 1$ рассмотрим произвольный базис $x_1, \ldots, x_n$. Пусть $f_1(x)$ — линейный функционал на одномерном пространстве $L(x_1)$, выбранный таким образом, что $f_1(x_1) = ||x_1||''$. Заметим, что $||f_1|| = ||x_1||''/||x_1||$. Оказывается, f_1 можно продолжить на двумерное пространство $L(x_1, x_2)$ с сохранением нормы: существует линейный функционал $f_2(x)$, $x \in L(x_1, x_2)$, такой, что $f_2(x) = f_1(x)$ при $x \in L(x_1)$ и $||f_2|| = ||f_1||$. Далее, из f_2 можно получить f_3 с более широкой областью определения $L(x_1, x_2, x_3)$ и той же нормой, и так далее. В итоге получается линейный функционал $f_n(x)$, определенный на всем $\mathbb{C}^n$, имеющий норму $||f_n|| = ||x_1||''$ и такой, что $f_n(x_1) = f_1(x_1) = ||x_1||''$. Значит,

$$||x_1||'' = \sup_{f\neq 0} \frac{|f(x_1)|}{||f||} \geq \frac{|f_n(x_1)|}{||f_n||} = ||x_1||.$$

Отсюда $||x_1||'' = ||x_1||$. Остается заметить, что базис можно начинать с любого вектора $x_1 \neq 0$.

Возможность продолжения линейного ограниченного функционала с сохранением нормы в достаточно общем случае — глубокий и не очень простой результат. Он относится к совокупности фактов, которые принято называть теоремами Хана–Банаха.

Теорема Хана–Банаха. *Пусть V — нормированное пространство, L — его подпространство, $w \notin L$ и $\widetilde{L} = L + L(w)$. Тогда любой линейный ограниченный функционал $f : L \to \mathbb{C}$ на L можно продолжить до линейного ограниченного функционала $\widetilde{f} : \widetilde{L} \to \mathbb{C}$ на $\widetilde{L}$ таким образом, что $\widetilde{f}(x) = f(x) \;\; \forall \;\; x \in L$ и при этом $||\widetilde{f}|| = ||f||$.*

Доказательство. Пусть $u \in L$ и $\alpha \in \mathbb{C}$. Тогда $\widetilde{f}(u + \alpha w) = f(u) + \alpha c$, где $c = \widetilde{f}(w)$. Таким образом, $\widetilde{f}$ определяется числом c. Будем считать, что $||f|| = 1$. Ясно, что $||\widetilde{f}|| \geq 1$ при любом выборе c. Поэтому нужно найти такое c, чтобы $|f(u) + \alpha c| \leq ||u + \alpha w||$ при всех $u \in L$ и $\alpha \in \mathbb{C}$.

Рассмотрим сначала более простой случай, когда все пространства и функционалы являются вещественными. [4] Все получается из вполне элементарного наблюдения:

$$f(u) - f(v) \leq ||u - v|| \leq ||u + w|| + ||v + w|| \quad \forall \;\; u, v \in L.$$

Но его нужно правильно проинтерпретировать. Запишем его в виде

$$f(u) - ||u + w|| \;\; \leq \;\; f(v) + ||v + w||,$$

[4] Независимо друг от друга, Хан и Банах рассмотрели именно это случай.

где левая часть зависит только от u, а правая только от v. Поэтому все числа слева и справа разделяются каким-то одним числом:

$$f(u) - ||u+w|| \;\le\; -c \;\le\; f(v) + ||v+w|| \quad \forall \;\; u,v \in L.$$

Теперь уже ясно, $|f(u)+c| \le ||u+w||$ для всех $u \in L$. Для любого вещественного $s \neq 0$ находим $|f(u)+sc| = |s||f(u/s)+c| \le |s|||u/s+w|| = ||u+sw||$. То же верно, конечно, и для $s=0$. Итак, вещественный линейный функционал $f(x)$ можно доопределить с помощью равенства $\widetilde{f}(w) = c$ на более широком пространстве $\widetilde{L}$ таким образом, что

$$|\widetilde{f}(u+sw)| = |f(u)+sc| \le ||u+sw|| \quad \forall \;\; u \in L, \quad \forall \;\; s \in \mathbb{R}.$$

Перейдем к общему случаю, когда пространства и функционалы комплексные. Выделив вещественную и мнимую части $f(x) = g(x) + \mathbf{i}h(x)$, заметим, что вещественные функционалы $g(x)$ и $h(x)$ уже не будут линейными. Тем не менее, $g(x)$ является вещественным линейным функционалом, если L рассматривать как линейное пространство над полем вещественных чисел. Выполнив подряд два описанных выше шага продолжения, получим вещественный линейный функционал $\widetilde{g}(x)$ на пространстве векторов вида $u + sw + t(\mathbf{i}\,w)$, где $u \in L$ и $s,t \in \mathbb{R}$. При этом будет выполняться неравенство $|\widetilde{g}(u+sw+t(\mathbf{i}\,w))| \le ||u+sw+t\mathbf{i}\,w||$. Отсюда понятно, что $\widetilde{g}(x)$ можно рассматривать как вещественный функционал, определенный на $\widetilde{L}$ и такой, что

$$|\widetilde{g}(x)| \le ||x|| \quad \forall \;\; x \in \widetilde{L}. \qquad (*)$$

Функционал $\widetilde{g}(x)$ обладает свойством линейности лишь при умножении на вещественные числа. Однако, с его помощью можно определить функционал

$$\widetilde{f}(x) = \widetilde{g}(x) - \mathbf{i}\widetilde{g}(\mathbf{i}x), \quad x \in \widetilde{L},$$

который, как можно убедиться, уже является комплексным линейным функционалом на $\widetilde{L}$. К тому же, при всех $x \in \widetilde{L}$ его вещественная часть $\mathtt{re}(\widetilde{f}(x))$ совпадает с $\widetilde{g}(x)$, а при всех $x \in L$ имеем $\widetilde{f}(x) = f(x)$.

Остается доказать, что $|\widetilde{f}(x)| \le ||x||$ при всех $x \in \widetilde{L}$. Пусть $\widetilde{f}(x) = |\widetilde{f}(x)|\xi$, где $\xi \in \mathbb{C}$ и, очевидно, $|\xi| = 1$. Согласно неравенству $(*)$, находим

$$|\widetilde{f}(x)| = |\widetilde{f}(\overline{\xi}x)| = |\widetilde{g}(\overline{\xi}x)| \le ||\overline{\xi}x|| = ||x||. \quad \square$$

Лекция 27

ОСНОВНАЯ ЧАСТЬ

27.1 Линейные операторы

Любую матрицу $A \in \mathbb{C}^{m\times n}$ можно естественным образом рассматривать как оператор, отображающий вектор $x \in \mathbb{C}^n$ в вектор $Ax \in \mathbb{C}^m$. Этот оператор очевидно обладает *свойством линейности* [1]

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay \quad \forall \ \alpha, \beta \in \mathbb{C}, \quad \forall \ x, y \in \mathbb{C}^n.$$

То же свойство линейности выполняется для многих очень важных отображений в линейных пространствах, элементами которых являются функции, объединенные каким-либо общим признаком (непрерывность, дифференцируемость и т.п.). Прежде всего, нужно сказать об отображениях, связанных с дифференцированием и интегрированием функций. Таким образом, поводов к тому, чтобы изучить свойство линейности с более общих позиций более чем достаточно.

Определение. Пусть V и W — произвольные линейные пространства над одним и тем же полем P. Отображение $\mathcal{A} : V \to W$ со свойством

$$\mathcal{A}(\alpha x + \beta y) = \alpha \mathcal{A}(x) + \beta \mathcal{A}(y) \quad \forall \ \alpha, \beta \in P, \quad \forall \ x, y \in V,$$

называется *линейным оператором* из V в W. В случае линейных операторов аргумент принято писать без скобок: $\mathcal{A}(x) = \mathcal{A}x$.

27.2 Непрерывность и ограниченность

Пусть V и W — нормированные пространства. Отображение $\mathcal{A} : V \to W$ называется *непрерывным в точке* $x \in V$, если для любой последовательности $x_k \in V$ такой, что $x_k \to x$ при $k \to \infty$, последовательность образов $\mathcal{A}(x_k)$ сходится к $\mathcal{A}(x)$:

$$||x_k - x||_V \to 0 \quad \Rightarrow \quad ||\mathcal{A}(x_k) - \mathcal{A}(x)||_W \to 0.$$

Отображение называется *непрерывным* на V, если оно непрерывно для всех $x \in V$.

Линейный оператор $\mathcal{A} : V \to W$ называется *ограниченным*, если для некоторой константы $c > 0$

$$||\mathcal{A}x||_W \le c||x||_V \quad \forall \ x \in V.$$

[1] Это свойство можно рассматривать также как свойство *сохранения операций* при отображении одного линейного пространства в другое пространство над тем же полем. Такие отображения называются *гомоморфизмами*.

Теорема. *Для непрерывности линейного оператора необходима и достаточна его ограниченность.*

Доказательство. Достаточность очевидна из неравенства

$$||\mathcal{A}x_k - \mathcal{A}x||_W = ||\mathcal{A}(x_k - x)||_W \le c||x_k - x||_V.$$

Чтобы доказать необходимость, рассмотрим множество значений нормы $||\mathcal{A}x||_W$ на единичной сфере $S = \{x : ||x||_V = 1\}$. Предположим, что это множество не ограничено. Тогда существует последовательность $x_k \in S$ такая, что $||\mathcal{A}x_k||_W \to \infty$. Положим $y_k = x_k/||\mathcal{A}x_k||_W$ и заметим, что

$$||y_k||_V = 1/||\mathcal{A}x_k||_W \to 0 \quad \Rightarrow \quad ||\mathcal{A}y_k||_W \to 0.$$

Последнее невозможно, так как $||\mathcal{A}y_k||_W = 1$ для всех k. Значит, для какого-то $c > 0$

$$||\mathcal{A}x||_W \le c \ \ \forall \ \ x \in S \quad \Rightarrow \quad ||\mathcal{A}x||_W \le c||x||_V \ \ \forall \ \ x \in V. \qquad \square$$

27.3 Операторная норма

Утверждение 1. *Множество $\mathcal{L}(V, W)$ всех ограниченных линейных операторов из V в W является линейным пространством (над общим для V и W полем).*

Доказательство. Пусть $||\mathcal{A}x||_W \le c_1||x||_V$, $||\mathcal{B}x||_W \le c_2||x||_V$. Тогда для любых чисел α и β

$$||(\alpha\mathcal{A} + \beta\mathcal{B})x||_W \le c||x||_V, \qquad c = |\alpha|c_1 + |\beta|c_2. \qquad \square$$

Утверждение 2. *Величина*

$$||\mathcal{A}|| \ \equiv \ \sup_{||x||_V=1} ||\mathcal{A}x||_W, \quad \mathcal{A} \in \mathcal{L}(V, W), \tag{$*$}$$

является нормой на линейном пространстве $\mathcal{L}(V, W)$.

Доказательство. Очевидно, величина $||\mathcal{A}||$ имеет конечное значение и, конечно, неотрицательна. Если $||\mathcal{A}|| = 0$, то $||\mathcal{A}x||_W = 0$ на единичной сфере $||x||_V = 1 \quad \Rightarrow \quad ||\mathcal{A}x||_W = 0 \ \ \forall \ \ x \in V \quad \Rightarrow \quad \mathcal{A}x = 0 \ \ \forall \ \ x \in V \quad \Rightarrow \quad \mathcal{A} = 0$. Положительная однородность следует из равенства

$$||\alpha\mathcal{A}x||_W = |\alpha| \, ||\mathcal{A}x||_W,$$

а неравенство треугольника — из неравенства

$$||(\alpha\mathcal{A} + \beta\mathcal{B})x||_W \le |\alpha| \, ||\mathcal{A}x||_W + |\beta| \, ||\mathcal{B}x||_W. \qquad \square$$

Определение. Норма $(*)$ для операторов $\mathcal{A} \in \mathcal{L}(V, W)$ называется *операторной нормой* или нормой, *подчиненной* векторным нормам $|| \cdot ||_V$, $|| \cdot ||_W$.

Утверждение 3. *Если V — конечномерное пространство, то любой линейный оператор $\mathcal{A} : V \to W$ является ограниченным и $||\mathcal{A}|| = ||\mathcal{A}x_0||_W$ для некоторого (зависящего от $\mathcal{A}$) вектора $x_0 \in V$ с нормой $||x_0||_V = 1$.*

Доказательство. Пусть $e_1, \ldots, e_n$ — базис в V и $x = \sum\limits_{i=1}^{n} \alpha_i e_i$. Тогда $||x||_{(e)} \equiv \sum\limits_{i=1}^{n} |\alpha_i|$ есть норма на V, эквивалентная любой другой норме, в том числе и норме $||x||_V$. Поэтому

для какого-то $c > 0$

$$||x||_{(e)} \le c||x||_V \quad \forall \ x \in V.$$

Следовательно,

$$||\mathcal{A}x||_W \le \sum_{i=1} |\alpha_i| \max_{1\le i\le n} ||\mathcal{A}e_i||_W = ||x||_{(e)} \max_{1\le i\le n} ||\mathcal{A}e_i||_W$$
$$\le (c \max_{1\le i\le n} ||\mathcal{A}e_i||_W)\ ||x||_V \qquad \forall \ x \in V.$$

Чтобы доказать существование x_0, достаточно учесть компактность единичной сферы в конечномерном пространстве, непрерывность на ней функции $||\mathcal{A}x||_W$ и теорему Вейерштрасса. $\square$

27.4 Матричная норма

Пусть каждой комплексной матрице A поставлено в соответствие неотрицательное число $f(A)$ таким образом, что:

(1) $f(A)$ является нормой на $\mathbb{C}^{m\times n}$ для всех m, n;

(2) $f(AB) \le f(A)f(B)$ для любых матриц A и B, допускающих умножение.

В таких случаях $f(A)$ называется *матричной нормой.*

Утверждение. *Пусть для каждого n задана векторная норма на $\mathbb{C}^n$, и пусть для каждых m, n и каждой матрицы $A \in \mathbb{C}^{m\times n}$ норма $||A||$ определена как операторная норма, порожденная данными векторными нормами. Тогда $||A||$ является матричной нормой.*

Доказательство. Пусть $||x||_*$ обозначает векторную норму для $x \in \mathbb{C}^n$ при любом n. Для любых матриц A и B, допускающих умножение, существует вектор x_0 единичной нормы такой,

$$||AB|| \ = \ ||ABx_0||_* \ \le \ ||A||\,||Bx_0||_* \ \le \ ||A||\,||B||\,||x_0||_* \ = \ ||A||\,||B||. \qquad \square$$

Задача. Дана обратимая матрица $A \in \mathbb{C}^{n\times n}$, выбирается произвольная матрица $X_0 \in \mathbb{C}^{n\times n}$ и строится последовательность матриц $X_{k+1} = 2X_k - X_kAX_k, \quad k = 0, 1, \ldots$. Доказать, что если для некоторой матричной нормы $||I - AX_0|| < 1$, то $X_k \to A^{-1}$ при $k \to \infty$.

27.5 Норма Фробениуса

Пусть $A = [a_{ij}]$ — матрица размеров $m \times n$. Величина

$$||A||_F \ = \ \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2}$$

называется *нормой Фробениуса* или *евклидовой нормой* матрицы A.

Утверждение. *Норма Фробениуса является матричной нормой.*

Доказательство. Для каждых m, n норма Фробениуса является нормой на линейном пространстве $\mathbb{C}^{m\times n}$ (как 2-норма на пространстве $\mathbb{C}^{mn}$, изоморфном $\mathbb{C}^{m\times n}$). Пусть $a_1, \ldots, a_n$ — столбцы матрицы A, а $b_1^\top, \ldots, b_n^\top$ — строки матрицы B. Тогда

$$AB \ = \ a_1b_1^\top + \ldots + a_nb_n^\top.$$

Используя неравенство треугольника, легко проверяемые равенства $||a_i b_i^\top||_F = ||a_i||_F||b_i||_F$ и неравенство Коши–Буняковского–Шварца, находим

$$||AB||_F \le \sum_{i=1}^{n} ||a_i b_i^\top||_F = \sum_{i=1}^{n} ||a_i||_F||b_i||_F$$

$$\le \left(\sum_{i=1}^{n} ||a_i||_F^2\right)^{1/2} \left(\sum_{i=1}^{n} ||b_i||_F^2\right)^{1/2} = ||A||_F||B||_F. \qquad \square$$

Замечание. Норма Фробениуса не может быть операторной нормой на $\mathbb{C}^{m\times n}$ ни при каком выборе векторных норм в пространствах $\mathbb{C}^n$ и $\mathbb{C}^m$ (докажите!).

27.6 Сохранение норм

Линейный ограниченный оператор $\mathcal{A} : V \to V$ со свойством

$$||\mathcal{A}x|| = ||x|| \quad \forall \ x \in V$$

называется *изометрическим* или *сохраняющим норму*. Сразу же заметим, что сохранение какой-то одной нормы не означает сохранение другой нормы.

Пусть в $\mathbb{C}^n$ задана какая-то норма, а матрица $A \in \mathbb{C}^{n\times n}$ (как линейный оператор из $\mathbb{C}^n$ в $\mathbb{C}^n$) ее сохраняет. Такую матрицу будем называть изометрической относительно данной нормы.

Утверждение. *Множество всех комплексных $n \times n$-матриц, изометрических относительно гельдеровской 2-нормы, совпадает с множеством унитарных матриц порядка n.*

Доказательство. Очевидно, 2-норма порождается естественным скалярным произведением в $\mathbb{C}^n$. Из наших исследований, связанных с тождеством параллелограмма, вытекает, что сохранение длин влечет за собой сохранение скалярных произведений:

$$(Ax, Ay) = (x, y) \quad \Leftrightarrow \quad y^*(A^*A)x = y^*x \quad \forall \ x, y \in \mathbb{C}^n.$$

Отсюда $y^*(A^*A - I)x = 0$ для все $x, y \in \mathbb{C}^n$. Выбирая в качестве x и y векторы стандартного базиса, приходим к выводу о том, что все элементы матрицы $A^*A - I$ равны нулю. Таким образом, сохранение 2-нормы равносильно условию $A^*A = I$, определяющем унитарную матрицу. $\square$

Замечание. Множество матриц, сохраняющих p-норму в случае $p \ne 2$, значительно беднее. Можно доказать, что для всех $p \ne 2$ оно одно и то же и совпадает с множеством матриц вида DP, где D — диагональная унитарная матрица, а P — матрица перестановки.

27.7 Унитарно инвариантные нормы

Матричная норма $||\cdot||$ называется *унитарно инвариантной*, если $||PAQ|| = ||A||$ для любой матрицы A и любых унитарных матриц P и Q, допускающих умножение.

Утверждение 1. *Норма Фробениуса является унитарно инвариантной.*

Доказательство. Пусть Q — унитарная матриц и $A = [a_1, \ldots, a_n]$. Тогда

$$||Qa_j||_2 = ||a_j||_2, \quad j = 1, \ldots, n.$$

Отсюда

$$||QA||_F^2 = \sum_{j=1}^{n} ||Qa_j||_2^2 = \sum_{j=1}^{n} ||a_j||_2^2 = ||A||_F^2. \qquad \square$$

Заметим, что при изучении метода вращений (в связи с упрощением вида уравнений для поверхностей 2-го порядка) мы уже использовали факт сохранения суммы квадратов элементов вещественной матрицы при умножении ее слева и справа на ортогональные матрицы.

Рассмотрим еще матричную норму, подчиненную гельдеровской 2-норме:

$$||A|| = \sup_{||x||_2=1} ||Ax||_2.$$

Данная норма называется *спектральной нормой* матрицы (смысл названия через некоторое время прояснится). Обозначение: $||A||_2$.

Утверждение 2. *Спектральная норма матрицы является унитарно инвариантной.*

Доказательство. Пусть Q — унитарная матриц и $A = [a_1, \ldots, a_n]$. По определению,

$$||A||_2 = \sup_{||x||_2=1} ||Ax||_2 = \sup_{||x||_2=1} ||(QA)x||_2 = ||QA||_2.$$

Кроме того,

$$||AQ||_2 = \sup_{||x||_2=1} ||(AQ)x||_2 = \sup_{||Q^*x||_2=1} ||(AQ)(Q^*x)||_2 = \sup_{||x||_2=1} ||(Ax)||_2 = ||A||_2. \qquad \square$$

27.8 Сингулярное разложение матрицы

В 70-х годах 19-го века независимо и почти одновременно Бельтрами (1873) и Жордан (1874) открыли, что любую квадратную матрицу можно привести к диагональному виду с помощью умножения слева и справа на унитарные матрицы. Различные вопросы, связанные с данным открытием, в том числе его обобщения, стали затем предметом целого ряда исследований. Не будет сильным преувеличением сказать, что данный факт оказался потрясающе полезным и одним из наиболее востребованных в теории матриц и приложениях линейной алгебры.

В действительности то же верно и для прямоугольной матрицы. С помощью умножения на унитарные матрицы она приводится к прямоугольной матрице тех же размеров, имеющей всюду нули, кроме элементов с индексами $i = j$. Такие матрицы будем называть *диагональными прямоугольными* матрицами. Итак, речь идет о разложении вида

$$A = V\Sigma U^*, \qquad (*)$$

где A — заданная $m \times n$-матрица, U и V — унитарные матрицы соответственно порядка m и n, а Σ — диагональная прямоугольная $m \times n$-матрица, имеющая при $i = j$ неотрицательные числа

$$\sigma_1 \ \geq \ \sigma_2 \ \geq \ \ldots \ \geq \ \sigma_{\min(m,n)}.$$

Разложение $(*)$ называется *сингулярным разложением* матрицы A. Числа σ_i называются *сингулярными числами* матрицы A.

Теорема. *Сингулярное разложение $A = V\Sigma U^*$ существует для любой комплексной*

прямоугольной матрицы A. Если A вещественная, то матрицы U и V можно выбрать вещественными.

Доказательство. Положим $\sigma_1 = ||A||_2 = \sup\limits_{x\neq 0} ||Ax||_2/||x||_2$. В силу компактности единичной сферы в $\mathbb{C}^n$, непрерывности нормы и теоремы Вейершрасса, найдется вектор x_1 такой, что $||Ax_1||_2 = ||A||_2$ и $||x_1||_2 = 1$. Пусть $y_1 = Ax_1/||Ax_1||_2$. Таким образом,

$$Ax_1 = \sigma_1 y_1, \quad ||x_1||_2 = ||y_1||_2. \qquad (\#)$$

Дополним x_1 и y_1 до ортонормированных базисов и образуем унитарные матрицы

$$U_1 = [x_1, x_2, \ldots, x_n], \quad V_1 = [y_1, y_2, \ldots, y_m].$$

Согласно (#), матрица $A_1 \equiv V_1^* A U_1$ имеет в первом столбце только один ненулевой элемент, равный σ_1:

$$A_1 = V_1^* A U_1 = \begin{bmatrix} \sigma_1 & z^* \\ 0 & A_2 \end{bmatrix}.$$

В силу унитарной инвариантности спектральной нормы, $||A_1|| = \sigma_1$. Поэтому

$$\sigma_1 \left\| \begin{bmatrix} \sigma_1 \\ z \end{bmatrix} \right\|_2 \geq \left\| A_1 \begin{bmatrix} \sigma_1 \\ z \end{bmatrix} \right\|_2 \geq \sigma_1^2 + ||z||_2^2 \;\Rightarrow\; \sigma_1^2 \geq \sigma_1^2 + ||z||_2^2 \;\Rightarrow\; z = 0 \;\Rightarrow\; A_1 = \begin{bmatrix} \sigma_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

Далее будем рассуждать по индукции. Если для A_2 уже имеется сингулярное разложение $A_2 = V_2^* \Sigma_2 U_2$, то сингулярное разложение для A находится с легкостью. Для этого достаточно взять

$$U = U_1 \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix}, \qquad V = V_1 \begin{bmatrix} 1 & 0 \\ 0 & V_2 \end{bmatrix},$$

заметить, что матрицы U и V унитарные (как произведение унитарных матриц), и убедиться в том, что выполняется равенство

$$V^* A U = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}.$$

Остается заметить, что индукция начинается с построения сингулярного разложения для матриц, представляющих собой один столбец либо одну строку.

Пусть $A = [a] \in \mathbb{C}^{m\times 1}$ — матрица-столбец. В этом случае найдем в $\mathbb{C}^m$ ортонормированный базис $v_1, \ldots, v_m$, начинающийся с $v_1 = a/||a||_2$. Тогда

$$A = V\Sigma U^*, \qquad V = [v_1, \ldots, v_m], \quad \Sigma = [||a||_2, 0, \ldots, 0]^\top, \quad U = [1] \in \mathbb{C}^{1\times 1}.$$

Для матрицы-строки сингулярное разложение получается транспонированием. □

Следствие 1. *Спектральная норма матрицы равна ее старшему сингулярному числу.*

Следствие 2. *Пусть матрица A обратима и σ_n — ее младшее сингулярное число. Тогда $||A^{-1}||_2 = 1/\sigma_n$.*

Теперь ясно, что старшее сингулярное число матрицы и младшее сингулярное число обратимой матрицы определены однозначно. То же верно для всего набора сингулярных чисел, но это мы докажем позже. Сингулярное разложение вместе еще с рядом важных следствий заслуживает более обстоятельного обсуждения, которое мы временно отложим — с тем, чтобы вернуться к нему на более подготовленной почве.

Лекция 28

ОСНОВНАЯ ЧАСТЬ

28.1 Матрица линейного оператора

Рассмотрим линейный оператор $\mathcal{A} : V_n \to V_m$, где V_n и V_m — линейные пространства размерности n и m (над общим полем P).

Фиксируем какой-нибудь базис $e_1, \ldots, e_n$ в V_n и какой-нибудь базис $f_1, \ldots, f_m$ в V_m. В силу линейности оператора $\mathcal{A}$,

$$\mathcal{A}(x_1e_1 + \ldots + x_ne_n) \;=\; x_1(\mathcal{A}e_1) \;+\; \ldots \;+\; x_n(\mathcal{A}e_n). \tag{1}$$

Поэтому $\mathcal{A}$ полностью определяется своим действием на базисных векторах $e_1, \ldots, e_n$. Разложим образы базисных векторов по базису пространства образов:

$$\mathcal{A}e_j \;=\; a_{1j}f_1 \;+\; \ldots \;+\; a_{mj}f_m, \quad j = 1, \ldots, n. \tag{2}$$

Из (1) и (2) получаем

$$\mathcal{A}(x_1e_1 + \ldots + x_ne_n) \;=\; (a_{11}x_1 + \ldots + a_{1n}x_n)\, f_1 \;+\; \ldots \;+\; (a_{m1}x_1 + \ldots + a_{mn}x_n)\, f_m.$$

Следовательно,

$$\mathcal{A}(x_1e_1 + \ldots + x_ne_n) = y_1f_1 + \ldots + y_mf_m \quad \Leftrightarrow \quad \begin{bmatrix} y_1 \\ \ldots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}.$$

Матрица, возникшая справа, называется *матрицей линейного оператора* $\mathcal{A}$ в паре базисов $\{e_j\}$ и $\{f_i\}$.

Таким образом, любая фиксированная пара базисов порождает три изоморфизма

$$V_n \leftrightarrow P^n, \qquad V_m \leftrightarrow P^m, \qquad \mathcal{L}(V_n, V_m) \leftrightarrow P^{m\times n},$$

где $\mathcal{L}(V_n, V_m)$ — линейное пространство всех линейных операторов, действующих из пространства V_n в пространство V_m, а $P^{m\times n}$ — линейное пространство всех $m \times n$-матриц с элементами из поля P.

Отсюда, в частности, видно, что размерность пространства линейных операторов $\mathcal{L}(V_n, V_m)$ равна mn.

28.2 Произведение линейных операторов

Произведение линейных операторов $\mathcal{A} : V_n \to V_m$ и $\mathcal{B} : V_m \to V_k$ определяется как композиция отображений: $\mathcal{BA}$ — это оператор из V_n в V_k, заданный правилом $(\mathcal{BA})(x) = \mathcal{B}(\mathcal{A}x)$. Элементарно проверяется, что произведение линейных операторов является линейным оператором.

Пусть A — матрица линейного оператора $\mathcal{A}$ в паре базисов $\{e_j\}$ и $\{f_i\}$, а B — матрица линейного оператора в паре базисов $\{f_i\}$ и $\{g_l\}$. Тогда *произведение матриц BA есть матрица произведения операторов $\mathcal{BA}$ в паре базисов $\{e_j\}$ и $\{g_l\}$.*

Доказательство сводится к прямой проверке. Заметим, что наш курс, собственно, начался с определения произведения матриц и фактически с обсуждения композиции линейных отображений!

28.3 Переход к другим базисам

Пусть A_{ef} — матрица линейного оператора $\mathcal{A}$ в паре базисов $e = \{e_j\}$ и $f = \{f_i\}$. Как найти матрицу A_{gh} того же оператора в другой паре базисов $g = \{g_j\}$ и $h = \{h_i\}$?

Рассмотрим равенства

$$\mathcal{A}(x_1e_1 + \ldots + x_ne_n) = y_1f_1 + \ldots + y_mf_m, \qquad \mathcal{A}(z_1g_1 + \ldots + z_ng_n) = u_1h_1 + \ldots + u_mh_m.$$

Согласно определению матриц A_{ef} и A_{gh}, находим

$$A_{ef}x = y, \quad A_{gh}z = u, \qquad x = \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}, \; y = \begin{bmatrix} y_1 \\ \ldots \\ y_m \end{bmatrix}, \quad z = \begin{bmatrix} z_1 \\ \ldots \\ z_n \end{bmatrix}, \; u = \begin{bmatrix} u_1 \\ \ldots \\ u_m \end{bmatrix},$$

Далее, запишем

$$g_j = s_{1j}e_1 + \ldots + s_{nj}e_n, \quad 1 \le j \le n, \qquad h_i = t_{1i}f_1 + \ldots + t_{mi}f_m, \quad 1 \le i \le m,$$

и введем матрицы перехода

$$S = \begin{bmatrix} s_{11} & \ldots & s_{1n} \\ \ldots & \ldots & \ldots \\ s_{n1} & \ldots & s_{nn} \end{bmatrix}, \quad T = \begin{bmatrix} t_{11} & \ldots & t_{1m} \\ \ldots & \ldots & \ldots \\ t_{m1} & \ldots & s_{mm} \end{bmatrix}.$$

Тогда $x = Sz$ и $y = Tu$. Следовательно, $A_{ef}(Sz) = Tu \;\; \Rightarrow \;\; (T^{-1}A_{ef}S)z = u \;\; \Rightarrow$

$$A_{gh} \;=\; T^{-1}A_{ef}S. \qquad (*)$$

Напомним определение эквивалентных матриц: A и B называются *эквивалентными*, если $B = PAQ$ для каких-то невырожденных P и Q.

Утверждение 1. *Матрицы эквивалентны в том и только том случае, когда они являются матрицами одного и того же линейного оператора в каких-то парах базисов.*

Доказательство. Очевидно, $(*)$ означает эквивалентность матриц A_{gh} и A_{ef}. Если $B = PAQ$, то P и Q можно рассматривать как матрицы перехода для разных пар базисов. □

Следствие. *Для того чтобы матрицы одинаковых размеров были матрицами одного и того же линейного оператора в каких-то парах базисов, необходимо и достаточно, чтобы они имели одинаковый ранг.*

Пусть A — матрица линейного оператора $\mathcal{A}$ в какой-то паре базисов. Если $r = \mathrm{rank}A$, то A эквивалентна матрице $B = [b_{ij}]$, в которой $b_{11} = \ldots = b_{rr} = 1$, а все остальные элементы равны 0. Следовательно, имеется пара "канонических" базисов, в которой $\mathcal{A}$ определяется матрицей B.

Таким образом, за счет выбора пары базисов матрица линейного оператора может приобрести вид настолько простой, чтобы оказаться почти бесполезной для изучения индивидуальных свойств данного оператора. Поэтому изучение оператора, вообще говоря, нельзя свести к изучению его матрицы.

Если $V_n = V_m$, то появляется возможность взять $e = f$. Вследствие того, что образы и прообразы рассматриваются в одном и том же базисе, теперь в матрице оператора есть все, что нужно для любого подробного его изучения. То же верно для любой другой пары базисов f и g, если только $f = g$. В этом случае, конечно, $T = S \quad \Rightarrow$

$$A_{gg} \;=\; S^{-1}A_{ee}S. \qquad (**)$$

Матрицы A и B называются *подобными*, если $B = S^{-1}AS$ для какой-то невырожденной матрицы S. Очевидно, справедливо следующее

Утверждение 2. *Матрицы подобны в том и только том случае, когда они являются матрицами одного и того же линейного оператора в каких-то базисах при условии выбора одинаковых базисов в общем пространстве образов и прообразов.*

28.4 Преобразование подобия

Пусть $\mathcal{A} : V_n \to V_n$ — линейный оператор в n-мерном пространстве V_n, и A — его матрица при выборе одного и того же базиса в пространстве образов и прообразов. В этом случае изучение оператора $\mathcal{A}$ полностью сводится к изучению его матрицы A.

Естественно попытаться выбрать базис таким образом, чтобы матрица A получила "наиболее простой" вид. Если оператор $\mathcal{A}$ задан своей матрицей A в каком-то базисе, то выбор нового базиса дает для того же оператора другую матрицу B, которая будет подобна заданной матрице. Переход от A к подобной ей матрице $B = S^{-1}AS$ называется *преобразованием подобия.*

Возникает такой вопрос: к какому "наиболее простому" виду можно привести заданную матрицу с помощью преобразования подобия? Фактически мы сейчас начинаем не очень простой путь к полному ответу на данный вопрос.

28.5 Инвариантные подпространства

Проблем нет, если $n = 1$. Кажется также, что проще изучать оператор в пространстве малой размерности. Поэтому давайте для начала поизучаем действие оператора $\mathcal{A}$ на подпространствах малой размерности.

Пусть L — подпространство в V_n. Чтобы изучать $\mathcal{A}$, используя только векторы из L, нужно потребовать, чтобы $\mathcal{A}x \in L$ для всех $x \in L$. Любое подпространство с таким свойством называется *инвариантным* относительно $\mathcal{A}$.

Пусть $v_1, \ldots, v_k$ — базис в подпространстве L. Тогда его можно дополнить какими-то векторами $v_{k+1}, \ldots, v_n$ до базиса в V_n.

Утверждение. *Пусть $v_1, \ldots, v_n$ — базис в V_n и $L = L(v_1, \ldots, v_k)$. Тогда L инвариантно относительно линейного оператора $\mathcal{A} : V_n \to V_n$ тогда и только тогда, когда*

матрица оператора $\mathcal{A}$ в базисе $v_1, \ldots, v_n$ имеет блочно треугольный вид

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \qquad (*)$$

где A_{11} — подматрица порядка k.

Доказательство. При изоморфизме $x = [x_1, \ldots, x_n]^\top \leftrightarrow x_1 v_1 + \ldots + x_n v_n$ векторам из L соответствуют те и только те столбцы x, для которых $x_{k+1} = \ldots = x_n = 0$. Если A имеет вид $(*)$, то, очевидно, для $y = [y_1, \ldots, y_n]^\top = Ax$ получаем $y_{k+1} = \ldots = y_k = 0$. Значит, L инвариантно относительно умножения на матрицу $A \quad \Rightarrow \quad L$ инвариантно относительно оператора $\mathcal{A}$.

Пусть известно, что L инвариантно относительно умножения на матрицу $A = [a_{ij}]$, и пусть $y = Ax$, где $x_{k+1} = \ldots = x_n = 0$. Тогда $y_{k+1} = \ldots = y_n = 0 \quad \Rightarrow \quad a_{ij} = 0$ при $1 \le j \le k$, $i \ge k+1$. $\square$

28.6 Ядро и образ линейного оператора

Множество $\ker\mathcal{A} \equiv \{x \in V_n : \mathcal{A}x = 0\}$ называется *ядром* линейного оператора $\mathcal{A}$, а множество $\mathrm{im}\mathcal{A} \equiv \{y \in V_n : y = \mathcal{A}x,\ x \in V_n\}$ — его *образом*.

Утверждение 1. *Ядро линейного оператора $\mathcal{A} : V \to W$ является подпространством в V, а его образ — подпространством в W.*

Доказательство. Пусть $x, y \in \ker\mathcal{A}$. Тогда $\mathcal{A}(\alpha x + \beta y) = \alpha\mathcal{A}x + \beta\mathcal{A}y = \alpha \cdot 0 + \beta \cdot 0 = 0 \quad \Rightarrow \quad \alpha x + \beta y \in \ker$. Пусть $x, y \in \mathrm{im}\mathcal{A}$. Тогда $x = \mathcal{A}u$ и $y = \mathcal{A}v$ для каких-то $u, v \in V \quad \Rightarrow \quad \alpha x + \beta y = \mathcal{A}(\alpha u + \beta v) \quad \Rightarrow \quad \alpha x + \beta y \in \mathrm{im}\mathcal{A}$. $\square$

Теорема о размерности ядра и образа. *Пусть V конечномерно. Тогда*

$$\dim\ker\mathcal{A} \ + \ \dim\mathrm{im}\mathcal{A} \ = \ \dim V.$$

Доказательство. Пусть $\dim\ker\mathcal{A} = k$ и $v_1, \ldots, v_k$ — базис в $\ker\mathcal{A}$. Достроим его какими-то векторами $v_{k+1} \ldots, v_n$ до базиса в V. Очевидно,

$$\mathrm{im}\mathcal{A} \ = \ L(\mathcal{A}v_{k+1}, \ldots, \mathcal{A}v_n).$$

Остается доказать, что векторы $\mathcal{A}v_{k+1}, \ldots, \mathcal{A}v_n$ линейно независимы.

Пусть $\alpha_{k+1}\mathcal{A}v_{k+1} + \ldots + \alpha_n\mathcal{A}v_n \ = \ 0 \quad \Rightarrow \quad \mathcal{A}(\alpha_{k+1}v_{k+1} + \ldots + \alpha_n v_n) \ = \ 0 \quad \Rightarrow \quad \alpha_{k+1}v_{k+1} + \ldots + \alpha_n v_n \in \ker\mathcal{A} \quad \Rightarrow \quad \alpha_{k+1}v_{k+1} + \ldots + \alpha_n v_n = \beta_1 v_1 + \ldots + \beta_k v_k$ для каких-то чисел $\beta_1, \ldots, \beta_k \quad \Rightarrow \quad \alpha_{k+1} = \ldots = \alpha_n = 0$. $\square$

Замечание. Данную теорему можно было бы и не доказывать, поскольку она есть следствие уже известного нам факта: для любой матрицы A сумма ее ранга и размерности ее ядра (линейного пространства решений однородной системы $Ax = 0$) равна числу ее столбцов. Мы знаем, что ранг совпадает с размерностью линейной оболочки столбцов матрицы, а последняя, очевидно, есть ее образ (как оператора умножения на данную матрицу).

Утверждение 2. *Пусть линейный оператор $\mathcal{A}$ действует из V в V. Тогда его ядро и образ инвариантны относительно $\mathcal{A}$.*

Доказательство. Инвариантность ядра очевидна, поскольку любой его вектор отображается в 0.

Пусть $x \in \mathrm{im}\mathcal{A}$. Тогда $x = \mathcal{A}u \quad \Rightarrow \quad \mathcal{A}x = \mathcal{A}(\mathcal{A}u) \quad \Rightarrow \quad \mathcal{A}x \in \mathrm{im}\mathcal{A}$. $\square$

28.7 Обратный оператор

Оператор $\mathcal{A} : V \to W$ называется *обратимым*, если существует оператор $\mathcal{B} : W \to V$ такой, что $\mathcal{A}(\mathcal{B}(y)) = y \quad \forall\ y \in W$ и $\mathcal{B}(\mathcal{A}(x)) = x \quad \forall\ x \in V$. При этом $\mathcal{B}$ называется *обратным оператором* для $\mathcal{A}$.

Утверждение. *Если линейный оператор обратим, то обратный оператор также является линейным.*

Доказательство. Любые векторы $y_1, y_2 \in W$ можно представить в виде $y_1 = \mathcal{A}x_1$, $y_2 = \mathcal{A}x_2$. Поэтому

$$\mathcal{B}(\alpha y_1 + \beta y_2) = \mathcal{B}(\alpha\mathcal{A}y_1 + \beta\mathcal{B}y_2) = \mathcal{B}(\mathcal{A}(\alpha x_1 + \beta x_2)) = \alpha x_1 + \beta x_2.$$

Остается учесть, что

$$x_1 = \mathcal{B}y_1, \quad x_2 = \mathcal{B}y_2. \qquad \square$$

Теорема. *Пусть $\mathcal{A} : V \to W$ — линейный оператор, а V и W — конечномерные пространства одинаковой размерности. Тогда $\mathcal{A}$ является обратимым оператором тогда и только тогда, когда* $\ker\mathcal{A} = \{0\}$.

Доказательство. Пусть $\dim V = \dim W = n$. Согласно теореме о размерности ядра и образа, если $\dim\ker\mathcal{A} = 0$, то $\dim\mathrm{im}\mathcal{A} = n$. Это означает, что для каждого вектора $y \in W$ существует $x \in V$ такой, что $\mathcal{A}x = y$. Более того, такой вектор x единствен (иначе ядро содержало бы ненулевой вектор). Определим оператор $\mathcal{B} : W \to V$ правилом $\mathcal{B}(y) = x$. Тогда $\mathcal{A}(\mathcal{B}(y)) = y$ и $\mathcal{B}(\mathcal{A}(x)) = x \quad \Rightarrow \quad \mathcal{B}$ является обратным оператором для $\mathcal{A}$.

Если же известно, что $\mathcal{A}$ — обратимый оператор, то его ядро может быть только нулевым (если для каких-то $x_1 \neq x_2$ выполнялось бы равенство $\mathcal{A}x_1 = \mathcal{A}x_2$, то это противоречило бы обратимости оператора $\mathcal{A}$). $\square$

Замечание. Если линейный оператор $\mathcal{A} : V \to W$ обратим, то непременно $W = \mathrm{im}\mathcal{A}$. В то же время, условие $W = \mathrm{im}\mathcal{A}$ недостаточно для обратимости $\mathcal{A}$.

28.8 Ортогональные дополнения ядра и образа

Дадим еще одно доказательство теоремы о размерности ядра и образа. Пусть $A \in \mathbb{C}^{m\times n}$. Если $x \in \ker A$, то для любого $y \in \mathbb{C}^m$ находим

$$0 = y^*Ax = (A^*y)^*x = (x, A^*y) \quad \Rightarrow \quad x \perp \mathrm{im}A^* \quad \Rightarrow \quad \ker A \subset (\mathrm{im}A^*)^\perp.$$

Пусть теперь $x \in (\mathrm{im}A^*)^\perp$. Тогда $(x, A^*y) = y^*Ax = 0 \ \ \forall\ \ y \in \mathbb{C}^m$. Взяв $y = Ax$, получаем $y^*Ax = (Ax)^*(Ax) = |Ax|^2 = 0 \quad \Rightarrow \quad Ax = 0 \quad \Rightarrow \quad x \in \ker A \quad \Rightarrow \quad (\mathrm{im}A^*)^\perp \subset \ker A$.

Итак, $\ker A = (\mathrm{im}A^*)^\perp$. Мы уже знаем, что размерность ортогогонального дополнения к $\mathrm{im}A^*$ равна $n - \dim\mathrm{im}A^* = n - \mathrm{rank}A = n - \dim\mathrm{im}A$. $\square$

В действительности нами обнаружено интересное общее свойство ядра матрицы и образа сопряженной матрицы.

Теорема. *Пусть $A \in \mathbb{C}^{m\times n}$. Тогда $\mathbb{C}^n$ и $\mathbb{C}^m$ представляются ортогональными суммами вида*

$$\mathbb{C}^n = \ker A \oplus \mathrm{im}A^*, \qquad \mathbb{C}^m = \ker A^* \oplus \mathrm{im}A.$$

Отметим два очевидных следствия. Они интересны, прежде всего, тем, что в тех же формулировках переносятся на важные классы операторных уравнений в гильбертовых пространствах и помогают получать там факты о существовании и единственности решений.

Теорема Фредгольма. *Для совместности системы $Ax = b$ необходимо и достаточно, чтобы правая часть b была ортогональна ко всем решениям y однородной сопряженной системы $A^*y = 0$.*

Альтернатива Фредгольма. *Либо система $Ax = b$ имеет единственное решение для любой правой части b, либо однородная сопряженная система $A^*y = 0$ имеет только нулевое решение.*

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

28.9 Выбор базиса

С точки зрения "точной математики" все базисы равноправны. Но при проведении вычислений разница между базисами огромна!

Пусть $e = \{e_1, \ldots, e_n\}$ — стандартный базис в $\mathbb{C}^n$, а $g = \{g_1, \ldots, g_n\}$ — какой-то другой базис. Пусть j-й столбец матрицы P состоит из координат вектора g_j в стандартном базисе e. Тогда координаты одного и того же вектора в базисах e и g связаны равенством $x = Pz$, где $x \in \mathbb{C}^n$ содержит координаты разложения вектора по стандартному базису e, а $z \in \mathbb{C}^n$ — координаты разложения того же вектора по базису g. $\Rightarrow$

$$z = P^{-1}x.$$

Типична ситуация, когда в ходе вычислений вместо x возник слабо *возмущенный* вектор $\widetilde{x} = x + \delta$. Тогда вместо z будет получен вектор

$$\widetilde{z} \equiv z + \Delta = P^{-1}(x + \delta) \quad \Rightarrow \quad \Delta = P^{-1}\delta.$$

Предположим, что $x \neq 0$ (тогда и $z \neq 0$). Используя спектральную норму, находим

$$||\Delta||_2 \;=\; \frac{||P^{-1}\delta||_2}{||x||_2}||Pz||_2 \;\leq\; \frac{||P^{-1}||_2||\delta||_2}{||x||_2}||P||_2||z||2 \quad \Rightarrow$$

$$\frac{||\Delta||_2}{||z||_2} \;\leq\; (||P^{-1}||_2||P||_2)\,\frac{||\delta||_2}{||x||_2}. \qquad (\#)$$

Таким образом, *относительная погрешность* $||\Delta||_2/||z||_2$ в векторе z не больше, чем относительная погрешность $||\delta||_2/||x||_2$ в векторе x, умноженная на число

$$\gamma(P) \equiv ||P^{-1}||_2||P||_2.$$

Величина $\gamma(P)$ называется спектральным *числом обусловленности* матрицы P.

К сожалению, число обусловленности может оказаться очень большим, а неравенство $(\#)$ для некоторых векторов x и δ может превращаться в равенство. В самом деле, пусть

$$P = V\Sigma U^* \;\text{ — сингулярное разложение матрицы } P,$$

u_1 и v_1 — первые столбцы матриц U и V, а u_n и v_n — последние столбцы тех же матриц. Тогда

$$Pu_1 = \sigma_1 v_1, \quad Pu_n = \sigma_n v_n.$$

Взяв $x = v_1$ и $\delta = \varepsilon v_n$, находим

$$\frac{||\Delta||_2}{||z||_2} \;=\; \frac{|\varepsilon|/\sigma_n}{1/\sigma_1} \;=\; |\varepsilon|\,\frac{\sigma_1}{\sigma_n} \;=\; ||P^{-1}||_2||P||_2\,\frac{||\delta||_1}{||x||_2}.$$

В отличие от произвольных базисов, ортонормированные базисы обладают замечательным достоинством. Для них матрица P унитарная, а для любой унитарной матрицы спектральное число обусловленности равно 1 (докажите!).

По этой причине математики-вычислители предпочитают, если возможно, иметь дело с ортонормированными базисами.

28.10 Базисы в пространстве многочленов

Пусть $\mathcal{P}_n$ — линейное пространство вещественных многочленов порядка n (степени $n-1$ и ниже). Естественный базис в $\mathcal{P}_n$ образуют одночлены 1, x, $\ldots$, x^{n-1}.

С точки зрения вычислений это "очень плохой" базис. Пусть, например, нужно найти многочлен $p(x) \in \mathcal{P}_n$, принимающий в заданных точках $a \le x_1 < x_2 < \ldots < x_n \le b$ заданные значения f_1, f_2, $\ldots$, f_n. Это могут быть значения какой-то функции $f(x)$ на отрезке $[a,b]$ — в этом случае $p(x)$ можно рассматривать как некоторое приближение к $f(x)$ на данном отрезке, выбираемое из условия совпадения значений $f(x)$ и $p(x)$ в точках x_i. Такая задача называется задачей *интерполяции*, а $p(x)$ — *интерполяционным многочленом* для функции $f(x)$ в узлах x_i. Решение вроде бы очевидно: если $p(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1}$, то

$$\begin{bmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ 1 & x_2 & \ldots & x_2^{n-1} \\ \ldots & \ldots & \ldots & \ldots \\ 1 & x_n & \ldots & x_n^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \ldots \\ a_{n-1} \end{bmatrix} \;=\; \begin{bmatrix} f_1 \\ f_2 \\ \ldots \\ f_n \end{bmatrix}.$$

Однако, *матрица коэффициентов данной системы имеет спектральное число обусловленности не меньше* $2^{n-2}/\sqrt{n}$ *независимо от выбора узлов* x_i. [1] Поэтому даже малые погрешности в значениях f_i могут привести к недопустимым погрешностям в коэффициентах интерполяционного многочлена $p(x)$.

Строить вычисления на основе коэффициентов интерполяционного многочлена — дело почти безнадежное. Но это не означает, что нужно отказаться от использования интерполяционных многочленов. Нужно лишь выбрать другой базис для их представления!

Одна из возможностей — записать $p(x)$ следующим образом:

$$p(x) \;=\; \sum_{i=1}^{n} f_i \prod_{\substack{1 \le j \le n \\ j \ne i}} \frac{x - x_j}{x_i - x_j}. \qquad (*)$$

В данном случае для разложения $p(x)$ используется базис из так называемых *элементарных многочленов Лагранжа*

$$l_i(x) \;=\; \prod_{\substack{1 \le j \le n \\ j \ne i}} \frac{x - x_j}{x_i - x_j}.$$

Легко проверить, что

$$l_i(x_k) \;=\; \begin{cases} 1, & i = k, \\ 0, & i \ne k. \end{cases}$$

[1] Простое доказательство (все же требующее техники, которую мы еще не успели развить), можно найти в статье: E. E. Tyrtyshnikov, How bad are Hankel matrices?, *Numer. Math.*, no. 67, 1994, pp. 261–269.

Поэтому $p(x)$ действительно удовлетворяет условиям $p(x_k) = f_k,\ 1 \le k \le n$. Формула $(*)$ называется *интерполяционной формулой Лагранжа*.

Другая возможность — ввести в $\mathcal{P}_n$ скалярное произведение и построить базис из ортогональных (ортонормированных) многочленов с помощью процесса ортогонализации Грама-Шмидта, примененного к системе многочленов $1,\ x,\ x^2,\ \ldots$. Например, для многочленов на отрезке $[-1,\ 1]$ можно определить скалярное произведение как интеграл

$$(f, g) \ = \int_{-1}^{1} f(x)g(x)\,dx, \quad f, g \in \mathcal{P}_n.$$

Тогда получатся ортогональные многочлены, известные как *многочлены Лежандра*.

В теории и вычислениях применяются и многие другие способы задания скалярного произведения в $\mathcal{P}_n$, приводящие к другим полезным системам ортогональных многочленов.

Лекция 29

ОСНОВНАЯ ЧАСТЬ

29.1 Диагонализуемые матрицы

Матрицы, подобные диагональным матрицам, называют *диагонализуемыми* или *матрицами простой структуры*.

Рассмотрим задачу о диагонализации 3×3-матрицы:

$$AP = P\Lambda \quad \Leftrightarrow \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{12} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{bmatrix}.$$

Данное равенство эквивалентно трем равенствам

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} p_{1j} \\ p_{2j} \\ p_{3j} \end{bmatrix} = \lambda_j \begin{bmatrix} p_{1j} \\ p_{2j} \\ p_{3j} \end{bmatrix}, \qquad j = 1,\ 2,\ 3.$$

Предположим, что значение λ_j известно. Тогда элементы j-го столбца матрицы P удовлетворяют однородной системе

$$\begin{bmatrix} a_{11} - \lambda_j & a_{12} & a_{13} \\ a_{11} & a_{12} - \lambda_j & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda_j \end{bmatrix} \begin{bmatrix} p_{1j} \\ p_{2j} \\ p_{3j} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (*)$$

Данная система должна иметь нетривиальное решение $\Leftrightarrow$ определитель матрицы коэффициентов равен нулю. Таким образом, λ_j удовлетворяет следующему уравнению относительно λ:

$$\det \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{11} & a_{12} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{bmatrix} = 0. \qquad (\#)$$

Это кубическое уравнение вида $\lambda^3 - s_2\lambda^2 + s_1\lambda - s_0 = 0$, где, как легко видеть,

$$s_2 = a_{11} + a_{22} + a_{33},$$

$$s_1 = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} + \det \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} + \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix},$$

$$s_0 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Можно вспомнить, что вопросом о диагонализации матриц порядка 3 мы занимались при изучении линий и поверхностей второго порядка — при поиске такой системы координат, в которой матрица квадратичной части (вещественная симметричная матрица порядка 3) становится диагональной. [1]

29.2 Собственные значения и собственные векторы

Пусть A — матрица порядка n и P — обратимая матрица со столбцами $p_1,\ \ldots,\ p_n$. Легко видеть, что равенство $AP = P\Lambda$ эквивалентно системе равенств

$$Ap_j \ = \ \lambda_j\, p_j, \quad j = 1,\ \ldots,\ n.$$

Эти равенства подводят нас к важным понятиям собственного значения матрицы и собственного вектора.

Определение. Пусть A — матрица порядка n. Число $\lambda \in \mathbb{C}$ и ненулевой столбец $x \in \mathbb{C}^n$, связанные соотношением $Ax = \lambda x$, называются *собственным значением* и *собственным вектором* матрицы A. Пара λ, x иногда называется *собственной парой* матрицы A.

Задача. Пусть λ — собственное значение матрицы A. Доказать, что для любой матричной нормы выполняется неравенство $|\lambda| \le ||A||$.

Теорема. *Матрица A порядка n диагонализуема тогда и только тогда, когда она обладает линейно независимой системой n собственных векторов.*

Доказательство. Пусть $p_1,\ \ldots,\ p_n$ — линейно независимая система собственных веткоров матрицы A, соответствующих собственным значениям $\lambda_1,\ \ldots,\ \lambda_n$:

$$Ap_j = \lambda_j p_j,\ \ j = 1,\ \ldots,\ n. \quad \Leftrightarrow \quad AP = P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad P = [p_1,\ \ldots,\ p_n].$$

Матрица A обратима как матрица с линейно независимыми столбцами. $\square$

Пример недиагонализуемой матрицы: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Допустим, что

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

[1] Если B и A — новая и исходная матрицы, то $B = P^\top AP$ и, более того, $P^\top = P^{-1}$ — поскольку исходная и новая системы координат предполагались декартовыми. Таким образом, A и B подобны и при этом преобразование подобия осуществляется с помощью унитарной матрицы. Такие матрицы называются *унитарно подобными*. В Лекции 20 было доказано, что *любая вещественная симметричная матрица унитарно подобна диагональной матрице*. Это же утверждение скоро появится как простое следствие более общих результатов.

Теперь ясно, что при поиске новой декартовой системы координат в случае поверхностей второго порядка можно обойтись без бесконечного процесса с использованием матриц вращения. Достаточно решить кубическое уравнение (#) (например, по формуле Кардано — более сложной, чем формула для корней квадратного уравнения, но в целом вполне аналогичной — см. раздел 14.8), а затем с уже известными $\lambda_1,\ \lambda_2,\ \lambda_3$ найти нетривиальные решения однородных систем линейных алгебраических уравнений (∗). В силу полученной в Лекции 20 теоремы о диагонализации вещественных симметричных матриц можно утверждать, что для вещественных симметричных матриц порядка 3 кубическое уравнение (#) имеет с учетом кратностей три вещественных корня, а из компонент нетривиальных решений соответствующих однородных систем (∗) можно составить ортогональную матрицу P.

Отсюда

$$\begin{bmatrix} p_{21} & p_{22} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} p_{11}\lambda_1 & p_{12}\lambda_2 \\ p_{21}\lambda_1 & p_{22}\lambda_2 \end{bmatrix}.$$

Хотя бы одно из чисел $\lambda_1,\ \lambda_2$ должно отличаться от нуля. Пусть для определенности $\lambda_1 \neq 0 \ \Rightarrow \ p_{21} = 0 \ \Rightarrow \ p_{11} = 0$. Получаем противоречие, поскольку матрица с нулевым столбцом не может быть обратимой. $\square$

29.3 Собственные векторы для различных собственных значений

Теорема. *Собственные векторы, соответствующие попарно различным собственным значениям матрицы, являются линейно независимыми.*

Пусть $x_1,\ \ldots,\ x_m$ — собственные векторы для попарно различных собственных значений $\lambda_1,\ \ldots,\ \lambda_m$ матрицы A. Пусть $\alpha_1 x_1 + \ldots + \alpha_m x_m \ = \ 0$. Умножим обе части слева на матрицу A:

$$\alpha_1\lambda_1 x_1 + \ldots + \alpha_m\lambda_m x_m \ = \ 0.$$

Из данного равенства вычтем предыдущее, умноженное на λ_m:

$$\alpha_1(\lambda_1 - \lambda_m)x_1 + \ldots + \alpha_{m-1}(\lambda_{m-1} - \lambda_m)x_{m-1} \ = \ 0.$$

Отсюда ясно, что из линейной независимости векторов $x_1,\ \ldots,\ x_{m-1}$ вытекала бы линейная независимость векторов $x_1,\ \ldots,\ x_m$. Доказательство завершается применением индукции. $\square$

Следствие. *Если матрица порядка n имеет n различных собственных значений, то она диагонализуема.*

29.4 Характеристическое уравнение

Пусть λ — произвольное собственное значение матрицы A. При фиксированном λ все соответствующие ему собственные векторы x удовлетворяют однородной системе линейных алгебраических уравнений

$$(A - \lambda I)x \ = \ 0.$$

Число λ является собственным значением матрицы $A \quad \Leftrightarrow \quad$ данная система имеет нетривиальное решение $\quad \Leftrightarrow \quad \det(A - \lambda I) \ = \ 0.$

Определение. Уравнение $\det(A - \lambda I) = 0$ относительно λ называется *характеристическим уравнением* матрицы A. Левая часть этого уравнения есть многочлен степени n от λ, называемый *характеристическим многочленом* матрицы A.

Утверждение. *Характеристический многочлен $f(\lambda) = \det(A - \lambda I)$ матрицы A имеет вид*

$$f(\lambda) = (-1)^n(\lambda^n - s_{n-1}\lambda^{n-1} + s_{n-2}\lambda^{n-2} - \ldots + (-1)^n s_0),$$

где s_k есть сумма всех миноров матрицы A порядка $n - k$, расположенных на пересечении столбцов и строк с одинаковыми номерами.

Доказательство. Чтобы получить коэффициент s_k, нужно среди $n!$ членов определителя

$$\det(A - \lambda I) \ = \ \sum_{\sigma \in S_n} d_\sigma$$

выбрать те и только те члены d_σ, которые содержат произведение ровно k диагональных членов вида $a_{ii} - \lambda$ (они и только они являются многочленами степени k от λ), в каждом из них выделить слагаемое старшей степени вида $(-\lambda)^k c_\sigma$ и просуммировать полученные коэффициенты c_σ. Очевидно, что сумма всех c_σ, отвечающих k диагональным элементам в фиксированных позициях $i_1 < \ldots < i_k$, будет равна минору матрицы A, расположенному на строках и столбцах, дополнительных к строкам и столбцам с номерами $i_1, \ldots, i_k$. $\square$

В частности, $s_{n-1} = a_{11} + \ldots + a_{nn}$ — величина, называемая *следом* матрицы A. Обозначение: $\operatorname{tr} A$. В силу формул Виета, след равен сумме всех собственных значений с учетом кратностей. Заметим также, что $s_0 = \det A$.

При $n \leq 4$ собственные значения (как корни многочлена степени $n \leq 4$) могут быть выражены в радикалах через коэффициенты характеристического многочлена и, следовательно, через элементы матрицы. При $n \geq 5$ таких формул уже не существует (знаменитый результат Абеля, Руффини и Галуа).

29.5 Алгебраическая кратность собственного значения

Кратность собственного значения как корня характеристического многочлена называется его *алгебраической кратностью*. Из основной теоремы алгебры сразу же вытекает следующая

Теорема. *Любая комплексная матрица A порядка n имеет n комплексных собственных значений с учетом алгебраических кратностей.*

29.6 Характеристический многочлен и подобие

Теорема. *Характеристические многочлены подобных матриц совпадают.*

Доказательство. Пусть $B = P^{-1}AP$, где P — обратимая матрица. Тогда

$$\begin{aligned} \det(B - \lambda I) &= \det(P^{-1}AP - \lambda P^{-1}P) = \det(P^{-1}(A - \lambda I)P) \\ &= \det P^{-1} \det P \det(A - \lambda I) = \det(P^{-1}P) \det(A - \lambda I) \\ &= \det(A - \lambda I). \quad \square \end{aligned}$$

Следствие. *Собственные значения и их алгебраические кратности для подобных матриц совпадают.*

Задача. Пусть A и B — квадратные матрицы одного и того же порядка. Докажите, что AB и BA имеют одинаковые характеристические многочлены.

29.7 Приведение к почти треугольной матрице

Таким образом, при вычислении собственных значений матрицы A можно использовать преобразования подобия для перехода к матрице более простого вида, имеющей те же собственные значения.

Например, от A можно перейти к подобной ей *верхней почти треугольной* матрице. Так называется матрица $H = [h_{ij}]$, в которой $h_{ij} = 0$ при $i \geq j + 2$. Такая матрица называется также *верхней хессенберговой*.

Утверждение. *Для произвольной матрицы A порядка n существует невырожденная матрица P такая, что матрица $B = PAP^{-1}$ является верхней почти треугольной.*

Матрицу P можно выбрать в виде $P = P_{n-2} \dots P_1$, где $P_k = Z_k \Pi_k$ — произведение матрицы перестановки Π_k и матрицы модификации строк Z_k.

Доказательство. Если $a_{21} \neq 0$, то $\Pi_1 = I$. Если $a_{21} = 0$, но $a_{i1} \neq 0$ при $i \geq 3$, то 2-ю и i-ю строки следует переставить — с помощью умножения на соответствующую матрицу перестановки Π_1. В случае $a_{21} \neq 0$ с помощью матрицы модификации строк Z_1 исключаем все элементы первого столбца в позициях $(i, 1)$ при $3 \leq i \leq n$. Проиллюстрируем первый шаг для $n = 4$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{-a_{31}}{a_{21}} & 1 & 0 \\ 0 & \frac{-a_{41}}{a_{21}} & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix},$$

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{a_{31}}{a_{21}} & 1 & 0 \\ 0 & \frac{a_{41}}{a_{21}} & 0 & 1 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ 0 & c_{32} & c_{33} & c_{34} \\ 0 & c_{42} & c_{43} & c_{44} \end{bmatrix}.$$

Важно, что при умножении справа на P_1^{-1} элементы первого столбца не изменяются $\Rightarrow$ нули, полученные там ранее, сохранятся.

Второй шаг направлен на получение нулей во втором столбце. Если $c_{32} \neq 0$, то исключение проводится таким образом:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-c_{42}}{c_{32}} & 1 \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ 0 & c_{32} & c_{33} & c_{34} \\ 0 & c_{42} & c_{43} & c_{44} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} \\ d_{21} & d_{22} & d_{23} & d_{24} \\ 0 & d_{32} & d_{33} & d_{34} \\ 0 & 0 & d_{43} & d_{44} \end{bmatrix},$$

$$\begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} \\ d_{21} & d_{22} & d_{23} & d_{24} \\ 0 & d_{32} & d_{33} & d_{34} \\ 0 & 0 & d_{43} & d_{44} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{c_{42}}{c_{32}} & 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ 0 & h_{32} & h_{33} & h_{34} \\ 0 & 0 & h_{43} & h_{44} \end{bmatrix}.$$

В случае $n \geq 5$ точно так же на третьем шаге получаем нули в позициях третьего столбца $(i, 3)$ при $i \geq 5$. И так далее. $\square$

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

29.8 Матрицы Фробениуса

Задача о вычислении собственных значений матрицы сводится к вычислению корней некоторого многочлена (характеристического многочлена данной матрицы). Верно ли обратное? Можно ли задачу о вычислении корней многочлена степени n свести к вычислению собственных значений некоторой матрицы? Ответ положительный. Пусть многочлен имеет вид

$$f(\lambda) \;=\; \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0.$$

Тогда интересующая нас матрица может быть, в частности, такой:

$$A_f \;=\; \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}.$$

Матрица A_f называется *матрицей Фробениуса* или *сопровождающей матрицей* многочлена $f(x)$.

Утверждение. *Характеристический многочлен матрицы Фробениуса A_f для многочлена $f(\lambda)$ имеет вид* $\det(A_f - \lambda I) = (-1)^n f(\lambda)$.

Доказательство. При вычислении определителя $\det(A_f - \lambda I)$ прибавим к первой строке 2-ю строку, умноженную на λ, затем 3-ю строку, умноженную на λ^2, и так далее. Вот что получается при $n = 4$:

$$\det \begin{bmatrix} -\lambda & 0 & 0 & -a_0 \\ 1 & -\lambda & 0 & -a_1 \\ 0 & 1 & -\lambda & -a_2 \\ 0 & 0 & 1 & -a_3 - \lambda \end{bmatrix} = \det \begin{bmatrix} 0 & -\lambda^2 & 0 & -a_0 - a_1\lambda \\ 1 & -\lambda & 0 & -a_1 \\ 0 & 1 & -\lambda & -a_2 \\ 0 & 0 & 1 & -a_3 - \lambda \end{bmatrix}$$

$$= \det \begin{bmatrix} 0 & 0 & -\lambda^3 & -a_0 - a_1\lambda - a_2\lambda^2 \\ 1 & -\lambda & 0 & -a_1 \\ 0 & 1 & -\lambda & -a_2 \\ 0 & 0 & 1 & -a_3 - \lambda \end{bmatrix} = \det \begin{bmatrix} 0 & 0 & 0 & -a_0 - a_1\lambda - a_2\lambda^2 - a_3\lambda^3 - \lambda^4 \\ 1 & -\lambda & 0 & -a_1 \\ 0 & 1 & -\lambda & -a_2 \\ 0 & 0 & 1 & -a_3 - \lambda \end{bmatrix}$$

$$= a_0 + a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + \lambda^4. \qquad \square$$

29.9 Вычисление характеристического многочлена

Как мы знаем, с помощью элементарных преобразований любую квадратную матрицу можно привести к подобной ей верхней почти треугольной матрице H. Поэтому достаточно научиться вычислять характеристический многочлен для H.

Для этого вложим верхнюю почти треугольную матрицу $H - \lambda I$ в верхнюю треугольную матрицу и рассмотрим следующую систему линейных алгебраических уравнений (пусть для простоты $n = 4$):

$$\begin{bmatrix} 1 & h_{11} - \lambda & h_{12} & h_{13} & h_{14} \\ 0 & h_{21} & h_{22} - \lambda & h_{23} & h_{24} \\ 0 & 0 & h_{32} & h_{33} - \lambda & h_{34} \\ 0 & 0 & 0 & h_{43} & h_{44} - \lambda \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f_1(\lambda) \\ f_2(\lambda) \\ f_3(\lambda) \\ f_4(\lambda) \\ f_5(\lambda) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Предположим, что поддиагональные элементы матрицы H отличны от нуля. Тогда матрица коэффициентов данной системы обратима $\Rightarrow$ система имеет единственное решение, в котором, очевидно, $f_k(\lambda)$ будет многочленом степени $n+1-k$ от λ. Согласно правилу Крамера,

$$f_1(\lambda) = \frac{(-1)^{n+1} \det(H - \lambda I)}{h_{21} \dots h_{n+1\,n}}.$$

В данном методе для вычисления всех коэффициентов характеристического многочлена матрицы порядка n выполняется $O(n^3)$ арифметических операций (проверьте!).

Лекция 30

ОСНОВНАЯ ЧАСТЬ

30.1 Одномерные инвариантные подпространства

Пусть L инвариантно относительно $\mathcal{A}$ и $\dim L = 1$. Пусть $x \in L$ и $x \neq 0$. Инвариантность означает, что $\mathcal{A}x = \lambda x$ для некоторого числа λ. В таких случаях λ и $x \neq 0$ называются *собственным значением* и *собственным вектором* оператора $\mathcal{A}$.

Если x — собственный вектор для $\mathcal{A}$, то линейная оболочка $L(x)$ будет инвариантным подпространством размерности 1: $z \in L(x) \quad \Rightarrow \quad z = \alpha x \quad \Rightarrow \quad Az = (\alpha\lambda)x \in L(x)$.

В дальнейшем будем считать, что оператор $\mathcal{A}$ действует на комплексном пространстве размерности n и задан своей матрицей $A \in \mathbb{C}^{n\times n}$ в произвольном фиксированном базисе. Таким образом, можно говорить о подпространствах в $\mathbb{C}^n$, инвариантных относительно умножения на матрицу A (или, короче, относительно матрицы A). Сохраним обозначения L и x для подпространства и столбца из $\mathbb{C}^n$, имеющих смысл упомянутых выше L и x. Мы уже знаем, что собственные значения λ матрицы A и только они суть корни характеристического уравнения $\det(A - \lambda I) = 0$. Из основной теоремы алгебры вытекает, что матрица A (оператор $\mathcal{A}$) имеет комплексное собственное значение. Отсюда получаем нужное нам

Утверждение. *Любая матрица $A \in \mathbb{C}^{n\times n}$ имеет инвариантное подпространство размерности 1.*

30.2 Геометрическая кратность собственного значения

Фиксируем собственное значение λ оператора $\mathcal{A}$ и рассмотрим множество L всех векторов x таких, что $\mathcal{A}x = \lambda x$.

Утверждение. *Множество L является подпространством, инвариантным относительно $\mathcal{A}$.*

Доказательство. Пусть $x, y \in L \ \Rightarrow \ \mathcal{A}x = \lambda x$, $\mathcal{A}y = \lambda y \ \Rightarrow \ \mathcal{A}(\alpha x+\beta y) = \lambda(\alpha x+\beta y) \ \Rightarrow \ \alpha x + \beta y \in L$. Инвариантность L очевидна: если $x \in L$, то $\mathcal{A}x = \lambda x \in L$. $\square$

Определение. Подпространство L называется *собственным подпространством*, а его размерность — *геометрической кратностью* собственного значения Λ.

30.3 Матричное выражение инвариантности

Теорема. *Пусть $L \subset \mathbb{C}^n$ инвариантно относительно $A \in \mathbb{C}^{n\times n}$ и $\dim L = k$. Тогда существуют матрицы $X \in \mathbb{C}^{n\times k}$ и $B \in \mathbb{C}^{k\times k}$ такие, что столбцы X образуют в L*

базис и выполняется равенство $AX = XB$. Характеристический многочлен матрицы B является делителем характеристического многочлена матрицы A.

Доказательство. Образуем X из базисных векторов $x_1, \ldots, x_k$ для L. Инвариантность означает, что Ax_j есть линейная комбинация векторов $x_1, \ldots, x_k$. Определим матрицу B таким образом, что ее j-й столбец b_j содержит коэффициенты данной линейной комбинации. Тогда $Ax_j = Xb_j \;\Rightarrow\; AX = XB$.

Дополним X какими-то столбцами до невырожденной матрицы $\widetilde{X} \in \mathbb{C}^{n\times n}$. Тогда

$$A\widetilde{X} \;=\; \widetilde{X}\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$$

для каких-то блоков C и D. Отсюда

$$\det(A-\lambda I) \;=\; \det(\widetilde{X}^{-1}A\widetilde{X}-\lambda I) \;=\; \det(B-\lambda I_k)\det(D-\lambda I_{n-k}). \qquad \square$$

Следствие. *Геометрическая кратность собственного значения не выше его алгебраической кратности.*

30.4 Сужение оператора на подпространство

Если подпространство L инвариантно относительно оператора $\mathcal{A}$, то можно определить линейный оператор $\mathcal{B} : L \to L$ правилом

$$\mathcal{B}x \;=\; \mathcal{A}x, \quad x \in L.$$

Оператор $\mathcal{A}$ имеет более широкую область определения, чем $\mathcal{B}$. Но $\mathcal{B}$ действует на векторы из L так же, как $\mathcal{A}$ — поэтому его называют *сужением оператора $\mathcal{A}$ на L*. Говорят также, что $\mathcal{A}$ *индуцирует* на L оператор $\mathcal{B}$ и называют $\mathcal{B}$ *индуцированным оператором*.

Если A — матрица оператора $\mathcal{A}$ в каком-то базисе, $x_1, \ldots, x_k$ — базис в L и $X = [x_1, \ldots, x_k]$, то равенство $AX = XB$ означает, что матрица B является матрицей сужения оператора $\mathcal{A}$ на L в базисе $x_1, \ldots, x_k$.

30.5 Инвариантные пространства и сдвиги

Утверждение. *Матрицы A и $A-\lambda I$ имеют общие инвариантные пространства для любого λ.*

Доказательство. Пусть L инвариантно относительно A. Если $x \in L$, то $Ax \in L \;\Rightarrow\; Ax-\lambda x \in L \;\Rightarrow\; L$ инвариантно относительно $A-\lambda I$. Заметим также, что $A = B-\lambda' I$, где $B = A-\lambda I,\;\; \lambda' = -\lambda$. $\square$

30.6 Треугольная форма матрицы

Лемма 1. *Для любой матрицы $A \in \mathbb{C}^{n\times n}$ существует инвариантное пространство размерности $n-1$.*

Доказательство. Мы уже знаем, что образ $\mathrm{im}A$ является инвариантным пространством. Если его размерность равна $n-1$, то все доказано.

Если она равна $k < n-1$, то $\mathrm{im}A$ заведомо принадлежит какому-то более широкому подпространству L размерности $n-1$, притом если $x \in L$, то $Ax \in \mathrm{im}A \subset L$. Значит, L инвариантно относительно A. Если $\dim \mathrm{im}A = n$, то перейдем к матрице $B = A - \lambda I$, где λ — какое-то собственное значение матрицы A. Ясно, что $\dim \ker B \geq 1 \quad \Rightarrow \quad \dim \mathrm{im}B \leq n-1 \quad \Rightarrow \quad B$ имеет инвариантное пространство размерности $n-1$. Оно же инвариантно относительно A. $\square$

Лемма 2. *Пусть L инвариантно относительно $A \in \mathbb{C}^{n\times n}$ и $\dim L = k > 1$. Тогда в L имеется инвариантное относительно A подпространство размерности $k-1$.*

Доказательство. Согласно матричному выражению инвариантности, $AX = XB$, где столбцы X образуют в L базис и $B \in \mathbb{C}^{k\times k}$. По лемме 1, матрица B имеет инвариантное пространство размерности $k-1$. Обозначим его через M и рассмотрим множество N векторов вида Xz, $z \in M$. Конечно, $N \subset \mathbb{C}^n$ есть подпространство размерности $k-1$. При этом $A(Xz) = X(Bz) \quad \Rightarrow \quad N$ инвариантно относительно A. $\square$

Следствие. *Для любой матрицы $A \in \mathbb{C}^{n\times n}$ существует цепочка вложенных подпространств*

$$L_1 \subset \ldots \subset L_n = \mathbb{C}^n,$$

каждое из которых инвариантно относительно A и притом $\dim L_k = k$.

Теорема о верхней треугольной форме. *Любая матрица $A \in \mathbb{C}^{n\times n}$ подобна верхней треугольной матрице.*

Доказательство. Построим базис $x_1, \ldots, x_n$ таким образом, что $L_k = L(x_1, \ldots, x_k)$ (достаточно взять $x_1 \in L_1$, дополнить его до базиса в L_2 вектором x_2, и так далее). Пусть $X = [x_1, \ldots, x_n]$. Тогда Ax_j есть линейная комбинация столбцов $x_1, \ldots, x_j \quad \Rightarrow \quad Ax_j = Xb_j$ для столбца b_j с нулями в позициях ниже j-й. Таким образом, матрица $B = [b_1, \ldots, b_n]$ — верхняя треугольная, и при этом $AX = XB \quad \Rightarrow \quad B = X^{-1}AX$. $\square$

Заметим, что если $B = X^{-1}AX$, то B и A имеют один и тот же характеристический многочлен. Поэтому B и A имеют один и тот же набор n собственных значений с учетом кратностей. Если матрица B треугольная, то ее собственные значения суть элементы главной диагонали.

30.7 Теорема Шура

Пусть $\lambda_1, \ldots, \lambda_n$ — полный набор n собственных значений матрицы $A \in \mathbb{C}^{n\times n}$ с учетом кратностей. Пусть фиксируется произвольная нумерация собственных значений.

Теорема Шура. *Для любой матрицы $A \in \mathbb{C}^n$ с произвольной предписанной нумерацией ее собственных значений $\lambda_1, \ldots, \lambda_n$ существует унитарная матрица $X \in \mathbb{C}^{n\times n}$ такая, что $B = [b_{ij}] = X^*AX$ есть верхняя треугольная матрица с диагональными элементами $b_{ii} = \lambda_i, \;\; i = 1, \ldots, n$.*

Доказательство. Пусть $Ax_1 = \lambda_1 x_1$, $|x_1| = 1$ (длина определяется естественным скалярным произведением). Построим ортонормированный базис $x_1, \ldots, x_n$, начинающийся с вектора x_1, и пусть $X = [x_1, \ldots, x_n]$. Легко проверить, что

$$AX \;=\; X \begin{bmatrix} \lambda_1 & u^\top \\ 0 & B \end{bmatrix}, \quad B \in \mathbb{C}^{(n-1)\times(n-1)}, \quad u \in \mathbb{C}^{n-1}.$$

Заметим, что $\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\ldots(\lambda_n - \lambda) = (\lambda_1 - \lambda)\det(B - \lambda I_{n-1})$. Значит, B имеет собственные значения $\lambda_2, \ldots, \lambda_n$.

Рассуждая по индукции, предположим, что $Y^*BY = T$, где Y — унитарная матрциа порядка $n-1$, а T — верхняя треугольная матрица порядка $n-1$ с диагональными элементами $\lambda_2, \ldots, \lambda_n$. В итоге

$$(\widetilde{Y}^* X^*) A (X\widetilde{Y}) = \begin{bmatrix} \lambda_1 & u^\top Y \\ 0 & T \end{bmatrix}, \qquad \widetilde{Y} = \begin{bmatrix} 1 & 0 \\ 0 & Y \end{bmatrix}.$$

Из унитарности матрицы Y следует, что $\widetilde{Y}$ — унитарная матрица. Матрица $X\widetilde{Y}$ унитарна как произведение унитарных матриц. □

Сформулированная выше теорема о треугольной форме матрицы является, конечно, следствием теоремы Шура. При этом в теореме Шура утверждается больше — треугольная форма с предписанным порядком собственных значений на диагонали достигается преобразованием подобия с помощью унитарной матрицы.

Отметим конструктивный характер приведенного доказательства теоремы Шура. Как только найдены собственное значение λ_1 и отвечающий ему собственный вектор x_1, задача определения остальных собственных значений сводится к аналогичной задаче порядка $n-1$. [1] Такого рода прием понижения размерности иногда называют *дефляцией*.

Задача. Докажите, что для любой комплексной матрицы A порядка 3 существует унитарная матрица Q такая, что матрица $B = Q^*AQ$ является трехдиагональной. (Матрица B называется *трехдиагональной*, если $b_{ij} = 0$ при $|i-j| > 1$.) [2]

30.8 Делители и подпространства

Вследствие матричного выражения инвариантности, любому инвариантному подпространству матрицы A соответствует некоторый делитель ее характеристического многочлена, являющийся характеристическим многочленом сужения A на данное подпространство. Из теоремы Шура легко вывести и обратное.

Теорема о делителях и подпространствах. *Пусть $A \in \mathbb{C}^{n\times n}$ и $f(\lambda) = \det(A - \lambda I)$ — характеристический многочлен. Предположим, что $f(\lambda)$ делится на многочлен $p(\lambda)$ степени k. Тогда A имеет инвариантное подпространство L размерности k такое, что $p(\lambda)$ есть характеристический многочлен сужения A на L.*

Доказательство. Упорядочим корни многочлена $f(\lambda)$ таким образом, что первые k корней будут также корнями делителя $p(\lambda)$. Согласно теореме Шура, существуют X и B такие, что в верхней треугольной матрице B первые k элементов главной диагонали будут корнями $p(\lambda)$. Пусть X_k — прямоугольная матрица, содержащая первые k столбцов X, а B_k — левый верхний блок порядка k в матрице B. Тогда $AX_k = X_kB_k$ и при этом $\det(B_k - \lambda I) = p(\lambda)$. □

[1] Ниоткуда, впрочем, не следует, что собственный вектор матрицы B автоматически соответствует какому-то собственному вектору матрицы A.

[2] Недавно было доказано, что то же верно для любой комплексной матрицы порядка 4 (V. Pati, 2001) и что существуют матрицы порядка 5, которые не приводятся к трехдиагональному виду преобразованием подобия с помощью унитарной матрицы.

Лекция 31

ОСНОВНАЯ ЧАСТЬ

31.1 Многочлены от матрицы

Если $f(\lambda) = a_0 + a_1\lambda + \ldots + a_m\lambda^m$ — многочлен от λ, то для любой квадратной матрицы A имеет смысл выражение

$$f(A) \;\equiv\; a_0 I + a_1 A + \ldots + a_m A^m.$$

Оно называется *многочленом от матрицы* A. [1] Ясно, что $f(A)$ — квадратная матрица того же порядка, что и A.

Если $f(A) = 0$, то говорят, что многочлен $f(\lambda)$ является *аннулирующим многочленом* для A. Пусть A — матрица порядка n. Тогда система матриц I, A, A^2, …, A^{n^2} будет линейно зависимой (почему?) $\Rightarrow$ для любой матрицы порядка n имеется аннулирующий многочлен степени не выше n^2.

В действительности всегда имеется аннулирующий многочлен степени n (мы скоро докажем, что характеристический многочлен для A является аннулирующим). Иногда можно найти аннулирующие многочлены еще меньшей степени. Аннулирующий многочлен минимальной степени называется *минимальным многочленом* для A.

При поиске инвариантных подпространств многочлены от матрицы A интересны тем, что $\ker f(A)$ и $\mathrm{im} f(A)$ всегда инвариантны относительно A (докажите!).

31.2 Корневые пространства

Предположим, что матрица $A \in \mathbb{C}^{n\times n}$ имеет m попарно различных собственных значений λ_1, …, λ_m алгебраической кратности k_1, …, k_m, соответственно. Это означает, что

$$f(\lambda) \equiv \det(A - \lambda I) \;=\; f_1(\lambda)\ldots f_m(\lambda), \quad f_i(\lambda) = (\lambda_i - \lambda)^{k_i}, \;\; 1 \le i \le m; \quad \lambda_i \ne \lambda_j,\; i \ne j.$$

Подпространства $\mathcal{K}_i \equiv \ker f_i(A) = \ker(A - \lambda_i I)^{k_i}$ называются *корневыми пространствами* матрицы A.

Лемма 1. *Корневое пространство $\mathcal{K}_i$ инвариантно относительно A и имеет размерность k_i. Характеристический многочлен сужения A на $\mathcal{K}_i$ есть $f_i(\lambda) = (\lambda_i - \lambda)^{k_i}$. Сужение $A - \alpha I$ на $\mathcal{K}_i$ при $\alpha \ne \lambda_i$ является обратимым оператором.*

Доказательство. Инвариантность: если $f_i(A)x = 0$, то $f_i(A)(Ax) = A(f_i(A)x) = 0$.

По теореме Шура, существует подобная A верхняя трегольная матрица $B = X^{-1}AX$ с элементами

$$b_{jj} = \lambda_i, \quad 1 \le j \le k_i, \qquad b_{jj} \ne \lambda_i, \quad k_i + 1 \le j \le n. \tag{$*$}$$

Очевидно, $C \equiv B - \lambda_i I = X^{-1}(A - \lambda_i I)X \;\;\Rightarrow\;\; C^{k_i} = (B - \lambda_i I)^{k_i} = X^{-1}(A - \lambda_i I)^{k_i}X$. Запишем C в блочном виде

$$C = \begin{bmatrix} P & Q \\ 0 & R \end{bmatrix},$$

где P и R — верхние треугольные матрицы порядка k_i и $n - k_i$. При этом P имеет нулевую главную диагональ $\;\Rightarrow\;\; P^{k_i} = 0$ (проверяется непосредственно: в матрице P^2 к нулевой главной диагонали добавляется еще одна диагональ, в P^3 — еще одна, и так далее).

[1] Многочлен от матрицы A имеет скалярные коэффициенты. Термин *матричный многочлен* обычно используется для обозначения многочлена от λ, коэффициенты которого являются матрицами.

Следовательно,

$$C^{k_i} = \begin{bmatrix} P^{k_i} & \widetilde{Q} \\ 0 & R^{k_i} \end{bmatrix} = \begin{bmatrix} 0 & \widetilde{Q} \\ 0 & R^{k_i} \end{bmatrix},$$

где все диагональные элементы верхнего треугольного блока R^{k_i} порядка $n-k_i$ отличны от нуля. Блок $\widetilde{Q}$ — какой-то блок размеров $k_i \times (n-k_i)$. Независимо от его вида, находим

$$\operatorname{rank} C^{k_i} = n - k_i \quad \Rightarrow \quad \operatorname{rank}(A-\lambda_i I)^{k_i} = n - k_i \quad \Rightarrow \quad \dim\ker(A-\lambda_i I)^{k_i} = k_i.$$

Матрица сужения A на $\mathcal{K}_i$ представляет собой левый верхний блок порядка k_i в матрице $B = X^{-1}AX$. Согласно $(*)$, все элементы его главной диагонали равны λ_i. Чтобы получить матрицу сужения $A-\alpha I$ на $\mathcal{K}_i$, нужно заменить диагональные элементы на $\lambda_i - \alpha$. При $\alpha \neq \lambda_i$ это будет невырожденная матрица. $\square$

Лемма 2. *Если L инвариантно относительно A и сужение A на L имеет своим характеристическим многочленом $f_i(\lambda)$, то $L = \mathcal{K}_i$.*

Доказательство. Пусть $M \in \mathbb{C}^{k_i \times k_i}$ — матрица сужения A на L в каком-то базисе. Согласно теореме Шура, этот базис можно выбрать так, чтобы M была верхней треугольной. Тогда $M - \lambda_i I$ — верхняя треугольная матрица с нулевой главной диагональю $\Rightarrow$ $(M-\lambda_i I)^{k_i} = 0$ $\Rightarrow$ $(A-\lambda_i I)^{k_i} x = 0$ $\forall$ $x \in L$ $\Rightarrow$ $L \subset \mathcal{K}_i$. Поскольку $\dim L = \dim \mathcal{K}_i$, получаем $L = \mathcal{K}_i$. $\square$

31.3 Нильпотентные операторы

Оператор $\mathcal{A} : V \to V$ называется *нильпотентным*, если $\mathcal{A}^k = 0$ для какого-то k. Так же называется матрица A такая, что $A^k = 0$.

Утверждение. *Матрица A порядка n нильтотентна тогда и только тогда, когда ее характеристический многочлен имеет вид* $\det(A-\lambda I) = (-\lambda)^n$.

Доказательство. Пусть $A^k = 0$, $Ax = \lambda x$, $x \neq 0$ $\Rightarrow$ $A^k x = \lambda^k x = 0$ $\Rightarrow$ $\lambda = 0$. Если A имеет собственное значение нуль кратности n, то, по теореме Шура, она подобна верхней треугольной матрице B с нулями на главной диагонали $\Rightarrow$ $B^n = 0$. $\square$

Следствие. *Сужение $A - \lambda_i I$ на корневое пространство $\mathcal{K}_i$ является нильпотентным оператором на $\mathcal{K}_i$.*

31.4 Корневое разложение

Теорема о корневом разложении. *Пусть матрица $A \in \mathbb{C}^{n\times n}$ имеет m попарно различных собственных значений алгебраической кратности $k_1, \ldots, k_m$, а $\mathcal{K}_1, \ldots, \mathcal{K}_m$ — отвечающие им корневые пространства. Тогда $\mathbb{C}^n$ разлагается в прямую сумму*

$$\mathbb{C}^n = \mathcal{K}_1 + \ldots + \mathcal{K}_m. \qquad (*)$$

Доказательство. Докажем, что сумма $\mathcal{K}_1 + \ldots + \mathcal{K}_m$ является прямой. Пусть

$$x_1 + \ldots + x_m = 0, \quad x_i \in \mathcal{K}_i, \quad 1 \le i \le m. \quad \Rightarrow$$

$$(A-\lambda_2 I)^{k_2} \ldots (A-\lambda_m I)^{k_m}(x_1 + \ldots + x_m) = (A-\lambda_2 I)^{k_2} \ldots (A-\lambda_m I)^{k_m} x_1 = 0.$$

Здесь мы используем то, что любые матричные многочлены от A коммутируют. В силу Леммы 1, сужение каждой из матриц $(A-\lambda_i I)^{k_i}$, $2 \le i \le m$, на $\mathcal{K}_1$ является обратимым оператором $\Rightarrow$ $x_1 = 0$. Аналогично доказывается, что $x_2 = \ldots = x_m = 0$. Остается учесть, что

$$\dim \mathcal{K}_1 + \ldots + \dim \mathcal{K}_m = n. \qquad \square$$

Разложение $(*)$ иногда называется *корневым разложением* матрицы A.

Пусть A рассматривается как матрица линейного оператора $\mathcal{A} : V_n \to V_n$ на комплексном n-мерном пространстве V_n. Собственные значения λ_i и их алгебраические кратности k_i не зависят от выбора базиса для представления оператора $\mathcal{A}$. Под корневыми пространствами оператора $\mathcal{A}$ понимаются

подпространства $\ker(\mathcal{A} - \lambda_i I)^{k_i} \subset V_n$ (здесь I — тождественный оператор). Полученной нами теореме можно дать и операторную формулировку.

Операторная формулировка теоремы о корневом разложении. *Сумма m корневых пространств оператора $\mathcal{A}$ является прямой и совпадает с V_n:*

$$V_n = \ker(\mathcal{A} - \lambda_1 I)^{k_1} \ + \ \ldots \ + \ \ker(\mathcal{A} - \lambda_m I)^{k_m}$$

31.5 Блочно диагональная форма матрицы

Согласно теореме о корневом разложении, базис в $\mathbb{C}^n$ можно выбрать как объединение базисов в корневых пространствах $\mathcal{K}_i$, $1 \le i \le m$. Пусть этот базис представлен столбцами матрицы X. Тогда, вследствие теоремы о корневом разложении,

$$X^{-1}AX \ = \ \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_m \end{bmatrix}.$$

Порядок блока B_i равен алгебраической кратности собственного значения λ_i.

Заметим, что, в силу теоремы Шура, X можно выбрать таким образом, что каждый блок B_i будет верхней треугольной матрицей.

31.6 Теорема Гамильтона–Кэли

Теорема Гамильтона–Кэли. *Пусть $A \in \mathbb{C}^{n\times n}$ — произвольная матрица и $f(\lambda) = \det(A - \lambda I)$ — ее характеристический многочлен. Тогда $f(A) = 0$.*

Доказательство. Пусть имеется m попарно различных собственных значений $\lambda_1, \ldots, \lambda_m$ алгебраической кратности $k_1, \ldots, k_m$. Тогда

$$f(A) \ = \ (A - \lambda_1 I)^{k_1} \ldots (A - \lambda_m I)^{k_m}.$$

Любой вектор $x \in \mathbb{C}^n$ имеет вид $x = x_1 + \ldots + x_m$, где $(A - \lambda_i I)^{k_i} x_i = 0$. Остается заметить, что матрицы $(A - \lambda_i I)^{k_i}$ и $(A - \lambda_j I)^{k_j}$ коммутируют. $\square$

Замечание. При доказательстве теоремы Гамильтона–Кэли было использовано каноническое разложение комплексного многочлена (характеристического многочлена матрицы A) и связанный с ним результат о расщеплении $\mathbb{C}^n$ в прямую сумму корневых пространств матрицы A. Однако, характеристический многочлен имеет смысл для матрицы над любым полем, причем это будет многочлен с коэффициентами именно из этого поля. В общем случае, правда, он может не иметь ни одного корня в заданном поле. Тем не менее, теорема Гамильтона–Кэли остается справедливой и в общем случае.

Лекция 32

ОСНОВНАЯ ЧАСТЬ

32.1 Минимальное инвариантное подпространство

Попробуем сделать более специальный выбор базиса в корневом пространстве $\mathcal{K}_i$, позволяющий расщепить $\mathcal{K}_i$ в прямую сумму инвариантных подпространств с максимально возможным числом слагаемых.

Поскольку инвариантные подпространства не меняются при сдвиге, их можно строить для $B = A - \lambda_i I$. Если A имеет попарно различные собственные значения $\lambda_1, \ldots, \lambda_m$, то B получает попарно различные собственные значения $\mu_1 = \lambda_1 - \lambda_i, \ldots, \mu_m = \lambda_m - \lambda_i$ с теми же алгебраическими кратностями. В частности, B имеет собственное значение $\mu_i = 0$ алгебраической кратности k_i.

Предположим, что $\mathcal{L} \subset \mathcal{K}_i$ инвариантно относительно B, и пусть $x \neq 0$, $x \in \mathcal{L}$. Тогда $\mathcal{L}$ содержит все векторы вида $x,\ Bx,\ B^2x,\ \ldots$. Поскольку $\mathcal{K}_i = \ker B^{k_i}$, заключаем, что

$$B^l x = 0 \ \text{ при } \ l \geq k_i.$$

Обозначим через $k = k(x)$ наименьший номер такой, что $B^k x = 0$. Будем называть k *высотой вектора* x в корневом пространстве $\mathcal{K}_i$.

Лемма о минимальном инвариантном подпространстве. *Пусть $x \in \mathcal{K}_i$ — вектор высоты k. Тогда*

$$L_k = L(x,\, Bx,\, \ldots,\, B^{k-1}x) \subset \mathcal{K}_i$$

является наименьшим инвариантным подпространством, содержащим x. При этом векторы x, Bx, $\ldots$, $B^{k-1}x$ линейно независимы.

Доказательство. Инвариантность очевидна. Пусть

$$\alpha_1 x + \alpha_2 Bx + \ldots + \alpha_k B^{k-1}x = 0. \qquad (\#)$$

Умножив обе части слева на B^{k-1}, находим $B^{k-1}Bx = B^{k-1}B^2x = \ldots = B^{k-1}B^{k-1}x = 0 \ \Rightarrow \ \alpha_1 B^{k-1}x = 0 \ \Rightarrow \ \alpha_1 = 0$. Далее, умножив обе части $(\#)$ слева на B^{k-2}, находим $\alpha_2 = 0$, и так далее. Таким образом, $\dim L_k = k$. $\square$

32.2 Жордановы цепочки

Занумеруем векторы $x, Bx, \ldots, B^{k-1}x$ в обратном порядке: $x_1 = B^{k-1}x$, $x_2 = B^{k-2}x$, $\ldots$, $x_{k-1} = Bx$, $x_k = x$. Тогда

$$Bx_1 = 0, \quad Bx_j = x_{j-1}, \ \ 2 \leq j \leq k. \qquad (*)$$

Система векторов $x_1, \ldots, x_k$, обладающих свойствами $(*)$, называется *жордановой цепочкой длины k, начинающейся с вектора x_1*. В силу определения B, равенства $(*)$ эквивалентны равенствам

$$Ax_1 = \lambda_i x_1, \quad Ax_j = \lambda_i x_j + x_{j-1}, \ \ 2 \leq j \leq k. \qquad (**)$$

Пусть $X = [x_1, \ldots, x_k]$ и J_k — матрица порядка k, определенная равенством (матрица сужения A на L_k в базисе $x_1, \ldots, x_k$)

$$AX = XJ_k.$$

В силу (**),

$$J_k = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}.$$

Матрица вида J_k называется *жордановой клеткой* (жордановым блоком, жордановым ящиком), отвечающей собственному значению λ_i.

32.3 Жорданова форма матрицы

Подпространство L_k, натянутое на жорданову цепочку векторов вида (*) или (**), иногда называется *циклическим подпространством* в $\mathcal{K}_i$, отвечающим собственному значению λ_i. Наша ближайшая цель — показать, что $\mathcal{K}_i$ можно представить в виде прямой суммы циклических подпространств.

Тогда, объединив базисы циклических подпространств, получаем в $\mathcal{K}_i$ такой базис, в котором матрица сужения A на $\mathcal{K}_i$ имеет блочно диагональный вид, где каждый блок есть жорданова клетка. Сделав то же для каждого корневого пространства, в результате объединения базисов всех циклических подпространств получаем так называемый *жорданов базис*: в нем матрица A получает свою *жорданову форму* — становится блочно диагональной матрицей, в которой каждый блок главной диагонали является жордановой клеткой для какого-то ее собственного значения.

Матрица J блочно диагонального вида с блоками $J_1, \ldots, J_N$ называется *прямой суммой* своих блоков $J_1, \ldots, J_N$. Обозначение:

$$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_N \end{bmatrix} = J_1 \oplus \ldots \oplus J_N.$$

В этой терминологии жорданова форма представляет собой прямую сумму жордановых клеток.

32.4 Индекс собственного значения

Очевидно, $\ker B \subset \ker B^2 \subset \ldots$. В конечномерном пространстве подпространства не могут расширяться неограниченно, поэтому для некоторой степени $\ker B^k = \ker B^{k+1}$. Минимальный номер k с таким свойством называется *индексом собственного значения* λ_i (напомним, что $B = A - \lambda_i I$).

Утверждение. *Если* $\ker B^k = \ker B^{k+1}$, *то* $\ker B^l = \ker B^{l+1}$ *при всех* $l \geq k$.

Доказательство. Пусть $x \in \ker B^{l+1} \quad \Rightarrow \quad B^{k+1}(B^{l-k}x) = 0 \quad \Rightarrow \quad B^k(B^{l-k}x) = 0 \quad \Rightarrow \quad x \in \ker B^l$. □

Следствие. *Индекс не больше алгебраической кратности данного собственного значения.*

Достаточно учесть, что $k \leq \dim \ker B^k$ и $\ker B^k = \ker B^{k+1} = \ldots = \ker B^{k_i}$.

32.5 Жорданов базис в корневом пространстве

Пусть k — индекс λ_i. Тогда $s \equiv \dim \ker B^k - \dim \ker B^{k-1} > 0$. Поэтому существуют s линейно независимых векторов $x_1, \ldots, x_s$, дополняющих какой-нибудь базис в $\ker B^{k-1}$ до базиса в $\ker B^k$:

$$\ker B^k = \ker B^{k-1} + L(x_1, \ldots, x_s).$$

(1) Векторы $x_1, \ldots, x_s$ имеют высоту k и порождают циклические подпространства

$$L_{1i} = L(x_i, Bx_i, \ldots, B^{k-1}x_i), \quad 1 \leq i \leq s.$$

(2) Сумма $L_{11} + \ldots + L_{1s}$ является прямой, поскольку векторы

$$x_1, Bx_1, \ldots, B^{k-1}x_1, \quad \ldots, \quad x_s, Bx_s, \ldots, B^{k-1}x_s$$

линейно независимы. В самом деле, пусть

$$\sum_{i=1}^{s} \sum_{j=1}^{k} \alpha_{ij} B^{j-1} x_i = 0.$$

Умножив обе части слева на B^{k-1}, находим

$$\sum_{i=1}^{s} \alpha_{i1} B^{k-1} x_i = 0 \quad \Rightarrow \quad \sum_{i=1}^{s} \alpha_{i1} x_i \in \ker B^{k-1} \quad \Rightarrow \quad \alpha_{i1} = 0, \quad 1 \le i \le s.$$

Умножив затем обе части слева на B^{k-2}, по той же причине получим $\alpha_{i2} = 0,\ 1 \le i \le s$, и так далее.

(3) Сумма $\ker B^{k-2} + L(Bx_1, \ldots, Bx_s)$ является прямой.

(4) Если она не совпадает с $\ker B^{k-1}$, то найдутся t линейно независимых векторов $y_1, \ldots, y_t$ таких, что

$$\ker B^{k-1} = \ker B^{k-2} + L(Bx_1, \ldots, Bx_s,\ y_1, \ldots, y_t),$$

причем сумма является прямой.

(5) Векторы $y_1, \ldots, y_t$ имеют высоту $k-1$ и порождают циклические подпространства

$$L_{2i} = L(y_i,\ By_i,\ \ldots,\ B^{k-2} y_i), \quad 1 \le i \le t.$$

(6) Сумма $L_{11} + \ldots + L_{1s} \ + \ L_{21} + \ldots + L_{2t}$ является прямой. Доказательство аналогично доказательству предложения **(2)**.

(7) Сумма $\ker B^{k-3} + L(B^2 x_1, \ldots, B^2 x_s,\ By_1, \ldots,\ By_t)$ является прямой.

(8) Если она не совпадает с $\ker B^{k-2}$, действуем по аналогии с шагом **(4)**.

И так далее.

Для наглядности построенные векторы расположим в виде следующей таблицы:

$$\begin{array}{cccccccccc}
x_1 & \ldots & x_s & & & & & & & \\
Bx_1 & \ldots & Bx_s & y_1 & \ldots & y_t & & & & \\
B^2x_1 & \ldots & B^2x_s & By_1 & \ldots & By_t & & & & \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & & & & \\
B^{k-1}x_1 & \ldots & B^{k-1}x_s & B^{k-2}y_1 & \ldots & B^{k-2}y_t & \ldots & z_1 & \ldots & z_r
\end{array}$$

Векторы последней строки образуют базис в ядре $\ker B$. Это собственные векторы, отвечающие собственному значению λ_i. Подпространство $\ker B = \ker(A - \lambda_i I)$ называется *собственным подпространством* для λ_i, а его размерность — *геометрической кратностью* собственного значения λ_i. По построению, общее число векторов таблицы равно алгебраической кратности λ_i.

Утверждение. *Все векторы указанной таблицы линейно независимы и образуют базис в $\mathcal{K}_i$.*

Доказательство. Рассмотрим равную нулю линейную комбинацию всех векторов таблицы. Умножив ее слева на B^{k-1}, заметим, что все векторы, кроме первой строки, обращаются в нуль. Остается лишь линейная комбинация векторов верхней строки, которую матрица B^{k-1} переводит в нуль. Вывод: линейная комбинация векторов верхней строки принадлежит $\ker B^{k-1}$. Значит, коэффициенты при векторах верхней строки равны нулю. С помощью умножения на B^{k-2} находим, что линейная комбинация векторов второй сверху строки принадлежит $\ker B^{k-2}$. Поэтому соответствующие коэффициенты равны нулю. И так далее. $\square$

Векторы каждого столбца данной таблицы образуют базис циклического подпространства. Соответствующие жордановы цепочки получаются при нумерации их в каждом столбце снизу вверх.

32.6 Существование и единственность жордановой формы

Теорема. *Любая матрица $A \in \mathbb{C}^{n\times n}$ подобна прямой сумме жордановых клеток*

$$J = J_1 \oplus \ldots \oplus J_N,$$

где число и размеры жордановых клеток для каждого собственного значения определяются однозначно по матрице A.

Доказательство. Мы только что установили, что корневое пространство $\mathcal{K}_i$ есть прямая сумма циклических подпространств. Каждый столбец полученной выше таблицы отвечает одной жордановой клетке. Из этой же таблицы можно найти число жордановых клеток заданного порядка.

Обозначим через m_j число жордановых клеток для λ_i порядка j. Заметим, в частности, что $m_k = s$, $m_{k-1} = t$ и $m_1 = r$. В общем случае

$$\begin{array}{rcl}
m_k & = & \dim \ker B^k - \dim \ker B^{k-1}, \\
m_{k-1} + m_k & = & \dim \ker B^{k-1} - \dim \ker B^{k-2}, \\
 & \ldots & \\
m_1 + \ldots + m_{k-1} + m_k & = & \dim \ker B.
\end{array}$$

Отсюда находим

$$m_j = 2\dim\ker B^j - \dim\ker B^{j-1} - \dim\ker B^{j+1}, \quad 1 \le j \le k.$$

Следовательно, число и порядки жордановых клеток для λ_i определяются размерностями ядер $\ker(A-\lambda_i I)^j$, а значит, и рангами матриц $(A-\lambda_i I)^j$. То же верно для жордановых клеток каждого корневого пространства. □

Следствие. *Матрицы подобны тогда и только тогда, когда они имеют одинаковую жорданову форму с точностью до перестановки жордановых клеток.*

32.7 Минимальный многочлен матрицы

По теореме Гамильтона—Кэли, матрица $A \in \mathbb{C}^{n\times n}$ аннулируется своим характеристическим многочленом: если $f(\lambda) = \det(A-\lambda I)$, то $f(A) = 0$. Многочлен минимальной степени с тем же свойством называется *минимальным многочленом* матрицы A.

Лемма. *Минимальный многочлен является делителем характеристического многочлена.*

Доказательство. Пусть $\phi(\lambda)$ и $f(\lambda)$ — минимальный и характеристический многочлены для A. Выполним деление с остатком: $f(\lambda) = q(\lambda)\phi(\lambda)+r(\lambda)$. Очевидно, $r(A) = 0$. Неравенство $\deg r(\lambda) < \deg\phi(\lambda)$ противоречило бы минимальности многочлена $\phi(\lambda)$. Поэтому $r(\lambda)$ — нулевой многочлен. □

Теорема. *Пусть A имеет попарно различные собственные значения $\lambda_1, \ldots, \lambda_m$. Степень минимального многочлена матрицы A равна сумме $n_1+\ldots+n_m$, где n_i — максимальный порядок жордановых клеток для собственного значения λ_i.*

Доказательство. Достаточно рассмотреть разложение произвольного вектора $x = \sum x_j$ по циклическим подпространствам L_j (последние в прямой сумме дают $\mathbb{C}^n$). Пусть подпространства $L_{j_1}, \ldots, L_{j_m}$ отвечают, соответственно, $\lambda_1, \ldots, \lambda_m$ и имеют размерности $n_1, \ldots, n_m$. Тогда $\ker(A-\lambda_i I)^{n_i} = \mathcal{K}_i \;\Rightarrow$

$$(A-\lambda_1 I)^{n_1} \ldots (A-\lambda_m I)^{n_m}\; x \;=\; 0.$$

Таким образом, степень минимального многочлена не выше $n_1+\ldots+n_m$.

В то же время, степень минимального многочлена не может быть меньше: жорданова клетка порядка n_i для λ_i не может быть аннулирована многочленом степени меньше n_i, при этом ее минимальный многочлен есть в точности $(\lambda_i-\lambda)^{n_i}$ и этот многочлен не может аннулировать ни одну из жордановых клеток, отвечающих другому собственному значению. □

32.8 Вычисление жордановой формы

ПРИМЕР 1. Выяснить диагонализуемость матрицы

$$A = \begin{bmatrix} -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

В силу блочно диагонального вида заданной матрицы, ее собственные значения можно искать по отдельности для блоков $A_1 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$ и $A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Матрица A имеет собственные значения $\lambda = 2$ кратности 1 и $\lambda = 0$ кратности 3.

Собственный вектор для $\lambda = 2$ есть нетривиальное решение системы $(A-2\cdot I)x = 0$. Ранг матрицы коэффициентов равен 3, поэтому фундаментальная система решений состоит из одного вектора. Собственные векторы для $\lambda = 0$ — это нетривиальные решения системы $(A-0\cdot I)x = 0$. В данном случае ранг матрицы коэффициентов равен 2, поэтому в фундаментальной системе 2 вектора $\Rightarrow$ имеется система ровно из двух линейно независимых собственных векторов для $\lambda = 0$. Таким образом, базиса из собственных векторов не существует, поэтому матрица A не может быть подобна диагональной матрице.

ПРИМЕР 2. Найти жорданову форму и соответствующий жорданов базис для

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Данная матрица имеет собственное значение $\lambda = 1$ кратности 4. Все пространство $\mathbb{C}^4$ является корневым для собственного значения $\lambda = 1$. С помощью сдвига перейдем к матрице $B = A - 1 \cdot I$ и поинтересуемся ее степенями:

$$B = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad B^2 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad B^3 = 0.$$

Значит, число жордановых клеток порядка 3 равно $\dim \ker B^3 - \dim \ker B^2 = 4 - 3 = 1$. Имеется также одна жорданова клетка порядка 1.

Принадлежность вектора $x = [x_1, x_2, x_3, x_4]^\top$ ядру $\ker B^2$ описывается уравнением $x_4 = 0$. Поэтому, взяв $x = [0, 0, 0, 1]^\top$, получаем прямую сумму

$$\ker B^3 = \ker B^2 + L(x).$$

Вектор x имеет высоту 3 и порождает циклическое подпространство, натянутое на векторы

$$x, \;\; Bx = [0, 0, 1, 0]^\top, \;\; B^2 x = [0, 1, 0, 0]^\top.$$

Принадлежность вектора $z = [z_1, z_2, z_3, z_4]^\top$ собственному подпространству $\ker B$ описывается системой уравнений $z_3 = z_4 = 0$. Поэтому, например, $z = [1, 0, 0, 0]^\top$ является собственным вектором, линейно независимым с уже найденным собственным вектором $B^2 x = [0, 1, 0, 0]^\top$. Окончательно,

$$J = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \qquad X = [B^2 x,\, Bx,\, x,\, z].$$

Обратим внимание на то, что жорданова форма J и матрица X жорданова базиса *должны соответствовать* друг другу: $AX = XJ$. Это значит, что даже из правильно найденных векторов имеется возможность составить неправильную матрицу X (за счет неверной их нумерации).

ПРИМЕР 3. Нильпотентная матрица J порядка $n = 10$ имеет две жордановы клетки порядка 3 и две жордановы клетки порядка 2. Требуется вычислить жорданову форму матрицы $A = J^2$.

Нильпотентность означает, что J имеет собственное значение $\lambda = 0$ кратности 10. То же верно и для матрицы $A = J^2$. Вычисляем размерности ядер: $\dim \ker A = 8$, $\dim \ker A^2 = 10$. Следовательно, жорданова форма матрицы A состоит из $m_2 = 10 - 8 = 2$ клеток порядка 2 и $m_1 = 2 \cdot 8 - 10 = 6$ клеток порядка 1.

Задача. Известно, что $A^{k+1} = A$. Докажите, что матрица A диагонализуема.

32.9 Инвариантные подпространства для вещественных матриц

Если матрица A порядка n вещественная, то можно потребовать, чтобы ее инвариантные подпространства выбирались только в $\mathbb{R}^n$. При данном ограничении может не найтись ни одного инвариантного подпространства размерности 1 (приведите пример!). Тем не менее, справедливо

Утверждение. *Матрица $A \in \mathbb{R}^{n \times n}$ имеет инвариантное подпространство $L \subset \mathbb{R}^n$ размерности 2.*

Доказательство. Пусть $\lambda = a + \mathbf{i}b$ — собственное значение с мнимой частью $b \neq 0$. Представим собственный вектор для λ в виде $x + \mathbf{i}y$, где $x,\; y \in \mathbb{R}^n$. Тогда

$$A(x + \mathbf{i}y) = (a + \mathbf{i}b)(x + \mathbf{i}y) \quad \Rightarrow \quad Ax = ax - by, \;\; Ay = bx + ay.$$

Отсюда получаем также, что $A(x - \mathbf{i}y) = (a - \mathbf{i}b)(x - \mathbf{i}y)$. Векторы $x + \mathbf{i}y$ и $x - \mathbf{i}y$ линейно независимы, так как отвечают разным собственным значениям матрицы A. Пусть $\alpha x + \beta y = 0 \;\; \Rightarrow$ $(\alpha - \mathbf{i}\beta)(x + \mathbf{i}y) + (\alpha + \mathbf{i}\beta)(x - \mathbf{i}y) = 2(\alpha x + \beta y) = 0 \;\; \Rightarrow \;\; \alpha - \mathbf{i}\beta = \alpha + \mathbf{i}\beta = 0 \;\; \Rightarrow \;\; \alpha = \beta = 0$. Следовательно, линейная оболочка $L(x, y)$ является двумерным инвариантным подпространством относительно A. Если комплексных собственных значений нет, то базис, очевидно, можно составить из вещественных векторов. $\square$

32.10 Вещественный аналог жордановой формы

Пусть $A \in \mathbb{R}^{n\times n}$ имеет жорданову клетку J порядка k для комплексного собственного значения $\lambda = a + \mathbf{i}b$ с мнимой частью $b \neq 0$. Это означает существование жордановой цепочки

$$Av_1 = \lambda v_1, \quad Av_j = \lambda v_j + v_{j-1}, \quad 2 \le j \le k.$$

Представим каждый вектор v_j в виде $v_j = x_j + \mathbf{i}y_j$, где $x_j,\ y_j \in \mathbb{R}^n$. Тогда находим

$$A[x_1, y_1,\ x_2, y_2,\ \ldots,\ x_k, y_k] = [x_1, y_1,\ x_2, y_2,\ \ldots,\ x_k, y_k]M_{2k},$$

где

$$M_{2k} = \begin{bmatrix} a & b & 1 & 0 & & & & \\ -b & a & 0 & 1 & & & & \\ & & a & b & 1 & 0 & & \\ & & -b & a & 0 & 1 & & \\ & & & \ddots & \ddots & & & \\ & & & & \ddots & \ddots & & \\ & & & & a & b & 1 & 0 \\ & & & & -b & a & 0 & 1 \\ & & & & & & a & b \\ & & & & & & -b & a \end{bmatrix} \in \mathbb{R}^{(2k)\times(2k)}. \qquad (*)$$

Заметим, что линейная оболочка $L(x_1, y_1,\ \ldots,\ x_k, y_k) \subset \mathbb{R}^n$ является инвариантным подпространством размерности $2k$, совпадающим с прямой суммой двух корневых пространств матрицы A — для собственного значения $\lambda = a + \mathbf{i}b$ и сопряженного собственного значения $\overline{\lambda} = a - \mathbf{i}b$ (в силу вещественности коэффициентов характеристического многочлена, λ и $\overline{\lambda}$ оба являются собственными значениями матрицы A одинаковой кратности). Из сказанного вытекает

Теорема. *Любая матрица $A \in \mathbb{R}^{n\times n}$ с помощью вещественного преобразования подобия приводится к прямой сумме вещественных жордановых блоков и вещественных блоков вида* $(*)$.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

32.11 Прямое доказательство по индукции

Путь к теореме о приведении квадратной комплексной матрицы к жордановой форме, очевидно, потребовал от нас изрядных усилий. Поэтому естественно желание как-то его “срезать” — в какой-то степени это удается с помощью следующего рассуждения, предложенного А. Ф. Филипповым.

Теорема. *Пусть L инвариантно относительно $A \in \mathbb{C}^{n\times n}$ и сужение A на L имеет единственное собственное значение λ кратности k. Тогда существует цепочка линейно независимых векторов $x_1, \ldots, x_k \in L$ такая, что для каждого j*

$$Ax_j = \lambda x_j \quad \textit{либо} \quad Ax_j = \lambda x_j + x_{j-1}.$$

Доказательство. Перейдем к матрице $B = A - \lambda I$ и будем доказывать существование цепочки со свойствами

$$Bx_j = 0 \quad \text{либо} \quad Bx_j = x_{j-1}.$$

При $k = 1$ это очевидно (в данном случае L — одномерное инвариантное подпространство). Рассуждая по индукции, предположим, что в случае, когда размерность инвариантного подпространства равна $r < k$, цепочка нужного вида существует. В качестве такого пространства возьмем $\mathrm{im}B \cap L$. Ясно, что $r \equiv \dim(\mathrm{im}B \cap L) < k$.

Итак, по индуктивному предположению, имеется цепочка линейно независимых векторов $y_1, \ldots, y_r$ таких, что

$$By_j = 0 \quad \text{либо} \quad By_j = y_{j-1}.$$

Ясно, что система $y_1, \ldots, y_k$ разбивается ка конечное число жордановых цепочек:

$$y_{i_1}, \ldots, y_{j_1};\ \ \ldots\ ;\ \ y_{i_l}, \ldots, y_{j_l}.$$

Таким образом, жордановых цепочек всего l, а векторы $y_{i_1}, \ldots, y_{i_l}$ и $y_{j_1}, \ldots, y_{i_l}$ — начальные и конечные векторы этих цепочек.

Все векторы, и в частности, конечные векторы жордановых цепочек, принадлежат $\mathrm{im}B$. Поэтому найдутся векторы $w_1, \ldots, w_l$ такие, что

$$Bw_1 = y_{j_1}, \ \ldots, \ Bw_l = y_{j_l}.$$

Заметим, что $w_1, \ldots, w_l \in \ker B^{r+1} \cap L$.

Начальные векторы жордановых цепочек линейно независимы (как часть линейно независимой системы) и принадлежат подпространству $\ker B \cap L$, но, возможно, их недостаточно для того, чтобы составить его базис. Пусть векторы $z_1, \ldots, z_s$ дополняют систему $y_{i_1}, \ldots, y_{i_l}$ до базиса в подпространстве $\ker B \cap L$.

Заметим, что $\dim L = \dim(\mathrm{im}B \cap L) + \dim(\ker B \cap L)$. Таким образом, цепочка векторов

$$z_1, \ldots, z_s, \quad y_{i_1}, \ldots, y_{j_1}, w_1, \quad \ldots \ , \quad y_{i_l}, \ldots, y_{j_l}, w_l \qquad (*)$$

имеет нужный вид, и в ней ровно $\dim L = r + (l + s)$ векторов. Остается лишь доказать, что система $(*)$ линейно независима. Запишем

$$\Big(\sum \alpha_i z_i\Big) + \Big(\sum \beta_i y_i\Big) + \Big(\sum \gamma_i w_i\Big) = 0.$$

Умножив обе части слева на B, получаем равную нулю линейную комбинацию части векторов y_i — без начальных векторов жордановых цепочек $y_{i_1}, \ldots, y_{i_l}$. Отсюда находим, что $\gamma_i = 0$ для всех $1 \le i \le l$ и $\beta_i = 0$ для всех $1 \le i \le r$, кроме $i = i_1, \ldots, i_l$. Таким образом,

$$\left(\sum_{i=1}^{s} \alpha_i z_i\right) + \left(\sum_{t=1}^{l} \beta_{i_t} y_{i_t}\right) = 0.$$

Данная система линейно независима по построению $\Rightarrow$ все α_i и β_{i_t} равны нулю. $\square$

Лекция 33

ОСНОВНАЯ ЧАСТЬ

33.1 Нормальные матрицы

Основу матричной техники составляют преобразования и разложения матриц общего вида, получаемые при помощи специальных классов матриц.

Квадратная комплексная матрица A называется *нормальной*, если $A^*A = AA^*$.

Теорема. *Матрица $A \in \mathbb{C}^{n\times n}$ нормальная тогда и только тогда, когда для некоторой унитарной матрицы $Q \in \mathbb{C}^{n\times n}$ матрица Q^*AQ является диагональной.*

Доказательство. По теореме Шура, существует унитарная матрица Q, приводящая A к верхнему треугольному виду $B = QAQ^*$. Равенство $A^*A = AA^*$ равносильно равенству $B^*B = BB^*$. Остается посмотреть, что оно означает в случае верхней треугольной матрицы B:

$$\begin{bmatrix} \bar{b}_{11} & & & \\ \bar{b}_{12} & \bar{b}_{22} & & \\ \dots & \dots & \dots & \\ \bar{b}_{1n} & \bar{b}_{2n} & \dots & \bar{b}_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ & b_{22} & \dots & b_{2n} \\ & & \dots & \dots \\ & & & b_{nn} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ & b_{22} & \dots & b_{2n} \\ & & \dots & \dots \\ & & & b_{nn} \end{bmatrix} \begin{bmatrix} \bar{b}_{11} & & & \\ \bar{b}_{12} & \bar{b}_{22} & & \\ \dots & \dots & \dots & \\ \bar{b}_{1n} & \bar{b}_{2n} & \dots & \bar{b}_{nn} \end{bmatrix}.$$

Приравнивая элементы в позиции $(1,1)$, получаем

$$|b_{11}|^2 = |b_{11}|^2 + |b_{12}|^2 + \ldots + |b_{1n}|^2 \quad \Rightarrow \quad b_{12} = \ldots = b_{1n} = 0.$$

Учитывая это, приравниваем элементы в позиции $(2,2)$:

$$|b_{22}|^2 = |b_{22}|^2 + |b_{23}|^2 + \ldots + |b_{2n}|^2 \quad \Rightarrow \quad b_{23} = \ldots = b_{2n} = 0.$$

И так далее. Вывод такой: верхняя треугольная матрица является нормальной тогда и только тогда, когда она является диагональной. Значит, равенство $A^*A = AA^*$ выполняется в том и только том случае, когда B — диагональная матрица. $\square$

Следствие. *Матрица является нормальной в том и только том случае, когда она обладает ортонормированным базисом из собственных векторов.*

Пусть $\Lambda = Q^*AQ$ —- диагональная матрица. Столбцы унитарной матрицы Q образуют ортонормированный базис и, в силу равенства $AQ = Q\Lambda$, являются собственными векторами матрицы A. $\square$

Как видим, любая нормальная матрица подобна диагональной, причем преобразование подобия реализуется с помощью унитарной матрицы. В таких случаях говорят об *унитарном подобии*.

Если $A^* = f(A)$ для некоторого многочлена $f(\lambda)$, то матрица A, очевидно, нормальная. Верно и обратное. Предположим, что A имеет m попарно различных собственных значений $\lambda_1, \ldots, \lambda_m$ и возьмем в качестве $f(\lambda)$ многочлен степени не выше $m-1$, принимающий при λ_i значение $\overline{\lambda}_i$. Тогда $\Lambda^* = f(\Lambda) \quad \Rightarrow \quad A^* = Q\Lambda^*Q^* = Qf(\Lambda)Q^* = f(A)$.

33.2 Унитарные матрицы

Пусть A — нормальная матрица, $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = Q^*AQ$ — диагональная матрица из ее собственных значений и Q — унитарная матрица из ее собственных векторов.

Напомним, что квадратная матрица A называется *унитарной*, если $A^*A = I$. Из определения ясно, что любая унитарная матрица является нормальной.

Утверждение. *Нормальная матрица является унитарной тогда и только тогда, когда все ее собственные значения по модулю равны 1.*

Доказательство. $A^*A = 1 \quad \Leftrightarrow \quad \Lambda^*\Lambda = I \quad \Leftrightarrow \quad |\lambda_i| = 1,\ 1 \le i \le n.$ $\square$

33.3 Матрицы отражения и вращения

Унитарные матрицы занимают, бесспорно, особое место в вычислительной алгебре: во-первых, они задают ортонормированные базисы; во-вторых, при умножении на них сохраняются длины столбцов (и даже их скалярные произведения). Среди них выделяются два очень полезных для вычислений подкласса: матрицы отражения и матрицы вращения.

Матрицей отражения (матрицей Хаусхолдера), порожденной вектором $v \in \mathbb{C}^n$ единичной длины, называется матрица вида

$$H = H(v) \ = \ I - 2vv^*, \qquad |v| = 1.$$

Очевидно, $H^* = H$ и $H^*H = H^2 = I - 4vv^* + 4v(v^*v)v^* = I$.

Название вполне оправдано. Пусть $x \perp v \quad \Rightarrow \quad v^*x = 0$. Тогда $Hx = x - 2v(v^*x) = x \quad \Rightarrow$ подпространство $(L(v))^\perp$ является собственным подпространством для собственного значения $\lambda = 1$ кратности $n-1$. Кроме того, $Hv = v - 2v(v^*v) = -v \quad \Rightarrow$ вектор v отражается относительно подпространства $(L(v))^\perp$ и определяет одномерное собственное подпространство для собственного значения $\lambda = -1$ кратности 1.

Таким образом, в некотором ортонормированном базисе матрица отражения имеет вид

$$\Lambda = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \end{bmatrix}$$

Вещественной матрицей вращения (матрицей Гивенса) порядка n, определяемой углом ϕ и номерами $1 \le k < l \le n$, называется матрица $W = W(\phi, k, l)$, отличающаяся от единичной лишь элементами 2×2-подматрицы на пересечении строк и столбцов с номерами k и l; данная подматрица имеет вид

$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

Под *комплексной матрицей вращения* можно понимать матрицу такого же вида, в которой указанная 2×2-подматрица может быть умножена справа и слева на произвольные диагональные унитарные матрицы.

Унитарность вещественных и комплексных матриц вращения проверяется непосредственно.

33.4 Эрмитовы матрицы

Напомним, что матрица A называется эрмитовой, если $A^* = A$. Очевидно, любая эрмитова матрица является нормальной.

Утверждение. *Нормальная матрица является эрмитовой тогда и только тогда, когда все ее собственные значения вещественны.*

Доказательство. $A^* = A \quad \Leftrightarrow \quad \Lambda^* = \Lambda \quad \Leftrightarrow \quad \overline{\lambda}_i = \lambda_i,\ 1 \le i \le n.$ $\square$

Задача. Известно, что $A^2 = A$ и $\ker A \perp \operatorname{im} A$ (ортогональность относительно естественного скалярного произведения). Докажите, что $A = A^*$.

33.5 Эрмитово разложение

Запись матрицы A в виде $A = H + \mathbf{i}K$, где $H^* = H$, $K^* = K$, называется ее *эрмитовым разложением*.

Теорема. *Для любой матрицы $A \in \mathbb{C}^{n\times n}$ эрмитово разложение существует и единственно.*

Доказательство. Единственность: $A = H + \mathbf{i}K \quad \Rightarrow \quad A^* = H - \mathbf{i}K \quad \Rightarrow$

$$H = \frac{1}{2}(A + A^*), \quad K = \frac{1}{2\mathbf{i}}(A - A^*). \qquad (*)$$

Существование: пусть H и K определяются формулами $(*)$; они, очевидно, эрмитовы и при этом $A = H + \mathbf{i}K$. $\square$

Заметим, что матрица $B = \mathbf{i}K$ является *косоэрмитовой* — так называются матрицы B со свойством $B^* = -B$.

33.6 Неотрицательная и положительная определенность

Матрица $A \in \mathbb{C}^{n\times n}$ называется *неотрицательно (положительно) определенной*, если $x^*Ax \geq 0$ $(x^*Ax > 0) \quad \forall \ x \in \mathbb{C}^n, \ x \neq 0$. Обозначение: $A \geq 0$ $(A > 0)$. Неотрицательно определенные матрицы называются также *положительно полуопределенными*.

Теорема. *Для неотрицательной (положительной) определенности матрицы $A \in \mathbb{C}^{n\times n}$ необходимо и достаточно, чтобы она была эрмитовой матрицей с неотрицательными (положительными) собственными значениями.*

Доказательство. Используя эрмитово разложение $A = H + \mathbf{i}K$, находим

$$x^*Ax = (x^*Hx) + \mathbf{i}(x^*Kx).$$

Число x^*Ax вещественно для любого $x \quad \Rightarrow \quad x^*Kx = 0$ для всех x. Отсюда вытекает, что эрмитова матрица K имеет только нулевые собственные значения: $Kx = \lambda x, \ x \neq 0 \quad \Rightarrow \quad x^*Kx = \lambda(x^*x) = 0$ $\Rightarrow \quad \lambda = 0$. Будучи подобна нулевой матрице, матрица K может быть только нулевой $\quad \Rightarrow \quad A = H$. Если $Hx = \lambda x, x \neq 0$, то $x^*Hx = \lambda(x^*x) \geq 0 \quad \Rightarrow \quad \lambda \geq 0$. Если $X^*Hx > 0$, то, конечно, $\lambda > 0$.

Теперь предположим, что A — эрмитова матрица с неотрицательными собственными значениями $\lambda_1, \ldots, \lambda_n$ и ортонормированным базисом собственных векторов $v_1, \ldots, v_n$. Пусть $x = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Тогда $Ax = \alpha_1\lambda_1 v_1 + \ldots + \alpha_n\lambda_n v_n$. Отсюда

$$x^*Ax = (Ax, x) = \lambda_1|\alpha_1|^2 + \ldots + \lambda_n|\alpha_n|^2 \geq 0.$$

В случае $\lambda_i > 0$ находим $x^*Ax > 0$ при $x \neq 0$. $\square$

Задача. Пусть заданы вещественная положительно определенная матрица A порядка n и вектор $b \in \mathbb{R}^n$. Доказать, что функционал $f(x) = (Ax, x) + (b, x)$ при $x \in \mathbb{R}^n$ ограничен снизу и существует единственная точка x_0, в которой $f(x_0)$ есть его минимальное значение.

33.7 Квадратный корень

Если $A = S^2$, то S естественно называть *квадратным корнем* из матрицы A.

Теорема. *Для любой неотрицательно определенной матрицы $A \in \mathbb{C}^{n\times n}$ существует единственная неотрицательно определенная матрица $S \in \mathbb{C}^{n\times n}$ такая, что $S^2 = A$.*

Доказательство. Матрица A эрмитова и поэтому унитарно подобна вещественной диагональной матрице Λ с диагональными элементами $\lambda_i \geq 0$ (вследствие неотрицательной определенности): $A = Q\Lambda Q^*$. Пусть D — диагональная матрица с элементами $\sqrt{\lambda_i}$. Тогда $D^2 = \Lambda$ и, очевидно, $S = QDQ^*$ — неотрицательно определенный квадратный корень из A.

Приведенное построение доказывает существование. Но единственность требует дополнительного рассуждения. Если $SQ = QD$, то $AQ = QD^2$. Пусть $Q = [q_1, \ldots, q_n]$ и D имеет диагональные элементы d_i. Пусть x — собственный вектор матрицы A для собственного значения λ. Тогда для некоторых коэффициентов α_i

$$x = \sum_{d_i = \sqrt{\lambda}} \alpha_i q_i \quad \Rightarrow \quad Sx = \sum_{d_i = \sqrt{\lambda}} \alpha_i \sqrt{\lambda} q_i = \sqrt{\lambda} x.$$

Таким образом, действие S однозначно определено на векторах любого базиса из собственных векторов матрицы A. $\square$

Задача. Матрицы A и B обе эрмитовы, при этом A положительно определенная. Докажите, что собственные значения матриц AB и BA вещественные.

33.8 Блочно диагональная форма вещественной нормальной матрицы

Пусть A — вещественная нормальная матрица. В силу нормальности, все жордановы клетки — порядка 1.

Предположим, что $\lambda = a + \mathbf{i}b$ — собственное значение с ненулевой мнимой частью b, и пусть

$$A(x + \mathbf{i}y) = (a + \mathbf{i}b)(x + \mathbf{i}y) = (ax - by) + \mathbf{i}(bx + ay), \quad x, y \in \mathbb{R}^n. \quad \Rightarrow$$

$$A[x, y] = [x, y] \begin{bmatrix} a & b \\ -b & a \end{bmatrix}. \qquad (*)$$

Заметим, что сопряженное число $\overline{\lambda} = a - \mathbf{i}b$ тоже будет собственным значением, отвечающим собственному вектору $x - \mathbf{i}y$. Для нормальной матрицы собственные векторы для различных собственных значений ортогональны $\Rightarrow$

$$(x + \mathbf{i}y, x - \mathbf{i}y) = (x, x) - (y, y) + \mathbf{i}2(x, y) = 0 \quad \Rightarrow \quad (x, y) = 0, \quad |x| = |y|.$$

Отсюда следует, что равенство $(*)$ сохранится при замене x и y на нормированные и ортогональные векторы x/s и y/s, $s = |x| = |y|$. Таким образом, имеет место

Теорема. *Для любой вещественной нормальной матрицы существует вещественный ортонормированный базис, в котором она является прямой суммой вещественных блоков порядка 1 и вещественных блоков порядка 2 вида* $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

33.9 Блочно диагональная форма ортогональной матрицы

Собственные значения ортогональной матрицы по модулю равны 1. Поэтому аналог жордановой формы в данном случае представляет собой прямую сумму блоков порядка 1, отвечающих вещественным собственным значениям, равным 1 или -1, и блоков порядка 2, отвечающих парам комплексно сопряженных собственных значений $\lambda = a + \mathbf{i}b$ и $\overline{\lambda} = a - \mathbf{i}b$, $b \neq 0$. Заметим, что $a^2 + b^2 = 1 \quad \Rightarrow$ согласно $(*)$, каждый блок порядка 2 в данном случае есть вещественная матрица вращения.

Теорема. *Для любой ортогональной матрицы существует вещественный ортонормированный базис, в котором она является произведением вещественных матриц отражения и вещественных матриц вращения.*

Доказательство. Из сказанного выше ясно, что в некотором ортонормированном базисе получается блочно диагональная матрица с вещественными блоками порядка 1 для собственных значений ± 1 и блоками порядка 2, которые оказываются вещественными матрицами вращения. Достаточно заметить, что

$$\begin{bmatrix} M_1 & & & \\ & M_2 & & \\ & & \ddots & \\ & & & M_k \end{bmatrix} = \begin{bmatrix} M_1 & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix} \begin{bmatrix} I & & & \\ & M_2 & & \\ & & \ddots & \\ & & & I \end{bmatrix} \cdots \begin{bmatrix} I & & & \\ & I & & \\ & & \ddots & \\ & & & M_k \end{bmatrix}. \quad \square$$

Теорему можно проинтерпретировать таким образом: *линейное отображение в $\mathbb{R}^n$, сохраняющее длины, сводится к композиции отражений и вращений.*

Задача. Докажите, что любая вещественная матрица вращения является произведением двух вещественных матриц отражения.

Лекция 34

ОСНОВНАЯ ЧАСТЬ

34.1 Матрица Фурье

Исключительно важный класс унитарных матриц в математике и приложениях — это специальные матрицы Вандермонда, построенные на корнях из единицы. Пусть

$$\varepsilon \;=\; \cos\left(-\frac{2\pi}{n}\right) \;+\; \mathbf{i}\,\sin\left(-\frac{2\pi}{n}\right).$$

Это первообразный корень из единицы степени n. [1] Матрица Вандермонда для чисел ε^0, ε^1, $\ldots$, ε^{n-1} называется также матрицей (прямого) *дискретного преобразования Фурье*, или, короче, *матрицей Фурье* порядка n. Обозначение:

$$F_n \;=\; \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \varepsilon^{1\cdot 1} & \varepsilon^{1\cdot 2} & \ldots & \varepsilon^{1\cdot(n-1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \varepsilon^{(n-2)\cdot 1} & \varepsilon^{(n-2)\cdot 2} & \ldots & \varepsilon^{(n-2)\cdot(n-1)} \\ 1 & \varepsilon^{(n-1)\cdot 1} & \varepsilon^{(n-1)\cdot 2} & \ldots & \varepsilon^{(n-1)\cdot(n-1)} \end{bmatrix}.$$

Утверждение. *Матрица Фурье обратима и при этом обратная матрица имеет вид*

$$F_n^{-1} \;=\; \frac{1}{n}\,F_n^*.$$

Доказательство. Элементы произведения матриц $F_n^* F_n$ легко вычисляются как суммы членов геометрической прогрессии:

$$(F_n^* F_n)_{ij} \;=\; \sum_{k=0}^{n-1} \bar{\varepsilon}^{ki}\varepsilon^{kj} \;=\; \sum_{k=0}^{n-1} \varepsilon^{k(j-i)} \;=\; \sum_{k=0}^{n-1} \left(\varepsilon^{(j-i)}\right)^k \;=\; \begin{cases} \frac{\varepsilon^{(j-i)n}-1}{\varepsilon^{j-i}-1} = 0, & i \neq j, \\ n, & i = j. \end{cases}$$

Таким образом, $F_n^* F_n \;=\; n\,I$. $\square$

Задача. Доказать, что $F_n^4 = n^2 I$.

34.2 Циркулянтные матрицы

Красивый и полезный класс нормальных матриц составляют матрицы вида

$$A = \begin{bmatrix} a_0 & a_{n-1} & a_{n-2} & \ldots & a_2 & a_1 \\ a_1 & a_0 & a_{n-1} & \ldots & a_3 & a_2 \\ a_2 & a_1 & a_0 & \ldots & a_4 & a_3 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n-2} & a_{n-3} & a_{n-4} & \ldots & a_0 & a_{n-1} \\ a_{n-1} & a_{n-2} & a_{n-3} & \ldots & a_1 & a_0 \end{bmatrix}.$$

[1] Минус — дань сложившейся традиции определения прямого и обратного преобразований Фурье: минус — для прямого, плюс — для обратного.

Матрица A называется *циркулянтной матрицей* или *циркулянтом*. В частности, при $n = 4$ получаем

$$A = \begin{bmatrix} a_0 & a_3 & a_2 & a_1 \\ a_1 & a_0 & a_3 & a_2 \\ a_2 & a_1 & a_0 & a_3 \\ a_3 & a_2 & a_1 & a_0 \end{bmatrix}.$$

Как видим, циркулянтная матрица полностью определяется элементами любой своей строки или любого столбца. Ее первый столбец есть $a = [a_0, a_1, \ldots, a_{n-1}]^\top$.

Чтобы найти собственные значения и собственные векторы матрицы A, возьмем произвольный корень ξ степени n из единицы ($\xi^n = 1$) и рассмотрим число

$$\lambda = \lambda(\xi) \equiv a_0 + \xi a_1 + \ldots + \xi^{n-1} a_{n-1}.$$

Последовательно умножая обе части на $1, \xi, \xi^2, \ldots, \xi^{n-1}$, находим

$$\begin{array}{lcllll} \lambda \cdot 1 & = & a_0 & + \xi\, a_1 & + \ldots + & \xi^{n-1}\, a_{n-1}, \\ \lambda \cdot \xi & = & a_{n-1} & + \xi\, a_0 & + \ldots + & \xi^{n-1}\, a_{n-2}, \\ \lambda \cdot \xi^2 & = & a_{n-2} & + \xi\, a_{n-1} & + \ldots + & \xi^{n-1}\, a_{n-3}, \\ & \ldots & & & & \\ \lambda \cdot \xi^{n-1} & = & a_1 & + \xi\, a_2 & + \ldots + & \xi^{n-1}\, a_0. \end{array}$$

Следовательно,

$$\lambda(\xi)\,[1, \xi, \ldots, \xi^{n-1}] \;=\; [1, \xi, \ldots, \xi^{n-1}]A. \qquad (*)$$

Выберем $\varepsilon = \cos(-2\pi/n) + \mathbf{i}\sin(-2\pi/n)$. Равенство $(*)$ справедливо при $\xi = 1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{n-1}$ и, следовательно, дает систему равенств, которая в матричной записи имеет вид

$$\Lambda F_n = F_n A,$$

где F_n — матрица Фурье порядка n, Λ — диагональная матрица вида

$$\Lambda = \begin{bmatrix} \lambda(1) & & & \\ & \lambda(\varepsilon) & & \\ & & \ddots & \\ & & & \lambda(\varepsilon^{n-1}) \end{bmatrix},$$

Итак, $AF_n^* = F_n^*\Lambda \;\Rightarrow\;$ столбцы матрицы F_n^* суть собственные векторы матрицы A, отвечающие собственным значениям, расположенным на диагонали матрицы Λ. Заметим, что F_n^* получается из F_n перестановкой столбцов: первый столбец остается на месте, а столбцы со второго по последний ставятся в обратном порядке. Поэтому можно утверждать, что базисом из собственных векторов циркулянтной матрицы A являются столбцы матрицы Фурье F_n. Полученные результаты сформулируем в виде теоремы.

Теорема о циркулянтах. *Пусть A — циркулянтная матрица с первым столбцом $a = [a_0, \ldots, a_{n-1}]^\top$. Тогда*

$$A = \frac{1}{n} F_n^* \Lambda F_n, \qquad (\#)$$

где F_n — матрица Фурье порядка n и Λ — диагональная матрица собственных значений вида

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}, \quad \begin{bmatrix} \lambda_1 \\ \ldots \\ \lambda_n \end{bmatrix} = F_n \begin{bmatrix} a_0 \\ \ldots \\ a_{n-1} \end{bmatrix}.$$

Несложно проверить, что для любых $\lambda_1, \ldots, \lambda_n$ матрица в правой части $(\#)$ является циркулянтной матрицей. Отсюда ясно, что произведение циркулянтных матриц остается циркулянтной матрицей.

Матрица, обратная к невырожденной циркулянтной матрице, также является циркулянтной.

34.3 Алгебры матриц

Любая линейная комбинация циркулянтов есть циркулянт. Таким образом, множество циркулянтов порядка n является n-мерным линейным пространством, на котором определена операция умножения элементов, которая вместе с операцией сложения превращает данное линейное пространство в кольцо.

Пусть в линейном пространстве V определена операция умножения элементов, которая делает его также кольцом с единицей, и пусть умножение произвольных элементов a и b и умножение на число α связаны аксиомой $\alpha(ab) = (\alpha a)b = a(\alpha b)$. В таких случаях пространство V называется *алгеброй*.

Заметим, что умножение циркулянтов коммутативно — поэтому они дают пример *коммутативной алгебры* матриц. Все множество матриц фиксированного порядка n — пример некоммутативной алгебры.

Теорема. *Пусть $\mathcal{M}$ — алгебра матриц и $A \in \mathcal{M}$ — невырожденная матрица. Тогда $A^{-1} \in \mathcal{M}$.*

Доказательство. Пусть $A \in \mathcal{M}$. По теореме Гамильтона–Кэли, A аннулируется своим характеристическим многочленом: $a_0 I + a_1 A + \ldots + a_{n-1}A^{n-1} + A^n = 0$. Если A — невырожденная матрица, то, умножая обе части на A^{-1} и учитывая, что $a_0 = (-1)^n \det A \neq 0$, получаем

$$A^{-1} \;=\; -\frac{1}{a_0}\left(a_1 I + a_2 A + \ldots a_{n-1}A^{n-2} + A^{n-1}\right) \;\in\; \mathcal{M}. \qquad \square$$

По аналогии с циркулянтами, можно построить много других коммутативных матричных алгебр.

Утверждение. *Для любой фиксированной невырожденной матрицы $Q \in \mathbb{C}^{n\times n}$ все матрицы вида $Q\Lambda Q^{-1}$, где Λ — произвольная диагональная матрица порядка n, образуют коммутативную алгебру.*

Доказательство. Указанное множество матриц обозначим через $\mathcal{M}$. Если $A_1, A_2 \in \mathcal{M}$, то $A_1 = Q\Lambda_1 Q^{-1}$, $A_2 = Q\Lambda_2 Q^{-1}$ для каких-то диагональных матриц Λ_1 и Λ_2. Тогда $\alpha A_1 + \beta A_2 = Q(\Lambda_1 + \Lambda_2)Q^{-1} \in \mathcal{M}$ и $A_1 A_2 = Q(\Lambda_1\Lambda_2)Q^{-1} \in \mathcal{M}$. $\square$

Замечание. Данное утверждение описывает не все возможные коммутативные алгебры матриц. Например, пусть $\mathcal{M}$ состоит из всех $n \times n$-матриц вида

$$A = \begin{bmatrix} a_0 & & & & & \\ a_1 & a_0 & & & & \\ a_2 & a_1 & a_0 & & & \\ \ldots & \ldots & \ldots & \ldots & & \\ a_{n-2} & a_{n-3} & \ldots & a_1 & a_0 & \\ a_{n-1} & a_{n-2} & a_{n-3} & \ldots & a_1 & a_0 \end{bmatrix}.$$

Несложно проверить, что $\mathcal{M}$ является коммутативной алгеброй, но в $\mathcal{M}$ имеются недиагонализуемые матрицы (докажите!). Еще один пример коммутативной алгебры — множество матриц A таких, что $A^\top \in \mathcal{M}$.

34.4 Одновременное приведение к треугольному виду

Теорема. *Для произвольной коммутативной алгебры матриц $\mathcal{M}$ существует обратимая матрица Q такая, что для любой $A \in \mathcal{M}$ матрица $Q^{-1}AQ$ является верхней треугольной.*

Доказательство. Пусть матрицы $A_1, \ldots, A_k \in \mathcal{M} \subset \mathbb{C}^{n\times n}$ образуют базис в линейном пространстве $\mathcal{M}$. Докажем, что они имеют общий собственный вектор.

Обозначим через L собственное подпространство матрицы A_1 для собственного значения λ_1. Пусть $A_1 x = \lambda_1 x$, $x \neq 0$. Тогда

$$A_1(A_2 x) = A_2(A_1 x) = \lambda_1(A_2 x). \qquad (*)$$

Следовательно, $A_2 x \in L$. Более того, $A_2^l x \in L$ для всех $l = 1, 2, \ldots$. Пусть M — минимальное подпространство, содержащее все векторы вида $A_2^l x$. Очевидно, это минимальное подпространство, инвариантное относительно A_2 и содержащее x. В силу $(*)$ заключаем, что $M \subset L$. В M обязательно имеется собственный вектор для A_2, он же будет собственным вектором и для A_1.

Далее по индукции. Пусть L — содержащее $x \neq 0$ пересечение собственных подпространств $L_1, \ldots, L_k$, отвечающих соответственно матрицам $A_1, \ldots, A_{k-1}$, а M — содержащее x минимальное

подпространство, инвариантное относительно A_k (очевидно, оно состоит из векторов вида $p(A_k)x$ для всевозможных многочленов p). Легко проверить, что M является (ненулевым!) подпространством для каждого из собственных подпространств $L_1, \ldots, L_k$. Поэтому $M \subset L$, а содержащийся в M собственный вектор для A_k является собственным вектором также для $A_1, \ldots, A_{k-1}$.

Итак, пусть x — общий собственный вектор для $A_1, \ldots, A_k$. Пусть P — любая обратимая матрица, первый столбец которой равен x. Тогда каждая из матриц $PA_1P^{-1}, \ldots, PA_kP^{-1}$ имеет блочный вид

$$P^{-1}A_iP = \begin{bmatrix} \lambda_i & v_i^\top \\ 0 & B_i \end{bmatrix}, \quad B_i \in \mathbb{C}^{(n-1)\times(n-1)}.$$

Непосредственно проверяется, что матрицы $B_1, \ldots, B_k$ коммутируют. Если они одновременно приводятся к верхнему треугольному виду с помощью обратимой матрицы Z порядка $n-1$ (каждая из матриц $Z^{-1}B_iZ$ является верхней треугольной), то матрица

$$Q = P\begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix}$$

одновременно приводит к треугольному виду каждую из матриц $A_1, \ldots, A_k$. То же верно и для произвольной линейной комбинации матриц $A_1, \ldots, A_k$. $\square$

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

34.5 Быстрое преобразование Фурье

Умножение матрицы Фурье F_n на вектор-столбец $x \in \mathbb{C}^n$ называется *прямым преобразованием Фурье* вектора x.

Классическое правило умножения матрицы на вектор дает алгоритм с числом операций порядка n^2. Однако, специальный вид матрицы F_n позволяет умножать ее на вектор с затратой лишь $O(n\log_2 n)$ арифметических операций!

Алгоритмы с таким свойством (*быстрое преобразование Фурье*) начали внедряться в практику вычислений в 60-х годах 20-го века и произвели буквально переворот в ряде разделов прикладной математики. [2] Так или иначе, быстрое преобразование Фурье стало основной компонентой многих быстрых алгоритмов в задачах линейной алгебры.

Предположим, что $n = 2^L$ и $m = n/2$. Будем нумеровать строки и столбцы матрицы F_n числами от 0 до $n-1$. От F_n перейдем к матрице $\widetilde{F}_n$, в которой сначала идут подряд все строки F_n с четными номерами, а затем — все строки с нечетными номерами (ясно, что $\widetilde{F}_n = P_nF_n$, где P_n — соответствующая матрица перестановки). Рассмотрим $\widetilde{F}_n$ как блочную 2×2-матрицу:

$$\widetilde{F}_n = \begin{bmatrix} [\varepsilon^{2\,k\,l}]_{m\times m} & [\varepsilon^{2\,k\,(m+l)}]_{m\times m} \\ [\varepsilon^{(2k+1)\,l}]_{m\times m} & [\varepsilon^{(2k+1)(m+l)}]_{m\times m} \end{bmatrix}, \qquad 0 \le k,l \le m-1.$$

Заметим, что

$$[\varepsilon^{2\,k\,l}]_{m\times m} = [\varepsilon^{2\,k\,(m+l)}]_{m\times m} = F_m,$$
$$[\varepsilon^{(2k+1)\,l}]_{m\times m} = F_mD_m, \quad [\varepsilon^{(2k+1)(m+l)}]_{m\times m} = -F_mD_m,$$

где

$$D_m = \begin{bmatrix} 1 & & & \\ & \varepsilon^1 & & \\ & & \ddots & \\ & & & \varepsilon^{m-1} \end{bmatrix}.$$

Следовательно,

$$F_n = P_n\begin{bmatrix} F_m & 0 \\ 0 & F_m \end{bmatrix}\begin{bmatrix} I_m & 0 \\ 0 & D_m \end{bmatrix}\begin{bmatrix} I_m & I_m \\ I_m & -I_m \end{bmatrix}, \qquad m = n/2.$$

[2] Начало "переворота" отсчитывается с 1965 года — со знаменитой работы американцев Кули и Тьюки. Впоследствии было выяснено, что быстрые алгоритмы были описаны Рунге еще в начале 20-го века; более того, Г. Стрэнг утверждает, что обнаружил их прототипы еще у Гаусса.

Таким образом, задача умножения матрицы F_n на вектор сводится к двум аналогичным задачам для матрицы $F_{n/2}$. Чтобы осуществить редукцию, требуется выполнить n сложений-вычитаний и $n/2$ умножений (на элементы диагональной матрицы D_n). Обозначим через $S_\pm(n)$ и $S_*(n)$ общее число сложений-вычитаний и умножений. Чтобы их оценить, нужно просуммировать затраты на редукцию задач для всех $L = \log_2 n$ шагов рекурсии:

$$S_\pm(n) \;=\; n \;+\; 2(n/2) \;+\; 2^2(n/2^2) \;+\; \ldots \;+\; 2^{L-1}(n/2^{L-1}) \;=\; nL \;=\; n\log_2 n,$$

$$S_*(n) \;=\; \frac{1}{2}\, n\log_2 n.$$

34.6 Свертки

Пусть циркулянтная матрица A определяется первым столбцом a. Вектор $y = Ax$ называется *периодической сверткой* векторов a и x. Обозначение: $y = a * x$.

Задача. Докажите, что $a * x = x * a$.

Согласно теореме о циркулянтах, вычисление периодической свертки векторов из $\mathbb{R}^n$ (умножение на циркулянтную матрицу) сводится к трем умножениям на матрицу Фурье. Последнее можно выполнить с помощью алгоритма быстрого преобразования Фурье за $O(n\log n)$ арифметических операций, если $n = 2^L$.

Решение линейных систем с циркулянтной матрицей осуществляется с теми же затратами (докажите!).

Пусть теперь $a = [a_{-n+1}, a_{-n+2}, \ldots, a_0, a_1, \ldots, a_{n-1}]^\top \in \mathbb{C}^{2n-1}$ и $x \in \mathbb{C}^n$. Под *апериодической сверткой* векторов a и x иногда понимается вектор $y = Ax$, где

$$A = \begin{bmatrix} a_0 & a_{-1} & \ldots & a_{-n+1} \\ a_1 & a_0 & \ldots & a_{-n+2} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1} & a_{n-2} & \ldots & a_0 \end{bmatrix}.$$

Матрица A такого вида называется *теплицевой матрицей*. [3] Заметим, что любой циркулянт является также теплицевой матрицей.

Утверждение. *Для любого n теплицева матрица порядка n может быть умножена на вектор с затратой $O(n\log_2 n)$ операций.*

Доказательство. Достаточно заметить, что теплицеву матрицу A порядка n можно "достроить" до циркулянта

$$C = \begin{bmatrix} A & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

порядка $N = 2^L \;<\; 4n$. Вот как это делается в случае $n = 3$:

$$C = \left[\begin{array}{ccc|ccccc} a_0 & a_{-1} & a_{-2} & 0 & 0 & 0 & a_2 & a_1 \\ a_1 & a_0 & a_{-1} & a_{-2} & 0 & 0 & 0 & a_2 \\ a_2 & a_1 & a_0 & a_{-1} & a_{-2} & 0 & 0 & 0 \\ \hline 0 & a_2 & a_1 & a_0 & a_{-1} & a_{-2} & 0 & 0 \\ 0 & 0 & a_2 & a_1 & a_0 & a_{-1} & a_{-2} & 0 \\ 0 & 0 & 0 & a_2 & a_1 & a_0 & a_{-1} & a_{-2} \\ a_{-2} & 0 & 0 & 0 & a_2 & a_1 & a_0 & a_{-1} \\ a_{-1} & a_{-2} & 0 & 0 & 0 & a_2 & a_1 & a_0 \end{array}\right].$$

Далее, пусть

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} A & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix}.$$

Отсюда ясно, что $u = Ax$. Таким образом, умножение на теплицеву матрицу сводится к умножению на циркулянтную матрицу порядка $N = 2^L$. Применение быстрого преобразования Фурье дает алгоритм с число операций $O(N\log_2 N) = O(n\log_2 n)$. $\square$

[3] В честь немецкого математика Отто Теплица.

34.7 Быстрые алгоритмы

Что можно сказать о сложности преобразования Фурье в случае $n \neq 2^L$?

Пусть элементы матрицы F_n нумеруются индексами от 0 до $n-1$. В позиции (k,l) находится число

$$\varepsilon^{kl} = \varepsilon^{(k^2+l^2-(k-l)^2)/2} = \varepsilon^{k^2/2}\ \varepsilon^{-(k-l)^2}\ \varepsilon^{l^2/2}.$$

Поэтому матрица Фурье расщепляется в произведение трех матриц

$$F_n = DAD, \quad D = \begin{bmatrix} \varepsilon^{0^2/2} & & & \\ & \varepsilon^{1^2/2} & & \\ & & \ddots & \\ & & & \varepsilon^{(n-1)^2/2} \end{bmatrix}, \quad A = [\varepsilon^{-(k-l)^2/2}], \ \ 0 \leq k,l \leq n-1.$$

Таким образом, умножение на матрицу Фурье произвольного порядка n сводится к умножению на теплицеву матрицу A того же порядка n. Последнее сводится к умножению на циркулянтную матрицу порядка $n \leq N = 2^L < 4n$.

В итоге все сводится к троекратному применению алгоритма быстрого преобразования Фурье специально выбранного порядка $N = 2^L$.

Описанная возможность получения быстрого преобразования Фурье без ограничений на его порядок является, вероятно, самой простой — но не единственной и не всегда наилучшей для практических вычислений.

Задача. Доказать, что два многочлена степени n можно перемножить с затратой $O(n \log_2 n)$ арифметических операций.

Лекция 35

ОСНОВНАЯ ЧАСТЬ

35.1 Сингулярные числа и сингулярные векторы

Пусть $A \in \mathbb{C}^{m\times n}$. Тогда $A^*A \in \mathbb{C}^{n\times n}$ — эрмитова неотрицательно определенная матрица:

$$(A^*A)^* = A^*(A^*)^* = A^*A; \qquad xA^*Ax = (Ax, Ax) = |Ax|^2 \geq 0 \ \ \forall \ \ x \in \mathbb{C}^n.$$

Поэтому все ее собственные значения неотрицательны.

Неотрицательные квадратные корни из собственных значений матрицы A^*A называются *сингулярными числами* матрицы A. Сингулярные числа $\sigma_i = \sigma_i(A)$ принято нумеровать по невозрастанию:

$$\sigma_1 \ \geq \ \sigma_2 \ \geq \ \ldots \ \geq \ \sigma_r \ > \ \sigma_{r+1} = \ldots = \sigma_n = 0.$$

Будем считать, что A имеет r ненулевых сингулярных чисел.

Пусть $u_1, \ldots, u_n$ — ортонормированный базис собственных векторов матрицы A^*A такой, что

$$A^*Au_i = \begin{cases} \sigma_i^2 u_i, & 1 \leq i \leq r, \\ 0\,, & r+1 \leq i \leq n. \end{cases}$$

Положим $v_i = Au_i/\sigma_i$, $1 \leq i \leq r$. Тогда $(v_i, v_j) = 0$ при $i \neq j$ и $(v_i, v_i) = 1$. Дополним систему $v_1, \ldots, v_r$ векторами $v_{r+1}, \ldots, v_m$ до ортонормированного базиса в $\mathbb{C}^m$. Заметим также, что при $j \geq r+1$

$$A^*Au_j = 0 \quad \Rightarrow \quad u_j^*A^*Au_j = 0 \quad \Rightarrow \quad (Au_j)^*(Au_j) = 0 \quad \Rightarrow \quad |Au_j| = 0 \quad \Rightarrow \quad Au_j = 0.$$

В итоге получаем

$$A[u_1, \ldots, u_n] = [v_1, \ldots, v_m] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix} \quad \Rightarrow \quad AU = V\Sigma,$$

где $U = [u_1, \ldots, u_n]$ и $V = [v_1, \ldots, v_m]$ — унитарные матрицы, а Σ — диагональная прямоугольная матрица тех же размеров, что и матрица A.

Столбцы матриц U и V образуют *сингулярные базисы* матрицы A. Столбцы U называются *правыми сингулярными векторами*, а столбцы V — *левыми сингулярными векторами* матрицы A. Связь между сингулярными векторами и ненулевыми сингулярными числами устанавливается соотношениями

$$Au_i = \sigma_i v_i, \quad A^*v_i = \sigma_i u_i, \quad 1 \leq i \leq r.$$

Кроме того,

$$Au_i = 0, \ \ r+1 \leq i \leq n, \qquad A^*v_i = 0, \ \ r+1 \leq i \leq m.$$

Итак, мы доказали, что *для любой матрицы $A \in \mathbb{C}^{m\times n}$ имеет место равенство*

$$AU = V\Sigma \tag{$*$}$$

для некоторых унитарных матриц $U \in \mathbb{C}^{n\times n}$, $V \in \mathbb{C}^{m\times m}$ и диагональной прямоугольной матрицы размеров $m \times n$ с числами $\sigma_i \geq 0$ при $i = j$. Записав $(*)$ в виде

$$A = V\Sigma U^*, \qquad (**)$$

получаем представление матрицы, называемое ее *сингулярным разложением.* [1]

Если каким-то способом получено разложение $(**)$ с унитарными матрицами U и V, то $A^*A = U(\Sigma^*\Sigma)U^*$. Поэтому если Σ — диагональная прямоугольная матрица с неотрицательными элементами, то ее ненулевые элементы определены однозначно.

35.2 Полярное разложение

Если $m = n$, то можно записать $(**)$ в виде

$$A = (V\Sigma V^*)(VU^*) = HQ,$$

где $H = V\Sigma V^*$ — неотрицательно определенная (поэтому также эрмитова) матрица, а $Q = VU^*$ — унитарная матрица (как произведение унитарных матриц). Представление матрицы A в виде $A = HQ$ с неотрицательно определенной H и унитарной Q называется ее *полярным разложением.*

Полярное разложение матрицы можно считать аналогом тригонометрической формы комплексного числа.

35.3 Выводы из сингулярного разложения

(1) *Число ненулевых сингулярных чисел r равно рангу матрицы A.*

(2) *Сингулярное разложение сопряженной матрицы имеет вид*

$$A^* = U\Sigma^\top V^*.$$

(3) $\mathrm{im}A = L(v_1, \ldots, v_r), \quad \ker A = L(u_{r+1}, \ldots, u_n).$

(4) $\mathrm{im}A^* = L(u_1, \ldots, u_r), \quad \ker A^* = L(v_{r+1}, \ldots, v_m).$

В качестве следствия можно получить представления пространств в виде ортогональных сумм

$$\mathbb{C}^n = \ker A \oplus \mathrm{im}A^*, \quad \mathbb{C}^m = \ker A^* \oplus \mathrm{im}A.$$

(5) $A = \sum\limits_{k=1}^{r} \sigma_k v_k u_k^*, \quad A^* = \sum\limits_{k=1}^{r} \sigma_k u_k v_k^*.$

(6) *Если $m = n = r$ (матрица A невырожденная), то*

$$A = \sum_{k=1}^{n} \sigma_k v_k u_k^*, \quad A^{-1} = \sum_{k=1}^{n} \frac{1}{\sigma_k} u_k v_k^*.$$

(7) *Пусть $\sigma_1 \geq \ldots \geq \sigma_n$ — сингулярные числа невырожденной матрицы A. Тогда $\sigma_n^{-1} \geq \ldots \geq \sigma_1^{-1}$ — сингулярные числа матрицы A^{-1}.*

(8) $||A||_2 = \sigma_1, \quad ||A||_F = \sqrt{\sigma_1^2 + \ldots + \sigma_r^2}.$

Спектральная и фробениусова нормы являются унитарно инвариантными. Поэтому $||A||_2 = ||\Sigma||_2$ и $||A||_F = ||\Sigma||_F$. Очевидно, $||\Sigma x||_2 \leq \sigma_1 ||x||_2$; равенство достигается, если x имеет 1 в первой позиции и 0 в остальных.

Ясно также, что $||\Sigma||_F = \sqrt{\sigma_1^2 + \ldots + \sigma_r^2}$. $\square$

[1] Оно было получено совершенно другим способом в Лекции 27.

35.4 Сингулярное разложение и решение систем

Утверждение. *Решение системы $Ax = b$ с невырожденной матрицей A имеет вид* $x = \sum\limits_{k=1}^{n} \frac{\beta_k}{\sigma_k} u_k$, *где* $\beta_k = v_k^* b = (v_k, b)$ — *коэффициенты разложения вектора правой части b по сингулярным векторам* $v_1, \ldots, v_m$.

Доказательство. Выражение для x сразу же получается из **(6)**. Если $b = \beta_1 v_1 + \ldots + \beta_n v_n$, то $(b, v_k) = \beta_k (v_k, v_k) = \beta_k$ (вследствие ортонормированности системы векторов $v_1, \ldots, v_n$). □

Данное утверждение проясняет роль направления возмущений при решении систем. Если коэффициент β_k заменяется на $\beta_k + \varepsilon$, то коэффициент при u_k в разложении x по базису $u_1, \ldots, u_n$ возмущается на величину ε/σ_k. Чем меньше σ_k, тем сильнее может измениться решение. При малом σ_n "особенно опасны" возмущения вектора правой части b в направлении вектора v_n.

35.5 Метод наименьших квадратов

Если система $Ax = b$ несовместна, то равенство $Ax = b$ не выполняется ни для одного вектора x. В этом случае, тем не менее, пытаются интересоваться такими x, при которых вектор $b - Ax$ (его называют *невязкой* для x) имеет минимально возможную длину. Вектор x называется *псевдорешением* системы $Ax = b$, если

$$||b - Ax||_2 = \min_z ||b - Az||_2.$$

В данном методе определения "обобщенного решения" в вещественном случае речь действительно идет о наименьшем значении суммы квадратов (отсюда название метода)

$$||b - Ax||_2^2 = \sum_{i=1}^{m} (b_i - a_{i1}x_1 - \ldots - a_{in}x_n)^2.$$

Утверждение. *Пусть A — матрица размеров $m \times n$ и ранга r. Множество псевдорешений системы $Ax = b$ есть линейное многообразие, размерность которого равна $n - r$.*

Доказательство. Пусть h — перпендикуляр, опущенный из вектора b на подпространство $\mathrm{im}A$, а $y \in \mathrm{im}A$ — соответствующая ортогональная проекция. Тогда система $Az = y$ совместна, и если z — ее произвольное решение, то $|h| = |b - Az| < |b - Ax|$ для всех x таких, что $Ax \neq y$. Значит, множество псевдорешений совпадает с множеством решений совместной системы $Az = y$. □

Среди всех псевдорешений выделяется псевдорешение $\hat{x}$ минимальной длины — оно называется *нормальным псевдорешением*. Геометрически ясно, что $\hat{x}$ есть перпендикуляр, опущенный на $\ker A$ из любого частного решения z совместной системы $Az = y$ (вектор y — ортогональная проекция вектора b на $\mathrm{im}A$). Таким образом, *нормальное псевдорешение существует и единственно.*

Сингулярное разложение позволяет дать явный вид нормального псевдорешения:

$$\hat{x} \;=\; \sum_{k=1}^{r} \frac{v_k^* b}{\sigma_k}\, u_k.$$

Для доказательства достаточно проверить, что $b - A\hat{x} \perp \mathrm{im}A$ и $\hat{x} \perp \ker A$.

Простота формулы не должна создавать впечатление об отсутствии проблем при вычислении $\hat{x}$. Главная проблема, собственно, в том, что в случае $r < \min(m, n)$ ранг r можно повысить сколь угодно малым возмущением элементов матрицы, а это означает, что нормальное псевдорешение, несмотря на факт существования и единственности, не является непрерывной функцией от элементов матрицы A. Например, пусть $m = n = 1$ и рассматривается система $0 \cdot x = 1$. Ее нормальное псевдорешение есть, очевидно, $\hat{x} = 0$, а нормальное псевдорешение возмущенной системы $\varepsilon \cdot x = 1$ есть $\hat{x}(\varepsilon) = 1/\varepsilon$. Как видим, $\hat{x}(\varepsilon)$ не стремится к $\hat{x}$ при $\varepsilon \to 0$. Сама задача о вычислении столь неустойчивого объекта не кажется очень уж осмысленной.

В то же время, задачи такого рода постоянно возникают в приложениях, и от нас требуются какие-то методы их решения. При построении таких методов следует иметь в виду, что это должны быть, прежде всего, *методы изменения самой постановки задачи.* Подобные вопросы связаны с так называемыми *методами регуляризации.* [2]

[2] Общую теорию методов регуляризации создал основатель факультета ВМиК академик Андрей Николаевич Тихонов.

35.6 Наилучшие аппроксимации с понижением ранга

В каждой матрице $\sigma_k v_k u_k^*$ элемент в позиции (i,j) может рассматриваться как функция от i и j с разделенными дискретными переменными i и j: $f(i,j) = f_1(i)f_2(j)$. Таким образом, запись A в виде $A = \sum\limits_{i=1}^{r} \sigma_i v_i u_i^*$ описывает некоторый специальный способ разделения переменных в каждом члене суммы или, в матричной терминологии, скелетное разложение матрицы A — причем с важным дополнительным свойством ортонормированности систем $u_1, \ldots, u_r$ и $v_1, \ldots, v_r$.

Особая ценность и широта применений сингулярного разложения вызваны, прежде всего, тем, что оно дает простой и надежный механизм исключения из матрицы "наименее значимой информации" — путем ее аппроксимации суммой *меньшего числа слагаемых* с разделенными переменными i и j. Речь идет о поиске элемента наилучшего приближения для заданной матрицы A на довольно сложном множестве — множестве матриц, ранг которых ограничен заданным числом.

Теорема о наилучших аппроксимациях с понижением ранга. *Пусть матрица $A \in \mathbb{C}^{m\times n}$ задана сингулярным разложением вида*

$$A = \sum_{l=1}^{r} \sigma_l v_l u_l^*,$$

и условимся считать, что $\sigma_{r+1} = 0$. Пусть задано целое $1 \le k \le r$. Тогда

$$\min_{\substack{\mathrm{rank}B \le k \\ B \in \mathbb{C}^{m\times n}}} ||A-B||_2 \;=\; \sigma_{k+1} \;=\; ||A-A_k||_2, \quad \textit{где} \quad A_k = \sum_{l=1}^{k} \sigma_l v_l u_l^*.$$

Доказательство. Пусть $\mathrm{rank}B \le k$. Тогда $\dim \ker B \ge n-k$. Рассмотрим линейную оболочку $L = L(u_1, \ldots, u_{k+1})$, натянутую на старшие сингулярные векторы. По теореме Грассмана,

$$\dim(\ker B \cap L) = \dim \ker B + \dim L - \dim(\ker B + L) \ge (n-k) + (k+1) - n = 1.$$

Поэтому существует ненулевой вектор $z \in \ker B \cap L$. Будем считать, что $||z||_2 = 1$. Учитывая, что

$$z = \sum_{l=1}^{k+1} \alpha_l u_l, \quad \sum_{l=1}^{k+1} |\alpha_l|^2 = 1,$$

находим

$$||A-B||_2 \ge ||(A-B)z||_2 = ||Az||_2 = \sqrt{\sum_{l=1}^{k+1} |\alpha_l|^2 \sigma_l} \ge \sigma_{k+1}.$$

В то же время,

$$A - A_k = \sum_{l=k+1}^{r} \sigma_l v_l u_l^* \quad \Rightarrow \quad ||A-A_k||_2 = \sigma_{k+1}. \qquad \square$$

35.7 Расстояние до множества вырожденных матриц

Если A — невырожденная матрица, то все матрицы $A+F$ при достаточно малой норме $||F||_2$ будут невырожденными (почему?). Под спектральным расстоянием между A и множеством вырожденных матриц понимается величина $\rho \equiv \inf\limits_{\det B = 0} ||A-B||_2$.

Из теоремы об аппроксимациях с понижением ранга вытекает, что

$$\rho \;=\; \inf_{\mathrm{rank}B \le n-1} ||A-B||_2 \;=\; \sigma_n(A).$$

Таким образом, *спектральное расстояние от заданной невырожденной матрицы до множества вырожденных матриц равно ее минимальному сингулярному числу.*

Этот результат подчеркивает значение ортонормированных базисов: *если матрица V унитарная,*

то матрица $V+F$ будет невырожденной для всех возмущений F при условии $||F||_2 < 1$ (докажите!). В частности, матрица $I+F$ будет невырожденной для всех возмущений F с нормой $||F||_2 < 1$.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

35.8 Общий вид унитарно инвариантных норм

При работе с матрицами мы активно используем две унитарно инвариантных нормы: спектральную норму $||A||_2$ и норму Фробениуса $||A||_F$. Другие нормы того же типа с огромной пользой применяются, например, в асимптотическом матричном анализе (при изучении последовательностей матриц, порядок которых стремится к бесконечности).

Полное описание унитарно инвариантных норм было дано Джоном фон Нейманом в 1937 году. [3]

Пусть $A = V\Sigma U^*$ — сингулярное разложение матрицы A. Тогда для любой унитарно инвариантной нормы имеем равенство $||A|| = ||\Sigma||$. Поэтому $||A||$ есть функция от сингулярных чисел матрицы A:

$$||A|| = \Phi(\sigma_1, \ldots, \sigma_k), \quad k = \min(m, n).$$

Ясно, что $\Phi(\sigma_1, \ldots, \sigma_k)$ можно рассматривать как функцию от вектора $\sigma = [\sigma_1, \ldots, \sigma_k]^\top \in \mathbb{R}^n$.

Конечно, сингулярные числа неотрицательны, но давайте предположим, что $\Phi(\sigma)$ определена при всех $\sigma \in \mathbb{R}^k$. Рассмотрим следующий список требований к функции Φ:

(1) $\Phi(\sigma)$ является векторной нормой на $\mathbb{R}^k$;

(2) $\Phi(\sigma)$ зависит только от модулей координат вектора $\sigma \in \mathbb{R}^k$;

(3) $\Phi(P\sigma) = \Phi(\sigma)$ для любой матрицы перестановки порядка k;

(4) если $\sigma = [1, 0, \ldots, 0]^\top$, то $\Phi(\sigma) = 1$.

Функция $\Phi(\sigma)$ с такими свойствами называется *симметричной калибровочной функцией* на $\mathbb{R}^k$.

Если $\Phi(\sigma)$ определяется унитарно инвариантной нормой как $||\Sigma||$, то эти свойства, очевидно, должны выполняться. Нетривиальная часть теоремы Джона фон Неймана — в том, что *любая симметричная калибровочная функция определяет унитарно инвариантную норму.* Единственную (но ощутимую) трудность доставляет получение неравенства треугольника.

[3] Любопытный исторический факт: данный результат был опубликован автором в Ученых записках Томского университета.

Лекция 36

ОСНОВНАЯ ЧАСТЬ

36.1 Квадратичные формы

Выражение $f = \sum\limits_{1 \le i,j \le n} a_{ij} x_i x_j$ называется *квадратичной формой* от переменных $x_1, \ldots, x_n$. При $i \neq j$ в сумме имеются два члена, для которых

$$a_{ij}\, x_i x_j \;+\; a_{ji}\, x_j x_i \;=\; \frac{a_{ij} + a_{ji}}{2}\,(x_i x_j + x_j x_i).$$

Поэтому, не ограничивая общности, всегда полагают, что $a_{ij} = a_{ji}$.

Квадратичные формы успешно изучались еще до введения понятия матрицы. Современный подход, конечно, использует матрицы — они возникают здесь естественным образом:

$$f = x^\top A x, \quad \text{где} \quad A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}.$$

Матрица A называется *матрицей квадратичной формы* f. Согласно нашей договоренности, $a_{ij} = a_{ji}$ — поэтому матрица A симметричная.

36.2 Конгруэнтность

Замена переменных $x = Py$ с помощью невырожденной матрицы P делает f квадратичной формой от новых переменных:

$$f = x^\top A x = (Py)^\top A (Py) = y^\top (P^\top A P) y.$$

Матрицы A и B, связанные равенством $B = P^\top A P$ для некоторой невырожденной матрицы P, называются *конгруэнтными*. Легко видеть, что отношение конгруэнтности есть отношение эквивалентности на множестве матриц фиксированного порядка.

Квадратичные формы от трех переменных нам уже встречались при изучении поверхностей второго порядка. В этом случае переменные были вещественными координатами, а матрица A — вещественной симметричной матрицей. Тогда нас особенно интересовали декартовы системы координат — поэтому требовалось, чтобы матрица P была ортогональной. Как следствие, переход от A к B в данном случае является одновременно преобразованием конгруэнтности и подобия.

36.3 Канонический вид квадратичной формы

Мы знаем, что любая вещественная симметричная матрица ортогонально подобна вещественной диагональной матрице:

$$\Lambda = P^\top A P, \qquad P^\top = P^{-1}, \qquad P \in \mathbb{R}^{n \times n}.$$

В новых переменных квадратичная форма f оказывается алгебраической суммой квадратов

$$f = \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2.$$

В общем случае от P можно требовать лишь невырожденности. Поиск соответствующей замены переменных (матрицы P) для заданной квадратичной формы называется *приведением к каноническому виду*. Если P — ортогональная матрица, то говорят о *приведении f к главным осям*.

Если $r = \mathrm{rank}\Lambda = \mathrm{rank}A$, то в данной сумме можно оставить только r членов, отвечающих $\lambda_i \neq 0$. Не ограничивая общности, можно считать, что

$$\lambda_1, \ldots, \lambda_k > 0, \qquad \lambda_{k+1}, \ldots, \lambda_r < 0, \qquad \lambda_{r+1} = \ldots = \lambda_n = 0.$$

Очевидно, k, $r-k$ и $n-r$ равны, соответственно, числу положительных, отрицательных и нулевых собственных значений матрицы A.

Тройка чисел $(k, r-k, n-r)$ называется *инерцией* вещественной симметричной матрицы A. Точно так же вводится понятие инерции для произвольной эрмитовой матрицы.

36.4 Закон инерции

Пусть все матрицы вещественные.

Теорема. *Вещественные симметричные матрицы конгруэнтны тогда и только тогда, когда они имеют одинаковую инерцию.*

Доказательство. Достаточно доказать совпадение инерций для конгруэнтных вещественных диагональных матриц. Пусть это матрицы Λ и $D = P^\top \Lambda P$, где P — вещественная невырожденная матрица. Конечно, D и Λ имеют общий ранг r. Пусть инерция D равна $(l, r-l, n-r)$, а инерция Λ равна $(k, r-k, n-r)$. Предположим, что

$$d_1, \ldots, d_l > 0, \quad d_{l+1}, \ldots, d_r < 0; \qquad \lambda_1, \ldots, \lambda_k > 0, \quad \lambda_{k+1}, \ldots, \lambda_r < 0.$$

Равенство $y^\top D y = x^\top \Lambda x$ при условии $x = Py$ означает, что

$$(d_1\, y_1 \ + \ \ldots \ + \ d_l\, y_l^2) \ + \ (d_{l+1}\, y_{l+1}^2 \ + \ \ldots \ + \ d_r\, y_r^2) \ =$$

$$(\lambda_1\, x_1 \ + \ \ldots \ + \ \lambda_k\, x_k^2) \ + \ (\lambda_{k+1}\, x_{k+1}^2 \ + \ \ldots \ + \ \lambda_r\, x_r^2). \qquad (*)$$

Рассмотрим два подпространства:

$$L = \{y \in \mathbb{R}^n : y_{l+1} = \ldots = y_r = 0\}, \qquad M = \{y \in \mathbb{R}^n : y = P^{-1}x, \ \ x_1 = \ldots = x_k = 0\}.$$

Легко видеть, что $\dim L = l$. Поскольку $y = P^{-1}x$, ясно, что $\dim M = n - k$. Если $l > k$, то $\dim L + \dim M > n \ \Rightarrow$ существует ненулевой вектор $y \in L \cap M$. Для этого вектора y левая часть в равенстве $(*)$ строго положительна, а правая часть отрицательна или равна нулю. Противоречие означает, что $l \leq k$. Противоположное неравенство тоже верно —достаточно поменять ролями x и y. $\square$

36.5 Эрмитова конгруэнтность

Комплексные матрицы A и B называются *эрмитово конгруэнтными*, если $B = P^* A P$ для некоторой невырожденной матрицы P. Это отношение эквивалентности на множестве $n \times n$-матриц (докажите!). Если матрица A эрмитова, то и B эрмитова.

Теорема. *Эрмитовы матрицы эрмитово конгруэнтны тогда и только тогда, когда они имеют одинаковую инерцию.*

Доказательство практически дословно повторяет предыдущее доказательство (надо лишь вместо x_i^2 и y_i^2 писать $|x_i|^2$ и $|y_i|^2$).

36.6 Канонический вид пары квадратичных форм

Если приходится одновременно иметь дело с парой поверхностей второго порядка в пространстве или с парой кривых второго порядка на плоскости, то разумно пытаться упростить их уравнения в одной и той же системе координат. В общем случае эта система координат будет аффинной.

Для простоты рассмотрим случай кривых на плоскости. Предположим, что одна из кривых является эллипсом. Тогда перейдем к такой декартовой системе, в который для нее получается уравнение $x^2/a^2 + y^2/b^2 = 1$. Уравнение второй кривой в этой системе может иметь самый общий вид. Изменив масштабы по осям, перейдем к аффинной системе, в которой уравнением эллипса будет уравнение окружности $(x')^2 + (y')^2 = 1$. Уравнение второй кривой в новой (аффинной) системе имеет все еще общий вид. Но с помощью поворота, как мы знаем, для его квадратичной части можно получить форму

$\lambda_1(x'')^2+\lambda_2(y'')^2$. При этом поворот системы координат не может изменить формы первого уравнения! В сущности это же рассуждение переносится на более общий случай.

Теорема 1. *Пусть A и B — вещественные симметричные матрицы и при этом A положительно определенная. Тогда существует вещественная невырожденная матрица P такая, что матрицы $P^\top AP$ и $P^\top BP$ обе диагональные.*

Доказательство. Вещественная симметричная матрица A ортогонально подобна (поэтому и конгруэнтна) диагональной матрице

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = Q^\top AQ, \quad Q^\top = Q^{-1}.$$

В силу положительной определенности, $\lambda_i > 0$ для всех i. Далее заметим, что A конгруэнтна единичной матрице (по определению, $\Lambda^{-1/2} \equiv (\Lambda^{1/2})^{-1}$):

$$I = \Lambda^{-1/2}Q^\top AQ\Lambda^{-1/2} = (Q\Lambda^{-1/2})^\top A(Q\Lambda^{-1/2}).$$

Пусть то же преобразование конгруэнтности в применении к B дает матрицу

$$C = (Q\Lambda^{-1/2})^\top B(Q\Lambda^{-1/2}).$$

Легко проверить, что C остается вещественной симметричной матрицей. Следовательно, с помощью ортогональной матрицы Z получаем диагональную матрицу $D = Z^\top CZ$. В то же время, $Z^\top IZ = I$. Окончательно,

$$I = P^\top AP, \quad D = P^\top BP, \quad \text{где} \quad P = Q\Lambda^{-1/2}Z. \qquad \square$$

Следствие. *Пусть $f(x)$ и $g(x)$ — вещественные квадратичные формы и $f(x) > 0$ для всех вещественных векторов $x \neq 0$. Тогда f и g можно привести к каноническому виду с помощью общей замены переменных.*

Вот вариант этой же теоремы в случае эрмитовых матриц и преобразования эрмитовой конгруэнтности — имеющееся доказательство модифицируется очевидным образом.

Теорема 2. *Пусть A и B — эрмитовы матрицы и A положительно определенная. Тогда существует невырожденная матрица P такая, что матрицы P^*AP и P^*BP обе диагональные.*

36.7 Метод Лагранжа

Простая идея, позволяющая получить канонический вид квадратичной формы, связана с выделением полных квадратов. В итоге вещественная симметричная матрица A приводится к конгруэнтной диагональной матрице $\Lambda = P^\top AP$ с помощью вещественной невырожденной матрицы P.

Эта идея ведет к так называемому методу Лагранжа. Чтобы понять его суть, рассмотрим квадратичную форму

$$f = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3.$$

Если $a_{11} \neq 0$, то полный квадрат выделяется следующим образом:

$$f = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}x_3\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 + \left(a_{33} - \frac{a_{13}^2}{a_{11}}\right)x_3^2 + 2\left(a_{23} - \frac{a_{12}a_{13}}{a_{11}}\right)x_2x_3$$

$$= b_{11}\,y_1^2 \;+\; b_{22}\,y_2^2 \;+\; b_{33}\,y_3^2 \;+\; 2b_{23}\,y_2y_3,$$

$$b_{11} = a_{11}, \quad b_{22} = a_{22} - \frac{a_{12}^2}{a_{11}}, \quad b_{33} = a_{33} - \frac{a_{13}^2}{a_{11}}, \quad b_{23} = a_{23} - \frac{a_{12}a_{13}}{a_{11}},$$

$$y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \frac{a_{13}}{a_{11}}, \quad y_2 = x_2, \quad y_3 = x_3.$$

Таким образом, A конгруэнтна матрице

$$B = \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & b_{23} \\ 0 & b_{23} & b_{33} \end{bmatrix} = P_1^\top AP_1, \qquad P_1 = \begin{bmatrix} 1 & a_{12}/a_{11} & a_{13}/a_{11} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Следующий шаг очевиден — с помощью выделения полного квадрата исключить произведение y_2y_3.

С помощью метода Лагранжа можно найти инерцию матрицы A. Если же нужно получить ортогональную матрицу P, то следует обратиться к другим методам — например, к методу вращений.

Мы не будем здесь заниматься формализацией метода Лагранжа для симметричных матриц общего вида. Вместо этого мы рассмотрим случай вещественных положительно определенных матриц и метод квадратного корня — с помощью преобразований того же типа он решает ту же задачу, что и метод Лагранжа.

36.8 Метод квадратного корня

Пусть дана матрица A порядка n и A_k — ее $k \times k$-подматрица, расположенная на пересечении первых k строк и столбцов. Подматрицы $A_1, \ldots, A_n = A$ называются *ведущими подматрицами*, а их определители — *ведущими минорами* матрицы A.

Для вещественной симметричной матрицы A, в которой все ведущие миноры положительны, имеет место разложение $A = R^\top R$, где R — вещественная верхняя треугольная матрица с положительными диагональными элементами. [1]

Предположим, что факт существования разложения уже доказан. Тогда нетрудно понять, как его можно вычислить. Для матрицы порядка $n = 3$ имеем

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} = \begin{bmatrix} r_{11} & & \\ r_{12} & r_{22} & \\ r_{13} & r_{23} & r_{33} \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ & r_{22} & r_{23} \\ & & r_{33} \end{bmatrix} \quad \Rightarrow$$

$$r_{11} = \sqrt{a_{11}}, \quad r_{12} = a_{12}/r_{11}, \quad r_{13} = a_{13}/r_{11},$$

$$r_{22} = \sqrt{a_{22} - r_{12}^2}, \quad r_{23} = (a_{23} - r_{13}r_{12})/r_{22}, \quad r_{33} = \sqrt{a_{33} - r_{13}^2 - r_{23}^2}.$$

Вычисления аналогичны и в случае произвольного n. Метод называется *методом квадратного корня*.

Интересно, что в данном случае "как бы" не используется идея исключения элементов, но именно "как бы": чтобы объяснить, почему можно извлекать корни, проще всего вернуться к идее метода Гаусса.

Теорема. *Пусть A — матрица порядка n, в которой все ведущие миноры отличны от нуля. Тогда существуют единственные нижняя треугольная матрица L с единицами на диагонали и верхняя треугольная матрица U такие, что $A = LU$.*

Доказательство. Пусть $n = 3$. Первый шаг метода Гаусса дает

$$\begin{bmatrix} 1 & 0 & 0 \\ -l_{21} & 1 & 0 \\ -l_{31} & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & b_{32} & b_{33} \end{bmatrix}, \qquad l_{21} = a_{21}/a_{11}, \quad l_{31} = a_{31}/a_{11}.$$

$$\Rightarrow \quad \left[\begin{array}{cc|c} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \hline a_{31} & a_{32} & a_{33} \end{array}\right] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ \hline l_{31} & 0 & 1 \end{array}\right] \left[\begin{array}{cc|c} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ \hline 0 & b_{32} & b_{33} \end{array}\right] \Rightarrow \det A_2 = a_{11}\, b_{22} \quad \Rightarrow \quad b_{22} \neq 0.$$

Поскольку $b_{22} \neq 0$, можно обойтись без перестановок строк и перейти ко второму шагу метода Гаусса:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -l_{31} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & c_{33} \end{bmatrix}, \quad l_{31} = b_{32}/b_{22}.$$

В итоге получаем

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & c_{33} \end{bmatrix}.$$

Заметим, что $\det A_3 = a_{11}b_{22}c_{33} \quad \Rightarrow \quad c_{33} \neq 0$ (это гарантирует возможность проведения третьего шага метода Гаусса без перестановок строк в случае $n > 3$). Единственность построенного LU-разложения

[1] В вычислительной алгебре разложение такого вида называют *разложением Холецкого*.

проверяется непосредственно: первая строка в U и первый столбец в L определены однозначно, отсюда то же самое получаем для второй строки в U и второго столбца в L, и так далее. Обобщение доказательства на случай произвольного n не представляет никакой трудности. $\square$

Следствие. *Для любой вещественной симметричной матрицы, в которой все ведущие миноры положительны, существует вещественная верхняя треугольная матрица R такая, что $A = R^\top R$. Элементы главной диагонали R могут быть выбраны положительными, при этом ограничении R единственна.*

Доказательство. Воспользуемся существованием и единственностью LU-разложения $A = LU$, в котором L имеет единицы на главной диагонали. Пусть D — диагональная матрица с главной диагональю, взятой из матрицы $U = [u_{ij}]$. Поскольку $\det A_k = u_{11} \dots u_{kk}$ для всех k, находим, что $u_{kk} > 0$ для всех k.

В силу симметричности матрицы A,

$$A = A^\top = LU = (U^\top D^{-1})(DL) \quad \Rightarrow \quad L = U^\top D^{-1}.$$

Отсюда $A = (D^{-1/2}U)^\top (D^{-1/2}U)$. Таким образом, $R = D^{-1/2}U$. Единственность проверяется непосредственно — так же, как в случае LU-разложения. $\square$

Замечание. Определитель вещественной симметричной положительно определенной матрицы положителен (как произведение положительных собственных значений). Легко показать, что свойство положительной определенности наследуется всеми ведущими подматрицами $\Rightarrow$ все ее ведущие миноры положительны. Поэтому метод квадратного корня можно применять для любой вещественной симметричной положительно определенной матрицы. Метод квадратного корня легко переносится также на случай комплексных положительно определенных матриц (они обязательно эрмитовы). Для таких матриц всегда имеет место разложение $A = R^* R$, где R — комплексная верхняя треугольная матрица с положительными диагональными элементами.

Задача. Доказать, что для любой положительно определенной матрицы $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ имеет место неравенство

$$\det A \ \le \ a_{11} a_{22} \dots a_{nn}.$$

36.9 Критерий положительной определенности

Докажем важный результат, известный как *критерий Сильвестра.*

Теорема. *Пусть дана эрмитова матрица. Для ее положительной определенности необходимо и достаточно, чтобы все ее ведущие миноры были положительны.*

Доказательство. Необходимость вытекает из того, что свойство положительной (и неотрицательной) определенности эрмитовой матрицы A порядка n наследуется ее ведущими подматрицами $A_1, \dots, A_n$ — нужно лишь учесть равенство

$$[x_1, \dots, x_k]\, A_k \begin{bmatrix} x_1 \\ \dots \\ x_k \end{bmatrix} \ = \ [x_1, \dots, x_k, 0, \dots, 0]\, A \begin{bmatrix} x_1 \\ \dots \\ x_k \\ 0 \\ \dots \\ 0 \end{bmatrix}.$$

Из положительной определенности матрицы A_k следует, что все ее собственные значения положительны $\Rightarrow$ $\det A_k > 0$ (как произведение положительных собственных значений). Достаточность получается из разложения $A = R^* R$, где R — верхняя треугольная матрица: для любого $x \ne 0$ получаем $x^* A x = x^*(R^* R)x = (Rx)^*(Rx) > 0$. $\square$

36.10 Гиперповерхности второго порядка

Рассмотрим в $\mathbb{R}^n$ множество точек S с координатами $x_1, \dots, x_n$, удовлетворяющими уравнению

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j - 2 \sum_{k=1}^{n} b_k x_k + c \ = \ 0,$$

или, в матричной форме,

$$f(x) \equiv (Ax, x) - 2(b, x) + c \;=\; 0, \qquad A = [a_{ij}], \quad b = \begin{bmatrix} b_1 \\ ... \\ b_n \end{bmatrix}, \quad (x, y) \equiv y^\top x.$$

Все коэффициенты предполагаются вещественными и, кроме того, $a_{ij} = a_{ji} \;\Rightarrow\; A = A^\top$. Если $A \neq 0$, то множество решений данного уравнения называется *гиперповерхностью второго порядка*.

Как и любая вещественная симметричная матрица, A конгруэнтна и даже ортогонально подобна диагональной матрице $\Lambda = P^\top AP$, где P — ортогональная матрица. Замена переменных $x = Py$ приводит уравнение $f(x) = 0$ к виду

$$(\Lambda y, y) - 2(d, y) + c \;=\; 0 \quad \Leftrightarrow \quad \lambda_1 y_1^2 + \ldots + \lambda_r y_r^2 \;-\; 2d_1 y_1 - \ldots - 2d_n y_n + c \;=\; 0,$$

где $d = P^\top b$, r — ранг матрицы Λ, а $\lambda_1, \ldots, \lambda_r$ — ее отличные от нуля элементы (ненулевые собственные значения матрицы A). Последнее уравнение с помощью сдвигов $z_i = y_i - d_i/\lambda_i, \quad 1 \le i \le r, \quad z_i = y_i, \quad r+1 \le i \le n$, приводится к виду

$$\lambda_1 z_1^2 + \ldots + \lambda_r z_r^2 \;-\; 2d_{r+1} z_{r+1} - \ldots - 2d_n z_n + h \;=\; 0,$$

$$h = c - d_1^2/\lambda_1^2 - \ldots - d_r^2/\lambda_r^2.$$

Если $d_{r+1} = \ldots = d_n = 0$, то данное уравнение имеет уже достаточно простой вид

$$\lambda_1 z_1^2 \;+\; \ldots \;+\; \lambda_r z_r^2 \;+\; h \;=\; 0. \tag{1}$$

В противном случае какое-то из чисел $d_{r+1}, \ldots, d_n$ отлично от нуля. Пусть $d_{r+1} \neq 0$. Тогда существует ортогональная матрица Q блочного вида

$$Q = \begin{bmatrix} I_r & 0 \\ 0 & \widetilde{Q} \end{bmatrix},$$

где $\widetilde{Q}$ — ортогональная матрица порядка $n - r$ и при этом

$$\widetilde{Q}^\top \begin{bmatrix} d_{r+1} \\ d_{r+2} \\ ... \\ d_n \end{bmatrix} \;=\; \mu \begin{bmatrix} 1 \\ 0 \\ ... \\ 0 \end{bmatrix}.$$

Матрицу $\widetilde{Q}^\top$ можно получить как произведение матриц вращения. Если $z = Qu$ и $\hat{d} = [0, ..., 0, d_{r+1}, ..., d_n]^\top$, то $(\hat{d}, y) = (Q^\top \hat{d}, u) = \mu\, u_{r+1} \quad \Rightarrow \quad$ замена $z = Qu$ дает уравнение вида

$$\lambda_1 u_1^2 \;+\; \ldots \;+\; \lambda_r u_r^2 \;-\; 2\mu\, u_{r+1} \;+\; h \;=\; 0.$$

Ясно, что $\mu \neq 0$ (почему?). Поэтому можно выполнить сдвиг $w_{r+1} = u_{r+1} - h/(2\mu), \;\; w_i = u_i, \;\; i \neq r+1$, и получить уравнение

$$\lambda_1 w_1^2 \;+\; \ldots \;+\; \lambda_r w_r^2 \;-\; 2\mu\, w_{r+1} \;=\; 0. \tag{2}$$

Уравнения (1) и (2) называются *приведенными уравнениями* гиперповерхности S. Из нашего обсуждения ясно, что они получаются с помощью перехода к другому ортономированному базису и сдвига начала координат. Отказавшись от ортонормированности, можно получить уравнения такого же вида, в которых $\lambda_i = \pm 1$. Выбор соответствующей системы координат связан с приведением квадратичной формы (Ax, x) к каноническому виду; в силу закона инерции число положительных и отрицательных коэффициентов не зависит от способа приведения.

В заключение обсудим интересную связь между геометрическими свойствами гиперповерхности S и множеством решений системы $Ax = b$. Фиксируем точку $x_0 \in \mathbb{R}^n$ и рассмотрим прямую $x_0 + tv, \;\; t \in \mathbb{R}$, с направляющим вектором $v \neq 0$. Ее точки пересечения с гиперповерхностью S определяются квадратным уравнением

$$(A(x_0 + tv),\, x_0 + tv) \;-\; 2(b,\, x_0 + tv) \;+\; c \;=\; 0 \quad \Leftrightarrow$$

$$(Av, v)\, t^2 \;-\; 2(b - Ax_0, v)\, t \;+\; f(x_0) \;=\; 0. \qquad (*)$$

Говорят, что вектор v имеет *асимптотическое* направление относительно S, если $(Av, v) = 0$, и *неасимптотическое* направление, если $(Av, v) \neq 0$.

Пусть v имеет неасимптотическое направление и $x_0 \in S$. В этом случае $f(x_0) = 0 \;\Rightarrow\;$ уравнение $(*)$ имеет два (возможно, совпадающих) решения: при $t = 0$ и $t = 2(b - Ax_0, v)/(Av, v)$. Точка

$$z = x_0 + ((b - Ax_0, v)/(Av, v))\, v \qquad (**)$$

является, очевидно, серединой отрезка, параллельного v и соединяющего две точки из S. Такой отрезок называется *хордой* для S с направляющим вектором v. Умножив $(**)$ скалярно на Av и заметив, что $(Av, z) = (Az, v)$, находим

$$(Az, v) \;=\; (b, v). \qquad (\#)$$

Вывод: *все точки z, являющиеся серединами всевозможных хорд для S с фиксированным неасимптотическим направлением v, принадлежат гиперплоскости* $(\#)$. Данная гиперплоскость называется *диаметральной гиперплоскостью*, сопряженной вектору v относительно гиперповерхности S.

Точка z называется *центром симметрии* для S, если $z + p \in S$ в том и только том случае, когда $z - p \in S$.

Утверждение. *Совместность системы $Ax = b$ с произвольной вещественной симметричной матрицей A равносильна существованию центра симметрии у гиперповерхности $f(x) = 0$. Множество всех центров симметрии совпадает с множеством всех решений системы $Ax = b$.*

Доказательство. Пусть $Az = b \;\Rightarrow\; (Av, z) = (b, v)$ для любого неасимптотического вектора $v \;\Rightarrow\;$ z принадлежит пересечению *всех* диаметральных гиперплоскостей $\;\Rightarrow\;$ z является серединой *любой* хорды (а значит, и центром симметрии) для S.

Теперь предположим, что z — центр симметрии для $S \;\Rightarrow$

$$(A(z + p), z + p) - 2(b, z + p) = (A(z - p), z - p) - 2(b, z - p) \quad\Rightarrow\quad (Az - b, p) = 0.$$

Легко показать (например, с помощью приведенных уравнений), что существуют n линейно независимых неасимптотических векторов $v_1, \ldots, v_n$. Тогда точки $x_0,\ x_1 = x_0 + v_1, \ldots, x_n = x_0 + v_n \in S$ будут аффинно независимыми (см. раздел 13.6). Пусть точка $x_0 \in S$ такова, что $b - Ax_0 \neq 0$. Из $(*)$ ясно, что v_i можно выбрать таким образом, что все x_i будут принадлежать S. Легко видеть, что векторы (точки) $x_i - z$, $0 \leq i \leq n$, будут аффинно независимыми. Поэтому из них можно выбрать подсистему из n линейно независимых векторов (см. задачу из раздела 13.6). Следовательно, существуют n линейно независимых векторов p таких, что $z + p \in S \;\Rightarrow\; Az = b$. $\square$

Лекция 37

ОСНОВНАЯ ЧАСТЬ

37.1 Собственные значения эрмитовой подматрицы

Пусть эрмитова матрица $A \in \mathbb{C}^{n\times n}$ записана в блочном виде

$$A \;=\; \begin{bmatrix} B & u \\ u^* & a_{nn} \end{bmatrix}, \quad B \in \mathbb{C}^{(n-1)\times(n-1)}, \;\; u \in \mathbb{C}^{n-1}. \tag{1}$$

Ясно, что подматрица B тоже эрмитова. Пусть $\mu_1 \geq \ldots \geq \mu_{n-1}$ — ее собственные значения, и пусть Q — унитарная матрица порядка $n-1$, приводящая ее к диагональному виду

$$Q^*BQ = \begin{bmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_{n-1} \end{bmatrix} \quad \Rightarrow$$

$$\begin{bmatrix} Q^* & \\ & 1 \end{bmatrix} \begin{bmatrix} B & u \\ u^* & a_{nn} \end{bmatrix} \begin{bmatrix} Q & \\ & 1 \end{bmatrix} = \begin{bmatrix} \mu_1 & & & s_1 \\ & \ddots & & \\ & & \mu_{n-1} & s_{n-1} \\ \bar{s}_1 & \ldots & \bar{s}_{n-1} & s_n \end{bmatrix}, \quad \begin{bmatrix} s_1 \\ \ldots \\ s_{n-1} \end{bmatrix} = Q^*u, \quad s_n = \bar{s}_n = a_{nn}.$$

Характеристический многочлен матрицы A легко вычисляется:

$$\det(A-\lambda I) = \begin{bmatrix} \mu_1-\lambda & & & s_1 \\ & \ddots & & \\ & & \mu_{n-1}-\lambda & s_{n-1} \\ \bar{s}_1 & \ldots & \bar{s}_{n-1} & s_n-\lambda \end{bmatrix}$$

$$= \prod_{i=1}^{n-1}(\mu_i-\lambda)\left(s_n - \lambda - \frac{|s_1|^2}{\mu_1-\lambda} - \ldots - \frac{|s_{n-1}|^2}{\mu_{n-1}-\lambda}\right).$$

Следовательно, если собственное значение λ матрицы A не совпадает ни с одним из собственных значений $\mu_1, \ldots, \mu_{n-1}$ ее подматрицы B, то оно удовлетворяет уравнению

$$\lambda \;=\; F(\lambda) \equiv \frac{|s_1|^2}{\lambda-\mu_1} + \ldots + \frac{|s_{n-1}|^2}{\lambda-\mu_{n-1}} + s_n.$$

Утверждение. *Пусть эрмитова матрица A порядка n с собственными значениями $\lambda_1 \geq \ldots \geq \lambda_n$ имеет блочное разбиение* (1)*, в котором B — ее эрмитова подматрица порядка $n-1$ с собственными значениями $\mu_1 \geq \ldots \geq \mu_{n-1}$. Тогда если*

$$\mu_1 > \mu_2 > \ldots > \mu_{n-1} \quad u \quad s_i \neq 0, \;\; 1 \leq i \leq n-1,$$

то имеют место соотношения разделения

$$\lambda_1 > \mu_1 > \lambda_2 > \mu_2 > \ldots > \lambda_{n-1} > \mu_{n-1} > \lambda_n. \tag{2}$$

Доказательство. Рассмотрим график функции $y = F(\lambda)$ (λ и y — переменные осей абсцисс и ординат). Очевидно, $F(\lambda)$ не определено при $\lambda = \mu_k$. Поскольку $F(\lambda) \to \infty$ при $\lambda \to \mu_k$, естественно говорить, что $F(\lambda)$ при $\lambda = \mu_k$ обращается в бесконечность. Изучим поведение функции $F(\lambda)$ на каждом из n интервалов

$$I_n = (-\infty, \mu_{n-1}), \quad I_{n-1} = (\mu_{n-1}, \mu_{n-2}), \quad \ldots, \quad I_2 = (\mu_2, \mu_1), \quad I_1 = (\mu_1, +\infty).$$

Пусть $\lambda \in I_k$, $2 \le k \le n-1$. Тогда

$$\frac{|s_k|^2}{\lambda - \mu_k} + \frac{|s_{k-1}|^2}{\lambda - \mu_{k-1}} \to \begin{cases} +\infty & \text{при} \quad \lambda \to \mu_k, \\ -\infty & \text{при} \quad \lambda \to \mu_{k-1}, \end{cases}$$

а остальные слагаемые в представлении $F(\lambda)$ являются ограниченными. Поэтому

$$F(\lambda) \to \begin{cases} +\infty & \text{при} \quad \lambda \to \mu_k, \\ -\infty & \text{при} \quad \lambda \to \mu_{k-1}. \end{cases}$$

В силу непрерывности $F(\lambda)$, прямая $y = \lambda$ имеет при $\lambda \in I_k$ точку пересечения с графиком функции $y = F(\lambda)$. Случаи $\lambda \in I_1$ и $\lambda \in I_n$ рассматриваются аналогично. Таким образом, уравнение $F(\lambda) = \lambda$ имеет n различных корней. Ни один из них не совпадает ни с одним из чисел μ_k и поэтому каждый из них является собственным значением матрицы A. $\square$

Если B имеет кратные собственные значения или $s_k = 0$ для каких-то k, строгие неравенства в соотношениях разделения (2) следует заменить на нестрогие неравенства. Можно было бы рассуждать таким образом: с помощью сколь угодно малых возмущений можно сделать $\mu_1, \ldots, \mu_{n-1}$ попарно различными, а все s_k ненулевыми, при этом для возмущенной матрицы A можно применить доказанное утверждение, а затем перейти к пределу. Чтобы это рассуждение сделать строгим, требуется факт непрерывной зависимости собственных значений матрицы от ее коэффициентов. Этот важный факт действительно имеет место. Но мы пойдем другим путем — случай нестрогих неравенств легко анализируется на основе вариационных свойств собственных значений эрмитовой матрицы.

37.2 Вариационные свойства собственных значений

Под вариационными свойствами понимаются свойства, связанные с минимальными или максимальными значениями каких-то функций. В случае эрмитовой матрицы $A \in \mathbb{C}^{n\times n}$ в качестве такой функции от векторов $x \in \mathbb{C}^n$ рассматривается так называемое *отношение Рэлея*

$$\Phi_A(x) = \frac{x^* A x}{x^* x}, \quad x \ne 0.$$

Лемма. *В любом подпространстве $L \subset \mathbb{C}^n$ существуют векторы $x_{\min}(L)$ и $x_{\max}(L)$, принадлежащие L и такие, что*

$$\Phi_A(x_{\min}) \le \Phi_A(x) \le \Phi_A(x_{\max}) \quad \forall \ \ x \in L, \ x \ne 0.$$

Доказательство. Функция $\Phi_A(x)$ непрерывна на единичной сфере $||x||_2 = 1$ конечномерного пространства L. По теореме Вейерштрасса, она принимает там наименьшее и наибольшее значение в каких-то точках $x_{\min}$ и $x_{\max}$. Легко проверить, что эти точки являются искомыми. $\square$

Теорема Куранта–Фишера. *Собственные значения $\lambda_1(A) \ge \ldots \ge \lambda_n(A)$ эрмитовой матрицы $A \in \mathbb{C}^{n\times n}$ связаны с отношением Рэлея $\Phi_A(x)$ следующим образом:*

$$\lambda_k(A) \ = \ \max_{\dim L = k} \ \min_{x\in L,\, x\ne 0} \Phi_A(x) \ = \ \min_{\dim L = n-k+1} \ \max_{x\in L,\, x\ne 0} \Phi_A(x). \tag{3}$$

Доказательство. Пусть $v_1, \ldots, v_n \in \mathbb{C}^n$ — ортонормированный базис собственных векторов матрицы A: $Av_i = \lambda_i v_i$, $1 \le i \le n$.

Пусть $L_k = L(v_1, \ldots, v_k)$ и $x = \alpha_1 v_1 + \ldots + \alpha_k v_k \in L_k$, $x \ne 0$. $\Rightarrow$

$$\Phi_A(x) = \frac{\lambda_1|\alpha_1|^2 + \ldots + \lambda_k|\alpha_k|^2}{|\alpha_1|^2 + \ldots + |\alpha_k|^2} \ge \lambda_k, \quad \Phi_A(v_k) = \lambda_k \quad \Rightarrow \quad \min_{x\in L_k,\, x\ne 0} \Phi_A(x) = \lambda_k.$$

Рассмотрим также подпространство $M_k = L(v_k, \ldots, v_n)$ размерности $n-k+1$. Пусть $x = \alpha_k v_k + \ldots + \alpha_n v_n \in M_k$, $x \neq 0 \quad \Rightarrow$

$$\Phi_A(x) = \frac{\lambda_k|\alpha_k|^2 + \ldots + \lambda_n|\alpha_n|^2}{|\alpha_k|^2 + \ldots + |\alpha_n|^2} \leq \lambda_k, \qquad \Phi_A(v_k) = \lambda_k \quad \Rightarrow \quad \max_{x\in M_k,\, x\neq 0} \Phi_A(x) = \lambda_k.$$

Пусть теперь L — произвольное подпространство размерности k. В силу теоремы Грассмана, $\dim(L \cap M_k) \geq 1 \quad \Rightarrow \quad$ существует ненулевой вектор $z \in (L \cap M_k)$. Тогда

$$\min_{x\in L,\, x\neq 0} \Phi_A(x) \;\leq\; \Phi_A(z) \;\leq \max_{x\in M_k,\, x\neq 0} \Phi_A(x) = \lambda_k.$$

Таким образом, первое из соотношений (3) доказано.

Чтобы получить второе соотношение, возьмем произвольное подпространство L размерности $n-k+1$. Тогда существует ненулевой вектор $z \in L \cap L_k \quad \Rightarrow$

$$\max_{x\in L,\, x\neq 0} \Phi_A(x) \;\geq\; \Phi_A(z) \;\geq \min_{x\in L_k,\, x\neq 0} \Phi_A(x) = \lambda_k. \qquad \square$$

37.3 Соотношения разделения

Теорема. *Пусть эрмитова матрица $A \in \mathbb{C}^{n\times n}$ имеет собственные значения*

$$\lambda_1 \geq \ldots \geq \lambda_n,$$

и пусть $B \in \mathbb{C}^{(n-1)\times(n-1)}$ — ее эрмитова подматрица в блочном разбиении вида (1)*, имеющая собственные значения*

$$\mu_1 \geq \ldots \geq \mu_{n-1}.$$

Тогда имеют место соотношения разделения

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \ldots \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n.$$

Доказательство. Обозначим через M подпространство векторов $x = [x_1, \ldots, x_n]^\top$, определяемое уравнением $x_n = 0$. Пусть отображение $\nu : \mathbb{C}^n \to \mathbb{C}^{n-1}$ задается правилом $\nu(x) = [x_1 \ \ldots, \ x_{n-1}]^\top$. Тогда очевидно, что если $x \in M$, то $\Phi_A(x) = \Phi_B(\nu(x))$.

Пусть $1 \leq k \leq n-1$. Согласно теореме Куранта–Фишера, находим

$$\lambda_k \quad = \quad \max_{\dim L = k} \min_{x\in L,\, x\neq 0} \Phi_A(x) \;\geq\; \max_{\dim L = k,\, L\subset M} \min_{x\in L,\, x\neq 0} \Phi_A(x) =$$

$$\max_{\dim L = k,\, L\subset M} \min_{x\in L,\, x\neq 0} \Phi_B(\nu(x)) \;=\; \max_{\dim L = k,\, L\subset \mathbb{C}^{n-1}} \min_{y\in L,\, y\neq 0} \Phi_B(y) \;=\; \mu_k.$$

Пусть теперь $2 \leq k \leq n$. Согласно той же теореме Куранта–Фишера,

$$\lambda_k \quad = \quad \min_{\dim L = n-k+1} \max_{x\in L,\, x\neq 0} \Phi_A(x) \;\leq\; \min_{\dim L = n-k+1,\, L\subset M} \max_{x\in L,\, x\neq 0} \Phi_A(x) =$$

$$\min_{\dim L = n-k+1,\, L\subset M} \max_{x\in L,\, x\neq 0} \Phi_B(\nu(x)) =$$

$$\min_{\substack{\dim L = (n-1)-(k-1)+1 \\ L\,\subset\, \mathbb{C}^{n-1}}} \max_{y\in L,\, y\neq 0} \Phi_B(y) \;=\; \mu_{k-1}. \quad \square$$

В качестве простого следствия можно получить еще одно доказательство достаточности уже известного нам критерия положительной определенности эрмитовой матрицы: *для положительной определенности необходимо и достаточно, чтобы все ее ведущие миноры были положительны.*

Пусть $\lambda_{1k} \geq \ldots \geq \lambda_{kk}$ — собственные значения ведущей подматрицы A_k порядка k. Достаточно доказать, что $\lambda_{kk} > 0$. Пусть известно, что

$$\det A_k = \lambda_{11} \ldots \lambda_{1k} > 0, \quad 1 \leq k \leq n.$$

Очевидно, $\lambda_{11} > 0$. Пусть уже доказано, что $\lambda_{k-1\,k-1} > 0$. В силу соотношений разделения, $\lambda_{k-1\,k} \geq \lambda_{k-1\,k-1} > 0$. Далее,

$$\det A_k \;=\; (\lambda_{1k} \ldots \lambda_{k-1\,k})\, \lambda_{kk} > 0 \quad \Rightarrow \quad \lambda_{kk} > 0. \qquad \square$$

37.4 Критерий неотрицательной определенности

Легко видеть, что ведущие подматрицы наследуют также свойство неотрицательной определенности. Поэтому для неотрицательной определенности эрмитовой матрицы необходимо, чтобы ее ведущие миноры были неотрицательными. Однако, пример матрицы $A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$ показывает, что этого уже не достаточно. Кроме ведущих миноров, теперь нужно вовлечь в рассмотрение также все *главные миноры* и *главные подматрицы* — так называются миноры и подматрицы, расположенные на пересечении строк и столбцов с одинаковой системой номеров. Заметим, что в эрмитовой матрице все главные подматрицы будут эрмитовы.

Лемма 1. Пусть $r = \mathrm{rank}A$. Тогда подматрица порядка r, расположенная на пересечении любых r линейно независимых строк и любых r линейно независимых столбцов, будет невырожденной.

Доказательство. Обозначим эту подматрицу через B, и пусть R — подматрица размеров $r \times n$, образованная заданными строками. Каждый столбец A есть линейная комбинация столбцов, на которых находится B. $\Rightarrow$ Каждый столбец R есть линейная комбинация столбцов B. Поэтому если $k \equiv \mathrm{rank}B < r$, то каждый столбец R есть линейная комбинация k базисных столбцов B $\Rightarrow$ $\mathrm{rank}R < r$ $\Rightarrow$ строки R линейно зависимы, а это противоречит предположению. $\square$

Лемма 2. *Среди отличных от нуля миноров порядка r эрмитовой матрицы ранга r имеется главный минор.*

Доказательство. Пусть $A = A^*$. Тогда если r строк (столбцов) линейно независимы, то r столбцов (строк) с теми же номерами также линейно независимы. По лемме 1, минор на их пересечении отличен от нуля. Он же, очевидно, главный. $\square$

Лемма 3. *Пусть A — невырожденная эрмитова матрица порядка $n \geq 2$, в которой главные миноры порядка k для всех k от 1 до $n-1$ равны нулю. Тогда $n = 2$ и $\det A < 0$.*

Доказательство. Пусть $\lambda_1 \geq \ldots \geq \lambda_n$ — собственные значения матрицы A. Если $\lambda_k > 0$ при каком-то k из промежутка от 2 до n, то из соотношений разделения следует, что все главные подматрицы порядка $k-1$ имеют положительные собственные значения и поэтому невырожденные. Если $\lambda_1 < 0$, то все главные миноры отличны от нуля. Таким образом,

$$\lambda_1 > 0 > \lambda_2 \geq \ldots \geq \lambda_n.$$

В то же время, если главные миноры первого и второго порядка равны нулю, то любая главная подматрица второго порядка нулевая:

$$\det \begin{bmatrix} 0 & a \\ \bar{a} & 0 \end{bmatrix} = -|a|^2 = 0 \;\Rightarrow\; a = 0.$$

Из соотношений разделения получаем $\lambda_2 \geq 0$. Поскольку противоречие возникает при $n > 2$, должно быть $n = 2$. В этом случае $\det A = \lambda_1 \lambda_2 < 0$. $\square$

Теорема. *Для неотрицательной определенности эрмитовой матрицы необходимо и достаточно, чтобы все ее главные миноры были неотрицательны.*

Доказательство. Необходимость ясна, так как свойство неотрицательной определенности наследуется любой главной подматрицей. Докажем достаточность.

Пусть $\lambda_1 \geq \ldots \geq \lambda_n$ — собственные значения матрицы A.

Пусть $r = \mathrm{rank}A$. По лемме 2, имеется невырожденная главная подматрица порядка r. Обозначим ее через B. По лемме 3, если $r > 2$, в B существует невырожденная главная подматрица порядка $r-1$. Отсюда ясно, что с помощью некоторой матрицы перестановки P из B можно получить эрмитову матрицу $P^\top BP$, в которой все ведущие миноры отличны от нуля и, следовательно, положительны. В силу критерия положительной определенности, B является положительно определенной матрицей $\Rightarrow$ все ее собственные значения положительны $\Rightarrow$ $\lambda_{r-1} > 0$. Если $\lambda_r < 0$, то и $\lambda_{r+1} < 0$ $\Rightarrow$ $\mathrm{rank}A > r$. Значит, $\lambda_r > 0$ и $\lambda_{r+1} = \ldots = \lambda_n = 0$. Неотрицательность всех собственных значений эрмитовой матрицы влечет за собой ее неотрицательную определенность. $\square$

37.5 Вариационные свойства сингулярных чисел

Теорема. *Пусть $A \in \mathbb{C}^{m\times n}$ имеет сингулярные числа*

$$\sigma_1(A) \geq \ldots \geq \sigma_{\min(m,n)}(A).$$

Тогда при всех $1 \le k \le \min(m,n)$

$$\sigma_k(A) \;=\; \max_{\dim L = k} \; \min_{x \in L,\, x \ne 0} \frac{||Ax||_2}{||x||_2} \;=\; \min_{\dim L = n-k+1} \; \max_{x \in L,\, x \ne 0} \frac{||Ax||_2}{||x||_2}.$$

Доказательство. Заметим, что $\sigma_k(A) = \sqrt{\lambda_k(A^*A)}$. Очевидно также, что

$$\frac{||Ax||_2}{||x||_2} \;=\; \sqrt{\frac{x^*(A^*A)x}{x^*x}}, \quad x \ne 0.$$

Таким образом, все сразу же следует из вариационных свойств собственных значений эрмитовой матрицы A^*A. $\square$

Задача. Пусть $A \in \mathbb{C}^{n\times n}$ и $f_k(A) = \sigma_1(A) + \ldots + \sigma_k(A)$. Докажите, что для любого $1 \le k \le n$ функция $f_k(A)$ является матричной нормой на $\mathbb{C}^{n\times n}$.

37.6 Разделение сингулярных чисел

Теорема. *Пусть* $A \in \mathbb{C}^{m\times n}$ *и* $B \in \mathbb{C}^{m\times(n-1)}$ *— подматрица, состоящая из первых* $n-1$ *столбцов матрицы* A*. Тогда для сингулярных чисел* A *и* B *имеют место соотношения разделения*

$$\sigma_1(A) \;\ge\; \sigma_1(B) \;\ge\; \sigma_2(A) \;\ge\; \ldots \;\ge\; \sigma_{n-1}(B) \;\ge\; \sigma_n(A).$$

Доказательство. Согласно условию теоремы, A имеет вид $A = [B, v]$, где v — ее последний столбец. Значит,

$$A^*A \;=\; \begin{bmatrix} B^* \\ v^* \end{bmatrix} \begin{bmatrix} B & v \end{bmatrix} \;=\; \begin{bmatrix} B^*B & B^*v \\ v^*B & v^*v \end{bmatrix}.$$

Искомые неравенства получаются из соотношений разделения для эрмитовой матрицы A^*A порядка n и ее ведущей подматрицы B^*B порядка $n-1$. $\square$

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

37.7 Эрмитово возмущение заданного ранга

Теорема. *Пусть* A *— эрмитова матрица порядка* n *и* $B = A + vv^*$ *— ее эрмитово возмущение ранга 1. Тогда*

$$\lambda_1(B) \;\ge\; \lambda_1(A) \;\ge\; \lambda_2(B) \;\ge\; \ldots \;\ge\; \lambda_{n-1}(A) \;\ge\; \lambda_n(B) \;\ge\; \lambda_n(A),$$
$$\lambda_k(A) + ||v||_2 \;\ge\; \lambda_k(B), \quad 1 \le k \le n.$$

Доказательство. Используя теорему Куранта–Фишера, находим

$$\lambda_k(A) \quad = \quad \max_{\dim L = k} \; \min_{x \in L,\, x \ne 0} \frac{x^*Ax}{x^*x} \;\le\; \max_{\dim L = k} \; \min_{x \in L,\, x \ne 0} \frac{x^*Ax + |v^*x|^2}{x^*x} =$$

$$\max_{\dim L = k} \; \min_{x \in L,\, x \ne 0} \frac{x^*Bx}{x^*x} \;=\; \lambda_k(B) \;\le\; \lambda_k(A) + ||v||_2.$$

Далее, пусть V — унитарная матрица с последним столбцом, равным $v/||v||_2$. Тогда $\lambda_k(V^*AV) = \lambda_k(A)$, $\lambda_k(B) = \lambda_k(V^*BV)$ и, как легко видеть,

$$V^*BV = V^*AV + \begin{bmatrix} 0 \\ \ldots \\ 0 \\ \bar{\gamma} \end{bmatrix} \begin{bmatrix} 0 & \ldots & 0 & \gamma \end{bmatrix}, \quad \gamma = ||v||_2.$$

Обозначим через C общую для V^*AV и V^*BV подматрицу порядка $n-1$ на пересечении первых $n-1$ строк и столбцов.

Пусть M — подпространство столбцов из $\mathbb{C}^n$ с последней координатой, равной нулю. По той же теореме Куранта–Фишера, при $2 \le k \le n$

$$\lambda_k(B) = \min_{\dim L = n-k+1} \max_{x \in L,\, x \neq 0} \frac{x^*(V^*BV)x}{x^*x} \le \min_{\dim L = n-k+1,\, L \subset M} \max_{x \in L,\, x \neq 0} \frac{x^*(V^*AV)x}{x^*x} =$$

$$\min_{\substack{\dim L = (n-1)-(k-1)+1 \\ L \subset \mathbb{C}^{n-1}}} \max_{y \in L,\, y \neq 0} \frac{y^*Cy}{y^*y} = \lambda_{k-1}(C) \le \lambda_{k-1}(V^*AV). \quad \square$$

Следствие. *Пусть A и B — эрмитовы матрицы порядка n и при этом*

$$B = V - U, \quad V = \sum_{i=1}^{k} v_i v_i^*, \quad U = \sum_{i=1}^{l} u_i u_i^*.$$

Тогда

$$\lambda_{i+l}(A) \le \lambda_i(A+B) \le \lambda_{i-k}(A),$$

где левое неравенство справедливо при $i + l \le n$, а правое при $1 \le i - k$.

Доказательство. Последовательное применение теоремы дает

$$\lambda_i(A) \le \lambda_i(A+V) \le \lambda_{i-k}(A),$$

$$\lambda_i(A+B) \le \lambda_i(A+V) \le \lambda_{i-l}(A+B).$$

Следовательно,

$$\lambda_{i+l}(A) \le \lambda_i(A+B) \le \lambda_{i-k}(A). \qquad \square$$

Часто бывает известно, что все собственные значения эрмитовой матрицы A принадлежат некоторому отрезку $[a, b]$. Полученный результат означает, что при всех эрмитовых возмущениях F ранга r матрица $A + F$ будет, по-прежнему, иметь все собственные значения на отрезке $[a, b]$, кроме, быть может, r "аутсайдеров".

37.8 Собственные значения и сингулярные числа

Есть много интересных сооотношений, связывающих собственные значения матрицы и ее сингулярные числа. Некоторые из них получаются очень просто.

Пусть $A \in \mathbb{C}^{n \times n}$ имеет сингулярные числа $\sigma_1 \ge \ldots \ge \sigma_n$, а ее собственные значения упорядочены по неубыванию модуля: $|\lambda_1| \ge \ldots \ge |\lambda_n|$.

Утверждение. $\sigma_n \le |\lambda_n|, \quad |\lambda_1| \le \sigma_1.$

Доказательство. Пусть $Ax = \lambda_i x$, $x \neq 0$. $\Rightarrow$ $|\lambda_i| ||x||_2 = ||Ax||_2 \le ||A||_2 ||x||_2 = \sigma_1 ||x||_2$ $\Rightarrow$ $|\lambda_i| \le \sigma_1$. Далее, если матрица A вырожденная, то $\lambda_n = 0$ и $\sigma_n = 0$. Если же A невырожденная, то A^{-1} имеет собственные значения λ_i^{-1} и $||A^{-1}||_2 = 1/\sigma_n$. $\square$

Данный простой факт имеет много обобщений. Например, такое.

Теорема. *Для всех $1 \le k \le n$ справедливы неравенства*

$$\sum_{i=1}^{k} |\lambda_i|^2 \le \sum_{i=1}^{k} \sigma_i^2.$$

Доказательство. В силу теоремы Шура, с помощью унитарной матрицы Q можно привести A к верхней треугольной матрице

$$Q^*AQ = R = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}, \qquad B = \begin{bmatrix} \lambda_1 & b_{12} & \ldots & b_{1k} \\ & \lambda_2 & \ldots & b_{2k} \\ & & \ddots & \\ & & & \lambda_k \end{bmatrix}.$$

Собственные значения для A^*A равны $\sigma_1^2 \geq \ldots \geq \sigma_n^2$ и совпадают с собственными значениями для

$$R^*R = \begin{bmatrix} B^* & 0 \\ C^* & D^* \end{bmatrix} \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = \begin{bmatrix} B^*B & B^*C \\ C^*B & C^*C + D^*D \end{bmatrix}.$$

Используя соотношения разделения для эрмитовых матриц B^*B и R^*R, находим

$$\sum_{i=1}^{k} |\lambda_i|^2 \leq \operatorname{tr}(B^*B) = \sum_{i=1}^{k} \lambda_i(B^*B) \leq \sum_{i=1}^{k} \lambda_i(R^*R) = \sum_{i=1}^{k} \sigma_i^2.$$

Задача. Доказать, что матрица A является нормальной тогда и только тогда, когда сумма квадратов ее сингулярных чисел равна сумме квадратов модулей собственных значений.

Неравенства Вейля. *Сингулярные числа и собственные значения, занумерованные по неубыванию модулей, удовлетворяют неравенствам*

$$\prod_{i=1}^{k} |\lambda_i| \leq \prod_{i=1}^{k} \sigma_i, \qquad 1 \leq k \leq n.$$

Доказательство. В обозначениях предыдущего доказательства,

$$\prod_{i=1}^{k} |\lambda_i|^2 = |\det B|^2 = \det(B^*B) = \prod_{i=1}^{k} \lambda_i(B^*B) \leq \prod_{i=1}^{k} \lambda_i(R^*R) = \prod_{i=1}^{k} \sigma_i^2. \quad \square$$

37.9 Мажоризация и неравенства

На базе неравенств Вейля можно получить целую серию полезных неравенств. Для этого их надо переписать в виде (давайте считать, что матрица A невырожденная)

$$\ln|\lambda_1| + \ldots + \ln|\lambda_k| \leq \ln\sigma_1 + \ldots + \ln\sigma_k, \quad 1 \leq k \leq n,$$

и заметить дополнительно, что

$$\ln|\lambda_1| + \ldots + \ln|\lambda_n| = \ln\sigma_1 + \ldots + \ln\sigma_n.$$

В данной форме неравенства Вейля оказываются частным случаем некоторого общего типа неравенств.

Говорят, что вектор $x = [x_1, \ldots, x_n]^\top \in \mathbb{R}^n$ *мажорируется* вектором $y = [y_1, \ldots, y_n]^\top \in \mathbb{R}^n$, если

(1) $x_1 \geq \ldots \geq x_n, \quad y_1 \geq \ldots \geq y_n$;

(2) $x_1 + \ldots + x_k \leq y_1 + \ldots + y_k, \quad 1 \leq k \leq n-1$;

(3) $x_1 + \ldots + x_n = y_1 + \ldots + y_n$.

Обозначение: $x \prec y$. Мажоризация всегда связана с равенством $x = Sy$, где S — матрица порядка n с неотрицательными элементами, суммы которых для каждой строки и для каждого столбца одинаковы и равны 1. Матрица с такими свойствами называется *двоякостохастической*.

Задача. Докажите, что множество всех двоякостохастических матриц порядка n является выпуклым и при этом матрицы перестановок и только они являются его угловыми точками.

Теорема. *Пусть $x_1 \geq \ldots \geq x_n$, $y_1 \geq \ldots \geq y_n$. Для того чтобы вектор $x = [x_1, \ldots, x_n]^\top$ мажорировался вектором $y = [y_1, \ldots, y_n]^\top$, необходимо и достаточно существование двоякостохастической матрицы S такой, что $x = Sy$.*

Доказательство. Достаточность: пусть $x = Sy$ для некоторой двоякостохастической матрицы $S = [s_{ij}]$, тогда

$$\sum_{i=1}^{k} x_i = \sum_{i=1}^{k}\sum_{j=1}^{n} s_{ij}y_j \leq \sum_{i=1}^{k} \left(\sum_{j=1}^{k-1} s_{ij}y_j + \left(1 - \sum_{j=1}^{k-1} s_{ij}\right) y_k \right) = ky_k - \sum_{j=1}^{k-1} \left(\sum_{i=1}^{k} s_{ij} \right) (y_k - y_j) =$$

$$\sum_{j=1}^{k} y_j \;+\; \sum_{j=1}^{k-1}\left(1-\sum_{i=1}^{k} s_{ij}\right)(y_k - y_j) \;\le\; \sum_{j=1}^{k} y_j.$$

Докажем необходимость. Пусть $x \prec y$. Очевидно,

$$nx_1 \ge x_1 + \ldots + x_n = y_1 + \ldots + y_n \ge ny_n \quad \Rightarrow \quad x_1 \ge y_n.$$

В случае $n = 2$ имеем $y_2 \le x_1 \le y_1 \quad \Rightarrow \quad x_1$ является выпуклой комбинацией чисел y_1 и y_2: $x_1 = sy_1 + ty_2,\ \ s, t \ge 0,\ \ s + t = 1$. Таким образом,

$$x = Sy, \quad S = \begin{bmatrix} s & t \\ t & s \end{bmatrix}.$$

В общем случае $y_n \le x_1 \le y_1$. Обозначим через k наименьший номер такой, что $y_k \le x_1 \le y_{k-1} \le y_1$. Поэтому $x_1 = sy_1 + ty_k,\ \ s, t \ge 0,\ \ s+t = 1$. Пусть матрица $F \in \mathbb{R}^{n\times n}$ задает преобразование $u \mapsto v = Fu$, определяемое следующим правилом:

$$v_1 = su_1 + tu_k, \quad v_k = tu_1 + su_k, \qquad v_i = u_i, \quad i \ne 1,\ k.$$

Легко видеть, что матрица F двоякостохастическая. Далее, положим $z = Fy$, рассмотрим векторы $x' = [x_2, \ldots, x_n]^\top,\ \ z' = [z_2, \ldots, z_n]^\top$ и докажем, что $x' \prec z'$. Согласно выбору номера k,

$$x_n \le \ldots \le x_1 \le y_{k-1} \le \ldots \le y_1.$$

Поэтому $\sum\limits_{i=2}^{l} x_i \le \sum\limits_{i=2}^{l} y_i$ для всех $1 \le l \le k-1$. При $k \le l \le n$ находим

$$\begin{aligned} \sum_{i=2}^{l} z_i \quad &= \quad (ty_1 + sy_k) \;+\; \sum_{i=2}^{k-1} y_i \;+\; \sum_{i=k+1}^{l} y_i \\ &= \quad \sum_{i=1}^{l} y_i \;-\; (sy_1 + ty_k) \;\ge\; \sum_{i=1}^{l} x_i \;-\; x_1 \;=\; \sum_{i=2}^{l} x_i. \end{aligned}$$

Рассуждая по индукции, предположим, что существует двоякостохастическая матрица T' порядка $n-1$ такая, что $x' = T'z'$. Тогда матрица

$$T = \begin{bmatrix} 1 & 0 \\ 0 & T' \end{bmatrix}$$

будет, очевидно, двоякостохастической. Учитывая, что $x_1 = z_1$, получаем $x = Tz$. Таким образом, $x = Sy$, где $S = TF$ есть произведение двух двоякостохастических матриц и поэтому, как легко проверить, тоже является двоякостохастической матрицей. $\square$

Следствие. *Пусть $[x_1, \ldots, x_n]^\top \prec [y_1, \ldots, y_n]^\top$. Тогда для любой выпуклой монотонно возрастающей функции $\phi(t)$ справедливы неравенства*

$$\phi(x_1) + \ldots + \phi(x_k) \le \phi(y_1) + \ldots + \phi(y_k), \quad 1 \le k \le n.$$

Доказательство. Согласно теореме, $x = Sy$ для некоторой двоякостохастической матрицы $S = [s_{ij}]$. Вследствие этого,

$$\sum_{i=1}^{k} \phi(x_i) \le \sum_{i=1}^{k}\sum_{j=1}^{n} s_{ij}\phi(y_j) \;\le\; \sum_{j=1}^{k-1}\left(\left(\sum_{i=1}^{k} s_{ij}\right)\phi(y_j) \;+\; \left(1 - \left(\sum_{i=1}^{k} s_{ij}\right)\right)\phi(y_k)\right) \;\le$$

$$\sum_{j=1}^{k} \phi(y_j) \;+\; \sum_{j=1}^{k-1}\left(1 - \sum_{i=1}^{k} s_{ij}\right)(\phi(y_k) - \phi(y_j)) \;\le\; \sum_{j=1}^{k} \phi(y_j). \quad \square$$

Теперь пусть A — невырожденная матрица с сингулярными числами $\sigma_1 \ge \ldots \ge \sigma_n$ и собственными значениями $\lambda_1, \ldots, \lambda_n$, упорядоченными по неубыванию модуля. Положим $x_i = \ln|\lambda_i|$ и $y_i = \ln\sigma_i$. Тогда из неравенств Вейля вытекает, что $x \prec y$. Возьмем, например, функцию $\phi(t) = e^t$. В силу того, что она является выпуклой и монотонно возрастающей, получаем неравенства

$$|\lambda_1| + \ldots + |\lambda_k| \le \sigma_1 + \ldots + \sigma_k, \quad 1 \le k \le n.$$

Лекция 38

ОСНОВНАЯ ЧАСТЬ

38.1 Сопряженный оператор

Пусть $\mathcal{A} : V \to W$ — произвольный оператор, а V и W — пространства со скалярными произведениями $(\cdot\,,\cdot)_V$ и $(\cdot\,,\cdot)_W$. Попробуем построить оператор $\mathcal{A}^* : W \to V$, обладающий свойством

$$(\mathcal{A}(x), y)_W = (x, \mathcal{A}^*(y))_V \quad \forall \;\; x \in V, \;\; \forall \;\; y \in W. \tag{$*$}$$

Утверждение. *Если оператор $\mathcal{A}^*$ существует, то он является линейным и единствен.*

Доказательство. $(\mathcal{A}^*(\alpha u + \beta v), x)_W = (\alpha u + \beta v, \mathcal{A}(x))_V = \alpha(u, \mathcal{A}(x))_V + \beta(v, \mathcal{A}(x))_V =$

$$\alpha(\mathcal{A}^*(u), x)_W + \beta(\mathcal{A}^*(v), x)_W = (\alpha\mathcal{A}^*(u) + \beta\mathcal{A}^*(v), x)_W.$$

Положим $z = \mathcal{A}^*(\alpha u + \beta v) - \alpha\mathcal{A}^*(u) - \beta\mathcal{A}^*(v)$. Мы доказали, что $(z, x)_V = 0 \;\forall\; x \in V$. Это верно, в частности, для $x = z \quad \Rightarrow \quad (z, z)_V = 0 \quad \Rightarrow \quad z = 0$.

Докажем единственность. Предположим, что для некоторого $y \in W$ имеем $(\mathcal{A}(x), y)_W = (x, z_1)_V = (x, z_2)_V \;\; \forall\; x \in V$. Тогда, взяв $x = z_1 - z_2$, находим $(x, x)_V = 0 \quad \Rightarrow \quad z_1 = z_2$. $\square$

Следствие. *Если операторы $\mathcal{A} : V \to W$ и $\mathcal{A}^* : W \to V$ связаны соотношением $(*)$, то они оба являются линейными.*

Типичная ситуация, в которой сопряженный оператор очень полезен, такая. Предположим, имеется операторное уравнение $\mathcal{A}(u) = f$ с обратимым оператором $\mathcal{A}$ и при этом для различных правых частей f требуется вычислить значение линейного функционала

$$\Phi(u) = (u, \phi)_V,$$

заданного одним и тем же вектором ϕ.

Определение сопряженного оператора $(\mathcal{A}(u), z)_W = (u, \mathcal{A}^*(z))$ приводит к следующей идее: вместо того чтобы многократно решать уравнение $\mathcal{A}(u) = f$ для различных f, рассмотреть сопряженное уравнение $\mathcal{A}^*(z) = \phi$, найти его решение z, а затем использовать формулу

$$\Phi(u) = (f, z)_W.$$

Замечательно, что $\Phi(u)$ можно найти, не вычисляя u. [1]

Теорема. *Пусть $\mathcal{A} : V \to W$ — линейный оператор. Если пространства V и W конечномерны, то оператор $\mathcal{A}^*$, удовлетворяющий равенству $(*)$, существует и единствен. При этом в паре ортонормированных базисов сопряженному оператору соответствует сопряженная матрица.*

Доказательство. Пусть $v_1, \ldots, v_n$ — ортонормированный базис в V, а $w_1, \ldots, w_m$ — ортонормированный базис в W. Обозначим через $A = [a_{ij}] \in \mathbb{C}^{m\times n}$ матрицу оператора $\mathcal{A}$ в данной паре базисов. В силу ортонормированности,

$$a_{ij} = (\mathcal{A}v_j, w_i), \quad 1 \le i \le m, \;\; 1 \le j \le n.$$

[1] Глубокое изучение сопряженных уравнений, во многом навеянное данной общей идеей, выполнил академик Гурий Иванович Марчук — последний президент Академии наук СССР.

Чтобы определить оператор $\mathcal{A}^*$, рассмотрим разложение $\mathcal{A}^* w_i = \alpha_1 v_1 + \ldots + \alpha_n v_n$. Умножая скалярно на v_j, находим $\alpha_j = (\mathcal{A}^* w_i, v_j) = (w_i, \mathcal{A} v_j), \quad 1 \le j \le n$. Таким образом, матрица $B = [b_{ji}]$ линейного оператора $\mathcal{A}^*$ в паре базисов $\{w_i\}$ и $\{v_j\}$ должна иметь элементы

$$b_{ji} = (w_i, \mathcal{A} v_j) = \overline{(\mathcal{A} v_j, w_i)} = \bar{a}_{ij} \quad \Rightarrow \quad B = A^*.$$

Ясно также, что мы получили единственность оператора $\mathcal{A}^*$. Существование доказывается так: рассмотрим оператор, заданный матрицей A^*, и проверим, что для него выполняется равенство $(*)$:

$$x = \sum_{j=1}^{n} x_j v_j, \quad y = \sum_{i=1}^{m} y_i w_i \quad \Rightarrow \quad (\mathcal{A} x, y)_W = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} x_j \bar{y}_i = (x, \mathcal{A}^* y)_V,$$

что и требовалось доказать. $\square$

38.2 Матрица сопряженного оператора

Пусть $V = \mathbb{C}^n$ и $W = \mathbb{C}^m$. Как мы знаем, произвольные скалярные произведения в $\mathbb{C}^n$ и $\mathbb{C}^m$ имеют вид

$$(p, q)_V = q^* S p, \quad (y, z)_W = z^* T y,$$

где $S \in \mathbb{C}^{n \times n}$ и $T \in \mathbb{C}^{m \times m}$— эрмитовы положительно определенные матрицы.

Пусть линейный оператор $\mathcal{A} : \mathbb{C}^n \to \mathbb{C}^m$ определяется умножением на матрицу $A \in \mathbb{C}^{m \times n}$, а сопряженный оператор — умножением на матрицу $B \in \mathbb{C}^{n \times m}$. Тогда для любых $x \in \mathbb{C}^n$ и $y \in \mathbb{C}^m$ должно быть

$$y^* T (A x) = (B y)^* S x \quad \Rightarrow \quad y^* (T A) x = y^* (B^* S) x \quad \Rightarrow \quad T A = B^* S \quad \Rightarrow \quad B = S^{-1} A^* T.$$

Разные скалярные произведения в $\mathbb{C}^n$ и $\mathbb{C}^m$ приводят, конечно, к разным сопряженным операторам — но, как видим, любой из них есть умножение на матрицу вида $S^{-1} A^* T$, где S и T — эрмитовы положительно определенные матрицы, задающие скалярные произведения.

Пусть A — матрица линейного оператора $\mathcal{A} : V \to W$, $\dim V = n$, $\dim W = m$, в какой-то паре базисов. Если x и y — вектор-столбцы из координат разложения прообраза и образа при действии $\mathcal{A}$, то получаем $y = Ax$. Пусть теперь $x = By$. Тогда $Sx = A^* T y \quad \Rightarrow \quad$ замена $\widetilde{x} = Sx$, $\widetilde{y} = Ty$ приводит к соотношению $\widetilde{x} = A^* \widetilde{y} \quad \Rightarrow \quad$ *в паре базисов, определенных столбцами матриц* T^{-1} *и* S^{-1}, *матрица оператора* $\mathcal{A}^*$ *имеет вид* A^*. Легко видеть, что это базисы, биортогональные (в скалярных произведениях пространств W и V, соответственно) для базисов, в которых получена матрица A (см. раздел 25.7).

38.3 Нормальный оператор

Пусть $\mathcal{A} : V \to V$ — линейный оператор, V — пространство со скалярным произведением $(\cdot\,, \cdot)_V$. Если $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$, то $\mathcal{A}$ называется *нормальным оператором*. Данное свойство зависит от скалярного произведения: в другом скалярном произведении $\mathcal{A}$ может не быть нормальным.

Задача. Пусть $\mathcal{A} : V \to V$ — линейный оператор в произвольном конечномерном унитарном пространстве V. Докажите, что существует ортонормированный базис, в котором матрица оператора $\mathcal{A}$ является верхней треугольной.

Изучение нормальных операторов легко сводится к изучению нормальных матриц: достаточно выбрать в V ортонормированный базис, тогда нормальность оператора равносильна нормальности его матрицы в данном базисе. Отсюда ясно, что нормальный оператор является оператором простой структуры. Заметим также,что любой оператор простой структуры можно сделать нормальным за счет выбора скалярного произведения (докажите!).

Важнейшие классы нормальных операторов: унитарные операторы ($\mathcal{A}^* = \mathcal{A}^{-1}$) и эрмитовы (самосопряженные) операторы ($\mathcal{A}^* = \mathcal{A}$). Пусть $\mathcal{A}$ — нормальный оператор. Легко доказывается, что *унитарность оператора* $\mathcal{A}$ *равносильна тому, что все его собственные значения по модулю равны* 1, *а эрмитовость равносильна вещественности собственных значений*. Подчеркнем, что унитарность и эрмитовость оператора зависят от скалярного произведения.

38.4 Самосопряженный оператор

Если $(\mathcal{A}x, y)_V = (x, \mathcal{A}y)_V \quad \forall \quad x, y \in V$, то, в силу единственности сопряженного оператора, $\mathcal{A}^* = \mathcal{A}$. В таких случаях $\mathcal{A}$ называется *самосопряженным* оператором. Если $(\mathcal{A}x, x) > 0$ при всех $x \in V,\ x \neq 0$, то оператор называется *положительно определенным.*

Если $V = \mathbb{C}^n$ и скалярное произведение $(x, y)_S = y^* S x$ определяется с помощью эрмитовой положительно определенной матрицы $S \in \mathbb{C}^{n\times n}$, то, согласно предыдущему разделу, самосопряженность оператора умножения на матрицу $A \in \mathbb{C}^{n\times n}$ означает, что

$$A = S^{-1} A^* S. \qquad (*)$$

Заметим, что равенство $S^{-1/2}\, S\, S^{-1/2} = I$ означает, что столбцы матрицы $S^{-1/2}$ образуют ортонормированный базис относительно скалярного произведения $(\cdot\ \cdot)_S$. Матрица B оператора умножения на A в данном базисе определяется равенством

$$AS^{-1/2} = S^{-1/2} B \quad \Rightarrow \quad B = S^{1/2} A S^{-1/2}.$$

Самосопряженность означает, что B должна быть эрмитовой матрицей — это легко также вывести непосредственно из $(*)$. Как видим, матрица A подобна эрмитовой матрице $B \quad \Rightarrow \quad$ все ее собственные значения вещественны.

38.5 Минимизация на подпространствах

Обсудим важную идею, позволяющую строить методы решения системы $Ax = b$, совсем не похожие на известный нам метод Гаусса. Пусть $A \in \mathbb{C}^{n\times n}$ — невырожденная матрица.

Рассмотрим так называемые *подпространства Крылова* [2]

$$L_k \ = \ L(b,\, Ab,\, \ldots,\, A^{k-1}b), \quad k = 1,\, 2,\, \ldots\,,$$

и определим $x_k \in L_k$ из следующего условия:

$$||b - Ax_k||_2 = \min_{z\in L_k} ||b - Az||_2.$$

Вектор $r(z) = b - Az$ называется *невязкой* вектора z. Очевидно, вычисление вектора x_k сводится к задаче о перпендикуляре, опущенном из вектора b на подпространство

$$M_k = AL_k = \{y \in \mathbb{C}^n : y = Az,\ z \in L_k\}.$$

Как решать такую задачу — мы уже знаем. Понятно также, что решение существенно облегчается наличием "удобного" базиса $p_1, \ldots, p_k$ в L_k (например, приводящего к ортогональной системе $Ap_1, \ldots, Ap_k$).

В условиях точных вычислений процесс всегда завершается получением решения x. Если $L_n = \mathbb{C}^n$, то, очевидно, $x_n = x$. Если на каком-то шаге $L_k = L_{k+1}$, то

$$AL_k \subset L_{k+1} = L_k \quad \Rightarrow \quad AL_k = L_k \quad \text{(в силу невырожденности матрицы } A\text{)}.$$

Поскольку $b \in L_k$, то должно быть $Az = b$ для какого-то $z \in L_k$. Невырожденность A означает, что $z = x \quad \Rightarrow \quad x \in L_k \quad \Rightarrow \quad x_k = x$. Заметим также, что если $x \in L_k$ (а значит, $x_k = x$), то $L_k = L_{k+1}$ (докажите!).

Обратим внимание на то, что x_k часто оказывается хорошим приближением к решению x при $k \ll n$. Описанная идея является ключевой в современных методах решения систем в многочисленных прикладных задачах.

[2] Заметим, что L_k есть подпространство минимального инвариантного подпространства, порожденного вектором b. В Лекции 32 было доказано, что если $A^k b = 0$, то отличие от нуля векторов $b, Ab, \ldots, A^{k-1}b$ влечет за собой их линейную независимость.

38.6 Метод сопряженных градиентов

Данная идея приобретает особенно элегантную форму в случае, когда A — эрмитова положительно определенная матрица.

Пусть x_0 — произвольный начальный вектор. Если $r_0 = b - Ax_0 = 0$, то решение найдено. Если $r_0 \neq 0$, начинаем строить подпространства Крылова

$$L_k = L(r_0,\, Ar_0,\, \ldots,\, A^{k-1}r_0) = L(p_1,\, \ldots,\, p_k),$$

последовательно получая в них базис $p_1, \ldots, p_k$ со следующим свойством:

$$(Ap_i, p_j) = 0, \quad i \neq j; \qquad p_1 = r_0.$$

Поскольку $(x, y)_A = (Ax, y)$ есть скалярное произведение, данное свойство называется *свойством A-ортогональности* векторов $p_1, \ldots, p_k$; *A-нормой* вектора x называется величина $||x||_A = \sqrt{(x,x)_A} = \sqrt{(Ax, x)}$.

Пусть x_k имеет вид $x_k = x_0 + y$, где $y \in L_k$ выбирается таким образом, чтобы минимизировать величину $||x - x_k||_A = ||(x - x_0) - y||_A$ (A-норму отклонения x_k от точного решения x). Ясно, что это задача о перпендикуляре в случае A-ортогональности. Поэтому y определяется из уравнений

$$((x - x_0) - y,\, p_i)_A = 0 \quad \Leftrightarrow \quad (r_0 - Ay,\, p_i) = 0, \quad 1 \leq i \leq k.$$

Записав $y = \alpha_1 p_1 + \ldots + \alpha_k p_k$, находим $\alpha_i = (r_0,\, p_i)/(Ap_i,\, p_i)$. Следовательно, векторы x_k можно вычислять по очень простой рекуррентной формуле

$$x_k = x_{k-1} + \alpha_k p_k, \quad \alpha_k = (r_0,\, p_k)/(Ap_k,\, p_k).$$

Отсюда видно, что невязки $r_k = b - Ax_k$ связаны рекуррентной формулой

$$r_k = r_{k-1} - \alpha_k A p_k.$$

Удивительно и приятно то, что для вычисления x_k требуется лишь один вектор p_k из базиса $p_1, \ldots, p_k$! Но еще более удивительно и приятно то, что p_{k+1} можно найти, используя лишь два вектора: p_k и r_k.

В самом деле, если $r_k = 0$, то решение найдено. [3]

Если же $r_k \neq 0$, то невязка $r_k = r_0 - Ay$ является ортогональной подпространству L_k и поэтому p_{k+1} можно записать в виде

$$p_{k+1} = r_k + \beta_1 p_1 + \ldots + \beta_k p_k.$$

Условие A-ортогональности дает равенства

$$(Ap_{k+1}, p_i) = 0 \quad \Rightarrow \quad \beta_i = (Ar_k, p_i)/((Ap_i, p_i), \quad 1 \leq i \leq k.$$

При этом $(Ar_k, p_i) = (r_k, Ap_i) = 0$ при $i \leq k-1$, так как вектор $Ap_i \in AL_i \subset L_{i+1}$. Таким образом, $\beta_i = 0$ при $1 \leq i \leq k-1 \quad \Rightarrow$

$$p_{k+1} = r_k + \beta_k p_k, \quad \beta_k = (r_k, Ap_k)/(Ap_k, p_k).$$

38.7 Двучленные формулы

Заметим, что для вычисления α_k совсем не обязательно использовать формулу $\alpha_k = (r_0,\, p_k)/(Ap_k,\, p_k)$. Поскольку $r_k \perp L_k$, находим $0 = (r_k,\, p_k) = (r_{k-1} - \alpha_k A p_k,\, p_k) \quad \Rightarrow$

$$\alpha_k \;=\; \frac{(r_{k-1}, p_k)}{(Ap_k, p_k)} \;=\; \frac{(r_{k-1}, r_{k-1} + \beta_{k-1} p_{k-1})}{(Ap_k, p_k)} \;=\; \frac{(r_{k-1}, r_{k-1})}{(Ap_k, p_k)}.$$

[3] В этом случае $L_k = L_{k+1}$ (докажите!).

Далее, если $r_{k-1} \neq 0$, то $\alpha_k \neq 0 \quad \Rightarrow \quad Ap_k = (r_{k-1} - r_k)/\alpha_k \quad \Rightarrow$

$$\beta_k \;=\; \frac{(r_k, r_{k-1} - r_k)}{\alpha_k\,(Ap_k, p_k)} \;=\; -\frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})}.$$

Окончательно, *метод сопряженных градиентов* сводится к итерациям, выполняемым по следующим двучленным формулам:

$$\begin{aligned} x_k &= x_{k-1} + \alpha_k p_k, \qquad & \alpha_k &= \frac{(r_{k-1}, r_{k-1})}{(Ap_k, p_k)}, \\ r_k &= r_{k-1} - \alpha_k A p_k, & & \\ p_{k+1} &= r_k + \beta_k p_k, \qquad & \beta_k &= -\frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})}. \end{aligned}$$

Теоретически итерации выполняются до тех пор, пока $r_k \neq 0$. На практике они останавливаются, когда $||r_k||_2$ становится достаточно малой.

Наиболее сложное действие на k-м шаге метода сопряженных градиентов — это умножение заданной матрицы A на вектор. При этом совсем не обязательно хранить все n^2 элементов матрицы в каком-то массиве — требуется лишь наличие какой-то процедуры умножения матрицы на вектор. Именно в этом плане итерационные методы существенно отличаются от метода Гаусса, это же обстоятельство делает их особенно полезными при решении систем с очень большим числом неизвестных.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

38.8 Число итераций

В методе сопряженных градиентов $x_k \in x_0 + L_k$, но $x_k \notin x_0 + L_{k-1}$. Значит, $x_k - x_0$ является линейной комбинацией векторов r_0, Ar_0, …, $A^{k-1}r_0$ с ненулевым коэффициентом при $A^{k-1}r_0 \quad \Rightarrow \quad x_k = x_0 + \psi_{k-1}(A)r_0$, где $\phi_{k-1}(\lambda)$ — многочлен степени $k-1$. Итак, $r_k = r_0 - A\psi_{k-1}(A)r_0 \quad \Rightarrow$

$$r_k \;=\; \phi_k(A)r_0, \quad \deg\phi_k(\lambda) = k, \quad \phi_k(0) = 1.$$

Утверждение. *Если A — эрмитова положительно определенная матрица, имеющая m попарно различных собственных значений, то число итераций в методе сопряженных градиентов при любом начальном векторе не больше m.*

Доказательство. Достаточно учесть, что степень минимального многочлена для эрмитовой матрицы A не больше m. □

38.9 Как убывают нормы невязок

Теоретически метод сопряженных градиентов требует не более n шагов для получения точного решения. Практически норма k-й невязки может оказаться достаточно малой при $k \ll n$. Получение оценок основано на следующем результате.

Лемма об оценке норм невязок. *Пусть $\lambda_{\min}$ и $\lambda_{\max}$ — минимальное и максимальное собственные значения эрмитовой положительно определенной матрицы A. Тогда k-я невязка в методе сопряженных градиентов при любом начальном векторе удовлетворяет неравенству*

$$||r_k||_2 \;\leq\; \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}\;\max_{\lambda_{\min}\leq\lambda\leq\lambda_{\max}} |\Phi_k(\lambda)|\;||r_0||_2,$$

где $\Phi_k(\lambda)$ — любой многочлен степени не выше k, подчиненный условию $\Phi_k(0) = 1$.

Доказательство. В методе сопряженных градиентов

$$||x - x_k||_A \;=\; \min_{y\in L_k} ||x - (x_0 + y)||_A.$$

Произвольный вектор $y \in L_k$ имеет вид $y = \Psi_{k-1}(A)r_0$, где $\Psi_{k-1}(\lambda)$ — многочлен степени $k-1$ или ниже $\Rightarrow$ $A(x-(x_0+y)) = r_0 - Ay = \Phi_k(A)r_0$, где $\Phi_k(\lambda)$ — многочлен степени не выше k со свободным членом $\Phi_k(0) = 1$. Таким образом,

$$||x - x_k||_A = ||A^{-1}r_k||_A \;\le\; ||A^{-1}\Phi_k(A)r_0||_A.$$

Пусть $\lambda_1 \ge \ldots \ge \lambda_n > 0$ — собственные значения матрицы A и $q_1, \ldots, q_n$ — ортонормированный базис из собственных векторов:

$$AQ = Q\Lambda, \quad Q = [q_1, \ldots, q_n], \quad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}.$$

$$r_k = \sum_{i=1}^{n} \zeta_i q_i \quad \Rightarrow \quad ||A^{-1}r_k||_A^2 = \sum_{i=1}^{n} \frac{|\zeta_i|^2}{\lambda_i} \;\ge\; \frac{||r_k||_2^2}{\lambda_1},$$

$$r_0 = \sum_{i=1}^{n} \xi_i q_i \;\Rightarrow\; ||A^{-1}\Phi_k(A)r_0||_A^2 = \sum_{i=1}^{n} \frac{|\Phi_k(\lambda_i)|^2\,|\xi_i|^2}{\lambda_i} \;\le\; \frac{1}{\lambda_n} \max_{\lambda_{\min}\le\lambda\le\lambda_{\max}} |\Phi_k(\lambda)|^2\;||r_0||_2^2. \quad \square$$

38.10 Оценка с помощью многочленов Чебышева

Таким образом, оценки для нормы k-й невязки можно получать с помощью многочленов. При этом нас интересует величина, уже известная нам как C-норма в пространстве непрерывных функций на отрезке $[\lambda_{\min}, \lambda_{\max}]$:

$$||\Phi_k||_C = \min_{\lambda_{\min}\le\lambda\le\lambda_{\max}} |\Phi_k(\lambda)|.$$

Как выбрать многочлен $\Phi_k(\lambda)$ с условием нормировки $\Phi_k(0) = 1$ и наименьшей C-нормой на отрезке $[\lambda_{\min}, \lambda_{\max}]$? Решение этой задачи дают *многочлены Чебышева.*

Многочлены Чебышева для отрезка $[-1, 1]$ определяются следующим образом:

$$T_0(t) = 1, \;\; T_1(t) = t, \quad T_{n+1}(t) = 2tT_n(t) - T_{n-1}(t), \;\; n = 1\,2, \ldots, .$$

Элементарно проверяется, что $T_n(t) = \cos(n \arccos t)$ при $-1 \le t \le 1$. Чтобы найти представление для $T_n(t)$ при $|t| > 1$, рассмотрим однородное реккуррентное уравнение

$$z_{n+1} - 2tz_n + z_{n-1} = 0$$

и попробуем искать его решение в виде $z_n = z^n$, $z \ne 0$. Тогда

$$z^2 - 2tz + 1 = 0 \;\; \Rightarrow \;\; z_{(\pm)} = t \pm \sqrt{t^2 - 1}.$$

Ясно, что $z_n = c_1 z_{(+)}^n + c_2 z_{(-)}^n$ будет решением данного рекуррентного уравнения при любых константах $c_1,\ c_2$. Выберем их так, чтобы $z_0 = T_0(t),\ z_1 = T_1(t)$. В итоге получаем

$$T_n(t) \;=\; \frac{1}{2}(t + \sqrt{t^2-1})^n \;+\; \frac{1}{2}(t - \sqrt{t^2-1})^n.$$

В случае многочленов от $\lambda \in [\lambda_{\min}, \lambda_{\max}]$ сделаем замену переменной

$$\lambda \;=\; \frac{\lambda_{\max} + \lambda_{\min}}{2} \;+\; t\frac{\lambda_{\max} - \lambda_{\min}}{2} \quad \Leftrightarrow \quad t = \frac{\lambda - \frac{\lambda_{\max}+\lambda_{\min}}{2}}{\frac{\lambda_{\max}-\lambda_{\min}}{2}}.$$

Положим

$$\Phi_k(\lambda) \;=\; \frac{T_k\left(\frac{\lambda - \frac{\lambda_{\max}+\lambda_{\min}}{2}}{\frac{\lambda_{\max}-\lambda_{\min}}{2}}\right)}{T_k(-\frac{\lambda_{\max}+\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}})} \quad \Rightarrow \quad \max_{\lambda_{\min}\le\lambda\le\lambda_{\max}} |\Phi_k(\lambda)| \;\le\; \frac{1}{T_k\left(-\frac{\lambda_{\max}+\lambda_{\min}}{\lambda_{\max}-\lambda_{\min}}\right)}.$$

Далее,

$$|t| = \frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}} \quad \Rightarrow \quad |t| + \sqrt{t^2-1} = \frac{\lambda_{\max} + \lambda_{\min} + 2\sqrt{\lambda_{\max}}\sqrt{\lambda_{\min}}}{\lambda_{\max} - \lambda_{\min}} \;=\; \frac{\sqrt{\lambda_{\max}} + \sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}}}.$$

При этом t получаем

$$|T_k(t)| \;\geq\; \frac{1}{2}\left(\frac{\sqrt{\lambda_{\max}}+\sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}}-\sqrt{\lambda_{\min}}}\right)^k.$$

Этот результат вместе с леммой об оценке норм невязок метода сопряженных градиентов доказывает следующую теорему.

Теорема. *В условиях леммы об оценке норм невязок метода сопряженных градиентов справедливы неравенства*

$$||r_k||_2 \;\leq\; 2\sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}\;\left(\frac{\sqrt{\lambda_{\max}}-\sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}}+\sqrt{\lambda_{\min}}}\right)^k\;||r_0||_2, \quad k=1,\,2,\,\ldots\,.$$

38.11 Предобусловленный метод сопряженных градиентов

Полученная теорема показывает, что нормы невязок в методе сопряженных градиентов убывают тем сильнее, чем меньше отношение $\lambda_{\max}/\lambda_{\min}$. Если это отношение велико, то можно попытаться найти "близкую" эрмитову положительно определенную матрицу B и решать равносильную *предобусловленную систему* $B^{-1}Ax = B^{-1}b$.

Проблема, однако, в том, что мы получили метод сопряженных градиентов для решения систем с эрмитовой положительно определенной матрицей, а произведение $B^{-1}A$ в общем случае не будет эрмитовой матрицей. Тем не менее, справедливо следующее

Утверждение. *Пусть A и B — эрмитовы положительно определенные матрицы. Тогда оператор умножения на матрицу $B^{-1}A$ является самосопряженным положительно определенным оператором относительно скалярного произведения $(x,y)_B = (Bx,y)$.*

Доказательство. $(B^{-1}Ax,y)_B = (Ax,y) = (x,Ay) = (x,B(B^{-1}A)y) = (Bx,B^{-1}Ay) = (x,B^{-1}Ay)_B$. Положительная определенность очевидна: $(B^{-1}Ax,x)_B = (Ax,x) > 0$ при $x \neq 0$. $\square$

Теперь мы можем повторить все рассуждения и выкладки, приведшие к двучленным формулам метода сопряженных градиентов, с заменой естественного скалярного произведения на $(\cdot\,,\cdot)_B$:

$$\begin{aligned}
&\hat{x}_k = \hat{x}_{k-1} + \alpha_k\hat{p}_k, && \alpha_k = \frac{(\hat{r}_{k-1},\hat{r}_{k-1})_B}{(B^{-1}A\hat{p}_k,\hat{p}_k)_B} = \frac{(B\hat{r}_{k-1},\hat{r}_{k-1})}{(A\hat{p}_k,\hat{p}_k)},\\
&\hat{r}_k = \hat{r}_{k-1} - \alpha_k B^{-1}A\hat{p}_k,\\
&\hat{p}_{k+1} = \hat{r}_k + \beta_k\hat{p}_k, && \beta_k = -\frac{(\hat{r}_k,\hat{r}_k)_B}{(\hat{r}_{k-1},\hat{r}_{k-1})_B} = -\frac{(B\hat{r}_k,\hat{r}_k)}{(B\hat{r}_{k-1},\hat{r}_{k-1})}.
\end{aligned}$$

Заметим, что $\hat{r}_k = B^{-1}(b - A\hat{x}_k)$ — это невязка предобусловленной системы. Соответствующая невязка исходной системы имеет вид $r_k = B\hat{r}_k$. Предобусловленный метод сопряженных градиентов, вычисляющий "настоящие" невязки r_k и те же векторы $x_k = \hat{x}_k$ и $p_k = \hat{p}_k$, принимает такую форму:

$$\begin{aligned}
&r_0 = b - Ax_0, \quad p_1 = B^{-1}r_0;\\
&x_k = x_{k-1} + \alpha_k p_k, && \alpha_k = \frac{(B^{-1}r_{k-1},r_{k-1})}{(Ap_k,p_k)},\\
&r_k = r_{k-1} - \alpha_k Ap_k,\\
&p_{k+1} = B^{-1}r_k + \beta_k p_k, && \beta_k = -\frac{(B^{-1}r_k,r_k)}{(B^{-1}r_{k-1},r_{k-1})}.
\end{aligned}$$

38.12 Обобщения метода сопряженных градиентов

В случае "больших" неэрмитовых матриц основным является метод минимизации нормы невязки на подпространствах Крылова. В отличие от метода сопряженных градиентов, в данном случае в подпространствах Крылова требуется строить и хранить полные базисы. Существуют ли методы с "короткими" рекуррентными соотношениями в неэрмитовом случае?

Прежде всего, уточним вопрос. Пусть $Ax = b$ — система с невырожденной и в общем случае неэрмитовой матрицей. Выбрав начальный вектор x_0, находим начальную невязку r_0, в случае $r_0 \neq 0$ полагаем $p_1 = r_0$ и последовательно дополняем базис $p_1,\,\ldots,\,p_k$ в пространствах Крылова

$$L_k = L(r_0, Ar_0, \ldots, A^{k-1}r_0) = L(p_1,\,\ldots,\,p_k),$$

причем таким образом, чтобы векторы удовлетворяли условиям *формальной A-ортогональности*

$$(Ap_i, p_j) = 0, \quad i \neq j, \quad 1 \leq i, j \leq k; \quad (Ap_i, p_i) \neq 0, \quad 1 \leq i \leq k.$$

Как только получено пространство L_k, ищем x_k в виде $x_k = x_0 + y,\ y \in L_k$. При этом откажемся от минимизации невязки $r_k = b - Ax_k$ в какой-либо норме и будем определять y *проекционным условием*

$$r_k \perp L_k.$$

Из сказанного вытекает, что

$$x_k = x_{k-1} + \alpha_k p_k, \quad r_k = r_{k-1} - \alpha_k A p_k,$$

где α_k определяется проекционным условием.

Если $r_k = 0$, то решение уже найдено. Если $r_k \neq 0$, то ищем p_{k+1} в виде

$$p_{k+1} = r_k + \gamma_{11} p_1 + \ldots + \gamma_{k1} p_k \quad \Rightarrow \quad \gamma_{jk} = -(r_k, A^* p_j)/(Ap_j, p_j).$$

Таким образом, если у нас есть формально A-ортогональный базис $p_1, \ldots, p_k$ в L_k, то мы можем найти вектор p_{k+1} такой, что $(Ap_{k+1}, p_j) = 0,\ 1 \leq j \leq k$.

В отличие от случая положительно определенной матрицы теперь, однако, *ниоткуда не следует*, что $(Ap_{k+1}, p_{k+1}) \neq 0$. Это свойство отнесем к *основным предположениям*; в частности, мы предполагаем, что $(Ar_0, r_0) \neq 0$. Если невязки $r_0, r_1, \ldots, r_{k-1}$ ненулевые и формально A-ортогональный базис $p_1, \ldots, p_k$ в L_k построен, то будем говорить, что процесс *не обрывается* на k-м шаге. Если при этом $r_k = 0$, то будем говорить, что процесс *успешно завершается* на k-м шаге.

Лемма 1. *Если процесс не обрывается на k-м шаге, то невязки $r_0, \ldots, r_{k-1}$ образуют ортогональный базис в L_k.*

Доказательство. Действительно, $r_j \in L_{j+1} \subset L_k$ при $0 \leq j \leq k-1$ и, в силу проекционного условия, $r_j \perp r_0, \ldots, r_{j-1}$. $\square$

Вопрос о "коротких" рекуррентных соотношениях поставим следующим образом. [4] Пусть фиксировано $1 \leq s \leq n-1$, и предположим, что всякий раз, когда процесс не обрывается на k-м шаге, имеют место равенства

$$\gamma_{jk} = (r_k, A^* p_j) = 0 \quad \text{при} \quad 1 \leq j \leq k - s. \tag{1}$$

Это означает, что p_{k+1} выражается через s последних векторов базиса:

$$p_{k+1} \ = \ r_k + \sum_{j=k-s+1}^{k} \gamma_{jk} p_j.$$

Какими свойствами при этом должна обладать матрица A?

Рассмотрим такие матрицы, для которых A^* есть многочлен от A вида

$$A^* \ = \ \sum_{j=0}^{s-1} a_j A^j. \tag{2}$$

[4] Данный вопрос усиленно дискутировался в конце 1970-х годов. Простое и ясное решение, которое мы здесь излагаем, основано на идеях статьи: В. В. Воеводин, Е. Е. Тыртышников, Об обобщении методов сопряженных направлений, *Численные методы алгебры*, Издательство Московского университета, 1981, с. 3–9. В 2004 г. Йорг Лиезен и Поль Сэйлор заметили, что использованное в этой статье дополнительное ограничение на порядок матрицы легко снимается. Заметим, что другое, причем весьма сложное, доказательство необходимости условия (2) было опубликовано в 1984 г. Фабером и Мантефелем (V. Faber, T. Manteuffel, Necessary and sufficient conditions for the existence of a conjugate gradient method, *SIAM J. Numer. Anal.*, vol. 21, no. 2, 1984, pp. 352–362).

Лемма 2. *Пусть имеет место* (2). *Тогда для любой начальной невязки* $r_0 \neq 0$, *не дающей обрыва процесса на* k*-м шаге, выполняются равенства* (1).

Доказательство. В силу (2), $A^* p_j$ есть линейная комбинация векторов $p_1, \ldots, p_{j+s}$. Согласно проекционному условию, $r_k \perp p_1, \ldots, p_{j+s}$ при $j + s \leq k \quad \Rightarrow \quad (1)$. $\square$

Лемма 3. *Предположим, что начальная невязка* $r_0 \neq 0$ *такова, что процесс не обрывается на* n*-м шаге и при этом выполняются равенства* (1) *для всех* $1 \leq k \leq n$. *Тогда для некоторых чисел* $\alpha_j = \alpha_j(r_0)$ *имеет место соотношение*

$$A^* r_0 \;=\; \sum_{j=0}^{s-1} \alpha_j A^j r_0.$$

Доказательство. То, что процесс не обрывается на n-м шаге, означает ортогональность невязок $r_0, \ldots, r_{n-1}$ и линейную независимость векторов $r_0, Ar_0, \ldots, A^{n-1} r_0$. Равенства $(Ar_k, p_j) = 0$ при $1 \leq j \leq k - s$ означают, что $(Ar_k, r_j) = 0$ при $0 \leq j \leq k - s - 1$. Следовательно, $A^* r_0 \perp r_k$ при $k \geq s - 1 \quad \Rightarrow$ $A^* r_0$ есть линейная комбинация векторов $r_0, \ldots, r_{s-2} \quad \Rightarrow \quad A^* r_0$ есть линейная комбинация векторов $r_0, Ar_0, \ldots, A^{s-1} r_0$. $\square$

Теорема. *Пусть* $1 \leq s < n$ *и матрица* A *такова, что хотя бы для одной начальной невязки* $r_0 \neq 0$ *процесс не обрывается на* n*-м шаге. Тогда для всех начальных невязок с тем же свойством для выполнения условия* (1) *необходимо и достаточно, чтобы матрица* A *удовлетворяла соотношению* (2).

Доказательство. Достаточность получена в лемме 2, поэтому перейдем сразу к доказательству необходимости. Линейная независимость векторов $r_0, Ar_0, \ldots, A^{n-1} r_0$ означает, что степень минимального многочлена матрицы A равна $n \quad \Rightarrow$ для каждого собственного значения имеется ровно одна жорданова клетка. Пусть $x = r_0$ и $y = Ar_0$. Ясно, что в случае начальной невязки, равной x или y, процесс не обрывается на n-м шаге. Более того, для начальной невязки вида $x + \gamma y$ процесс может обрываться ранее n-го шага лишь для какого-то конечного числа значений γ (не более числа жордановых клеток для A). Согласно лемме 3, имеем

$$A^* x = \sum_{j=0}^{s-1} \alpha_j A^j x, \qquad A^* y = \sum_{j=0}^{s-1} \beta_j A^j y, \qquad A^*(x + \gamma y) = \sum_{j=0}^{s-1} \phi_j A^j (x + \gamma y).$$

Отсюда, с учетом равенства $y = Ax$,

$$\alpha_0 x \;+\; \sum_{j=1}^{s-1} (\alpha_j + \gamma \beta_{j-1}) A^j x \;+\; \beta_{s-1} A^s x \;=\; \phi_0 x \;+\; \sum_{j=1}^{s-1} (\phi_j + \gamma \phi_{j-1}) A^j x \;+\; \phi_{s-1} A^s x \quad \Rightarrow$$

$$\phi_0 = \alpha_0; \qquad \phi_j + \gamma \phi_{j-1} = \alpha_j + \gamma \beta_{j-1}, \quad 1 \leq j \leq s - 1; \qquad \phi_{s-1} = \beta_{s-1}.$$

Вычтем из второго равенства первое, умноженное на γ: $\phi_1 = \alpha_1 + \gamma(\beta_0 - \alpha_0)$. Это равенство умножим на γ и вычтем из третьего равенства: $\phi_2 = \alpha_2 + \gamma(\beta_1 - \alpha_1) - \gamma^2(\beta_0 - \alpha_0)$. И так далее. В итоге получаем

$$\phi_{s-1} = \beta_{s-1} = \alpha_{s-1} + \gamma(\beta_{s-2} - \alpha_{s-2}) - \gamma^2(\beta_{s-3} - \alpha_{s-3}) + \ldots + (-1)^s \gamma^{s-2}(\beta_0 - \alpha_0) \quad \Rightarrow$$

$$\sum_{j=0}^{s-2} \gamma^{s-2-j} (\beta_j - \alpha_j)(-1)^{s-j} = 0.$$

Последнее соотношение должно выполняться для бесконечного числа значений $\gamma \quad \Rightarrow \quad \alpha_j = \beta_j$ для всех $0 \leq j \leq s - 1$. Следовательно, равенство

$$A^* z = \sum_{j=0}^{s-1} \alpha_j A^j z$$

выполняется с одними и теми же числами α_j для каждого из векторов $z = x, Ax, \ldots, A^{n-1} x$, образующих базис в $\mathbb{C}^n$. Поэтому получаем матричное равенство (2), в котором $a_j = \alpha_j$. $\square$

Лекция 39

ОСНОВНАЯ ЧАСТЬ

39.1 Спектральные задачи

Множество собственных значений матрицы называется также ее *спектром*, а любые задачи и свойства, связанные с собственными значениями и векторами, называются спектральными. В этом плане термин "спектральная норма матрицы" вполне понятен: норма $||A||_2$ равна старшему сингулярному числу матрицы A, которое есть корень квадратный из старшего собственного значения матрицы AA^*.

Методы решения спектральных задач обычно основаны на редукции задачи к аналогичной задаче для матрицы "простого вида", для которой задача решается уже очевидным образом. Существенное отличие от задач, связанных с системами линейных алгебраических уравнений, заключается в том, что в спектральных задачах редукция почти всегда содержит бесконечное число шагов. На практике это означает, что с помощью конечного числа шагов исходная матрица приводится к матрице все еще достаточно общего вида, но такой, что путем замены "малых" элементов на нули из нее получается искомая матрица "простого вида".

Таким образом, при решении спектральных задач очень важно знать, как изменяются спектральные свойства при малых возмущениях элементов матрицы. Прежде всего, что будет с собственными значениями? Этот вопрос, очевидно, сводится к вопросу об изменении корней многочлена при изменении коэффициентов.

Пусть $x_1, \ldots, x_n$ и $y_1, \ldots, y_n$ — полные системы корней (с учетом кратностей) двух многочленов степени n. Базой для изучения "близких" систем корней может служить разумным образом определенное расстояние между n-элементными системами. Например, такое:

$$\rho_p(x,\ y)\ =\ \min_Q ||x - Qy||_p, \qquad \lambda = \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}, \quad \mu = \begin{bmatrix} y_1 \\ \ldots \\ y_n \end{bmatrix},$$

минимум берется по всем матрицам перестановок Q порядка n, $p \geq 1$.

39.2 Непрерывность корней многочлена

Лемма 1. *Любой корень ζ многочлена $f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n$ удовлетворяет неравенству*

$$|\zeta|\ \leq\ \max\left(||a||_1,\ \sqrt[n]{||a||_1}\right), \qquad ||a||_1\ =\ |a_0| + |a_1| + \ldots + |a_{n-1}|.$$

Доказательство. Пусть $f(\zeta) = 0 \ \Rightarrow\ |\zeta|^n \leq |a_0| + |a_1||\zeta| + \ldots + |a_{n-1}||\zeta|^{n-1}$. Если $|\zeta| \leq 1$, то получаем $|\zeta|^n \leq ||a||_1$. Если $|\zeta| > 1$, то $|\zeta|^n \leq ||a||_1 |\zeta|^{n-1} \ \Rightarrow\ |\zeta| \leq ||a||_1$. $\square$

Если даны многочлены $f(z) = \sum\limits_{i=0}^{n} a_i z^i$, $g(z) = \sum\limits_{i=0}^{n} b_i z^i$, то пусть, по определению,

$$||f - g||_1\ =\ \sum_{i=0}^{n} |a_i - b_i|.$$

Предположим далее, что $a_n = b_n = 1$. Корни $f(z)$ и $g(z)$ обозначим через $x_1, \ldots, x_n$ и $y_1, \ldots, y_n$ и составим из них векторы-столбцы $x = [x_1, \ldots, x_n]^\top$ и $y = [y_1, \ldots, y_n]^\top$.

Лемма 2. *Существует перестановка $i_1, \ldots, i_n$ номеров $1, \ldots, n$ такая, что*

$$\sum_{k=1}^{n} |x_k - y_{i_k}| \;\le\; \sum_{k=1}^{n} |g_k(x_k)|^{1/(n+1-k)},$$

где $g_1(z) = g(z)$ и $g_k(z) = g_{k+1}(z)(z - y_{i_k})$, $1 \le k \le n-1$.

Доказательство. Если $|x_1 - y_{i_1}| = \min\limits_{1\le i\le n} |x_1 - y_i|$, то $|g(x_1)| = \left|\prod\limits_{i=1}^{n}(x_1 - y_i)\right| \ge |x_1 - y_{i_1}|^n \;\Rightarrow$ $|x_1 - y_{i_1}| \le |g(x_1)|^{1/n}$. Пусть $f_1(z) = f(z)$ и $f_k(z) = f_{k+1}(z)(z - x_k)$, $1 \le k \le n-1$. Тогда если $|x_2 - y_{i_2}| = \min\limits_{i\ne i_1} |x_2 - y_i|$, то $|g_2(x_2)| = \left|\prod\limits_{i\ne i_1}(x_2 - y_i)\right| \ge |x_2 - y_{i_2}|^{n-1} \;\Rightarrow\; |x_2 - y_{i_2}| \le |g_2(x_2)|^{1/(n-1)}$. И так далее. $\square$

Лемма 3. *Пусть ζ и η — корни многочленов $f(z)$ и $g(z)$, и пусть многочлены $\phi(z)$ и $\psi(z)$ определены равенствами $f(z) = \phi(z)\,(z - \zeta)$ и $g(z) = \psi(z)\,(z - \eta)$. Тогда*

$$||\phi - \psi||_1 \;\le\; \gamma(\alpha,\beta)(\alpha + \beta), \qquad \alpha = ||f - g||_1, \quad \beta = |\zeta - \eta|,$$

где $\gamma(\alpha,\beta)$ — непрерывная функция от α и β.

Доказательство. Если $f(z) = \sum\limits_{i=0}^{n} a_i z^i$, $\phi(z) = \sum\limits_{i=0}^{n-1} c_i z^i$ и $g(z) = \sum\limits_{i=0}^{n} b_i z^i$, $\psi(z) = \sum\limits_{i=0}^{n-1} d_i z^i$, то

$$a_i = c_{i-1} - c_i\zeta, \quad b_i = d_{i-1} - d_i\eta, \qquad 0 \le i \le n,$$

если условиться, что $c_{-1} = c_n = 0 = d_{-1} = d_n$. Отсюда получаем

$$c_{i-1} - d_{i-1} = (a_i - b_i) + (c_i - d_i)\zeta + d_i(\zeta - \eta), \quad 1 \le i \le n.$$

Остается учесть оценку леммы 1 для ζ. $\square$

Теорема. *Для любого достаточно малого $\varepsilon > 0$ существует $\delta > 0$ такое, что если $||f - g||_1 \le \delta$, то $\rho_1(x, y) \le \varepsilon$.*

Доказательство. Рассмотрим многочлены $g_k(z)$ и $f_k(z)$, возникшие в формулировке и доказательстве леммы 2. Очевидно, $f_k(x_k) = 0$. Поэтому

$$\rho_1(x, y) \;\le\; \sum_{k=1}^{n} |x_k - y_{i_k}| \;\le\; \sum_{k=1}^{n} |f_k(x_k) - g_k(x_k)|^{1/(n+1-k)}.$$

Фиксируем $f(z)$ и рассмотрим многочлены $g(z)$ с достаточно малой нормой $||f - g||_1$ (старшие коэффициенты многочленов равны 1). Согласно лемме 1, все корни многочленов $g(z)$ ограничены. Ясно, что $|f_k(x_k) - g_k(x_k)| \le c||f_k - g_k||_1$ с некоторой константой $c > 0$. Применяя лемму 3, находим, что $||f_{k+1} - g_{k+1}||_1$ стремится к нулю, если $||f_k - g_k||_1$ стремится к нулю. Поэтому $\max\limits_{1\le k\le n} ||f_k - g_k||_1$ стремится к нулю, если $||f - g||_1$ стремится к нулю. $\square$

Замечание. Более тонкое рассуждение позволяет получить оценку

$$\rho_1(x, y) \le cn||f - g||_1^{1/n},$$

в которой показатель $1/n$ улучшить нельзя. Например, пусть $f(z) = (z-\zeta)^n$ и $g(z) = (z-\zeta)^n - \varepsilon$, $\varepsilon > 0$. Тогда если η — корень $g(z)$, то $|\eta - \zeta| = \varepsilon^{1/n}$. Даже при малом ε величина $\varepsilon^{1/n}$ может оказаться не такой уж малой. Например, если $\varepsilon = 10^{-10}$, то при $n = 10$ получаем $\varepsilon^{1/n} = 0.1$, а при $n = 100$ и $n = 1000$ это будет ≈ 0.79 и ≈ 0.98.

Пример Дж. Х. Уилкинсона. Многочлен $f(z) = \prod\limits_{i=1}^{20}(z - i)$ имеет $n = 20$ различных вещественных корней. Несмотря на доказанный нами факт непрерывной зависимости корней от коэффициентов, при

практически малых возмущениях корни могут измениться очень сильно. В данном случае ситуацию легко проанализировать, воспользовавшись теоремой математического анализа о неявной функции. Пусть $x = x(t)$ — корень многочлена $g_t(z) = f(z) + tz^{19}$, являющийся возмущением корня $x(0) = 20$ при возмущении лишь одного коэффициента исходного многочлена — при z^{19}. Функция $x = x(t)$ — типичный пример неявной функции, заданной уравнением

$$F(x,t) = 0, \quad \text{где} \quad F(x,t) = f(x) + tx^{19}.$$

Отсюда находим $\frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial t} = 0 \quad \Rightarrow \quad \frac{dx}{dt} = -\frac{\partial F}{\partial t} / \frac{\partial F}{\partial x}$. В нашем случае

$$\frac{\partial F}{\partial x} \;=\; \sum_{j=1}^{20} \prod_{\substack{1 \le i \le 20 \\ i \ne j}} (x-i) \;+\; 19\,t\,x^{19} \qquad \Rightarrow \qquad \left.\frac{\partial F}{\partial x}\right|_{x=20,t=0} = 19!.$$

Ясно также, что $\left.\frac{\partial F}{\partial t}\right|_{x=20} = 20^{19}$. Следовательно, при условии $x(0) = 20$ находим

$$\left.\frac{dx}{dt}\right|_{t=0} = -\frac{20^{19}}{19!} \;\approx\; -4.3 \cdot 10^7.$$

39.3 Возмущение спектра матрицы

Любые примеры чувствительности корней многочлена к возмущениям коэффициентов дают примеры чувствительности собственных значений (спектра) матрицы к возмущениям ее элементов — достаточно рассмотреть матрицу Фробениуса для данного многочлена.

При вычислении собственных значений, способных сильно измениться при малых возмущениях элементов матрицы, следует задуматься о том, в какой степени можно доверять полученному ответу. Современная точка зрения на решение спектральных задач [1] связана с изучением так называемых *спектральных портретов*: для заданной матрицы A и параметра $\varepsilon > 0$ это множества вида

$$S(\varepsilon) = \{z \in \mathbb{C} : \sigma_{\min}(A - zI) \le \varepsilon\},$$

где $\sigma_{\min}(B)$ обозначает минимальное сингулярное число матрицы B.

Очевидно, спектр матрицы A содержится в $S(\varepsilon)$. Однако, возмущения порядка ε могут дать матрицу с собственными значениями, изменяющимися в пределах множества $S(\varepsilon)$. Во многих задачах не следует ожидать сколь-нибудь точного вычисления *отдельных* собственных значений, но, в то же время, при фиксированном ε сами множества $S(\varepsilon)$ — объект, мало изменяющийся при малых вариациях элементов матрицы A.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

39.4 Преобразования отражения и вращения

При решении спектральных задач для упрощения вида исходной матрицы A обычно используют унитарное подобие — подобие сохраняет спектр, а унитарность сохраняет сингулярные числа и, следовательно, не меняет спектральные портреты.

На практике унитарное подобие реализуется с помощью последовательности матриц отражения или (комплексных) матриц вращения. Выбор матриц отражения или вращения связан с желанием исключить те или иные элементы. При этом одна матрица вращения позволяет получить один нуль, а одна матрица отражения — нули сразу во всех, кроме одной, позициях столбца или строки.

Исключение с помощью вращений. Всегда существуют комплексные числа $\xi, \eta,\ |\xi| = |\eta| = 1$, и вещественное число ϕ такие, что для заданных комплексных чисел x_1, x_2 получаем

$$\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \xi & 0 \\ 0 & \eta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;=\; \begin{bmatrix} y_1 \\ 0 \end{bmatrix}.$$

[1] Описание и развитие данной точки зрения можно найти в книге: С. К. Годунов, *Современные аспекты линейной алгебры*, Научная книга, Новосибирск, 1997.

Если $x_1 = 0$, положим $\xi = 1$, в противном случае пусть $\xi = |x_1|/x_1$. Аналогично, если $x_2 = 0$, то $\zeta = 1$, иначе пусть $\eta = |x_2|/x_2$. Таким образом, числа ξx_1 и ηx_2 вещественные и даже неотрицательные. Угол ϕ выбирается из условия $(\xi x_1)\cos\phi + (\eta x_2)\sin\phi = 0$.

Исключение с помощью отражений. Всегда существует вектор $v = [v_1, \ldots, v_n]^\top \in \mathbb{C}^n$, $||v||_2 = 1$, такой, что для заданных комплексных чисел x_1, …, x_n получаем

$$(I - 2vv^*)\begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \\ \ldots \\ 0 \end{bmatrix}.$$

Докажем более общее предложение: *если* $x = [x_1, \ldots, x_n]^\top$, $y = [y_1, \ldots, y_n]^\top$ *и* $||x||_2 = ||y||_2$, *то найдется вектор* v, $||v||_2 = 1$, *такой, что* $(I - 2vv^*)x = \gamma y$ *для некоторого* γ, $|\gamma| = 1$.

Если $x \neq \gamma y$, положим $u = x - \gamma y$, $\quad v = u/||u||_2$. Тогда

$$x - 2v(v^*x) = \gamma y \quad \Rightarrow \quad 2v(v^*x) = u \quad \Rightarrow \quad 2(u^*x) = ||u||_2^2.$$

Последнее уравнение позволяет найти γ:

$$2(x^*x - \bar\gamma y^*x) = ||x||_2^2 + ||y||_2^2 - 2\mathrm{Re}\,(\bar\gamma y^*x).$$

Поскольку $||x||_2 = ||y||_2$, отсюда вытекает, что число $\bar\gamma y^*x$ вещественное. Если $y^*x = 0$, то можно взять любое γ с модулем 1. В противном случае у нас ровно две возможности: $\gamma = y^*x/|y^*x|$ или $\gamma = -y^*x/|y^*x|$. $\square$

39.5 Приведение к треугольному виду

Матрицу можно привести к треугольному виду путем последовательного исключения элементов с помощью умножения ее слева на матрицы отражения или вращения. При использовании отражений умножений будет максимум $n - 1$, в случае вращений их не более $(n^2 - n)/2$.

Вот три шага исключения при $n = 4$ в случае отражений:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \mapsto \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix} \mapsto \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & c_{22} & c_{23} & c_{24} \\ 0 & 0 & c_{33} & c_{34} \\ 0 & 0 & c_{43} & c_{44} \end{bmatrix} \mapsto \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & c_{22} & c_{23} & c_{24} \\ 0 & 0 & d_{33} & d_{34} \\ 0 & 0 & 0 & d_{44} \end{bmatrix}.$$

Данное построение является конструктивным доказательством существования QR-разложения матрицы. Оно полезно при решении линейных систем, особенно в задачах, связанных с методом наименьших квадратов.

39.6 Приведение к почти треугольному виду

Унитарно подобное преобразование матрицы к треугольному виду за конечное число шагов невозможно — иначе существовал бы конечный алгоритм получения собственных значений матрицы и корней многочленов. Однако, за конечное число шагов исключения элементов можно получить унитарно подобную почти треугольную матрицу.

Например, при $n = 4$ преобразования выглядят так:

$$Q_1 \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} Q_1^* = \begin{bmatrix} a_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix},$$

$$Q_2 \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix} Q_2^* = \begin{bmatrix} b_{11} & b_{12} & c_{13} & c_{14} \\ b_{21} & b_{22} & c_{23} & c_{24} \\ 0 & c_{32} & c_{33} & c_{34} \\ 0 & 0 & c_{43} & c_{44} \end{bmatrix}.$$

При умножении слева на матрицу отражения Q_1 первая строка не изменяется, а в первом столбце появляются два нуля. При умножении на Q_1^* справа сохраняется первый столбец, а значит, и два полученных в нем нуля. Далее, умножение слева на Q_2 дает еще один нуль и не меняет первые две строки. Умножение справа на Q_2^* сохраняет первые два столбца, и следовательно, все ранее полученные в них нули.

Заметим, что если исходная матрица A эрмитова, то такой же будет и полученная в итоге верхняя почти треугольная матрица. Ее эрмитовость означает, очевидно, что она в данном случае оказывается *трехдиагональной матрицей*.

39.7 Приведение к двухдиагональному виду

Используя для умножений слева и справа разные матрицы отражения или вращения, любую заданную матрицу можно привести к верхнему двухдиагональному виду.

При $n = 4$ это делается таким образом:

$$U_1 \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix}, \qquad \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & b_{22} & b_{23} & b_{24} \\ 0 & b_{32} & b_{33} & b_{34} \\ 0 & b_{42} & b_{43} & b_{44} \end{bmatrix} V_1^* = \begin{bmatrix} b_{11} & c_{12} & 0 & 0 \\ 0 & c_{22} & c_{23} & c_{24} \\ 0 & c_{32} & c_{33} & c_{34} \\ 0 & c_{42} & c_{43} & c_{44} \end{bmatrix},$$

$$U_2 \begin{bmatrix} b_{11} & c_{12} & 0 & 0 \\ 0 & c_{22} & c_{23} & c_{24} \\ 0 & c_{32} & c_{33} & c_{34} \\ 0 & c_{42} & c_{43} & c_{44} \end{bmatrix} = \begin{bmatrix} b_{11} & c_{12} & 0 & 0 \\ 0 & d_{22} & d_{23} & d_{24} \\ 0 & 0 & d_{33} & d_{34} \\ 0 & 0 & d_{43} & d_{44} \end{bmatrix}, \qquad \begin{bmatrix} b_{11} & c_{12} & 0 & 0 \\ 0 & d_{22} & d_{23} & c_{24} \\ 0 & 0 & d_{33} & d_{34} \\ 0 & 0 & d_{43} & d_{44} \end{bmatrix} V_2^* = \begin{bmatrix} b_{11} & c_{12} & 0 & 0 \\ 0 & d_{22} & e_{23} & 0 \\ 0 & 0 & e_{33} & e_{34} \\ 0 & 0 & e_{43} & e_{44} \end{bmatrix}.$$

Умножение слева на U_1 дает три нуля в первом столбце. После этого умножение справа на V_1^* добавляет два нуля в первой строке и не изменяет первый столбец. Важно, что при каждом преобразовании сохраняются все нули, полученные ранее.

39.8 Вычисление сингулярных чисел

Унитарное приведение к двухдиагональному виду дает возможность свести задачу о вычислении сингулярного разложения матрицы к аналогичной задаче для двухдиагональной матрицы. Более того, можно считать, что все ее элементы неотрицательны (этого можно добиться умножением слева и справа на диагональные унитарные матрицы). Итак, пусть

$$A_0 = \begin{bmatrix} a_1^0 & b_1^0 & & & \\ & a_2^0 & b_2^0 & & \\ & & \cdots & \cdots & \\ & & & a_{n-1}^0 & b_{n-1}^0 \\ & & & & a_n^0 \end{bmatrix}.$$

Рассмотрим следующий бесконечный процесс исключения элементов, начинающийся с вещественной двухдиагональной матрицы A_0 и использующий вещественные матрицы вращения. Последовательно исключая наддиагональные элементы $b_1^0, \ldots, b_{n-1}^0$ с помощью умножения на матрицы вращения справа, преобразуем A_0 в нижнюю двухдиагональную матрицу

$$A_1 = \begin{bmatrix} a_1^1 & & & & \\ b_1^1 & a_2^1 & & & \\ & b_2^1 & \cdots & & \\ & & \cdots & a_{n-1}^1 & \\ & & & b_{n-1}^1 & a_n^1 \end{bmatrix}.$$

Далее будем последовательно исключать поддиагональные элементы и вновь будем иметь верхнюю двухдиагональную матрицу A_2. Затем из A_2 умножениями справа получим нижнюю двухдиагональную матрицу A_3, и так далее.

Данный процесс описывается равенствами

$$A_k Q_k = A_{k+1}, \quad Q_{k+1} A_{k+1} = A_{k+2}, \quad k = 0, 2, \ldots,$$

где матрицы Q_k являются унитарными. Ясно, что для любого k матрица A_k унитарно подобна A_0.

Обозначим через $a_1^k, \ldots, a_n^k$ элементы главной диагонали A_k, а через $b_1^k, \ldots, b_{n-1}^k$ элементы второй диагонали (верхней или нижней в зависимости от четности k). Все числа вещественные. Сохранение длин столбцов при умножении на Q_k слева и сохранение длин строк при умножении на Q_k справа дает следующую систему равенств:

$$\begin{array}{lll} (b_1^k)^2 + (a_1^k)^2 = (a_1^{k+1})^2, & & \\ (b_2^k)^2 + (a_2^k)^2 = (a_2^{k+1})^2 & + (b_1^{k+1})^2, & \\ \cdots \quad \cdots & \cdots & \\ (a_n^k)^2 = (a_n^{k+1})^2 & + (b_{n-1}^{k+1})^2. & \end{array}$$

Поскольку эти равенства имеют место для всех k, находим, в частности, что

$$(a_n^0)^2 \geq \sum_{i=2}^{k} (b_{n-1}^i)^2 \quad \forall \ k \quad \Rightarrow \quad b_{n-1}^k \to 0 \ \text{ при } \ k \to \infty.$$

Отсюда вытекает также существование предела при $k \to \infty$ для (монотонно убывающей) последовательности a_n^k. Аналогичным образом можно доказать, что

$$b_j^k \to 0 \quad \text{при} \quad k \to \infty, \qquad 1 \le j \le n-1,$$

а также и существование пределов при $k \to \infty$ для последовательностей диагональных элементов a_j^k. Эти пределы, конечно же, будут равны сингулярным числам исходной матрицы A_0.

Данный процесс дает некоторое общее представление о том, как могут строиться алгоритмы для вычисления сингулярных чисел. Некоторые черты того же процесса можно обнаружить и в алгоритмах вычисления собственных значений. Следует заметить, однако, что эффективность алгоритмов, используемых в современных пакетах и библиотеках программ, связана с определенным числом очень важных деталей и идей, которые мы обсудить здесь не имели возможности.

39.9 Локализация собственных значений

Пусть $A = [a_{ij}] \in \mathbb{C}^{n\times n}$. Если $Ax = \lambda x,\ x \ne 0$, то $||Ax|| = ||\lambda x|| \le ||A||||x|| \quad \Rightarrow \quad |\lambda| \le ||A||$. Полученное неравенство справедливо при использовании *любой матричной нормы.*

Чтобы получить более детальную локализацию собственных значений матрицы A, рассмотрим на комплексной плоскости так называемые *круги Гершгорина*

$$D_i(A) \ = \ \{z \in \mathbb{C} : |z - a_{ii}| \le \sum_{1\le j\le n,\ j\ne i} |a_{ij}|\}, \quad 1 \le i \le n.$$

Первая теорема Гершгорина. *Любое собственное значение матрицы $A \in \mathbb{C}^{n\times n}$ принадлежит объединению кругов Гершгорина для A и одновременно объединению кругов Гершгорина для $A^\top$.*

Доказательство. Предположим, что $|a_{ii} - \lambda| > \sum\limits_{1\le j\le n,\, j\ne i} |a_{ij}|, \quad 1 \le i \le n$. Это означает, что $A - \lambda I$ является матрицей с диагональным преобладанием по строкам и поэтому обратима (см. раздел 8.10). Значит, никакое комплексное число $\lambda \notin \bigcup\limits_{1\le i\le n} D_i(A)$ не может быть собственным значением для A. $\square$

Вторая теорема Гершгорина. *Предположим, что объединение k кругов Гершгорина $\mathcal{D} = D_{i_1} \cup \ldots \cup D_{i_k}$ для матрицы A не имеет общих точек с остальными кругами Гершгорина. Тогда $\mathcal{D}$ содержит ровно k собственных значений матрицы A.*

Доказательство. Обозначим через $B = [b_{ij}]$ диагональную матрицу порядка n с элементами $b_{ii} = a_{ii}$ и рассмотрим семейство матриц $A(t) = At + (1-t)B$ при $0 \le t \le 1$. Очевидно, $A(0) = B$ и $A(1) = A$. Обозначим через $\lambda(t) = [\lambda_1(t), \ldots, \lambda_n(t)]^\top$ вектор-столбец, составленный из собственных значений матрицы $A(t)$, и через $\nu(t)$ число компонент $\lambda(t)$, принадлежащих $\mathcal{D}$. Зафиксируем t_0. Тогда при всех t, достаточно близких к t_0, должно быть $\nu(t) = \nu(t_0)$. Если это не так, то существует последовательность t_m, сходящаяся к t_0 при $m \to \infty$ и такая, что для любой матрицы перестановки P

$$\rho_1(\lambda(t_m),\, \lambda(t_0)) \ge ||\lambda(t_0) - P\lambda(t_m)||_1 \ge d \equiv \inf_{u\in\mathcal{D},\ v\in\mathcal{D}'} |u - v|,$$

где $\mathcal{D}'$ — объединение кругов Гершгорина, не входящих в $\mathcal{D}$. Данное неравенство противоречит теореме о непрерывной зависимости корней многочлена от коэффициентов (а значит, и собственных значений матрицы от ее элементов). Таким образом, функция $\nu(t)$ непрерывна по t и принимает целочисленные значения $\quad \Rightarrow \quad \nu(t)$ — константа. При этом $\nu(0) = k \quad \Rightarrow \quad \nu(t) = k$ для всех $0 \le t \le 1$. $\square$

Отметим еще одно простое утверждение, приводящее к серии результатов по локализации собственных значений при возмущениях заданной матрицы.

Теорема Бауэра–Файка. *Если μ является собственным значением матрицы $B = A + F$, но не является собственным значением матрицы A, то $1/||(A - \mu I)^{-1}||_2 \ \le \ ||F||_2$.*

Доказательство. Матрица $B - \mu I = (A - \mu I) + F$ вырожденная $\quad \Rightarrow \quad$ матрица $I + (A - \mu I)^{-1}F$ вырожденная $\quad \Rightarrow \quad ||(A - \mu I)^{-1}||_2||F||_2 \ge ||(A - \mu I)^{-1}F||_2 \ge 1$. $\square$

Следствие. *Пусть A — диагонализуемая матрица, и предположим,что $AX = X\Lambda$, где X — матрица из собственных векторов, Λ — диагональная матрица собственных значений $\lambda_1, \ldots, \lambda_n$ матрицы A. Тогда собственные значения матрицы $B = A + F$ принадлежат объединению кругов вида*

$$K_i \ = \ \{z \in \mathbb{C} : |z - \lambda_i| \ \le \ ||X||_2\, ||X^{-1}||_2\, ||F||_2\}, \quad 1 \le i \le n.$$

Доказательство. Пусть μ — собственное значение для B, но не для A. Тогда μ есть собственное значение для $\Lambda + X^{-1}FX$, но не для Λ. Остается применить теорему Бауэра–Файка. $\square$

39.10 Расстояние между спектрами нормальных матриц

Теорема Виландта–Хоффмана. *Пусть A и B — нормальные матрицы с собственными значениями $\lambda_1(A), \ldots, \lambda_n(A)$ и $\lambda_1(B), \ldots, \lambda_n(B)$. Тогда для некоторой подстановки $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$*

$$\sum_{i=1}^{n} |\lambda_i(A) - \lambda_{\sigma(i)}(B)|^2 \;\le\; ||A - B||_F^2.$$

Доказательство. Запишем $A = Q\Phi Q^*$, $B = Z\Psi Z^*$, где Q, Z — унитарные матрицы, а Φ и Ψ — диагональные матрицы из собственных значений $\phi_i = \lambda_i(A)$ и $\psi_i = \lambda_i(B)$. В силу унитарной инвариантности нормы Фробениуса, $||A - B||_F = ||\Phi - V\Psi V^*||_F$, где $V = Q^*Z$ — унитарная матрица. Далее,

$$\begin{aligned} ||\Phi - V\Psi V^*||_F^2 &= \operatorname{tr}(\Phi^* - V\Psi^* V^*)(\Phi - V\Psi V^*) = \operatorname{tr}(\Phi^*\Psi) + \operatorname{tr}(\Phi^*\Psi) - 2\mathrm{Re}\,(\operatorname{tr}(\Phi^* V)(\Psi V^*)) \\ &= \sum_{i=1}^{n} |\phi_i|^2 + \sum_{i=1}^{n} |\psi_i|^2 - 2\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij} s_{ij}, \qquad \alpha_{ij} = \mathrm{Re}(\overline{\phi_i}\psi_j), \quad s_{ij} = |v_{ij}|^2. \end{aligned}$$

Легко проверить, что матрица $S = [s_{ij}]$ является двоякостохастической (см. раздел 37.9). Поэтому при фиксированных вещественных числах α_{ij} функционал

$$f(S) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij} s_{ij}$$

можно рассматривать как линейный функционал на множестве двоякостохастических матриц. Это замкнутое ограниченное выпуклое множество $\Rightarrow$ максимум линейного функционала на нем достигается в какой-то угловой точке (см. раздел 26.7). Нетрудно убедиться в том, что угловыми точками множества двоякостохастических матриц являются матрицы перестановок и только они $\Rightarrow$ для некоторой матрицы перестановки P и соответствующей ей подстановке σ

$$\max_S f(s) \;\le\; f(P) = \sum_{i=1}^{n} \alpha_{i\,\sigma(i)} = \sum_{i=1}^{n} \mathrm{Re}\,(\overline{\phi_i}\psi_{\sigma(i)}) \quad \Rightarrow$$

$$||A - B||_F^2 \ge \sum_{i=1}^{n} \left(|\phi_i|^2 + |\psi_{\sigma(i)}|^2 - 2\mathrm{Re}\,(\overline{\phi_i}\psi_{\sigma(i)})\right) = \sum_{i=1}^{n} |\phi_i - \psi_{\sigma(i)}|^2. \quad \square$$

Замечание. Теорема о непрерывной зависимости корней многочлена от коэффициентов в данном доказательстве не использовалась. Поэтому теорема Виландта–Хоффмана дает еще одно доказательство факта непрерывной зависимости собственных значений матрицы от ее коэффициентов для специального класса матриц — для нормальных матриц.

Следствие. *Пусть A и B — эрмитовы матрицы с собственными значениями $\lambda_1(A) \ge \ldots \ge \lambda_n(A)$ и $\lambda_1(B) \ge \ldots \ge \lambda_n(B)$. Тогда*

$$\sum_{i=1}^{n} (\lambda_i(A) - \lambda_i(B))^2 \;\le\; ||A - B||_F^2.$$

Доказательство. Пусть $\phi_i = \lambda_i(A)$ и $\psi_i = \lambda_i(B)$. Достаточно заметить, что если $\phi_{\sigma(i_1)} < \phi_{\sigma(i_2)}$ при $i_1 < i_2$, то

$$(\phi_{i_1} - \psi_{\sigma(i_1)})^2 + (\phi_{i_2} - \psi_{\sigma(i_2)})^2 \;\ge\; (\phi_{i_1} - \psi_{\sigma(i_2)})^2 + (\phi_{i_2} - \psi_{\sigma(i_1)})^2. \quad \square$$

Лекция 40

ОСНОВНАЯ ЧАСТЬ

40.1 Многомерные массивы и матрицы

Матрицу можно рассматривать как способ задания числовой функции от дискретных переменных i, j или, в терминологии некоторых языков программирования, как двумерный массив. Данная точка зрения приводит к такому естественному обобщению как m-мерный массив (m-мерная матрица) с элементами $x_{i_1 \dots i_m}$ или функция от m индексов $i_1, \dots, i_m$.

Существенная часть понятий и фактов теории матриц в случае m-мерных массивов при $m \geq 3$ уже утрачивается. А для понятий, возникающих по прямой аналогии, оказывается, что отличий больше, чем сходства.

Так обстоит дело с исключительно важным обобщением понятия ранга. Как и в случае матриц, оно связано с разделением переменных $i_1, \dots, i_m$, приводящим к *m-линейному разложению*

$$x_{i_1 \dots i_m} = \sum_{s=1}^{r} u_{i_1 s} \dots u_{i_m s}, \qquad 1 \leq i_1 \leq n_1, \dots, 1 \leq i_m \leq n_m.$$

Наименьшее число слагаемых r в разложениях такого вида называется *тензорным рангом* m-мерного массива $X = [x_{i_1 \dots i_m}]$.

Как обычно, предполагается, что элементы массивов $x_{i_1, \dots, i_m}$ и разложений $u_{i_1 s}, \dots, u_{i_m s}$ принадлежат некоторому общему числовому полю. В отличие от ранга матриц, тензорные ранги могут зависеть от этого поля. Поэтому скажем сразу, что в дальнейшем таким полем является поле вещественных чисел.

Матричные методы могут быть полезны и для многомерных массивов — простой прием позволяет ассоциировать их с некоторыми прямоугольными матрицами. Разобьем систему индексов $i_1, \dots, i_m$ на две непересекающиеся подсистемы

$$i'_1, \dots, i'_p \quad \text{и} \quad j'_1, \dots, j'_q, \qquad p + q = m,$$

и пусть $y_{(i'_1, \dots, i'_m), (j'_1, \dots, j'_m)} = x_{i_1 \dots i_m}$. Тогда $Y = [y_{(i'_1, \dots, i'_p), (j'_1, \dots, j'_q)}]$ есть матрица, в которой роль строчного и столбцового индексов играют $(i'_1, \dots, i'_p)$ и $(j'_1, \dots, j'_q)$.

40.2 Трехмерные массивы и трилинейные разложения

Остановимся подробнее на случае трехмерных массивов. Под *трилинейным разложением* трехмерного массива $X = [x_{ijk}]$ понимается разложение вида

$$x_{ijk} = \sum_{s=1}^{r} a_{is} b_{js} c_{ks}.$$

Обозначение: $X = (A, B, C)$, где A, B, C — матрицы вида

$$A = [a_{is}] = [a_1, \dots, a_r], \qquad B = [b_{js}] = [b_1, \dots, b_r], \qquad C = [c_{ks}] = [c_1, \dots, c_r].$$

Число столбцов для матриц A, B, C одно и то же и равно r, число строк для них определяется границами для индексов i, j и k — пусть это будут n_1, n_2 и n_3.

Таким образом, любые три матрицы с одним и тем же числом столбцов r порождают трилинейное разложение (A, B, C) некоторого трехмерного массива. Общее число столбцов называется рангом данного трилинейного разложения. Среди всех трилинейных разложений трехмерного массива X имеется, конечно, разложение с минимальным числом столбцов. Его ранг (число столбцов) и называется тензорным рангом трехмерного массива X. Обозначение: $\operatorname{Rank} X$.

40.3 Сечения трехмерного массива

С трехмерным массивом $X = [x_{ijk}]$ ассоциируем три *матрицы сечений*

$$Y = [y_{(i),(jk)}], \quad Z = [z_{(j),(ik)}], \quad W = [w_{(k),(ij)}], \qquad y_{(i),(jk)} = z_{(j),(ik)} = w_{(k),(ij)} = x_{ijk},$$

и положим

$$\dim_1 X \equiv \operatorname{rank} Y, \quad \dim_2 X \equiv \operatorname{rank} Z, \quad \dim_3 X \equiv \operatorname{rank} W.$$

Строки матриц Y, Z, W соответствуют "векторизованным" сечениям трехмерного массива X по осям i, j, k, соответственно.

Каждое сечение по оси i представляет собой прямоугольную матрицу $[x_{ijk}]_{i=i_0}$. Очевидно, $\dim_1 X$ есть размерность линейной оболочки, натянутой на матрицы сечений при $i = 1, \ldots, n_1$. Аналогичный смысл имеют величины $\dim_2 X$ и $\dim_3 X$.

Утверждение. $\max(\dim_1 X, \dim_2 X, \dim_3 X) \;\le\; \operatorname{Rank} X \;\le\; \min(n_1 n_2, n_2 n_3, n_1 n_3)$.

Доказательство. Докажем для определенности, что $\dim_1 X \le \operatorname{Rank} X \le n_2 n_3$. Если $r = \operatorname{rank} X$, то существует трилинейное разложение с числом столбцов r:

$$X = (A, B, C) \quad \Rightarrow \quad [x_{ijk}]_{i=i_0} \in L(b_1 c_1^\top, \ldots, b_r c_r^\top) \quad \Rightarrow \quad \dim_1 X \le r.$$

Далее, ранг матрицы W не больше n_3. Поэтому для нее существует разложение вида

$$w_{(k),(ij)} \;=\; \sum_{s=1}^{n_3} \Phi_{ks} \Psi_{(ij),s}.$$

Для каждого s ранг матрицы $[\Psi_{(ij),s}]$ не больше $n_2 \quad \Rightarrow$

$$\Psi_{(ij),s} \;=\; \sum_{t=1}^{n_2} U_{ist} V_{jst} \quad \Rightarrow \quad x_{ijk} = w_{(k),(ij)} = \sum_{s=1}^{n_3} \sum_{t=1}^{n_2} U_{ist} V_{jst} \Phi_{ks}. \qquad \square$$

Аналог сечений для обычных матриц — запись их в виде системы строк или столбцов. В отличие от матриц, для которых строчный и столбцовый ранги совпадают и равны рангу матрицы, четыре числа $\operatorname{rank} X,\ \dim_1 X,\ \dim_2 X,\ \dim_3 X$, вообще говоря, разные.

40.4 Примеры трилинейных разложений

Любой трехмерный $2 \times 2 \times 2$-массив $X = [x_{ijk}]$ определяется двумя сечениями

$$X_1 = [x_{1jk}], \; X_2 = [x_{2jk}].$$

ПРИМЕР 1. $X_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$

Ясно, что $\dim_1 X = 2 \quad \Rightarrow \quad \operatorname{Rank} X \ge 2$. Нетрудно проверить, что

$$\begin{array}{rcl} X_1 & = & \frac{1}{2} b_1 c_1^\top \;+\; \frac{1}{2} b_2 c_2^\top, \\ X_2 & = & \frac{1}{2} b_1 c_1^\top \;-\; \frac{1}{2} b_2 c_2^\top, \end{array} \qquad b_1 = c_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \;\; b_2 = c_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Таким образом, следует взять $a_1 = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix}$. Тогда

$$X = (A, B, C), \quad \text{где} \quad A = [a_1, a_2], \; B = [b_1, b_2], \; C = [c_1, c_2].$$

ПРИМЕР 2. $X_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $X_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Используя трилинейное разложение предыдущего примера, для данного массива мы можем с легкостью получить разложение ранга 3 (сделайте это!). Но верно ли, что разложение меньшего ранга не существует? Допустим, что

$$\begin{aligned} X_1 &= a_{11}b_1c_1^\top + a_{12}b_2c_2^\top, \\ X_2 &= a_{21}b_1c_1^\top + a_{22}b_2c_2^\top. \end{aligned}$$

Каждая из матриц X_1 и X_2 имеет ранг 2 $\Rightarrow$ коэффициенты $a_{11}, a_{12}, a_{21}, a_{22}$ отличны от нуля. Рассмотрим линейную комбинацию

$$V = a_{21}X_1 - a_{11}X_2 = \begin{bmatrix} -a_{21} & -a_{11} \\ -a_{11} & a_{21} \end{bmatrix} \Rightarrow \det V = -a_{21}^2 - a_{11}^2 \neq 0 \Rightarrow \operatorname{rank} V = 2.$$

Преобразуя правые части выражений для X_1 и X_2, находим

$$V = (a_{21}a_{12} - a_{11}a_{22})\, b_2c_2^\top \Rightarrow \operatorname{rank} V \leq 1.$$

Противоречие означает, что $\operatorname{Rank} X \geq 3$.

Замечание. В только что законченном рассуждении предполагалось, что все числа вещественные. Если допустить к рассмотрению трилинейные разложения с комплексными числами, то в данном случае оказывается, что тензорный ранг равен 2.

40.5 Все не так, как всегда

Итак, свойства тензорных рангов трехмерных массивов и рангов матриц различаются коренным образом.

1. Тензорный ранг трехмерных массивов существенно *зависит от числового поля*, которому принадлежат элементы трилинейных разложений.

В дальнейшем всюду полагаем, что числовое поле есть поле вещественных чисел.

2. Для тензорного ранга не известны какие-либо конечные алгоритмы его вычисления — в отличие от ранга матрицы, который в точной арифметике легко находится с помощью конечного числа элементарных преобразований.

3. В общем случае при фиксированных размерах трехмерного массива до сих пор не получены точные значения максимального значения тензорного ранга.

Кое-что, правда, известно. В 1970-х годах Йозеф Крускал доказал, что тензорный ранг произвольного вещественного $2 \times 2 \times 2$-массива не превышает 3. Соединив этот факт с разобранным выше примером, приходим к выводу о том, что максимальное значение тензорного ранга в данном частном случае равно 3.

4. Обратим внимание на специфические "вероятностные" свойства тензорных рангов (при этом оставим строгие определения в стороне и доверимся интуиции): среди всего множества вещественных $2 \times 2 \times 2$-массивов имеется примерно 79% массивов тензорного ранга 2 и примерно 21% массивов тензорного ранга 3.

Это экпериментальные данные, полученные Крускалом. В случае двумерных массивов (матриц) все проще: *почти любая* матрица имеет максимально возможный ранг (равный минимальному из ее размеров).

40.6 Эквивалентные трилинейные разложения

В буквальном смысле трилинейное разложение, конечно, не может быть единственным. Если $X = (A, B, C)$, где $A = [a_1, \ldots, a_r]$, $B = [b_1, \ldots, b_r]$, $C = [c_1, \ldots, c_r]$, то формально другое разложение для того же X легко строится с помощью двух приемов:

(1) можно произвольным, но одинаковым образом переставить столбцы в матрицах A, B, C;

(2) взяв любые числа α_s, β_s, γ_s такие, что $\alpha_s\beta_s\gamma_s = 1$, можно заменить столбцы a_s, b_s, c_s на $\alpha_s a_s$, $\beta_s b_s$, $\gamma_s c_s$.

Эти два приема приводят к разложению $X = (\widetilde{A}, \widetilde{B}, \widetilde{C})$, где

$$\widetilde{A} = APD_A, \quad \widetilde{B} = BPD_B, \quad \widetilde{C} = CPD_C, \qquad (*)$$

P — матрица перестановки, D_A, D_B, D_C — диагональные матрицы такие, что $D_A D_B D_C = I$. Трилинейные разложения (A, B, C) и $(\widetilde{A}, \widetilde{B}, \widetilde{C})$, связанные соотношениями $(*)$, называются *эквивалентными*.

Аналогичным образом вводится понятие эквивалентности для билинейных (скелетных) разложений матриц и m-линейных разложений произвольных m-мерных массивов.

40.7 Единственность с точностью до эквивалентности

Множество билинейных (скелетных) разложений заданной матрицы весьма широко, и его описание не сводится к эквивалентности разложений.

Например, пусть $X = [x_1, x_2]$ — матрица размеров $n \times 2$ с линейно независимыми столбцами x_1, x_2. Для произвольной невырожденной 2×2-матрицы $G = [g_1, g_2]$ запишем $XG^{-1} = [x_1^G, x_2^G]$. Тогда, очевидно,

$$X \;=\; x_1^G g_1^\top + x_2^G g_2^\top . \qquad (*)$$

Высокая степень произвола в компонентах билинейных аппроксимаций матрицы заставляет вводить при их построении различные ограничения — обычно типа ортогональности. Например, сингулярное разложение матрицы X имеет тот же вид $(*)$, но если сингулярные числа различны, то сингулярные векторы будут определены однозначно с точностью до множителя. Это обстоятельство очень важно — оно позволяет использовать сингулярные векторы как носители существенной информации о данных, представленных элементами матрицы.

В случае трехмерных массивов ситуация одновременно и проще, и сложнее. Почему сложнее — понятно: теория и алгоритмы вычисления трилинейных разложений и аппроксимаций далеки от стадии завершенности. А проще вот по какой причине.

Пусть $X = (A, B, C)$ — трилинейное раложение ранга r. Это означает, что каждая из матриц A, B, C имеет r столбцов. Предположим, что каждая их этих матриц имеет линейно независимую систему столбцов. Допустим, что $X = (\widetilde{A}, \widetilde{B}, \widetilde{C})$ — еще одно разложение ранга r с линейно независимыми столбцами в матрицах $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$.

Пусть для ясности $r = 2$. Тогда

$$a_{i1}b_{j1}c_{k1} + a_{i2}b_{j2}c_{k2} = \widetilde{a}_{i1}\widetilde{b}_{j1}\widetilde{c}_{k1} + \widetilde{a}_{i2}\widetilde{b}_{j2}\widetilde{c}_{k2}. \qquad (\#)$$

Выберем вектор $p = [p_1, \ldots, p_{n_1}]^\top$ таким образом, чтобы $p \in \{a_2\}^\perp$, но $p \notin \{a_1\}^\perp$ (вектор p ортогонален a_2, но не a_1) — в смысле естественного скалярного произведения в пространстве $\mathbb{R}^{n_1}$. Умножим равенства $(\#)$ на коэффициенты p_i и просуммируем их по i от 1 до n_1:

$$(p^\top a_1)b_1c_1^\top \;=\; (p^\top \widetilde{a}_1)\,\widetilde{b}_1\widetilde{c}_1^\top \;+\; (p^\top \widetilde{a}_2)\,\widetilde{b}_2\widetilde{c}_2^\top .$$

В силу выбора p, $p^\top a_1 \neq 0 \;\Rightarrow\;$ ранг матрицы в левой части равен 1 $\;\Rightarrow\;$ $p^\top \widetilde{a}_1 = 0$ либо $p^\top \widetilde{a}_2 = 0$, иначе ранг матрицы в правой части был бы равен 2:

$$V \equiv t_1\widetilde{b}_1\widetilde{c}_1^\top + t_2\widetilde{b}_2\widetilde{c}_2^\top = [\widetilde{b}_1, \widetilde{b}_2]\begin{bmatrix} t_1 & \\ & t_2 \end{bmatrix}[\widetilde{c}_1, \widetilde{c}_2]^\top \;\Rightarrow\; \det V = t_1 t_2 \,\det\widetilde{B}\,\det\widetilde{C}.$$

Пусть для определенности $p^\top\widetilde{a}_2 = 0$. Тогда $\widetilde{b}_1\widetilde{c}_1^\top = t\, b_1c_1^\top$, $t \neq 0$. Поскольку все векторы ненулевые, отсюда вытекает, что $\widetilde{b}_1 = \beta_1 b_1$, $\widetilde{c}_1 = \gamma_1 c_1$ для каких-то ненулевых коэффициентов β_1, γ_1.

Далее, мы можем выбрать вектор $q = [q_1, \ldots, q_{n_3}]^\top$, ортогональный c_2, но не ортогональный c_1. Те же равенства $(\#)$ можно умножить на коэффициенты q_k и просуммировать по k от 1 до n_3:

$$(q^\top c_1)a_1b_1^\top \;=\; (q^\top \widetilde{c}_1)\,\widetilde{a}_1\widetilde{b}_1^\top \;+\; (q^\top \widetilde{c}_2)\,\widetilde{a}_2\widetilde{b}_2^\top .$$

Если $q^\top\widetilde{c}_2 \neq 0$, то окажется, что $\widetilde{b}_2 = hb_1$, $h \neq 0 \;\Rightarrow\;$ столбцы $\widetilde{b}_1$, $\widetilde{b}_2$ линейно зависимы. Это противоречит исходным предположениям. Значит, $q^\top\widetilde{c}_2 = 0$. Но тогда, повторяя предыдущие рассуждения, находим $\widetilde{a}_1 = \alpha_1 a_1$, $\widetilde{b}_2 = \beta_2 b_2$. В итоге

$$(q^\top \widetilde{c}_1)\,\widetilde{a}_1\widetilde{b}_1^\top \;=\; (\alpha_1\beta_1\gamma_1)\,(q^\top c_1)a_1b_1^\top \;\;\Rightarrow\;\; \alpha_1\beta_1\gamma_1 = 1.$$

Теперь предположим, что при использовании вектора p оказалось, что $p^\top \widetilde{a}_1 = 0$. Чтобы оставить в силе последовавшие рассуждения, достаточно переставить столбцы в матрицах $\widetilde{A}$, $\widetilde{B}$, $\widetilde{C}$. Таким образом, мы доказали, что трилинейные разложения (A, B, C) и $(\widetilde{A}, \widetilde{B}, \widetilde{C})$ эквивалентны. Легко видеть также, как вести рассуждение в случае $r > 2$. Итак, полностью доказана следующая

Теорема единственности. *Пусть $X = (A, B, C)$ и столбцы в каждой из матриц A, B, C линейно независимы. Тогда трилинейное разложение (A, B, C) определено однозначно с точностью до эквивалентности: если трилинейное разложение $X = (\widetilde{A}, \widetilde{B}, \widetilde{C})$ таково, что каждая из матриц $\widetilde{A}, \widetilde{B}, \widetilde{C}$ с общим числом столбцом $\widetilde{r}$ имеет линейно независимые столбцы, то $\widetilde{r} = r$ и разложения (A, B, C) и $(\widetilde{A}, \widetilde{B}, \widetilde{C})$ эквивалентны.*

Данный факт имеет огромное (возможно, основное) значение в многочисленных применениях трилинейных аппроксимаций к анализу данных (например, при изучении химического состава смесей в спектрометрии или психометрических и социометрических данных при изучении особенностей личности и общества).

Замечание. Единственность с точностью до эквивалентности имеет место и при более слабых предположениях, чем в доказанной нами теореме единственности. В 1970-х годах Крускал доказал следующую теорему: *пусть ранги матриц A, B, C равны r_A, r_B, r_C и пусть любые r_A столбцов из A, любые r_B столбцов из B и любые r_C столбцов из C являются линейно независимыми; если $r_A + r_B + r_C \geq 2r + 2$, где r — общее число столбцов для A, B и C, то трилинейное разложение (A, B, C) определено однозначно с точностью до эквивалентности.* Возможно какое-то ослабление и этих условий.

40.8 Тензорный ранг и умножение матриц

Трилинейные разложения имеют глубокую связь с теорией сложности вычислений. К компетенции данной теории относится, например, вопрос, интересующий каждого, кто имеет дело с матрицами: какова *истинная* сложность умножения двух $n \times n$-матриц? Эпитет подчеркивает, что нас интересует сложность (число операций) самого быстрого алгоритма.

Ответ на этот вопрос до сих пор не получен. Для большинства лиц, когда-то знакомившихся с линейной алгеброй, в памяти остается правило "строка на столбец", дающее $O(n^3)$ операций. Однако, мы можем утверждать, что "истинное" число операций не превышает $O(n^{\log_2 7})$. Именно столько операций дает алгоритм Штрассена, который мы обсуждали в самой первой лекции нашего курса.

Откуда же берется оригинальный способ умножения 2×2-матриц, на котором там все основано? Теперь, в заключительной лекции, мы имеем возможность раскрыть тайну алгоритма Штрассена.

Итак, пусть

$$\begin{bmatrix} u_1 & u_3 \\ u_2 & u_4 \end{bmatrix} \begin{bmatrix} v_1 & v_3 \\ v_2 & v_4 \end{bmatrix} = \begin{bmatrix} w_1 & w_3 \\ w_2 & w_4 \end{bmatrix}.$$

Равенства, выражающие w_k через u_i и v_j, можно, очевидно, записать в такой форме:

$$\begin{cases} w_1 = u_1 v_1 + u_3 v_2, \\ w_2 = u_2 v_1 + u_4 v_2, \\ w_3 = u_1 v_3 + u_3 v_4, \\ w_4 = u_2 v_3 + u_4 v_4. \end{cases} \quad \Leftrightarrow \quad w_k = \sum_{i=1}^{4} \sum_{j=1}^{4} x_{ijk}\, u_i v_j, \quad k = 1, 2, 3, 4.$$

Возникший здесь трехмерный массив $X = [x_{ijk}]$ имеет размеры $4 \times 4 \times 4$, его элементы x_{ijk} равны 0 либо 1. Вот сечения X по оси k:

$$X_{k=1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad X_{k=2} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad X_{k=3} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad X_{k=4} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

В данном случае ясно, что тензорный ранг массива X не больше 8 (докажите!). Пусть он равен r. Тогда имеется трилинейное разложение

$$x_{ijk} = \sum_{s=1}^{r} a_{is} b_{js} c_{ks} \quad \Rightarrow \quad w_k = \sum_{s=1}^{r} c_{ks} \left(\sum_{i=1}^{4} a_{is} u_i \right) \left(\sum_{j=1}^{4} b_{js} v_j \right), \quad k = 1, 2, 3, 4.$$

Как видим, трилинейное разложение ранга r порождает специальный алгоритм вычисления величин w_k, в котором всего r *активных* умножений — так называются умножения, в которых оба множителя существенно зависят от входных переменных u_i и v_j (числа a_{is}, b_{js}, c_{ks} не зависят от u_i, v_j; их называют константами алгоритма — умножение на константу не считается активным умножением).

Чтобы получить открытие Штрассена, достаточно решить задачу о вычислении тензорного ранга данного конкретного массива X. Можно ограничиться и более скромной задачей: найти какое-нибудь трилинейное разложение ранга 7 (разложение ранга 8 связано с правилом "строка на столбец"). Несмотря на отсутствие конечных алгоритмов точного вычисления тензорного ранга, разработка алгоритмов трилинейной аппроксимации заданного ранга является посильной задачей.

Эффективные методы для данной задачи являются одной из крупных исследовательских проблем. Однако, в отдельных случаях можно добиться нужного результата и с помощью каких-либо эвристических и, возможно, "медленных" вычислений: чтобы построить быстрый алгоритм умножения матриц, мы вполне готовы потратить *очень много* времени на поиск тензорного ранга массива X. Вот как выглядит трилинейное разложение $X = (A, B, C)$ ранга 7 в нашем случае: [1]

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 & 1 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Данное разложение найдено с помощью компьютера. Таким образом, компьютер может использоваться не только как инструмент решения вычислительных задач, но также и как инструмент *получения алгоритмов* для решения этих задач.

ДОПОЛНИТЕЛЬНАЯ ЧАСТЬ

40.9 Преобразования массивов с помощью матриц

После обсуждения проблем и трудностей, связанных с многомерными массивами, особенно приятно закончить тему одним "положительным" результатом, легко получаемым с помощью изученной нами матричной техники. Речь идет о так называемом *разложении Таккера* — о нем нередко говорят как о многомерном обобщении сингулярного разложения.

Формулировка результата требует небольшой подготовки. Пусть $X = [x_{ijk}]$ — трехмерный массив размеров $n_1 \times n_2 \times n_3$, и пусть $P = [p_{i'i}]$, $Q = [q_{j'j}]$, $R = [r_{k'k}]$ — матрицы размеров $n'_1 \times n_1$, $n'_2 \times n_2$, $n'_3 \times n_3$, соответственно. Определим новый трехмерный массив $X' = [x'_{i'j'k'}]$ следующим образом:

$$x'_{i'j'k'} \;=\; \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} p_{i'i}\, q_{j'j}\, r_{k'k}\; x_{ijk}.$$

Иногда говорят, что X' есть свертка массива X с матрицами P, Q, R. Обозначение: $X' = X \odot \{P, Q, R\}$. Кроме того, по определению,

$$\begin{aligned} X \odot_1 P &= X \odot \{P,\ I_{n_2 \times n_2},\ I_{n_3 \times n_3}\}, \\ X \odot_2 Q &= X \odot \{I_{n_1 \times n_1},\ Q,\ I_{n_3 \times n_3}\}, \\ X \odot_3 R &= X \odot \{I_{n_1 \times n_1},\ I_{n_2 \times n_2},\ R\}. \end{aligned}$$

Согласно данным определениям,

$$\begin{aligned} X' &= X \odot \{P, Q, R\} \\ &= ((X \odot_1 P) \odot_2 Q) \odot_3 R = ((X \odot_2 Q) \odot_3 R) \odot_1 P = ((X \odot_3 R) \odot_1 P) \odot_2 Q \\ &= ((X \odot_3 R) \odot_2 Q) \odot_1 P = ((X \odot_2 Q) \odot_1 P) \odot_3 R = ((X \odot_1 P) \odot_3 R) \odot_2 Q. \end{aligned}$$

[1] Условия доказанной нами теоремы единственности в данном случае не выполнены. Поэтому можно найти и другое, неэквивалентное данному, разложение.

40.10 Ортогональные преобразования массивов

Обозначим через X_1, X_2, X_3 и X'_1, X'_2, X'_3 матрицы сечений массивов X и X' по осям i, j, k. Тогда легко проверяется, что

$$X' = X \odot_1 P \quad \Leftrightarrow \quad X'_1 = PX_1, \quad X' = X \odot_2 Q \quad \Leftrightarrow \quad X'_2 = QX_2, \quad X' = X \odot_3 R \quad \Leftrightarrow \quad X'_3 = RX_3.$$

Лемма. *Пусть матрицы P, Q, R ортогональные. Тогда если $X' = (X \odot_1 P) \odot_2 Q$, то скалярные произведения строк с одинаковыми номерами в матрицах X'_3 и X_3 одинаковы. Аналогично, если $X' = (X \odot_1 P) \odot_3 R$, то одинаковы скалярные произведения строк в матрицах X'_2 и X_2; если $X' = (X \odot_2 Q) \odot_3 R$, то одинаковы скалярные произведения строк в матрицах X'_1 и X_1.*

Доказательство. Пусть $X' = (X \odot_1 P) \odot_2 Q$. Это означает, что

$$x'_{i'j'k} = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} p_{i'i}\, q_{j'j}\; x_{ijk}.$$

Рассмотрим скалярные произведения строк матрицы X'_3 с номерами k_1 и k_2:

$$\sum_{i'} \sum_{j'} x'_{i'j'k_1}\; x'_{i'j'k_2} = \sum_{i'} \sum_{j'} \left(\sum_{i_1} \sum_{j_1} p_{i'i_1}\, q_{j'j_1}\; x_{i_1j_1k_1} \right) \left(\sum_{i_2} \sum_{j_2} p_{i'i_2}\, q_{j'j_2}\; x_{i_2j_2k_2} \right) =$$

$$\sum_{i_1} \sum_{j_1} \sum_{i_2} \sum_{j_2} \left(\sum_{i'} p_{i'i_1}\, p_{i'i_2} \right) \left(\sum_{j'} q_{j'j_1}\, q_{j'j_2} \right)\; x_{i_1j_1k_1}\; x_{i_2j_2k_2} =$$

$$\sum_{i_1} \sum_{j_1} \sum_{i_2} \sum_{j_2} \delta_{i_1i_2} \delta_{j_1j_2}\; x_{i_1j_1k_1}\; x_{i_2j_2k_2} = \sum_{i_1} \sum_{j_1} x_{i_1j_1k_1}\; x_{i_1j_1k_2}.$$

Здесь мы использовали так называемый символ Кронекера:

$$\delta_{\alpha\beta} = \begin{cases} 0, & \alpha \neq \beta, \\ 1, & \alpha = \beta. \end{cases}$$

Получено первое утверждение леммы. Остальные два утверждения устанавливаются аналогичным образом. $\square$

40.11 Разложение Таккера

Теорема. *Для любого трехмерного массива $X = [x_{ijk}]$ размеров $n_1 \times n_2 \times n_3$ существуют ортогональные матрицы P, Q, R такие, что трехмерный массив*

$$S = [s_{ijk}] \equiv X \odot \{P, Q, R\}$$

обладает следующими свойствами:

(1) *каждая из трех матриц сечений для S имеет попарно ортогональные строки*;

(2) $\sum\limits_{j,k} s^2_{1jk} \geq \sum\limits_{j,k} s^2_{2jk} \geq \ldots \geq \sum\limits_{j,k} s^2_{n_1jk}$;

(3) $\sum\limits_{i,k} s^2_{i1k} \geq \sum\limits_{i,k} s^2_{i2k} \geq \ldots \geq \sum\limits_{i,k} s^2_{in_2k}$;

(4) $\sum\limits_{i,j} s^2_{ij1} \geq \sum\limits_{i,j} s^2_{ij2} \geq \ldots \geq \sum\limits_{i,j} s^2_{ijn_3}$.

Доказательство. Обозначим через X_1, X_2, X_3 матрицы сечений массива X по осям i, j, k и рассмотрим их сингулярные разложения:

$$X_1 = P^\top \Sigma_1 V_1, \qquad X_2 = Q^\top \Sigma_2 V_2, \qquad X_3 = R^\top \Sigma_3 V_3,$$

где матрицы P, Q, R, V_1, V_2, V_3 ортогональные, а $\Sigma_1, \Sigma_2, \Sigma_3$ — диагональные прямоугольные матрицы, в которых сингулярные числа занумерованы по невозрастанию. Отсюда вытекает, что в каждой из преобразованных матриц сечений

$$X_1 \odot_1 P = \Sigma_1 V_1, \qquad X_2 \odot_2 Q = \Sigma_2 V_2, \qquad X_3 \odot_3 R = \Sigma_3 V_3$$

строки попарно ортогональны и расположены в порядке невозрастания их длин.

Далее, согласно доказанной выше лемме, скалярные произведения строк в матрице сечений по оси i для массива $S = X \odot \{P, Q, R\}$ те же самые, что и в матрице сечений по той же оси для массива $X \odot_1 P$. То же верно в отношении скалярных произведений строк для матриц сечений по оси j для массивов S и $X \odot_2 Q$, а также и для матриц сечений по оси k для массивов S и $X \odot_3 R$. Тем самым доказаны свойства (1)–(4). $\square$

Разложение $S = X \odot \{P, Q, R\}$ с указанными свойствами (1)–(4) называется *разложением Таккера*. Корни квадратные из сумм в (1)–(4) суть сингулярные числа матриц сечений массива X по осям i, j, k, соответственно.

Важное практическое значение разложения Таккера заключается в том, что оно дает надежную базу для построения приближений массива X суммами с малым числом членов с разделением индексов i, j, k: для этого достаточно заменить строки с относительно малыми длинами на нули. Полученная от такой замены погрешность легко оценивается.

В задачах о вычислении аппроксимаций малого тензорного ранга разложение Таккера часто используется, чтобы получить начальное приближение.

Заметим, что разложение Таккера может быть построено с помощью матричных методов вычисления сингулярного разложения. В принципе, аналогичные построения можно выполнить и на основе каких-либо других методов аппроксимации с понижением ранга, применяемых к матрицам сечений массива X.

Несмотря на то, что мы ограничились обсуждением трехмерных массивов, разложение Таккера легко переносится и на случай произвольных многомерных массивов. То же можно сказать и о других построениях данной лекции, в частности о факте единственности полилинейных аппроксимаций с точностью до эквивалентности.

Литература

1. Б. Л. ван дер Варден, *Алгебра*, М., Наука, 1976.
2. Э. Б. Винберг, *Курс алгебры*, М., Издательство "Факториал Пресс", 2002.
3. В. В. Воеводин, *Численные методы алгебры (теория и алгорифмы)*, М., Наука, 1966.
4. В. В. Воеводин, *Линейная алгебра*, М., Наука, 1980.
5. В. В. Воеводин, *Вычислительные основы линейной алгебры*, М., Наука, 1977.
6. В. В. Воеводин, Е. Е. Тыртышников, *Вычислительные процессы с теплицевыми матрицами*, М., Наука, 1987.
7. Ф. Р. Гантмахер, *Теория матриц*, М., Физматлит, 1967.
8. С. К. Годунов, *Современные аспекты линейной алгебры*, Новосибирск, Научная книга, 1997.
9. Дж. Голуб, Ч. Ван Лоун, *Матричные вычисления*, М., Мир, 1999.
10. Х. Д. Икрамов, *Задачник по линейной алгебре*, М., Наука, 1975.
11. Х. Д. Икрамов, *Численное решение матричных уравнений*, М., Наука, 1984.
12. Х. Д. Икрамов, *Численные методы для симметричных линейных систем*, М., Наука, 1988.
13. Х. Д. Икрамов, *Несимметричная проблема собственных значений*, М., Наука, 1991.
14. В. А. Ильин, Э. Г. Поздняк, *Аналитическая геометрия*, М., Наука, 1981.
15. В. А. Ильин, Г. Д. Ким, *Линейная алгебра и аналитическая геометрия*, Издательство МГУ, 1998.
16. В. Г. Карманов, *Математическое программирование*, М., Наука, 1975.
17. А. И. Кострикин, *Введение в алгебру*, М., Наука, 1977.
18. А. Г. Курош, *Курс высшей алгебры*, М., Наука,1971.
19. М. М. Постников, *Линейная алгебра и дифференциальная геометрия*, М., Наука, 1979.
20. М. М. Постников, *Основы теории Галуа*, М., Физматлит, 1960.
21. В. В. Прасолов, *Многочлены*, М., МЦНМО, 2001.
22. И. В. Проскуряков, *Сборник задач по линейной алгебре*, М., Наука, 1984.
23. Г. Стрэнг, *Линейная алгебра и ее применения*, М., Мир, 1980.
24. Е. Е. Тыртышников, *Теплицевы матрицы, некоторые их аналоги и приложения*, Отдел вычислительной математики АН СССР, М., 1989.
25. Е. Е. Тыртышников, *Краткий курс численного анализа*, М., ВИНИТИ, 1994.
26. Дж. Х. Уилкинсон, *Алгебраическая проблема собственных значений*, М. Физматлит, 1970.
27. Д. К. Фаддеев, *Лекции по алгебре*, М., Наука, 1984.

28. Д. К. Фаддеев, В. Н. Фаддеева, *Вычислительные методы линейной алгебры*, М.-Л., Физматлит, 1963.

29. Дж. Форсайт, М. Малькольм, К. Молер, *Машинные методы математических вычислений*, Мир, М., 1980.

30. П. Халмош, *Конечномерные векторные пространства*, М., Физматлит, 1963.

31. Р. Хорн, Ч. Джонсон, *Матричный анализ*, М., Мир, 1989.

32. R. Bhatia, *Matrix Analysis*, Springer–Verlag, New York, 1996.

33. G. W. Stewart, J. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.