Алгебра и геометрия (1 поток)

Лекция 1

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 $y_1, \dots, y_m$ выражаются через $x_1, \dots, x_n$ следующим образом:

$$\begin{cases} y_1 = a_{11}x_1 + \dots + a_{1n}x_n, \\ \dots \\ y_m = a_{m1}x_1 + \dots + a_{mn}x_n. \end{cases} \qquad (*)$$

Коэффициенты a_{ij} считаются заданными постоянными величинами. Соберем их в прямоугольную таблицу и обозначим ее буквой A, составим также таблицы-столбцы из величин $x_1, \dots, x_n$ и $y_1, \dots, y_m$:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \qquad x = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}, \qquad y = \begin{bmatrix} y_1 \\ \dots \\ y_m \end{bmatrix}.$$

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

Матрица A может умножаться справа не только на столбцевые матрицы. Естественным образом можно ввести также произведение $C = AB$ матриц размеров $m \times n$ и $n \times k$. При умножении A на B возникает матрица C размеров $m \times k$.

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

$$\begin{cases} y_1 = a_{11}x_1 + \dots + a_{1n}x_n, \\ \dots \\ y_m = a_{m1}x_1 + \dots + a_{mn}x_n, \end{cases} \qquad \begin{cases} x_1 = b_{11}z_1 + \dots + b_{1k}z_k, \\ \dots \\ x_n = b_{n1}z_1 + \dots + b_{nk}z_k. \end{cases}$$

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 $C = AB$.

Утверждение. $y = A(Bz) = (AB)z$.

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.

Задача 1. *Известно, что произведение матриц A и B существует и при этом сумма элементов в каждой строке матрицы B равна нулю. Докажите, что матрица AB обладает тем же свойством.*

1.3 Ассоциативность умножения матриц

Теорема. $(AB)C = A(BC)$.
Доказательство. Пусть A — $m \times n$, B — $n \times k$, C — $k \times l$. Тогда

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

Ассоциативность полезно учитывать при вычислениях. Пусть нужно найти произведение трех прямоугольных матриц размеров $1 \times n$, $n \times 1$ и $1 \times n$:

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

В данном случае есть два варианта расстановки скобок:

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

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

Варианты (1) и (2) приводят к двум разным алгоритмам вычисления матрицы A. Согласно 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).

Задача 2. *Известно, что для матриц A и B оба произведения AB и BA существуют. Докажите, что если $(AB)^k = 0$, то $(BA)^{k+1} = 0$.*

1.4 Некоммутативность умножения матриц

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

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

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 $c_{ij} = \alpha a_{ij}$.

Задача 3. *Квадратные матрицы A и B коммутируют, т.е. $AB = BA$. Докажите, что если $A^k = 0$ и $B^l = 0$ для каких-то натуральных чисел k и l, то $(A + B)^{k+l} = 0$. Покажите, что утверждение теряет силу, если $AB \neq BA$.*

Задача 4. *Известно, что A и B — вещественные матрицы порядка n и при этом $(A + \lambda B)^k = 0$ для некоторого натурального числа k и любого вещественного числа λ. Докажите, что $B^k = 0$.*

1.6 Умножение блочных матриц

Предположим, что матрицы $A = [A_{il}]$ и $B = [B_{lj}]$ составлены из матриц-блоков:

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

где A_{il} — $m_i \times n_l$, B_{lj} — $n_l \times k_j$. Тогда произведение $C = AB$ существует и его можно вычислять, используя операции умножения и сложения матриц-блоков (докажите!):

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

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 нужно найти 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}.

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?

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.

1.9 Метод Винограда

Попробуйте-ка перемножить матрицы как-либо иначе — не по классическому алгоритму. Вероятно, впервые это сделал Виноград (в начале 1960-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 $n^2 m = \frac{1}{2}n^3$ умножений и $3n^2 m = \frac{3}{2}n^3$ 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.

1.10 Метод Штрассена

В 1969 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 *трехмерных матриц* — таблиц для величин с тремя индексами. Об этом мы поговорим позже. А пока давайте посмотрим на изобретение Штрассена "без комментариев":

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

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

1.11 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 $O(n^{\log_2 7})$ операций. Поскольку

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

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 и, более того, вредно раскручивать до конца: $7\,n^{\log_2 7}>2n^3$ 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 $\mathcal{O}\left(n^\alpha\right)$, где $\alpha<2.42$. Никто не знает, каков минимальный показатель в таких оценках. Ясно лишь, что $\alpha\ge 2$.

1.12 Применение трехмерных матриц

Трехмерные матрицы — это таблицы, составленные из элементов, которые расположены в ячейках параллелепипеда и для указания на их место используются три индекса. Они имеют прямое отношение к получению быстрых алгоритмов умножения матриц.

В случае 2×2-матриц естественным образом возникает некоторая трехмерная матрица размеров $4\times 4\times 4$. Нас интересует равенство

$$\begin{bmatrix}c_1 & c_2\\ c_3 & c_4\end{bmatrix}=\begin{bmatrix}a_1 & a_2\\ a_3 & a_4\end{bmatrix}\begin{bmatrix}b_1 & b_2\\ b_3 & b_4\end{bmatrix}.$$

Используя правило умножения матриц, нетрудно понять, что имеют место соотношения

$$c_k=\sum_{i=1}^{4}\sum_{j=1}^{4}t_{ijk}a_ib_j,\quad 1\le k\le 4,$$

в которых числа t_{ijk} не зависят от элементов перемножаемых матриц и принимают значения 0 и 1. Пусть каким-то образом удалось найти разложение вида

$$t_{ijk}=\sum_{\alpha=1}^{r}u_{i\alpha}v_{j\alpha}w_{k\alpha},\quad 1\le i,j,k\le 4,$$

с целыми числами $u_{i\alpha}$, $v_{j\alpha}$, $w_{k\alpha}$. Тогда

$$c_k=\sum_{\alpha=1}^{r}w_{k\alpha}\left(\sum_{i=1}^{4}u_{i\alpha}a_i\right)\left(\sum_{j=1}^{4}v_{j\alpha}b_j\right),\quad 1\le k\le 4,$$

и у нас сразу же появляется алгоритм умножения 2×2-матриц с числом умножений r. Замечательно и то, что этот алгоритм можно применять также для умножения блочных 2×2-матриц, и в этому случае r будет числом умножений блоков. Фактически Штрассену удалось найти одно из таких разложений, в котором $r=7$. В действительности разложений такого типа много, и читатель имеет возможность придумать свой собственный алгоритм с 7-ю умножениями.

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 способа расстановки скобок*, и 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: сначала находим парные произведения A_nA_{n-1}, $A_{n-2}A_{n-3}$, $\ldots,$, A_2A_1, затем парные произведения полученных результатов, и так далее. Потребуется всего лишь $O(\log_2 n)$ параллельных шагов!

Задача 5. *Числа Фибоначчи f_1, f_2, f_3, $\ldots$ определяются условиями $f_1 = f_2 = 1$ и рекуррентным соотношением $f_n = f_{n-1} + f_{n-2}$ при $n \geq 3$. Докажите, что число f_n при $n = 2^k$ можно вычислить за $O(\log_2 n)$ арифметических операций.*

Задача 6. *Докажите, что значение многочлена $f_n(x) = 1 + x + \ldots + x^n$ при $n = 2^k$ в любой точке x можно найти за $O(\log_2 n)$ арифметических операций.*

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.

Алгебра и геометрия (1 поток)

Лекция 2

2.1 Система линейных алгебраических уравнений

Система линейных алгебраических уравнений — это система уравнений вида

$$\begin{cases} a_{11}x_1 & + \ldots + & a_{1n}x_n & = & b_1, \\ & & & \ldots & \\ a_{m1}x_1 & + \ldots + & a_{mn}x_n & = & b_m \end{cases} \tag{1}$$

относительно неизвестных величин $x_1, \ldots, x_n$. Если система имеет решение, то она называется *совместной*. С помощью матричных обозначений ее можно записать в виде

$$Ax = b, \qquad A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ \ldots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \ldots \\ b_m \end{bmatrix}.$$

Множество матриц размеров $m \times n$ с элементами $a_{ij} \in \mathbb{R}$, где $\mathbb{R}$ — множество вещественных чисел, обозначается $\mathbb{R}^{m\times n}$. Часто используются сокращенные обозначения $\mathbb{R}^m = \mathbb{R}^{m\times 1}$ и $\mathbb{R}^n = \mathbb{R}^{n\times 1}$, а матрицы-столбцы называются также *векторами*.

Матрица $A \in \mathbb{R}^{m\times n}$ называется *матрицей коэффициентов*, вектор $b \in \mathbb{R}^m$ — *правой частью*, а вектор $x \in \mathbb{R}^n$ — *решением* системы (1). Если $b = 0$, то система называется *однородной*.

2.2 Линейные комбинации

Для понимания сути дела полезна следующая интерпретация системы (1). Пусть $a_1, \ldots, a_n$ — столбцы матрицы A. Тогда соотношения (1) равносильны равенству между векторами

$$x_1a_1 + \ldots + x_na_n = b. \tag{2}$$

Вектор $x_1a_1 + \ldots + x_na_n$ называется *линейной комбинацией* векторов $a_1, \ldots, a_n$, а числа $x_1, \ldots, x_n$ — ее *коэффициентами*. Множество

$$L(a_1, \ldots, a_n) = \{\alpha_1a_1 + \ldots + \alpha_na_n : \alpha_1, \ldots, \alpha_n \in \mathbb{R}\}$$

всевозможных линейных комбинаций векторов $a_1, \ldots, a_n$ называется *линейной оболочкой* векторов $a_1, \ldots, a_n$. Таким образом, равенство (2) означает, что

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

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

2.3 Арифметические векторы

При изучении линейных комбинаций от матриц-столбцов естественно перейти к несколько более общим объектам. *Арифметическим вектором* x размера n называется упорядоченная последовательность n чисел, записанная, например, в виде $x = (x_1, \ldots, x_n)$. Эти числа могут быть элементами матрицы-столбца или матрицы-строки. Более того, они могут быть каким-то образом упорядоченными элементами прямоугольной матрицы.

Суммой арифметических векторов $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ называется арифметический вектор $x+y = (x_1+y_1, \ldots, x_n+y_n)$. Если α – число, то, по определению, $\alpha x = (\alpha x_1, \ldots, \alpha x_n)$. Для краткости арифметические векторы часто называются просто векторами. Вектор, все элементы которого равны нулю, называется *нулевым вектором* и обозначается символом 0.

2.4 Векторное пространство

Пусть V – некоторое непустое множество арифметических векторов размера n, замкнутое относительно операций сложения векторов и умножения вектора на число. Любое такое множество будем называть *векторным пространством*. Векторное пространство, содержащее только один вектор (и это, конечно, нулевой вектор) называется *нулевым* или *тривиальным*.

Очевидно, векторным пространством является линейная оболочка любой системы векторов. Мы скоро также увидим, что верно и обратное, то есть любое векторное пространство, составленное из арифметических векторов, можно представить как линейную оболочку некоторой системы векторов.

Векторные пространства имеют ключевое значение при описании строения множества всех решений совместной системы линейных уранения $Ax = b$.

Утверждение. *Множество всех решений однородной системы $Ax = 0$ представляет собой векторное пространство.*

Доказательство. Если $Ax = 0$ и $Ay = 0$, то $A(x+y) = Ax+Ay = 0$. Для произвольного числа α находим $A(\alpha x) = \alpha Ax = 0$. □

Теорема. *Пусть система $Ax = b$ совместна и z – какое-то ее решение. Тогда множество всех решений имеет вид $z+V := \{z+y : y \in V\}$, где V – векторное пространство всех решений соответствующей однородной системы $Ax = 0$.*

Доказательство. Если $Ay = 0$, то, очевидно, $A(z + y) = Az + Ay = b + 0 = b$. Далее, пусть x – произвольное решение системы $Ax = b$. Тогда $A(x - z) = Ax - Az = b - b = 0$, и значит, $x = z + y$, где $y = x - z \in V$. □

2.5 Линейная зависимость

Линейная комбинация векторов, все коэффициенты которой равны нулю, называется *тривиальной*, а в противном случае – *нетривиальной*. Под *системой векторов* обычно подразумевается упорядоченная совокупность конечного числа векторов. Система

векторов называется *линейно зависимой*, если для них существует нетривиальная линейная комбинация, равная нулевому вектору.

Лемма о линейной зависимости. *Пусть $a_1, \ldots, a_n$ — линейно зависимая система и a_1 – ненулевой вектор. Тогда $n \geq 2$ и для некоторого $2 \leq k \leq n$ вектор a_k является линейной комбинацией векторов $a_1, \ldots, a_{k-1}$.*

Доказательство. Рассмотрим нетривиальную линейную комбинацию

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = 0$$

и предположим, что $\alpha_{k+1} = \ldots = \alpha_n = 0$, но $\alpha_k \neq 0$. Ясно, что $k \geq 2$ (иначе находим $\alpha_1 a_1 = 0$ и, поскольку $a_1 \neq 0$, в итоге $\alpha_1 = 0$, и значит, комбинация является тривиальной). Далее,

$$\alpha_1 a_1 + \ldots + \alpha_k a_k = 0 \quad \Rightarrow \quad a_k = \left(-\frac{\alpha_1}{\alpha_k}\right) a_1 + \ldots + \left(-\frac{\alpha_{k-1}}{\alpha_k}\right) a_{k-1}. \ \square$$

Заметим также, что любая линейно зависимая система остается таковой при добавлении любого количества векторов.

2.6 Линейная независимость

Система векторов называется *линейно независимой*, если она не является линейно зависимой. Таким образом, если векторы $a_1, \ldots, a_n$ линейно независимы, то в этом и только этом случае

$$\alpha_1 a_1 + \ldots + \alpha_n a_n = 0 \quad \Rightarrow \quad \alpha_1 = \ldots = \alpha_n = 0.$$

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

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

$$b \ = \ \alpha_1 a_1 \ + \ \ldots, \ + \alpha_n a_n \ = \ \beta_1 a_1 \ + \ \ldots, \ + \beta_n a_n.$$

Отсюда

$$(\alpha_1 - \beta_1) a_1 \ + \ \ldots \ + \ (\alpha_n - \beta_n) a_n \ = \ 0 \quad \Rightarrow \quad \alpha_1 - \beta_1 = \ldots = \alpha_n - \beta_n = 0. \qquad \square$$

Задача 1. *Матрица размеров 4×3 имеет элементы $a_{ij} > 0$ при $i = j$ и $a_{ij} < 0$ при $i \neq j$. Докажите, что ее столбцы линейно независимы.*

2.7 Транзитивность линейной зависимости

Пусть имеются три системы векторов:

$$c_1, \ldots, c_k, \qquad b_1, \ldots, b_m, \qquad a_1, \ldots, a_n,$$

и предположим, что каждый вектор первой системы есть линейная комбинация векторов второй системы, а каждый вектор второй системы есть линейная комбинация векторов третьей системы. Тогда каждый вектор первой системы является также линейной комбинацией векторов третьей системы.

Это важное утверждение совершенно очевидно:

$$L(c_1, \ \ldots, \ c_k) \subseteq L(b_1, \ \ldots, \ b_m) \subseteq L(a_1, \ \ldots, \ a_n) \quad \Rightarrow \quad L(c_1, \ \ldots, \ c_k) \subseteq L(a_1, \ \ldots, \ a_n).$$

2.8 Монотонность числа линейно независимых векторов

Лемма о монотонности. *Пусть система векторов $b_1, \ldots, b_m$ линейно независима и каждый ее вектор является линейной комбинацией векторов системы $a_1, \ldots, a_n$. Тогда $m \le n$.*

Доказательство. Система $b_1, a_1, \ldots, a_n$ линейно зависима и заведомо $b_1 \neq 0$. Согласно лемме о линейной зависимости, в этой системе имеется вектор, являющийся линейной комбинацией предыдущих векторов. Для определенности пусть это будет вектор a_n. Ясно, что (транзитивность линейно зависимости)

$$L(b_2, \ldots, b_m) \subseteq L(a_1, \ldots, a_n) \subseteq L(b_1, a_1, \ldots, a_{n-1}).$$

Поэтому условия леммы выполнены также для систем $b_2, \ldots, b_m$ и $b_1, a_1, \ldots, a_{n-1}$.

Теперь мы рассмотрим систему $b_2, b_1, a_1, \ldots, a_{n-1}$ и можем утверждать, что один из ее векторов является линейной комбинацией предыдущих векторов. Понятно, что это может быть только какой-то из векторов $a_1, \ldots, a_{n-1}$. Для определенности пусть это будет вектор a_{n-1}. Две новых системы $b_3, \ldots, b_m$ и $b_2, b_1, a_1, \ldots, a_{n-2}$ удовлетворяют условиям леммы, и мы можем сделать еще один шаг такого же типа.

Если $m > n$, то на n-м шаге мы получим системы $b_{n+1}, \ldots, b_m$ и $b_n, \ldots, b_1$, удовлетворяющие условиям леммы, а это противоречит условию линейной независимости системы векторов $b_1, \ldots, b_m$. $\square$

Следствие. *Число векторов в максимальной линейно независимой подсистеме векторов $a_1, \ldots, a_n$ не меньше m.*

2.9 Размерность векторного пространства

Размерность векторного пространства V определяется как максимальное число его линейно независимых векторов и обозначается $\dim V$. Размерность нулевого пространства равна 0, так как в нем линейно независимые системы отсутствуют.

Таким образом, в ненулевом пространстве V имеется линейно независимая система с числом векторов $d = \dim V$, а любая система, в которой больше чем d векторов, должна быть линейно зависимой.

Утверждение. *Пусть векторное пространство V является линейной оболочкой каких-то n векторов. Тогда* $\dim V \le n$.

Доказательство. Пусть $b_1, \ldots, b_m$ – произвольная линейно независимая система в линейно оболочке векторов $a_1, \ldots, a_n$. Согласно лемме о монотонности, $m \le n$. $\square$

2.10 Базис и размерность

Базисом векторного пространства V называется любая его линейно независимая система, линейная оболочка которой совпадает с V.

Теорема о базисах. *Все базисы векторного пространства содержат одно и то же число векторов, равное его размерности.*

Доказательство. Пусть $b_1, \ldots, b_m$ и $a_1, \ldots, a_n$ — два базиса. Тогда, в силу леммы о монотонности, получаем сразу два неравенства: $m \le n$ и $n \le m \;\Rightarrow\; m = n$.

Если $a_1, \ldots, a_n$ – максимальная линейно независимая система, то для любого вектора v система $a_1, \ldots, a_n, v$ будет линейно зависимой. Поэтому, согласно лемме о линейной зависимости, какой-то вектор в ней линейно выражается через предыдущие и очевидно, что таковым будет именно вектор v. Значит, максимальная линейно независимая система является базисом. $\square$

Утверждение. *В качестве базиса в линейной оболочке $V = L(a_1, \ldots, a_n)$ всегда можно выбрать максимальную линейно независимую подсистему векторов $a_1, \ldots, a_n$.*

Доказательство. Не ограничивая общности, будем считать, что такую подсистему образуют векторы $a_1, \ldots, a_d$. Тогда любая система $a_1, \ldots, a_d, a_k$ при $d+1 \le k \le n$ будет линейно зависимой и, по лемме о линейной зависимости, вектор a_k является линейной комбинацией векторов $a_1, \ldots, a_d$. В силу транзитивности линейной зависимости, любой вектор пространства V будет принадлежать линейной оболочке данной подсистемы. Значит, подсистема является базисом пространства V. $\square$

Задача 2. *Пусть V – произвольное векторное пространство, вложенное в векторное пространство W. Докажите, что если $\dim V < \dim W$, то в пространстве W существует базис, не содержащий ни одного вектора из V.*

Задача 3. *Пусть A – матрица порядка n. Докажите, что если $A^{n+1} = 0$, то $A^n = 0$.*

2.11 Дополнение до базиса

Лемма о дополнении до базиса. *Пусть пространство V является линейной оболочкой векторов $v_1, \ldots, v_n$ Тогда любая линейно независимая система его векторов, не являющаяся базисом, может быть дополнена до базиса какими-ми то из векторов $v_1, \ldots, v_n$.*

Доказательство. Пусть $a_1, \ldots a_k$ – линейно независимая система в пространстве V. Если $V = L(a_1, \ldots, a_k)$, то система является базисом. В противном случае имеется вектор, скажем v_{i_1}, добавление которого к данной системе сохраняет линейную независиомсть. Если $V = L(a_1, \ldots, a_k, v_{i_1})$, то все доказано. Если нет, то продолжаем в том же духе. Поскольку $\dim V \le n$, до получения базиса мы будем добавлять не более $n - k$ векторов. $\square$

Задача 4. *Система $a_1, \ldots, a_k$ линейно независима и b — ненулевой вектор. Докажите, что после замены некоторого a_i на b система останется линейно независимой.*

2.12 Существование базиса

В нулевом пространстве базиса нет. Существование базиса в ненулевом векторном пространстве V, составленном из арифметических векторов размера n, вытекает из следующего наблюдения.

Утверждение. *Пространство V является линейной оболочкой какой-то линейно независимой системы своих векторов.*

Доказательство. Пусть a_1 – произвольный ненулевой вектор. Если $V = L(a_1)$, то все доказано. В противном случае имеется вектор $a_2 \in V$, дающий нам линейно независимую систему a_1, a_2. Если $V = L(a_1, a_2)$, то все доказано. Предположим, что уже

построена линейно независимая система $a_1, \ldots, a_k \in V$. Если $V = L(a_1, \ldots, a_k)$, то данная система является базисом пространства V. В противном случае ее можно дополнить до большей линейно независимой системы. Этот процесс не может продолжаться бесконечно, так как V имеет конечную размерность. $\square$

2.13 Матрицы с линейно независимыми столбцами

Пусть A — матрица размеров $m \times n$ и $a_1, \ldots, a_n$ — ее столбцы.

Утверждение. *Предположим, что столбцы $a_1, \ldots, a_n$ линейно независимы и $m = n$. Тогда система $Ax = b$ имеет единственное решение для любой правой части b.*

Доказательство. Очевидно, $a_1, \ldots, a_n \in V = L(e_1, \ldots, e_n)$. В силу линейной независимости векторов $a_1, \ldots, a_n$ находим

$$n = \dim L(a_1, \ldots, a_n) = \dim V \quad \Rightarrow \quad L(a_1, \ldots, a_n) = V.$$

Единственность вытекает из леммы о единственности разложения по линейно независимой системе. $\square$

Задача 5. *Любые k столбцов матрицы A линейно независимы. Докажите, что решение системы $Ax = b$, в котором число ненулевых элементов вектора x меньше $k/2$, определено однозначно.*

2.14 Обратимые матрицы

Матрица A порядка n называется *обратимой*, если существует матрица B такая, что $AB = BA = I$, где I – единичная матрица порядка n. Матрица B определяется однозначно ($AC = CA = I$, $AB = BA = I \;\Rightarrow\; B = BAC = C$), называется *обратной матрицей* к матрице A и обозначается A^{-1}.

Очевидно, матрица A^{-1} будет обратимой. Произведение обратимых матриц является обратимой матрицей: $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$. Подмножество всех обратимых матриц в пространстве $\mathbb{R}^{n\times n}$ называется *полной линейной группой* вещественных матриц порядка n.

Утверждение. *Для обратимости квадратной матрицы необходима и достаточна линейная независимость ее столбцов.*

Доказательство. Необходимость: $BA = I$, $Ax = 0 \;\Rightarrow\; 0 = B(Ax) = (BA)x = I\,x = x$. Достаточность: согласно утверждению, полученному в предыдущем разделе, система $Ax = b$ имеет решение для любой правой части, а значит, и для каждого столбца матрицы $I \quad \Rightarrow \quad$ для некоторой матрицы B имеем $AB = I$. Остается лишь понять, почему $X := BA = I$. В силу ассоциативности $XB = B(AB) = B\,I = B$. Поскольку столбцы матрицы B тоже линейно независимы (проверьте!), для некоторой матрицы C находим $BC = I$. Следовательно, $(XB)C = X(BC) = X = BC = I$. $\square$

Задача 6. *Докажите, что из линейной независимости столбцов квадратной матрицы вытекает линейная независимость ее строк.*

Задача 7. *В квадратной матрице элементы главной диагонали равны 2, элементы соседних диагоналей равны -1, а все остальные элементы равны 0. Докажите, что матрица является обратимой.*

Задача 8. *Докажите, что обратимая трехдиагональная матрица останется обратимой, если каждый поддиагональный элемент поделить, а каждый наддиагональный элемент умножить на одно и то же число.*

Задача 9. *Докажите, что после любой перестановки строк обратимая матрица остается обратимой, а ее обратная матрица получается из исходной обратной матрицы точно такой же перестановкой столбцов.*

Задача 10. *Матрица обратима и при этом все ее элементы и все элементы обратной матрицы неотрицательны. Докажите, что перестановкой строк данная матрица приводится к диагональному виду.*

Задача 11. *Докажите, что для любых квадратных матриц A и B одного и того же порядка имеет место равенство*

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

Задача 12. *Пусть для любого n имеется алгоритм обращения произвольной обратимой матрицы порядка n не более чем за cn^α арифметических операций, где α и c не зависят от n. Докажите, что в этом случае существует алгоритм умножения двух матриц порядка n с числом операций не более $c_1 n^\alpha$, где c_1 не зависит от n.*

2.15 Матрицы с диагональным преобладанием

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

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

В таких случаях говорят, что матрица A имеет *диагональное преобладание по строкам*, а транспонированная к ней матрица — *диагональное преобладание по столбцам*.

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

Доказательство. Пусть матрица A имеет диагональное преобладание по строкам. От противного, допустим, что $Ax = 0$ и x_i – по модулю самый большой элемент ненулевого вектора x. Тогда из равенства $(Ax)_i = 0$ следует

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

Задача 13. *Матрица A обратима, а B — произвольная матрица такого же порядка. Докажите, что матрицы $A + \varepsilon B$ обратимы при всех вещественных ε, достаточно малых по модулю.*

Задача 14. *Докажите, что любую квадратную вещественную матрицу можно представить в виде суммы двух обратимых матриц.*

Задача 15. *Матрица имеет диагональное преобладание по строкам, все элементы ее главной диагонали положительны, а все остальные элементы отрицательны. Докажите, что матрица обратима и все элементы обратной матрицы положительны.*

2.16 Матрицы с линейно зависимыми столбцами

Утверждение. *Если $m < n$, то столбцы $m \times n$-матрицы линейно зависимы.*

Доказательство. Пусть $e_1, \ldots, e_m$ – столбцы единичной матрицы порядка m. Тогда каждый из столбцов нашей матрицы принадлежит их линейной оболочке, и, как следствие леммы о монотонности, размерность линейной оболочки столбцов не превосходит m. $\square$

Данное утверждение в некоторых учебниках используется непосредственно для доказательства леммы о монотонности. В таких случаях, конечно, нужно предъявлять доказательство, не опирающееся на лемму о монотонности. Его нетрудно получить, например, индукцией по m (попробуйте!).

Наше доказательство леммы о монотонности (восходящее, по-видимому, к Штейницу) замечательно тем, что оно практически без изменений работает при получении аналогичного результата, связанного с более общим понятием *алгебраической зависимости*.

Чтобы показать суть упомянутого выше обобщения, рассмотрим здесь некоторый частный случай понятия алгебраической зависимости. Числа $a_1, \ldots, a_n$ называются *алгебраически зависимыми*, если они удовлетворяют уравнению $f(a_1, \ldots, a_n) = 0$, где f является нетривиальным (не равным нулю тождественно) многочленом от n переменных с рациональными коэффициентами. В противном случае числа называются *алгебраически независимыми*. Говорят, что число *алгебраически зависит* от чисел $a_1, \ldots, a_n$, если оно является корнем нетривиального многочлена от одной переменной с коэффициентами, записанными как значения каких-то многочленов от $a_1, \ldots, a_n$ с рациональными коэффициентами. Если число не является корнем нетривиального многочлена от одной переменной с рациональными коэффициентами (система, состоящая из одного этого числа, является алгебраически независимой), то оно называется *трансцендентным*.

Аналог леммы о линейной зависимости в данном случае выглядит таким образом: если числа $a_1, \ldots, a_n$ алгебраически зависимы, а число a_1 трансцендентно, то для некоторого $2 \leq k \leq n$ число a_k алгебраически зависит от чисел $a_1, \ldots, a_{k-1}$. Аналог леммы о монотонности будет таким: если числа $a_1, \ldots, a_m$ алгебраически независимы и каждое из них алгебраически зависит от чисел $b_1, \ldots, b_n$, то $m \leq n$.

Алгебра и геометрия (1 поток)

Лекция 3

3.1 Билинейные функции

Пусть V – векторное пространство, составленное из n-координатных арифметических векторов. Функция $f(x, y)$, определенная при всех $x, y \in V$, называется *билинейной*, если она линейна по каждому аргументу при любом фиксированном значении другого аргумента:

$$f(\alpha u + \beta v,\, y) = \alpha f(u,\, y) + \beta f(v,\, y),$$

$$f(x,\, \alpha u + \beta v) = \alpha f(x, u) + \beta f(x,\, v)$$

для любых чисел α, β и векторов $u, v, x, y \in V$.

Билинейные функции с числовыми значениями называются также *билинейными формами*. Обозначим через e_i арифметический вектор, отличающийся от нулевого только единицей в i-й позиции. Тогда

$$x = (x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i e_i, \quad y = (y_1, \ldots, y_n) = \sum_{j=1}^{n} y_j e_j \quad \Rightarrow$$

$$f(x,\, y) = f\left(\sum_{i=1}^{n} x_i e_i,\, \sum_{j=1}^{n} y_j e_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j f(e_i,\, e_j) = \begin{bmatrix} x_1 & \ldots x_n \end{bmatrix} A \begin{bmatrix} y_1 \\ \ldots \\ y_n \end{bmatrix},$$

где $A = [f(e_i,\, e_j)]$ – матрица порядка n, которую принято называть *матрицей билинейной формы* $f(x,\, y)$.

Задача 1. *Билинейная форма $f(x, y) = x^\top A y$ называется* `невырожденной`*, если для любого вектора x существует вектор y такой, что $f(x, y) \neq 0$. Докажите, что для невырожденности билинейной формы необходима и достаточна обратимость ее матрицы A.*

3.2 Полилинейные функции

Функция $f(x^{(1)},\, x^{(2)},\, \ldots\, ,\, x^{(d)})$, определенная на всех векторах $x^{(1)},\, x^{(2)},\, \ldots,\, x^{(d)}$ из пространства V арифметических векторов размера n, называется *полилинейной*, если она линейна по каждому аргументу при фиксации остальных аргументов. Полилинейные функции с числовыми значениями обычно называются *полилинейными формами*. По аналогии с билинейными формами, можно ввести величины

$$a_{i_1,\ldots,i_d} = f(e_{i_1},\, \ldots\, ,\, e_{i_d}), \quad 1 \leq i_1, \ldots, i_d \leq n,$$

и заметить, что (докажите!)

$$f(x^{(1)},\dots,x^{(d)}) \;=\; \sum_{i_1=1}^{n}\dots\sum_{i_d=1}^{n} x^{(1)}_{i_1}\dots x^{(d)}_{i_d}\, a_{i_1,\dots,i_d}.$$

Величины $a_{i_1,\dots,i_d}$ можно рассматривать как элементы d-мерной таблицы. Такие таблицы называют также *d-мерными матрицами* или *тензорами* — в нашем случае размеров $n\times\dots\times n$.

3.3 Обнуление и знакопеременность

Функция $f(x^{(1)},\dots,x^{(d)})$ называется *знакопеременной*, если она меняет знак при перестановке любой пары разных аргументов.

Утверждение. *Пусть полилинейная форма $f(x^{(1)},\dots,x^{(d)})$ обнуляется каждый раз, когда два разных аргумента принимают одно и то же значение. Тогда она является знакопеременной.*

Доказательство. Достаточно рассмотреть случай билинейной формы $f(x,y)$:

$$0 = f(x+y,x+y) = f(x,x)+f(x,y)+f(y,x)+f(y,y) = f(x,y)+f(y,x)$$

$$\Rightarrow\quad f(y,x) = -f(x,y). \quad \square$$

3.4 Перестановка аргументов

Что происходит со знакопеременной функцией при *произвольной* перестановке аргументов? Если переставить только два аргумента, то, согласно определению, меняется знак.

В общем случае от естественной системы номеров $(1,2,\dots,d)$ можно перейти к произвольной их перестановке $(i_1,i_2,\dots,i_d)$, последовательно переставляя только пары номеров – такие перестановки принято называть *транспозициями*. Например, с помощью перестановок пар соседних номеров можно сначала поставить i_1 на первое место, затем i_2 на второе место и так далее. Если p – общее число таких транспозиций, то должно выполняться равенство

$$f(x^{(i_1)},\dots,x^{(i_d)}) \;=\; (-1)^p\, f(x_1,x_2,\dots,x_d).$$

Здесь, заметим, возникает вопрос вообще о *существовании* знакопеременной функции. От набора $(1,2,\dots,d)$ к перестановке $(i_1,i_2,\dots,i_d)$ можно перейти с помощью многих различных последовательностей транспозиций. Поэтому нам нужно понять, почему *в каждой такой последовательности число транспозиций имеет одну и ту же четность*.

3.5 Четность подстановки

Взаимно-однозначное отображение $\sigma:\{1,\dots,n\}\to\{1,\dots,n\}$ называется *подстановкой степени n* или *перестановкой* и задается таблицей вида

$$\sigma = \begin{pmatrix} i_1 & i_2 & \dots & i_n \\ \sigma(i_1) & \sigma(i_2) & \dots & \sigma(i_n) \end{pmatrix}, \quad \{i_1,\dots,i_n\} = \{\sigma(i_1),\dots,\sigma(i_n)\} = \{1,2,\dots,n\}.$$

Множество всех подстановок степени n обозначается через S_n. На этом множестве естественным образом определена операции композиции (последовательного выполнения) подстановок. Эту операцию обычно называют *умножением*, а ее результат — *произведением* подстановок: если $a, b \in S_n$, то произведение $c = ab$ определяется правилом $c(i) := a(b(i))$. Подстановка называется *транспозицией* номеров $k \neq l$, если

$$\sigma(k) = l, \quad \sigma(l) = k \quad \text{и} \quad \sigma(i) = i \ \text{ при } \ i \notin \{k, l\}.$$

Для обозначения транспозиции будем использовать запись $\sigma = (k, l)$.

Утверждение. *Любая подстановка разлагается в произведение транспозиций.*

Доказательство. Проведем индукцию по числу переставляемых номеров. Пусть $\sigma(1) = k$. Тогда для подстановки $f = (k, 1)$ находим $f(1) = \sigma(1) = k$. Следовательно, подстановка $g := \sigma f^{-1}$ обладает свойством $g(k) = k$ и может рассматриваться как подстановка на меньшем множестве номеров. $\square$

Пусть $i < j$, но $\sigma(i) > \sigma(j)$. В таких случаях пару (i, j) называют *инверсией* подстановки σ. Число инверсий подстановки σ будем обозначать через $\delta(\sigma)$.

Теорема о четности. *Пусть подстановка σ является произведением p транспозиций. Тогда $(-1)^p = (-1)^{\delta(\sigma)}$.*

Доказательство. Пусть f – произвольная транспозиция и $\tau = \sigma f$. Достаточно показать, что разность $\delta(\tau) - \delta(\sigma)$ является числом нечетным. Пусть $f = (k, k+1)$. Если пара $(k, k+1)$ была инверсией для σ, то для τ она уже не будет инверсией, а если не была, то будет. Любые другие пары при переходе от σ к τ сохранят свойство быть или не быть инверсией. В случае $f = (k, k + s)$ при $s \geq 2$ нужно заметить, что f представляется произведением нечетного числа транспозиций соседних номеров:

$$f = \underbrace{(k, k+1) \dots (k+s-2, k+s-1)}_{s-1} \underbrace{(k+s, k+s-1) \dots (k+2, k+1)(k+1, k)}_{s}. \quad \square$$

Под *четностью подстановки* понимается четность ее числа инверсий. Согласно доказанной нами теореме, четность числа транспозиций в любом разложении данной подстановки в произведение транспозиций совпадает с четностью подстановки.

3.6 Транспозиции и циклы

Транспозиции относятся к специальному классу подстановок, называемых *циклами*. Под *циклом длины k* понимается подстановка σ, оставляющая на месте все номера, кроме выделенных k номеров, скажем $i_1, \dots, i_k$, и действующая на этих номерах циклически:

$$\sigma : \ i_1 \ \to \ i_2 \ \to \ \dots \ \to \ i_{k-1} \ \to \ i_k \ \to \ i_1.$$

Для обозначения цикла используется запись $\sigma = (i_1, i_2, \dots, i_k)$.

Таким образом, транспозиция — это цикл длины 2. Цикл длины k можно представить произведением $k - 1$ транспозиции: $(i_1, i_2, \dots, i_k) = (i_1, i_2)(i_2, i_3) \dots (i_{k-1}, i_k)$. Четность цикла длины k равна четности числа $k-1$. В частности, любой *тройной цикл* (цикл длины 3) является четным.

Любую подстановку можно разложить в произведение *независимых* циклов — так называются циклы, в которых множества переставляемых номеров не пересекаются.

Достаточно построить цикл, содержащий номер 1, затем цикл с номером, который не переставляется первым циклом, и так далее. Например,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 5 & 7 & 1 & 2 & 2 & 4 \end{pmatrix} \;=\; (1,3,7,4)\,(2,5).$$

Можно поставить также вопрос о минимально возможном числе транспозиций в разложениях заданной подстановки. Следующее утверждение показывает, что для цикла длины k оно равно $k-1$.

Утверждение. *Если произведение k транспозиций является циклом, то длина этого цикла не больше $k+1$.*

Доказательство. Проводим индукцию по k. Пусть цикл $\sigma = \pi_1 \dots \pi_{k-1}\pi_k = \tau\pi_k$ разложен в произведение минимально возможного числа транспозиций $\pi_1, \dots, \pi_k$. Тогда его длина удовлетворяет неравенству $|\sigma| \geq k+1$ (почему?), а для подстановки $\tau = \pi_1 \dots \pi_{k-1}$ есть две возможности:

- τ является циклом $\;\Rightarrow\;$ в силу индуктивного предположения $|\tau| \leq k$ и в силу минимальности числа k находим $|\tau| \geq k \;\Rightarrow\; |\tau| = k$. Если при умножении цикла на транспозицию получается цикл, то его длина не больше чем на единицу отличается от длины исходного цикла $\;\Rightarrow\; |\sigma| = |\tau\pi_k| \leq k+1$.
- $\tau = \tau_1\tau_2$ есть произведение двух независимых циклов, из которых каждый содержит по одному номеру из транспозиции π_k. Заметим, что каждый из этих циклов представляется произведением каких-то транспозиций из $\pi_1, \dots, \pi_{k-1}$. Согласно индуктивному предположению,

$$(|\tau_1|-1) + (|\tau_2|-1) \leq k-1 \quad\Rightarrow\quad |\sigma| = |\tau_1| + |\tau_2| \leq k+1. \;\square$$

Задача 2. *Докажите, что в множестве S_n при $n \geq 4$ любое произведение пары независимых транспозиций можно представить произведением тройных циклов.*

Задача 3. *Докажите, что произведение двух независимых циклов длины k_1 и k_2 нельзя разложить в произведение транспозиций, число которых меньше $k_1 + k_2 - 2$.*

3.7 Определитель

Будем рассматривать функции $f(a_1, \dots, a_n)$, определенные на арифметических векторах $a_1 = (a_{11}, \dots, a_{n1}), \quad \dots\,, \quad a_n = (a_{1n}, \dots, a_{nn})$ и принимающие числовые значения. Можно считать, что эти векторы соответствуют столбцам $n \times n$-матрицы A и для краткости писать просто $f(A)$. Как и раньше, e_i — это специальный арифметический вектор, отличающийся от нулевого только единицей в i-й позиции. Тогда $f(e_1, \dots, e_n) = f(I)$, где I — единичная матрица порядка n.

Основная теорема об определителе. *Полилинейная форма $f(A) = f(a_1, \dots, a_n)$, обнуляющаяся при совпадении значений любой пары разных аргументов, однозначно определяется своим значением $f(I)$ и имеет вид $f(A) = f(I)\det(A)$, где*

$$\det(A) \;=\; \sum_{\sigma \in S_n} A_\sigma\,(-1)^{\delta(\sigma)}, \quad A_\sigma = a_{\sigma(1),1} \dots a_{\sigma(n),n}.$$

Доказательство. Используя разложения

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

и полилинейность, находим

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

После исключения из суммирования заведомых нулей, возникающих при совпадении аргументов, приходим к равенству

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

В последнем равенстве учитывается знакопеременность, вытекающая из свойства обнуления в случае пары одинаковых аргументов, и применяется теорема о четности. □

Функция $\det(A)$ называется *определителем* или *детерминантом* матрицы A. Она задается суммой $n!$ слагаемых вида $A_\sigma(-1)^{\delta(\sigma)}$, называемых *членами определителя*. Каждый член представляет собой взятое с определенным знаком произведение n элементов матрицы, причем никакие два элемента одного члена не принадлежат одной строке или одному столбцу матрицы A. Часто применяются также обозначения

$$\det(A)=|A|=\begin{vmatrix}a_{11} & \ldots & a_{1n}\\ \ldots & \ldots & \ldots\\ a_{n1} & \ldots & a_{nn}\end{vmatrix}.$$

Можно говорить о столбцах или строках определителя, опуская слово “матрица” (тем более, что определители возникли в исследованиях Крамера задолго до появления термина “матрица”).

Задача 4. *Докажите, что если матрица $A=[a_{ij}]$ верхняя треугольная ($a_{ij}=0$ при $i>j$), то $|A|=a_{11}\ldots a_{nn}$.*

3.8 Частные случаи

При $n=2$ и $n=3$ получаются такие формулы:

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

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

3.9 Определитель транспонированной матрицы

Матрица $B=[b_{ji}]$ размеров $n\times m$ называется *транспонированной* к матрице $A=[a_{ij}]$ размеров $m\times n$, если $b_{ji}=a_{ij}$, и обозначается $B=A^\top$. При транспонировании столбцы становятся строками, а строки – столбцами. Квадратная матрица A называется *симметричной*, если $A=A^\top$. Что происходит с определителем квадратной матрицы при транспонировании?

Утверждение. $\det(A)=\det(A^\top)$.

Доказательство.

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

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

В данной выкладке используется равенство $\delta(\sigma) = \delta(\sigma^{-1})$ (проверьте!) и взаимная однозначность отображения $\sigma \to \sigma^{-1}$. $\square$

3.10 Характеристическое свойство определителя

Из основной теоремы об определителе, на первый взгляд, должно вытекать, что он является полилинейной функцией своих столбцов и обнуляется при совпадении пары столбцов — но только *при условии, что уже доказано*, что рассматриваемая в этой теореме полилинейная функция со свойством обнуления существует! Следующее утверждение по существу является *теоремой о существовании* такой функции.

Теорема. *Определитель $|A|$ является полилинейной функцией столбцов (строк) матрицы A, которая обнуляется на любой матрице с парой одинаковых столбцов (строк).*

Доказательство. Чтобы установить линейность определителя $|A|$ по j-му столбцу, рассмотрим матрицы B и C, отличающиеся от A только j-м столбцом, и предположим, что $a_j = \alpha b_j + \beta c_j$. Тогда

$$|A| = \sum_{\sigma\in S_n} a_{\sigma(1),1}\ldots(\alpha b_{\sigma(j),j} + \beta c_{\sigma(j),j})\ldots a_{\sigma(n),n}(-1)^{\delta(\sigma)} =$$

$$= \alpha\sum_{\sigma\in S_n} b_{\sigma(1),1}\ldots b_{\sigma(j),j}\ldots b_{\sigma(n),n}(-1)^{\delta(\sigma)} + \beta\sum_{\sigma\in S_n} c_{\sigma(1),1}\ldots c_{\sigma(j),j}\ldots c_{\sigma(n),n}(-1)^{\delta(\sigma)}$$

$$= \alpha|B| + \beta|C|.$$

Теперь предположим, что $a_i = a_j$ для каких-то заданных номеров $i < j$. Фиксируем транспозицию $\tau = (i,\, j)$ и заметим, что все множество подстановок степени n есть объединение множества A_n четных подстановок и множества $A_n\tau = \{\sigma\tau : \sigma \in A_n\}$ нечетных подстановок. Рассмотрим произведения $A_\sigma = a_{\sigma(1),1}\ldots a_{\sigma(n),n}$, $A_{\sigma\tau} = a_{(\sigma\tau)(1),1}\ldots a_{(\sigma\tau)(n),n}$ и заметим, что (в силу равенства $a_i = a_j$) $a_{(\sigma\tau)(i),i} = a_{\sigma(j),i} = a_{\sigma(j),j}$, $a_{(\sigma\tau)(j),j} = a_{\sigma(i),j} = a_{\sigma(i),i}$ $\Rightarrow$ $a_{(\sigma\tau)(i),i}a_{(\sigma\tau)(j),j} = a_{\sigma(i),i}a_{\sigma(j),j}$ $\Rightarrow$ $A_\sigma = A_{\sigma\tau}$. Следовательно,

$$|A| = \sum_{\sigma\in A_n} A_\sigma(-1)^{\delta(\sigma)} + \sum_{\sigma\in A_n} A_{\sigma\tau}(-1)^{\delta(\sigma\tau)} = \sum_{\sigma\in A_n} A_\sigma(-1)^{\delta(\sigma)} - \sum_{\sigma\in A_n} A_\sigma(-1)^{\delta(\sigma)} = 0.$$

Аналогичные утверждения для строк вытекают из равенства $|A| = |A^\top|$. $\square$

Следствие. *Определитель меняет знак при перестановке любой пары столбцов (строк) и сохраняется при прибавлении к столбцу (строке) произвольной линейной комбинации остальных столбцов (строк).*

Доказательство. Мы уже знаем, что знакопеременность полилинейной функции вытекает из свойства обнуления при совпадении значений двух разных аргументов. Вторая часть утверждения есть прямое следствие полилинейности и свойства обнуления. $\square$

Задача 5. *Пусть $u, v \in \mathbb{R}^n$ и I — единичная матрица. Докажите, что* $\det(I + uv^\top) = 1 + v^\top u$.

3.11 Определитель произведения матриц

Теорема. $|AB| = |A|\,|B|$.

Доказательство. Функция $f(B) := |AB|$ является полилинейной функцией столбцов матрицы B и обнуляется при совпадении любой пары столбцов (проверьте!). Поэтому, согласно основной теореме об определителе, она полностью определяется своим значением на единичной матрице:

$$f(B) = |AB| = f(I)|B| = |A||B|. \quad \square$$

Задача 6. *Докажите, что определитель трехдиагональной матрицы не изменится, если каждый поддиагональный элемент поделить, а каждый наддиагональный элемент умножить на одно и то же число.*

3.12 Миноры и их алгебраические дополнения

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. В связи с определителями нас будут интересовать $k \times k$-подматрицы $n \times n$-матрицы A, их определители называются *минорами* порядка k.

Если подматрица порядка k расположена на строках с номерами $i_1 < \ldots < i_k$ и столбцах с номерами $j_1 < \ldots < j_k$, то она обозначается $A(I, J)$, где $I = (i_1, \ldots, i_k)$ и $J = (j_1, \ldots, j_k)$. Через $A'(I, J)$ будем обозначать подматрицу порядка $n - k$, которая получается после вычеркивания столбцов и строк с указанными номерами. Такая подматрица называется *дополнительной* к подматрице $A(I, J)$, а ее определитель – *дополнительным минором* к минору $|A(I, J)|$. Величина

$$(-1)^{c(I,J)}\,|A'(I, J)|,$$

где $c(I, J)$ – сумма всех номеров систем I и J, называется *алгебраическим дополнением* минора $|A(I, J)|$.

3.13 Теорема Лапласа

Теорема Лапласа. *Пусть в матрице порядка n выбраны k строк (столбцов). Тогда ее определитель равен сумме всех расположенных на выбранных строках (столбцах) миноров, умноженных на свои алгебраические дополнения.*

Доказательство. Пусть фиксирована система строчных номеров I и каждый минор на выбранных строках определяется системой столбцовых номеров J. Тогда произведение каждого минора на его алгебраическое дополнение представляет собой некоторую сумму членов определителя заданной матрицы и каждый член определителя входит в одно и только одно такое произведение.

Чтобы это увидеть, сначала рассмотрим систему номеров $N = \{1, \ldots, k\}$. Тогда алгебраическое дополнение минора $|A(N, N)|$ совпадает с дополнительным минором, а то, что произведение $|A(N, N)||A'(N, N)|$ есть сумма некоторых членов определителя $|A|$, в данном случае проверяется совсем легко.

В общем случае подматрицу $A(I,J)$ можно перевести в левый верхний угол с помощью транспозиций строк и столбцов исходной матрицы. Для перемещения строки i_1 на первое место нужно выполнить i_1-1 транспозицию, затем для сдвига строки i_2 на второе место понадобится еще i_2-2 транспозиции, и так далее. Общее число строчных и столбцовых транспозиций будет равно

$$(i_1-1)+(i_2-2)+\ldots(i_k-k)+(j_1-1)+(j_2-2)+\ldots(j_k-k)=c(I,J)+k(k+1),$$

при этом число $k(k+1)$ заведомо четное. Напомним, что $c(I,J)$ обозначает сумму всех номеров, входящих в множества I и J. Произведение $|A(I,J)||A'(I,J)|$ будет суммой членов определителя новой матрицы, который равен $(-1)^{c(I,J)}|A|$. Следовательно, величина $|A(I,J)||A'(I,J)|(-1)^{c(I,J)}$ будет суммой членов определителя исходной матрицы.

Остается заметить, что каждое произведение минора на алгебраическое дополнение дает $k!(n-k)!$ членов определителя $|A|$, а общее число таких произведений равно $n!/(k!(n-k)!)$ — число сочетаний из n по k. Таким образом, будут получены все $n!$ членов определителя $|A|$. $\square$

Задача 7. *Докажите, что если сумма алгебраических дополнений всех отдельных элементов квадратной матрицы равна нулю, то ее определитель не меняется при прибавлении к каждому элементу одного и того же числа.*

3.14 Присоединенная матрица

Пусть A — матрица порядка n с элементами a_{ij}. Через A_{ij} обозначается алгебраическое дополнение элемента a_{ij} как минора первого порядка. Матрица $\widehat{A}$ с элементами $\widehat{a}_{ij}=A_{ji}$ называется *присоединенной* к матрице A.

Теорема о присоединенной матрице. *$A\widehat{A}=\widehat{A}A=|A|\cdot I$, где I — единичная матрица порядка n.*

Доказательство. По теореме Лапласа, $(A\widehat{A})_{ij}=\sum\limits_{k=1}^{n}a_{ik}A_{jk}=|B|$, где B — матрица, полученная из A заменой j-й строки на i-ю. Если $i=j$, то, очевидно, $B=A$. Если $i\neq j$, то матрица B имеет пару одинаковых строк и поэтому $|B|=0$. Аналогично, $(\widehat{A}A)_{ij}=\sum\limits_{k=1}^{n}A_{ki}a_{kj}=|C|$, где C получается из A заменой i-го столбца на j-й. $\square$

3.15 Обратимость и невырожденность

Квадратная матрица A называется *невырожденной*, если $|A|\neq 0$.

Теорема. *Матрица обратима в том и только том случае, когда она невырождена.*

Доказательство. Если $|A|\neq 0$, то, согласно теореме о присоединенной матрице, для матрицы $X=(1/|A|)\widehat{A}$ находим $AX=XA=I$. Следовательно, матрица A обратима и $X=A^{-1}$. Если $AA^{-1}=I$, то, по теореме об определителе произведения матриц, $|A||A^{-1}|=|I|=1\quad\Rightarrow\quad|A|\neq 0$. $\square$

Задача 8. *Докажите, что любую невырожденную матрицу можно сделать вырожденной, изменив лишь один из ее элементов.*

Задача 9. *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 $I_n - BA$. Докажите также, что обратимость каждой из этих матриц равносильна обратимости матрицы порядка $m + n$ с блочным разбиением вида*

$$\begin{bmatrix} I_m & A \\ B & I_n \end{bmatrix}.$$

Задача 10. *Даны числа a и b такие, что $1 - ab \neq 0$. Докажите, что матрица*

$$A = \begin{bmatrix} 1 & a & a^2 & \dots & a^n \\ b & 1 & a & \dots & a^{n-1} \\ b^2 & b & 1 & \dots & a^{n-2} \\ \dots & \dots & \dots & \dots & \dots \\ b^n & b^{n-1} & b^{n-2} & \dots & 1 \end{bmatrix}$$

обратима и обратная к ней матрица является трехдиагональной.

Задача 11. *Пусть в условии предыдущей задачи a и b — квадратные блоки порядка n, а 1 заменяется единичной матрицей I того же порядка. Докажите, что если блок $I - ab$ является обратимой матрицей, то блочная матрица A с блоками порядка n обратима и при этом обратная к ней матрица является блочно-трехдиагональной.*

Задача 12. *Рассматриваются матрицы порядка n, элементы которых равны 0 или 1. Среди матриц такого вида найдите число матриц с нечетным определителем.*

3.16 Правило Крамера

Теорема. *Если вектор-столбец x с элементами $x_1, \dots, x_n$ является решением системы $Ax = b$ с квадратной матрицей A, то*

$$x_j \, |A| = |A_j(b)|,$$

где матрица $A_j(b)$ получается из A заменой j-го столбца на вектор b.

Доказательство. Равенство $Ax = b$ означает, что вектор b является линейной комбинацией столбцов матрицы A с коэффициентами $x_1, \dots, x_n$:

$$b = x_1 a_1 + \dots + a_n a_n.$$

Используя свойство линейности определителя как функции j-го столбца при фиксированных значениях остальных столбцов, находим

$$|A_j(b)| = \sum_{i=1}^{n} x_i |A_j(a_i)| = x_j |A_j(a_j)| = x_j |A|. \quad \square$$

Следствие. *Если A — невырожденная матрица порядка n, то система $Ax = b$ совместна и обладает единственным решением, которое определяется по формулам Крамера*

$$x_j \;=\; \frac{|A_j(b)|}{|A|}, \quad 1 \leq j \leq n.$$

Исторически именно эти формулы привели Крамера к понятию определителя. Понятно, что решение системы $Ax = b$ можно вычислить с помощью арифметических операций, и это означает, что каждая компонента решения является *рациональной функцией* от элементов матрицы и правой части, т.е. x_j представляется несократимой дробью,

в которой числитель и знаменатель суть многочлены от элементов матрицы и правой части, которые можно рассматривать как переменные или, что вообще типично для алгебры, просто как *буквы* с индексами, используемые для формальной записи многочлена. Крамеру удалось найти явные выражения для многочленов указанных дробей.

Задача 13. *Множество всех целочисленных линейных комбинаций системы n линейно независимых арифметических векторов n-мерного пространства называется* `решеткой`. *Пусть L — некоторая решетка. Докажите, что для любой решетки $M \subset L$ найдется целое число c, зависящее только от M и такое, что $cx \in M$ для любого вектора $x \in L$.*

Алгебра и геометрия (1 поток)

Лекция 4

4.1 Разделение переменных и ранг

При изучении функций от двух или большего числа переменных особую роль играют произведения функций, каждая из которых зависит только от одной переменной: $f(x, y, \ldots, z) = u(x)v(y)\ldots w(z)$. Они называются *функциями с разделенными переменными* или *разложимыми функциями*. При достаточно общих условиях многие функции приближаются конечными суммами разложимых функций.

Нас интересуют прежде всего функции, определенные на конечных множествах. Они всегда представимы *конечной суммой* разложимых функций (докажите!):

$$f(x, y, \ldots, z) = \sum_{\alpha=1}^{r} u_\alpha(x)v_\alpha(y)\ldots w_\alpha(z).$$

Разложение с минимальным числом слагаемых r называется *минимальным*, а само число слагаемых — *рангом разделения переменных* данной функции. Ранг тождественно нулевой функции считается равным нулю.

Матрицу $A = [a_{ij}]$ размеров $m \times n$ можно рассматривать как таблицу значений функции $(i, j) \to a_{ij}$, аргументами которой служат индексы. Напомним, что под d-мерной матрицей или тензором понимается таблица величин $a_{i_1,\ldots,i_d}$, снабженных d индексами. Элементы $a_{i_1,\ldots,i_d}$ можно считать значениями отображения $(i_1, \ldots, i_d) \to a_{i_1,\ldots,i_d}$. Его ранг разделения переменных называется *рангом тензора* с элементами $a_{i_1,\ldots,i_d}$. Тензор, определяемый разложимой функцией, называется *разложимым тензором*.

Задача 1. *Докажите, что ранг разделения переменных функции $f(x, y) = \sin(x + y)$, определенной для всех вещественных x и y, равен 2.*

4.2 Разложимые матрицы

Разложимая функция от двух переменных имеет вид $a_{ij} = u_i v_j$. Обозначим через A матрицу с элементами a_{ij}, а через u и v — столбцы с элементами u_i и v_j. Тогда

$$A = uv^\top = \begin{bmatrix} u_1 \\ \ldots \\ u_m \end{bmatrix} \begin{bmatrix} v_1 & \ldots & v_n \end{bmatrix}.$$

Матрица такого вида называется *внешним произведением векторов* u и v. Такие матрицы мы будем называть *разложимыми матрицами*.

4.3 Скелетное разложение и сжатие информации

Рассмотрим матрицы

$$U = \begin{bmatrix} u_1 & \dots & u_r \end{bmatrix} = \begin{bmatrix} u_{11} & \dots & u_{1r} \\ u_{21} & \dots & u_{2r} \\ \dots & \dots & \dots \\ u_{m1} & \dots & u_{mr} \end{bmatrix}, \quad V = \begin{bmatrix} v_1 & \dots & v_r \end{bmatrix} = \begin{bmatrix} v_{11} & \dots & v_{1r} \\ v_{21} & \dots & v_{2r} \\ \dots & \dots & \dots \\ v_{n1} & \dots & v_{nr} \end{bmatrix}.$$

Тогда (проверьте!)

$$UV^\top = \begin{bmatrix} u_{11} & \dots & u_{1r} \\ u_{21} & \dots & u_{2r} \\ \dots & \dots & \dots \\ u_{m1} & \dots & u_{mr} \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} & \dots & v_{n1} \\ \dots & \dots & \dots & \dots \\ v_{1r} & v_{2r} & \dots & v_{nr} \end{bmatrix} = u_1 v_1^\top + \ldots + u_r v_r^\top.$$

Представление матрицы A в виде произведения $A = UV^\top$ называется ее *скелетным разложением*. Поэлементные равенства $a_{ij} = \sum\limits_{k=1}^{r} u_{ik} v_{jk}$ показывают, что функция $(i,j) \to a_{ij}$ является суммой разложимых функций $f_k(i,j) = u_{ik} v_{jk}$, а матрица A является суммой r разложимых матриц. Скелетное разложение называется *минимальным*, если r совпадает с рангом разделения переменных для матрицы A.

Очевидно, что скелетное разложение существует для любой матрицы — например, при $r = n$ можно взять $U = A$ и $V = I$. Такое тривиальное разложение, правда, малоинтересно. В приложениях часто приходится работать с матрицами больших размеров, для которых имеются скелетные разложения с достаточно малым значением r по сравнению с размерами. В таких случаях вместо элементов матрицы A в памяти компьютера можно хранить элементы матриц U и V ее скелетного разложения. Вместо mn ячеек для размещения всех элементов $m \times n$-матрицы понадобится всего лишь $(m+n)r \ll mn$ ячеек. Таким образом, скелетные разложения могут использоваться как инструмент *сжатия информации*, представленной в виде матриц.

4.4 Минимальное скелетное разложение

Как получить минимальное скелетное разложение?

Выберем в матрице A максимальную линейно независимую подсистему столбцов $u_1 = a_{j_1}, \ \dots \ , \ u_s = a_{j_s}$. Тогда каждый столбец является их линейной комбинацией

$$a_j = \sum_{k=1}^{s} v_{jk} u_k \quad \Rightarrow \quad A = UV^\top, \ \text{ где } \ U = [u_{ik}]_{m \times s}, \ \ V = [v_{jk}]_{n \times s}.$$

Полученное скелетное разложение будет минимальным. В самом деле, рассмотрим произвольное скелетное разложение $A = PQ^\top$, в котором матрицы P и Q составлены из r столбцов. Ясно, что линейная оболочка столбцов матрицы A, а значит и каждый из векторов $u_1, \ \dots \ , \ u_s$, содержится в линейной оболочке столбцов матрицы P этого разложения. Из теоремы о монотонности вытекает, что $s \le r$.

Аналогичное построение можно выполнить, взяв максимальную линейно независимую подсистему строк матрицы A. Этими строками определяется матрица $V^\top$, а матрица U составляется из коэффициентов линейных комбинаций этих строк для представления каждой строки матрицы A. Таким образом, справедливо следующее

Утверждение. *Размерность линейной оболочки столбцов матрицы A совпадает с размерностью линейной оболочки ее строк и равна минимальному числу разложимых матриц в ее представлениях в виде суммы таких матриц.*

4.5 Ранг и миноры

В курсах линейной алгебры обычно дается такое определение ранга матрицы: рассматриваются порядки всех ненулевых миноров матрицы A, а ее *рангом* называется наибольший среди них и обозначается через $\mathrm{rank}(A)$. Ранг нулевой матрицы считается равным нулю.

Однако, у нас уже есть понятие ранга разделения для матрицы, который определяется как минимальное число слагаемых в ее представлениях в виде суммы разложимых матриц (равное размерности линейной оболочки столбцов и одновременно размерности линейной оболочки строк данной матрицы). Теперь нам нужно понять, что речь идет об одном и том же числе: *наивысший порядок ненулевых миноров матрицы равен минимальному числу разложимых матриц во всевозможных ее представлениях в виде суммы таких матриц.* Это важное утверждение вытекает из теоремы о базисном миноре.

4.6 Теорема о базисном миноре

Невырожденная подматрица порядка r называется *базисной* в матрице A, если любая подматрица порядка $r+1$, полученная из нее окаймлением с помощью одной строки и одного столбца, является вырожденной. Определитель базисной подматрицы называется *базисным минором.* Столбцы и строки, на которых расположен базисный минор, часто называются *базисными.*

Теорема о базисном миноре. *Столбцы (строки) матрицы, на которых расположен ее базисный минор, являются линейно независимыми, а любой столбец (любая строка) данной матрицы представляется их линейной комбинацией.*

Доказательство. Предположим, что базисная подматрица матрицы A расположена на строках с номерами $i_1 < \ldots < i_r$ и столбцах с номерами $j_1 < \ldots < j_r$. Пусть базисный минор равен равен $\delta \neq 0$.

Уравнение $x_1 a_{j_1} + \ldots + x_r a_{j_r} = b$ относительно неизвестных коэффициентов $x_1, \ldots, x_r$ равносильно системе линейных алгебраических уравнений

$$\begin{bmatrix} a_{1,j_1} & \ldots & a_{1,j_r} \\ a_{2,j_1} & \ldots & a_{2,j_r} \\ \ldots & \ldots & \ldots \\ a_{m,j_1} & \ldots & a_{m,j_r} \end{bmatrix} \begin{bmatrix} x_1 \\ \ldots \\ x_r \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} a_{i_1,j_1} & \ldots & a_{i_1,j_r} \\ \ldots & \ldots & \ldots \\ a_{i_r,j_1} & \ldots & a_{i_r,j_r} \end{bmatrix} \begin{bmatrix} x_1 \\ \ldots \\ x_r \end{bmatrix} = \begin{bmatrix} b_{j_1} \\ \ldots \\ b_{j_r} \end{bmatrix}.$$

Определитель матрицы коэффициентов второй системы равен $\delta \neq 0$. Поэтому вторая система совместна и обладает единственным решением. Значит, величины $x_1, \ldots, x_r$ определяются из нее однозначно, а первая система в случае ее совместности имеет единственное решение. Если $b = 0$, то она, очевидно, совместна (как любая однородная система) и, следовательно, $x_1 = \ldots = x_r = 0$. Таким образом, система столбцов $a_{j_1}, \ldots, a_{j_r}$ линейно независима.

Теперь докажем совместность первой линейной системы в случае $b = a_j$. Нетривиален, конечно, лишь случай $j \notin \{j_1, \ldots, j_r\}$. Возьмем $i \notin \{i_1, \ldots, i_r\}$ и рассмотрим

определитель подматрицы, полученной из базисной окаймлением с помощью i-й строки и j-го столбца матрицы A:

$$\begin{vmatrix} a_{i_1,j_1} & \dots & a_{i_1,j_r} & a_{i_1,j} \\ \dots & \dots & \dots & \\ a_{i_r,j_1} & \dots & a_{i_r,j_r} & a_{i_r,j} \\ a_{i,j_1} & \dots & a_{i,j_r} & a_{i,j} \end{vmatrix} = \begin{vmatrix} a_{i_1,j_1} & \dots & a_{i_1,j_r} & 0 \\ \dots & \dots & \dots & \\ a_{i_r,j_1} & \dots & a_{i_r,j_r} & 0 \\ a_{i,j_1} & \dots & a_{i,j_r} & \gamma_{i,j} \end{vmatrix} = \delta\,\gamma_{i,j},$$

$$\gamma_{i,j} = a_{i,j} - x_1 a_{i,j_1} - \ldots - x_r a_{i,j_r}.$$

Здесь мы использовали то, что определитель сохраняется при вычитании из последнего столбца линейной комбинации предыдущих столбцов — в нашем случае с коэффициентами $x_1, \ \ldots \ , \ x_r$, и применили теорему Лапласа для разложения второго определителя по последнему столбцу. Согласно определению базисного минора, любой окаймляющий его минор порядка $r+1$ равен нулю. Следовательно, $\delta\,\gamma_{i,j} = 0 \quad \Rightarrow \quad \gamma_{i,j} = 0 \quad \Rightarrow$

$$x_1 a_{i,j_1} + \ldots x_r a_{i,j_r} = a_{i,j}, \quad 1 \le i,j \le n.$$

Утверждение теоремы относительно строк доказывается переходом к транспонированной матрице. □

Следствие. *Матрица имеет ранг r тогда и только тогда, когда некоторый минор порядка r отличен от нуля, а все окаймляющие его миноры порядка $r+1$ равны нулю.*

Задача 2. *Докажите, что подматрица $\widehat{A}$, расположенная на пересечении r линейно независимых строк и r линейно независимых столбцов матрицы A ранга r, является невырожденной и при этом $A = C\widehat{A}^{-1}R$, где C — подматрица, составленная из данных столбцов, а R — подматрица, составленная из данных строк.*

4.7 Минимальное число разложимых матриц

Теорема. *Ранг матрицы равен минимальному числу разложимых матриц во всевозможных ее представлениях в виде суммы таких матриц.*

Доказательство. В разделе 4.4 было доказано, что минимальное число разложимых матриц в представлениях матрицы равно размерности линейной оболочки ее столбцов, которая, согласно теореме о базисном миноре, равна рангу этой матрицы. □

4.8 Ранги суммы и произведения матриц

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

Доказательство. Линейная оболочка столбцов матрицы $A+B$ содержится в линейной оболочке объединенной системы столбцов матриц A и B, а значит и в линейной оболочке, натянутой на объединение базисов линейных оболочек столбцов матриц A и B. Число векторов в каждом базисе равно рангу соответствующей матрицы. Таким образом, размерность линейной оболочки столбцов матрицы $A+B$ не превосходит суммы размерностей линейных оболочек столбцов матриц A и B, которые равны рангам этих матриц. □

Следствие. *При изменении, добавлении или изъятии k строк (столбцов) матрицы ее ранг не может измениться больше чем на k.*

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

Доказательство. 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, $\operatorname{rank}(AB) = \operatorname{rank}(AB)^\top = \operatorname{rank}(B^\top A^\top) \le \operatorname{rank} B^\top = \operatorname{rank} B$. □

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. В силу утверждения 2, $\operatorname{rank} B \le \operatorname{rank} A$. В то же время $A = P^{-1}BQ^{-1} \Rightarrow \operatorname{rank} A \le \operatorname{rank} B$. □

Обычно ранг вычисляется с помощью *элементарных преобразований* строк и столбцов, упрощающих вид матрицы путем исключения ее элементов — эти преобразования сводятся к умножению матрицы слева и справа на некоторые обратимые матрицы (специального вида) и поэтому сохраняют ранг.

Задача 3. *Пусть A и B — матрицы ранга 1. Докажите, что если $AB = BA \ne 0$, то ранг матрицы $A + B$ не больше 1.*

Задача 4. *Заданы столбцы $x, y \in \mathbb{R}^n$, причем $x \ne 0$. Докажите, что существует симметричная матрица $A \in \mathbb{R}^{n\times n}$ такая, что* $\operatorname{rank} A \le 2$ *и* $Ax = y$.

Задача 5. *Матрица A имеет r столбцов, а матрица B имеет r строк. Докажите следующее* `неравенство Сильвестра`*:* $r \ge \operatorname{rank}(A) + \operatorname{rank}(B) - \operatorname{rank}(AB)$.

4.9 Теорема Кронекера–Капелли

Как мы уже знаем, система $Ax = b$ совместна в том и только том случае, когда правая часть b принадлежит линейной оболочке стольцов матрицы A. Это условие можно сформулировать в терминах рангов матрицы A и расширенной матрицы $\begin{bmatrix} A & b \end{bmatrix}$.

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

Доказательство. Пусть $a_{j_1}, \ldots a_{j_r}$ — базисная система столбцов матрицы A. Равенство рангов матрицы A и расширенной матрицы равносильно тому, что эта же система будет базисной в расширенной матрице. □

4.10 Ранг и дефект

С матрицей A связаны два очень полезных векторных пространства: линейная оболочка ее столбцов, называемая также *образом матрицы*, и множество всех решений однородной системы $Ax = 0$, называемое также *ядром матрицы*. Для образа и ядра употребляются обозначения $\operatorname{im}(A)$ и $\ker(A)$.

Размерность образа матрицы равна ее рангу. Размерность ядра матрицы называется ее *дефектом* и обозначается $\operatorname{def}(A) = \dim \ker(A)$.

Теорема. *Для любой матрицы сумма ранга и дефекта равна числу ее столбцов.*

Доказательство. Среди столбцов матрицы A выберем базисные столбцы $a_{j_1}, \ldots, a_{j_r}$, $r = \operatorname{rank}(A)$. Пусть $k_1 < \ldots < k_{n-r}$ — номера остальных столбцов, которые мы будем

называть *свободными*. Соответствующие им неизвестные $x_{k_1},\ \ldots,\ x_{k_{n-r}}$ будем также называть свободными. Тогда уравнение $Ax = 0$ можно записать в виде равенства линейных комбинаций базисных и свободных столбцов:

$$x_{j_1}a_{j_1} + \ldots + x_{j_r}a_{j_r} = -x_{k_1}a_{k_1} - \ldots - x_{k_{n-r}}a_{k_{n-r}}.$$

Для любых значений свободных неизвестных равенство выполняется для некоторых и притом однозначно определенных значений базисных неизвестных. Ясно, что любое решение однородной системы $Ax = 0$ можно получить именно таким образом. Теперь выберем $n - r$ специальных решений $x^{(1)},\ \ldots,\ x^{(n-r)}$, в каждом из которых все свободные неизвестные равны 0, кроме одной из них, равной 1. Такую систему традиционно называют *фундаментальной системой решений*. Легко понять, что фундаментальные решения линейно независимы, а любое решение выражается в виде их линейной комбинации (проверьте!). Таким образом, $\ker(A) = L(x^{(1)},\ \ldots,\ x^{(n-r)})$. $\square$

Задача 6. *Даны квадратные матрицы $A = I_m + UU^\top$ и $B = I_n + U^\top U$, где I_m и I_n — единичные матрицы порядка m и n. Найти ранг матрицы A, если ранг матрицы B равен r.*

Задача 7. *Даны матрицы A и B порядка n такие, что $AB = 0$ и при этом матрица $A + B$ невырожденная. Доказать, что* $\mathrm{rank}A + \mathrm{rank}B = n$.

Задача 8. *Что можно сказать о матрице и правой части системы $Ax = b$ относительно вектора $x \in \mathbb{R}^n$, если ее решением является любой вектор из $\mathbb{R}^n$?*

4.11 Элементарные преобразования

Чтобы выяснить совместность и найти общее решение заданной системы линейных алгебраических уравнений, особых знаний (понятие ранга, теорема Кронекера–Капелли, правило Крамера и т.п.), вообще говоря, не нужно. Задача решается с помощью последовательного исключения неизвестных (вероятно, первое, что приходит на ум) и во многих случаях, по крайней мере при решении "на бумаге", не видно каких-либо проблем, которые требуют развитой нами науки. Теория матриц, однако, *совершенно необходима* для анализа алгоритма исключения неизвестных и в особенности его компьютерных реализаций.

Процесс исключения неизвестных можно интерпретировать как переход от системы $Ax = b$ к системе $PAx = Pb$, в которой матрица коэффициентов PA в каких-то позициях приобретает нули. Другими словами, преобразование $A \to PA$ обнуляет какие-то элементы матрицы A. Матрица P должна быть, конечно, обратимой (почему?). Последовательность преобразований такого типа должна привести к системе специального вида, которая решается вообще очевидным образом (например, к системе с треугольной матрицей).

Чтобы показать, как происходит исключение элементов, рассмотрим матрицу A размеров 4×6, заведомо нулевые элементы будем обозначать "ноликом", произвольные элементы — "крестиком", а заведомо ненулевые элементы — "крестиком в рамочке". Если $a_{11} \neq 0$, то из каждой строки с номером $i \geq 2$ можно вычесть первую с коэффициентом, дающим нуль в позиции $(i, 1)$:

$$\begin{bmatrix} \boxtimes & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} \boxtimes & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{bmatrix}.$$

В дальнейшем первые строка и столбец остаются без изменений. Если во втором столбце имеется не нуль, делаем соответствующую строку второй и с ее помощью исключаем элементы второго столбца:

$$\begin{bmatrix} \boxtimes & \times & \times & \times & \times & \times \\ 0 & \boxtimes & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} \boxtimes & \times & \times & \times & \times & \times \\ 0 & \boxtimes & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times & \times \end{bmatrix}.$$

Первые два столбца и первые две строки больше не меняются. Может случиться так, что у нас получились "внеплановые" нулевые столбцы. Например, если оставшиеся "крестики" в третьем и четвертом столбцах оказались нулями, а ненулевой элемент обнаружен только в пятом столбце, то исключения проводятся сразу в пятом столбце:

$$\begin{bmatrix} \boxtimes & \times & \times & \times & \times & \times \\ 0 & \boxtimes & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & \boxtimes & \times \\ 0 & 0 & 0 & 0 & \times & \times \end{bmatrix} \quad \mapsto \quad \begin{bmatrix} \boxtimes & \times & \times & \times & \times & \times \\ 0 & \boxtimes & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & \boxtimes & \times \\ 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix}.$$

В общем случае, чтобы получить нуль в позиции (k, l), надо из k-й строки вычесть l-ю строку, взятую с подходящим множителем:

$$a_{kj} \;\to\; a_{kj} - \frac{a_{kl}a_{lj}}{a_{ll}}.$$

Именно так действует преобразование $A \to PA$ с обратимой матрицей вида

$$P = I + \gamma_{kl}E^{(kl)}, \quad \gamma_{kl} = -\frac{a_{kl}}{a_{ll}},$$

где $E^{(kl)}$ — матрица, отличающаяся от нулевой только одним элементом, расположенным в позиции (k, l) и равным 1. Деление происходит на элемент $a_{ll} \neq 0$, который принято называть *ведущим элементом*. Преобразования и матрицы такого типа в контексте исключения элементов называются *элементарными*. В набор элементарных матриц включаются также матрицы перестановки, которые получаются из единичной матрицы перестановкой строк или столбцов, и невырожденные диагональные матрицы.

Элементарные преобразования могут применяться и к столбцам матрицы. В таких случаях нуль в каком-либо столбце получается при вычитании из него другого столбца, взятого с подходящим множителем. Такие операции, а также перестановки столбцов и умножения столбцов на ненулевые числа, реализуются умножением на элементарную матрицу справа.

4.12 Ступенчатые матрицы

Матрица называется *ступенчатой*, если для каждого ее ненулевого элемента все элементы слева и снизу от него равны нулю, а при наличии нулевой строки все строки ниже нее тоже нулевые. Число ненулевых строк ступенчатой матрицы равно ее рангу (проверьте!).

Теорема. *Любая матрица путем умножения слева на последовательность элементарных матриц приводится к ступенчатому виду с числом ненулевых строк, равным рангу исходной матрицы.*

Доказательство. Пусть A — матрица размеров $m \times n$. Если она нулевая, то доказывать нечего. Если нет, то проведем индукцию по числу строк. Случай $m = 1$ тривиален: любая матрица, состоящая из одной строки, является ступенчатой. Пусть $m \geq 2$.

Найдем первый ненулевой столбец и переставим строки, сделав ненулевой элемент первым в этом столбце. Такое преобразование выполняется с помощью умножения слева на некоторую матрицу перестановки. В преобразованной матрице первый элемент первого ненулевого столбца отличен от нуля. Далее мы используем его как ведущий элемент для зануления расположенных под ним элементов. Это делается также путем умножения слева на последовательность элементарных матриц. В итоге возникает матрица следующего вида:

$$B = \begin{bmatrix} 0 & \dots 0 & \boxtimes & \times & \dots & \times \\ 0 & \dots 0 & 0 & \times & \dots & \times \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots 0 & 0 & \times & \dots & \times \end{bmatrix}.$$

Обозначим через $\widehat{B}$ подматрицу, полученную из B вычеркиванием первой строки. По индуктивному предположению, существует последовательность элементарных матриц $\widehat{P}_1$, …, $\widehat{P}_s$ такая, что матрица $\widehat{C} = \widehat{P}_s \dots \widehat{P}_1 \widehat{B}$ является ступенчатой. Остается заметить, что матрица

$$P_s \dots P_1 B = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \dots & & \widehat{C} & \\ 0 & & & \end{bmatrix}, \quad \text{где} \quad P_k = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \dots & & \widehat{P}_k & \\ 0 & & & \end{bmatrix},$$

будет ступенчатой, а каждая матрица P_k — элементарной. □

4.13 Эквивалентные матрицы

Матрица A называется *эквивалентной* матрице B, если существуют обратимые матрицы P и Q такие, что $B = PAQ$. При этом матрица B будет также эквивалентной матрице A (в силу равенства $A = P^{-1}BQ^{-1}$). Поэтому в таких случаях мы говорим просто об *эквивалентных матрицах*.

Теорема об элементарных преобразованиях. *Любую ненулевую матрицу A размеров $m \times n$ с помощью элементарных преобразований строк и столбцов можно привести к виду*

$$PAQ = \begin{bmatrix} I_r & 0_{r\times(n-r)} \\ 0_{m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix},$$

где r – ранг матрицы A, I_r – единичная матрица порядка r, P и Q представляются произведениями элементарных матриц,

Доказательство. С помощью элементарных преобразований строк матрицу A можно привести к ступенчатой матрице $B = PA$, в которой первые элементы ненулевых строк равны 1. После этого заявленный в теореме результат получается с помощью вполне очевидных элементарных преобразований столбцов матрицы B. □

Теорема об эквивалентных матрицах. *Матрицы эквивалентны тогда и только тогда, когда они имеют одинаковые размеры и одинаковые ранги.*

Доказательство. Если A и B эквивалентны, то совпадение размеров вытекает непосредственно из определения произведения матриц, а равенство рангов – из свойства сохранения ранга при умножении на невырожденные матрицы. Теперь предположим, что размеры и ранги матриц A и B сопадают. Тогда, согласно теореме об элементарных преобразованиях, обе матрицы эквивалентны одной и той же матрице, а значит, и сами являются эквивалентными. $\square$

Задача 9. *Даны $n \times n$-матрицы A и B ранга r, а элементы $n \times n$-матриц X и Y удовлетворяют системе линейных алгебраических уравнений, записанной в виде матричного уравнения $AX+YB = 0$. Найдите дефект матрицы коэффициентов данной системы.*

4.14 LU-разложение и строго регулярные матрицы

Если матрица невырождена, то ее ступенчатая форма, получаемая с помощью элементарных преобразований строк, является верхней треугольной матрицей (проверьте!).

Если перестановки строк не потребовались, то это значит, что в процессе исключения использовались только элементарные матрицы вида $I + \gamma_{kl}E^{(kl)}$ при $k > l$. Такие матрицы являются нижними треугольными, а их главная диагональ состоит только из единиц. Треугольные матрицы с таким свойством называются *унитреугольными*. При их перемножении и при обращении получаются также унитреугольные матрицы (проверьте!). В итоге для матрицы A возникает разложение $A = LU$, где матрица L нижняя унитреугольная, а матрица U – невырожденная верхняя треугольная. Разложение такого типа называется *LU-разложением*.

Утверждение. *Если LU-разложение существует, то оно единственно.*

Доказательство. $L_1U_1 = L_2U_2 \;\Rightarrow\; L_2^{-1}L_1 = U_2U_1^{-1} \;\Rightarrow\; L_2^{-1}L_1 = U_2U_1^{-1} = I$. $\square$

Существование LU-разложения связано со свойством *строгой регулярности* матрицы. Оно означает, что все ее *ведущие подматрицы* (т.е. квадратные подматрицы, расположенные в левом верхнем углу), включая саму матрицу, невырождены.

Теорема. *Для существования LU-разложения квадратной матрицы необходимо и достаточно, чтобы она была строго регулярной.*

Доказательство. Предположим сначала, что LU-разложение существует, и запишем его в блочной форме:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} \;\Rightarrow\; A_{11} = L_{11}U_{11} \;\Rightarrow\; |A_{11}| = |L_{11}||U_{11}| = |U_{11}| \neq 0.$$

Перейдем к доказательству достаточности. В строго регулярной матрице $a_{11} \neq 0$. Для получения нулей в первом столбце используем этот элемент как ведущий:

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ \gamma_{21} & & & \\ \dots & & I_{n-1} & \\ \gamma_{n1} & & & \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & & & \\ \dots & & A_{22} & \\ a_{n1} & & & \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & & & \\ \dots & & H & \\ 0 & & & \end{bmatrix}, \quad \gamma_{i1} = -\frac{a_{i1}}{a_{11}}.$$

Отсюда видно, что матрица H наследует свойство строгой регулярности (проверьте!). Доказательство завершается индукцией по n. $\square$

Следствие. *Для произвольной невырожденной матрицы A существует разложение вида $A = PLU$, где P – матрица перестановки, L – нижняя унитреугольная матрица, U – невырожденная верхняя треугольная матрица.*

Доказательство. Достаточно убедиться в том, что с помощью перестановки строк любую невырожденную матрицу можно преобразовать в строго регулярную. □

Задача 10. *Пусть A_{11} – ведущая подматрица в невырожденной матрице A, а B_{11} – ведущая подматрица такого же порядка в матрице A^{-1}. Рассмотреть блочные разбиения*

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

и доказать, что $|A_{11}| = |B_{22}||A|$.

Задача 11. *Докажите, что для симметричной строго регулярной матрицы A существует и единственно разложение вида $A = LDL^\top$, где матрица L унитреугольная, а D — диагональная.*

Задача 12. *Докажите, что любую вещественную матрицу, определитель которой равен единице, можно разложить в произведение элементарных матриц вида $I + \alpha E^{kl}$ для каких-то вещественных значений α и номеров $k \neq l$. По определению, $(E^{kl})_{ij} = 1$ при $i = k$ и $j = l$, все остальные элементы равны 0.*

4.15 Выбор ведущего элемента

С теоретической точки зрения важно только то, что ведущий элемент в процессе исключения элементов не равен нулю. С точки зрения практических вычислений этого мало.

Дело в том, что компьютер оперирует с конечным набором вещественных чисел — так называемых *машинных чисел*. При использовании p-ичной системы счисления любое вещественное число можно записать в виде

$$x = p^\alpha \cdot \beta, \quad 0 \leq \beta < 1, \qquad (*)$$

где α — целое число, называемое *порядком* числа x, а β — вещественное число, называемое *мантиссой* числа x (конечно, порядок и мантисса для x зависят от p). [1] На компьютере для представления порядка и мантиссы отводится лишь конечное число разрядов. Поэтому при выполнении операций с машинными числами приходится делать *округление* — замену точного результата каким-то близким машинным числом. Предположим, например, что мантисса имеет $t = 5$ разрядов. Тогда при сложении чисел $a = 10^2 \cdot 0.11111$ и $b = 10^{-4} \cdot 0.11111$ сначала "выравниваются" порядки — это означает изменение мантиссы числа с меньшим порядком и *потерю знаков*, оказавшихся за пределами отведенных для представления мантисс разрядов: $10^{-4} \cdot 0.11111 = 10^2 \cdot 0.00000011111 \mapsto 10^2 \cdot 0.00000$. Далее модифицированные мантиссы складываются, после чего результат приводится к виду $(*)$. В данном случае $10^2 \cdot 0.11111 + 10^2 \cdot 0.00000 = 10^2 \cdot 0.11111$. Как видим, сумма положительных чисел a и b оказалась равной a.

Пусть на этом же компьютере решается система $\begin{bmatrix} 10^{-5} & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Легко видеть, что точное решение имеет вид

$$x = \frac{-1}{1 - 10^{-5}}, \qquad y = \frac{2 - 10^{-5}}{1 - 10^{-5}}.$$

В то же время, при исключении элемента в позиции $(2, 1)$ получаем

$$\begin{bmatrix} 1 & 0 \\ -10^5 & 1 \end{bmatrix} \begin{bmatrix} 10^{-5} & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 10^{-5} & 1 \\ 0 & 1 - 10^5 \end{bmatrix} \mapsto \begin{bmatrix} 10^{-5} & 1 \\ 0 & -10^5 \end{bmatrix},$$

так как

$$1 - 10^5 = 10^1 \cdot 0.10000 - 10^6 \cdot 0.10000 = 10^6 \cdot 0.000001 - 10^6 \cdot 0.10000$$
$$\mapsto 10^6 \cdot 0.00000 - 10^6 \cdot 0.10000 = -10^5.$$

[1] Обычно $p = 2$, но есть и компьтер, для которого $p = 3$ – это электронно-вычислительная машина (ЭВМ) "Сетунь", разработанная в Московском университете в 1960-х годах.

Аналогично, при соответствующем преобразовании правой части находим

$$\begin{bmatrix} 1 & 0 \\ -10^5 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \;\mapsto\; \begin{bmatrix} 2 \\ -2 \cdot 10^5 \end{bmatrix}.$$

В итоге компьютерное решение $\tilde{x}$, $\tilde{y}$ будет точным решением системы

$$\begin{bmatrix} 10^{-5} & 1 \\ 0 & -10^5 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \cdot 10^5 \end{bmatrix} \quad \Rightarrow \quad \tilde{x} = 0, \quad \tilde{y} = 2.$$

Как видим, компьютерное решение далеко от истинного.

Причина чудовищно большой погрешности — в относительно малой величине ведущего элемента, приводящей к *росту элементов* в преобразованной матрице. Чтобы снизить неприятный эффект, вызванный ростом элементов, обычно рекомендуется в каждом столбце выбирать в качестве ведущего элемент, максимальный по модулю.

Задача 13. *Докажите, что для любой квадратной матрицы A с диагональным преобладанием по строкам существует LU-разложение, в котором верхняя треугольная матрица U имеет диагональное преобладание по строкам, а ее максимальный по модулю элемент не превосходит удвоенного максимального по модулю элемента матрицы A.*

4.16 Вычисление обратной матрицы

Подсчитаем число арифметических операций при приведении строго регулярной матрицы A к верхней треугольной матрице U. При работе с i-м столбцом на i-м шаге требуется получить $n - i$ нулей. Для этого понадобится $(n - i)^2$ умножений и столько же вычитаний. Общее число умножений при вычислении LU-разложения совпадает с числом вычитаний и будет равно

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

где через $O(n^2)$ обозначен многочлен от n степени 2. Далее решение системы $Ax = b$ проводится таким образом: решается система $Ly = b$, а затем система $Ux = y$. Благодаря треугольному виду матриц решение этих задач требует лишь $O(n^2)$ арифметических операций — на порядок меньше, чем приведение к верхнему треугольному виду.

Обратную матрицу можно вычислить, используя конструкции того же метода исключения. Если получено разложение $A = LU$, то, поскольку $A^{-1} = U^{-1}L^{-1}$, достаточно научиться вычислять обратные к верхней и нижней треугольным матрицам. Другой подход: искать отдельные столбцы обратной матрицы как решения систем, в которых правыми частями служат столбцы единичной матрицы. При этом, конечно. нужно использовать уже найденное LU-разложение. Общее число арифметических операций будет $O(n^3)$.

Однако, в 1969 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 $\leq c\, n^\gamma$, где $2 \leq \gamma < 3$ (нижняя оценка следует из того, что ответ зависит от n^2 независимых переменных, а верхняя оценка удовлетворяется, например, для рассмотренного в Лекции 1 алгоритма Штрассена, в котором $\gamma = \log_2 7 < 3$). 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:

$$\left[\begin{array}{cc} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{array}\right] \left[\begin{array}{cc} A_{11} & A_{12} \\ A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{cc} A_{11} & A_{12} \\ 0 & H \end{array}\right], \qquad H = A_{22} - A_{21}A_{11}^{-1}A_{12}.$$

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

$$\left[\begin{array}{cc} A_{11} & A_{12} \\ 0 & H \end{array}\right]^{-1} = \left[\begin{array}{cc} A_{11}^{-1} & -A_{11}^{-1}A_{12}H^{-1} \\ 0 & H^{-1} \end{array}\right].$$

Таким образом,

$$A^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}H^{-1} \\ 0 & H^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}.$$

Отсюда получаются выражения для блоков матрицы A^{-1} — так называемые *формулы Фробениуса*. Они показывает, как обращение матрицы порядка n сводится к двум аналогичным задачам для матриц A_{11} и H порядка $n/2$. Для реализации указанной редукции требуется выполнить несколько умножений матриц порядка $n/2$. Общие затраты на всех шагах редукции пропорциональны

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

и при $\gamma > 2$ имеют вид $O(n^\gamma)$. Если каким-то чудесным образом появится алгоритм умножения матриц, в котором $\gamma = 2$, то мы получим алгоритм обращения с числом операций $O(n^\gamma \log_2 n)$.

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.

От требования строгой регулярности можно избавиться, перейдя к обращению матрицы $B = A^\top A$. Эта матрица является строго регулярной для любой вещественной невырожденной матрицы A (докажите!). В то же время $A^{-1} = B^{-1}A^\top$.

Задача 14. *Докажите, что определитель строго регулярной матрицы порядка n можно вычислить за $O(n^{\log_2 7})$ арифметических операций.*

Алгебра и геометрия (1 поток)

Лекция 5

5.1 Геометрическое пространство и метод координат

В геометрическом пространстве, которое считается состоящим из точек, изучаются различные множества, образованные из его точек. Прежде всего это прямые, отрезки, треугольники, плоскости, тетраэдры и т.д. Идея метода координат заключается в том, что каждой точке ставится во взаимно-однозначное соответствие вектор, элементы которого называются *координатами* точки. В случае трехмерного пространства это вектор с тремя координатами. В алгебре геометрическое пространство точек рассматривается как частный случай *аффинного пространства*.

Главный факт в обосновании метода координат — это взаимно-однозначное соответствие $t \leftrightarrow P(t)$ между вещественными числами и точками прямой при фиксации произвольной пары различных точек, которым ставятся в соответствие числа 0 и 1. Число t считается координатой точки $P(t)$ на данной прямой. Отрезок $P(a)P(b)$ состоит из точек $P(t)$ при $\min(a,b) \le t \le \max(a,b)$.

Аффинная система координат в пространстве определяется тройкой прямых, проходящих через общую точку O и не лежащих в одной плоскости. Точка O называется *началом координат*, на каждой прямой ей соответствует число 0. Фиксированные прямые принято называть *осями координат*. Координаты произвольной точки пространства определяются как числа, которые соответствуют точкам осей, возникающим при их пересечении плоскостью, проходящей через заданную точку параллельно двум другим осям. Чтобы показать, что точка M имеет координаты x_1, x_2, x_3, пишут $M(x_1, x_2, x_3)$ или $M = (x_1, x_2, x_3)$.

Если углы между осями координат прямые и на каждой оси длины отрезков, концы которых соответствуют числам 0 и 1, равны единице, то система координат называется *декартовой*.

Метод координат заключается в описании геометрических свойств фигур в пространстве с помощью соотношений между координатами входящих в них точек. Очевидно, что это мощный инструмент для изучения геометрических объектов. Важно, однако, понимать, что геометрические свойства — это как раз такие свойства фигур, которые от выбора системы координат не зависят.

5.2 Направленные отрезки и свободные векторы

Направленным отрезком $\overrightarrow{AB}$ называется упорядоченная пара точек A и B.

Любой направленный отрезок порождает целый класс направленных отрезков, которые мы намерены с ним отождествить. Пусть символ "$\sim$" временно означает, что

нечто "эквивалентно" чему-то. По определению, $\overrightarrow{AB} \sim \overrightarrow{CD}$, если середина отрезка AD совпадает с серединой отрезка BC. Такие направленные отрезки будем называть *эквивалентными*. Если точки A, B, C, D не лежат на одной прямой, то эквивалентность означает, что четырехугольник $ABDC$ является параллелограммом.

Утверждение. *Из определения эквивалентности вытекают такие свойства:*

- $\overrightarrow{AB} \sim \overrightarrow{AB}$ *(рефлексивность);*
- $\overrightarrow{AB} \sim \overrightarrow{CD} \;\Rightarrow\; \overrightarrow{CD} \sim \overrightarrow{AB}$ *(симметричность);*
- $\overrightarrow{AB} \sim \overrightarrow{CD},\ \overrightarrow{CD} \sim \overrightarrow{EF} \;\Rightarrow\; \overrightarrow{AB} \sim \overrightarrow{EF}$ *(транзитивность).*

Доказательство. Рефлексивность и симметричность очевидны. Докажем транзитивность в случае, когда точки A, B, C, D не лежат на одной прямой. Рассмотрим треугольники $\triangle ADE$ и $\triangle BCF$ и заметим, что точка X пересечения отрезков AD и BC и точка Y пересечения отрезков DE и CF делят соответствующие отрезки пополам. Поэтому отрезок XY соединяет середины сторон AD и DE в треугольнике $\triangle ADE$ и середины сторон BC и CF в треугольнике $BCF \;\Rightarrow\; AE \parallel BF$. Кроме того, $AB \parallel CD$, $CD \parallel EF \;\Rightarrow\; AB \parallel EF$. Значит, четырехугольник $ABFE$ является параллелограммом $\Rightarrow\; \overrightarrow{AB} \sim \overrightarrow{EF}$. Разбор случая, когда точки A, B, C, D оказались на одной прямой, оставляем читателю. □

Если среди элементов совершенно абстрактного множества выделены пары (a, b), а запись $a \sim b$ означает присутствие такой пары, то в случае выполнения приведенных выше свойств рефлексивности, симметричности и транзитивности говорят, что на множестве задано *отношение эквивалентности*. В таких случаях все множество разбивается на непересекающиеся классы таким образом, что все элементы каждого отдельного класса эквивалентны, а элементы разных классов не являются эквивалентными. Знакомые нам примеры отношения эквивалентности — это эквивалентность матриц, параллельность прямых, равенство треугольников, равенство целых чисел по модулю и т.п.

Классы эквивалентности возникают и на множестве направленных отрезков. Обозначим через $v(\overrightarrow{AB})$ множество всех направленных отрезков, которые эквивалентны направленному отрезку $\overrightarrow{AB}$. Это множество принято называть *свободным вектором* или просто *вектором*, порожденным направленным отрезком $\overrightarrow{AB}$. Из транзитивности следует, что если $\overrightarrow{CD} \in v(\overrightarrow{AB})$, то $v(\overrightarrow{CD}) = v(\overrightarrow{AB})$ (проверьте!).

Запись $v = v(\overrightarrow{AB})$ показывает, что класс v определяется своим представителем $\overrightarrow{AB} \in v$, но при этом он определяется *любым* своим представителем. Это обстоятельство объясняет традицию при работе со свободными векторами обозначать их так же, как направленные отрезки. Таким образом, запись $\overrightarrow{AB}$ в зависимости от контекста может означать как направленный отрезок, так и порождаемый им свободный вектор.

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 $\overrightarrow{OA}$ и соответствующего свободного вектора. Свободные векторы можно складывать и умножать на вещественные числа. Эти операции соответствуют сложению и умножению на числа арифметических векторов, составленных из их координат.

5.3 Аффинные пространства

Заметим, что для любой точки C существует направленный отрезок $\overrightarrow{CD}$, порождающий заданный свободный вектор v. Фактически мы имеем операцию "сложения" точки и вектора: $C + v \to D$. Если вектор v фиксирован, то возникает взаимно-однозначное отображение геометрического пространства на себя, которое называется *параллельным переносом* или *сдвигом* точек на заданный вектор.

В общем случае под *аффинным пространством* $\mathcal{A}$, ассоциированным с векторным пространством V, подразумевается непустое множество $\mathcal{A}$ точек, на котором определена операция сложения точки $A \in \mathcal{A}$ и вектора $v \in V$, которая обладает следующими свойствами:

- $(A + u) + v = A + (u + v)$ для любой точки A и любых векторов u и v,
- если 0 – нулевой вектор, то $A + 0 = A$ для любой точки A,
- для любых точек A и B существует единственный вектор v такой, что $A + v = B$.

Под *размерностью аффинного пространства* понимается размерность ассоциированного с ним векторного пространства.

В этой книге в качестве точек аффинного пространства мы всегда используем векторы ассоциируемого с ним векторного пространства. В этом случае операция сложения точки и вектора определяется уже существующей операцией сложения векторов, а приведенные выше свойства выполняются очевидным образом.

5.4 Линейные многообразия

Множество точек, координаты которых удовлетворяют системе линейных алгебраических уравнений, называется *линейным многообразием*. Это частный случай *алгебраического многообразия*, в котором координаты точек удовлетворяют системе полиномиальных уравнений. Из теории систем линейных алгебраических уравнений мы знаем, что непустое линейное многообразие имеет вид $M = x^{(0)} + L$, где L – векторное пространство, состоящее из решений однородной системы $Ax = 0$. Оно однозначно определяется по множеству M и называется *направляющим пространством* для данного линейного многообразия.

Утверждение. *В трехмерном геометрическом пространстве непустое линейное многообразие представляет собой точку, прямую, плоскость либо совпадает со всем пространством.*

Доказательство. Если система $Ax = b$ совместна, то множество ее решений имеет вид $M = x^{(0)} + L$, где $L = \ker(A)$, а $x^{(0)}$ – фиксированное частное решение. В случае $\dim L = 0$ множество M состоит из одной точки. Если $\dim L = 3$, то M совпадает со всем пространством.

Если $\dim L = 1$ и v – ненулевой вектор из L, то получаем множество

$$M = \{x = x^{(0)} + tv, \quad t \in \mathbb{R}\},$$

состоящее из точек прямой, проходящей через точку $x^{(0)}$ параллельно вектору v, который обычно называется ее *направляющим вектором*. Равенство $x = x^{(0)} + tv$ принято называть *векторным уравнением прямой*.

Если $\dim L = 2$ и пространство L натянуто на линейно независимы векторы u и v, то множество

$$M = \{x = x^{(0)} + \alpha u + \beta v, \quad \alpha, \beta \ \in \ \mathbb{R}\}$$

состоит из точек плоскости, проходящей через точку $x^{(0)}$ параллельно векторам u и v, которые называются ее *направляющими векторами*. Равенство $x = x^{(0)} + \alpha u + \beta v$ называется *векторным уравнением плоскости*. $\square$

Векторные уравнения прямой или плоскости, записанные покоординатно, называются их *параметрическими уравнениями*:

$$x_i = x_i^{(0)} + tv_i, \quad 1 \le i \le 3 \quad \text{(прямая)}; \quad x_i = x_i^{(0)} + \alpha u_i + \beta v_i, \quad 1 \le i \le 3 \quad \text{(плоскость)}.$$

Следуя геометрической аналогии, в произвольных аффинных пространствах множества вида $M = x^{(0)} + L$, где L — некоторое векторное пространство, часто называют *плоскостями*. Пространство L однозначно определяется по множеству M и называется *направляющим пространством* плоскости M. Под *размерностью плоскости* понимается размерность ее направляющего пространства. В n-мерном пространстве плоскость размерности $n-1$ называется *гиперплоскостью*.

Задача 1. *Докажите, что любая плоскость является линейным многообразием.*

5.5 Общие уравнения прямой и плоскости

В двумерном геометрическом пространстве прямая является гиперплоскостью и задается одним уравнением, в котором хотя бы один из коэффициентов при координатах отличен от нуля. Оно обычно записывается в виде

$$A_0 + A_1 x_1 + A_2 x_2 = 0$$

и называется *общим уравнением прямой на плоскости*.

Общим уравнением можно задать *любую* прямую на плоскости. В самом деле, из параметрических уравнений общее уравнение получается исключением параметра t. Можно также заметить, что вектор $(x_1 - x_1^{(0)}, x_2 - x_2^{(0)})$ пропорционален направляющему вектору (v_1, v_2) (такие векторы иногда называют *коллинеарными*) и получить общее уравнение, раскрывая определитель

$$\begin{vmatrix} x_1 - x_1^{(0)} & v_1 \\ x_2 - x_2^{(0)} & v_2 \end{vmatrix} = 0.$$

Аналогично, уравнение вида

$$A_0 + A_1 x_2 + A_2 x_2 + A_3 x_3 = 0,$$

в котором хотя бы один из коэффициентов при координатах отличен от нуля, называется *общим уравнением плоскости в трехмерном пространстве*. Из параметрических уравнений оно получается исключением параметров α и β или раскрытием определителя третьего порядка

$$\begin{vmatrix} x_1 - x_1^{(0)} & u_1 & v_1 \\ x_2 - x_2^{(0)} & u_2 & v_2 \\ x_3 - x_3^{(0)} & u_3 & v_3 \end{vmatrix} = 0.$$

5.6 Гиперплоскости и полупространства

Утверждение 1. *В n-мерном пространстве любая гиперплоскость задается уравнением вида*

$$A_0 + A_1x_1 + \ldots + A_nx_n = 0, \qquad (*)$$

в котором хотя бы один из коэффициентов при координатах отличен от нуля.

Доказательство. Пусть $M = x^{(0)} + L$, где L – векторное пространство с базисом $p_1, \ldots, p_{n-1}$. Тогда равенство $x = x^{(0)} + \alpha_1 p_1 + \ldots + \alpha_{n-1}p_{n-1}$ равносильно условию

$$x - x^{(0)} \in L(p_1, \ldots, p_{n-1}) \quad \Leftrightarrow \quad A_0 + A_1x_1 + \ldots A_nx_n := \det[x - x^{(0)}, p_1, \ldots, p_{n-1}] = 0. \quad \square$$

Вектор $\mathbf{n} := (A_1, \ \ldots \ , \ A_n)$ называется *нормальным вектором* или *нормалью* гиперплоскости, заданной уравнением $(*)$. Вектор v называется *параллельным* гиперплоскости, если он принадлежит ее направляющему пространству.

Утверждение 2. *Вектор $v = (v_1, \ldots, v_n)$ параллелен гиперплоскости $(*)$ тогда и только тогда, когда $A_1v_1 + \ldots + A_nv_n = 0$.*

Доказательство. Достаточно учесть, что направляющее пространство совпадает с множеством решений однородной системы, соответствующей уравнению $(*)$. $\square$

Теперь положим $f(x) = f(x_1, \ldots, x_n) := A_0 + A_1x_1 + \ldots + A_nx_n$ и заметим, что гиперплоскость Π, заданная уравнением $f(x) = 0$, разделяет n-мерное пространство на два полупространства – *положительное* и *отрицательное*:

$$\Pi_+ = \{x : f(x) \geq 0\} \quad \text{и} \quad \Pi_- + \{x : f(x) \leq 0\}, \quad \Pi_+ \cap \Pi_- = \Pi.$$

Пусть точки $P \neq Q$ не принадлежат гиперплоскости Π. Говорят, что они лежат *по одну сторону* от нее, если среди точек отрезка PQ нет точек гиперплоскости. Если есть, то такие точки считаются лежащими *по разные стороны* от гиперплоскости.

Утверждение 3. *Две разные точки, не лежащие на заданной гиперплоскости, находятся по разные стороны от нее в том и только том случае, когда они принадлежат разным полупространствам.*

Доказательство. Пусть точки $p = (p_1, \ldots, p_n)$ и $q = (q_1, \ldots, q_n)$ разные и не принадлежат гиперплоскости. Проведем через них прямую, заданную векторным уравнением $x = p + t(q - p)$, и посмотрим, при каких значения параметра t она может пересекаться с гиперплоскостью. Если $f(p) = f(q)$, то прямая параллельна гиперплоскости. В случае $f(p) \neq f(q)$ находим

$$f(p + t(q - p)) \ = \ 0 \quad \Leftrightarrow \quad t = \frac{f(p)}{f(p) - f(q)}.$$

Ненулевые числа $f(p) \neq f(q)$ имеют разные знаки в том и только том случае, когда $0 < t < 1$. $\square$

Задача 2. *Докажите, что для любой конечной системы точек в n-мерном пространстве существует гиперплоскость, которая не содержит ни одну из них.*

5.7 Пересечение гиперплоскостей

Теорема о пересечении гиперплоскостей. *Плоскость размерности $0 \le k \le n-1$, расположенная в n-мерном пространстве, является пересечением $n-k$ гиперплоскостей.*

Доказательство. Пусть P – матрица размеров $n \times k$, столбцы которой образуют базис направляющего пространства k-мерной плоскости M. Тогда $\mathrm{def}(P^\top) = n-k$. Из базисных векторов-столбцов пространства решений однородной системы $P^\top x = 0$ составим $n \times (n-k)$-матрицу и обозначим через A ее транспонированную матрицу. Тогда $AP = 0$. Зафиксируем произвольный вектор $x^{(0)} \in M$. Тогда $M = \{x = x^{(0)} + Py : y \in \mathbb{R}^k\}$. Отсюда вытекает, что любой вектор $x \in M$ удовлетворяет системе линейных алгебраических уравнений $Ax = b$ с правой частью $b = Ax^{(0)}$. Понятно также, что любое решение этой системы принадлежит M (почему?). Каждое уравнение системы соответствует отдельной строке матрицы A и определяет гиперплоскость, при этом число строк матрицы A равно $n-k$. $\square$

5.8 Аффинные комбинации и выпуклые множества

Рассмотрим плоскость, проходящую через точки $x_0, x_1, \ldots, x_k$. Любая точка x на ней имеет вид $x = x_0 + \alpha_1(x_1 - x_0) + \ldots + \alpha_k(x_k - x_0)$, где $\alpha_1, \ldots, \alpha_k$ — вещественные числа. Мы полагаем, что точки аффинного пространства отождествляются с векторами ассоциированного векторного пространства, и видим, что вектор x выражается в виде линейной комбинации векторов $x_0, x_1, \ldots, x_k$:

$$x = \alpha_0 x_0 + \alpha_1 x_1 + \ldots + \alpha_k x_k, \quad \alpha_0 = 1 - \alpha_1 - \ldots - \alpha_k.$$

Такого типа линейные комбинации принято называть *аффинными комбинациями*. Вот формальное и более симметричное определение: для любой системы векторов их линейная комбинация называется *аффинной комбинацией*, если сумма всех ее коэффициентов равна единице. Множество всевозможных аффинных комбинаций всевозможных конечных систем векторов, принадлежащих множеству M, называется его *аффинной оболочкой* и обозначается $\mathrm{aff}(M)$.

Аффинная комбинация с неотрицательными коэффициентами называется *выпуклой комбинацией*. Множество всевозможных выпуклых комбинаций всевозможных конечных систем векторов из множества M называется *выпуклой оболочкой* множества M и обозначается $\mathrm{conv}(M)$. Для точек x и y множество $V(x,y) = \{v = x + t(y-x) : 0 \le t \le 1\}$ называется *отрезком* с концевыми точками x и y. При $0 < t < 1$ точки называются *внутренними*. Множество называется *выпуклым*, если вместе с любой парой точек оно содержит соединяющий их отрезок. Поскольку в качестве точек мы используем векторы, можно дать и такое определение: множество называется выпуклым, если для любой пары своих векторов оно содержит также все их *выпуклые комбинации*.

Пусть собрана коллекция *всех* множеств, обладающих некоторым свойством, и известно, что их пересечение также обладает этим свойством. Тогда пересечение называется *минимальным множеством*, которое обладает данным свойством. Заметим, что пересечение любого количества плоскостей или выпуклых множеств является соответственно плоскостью или выпуклым множеством (проверьте!).

Теорема об аффинной оболочке. *Аффинная оболочка непустого множества является минимальной плоскостью, содержащей это множество.*

Доказательство. Пусть A – аффинная оболочка множества M. Ее точки принадлежат любой плоскости, содержащей M. Поэтому нам нужно лишь понять, почему множество A является плоскостью. Если M – конечное множество векторов $v_1, \ldots, v_s$, то это очевидно:

$$A = \{\alpha_1 v_1 + \ldots + \alpha_s v_s \,:\, \alpha_1 + \ldots + \alpha_s = 1, \quad \alpha_1, \ldots, \alpha_s \in \mathbb{R}, \quad s \in \mathbb{N}\} =$$

$$\{v_1 + \alpha_2(v_2 - v_1) + \ldots + \alpha_s(v_s - v_1) \,:\, \alpha_2, \ldots, \alpha_s \in \mathbb{R}, \quad s \in \mathbb{N}\}.$$

В общем случае можно доказать, что множество $L = \{x - y \,:\, x - y \in A\}$ является векторным пространством и $A = x^{(0)} + L$, где $x^{(0)}$ – произвольный фиксированный вектор из M (проверьте!). $\square$

Теорема о выпуклой оболочке. *Выпуклая оболочка непустого множества является минимальным выпуклым множеством, содержащим это множество.*

Доказательство. Заметим, что выпуклое множество содержит любые выпуклые комбинации своих векторов. Проведем индукцию по числу векторов k. Если $k = 2$, то утверждение следует из определения выпуклого множества. Если $k \geq 3$, то для произвольной выпуклой комбинации с коэффициентом $0 < \alpha_1 < 1$ находим

$$\sum_{i=1}^{k} \alpha_i a_i = \alpha_1 a_1 + (1 - \alpha_1) b_1, \quad b_1 = \sum_{i=2}^{k} \frac{\alpha_i}{1 - \alpha_1} a_i.$$

Вектор b_1 является выпуклой комбинацией меньшего числа векторов и, согласно индуктивному предположению, принадлежит выпуклому множеству. Чтобы доказать теорему, нам достаточно установить выпуклость выпуклой оболочки множества M. Для этого рассмотрим выпуклую комбинацию произвольной пары его точек, представленных выпуклыми комбинациями конечной системы векторов из множества M:

$$\alpha \left(\sum_{i=1}^{s} \alpha_i v_i \right) + \beta \left(\sum_{i=1}^{s} \beta_i v_i \right) = \sum_{i=1}^{s} (\alpha\alpha_i + \beta\beta_i) v_i, \quad \sum_{i=1}^{s} (\alpha\alpha_i + \beta\beta_i) = 1. \quad \square$$

Задача 3. *Пусть M — непустое множество векторов и N — множество всех векторов вида $u - v$, где $u, v \in M$. Докажите, что разность любых двух векторов из выпуклой оболочки множества M принадлежит выпуклой оболочке множества N.*

5.9 Полиэдры и выпуклые многогранники

Геометрическое место точек в $\mathbb{R}^n$, координаты которых удовлетворяют конечной системе линейных неравенств, принято называть *полиэдром*. В тривиальном случае полиэдр может совпадать с $\mathbb{R}^n$ (приведите пример соответствующей системы неравенств). В нетривиальных случаях он представляет собой пересечение конечного числа полупространств в $\mathbb{R}^n$. Гиперплоскости, связанные с полупространствами, будем называть *определяющими*. Примеры: треугольники, тетраэдры и т.п.

Теорема о проекциях полиэдра. *Пусть M – полиэдр в пространстве $\mathbb{R}^n$ и пусть отображение $p : \mathbb{R}^n \to \mathbb{R}^{n-1}$ задано правилом*

$$p : (x_1, \ldots, x_{n-1}, x_n) \to (x_1, \ldots, x_{n-1}).$$

Тогда множество $p(M) := \{p(x) : x \in M\}$ является полиэдром в пространстве $\mathbb{R}^{n-1}$.

Доказательство. Пусть полиэдр M задается системой неравенств

$$\begin{gathered} a_{11}x_1 + \ldots + a_{1n}x_n \le b_1, \\ \ldots \quad \ldots \quad \ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \le b_m. \end{gathered}$$

Номера неравенств разобьем на три части: $i \in I_0$, если $a_{in} = 0$, $i \in I_+$, если $a_{in} > 0$, $i \in I_-$, если $a_{in} < 0$. Положим $\hat{x} = (x_1, \ldots, x_{n-1})$ и рассмотрим функции

$$L_i(\hat{x}) = a_{i1}x_1 + \ldots + a_{i,n-1}x_{n-1} - b_i, \quad 1 \le i \le m.$$

Исходная система неравенств может быть записана в следующем виде:

$$L_i(\hat{x}) \le 0 \ \text{ при } \ i \in I_0, \quad x_n \le -L_i(\hat{x}) \ \text{ при } \ i \in I_+, \quad L_i(\hat{x}) \le x_n \ \text{ при } i \in I_-.$$

Отсюда следует, что если $(\hat{x}, x_n) \in M$, то выполняются неравенства

$$\underbrace{\max_{i \in I_-} L_i(\hat{x})}_{=\alpha(\hat{x})} \ \le \ \underbrace{\min_{j \in I_+} L_j(\hat{x})}_{=\beta(\hat{x})}, \qquad \max_{i \in I_0} L_i(\hat{x}) \le 0. \tag{$*$}$$

Неравенства $(*)$, очевидно, определяют полиэдр в пространстве $\mathbb{R}^{n-1}$. Остается только заметить, что для каждой его точки $\hat{x}$ найдется число x_n такое, что $\alpha(\hat{x}) \le x_n \le \beta(\hat{x})$, а это означает, что $(\hat{x}, x_n) \in M$. □

Система двух и большего числа векторов называется *аффинно зависимой*, если какой-то ее вектор выражается в виде аффинной комбинации остальных векторов. Система называется *аффинно независимой*, если она не является аффинно зависимой. Аффинная независимость векторов имеет место в том и только том случае, когда из равенства нулю их линейной комбинации с нулевой суммой коэффициентов вытекает, что все коэффициенты равны нулю (проверьте!). О таких векторах также говорят, что они *находятся в общем положении*. Выпуклая оболочка конечного числа векторов называется *выпуклым многогранником*. Если система с числом векторов $k+1$ аффинно независима, то их выпуклая оболочка называется *симплексом размерности* k.

Теорема о выпуклом многограннике. *Выпуклый многогранник является полиэдром.*

Доказательство. Пусть V – выпуклая облочка столбцов $n \times k$-матрицы A. Тогда

$$V = \{Az : \ \ z = \begin{bmatrix} z_1 & \ldots & z_k \end{bmatrix}^\top, \ \ z_1 \ge 0, \ \ldots, \ z_k \ge 0, \ \ z_1 + \ldots + z_k = 1\}.$$

Теперь перейдем к пространству $\mathbb{R}^{n+k}$ и рассмотрим в нем множество

$$M = \left\{ \begin{bmatrix} x \\ z \end{bmatrix} \ : \ x = Az, \ \ z_1 \ge 0, \ \ldots, \ z_k \ge 0, \ \ z_1 + \ldots + z_k = 1 \right\}.$$

Очевидно, M является полиэдром в пространстве $\mathbb{R}^{n+k}$ (почему?). Применяя k раз теорему о проекциях полиэдра для исключения координат вектора z, мы приходим к выводу о том, что множество V в пространстве $\mathbb{R}^n$ также является полиэдром. □

Заметим, что обратное неверно: произвольный полидр не обязан быть выпуклым многогранником.

Задача 4. *Докажите, что в n-мерном пространстве любые $k \ge n+2$ векторов аффинно зависимы.*

Задача 5. *Докажите, что в n-мерном пространстве произвольное множество из $k \ge n+2$ векторов можно разбить на два непересекающихся подмножества, выпуклые оболочки которых имеют общую точку (утверждение известно как* `теорема Радона`*).*

5.10 Преобразование координат

Если $n \times n$-матрицы A и B невырождены, то их столбцы образуют два базиса пространства $\mathbb{R}^n$. Равенства $x = Au = Bv$ означают, что элементы векторов u и v являются координатами вектора x при разложении по этим двум базисам в двух системах координат с общим началом. Очевидно,

$$v = Pu, \quad u = P^{-1}v, \quad \text{где} \quad P = B^{-1}A.$$

Матрица P называется *матрицей перехода* от одной системы координат к другой.

5.11 Скалярное произведение векторов

Пусть $a = \overrightarrow{OA}$ и $b = \overrightarrow{OB}$ – векторы, порожденные направленными отрезками. Их длины $|a|$ и $|b|$ определяются как длины отрезков: $|a| = |OA|$, $|b| = |OB|$. В случае ненулевых векторов *углом* $\phi(a,\, b)$ между ними называется угол между сторонами OA и OB в треугольнике OAB. *Скалярным произведением* ненулевых векторов a и b называется число $|a||b| \cos \phi(a,\, b)$. Если хотя бы один из векторов нулевой, то скалярное произведение по определению равно нулю. Скалярное произведение обозначается с помощью круглых или угловых скобок $(a, b) = \langle a, b \rangle$.

Вот *основные свойства скалярного произведения*:

- $(a, a) \geq 0$ для любого вектора a, $(a, a) = 0 \quad \Leftrightarrow \quad a = 0$;
- $(a, b) = (b, a)$ для любых векторов a и b;
- $(a + b, c) = (a, c) + (b, c)$ для любых векторов a, b, c;,
- $(\alpha a, c) = \alpha(a, c)$ для любых векторов a, c и любого вещественного числа α.

Последние два свойства доказываются следующим образом. Если $c = 0$, то они очевидны. Если $c = \overrightarrow{OC} \neq 0$, то рассматриваем такую декартову систему координат с началом в точке O, в которой точка C лежит на первой оси и имеет положительную координату. Тогда для любого вектора v величина $|v| \cos \phi(v,\, c)$ является координатой вектора v на первой оси. Остается принять во внимание то, что при сложении векторов координаты складываются, а при умножении на число умножаются на это число.

Векторы называются *ортогональными*, если их скалярное произведение равно нулю. Пусть векторы e_1, e_2, e_3 имеют единичную длину и попарно ортогональны. Тогда из вышеприведенных свойств легко выводится равенство

$$(x_1 e_1 + x_2 e_2 + x_3 e_3,\, y_1 e_1 + y_2 e_2 + y_3 e_3) = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

Оно подсказывает, как можно ввести скалярное произведение произвольных арифметических векторов $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$. Если его определить формулой

$$(x, y) := x_1 y_1 + \ldots + x_n y_n,$$

то перечисленные выше основные свойства будут очевидно выполнены.

5.12 Длины и углы

Для геометрических векторов скалярное произведение определялось на основе таких понятий, как длина вектора и угол между векторами. В то же время длина и угол легко выражаются через скалярное произведение:

$$|a| = \sqrt{(a,\, a)}, \qquad \cos \phi(a,\, b) = \frac{(a,\, b)}{|a|\ |b|}. \qquad (*)$$

В произвольных векторных пространствах скалярные произведения определяются аксиоматически – как функции векторов a и b, обладающие отмеченными выше свойствами. Длина и угол определяются по формулам $(*)$. При этом для корректного определения угла важно иметь неравенство

$$|(a,\, b)| \leq |a||b|,$$

известное как *неравенство Коши–Буняковского–Шварца.*

Утверждение. *Для любой функции (a, b) векторов a и b, обладающей основными свойствами скалярного произведения, выполняется неравенство Коши–Буняковского–Шварца. Равенство имеет место в том и только том случае, когда векторы a и b линейно зависимы.*

Доказательство. Предположим, что $b \neq 0$. Тогда квадратный трехчлен

$$f(t) = (a + tb, a + tb) = (b, b)t^2 + 2(a, b)t + (a, a)$$

неотрицателен при всех вещественных значениях t. Значит, его дискриминант неположителен: $(a, b)^2 - (a, a)(b, b) \leq 0$. Дискриминант равен нулю в том и только том случае, когда $f(t) = 0$ при некотором t. $\square$

5.13 Расстояние от точки до гиперплоскости

Пусть $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$. Будем считать, что скалярное произведение задано формулой $(x, y) = x_1y_1 + \ldots + x_ny_n$.

Как найти расстояние $\rho(z, \Pi)$ от точки $z = (z_1, \ldots, z)$ до гиперплоскости Π, заданной уравнением $f(x) := A_0 + A_1x_1 + \ldots + A_nx_n = 0$? В силу специального вида скалярного произведения, нормальный вектор $\mathbf{n} = (A_1, \ldots, A_n)$ гиперплоскости Π будет ортогонален ее направляющему пространству. Следовательно, прямая с векторным уравнением $x = z + t\mathbf{n}$ будет ортогональна плоскости Π. Пусть $h = (h_1, \ldots, h_n)$ – точка пересечения этой прямой с гиперплоскостью Π. Тогда $\rho(z, \Pi) = |z - h|$. Условие $z + t\mathbf{n} \in \Pi$ приводит к следующему уравнению относительно t: $f(z + t\mathbf{n}) = 0 \quad \Leftrightarrow \quad t = f(z)/|\mathbf{n}|^2$. В итоге

$$\rho(z, \Pi) \;=\; \frac{|f(z)|}{|\mathbf{n}|} \;=\; \frac{|A_0 + A_1z_1 + \ldots + A_nz_n|}{\sqrt{A_1^2 + \ldots + A_n^2}}.$$

5.14 Ориентация системы векторов

Понятие ориентации для тройки линейно независимых векторов трехмерного геометрического пространства часто вводится в буквальном смысле “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, и здесь остается чувство неудовлетворения. Но оно имеет неустранимый характер — в силу фундаментальных причин. Дело в том, что любые тройки линейно независимых векторов могут иметь ровно два типа ориентации, а фиксация одного из них, вообще говоря, произвольна. Можно выбрать произвольную декартову систему координат и объявить, что тройка ее базисных векторов имеет, скажем, “правильную

[1] Если указательный палец согнуть, то получится совсем не то.

ориентацию". Пусть $a = (a_1,\ a_2,\ a_3)$, $b = (b_1,\ b_2,\ b_3)$, $c = (c_1,\ c_2,\ c_3)$. Тройку $a,\ b,\ c$ можно назвать тройкой *положительной ориентации*, если

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} > 0.$$

Если определитель меньше нуля, то это будет тройка *отрицательной ориентации*. Таким образом, определение ориентации зависит от объявления типа ориентации для исходной системы координат.

5.15 Векторное произведение векторов

Пусть a и b – линейно независимые векторы трехмерного геометрического пространства. Их *векторным произведением* называется вектор c, имеющий длину $|c| = |a||b|\sin\phi(a,b)$, ортогональный векторам a и b и направленный таким образом, что тройка a, b, c имеет положительную ориентацию. Если векторы a и b линейно зависимы, то по определению $c = 0$. Обозначение: $c = [a,\ b]$. Число $(a, b, c) := ([a,\ b], c)$ называется *смешанным произведением* векторов a, b, c.

Теорема о смешанном произведении. *Пусть радиус-векторы a, b, c линейно независимы и V — объем натянутого на них параллелепипеда. Тогда смешанное произведение (a, b, c) равно $\pm V$, знак плюс берется, если тройка a, b, c ориентирована положительно, и минус, если отрицательно.*

Доказательство. Пусть $a = \overrightarrow{OA},\ b = \overrightarrow{OB},\ c = \overrightarrow{OC}$, и пусть $\overrightarrow{OD} = [a,\ b]$. Согласно определению смешанного произведения, $(a,\ b,\ c) = |\overrightarrow{OD}|\ h$, где $h = |\overrightarrow{OC}|\cos\phi(\overrightarrow{OD},\ \overrightarrow{OC})$. Ясно, что $|h|$ есть длина перпендикуляра, опущенного из точки C на плоскость OAB (высота параллелепипеда). При этом $h > 0$, если точки D и C находятся по одну сторону от плоскости OAB, и $h < 0$, если по разные стороны. В первом случае тройка векторов $\overrightarrow{OA},\ \overrightarrow{OB},\ \overrightarrow{OC}$ ориентирована положительно, во втором — отрицательно. □

Следствие. *Смешанное произведение (a, b, c) линейно по каждому аргументу.*

Доказательство. Достаточно заметить, что $(a,\ b,\ c) = (b,\ c,\ a) = (c,\ a,\ b) = -(b,\ a,\ c) = -(a,\ c,\ b) = -(c,\ b,\ a)$. Первые три тройки ориентированы одинаково, а остальные три тройки имеют противоположную ориентацию. □

Теорема о векторном произведении. *Векторное произведение является векторнозначной функцией, линейной по каждому аргументу и меняющей знак при перестановке аргументов.*

Доказательство. Докажем, что $[a+b,\ c] = [a,\ c]+[b,\ c]$. Для этого рассмотрим вектор $d = [a+b,\ c]-[a,\ c]-[b,\ c]$. Используя линейность смешанного произведения по каждому аргументу, находим $(d,\ d) = (a+d,\ c,\ d)-(a,\ c,\ d)-(b,\ c,\ d) = 0 \ \Rightarrow\ d = 0$. Аналогично, если $d = [\alpha\, a,\ b]-\alpha\,[a,\ b]$, то $(d,\ d) = (\alpha\, a,\ b,\ d)-\alpha\,(a,\ b, d) = 0 \ \Rightarrow\ d = 0$. Свойство знакопеременности вытекает непосредственно из определения векторного произведения. Оно влечет за собой и свойство линейности по второму аргументу. □

5.16 Формулы в координатах

Пусть $e_1,\ e_2,\ e_3$ — положительно ориентированная тройка базисных векторов декартовой системы координат. Легко проверить, что $[e_1,\, e_2] = e_3$, $[e_2,\, e_3] = e_1$ и $[e_3,\, e_1] = e_2$. Для векторов $a = a_1e_1 + a_2e_2 + a_3e_3$ и $b = b_1e_1 + b_2e_2 + b_3e_3$ получаем

$$\begin{aligned}
[a,\, b] &= a_1b_1[e_1,\, e_1] + a_1b_2[e_1,\, e_2] + a_1b_3[e_1, e_3] \\
&+ a_2b_1[e_2,\, e_1] + a_2b_2[e_2,\, e_2] + a_2b_3[e_2, e_3] \\
&+ a_3b_1[e_3,\, e_1] + a_3b_2[e_3,\, e_2] + a_3b_3[e_3, e_3] \\
&= (a_2b_3 - a_3b_2)\, e_1 \,-\, (a_1b_3 - a_3b_1)\, e_2 \,+\, (a_1b_2 - a_2b_1)\, e_3.
\end{aligned}$$

Полученную формулу легче всего запомнить, увидев в ней применение теоремы Лапласа для разложения по первой строке следующего определителя третьего порядка:

$$[a,\, b] \;=\; \begin{vmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Теперь рассмотрим смешанное произведение (a, b, c) с вектором $c = c_1e_1 + c_2e_2 + c_3e_3$.

Применяя правило вычисления скалярного произведения в декартовой системе координат и теорему Лапласа о разложении определителя по последнему столбцу, находим

$$(a, b, c) = ([a, b], c) =$$

$$((a_2b_3 - a_3b_2)\, e_1 \,-\, (a_1b_3 - a_3b_1)\, e_2 \,+\, (a_1b_2 - a_2b_1)\, e_3,\;\; c_1e_1 + c_2e_2 + c_3e_3) =$$

$$c_1(a_2b_3 - a_3b_2) \,-\, c_2(a_1b_3 - a_3b_1) \,+\, c_3(a_1b_2 - a_2b_1) = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

Следствие. *Определитель по абсолютной величине — это объем параллелепипеда, натянутого на векторы, определяемые его столбцами или строками.*

Замечание. Вывод о том, что величина (a, b, c) совпадает с определителем, можно сделать и без вычислений: надо лишь заметить, что это есть полилинейная функция со свойством обнуления при совпадении пары аргументов и, кроме того, $(e_1, e_2, e_3) = 1$.

Задача 6. *Докажите, что для произвольных геометрических векторов a, b, c выполняется равенство $[a,\, [b,\, c]] \;=\; (a, c)\, b \,-\, (a, b)\, c$.*

Задача 7. *Докажите, что уравнение $[a, x] + [b, x] = [a, b]$ имеет решение для любых векторов a, b и найдите все решения.*

Задача 8. *Докажите тождество $[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$.*

Алгебра и геометрия (1 поток)

Лекция 6

6.1 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 $\varnothing$. По определению, $\varnothing \subset M \ \ \forall\, M$.

Следует иметь в виду, что при описании множеств как совокопности элементов, обладающих неким общим свойством, иногда возникают логические противоречия. Например, что можно сказать о множестве M всех множеств, которые не содержат себя в качестве элемента? Допустим, что оно существует. Тогда верно ли, что $M \notin M$? Если верно, то M должно быть элементом множества M, и мы получаем то, что известно как *парадокс Рассела*: $M \notin M \implies M \in M$. Еще один забавный пример: 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, и 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 задаются перечислением своих элементов. Например, $M = \{1, 2, 3\}$ — множество, состоящее из трех чисел 1, 2, 3. Отметим также, что 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.

6.2 Отображения, функции, операторы

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

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

называемого *графиком* отображения (функции, оператора) f. Запись $f : \ X \to Y$ обычно означает, что отображение f определено для каждого элемента множества X, а его значения принадлежат множеству Y.

Элемент $y = f(x)$ называется *образом* элемента x, а x — *прообразом* элемента y. Множество $f(X) := \{y : \ y = f(x)$ для некоторого $x \in X\}$ называется *образом* (множеством значений) отображения f. Если $M \subset Y$, то множество $f^{-1}(M) := \{x : f(x) \in M\}$ называется *полным прообразом* множества M.

Отображение $f : \ X \to Y$ называется *сюръективным* (или отображением X *на* множество Y), если $f(X) = Y$, и *инъективным* (или *вложением*), если разным элементам из X соответствуют разные образы. Отображение f называется *взаимно-однозначным*, если оно одновременно сюръективно и инъективно – в этом и только этом случае для любого $y \in Y$ полный прообраз $f^{-1}(\{y\})$ состоит ровно из одного элемента x, такого что $f(x) = y$. В таких случаях отображение $f^{-1} : y \to x$ называется *обратным* к отображению $f : x \to y$. Можно дать и такое определение: отображение называется *обратимым*, если существует отображение $g : \ Y \to X$ такое, что $f(g(y)) = y \ \forall \ y \in Y$ и $g(f(x)) = x \ \forall \ x \in X$. Легко показать, что обратимость равносильна взаимной однозначности.

6.3 Бинарные операции

Отображение $f : \ X \times X \to Y$ называется *бинарной операцией* на X. Пусть для обозначения такой операции используется символ $*$. Тогда запись $c = a * b$ означает, что $(a, b) \in X \times X$ и $c = f((a, b)) \in Y$. Если $Y = X$, то бинарная операция называется *замкнутой* или *внутренней*.

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.

6.4 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, $\ldots$, $x_{n-1}x_n$. 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 u и v получаются при разных расстановках скобок. В любом случае имеем

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

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

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

6.5 Группы

Непустое множество G с бинарной операцией $*$ называется *группой*, если:

(1) $a * b \in G$ для любых $a, b \in G$ (замкнутость);

(2) $a * (b * c) = (a * b) * c$ для любых $a, b, c \in G$ (ассоциативность);

(3) существует элемент $e \in G$ такой, что $a * e = e * a = a$ для каждого $a \in G$;

(4) для любого элемента $a \in G$ существует элемент $b \in G$ такой, что $a * b = b * a = e$.

Элемент e называется *нейтральным* и определяется свойством (3) однозначно: если e_1 и e_2 — два таких элемента, то $e_1 = e_1 * e_2 = e_2$. Элемент b из свойства (4) однозначно определяется по a и называется *обратным* (или *противоположным*, если операция обозначается символом $+$) к элементу a: если b_1 и b_2 — два таких элемента, то $b_1 = b_1 * (a * b_2) = (b_1 * a) * b_2 = b_2$.

Для любых фиксированных $a, b \in G$ можно рассмотреть уравнение $a * x = b$ относительно x и уравнение $y * a = b$ относительно y. В любой группе оба уравнения однозначно разрешимы. Более того, при наличии замкнутой ассоциативной операции множество является группой тогда и только тогда, когда все такие уравнения разрешимы (докажите!).

Группа называется *абелевой* (коммутативной), если $a * b = b * a$ для всех $a, b \in G$. Хорошо знакомый школьный пример абелевой группы – это множество целых чисел с операцией сложения. В качестве примера неабелевой группы можно взять множество всех взаимно-однозначных отображений множества на себя с операцией композиции (последовательного выполнения) отображений – при условии что множество содержит не менее трех элементов.

6.6 Конечные группы

Группа G с конечным числом элементов называется *конечной*. Число элементов называется *порядком* конечной группы и обозначается $|G|$. Вот три примера.

Группа $\mathbb{Z}_n$ вычетов по модулю n. Пусть n – натуральное число. Целые числа a и b называются *сравнимыми по модулю* n, если $a-b$ делится на n. Множество всех целых чисел, сравнимых с числом a по модулю n, называется вычетом по модулю n, порожденным числом a, и обозначается через $[a]_n$. Сложение вычетов определяется следующим образом: $[a]_n + [b]_n := [a+b]_n$.

Определение такого рода требует проверки *корректности* – нужно доказать, что если $[a_1]_n = [a]_n$ и $[b_1]_n = [b]_n$, то $[a_1+b_1]_n = [a+b]_n$. В нашем случае это, к счастью, очень легко: число $(a_1+b_1)-(a+b) = (a_1-a)+(b_1-b)$ делится на n, так как является суммой чисел, делящихся на n. Нулевым элементом является вычет $[0]_n$, противоположным элементом для вычета $[a]_n$ является вычет $[-a]_n$. Нетрудно понять, что $|\mathbb{Z}_n| = n$.

Группа подстановок степени n. Пусть $X = \{1, 2, \ldots, n\}$ и S_n – множество всех взаимно-однозначных отображений $\sigma : X \to X$. Они обычно называются подстановками или перестановками, поскольку задаются таблицами вида

$$\sigma = \begin{pmatrix} i_1 & i_2 & \ldots & i_n \\ \sigma(i_1) & \sigma(i_2) & \ldots & \sigma(i_n) \end{pmatrix}, \quad \{i_1, i_2, \ldots, i_n\} = \{1, 2, \ldots, n\}$$

В качестве операции рассматривается композиция отображений: если $a, b \in S_n$, то произведение $c = ab$ определяется правилом $c(i) := a(b(i))$. Нетрудно проверить, что $|S_n| = n!$. Группа S_n называется *симметрической группой степени* n. Множество четных подстановок степени n образуют группу, которая является частью группы S_n (подгруппой) и называется *знакопеременной группой степени* n и обозначается A_n.

Группа вращений куба. Рассмотрим всевозможные прямые, проходящие через общую точку, и преобразования трехмерного пространства, представляющие собой вращения вокруг них на всевозможные заданные углы. Композиция (последовательное выполнение) таких преобразований будет также вращением вокруг некоторой прямой и все они образуют группу (докажите!). Теперь рассмотрим прямые, проходящие через центр куба, и выберем только такие вращения, которые переводят куб в себя. Это и будет *группа вращений куба*. В ней ровно 24 элемента (попробуйте это доказать и найти соответствующие вращения).

Задача 1. *Даны две системы чисел $x_1 \le x_2 \le \ldots \le x_n$ и $y_1 \le y_2 \le \ldots \le y_n$. Доказать, что для любой подстановки $\sigma \in S_n$ выполняется неравенство $\sum\limits_{i=1}^{n} |x_i - y_i| \le \sum\limits_{i=1}^{n} |x_i - y_{\sigma(i)}|$.*

Задача 2. *Докажите, что в группе вращений куба ровно 24 элемента.*

Задача 3. *Сколько элементов в группе вращений правильного тетраэдра?*

Задача 4. *Докажите, что в группах вращений правильного икосаэдра и правильного додекаэдра ровно 60 элементов.* [2]

[2] Правильный икосаэдр и правильный додекаэдр— это выпуклые многогранники. Икосаэдр имеет 20 граней в виде равносторонних треугольников, 30 ребер и 12 вершин, в каждой из которых сходятся 5 ребер. Додекаэдр имеет 12 граней в виде правильных пятиугольников, 30 ребер и 20 вершин, в каждой из которых сходятся 3 ребра.

6.7 Мультипликативные и аддитивные группы

Группа называется *мультипликативной*, если ее операция называется умножением и обозначается так же, как умножение чисел. Нейтральный элемент мультипликативной группы называется *единичным* и нередко обозначается символом 1. Обратный элемент к a обозначается через a^{-1}. Если n – натуральное число, то, по определению,

$$a^n := \underbrace{a * \ldots * a}_{n \text{ раз}}, \quad a^0 := 1, \quad a^{-n} := (a^{-1})^n.$$

Для любых целых чисел m и n имеет место тождество $a^{m+n} = a^m a^n$ (проверьте!).

Группа называется *аддитивной*, если ее операция называется сложением и обозначается так же, как сложение чисел. Нейтральный элемент аддитивной группы называется *нулевым* и обозначается символом 0, а обратный элемент для a называется *противоположным* и обозначается через $-a$. Вместо степеней a^n в такой группе рассматривают кратные na. По определению,

$$na := \underbrace{a + \ldots + a}_{n \text{ раз}}, \quad 0\,a := 0, \quad (-n)a := n(-a).$$

Обычно аддитивная группа считается абелевой. Поэтому для большей общности в дальнейшем мы будем рассматривать преимущественно мультипликативные группы.

6.8 Подгруппы и смежные классы

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 $a, b \in H$ и $a^{-1} \in H$ для любого элемента $a \in H$.

Пример 1. Множества целых чисел, кратных одному фиксированному натуральному числу, являются подгруппами аддитивной группы всех целых чисел.

Пример 2. Пусть G — мультипликативная группа и a — ее элемент. Множество $Z(a)$ всех элементов $g \in G$ таких, что $ga = ag$, называется *централизатором* элемента a и является подгруппой группы G. Множество $Z = Z(G)$ всех элементов группы G, коммутирующих с каждым ее элементом, называется *центром* группы G и является подгруппой любого централизатора $Z(a)$ и, конечно, подгруппой самой группы G.

Пусть $a \in G$. Тогда множества $Ha := \{ha : h \in H\}$ и $aH := \{ah : h \in H\}$ называются *правым и левым смежными классами* группы G по подгруппе H. Если подгруппа H конечна и ее порядок равен m, то каждый смежный класс по H состоит из одного и того числа элементов, равного m (проверьте!).

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

Доказательство. Пусть H – подгруппа конечной группы G. Заметим, что смежные классы Ha и Hb либо совпадают, либо не пересекаются: если $c \in Ha \bigcap Hb$, то $Hc = Ha = Hb$. Значит, конечная группа G разбивается на непересекающиеся подмножества (смежные классы по подгруппе H), в каждом их которых одинаковое число элементов. □

Число смежных классов по подгруппе H группы G называется ее *индексом* в содержащей ее группе G. В случае бесконечной группы G индекс подгруппы может быть как бесконечным, так и конечным. Например, индекс подгруппы четных чисел в аддитивной группе целых чисел равен 2.

Задача 5. *Докажите, что для любой группы и любой ее погруппы число левых смежных классов всегда равно числу правых смежных классов.*

Задача 6. *Найдите все подгруппы аддитивной группы целых чисел.*

6.9 Циклические группы

Пусть G – мультипликативная группа и $a \in G$. Обозначим через $\langle a \rangle$ множество всех целых степеней элемента a:

$$\langle a \rangle := \{a^n : n \in \mathbb{Z}\}.$$

Нетрудно проверить, что это множество является подгруппой – причем минимальной подгруппой, содержащей элемент a. Любая подгруппа такого вида называется *циклической*.

Элемент a называется *элементом конечного порядка* или *элементом кручения*, если группа $\langle a \rangle$ конечна. *Порядком* элемента a называется порядок этой группы. Заметим, что это есть целое положительное k такое, что $a^k = 1$ и при этом $a^l \neq 1$ при $1 \leq l \leq k-1$. Если $a^k \neq 1$ при всех $k > 0$, то a называется элементом *бесконечного порядка*. Если единственным элементом кручения является единичный элемент, то группа называется *группой без кручения*. В противном случае группа называется *группой с кручением*.

Теорема. *Любая подгруппа циклической группы является циклической.*

Доказательство. Пусть подгруппа имеет вид $H = \{a^{i_1}, a^{i_2}, \dots\}$, и 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 7. *Докажите, что группа простого порядка является циклической.*

Задача 8. *Докажите, что в циклической группе порядка n для любого делителя d числа n существует ровно одна подгруппа порядка d.*

Задача 9. *В бесконечной циклической группе найдите число подгрупп, индекс которых ограничен заданным натуральным числом n.*

6.10 Гомоморфизмы и изоморфизмы

Рассмотрим группы G и $\widehat{G}$ с операциями $*$ и $\widehat{*}$ и отображение $f : G \to \widehat{G}$, обладающее свойством *сохранения операции*:

$$f(a * b) = f(a) \,\widehat{*}\, f(b) \quad \forall\, a, b \in G.$$

Любое такое отображение называется также *гомоморфизмом* группы G в группу $\widehat{G}$.

Взаимно-однозначный гомоморфизм $f : G \to \widehat{G}$ называется *изоморфизмом*. В этом случае обратное отображение $f^{-1} : \widehat{G} \to G$ также будет изоморфизмом, группы G и $\widehat{G}$ называются *изоморфными* и применяется обозначение $G \cong \widehat{G}$. Несмотря на формальные различия в определении элементов и операций, изоморфные группы можно считать одинаковыми с точки зрения свойств операций.

Изоморфизм группы на себя называется *автоморфизмом*. Для определенности будем рассматривать мультипликативные группы. Отображение $g \to aga^{-1}$ при фиксированном элементе $a \in G$ является автоморфизмом (докажите!), который называется *внутренним автоморфизмом* данной группы.

Теорема. *Любые циклические группы одного порядка изоморфны.*

Доказательство. Пусть $G = \langle a \rangle$ и $\widehat{G} = \langle \widehat{a} \rangle$. Формально мы можем рассмотреть отображение $f(a^n) := \widehat{a}^n$. В случае бесконечных групп это и есть искомый изоморфизм. Однако, если группы конечны, то возможно равенство $a^k = a^l$ при $k \neq l$, а из определения не видно, повлечет ли оно за собой равенство образов. В общем случае этого может и не случиться. Но при равенстве порядков $|G| = |\hat{G}| = n$ находим

$$a^k = a^l \quad \Rightarrow \quad k - l \vdots n \quad \Rightarrow \quad \widehat{a}^k = \widehat{a}^l,$$

то есть, отображение f определено корректно, взаимно-однозначно и сохраняет операцию. $\square$

Примеры гомоморфизмов.

- Отображение, переводящее любой элемент группы G в единицу группы $\widehat{G}$.
- Отображение мультипликативной группы S_n в аддитивную группу $\mathbb{Z}_2$, переводящее каждую четную подстановку в вычет $[0]_2$, а нечетную — в вычет $[1]_2$.
- Отображение $A \to |A|$ мультипликативной группы обратимых вещественных матриц порядка n в мультипликативную группу ненулевых вещественных чисел (определитель произведения матриц равен произведению их определителей).
- Отображение $\sigma \to P_\sigma$ группы S_n в мультипликативную группу обратимых матриц порядка n, переводящее перестановку σ в *матрицу перестановки* P_σ — в каждом ее столбце все элементы нулевые, кроме одного, который равен единице и расположен в позиции (i, j), где $i = \sigma(j)$.

Пример изоморфных групп. Аддитивная группа вещественных чисел изоморфна мультипликативной группе положительных вещественных чисел. В качестве изоморфизма можно взять, например, экспоненту $f(x) = \mathrm{e}^x$.

Пример неизоморфных групп. Аддитивная группа вещественных чисел неизоморфна мультипликативной группе ненулевых вещественных чисел (докажите!).

Задача 10. *Докажите, что множество автоморфизмов группы образует группу относительно операции композиции автоморфизмов.*

Задача 11. *Найдите все автоморфизмы группы целых чисел по сложению.*

Задача 12. *Приведите пример автоморфизма группы, который не является внутренним.*

Задача 13. *Приведите пример группы, имеющей бесконечно много автоморфизмов.*

Задача 14. *Докажите, что число внутренних автоморфизмов конечной группы равно числу смежных классов этой группы по ее центру.*

6.11 Нормальные подгруппы и фактор-группы

Пусть G – мультипликативная группа, H – ее подгруппа и P – множество всех правых смежных классов группы G по подгруппе H. Попробуем на P ввести операцию умножения, используя правило $Ha \cdot Hb := H(ab)$. Будет ли такое определение корректным?

Нам нужно, чтобы из равенств $Ha_1 = Ha$, $Hb_1 = Hb$ вытекало равенство $H(a_1b_1) = H(ab)$. Другими словами, если $a_1a^{-1}, b_1b^{-1} \in H$, то $z := (a_1b_1)(ab)^{-1} \in H$. Заметим, что $z = a_1ha^{-1}$, где $h = b_1b^{-1} \in H$. Поскольку $aa_1^{-1} \in H$, находим $(aa_1^{-1})z = aha^{-1} \in H$. Таким образом, мы получаем некоторое *условие на подгруппу* H, а именно:

$$aha^{-1} \in H \quad \forall\, a \in G, \quad \forall\, h \in H.$$

Такие подгруппы называются *нормальными подгруппами* или *нормальными делителями* группы G.

Элементы a и b группы G называются *сопряженными*, если $a = gbg^{-1}$ для некоторого элемента $g \in G$. Таким образом, подгруппа H группы G называется нормальной, если вместе с каждым своим элементом она содержит также все сопряженные к нему в группе G. Нормальность подгруппы H равносильна выполнению равенств $Ha = aH$ для всех $a \in G$ (проверьте!).

Главное следствие нормальности — это корректность операции умножения смежных классов (правых или левых – теперь неважно, так как $Ha = aH$ в силу нормальности). Эта операция превращает множество P в группу, которая называется *фактор-группой* группы G по нормальной подгруппе H и обозначается $P = G/H$.

Задача 15. *Докажите, что любая подгруппа, для которой имеются ровно два смежных класса, является нормальной.*

Задача 16. *Пусть G — мультипликативная группа и H — множество всех ее элементов, которые являются произведениями конечного числа* коммутаторов, *т.е. элементов вида $aba^{-1}b^{-1}$ для всевозможных элементов $a, b \in G$. Докажите, что H является нормальной подгруппой группы G (эта подгруппа называется* коммутантом *группы G).*

Задача 17. *Пусть H — произвольная подгруппа мультипликативной группы G и $N = N(H)$ — множество всех элементов $g \in G$ таких, что $ghg^{-1} \in H\ \forall\ h \in H$. Докажите, что N является подгруппой группы G, а группа H является нормальной подгруппой группы N (по этой причине N называется* `нормализатором` *подгруппы H в группе G).*

Задача 18. *Докажите, что в любой абелевой группе число подгрупп, содержащих фиксированную подгруппу конечного индекса, конечно.*

6.12 Сопряженные элементы конечной группы

Напомним, что элементы a и b мультипликативной группы G называются сопряженными, если хотя бы для одного элемента $g \in G$ выполняется соотношение $a = gbg^{-1}$. Бинарное отношение "a сопряжен b" является отношением эквивалентности, и поэтому все множество элементов группы разбивается на непересекающиеся классы сопряженности — все элементы одного класса сопряжены между собой и не сопряжены никакому элементу из другого класса. Отдельный класс сопряженности определяется любым своим представителем и состоит из всех сопряженных ему элементов группы G. Для класса сопряженности, в который входит элемент $a \in G$, примем обозначение $C(a)$.

Лемма о числе сопряженных элементов. *Пусть G — конечная группа, a — ее элемент и $Z(a) = \{g \in G : ga = ag\}$ — централизатор элемента a. Тогда число элементов класса сопряженности $C(a)$ равно $|G|/|Z(a)|$.*

Доказательство. Каждый элемент b, сопряженный элементу a, имеет вид $b = gag^{-1}$, т.е. определяется некоторым элементом $g \in G$. При этом возможно, что разные элементы g_1 и g_2 задают один и тот же сопряженный элемент $b = g_1ag_1^{-1} = g_2ag_2^{-1} \quad \Leftrightarrow \quad (g_2^{-1}g_1)a = a(g_1^{-1}g_1) \quad \Leftrightarrow \quad g_2^{-1}g_1 \in Z(a) \quad \Leftrightarrow \quad g_1Z(a) = g_2Z(a)$. Таким образом, g_1 и g_2 опредеделяют один и тот же сопряженный элемент в том и только том случае, когда они принадлежат одному и тому же левому смежному классу группы G по подгруппе $Z(a)$. Значит, число разных элементов группы G, сопряженных одному и тому элементу a, равно числу этих смежных классов. Остается вспомнить теорему Лагранжа. $\square$

Если p — простое число, то группа порядка p^l называется *примарной группой* или *p-группой*.

Теорема о центре примарной группы. *Число элементов центра p-группы делится на p.*

Доказательство. Пусть центр группы G состоит из элементов $a_1 = e$, a_2, …, a_k. Заметим, что $C(a_1) = \{a_1\}$, …, $C(a_k) = \{a_k\}$ — это разные классы сопряженности, содержащие ровно по одному элементу. Пусть $C(b_1), \ldots, C(b_l)$ — другие непересекающиеся классы сопряженности, в которых число элементов не меньше двух. Согласно лемме о числе сопряженных элементов, каждый класс $C(b_j)$ содержит ровно $|G|/|Z(b_j)|$ элементов. Поскольку $|G| = p^l$, число элементов класса $C(b_j)$ делится на p. Множества, входящие в разбиение $G = C(a_1) \cup \ldots C(a_k) \cup C(b_1) \cup \ldots \cup C(b_l)$, не пересекаются $\Rightarrow$ k делится p. $\square$

Задача 19. *Докажите, что в любой примарной группе есть отличный от единичного элемент, который коммутирует с каждым элементом данной группы.*

Задача 20. *Докажите, что в любой p-группе есть нормальная подгруппа порядка p.*

6.13 Теорема о гомоморфизме

С каждым гомоморфизмом $f : G \to \widehat{G}$ связываются два множества: его *ядро* $H = \ker f$, состоящее из всех элементов группы G, отображаемых в единицу группы $\widehat{G}$, и *образ* $f(G)$, состоящий из всех элементов группы $\widehat{G}$, в которые отображается хотя бы один элемент группы G.

Утверждение. *Любой гомоморфизм $f : G \to \widehat{G}$ переводит единицу группы G в единицу группы $\widehat{G}$, а обратный к $a \in G$ в обратный к $f(a) \in \widehat{G}$.*

Доказательство.

- $f(1) = f(1 \cdot 1) = f(1)f(1) \;\Rightarrow\; f(1) = 1$.
- $1 = f(1) = f(aa^{-1}) = f(a)f(a^{-1}) \;\Rightarrow\; f(a^{-1}) = (f(a))^{-1}$.

Теорема о гомоморфизме. *Пусть $f : G \to \widehat{G}$ — гомоморфизм группы G в группу $\widehat{G}$. Тогда его ядро $H = \ker f$ является нормальной подгруппой в G, а образ $f(G)$ является подгруппой в $\widehat{G}$, изоморфной фактор-группе G/H.*

Доказательство.

- Если $f(a) = f(b) = 1$, то, в силу сохранения операции, $f(ab) = f(a)f(b) = 1$. Кроме того, $f(a^{-1}) = (f(a))^{-1} = 1$. Проверка нормальности: если $f(h) = 1$, то $f(ghg^{-1}) = f(g)f(h)f(g^{-1}) = f(g)f(g^{-1}) = f(gg^{-1}) = f(1) = 1$.
- $f(a)f(b) = f(ab),\ (f(a))^{-1} = f(a^{-1}) \ \Rightarrow\ f(a)f(b),\ f(a^{-1}) \in f(G)$.
- Рассмотрим отображение $\phi : G/H \to f(G)$, определенное следующим образом:

  $$\phi(aH) = f(a), \quad a \in H.$$

  Проверим корректность: $aH = a_1H \ \Rightarrow\ a^{-1}a_1 \in H \ \Rightarrow\ f(a^{-1}a_1) = 1 \ \Rightarrow\ f(a) = f(a_1)$. Обратно, если $f(a_1) = f(a)$, то $f(a_1a^{-1}) = 1 \ \Rightarrow\ a_1a^{-1} \in H$. Таким образом, отображение определено корректно, взаимно-однозначно и, как легко видеть, сохраняет операцию:

  $$\Phi((Ha)(Hb)) = \Phi(H(ab)) = f(ab) = f(a)f(b) = \phi(Ha)\phi(Hb). \qquad \square$$

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

6.14 Отображение Кэли

Пусть G – мультипликативная группа с элементами $g_1, \ldots, g_n$ и S_n – группа подстановок степени n. Отображение $f : G \to S_n$, заданное правилом

$$a \to f_a = \begin{pmatrix} g_1 & \cdots & g_n \\ ag_1 & \ldots & ag_n \end{pmatrix},$$

называется *отображением Кэли*. Согласно определению, $f_a(g_i) = ag_i$.

Теорема. *Отображение Кэли является инъективным и сохраняет операцию.*

Доказательство. Инъективность очевидна. Сохранение операции:

$$(f_af_b)(g_i) = f_a(f_b(g_i)) = f_a(bg_i) = abg_i = f_{ab}(g_i). \quad \square$$

Следствие. *Любая конечная группа порядка n изоморфна некоторой подгруппе симметрической группы степени n.*

Задача 21. *Пусть a — элемент порядка k в мультипликативной группе порядка n. Докажите, что при отображении Кэли ему соответствует подстановка, которая разлагается в произведение n/k независимых циклов длины k.*

6.15 Группа обратимых матриц и ее подгруппы

Множество всех обратимых вещественных матриц порядка n относительно операции умножения образует группу, которая называется *полной линейной группой* и обозначается $GL(n) = GL(n, \mathbb{R})$. Ее подгруппы дают много примеров групп с теми или иными свойствами. В частности, любая конечная группа порядка n изоморфна некоторой подгруппе полной линейной группы вещественных матриц порядка n (докажите!).

Рассмотрим некоторые множества обратимых матриц, которые являются подгруппами полной линейной группы (проведите проверку для каждого примера).

- Множество вещественных матриц порядка n с определителем, равным единице. Эта подгруппа иногда называется *специальной линейной группой* и обозначается $SL(n) = SL(n, \mathbb{R})$. Группа $SL(n)$ является нормальной подгруппой группы $GL(n)$.

- Множество невырожденных матриц порядка n, элементами которых являются рациональные числа.

- Множество матриц порядка n с целочисленными элементами и определителем, равным единице.

- Множество невырожденных диагональных матриц. Напомним, что матрица $A = [a_{ij}]$ порядка n называется *диагональной*, если $a_{ij} = 0$ при $i \neq j$. Невырожденность означает, что $a_{ii} \neq 0$ при всех $1 \leq i \leq n$. Роль единичного элемента играет матрица

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

  которая называется *единичной матрицей*. Матрицы вида λI, где λ – число, называются *скалярными*.

- Множество невырожденных *нижних треугольных* (*верхних треугольных* матриц порядка n. Матрица $A = [a_{ij}]$ размеров $n \times n$ называется *нижней треугольной*, если $a_{ij} = 0$ при $i < j$, и *верхней треугольной*, если $a_{ij} = 0$ при $i > j$. Невырожденность треугольной матрицы означает, что $a_{ii} \neq 0$ при всех $1 \leq i \leq n$. Роль единичного элемента играет единичная матрица I, которая является одновременно нижней и верхней треугольной. Чтобы доказать, что невырожденные нижние (верхние) треугольные матрицы образуют подгруппу, нужно сделать два вещи:

  – проверить, что произведение невырожденных нижних (верхних) треугольных матриц остается нижней (верхней) треугольной матрицей;

  – проверить, что матрица, обратная к невырожденной нижней (верхней) треугольной матрице, остается нижней (верхней) треугольной.

Задача 22. *Докажите, что если матрица A порядка n коммутирует со всеми матрицами порядка n, то она является скалярной.*

Задача 23. *Докажите, что множество верхних треугольных матриц порядка n с единицами на главной диагонали является нормальной подгруппой мультипликативной группы невырожденных верхних треугольных матриц порядка n.*

Задача 24. *Пусть G — мультипликативная группа невырожденных верхних треугольных матриц порядка n с единицами на главной диагонали и H — множество матриц, которые отличаются от единичной матрицы порядка n только одним элементом в позиции $(1, n)$. Докажите, что H является нормальной подгруппой группы G.*

Задача 25. *Пусть G — множество верхних треугольных матриц вида*

$$\begin{bmatrix} a_0 & a_1 & \dots & a_{n-2} & a_{n-1} \\ & a_0 & a_1 & \dots & a_{n-2} \\ & & \ddots & \ddots & \vdots \\ & & & a_0 & a_1 \\ & & & & a_0 \end{bmatrix}, \quad a_0 \neq 0, \quad a_0, a_1, \dots, a_{n-1} \in \mathbb{R}.$$

Докажите, что G является абелевой группой относительно умножения матриц.

Задача 26. *Является ли отображение $A \to A^\top$ автоморфизмом полной линейной группы?*

Задача 27. *Пусть $u, v \in \mathbb{R}^n$ — фиксированные векторы-столбцы такие, что $v^\top u = 0$. Докажите, что множество матриц вида $I + \alpha uv^\top$ при всех вещественных α является подгруппой полной линейной группы.*

6.16 Конечные абелевы группы

Аддитивная абелева группа G называется *прямой суммой* своих подгрупп $G_1, \ldots, G_t$, если каждый ее элемент g однозначно представляется в виде суммы $g = g_1 + \ldots + g_t$ элементов $g_i \in G_i$. Обозначение: $G = G_1 \oplus \ldots \oplus G_t$.

Теорема о конечных абелевых группах. *В любой конечной аддитивной абелевой группе G существуют циклические подгруппы $G_1, \ldots, G_t$ такие, что $G = G_1 \oplus \ldots \oplus G_t$ и при этом $|G_{i+1}|$ делится на $|G_i|$ при $1 \le i \le t-1$.*

Доказательство. Рассмотрим всевозможные системы элементов $a_1, \ldots, a_t$, порождающие группу G как множество элементов вида $m_1a_1 + \ldots + m_ta_t$ с целыми коэффициентами $m_1, \ldots, m_t$ с минимально возможным значением t. Такие системы естественно называть *минимальными*. Доказательство проведем индукцией по t.

Среди всевозможных минимальных систем рассмотрим всевозможные равенства вида

$$n_1a_1 + \ldots + n_ta_t = 0$$

и выберем такую систему, для которой n_1 – минимальное целое положительное число для всех таких систем и такого типа равенств. Такое число n_1 будем называть *минимальным коэффициентом* группы G. Тогда (проверьте!):

- если $m_1a_1 + \ldots + m_ta_t = 0$, то m_1 делится на n_1,
- $n_2 = n_1q_2, \quad \ldots, \quad n_t = n_1q_t \quad$ для некоторых целых чисел $\quad q_2, \ldots, q_t$.

Положим $g_1 = a_1 + q_2a_2 + \ldots + q_ta_t$ и обозначим через G_1 циклическую группу кратных элемента g_1. Заметим, что $|G_1| = n_1$.

Пусть $\widehat{G}_2$ – конечная абелева группа, порожденная элементами $a_2, \ldots, a_t$. Тогда

$$G = G_1 \oplus \widehat{G}_2.$$

В самом деле, пусть $m_1, \ldots, m_t$ – произвольные целые число. Тогда

$$z := m_1a_1 + \ldots + m_ta_t = a + b, \quad a = m_1g_1, \quad b = (m_2 - m_1q_2)a_2 + \ldots + (m_t - m_tq_t)a_t.$$

Если $z = 0$, то m_1 делится на n_1 и поэтому $m_1g_1 = 0$. Значит, a и b определяются однозначно.

Группа $\widehat{G}_2$ порождается меньшим числом элементов, чем исходная группа G. Поэтому для нее теорему можно считать уже доказанной. Минимальный коэффициент группы $\widehat{G}$ не может быть меньше минимального коэффициента группы G, и, как следует из вышеприведенных рассуждений, должен делиться на n_1. $\square$

Напомним, что *примарной группой* или *p-группой* называется группа с числом элементов, равным степени простого числа p.

Следствие. *Любая конечная аддитивная абелева группа является прямой суммой циклических примарных подгрупп.*

Доказательство. Покажем, что любая аддитивная циклическая группа G порядка n представляется прямой суммой примарных подгрупп. Пусть $n = p_1^{l_1} \ldots p_s^{l_s}$ — разложение в произведение простых чисел, и пусть группа $G = \{mg : m \in \mathbb{Z}\}$ порождается своим элементом g. Тогда порядок элемента $a_i := (n/p_i^{l_i})g$ равен $n_i := p_i^{l_i}$ (проверьте!). Докажем, что порядок элемента $a := a_1 + \ldots + a_s$ равен n. Пусть k — минимальное натуральное число такое, что $ka = 0 \quad \Rightarrow \quad (n/n_i)ka = (n/n_i)ka_i = 0 \quad \Rightarrow$

$(n/n_i)k \vdots n_i$. Отсюда, в силу взаимной простоты чисел n/n_i и n_i, заключаем, что $k \vdots n_i$. Следовательно, число $k \le n$ делится на произведение попарно взаимно простых чисел $n = n_1 \dots n_s \quad \Rightarrow \quad k = n$. $\square$

Замечание. В случае конечной мультипликативной группы вместо прямой суммы рассматривается разложение в *прямое произведение* подгрупп. При этом используется обозначение $G = G_1 \times \dots \times G_t$.

Пример. Пусть $\mathbb{Z}_m$ и $\mathbb{Z}_n$ — аддитивные группы вычетов по модулю m и n. Рассмотрим множество упорядоченных пар $G = \{(a,b) : \ a \in \mathbb{Z}_m, \ b \in \mathbb{Z}_n\}$ и введем на нем операцию $(a,b)+(c,d) := (a+c, b+d)$. Нетрудно проверить, что G оказывается аддитивной абелевой группой и к тому же прямой суммой своих подгрупп $H_1 = \{(a,0) : a \in \mathbb{Z}_m\}$ и $H_2 = \{(0,b) : b \in \mathbb{Z}_n\}$. Очевидно, $H_1 \cong \mathbb{Z}_m$ и $H_2 \cong \mathbb{Z}_n$. Значит, H_1 и H_2 — циклические группы порядка m и n. Тем не менее, разложение $G = H_1 \oplus H_2$ не обязано совпадать с тем разложением, о котором идет речь в теореме о конечных абелевых группах, так как свойство $|H_2| \vdots |H_1|$ в общем случае не выполняется. Более того, если числа m и n взаимно просты, то группа G сама оказывается циклической (докажите!). В данном случае минимальная порождающая система группы G состоит из одного элемента. Теперь предположим, что m и n не являются взаимно простыми, и пусть $d > 1$ — их нетривиальный общий делитель. Тогда порядок любого элемента $g \in G$ является делителем наименьшего общего кратного чисел m и n, равного mn/d. В этом случае минимальные порождающие системы для группы G состоят из двух элементов, а минимальный коэффициент равен d. Пусть $m = du$ и $n = dv$, где числа u и v взаимно просты. Тогда существуют целые числа a и b такие, что $ua + vb = 1$, а группа G будет прямой суммой циклических подгрупп G_1 и G_2, порожденных элементами $([u]_m, [v]_n)$ и $([-b]_m, [a]_n)$. Как и утверждается в теореме, порядок подгруппы G_2 делится на порядок подгруппы G_1: $\quad |G_1| = d, \quad |G_2| = \frac{mn}{d}$ $\qquad$ (проверьте!).

Задача 28. *Докажите, что если A и B — произвольные мультипликативные группы, то декартово произведение $A \times B = \{(a,b) : a \in A, \ b \in B\}$ превращается в группу, если операцию умножения пар ввести как покомпонентное умножение: $((a,b)(c,d) := (ac, bd)$. Докажите, что если A и B — нормальные подгруппы группы G (возможно, некоммутативной), состоящей из элементов ab, где $a \in A$ и $b \in B$, то при условии $A \cap B = \{1\}$ группа $A \times B$ изоморфна группе G.*

.

6.17 Конечно порожденные группы

Группа G называется *конечно порожденной*, если в ней можно выделить конечную систему элементов $g_1, \dots, g_t$, для которых минимальная содержащая их подгруппа $\langle g_1, \dots, g_t \rangle$ совпадает с G. Любая конечная группа очевидно является конечно порожденной. Если аддитивная абелева группа G порождается своими элементами $g_1, \dots, g_t$, то это означает, что она состоит из всевозможных линейных комбинаций вида

$$g = n_1 g_1 + \dots + n_t g_t, \quad n_1, \dots, n_t \in \mathbb{Z}.$$

Если для любого элемента $g \in G$ такое представление единственно, то G называется *свободной абелевой группой ранга t*. Образующие элементы свободной абелевой группы *линейно независимы* в том смысле, что из равенства $n_1 g_1 + \dots + n_t g_t = 0$ вытекает $n_1 = \dots = n_t = 0$.

Утверждение 1. *Ранг свободной абелевой группы не зависит от выбора линейно независимой системы образующих ее элементов: если $\langle g_1, \dots, g_t \rangle = \langle h_1, \dots, h_s \rangle$ и каждая из систем $g_1, \dots, g_t$ и $h_1, \dots, h_s$ линейно независима, то $s = t$.*

Доказательство. Рассмотрим разложения $h_j = \sum\limits_{i=1}^{t} a_{ij} g_i$ и составим из их коэффициентов матрицу $A = [a_{ij}] \in \mathbb{Z}^{t \times s}$. Из линейно независимости образующих элементов $h_1, \dots, h_s$ вытекает линейная независимость столбцов данной матрицы (проверьте!) $\quad \Rightarrow \quad s \ge t$. Аналогичным образом получается противоположное неравенство. Поэтому $s = t$. $\square$

Утверждение 2. *Любая подгруппа конечно порожденной абелевой группы является конечно порожденной.*

Доказательство. Будем вести индукцию по числу t образующих абелевой группы $G = \langle g_1, \dots, g_t \rangle$. Пусть H — подгруппа, а $v = n_1 g_1 + \dots + n_t g_t$ — ее элемент с наименьшим целым положительным значением коэффициента n_1 среди всех разложений всех элементов данной подгруппы H по образующим

$g_1, \ldots, g_t$ группы G. Тогда для любого $h \in H$ найдется целое число m такое, что $h - mv \in H_1$, где $H_1 = G_1 \cap H$ — подгруппа группы $G_1 = \langle g_2, \ldots, g_t \rangle$ с числом образующих $t-1$. $\square$

Изучение конечно порожденных абелевых групп можно связать с изучением *унимодулярных преобразований* целочисленной матрицы, приводящих ее к специальному диагональному виду. Матрица называется *унимодулярной*, если ее элементы являются целыми числами, а определитель равен ± 1. Унимодулярное преобразование матрицы заключается в умножении ее слева и справа на унимодулярные матрицы. К *элементарным унимодулярным преобразованиям* относятся перестановки строк (столбцов), прибавление к строке (столбцу) целочисленной линейной комбинации других строк (столбцов) и умножение строк (столбцов) на ± 1.

Теорема об унимодулярной диагонализации. *Любая ненулевая целочисленная матрица A с помощью элементарных унимодулярных преобразований строк и столбцов приводится к однозначно определенной целочисленной прямоугольной диагональной матрице тех же размеров вида*

$$\mathcal{D} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_r \end{bmatrix}, \quad r = \operatorname{rank}(A),$$

где целые числа $d_1, \ldots, d_r$ положительны и для каждого $1 \le i \le r-1$ число d_{i+1} делится на d_i.

Доказательство. Рассмотрим множество всех $m \times n$-матриц, которые получаются из заданной целочисленной матрицы A с помощью элементарных унимодулярных преобразований строк и столбцов, и выберем в этом множестве матрицу B с наименьшим целым положительным числом в позиции $(1,1)$. Обозначим это число через d_1 и заметим, что любой элемент матрицы B должен делиться на d_1, иначе с помощью дополнительных унимодулярных преобразований можно перейти к матрице с меньшим целым положительным числом в позиции $(1,1)$. С помощью элементарных унимодулярных преобразований из B можно получить матрицу C с нулями во всех позициях первой строки и первого столбца, за исключением числа d_1 в позиции $(1,1)$. Заметим, что наибольший общий делитель всех элементов матрицы C равен d_1 и совпадает с наибольшим общим делителем всех элементов исходной матрицы A, и более того, при любом фиксированном $1 \le k \le \operatorname{rank}(A)$ при элементарных унимодулярных преобразованиях сохраняется наибольший общий делитель всех миноров k-го порядка (проверьте!). Ясно, что ранг подматрицы, полученной из C вычеркиванием первой строки и первого столбца, на единицу меньше ранга матрицы C. Доказательство завершается индукцией по рангу матрицы. $\square$

Теорема о конечно порожденных абелевых группах. *Любая конечно порожденная абелева группа является прямой суммой конечной абелевой группы и свободной абелевой группы.*

Доказательство. Пусть абелева группа G порождается своими элементами $g_1, \ldots, g_n$, где n — минимально возможное число образующих. Рассмотрим всевозможные равенства вида $\alpha_1 g_1 + \ldots + \alpha_n g_n = 0$ с целочисленными коэффициентами $\alpha_1, \ldots, \alpha_n$ и множество H соответствующих им векторов-строк $(\alpha_1, \ldots, \alpha_n)$. Относительно сложения векторов множество H является абелевой группой в свободной абелевой группе целочисленных строк с n элементами. Согласно утверждению 2, группа H является конечно порожденной и, очевидно, свободной (почему?). Рассмотрим $m \times n$-матрицу A, составленную из строк, порождающих группу H, и применим к ней теорему об унимодулярной диагонализации. Пусть $PAQ = \mathcal{D}$, где P и Q — унимодулярные матрицы, а $\mathcal{D}$ — диагональная прямоугольная матрица с положительными элементами $d_1, \ldots, d_r$ такими, что d_{i+1} делится на d_i. Очевидно, G порождается элементами $v_i = \sum\limits_{j=1}^{n} (Q^{-1})_{ij} g_j$ и является прямой суммой подгрупп $G_1 = \langle v_1, \ldots v_r \rangle$ и $G_2 = \langle v_{r+1}, \ldots, v_n \rangle$.

Равенства $d_1 v_1 = 0, \ \ldots, \ d_r v_r = 0$ означают, что порядки элементов $v_1, \ldots, v_r$ являются делителями чисел $d_1, \ldots, d_r$ (в действительности они равны числам $d_1, \ldots, d_r$ — докажите!). Поэтому целочисленные линейные комбинации $\alpha_1 v_1 + \ldots + \alpha_r v_r$ дают лишь конечное число различных элементов. Следовательно, G_1 является конечной группой.

Докажем, что группа G_2 является свободной. Пусть $z_{r+1} v_{r+1} + \ldots + z_n v_n = 0$, где $z_{r+1}, \ldots, z_n$ — некоторые целые числа. Тогда строка $(0, \ldots, 0, z_{r+1}, \ldots, z_n)$ является целочисленной линейной комбинацией строк матрицы $\mathcal{D} \quad \Rightarrow \quad z_{r+1} = \ldots = z_n = 0$. $\square$

Задача 29. *Докажите, что множество элементов конечного порядка в конечно порожденной абелевой группе является нормальной подгруппой, а фактор-группа G/H является свободной.*

Задача 30. *Докажите, что любая унимодулярная матрица представляется в виде произведения элементарных унимодулярных матриц.*

Алгебра и геометрия (1 поток)

Лекция 7

7.1 Определение кольца

Довольно часто на одном и том же множестве определяются не одна, а две разные операции. Например, целые числа или $n \times n$-матрицы с операциями сложения и умножения. Конечно, эти две операции должны быть между собой связаны, иначе нет смысла изучать их совместно.

В абстрактной ситуации будем считать, что есть непустое множество K с двумя внутренними бинарными операциями, которые называются сложением и умножением и связаны между собой двумя законами дистрибутивности [1]

$$a(b+c) = ab + ac, \quad (b+c)a = ba + ca \qquad \forall \;\; a, b, c \in K, \tag{$*$}$$

и, кроме того, относительно сложения множество K является аддитивной абелевой группой. В таких случаях множество K называется *кольцом*.

Утверждение. $0 \cdot a = a \cdot 0 = 0 \quad \forall \;\; a \in K$.

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: $0 = b + (0 \cdot a) = (b + 0 \cdot a) + (0 \cdot a) = 0 + (0 \cdot a) = 0 \cdot a$. $\square$

Часто умножение оказывается ассоциативным — такие кольца называются *ассоциативными*. В этой книге мы будем рассматривать только такие кольца. Если умножение коммутативно (например, как в кольце целых чисел), то кольцо называется *коммутативным*. Уже знакомый нам пример *некоммутативного* кольца — это кольцо вещественных $n \times n$-матриц при $n \geq 2$.

Приведем еще довольно абстрактный пример кольца. Пусть $G = \{g_1, \ldots, g_n\}$ — конечное множество, на котором введена внутренняя бинарная операция, для которой используется символика умножения. Пусть $g_i g_j = g_k$, где $k = f(i,j)$ — некоторая функция от i и j. Пусть K — множество *формальных выражений* вида $a := \sum_{i=1}^{n} a_i g_i$, где $a_i \in \mathbb{Z}$. Пусть $b = \sum_{j=1}^{n} b_j g_j$ — еще одно формальное выражение. Множество K становится кольцом (проверьте!), если операции сложения и умножения на K определить следующим образом: $a + b := \sum\limits_{i=1}^{n} (a_i + b_i) g_i, \quad ab := \sum\limits_{k=1}^{n} \left(\sum\limits_{f(i,j)=k} a_i b_j \right) g_k$. Если операция умножения на G ассоциативна, то кольцо K будет ассоциативным.

7.2 Делители нуля и целостные кольца

Если $ab = 0$ для каких-то элементов a и b, то каждый из них называется *делителем нуля*. Вот пример делителей нуля в кольце 2×2-матриц: $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = 0$.

[1] Более точно, соотношения $(*)$ означают, что *умножение дистрибутивно относительно сложения*.

Ассоциативное коммутивное кольцо без делителей нуля называется *целостным кольцом* или *областью целостности*. Стандартный пример — это, конечно, кольцо $\mathbb{Z}$ целых чисел. Кроме того, при фиксированном натуральном числе n множество $n\mathbb{Z}$ целых чисел, кратных n, также будет целостным кольцом. Подмножество кольца, на котором те же операции делают его кольцом, называется *подкольцом* исходного кольца. Понятно, конечно, что подкольцо любого целостного кольца будет также целостным.

7.3 Кольцо с единицей

Если умножение обладает единицей ($1 \cdot a = a \cdot 1 = a \ \ \forall\ a \in K$), то кольцо называется *кольцом с единицей* или *унитальным кольцом*. Единица в унитальном кольце всегда только одна (достаточно перемножить две единицы: $e_1 = e_1 e_2 = e_2$). Есть одно малоинтересное унитальное кольцо, для которого единица совпадает с нулем, – это кольцо состоит из одного единственного элемента. Если в унитальном кольце есть хотя бы два элемента, то заведомо $1 \neq 0$ (почему?).

Утверждение. Для любого элемента a в кольце с единицей выполняется равенство $(-1) \cdot a = a \cdot (-1) = -a$.

Доказательство. $0 = (1 + (-1)) \cdot a = 1 \cdot a + (-1) \cdot a = a + (-1) \cdot a$. $\square$

Если $ab = ba = 1$, то каждый из элементов a и b называется *делителем единицы* или *обратимым элементом*.

Задача 1. *Пусть a и b — элементы ассоциативного кольца с единицей e. Докажите, что из обратимости элемента $e - ab$ в данном кольце вытекает обратимость элемента $e - ba$.*

Задача 2. *Пусть a и b — элементы ассоциативного кольца с единицей e. Верно ли, что из равенства $ab = e$ следует $ba = e$?*

7.4 Определение поля

Пусть K — ассоциативное кольцо, для которого множество $K_* = K \backslash \{0\}$ относительно операции умножения является абелевой группой. В таких случаях кольцо K называется *полем*. Группа K_* называется *мультипликативной группой* поля K. Группа K по сложению называется *аддитивной группой поля*.

Утверждение. *В поле делителей нуля нет.*

Доказательство. От противного, пусть $ab = 0$ с ненулевыми a и b. Каждый ненулевой элемент поля обратим. Поэтому $0 = a^{-1} \cdot 0 = a^{-1}(ab) = (a^{-1}a)b = 1 \cdot b = b$. Противоречие. $\square$

Хорошо знакомые примеры полей — это поле $\mathbb{Q}$ рациональных чисел и поле $\mathbb{R}$ вещественных чисел. При этом, конечно, $\mathbb{Q} \subset \mathbb{R}$ и операции на $\mathbb{Q}$ можно рассматривать как применение операций, действующих на $\mathbb{R}$. В таких случаях говорят, что меньшее поле является *подполем* большего поля, а большее поле — *расширением* меньшего поля. Если $K = L$, то поле K называется *тривиальным подполем* поля L, а поле L — *тривиальным расширением* поля K.

Поле $\mathbb{R}$ имеет много подполей, отличных от $\mathbb{Q}$. Например, множество всех вещественных чисел вида $a + b\sqrt{2}$ при $a, b \in \mathbb{Q}$ является полем (докажите!). Заметим также, что любое подполе поля $\mathbb{R}$ является расширением поля $\mathbb{Q}$ (докажите!).

7.5 Кольцо вычетов

Целое число b называется сравнимым с целым числом a по модулю n, если $b-a$ делится на n. Сравнимость означает, что a и b при делении на n имеют одинаковый остаток. Вычет $[a]_n$ — это множество всех целых чисел, сравнимых с a. Множество $\mathbb{Z}_n$ всех вычетов по модулю n состоит ровно из n вычетов. Сложение и умножение вычетов определяются следующим образом:

$$[a]_n + [b]_n := [a+b]_n, \quad [a]_n[b]_n := [ab]_n.$$

Поскольку операции над вычетами $[a]_n$ и $[b_n]$ (множествами) определяются через их представителей a и b, нужно проверить, что получаемые в результате вычеты (множества) не зависят от выбора представителей. Пусть $[a]_n = [a_1]_n$ и $[b]_n = [b_1]_n$. Тогда

$$a - a_1 \vdots n, \quad b - b_1 \vdots n \quad \Rightarrow \quad (a+b) - (a_1 + b_1) = (a - a_1) + (b - b_1) \vdots n,$$

$$ab - a_1 b_1 = a(b - b_1) + (a - a_1)b_1 \vdots n.$$

Ассоциативность сложения и дистрибутивность умножения относительно сложения следуют из аналогичных свойств для сложения и умножения целых чисел:

$$[a]_n+([b]_n+[c]_n) = [a]_n+[b+c]_n = [a+(b+c)]_n = [(a+b)+c]_n = [a+b]_n+[c]_n = ([a]_n+[b]_n)+[c]_n,$$

$$[a]_n([b]_n + [c]_n) = [a]_n[b + c]_n = [a(b + c)]_n = [ab + ac]_n = [ac]_n + [ac]_n = [a]_n[c]_n + [b]_n[c]_n.$$

Таким образом, $\mathbb{Z}_n$ – это кольцо. Более того, оно ассоциативно, коммутативно и обладает единицей $[1]_n$. Если n составное, то есть делители нуля: $n = ab \;\Rightarrow\; [a]_n\,[b]_n = [n]_n = [0]_n$.

7.6 Обратимые элементы кольца вычетов

Обозначим через $\mathbb{Z}_n^*$ множество обратимых элементов кольца $\mathbb{Z}_n$. Если числа a и n имеют нетривиальный общий делитель, то вычет $[a]_n$ является делителем нуля и по этой причине не может быть обратимым (проверьте!).

Теорема об обратимых вычетах. Вычет $[a]_n$ является обратимым в кольце $\mathbb{Z}_n$ тогда и только тогда, когда a и n взаимно просты.

Доказательство. Если a и n взаимно просты, то существуют целые числа x, y такие, что $ax + ny = 1$. В самом деле, пусть d — минимальное положительное значение для чисел вида $ax + ny$. Если $d > 1$, то поделим a на d с остатком: $a = dq + r$. Тогда $(ax+ny)q+r = dq+r = a \;\Rightarrow\; a(1-xq)+(-yq)n = r$. Если $r > 0$, то имеем противоречие с минимальностью значения d. Значит, a делится на d и в полной аналогии n делится на $d \;\Rightarrow\; d = ax + ny = 1 \quad\Rightarrow\quad [a]_n\,[x]_n + [n]_n\,[y]_n = [a]_n\,[x]_n = [1]_n$. $\square$

Теорема о мультипликативной группе вычетов. *Обратимые элементы кольца вычетов $\mathbb{Z}_n$ образуют группу по умножению.*

Доказательство. В коммутативном кольце произведение двух обратимых элементов очевидно является обратимым. $\square$

Таким образом, множество $\mathbb{Z}_n^*$ является мультипликативной группой. Ее порядок обозначается через $\phi(n)$, где $\phi(n)$ называется *функцией Эйлера*. При $n = 1$ кольцо $\mathbb{Z}_1$

состоит ровно из одного элемента, который является одновременно нулевым и единичным. Поэтому $\phi(1) = 1$. Если $n > 1$, то значение $\phi(n)$ равно числу натуральных чисел от 1 до n, взаимно простых с числом n.

Основным свойством функции Эйлера является мультипликативность: если a и b взаимно просты, то $\phi(ab) = \phi(a)\phi)b)$. Для доказательства нужно рассмотреть отображение $(x, y) \to ax + by$ при целых $0 \le x \le b-1$ и $0 \le y \le a-1$ и убедиться в том, что набор значений данного отображения дает полную систему остатков при делении на число ab (китайская теорема об остатках). Затем нужно проверить, что число $d = ax + by$ взаимно просто с числом ab тогда и только тогда, когда x взаимно прост с b и y взаимно прост с a. Свойство мультипликативности приводит к следующей формуле: если p_1, ..., p_t — различные простые числа и $n_1, \ldots, n_t$ — натуральные числа, то

$$\phi(p_1^{n_1} \ldots p_t^{n_t}) = (p_1^{n_1} - p_1^{n_1-1}) \ldots (p_t^{n_t} - p_t^{n_t-1}).$$

Пусть $a_1, \ldots, a_{\phi(n)}$ — все натуральные числа от 1 до n, взаимно простые с числом n. Возьмем любое число a, взаимно простое с числом n, и обозначим через r_k остаток при делении числа aa_k на n. Допустим, что $r_k = r_l$. Тогда $a(a_k - a_l)$ делится на n и, в силу взаимной простоты чисел a и n, отсюда следует, что $a_k - a_l$ делится на n. Ясно, что $-(n-1) \le a_k - a_l \le n-1$ и в этом диапазоне на n делится лишь число 0. Таким образом, остатки $r_1, \ldots, r_{\phi(n)}$ попарно различны и каждый из них взаимно прост с числом n. Значит, среди этих остатков есть число 1, и мы получаем еще одно доказательство теоремы об обратимых элементах. Кроме того, мы приходим к следующему утверждению, известному как

Теорема Эйлера. *Если a и n взаимно просты, то $a^{\phi(n)} - 1$ делится на n.*

Доказательство. Заметим, что $a_1 \ldots a_{\phi(n)} = r_1 \ldots r_{\phi(n)}$. Перемножая равенства

$$aa_1 = nq_1 + r_1, \quad \ldots \quad aa_{\phi(n)} = nq_{\phi(n)} + r_{\phi(n)},$$

возникающие при делении с остатком на n, находим, что число $(a^{\phi(n)} - 1)(a_1 \ldots a_{\phi(n)})$ делится на n $\Rightarrow$ $a^{\phi(n)} - 1$ делится на n. $\square$

Задача 3. *Докажите равенство $n = \sum\limits_{d|n} \phi(d)$, где сумма берется по всем делителям d числа n.*

Задача 4. *Элемент $a \in \mathbb{Z}_n^*$ называется* `квадратичным вычетом`*, если $a = x^2$ для некоторого $x \in \mathbb{Z}_n^*$. Докажите, что в случае простого n число квадратичных вычетов равно $(n-1)/2$ и, кроме того, $a^{(n-1)/2} = 1$, если a является квадратичным вычетом, и $a^{(n-1)/2} = -1$ в противном случае.*

7.7 Кольца вычетов в криптографии

Кольца вычетов встречаются в нашей жизни не только в математических книгах. Например, мы имеем с ними дело, снимая деньги в банкоматах или соединяясь с удаленными компьютерами по сети. Чтобы сохранить конфиденциальность информации при ее передаче по открытыми каналам связи, эта информация каким-то образом кодируется. При этом способ кодирования не является секретом (если банкомат узнает его от банка, то и злоумышленник, в принципе, тоже может им завладеть). Секретом, однако, является способ декодирования. Система шифрования должна быть такой, чтобы раскрытие способа расшифровки требовало серьезных усилий и достаточно большого времени.

В качестве алфавита ничто не мешает использовать произвольное конечное множество M. Способ кодирования задается некоторым отображением $f : M \to M$. Для декодирования нужно иметь обратное отбражение.

В 1977 году появилась система шифрования RSA [2], в которой $M = \mathbb{Z}_n$ и $f(x) = x^e$, где e — заданное натуральное число. Числа n и e подбираются специальным образом, а именно: $n = pq$, где p и q — различные достаточно большие простые числа, e — целое число, взаимно простое с числом $\phi(n) = (p-1)(q-1)$. Тогда для некоторых натуральных чисел d и q выполняется равенство $ed - \phi(n)q = 1$. Если a взаимно просто с n, то, согласно теореме Эйлера, $[a^{\phi(n)}]_n = [1]_n$. Следовательно,

$$[a]_n^e)^d \;=\; [a^{ed}]_n \;=\; [a^{1+\phi(n)q}]_n \;=\; [a]_n\,[a^{\phi(n)}]_n = [a]_n.$$

[2] Это заглавные буквы фамилий трех авторов: R. L. Rivest, A. Shamir, L. Adleman.

Можно убедиться и в том, что полученное равенство вычетов остается в силе и без требования взаимной простоты a и n, т.е для любых целых чисел a (проверьте!). Таким образом, если $y = x^e$, то $x = y^d$, т.е правило декодирования имеет вид $y \to y^d$.

Зная n и e, легко проводить шифрование. Но для дешифровки нужно знать d. Конечно, d определяется по n и e — но если p и q держатся в секрете, то получение d при больших n оказывается алгоритмически очень трудной задачей, сводящейся к разложению n на простые множители.

7.8 Поле вычетов

Теорема о полях вычетов. *В случае простого p и только в этом случае кольцо вычетов по модулю p является полем.*

Доказательство. Для простого p имеем $\phi(p) = p - 1$. Согласно полученной выше теореме об обратимых элементах кольца вычетов, в данном случае любой ненулевой вычет является обратимым элементом. □

Таким образом, теперь у нас есть примеры полей с конечным числом элементов. Такие поля называются *конечными полями*. В честь Эвариста Галуа их часто называют также *полями Галуа* и используют обозначение $\mathrm{GL}(n)$, где n — число элементов поля. В случае полей вычетов число n должно быть простым.

Каким вообще может быть число элементов в конечном поле? Чтобы ответить на этот вопрос, необходимо иметь более общее понятие векторного пространства, в котором векторы будут уже абстрактными объектами, над которыми производятся *абстрактные операции* сложения и умножения на элементы поля, но *свойства операций* остаются такими же, как у изученных нами операций над арифметическими векторами.

7.9 Линейные пространства

Пусть $\mathbb{P}$ — произвольное поле, элементы которого называются *числами*, V — непустое множество, элементы которого называются *векторами*. Мы будем считать, что на множестве V определены две операции: *сложение векторов* и *умножение вектора на число*. Результатом каждой операции является вектор.

Множество V называется *линейным пространством над полем* $\mathbb{P}$, если V является аддитивной абелевой группой относительно операции сложения векторов, а операция умножения вектора на число обладает следующим наборов свойств:

- $(\alpha\beta)v = \alpha(\beta v)$ (ассоциативность умножения на число);
- $(\alpha + \beta)v = \alpha v + \beta v$ (дистрибутивность относительно сложения чисел);
- $\alpha(u + v) = \alpha u + \alpha u$ (дистрибутивность относительно сложения векторов);
- $1 \cdot v = v$ (аксиома сюръективности).

Эти свойства выполняются для любых чисел $\alpha, \beta \in \mathbb{P}$ и любых векторов $u, v \in V$. Линейные пространства часто называются также векторными пространствами.

Аксиома $1 \cdot v = v$ позволяет каждый вектор u считать вектором вида $u = \alpha v$ (достаточно взять $\alpha = 1$ и $v = u$) и сама является следствием сюръективности отображения $(\alpha, v) \to \alpha v$ (в предположении, что свойства ассоциативности и дистрибутивности для умножения на число выполнены): $1 \cdot u = 1 \cdot (\alpha v) = (1 \cdot \alpha)v = \alpha v = u$.

Заметим, что одно и то же множество векторов V может рассматриваться как линейное пространство над разными полями. Линейные пространства для разных полей должны считаться разными.

Задача 5. *Докажите, что группа целых чисел с операцией сложения не может быть аддитивной группой линейного пространства над каким-либо полем.*

7.10 Операции с нулевым вектором

Нулевой вектор иногда обозначается символом $\mathbf{0}$, но чаще просто как 0, т.е. так же, как нуль в поле $\mathbb{P}$. Чтобы понять, идет ли речь о числе или о векторе, нужно учитывать контекст. Совершенно очевидные свойства операций с участием нулевого арифметического вектора теперь *нуждаются в доказательствах.*

Утверждение 1. $0 \cdot a = \mathbf{0} \quad \forall \ a \in V$.

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 $\mathbf{0} = b + (0 \cdot a) = (b + (0 \cdot a)) + (0 \cdot a) \quad \Rightarrow \quad \mathbf{0} = 0 \cdot a$. $\square$

Утверждение 2. $\alpha \cdot \mathbf{0} = \mathbf{0} \quad \forall \ \alpha \in \mathbb{P}$.

Доказательство. $\alpha \cdot \mathbf{0} = \alpha(\mathbf{0} + \mathbf{0}) = \alpha \cdot \mathbf{0} + \alpha \cdot \mathbf{0} \quad \Rightarrow \quad \alpha \cdot \mathbf{0} = \mathbf{0}$. $\square$

Утверждение 3. *Если* $\alpha \cdot a = \mathbf{0}$, *то либо* $\alpha = 0$, *либо* $a = \mathbf{0}$.

Доказательство. [3] Пусть $\alpha \neq 0$. Тогда

$$a = 1 \cdot a = \left(\left(\frac{1}{\alpha}\right)\alpha\right) \cdot a = \left(\frac{1}{\alpha}\right)(\alpha \cdot a) = \left(\frac{1}{\alpha}\right) \cdot \mathbf{0} = \mathbf{0}. \qquad \square$$

7.11 Линейные комбинации, размерность, базисы

Как и раньше, вектор вида $b = \alpha_1 a_1 + \ldots + \alpha_n a_n$ называется линейной комбинацией векторов $a_1, \ldots a_n$, только теперь коэффициенты $\alpha_1, \ldots, \alpha_n$ принадлежат полю $\mathbb{P}$, а векторы — абстрактному множеству V.

Понятия линейной оболочки, линейной зависимости и независимости систем векторов, размерности определяются дословно так же, как это было сделано для вещественных арифметических векторов. Сохраняются не только все изученные ранее свойства, связанные с этими понятиями, включая лемму о монотонности, но и буквально сами доказательства. Единственное изменение — это более абстрактный *смысл* ранее использовавшихся терминов.

Если размерность линейного пространства конечна, то оно называется *конечномерным*. Если существуют сколь угодно большие системы линейно независимых векторов, то пространство называется *бесконечномерным*. Базис в конечномерном пространстве определяется как линейно независимая система, линейная оболочка которой совпадает с

[3] 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.

заданным пространством. Дословно сохраняются та же формулировка и те же наблюдения, как и в случае вещественных арифметических векторов, в том числе возможность достраивания любой линейно независимой системы до базиса.

В бесконечномерных пространствах иногда рассматривают так называемые *линейные базисы* или *базисы Гамеля*, представляющие собой бесконечные множества с двумя свойства:

- множество должно быть *линейно независимым* — в том смысле, что любая конечная система его векторов линейно независима;
- *линейная оболочка* множества, понимаемая как совокупность всевозможных линейных комбинаций всевозможных конечных систем его векторов, должна совпадать с заданным пространством.

Существование базиса Гамеля можно вывести из так называемой *аксиомы выбора*, утверждающей, что для любого непустого множества существует отображение, определенное на множестве его непустых подмножеств и притом такое, что его значение на любом отдельном подмножестве является элементом этого подмножества. Заметим, что факт существования базиса Гамеля является в высшей степени неконструктивным, так как никак не связан с каким-либо алгоритмом его построения.

7.12 Изоморфные пространства

Пусть $\mathbb{P}$ — произвольное поле. Рассмотрим следующие два примера линейных пространств над полем $\mathbb{P}$.

1. V — множество $m \times n$-матриц с элементами из $\mathbb{P}$. По аналогии с вещественными матрицами принимается обозначение $V = \mathbb{P}^{m\times n}$. Операции сложения матриц и умножения на числа из $\mathbb{P}$ определяются поэлементно. Линейное пространство V конечномерно и его размерность равна mn (проверьте!).

2. V — множество всевозможных функций, определенных на непустом множестве M и принимающих значения в поле $\mathbb{P}$. Сумма функций и умножение функций на числа из $\mathbb{P}$ определяются поточечно:

$$(f + g)(x) := f(x) + f(y), \quad (\alpha f)(x) := \alpha f(x).$$

Роль нулевого вектора выполняет функция, тождественно равная нулю. Если множество M бесконечно, то данное линейное пространство бесконечномерно (докажите!).

Второй пример в некотором смысле включает первый, так как элементы матрицы можно рассматривать как значения функции $(i, j) \to a_{ij}$.

В более общей ситуации формально разные линейные пространства V и W над общим полем отождествляются с помощью понятия изоморфизма. Как и в случае групп, под изоморфизмом понимается взаимно-однозначное отображение $f : V \to W$, сохраняющее операции. В отличие от групп, в случае линейных пространств надо заботиться о сохранении двух операций:

$$f(x + y) = f(x) + f(y), \quad f(\alpha x) = \alpha f(x).$$

Если $W = V$, то изоморфизм называется *автоморфизмом*. Если отображение сохраняет операции, но, возможно, не является взаимно-однозначным, то оно называется *гомоморфизмом*. Если $W = V$, то гомоморфизм называется *эндоморфизмом*.

Тоерема. *Конечномерные линейные пространства над общим полем изоморфны в том и только том случае, когда равны их размерности.*

Доказательство. Пусть $f : V \to W$ — изоморфизм конечномерных пространств. Тогда любая линейно независимая система $a_1, \ldots, a_k \in V$ переводится в линейно независимую систему $f(a_1), \ldots, f(a_k) \in W$ (почему?) $\Rightarrow \dim V \le \dim W$. Обратное отображение $f^{-1} : W \to V$ также является изоморфизмом $\Rightarrow \dim W \le \dim V \Rightarrow \dim V = \dim W$. Теперь предположим, что $\dim V = \dim W = n$. Фиксируем произвольные базисы $e_1, \ldots, e_n$ и $g_1, \ldots, g_n$ в пространствах V и W и рассмотрим отображение $f(\alpha_1 e_1 + \ldots + \alpha_n e_n) := \alpha_1 g_1 + \ldots + \alpha_n g_n$. Оно и будет искомым изоморфизмом (проверьте!). $\square$

Задача 6. *Пусть A и B — обратимые $n \times n$-матрицы с элементами из поля $\mathbb{P}$. Докажите, что отображение $f(X) = AXB$ является автоморфизмом пространства $\mathbb{P}^{n\times n}$. Любой ли автоморфизм представляется таким образом?*

Задача 7. *Докажите, что для линейной независимости функций $f_1(x)$, …, $f_n(x)$ необходимо и достаточно, чтобы для некоторых точек x_1, …, x_n матрица $[f_i(x_j)]_{ij=1}^n$ была обратимой.*

Задача 8. *Докажите линейную независимость функций $\sin x$, $\sin 2x$, …, $\sin nx$ как элементов вещественного линейного пространства функций на произвольном фиксированном нетривиальном отрезке $[a, b]$.*

7.13 Сумма и пересечение подпространств

Непустое подмножество $L \subseteq V$ называется *подпространством* линейного пространства V, если оно само является линейным пространством относительно операций, действующих в V. Для этого необходимо и достаточно, чтобы результаты этих операций над векторами из L оставались в L.

Сумма подпространств $L = L_1 + \ldots + L_s$ пространства V определяется как множество вида $L = \{x_1 + \ldots + x_s \ : \ x_1 \in L_1, \ \ldots \ , \ x_s \in L_s\}$. Тривиально проверяется, что множество L является подпространством. Представление $x = x_1 + \ldots + x_s$ называется *разложением вектора по подпространствам*. Если при этом для каждого $x \in L$ компоненты разложения $x_i \in L_i$ определены однозначно, то L называется *прямой суммой* подпространств $L_1, \ldots, L_s$ и принимается обозначение $L = L_1 \oplus \ldots \oplus L_s$. Заметим также, что пересечение любого количества подпространств (как пересечение множеств) является подпространством (почему?).

Теорема Грассмана. *Пусть L и M — конечномерные подпространства некоторого линейного пространства. Тогда* $\dim(L + M) = \dim L + \dim M - \dim(L \cap M)$.

Доказательство. Рассмотрим базис $g_1, \ldots, g_r$ подпространства $L \cap M$ и дополним его сначала до базиса L, а затем до базиса M:

$$g_1, \ldots, g_r,\ p_1, \ldots, p_k \quad \text{(базис } L\text{)}, \qquad g_1, \ldots, g_r,\ q_1, \ldots, q_m \quad \text{(базис } M\text{)}.$$

Очевидно, $L+M$ является линейной оболочкой векторов $g_1, \ldots, g_r,\ p_1, \ldots, p_k,\ q_1, \ldots, q_m$, и нам остается лишь установить их линейную независимость. Пусть

$$\alpha_1\, g_1 + \ldots + \alpha_r\, g_r \ + \ \beta_1\, p_1 + \ldots + \beta_k\, p_k \ + \ \gamma_1\, q_1 + \ldots + \gamma_m\, q_m \ = \ 0 \quad \Rightarrow$$

$$z := \alpha_1\, g_1 + \ldots + \alpha_r\, g_r \ + \ \beta_1\, p_1 + \ldots + \beta_k\, p_k \ = \ -(\gamma_1\, q_1 + \ldots + \gamma_m\, q_m) \in L \cap M.$$

Будучи элементом из $L \cap M$, вектор z представляется в виде $z = \delta_1 g_1 + \ldots + \delta_r g_r \quad \Rightarrow$

$$\delta_1\, g_1 + \ldots + \delta_r\, g_r + \gamma_1\, q_1 + \ldots + \gamma_m\, q_m = 0 \quad \Rightarrow \quad \delta_1 = \ldots = \delta_r = \gamma_1 = \ldots = \gamma_m = 0. \ \square$$

Задача 9. *Найдите размерность суммы подпространства $n \times n$-матриц с нулевой суммой элементов в каждой строке и подпространства $n \times n$-матриц с нулевой суммой элементов в каждом столбце.*

Задача 10. *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.*

7.14 Матрицы над полями и кольцами

Пусть K — произвольное ассоциативное кольцо с единицей и пусть $K^{m\times n}$ — множество матриц с элементами из K. Такие матрицы можно складывать и умножать по тем же правилам, что и вещественные матрицы. При этом множество $K^{n\times n}$ становится ассоциативным унитальным кольцом (проверьте!).

Если кольцо K некоммутативно, то уже трудно говорить о разумных обобщениях таких важных понятий, как определитель и ранг. Если же K является полем, то ситуация другая — все наши построения остаются в силе! Понятие определителя переносится на квадратные матрицы с элементами из поля K без каких-либо изменений. Сохраняется основная теорема об определителе и вообще все его свойства. На прямоугольные матрицы с элементами из K дословно переносится понятие ранга, при этом остаются в силе все свойства ранга, включая теорему о базисном миноре, свойства минимального скелетного разложения и связь с матричными операциями. Заметим также, что если K — ассоциативное коммутативное унитальное кольцо, то любая полилинейная функция столбцов $n \times n$-матрицы, обнуляющаяся при совпадении пары столбцов, также оказывается определителем, умноженным на значение этой функции на столбцах единичной матрицы. Как и в случае поля, определитель сохраняется при прибавлении к любой строке (любому столбцу) линейной комбинации остальных строк (столцов) с коэффициентами, принадлежащими кольцу K. Остаются в силе теорема Лапласа и очень полезное утверждение о том, произведение присоединенной и исходной матриц равно скалярной матрице с элементом главной диагонали, равным определителю исходной матрицы.

Матрицы над конечными полями играют важную роль в прикладных вопросах математики — например, в теории кодирования. В частности, *коды Хэмминга* представляют собой решения однородной системы уравнений $Ax = 0$, где $A \in \mathbb{Z}_2^{m\times n}$ и $x \in \mathbb{Z}_2^n$. Если $m = 3$ и $n = 7$, то пространство $\mathbb{Z}_2^m$ содержит ровно $2^3 - 1$ ненулевых векторов — из них составляются столбцы матрицы $A \;\Rightarrow\;$ базис пространства решений содержит 4 вектора. Кодирование происходит таким образом: 4-битовое слово (c_1, c_2, c_3, c_4) с компонетами $c_i \in \mathbb{Z}_2$ кодируется 7-битовым словом x, получаемым как линейная комбинация фиксированных векторов базиса с коэффициентами c_1, c_2, c_3, c_4. Замечательно то, что ошибка в произвольном — но только одном! — элементе x_i легко обнаруживается: достаточно проверить равенство $Ax = 0$. Более того, любая такая ошибка легко исправляется!

7.15 Линейные пространства и расширения полей

Пусть $L \supseteq K$ — расширение поля K. Тогда можно рассматривать элементы поля L как векторы, а элементы из K как числа, на которые эти векторы умножаются. Операция сложения векторов определяется операцией сложения элементов поля L. Операция умножения векторов на числа определяется умножением элементов поля L на элементы

его подполя K. Таким образом, поле L можно рассматривать как линейное пространство над полем K. Если это пространство конечномерно, то его размерность в алгебраической литературе принято называть *степенью расширения* и обозначать символом $(L : K)$. В таких случаях расширение $L \supseteq K$ называется *конечным расширением*.

Теорема о степенях расширений. *Пусть имеются три поля $M \supseteq L \supseteq K$ и пусть оба расширения $M \supseteq L$ и $L \supseteq K$ являются конечными. Тогда расширение $M \supseteq K$ тоже является конечным, а размерности трех линейных пространств связаны соотношением $(M : L)(L : K) = (M : K)$.*

Доказательство. Пусть $a_1, \ldots, b_m$ — базис линейного пространства M над полем L, и пусть $b_1, \ldots, b_l$ — базис линейного пространства L над полем K. Числа $a_i b_j$ можно рассматривать как векторы линейного пространства M над полем K, и очевидно, что M состоит из их линейных комбинаций с коэффициентами из поля K. Кроме того, эти векторы линейно независимы (проверьте!). Таким образом, линейное пространство M над полем K обладает базисом, состоящим из ml векторов. □

Задача 11. *Докажите, что расширение $\mathbb{R} \supset \mathbb{Q}$ не является конечным.*

Задача 12. *Пусть заданы простые числа $p_1 < \ldots < p_s$, а числа $\sqrt{p_1} < \ldots < \sqrt{p_s}$ рассматриваются как элементы линейного пространства над полем рациональных чисел. Докажите, что эти элементы линейно независимы.*

7.16 Линейные пространства и алгебры

Линейное пространство V над полем $\mathbb{P}$ называется *алгеброй*, если на нем определена также операция умножения элементов ab, обладающая *свойством билинейности* как функция от a и b:

$$(\alpha u + \beta v)b = \alpha(ub) + \beta(vb), \quad a(\alpha u + \beta v) = \alpha(au) + \beta(av) \ \ \forall\, a, b, u, v \in V, \ \ \alpha, \beta \in \mathbb{P}.$$

Таким образом, алгебра является одновременно линейным пространством и кольцом с билинейным умножением. В случае конечномерности линейного пространства алгебра называется *конечномерной*. Если умножение ассоциативно или коммутативно, то алгебра называется ассоциативной или коммутативной. Если для умножения есть единица, то алгебра называется унитальной или алгеброй с единицей. Вот два очень распространенных примера:

- Алгебра $n \times n$-матриц над полем $\mathbb{P}$. В качестве билинейной операции умножения рассматривается обычная операция умножения матриц. Эта алгебра является ассоциативной алгеброй размерности n^2. При $n \geq 2$ это пример некоммутативной алгебры.

- Алгебра функций, определенных на непустом множестом множестве M и принимающих значения в поле $\mathbb{P}$. В качестве билинейной операции умножения рассматривается поточечное умножение: $(fg)(x) := f(x)g(x)$. Это пример ассоциативной коммутативной алгебры.

Алгебра и геометрия (1 поток)

Лекция 8

8.1 Определение многочлена

Во многих вопросах математики о многочленах можно думать просто как о функциях специального вида. Однако, в алгебре многочлены определяются как *формальные выражения*.

Многочлен $f(x_1, \ldots, x_n)$ от переменных $x_1, \ldots, x_n$ над полем $\mathbb{P}$ — это *формальная конечная сумма*, каждый член которой имеет вид $\alpha\, x_1^{i_1} \ldots x_n^{i_n}$, представляя собой формальное произведение *числового коэффициента* $\alpha \in \mathbb{P}$ и *монома* $x_1^{i_1} \ldots x_n^{i_n}$, в котором переменные считаются просто буквами, формально возведенными в целые неотрицательные степени. Сумма $i_1 + \ldots + i_n$ называется *степенью монома*. Можно также думать о мономе как о слове, составленном из букв некоторого алфавита, в котором могут встречаться одинаковые буквы, а порядок букв не имеет значения, — и тогда степенью будет общее число букв. Член с нулевым коэффициентом называется нулевым. *Степень ненулевого члена* — это степень входящего в него монома. Нулевые члены и ненулевые члены нулевой степени называются также *константами* и отождествляются с числами поля $\mathbb{P}$.

Члены, входящие в состав многочлена $f(x)$, разрешается ставить в любом порядке. Кроме того, члены с одним и тем же мономом называются *подобными членами*, и их разрешается заменить на один член с тем же мономом и коэффициентом, равным сумме коэффициентов подобных членов. Такое преобразование называется *приведением подобных членов*. Оно включает в себя также замену констант на одну константу, равную их сумме, и замену любого нулевого члена на константу, равную нулю. Многочлены считаются равными, если один из другого можно получить перестановками членов и приведением подобных членов.

Понятно, что каждый многочлен имеет *минимальное представление* в виде формальной суммы минимального количества членов. Если в результате получилась константа, равная нулю, то многочлен называется *нулевым*. *Степень ненулевого многочлена* $f(x)$ определяется как максимальная степень мономов в его минимальном представлении и обозначается $\deg f(x)$. Многочлен, в котором все ненулевые члены имеют одну и ту же степень, называется *однородным многочленом*. Для нулевого многочлена понятие степени не определяется.

Множество всех многочленов от переменных $x_1, \ldots, x_n$ над полем $\mathbb{P}$ обозначается через $\mathbb{P}[x_1, \ldots, x_n]$.

8.2 Сумма и произведение многочленов

Суммой двух многочленов называется формальная сумма всех членов, входящих в состав каждого многочлена.

Чтобы ввести произведение многочленов $f(x)$ и $g(x)$, сначала определяем правило перемножения одночленов: $(\alpha\, x_1^{i_1} \dots x_n^{i_n})\,(\beta\, x_1^{j_1} \dots x_n^{j_n}) := (\alpha\beta)\, x_1^{i_1+j_1} \dots x_n^{i_n+j_n}$. Произведением многочленов $f(x)$ и $g(x)$ называется сумма произведений принадлежащих им одночленов.

Таким образом, для многочленов

$$f(x_1,\dots,x_n) = \sum_{i_1=1}^{d_1} \dots \sum_{i_n=1}^{d_n} a_{i_1\dots i_n} x_1^{i_1} \dots x_n^{i_n}, \quad g(x_1,\dots,x_n) = \sum_{j_1=1}^{d_1} \dots \sum_{j_n=1}^{d_n} b_{j_1\dots j_n} x_1^{j_1} \dots x_n^{j_n}$$

сумма и произведение имеют вид

$$f(x) + g(x) = \sum_{i_1=1}^{d_1} \dots \sum_{i_n=1}^{d_n} (a_{i_1\dots i_n} + b_{i_1\dots i_n}) x_1^{i_1} \dots x_n^{i_n},$$

$$f(x)g(x) = \sum_{i_1=1}^{d_1} \dots \sum_{i_n=1}^{d_n} \sum_{j_1=1}^{d_1} \dots \sum_{j_n=1}^{d_n} (a_{i_1\dots i_n} b_{j_1\dots j_n}) x_1^{i_1+j_1} \dots x_n^{i_n+j_n}.$$

В частности, для коэффициентов произведения многочленов от одной переменной получаем такую формулу:

$$\left(\sum_{i=0}^{M} a_i x^i\right)\left(\sum_{j=0}^{N} b_j x^j\right) \;=\; \sum_{k=0}^{M+N} \left(\sum_{i=0}^{k} a_i b_{k-i}\right) x^k.$$

8.3 Старшие члены

Старшим членом ненулевого многочлена называется член максимальной степени в минимальном представлении данного многочлена.

Очевидно, для многочлена от одной переменной старший член определен однозначно. Для многочленов от нескольких переменных старших членов может оказаться несколько. Полезно иметь в виду, что минимальное представление многочлена $f(x_1,\dots,x_n)$ степени d можно записать в виде суммы однородных многочленов

$$f(x_1,\dots,x_n) = \sum_{k=0}^{d} F_k(x_1,\dots,x_n), \quad \text{где} \quad F_k(x_1,\dots,x_n) = \sum_{i_1+\dots+i_n=k} a_{i_1\dots i_k} x_1^{i_1} \dots x_n^{i_n}.$$

Многочлен F_d называется *старшей однородной компонентой* многочлена f, и ясно, что каждый ненулевой член в составе старшей однородной компоненты многочлена $f(x)$ является его старшим членом.

Полезно среди старших членов выбрать какой-то один, который будет именоваться, скажем, *главным членом*. Для этого нужен какой-то принцип упорядочивания старших членов. Часто используется следующий *лексикографический* или *словарный* принцип: моном $x_1^{i_1} \dots x_n^{i_n}$ считается главнее монома $x_1^{j_1} \dots x_n^{j_n}$, если для некоторого номера k

имеет место неравенство $i_k > j_k$ и при этом $i_l = j_l$ для всех $l < k$. Для краткости в таких случаях можно писать $x_1^{i_1} \dots x_n^{i_n} \succ x_1^{j_1} \dots x_n^{j_n}$. Старший член считается главнее другого старшего члена, если его моном главнее.

Утверждение. *Главный член произведения многочленов равен произведению их главных членов.*

Доказательство. Достаточно принять во внимание, что если один моном главнее другого, то это свойство сохраняется при умножениии обоих мономов на один и тот же моном. □

Следствие. *Степень произведения многочленов равна сумме степеней перемножаемых многочленов.*

Замечание. Лексикографический принцип можно использовать для упорядочивания не только старших членов, но вообще всех ненулевых членов минимального представления многочлена. В этом случае у нас появляется понятие однозначно определенного *лексикографически старшего члена*. Вообще говоря, это будет член, отличающийся от ранее определенного главного члена. Однако, свойство перемножения при умножении многочленов для лексикографически старших членов сохраняется: их произведение дает такой член для произведения (проверьте!).

8.4 Кольцо многочленов

Утверждение. *Множество $\mathbb{P}[x_1, \dots, x_n]$ вместе с операциями сложения и умножения многочленов является целостным кольцом с единицей.*

Доказательство. Ассоциативность и коммутативность каждой операции и дистрибутивность умножения относительно сложения прямо вытекают из аналогичных свойств сложения и умножения в поле. Роль единицы играет константа, равная единице.

Теперь докажем целостность. От противного, допустим, что есть делители нуля, т.е. $f(x)g(y) = 0$ и оба многочлена ненулевые. Тогда они обладают главными членами с ненулевыми числовыми коэффициентами, произведение которых равно нулю. Противоречие с тем, что в поле делителей нуля нет. □

Кольцо многочленов над полем является примером *градуированного кольца* — так называются кольца, ненулевые элементы которых можно разбить на непересекающиеся подмножества K_1, K_2, ... такие, что произведение элементов из K_m и K_n принадлежит K_{m+n}. Для кольца многочленов к подмножеству K_n можно отнести, например, все многочлены степени n.

Заметим, что операцию умножения многочлена над полем $\mathbb{P}$ на число из этого поля можно естественным образом определить как умножение данного многочлена на многочлен-константу. Таким образом, кольцо $\mathbb{P}[x_1, \dots, x_n]$ можно считать также линейным пространством над полем $\mathbb{P}$, а значит и алгеброй над этим же полем.

Пространство всех многочленов — это хороший пример бесконечномерного линейного пространства (докажите!) и бесконечномерной ассоциативной коммутативной алгебры с единицей. Если в пространстве всех многочленов взять только многочлены степени не выше d, то получится конечномерное подпространство. В случае одной переменной это будет подпространство размерности $d + 1$ (докажите!).

Задача 1. *Докажите, что для многочленов над полем вычетов по простому модулю p имеет место равенство $(z-1)^p = z^p - 1$.*

Задача 2. *Докажите, что размерность линейного пространства многочленов от n переменных степени не выше k равна числу сочетаний из $n+k$ по k.*

8.5 Значения и корни

Если $f(x) \in \mathbb{P}[x]$, то для любого числа $\theta \in \mathbb{P}$ естественным образом определяется число $f(\theta) \in \mathbb{P}$. Оно называется *значением многочлена* в точке θ или при $x = \theta$. Если $f(\theta) = 0$, то число θ называется *нулем* или *корнем* многочлена $f(x)$.

Верно ли, что ненулевой многочлен хотя бы в одной точке принимает ненулевое значение? В общем случае нет. Например, для многочлена $f(x) = x + x^2$ над полем $\mathbb{Z}_2$ находим $f(\theta) = 0 \quad \forall\ \theta \in \mathbb{Z}_2$. Этот же пример показывает, что совпадение функций, заданных двумя многочленами, в общем случае не означает, что эти многочлены совпадают как формальные выражения.

В случае n переменных часто вводят векторную переменную $x = (x_1, \ldots, x_n)$, рассматривают многочлен $f(x) = f(x_1, \ldots, x_n)$ как функцию точки $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{P}^n$ и называют число $f(\theta) = f(\theta_1, \ldots, \theta_n) \in \mathbb{P}$ его значением в точке θ.

8.6 Делимость в целостном унитальном кольце

Чрезвычайно важное понятие делимости многочленов кольца $\mathbb{P}[x_1, \ldots, x_n]$ проще исследовать в более общей ситуации, рассматривая понятие делимости в целостном унитальном кольце R. Нам понадобятся также многочлены, в которых коэффициенты берутся из кольца R. Операции с такими многочленами вводятся так же, как и в случае многочленов над полем. В результате возникает кольцо $R[x_1, \ldots, x_n]$, которое наследует целостность и унитальность кольца R (докажите!).

Говорят, что элемент $f \in R$ *делится* на ненулевой элемент $g \in R$, если существует элемент $q \in R$ такой, что, что $f = gq$. В таких случаях говорят также, что g *является делителем* f или, короче, g *делит* f. Фраза "f делится на g" записывается с помощью вертикального троеточия в виде $f \vdots g$, а для того же самого факта, выраженного фразой "g делит f", используется вертикальная черта: $g \mid f$.

При изучении делимости обычно нас интересуют все делители заданного ненулевого элемента f. Среди них заведомо есть все делители единицы кольца R и все элементы, полученные умножением f на произвольный делитель единицы. Такие делители элемента f называются *тривиальными*. Любой ненулевой элемент, имеющий только тривиальные делители и отличный от делителя единицы, называется *неприводимым*. Например, в кольце целых чисел множество делителей единицы состоит из двух чисел ± 1, а неприводимые элементы имеют вид $\pm p$, где p — простое число. В кольце многочленов над полем множество делителей единицы совпадает с множеством ненулевых констант (докажите!).

Целостное унитальное кольцо называется *факториальным* или *гауссовым*, если в нем любой ненулевой элемент, отличный от делителя единицы, имеет разложение в произведение неприводимых множителей со свойством единственности с точностью до перестановки множителей и их умножения на делители единицы.

Пусть f и g — два ненулевых элемента кольца R. Они называются *взаимно простыми*, если любой их общий делитель является делителем единицы. Если элемент $a \in R$

является общим делителем элементов f и g и делится на любой их общий делитель, то он называется *наибольшим общим делителем* элементов f и g. Если элемент $b \in R$ делится одновременно на f и на g и делит любой элемент с таким же свойством, то он называется *наименьшим общим кратным* элементов a и b. Если кольцо факториально, то $ab = fg$ с точностью до умножения на делитель единицы (докажите!).

Свойство факториальности кольца целых чисел составляет содержание известной школьникам *основной теоремы арифметики*. Очень полезно знать, что этим свойством обладает также любое кольцо многочленов от нескольких переменных над произвольным полем. Чтобы установить этот (не очень простой) факт, нам потребуется теория деления с остатком в случае многочленов от одной переменной над полем. Кроме того, нам нужно будет научиться строить поле дробей (частных), содержащее в себе заданное целостное кольцо, и познакомиться с красивой леммой Гаусса о делимости многочленов с целыми коэффициентами, а также ее обобщениями и следствиями.

8.7 Деление с остатком

Пусть $f(x)$ и $g(x) \neq 0$ — многочлены от одной переменной. Если имеет место представление

$$f(x) = g(x)q(x) + r(x), \quad \text{где} \quad r(x) = 0 \quad \text{либо} \quad \deg r(x) < \deg g(x),$$

то говорят, что $f(x)$ *делится с остатком* на $g(x)$. Многочлен $r(x)$ называется *остатком*, а многочлен $q(x)$ — *неполным частным*.

В общем случае можно считать, что коэффициенты всех многочленов принадлежат фиксированному кольцу, и нужно заметить, что деление с остатком не всегда можно выполнить. В самом деле, старший член делимого $f(x)$ должен делиться на старший член делителя $g(x)$, а над кольцом $\mathbb{Z}$ заведомо есть многочлены, для которых это свойство не выполняется. Например, $f(x) = x^3 + 1$ и $g(x) = 2x - 1$.

В то же время, *если делитель является приведенным многочленом, то деление с остатком всегда выполнимо* и, более того, реализуется школьным методом деления столбиком (докажите!). В особенно интересном для нас случае многочленов над полем справедлива следующая

Теорема о делении с остатком. *Для любой пары многочленов $f(x)$ и $g(x) \neq 0$ над полем деление с остатком выполнимо и при этом остаток и неполное частное определяются однозначно.*

Доказательство. Пусть $f(x) = a_n x^n + \ldots + a_0$, $g(x) = b_m x^m + \ldots + b_0$, $b_m \neq 0$. Если $\deg f(x) < \deg g(x)$, то полагаем $q(x) = 0$ и $r(x) = f(x)$. Если $\deg f(x) \geq \deg g(x)$, то реализуем деление столбиком:

$$f_1(x) := f(x) - \left(\frac{a_n}{b_m}\, x^{n-m}\right) g(x) \quad \Rightarrow \quad \deg f_1(x) < \deg f(x) \quad \text{либо} \quad f_1(x) = 0.$$

Предположим (индукция по степени делимого), что для $f_1(x)$ уже найдено разложение

$$f_1(x) = g(x)q_1(x) + r_1(x), \qquad \deg r_1(x) < \deg g(x) \quad \text{либо} \quad r_1(x) = 0.$$

Тогда искомое деление с остатком 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 $g(x)q(x)+r(x) = g(x)q_1(x)+r_1(x)$ — две реализации деления с остатком. Тогда $g(x)(q(x)-q_1(x)) = r(x)-r_1(x)$. Если $q(x) \neq q_1(x)$, то степень многочлена в левой части не меньше степени $g(x)$. Поэтому $\deg(r(x)-r_1(x)) \geq \deg g(x)$, а так быть не может, потому что степень разности многочленов не выше степени каждого из них. Следовательно, $q(x) = q_1(x)$ и $r(x) = r_1(x)$. $\square$

Задача 3. *Докажите, что в кольце многочленов над любым полем существует бесконечно много неприводимых многочленов.*

8.8 Теорема Безу

Теорема Безу. *Пусть $\mathbb{P}$ — произвольное поле, $\theta \in \mathbb{P}$ и $r(x)$ — остаток при делении многочлена $f(x)$ на многочлен $x-\theta$. Тогда $r(\theta) = f(\theta)$.*

Доказательство. Выполнив деление с остатком, находим $f(x) = (x-\theta)q(x) + r(x)$ и сравниваем значения левой и правой частей в точке $x = \theta$. $\square$

Следствие 1. *Если θ — корень многочлена $f(x)$, то $f(x)$ делится на $x-\theta$.*

Следствие 2. *Многочлен $f(x)$ степени n не может иметь более чем n корней.*

8.9 Алгоритм Евклида

Алгоритм Евклида получает на вход ненулевые многочлены $f(x)$ и $g(x)$ с коэффициентами из заданного поля и вычисляет их наибольший делитель. Он представляет собой последовательность делений с остатком:

$$\begin{array}{rcrcll}
f(x) & = & g(x)\,q_1(x) & + & r_1(x), & \deg r_1(x) < \deg g(x), \\
g(x) & = & r_1(x)\,q_2(x) & + & r_2(x), & \deg r_2(x) < \deg r_1(x), \\
r_1(x) & = & r_2(x)\,q_3(x) & + & r_3(x), & \deg r_3(x) < \deg r_2(x), \\
\dots & \dots & \dots & & & \\
r_{k-2}(x) & = & r_{k-1}(x)\,q_k(x) & + & r_k(x), & \deg r_k(x) < \deg r_{k-1}(x), \\
r_{k-1}(x) & = & r_k(x)\;q_{k+1}(x). & & &
\end{array}$$

На каждом шаге степень остатка понижается, $r_k(x)$ — последний ненулевой остаток.

Утверждение. *Последний ненулевой остаток в алгоритме Евклида, примененном к многочленам $f(x)$ и $g(x)$, является их наибольшим общим делителем.*

Доказательство. Последнее равенство показывает, что $r_{k-1}(x) \vdots r_k(x)$. Из предпоследнего видно, что $r_{k-2}(x) \vdots r_k(x)$. Просматривая равенства снизу вверх, приходим к выводу о том, что $r_k(x)$ является общим делителем для $f(x)$ и $g(x)$. Пусть $d(x)$ — любой их общий делитель. Просматривая те же равенства сверху вниз, получаем, что $d(x)$ делит $r_k(x)$. $\square$

8.10 Теорема о наибольшем общем делителе

Теорема о наибольшем общем делителе. *Для любых ненулевых многочленов $f(x)$ и $g(x)$ над некоторым полем существуют многочлены $u(x)$ и $v(x)$ над тем же полем, для которых многочлен $d(x) = f(x)u(x)+g(x)v(x)$ является наибольшим общим делителем исходных многочленов и, кроме того, если многочлены отличны от констант, то среди них есть такие, для которых $\deg u(x) < \deg g(x)$, $\deg v(x) < \deg f(x)$.*

Доказательство. Первое и второе равенства алгоритма Евклида можно, очевидно, записать в виде

$$r_1(x) = f(x)u_1(x) + g(x)v_1(x), \quad r_2(x) = f(x)u_2(x) + g(x)v_2(x),$$

$$u_1(x) = 1, \quad v_1(x) = -q_1(x), \quad u_2(x) = -q_2(x), \quad v_2(x) = 1 + q_1(x)q_2(x).$$

Пусть уже получены равенства

$$r_{i-2}(x) = f(x)\,u_{i-2}(x) + g(x)\,v_{i-2}(x), \quad r_{i-1}(x) = f(x)\,u_{i-1}(x) + g(x)\,v_{i-1}(x).$$

Тогда, учитывая их при получении остатка $r_i(x)$, находим

$$r_i(x) = f(x)\,u_i(x) + g(x)\,v_i(x),$$

$$u_i(x) = u_{i-2}(x) - u_{i-1}(x)\,q_i(x), \quad v_i(x) = v_{i-2}(x) - v_{i-1}(x)\,q_i(x).$$

Искомое равенство возникает при $i = k$. Подсчитывая степени, находим

$$\deg f(x) = \sum_{i=1}^{k+1} \deg q_i(x) + \deg r_k(x), \quad \deg g(x) = \sum_{i=2}^{k+1} \deg q_i(x) + \deg r_k(x),$$

$$\deg u_k(x) \le \sum_{i=2}^{k} q_i(x) \le \deg g(x) - \deg q_{k+1}(x),$$

$$\deg v_k(x) \le \sum_{i=1}^{k} q_i(x) \le \deg f(x) - \deg q_{k+1}(x).$$

Остается заметить, что $\deg q_{k+1}(x) \ge 1$, иначе не могло бы выполняться неравенство $\deg r_k(x) < \deg r_{k-1}(x)$. Таким образом, многочлены $u(x) = u_k(x)$ и $v(x) = v_k(x)$ являются искомыми. $\square$

8.11 Факториальность кольца $\mathbb{P}[x]$

Лемма о делимости. *Пусть $a(x), b(x), c(x)$ — ненулевые многочлены над некоторым полем, и пусть известно, что $c(x)$ взаимно прост с $b(x)$ и делит произведение $a(x)b(x)$. Тогда $c(x)$ делит $a(x)$.*

Доказательство. Теорема о наибольшем общем делителе гарантирует равенство вида $b(x)u(x) + c(x)v(x) = 1$. После умножения обеих части на $a(x)$ находим

$$a(x) = (a(x)b(x))\,u(x) \;+\; c(x)\,(v(x)a(x)) \;\vdots\; c(x). \quad \square$$

Теорема о факториальности кольца многочленов от одной переменной. *Кольцо многочленов от одной переменной над полем факториально.*

Доказательство. Пусть $f(x)$ — произвольный ненулевой многочлен. Если он неприводим, то разложение на неприводимые множители уже есть и состоит из одного множителя. В противном случае повторяем аналогичное рассуждение для каждого делителя многочлена $f(x)$. Таким образом, существование разложения на неприводимые множители достаточно очевидно.

Докажем единственность. Будем вести индукцию по числу s неприводимых множителей. Если $s = 1$, то все разложения того же многочлена имеют только один множитель (почему?). Теперь предположим, что $s \geq 2$ и для всех многочленов, обладающих разложением с числом неприводимых множителей $\leq s-1$ единственность уже доказана. Рассмотрим два разложения $f(x) = p_1(x)\dots p_s(x) = q_1(x)\dots q_t(x)$. Согласно доказанной выше лемме, какой-то из многочленов $q_1(x), \dots, q_t(x)$ должен делиться на $p_s(x)$. Для определенности пусть это будет $q_t(x)$. В силу неприводимости многочленов мы получаем равенство $q_t(x) = cp_s(x)$, где $c \in \mathbb{P}$. Таким образом,

$$p_s(x)(p_1(x)\dots p_{s-1}(x) - cq_1(x)\dots q_{t-1}(x)) = 0 \quad \Rightarrow \quad p_1(x)\dots p_{s-1}(x) = (cq_1(x))\dots q_{t-1}(x).$$

Согласно индуктивному предположению, два последних разложения обязаны иметь одно и то же число неприводимых множителей, которые могут отличаться только порядком и ненулевыми числовыми коэффициентами. $\square$

8.12 Поле дробей

Мы докажем здесь, что любое целостное кольцо изоморфно подкольцу некоторого поля. В таких случаях обычно говорят, что кольцо можно *расширить* до поля или *вложить* в поле. Под изоморфизмом, как обычно, понимается взаимно-однозначное отображение, сохраняющее операции. В случае колец должны сохраняться операции сложения и умножения.

Заметим, что кольцо K, составленное из части элементов поля $\mathbb{P}$ и наследующее операции данного поля, обязано быть целостным (почему?). Пусть L — минимальное подполе, содержащее кольцо K (пересечение всех подполей поля $\mathbb{P}$, содержащих K). Тогда в L должны находиться все элементы вида a/b, где $a, b \in K$ и $b \neq 0$. Нетрудно проверить, что множество дробей a/b образует поле, и это оначает, что никаких других элементов в поле L нет. Это наблюдение подсказывает, как нужно строить минимальное поле в том случае, когда задано лишь кольцо K.

Пусть K — произвольное целостное кольцо. *Формальной дробью* или *формальным частным* называется пара его элементов, записываемая в виде a/b и при условии $b \neq 0$. Элемент a называется *числителем*, а элемент b — *знаменателем*. Символ деления здесь носит формальный характер и не является обозначением операции деления.

Прежде всего нам нужно ввести правило отождествления разных дробей. По определению, $a/b = c/d$ означает, что $ad = bc$. Это отношение на множестве дробей рефлексивно ($a/b = a/b$), симметрично ($a/b = c/d \;\Rightarrow\; c/d = a/b$) и транзитивно ($a/b = c/d,\ c/d = p/q \;\Rightarrow\; a/b = p/q$). Проверка транзитивности: пусть $ad = bc,\ \ cq = dp \;\Rightarrow\; adcq = bcdp \;\Rightarrow\; cd(aq - bp) = 0$. Отсюда, в силу отсутствия делителей нуля, $aq = bp \;\Rightarrow\; a/b = p/q$. Как и должно быть при отождествлении, мы получаем отношение эквивалентности, и все множество дробей разбивается на непересекающие классы дробей, который считаются равными.

Так же, как в случае направленных отрезков, для указания на класс эквивалентности испольуется любая принадлежащая ему дробь, а сложение и умножение классов определяются через их представителей:

$$a/b + c/d := (ad + bc)/(bd), \quad (a/b)(c/d) := (ac)/(bd),$$

и поэтому здесь необходима *проверка корректности*. Она заключается в доказательстве следующего свойства:

$$a/b = A/B,\ \ c/d = C/D \ \ \Rightarrow \ \ (AD+BC)/(BD) = (ad+bc)/(bd),\ \ (AC)/(BD) = (ac)/bd),$$

и выполняется вполне рутинно:

$$(AD + BC)(bd) - (BD)(ad + bc) = ADbd + BCbd - BDad - BDbc =$$

$$(Ab - Ba)Dd + (Cd - Dc)Bb = 0,$$

$$(AC)(bd) - (BD)(ac) = AbCd - aBCd + aBCd - BcDa =$$

$$(Ab - Ba)(Cd) + (Cd - Dc)(Ab) = 0.$$

Теорема. *Множество формальных дробей над целостным кольцом является полем.*

Доказательство. Коммутативность сложения и умножения очевидна. Ассоциативность сложения:

$$(a/b+c/d)+p/q = (ad+bc)/(bd)+p/q = ((ad+bc)q+bdp)(bdq) = (adq+bcq+bdp)/(bdq),$$

$$a/b+(c/d+p/q) = a/b+(cq+dp)/(dq) = (adq+b(cq+dp))/(bdq) = (adq+bcq+bdp)/(bdq).$$

Нулевой элемент: $0 = 0/c,\ a/b+0/c = (ac+0\cdot b)/(bc) = a/b$. Противоположный элемент: $a/b + (-a)/b = (ab - ab)/b^2 = 0/b^2 = 0$. Дистрибутивность умножения относительно сложения:

$$(a/b)(c/d + p/q) = (a(cq + dp))/(bdq) = (acq + adp)/(bdq),$$

$$(a/b)(c/d) + (a/b)(p/q) = (ac)/(bd) + (ap)/(bq) = (acbq + bdap)/(bdbq) = (acq + adp)/(bdq).$$

Единичный элемент: $1 = c/c,\ \ (a/b)(c/c) = (ac)/bc) = a/b$. Обратный элемент: $(a/b)(b/a) = (ab)/(ab) = 1$. $\square$

Заметим, что элементы x исходного кольца K можно отождествить с дробями вида $(xa)/a$, где a — фиксированный ненулевой элемент кольца K.

Поле дробей для кольца $\mathbb{Z}$ целых чисел называется *полем рациональных чисел* и обозначается через $\mathbb{Q}$. Поле дробей для кольца многочленов $\mathbb{P}[x_1, \ldots, x_n]$ называется *полем рациональных функций* и обозначается через $\mathbb{P}(x_1, \ldots, x_n)$.

Обратим внимание на то, что мы имеем право рассматривать систему $Ax = b$ как связь между элементами поля дробей для кольца многочленов от букв, обозначающих элементы матрицы и правой части. В этом случае предполагается, что искомые элементы вектора x — это какие-то элементы того же поля дробей. Если матрица A квадратная порядка n, то ее определитель будет однородным многочленом степени n, и это значит, что мы имеем систему с невырожденной матрицей коэффициентов, решение которой выражается по формулам Крамера.

Задача 4. *Дана отличная от константы рациональная функция $f(x)$ от переменной x. Докажите, что рациональная функция $g(u, v) := f(u/v)$ от переменных u и v представима в виде отношения двух однородных многочленов одной и той же степени.*

8.13 Лемма Гаусса

Лемма Гаусса для кольца целых чисел. *Пусть $a(x), b(c), c(x)$ — многочлены над кольцом целых чисел, связанные равенством $c(x) = a(x)b(x)$, и пусть p — простое число. Тогда если $c(x) \vdots p$, то $a(x) \vdots p$ или $b(x) \vdots p$.*

Доказательство. Пусть $a(x) = a_0 + a_1x + \ldots + a_mx^m$, $b(x) = b_0 + b_1x + \ldots + b_nx^n$, $c(x) = c_0 + c_1x + \ldots + c_{m+n}x^{m+n}$. От противного, пусть ни один из многочленов $a(x)$ и $b(x)$ не делится на p. Тогда существуют индексы k и l такие, что a_k и b_l не делятся на p, а каждый из коэффициентов a_i при $i < k-1$ и b_j при $j < l$ делится на p. Следовательно,

$$c_{k+l} = a_kb_l + (a_{k-1}b_{l+1} + \ldots a_0b_{l+k}) + (a_{k+1}b_{l-1} + \ldots + a_{k+l}b_0) \quad \not\vdots p \quad \Rightarrow \quad c(x) \quad \not\vdots p. \quad \square$$

Такое же утверждение справедливо в более общей ситации — для многочленов над произвольным факториальным кольцом, в котором роль простых чисел играют неприводимые элементы.

Лемма Гаусса для факториального кольца. *Пусть $a(x), b(c), c(x)$ — многочлены над произвольным факториальным кольцом, связанные равенством $c(x) = a(x)b(x)$, и пусть p — неприводимый элемент кольца. Тогда если $c(x) \vdots p$, то $a(x) \vdots p$ или $b(x) \vdots p$.*

Доказательство. Рассуждение, приведенное в случае кольца чисел чисел, сохраняется дословно с единственным отличием: теперь под p понимается неприводимый элемент кольца, из которого берутся коэффициенты многочленов. $\square$

Следствие. *Пусть K — произвольное факториальное кольцо, K' — его поле дробей и $f(x) \in K[x] \subset K'[x]$. Тогда разложение $f(x) = a(x)b(x)$ над полем K' имеет место в том и только том случае, когда существует разложение $f(x) = A(x)B(x)$ над кольцом K и при этом $A(x) = \alpha a(x)$ и $B(x) = \beta b(x)$ для каких-то $\alpha, \beta \in K'$.*

Доказательство. Пусть $f(x) = a(x)b(x)$, где каждый коэффициент многочленов $a(x)$ и $b(x)$ является дробью, в которой числитель и знаменатель принадлежат кольцу K. Обозначим через u и v произведения знаменателей соответственно для $a(x)$ и $b(x)$. Тогда

$$uvf(x) = A_0(x)B_0(x), \quad A_0(x) := ua(x) \in K[x], \quad B_0(x) := vb(x) \in K[x].$$

Разложим uv в произведение неприводимых элементов $uv = p_1 \ldots p_s$ и применим лемму Гаусса сначала для $p = p_s$. Если $A_0(x) \vdots p$, то полагаем $A_1(x) := A_0(x)/p \in K[x]$ и $B_1(x) := B_0(x)$. Если $B_0(x) \vdots p$, то полагаем $A_1(x) := A_0(x)$ и $B_1(x) := B_0(x)/p \in K[x]$. В итоге получаем разложение над K вида

$$p_1 \ldots p_{s-1}f(x) = A_1(x)B_1(x).$$

Теперь применим лемму Гаусса для $p = p_{s-1}$ и найдем разложение над K вида

$$p_1 \ldots p_{s-2}f(x) = A_2(x)B_2(x).$$

Продолжая в том же духе, на s-м шаге мы получим разложение $f(x) = A_s(x)B_s(x)$, которое и является искомым разложением над K. $\square$

8.14 Факториальность кольца $\mathbb{P}[x_1, \ldots, x_n]$

Теорема о наследовании факториальности. *Кольцо многочленов от одной переменной над факториальным кольцом является факториальным.*

Доказательство. Пусть K — факториальное кольцо и $f(x) \in K[x]$ — ненулевой многочлен. Заметим, что $f(x)$ можно рассматривать также как многочлен над полем дробей K' кольца K. Поскольку мы уже знаем, что кольцо многочленов от одной переменной над полем факториально, для $f(x)$ существует разложение $f(x) = f_1(x) \ldots f_s(x)$, где $f_1(x), \ldots, f_s(x)$ — неприводимые многочлены над K', и значит, согласно следствию леммы Гаусса, существует также разложение $f(x) = F_1(x) \ldots F_s(x)$ в произведение многочленов над K.

Пусть d_i — наибольший общий делитель коэффициентов многочлена $F_i(x)$. Тогда $F_i(x) = d_i \phi_i(x)$, где $\phi_i(x)$ — неприводимый многочлен над K, отличный от константы. Произведение $d = d_1 \ldots d_s$ — это элемент кольца K, и, в силу факториальности, существует разложение $d = p_1 \ldots p_t$ на неприводимые элементы кольца K. Таким образом, разложение $f(x) = p_1 \ldots p_t \, \phi_1(x) \ldots \phi_s(x)$ содержит неприводимые над K многочлены $p_1, \ldots, p_t$ нулевой степени (константы) и неприводимые над K многочлены $\phi_1(x), \ldots, \phi_s(x)$ ненулевой степени.

Докажем единственность. Пусть имеется еще одно разложение с неприводимыми многочленами-константами $q_1, \ldots, q_l$ и неприводимыми многочленами $\psi_1(x), \ldots, \psi_k(x)$ степени выше нулевой (коэффициенты каждого из них взаимно просты):

$$p_1 \ldots p_t \, \phi_1(x) \ldots \phi_s(x) \;=\; q_1 \ldots q_l \, \psi_1(x) \ldots \psi_k(x).$$

Согласно лемме Гаусса, $q_1 \ldots q_l \vdots p_s$. После перенумерации можно считать, что $q_l \vdots p_s$. Теперь нужно заметить, что $q_1 \ldots q_{l-1} \vdots p_{s-1}$, после перенумерации можно считать,что $q_{l-1} \vdots p_{s-1}$, и так далее.

В итоге ясно, что $l = s$, а константы $q_1, \ldots, q_s$ получаются из констант $p_1, \ldots, p_s$ перестановкой и умножением на делители единицы кольца K.

Таким образом, у нас возникает равенство $\varepsilon \phi_1(x) \ldots \phi_s(x) = \psi_1(x) \ldots \phi_k(x)$, в котором ε есть делитель единицы кольца K. Факториальность кольца многочленов от одной переменной над полем нам уже известна. Поэтому $k = s$, и ясно, что после переупорядочивания можно полагать, что $\phi_i(x) = (u_i/v_i)\psi_i(x)$, где u_i и v_i — взаимно простые элементы кольца K. Согласно лемме Гаусса, из равенства $v_i \phi_i(x) = u_i \psi_i(x)$ следует, что $\psi_i(x) \vdots v_i$ и $\phi_i(x) \vdots u_i$, и, в силу взаимной простоты коэффициентов каждого из этих многочленов, u_i и v_i должны быть делителями единицы кольца K. $\square$

Теорема о факториальности кольца многочленов над полем. *Для произвольного поля $\mathbb{P}$ кольцо многочленов $\mathbb{P}[x_1, \ldots, x_n]$ факториально.*

Доказательство. Индукция по n: факториальность в случае $n = 1$ уже установлена, а при $n \geq 2$ нужно лишь заметить, что $K[x_1, \ldots, x_n] = (K[x_1, \ldots, x_{n-1}])\,[x_n]$. $\square$

Задача 5. *Поскольку каждое поле содержит числа ± 1, определитель $n \times n$-матрицы можно рассматривать как многочлен от ее элементов над произвольно выбранным полем. Докажите, что определитель является неприводимым многочленом.*

Задача 6. *О многочленах $a, b, c \in \mathbb{P}[x_1, \ldots, x_n]$ над произвольным полем $\mathbb{P}$ известно, что $ab \vdots c$ и при этом b и c взаимно просты. Докажите, что $a \vdots c$.*

8.15 Определитель Вандермонда

Теорема о факториальности имеет много полезных применений. Например, при вычислении *определителя Вандермонда*

$$V(x_1,\dots,x_n)=\begin{vmatrix}1 & x_1 & x_1^2 & \dots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1}\\ \dots & \dots & \dots & \dots & \dots\\ 1 & x_n & x_n^2 & \dots & x_n^{n-1}\end{vmatrix}$$

можно считать, что это есть многочлен от переменных $x_1,\dots,x_n$. Кроме того, для каждого i мы можем рассматривать V как многочлен от одной переменной x_i над полем рациональных функций от остальных переменных. Ясно, что $V=0$ при $x_i=x_j$ для любого $j\neq i$, и значит, по теореме Безу, $V(x_1,\dots,x_n)\,\vdots\, x_i-x_j$. Многочлены x_i-x_j при $i>j$ попарно взаимно просты. Поэтому

$$V(x_1,\dots,x_n)\,\vdots\,\prod_{i>j}(x_i-x_j).$$

Степень делителя равна $n(n-1)/2$ и, как нетрудно убедиться, совпадает со степенью делимого (проверьте!). Один из старших мономов многочлена V имеет вид $x_n^{n-1}x_{n-1}^{n-2}\dots x_3^2x_2$, а соответствующий этому моному коэффициент равен 1 (проверьте!). Произведение одночленов x_i-x_j при всех $n\geq i>j\geq 1$ имеет в точности такой же старший член. Следовательно,

$$V(x_1,\dots,x_n)\;=\;\prod_{i>j}(x_i-x_j).$$

Задача 7. *Пусть заданы переменные $x_1,\dots,x_n$, $y_1,\dots,y_n$ и A — матрица порядка n с элементами $a_{ij}=1/(x_i+y_j)$ (такая матрица называется* `матрицей Коши`*). Докажите, что*

$$|A|=\frac{\prod\limits_{i>j}(x_i-x_j)(y_i-y_j)}{\prod\limits_{i,j}(x_i+y_j)}.$$

Задача 8. *Пусть заданы вещественные числа $0<x_1<\ldots<x_n$ и $\alpha_1<\ldots<\alpha_n$. Докажите, что $n\times n$-матрица с элементами $a_{ij}=x_i^{\alpha_j}$ невырождена.*

Задача 9. *Докажите, что при условии $0<x_1<\ldots<x_n$ все миноры матрицы Вандермонда $V(x_1,\dots,x_n)$ положительны.*

8.16 Формулы Виета и симметрические многочлены

Пусть $\mathbb{P}$ — произвольное поле и $K=\mathbb{P}(x_1,\dots,x_n)$ — поле рациональных функций от переменных $x_1,\dots,x_n$.

Утверждение. *Если многочлен $a(x)\in K[x]$ имеет вид*

$$a(x)=(x-x_1)(x-x_2)\dots(x-x_n)=a_0+a_1x+\ldots+a_nx^n.$$

то его коэффициенты выражаются следующим образом:

$$a_{n-k}=(-1)^kE_k(x_1,\dots,x_n),\quad 0\leq k\leq n, \tag{1}$$

$$E_k(x_1,\dots,x_n)=\sum_{1\leq i_1<\ldots<i_k\leq n}x_{i_1}\dots x_{i_k},\quad 1\leq k\leq n,\quad E_0(x_1,\dots,x_n)=1. \tag{2}$$

Доказательство проводится раскрытием скобок. Формулы (1), (2) называются *формулами Виета*, а многочлены $E_k(x_1,\dots,x_n)$ — *элементарными симметрическими многочленами*. Многочлен $f(x_1,\dots,x_n)$ называется *симметрическим*, если он не изменяется при любой перестановке переменных:

$$f(x_1,\dots,x_n)=f(x_{\sigma(1)},\dots,x_{\sigma(n)})\;\;\forall\;\sigma\in S_n.$$

Заметим, что элементарные симметрические многочлены можно рассматривать над любым коммутативным ассоциативным кольцом с единицей.

Теорема о симметрических многочленах. *Для любого симметрического многочлена $f(x_1, \ldots, x_n)$ с коэффициентами из коммутативного ассоциативного кольца с единицей существует и единствен многочлен $g(y_1, \ldots, y_n)$ над тем же кольцом, такой что*

$$f(x_1, \ldots, x_n) \;=\; g(E_1(x_1, \ldots, x_n), \ldots, E_n(x_1, \ldots, x_n)).$$

Доказательство. Пусть $a\,x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ — лексикографически старший член многочлена $f(x_1, \ldots, x_n)$. Тогда для симметрического многочлена обязательно выполняются неравенства $\alpha_1 \geq \ldots \geq \alpha_n$ — вместе с мономом $x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ в состав симметрического многочлена входят все мономы, получаемые произвольной перестановкой того же набора степеней. Рассмотрим многочлен $\phi(y_1, \ldots, y_n) \;=\; a\,y_1^{\alpha_1-\alpha_2} \ldots y_{n-1}^{\alpha_{n-1}-\alpha_n} y_n^{\alpha_n}$ и заметим, что после замены переменных на элементарные симметрические многочлены он превращается в многочлен

$$\phi(E_1, \ldots, E_n) \;=\; a\,E_1^{\alpha_1-\alpha_2} \ldots E_{n-1}^{\alpha_{n-1}-\alpha_n} E_n^{\alpha_n} \tag{3}$$

от $x_1, \ldots, x_n$ с лексикографически старшим членом, равным

$$a\,x_1^{\alpha_1-\alpha_2} (x_1x_2)^{\alpha_2-\alpha_3} \ldots (x_1 \ldots x_{n-1})^{\alpha_{n-1}-\alpha_n} (x_1 \ldots x_{n-1}x_n)^{\alpha_n} \;=\; a\,x_1^{\alpha_1} \ldots x_n^{\alpha_n}. \tag{4}$$

Поэтому лексикографически старший член многочлена $f_1 = f - \phi(E_1, \ldots, E_n)$ будет младше лексикографически старшего члена для $f(x_1, \ldots, x_n)$. Аналогичным образом от f_1 можно перейти к многочлену с еще меньшим лексикографически 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 $g(y_1, \ldots, y_n) \neq 0$, то после замен $y_i = E_i(x_1, \ldots, x_n)$ и приведения подобных членов останется хотя бы один ненулевой член. Любой многочлен вида $aE_1^{\beta_1} \ldots E_n^{\beta_n}$ можно записать в виде (3), взяв $\alpha_i = \beta_i + \beta_{i+1} + \ldots + \beta_n$, $1 \leq i \leq n$. Как многочлен от $x_1, \ldots, x_n$, многочлен ϕ имеет в своем составе член вида (4), который будет для него лексикографически старшим. Для g как многочлена от $x_1, \ldots, x_n$ лексикографически старшим будет наивысший из членов такого вида. Он определен однозначно и поэтому не может сократиться при приведении подобных членов. □

Следствие. *Пусть числа $\xi_1, \ldots, \xi_n$ принадлежат некоторому расширению поля $\mathbb{P}$, а коэффициенты многочлена $f(x) = (x - \xi_1) \ldots (x - \xi_n)$ принадлежат полю $\mathbb{P}$. Тогда значение любого симметрического многочлена $\phi(x_1, \ldots, x_n)$ над полем $\mathbb{P}$ в точке $(\xi_1, \ldots, \xi_n)$ является элементом поля $\mathbb{P}$.*

Доказательство. В силу формул Виета, значения элементарных симметрических многочленов в точке $(\xi_1, \ldots, \xi_n)$ с точностью до знака совпадают с коэффициентами многочлена $f(x)$ и поэтому принадлежат полю $\mathbb{P}$, а любой симметрический многочлен является многочленом от элементарных симметрических многочленов. □

Одно из применений следствия связано с понятием *дискриминанта*. Для приведенного многочлена $f(x) = (x - \xi_1) \ldots (x - \xi_n)$ дискриминантом называется величина

$$D = \prod_{1 \leq i < j \leq n} (\xi_i - \xi_j)^2.$$

Согласно следствию, дискриминант является элементом поля $\mathbb{P}$, которое содержит коэффициенты многочлена, — даже в том случае, когда корни принадлежат более широкому полю. Понятно, что дискриминант отличен от нуля тогда и только тогда, когда все n корней многочлена попарно различны. Для квадратного многочлена $x^2 + a_1x + a_0 = (x - \xi_1)(x - \xi_2)$ получается всем знакомое выражение $D = (\xi_1 - \xi_2)^2 = (\xi_1 + \xi_2)^2 - 4\xi_1\xi_2 = a_1^2 - 4a_0$.

Задача 10. *Найти многочлен 3-й степени, корнями которого являются квадраты корней многочлена $z^3 - 2z - 5$.*

Задача 11. *Докажите, что* ньютоновы суммы $S_k(x_1, \ldots, x_n) := \sum\limits_{i=1}^{n} x_i^k$ *и элементарные симметрические многочлены $E_k(x_1, \ldots, x_n)$ связаны соотношениями*

$$\sum_{k=0}^{t-1} (-1)^k S_{t-k} E_k + (-1)^t t E_t = 0 \quad \textit{при} \quad 1 \leq t \leq n; \qquad \sum_{k=0}^{n} (-1)^k S_{t-k} E_k = 0 \quad \textit{при} \quad t > n.$$

8.17 Результант

Пусть заданы два многочлена

$$a(x) = a_0 + a_1x + \ldots + a_mx^m = a_m\prod_{i=1}^{m}(x - x_i), \quad a_m \neq 0,$$

$$b(x) = b_0 + b_1x + \ldots + b_nx^n = b_n\prod_{j=1}^{n}(x - y_j), \quad b_n \neq 0.$$

Матрица порядка $m + n$ вида (полагаем для определенности, что $m \geq n$)

$$S_{a,b} = \begin{bmatrix} a_m & a_{m-1} & \ldots & \ldots & \ldots & a_1 & a_0 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & a_m & a_{m-1} & \ldots & \ldots & \ldots & a_1 & a_0 \\ b_n & \ldots & b_1 & b_0 & & & & & \\ & \ddots & \ddots & \ddots & \ddots & & & & \\ & & b_n & \ldots & b_1 & b_0 & & & \\ & & & b_n & \ldots & b_1 & b_0 & & \\ & & & & \ddots & \ddots & \ddots & \ddots & \\ & & & & & b_n & \ldots & b_1 & b_0 \end{bmatrix}$$

известна как *матрица Сильвестра*. Ее определитель $R_{a,b} = |S_{a,b}|$ называется *результантом* многочленов $a(x)$ и $b(x)$.

Лемма. *Для многочленов a и b, рассматриваемых как многочлены от переменных $x, a_0, \ldots, a_m$, $b_0, \ldots, b_n$, существуют многочлены u и v от тех же переменных такие, что $au + bv = R_{a,b}$, и, кроме того, степени многочленов u и v по переменной x строго меньше n и m, соответственно.*

Доказательство. Умножим матрицу $S_{a,b}$ справа на столбец z с элементами x^{m+n-1}, x^{m+n-1}, …, x, 1 и заметим, что столбец $S_{a,b}\,z$ будет иметь элементы

$$x^{n-1}a(x),\ x^{n-2}a(x),\ \ldots,\ a(x),\ x^{m-1}b(x),\ x^{m-2}b(x),\ \ldots, b(x).$$

Вычисляя последнюю компоненту вектора-столбца z по правилу Крамера, получаем, что результант равен определителю такого же вида, в котором последний столбец заменен на столбец $S_{a,b}\,z$. □

Теорема. *Результант многочленов отличен от нуля тогда и только тогда, когда они взаимно просты.*

Доказательство. Считая коэффициенты многочленов $a(x)$ и $b(x)$ числами из некоторого поля, получаем равенство $a(x)u(x) + b(x)v(x) = R_{a,b}$. Если $R_{a,b} \neq 0$, то, очевидно, $a(x)$ и $b(x)$ взаимно просты. Если $R_{a,b} = 0$, то нужно учесть, что $\deg u(x) < \deg b(x)$ и $\deg v(x) < \deg a(x)$, и заметить, что эти условия гарантируют, что $a(x)$ и $b(x)$ имеют общий делитель ненулевой степени (проверьте!). □

Следствие. $R_{a,b} = a_m^n b_n^m \prod_{i=1}^{m}\prod_{j=1}^{n}(x_i - y_j)$.

Доказательство. Если $x_i = y_j$, то, очевидно, результант (как многочлен от $x_1, \ldots, x_m$, $y_1, \ldots, y_n$) делится на $x_i - y_j$. В силу попарной взаимной простоты многочленов вида $x_i - y_j$ получаем, что результант делится на их произведение. В случае $a_m = b_n = 1$, как легко видеть, имеет место равенство, а общий случай сводится к данному делением первой части строк результанта на a_m и второй части на b_n. □

Задача 12. *Многочлены $f(x, y)$ и $g(x, y)$ с коэффициентами из некоторого поля взаимно просты. Докажите, что система уравнений $f(x, y) = 0$, $g(x, y) = 0$ может иметь лишь конечное число решений.*

Алгебра и геометрия (1 поток)

Лекция 9

9.1 Умножение арифметических векторов

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 вещественные арифметические векторы вида $a = (a_1, a_2)$. Мы знаем, что они образуют вещественное линейное пространство с естественными операциями сложения векторов и умножения вектора на вещественное число. Мы бы хотели ввести еще и операцию умножения арифметических векторов, которая превращала бы это пространство в алгебру над полем вещественных чисел с такими приятными свойствами, как:

- наличие единицы и разрешимость уравнений вида $ax = b$ и $ya = b$ для любых $a \neq 0$ и b (такие алгебры называются *алгебрами с делением*),
- ассоциативность умножения,
- коммутативность умножения.

Условимся вектор вида $a = (a_1, 0)$ отождествлять с вещественным числом a_1 и писать $(a_1, 0) = a_1$. Естественно потребовать также, чтобы умножение вектора на вектор вида $(a_1, 0)$ было *согласовано* с умножением того же вектора на вещественное число a_1:

$$(a_1, 0) \cdot (b_1, b_2) = (b_1, b_2)(a_1, 0) = (a_1 b_1, a_1 b_2).$$

Таким образом, вектор $1 = (1, 0)$ будет единицей алгебры.

Вектор вида $\mathbf{i} = (0, 1)$ назовем *мнимой единицей*. Единица и мнимая единица образуют базис, и чтобы определить умножение, достаточно сказать, как перемножаются векторы базиса: $(a_1 + a_2\mathbf{i})(b_1 + b_2\mathbf{i}) = a_1 b_1 + (a_2 b_1 + a_1 b_2)\mathbf{i} + a_2 b_2 \mathbf{i}^2$. Как видим, достаточно научиться вычислять квадрат мнимой единицы. Если $\mathbf{i}^2 = u + v\mathbf{i}$, то

$$(a_1 + a_2\mathbf{i})(b_1 + b_2\mathbf{i}) = (a_1 b_1 + a_2 b_2 u) + (a_2 b_1 + a_1 b_2 + a_2 b_2 v)\mathbf{i}. \qquad (*)$$

Коммутативность имеет место при любом выборе u и v. Для проверки дистрибутивности и ассоциативности заметим, что отображение

$$(a_1, a_2) \to \begin{bmatrix} a_1 & a_2 \\ a_2 u & a_1 + a_2 v \end{bmatrix} = a_1 I + a_2 P, \quad \text{где} \quad P = \begin{bmatrix} 0 & 1 \\ u & v \end{bmatrix}, \quad P^2 = uI + vP,$$

является взаимно-однозначным отображением двумерного пространства вещественных арифметических векторов на подпространство вещественных матриц $L = \{a_1 I + a_2 P \ : \ a_1, a_2 \in \mathbb{R}\}$. Сумма и, кроме того, произведение любых двух матриц из L также принадлежат L, а сумма и произведение векторов по правилу $(*)$ отображаются в сумму и произведение матриц из L, соответствующих этим векторам (проверьте!). Отсюда сразу следует, что наша операция умножения при любом выборе вещественных значений u и v является дистрибутивной относительно сложения и ассоциативной (в силу того, что умножение матриц дистрибутивно относительно сложения и ассоциативно).

Утверждение. *Чтобы получилась алгебра с делением, необходимо и достаточно условие* $v^2 + 4u < 0$.

Доказательство. Рассмотрим уравнение $(a_1, a_2)(x_1, x_2) = (b_1, b_2)$ относительно вещественных чисел x_1, x_2 и заметим, что оно равносильно системе линейных алгебраических уравнений вида $\begin{bmatrix} a_1 & a_2 u \\ a_2 & a_1 + a_2 v \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Для разрешимости необходимо и достаточно, чтобы определитель $\Delta = a_1^2 + a_1 a_2 v - a_2^2 u$ был отличен от нуля всякий раз, когда $(a_1, a_2) \neq 0$. Если $a_2 = 0$, то условие разрешимости принимает вид $a_1 \neq 0 \ \Leftrightarrow \ \Delta = a_1^2 \neq 0$. Если $a_2 \neq 0$, то положим $t = a_1/a_2$ и заметим, что квадратный трехчлен $f(t) = t^2 - vt - u$ не обращается в нуль при вещественных t, а это равносильно отрицательности его дискриминанта: $v^2 + 4u < 0$. $\square$

Введем еще одно условие на умножение арифметических векторов — потребуем, чтобы длина вектора, получаемого при перемножении векторов, равнялась произведению длин векторов-сомножителей. Под длиной вещественного вектора (a_1, a_2) понимается величина $\sqrt{a_1^2 + a_2^2}$. Это дополнительное условие определяет u и v уже однозначно: $u = -1$ и $v = 0$ (докажите!).

9.2 Комплексные числа и комплексная плоскость

Комплексные числа — это вещественные арифметические векторы $z = (z_1, z_2)$ или точки на *комплексной плоскости* с координатами z_1 и z_2 в декартовой системе координат. Координаты называются соответственно *вещественной частью* и *мнимой частью* комплескного числа z и обозначаются $z_1 = \mathrm{Re}(z)$ и $z_2 = \mathrm{Im}(z)$. Величина $|z| := \sqrt{z_1^2 + z_2^2}$ называется *модулем* комплексного числа z. Если $z \neq 0$, то угол ϕ между положительным направлением первой оси и радиус-вектором z называется *аргументом* комплексного числа z и обозначается $\phi = \arg(z)$. Аргумент определен с точностью до прибавления числа, кратного 2π. Числу $z = 0$ можно приписать любое значение аргумента. *Тригонометрической формой* комплексного числа z называется его запись в виде $z = |z|(\cos\phi + \mathbf{i}\sin\phi)$. Если $z \neq 0$, то $\phi = \arg(z)$.

Множество всех комплексных чисел обозначается через $\mathbb{C}$ и содержит в себе множество вещественных чисел $\mathbb{R}$, которые отождествляются с векторами вида $z = (z_1, 0)$. Сложение комплексных чисел определяется как сложение арифметических векторов, а умножение обеспечивает равенство $\mathbf{i}^2 = -1$ и выполняется по правилу

$$(a_1 + a_2\mathbf{i})(b_1 + b_2\mathbf{i}) = (a_1 b_1 - a_2 b_2) + (a_1 b_2 + a_2 b_1)\mathbf{i}.$$

Множество $\mathbb{C}$ вместе с введенными на нем операциями сложения и умножения является полем, которое рассматривается как расширение поля $\mathbb{R}$ (проверьте!).

Заметим, что сумме комплексных чисел соответствует сумма соответствующих радиус-векторов. Отсюда возникает очень полезное *неравенство треугольника* $|u+v| \le |u|+|v|$ и не менее полезное его следствие $||u|-|v|| \le |u-v|$, имеющие место для любых комплексных чисел u и v.

Утверждение. *При умножении ненулевых комплексных чисел их модули перемножаются, а аргументы складываются.*

Доказательство. Пусть $u = |u|(\cos\phi + \mathbf{i}\sin\phi)$ и $v = |v|(\cos\phi + \mathbf{i}\sin\phi)$. Тогда $uv =$

$|u||v|((\cos\phi\cos\psi - \sin\phi\sin\psi) + \mathbf{i}(\sin\phi\cos\psi + \cos\phi\sin\psi)) = |u||v|(cos(\phi+\psi) + \mathbf{i}\sin(\phi+\psi)$. $\square$

Замечание. Есть очень удобное *обозначение Эйлера*: $\mathrm{e}^{\mathbf{i}\phi} = \cos\phi + \mathbf{i}\sin\phi$. В полном согласии с формальным применением известных свойств экспоненциальной функции оно приводит к равенству $\mathrm{e}^{\mathbf{i}\phi}\mathrm{e}^{\mathbf{i}\psi} = \mathrm{e}^{\mathbf{i}(\phi+\psi)}$.

Комплексное число $\bar{z} := z_1 - \mathbf{i}z_2$ называется *сопряженным* числу $z = z_1 + \mathbf{i}z_2$. На комплексной плоскости радиус-вектор для $\bar{z}$ получается из радиус-вектора для z симметричным отражением относительно первой оси. Заметим, что $\bar{z}\,z = |z|^2$.

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 обозначается через $\bar{A} = [\bar{a}_{ij}]$. Матрица $A^* := \left(\bar{A}\right)^\top$ называется *сопряженной матрицей* для A. Полезные свойства: $(AB)^* = B^*A^*$, $\det A^* = \overline{\det A}$, матрица A обратима тогда и только тогда, когда обратима сопряженная матрица A^*, при этом $(A^*)^{-1} = (A^{-1})^*$ (проверьте!).

Матрица $H \in \mathbb{C}^{n\times n}$ называется *самосопряженной* или *эрмитовой*, если $H^* = H$. Заметим, что для любой прямоугольной комплексной матрицы A произведения A^*A и AA^* будут эрмитовыми матрицами (проверьте!).

Применение комплексных чисел иногда упрощает получение интересных формул для вещественных чисел. Например, чтобы вычислить сумму $S_n = \sum\limits_{k=1}^{n} \cos k\phi$, заметим, что $S_n = \mathrm{Re}\left(\sum\limits_{k=1}^{n} z^k\right)$, где $z = \cos\phi + \mathbf{i}\sin\phi$. Таким образом, задача сводится с суммированию геометрической прогрессии: $S_n = \mathrm{Re}\left(\frac{z^{n+1}-z}{z-1}\right)$.

9.3 Преобразования комплексной плоскости

С помощью комплекных чисел можно задавать взаимно-однозначные отображения комплексной плоскости на себя.

Например, отображение $z \to z + w$ представляет собой параллельный перенос (сдвиг) точек на вектор, определенный комплексным числом w. Отображение $z \to wz$ при условии $|w| = 1$ 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). Если $w \ne 0$, то можно записать $w = |w|\,v$, где $|v| = 1$, и, следовательно, умножение на произвольное комплексное число $w \ne 0$ 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: $(\Phi\Psi)(z) = \Phi(\Psi(z))$. Полагаем, что оба отображения $\Phi(z) = a + bz$ и $\Psi(z) = c + dz$ принадлежат множеству $\mathcal{T}$. Тогда $|b| = |d| = 1 \;\Rightarrow\; |bd| = |b||d| = 1 \;\Rightarrow$ отображение $\Phi(\Psi(z)) = a+b(c+dz) = (c+bc)+(bd)z$ также принадлежит $\mathcal{T}$. Роль единичного элемента выполняет тождественное отображение $z \;\to\; z$, которое, очевидно, принадлежит $\mathcal{T}$. Далее, если $w = \Phi(z) = a+bz$, то обратное отображение имеет вид $z = \Phi^{-1}(w) = a-\bar{b}w$, и, поскольку $|-\bar{b}| = 1$, оно также принадлежит $\mathcal{T}$. При замене Φ на $\bar{\Phi}$ или Ψ на $\bar{\Psi}$ получаются отображения, композиция которых и обратные к ним также принадлежат множеству $\mathcal{T}$ (проверьте!). $\square$

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 рассуждений 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 (попробуйте доказать!).

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, связанные условием $ad - bc \neq 0$, которое необходимо и достаточно для обратимости отображения $\Phi(x)$. Дробно-линейные отображения относительно композиции образуют группу (проверьте!).

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 — докажите!) и играют важную роль в теории функций комплексной переменной.

Задача 1. *Доказать, что дробно-линейное отображение $z \;\mapsto\; a\frac{z-\bar{b}}{z-b}, \quad |a| = 1, \;\; Im(b) < 0$, переводит точки (комплексные числа) верхней полуплоскости в точки единичного круга с центром в начале координат.*

Задача 2. *Доказать, что дробно-линейное отображение $z \;\mapsto\; a\frac{z-\bar{b}}{1-z\bar{b}}, \quad |a| = 1, \;\; |b| < 1$, переводит точки (комплексные числа) единичного круга с центром в начале координат в точки того же множества.*

9.4 Корни из единицы

Комплексное число z называется корнем степени n из единицы , 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 1. *Существует ровно* n *различных корней из единицы степени* n. *Это комплексные числа, расположенные в вершинах правильного* n*-угольника, вписанного в единичную окружность*:

$$z_k = \mathrm{e}^{\mathbf{i}\frac{2\pi k}{n}} = \cos\left(\frac{2\pi\,k}{n}\right) \;+\; \mathbf{i}\,\sin\left(\frac{2\pi\,k}{n}\right), \qquad k = 0,\,1,\,\ldots,\,n-1.$$

Доказательство. Уравнение $z^n = 1$ относительно $z = |z|\mathrm{e}^{\mathbf{i}\phi}$ принимает вид $|z|^n\mathrm{e}^{\mathbf{i}n\phi} = 1$ $\Rightarrow$ $|z| = 1$ и $\mathrm{e}^{\mathbf{i}n\phi} = 1$ $\Rightarrow$ $\cos(n\phi) = 1,\ \sin(n\phi) = 0$. Отсюда $\phi = \frac{2\pi\,k}{n},\ \ k = 0,\,\pm1,\,\pm2,\,\ldots$. В силу периодичности синуса и косинуса, $z_k = z_l \;\Leftrightarrow\; k - l \,\vdots\, n$. $\square$

Следствие 2. *Множество* $\mathcal{K}_n = \{z \in \mathbb{C} : z^n = 1\}$ *является циклической мультипликативной группой, состоящей из* n *элементов.*

Корень $z \in \mathcal{K}_n$ называется *первообразным корнем* степени n из единицы, если его порядок как элемента группы $\mathcal{K}_n$ равен n. Очевидно, комплексное число $\varepsilon = \mathrm{e}^{\frac{2\pi}{n}\mathbf{i}}$ является первообразным корнем.

Утверждение. *Комплексное число* $\varepsilon^m \in \mathcal{K}_n$ *является первообразным корнем степени* n *из единицы тогда и только тогда, когда целые числа* m *и* n *взаимно просты.*

Доказательство. Предположим, что $d > 1$ — нетривиальный общий делитель чисел m и n. Тогда $(\varepsilon^m)^{n/d} = (\varepsilon^n)^{m/d} = 1$, и значит, порядок элемента ε^m группы $\mathcal{K}_n$ строго меньше n. Условие $(\varepsilon^m)^k = 1$ означает, что $mk \,\vdots\, n$. Если m и n взаимно просты, то отсюда следует, что $k \,\vdots\, n$. Ясно, что в диапазоне $1 \le k \le n$ на n делится только одно число $k = n$, и, следовательно, в случае взаимной простоты чисел m и n порядок элемента ε^m равен n. $\square$

Задача 3. *Докажите, что сумма всех* n *корней из единицы степени* $n \ge 2$ *равна нулю.*

Задача 4. *Докажите, что* $\sum\limits_{k=0}^{n-1}(x + \varepsilon^k y)^n = n(x^n + y^n)$, *где* $\varepsilon = \cos(2\pi/n) + \mathbf{i}\sin(2\pi/n)$.

Задача 5. *Используя комплексные числа, докажите, что*

$$(a)\quad x^{2n} - 1 = (x^2 - 1)\prod_{k=1}^{n-1}(x^2 - 2x\cos(\pi k/n) + 1), \qquad (b)\quad \prod_{k=1}^{n-1}\sin\left(\frac{\pi k}{2n}\right) = \frac{\sqrt{n}}{2^{n-1}}.$$

9.5 Комплексные многочлены

Мы только что установили, что многочлен $f(z) = z^n - 1$ имеет n комплексных корней. В действительно это частный случай очень общего утверждения: комплексные корни

есть у *любого* комплексного многочлена степени $n \geq 1$. Это 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. Оно знакомит нас с интересным свойством модуля многочлена и требует 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 комплексных многочленов как функций влечет за собой также их равенство как формальных выражений от степеней буквы z. В самом деле, согласно теореме Безу, многочлен степени n не может иметь более, чем n корней, а если $f(z) = g(z)$ для всех z, то многочлен $f(z) - g(z)$ имеет бесконечно много корней и поэтому обязан быть нулевым многочленом.

9.6 Последовательности комплексных чисел

Последовательность комплексных чисел z_k называется *сходящейся к* комплексному числу z_0, если $\lim\limits_{k\to\infty} |z_k - z_0| = 0$. Таким образом, понятие сходимости для комплекных последовательностей определяется через сходимость к нулю вещественной последовательности $|z_k - z|$. Записи $z_0 = \lim\limits_{k\to\infty} z_k$ или $z_k \to z_0$ означают, что z_0 является *пределом* последовательности z_k (в терминологии и обозначениях сохраняется преемственность с вещественным случаем).

Теорема Больцано-Вейерштрасса на комплексной плоскости. *Для произвольной последовательности z_k точек, принадлежащих кругу $|z| \leq R$, существует подпоследовательность, сходящаяся к некоторой точке этого круга.*

Доказательство. Пусть $|z_k| \leq R$. Тогда $-R \leq \mathrm{Re}(z_k) \leq R$ и, по теореме Больцано-Вейерштрасса для вещественных последовательностей на отрезке $[-R,\, R]$, существует сходящаяся подпоследовательность $\mathrm{Re}(z_{k_l}) \to x_0 \in [-R,\, R]$. В силу той же теоремы, подпоследовательность $\mathrm{Im}(z_{k_l}) \in [-R,\, R]$ обладает сходящейся подпоследовательностью $\mathrm{Im}(z_{k_{l_m}}) \to y_0 \in [-R,\, R]$. При этом $\mathrm{Re}(z_{k_{l_m}}) \to x_0$ (как подпоследовательность сходящейся последовательности). Положим $z_0 = x_0 + y_0\mathbf{i}$. Тогда

$$|z_{k_{l_m}} - z_0| = \sqrt{|\mathrm{Re}(z_{k_{l_m}}) - x_0|^2 + |\mathrm{Im}(z_{k_{l_m}}) - y_0|^2} \;\to\; 0 \quad\Rightarrow\quad z_{k_{l_m}} \to z_0.$$

Кроме того, после перехода к пределу в неравенствах,

$$|z_{k_{l_m}}| = \sqrt{|\mathrm{Re}(z_{k_{l_m}})|^2 + |\mathrm{Im}(z_{k_{l_m}})|^2} \;\leq\; R \quad\Rightarrow\quad |z_0| = \sqrt{|x_0|^2 + |y_0|^2} \;\leq\; R. \quad \square$$

9.7 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 на комплексной плоскости. *Пусть функция $\Phi(z)$ принимает вещественные значения, определена на комплексной плоскости и непрерывна*

во всех точках круга $D = \{z \in \mathbb{C} : |z| \le R\}$. *Тогда существуют точки* $z_{\min},\ z_{\max} \in D$ *такие, что* $\Phi(z_{\min}) \le \Phi(z) \le \Phi(z_{\max})$ *для всех* $z \in D$.

Доказательство. Докажем существование точки $z_{\max}$. Для этого сначала установим ограниченность функции $\Phi(z)$ сверху. От противного, допустим, что есть последовательность $z_k \in D$ со свойством $\Phi(z_k) > k$. По теореме Больцано-Вейерштрасса на комплексной плоскости, есть сходящаяся подпоследовательность $z_{k_l} \to z_0 \in D$. В силу непрерывности, $\Phi(z_{k_l}) \to \Phi(z_0)$, а это противоречит неравенствам $\Phi(z_{k_l}) > k_l \ge l$. Таким образом, множество вещественных чисел $\Phi(D) = \{\xi : \xi = \Phi(z),\ z \in D\}$ ограничено сверху, и, следовательно, для него существует *точная верхняя грань*, которая определяется как верхняя граница для всех чисел множества $\Phi(D)$, к которой сходится некоторая последовательность чисел этого множества. Пусть M — точная верхняя грань для $\Phi(D)$ и $\Phi(z_k) \to M$ для какой-то последовательности $z_k \in D$. В силу теоремы Больцано-Вейерштрасса на комплексной плоскости, имеется подпоследовательность $z_{k_l} \to z_0 \in D \ \Rightarrow\ \Phi(z_{k_l}) \to \Phi(z_0)$. Поскольку предел определяется однозначно, находим $\Phi(z_0) = M$ и можем взять $z_{\max} = z_0$. Существование точки $z_{\min}$ доказывается переходом к функции $-\Phi(z)$. $\square$

9.8 Свойства модуля многочлена

Рассмотрим произвольный комплексный многочлен вида

$$f(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n, \quad n \ge 1.$$

Лемма о непрерывности модуля многочлена. *Функция* $\Phi(z) = |f(z)|$ *непрерывна во всех точках* z *комплексной плоскости.*

Доказательство. Поделим $f(z)$ с остатком на $h = z - z_0$:

$$f(z_0 + h) = f(z_0) + b_1 h + \ldots + b_{n-1} h^{n-1} + h^n \quad \Rightarrow$$

$$|\Phi(z_0 + h) - \Phi(z_0)| \ = \ ||f(z_0 + h)| - |f(z_0)|| \ \le \ |f(z_0 + h) - f(z_0)| \le$$
$$|b_1 h \ + \ \ldots \ + \ b_{n-1} h^{n-1} \ + \ h^n| \ \le \ |b_1||h| \ + \ \ldots \ + \ |b_{n-1}||h|^{n-1} \ + \ |h|^n. \qquad \square$$

Лемма о росте модуля многочлена. *Для любого числа* $M > 0$ *существует* $R > 0$ *такое, что из неравенства* $|z| > R$ *вытекает, что* $|f(z)| > M$.

Доказательство. Учитывая, что $|z^i| = |z|^i$, получаем

$$|f(z)| \ \ge \ |z^n| - |a_0 + a_1 z + \ldots + a_{n-1} z^{n-1}| \ \ge \ |z|^n \ - \ |a_0| - |a_1||z| - \ldots - |a_{n-1}||z|^{n-1}.$$

Положим $A = |a_0| + \ldots + |a_{n-1}|$. Тогда при $|z| > 1$ находим $|f(z)| > |z|^n \left(1 - \frac{A}{|z|}\right)$, и, очевидно, если $|z| > 2A$, то $|f(z)| > |z|^n/2$. Значит, при $|z| > R = \max\{1,\ 2A,\ \sqrt[n]{2M}\}$ будет выполняться неравенство $|f(z)| > M$. $\square$

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. Игнорируя нулевые члены, запишем $f(z + h)$ в виде

$$f(z + h) = f(z) + b_m h^m + \underbrace{b_{m+1} h^{m+1} + \ldots + b_{n-1} h^{n-1} + b_n h^n}_{= g(h) h^{m+1}}, \quad b_m \ne 0,$$

выберем h_0 как одно из решений уравнения $b_m h_0^m = -f(z)$ и положим $h = th_0$, где вещественный параметр t подчинен условию $0 < t < 1$. Тогда

$$|f(z+th_0)| \le |f(z)|(1-t^m) + ct^{m+1} = |f(z)|(1-t^m(1-ct)), \quad c = \left(\max_{0\le t\le 1} |g(th_0)|\right)|h_0|^{m+1}.$$

При $0 < t < \min(1/(2c),\ 1)$ получаем $|f(z+th_0)| \le |f(z)|(1-t^m/2)$. $\square$

9.9 Основная теорема алгебры

Основная теорема алгебры. *Любой комплексный многочлен ненулевой степени 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 $|z| > R$ имеем $|f(z)| > M$. Функция $|f(z)|$ принимает вещественные значения и непрерывна на всей комплексной плоскости, в том числе и во всех точках круга $|z| \le R$. По теореме Вейерштрасса на комплексной плоскости, в некоторой точке $z_{\min}$ этого круга она достигает своего минимального значения на этом круге. Очевидно, $|f(z_{\min})| \le |f(0)| = M$. Таким образом, $|f(z_{\min})|$ является минимальным значением модуля *на всей комплексной плоскости*. Если $|f(z_{\min})| > 0$, то это утверждение входит в противоречие с леммой Даламбера. Следовательно, $|f(z_{\min})| = 0$ и $z_{\min}$ — корень многочлена $f(z)$. $\square$

Поле называется *алгебраически замкнутым*, если любой многочлен положительной степени с коэффициентами из этого поля имеет корень в этом же поле. Основная теорема алгебры утверждает, что поле комплексных чисел алгебраически замкнуто.

Задача 6. *Докажите, что любое алгебраически замкнутое поле содержит бесконечно много элементов.*

Задача 7 *Задан отличный от константы многочлен $f(x,y)$ от переменных x и y над алгебраически замкнутым полем $\mathbb{K}$. Докажите, что он имеет бесконечно много* `нулей`*, т. е. точек $(a,b) \in \mathbb{K}^2$ таких, что $f(a,b) = 0$.*

Задача 8. *Множество точек $(a,b) \in \mathbb{C}^2$ является геометрическим местом точек, удовлетворяющих уравнению $f(x,y) = 0$, где $f(x,y)$ — неприводимый комплексный многочлен от переменных x и y. Докажите, что если многочлен $g(x,y)$ является неприводимым многочленом, обращающимся в нуль в тех же точках, то $g(x,y) = cf(x,y)$ для некоторого комплексного числа c.*

9.10 Разложение на линейные множители

Многочлены первой степени называют также *линейными многочленами*.

Теорема. *В кольце $\mathbb{C}[z]$ множество неприводимых многочленов состоит из линейных многочленов и только из них, а любой приведенный многочлен $f(z)$ степени $n > 0$ имеет разложение вида*

$$f(z) \ = \ (z - z_1) \ldots (z - z_n), \qquad (*)$$

где $z_1, \ldots,\ z_n$ — комплексные числа.

Доказательство. Проведем индукцию по n. Случай $n = 1$ очевиден, Далее, 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 $z - z_1$. Значит, $f(z) = (z - z_1)f_1(z)$ и, поскольку $\deg f_1(z) = n - 1$, для многочлена $f_1(z)$ искомое разложение существует согласно индуктивному предположению. $\square$

Таким образом, равенство $(*)$ представляет собой разложение комплексного многочлена на неприводимые множителя и, следовательно, как вообще для многочленов над любым полем, единственно с точностью до переупорядочивания множителей и умножения их на ненулевые комплексные константы (делители единицы в кольце $\mathbb{C}[z]$).

Если ζ — корень многочлена $f(z)$, то $f(z)$ делится на $z - \zeta$. Обозначим через k максимальное натуральное число такое, что $f(z)$ делится на $(z - \zeta)^k$. Таким образом, имеет место равенство $f(z) = (z - \zeta)^k g(z)$, в котором $g(\zeta) \neq 0$. Число k называется *кратностью* корня ζ. Если $k = 1$, то корень называется *простым*, а при $k > 1$ — *кратным*. Полученные нами результаты удобно рассматривать как утверждение о том, что *любой комплексный многочлен степени $n > 0$ имеет ровно n комплексных корней с учетом кратностей.*

Следствие. *Любой комплексный приведенный многочлен $f(x)$ степени $n > 0$ имеет единственное разложение вида $f(z) = (z - \zeta_1)^{k_1} \ldots, (z - \zeta_s)^{k_s}$, где комплексные корни $z_1, \ldots, z_s$ попарно различны, а $k_1, \ldots, k_s$ — их кратности.*

9.11 Разложение вещественных многочленов

Мы знаем, что вещественный многочлен может и не иметь вещественных корней. Однако, вещественный многочлен можно рассматривать также как комплексный, и, согласно основной теореме алгебры, существование комплексных корней для него гарантировано.

Пусть $f(x) \in \mathbb{R}[x]$ и $f(z) = 0$ для некоторого $z \in \mathbb{C}$. Поскольку коэффициенты $f(x)$ вещественны, операция комплексного сопряжения оставляет их на месте и, следовательно, $f(\bar{z}) = \overline{f(z)} = 0$. Условие $\bar{z} \neq z$ равносильно невещественности корня z. Таким образом, вместе с любым невещественным корнем z вещественного многочлена присутствует также сопряженный ему корень $\bar{z}$. Заметим, что квадратный трехчлен

$$(x - z)(x - \bar{z}) = x^2 + (z + \bar{z})x + |z|^2$$

имеет вещественные коэффициенты и при условии $\bar{z} \neq z$ неприводим в кольце $\mathbb{R}[x]$. Следовательно, справедлива следующая

Теорема. *В кольце $\mathbb{R}[x]$ множество неприводимых многочленов состоит из линейных многочленов и квадратных трехчленов без вещественных корней и только из них, а любой приведенный многочлен $f(x)$ степени $n > 0$ имеет разложение вида*

$$f(x) = (x - x_1) \ldots (x - x_k)\ \phi_1(x) \ldots \phi_l(x), \quad k + 2l = n,$$

где $x_1, \ldots, x_k$ — вещественные корни, а $\phi_1(x), \ldots, \phi_l(x)$ — неприводимые квадратные трехчлены, имеющие пару различных комплексно сопряженных корней.

Следствие. *Любой вещественный многочлен нечетной степени имеет хотя бы один вещественный корень.*

Замечание. Последнее утверждение можно было бы доказать и непосредственно, не опираясь на основную теорему алгебры. При $x \to +\infty$ значение $f(x)$ будет положительным, а при $x \to -\infty$ отрицательным. Согласно теореме Ролля, если непрерывная функция принимает значения разных знаков на концах отрезка, то в какой-то внутренней точке отрезка она обязана обратиться в нуль.

9.12 Кратные корни и производные

Пусть $f(x) = a_0 + a_1x + \ldots + a_nx^n$ — многочлен над произвольным полем $\mathbb{K}$. Его *производной* называется многочлен $f'(x) \in \mathbb{K}[x]$ следующего вида:

$$f'(x) = a_1 \ + \ 2a_2x \ + \ \ldots \ + \ na_nx^{n-1}.$$

В этом определении *понятия математического анализа не используются.*

Однако, вещественный многочлен можно рассматривать как функцию от x и тогда производная этой функции, найденная по правилам анализа, будет именно такой, как и в нашем определении. Полезно иметь в виду, что и в нашем случае сохраняются известные в анализе свойства дифференцирования:

- $(f(x) + g(x))' = f'(x) + g'(x), \quad (cf(x))' = cf'(x) \quad (c$ — произвольная константа),
- $(f(x)g(x))' = f'(x)g(x) + f(x)g'(x) \quad$ (правило Лейбница).

Первое свойство очевидно, а второе достаточно проверить в случае одночленов, а затем использовать первое свойство. Заметим, что мы не можем взять эти свойства как факты дифференциального исчисления, так как наше определение производной многочлена носит совершенно формальный характер.

Теорема о кратном корне. *Пусть корень ζ многочлена $f(x) \in \mathbb{K}$ принадлежит более широкому полю $\mathbb{L} \supseteq \mathbb{K}$. Тогда если кратность корня ζ равна k, то ζ является корнем производной $f'(x) \in \mathbb{K}$ кратности не меньше $k-1$.*

Доказательство. Рассматривая многочлен $f(x)$ как многочлен над более широким полем $\mathbb{L}$, запишем $f(x) = (x-\zeta)^k g(x)$ и предположим, что $g(\zeta) \neq 0$. Вычисляя производную, находим $f'(x) = k(x-\zeta)^{k-1}g(x) + (x-\zeta)^k g'(x) \vdots (x-\zeta)^{k-1}$. $\square$

Следствие. *Многочлен имеет только простые корни в том и только том случае, когда он взаимно прост со своей производной.*

Доказательство. Если $f(x)$ и $f'(x)$ взаимно просты, то кратного корня быть не может в силу теоремы о кратном корне. Если все корни приведенного многочлена $f(x)$ простые, то он имеет вид $f(x) = \prod\limits_{i=1}^{n}(x-\zeta_i)$, где $\zeta_1, \ldots, \zeta_n$ — попарно различные числа некоторого поля, содержащего поле коэффициентов данного многочлена. При дифференцировании получаем $f'(x) = \sum\limits_{i=1}^{n}\prod\limits_{j\neq i}(x-\zeta_j) \quad \Rightarrow \quad f'(\zeta_i) = \prod\limits_{j\neq i}(\zeta_i - \zeta_j) \neq 0$. Таким образом, $f(x)$ и $f'(x)$ не имеют ни одного общего корня. $\square$

Совершенно замечательно то, что мы можем понять, будут ли корни простыми, не зная самих корней. Более того, это делается за конечное число арифметических операций над полем коэффициентов данного многочлена: достаточно применить алгоритм Евклида и найти наибольший общий делитель многочлена и его производной.

Задача 9. *Докажите, что многочлен $f(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots. + \frac{x^n}{n!}$ не имеет кратных корней.*

Задача 10. *Многочлен $f(z) = \sum\limits_{l=0}^{n} a_l z^l$ степени n имеет корень ζ кратности m. Докажите, что выполняются равенства $\sum\limits_{l=0}^{n} a_l\, l^k\, \zeta^k = 0, \quad 1 \leq k \leq m-1$.*

Задача 11. *Докажите, что результант многочленов $f(x)$ и $f'(x)$ лишь числовым коэффициентом отличается от дискриминанта приведенного многочлена $f(x)$.*

Задача 12. *Докажите, что дискриминант многочлена $x^n - 1 \in \mathbb{C}[x]$ равен $D = (-1)^{(n-1)(n+2)/2}n^n$.*

9.13 Непрерывность корней многочлена

Будем говорить, что последовательность комплексных приведенных многочленов $f_k(x)$ степени n сходится к комплексному приведенному многочлену $f(x)$ степени n, если $f_k(z) \to f(z)$ для всех комплексных чисел z. Для этого необходимо и достаточно, чтобы последовательности коэффициентов многочленов $f_k(x)$ при одинаковых степенях переменной x сходились к соответствующему коэффициенту многочлена $f(x)$ (докажите!).

Теорема о непрерывности корней многочлена. *Пусть последовательность комплексных приведенных многочленов $f_k(x)$ сходится к комплексном приведенному многочлену $f(x) = \prod_{i=1}^{n}(x - x_i)$. Тогда для многочленов $f_k(x)$ можно выбрать разложения $f_k(x) = \prod_{i=1}^{n}(x - x_{k,i})$ таким образом, что $\lim_{k\to\infty} x_{k,i} = x_i, \quad 1 \le i \le n.$*

Доказательство. Рассмотрим разложения $f_k(x) = \prod_{i=1}^{n}(x - y_{k,i})$ и заметим, что

$$|f_k(x_n)| = \left|\prod_{i=1}^{m}(x_n - y_{k,i})\right| \ge \left(\min_{1\le i\le n} |x_n - y_{k,i}|\right)^n = |x_n - y_{k,j}|^n.$$

Положим $x_{k,n} := y_{k,j}$. Тогда

$$|x_{k,n} - x_n| \;\le\; |f_k(x_n)|^{1/n} \;\to\; 0 \qquad \text{и} \qquad \frac{f_k(x)}{x - x_{k,n}} \;\to\; \frac{f(x)}{x - x_n}.$$

Задача сведена к аналогичной задаче для многочленов степени $n - 1$. □

9.14 Разностные уравнения с постоянными коэффициентами

В разных задачах возникают последовательности комплексных чисел $x_1, x_2, \ldots$, удовлетворяющие рекуррентным соотношениям вида

$$a_0 x_n + a_1 x_{n+1} + \ldots + a_p x_{n+p} = 0, \quad n = 0, 1 \ldots, \tag{1}$$

с заданными комплексными коэффициентами $a_0, \ldots, a_p$. Полагаем, что $a_0 \ne 0$ и $a_p \ne 0$. В этом случае уравнение (1) называется *разностным уравнением порядка p*. При любых фиксированных *начальных значениях* $x_0, x_1, \ldots, x_{p-1}$ оно однозначно определяет значения $x_p, x_{p+1}, \ldots$.

Основная теорема алгебры позволяет дать для x_n полезную *явную формулу*. Чтобы ее получить, будем искать нетривиальное решение разностного уравнения в виде $x_n = z^n$, где $z \ne 0$. Тогда для z получаем следующие соотношения:

$$a_0 z^n + a_1 z^{n+1} + \ldots + a_p z^{n+p} = 0 \quad \Leftrightarrow \quad a_0 + a_1 z + \ldots + a_p z^p = 0.$$

Таким образом, $x_n = z^n$ будет решением уравнения (1) в том и только том случае, когда z есть корень многочлена $f(x) = a_0 + a_1 x + \ldots + a_p x^p$.

Случай простых корней. Если $f(x)$ имеет p попарно различных корней $z_1, \ldots, z_p$ (в общем случае комплексных), то для произвольных констант $c_1, \ldots, c_p$ последовательность вида

$$x_n = c_1 z_1^n + \ldots + c_p z_p^n \tag{2}$$

будет, очевидно, решением уравнения (1). Более того, любое решение представляется в виде (2), так как x_n однозначно определяется по начальным значениям $x_0, \ldots, x_{p-1}$, а константы $c_1, \ldots, c_p$ однозначно определяются системой линейных уравнений

$$c_1 z_1^n + \ldots + c_p z_p^n = x_n, \quad n = 0, 1, \ldots, p-1,$$

для которой матрица коэффициентов является транспонированной матрицей Вандермонда для попарно различных узлов $z_1, \ldots, z_p$.

Случай кратных корней. Если многочлен $f(x)$ имеет кратные корни, то формула (2) уже не описывает все решения уравнения (1). Чтобы получить p линейно независимых решений в случае кратных корней, можно использовать следующие утверждения.

Лемма 1. *Пусть z — корень $f(x)$ кратности k. Тогда при любом фиксированном $0 \le t \le k-1$ последовательности вида*

$$x_{tn} \;=\; n^t z^n, \quad n = 0,\, 1,\, \ldots,$$

являются решениями уравнения (1).

Доказательство. Применяем индукцию по t. При $t \ge 1$ находим

$$a_0 n^t z^n + a_1 (n+1)^t z^{n+1} + \ldots + a_p (n+p)^t z^{n+p} =$$

$$n(a_0 n^{t-1} z^n + a_1 (n+1)^{t-1} z^{n+1} + \ldots + a_p (n+p)^{t-1} z^{n+p}) \;+\; z^{n+1}(a_1 + 2a_2 z + \ldots + p a_p z^{p-1}).$$

Выражение во второй скобке — это значение производной $f'(z)$. Поскольку z — кратный корень, получаем $f'(z) = 0$. $\square$

Лемма 2. *Пусть даны попарно различные ненулевые числа $z_1, \ldots, z_s$ и натуральные числа $k_1, \ldots, k_s$ такие, что $k_1 + \ldots + k_m = p$. Тогда p векторов*

$$\begin{gathered} [z_1^n]_{n=0}^{p-1}, \quad [n z_1^n]_{n=0}^{p-1}, \quad \ldots, \quad [n^{k_1-1} z_1^n]_{n=0}^{p-1}, \\ \ldots \quad \ldots \quad \ldots \\ [z_s^n]_{n=0}^{p-1}, \quad [n z_s^n]_{n=0}^{p-1}, \quad \ldots, \quad [n^{k_s-1} z_s^n]_{n=0}^{p-1} \end{gathered} \tag{3}$$

образуют линейно независимую систему.

Доказательство. Данная система состоит из s подсистем для попарно различных чисел $z_1, \ldots, z_s$, при этом в подсистеме для z_i имеется k_i векторов. Можно проверить, что линейная оболочка, натянутая на столбцы подсистемы для z_i совпадает с линейной оболочкой векторов

$$[z_i^n]_{n=0}^{p-1}, \quad [n\, z_i^n]_{n=0}^{p-1}, \quad [n(n-1)\, z_i^n]_{n=0}^{p-1}, \quad \ldots, \quad [n(n-1)\ldots(n-k_i+2)\, z_i^n]_{n=0}^{p-1}. \tag{4}$$

Поэтому линейная независимость векторов вида (3) равносильна линейной независимости системы, составленной из векторов вида (4) при $i = 1, \ldots, s$. Пусть A_p — матрица порядка p, составленная из векторов-столбцов вида (4). Чтобы вычислить определитель матрицы A_p, вычтем из каждой ее строки, кроме первой, предыдущую строку, умноженную на z_1. Несложные, хотя и громоздкие, выкладки приводят к соотношению $\det A_p = c \det A_{p-1}$, где $c \ne 0$, а A_{p-1} обозначает матрицу порядка $p-1$, вид которой аналогичен виду матрицы A_p с той лишь разницей, что k_1 следует заменить на $k_1 - 1$. Далее по индукции. $\square$

9.15 Алгебры с делением и теорема Фробениуса

Комплексные числа появились как результат попытки ввести операцию умножения, которая превратила двумерное пространство вещественных арифметических векторов в ассоциативную алгебру с делением. Можно ли ввести умножение для трехмерного пространства? Гамильтон пытался это сделать, но безуспешно. Ему, однако, удалось ввести умножение для четырехмерного пространства и построить алгебру *кватернионов* — пример ассоциативной алгебры с делением без свойства коммутативности.

Предположим, что на n-мерном пространстве $\mathbb{V}$ вещественных арифметических векторов введена билинейная операция умножения, которая делает его алгеброй с делением. Последнее означает наличие единицы $\mathbf{1}$ и разрешимость уравнений $ax = b$ и $ya = b$ для любых $a \ne 0$ и b из $\mathbb{V}$. При вещественном α элемент вида $\alpha \cdot \mathbf{1}$ отождествляется с α. Таким образом, мы полагаем, что $\alpha = \alpha \cdot \mathbf{1}$ и $1 = \mathbf{1}$. Множество вещественных чисел $\mathbb{R}$ (линейная оболочка единицы) рассматривается как часть множества $\mathbb{V}$.

Пусть $z \in \mathbb{V} \backslash \mathbb{R}$ и пусть k — минимальное натуральное k, для которого векторы $1, z, z^2, \ldots, z^k$ линейно зависимы. Это означает, что $f(z) = 0$ для вещественного многочлена $f(x)$ степени k. Если $f(x) = u(x)v(x)$, то $f(z) = \phi(z)\psi(z) = 0 \;\Rightarrow\; \phi(z) = 0$ или $\psi(z) = 0$ (здесь мы опираемся на то, что $\mathbb{V}$ —

алгебра с делением). В силу минимальности k вещественный многочлен $f(x)$ является неприводимым $\Rightarrow$ $\deg f(x) = 2$. Таким образом,

$$z^2 = \alpha + \beta z, \quad \alpha, \beta \in \mathbb{R}, \quad \beta^2 + 4\alpha < 0 \quad \Rightarrow \quad (z - \beta/2)^2 = \beta^2/4 + \alpha < 0.$$

Отрицательность дискриминанта $\beta^2 + 4\alpha$ необходима и достаточна для неприводимости вещественного квадратного трехчлена $f(x) = x^2 - \beta x - \alpha$. То, что у нас получилось, сформулируем следующим образом: *любой невещественный элемент z алгебры $\mathbb{V}$ является суммой $z = \rho + v$, где $\rho,\ v^2 \in \mathbb{R}$ и $v^2 < 0$ ($\rho = \beta/2,\ v = z - \beta/2$).*

Лемма. *Множество $\mathbb{D} = \{v \in \mathbb{V} : v^2 \in \mathbb{R},\ v^2 \le 0\}$ является линейным подпространством в $\mathbb{V}$.*

Доказательство. Пусть $v \in \mathbb{D}$ и $\alpha \in \mathbb{R}$. Тогда, очевидно, $(\alpha v)^2 = \alpha^2 v^2 \le 0 \ \Rightarrow\ \alpha v \in \mathbb{D}$. Теперь возьмем еще один вектор $u \in \mathbb{D}$ и докажем, что $u + v \in \mathbb{D}$. Если u и v линейно зависимы, то утверждение очевидно. Поэтому будем считать их линейно независимыми и заметим, что система $1,\ u,\ v$ должна быть линейно независимой. В противном случае $\alpha u + \beta v = 1$ для каких-то $\alpha, \beta \in \mathbb{R}$. При этом ясно, что числа α и β не могут быть нулями одновременно. Более того, ни одно из них не равно нулю: если, скажем, $\alpha = 0$, то $v \in \mathbb{R} \cap \mathbb{D} \ \Rightarrow\ v = 0$. После возведения в квадрат обеих частей равенства $\alpha u = 1 - \beta v$ находим $\alpha^2 u^2 = (1 - \beta v)^2 = 1 - 2\beta v + u^2 \ \Rightarrow\ v \in \mathbb{R} \cap \mathbb{D} \ \Rightarrow\ v = 0$. Итак, векторы $1,\ u,\ v$ линейно независимы. Это означает, что $u \pm v \notin \mathbb{R}$. Как было замечено ранее, в этом случае

$$(u+v)^2 = \alpha_1 + \beta_1(u+v), \quad (u-v)^2 = \alpha_2 + \beta_2(u-v), \quad \beta_1^2 + 4\alpha_1 < 0, \quad \beta_2^2 + 4\alpha_2 < 0.$$

Отсюда $\alpha + (\beta_1 + \beta_2)u + (\beta_1 - \beta_2)v = 0$, где $\alpha = \alpha_1 + \alpha_2 - (u+v)^2 - (u-v)^2 = \alpha_1 + \alpha_2 - 2u^2 - 2v^2 \in \mathbb{R}$. В силу линейной независимости системы $1,\ u,\ v$ получаем $\alpha = \beta_1 = \beta_2 = 0$. $\square$

Следствие. *Линейное пространство $\mathbb{V}$ является прямой суммой своих подпространств $\mathbb{R}$ и $\mathbb{D}$.*

Теорема Фробениуса. *Ассоциативные алгебры с делением размерности n над полем вещественных чисел существуют только для $n = 1,\ 2,\ 4$.*

Доказательство. Заметим, что функция

$$\Phi(u, v) = \frac{u^2 + v^2 - (u+v)^2}{2} = -(uv + vu)/2$$

обладает всеми свойствами скалярного произведения векторов $u, v \in \mathbb{D}$: скалярный квадрат неотрицателен ($\Phi(u,u) = -u^2 \ge 0$) и равен нулю только для нулевого вектора ($\Phi(u,u) = 0 \ \Leftrightarrow\ u = 0$), симметричность ($\Phi(u,v) = \Phi(v,u)$) очевидна, линейность по первому аргументу следует из равенства $\Phi(u,v) = -(uv + vu)/2$. Пусть $\dim \mathbb{D} = d$. Будем говорить, что базис $v_1, \ldots, v_d$ является ортонормированным, если $\Phi(v_i, v_j) = \delta_{ij}$ (0 при $i \ne j$ и 1 при $i = j$), и построим в $\mathbb{D}$ ортонормированный базис. Чтобы это сделать, возьмем произвольный базис $u_1, \ldots, u_d$ и проведем следующий процесс его ортогонализации:

$$p_1 := u_1, \quad p_2 := u_2 - \frac{\Phi(u_2, p_1)}{\Phi(p_1, p_1)} p_1, \quad \ldots, \quad p_d := u_d - \frac{\Phi(u_d, p_1)}{\phi(p_1, p_1)} p_1 - \ldots - \frac{\Phi(u_d, p_d)}{\Phi(p_d, p_d)} p_d.$$

Полученные векторы $p_1, \ldots, p_d$ ортогональны, и искомый базис получается их нормировкой:

$$v_1 := p_1/\sqrt{\Phi(p_1, p_1)}, \quad \ldots, \quad v_d := p_d/\sqrt{\Phi(p_d, p_d)}.$$

Ортонормированность в данном случае означает, что

$$v_i v_j = -v_j v_i \quad \text{при} \quad i \ne j \quad \text{и} \quad v_i^2 = -1. \qquad (*)$$

Заметим, что $\Phi(v_1 v_2, v_1) = \Phi(v_1 v_2, v_2) = 0$ и, кроме того, $\Phi(v_1 v_2, v_1 v_2) = 1$. Поэтому мы будем считать, что $v_3 = v_1 v_2$.

Допустим, что $d \ge 4$. Тогда, используя выбор $v_3 = v_1 v_2$ и равенства $(*)$, находим $((v_1 v_2) v_3) v_4 = -v_4$ и, в то же время, при другой расстановке скобок и с учетом тех же равенств $(*)$, можно получить $(v_1 v_2 v_3) v_4 = -v_4 (v_1 v_2 v_3) = v_4$. Таким образом, условие $d \ge 4$ входит в противоречие с ассоциативностью умножения. Следовательно, $d \le 3$.

Если $d = 2$, то должно иметь место разложение $v_1 v_2 = c_0 + c_1 v_1 + c_2 v_2, \quad c_0, c_1, c_2 \in \mathbb{R}$. Отсюда $-v_1 = c_0 v_2 + c_1 v_1 v_2 - c_2$ (умножаем справа на v_2) и $c_1 v_1 v_2 = c_0 c_1 + c_1^2 v_1 + c_1 c_2 v_2$ (умножаем на c_1). Сложив полученные равенства, находим $\quad -v_1 = c_0 v_2 - c_2 + c_0 c_1 + c_1^2 v_1 + c_1 c_2 v_2 \quad \Rightarrow$

$$(c_0 c_1 - c_2) \ + \ (1 + c_1^2) v_1 \ + \ (c_0 + c_1 c_2) v_2 \ = \ 0.$$

Противоречие с линейной независимостью векторов $1,\ v_1,\ v_2$. Значит, случай $d = 2$ невозможен.

Случаи $d = 0$ и $d = 1$ дают известные нам коммутативные алгебры $\mathbb{V} = \mathbb{R}$ и $\mathbb{V} = \mathbb{C}$. В случае $d = 3$ возникает некоммутативная алгебра кватернионов, открытая Гамильтоном. Умножение в ней задается соотношениями $(*)$. $\square$

9.16 Кватернионы

Базисные кватернионы обычно обозначаются через $\mathbf{i} = v_1,\ \mathbf{j} = v_2,\ \mathbf{k} = v_3$. Произвольный кватернион является линейной комбинацией вида $a = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ с вещественными коэффициентами a_0, a_1, a_2, a_3. Он имеет *скалярную часть* a_0 и *векторную часть* $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$. Кватернион $\bar{a} = a_0 - \mathbf{a}$ называется сопряженным кватерниону a. Можно проверить, что $a\bar{a} = \bar{a}a = |a|^2$, где $|a| = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2}$ называется модулем (длиной) кватерниона a. С помощью сопряженных кватернионов уравнения $ax = b$ и $ya = b$ решаются следующим образом:

$$x = \bar{a}b/|a|^2, \quad y = b\bar{a}/|a|^2.$$

При умножении кватернионов нетривиальная часть операции связана с правилом перемножения их векторных частей. Оно задается следующей таблицей умножения:

	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{k}$
$\mathbf{i}$	-1	$\mathbf{k}$	$-\mathbf{j}$
$\mathbf{j}$	$-\mathbf{k}$	-1	$\mathbf{i}$
$\mathbf{k}$	$\mathbf{j}$	$-\mathbf{i}$	-1

Пусть $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ и $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$. Тогда (проверьте!)

$$\mathbf{ab} = -(a_1b_1 + a_2b_2 + a_3b_3) + (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_2b_3 - a_3b_2)\mathbf{k}.$$

Здесь трудно не обратить внимание на связь со скалярным и векторным умножением трехмерных векторов (a_1, a_2, a_3) и (b_1, b_2, b_3). Имея в виду эту связь, произведение векторных частей кватернионов записывают также в виде

$$\mathbf{ab} = -(\mathbf{a}, \mathbf{b}) + [\mathbf{a}, \mathbf{b}].$$

Подобно тому, как комплексные числа сопоставляются со специальными матрицами порядка 2, кватернионы можно представлять специальными матрицами порядка 4. Из записывают в блочном виде с помощью 2×2-блоков

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Вот матрицы, которые ставятся в соответствие базисным кватернионам:

$$1 \to \mathbb{E} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \quad \mathbf{i} \to \mathbb{I} = \begin{bmatrix} J & 0 \\ 0 & -J \end{bmatrix}, \quad \mathbf{j} \to \mathbb{J} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}, \quad \mathbf{k} \to \mathbb{K} = \begin{bmatrix} 0 & J \\ J & 0 \end{bmatrix}.$$

Отображение $a = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} \to a_0\mathbb{E} + a_1\mathbb{I} + a_2\mathbb{J} + a_3\mathbb{K}$ сохраняет операции и реализует изоморфизм на алгебру матриц, составляющих линейную оболочку матриц $\mathbb{E}$, $\mathbb{I}$, $\mathbb{J}$, $\mathbb{K}$. Для алгебры матриц билинейность и ассоциативность умножения, конечно, очевидны. Изоморфизм позволяет утверждать, что операции алгебры кватернионов обладают теми же свойствами.

Задача 13. *Докажите, что при перемножении кватернионов их модули перемножаются.*

Алгебра и геометрия (1 поток)

Лекция 10

10.1 Минимальные расширения

Пусть $\mathbb{P}$ — произвольное поле и $\mathbb{F}$ — его расширение. Возможно, что в поле $\mathbb{F}$ имеются подполя, которые также являются расширениями поля $\mathbb{P}$. Рассмотрим все такие подполя, которые содержат также фиксированное число $\theta \in \mathbb{F}$. Их пересечение будет полем (проверьте!). Оно называется *минимальным расширением* поля $\mathbb{P}$, содержащим число θ, и обозначается $\mathbb{P}(\theta)$. Говорят также, что поле $\mathbb{P}(\theta)$ получено из поля $\mathbb{P}$ *присоединением* числа θ.

В более общем случае, если выбраны числа $\theta_1, \ldots, \theta_k \in \mathbb{F}$, то через $\mathbb{P}(\theta_1, \ldots, \theta_k)$ обозначается минимальное поле, содержащее $\mathbb{P}$ и числа $\theta_1, \ldots, \theta_k$. Минимальность означает, что данное поле вложено в любое поле, содержащее $\mathbb{P}$ и $\theta_1, \ldots, \theta_k$ (и значит, является пересечением всех полей с таким свойством).

10.2 Алгебраические числа и минимальные многочлены

Число θ называется *алгебраическим над полем* $\mathbb{P}$, если оно является корнем ненулевого многочлена над $\mathbb{P}$. Среди таких многочленов имеется многочлен минимальной степени, он называется *минимальным многочленом* числа θ над полем $\mathbb{P}$. В таких случаях поле $\mathbb{P}(\theta)$ называется *простым алгебраическим* расширением поля $\mathbb{P}$. В более общем случае, если каждое число поля $\mathbb{K} \supseteq \mathbb{P}$ является алгебраическим над $\mathbb{P}$, то $\mathbb{K}$ называется *алгебраическим расширением поля* $\mathbb{P}$.

Примеры.

- $\mathbb{C} = \mathbb{R}(\mathbf{i})$: поле комплексных чисел получается из поля вещественных чисел присоединением мнимой единицы $\mathbf{i}$. Число $\mathbf{i}$ является алгебраическим над $\mathbb{R}$, а его минимальный многочлен имеет вид $f(x) = x^2 + 1 \in \mathbb{R}[x]$. Поле $\mathbb{C}$ можно рассматривать как линейное пространство над полем $\mathbb{R}$ — оно состоит из линейных комбинаций $a \cdot 1 + b \cdot \mathbf{i}$, где $a, b \in \mathbb{R}$. Напомним, что при расширении полей $\mathbb{K} \supseteq \mathbb{P}$ размерность называется степенью расширения и обозначается $(\mathbb{K} : \mathbb{P})$. В нашем случае $(\mathbb{R}(\mathbf{i}) : \mathbb{R}) = 2$.
- Число $\sqrt[3]{2}$ является алгебраическим над полем рациональных чисел $\mathbb{Q}$, его минимальный многочлен имеет вид $f(x) = x^3 - 2 \in \mathbb{Q}[x]$. Поле $\mathbb{Q}\left(\sqrt[3]{2}\right)$ состоит из чисел вида $a + b\sqrt[3]{2} + c\left(\sqrt[3]{2}\right)^2$, где $a, b, c \in \mathbb{Q}$, и $\left(\mathbb{Q}\left(\sqrt[3]{2}\right) : \mathbb{Q}\right) = 3$.

Лемма о минимальном многочлене. *Минимальный многочлен алгебраического числа неприводим и определен однозначно с точностью до ненулевого числового множителя.*

Доказательство. Приводимость минимального многочлена означала бы, что алгебраическое число является корнем ненулевого многочлена меньшей степени. Пусть $f(x)$ и $g(x)$ — два минимальных многочлена для θ, оба одной и той же степени n. Тогда их наибольший общий делитель можно записать в виде $d(x) = f(x)\phi(x) + g(x)\psi(x) \;\Rightarrow\; d(\theta) = 0 \;\Rightarrow\; \deg d(x) = n \;\Rightarrow$ каждый из многочленов $f(x)$ и $g(x)$ отличается от $d(x)$ лишь ненулевым числовым множителем. □

Лемма об алгебраичности конечных расширений. *Любое конечное расширение является алгебраическим.*

Доказательство. Напомним, что расширение $\mathbb{K}$ поля $\mathbb{P}$ называется конечным, если линейное пространство $\mathbb{K}$ над полем $\mathbb{P}$ конечномерно. Возьмем $\theta \in \mathbb{K}$ и обозначим через n минимальное натуральное n такое, что числа 1, θ, θ^2, $\ldots$, θ^n линейно зависимы как элементы конечномерного линейного пространства над полем $\mathbb{P}$. Число θ, очевидно, будет корнем многочлена $f(x)$, составленного из коэффициентов равной нулю нетривиальной линейной комбинации этих элементов. Можно даже заметить, что это будет минимальный многочлен числа θ. Таким образом, любое число поля $\mathbb{K}$ является алгебраическим над полем $\mathbb{P}$. □

10.3 Присоединение корня

Теорема о присоединении корня. *Пусть число θ алгебраично над полем $\mathbb{P}$ и степень его минимального многочлена равна n. Тогда $(\mathbb{P}(\theta) : \mathbb{P}) = n$, а поле $\mathbb{P}(\theta)$ состоит из чисел вида $a_0 + a_\theta + \ldots + a_{n-1}\theta^{n-1}$, где $a_0, a_1, \ldots, a_{n-1} \in \mathbb{P}$.*

Доказательство. Множество линейных комбинаций $a_0 + a_1\theta + \ldots + a_{n-1}\theta^{n-1}$ с коэффициентами из $\mathbb{P}$, конечно, содержится в поле $\mathbb{P}(\theta)$. При этом оно само является полем. Его элементы удобно рассматривать как значения многочленов $a(x) \in \mathbb{P}[x]$ степени не выше $n-1$. Пусть $b(x)$ — еще один многочлен такого вида. Тогда $a(\theta) + b(\theta)$ есть значение многочлена $a(x) + b(x)$, степень которого не выше $n-1$. Произведение $a(\theta)b(\theta)$ есть значение многочлена степени не выше $n-1$, который является остатком при делении $a(x)b(x)$ на минимальный многочлен $f(x)$. Линейная зависимость чисел $1, \theta, \ldots, \theta^{n-1}$ как векторов линейного пространства над полем $\mathbb{P}$ означала бы, что $a(\theta) = 0$ для некоторого ненулевого многочлена степени меньше n. В то же время ясно, что $a(\theta) = 0$ тогда и только тогда, когда $a(x)$ — нулевой многочлен (иначе степень минимального многочлены была бы меньше n). Главный вопрос теперь такой: почему $(a(\theta))^{-1} = b(\theta)$ для какого-то многочлена $b(x) \in \mathbb{P}[x]$ степени не выше $n-1$? Если $a(\theta) \neq 0$, то многочлен $a(x)$ взаимно прост с минимальным многочленом $f(x)$. Согласно теореме о наибольшем общем делителе, можно получить равенство

$$a(x)b(x) + f(x)g(x) = 1, \quad b(x), g(x) \in \mathbb{P}[x], \quad \deg b(x) < \deg f(x) = n, \quad \deg g(x) < \deg a(x).$$

Полагая $x = \theta$, находим $a(\theta)b(\theta) = 1$. □

Следствие. *Для любого конечного набора алгебраических чисел $\theta_1, \ldots, \theta_n$ поле $\mathbb{P}(\theta_1, \ldots, \theta_n)$ является алгебраическим расширением поля $\mathbb{P}$.*

Доказательство. Заметим, что числа $\theta_1, \ldots, \theta_n$ можно присоединять последовательно. Тогда, согласно теореме о присоединении корня, каждое отдельное расширение в этой последовательности будет конечным, и, согласно теореме о степенях расширений, итоговое расширение также будет конечным. Из леммы об алгебраичности конечного расширения следует, что оно будет алгебраическим. □

Задача 1. *Докажите, что если число является корнем неприводимого многочлена, то он является минимальным многочленом этого числа.*

Задача 2. *Докажите, что поле $\mathbb{R}$ не может быть получено из поля $\mathbb{Q}$ присоединением конечного набора чисел.*

Задача 3. *Пусть $p_1 < \ldots < p_n$ — простые числа. Докажите, что $(\mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_n}) : \mathbb{Q}) = 2^n$.*

10.4 Поле алгебраических чисел

Теорема. *Множество всех алгебраических чисел поля $\mathbb{F}$ над его подполем $\mathbb{P}$ относительно операций сложения и умножения поля $\mathbb{F}$ является подполем в $\mathbb{F}$.*

Доказательство. Пусть $\alpha, \beta \in \mathbb{F}$ — корни ненулевых многочленов над $\mathbb{P}$, и пусть поле $\mathbb{P}(\alpha)$ получено из $\mathbb{P}$ присоединением числа α, а поле $\mathbb{P}(\alpha, \beta) = \mathbb{P}(\alpha)(\beta)$ получено из $\mathbb{P}(\alpha)$ присоединением числа β. В силу следствия из теоремы о присоединении корня, поле $\mathbb{P}(\alpha, \beta)$ является алгебраическим расширением поля $\mathbb{P}$. Таким образом, числа $\alpha \pm \beta$, $\alpha\beta$ и α/β при $\beta \neq 0$ являются алгебраическими, и, следовательно, все алгебраические над $\mathbb{P}$ числа образуют подполе в $\mathbb{F}$. $\square$

Неалгебраические числа называют *трансцендентными*. Они заведомо есть, например, в поле вещественных чисел, когда оно рассматривается как расширение поля рациональных чисел (достаточно заметить, что множество многочленов с рациональными коэффициентами счетно). Более трудная задача — установить трансцендентность некоторых конкретных чисел, в частности, e и π.

10.5 Поле разложения

Пусть числа $\theta_1, \ldots, \theta_n$ принадлежат некоторому расширению поля $\mathbb{P}$ и являются корнями приведенного многочлена $f(x) = \prod_{i=1}^{n}(x - \theta_i)$ с коэффициентами из поля $\mathbb{P}[x]$. Поле

$$\mathbb{P}_f = \mathbb{P}(\theta_1, \ldots, \theta_n)$$

называется *полем разложения* многочлена $f(x)$.

Обратим внимание на то, что поле разложения определяется не только самим многочленом, но и полем для коэффициентов многочлена. При выборе поля для коэффициентов всегда есть много возможностей, и они приводят к *разным* полям разложения. Например, многочлен с рациональными коэффициентами можно также рассматривать как многочлен с вещественными коэффициентами. В первом случае поле разложения обязано содержать поле рациональных чисел, а во втором оно не меньше, чем поле вещественных чисел. А если тот же многочлен рассматривать как многочлен над полем комплексных чисел, то его поле разложение совпадет с полем комплексных чисел.

Линейное пространство $\mathbb{P}_f$ над полем $\mathbb{P}$ всегда конечномерно. Но задача о вычислении его размерности часто не очень простая. Например, многочлен $f(x) = x^5 - 4x + 2$ можно рассматривать как многочлен над полем рациональных чисел $\mathbb{Q}$ и в этом случае известно, что $(\mathbb{Q}_f : \mathbb{Q}) = 120$. Но попробуйте-ка понять, почему размерность именно такая!

Задача 4. *Пусть $f(x)$ — многочлен степени n над полем $\mathbb{P}$. Докажите, что $(\mathbb{P}_f : \mathbb{P}) \leq n!$.*

10.6 Вычеты в кольце многочленов

Существование поля разложения очевидно в случае, когда известно достаточно широкое поле, в котором многочлен распадается на линейные множители. Согласно основной теореме алгебры, для комплексных многочленов таким полем является поле комплексных чисел. В случае многочленов над совершенно произвольным полем нам понадобятся *вычеты в кольце многочленов* над этим полем.

Пусть $\mathbb{P}$ — произвольное поле и $f(x) \in \mathbb{P}[x]$ — многочлен степени $n \geq 1$. *Вычетом по модулю* $f(x)$ называется множество многочленов кольца $\mathbb{P}[x]$, имеющих один и тот же остаток при делении на $f(x)$. Каждый вычет определяется любым своим представителем $a(x)$ и обозначается $[a(x)]_f = \{b(x) \in \mathbb{P}[x] : \quad b(x) - a(x) \vdots f(x)\}$. Сложение и умножение вычетов определяются через их представителей:

$$[a(x)]_f + [b(x)]_f := [a(x) + b(x)]_f, \quad [a(x)]_f[b(x)]_f := [a(x)b(x)]_f,$$

но нетрудно проверить, что сумма и произведение вычетов не зависят от выбора представителей. Таким образом, операции определены корректно и превращают множество вычетов по модулю $f(x)$ в коммутативное ассоциативное кольцо с единицей.

Вычеты, порожденные многочленами-константами, естественным образом отождествляются с числами поля $\mathbb{P}$. Таким образом, мы полагаем, что построенное кольцо вычетов содержит в себе поле $\mathbb{P}$. Если кольцо вычетов окажется полем, то это поле будет расширением поля $\mathbb{P}$.

Теорема о вычетах кольца многочленов. *Кольцо вычетов по модулю многочлена $f(x)$ над полем $\mathbb{P}$ является полем в том и только том случае, когда многочлен $f(x)$ неприводим над $\mathbb{P}$.*

Доказательство. Если $f(x) = u(x)v(x)$, то $[u(x)]_f[v(x)]_f = [u(x)v(x)]_f = [f(x)]_f = [0]_f$. В случае нетривиального разложения возникают делители нуля. Значит, для получения поля неприводимость многочлена $f(x)$ необходима. Теперь предположим, что многочлен $f(x)$ неприводим. Роль единичного элемента, очевидно, выполняет вычет $[1]_f$. Рассмотрим ненулевой вычет $[a(x)]_f$. Из неприводимости $f(x)$ следует, что $a(x)$ и $f(x)$ взаимно просты. По теореме о наибольшем общем делителе, можно получить равенство вида $a(x)b(x) + f(x)g(x) = 1 \quad \Rightarrow \quad [a(x)]_f[b(x)]_f = [1]_f$. $\square$

10.7 Существование поля разложения

Пусть задан многочлен над абстрактным полем $\mathbb{P}$. Он может не иметь корней в $\mathbb{P}$, но получить их в каком-то более широком поле $\mathbb{F}$. Всегда ли найдется такое поле? Положительный ответ напоминает основную теорему алгебры, но в последней утверждается, что $\mathbb{F} = \mathbb{P} = \mathbb{C}$. В абстрактном случае теорема о существовании корня, с одной стороны, относится к более общей ситуации, а с другой стороны, доказывается намного легче, так как является более слабым утверждением.

Теорема о существовании корня. *Любой многочлен ненулевой степени над произвольным полем $\mathbb{P}$ имеет корень в некотором расширении поля $\mathbb{P}$*

Доказательство. Не ограничивая общности, будем считать, что $f(x)$ неприводим над полем $\mathbb{P}$. Искомым полем $\mathbb{F}$ является поле вычетов по модулю $f(x)$. Пусть $f(x) = \sum\limits_{i=0}^{n} a_i x^i$,

$a_i \in \mathbb{P}$. При отождествлении $[a_i]_f = a_i$ тот же многочлен можно рассматривать как многочлен над $\mathbb{F}$. Искомым корнем является вычет $[x]_f$:

$$f([x]_f) = \sum_{i=1}^{n} [a_i]([x]_f)^i = \sum_{i=1}^{n} [a_i x^i]_f = \left[\sum_{i=1}^{n} a_i x^i\right]_f = [f(x)]_f = [0]_f. \quad \square$$

Следствие. *Любой приведенный многочлен $f(x)$ степени $n \geq 1$ над произвольным полем $\mathbb{P}$ может быть записан в виде $f(x) = \prod_{i=1}^{n}(x - \theta_i)$, где числа $\theta_1, \ldots, \theta_n$ принадлежат некоторому расширению $\mathbb{F}$ поля $\mathbb{P}$.*

Таким образом, *для любого многочлена существует поле разложения*, которое оказывается минимальным подполем поля $\mathbb{F}$, которое содержит поле $\mathbb{P}$ и все корни многочлена с учетом их кратностей.

Следствие из теоремы о существовании корня утверждает, что содержащее все корни поле $\mathbb{F}$ существует, но, вообще говоря, для одного и того же многочлена можно получать формально разные поля $\mathbb{F}$, а значит и формально разные поля разложения. Тем не менее, все поля разложения для одного и того же многочлена изоморфны — важный факт, который еще требует обстоятельного доказательства.

10.8 Единственность поля разложения

Здесь нам понадобится понятие *продолжения отображения* — так называется отображение с более широкой областью определения, чем исходное отображение. В частности, для полей $\mathbb{P}$ и $\widehat{\mathbb{P}}$ отображение $\Phi : \mathbb{P} \to \widehat{\mathbb{P}}$ продолжается на кольцо многочленов $\mathbb{P}[x] \to \widehat{\mathbb{P}}[x]$ следующим образом:

$$a(x) = a_0 + a_1 x + \ldots + a_n x^n \;\to\; \widehat{a}(x) = \widehat{a}_0 + \widehat{a}_1 x + \ldots + \widehat{a}_n x^n, \quad \widehat{a}_i = \Phi(a_i).$$

При продолжении отображения мы пытаемся сохранить какое-то полезное свойство. Если оно было изоморфизмом, то и при продолжении должно остаться изоморфизмом, но уже с более широкой областью определения. Приведенное выше правило, очевидно, продолжает изоморфизм полей до изоморфизма их колец многочленов. Заметим, что при изоморфизме неприводимый многочлен $a(x)$ соответствует неприводимому многочлену $\widehat{a}(x)$.

Лемма. *Пусть поле $\mathbb{P}$ изоморфно полю $\widehat{\mathbb{P}}$ и неприводимый многочлен $f(x) \in \mathbb{P}[x]$ соответствует неприводимому многочлену $\widehat{f}(x) \in \widehat{\mathbb{P}}[x]$. Тогда при выборе произвольных корней θ и $\widehat{\theta}$ многочленов $f(x)$ и $\widehat{f}(x)$ из каких-то полей разложения этих многочленов изоморфизм полей $\mathbb{P}$ и $\widehat{\mathbb{P}}$ может быть продолжен до изоморфизма их расширений $\mathbb{P}(\theta)$ и $\widehat{\mathbb{P}}(\widehat{\theta})$.*

Доказательство. Пусть $\deg f(x) = \deg \widehat{f}(x) = n$ и $a(x) \in \mathbb{P}[x]$ — многочлен степени не выше $n-1$. Искомое продолжение определим правилом $a(\theta) \to \widehat{a}(\widehat{\theta})$. Взаимная однозначность вытекает из неприводимости многочленов $f(x)$ и $\widehat{f}(x)$.

Проверим сохранение операций. Пусть $b(x) \in \mathbb{P}[x]$ — еще один многочлен степени не выше $n-1$. Рассмотрим сумму $c(x) = a(x) + b(x)$. Тогда $\widehat{c}(x) = \widehat{a}(x) + \widehat{b}(x) \quad \Rightarrow$

$$a(\theta) + b(\theta) = c(\theta) \;\to\; \widehat{c}(\widehat{\theta}) = \widehat{a}(\widehat{\theta}) + \widehat{b}(\widehat{\theta}).$$

Теперь пусть $c(x) = a(x)b(x) = f(x)q(x) + r(x)$, где $r(x)$ — остаток от деления $c(x)$ на $f(x)$. Тогда $\widehat{c}(x) = \widehat{a}(x)\widehat{b}(x) = \widehat{f}(x)\widehat{q}(x) + \widehat{r}(x) \quad \Rightarrow$

$$a(\theta)b(\theta) = c(\theta) = r(\theta) \;\to\; \widehat{r}(\widehat{\theta}) = \widehat{c}(\widehat{\theta}) = \widehat{a}(\widehat{\theta})\widehat{b}(\widehat{\theta}). \quad \square$$

Теорема о продолжении изоморфизма. *Изоморфизм полей можно продолжить до изоморфизма заданных полей разложения многочленов над этими полями, которые соответствуют друг другу при изоморфизме полей.*

Доказательство. Пусть поле $\mathbb{P}$ изоморфно полю $\widehat{\mathbb{P}}$, многочлен $f(x) \in \mathbb{P}[x]$ соответствует многочлену $\widehat{f}(x) \in \widehat{\mathbb{P}}[x]$, и пусть $\mathbb{P}_f$ и $\widehat{\mathbb{P}}_{\widehat{f}}$ — какие-то поля разложения многочленов $f(x)$ и $\widehat{f}(x)$. Обозначим через m число корней многочлена $f(x)$, не принадлежащих полю $\mathbb{P}$, и будем вести индукцию по m.

Если $m = 0$, то все очевидно, так как $\mathbb{P}_f = \mathbb{P}$ и $\widehat{\mathbb{P}}_{\widehat{f}} = \widehat{\mathbb{P}}$. В случае $m \geq 1$ рассмотрим разложение на неприводимые множители $f(x) = f_1(x)f_2(x)\dots f_s(x)$. Для определенности будем считать, что $\phi(x) := f_1(x)$ имеет корень θ, не принадлежащий полю $\mathbb{P}$. Согласно доказанной выше лемме, изоморфизм исходных полей продолжается до изоморфизма их расширений $\mathbb{P}(\theta)$ и $\widehat{\mathbb{P}}(\widehat{\theta})$, и очевидно, что многочлены $f(x)$ и $\widehat{f}(x)$ можно рассматривать как многочлены над этими расширении и при этом число корней многочлена $f(x)$ вне поля $\mathbb{P}$ строго меньше m. Остается применить индуктивное предположение. $\square$

Теорема о единственности поля разложения. *Все поля разложения одного и того же многочлена над одним и тем же полем изоморфны.*

Доказательство. Положим $\widehat{\mathbb{P}} = \mathbb{P}$ и в качестве исходного изоморфизма возьмем тождественное отображение. В этой ситуации результат сразу следует из теоремы о продолжении изоморфизма. $\square$

10.9 Алгебраическое доказательство основной теоремы алгебры

Мы приведем доказательство, которое опирается на факт существования поля разложения и использует понятие непрерывности "минимальным" образом.

(1) Пусть $f(x)$ — многочлен степени $n > 0$ с вещественными коэффициентами. Будем рассматривать его как многочлен над полем комплексных чисел и обозначим через $\mathbb{F}$ его поле разложения, содержащее, в частности, все комплексные числа. В поле $\mathbb{F}$ многочлен $f(x)$ разлагается на линейные множители и, следовательно, имеет n корней $x_1, \dots, x_n$ 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

$$\mathcal{F}_c(x) \;=\; \prod_{1 \leq i < j \leq n} (x - x_{ij}^c), \qquad x_{ij}^c \;=\; c\,x_i x_j + x_i + x_j.$$

В силу формул Виета и определения x_{ij}, коэффициенты $\mathcal{F}_c(x)$ — симметрические функции от корней вещественного многочлена $f(x) \quad \Rightarrow \quad$ они вещественны. Степень $\mathcal{F}_c(x)$ равна $(n^2 - n)/2 = 2^{k-1}q$, где $q = (2^k p - 1)p$ — 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 $i = i(c_1) = i(c_2)$, $j = j(c_1) = j(c_2)$. $\Rightarrow$

$$\left\{ \begin{array}{rcl} c_1\, x_i x_j + x_i + x_j & = & z_1 \in \mathbb{C} \\ c_2\, x_i x_j + x_i + x_j & = & z_2 \in \mathbb{C} \end{array} \right. \quad \Rightarrow \quad x_i x_j = \frac{z_1 - z_2}{c_1 - c_2} \in \mathbb{C} \quad \Rightarrow \quad x_i + x_j \in \mathbb{C}.$$

Следовательно, $x_i x_j$ и $x_i + x_j$ являются корнями квадратного уравнения с комплексными коэффициентами $\Rightarrow$ $x_i, x_j \in \mathbb{C}$.

(5) Итак, доказано,что любой *вещественный* многочлен степени $n > 0$ имеет хотя бы один комплексный корень. Теперь рассмотрим произвольный комплексный многочлен $f(x) = a_0 + a_1 x + \ldots a_{n-1} x^{n-1} + x^n$ и введем "сопряженный" многочлен $\bar{f}(x) = \bar{a_0} + \bar{a}_1 x + \ldots + \bar{a}_{n-1} x^{n-1} + x^n$. Легко проверить, что все коэффициенты многочлена $g(x) = f(x)\bar{f}(x)$ вещественные. Значит, он имеет комплексный корень z. Таким образом, $g(z) = f(z)\bar{f}(z) = f(z)\overline{f(\bar{z})} = 0 \quad \Rightarrow \quad f(z) = 0$ или $f(\bar{z}) = 0$. $\square$

10.10 Характеристика поля

Среди полей выделяются поля, которые можно рассматривать как расширение поля рациональных чисел. В таком поле с натуральным числом p отождествляется число вида

$$p \cdot 1 \;=\; \underbrace{1 + \ldots + 1}_{p \text{ раз}},$$

и очевидно, что все числа такого вида должны быть разными. В противном случае поле не может быть расширением поля рациональных чисел, а наименьшее натуральное p такое, что $p \cdot 1 = 0$ называется *характеристикой* такого поля. Если все числа вида $p \cdot 1$ разные, то характеристика поля считается равной нулю.

Утверждение 1. *Если характеристика поля отлична от нуля, то она является простым числом.*

Доказательство. Нетривиальное разложение $p = kl$ сразу приводит к делителям единицы: $(k \cdot 1)(l \cdot 1) = 0$. $\square$

Утверждение 2. *Поле характеристики $p > 0$ можно рассматривать как расширение поля $\mathbb{Z}_p$.*

Доказательство. Отображение $[k]_p \to k \cdot 1$ является инъективным и сохраняет операции. Поэтому его образ есть поле, изоморфное полю $\mathbb{Z}_p$. $\square$

Любое конечное поле имеет положительную характеристику. В качестве примера бесконечного поля положительной характеристики можно взять поле рациональных функций $\mathbb{Z}_p(x)$, которое определяется как поле дробей кольца $\mathbb{Z}_p[x]$.

Утверждение 3. *Для существования конечного поля характеристики p с числом элементов n необходимо, чтобы $n = p^d$ для некоторого натурального d.*

Доказательство. Если $\mathbb{P}$ — конечное поле характеристики p, то его можно рассматривать как расширение поля $\mathbb{Z}_p$, а значит и как линейное пространство над полем $\mathbb{Z}_p$. Пусть $d = (\mathbb{P} : \mathbb{Z}_p)$ — размерность этого пространства и $z_1, \ldots, z_d$ — его базис. Тогда каждый элемент поля $\mathbb{P}$ однозначно определяется разложением $\alpha_1 z_1 + \ldots + \alpha_d z_d$, где каждый коэффициент принадлежит полю $\mathbb{Z}_p$ и поэтому может принимать p различных значений. Всего элементов такого вида будет ровно p^d. $\square$

10.11 Полное описание конечных полей

Теорема. *Для любого простого p и любого натурального d поле с числом элементов $n = p^d$ существует и определяется однозначно с точностью до изоморфизма.*

Доказательство. Любой элемент a искомого поля должен удовлетворять уравнению $a^n - a = 0$. В самом деле, для $a = 0$ это очевидно, а любой элемент $a \neq 0$ принадлежит мультипликативной группе порядка $n - 1$ и, следовательно, $a^{n-1} = 1 \;\Rightarrow\; a^n = a$.

Рассмотрим многочлен $f(x) = x^n - x \in \mathbb{Z}_p$ и его поле разложения $\mathbb{F} = (\mathbb{Z}_p)_f$. Поскольку n делится на характеристику поля, $n \cdot a = 0$ для любого элемента $a \in \mathbb{F}$. Поэтому производная многочлена $f(x)$ имеет вид $f'(x) = nx^{n-1} - 1 = -1$ и, следовательно, является многочленом, взаимно простым с $f(x)$. Отсюда вытекает, что $f(x)$ имеет n различных корней, принадлежащих полю $\mathbb{F}$. Докажем, что множество $\mathbb{P}$ этих корней является подполем поля $\mathbb{F}$. Условие $a, b \in \mathbb{P}$ равносильно равенствам $a^n = a$ и $b^n = b$. Ясно, что $(ab)^n = a^n b^n = ab \;\Rightarrow\; ab \in \mathbb{P}$, и, кроме того, если $a \neq 0$, то $(a^{-1})^n = a^{-1}$. Чтобы понять, почему $(a + b)^n = a + b$, будем использовать тождество

$$(u + v)^p = u^p + v^p \quad \forall\ u,\ v \in \mathbb{F}.$$

Для его проверки нужно заметить, что при разложении в бином Ньютона все коэффициенты, кроме первого и последнего, делятся на p. Таким образом, $(a + b)^p = a^p + b^p$, $(a^p + b^p)^p = a^{p^2} + b^{p^2}$, $\ldots$, $(a^{p^{d-1}} + b^{p^{d-1}})^p = a^{p^d} + b^{p^d} \;\Rightarrow\; (a + b)^{p^d} = a^{p^d} + b^{p^d}$. Для противоположного элемента к a находим $(-a)^n = (-1)^n a = -a$. Действительно, если $p > 2$, то n нечетно и $(-1)^n = -1$, а в случае $p = 2$ имеет место тождество $-a = a$ для любого элемента $a \in \mathbb{F}$.

Таким образом, n-элементное множество $\mathbb{P}$ является полем, содержащим поле $\mathbb{Z}_p$ и все n попарно различных корней многочлена $f(x) = x^n - x$. Согласно определению поля разложения, в данном случае оно совпадает с $\mathbb{P} = \mathbb{F}$. Однозначность с точностью до изоморфизма вытекает из теоремы о единственности поля разложения. □

10.12 Таблица умножения для конечного поля

Конечное поле $\mathbb{P}$ с числом элементом $n = p^d$ является линейным пространством и алгеброй над полем $\mathbb{Z}_p$ размерности d. Элементы $z \in \mathbb{P}$ определяются разложениями по базису $z = \alpha_1 z_1 + \ldots + \alpha_d z_d$ или, другими словами, вектором $(\alpha_1, \ldots, \alpha_d) \in \mathbb{Z}_m^d$. Чтобы перемножать произвольные элементы, достаточно знать разложения произведений $z_i z_j$. Оказывается, при специальном выборе базиса эти разложения находятся совсем просто. Этот выбор опирается на следующий результат.

Теорема о мультипликативной группе конечного поля. *Мультипликативная группа конечного поля является циклической.*

Доказательство. Пусть G — мультипликативная группа поля с числом элементов n. Согласно теореме о конечных абелевых группах, переформулированной для мультипликативной группы, G является прямым произведением своих циклических подгрупп $G = G_1 \times \ldots \times G_s$, в котором порядок группы G_i делится на порядок группы G_{i-1} для всех $2 \leq i \leq s$. Предположим, что $s \geq 2$. Тогда $m := |G_1| < |G| = n - 1$. Любой элемента $a \in G$ имеет однозначное разложение $a = a_1 \ldots a_s$, где $a_i \in G_i$, и, следовательно, $a^m = a_1^m \ldots a_s^m = 1$. Возникает противоречие с тем, что многочлен $x^m - 1$ степени m в данном конечном поле имеет $n - 1 > m$ различных корней. Следовательно, $s = 1$. □

Следствие. *Конечное поле $\mathbb{P}$ с числом элементов $n = p^d$ получается из $\mathbb{Z}_p$ присоединением всего лишь одного элемента θ, при этом степень минимального многочлена для θ равна d.*

Доказательство. По определению, циклическая группа состоит из целых степеней какого-то одного элемента. Пусть θ — элемент, порождающий мультипликативную группу поля $\mathbb{P}$. Тогда, очевидно, $\mathbb{P} = \mathbb{Z}_p(\theta)$. Если s — степень минимального многочлена элемента θ, то, согласно теореме о присоединении корня, $(\mathbb{Z}_p(\theta) : \mathbb{Z}_p) = s \;\Rightarrow\; s = d$. $\square$

Хороший выбор базиса $z_1, \ldots, z_d$ — это элементы $1, \theta, \ldots, \theta^{d-1}$. Таблица умножения для них заполняется тривиальным образом.

Задача 5. *Докажите, что для любого простого p и натурального d над полем $\mathbb{Z}_p$ существует неприводимый многочлен степени d.*

10.13 Расширения полей при построениях циркулем и линейкой

Изучающим геометрию обычно предлагают задачи на построение циркулем и линейкой: построить треугольник по трем медианам или по трем высотам и т.п. Вот примеры более сложных задач, ставших знаменитыми в истории математики: (1) построить ребро куба, объем которого в два раза больше объема заданного куба (удвоение куба); (2) разделить угол на три равные части (трисекция угла); (3) 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. Успехи появились после осознания связи с изучением специальных расширений полей и вычислением минимальных многочленов.

Пусть на плоскости задан отрезок единичной длины. Опираясь на теорему Фалеса, с помощью циркуля и линейки мы можем построить любой отрезок рациональной длины. Мы можем считать, что имеется система координат и нам доступны точки с рациональными координатами. Если берется случайная точка, то ничто не мешает полагать, что она тоже имеет рациональные координаты. Пусть на каком-то этапе построений нам доступны любые точки, координаты которых принадлежат некоторому полю $\mathbb{P} \supseteq \mathbb{Q}$. Используя линейку, мы можем провести прямую через пару точек с координатами из поля $\mathbb{P}$. Кроме того, мы можем построить окружность, закрепив циркуль в двух точках с координатами из поля $\mathbb{P}$. Новые точки могут появиться как: (a) пересечение прямых; (b) пересечение прямой и окружности; (c) пересечение двух окружностей; (d) случайные точки на прямых и окружностях.

В случае (a) координаты новой точки будут принадлежать тому же полю $\mathbb{P}$. В случаях (b) и (c) они удовлетворяют квадратному уравнению с коэффициентами из поля $\mathbb{P}$ (проверьте!). В случае (d) при выборе случайной точки на прямой $(x, y) = (x_0, y_0) + t(x_1, y_1)$, где $x_0, y_0, x_1, y_1 \in \mathbb{P}$ можно считать, что $t \in \mathbb{P}$, и значит, случайная точка получает координаты в том же поле $\mathbb{P}$. При выборе случайной точки на окружности $(x_0 - x)^2 + (y_0 - y)^2 = R^2$, где $x_0, y_0, R^2 \in \mathbb{P}$, можно считать, что, скажем, $x \in \mathbb{P}$, и тогда для y получается квадратное уравнение с коэффициентами из $\mathbb{P}$. В любом случае новые точки имеют координаты в некотором расширении поля $\mathbb{P}$, которое возникает при присоединении корней квадратных трехчленов над полем $\mathbb{P}$. Таким образом, алгоритм построения с помощью циркуля и линейки порождает цепочку расширений

$$\mathbb{Q} = \mathbb{Q}_0 \;\subset\; \mathbb{Q}_1 \;\subset\; \ldots \;\subset\; \mathbb{Q}_{s-1} \;\subset\; \mathbb{Q}_s, \qquad \mathbb{Q}_i = \mathbb{Q}_{i-1}(\sqrt{D_i}), \quad D_i \in \mathbb{Q}_{i-1}, \quad \sqrt{D_i} \notin \mathbb{Q}_{i-1}.$$

Теперь совсем просто можно получить, например, такой результат.

Утверждение. *Задача об удвоении куба неразрешима с помощью циркуля и линейки.*

Доказательство. В данном случае цель построений — отрезок длины $\sqrt[3]{2}$. Если построение возможно, то существует такая цепочка расширений, в которой $\sqrt[3]{2} \in \mathbb{Q}_s \backslash \mathbb{Q}_{s-1}$. Следовательно,

$$\sqrt[3]{2} = a + b\sqrt{D_s}, \quad a, b \in \mathbb{Q}_{s-1}, \quad b \neq 0.$$

Возводя в куб, находим $2 = a^3 + 3a^2 b\sqrt{D_s} + 3ab^2 D_s + b^3 D_s\sqrt{D_s} \;\Rightarrow\; 2 - a^3 - 3ab^2 D_s = (3a^2 + b^2 D_s) b\sqrt{D_s}$. Учитывая, что $b \neq 0$ и $3a^2 + b^2 D_s > 0$, получаем $\sqrt{D_s} = \frac{2 - a^3 - 3ab^2 D_s}{(3a^2 + b^2 D_s) b} \in \mathbb{Q}_{s-1}$, что противоречит предположению о том, что $\sqrt{D_s} \notin \mathbb{Q}_{s-1}$. $\square$

Попробуем посмотреть на ту же задачу с позиций некоторого общего метода. Пусть степень минимального многочлена числа $\theta \in \mathbb{Q}_s$ над полем $\mathbb{Q}$ равна m. Тогда, в силу теоремы о степенях расширений, число $m = (\mathbb{Q}(\theta) : \mathbb{Q})$ должно быть делителем числа $(\mathbb{Q}_s : \mathbb{Q}) = 2^s$. В случае $\theta = \sqrt[3]{2}$ минимальный многочлен имеет вид $x^3 - 2 \;\Rightarrow\; m = 3$. Построение невозможно, так как $2^s \;\vdots\!\!\!/\; 3$.

10.14 Неприводимость многочленов с целыми коэффициентами

Минимальный многочлен над полем $\mathbb{Q}$ для числа θ — это то же самое, что неприводимый над $\mathbb{Q}$ многочлен с корнем θ. Такой многочлен с точностью до ненулевого множителя совпадает с многочленом из кольца многочленов с целыми коэффициентами, который, согласно лемме Гаусса, будет в этом кольце неприводимым.

В задаче о построении правильного n-угольника нужно построить отрезок длины $\alpha = \cos(2\pi/n)$. Однако, для анализа задачи удобнее ввести комплексное число

$$\varepsilon = \cos(2\pi/n) + \mathbf{i}\sin(2\pi/n) = \mathrm{e}^{\frac{2\pi}{n}\mathbf{i}}$$

и рассмотреть еще одно расширение $\mathbb{Q}_s \subset \mathbb{Q}_s(\varepsilon)$. Заметим, что ε является корнем квадратного трехчлена $x^2 - 2\alpha x + 1$ с коэффициентами в поле $\mathbb{Q}_s$, который, очевидно, неприводим над $\mathbb{Q}_s$, и, следовательно, $(\mathbb{Q}_s(\varepsilon) : \mathbb{Q}_s) = 2$. Если построение возможно, то степень m минимального многочлена должна быть делителем числа 2^{s+1}.

Число $\varepsilon = (2\pi/2)\mathbf{i}$ удовлетворяет уравнению $x^n - 1 = 0$. Однако, многочлен $x^n - 1$ не является минимальным для ε, так как он приводим: $x^n - 1 = (x-1)f(x)$, где $f(x) = x^{n-1} + \ldots + x + 1$. Очевидно, ε будет корнем многочлена $f(x)$. Но будет ли $f(x)$ искомым минимальным многочленом?

Если n простое, то ответ положительный.

Утверждение. *Для простого n многочлен $f(x) = x^{n-1} + \ldots + x + 1$ неприводим.*

Первое доказательство. От противного, предположим, что есть нетривиальное разложение $f(x) = a(x)b(x)$. Обозначим через $\bar{f}(x), \bar{a}(x), \bar{b}(x)$ многочлены над полем $\mathbb{Z}_n$, коэффициенты которых суть вычеты, порожденные соответствующими коэффициентами многочленов $f(x), a(x), b(x)$. Для краткости можно говорить, что это те же многочлены по модулю n. Каждый из многочленов $\bar{a}(x)$, $\bar{b}(x)$ является делителем многочлена $x^n - 1 \in \mathbb{Z}_n[x]$. Однако, в поле $\mathbb{Z}_n$ имеет место разложение $x^n - 1 = (x-1)^n$. Поэтому в силу факториальности кольца многочленов над полем, многочлены $\bar{a}(x)$ и $\bar{b}(x)$ имеют вид

$$\bar{a}(x) = (x-1)^k, \quad \bar{b}(x) = (x-1)^l \quad \Rightarrow \quad a(1) \;\vdots\; n, \;\; b(1) \;\vdots\; n \quad \Rightarrow \quad f(1) = a(1)b(1) \;\vdots\; n^2.$$

Однако $f(1) = n$ делится на n^2 только при $n = 1$. $\square$

Второе доказательство. Неприводимость $f(x)$, очевидно, равносильна неприводимости многочлена

$$F(x) = f(x+1) = \frac{(x+1)^n - 1}{x}.$$

Старший коэффициент $F(x)$ равен 1, а все остальные коэффициенты делятся на простое число $p = n$ и при этом свободный член равен n, и значит, не делится на p^2. Многочлен с такими свойствами неприводим согласно следующему признаку неприводимости.

Признак Эйзенштейна. *Пусть старший коэффициент многочлена $F(x)$ с целыми коэффициентами не делится на простое число p, а все остальные коэффициенты делятся на p, но при этом свободный член не делится на p^2. Тогда $F(x)$ неприводим над кольцом целых чисел.*

От противного, пусть имеется нетривиальное разложение $F(x) = (a_0 + \ldots + a_k x^k)(b_0 + \ldots + b_l x^m)$. Тогда $a_0 b_0$ делится на p, но не на p^2. Поэтому одно и только одно из чисел a_0, b_0 делится на p. Пусть это будет b_0. Среди коэффициентов $b_0, \ldots, b_l$ должен быть не делящийся на p, иначе старший коэффициент $F(x)$ будет делиться на p. Пусть b_i — первый такой коэффициент. Тогда i-й коэффициент $F(x)$ имеет вид $a_0 b_i + (a_1 b_{i-1} + \ldots a_i b_0)$ и не может делиться на p, так как число в скобках делится на p, а произведение $a_0 b_i$ не делится. Значит, i равно степени многочлена $F(x)$ и рассмотренное разложение является тривиальным. $\square$

10.15 Многочлены деления круга

В случае составного n многочлен $f(x) = x^{n-1} + \ldots + x + 1$ оказывается приводимым. Минимальный многочлен для $\varepsilon = \mathrm{e}^{(2\pi/n)\mathbf{i}}$ строится следующим образом:

$$\Phi_n(x) = \prod_{k \in E_n} (x - \varepsilon^k), \quad E_n = \{k : \ 1 \le k \le n, \ k \text{ и } n \text{ взаимно просты}\}.$$

Линейные множители в составе $\Phi_n(x)$ отвечают первообразным корням степени n из единицы, $\deg \Phi_n(x) = \phi(n)$ (функция Эйлера). Многочлены $\Phi_n(x)$ называются *многочленами деления круга* или *круговыми многочленами*.

Теорема о многочленах деления круга. *Многочлены деления круга имеют целые коэффициенты и являются неприводимыми над кольцом целых чисел.*

Доказательство. Прямое вычисление дает

$$\Phi_1(x) = x - 1, \quad \Phi_2(x) = x + 1, \quad \Phi_3(x) = x^2 + x + 1, \quad \Phi_4(x) = x^2 + 1, \quad \Phi_5(x) = x^4 + x^3 + x^2 + x + 1.$$

В общем случае нужно заметить, что

$$x^n - 1 \ = \ \prod_{d \,|\, n} \Phi_d(x).$$

Объяснение очень простое: каждый корень степени n из единицы является корнем одного и только одного многочлена $\Phi_d(x)$, для которого d есть порядок этого корня как элемента мультипликативной циклической группы корней степени n из единицы. Если уже доказано, что при всех $d < n$ многочлены $\Phi_d(x)$ являются приведенными многочленами с целыми коэффициентами, то $\Phi_n(x)$ получается делением $x^n - 1$ на некоторый приведенный многочлен с целыми коэффициентами по алгоритму деления столбиком, в котором могут возникать только многочлены с целыми коэффициентами.

Неприводимость доказывается несколько сложнее. Обозначим через $F(x)$ минимальный многочлен для ε, имеющий целые коэффициенты. Он делит многочлен $x^n - 1$ и поэтому его старший коэффициент равен ± 1. Будем считать, что старший коэффициент равен 1. Возьмем любое простое число p, которое не делит n. Тогда ε^p также будет первообразным корнем степени n из единицы. Его минимальный многочлен с целыми коэффициентами обозначим через $G(x)$. Как и для $F(x)$, будем считать, что старший коэффициент $G(x)$ равен 1. Оба многочлена $F(x)$ и $G(x)$ неприводимы и поэтому если $F(x) \ne G(x)$, то они взаимно просты. Допустим, что $F(x) \ne G(x)$. Тогда $x^n - 1$ делится на их произведение: $x^n - 1 = F(x)G(x)Q(x)$, где $Q(x)$ — приведенный многочлен с целыми коэффициентами. Полученное равенство рассмотрим как равенство многочленов по модулю p: $x^n - 1 = \bar{F}(x)\bar{G}(x)\bar{Q}(x)$, где черта означает переход к многочлену, в котором коэффициенты заменяются на порожденные ими вычеты по модулю p. Далее, ε является общим корнем многочленов $G(x^p)$ и $F(x)$. Поскольку $F(x)$ неприводим, $G(x^p)$ должен делиться на $F(x)$: $G(x^p) = F(x)H(x)$, где $H(x)$ — приведенный многочлен с целыми коэффициентами. Переходя к многочленам по модулю p, находим $\bar{G}(x^p) = (\bar{G}(x))^p = \bar{F}(x)\bar{H}(x)$. Значит, многочлены $\bar{G}(x)$ и $\bar{F}(x)$ в каком-то расширении поля $\mathbb{Z}_p$ обладают общим корнем, и, следовательно, многочлен $x^n - 1 \in \mathbb{Z}_p[x]$ имеет кратный корень. Этот корень должен быть также корнем производной $\bar{F}'(x) = [n]_p x^{n-1}$ и, значит, равен 0. В то же время нетрудно понять, что свободный член многочлена $F(x)$ равен ± 1 и таким же будет свободный член многочлена $\bar{F}(x)$. Полученное противоречие означает, что $F(x) = G(x)$.

Теперь возьмем произвольное число $k \in E_n$ и рассмотрим его разложение на простые множители $k = p_1 p_2 \ldots p_s$. Ясно, что каждый из них не делит n. Значит, вместе с ε корнями многочлена $F(x)$ должны быть ε^{p_1}, $\varepsilon^{p_1 p_1}$, $\ldots$, $\varepsilon^{p_1 \cdots p_s} = \varepsilon^k$. Поэтому $F(x)$ делится на каждый из попарно взаимно простых линейных многочленов $x - \varepsilon^k$ при $k \in E_k$, и следовательно, он делится на их произведение $\Phi_n(x)$. В то же время $F(x)$ является неприводимым многочленом, и это означает, что $F(x) = \Phi_n(x)$. $\square$

10.16 Задача о построении правильных многоугольников

Если правильный n-угольник строится с помощью циркуля и линейки, то степень m минимального многочлена числа $\varepsilon = \mathrm{e}^{(2\pi/n)\mathbf{i}}$ должна быть степенью двойки. В нашем случае $m = \phi(n)$.

Теорема. *Для того чтобы правильный n-угольник строился с помощью циркуля и линейки необходимо, чтобы разложение числа n на простые множители имело вид*

$$n = 2^k p_1 \dots p_s, \quad p_i = 2^{k_i} + 1, \quad 1 \le i \le s. \qquad (*)$$

Доказательство. Пусть разложение n на простые множители имеет вид $n = q_1^{l_1} \dots q_t^{l_t}$, где $q_1, \dots, q_t$ — попарно различные простые числа. Тогда число $\phi(n) = \prod\limits_{i=1}^{t} (q_i^{l_i} - q_i^{l_i - 1})$ делит степень двойки. Значит, $q_i^{l_i - 1}(q_i - 1)$ есть степень двойки. Если $q_i > 2$, то число $q_i - 1$ должно быть степенью двойки. $\square$

Замечание. Условие $(*)$ является также и достаточным, но этот факт мы оставляем без доказательства.

10.17 Построение правильного 17-угольника

Одно из самых ранних достижений Гаусса — это его метод построения правильного 17-угольника ($17 = 2^4 + 1$) и яркая демонстрация пользы комплексных чисел. В данном случае наша цель — отрезок длины $2\cos(2\pi/17) = \varepsilon + \varepsilon^{-1}$, где $\varepsilon = \mathrm{e}^{(2\pi/17)\mathbf{i}}$.

Преобразование $\varepsilon \to \varepsilon^2$ разбивает множество отличных от единицы корней на два непересекающихся подмножества. Суммы корней в них равны

$$\alpha_1 = \varepsilon + \varepsilon^2 + \varepsilon^4 + \varepsilon^8 + \varepsilon^{-1} + \varepsilon^{-2} + \varepsilon^{-4} + \varepsilon^{-8}, \quad \alpha_2 = \varepsilon^3 + \varepsilon^6 + \varepsilon^{-5} + \varepsilon^7 + \varepsilon^{-3} + \varepsilon^{-6} + \varepsilon^5 + \varepsilon^{-7}.$$

Наличие комплексно сопряженных пар показывает, что числа α_1 и α_2 вещественны, и нетрудно проверить, что $\alpha_1 + \alpha_2 = -1$, $\alpha_1\alpha_2 = -4 \;\Rightarrow\;$ числа α_1 и α_2 являются корнями квадратного уравнения с целыми коэффициентами, и значит, строятся с помощью циркуля и линейки.

Преобразования $\varepsilon \to \varepsilon^4$ и $\varepsilon \to \varepsilon^{-1}$ дают соответственно четыре и восемь непересекающихся подмножеств с суммами

$$\beta_1 = \varepsilon + \varepsilon^4 + \varepsilon^{-1} + \varepsilon^{-4}, \quad \beta_2 = \varepsilon^2 + \varepsilon^8 + \varepsilon^{-2} + \varepsilon^{-8}, \quad \beta_3 = \varepsilon^3 + \varepsilon^{-5} + \varepsilon^{-3} + \varepsilon^5, \quad \beta_4 = \varepsilon^6 + \varepsilon^7 + \varepsilon^{-6} + \varepsilon^{-7}.$$

$$\gamma_1 = \varepsilon + \varepsilon^{-1}, \quad \gamma_2 = \varepsilon^4 + \varepsilon^{-4}, \quad \gamma_3 = \varepsilon^2 + \varepsilon^{-2}, \quad \gamma_4 = \varepsilon^8 + \varepsilon^{-8},$$
$$\gamma_5 = \varepsilon^3 + \varepsilon^{-3}, \quad \gamma_6 = \varepsilon^5 + \varepsilon^{-5}, \quad \gamma_7 = \varepsilon^6 + \varepsilon^{-6}, \quad \gamma_8 = \varepsilon^7 + \varepsilon^{-7}.$$

Все эти числа вещественные и удовлетворяют следующим соотношениям:

$$\beta_1 + \beta_2 = \alpha_1, \;\; \beta_1\beta_2 = -1, \qquad \beta_3 + \beta_4 = \alpha_2, \;\; \beta_3\beta_2 = -1, \qquad \gamma_1 + \gamma_2 = \beta_1, \;\; \gamma_1\gamma_2 = \beta_3.$$

Отрезок длины γ_1 является искомым и строится через решения квадратных уравнений.

Алгебра и геометрия (1 поток)

Лекция 11

11.1 Квадратные уравнения

Основная теорема алгебры — это теорема существования корней. Можно ли для корней получить удобные для вычислений формулы?

В школьной математике много внимания уделяется формуле для корней квадратных трехчленов с неотрицательным дискримантом. В случае комплексных коэффициентов все делается по той же схеме. Выделением полного квадрата

$$z^2+az+b \;=\; \left(z^2 \;+\; 2\,\frac{a}{2}\,z \;+\; \left(\frac{a}{2}\right)^2\right) \;+\; \left(b-\left(\frac{a}{2}\right)^2\right) \;=\; \left(z+\frac{a}{2}\right)^2 \;+\; \left(b-\frac{a^2}{4}\right)$$

уравнение $z^2+az+b=0$ сводится к равносильному уравнению $\left(z+\frac{a}{2}\right)^2=\frac{D}{4}$, где $D=a^2-4b$ — дискриминант многочлена z^2+az+b. После записи дискриминанта в тригонометрической форме $D=|D|(\cos\phi+\mathbf{i}\sin\phi)\neq 0$ у нас возникает формула

$$z_\pm=-\frac{a\pm\sqrt{|D|}\left(\cos\frac{\phi}{2}+\mathbf{i}\sin\frac{\phi}{2}\right)}{2}.$$

Если $D=0$, то единственное решение имеет вид $z=-a/2$. В этом и только этом случае квадратный трехчлен является квадратом линейного двучлена: $z^2+az+b=\left(z+\frac{a}{2}\right)^2$. Если $D\neq 0$, то формула дает пару комплексных корней. В случае $D>0$ корни оказываются вещественными.

11.2 Кубические уравнения

Кубическое уравнение $z^3+a_2z^2+a_1z+a_0=0$ с помощью замены $z=x-a_2/3$ приводится к более удобному виду $x^3+px+q=0$. Запишем решение 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, если

$$\begin{cases} u^3+v^3 &= -q,\\ uv &= -p/3. \end{cases} \qquad\Rightarrow\qquad \begin{cases} u^3+v^3 &= -q,\\ u^3v^3 &= -p^3/27. \end{cases}$$

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 u и v 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.

Как видим, формула для корней кубических уравнений у нас имеется, но вряд ли ее можно считать удобной для вычислений.

11.3 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 $x^4 + px^2 + qx + r = 0$.

Редуцированное уравнение может быть сведено к кубическому. Наиболее простой способ для этого был предложен итальянским математиком Феррари. Идея состоит в том, чтобы представить его левую часть как разность двух квадратов:

$$x^4 + px^2 + qx + r \;=\; (x^2 + y/2)^2 - ((y-p)x^2 - qx + (y^2/4 - r)) \;=\; (x^2 + y/2)^2 - (\alpha x + \beta)^2,$$

где α и β — многочлены от p, q, r, y. 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 $\alpha,\ \beta$ находим

$$x^4 + px^2 + qx + r \;=\; (x^2 + y/2)^2 - (\alpha x + \beta)^2 \;=\; (x^2 + y/2 + \alpha x + \beta)(x^2 + y/2 - \alpha x - \beta).$$

Таким образом, получение решений для редуцированного уравнения сводится к решению одного кубического и нескольких квадратных уравнений.

11.4 Общее алгебраическое уравнение

Пусть $\mathbb{L} = \mathbb{C}(x_1, \ldots, x_n)$ — поле рациональных функций от переменных $x_1, \ldots, x_n$ и $\mathbb{K} = \mathbb{C}(E_1, \ldots, E_n)$ есть минимальное расширение поля $\mathbb{C}$, содержащее элементарные симметрические многочлены $E_1, \ldots, E_n$ от переменных $x_1, \ldots, x_n$. Будем говорить, что общее алгебраическое уравнение степени n *разрешимо в радикалах*, если для некоторого поля $\mathbb{M} \supseteq \mathbb{L}$ существует цепочка расширений с такими свойствами:

$$\mathbb{K} = K_0 \subset K_1 \subset \ldots \subset K_{s-1} \subset K_s, \quad \mathbb{L} \subseteq K_s \subseteq \mathbb{M},$$

$$K_i = K_{i-1}(\theta_i), \quad \theta_i \in \mathbb{M} \setminus K_{i-1}, \quad \theta_i^{n_i} \in K_{i-1}, \quad n_i \text{ — натуральное число.}$$

[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.

Теорема Руффини. *Если $\mathbb{M} = \mathbb{L}$, то при $n \geq 5$ цепочки расширений с указанными свойствами не существует.*

Доказательство. От противного, предположим, что такая цепочка расширений существует. Тогда

$$x_1 = f(E_1, \ldots, E_n, \theta_1, \ldots, \theta_s),$$

$$\theta_1^{n_1} = f_1(E_1, \ldots, E_n), \quad \theta_2^{n_2} = f_2(E_1, \ldots, E_n, \theta_1), \quad \ldots, \quad \theta_s^{n_s} = f_s(E_1, \ldots, E_n, \theta_1, \ldots, \theta_{s-1}),$$

где f, f_1, …, f_s — рациональные функций своих аргументов и одновременно рациональные функции от переменных $x_1, \ldots, x_s$. Пусть $\sigma \in S_n$ — подстановка n-й степени и $F(x_1, \ldots, x_n)$ — рациональная функция. Будем считать, что σF — это новая рациональная функция, определенная правилом $(\sigma F)(x_1, \ldots, x_N) := F(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. Заметим, что $\sigma\theta^{n_1} = (\sigma\theta)^{n_1} = \sigma f_1 = f_1$, так как $E_1, \ldots, E_n$ переходят в себя при действии любых подстановок. Следовательно, $(\sigma\theta_1/\theta_1)^{n_1} = 1 \;\Rightarrow\; \sigma\theta_1 = \varepsilon\theta_1$, где ε — некоторый корень степени n_1 из единицы.

Пусть $\sigma = (1, 2, 3)$ и $\tau = (3, 4, 5)$ — два тройных цикла. Тогда $\tau\sigma = (1, 2, 3, 4, 5)$ и $\tau\sigma^2 = (1, 3, 4, 5, 2)$ — два пятерных цикла. Если $\sigma\theta_1 = \varepsilon_\sigma\theta_1$ и $\tau\theta_1 = \varepsilon_\tau\theta_1$ для каких-то корней из единицы ε_σ и ε_τ, то $(\tau\sigma)\theta_1 = \varepsilon_\tau\varepsilon_\sigma\theta_1$ и $(\tau\sigma^2)\theta_1 = \varepsilon_\tau\varepsilon_\sigma^2\theta_1$. Следовательно,

$$\sigma^3\theta_1 = \theta_1 = \varepsilon_\sigma^3\theta_1, \quad \tau^3\theta_1 = \theta_1 = \varepsilon_\tau^3\theta_1,$$

$$(\tau\sigma)^5\theta_1 = \theta_1 = (\varepsilon_\tau\varepsilon_\sigma)^5\theta_1, \quad (\tau\sigma^2)^5\theta_1 = \theta_1 = (\varepsilon_\tau\varepsilon_\sigma^2)^5\theta_1 \quad \Rightarrow$$

$$\varepsilon_\sigma^3 = \varepsilon_\tau^3 = \varepsilon_\tau^5\varepsilon_\sigma^5 = \varepsilon_\tau^5\varepsilon_\sigma^{10} = 1 \quad \Rightarrow \quad \varepsilon_\tau = \varepsilon_\sigma = 1.$$

Таким образом, мы доказали, что из равенств $\sigma f_1 = \tau f_1 = f_1$ вытекают равенства $\sigma\theta_1 = \tau\theta_1 = \theta_1 \;\Rightarrow\; \sigma f_2 = \tau f_2 = f_2$. Повторяя аналогичные рассуждения, мы придем к равенствам $\sigma f = \tau f = f$ и в итоге к абсурдному равенству $x_2 = x_1$. $\square$

Условие $\mathbb{M} = \mathbb{L}$ можно убрать, но это требует существенно большей изобретательности. Впервые это сделал Абель. Полученная им теорема утверждает, что *при $n \geq 5$ общее алгебраическое уравнение не решается в радикалах.*

11.5 Радикальные расширения

Рассмотрим конкретное уравнение, например $x^5 - 4x + 2 = 0$. Выражаются ли его корни в радикалах? Теорема Абеля утверждает, что нет единой формулы для *всех* уравнений 5-й степени, но не дает ответа для конкретного уравнения с заданными числовыми коэффициентами. Вопрос ставится следующим образом. Пусть задано уравнение

$$f(x) \;=\; a_0 \;+\; a_1 x \;+\; \ldots \;+\; a_{n-1}x^{n-1} \;+\; x^n \;\;=\;\; 0$$

с коэффициентами из заданного числового поля $\mathbb{K} \subsetneq \mathbb{C}$, которое мы будем называть *основным полем*, и пусть $\mathbb{L} \subseteq \mathbb{C}$ — поле разложения многочлена $f(x)$ над $\mathbb{K}$. Нас интересует существование цепочки расширений

$$\mathbb{K} = \mathbb{K}_0 \subset \mathbb{K}_1 \subset \ldots \subset \mathbb{K}_{s-1} \subset \mathbb{K}_s, \quad \mathbb{L} \subseteq \mathbb{K}_s \subseteq \mathbb{C},$$

в которой каждое отдельное расширение является *радикальным расширением*:

$$\mathbb{K}_i = \mathbb{K}_{i-1}(\theta_i), \quad \theta_i \in \mathbb{C}\setminus\mathbb{K}_{i-1}, \quad \theta_i^{n_i} \in \mathbb{K}_i, \quad n_i \text{ — натуральное число.}$$

Можно доказать, что от интересующей нас цепочки радикальных расширений всегда можно перейти к цепочке радикальных расширений, к которой последнее поле будет полем разложения $\mathbb{M}$ некоторого многочлена над полем $\mathbb{K}$ (вообще говоря, не совпадающего с $f(x)$). В теории Галуа вопрос о существовании радикальных расширений для заданного уравнения $f(x) = 0$ сводится к изучению группы автоморфизмов поля $\mathbb{M}$, оставляющих на месте каждый элемент поля $\mathbb{K}$. Такие автоморфизмы называются *автоморфизмами поля $\mathbb{M}$ над полем $\mathbb{K}$*. То, что они образуют группу относительно операции композиции автоморфизмов, вполне очевидно (докажите!). Для этой группы используется обозначение $\mathrm{Aut}(\mathbb{M} : \mathbb{K})$.

11.6 Примитивный элемент

Пусть $\mathbb{M}$ — поле разложения некоторого многочлена над полем $\mathbb{K}$. Как найти число автоморфизмов в группе автоморфизмов поля $\mathbb{M}$ над $\mathbb{K}$?

Чтобы это понять, проще всего опираться на то, что в наиболее интересных случаях поле $\mathbb{M}$ можно получить присоединением всего лишь одного числа $\theta \in \mathbb{M}$ — вообще говоря, отличного от корней многочлена, для которого строится поле разложения. Такое число θ называется *примитивным элементом* расширения $\mathbb{M} \supseteq \mathbb{K}$.

Примитивный элемент заведомо существует в любом конечном расширении конечного поля, так как любое конечное поле (а значит, и любое конечное расширение конечного поля) порождается присоединением одного элемента к полю вычетов по простому модулю. Поэтому будем далее считать, что поле $\mathbb{K}$ есть поле нулевой характеристики, или, что одно и то же, содержит в себе поле рациональных чисел.

Теорема о примитивном элементе. *Пусть поле $\mathbb{K}$ является расширением поля рациональных чисел и $\theta_1, \ldots, \theta_n$ — алгебраические числа над полем $\mathbb{K}$. Тогда в поле $\mathbb{K}(\theta_1, \ldots, \theta_n)$ существует примитивный элемент вида $\theta = \theta_1 + c_2\theta_2 + \ldots + c_n\theta_n$, где $c_2, \ldots, c_n$ — целые числа.*

Доказательство. Достаточно изучить случай $n = 2$. Пусть α и β — корни своих минимальных многочленов $F(x)$ и $G(x)$ с коэффициентами из поля $\mathbb{K}$. Будем считать их приведенными многочленами степени k и l, соответственно. Случаи $k = 1$ или $l = 1$ тривиальны. Поэтому полагаем, что $k, l \geq 2$.

Мы будем опираться на то, что все корни минимальных многочленов $F(x)$ и $G(x)$ принадлежат некоторому расширению поля $\mathbb{K}$ и являются простыми. Если бы имелся кратный корень, то он был бы также корнем производной минимального многочлена. Однако, если поле $\mathbb{K}$ содержит поле рациональных чисел, то производная многочлена ненулевой степени не может быть нулевым многочленом. Поэтому наличие кратного корня у многочлена означает, что он не может быть минимальным. Пусть

$$F(x) = \prod_{i=1}^{k}(x - \alpha_i), \quad G(x) = \prod_{j=1}^{l}(x - \beta_j), \quad \alpha_1 = \alpha, \quad \beta_1 = \beta.$$

Выберем ненулевое целое число

$$c \neq \frac{\alpha - \alpha_i}{\beta_j - \beta} \quad \text{для всех пар } (i,\, j) \neq (1,\, 1)$$

и положим $\theta = \alpha + c\beta$. Тогда $(\theta - \alpha_i)/c \neq \beta_j$ при всех i, j, кроме $i = j = 1$. Следовательно, многочлен $\Phi(x) := G((\theta - x)/c)$ имеет своим корнем α, но ни один из его корней не равен α_i при $2 \leq i \leq k$. Значит, $\Phi(x)$ и $F(x)$ имеют в точности один общий корень α. Поэтому их наибольший общий делитель равен $x - \alpha$ с точностью до числового множителя из поля $\mathbb{K}(\theta)$, которому принадлежат коэффициенты обоих многочленов, и может быть найден с помощью алгоритма Евклида, в котором все многочлены имеют коэффициенты из поля $\mathbb{K}(\theta)$. Таким образом, $\alpha \in \mathbb{K}(\theta)$, а значит и $\beta \in \mathbb{K}(\theta)$. □

11.7 Число автоморфизмов

Теорема о числе автоморфизмов. *Пусть поле $\mathbb{K}$ содержит в себе поле рациональных чисел. Тогда для любого конечного расширения $\mathbb{L} \supseteq \mathbb{K}$ число автоморфизмов поля $\mathbb{L}$ над $\mathbb{K}$ не превышает степени расширения.*

Доказательство. Согласно теореме о примитивном элементе, $\mathbb{L} = \mathbb{K}(\theta)$. По теореме о присоединении корня, степень расширения $(\mathbb{L} : \mathbb{K})$ равна степени минимального многочлена $F(x) \in \mathbb{K}[x]$ числа θ. Рассмотрим разложение $F(x) = \prod_{i=1}^{m}(x - \theta_i)$, в котором $\theta_1 = \theta$. Мы уже знаем, что при выборе любого θ_i существует изоморфизм $\Phi_i : \mathbb{K}(\theta) \to \mathbb{K}(\theta_i)$, переводящий θ в θ_i. Если $\theta_i \in \mathbb{L}$, то Φ_i будет автоморфизмом поля $\mathbb{L}$. Остается заметить, что любой автоморфизм поля $\mathbb{L}$ переводит θ в один из корней того же многочлена $F(x)$ (докажите!). Следовательно, число автоморфизмов не больше числа корней многочлена $F(x)$, принадлежащих полю $\mathbb{L}$. □

Замечание. В действительности теорема справедлива для *любого* поля $\mathbb{K}$. Условие на поле $\mathbb{K}$ позволило нам применить теорему о примитивном элементе и получить очень короткое доказательство.

Число автоморфизмов $\mathbb{L}$ над $\mathbb{K}$ *может оказаться меньше* степени расширения. Пусть, например, $\mathbb{K} = \mathbb{Q}$ и $\mathbb{L} = \mathbb{Q}(\sqrt[3]{2})$. Тогда единственным автоморфизмом $\mathbb{L}$ над $\mathbb{K}$ будет тождественное отображение, а степень расширения равна 3 (докажите!). Если число автоморфизмов равно степени расширения, то такое расширение называется *расширением Галуа*, а группа автоморфизмов называется *группой Галуа*. Из нашего доказательства теоремы о числе автоморфизмов сразу возникает следующее

Утверждение. *Поле $\mathbb{K}(\theta) \supseteq \mathbb{K}$ является конечным расширением Галуа в том и только том случае, когда все корни минимального многочлена числа θ принадлежат $\mathbb{L}$.*

11.8 Характеристическое свойство расширений Галуа

При знакомстве с идеями совсем не обязательно стремиться к максимально абстрактным формулировкам. Поэтому при обсуждении расширений Галуа мы будем полагать, что все поля вложены в поле комплексных чисел.

Теорема о расширениях Галуа. *Для того чтобы поле $\mathbb{L} \supseteq \mathbb{K}$ было расширением Галуа над $\mathbb{K}$, необходимо и достаточно, чтобы оно было полем разложения какого-то многочлена над полем $\mathbb{K}$.*

Доказательство. Необходимость следует из доказательства теоремы о числе автоморфизмов. Достаточность вытекает из следующего замечательного свойства полей разложения: *для любого неприводимого над $\mathbb{K}$ многочлена из принадлежности полю разложения $\mathbb{L}$ хотя бы одного его корня вытекает, что все остальные корни также принадлежат $\mathbb{L}$.*

Пусть поле $\mathbb{L}$ получено присоединением к полю $\mathbb{K}$ всех корней многочлена $F(x) = \prod_{i=1}^{n} (x - \theta_i)$. Тогда из теоремы о присоединении корня следует, что любой элемент поля $\mathbb{L}$ имеет вид $g(\theta_1, \ldots, \theta_n)$, где $g(x_1, \ldots, x_n)$ — многочлен от n переменных с коэффициентами из поля $\mathbb{K}$. В силу формул Виета и теоремы о симметрических многочленах все коэффициенты многочлена

$$\Psi(x) = \prod_{\sigma \in S_n} \bigl(x - g(\theta_{\sigma(1)},\ \ldots,\ \theta_{\sigma(n)})\bigr)$$

принадлежат полю $\mathbb{K}$. Пусть $\alpha = g(\theta_1, \ldots, \theta_n) \in \mathbb{L}$ — корень неприводимого многочлена $\phi(x) \in \mathbb{K}[x]$ 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 неприводимости многочлена $\phi(x)$, он не имеет нетривиальных делителей и, следовательно, $\Psi(x) \,\vdots\, \phi(x) \quad \Rightarrow \quad \beta = g(\theta_{\sigma(1)}, \ldots, \theta_{\sigma(n)})$ для какой-то подстановки $\sigma \in S_n$ $\Rightarrow \quad \beta \in \mathbb{L}$. $\square$

Заметим также, что если расширение $\mathbb{L} \supseteq \mathbb{K}$ конечно и для любого неприводимого многочлена над $\mathbb{K}$ из принадлежности хотя бы одного корня полю $\mathbb{L}$ вытекает, что все остальные его корни тоже содержатся в $\mathbb{L}$ (такие расширения называются *нормальными*), то $\mathbb{L}$ есть поле разложения какого-то многочлена над $\mathbb{K}$. В самом деле, если $\mathbb{L} = \mathbb{K}(\theta_1, \ldots, \theta_s)$ и $f_1, \ldots, f_s$ — минимальные многочлены для $\theta_1, \ldots, \theta_s$, то $\mathbb{L}$ будет полем разложения многочлена $f = f_1 \ldots f_s$.

11.9 Промежуточные поля и подгруппы

Пусть H — подгруппа группы $G = \mathrm{Aut}(\mathbb{L} : \mathbb{K})$. Обозначим через $\mathbb{L}^H$ множество всех элементов $a \in \mathbb{L}$ таких, что $h(a) = a$ для всех автоморфизмов $h \in H$. Очевидно, что множество $\mathbb{L}^H$ является *промежуточным полем*: $\mathbb{K} \subseteq \mathbb{L}^H \subseteq \mathbb{L}$ (почему?).

Что можно сказать о поле $\mathbb{L}^G$? По определению, группа G оставляет на месте каждый элемент поля $\mathbb{K}$. Но может ли случиться так, что она оставляет неподвижными также элементы более широкого поля?

Теорема о неподвижном поле. *Пусть $\mathbb{L}$ — поле разложения какого-то многочлена над полем $\mathbb{K}$ и G — какая-то группа, состоящая из автоморфизмов поля $\mathbb{L}$ над $\mathbb{K}$. Тогда*

$$G = \mathrm{Aut}(\mathbb{L} : \mathbb{K}) \quad \Leftrightarrow \quad \mathbb{L}^G = \mathbb{K}.$$

Доказательство. Пусть $G = \mathrm{Aut}(\mathbb{L} : \mathbb{K}) \ \Rightarrow \ |G| = (\mathbb{L} : \mathbb{K})$. Поле $\mathbb{L}$ есть поле разложения некоторого многочлена над $\mathbb{K} \ \Rightarrow$ поле $\mathbb{L}$ есть поле разложения того же самого многочлена над полем $\mathbb{L}^G$. Следовательно, $(\mathbb{L} : \mathbb{L}^G) \geq |G| = (\mathbb{L} : \mathbb{K})$ и, учитывая равенство $(\mathbb{L} : \mathbb{L}^G)(\mathbb{L}^G : \mathbb{K}) = (\mathbb{L} : \mathbb{K})$, находим $(\mathbb{L}^G : \mathbb{K}) = 1 \ \Rightarrow \ \mathbb{L}^G = \mathbb{K}$.

Теперь предположим, что есть некоторая группа автоморфизмов поля $\mathbb{L}$ над $\mathbb{K}$, для которой $\mathbb{L}^G = \mathbb{K}$. Пусть $\mathbb{L} = \mathbb{K}(\theta)$ и пусть $g_1, \ldots, g_m$ — максимальный набор автоморфизмов группы G, для которых числа $g_1(\theta)$, …, $g_m(\theta)$ попарно различны. Тогла коэффициенты многочлена $\phi(x) = \prod\limits_{i=1}^{m}(x - g_i(\theta))$ остаются на месте при действии любого автоморфизма группы G (почему?) и, следовательно, принадлежат полю $\mathbb{L}^G$. Кроме того, многочлен $\phi(x)$ неприводим над $\mathbb{L}^G$ (докажите!). Значит, $(\mathbb{L} : \mathbb{L}^G) = m \leq |G| \leq (\mathbb{L} : \mathbb{L}^G) \quad \Rightarrow \quad m = |G| = (\mathbb{L} : \mathbb{L}^G)$. Условие $\mathbb{L}^G = \mathbb{K}$ означает, что группа G содержит *все* автоморфизмы поля G над $K \ \Rightarrow \ G = \mathrm{Aut}(\mathbb{L} : \mathbb{K})$. □

11.10 Теория Галуа

Основная теорема теории Галуа. *Пусть $\mathbb{L}$ — поле разложение какого-то многочлена над полем $\mathbb{K}$ и $G = \mathrm{Aut}(\mathbb{L} : \mathbb{K})$. Пусть H — произвольная подгруппа группы G. Тогда*

- *отображение $H \to \mathbb{L}^H$ является взаимно-однозначным отображением множества подгрупп группы G на множество промежуточных полей $\mathbb{P}$ между полем $\mathbb{K}$ и полем $\mathbb{L}$,*
- *обратное отображение определяется правилом $\mathbb{P} \to \mathrm{Aut}(\mathbb{L} : \mathbb{P})$,*
- *для того чтобы промежуточное поле $\mathbb{P}$ было полем разложения некоторого многочлена над $\mathbb{K}$, необходимо и достаточно, чтобы подгруппа $H = \mathrm{Aut}(\mathbb{L} : \mathbb{P})$ группы G была нормальной, и в этом случае группа $\mathrm{Aut}(\mathbb{P} : \mathbb{K})$ изоморфна фактор-группе G/H.*

Доказательство. Из теоремы о неподвижном поле следует, что если $\mathbb{P} = \mathbb{L}^H$, то $H = \mathrm{Aut}(\mathbb{L} : \mathbb{P})$. Поле $\mathbb{L}$ является полем разложения какого-то многочлена над $\mathbb{K}$, и этот многочлен можно рассматривать как многочлен над любым промежуточным полем, а значит, $\mathbb{L}$ есть поле разложения этого многочлена над любым промежуточным полем. Поэтому $H = \mathrm{Aut}(\mathbb{L} : \mathbb{P}) \to \mathbb{P}$.

Теперь предположим, что $\mathbb{P}$ есть поле разложения какого-то многочлена над $\mathbb{K}$. Тогда имеется примитивный элемент θ, для которого $\mathbb{L} = \mathbb{K}(\theta)$, и все корни минимального многочлена числа θ принадлежат полю $\mathbb{P}$. Отсюда следует, что любой автоморфизм группы G переводит число θ в какой-то корень того же самого многочлена, и значит, поле $\mathbb{P}$ *инвариантно* относительно действий группы G — в том смысле, что $g(a) \in \mathbb{P}$ для любого автоморфизма $g \in G$ и любого числа $a \in \mathbb{P}$. Пусть $H = \mathrm{Aut}(\mathbb{L} : \mathbb{P})$. Тогда

$$a \in \mathbb{P} \quad \Leftrightarrow \quad h(a) = a \ \ \forall\, h \in H \quad \Rightarrow \quad g(a) \in \mathbb{P} \ \ \forall\, g \in G \quad \Leftrightarrow$$

$$h(g(a)) = g(a) \ \ \forall\, h \in H, \ \ \forall\, g \in G \quad \Leftrightarrow \quad (g^{-1}hg)(a) = a \ \ \forall\, g \in G, \ \ \forall\, h \in H, \ \ \forall\, a \in \mathbb{P}.$$

Таким образом, $g^{-1}hg \in H \ \ \forall\, h \in H, \ \ \forall\, g \in G$ — именно так определяется нормальность подгруппы H в группе G.

Пусть известно, что подгруппа $H = \mathrm{Aut}(\mathbb{L} : \mathbb{P})$ является нормальной в G. Тогда $(g^{-1}hg)(a) = a$ $\forall\, h \in H, \ \ \forall\, g \in G, \ \ \forall\, a \in \mathbb{P}$. Отсюда

$$h(g(a)) = g(a) \ \ \forall\, h \in H \quad \Rightarrow \quad g(a) \in \mathbb{P}.$$

Таким образом, каждый автоморфизм $g \in G$ при действии на числа $a \in \mathbb{P}$ переводит их в числа из $\mathbb{P}$, порождая тем самым автоморфизм поля $\mathbb{P}$ над $\mathbb{K}$. При этом все автоморфизмы вида hg, где $h \in H$, порождают один и тот же автоморфизм поля $\mathbb{P}$ над $\mathbb{K}$. Автоморфизмы $g_1, g_2 \in G$ оставляют разные *следы* (те же отображения с меньшей областью определения) на $\mathbb{P}$ тогда и только тогда, когда $g_1 g_2^{-1} \notin H$. Следовательно, число автоморфизмов $\mathbb{P}$ над $\mathbb{K}$ равно числу различных смежных классов группы G по нормальной подгруппе H. Согласно теореме Лагранжа, это число равно $|G|/|H| = (\mathbb{L} : \mathbb{K})/(\mathbb{L} : \mathbb{P}) = (\mathbb{P} : \mathbb{K})$. Поскольку число автоморфизмов поля $\mathbb{P}$ над $\mathbb{K}$ совпадает со степенью расширения, $\mathbb{P}$ является полем разложения некоторого многочлена над $\mathbb{K}$. Искомый изоморфизм $G/H \to \mathrm{Aut}(\mathbb{P} : \mathbb{K})$ строится следующим образом: смежному классу gH ставится в соответствие след отображения g на множестве элементов поля $\mathbb{P}$. В силу установленной нами инвариантности образ любого элемента из $\mathbb{P}$ остается в том же множестве $\mathbb{P}$, а след не зависит от выбора представителя в смежном классе gH. □

11.11 Два типа радикальных расширений

В связи с разрешимостью алгебраических уравнений в радикалах нас интересуют радикальные расширения вида $\mathbb{P}(\theta) \supset \mathbb{P}$, где $\theta^n \in \mathbb{P}$, и ясно, что от произвольных радикальных расширений мы всегда можем перейти к цепочке радикальных расширений, для которых n будет простым числом.

К первому типу таких расширений отнесем случай $\theta^n = 1$. Не ограничивая общности мы можем считать, что $\theta = \varepsilon = \mathrm{e}^{(2\pi/n)\mathbf{i}}$. Тогда поле $\mathbb{P}(\varepsilon)$ будет полем разложения многочлена $x^n - 1$. Как устроена группа Галуа в этом случае? Мы уже знаем, что минимальным многочленом для ε над полем $\mathbb{Q}$ в случае простого n будет многочлен деления круга $\Phi_n(x) = x^{n-1} + \ldots + x + 1$. Нетрудно понять, что группа $\mathrm{Aut}(\mathbb{Q}(\varepsilon) : \mathbb{Q})$ будет изоморфна мультипликативной группе $\mathbb{Z}_n^*$ поля $\mathbb{Z}_n$. Действительно, автоморфизмы f, g поля $\mathbb{Q}(\varepsilon)$ над $\mathbb{Q}$ определяются своими значениеми $f(\varepsilon) = \varepsilon^{k_f}$, $g(\varepsilon) = \varepsilon^{k_g}$, а их композиция fg задается значением $(fg)(\varepsilon) = f(\varepsilon^{k_g}) = \varepsilon^{k_f k_g} = \varepsilon^{k_{fg}}$, где $k_{fg} = k_f k_g$. Таким образом, отображение $f \to [k_f]_n$ будет изоморфизмом группы $\mathrm{Aut}(\mathbb{Q}(\varepsilon) : \mathbb{Q})$ на группу $\mathbb{Z}_n^*$. Пусть теперь $\mathbb{P}$ — произвольное поле, и пусть F — автоморфизм поля $\mathbb{P}(\varepsilon)$. Конечно, F определяется своим значением $F(\varepsilon) = \varepsilon^{k_F}$. Рассмотрим отображение $F \to f$, где f — автоморфизм поля $\mathbb{Q}(\varepsilon)$ над $\mathbb{Q}$, для которого $k_f = k_F$. Это отображение будет изоморфизмом группы $\mathrm{Aut}(\mathbb{P}(\varepsilon) : \mathbb{P})$ на какую-то подгруппу группы $\mathrm{Aut}(\mathbb{Q}(\varepsilon) : \mathbb{Q})$ (проверьте!). Следовательно, группа Галуа расширения $\mathbb{P}(\varepsilon)$ над $\mathbb{P}$ изоморфна подгруппе циклической группы и поэтому является циклической.

Ко второму типу радикальных расширений отнесем случай $\theta^n = a \neq 1$. Если $a = b^n$ для какого-то числа $b \in \mathbb{P}$, то $x^n - a \mathrel{\vdots} x - b$ (проверьте!). Если $a \neq b^n$ при любом $b \in \mathbb{P}$, то для простого n многочлен $x^n - a$ оказывается неприводимым над $\mathbb{P}$. От противного, пусть имеется нетривиальное разложение $x^n - a = (a_0 + \ldots + a_k x^k)(b_0 + \ldots + b_l x^l)$. Тогда $a_0 = \theta^k \varepsilon^s$, $b_0 = \theta^l \varepsilon^t$, $\theta^n = a$, s и t — целые числа (почему?). Поскольку $k + l = n$, из простоты n следует взаимная простота чисел k и l. Поэтому для некоторых целых u и v имеет место равенство $ku + lv = 1$. Отсюда $a_0^u b_0^v = \theta \varepsilon^{su+tv}$ и для числа $b := a_0^u b_0^v \in \mathbb{P}$ находим $b^n = a$.

В случае радикальных расшиерений второго типа мы будем считать, что поле $\mathbb{P}$ уже содержит число $\varepsilon = \mathrm{e}^{(2\pi/2)\mathbf{i}}$ (любую заданную цепочку можно перестроить, чтобы получить это свойство). Если $a \neq b^n$ для любого $b \in \mathbb{P}$, то многочлен неприводим над $\mathbb{P}$ и поэтому является минимальным многочленом числа θ. Условие $\varepsilon \in \mathbb{P}$ позволяет утверждать, что поле $\mathbb{P}(\theta)$ будет его полем разложения. Пусть автоморфизмы f и g задаются своими значениями $f(\theta) = \varepsilon^{k_f}\theta$ и $g(\theta) = \varepsilon^{k_g}\theta$. Тогда их композиция определяется значением $(fg)(\theta) = f(\varepsilon^{k_g}\theta) = \varepsilon^{k_g} f(\theta) = \varepsilon^{k_f + k_g}\theta$. Таким образом, для второго типа радикального расширения группа Галуа изоморфна группе вычетов по модулю n относительно сложения и также является циклической.

Теперь мы понимаем, что вопрос о разрешимости в радикалах сводится к вопросу о существовании цепочки расширений с циклическими группами автоморфизмов каждого поля цепочки над предшествующим ему полем. Теория Галуа сводит это вопрос целиком к изучению свойств группы автоморфизмов поля разложения интересующего нас многочлена.

11.12 Разрешимость алгебраических уравнений

Пусть $\mathbb{M}$ — поле разложения некоторого многочлена над полем $\mathbb{K}$ и пусть G — группа Галуа, состоящая из всех автоморфизмов поля $\mathbb{M}$ над $\mathbb{K}$. Тогда, согласно теории Галуа, цепочке радикальных расширений

$$\mathbb{K} = \mathbb{K}_0 \subset \mathbb{K}_1 \subset \ldots \mathbb{K}_{s-1} \subset \mathbb{K}_s = \mathbb{M}$$

соответствует ряд подгрупп

$$G = G_0 \supset G_1 \supset \ldots \supset G_{s-1} \supset G_s = \{e\},$$

в котором последняя подгруппа состоит из одного единичного элемента и каждая подгруппа G_i является нормальной подгруппой в подгруппе G_{i-1} с циклической фактор-группой G_{i-1}/G_i. Такой ряд подгрупп группы G будем называть ее *циклическим рядом*.

При изучении вопроса о разрешимости в радикалах уравнения $f(x) = 0$ поле $\mathbb{M}$ должно содержать поле разложения многочлена $f(x)$. При этом оно само должно быть полем разложения некоторого многочлена, который, вообще говоря, отличен от $f(x)$. Можно, однако, доказать, что наличие цепочки

радикальных расширений, в которой последнее поле $\mathbb{M}$ *совпадает* с полем разложения многочлена $f(x)$, является необходимым и достаточным условием для разрешимости в радикалах. Чтобы не отвлекаться от демонстрации *всей схемы* рассуждений, мы примем этот факт без доказательства.

То, что $\mathbb{M}$ есть поле разложения именно интересующего нас многочлена, важно потому, что для конкретного многочлена мы можем *найти* группу Галуа расширения $\mathbb{M} \supset \mathbb{K}$. Она же при этом наывается просто *группой Галуа данного многочлена* или *данного алгебраичекого уравнения* и может рассматриваться как некоторая подгруппа группы все подстановок его корней. Если n — степень многочлена, то G будет подгруппой симметрической группы S_n.

Примеры неразрешимых уравнений n-й степени строятся таким образом, чтобы их группа Галуа оказалась равной S_n. При $n \geq 5$ группа S_n не имеет циклических рядов по следующей причине.

Утверждение 1. *При $n \geq 5$ единственной нетривиальной нормальной подгруппой в симметрической группе S_n с абелевой фактор-группой является знакопеременная группа A_n, а группа A_n нетривиальных нормальных подгрупп с абелевой фактор-группой не имеет.*

Доказательство. По условию любые смежные классы по подгруппе H коммутируют: $H(ab) = H(ba) \Rightarrow aba^{-1}b^{-1} \in H \quad \forall\, a, b \in A_n$. Возьмем два тройных цикла $a = (ijk)$ и $b = (ijm)$. 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: $(ij)(kl)\,(kl)(jm) = (ijm)$. Тройные циклы и произведения пар независимых транспозиций порождают все четные подстановки. □

Замечание. В действительности группа S_n при $n \geq 5$ имеет вообще только одну нетривиальную нормальную подгруппу — это знакопеременная группа A_n, а группа A_n при этом не имеет нетривиальных нормальных подгрупп.

Подгруппа G группы S_n называется *транзитивной*, если для любых номеров $1 \leq i, j \leq n$ существует подстановка $\sigma \in G$ такая, что $\sigma(i) = j$.

Утверждение 2. *Группа Галуа неприводимого многочлена степени n изоморфна транзитивной подгруппе группы S_n.*

Доказательство. Все корни неприводимого многочлена над $\mathbb{K}$ являются простыми и принадлежат полю разложения $\mathbb{M}$ для этого многочлена. Мы уже знаем, что существует автоморфизм поля $\mathbb{M}$ над $\mathbb{K}$, переводящий любой заданный корень этого многочлена в любой другой его корень. □

Утверждение 3. *Любая транзитивная подгруппа G группы S_n, содержащая хотя бы одну транспозицию, при простом n совпадает с S_n.*

Доказательство. 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. □

Пусть $K = \mathbb{Q}$. Многочлен $f(x) = x^5 - 4x + 2$ неприводим над $\mathbb{Q}$ и имеет три различных вещественных корня и два комплексно сопряженных корня ζ и $\overline{\zeta}$ (докажите!). В данном случае группа Галуа содержит транспозицию — это автоморфизм, переводящий ζ в $\overline{\zeta}$ и оставляющий на месте вещественные корни. Будучи транзитивной, группа Галуа данного многочлена совпадает с S_5.

11.13 Циклический ряд для примарной группы

Теория Галуа сводит вопрос о возможности построения правильного многоугольника с помощью циркуля и линейки к вопросу о том, обладает ли группа G порядка 2^k циклическим рядом

$$G = G_0 \supset G_1 \ldots \supset G_{k-1} \supset G_k = \{e\},$$

в котором каждая подгруппа G_i является нормальной в группе G_{i-1} с циклической фактор-группой G_{i-1}/G_i порядка 2. Конечные группы порядка p^k, где p — простое число, называются *примарными*. Существование циклического ряда можно вывести, опираясь на следующую теорему (попробуйте это сделать!).

Теорема о примарной группе. *Примарная группа порядка p^k 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 $G = K_1 \cup \ldots \cup K_m$.

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 элементом 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 $H(a)$. $\square$

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 K_i сопряженных элементов. Согласно лемме 1, число элементов в классе K_i имеет вид p^{k_i} (делитель числа p^k). Отсюда ясно, что число классов K_i, состоящих из одного элемента, должно делиться на $p \quad \Rightarrow$ 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 о примарной группе. 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. $\square$

11.14 Приращение аргумента и непрерывные деформации

Исследование свойств корней многочленов можно сделать менее абстракным и даже весьма наглядным, рассматривая непрерывные кривые на комплексной плоскости и их непрерывные деформации.

Непрерывная кривая на комплексной плоскости определяется непрерывной комплекснозначной функцией $z(t)$ от вещественного параметра $t_1 \leq t \leq t_2$. Непрерывность означает, что вещественнозначные функции $\mathrm{Re}(z(t))$ и $\mathrm{Im}(z(t))$ являются непрерывными. Если $z(t_1) = z(t_2)$, то кривая называется *замкнутой кривой*, *петлей* или *контуром*.

Довольно часто нас интересует семейство кривых $z_r(t) = z(t, r)$, зависящих от вещественного параметра $r_1 \leq r \leq r_2$. Предполагается, что функция $z(t, r)$ непрерывна по t и r. В таких случаях кривые данного семейства можно рассматривать как *непрерывные деформации* любой фиксированной кривой данного семейства.

Будем рассматривать только кривые, не проходящие через нуль. Для них каждая точка имеет аргумент $\arg(z(t))$, определенный с точностью до прибавления числа вида $2\pi k$, где k — целое число. Мы примем без доказательства следующие утверждения.

Утверждение 1. *Пусть $z(t)$, $t_1 \leq t \leq t_2$, — непрерывная кривая, не проходящая через нуль. Тогда существует непрерывная функция $\phi(t)$ такая, что $\phi(t) = \arg(z(t))$, и при этом для любой непрерывной функции с таким свойством приращение аргумента $\phi(t_2) - \phi(t_1)$ одинаково.*

Величину $\phi(t_2) - \phi(t_1)$ будем называть *приращением аргумента на кривой* $z(t)$.

Утверждение 2. *Пусть имеется семейство кривых $z_r(t)$, $t_1 \leq t \leq t_2$, где функция $z_r(t) = z(t, r)$ непрерывна при $t_1 \leq t \leq t_2$, $r_1 \leq r \leq r_2$. Пусть ни одна из кривых не проходит через нуль и пусть*

$\phi_r(t)$ — функция, определенная в утверждении 1 для каждой отдельной кривой данного семейства. Тогда приращение аргумента $\delta(r) = \phi_r(t_2) - \phi_r(t_1)$ является непрерывной функцией при $r_1 \le r \le r_2$.

Утверждение 3. *Пусть $z(t)$ — непрерывная петля, не проходящая через нуль. Тогда для любой непрерывной петли $w(t)$ такой, что $|w(t)| < |z(t)|$ для всех $t_1 \le t \le t_2$, петля $z(t) + w(t)$ не проходит через нуль и имеет такое же приращение аргумента, как исходная петля $z(t)$.*

Используя эти утверждения, можно получить еще одно, не очень "алгебраическое", но, тем не менее, лапидарное и весьма элегантное

Доказательство основной теоремы алгебры. Пусть имеется многочлен

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$$

с комплексными коэффициентами. Если $a_0 = 0$, то существование корня очевидно. Поэтому полагаем, что $a_0 \ne 0$. Рассмотрим семейство замкнутых кривых

$$z_r(t) = f(x_r(t)), \quad \text{где} \quad x_r(t) = r\mathrm{e}^{\mathbf{i}t}, \quad 0 \le t \le 2\pi, \quad 0 < r_1 \le r \le r_2.$$

При достаточно малом r_1 для $x = x_{r_1}(t)$ находим $|a_0| > |x^n + a_{n-1}x^{n-1} + \ldots + a_1x|$, поэтому кривая $z_{r_1}(t)$ и кривая $z(t) = a_0$ имеют одно и то же приращение аргумента, которое, очевидно, равно нулю. При достаточно большом r_2 для $x = x_{r_2}(t)$ находим $|x^n| > |a_{n-1}x^{n-1} + \ldots + a_1x + a_0|$, поэтому кривая $z_{r_2}(t)$ и кривая $z(t) = x^n$ имеют одно и то же приращение аргумента, которое, как нетрудно подсчитать, равно $2\pi n$.

Теперь предположим, что ни одна из кривых $z_r(t)$ не проходит через нуль. Тогда, согласно утверждению 2, приращение аргумента $\delta(r)$ является непрерывной функцией на отрезке $r_1 \le r \le r_2$. Однако очевидно, что приращение аргумента на замкнутом контуре может принимать лишь дискретные значения, кратные 2π. Таким образом, функция $\delta(r)$ должна быть константой, а мы только что нашли, что $\delta(r_1) = 0$ и $\delta(r_2) = 2\pi n$. $\square$

11.15 Визуализация теоремы Абеля

У нас уже есть теорема о непрерывной зависимости корней приведенного многочлена от его коэффициентов. Используя ее и некоторые факты математического анализа, можно доказать, что если коэффициенты многочлена $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$ непрерывно зависят от параметра $t_1 \le t \le t_2$, то имеются непрерывные кривые $z_1(t), \ldots, z_n(t)$ такие, что

$$f(x) = (x - z_1(t)) \ldots (x - z_n(t)).$$

Будем считать, что при $t = t_1$ все корни простые. Тогда каждая кривая $z_i(t)$ ассоциируется с одним и только одним корнем $z_i(t_1)$ — при условии, однако, что *разные кривые не попадают в одну точку при одном и том же значении* $t_1 \le t < t_2$. Пусть кривые $a_i(t)$ являются петлями. Тогда значения $z_1(t_2), \ldots, z_n(t_2)$ будут, вообще говоря, *перестановкой* исходных значений $z_1(t_1), \ldots, z_n(t_1)$. Движение корней по своим кривым и то, как получаются перестановки, можно увидеть, рисуя эти кривые на комплексной плоскости средствами компьютерной графики.

В связи с теоремой Абеля достаточно изучить кривые для корней уравнения[2]

$$x^5 - x + a(t) = 0,$$

в котором комплексное число $a(t)$ при изменении параметра t пробегает петли, начинающиеся в одной и той же точке. На множестве таких петель естественным образом определяется операция умножения петель — в результате появляется петля, составленная из двух петель. Петля, для которой $a(t) = \mathrm{const}$, играет роль единицы. Обратным элементом будет петля с противоположным направления обхода. Множество петель превращается в группу, которая гомоморфно отображается на некоторую подгруппу, состоящую из перестановок корней. Оказывается, для данного уравнения существуют петли, реализующие *любую перестановку* корней. В конечном счете именно это становится причиной неразрешимости данного уравнения в радикалах.

[2]Табачников С.Л., Фукс Д.Б., Математический дивертисмент, М.: МЦНМО, 2011.

Алгебра и геометрия (1 поток)

Лекция 12

12.1 Квадратичное многообразие

Под *алгебраическим многообразием* понимается геометрическое место точек, координаты которых удовлетворяют некоторой системе полиномиальных уравнений. Если уравнение только одно и определяется ненулевым многочленом, то многообразие обычно называют *гиперповерхностью*.

Попробуем понять, как устроена гиперповерхность, представляющая собой *квадратичное многообразие*, т. е. множество точек $x = (x_1, \ldots, x_n)$, удовлетворяющих уравнению $f(x) = f(x_1, \ldots, x_n) = 0$ в случае, когда $f(x)$ является вещественным квадратичным многочленом (многочленом второй степени). Такой многочлен принято записывать в виде

$$f(x_1, \ldots . x_n) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j \;\;+\;\; 2\sum_{i=1}^{n} b_i x_i \;+\; c \;=\; x^\top A x + 2x^\top b + c,$$

где $A = [a_{ij}]$ — матрица порядка n, а x и b — вектор-столбцы с элементами $x_1, \ldots, x_n$ и $b_1, \ldots, b_n$. Легко проверить, что после замены $a_{ij} = a_{ji} := (a_{ij} + a_{ji})/2$ значение $f(x)$ не изменяется. Поэтому в дальнейшем мы будем считать, что матрица A симметрична. Полагаем также, что $A \neq 0$, иначе мы имели бы дело с уже хорошо изученным нами множеством решений линейного алгебраического уравнения.

Функция $F(x) = x^\top A x$ называется *квадратичной формой* или *квадратичной частью многочлена* $f(x)$, а матрица $A = A^\top$ называется *матрицей квадратичной формы* или *основной матрицей* квадратичного многочлена $f(x)$. В составе $f(x)$ есть также *линейная часть* $2\,x^\top b$ и *свободный член* c. Коэффициент 2 в линейной части позволяет записать уравнение $f(x) = 0$ следующим образом (проверьте!):

$$\begin{bmatrix} x^\top & 1 \end{bmatrix} \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = 0.$$

Матрицу порядка $n+1$ в этом уравнении будем называть *расширенной матрицей* квадратичного многочлена $f(x)$.

Заметим также, что если скалярное произведение векторов x и y определяется формулой $(x, y) = x_1 y_1 + \ldots + x_n y_n = y^\top x$ (такое скалярное произведение принято называть *естественным*), то квадратичное уравнение приобретает вид

$$(Ax, x) + 2(b, x) + c = 0.$$

12.2 Замена переменных и конгруэнтные матрицы

При изучении квадратичных многообразий вполне естественна идея упростить вид уравнения с помощью замены переменных или перехода к другой системе координат.

Положим $x = Py$, где P — невырожденная матрица. Тогда уравнение

$$(Ax, x) + 2(b, x) + c = 0$$

преобразуется к виду $(APy, Py) + 2(b, Py) + c = 0$ и в итоге

$$(APy, Py) = (Py)^\top APy = y^\top (P^\top AP)y = ((P^\top AP)y, y),$$
$$(b, Py) = (Py)^\top b = y^\top (P^\top b) = (P^\top b, y) \quad \Rightarrow$$
$$(\Lambda y, y) + 2(g, y) + c = 0, \quad \text{где} \quad \Lambda = P^\top AP, \quad g = P^\top b.$$

В случае невырожденности вещественной матрицы P матрица $\Lambda = P^\top AP$ называется *конгруэнтной* матрице A, а преобразование $A \to P^\top AP$ — *преобразованием конгруэнтности* или просто *конгруэнцией*. На множестве вещественных матриц порядка n конгруэнция определяет бинарное отношение со свойствами рефлексивности, симметричности и транзитивности, т. е. отношение эквивалентности.

12.3 Конгруэнтная диагонализация

Теорема о конгруэнтной диагонализации. *Любая вещественная симметричная матрица конгруэнтна диагональной матрице.*

Доказательство. Пусть A — вещественная симметричная матрица порядка n. Проведем индукцию по n. При $n = 1$ утверждение очевидно. При $n \geq 2$ рассмотрим два случая.

1. Пусть хотя бы один элемент главной диагонали матрицы A отличен от нуля. Если $a_{ii} \neq 0$, то после перестановки первой и i-й строк и первого и i-го столбца мы получаем матрицу, в которой ненулевой элемент находится в позиции $(1, 1)$. Ясно, что такое преобразование является конгруэнцией. Теперь можно считать, что $a_{11} \neq 0$. Тогда

$$\begin{bmatrix} 1 & & & \\ -a_{21}/a_{11} & 1 & & \\ \vdots & & \ddots & \\ -a_{n1}/a_{11} & 0 & \dots & 1 \end{bmatrix} A \begin{bmatrix} 1 & -a_{21}/a_{11} & \dots & -a_{n1}/a_{11} \\ & 1 & & 0 \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \widetilde{A} & \\ 0 & & & \end{bmatrix}.$$

Матрица $\widetilde{A}$ является вещественной симметричной матрицей порядка $n - 1$, и к ней можно применить индуктивное предположение.

2. Пусть все элементы главной диагонали равны нулю. Если $A = 0$, то доказывать нечего. В противном случае A имеет ненулевой элемент, скажем a_{ij} при $i < j$. Заметим, что

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & a_{ij} \\ a_{ij} & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2a_{ij} & 0 \\ 0 & -2a_{ij} \end{bmatrix}.$$

Пусть P отличается от единичной матрицы лишь в позициях $(i, i), (i, j), (j, i), (j, j)$ и пусть в этих позициях размещается 2×2-матрица вида $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. Тогда матрица $P^\top AP$ получает главную дигональ с ненулевыми элементами в позициях (i, i) и (j, j). Таким образом, случай 2 сводится к случаю 1. $\square$

12.4 Аффинные инварианты

Аффинными инвариантами квадратичного уравнения называются величины или свойства, сохраняющиеся при переходе к любой другой аффинной системе координат.

Теорема об аффинных инвариантах. *Ранги основной и расширенной матриц квадратичного многочлена и знаки определителей этих матриц являются аффинными инвариантами.*

Доказательство. При замене $x = Py$ новая матрица квадратичной части конгрунтна старой матрице квадратичной части, а новая расширенная матрица конгруэнтна старой расширенной матрице:

$$\begin{bmatrix} P^\top AP & P^\top b \\ b^\top P & c \end{bmatrix} = \begin{bmatrix} P^\top & \\ & 1 \end{bmatrix} \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \begin{bmatrix} P & \\ & 1 \end{bmatrix}.$$

При замене $x = y + r$ получается уравнение

$$(A(y+r), y+r) + 2(b, y+r) + c = (Ay, y) + 2(Ar+b, y) + (Ar, r) + 2(b, r) + c = 0.$$

Мы видим, что матрица квадратичной части не меняется, а новая расширенная матрица оказывается конгруэнтной старой расширенной матрице:

$$\begin{bmatrix} A & Ar+b \\ (Ar+b)^\top & r^\top Ar + 2r^\top b + c \end{bmatrix} = \begin{bmatrix} I & \\ r^\top & 1 \end{bmatrix} \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \begin{bmatrix} I & r \\ & 1 \end{bmatrix}.$$

Остается заметить, что при конгруэнции ранг и знак определителя сохраняются. □

12.5 Приведенные уравнения

Теорема о приведенных уравнениях. *Любое квадратичное уравнение в некоторой аффинной системе координат приводится к уравнению одного из следующих видов*:

$$(1) \quad \lambda_1 x_1^2 + \ldots + \lambda_k x_k^2 + c = 0 \qquad \textit{либо} \qquad (2) \quad \lambda_1 x_1^2 + \ldots + \lambda_k x_k^2 + 2\beta x_{k+1} = 0,$$

где все коэффициенты при переменных отличны от нуля. Тип уравнения, определяемый его видом и значением k, является аффинным инвариантом.

Доказательство. Замена $x := Px$ позволяет перейти к уравнению, в котором матрица квадратичной части диагональная: $(\Lambda x, x) + 2(g, x) + c = \sum\limits_{i=1}^{n} (\lambda_i^2 x_i + 2 g_i x_i) + c = 0$. Пусть $\lambda_1, \ldots, \lambda_k \neq 0$ и $\lambda_{k+1} = \ldots = \lambda_n = 0$. Тогда естестественным образом выделяются полные квадраты:

$$\sum_{i=1}^{k} (\lambda_i x_i^2 + 2 g_i x_i) + \sum_{i=k+1}^{n} 2 g_i x_i + c \;=\; \sum_{i=1}^{k} \lambda_i (x_i + g_i/\lambda_i)^2 \;+\; \sum_{i=k+1}^{n} 2 g_i x_i \;+\; c - \sum_{i=1}^{k} g_i^2/\lambda_i.$$

Положим $x := x - r$, где $r_i = r_i/\lambda_i$ при $1 \leq i \leq k$ и $r_i = 0$ при $k+1 \leq i \leq n$. Пусть векторы u и v составлены из компонент $x_1, \ldots, x_k$ и $x_{k+1}, \ldots, x_n$ вектора x, диагональ матрицы Λ_k содержит ненулевые значения $\lambda_1, \ldots, \lambda_k$, а вектор h собран из чисел $g_{k+1}, \ldots, g_n$. Тогда новое уравнение приобретает вид $(\Lambda_k u, u) + 2(h, v) + c = 0$, где $c := c - \sum\limits_{i=1}^{k} g_i^2/\lambda_i$.

Если $h = 0$, то мы имеем уравнение вида (1). Если $h \neq 0$, то для любого значения $\beta \neq 0$ можно найти невырожденную матрицу Q порядка $n-k$, для которой вектор Qh будет отличаться от нулевого вектора только первой координатой, равной β. Еще одна замена $x := \Pi x$ с матрицей $\Pi = \begin{bmatrix} I & \\ & Q^\top \end{bmatrix}$ приводит к уравнению $\lambda_1 x_1^2 + \ldots + \lambda_k x_k^2 + 2\beta x_{k+1} + c = 0$, и, наконец, замена $x_{k+1} := x_{k+1} - c/(2\beta)$ дает уравнение вида (2).

Остается прояснить, почему с помощью перехода к другой системе координат уравнение одного вида не может преобразоваться в уравнение другого вида. Если $k = n$, то ранг матрицы квадратичной части равен n и является аффинным инвариантом, а в случае уравнения вида (2) ранг матрицы квадратичной части равен $k < n$. Если $k < n$, то ранг расширенной матрицы уравнения вида (1) не больше $k+1$, в то время как ранг расширенной матрицы уравнения вида (2) равен $k+2$. $\square$

Уравнения (1) и (2) часто называют *приведенными уравнениями* квадратичной гиперповерхности. В случае (1) для k имеется n возможных значений, в случае (2) различных значений $n-1$. Таким образом, в n-мерном пространстве выделяются $2n-1$ типов приведенных уравнений.

12.6 Параллельные хорды и центры симметрии

Пусть квадратичное многообразие задано уравнением $(Ax, x) + 2(b, x) + c = 0$, и предположим, что оно непустое. Под *хордой* многообразия понимается отрезок, соединяющий пару его точек. Рассмотрим семейство хорд, параллельных вектору $h \neq 0$. Дополнительно потребуем, чтобы выполнялось неравенство $(Ah, h) \neq 0$. В таких случаях принято говорить, что h имеет *неасимптотическое направление* по отношению к данному многообразию.

Пусть хорда соединяет точки x и $x+th$. Тогда $(Ax,\, x) + 2(b,\, x) + c = 0$ и, кроме того, $(A(x+th),\, x+th) + 2(b,\, x+th) + c = 0 \;\Rightarrow\; t^2(Ah, h) + 2t(Ax+b, h) = 0 \;\Rightarrow\; t = 0$ или $t = t' := -2(Ax+b, h)/(Ah, h)$. Рассмотрим середину хорды

$$z = x + (t'/2)h = x - ((Ax+b, h)/(Ah, h))h \quad \Rightarrow$$

$$(Ah, z) = (Ah, x) - (Ax+b, h) = -(b, h).$$

Положим $A_0 = (b, h)$, и обозначим координаты заведомо ненулевого векторы Ah через $A_1, \ldots, A_n$. Тогда координаты середин всех хорд, параллельных неасимптотическому вектору h, удовлетворяют уравнению

$$(Az + b, h) = 0 \quad \Leftrightarrow \quad A_0 + A_1 z_1 + \ldots + A_n z_n = 0,$$

которое, как мы знаем, является уравнением некоторой гиперплоскости в n-мерном пространстве. Таким образом, мы доказали следующее интересное утверждение.

Теорема о серединах хорд. *Пусть квадратичное многообразие задано уравнением $(Ax, x) + 2(b, x) + c = 0$ и непусто. Тогда для любого вектора h с неасимптотическим направлением середины всех хорд, параллельных вектору h, лежат на одной гиперплоскости.*

Непустое множество M называется *центрально симметричным*, если существует точка z такая, что для любой точки x из множества M точка $x + 2(z-x)$ также принадлежит M. Такая точка z называется *центром симметрии* множества M и, вообще говоря, не обязана принадлежать M.

Теорема о центрах симметрии. *Пусть квадратичное многообразие задано уравнением* $(Ax, x) + 2(b, x) + c = 0$ *и непусто. Тогда оно является центрально симметричным в том и только том случае, когда система линейных алгебраических уравнений* $Az + b = 0$ *совместна. При этом множество центров симметрии совпадает с множеством решений данной системы.*

Доказательство. Понятно, что центр симметрии z, если он есть, обязан быть серединой любой хорды. Следовательно, $(Az + b, h) = 0$ для любого вектора h с неасимптотическим направлением. Теперь заметим, что ненулевая симметричная матрица A обязательно обладает хотя бы одним неасимптотическим вектором. В самом деле, она конгруэнтна диагональной матрице $\Lambda = P^\top AP$ и условие $(Ah, h) \neq 0$ равносильно условию $(\Lambda q, q) \neq 0$, где $q = Ph$. Пусть λ_i — один из ненулевых элементов диагонали матрицы Λ. Тогда вектор q с единицей в i-й позиции и нулями в остальных будет неасимпотическим для Λ, и значит, вектор $h = P^{-1}q$ будет неасимптотическим для A. Возьмем неасимптотический вектор v_1 и построим линейно независимую систему $v_1, v_2, \ldots, v_n$. Тогда при любом $\varepsilon \neq 0$ векторы

$$h_1 = v_1, \quad h_2 = v_1 + \varepsilon v_2, \quad \ldots, \quad h_n = v_n + \varepsilon v_1$$

будут линейно независимы (почему?) и в то же время при достаточно малом $\varepsilon > 0$ все они являются неасимптотическими векторами для матрицы A (почему?). Таким образом, $(Az+b, h) = 0$ при $h = h_i$ для всех $1 \leq i \leq n$, а следовательно, и для любого вектора h. В том числе для $h = Az+b \;\Rightarrow\; Az+b = 0$. Теперь предположим, что z удовлетворяет уравнению $Az + b = 0$, и пусть x — произвольная точка квадратичного многообразия. Покажем, что точка $x + 2(z - x) = 2z - x$ также принадлежит многообразию:

$$(A(2z-x), 2z-x)+2(b, 2z-x)+c = ((Ax, x)+2(b, x)+c)+4(Az+b, z)-4(Az+b, x) = 0. \;\square$$

12.7 Неособые точки и касательные плоскости

Пусть $f(x)$ — ненулевой вещественный многочлен от координат точки $x = (x_1, \ldots, x_n)$. Для любой точки $x_0 = (x_{01}, \ldots, x_{0n})$ его можно записать как многочлен от координат вектора $\delta = x - x_0$ в виде

$$f(x) = f(x_0 + \delta) = f(x_0) + (h, \delta) + R(\delta),$$

где $R(\delta)$ — многочлен, в составе которого нет членов степени меньше двух. Вектор h зависит от x_0 и однозначно определяется точкой x_0.

Множество точек x, удовлетворяющих полиномиальному уравнению $f(x) = 0$ с ненулевым многочленом $f(x)$, принято называть *алгебраической гиперповерхностью*. Если $f(x_0) = 0$ и $h \neq 0$, то точка x_0 данной гиперповерхности называется *неособой* или *регулярной*. Гиперплоскость, заданная уравнением $(h, x - x_0) = 0$, называется *касательной плоскостью* к данной гиперповерхности в точке x_0.

В случае квадратичного многочлена находим

$$f(x) = (Ax, x) + 2(b, x) + c = f(x_0) + 2(Ax_0 + b, x - x_0) + (A(x - x_0), x - x_0).$$

Таким образом, для квадратичного многообразия $(Ax, x) + 2(b, x) + c = 0$ *неособая* или *регулярная* точка определяется условием $Ax_0 + b \neq 0$, а уравнение касательной

плоскости имеет вид $(Ax_0+b, x-x_0)=0$. Обратим внимание на то, что свойство точки быть неособой является аффинным инвариантом (почему?).

Теорема о касательной плоскости. *Пусть непустое квадратичное многообразие задано уравнением $(Ax,x)+2(b,x)+c=0$ и x_0 — его неособая точка. Тогда любая принадлежащая касательной плоскости и проходящая через точку x_0 прямая неасимптотического направления пересекается с многообразием только в одной точке, а прямая с асимптотическим направлением целиком содержится в многообразии.*

Доказательство. Пусть прямая имеет вид $x=x_0+tp$ и принадлежит касательной плоскости $(Ax_0+b, x-x_0)=0$. Тогда, очевидно, $(Ax_0+b,p)=0$. Значения параметра t, при котором точки прямой попадают на многообразие, удовлетворяют уравнению

$$(A(x_0+tp), x_0+tp)+2(b, x_0+tp)+c=(Ap,p)t^2+2(Ax_0+b,p)t=0.$$

Если $(Ap,p)\neq 0$, то $t=0$ является единственным решением. Если $(Ap,p)=0$, то в качестве t можно взять любое вещественное число и, следовательно, все точки прямой $x=x_0+tp$ принадлежат многообразию. $\square$

12.8 Декартовы системы и ортогональные матрицы

При изучении многообразий декартовы системы координат упрощают вычисление длин и углов. Поэтому при замене $x=Py$ разумно переходить от декартовой системы также к декартовой. Базисные векторы исходной системы — это столбцы единичной матрицы, а базисные векторы новой системы — это столбцы матрицы P. В декартовой системе базисные векторы должны быть ортонормированными. Значит, если новая система остается декартовой, то столбцы $p_1,\ldots,p_n$ матрицы P должны быть связаны соотношениями

$$(p_i,p_j)=p_j^\top p_i=\delta_{ij}=\begin{cases}0 & \text{при } i\neq j,\\ 1 & \text{при } i=j,\end{cases}$$

которые, как нетрудно проверить, равносильны равенству $P^\top P=I$. Квадратная вещественная матрица с таким свойством называется *ортогональной матрицей*.

Обращение ортогональной матрицы сводится к ее транспонированию: $P^{-1}=P$. Кроме того, произведение ортогональных матриц сохраняет ортогональность:

$$P^\top P=Q^\top Q=I \quad\Rightarrow\quad (PQ)^\top(PQ)=Q^\top(P^\top P)Q=Q^\top Q=I.$$

Единичная матрица очевидно является ортогональной. Таким образом, *множество всех ортогональных матриц одного и того же порядка относительно операции умножения образует группу*. Если матрица P ортогональная, то матрица $\Lambda=P^\top AP$ называется *ортогонально конгруэнтной* матрице A, а преобразование $A\to P^\top AP$ — *ортогональной конгруэнцией* или, короче, *ортоконгруэнцией*.

Очень полезное свойство ортогональной конгруэнции — это *сохранение суммы квадратов всех элементов матрицы*. Это очевидное следствие несколько более общего наблюдения.

Лемма о сохранении сумм квадратов. *Сумма квадратов элементов вещественной матрицы не меняется при умножения ее слева или справа на ортогональные матрицы.*

Доказательство. Достаточно заметить, что сумма квадратов вектора-столбца a равна естественному скалярному произведению $(a, a) = a^\top a$. Если P — ортогональная матрица, то $(Pa)^\top (Pa) = a^\top (P^\top P)a = a^\top a$. $\square$

Нас, конечно, интересует такой вопрос: можно ли от вещественной симметричной матрицы перейти к диагональной *посредством ортогональной конгруэнции*? Ответ положительный, но это уже не такое простое утверждение, как теорема о конгруэнтной диагонализации. Более того, можно с полным правом говорить о том, что среди многочисленных и разнообразных применений матричного анализа это один из самых востребованных его фактов.

12.9 Ортоконгруэнтная диагонализация в двумерном случае

Пусть e_1 и e_2 — базисные векторы декартовой системы координат в двумерном пространстве. При повороте системы на угол ϕ против часовой стрелки векторы преобразуются следующим образом: $e_1 \to \cos\phi\, e_1 + \sin\phi\, e_2$, $\quad e_2 \to -\sin\phi\, e_1 + \cos\phi\, e_2$. Матрица перехода в данном случае — это ортогональная матрица вида

$$P = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}.$$

Такие матрицы называются *матрицами вращения*. При ортогональной конгруэнции с матрицей вращения симметричная матрица A переходит в матрицу

$$P^\top A P = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = \begin{bmatrix} \widetilde{a}_{11} & \widetilde{a}_{12} \\ \widetilde{a}_{12} & \widetilde{a}_{22} \end{bmatrix},$$

$$\begin{aligned} \widetilde{a}_{11} &= \cos^2\phi\, a_{11} + 2\cos\phi\sin\phi\; a_{12} + \sin^2\phi\, a_{22}, \\ \widetilde{a}_{22} &= \sin^2\phi\, a_{11} - 2\cos\phi\sin\phi\; a_{12} + \cos^2\phi\, a_{22}, \\ \widetilde{a}_{12} &= (\cos^2\phi - \sin^2\phi)\; a_{12} - \sin\phi\cos\phi\; (a_{11} - a_{22}). \end{aligned}$$

Как видим, если $a_{12} \neq 0$, то внедиагональные элементы преобразованной матрицы зануляются при выборе угла ϕ из уравнения

$$\operatorname{ctg} 2\phi = \frac{\cos^2\phi - \sin^2\phi}{2\sin\phi\cos\phi} = \frac{a_{11} - a_{22}}{2a_{12}} \quad \Rightarrow \quad \widetilde{a}_{12} = 0.$$

При таком выборе матрица $\Lambda = P^\top A P$ будет диагональной, а для получения ее элементов λ_1 и λ_2 можно обойтись без тригонометрии — легко вычисляются их сумма и произведение: $\lambda_1 + \lambda_2 = a_{11} + a_{22}$, $\quad \lambda_1\lambda_2 = |A|$, поэтому λ_1 и λ_2 можно найти как корни квадратного уравнения $\lambda^2 - (a_{11} + a_{22})\lambda + |A| = 0$.

12.10 Метод вращений

Чтобы провести ортоконгруэнтную диагонализацию вещественной симметричной матрицы A порядка $n \geq 3$, матрицы вращения тоже полезны, хотя их употребление в данном случае менее очевидно.

Будем строить, вообще говоря, бесконечную последовательность симметричных матриц

$$A = A_0, \quad A_1 = P_1^\top A_0 P_2, \quad \ldots, \quad A_k = P_k^\top A_{k-1} P_k, \quad \ldots .$$

Матрица $P_k = P_k(i_k, j_k, \phi_k)$ определяется парой номеров $1 \le i_k < j_k \le n$ и углом ϕ_k. Она отличается от единичной матрицы только в четырех позициях $(i_k, j_k), (i_k, j_k), (j_k, i_k), (j_k, j_k)$, в которых размещается матрица вращения для угла $\phi = \phi_k$. Угол выбирается с целью получения нуля в симметричной паре позиций (i_k, j_k) и (j_k, i_k).

Основная неприятность при проведении этого процесса заключается в том, что ранее полученные нули исчезают при полученииии новых нулей. Тем не менее, есть довольно простой способ *выбора позиции* (i_k, j_k) для зануления, гарантирующий сходимость к нулю всех внедиагональных элементов матриц A_k при $k \to \infty$. Пусть на k-м шаге для исключения выбирается внедиагональный элемент, *наибольший по модулю*.

Обозначим через d_k и f_k суммы квадратов элементов матрицы A_k на главной диагонали и вне главной диагонали. Согласно лемме о сумме квадратов, $d_k + f_k = d_{k-1} + f_{k-1}$. В силу нашего правила выбора исключаемого элемента, $|a^{(k-1)}_{i_k,j_k}|^2 \ge f_{k-1}/(n^2 - n)$. Кроме того, как следствие той же леммы о сумме квадратов, примененной к подматрице в позициях $(i_k, j_k), (i_k, j_k), (j_k, i_k), (j_k, j_k)$, находим $d_k = d_{k-1} + 2|a^{(k-1)}_{i_k,j_k}|^2$. Собирая все вместе, получаем неравенство

$$f_k = f_{k-1} - (d_k - d_{k-1}) \;\le\; \gamma f_{k-1}, \quad \gamma = 1 - \frac{2}{n^2 - n}.$$

Если $n = 2$, то $\gamma = 0$, что вполне согласуется с уже разобранным нами случаем. При $n \ge 3$ находим $0 < \gamma < 1 \quad \Rightarrow \quad f_k \;\le\; \gamma^k f_0 \;\to\; 0$.

Теорема о диагонализации вещественных симметричных матриц. *Любая вещественная симметричная матрица ортогонально конгруэнтна некоторой диагональной матрице.*

Доказательство. Заметим, что $A_k = Q_k^\top A Q_k$, где матрица $Q_k = P_1 \dots P_k$ является ортогональной. Все элементы любой ортогональной матрицы принадлежат отрезку $[-1, 1]$. Поэтому, используя теорему Больцано–Вейерштрасса и последовательно выбирая сходящиеся подпоследовательности, мы можем выделить сходящуюся подпоследовательность $Q_{k_l} \to Q \quad \Rightarrow \quad A_{k_l} = Q_{k_l}^\top A Q_{k_l} \to \Lambda$ (сходимость последовательности матриц определяется как сходимость последовательностей их элементов). Предельная матрица Λ является диагональной, так как все внедиагональные элементы матриц A_{k_l} сходятся к нулю. $\square$

Доказанная теорема сформулирована как теорема существования. Но приведенное доказательство замечательно своей конструктивностью: оно дает одновременно и метод приближенного вычисления диагональной матрицы и соответствующей ортогональной матрицы. Это один из ранних практических методов вычислительной алгебры, предложенный К. Якоби в 1846 году. В 1990-х годах были обнаружены особые возможности метода вращений, связанные с высокоточным вычислением малых по модулю элементов искомой диагональной матрицы.

Задача 1. *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.*

12.11 Кривые второго порядка в декартовых координатах

Квадратичные многообразия, заданные квадратичным уравнением в двумерном пространстве, называются *кривыми второго порядка*. Согласно общей теореме о приведенных уравнениях, здесь возникают три возможности:

$$(1)\ \ \lambda_1 x_1^2 + \lambda_2 x_2^2 + c = 0, \quad (2)\ \ \lambda_1 x_1^2 + c = 0, \quad (3)\ \ \lambda_1 x_1^2 + 2\beta x_2 = 0.$$

Напомним, что λ_1, λ_2 и β отличны от нуля, а c разрешается принимать любые значения.

В теореме о конгруэнтной диагонализации рассматривались, однако, аффинные системы. Теперь мы можем также утверждать, что существует и декартова система координат, в которой уравнение кривой второго порядка имеет один из этих трех типов. Новая система координат получается с помощью *поворота* исходной декартовой системы и *сдвига* ее начала.

В случае (1) ранг матрицы квадратичной части равен 2. Здесь выделяются две возможности: когда числа λ_1, λ_2 одного знака (*эллиптический случай*) и когда они разных знаков (*гиперболический случай*). О том, какая ситуация имеет место, можно судить по знаку определителя матрицы квадратичной формы — плюс в эллиптическом случае и минус в гиперболическом. Для уравнений типа (2) ранг матрицы квадратичной части равен 1, а ранг расширенной матрицы не больше 2. В случае (3) ранг матрицы квадратичной формы равен 1, а ранг расширенной матрицы равен 3 (*параболический случай*).

Для непустого многообразия его тип легко определяется по наличию центра симметрии. Когда центра симметрии нет, то это заведомо параболический случай. Если центр симметрии есть и только один, то мы имеем дело с уравнением типа (1), а когда центров симметрии много, то это непременно уравнение типа (2).

12.12 Эллипс

В эллиптическом случае числа λ_1 и λ_2 либо оба положительны, либо оба отрицательны. Многообразие может оказаться пустым — так будет, если числа λ_1, λ_2 и c имеют один и тот же знак. Если $c = 0$, то в нем только одна точка. В остальных, самых интересных случаях после некоторых переобозначений возникает декартова система координат x и y, в которой уравнение кривой получает вид

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a \geq b > 0.$$

Кривая, заданная этим уравнением, называется *эллипсом*. Числа a и b называются *большей и меньшей полуосями эллипса*. При $a = b$ получается, конечно, окружность радиуса a.

Точки $F_\pm = (\pm c,\ 0)$, где $c = \sqrt{a^2 - b^2} \geq 0$, называются правым и левым *фокусами эллипса*. Число $e = c/a$ называется *эксцентриситетом* эллипса ($0 \leq e < 1$). Если $e \neq 0$, то прямые $l_\pm : x = \pm a/e$ называются правой и левой *директрисами* эллипса.

Теорема об эллипсе. *Сумма расстояний от любой точки эллипса до его фокусов постоянна и равна* $2a$. *Если эксцентриситет отличен от нуля, то отношение расстояний от любой точки эллипса до правого (левого) фокуса и до правой (левой) директрисы постоянно и равно его эксцентриситету.*

Доказательство. Пусть точка $M = (x,\ y)$ принадлежит эллипсу. Учитывая, что $y^2 = b^2 - (b^2/a^2)x^2$, находим (проверьте!)

$$|MF_-| + |MF_+| = \sqrt{(x+c)^2 + y^2}\ +\ \sqrt{(x-c)^2 + y^2} =$$

$$|a + ex| + |a - ex| = (a + ex)\ +\ (a - ex) = 2a.$$

$$|MF_-| = |a + ex| = e|x + (a/e)| \quad \Rightarrow \quad \frac{|MF_-|}{|x + (a/e)|} = e,$$

$$|MF_+| = |a - ex| = e|x - (a/e)| \quad \Rightarrow \quad \frac{|MF_+|}{|x - (a/e)|} = e.\ \square$$

Эллипс обладает интересным *оптическим свойством*: прямая, ортогональная к касательной к эллипсу в точке M, является биссектрисой угла F_-MF_+. Данное утверждение можно доказать с помощью вычислений, выписав уравнение касательной, но оно почти очевидно получается из простых геометрических соображений. Достаточно заметить, что минимальная сумма расстояний от двух заданных точек A и B до точек заданной прямой достигается в точке M, которая лежит на пересечении данной прямой с прямой $A'B$, где точка A' получается из A симметричным отражением относительно заданной прямой. Отсюда ясно, что биссектриса угла AMB ортогональна данной прямой. В качестве точек A и B нужно взять фокусы, тогда на касательной в точке M минимум суммы расстояний до фокусов реализуется именно в точке M: любая точка $P \neq M$ будет внешней для области, ограниченной эллипсом, а для суммы расстояний от нее до фокусов выполняется неравенство $|AP| + |PB| > |AM| + |MB|$ (докажите!).

Задача 2. *Докажите, что геометрическое место точек, сумма расстояний от которых до двух заданных точек постоянна, является отрезком прямой либо эллипсом.*

Задача 3. *Докажите, что геометрическое место точек, для которых отношение расстояний до заданной точки и заданной прямой постоянно и равно $0 < e < 1$, является эллипсом.*

Задача 4 *Докажите, что преобразование $z \mapsto \frac{1}{2}(z+z^{-1})$ комплексной плоскости [1] переводит точки окружности радиуса $r > 1$ с центром в начале координат в точки некоторого эллипса.*

Задача 5. *Докажите, что точки эллипса $x^2/a^2 + y^2/b^2 = 1$ допускают параметрическое представление $x = a\cos\phi,\ y = b\sin\phi,\quad 0 \leq \phi < 2\pi$.*

Задача 6. *Докажите, что через каждую точку, лежащую вне ограниченной эллипсом замкнутой области, проходят ровно две прямые, касающиеся данного эллипса.*

Задача 7 *Пусть A и B — фокусы эллипса, а его касательные в точках M и N пересекаются в точке P. Докажите, что угол MPA равен углу BPN.*

Задача 8. *Докажите, что геометрическое место точек пересечения взаимно ортогональных касательных к эллипсу $x^2/a^2 + y^2/b^2 = 1$ есть окружность $x^2 + y^2 = a^2 + b^2$.*

[1] Функция $z \mapsto \frac{1}{2}(z + z^{-1})$ называется *функцией Жуковского* и широко применяется при решении задач гидро- и аэродинамики.

12.13 Гипербола

В гиперболическом случае числа λ_1 и λ_2 имеют разные знаки. При $c = 0$ получаем пару прямых, проходящих через начало координат. Если $c \neq 0$, то в некоторой декартовой системе координат (после переобозначений) уравнение может быть записано в виде

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

где a, b —положительные числа (неравенство $a \geq b$ здесь не предполагается). Кривая, заданная этим уравнением, называется *гиперболой*. Точки гиперболы находятся в объединении двух непересекающихся областей плоскости (распадаются на *две ветви*) в соответствии с неравенством $|y| \leq (b/a)|x|$ при $x \geq a$ и при $x \leq -a$.

Прямые $h_\pm : y = (b/a)x$ называются *асимптотами* гиперболы. Рассмотрим ее точки $M(x, y)$ при $x > 0$ и $y > 0$. Очевидно, расстояние от точки $M(x, y)$ до асимптоты h_+ не превышает

$$|(b/a)x - y(x)| = \frac{b^2}{|(b/a)x + y(x)|} \to 0 \quad \text{при} \quad x \to +\infty.$$

Точки гиперболы при $x > 0$ и $y < 0$ приближаются к асимптоте h_-. Аналогичные наблюдения справедливы также для точек гиперболы при $x < 0$.

Точки $F_\pm = (\pm c,\ 0)$, где $c = \sqrt{a^2 + b^2} > 0$, называются правым и левым *фокусами* гиперболы. Число $e = c/a$ называется *эксцентриситетом* гиперболы ($e > 1$). Прямые $l_\pm : x = -a/e$ называются правой и левой *директрисами* гиперболы.

Теорема о гиперболе. *Разность расстояний от произвольной точки гиперболы до ее левого и правого фокусов для каждой ветви гиперболы постоянна и равна $2a$ для правой ветви и $-2a$ для левой. Отношение расстояний от произвольной точки гиперболы до ее правого (левого) фокуса и до правой (левой) директрисы постоянно и равно эксцентриситету гиперболы.*

Доказательство. Пусть $M = (x, y)$ — точка гиперболы. Как и в случае эллипса, при вычислении расстояний под радикалами удачно возникают полные квадраты и в результате

$$|MF_-| = \sqrt{(x + c)^2 + y^2} = |a + ex|,$$
$$|MF_+| = \sqrt{(x - c)^2 + y^2} = |a - ex|.$$

Поскольку $|x| \geq a$ и $e > 1$, получаем

$$|a + ex| - |a - ex| = \begin{cases} (ex + a) - (ex - a) = 2a, & x > 0, \\ -(ex + a) + (ex - a) = -2a, & x < 0. \end{cases}$$

Постоянство отношения расстояний от точки гиперболы до фокуса и до соответствующей директрисы проверяется так же, как в случае эллипса. □

Задача 9. *Написать общее уравнение касательной прямой, проходящей через точку $M(x_0, y_0)$ гиперболы $x^2/a^2 - y^2/b^2 = 1$ и доказать, что данная прямая является биссектрисой угла AMB, где A и B — фокусы гиперболы.*

Задача 10. *Докажите, что никакая прямая не может пересекаться с каждой ветвью гиперболы ровно в одной точке.*

Задача 11. *Докажите, что точки ветви гиперболы $x^2/a^2 - y^2/b^2 = 1$ при $x > 0$ допускают параметрическое представление $x = a\,ch(\phi)$, $y = b\,sh(\phi)$, $-\infty < \phi < +\infty$. По определению,*

$$ch(\phi) = \frac{1}{2}(e^\phi + e^{-\phi}), \quad sh(\phi) = \frac{1}{2}(e^\phi - e^{-\phi}).$$

12.14 Парабола

В параболическом случае кривая называется *параболой* и ее уравнение принято записывать в виде

$$y^2 \;=\; 2px, \quad p > 0.$$

Соответствующая декартова система возникает после вполне очевидных переобозначений. Число p называется *фокальным параметром* параболы. Точка $F = (p/2,\ 0)$ называется *фокусом* параболы, а прямая $l : x = -p/2$ — *директрисой* параболы.

Теорема о параболе. *Для каждой точки параболы расстояние до фокуса равно расстоянию до директрисы.*

Доказательство. Пусть $M = (x, y)$ — произвольная точка параболы. Тогда

$$\begin{aligned} |MF| \;=\; \sqrt{(x-p/2)^2+y^2} \quad &= \quad \sqrt{x^2-px+(p/2)^2+2px} \\ &= \quad \sqrt{x^2+2x(p/2)+(p/2)^2} \;=\; |x+p/2|. \end{aligned}$$

Задача 12. *Написать общее уравнение касательной прямой, проходящей через точку $M(x_0, y_0)$ параболы $y^2 = 2px$, и доказать, что перпендикуляр к этой прямой в точке M делит пополам угол между прямой MF, где точка F — фокус параболы, и прямой, проходящей через точку M параллельно оси x.*

Задача 13. *Докажите, что геометрическое место точек пересечения взаимно ортогональных касательных к параболе совпадает с ее директрисой.*

Задача 14. *Кривая S — это эллипс, одна из ветвей гиперболы или парабола, E — произвольная фиксированная точка на S. Для произвольных точек A и B на S их суммой называется точка C, получаемая при пересечении S прямой, проходящей через точку E параллельно отрезку AB при $A \neq B$ и касательной к S в точке $A = B$ в противном случае. Докажите, что множество S относительно этой операции является абелевой группой.*

12.15 Поверхности второго порядка в декартовых координатах

Квадратичные многообразия, заданные квадратичным уравнением в трехмерном пространстве, называются *поверхностями второго порядка*. Согласно общей теореме о приведенных уравнениях, здесь возникают уравнения пяти типов:

$$(1)\ \ \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 + c = 0, \quad (2)\ \ \lambda_1 x_1^2 + \lambda_2 x_2^2 + c = 0, \quad (3)\ \ \lambda_1 x_1^2 + c = 0,$$

$$(4)\ \ \lambda_1 x_1^2 + \lambda_2 x_2^2 + 2\beta x_3 = 0, \quad (5)\ \ \lambda_1 x_1^2 + 2\beta x_2 = 0.$$

Числа $\lambda_1, \lambda_2, \lambda_3$ и β отличны от нуля, а c может принимать любые значения.

В дополнение к теореме о приведенных уравнениях, мы можем утверждать, что *эти типы уравнений могут быть получены в декартовых системах координат*. Упрощение вида квадратичной части реализуется с помощью ортоконгруэнции — ее существование гарантировано теоремой о диагонализации вещественной симметричной матрицы, применяемой к матрицам порядка $n = 3$. Сдвиги начала системы координат, очевидно, сохраняют декартовость.

Единственное место, которое требует пояснения, возникает при получении уравнения типа (5). В данном случае после ортоконгруэнции и выделения полного квадрата для первой координаты появляется уравнение вида $\lambda_1 x_1^2 + 2b_2 x_2 + 2b_3 x_3 + c = 0$. При

выводе теоремы о приведенных уравнений для аффинных систем кооординат строилась невырожденная матрица Q, для которой $Q\begin{bmatrix} b_1 \\ b_3 \end{bmatrix} = \begin{bmatrix} \beta \\ 0 \end{bmatrix}$. Теперь нужно заметить, что в качестве Q можно выбрать некоторую матрицу вращения (докажите!). Тогда конгруэнция с помощью матрицы вида $\Pi = \begin{bmatrix} 1 & 0 \\ 0 & Q^\top \end{bmatrix}$ будет ортоконгруэнцией.

12.16 Примеры поверхностей второго порядка

Здесь мы рассмотрим наиболее интересные поверхности второго порядка, возникающие при исследовании полученных нами пяти типов приведенных уравнений.

Эллипсоид.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad a,\ b,\ c > 0.$$

Уравнение такого вида получается с помощью переобозначений из уравнения типа (1) в случае, когда числа $\lambda_1, \lambda_2, \lambda_3$ одинаковый знак, противоположный знаку числа c. Название отражает тот факт, что в любом сечении эллипсоида плоскостью возникает эллипс (вырождающийся в точку, когда плоскость касается эллипсоида). Заметим, что эллипсоид целиком содержится в параллелепипеде $|x| \le a,\ |y| \le b,\ |z| \le c$.

Задача 15. *Написать общее уравнение касательной плоскости эллипсоида в точке $M(x_0, y_0, z_0)$.*

Задача 16. *Докажите, что в сечении эллипсоида некоторой плоскостью можно получить окружность.*

Задача 17. *Доказать, что геометрическое место точек пересечения трех взаимно ортогональных касательных плоскостей к эллипсоиду $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ есть сфера $x^2 + y^2 + z^2 = a^2 + b^2 + c^2$.*

Задача 18. *В n-мерном вещественном арифметическом пространстве расположен n-мерный эллипсоид $x_1^2/a_1^2 + \ldots + x_n^2/a_n^2 = 1$. Докажите, что через любую точку $z = (z_1, \ldots, z_n)$ такую, что $z_1^2/a_1^2 + \ldots + z_n^2/a_n^2 > 1$, проходит гиперплоскость, которая касается данного эллипсоида в некоторой точке.*

Задача 19. *В n-мерном вещественном арифметическом пространстве скалярное произведение векторов $x = (x_1, \ldots, x_n)$ и $y = (y_1, \ldots, y_n)$ определяется формулой $(x, y) = x_1y_1 + \ldots + x_ny_n$. Докажите, что геометрическое место точек пересечения n попарно ортогональных касательных плоскостей к n-мерному эллипсоиду $x_1^2/a_1^2 + \ldots + x_n^2/a_n^2 = 1$ задается уравнением $x_1^2 + \ldots + x_n^2 = a_1^2 + \ldots + a_n^2$.*

Задача 20. *Отрезок AB является диаметром сферы S, плоскость L касается сферы в точке A. Отображение $f : S\backslash\{B\} \to L$, называемое* `стереографической проекцией`*, определяется следующим правилом: $f(X)$ — это точка пересечения прямой BX с плоскостью L. Докажите, что точки любой окружности на сфере S переходят в точки окружности на плоскости L, если окружность на сфере не проходит через точку B, и в точки некоторой прямой на плоскости L в противном случае.*

Однополостный гиперболоид.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1, \qquad a,\ b,\ c > 0.$$

Это уравнение получается из уравнения типа (1), когда одно из трех чисел $\lambda_1, \lambda_2, \lambda_2$ имеет знак, противоположный общему знаку двух других, но совпадающий со знаком числа c. Название связано с тем, что, во-первых, в любом сечении однополостного

гиперболоида плоскостью $x = \text{const}$ или $y = \text{const}$ возникает гипербола, и, во-вторых, точки с координатами x, y, z, подчиненными неравенству

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} < 1,$$

образуют *связное* множество (полость): вместе с любыми двумя точками оно целиком содержит все точки некоторой соединяющей их ломаной линии, состоящей из конечного числа отрезков.

Двуполостный гиперболоид.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1, \quad a,\ b,\ c > 0.$$

Данная поверхность также связана с уравнением типа (1). Она не имеет точек в полосе $|z| < c$, а множество точек, определенное неравенством

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} < -1,$$

разбивается на два связных множества — две полости.

Эллиптический конус.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0.$$

Это еще одна форма уравнения типа (1) в случае $c = 0$. Название поверхности связано с тем, что при ее сечении плоскостью $z = \text{const}$ получаются эллипсы.

Задача 21. *Дана плоскость $Ax+By+Cz+D = 0$ при условии $D \neq 0$ и круговой конус $x^2+y^2-z^2 = 0$. Докажите, что в сечении конуса данной плоскостью получается эллипс, гипербола или парабола в том и только том случае, когда, соответственно, $A^2 + B^2 < C^2$, $A^2 + B^2 > C^2$, $A^2 + B^2 = C^2$.*

Эллиптический параболоид.

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = z, \qquad a,\ b > 0.$$

Это один из случаев, связанных с уравнением типа (4). Название навеяно формой сечений в плоскостях $z = \text{const}$ (эллипсы) и в плоскостях $x = \text{const}$ и $y = \text{const}$ (параболы).

Гиперболический параболоид.

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = z, \qquad a,\ b > 0.$$

Название объясняется видом кривых, получаемых в сечениях плоскостями $z = \text{const}$ (гиперболы) и плоскостями $x = \text{const}$ и $y = \text{const}$ (параболы).

Цилиндрические поверхности. Приведенные уравнения типов (2), (3), (5) не зависят от z. Поэтому кривые в сечениях любой плоскостью вида $z = \text{const}$ одинаковы. Такие поверхности называются *цилиндрическими*.

Задача 22. *Сфера вписана в круговой цилиндр. Докажите, что любая ее точка, не принадлежащая поверхности цилиндра, является фокусом эллипса, который получается в сечении цилиндра плоскостью, касательной к сфере в данной точке.*

12.17 Линейчатые поверхности

Однополостный гиперболоид и гиперболический параболоид являются примерами *линейчатой поверхности* — так называются поверхности, составленные из целиком принадлежащих им прямых.

Утверждение. *Через каждую точку однополостного гиперболоида и гиперболического параболоида проходят в точности две прямые, все точки которых принадлежат данной поверхности.*

Доказательство. Мы ограничимся случаем однополостного гиперболоида. Изменив масштаб, перейдем к аффинной системе координат, в которой он получает уравнение вида $x^2 + y^2 - z^2 = 1$. Рассмотрим прямую

$$x = x_0 + p_1 t, \quad y = y_0 + p_2 t, \quad z = z_0 + p_3 t,$$

проходящую через точку $M(x_0, y_0, z_0)$ на поверхности, и запишем условие, при котором все ее точки принадлежат этой же поверхности:

$$(x_0 + p_1 t)^2 + (y_0 + p_2 t)^2 - (z_0 + p_3 t)^2 = 1 \quad \forall \ \ t \in \mathbb{R} \quad \Leftrightarrow$$

$$\left\{ \begin{array}{ccccccc} p_1^2 & + & p_2^2 & - & p_3^2 & = & 0, \\ p_1 x_0 & + & p_2 y_0 & - & p_3 z_0 & = & 0, \\ x_0^2 & + & y_0^2 & - & z_0^2 & = & 1. \end{array} \right.$$

Нас интересуют только ненулевые векторы $(p_1, p_2, p_3) \ \Rightarrow \ p_3 \neq 0$. Поскольку направляющий вектор прямой определяется с точностью до множителя, мы нормируем его условием $p_3 = 1 \quad \Rightarrow \quad p_1^2 + p_2^2 = 1, \ \ p_1 x_0 + p_2 y_0 = z_0$.

Очевидно, значения x_0 и y_0 не могут быть нулями одновременно. Рассмотрим случай

$$y_0 \neq 0 \quad \Rightarrow \quad p_2 = (z_0 - p_1 x_0)/y_0 \quad \Rightarrow \quad p_1^2 + (z_0 - p_1 x_0)^2/y_0^2 = 1 \quad \Rightarrow$$

$$(x_0^2 + y_0^2) p_1^2 \ - \ 2(x_0 z_0) p_1 \ + \ (z_0^2 - y_0^2) \ = \ 0.$$

Вычисляем дискриминант:

$$x_0^2 z_0^2 - (x_0^2 + y_0^2)(z_0^2 - y_0^2) = y_0^2 (x_0^2 + y_0^2 - z_0^2) = y_0^2 > 0.$$

Таким образом, для p_1 получаем в точности два различных значения. Поскольку $p_3 = 1$, соответствующие направляющие векторы, очевидно, линейно независимы и порождают две разные прямые, проходящие через точку (x_0, y_0, z_0) и целиком принадлежащие поверхности. Случай $x_0 \neq 0$ разбирается аналогично. $\square$

Замечание. Для поиска тех же самых прямых на поверхности однополостного гиперболоида можно записать уравнение поверхности 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

$$\alpha\left(\frac{x}{a} - \frac{z}{c}\right) = \beta\left(1 - \frac{y}{b}\right), \ \beta\left(\frac{x}{a} + \frac{z}{c}\right) = \alpha\left(1 + \frac{y}{b}\right);$$

$$\gamma\left(\frac{x}{a} - \frac{z}{c}\right) = \delta\left(1 + \frac{y}{b}\right), \ \delta\left(\frac{x}{a} + \frac{z}{c}\right) = \gamma\left(1 - \frac{y}{b}\right),$$

определяемых парами не равных одновременно нулю параметров α, β и γ, δ. Ясно, что для каждой пары плоскостей в пересечении получается прямая, целиком принадлежащая поверхности. Но надо еще как-то обосновать, почему получаются ровно две прямые и почему других прямых на этой поверхности быть не может.

Задача 23. *Докажите, что через каждую точку гиперболического параболоида проходят в точности две прямые, все точки которых принадлежат данной поверхности.*