§8. Матрицы

8.1. Алгебры над полем. Векторное пространство A над полем $\Bbbk$ называется *алгеброй* над $\Bbbk$ (или $\Bbbk$-*алгеброй*), если на нём имеется такая операция умножения $A \times A \to A$, что при каждом $a \in \Bbbk$ операторы левого и правого умножения на a

$$\lambda_a : A \to A\,,\ v \mapsto av\,, \quad \text{и} \quad \varrho_a : A \to A\,,\ v \mapsto va\,, \tag{8-1}$$

линейны. Это включает в себя перестановочность умножения векторов на константы с умножением в алгебре: $\forall\,\lambda \in \Bbbk\,,\ \forall\,a, b \in A\ \ (\lambda a)b = \lambda(ab) = a(\lambda b)$ и стандартное правило раскрытия скобок: $\forall\,\lambda_1, \lambda_2, \mu_1, \mu_2 \in \Bbbk$ и $\forall\,a_1, a_2, b_1, b_2 \in A$

$$(\lambda_1 a_1 + \mu_1 b_1)(\lambda_2 a_2 + \mu_2 b_2) = \lambda_1\lambda_2 a_1 a_2 + \lambda_1\mu_2 a_1 b_2 + \mu_1\lambda_2 b_1 a_2 + \mu_1\mu_2 b_1 b_2\,.$$

Алгебра A называется *ассоциативной*, если $\forall\,a, b, c \in A\ (ab)c = a(bc)$. Алгебра A называется *коммутативной*, если $\forall\,a, b \in A\ ab = ba$. Алгебра A называется *алгеброй с единицей*, если в ней есть нейтральный элемент по отношению к умножению (или *единица*) — такое $e \in A$, что $ea = ae = a$ для всех $a \in A$.

Упражнение 8.1. Покажите, что $0 \cdot a = 0$ для всех a в любой алгебре A и что единичный элемент единственен (если существует).

Примерами *коммутативных* ассоциативных алгебр с единицами являются алгебра многочленов $\Bbbk[x_1, x_2, \dots, x_n]$ и прочие коммутативные $\Bbbk$-алгебры в смысле n° 5.2.1 на стр. 74. Модельный пример некоммутативной ассоциативной алгебры — это линейные эндоморфизмы векторного пространства.

Пример 8.1 (алгебра $\operatorname{End} V$ линейных эндоморфизмов пространства V)
Композиция линейных отображений $G : U \to V$ и $F : V \to W$ тоже является линейным отображением, поскольку $FG(\lambda u + \mu w) = F(\lambda G(u) + \mu G(w)) = \lambda FG(u) + \mu FG(w)$. При этом отображение композиции $\operatorname{Hom}(V, W) \times \operatorname{Hom}(U, V) \to \operatorname{Hom}(U, W)$, $(F, G) \mapsto FG$, линейно по каждому из аргументов при фиксированном втором:

$$(\lambda_1 F_1 + \lambda_2 F_2)G = \lambda_1 F_1 G + \lambda_2 F_2 G \quad \text{и} \quad F(\mu_1 G_1 + \mu_2 G_2) = \mu_1 FG_1 + \mu_2 FG_2\,.$$

Таким образом, линейные эндоморфизмы $\operatorname{End} V = \operatorname{Hom}(V, V)$ любого пространства V образуют алгебру с единицей $e = \mathrm{Id}_V$.

Упражнение 8.2. Составьте таблицу умножения базисных операторов[1] $E_{ij} \in \operatorname{End}(\Bbbk^n)$ и покажите, что при $\dim V \geqslant 2$ композиция в $\operatorname{End}(\Bbbk^n)$ не коммутативна.

Поскольку композиция отображений всегда ассоциативна (когда определена):

$$F(GH) = (FG)H : u \mapsto F(G(H(u)))\,,$$

алгебра $\operatorname{End} V$ ассоциативна.

[1]напомним (см. предл. 6.2), что линейный оператор $E_{ij} : \Bbbk^n \to \Bbbk^n$ переводит e_j в e_i, а все остальные стандартные базисные векторы — в нуль

8.1.1. Обратимые элементы. Элемент a алгебры A с единицей $e \in A$ называется *обратимым*, если существует $a^{-1} \in A$, такой что $aa^{-1} = a^{-1}a = e$. В ассоциативной алгебре A это требование можно ослабить до существования левого и правого обратных к a элементов $a', a'' \in A$, таких что $a'a = aa'' = e$ — при этом они автоматически совпадут друг с другом: $a' = a'e = a'(aa'') = (a'a)a'' = ea'' = a''$. Эта же выкладка показывает, что обратный к a элемент однозначно определяется по a.

Пример 8.2 (полная линейная группа $\mathrm{GL}\,V \subset \mathrm{End}\,V$)
Согласно предл. 1.4 обратимыми элементами алгебры $\mathrm{End}\,V$ являются линейные изоморфизмы $V \xrightarrow{\sim} V$. Они образуют группу преобразований[1] пространства V. Эта группа обозначается $\mathrm{GL}\,V$ и называется *полной линейной группой* пространства V.

8.1.2. Алгебраические и трансцендентные элементы. Любой элемент ξ любой ассоциативной $\Bbbk$-алгебры A с единицей определяет гомоморфизм вычисления

$$\mathrm{ev}_\xi : \Bbbk[t] \to A\,, \quad f(x) \mapsto f(\xi) \in A \tag{8-2}$$

который переводит многочлен $f(x) = a_0x^m + a_1x^{m-1} + \dots + a_{m-1}x + a_m$ в результат подстановки в него[2] $x = \xi$.

Если гомоморфизм (8-2) инъективен, то элемент ξ называется *трансцендентным* над $\Bbbk$. Отметим, что в этом случае алгебра A обязательно бесконечномерна как векторное пространство над $\Bbbk$, поскольку все степени элемента ξ линейно независимы.

Если гомоморфизм (8-2) имеет ненулевое ядро, то элемент ξ называется *алгебраическим* над $\Bbbk$. В этом случае ядро $\ker \mathrm{ev}_\xi = (\mu_\xi)$ является главным идеалом[3] в $\Bbbk[x]$. Приведённый многочлен, порождающий этот идеал, называется *минимальным многочленом* элемента ξ и обозначается $\mu_\xi(x)$. Таким образом, минимальный многочлен — это многочлен наименьшей степени с единичным старшим коэффициентом, такой что $\mu_\xi(\xi) = 0$. Отметим, что все остальные многочлены, аннулирующие ξ, делятся на минимальный.

Пример 8.3 (аннулирующий многочлен линейного оператора)
Поскольку алгебра эндоморфизмов $\mathrm{End}\,V$ n-мерного векторного пространства V имеет размерность $\dim \mathrm{End}\,V = n^2$, всякий линейный оператор $F : V \to V$ алгебраичен над $\Bbbk$: n^2+1 векторов $F^k \in \mathrm{End}\,V$ с $0 \leqslant k \leqslant n^2$ линейно зависимы, и равная нулю нетривиальная линейная комбинация этих элементов представляет собой многочлен степени не выше n^2, аннулирующий оператор F. На самом деле эта степень сильно завышена: в n° 10.1.4 на стр. 148 мы покажем, что любой оператор $F : V \to V$ аннулируется многочленом степени $\dim V$ (см. также прим. 8.4 на стр. 120).

8.2. Алгебра матриц. Рассмотрим три координатных пространства $\Bbbk^n$, $\Bbbk^s$, $\Bbbk^n$ и обозначим через $u_1, u_2, \dots, u_n \in \Bbbk^n$, $v_1, v_2, \dots, v_s \in \Bbbk^s$, $w_1, w_2, \dots, w_m \in \Bbbk^m$ их стандартные базисы. Пусть линейные операторы $B : \Bbbk^n \to \Bbbk^s$ и $A : \Bbbk^s \to \Bbbk^m$ имеют в этих базисах матрицы $A = (a_{ij})$ и $B = (b_{ij})$. Матрица $P = (p_{ij})$ их композиции $P = AB : \Bbbk^n \to \Bbbk^m$

[1] см. n° 1.6 на стр. 14
[2] т. е. в $a_0\xi^m + a_1\xi^{m-1} + \dots + a_{m-1}\xi + a_m\xi^0 \in A$, где $\xi^0 \stackrel{\text{def}}{=} e$ — единица алгебры A
[3] ибо $\Bbbk[x]$ — это кольцо главных идеалов

называется *произведением*[1] матриц A и B. Таким образом, для каждой упорядоченной пары матриц, в которой ширина первой матрицы совпадает с высотой второй, определена матрица-произведение, имеющая столько же строк, сколько первый сомножитель, и столько же столбцов, сколько второй. Элемент p_{ij} в пересечении i-той строки и j-того столбца произведения равен коэффициенту при w_i в разложении $AB(u_j) = A\Big(\sum_k v_k b_{kj}\Big) = = \sum_k A(v_k) b_{kj} = \sum_i \sum_k w_i a_{ik} b_{kj}$, т. е. $p_{ij} = \sum_k a_{ik} b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{is}b_{sj}$. Это правило для вычисления произведения можно переговорить несколькими эквивалентными способами, каждый из которых по-своему полезен при практических вычислениях.

Во-первых, произведение матриц полностью определяется правилом умножения строки ширины s на столбец высоты s:

$$(a_1, a_2, \dots, a_s) \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{pmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_s b_s\,,$$

и результат умножения матрицы A из m строк ширины s на матрицу B из n столбцов высоты s представляет собою таблицу, в (i, j)-той клетке которой стоит произведение i-той строки A на j-тый столбец B :

$$p_{ij} = \big(a_{i1},\, a_{i2},\, \dots\,,\, a_{is}\big) \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj} \end{pmatrix}.$$

Упражнение 8.3. Убедитесь, что операция транспонирования матриц $A \mapsto A^t$ (см. n° 7.3.1) взаимодействует с умножением матриц по правилу $(AB)^t = B^t A^t$.

Второе описание таково: в j-том столбце произведения AB стоит линейная комбинация s столбцов матрицы A, рассматриваемых как векторы координатного пространства $\Bbbk^m$, с коэффициентами, стоящими в j-том столбце матрицы B. Если, к примеру, в матрице

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \tag{8-3}$$

хочется написать вместо второго столбца сумму первого и третьего, а первый и третий столбец заменить на их суммы со вторым, умноженным, соответственно, на λ и на μ, после чего добавить к полученной матрице ещё один, четвёртый столбец, равный сумме столбцов матрицы A, умноженных на их номера, то это достигается умножением A справа на матрицу

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ \lambda & 0 & \mu & 2 \\ 0 & 1 & 1 & 3 \end{pmatrix}$$

Упражнение 8.4. Проверьте это прямым вычислением по первому способу.

[1]обратите внимание, что сомножители стоят в том же порядке, что и операторы в композиции

Третье описание произведения двойственно второму и получается из во второго описания произведения транспонированных матриц $B^tA^t = (AB)^t$ заменой слова «столбец» на слово «строка». А именно, в i-той строке матрицы AB стоит линейная комбинация s строк матрицы B, рассматриваемых как векторы координатного пространства $\Bbbk^n$, с коэффициентами, стоящими в i-той строке матрицы A. Например, если в той же матрице (8-3) хочется поставить вторую строку на место первой, а вместо второй написать её сумму с первой строкой, умноженной на λ, то это достигается умножением слева на матрицу

$$\begin{pmatrix} 0 & 1 \\ \lambda & 1 \end{pmatrix}$$

Упражнение 8.5. Проверьте это прямым вычислением по первому способу.

Поскольку композиция операторов ассоциативна и линейна по каждому сомножителю, произведение матриц также ассоциативно и линейно по каждому сомножителю, т. е.

$$(FG)H = H(FG) \quad \forall\, F \in \mathrm{Mat}_{m\times k}\,,\; G \in \mathrm{Mat}_{k\times\ell}\,,\; H \in \mathrm{Mat}_{\ell\times n}$$
$$(\lambda_1F_1 + \mu_1G_1)(\lambda_2F_2 + \mu_2G_2) = \lambda_1\lambda_2F_1F_2 + \lambda_1\mu_2F_1G_2 + \mu_1\lambda_2G_1F_2 + \mu_1\mu_2G_1G_2$$

Таким образом, пространство $\mathrm{Mat}_n(\Bbbk) \overset{\text{def}}{=} \mathrm{Mat}_{n\times n}(\Bbbk) \simeq \mathrm{End}(\Bbbk^n)$ квадратных матриц размера $n\times n$ является ассоциативной $\Bbbk$-алгеброй с единицей

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

(по диагонали стоят единицы, в остальных местах — нули). При $n \geqslant 2$ алгебра $\mathrm{Mat}_n(\Bbbk)$ некоммутативна. Например:

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 10 \\ 12 & 15 \end{pmatrix}$$
$$\begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 4 & 23 \end{pmatrix}$$

Пример 8.4 (аннулирующий многочлен 2×2-матрицы)
Поскольку $\dim \mathrm{Mat}_n(\Bbbk) = n^2 < \infty$, все матрицы алгебраичны над $\Bbbk$. Покажем, что каждая 2×2-матрица $F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ удовлетворяет квадратному уравнению. В самом деле,

$$F^2 = \begin{pmatrix} a^2 + bc & ab + bd \\ ca + dc & cb + d^2 \end{pmatrix} = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & cb + d^2 \end{pmatrix}\,.$$

Поэтому

$$F^2 - (a+d)\cdot F = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & cb + d^2 \end{pmatrix} - \begin{pmatrix} a(a+d) & b(a+d) \\ c(a+d) & d(a+d) \end{pmatrix} =$$
$$= \begin{pmatrix} (bc - ad) & 0 \\ 0 & (bc - ad) \end{pmatrix} = (bc - ad)\cdot E$$

и F удовлетворяет квадратному уравнению $F^2 - (a+b)F + (ad-bc)E = 0$. Числа

$$\det F \overset{\text{def}}{=} ad - bc \quad \text{и} \quad \operatorname{tr} F \overset{\text{def}}{=} a + b$$

называются, соответственно, *определителем* и *следом* матрицы F. В этих обозначениях квадратное уравнение на матрицу F имеет вид

$$F^2 - \operatorname{tr}(F) \cdot F + \det(F) \cdot E = 0\,. \tag{8-4}$$

8.3. Обратимые матрицы. Обратимые элементы алгебры $\mathrm{Mat}_n(\Bbbk)$ называются *обратимыми матрицами*. Это в точности матрицы линейных изоморфизмов координатного пространства $\Bbbk^n$, записанные в стандартном базисе. Группа обратимых матриц обозначается $\mathrm{GL}_n(\Bbbk) \subset \mathrm{Mat}_n(\Bbbk)$.

Пример 8.5 (обратимые 2×2-матрицы)
Формулу (8-4) можно переписать в виде $\det(F) \cdot E = \operatorname{tr}(F) \cdot F - F^2 = F \cdot \big(\operatorname{tr}(F)\,E - F\big)$. Если матрица $F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ обратима, то, умножая обе части слева на F^{-1}, получаем

$$\det(F) \cdot F^{-1} = \operatorname{tr}(F) \cdot E - F = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{8-5}$$

При $\det F = 0$ в левой части стоит нулевая матрица, откуда $a = b = c = d = 0$, и т. к. нулевая матрица не обратима, мы заключаем, что матрицы F с $\det F = 0$ необратимы. Наоборот, при $\det F \neq 0$ формула (8-5) явно вычисляет F^{-1}:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = (ad-bc)^{-1} \begin{pmatrix} a & -b \\ -c & d \end{pmatrix} \tag{8-6}$$

Упражнение 8.6. Проверьте прямым вычислением, что эта матрица обратна к F.

8.3.1. Обращение матриц методом Гаусса. Выяснить, обратима ли данная матрица $A \in \mathrm{Mat}_n(\Bbbk)$, и если да, то явно вычислить A^{-1}, можно умножая матрицу A слева на *заведомо обратимые* матрицы с таким расчётом, чтобы в результате линейных преобразований строк, которые матрица A при этом будет испытывать, в конце концов получилась либо единичная матрица, либо матрица с нулевой строкой.

Упражнение 8.7. Покажите, что матрица с линейно зависимыми строками или столбцами (в частности, матрица, содержащая нулевую строку или нулевой столбец) необратима.

Если после k последовательных умножений слева на обратимые матрицы $S_1, S_2, \dots, S_k$ получится заведомо необратимая матрица $N = S_k S_{k-1} \cdots S_2 S_1 A$, то матрица A тоже не обратима, ибо существование A^{-1} влечёт существование $N^{-1} = A^{-1} S_1^{-1} S_2^{-1} \,\dots\, S_k^{-1}$. Если же получится единичная матрица $S_k S_{k-1} \cdots S_2 S_1 A = E$, то умножая это равенство слева на $S_1^{-1} S_2^{-1} \,\dots\, S_k^{-1}$, мы приходим к соотношению $A = S_1^{-1} S_2^{-1} \,\dots\, S_k^{-1}$, из которого вытекает, что A обратима, и обратная к A матрица $A^{-1} = S_k S_{k-1} \cdots S_2 S_1 E$ получается применением к единичной матрице E ровно той же цепочки преобразований, которая позволила получить из матрицы A матрицу E.

Вычисление удобно организовать следующим образом. Припишем к матрице A справа единичную матрицу, чтобы получилась матрица $\boxed{A\,|\,E}$ размера $n \times 2n$. Если в результате обратимых линейных преобразований строк этой большой матрицы мы придём к матрице вида $\boxed{E\,|\,B}$, то $A^{-1} = B$. Если же мы придём к матрице вида $\boxed{N\,|\,C}$, в которой N необратима, то матрица A тоже необратима. Коль скоро мы знаем все обратимые 2×2-матрицы, проще всего на каждом шагу изменять только какие-нибудь две строки матрицы $M = \boxed{A\,|\,E}$, а все остальные строки оставлять без изменения. Умножение пары строк ϱ_1 и ϱ_2 слева на обратимую матрицу $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ приведёт к замене этих строк[1] на их линейные комбинации

$$\begin{pmatrix} \varrho_1 \\ \varrho_2 \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \varrho_1 \\ \varrho_2 \end{pmatrix} = \begin{pmatrix} a\varrho_1 + b\varrho_2 \\ c\varrho_1 + d\varrho_2 \end{pmatrix} .$$

Подчеркнём, что $ad - bc$ должно быть отлично от нуля, т. е. коэффициенты используемых двух линейных комбинаций не должны быть пропорциональны. Классический метод Гаусса из n° 7.4 на стр. 110 ограничивался тремя специальными типами таких преобразований, отвечающих умножению на обратимые 2×2 матрицы S вида:

1) $S = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}$ с $S^{-1} = \begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix}$ или $S = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ с $S^{-1} = \begin{pmatrix} 1 & -\lambda \\ 0 & 1 \end{pmatrix}$

2) $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ с $S^{-1} = S$

3) $S = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ с $S^{-1} = \begin{pmatrix} \lambda^{-1} & 0 \\ 0 & \mu^{-1} \end{pmatrix}$, где $\lambda, \mu \in \Bbbk$ отличны от нуля.

Следствие 8.1
Приведение матрицы $\boxed{A\,|\,E}$ к строгому ступенчатому виду методом Гаусса позволяет за конечное число шагов либо найти A^{-1}, либо убедиться, что A необратима. □

Пример 8.6
Выясним обратима ли матрица

$$A = \begin{pmatrix} 6 & 3 & -2 & 1 \\ 1 & 4 & 1 & 1 \\ 1 & 1 & 3 & -1 \\ -1 & 0 & -2 & 1 \end{pmatrix}$$

Для этого припишем к ней справа единичную матрицу и применим метод Гаусса

$$\left(\begin{array}{rrrr|rrrr} 6 & 3 & -2 & 1 & 1 & 0 & 0 & 0 \\ 1 & 4 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 3 & -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & -2 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

[1] соответственно, чтобы проделать такое преобразование с i-той и j-той строками $n \times 2n$-матрицы $\boxed{A\,|\,E}$, мы должны умножить эту матрицу слева на $n \times n$-матрицу S', содержащую 2×2-подматрицу S в пересечениях i-той и j-той строк с i-тым и j-тым столбцами и имеющую $s'_{kk} = 1$ при $k \neq i, j$ и нули в остальных местах

меняем знак нижней строки, потом меняем её местами с верхней

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 2 & -1 & 0 & 0 & 0 & -1 \\ 1 & 4 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 3 & -1 & 0 & 0 & 1 & 0 \\ 6 & 3 & -2 & 1 & 1 & 0 & 0 & 0 \end{array}\right)$$

зануляем первый столбец под первой строкой, отнимая из всех строк надлежащие кратности первой строки

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 2 & -1 & 0 & 0 & 0 & -1 \\ 0 & 4 & -1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 3 & -14 & 7 & 1 & 0 & 0 & 6 \end{array}\right)$$

меняем вторую и третью строки местами и зануляем нижние два элемента второго столбца

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 2 & -1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & -5 & 2 & 0 & 1 & -4 & -3 \\ 0 & 0 & -17 & 7 & 1 & 0 & -3 & 3 \end{array}\right) \tag{8-7}$$

Теперь, чтобы избежать вычислений с дробями, отклонимся от классического метода Гаусса и умножим нижние две строки на матрицу[1]

$$\begin{pmatrix} -5 & 2 \\ -17 & 7 \end{pmatrix}^{-1} = \begin{pmatrix} 7 & -2 \\ 17 & -5 \end{pmatrix}$$

Получим

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 2 & -1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 2 & -7 & 22 & 27 \\ 0 & 0 & 0 & 1 & 5 & -17 & 53 & 66 \end{array}\right)$$

Остаётся вычесть из 2-й строки 3-ю, а из 1-й — 4-ю и удвоенную 3-ю

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -3 & 9 & 11 \\ 0 & 1 & 0 & 0 & -2 & 7 & -21 & -26 \\ 0 & 0 & 1 & 0 & 2 & -7 & 22 & 27 \\ 0 & 0 & 0 & 1 & 5 & -17 & 53 & 66 \end{array}\right)$$

Итак, A обратима и

$$A^{-1} = \begin{pmatrix} 1 & -3 & 9 & 11 \\ -2 & 7 & -21 & -26 \\ 2 & -7 & 22 & 27 \\ 5 & -17 & 53 & 66 \end{pmatrix}$$

[1]что соответствует умножению всей матрицы слева на $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 7 & -2 \\ 0 & 0 & 17 & -5 \end{pmatrix}$

Пример 8.7 (решение системы линейных уравнений)
Система из n (неоднородных) линейных уравнений с n неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n = b_3 \\ \dots\dots\dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

в матричных обозначениях сворачивается до одного линейного уравнения $Ax = b$, где $A = (a_{ij})$ есть матрицы коэффициентов, а x и b суть матрицы-столбцы размеров $n \times 1$, представляющие собою столбец переменных и столбец правых частей. Если матрица A обратима, то решение задаётся формулой $x = A^{-1}b$, причём вместо поиска $A^{-1} = S_kS_{k-1} \cdots S_2S_1$ методом Гаусса, можно искать решение конкретной системы: поскольку $A^{-1}b = S_kS_{k-1} \cdots S_2S_1b$ получается применением к столбцу b той же цепочки преобразований, что приводит от A к E, преобразовав по Гауссу $n \times (n+1)$-матрицу $\boxed{A \mid b}$ к виду $\boxed{E \mid s}$, мы получаем в правом столбце решение s. Однако, если требуется искать решения многих уравнений с одной и той же матрицей A и меняющимися правыми частями, то может оказаться выгоднее всё-таки вычислить A^{-1}, а потом находить решения умножая правые части на A^{-1}.

8.4. Матрицы перехода. Пусть некий вектор v линейно выражается через какие-то ещё векторы w_i

$$v = \sum_{i=1}^{m} x_iw_i = w_1x_1 + w_2x_2 + \cdots + w_mx_m\,. \tag{8-8}$$

Организуем коэффициенты $x_i \in \Bbbk$ в матрицу-столбец размера $m \times 1$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \tag{8-9}$$

а векторы w_i — в матрицу-строку $w = (w_1, w_2, \dots, w_m)$ размера $1 \times n$ с элементами $w_i \in V$. Тогда формула (8-8) свернётся в матричное равенство $v = wx$, в котором v рассматривается как матрица размера 1×1 с элементом из V. Такая матричная запись позволяет упростить многие вычисления, связанные с линейным выражением одних векторов через другие.

Пусть, например, даны два набора векторов $u = (u_1, u_2, \dots, u_n)$, $w = (w_1, w_2, \dots, w_m)$ и пусть каждый из векторов u_j линейно выражен через векторы w_i

$$u_j = \sum_{\nu=1}^{m} c_{\nu j}w_\nu = w_1 \cdot c_{1j} + w_2 \cdot c_{2j} + \cdots + w_m \cdot c_{mj}\,.$$

Эти n равенств сокращённо записывается одной матричной формулой $u = w \cdot C_{wu}$, в которой $u = (u_1, u_2, \dots, u_n)$, $w = (w_1, w_2, \dots, w_m)$, а матрица

$$C_{wu} = \left(c_{ij}\right) = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{pmatrix} \tag{8-10}$$

получается подстановкой в матрицу u вместо каждого из векторов u_j столбца коэффициентов его линейного выражения через векторы w_i.

Матрица (8-10) называется *матрицей перехода* от векторов u к векторам w. Отметим, что столбец (8-9) коэффициентов линейного выражения вектора v через векторы u_j является частным случаем матрицы перехода: $x = C_{uv}$. Название «матрица перехода» вызвано тем, что C_{uw} позволяет переходить от линейных выражений векторов $v \in V$ через векторы u_j к их линейным выражениям через векторы w_i: $v = uC_{uv} \Rightarrow v = wC_{wu}C_{uv}$, т. е. произведение матрицы перехода от векторов u к векторам w и матрицы перехода от векторов v к векторам u является матрицей перехода от векторов v к векторам w:

$$C_{wu}C_{uv} = C_{wv}\,. \tag{8-11}$$

Замечание 8.1. Если набор векторов $w = (w_1, w_2, \dots, w_m)$ линейно зависим, то каждый вектор v из их линейной оболочки допускает много *различных* линейных выражений[1] через векторы w_j. Поэтому обозначение C_{wv} не корректно в том смысле, что элементы матрицы C_{wv} определяются по векторам w и v не однозначно. Тем не менее, равенство (8-11) содержательно и означает, что имея какие-нибудь линейные выражения C_{wu} и C_{uv} векторов u через v и векторов v через w, мы можем предъявить явное линейное выражение C_{wv} векторов u через w *перемножив матрицы* C_{wu} и C_{uv}.

Если набор векторов $e = (e_1, e_2, \dots, e_n)$ является базисом, то матрица перехода C_{ew}, выражающая произвольный набор векторов $w = (w_1, w_2, \dots, w_m)$ через базис e, однозначно определяется по e и w, и два набора векторов u и w совпадают тогда и только тогда, когда совпадают матрицы перехода $C_{eu} = C_{ew}$ от них к базису e.

Лемма 8.1
Пусть набор векторов $v = (v_1, v_2, \dots, v_n)$ образует базис пространства V. Для того, чтобы набор векторов $u = vC_{vu}$ тоже составлял базис, необходимо и достаточно, чтобы матрица C_{vu} была обратима, и в этом случае $C_{vu}^{-1} = C_{uv}$.

Доказательство. Если u базис, то векторы e линейно выражается через u и по (8-11) выполнены равенства $C_{ee} = C_{eu}C_{ue}$ и $C_{uu} = C_{ue}C_{eu}$. Так как каждый набор векторов (в том числе, и базис) имеет единственное выражение через базис, $C_{ee} = C_{uu} = E$, откуда $C_{ue}C_{eu} = C_{ue}C_{eu} = E$. Наоборот, если u не базис, то это линейно зависимая система векторов, и $u\lambda = 0$ для некоторого *ненулевого* столбца коэффициентов λ. Тогда $eC_{eu}\lambda = 0$, откуда $C_{eu}\lambda = 0$. Такое равенство невозможно с обратимым C_{eu} и ненулевым λ, поскольку умножение обеих частей слева на C_{eu}^{-1} даёт $\lambda = 0$. □

[1] как мы видели в n° 7.4 эти выражения представляют собою смежный класс подпространства линейных зависимостей $U \subset \Bbbk^m$ между векторами w_j

Пример 8.8 (замена координат при смене базиса)
Пусть набор векторов $w = (w_1, w_2, \dots, w_m)$ выражается через базис $e = (e_1, e_2, \dots, e_n)$ как $w = eC_{ew}$. Если $v = eC_{ev}$ — другой базис, то в выражении $w = vC_{vw}$ векторов w через базис v матрица $C_{vw} = C_{ve}C_{ew} = C_{ev}^{-1}C_{vw}$. В частности столбец координат произвольного вектора w в базисе v получаются из столбца его координат в базисе e умножением слева на матрицу C_{ev}^{-1}, обратную к матрице координат векторов базиса v в базисе e.

Пример 8.9 (замена матрицы оператора при смене базиса)
Для произвольных линейного оператора $F : U \to W$ и строки векторов $v = (v_1, v_2, \dots, v_r)$ будем обозначать через $F(v)$ строку значений оператора F на этих векторах

$$F(v) \stackrel{\text{def}}{=} \big(F(v_1),\ F(v_2),\ \dots\ ,\ F(v_r)\big)\,.$$

В силу линейности оператора F для любой числовой матрицы $M \in \mathrm{Mat}_{r\times s}(\Bbbk)$ выполняется равенство $F(vM) = F(v)M$.

Упражнение 8.8. Убедитесь в этом.

В таких обозначениях матрица F_{wu} оператора F, записанная в базисах u и w пространств U и W, однозначно определяется равенством[1] $F(u) = wF_{wu}$. При переходе к другим базисам $\widetilde{u} = uC_{u\widetilde{u}}$ и $\widetilde{w} = wC_{w\widetilde{w}}$ она меняется по правилу

$$F_{\widetilde{w}\widetilde{u}} = C_{w\widetilde{w}}^{-1}F_{wu}C_{u\widetilde{u}}\,. \tag{8-12}$$

ибо $F(\widetilde{u}) = F(uC_{u\widetilde{u}}) = F(u)\,C_{u\widetilde{u}} = w\,F_{wu}C_{u\widetilde{u}} = \widetilde{w}\,C_{\widetilde{w}w}F_{wu}C_{u\widetilde{u}} = \widetilde{w}\,C_{\widetilde{w}w}^{-1}F_{wu}C_{u\widetilde{u}}$.

В частности, если линейный эндоморфизм $F : V \to V$ задаётся матрицей $F_e = F_{ee}$, j-тый столбец которой есть столбец координат $F(e_j)$ в *том же самом* базисе e, то при замене базиса e на базис $u = eC_{eu}$ матрица оператора F в новом базисе будет равна

$$F_u = C_{eu}^{-1}F_eC_{eu}\,. \tag{8-13}$$

8.5. Некоммутативные кольца. Абелева группа R с операцией умножения $R \times R \to R$ называется *кольцом*, если умножение ассоциативно, т. е. $\forall\ f,g,h \in R\ \ f(gh) = (fg)h$ и двусторонне дистрибутивно, т. е. $\forall f,g,h \in R\ \ f(g+h) = fg + fh$ и $(f+g)h = fh + gh$. Если в кольце R существует элемент e, такой что $ef = fe = f$ для всех $f \in R$, этот элемент называется *единицей* и кольцо называется *кольцом с единицей*.

Упражнение 8.9. Покажите, что $0 \cdot f = 0$ для всех f в любом кольце R и что единичный элемент единственен (если существует).

Всякая (некоммутативная) алгебра является одновременно (некоммутативным) кольцом, так что рассмотренные выше алгебра эндоморфизмов векторного пространства и алгебра матриц с элементами из поля доставляют примеры некоммутативных колец. Последний из них можно обобщить.

[1] напомним (см. формулу (6-19) на стр. 93), что j-тый столбец матрицы F_{wu} есть столбец координат вектора $F(u_j)$ по базису w

8.5.1. Матрицы над некоммутативным кольцом. Квадратные $n \times n$-матрицы с элементами из произвольного кольца R образуют кольцо $\mathrm{Mat}_n(R)$, сложение и умножение в котором задаются теми же правилами, что и сложение и умножение матриц с элементами из поля: сумма $S = F + G$ и произведение $P = FG$ матриц $F = \left(f_{ij}\right)$ и $G = \left(g_{ij}\right)$ имеют в качестве матричных элементов

$$s_{ij} = f_{ij} + g_{ij} \quad \text{и} \quad p_{ij} = \sum_{\nu} f_{i\nu} g_{\nu j}$$

Упражнение 8.10. Проверьте выполнение свойств ассоциативности и дистрибутивности для умножения матриц с элементами из произвольного кольца.

Замечание 8.2. Вычисления с матрицами, элементы которых лежат в некоммутативном кольце отличаются от вычислений с матрицами, элементы которых лежат в поле, двумя существенными особенностями: сомножители в произведениях нельзя переставлять друг с другом (последствие некоммутативности) и не на все ненулевые элементы можно делить (последствие того, что не все элементы кольца обратимы).

Например, формула (8-4) перестаёт быть верной над некоммутативным кольцом, поскольку при её выводе мы переставили сомножители, когда выделили на побочной диагонали матрицы F^2 общий множитель $(a + d)$ — над некоммутативным кольцом этот множитель, вообще говоря, не выносится.

Аналогично, критерий обратимости матрицы размера 2×2 и формула (8-6) для обратной матрицы над некоммутативным кольцом, вообще говоря, неверны, а над коммутативным кольцом, не являющимся полем, нуждаются в уточнении: 2×2-матрица над коммутативным кольцом обратима тогда и только тогда, когда её определитель $\det F$ обратим, и если это так, то имеет место формула (8-6) для обратной матрицы.

Упражнение 8.11. Докажите последнее утверждение.

8.5.2. Примеры обратимых матриц 2×2. Матрица вида

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

с элементами из произвольного (некоммутативного) кольца R обратима тогда и только тогда, когда обратимы её диагональные элементы. В самом деле, из равенства

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} ax + bz & ay + bw \\ dz & dw \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

вытекает, что $dw = 1$ и $dz = 0$, откуда d обратим, а $w = d^{-1}$ и $z = 0$. Поэтому $ax = 1$, откуда a обратим, а $x = a^{-1}$. Тогда в правом верхнем углу получаем соотношение $ay + bd^{-1} = 0$, из которого $y = -a^{-1}bd^{-1}$. Таким образом,

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & -a^{-1}bd^{-1} \\ 0 & d^{-1} \end{pmatrix}$$

Аналогичные рассуждения показывают, что обратимость матрицы вида

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$$

равносильна обратимости диагональных элементов a, d, и в этом случае

$$\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & 0 \\ -d^{-1}ca^{-1} & d^{-1} \end{pmatrix}$$

Упражнение 8.12. Покажите, что матрицы $\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$ и $\begin{pmatrix} 0 & b \\ c & d \end{pmatrix}$ обратимы тогда и только тогда, когда обратимы оба элемента c и b, и в этом случае

$$\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & c^{-1} \\ b^{-1} & -b^{-1}ac^{-1} \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 0 & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} -c^{-1}db^{-1} & c^{-1} \\ b^{-1} & 0 \end{pmatrix}$$

Из проделанных вычислений вытекает, что гауссовы элементарные преобразования строк задаются умножениями на обратимые матрицы и, стало быть, могут применяться для обращения матриц методом Гаусса над произвольным некоммутативным кольцом с единицей.

8.5.3. Обратимость унитреугольных матриц Диагонали

$$\begin{pmatrix} * & & \\ & * & \\ & & * \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} & & * \\ & * & \\ * & & \end{pmatrix}$$

квадратной матрицы называются, соответственно, *главной* и *побочной*. Квадратная матрица называется *верхней* (соотв. *нижней*) *треугольной*, если у неё обращаются в нуль все элементы, стоящие под (соотв. над) *главной* диагональю.

Упражнение 8.13. Проверьте, что над любым (в том числе некоммутативным) кольцом R верхние и нижние треугольные матрицы составляют подкольца в $\mathrm{Mat}_n(R)$.

Если в кольце R есть единица, то треугольные матрицы с единицами на главной диагонали называются *унитреугольными*.

Лемма 8.2
Любая верхняя унитреугольная матрица $A = (a_{ij})$ над произвольным (в том числе, некоммутативным) кольцом с единицей обратима, причём $B = A^{-1}$ тоже верхняя унитреугольная с наддиагональными элементами

$$\begin{aligned} b_{ij} &= \sum_{s=0}^{j-i-1} (-1)^{s+1} \sum_{i<\nu_1<\dots<\nu_s<j} a_{i\nu_1}a_{\nu_1\nu_2}a_{\nu_2\nu_3} \dots a_{\nu_{s-1}\nu_s}a_{\nu_s j} = \\ &= -a_{ij} + \sum_{i<k<j} a_{ik}a_{kj} - \sum_{i<k<\ell<j} a_{ik}a_{k\ell}a_{\ell j} + \sum_{i<k<\ell<m<j} a_{ik}a_{k\ell}a_{\ell m}a_{mj} - \cdots . \end{aligned} \tag{8-14}$$

Доказательство. Прямое вычисление методом Гаусса. Для матрицы 4×4

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} & a_{14} \\ 0 & 1 & a_{23} & a_{24} \\ 0 & 0 & 1 & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

оно выглядит так: приписываем справа единичную матрицу

$$\left(\begin{array}{cccc|cccc} 1 & a_{12} & a_{13} & a_{14} & 1 & 0 & 0 & 0 \\ 0 & 1 & a_{23} & a_{24} & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & a_{34} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

зануляем 1-й столбец над главной диагональю используя 2-ю строку

$$\left(\begin{array}{cccc|cccc} 1 & 0 & a_{13}-a_{12}a_{23} & a_{14}-a_{12}a_{24} & 1 & -a_{12} & 0 & 0 \\ 0 & 1 & a_{23} & a_{24} & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & a_{34} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

зануляем 2-й столбец над главной диагональю используя 3-ю строку

$$\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & a_{14}-a_{12}a_{24}-a_{13}a_{34}+a_{12}a_{23}a_{34} & 1 & -a_{12} & -a_{13}+a_{12}a_{23} & 0 \\ 0 & 1 & 0 & a_{24}-a_{23}a_{34} & 0 & 1 & -a_{23} & 0 \\ 0 & 0 & 1 & a_{34} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

наконец, зануляем последний столбец, используя 4-ю строку, получая справа

$$A^{-1} = \begin{pmatrix} 1 & -a_{12} & -a_{13}+a_{12}a_{23} & -a_{14}+a_{12}a_{24}+a_{13}a_{34}-a_{12}a_{23}a_{34} \\ 0 & 1 & -a_{23} & -a_{24}+a_{23}a_{34} \\ 0 & 0 & 1 & -a_{23} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

В общем случае удобно нарисовать n различных точек $1, 2, \dots, n$ и воспринимать матричный элемент a_{ij} как стрелку, ведущую из j в i, а левое умножение на a_{ij} — как проход из j в i по этой стрелке. Тогда формула (8-14) гласит, что b_{ij} равен сумме всех маршрутов, ведущих из j в i, в которую все маршруты, состоящие из $s+1$ стрелок, входят со знаком $(-1)^{s+1}$. По индукции, умножая $n \times (2n)$-матрицу $\boxed{A \mid E}$ слева на матрицу

$$S = \begin{pmatrix} 1 & b_{12} & b_{13} & \dots & b_{1\,(n-1)} & 0 \\ 0 & 1 & b_{23} & \dots & b_{2\,(n-1)} & 0 \\ & & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & b_{(n-2)\,(n-1)} & 0 \\ 0 & \dots & \dots & 0 & 1 & 0 \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

в левом верхнем углу которой стоит матрица размера $(n-1)\times(n-1)$, обратная к верхней левой угловой подматрице в A, образованной первыми $(n-1)$ строками и столбцами, мы получим в последнем n-том столбце левой половины матрицы

$$S \cdot \boxed{A \mid E}$$

в позиции (i,n) сумму $a_{in} + b_{i2}a_{2n} + b_{i3}a_{3n} + \cdots + b_{i\,(n-1)}a_{(n-1)\,n}$ всех маршрутов, ведущих из n в i, в которую каждый маршрут длины $s+1$ входит со знаком $(-1)^s$. Обнуляя этот столбец методом Гаусса, получаем в n-м столбце правой половины матрицы требуемые значения b_{in}. □

Ответы и указания к некоторым упражнениям

Упр. 8.1. Первое доказывается выкладкой $0 \cdot a = (b + (-1) \cdot b)a = ba + (-1)ba = 0$, второе — выкладкой $e' = e' \cdot e'' = e''$.

Упр. 8.2. $E_{ij}E_{k\ell} = \begin{cases} E_{i\ell} & \text{при } j = k \\ 0 & \text{в остальных случаях} \end{cases}$. В частности, $E_{12}E_{21} \neq E_{21}E_{12}$. Полный список коммутационных соотношений таков:

$$[E_{ij}, E_{k\ell}] \overset{\text{def}}{=} E_{ij}E_{k\ell} - E_{k\ell}E_{ij} = \begin{cases} E_{ii} - E_{jj} & \text{при } j = k \text{ и } i = \ell \\ E_{i\ell} & \text{при } j = k \text{ и } i \neq \ell \\ -E_{kj} & \text{при } j \neq k \text{ и } i = \ell \\ 0 & \text{в остальных случаях.} \end{cases}$$

Упр. 8.3. Пусть $AB = C$, $B^tA^t = D$, тогда $c_{ij} = \sum\limits_k a_{ik}b_{kj} = \sum\limits_k a^t_{ki}b^t_{jk} = \sum\limits_k b^t_{jk}a^t_{ki} = d_{ji}$.

Упр. 8.7. По по теореме о ранге матрицы[1] линейная зависимость строк $n \times n$-матрицы A равносильна линейной зависимости её столбцов и означает, что размерность образа линейного оператора $A : \Bbbk^n \to \Bbbk^n$ с матрицей A меньше n. Поэтому $\ker A \neq 0$, и стало быть оператор A не биективен, а значит, не обратим.

Упр. 8.9. См. указания к упр. 8.1

Упр. 8.11. Легко видеть, что $\det(FG) = \det F \cdot \det G$. Поэтому, если матрица F обратима, то $\det F \cdot \det F^{-1} \det(FF^{-1}) = \det E = 1$, и тем самым $\det F$ обратим. То, что формула (8-6) при обратимом $\det F$ даёт обратную матрицу, устанавливается прямым вычислением.

Упр. 8.12. Можно воспользоваться тем, что

$$\begin{pmatrix} a & b \\ c & 0 \end{pmatrix} = \begin{pmatrix} b & a \\ 0 & c \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} = \begin{pmatrix} b & 0 \\ d & c \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

[1]см. сл. 7.4 на стр. 109