§7. Матрицы

Всюду в этом параграфе K по умолчанию обозначает коммутативное кольцо с единицей, а $\Bbbk$ — произвольное поле.

7.1. Матричный формализм. Таблица из m строк и n столбцов, заполненная элементами множества $\mathcal{A}$, называется *$m \times n$ матрицей с элементами из $\mathcal{A}$*. Множество всех таких матриц обозначается $\mathrm{Mat}_{m\times n}(\mathcal{A})$. Элемент матрицы A, расположенный в i-й строке и j-м столбце, обозначается a_{ij}. Запись $A = (a_{ij})$ означает, что матрица A состоит из таких элементов a_{ij}. Например, матрица $A \in \mathrm{Mat}_{3\times 4}(\mathbb{Z})$ с элементами $a_{ij} = i - j$ имеет вид

$$\begin{pmatrix} 0 & -1 & -2 & -3 \\ 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 \end{pmatrix}.$$

Если множество $\mathcal{A}$ является модулем над коммутативным кольцом K, то множество матриц $\mathrm{Mat}_{m\times n}(\mathcal{A})$ тоже является K-модулем относительно операций поэлементного сложения таблиц и умножения их на числа: сумма $S = (s_{ij})$ матриц $A = (a_{ij})$ и $B = (b_{ij})$ имеет $s_{ij} = a_{ij} + b_{ij}$, а матрица $P = \lambda A$, где $\lambda \in K$, имеет $p_{ij} = \lambda a_{ij}$. Таким образом, K-модуль $\mathrm{Mat}_{m\times n}(\mathcal{A}) \simeq \mathcal{A}^{mn}$ представляет собою прямую сумму mn экземпляров модуля $\mathcal{A}$, слагаемые которой выписаны не в строку, а в таблицу размера $m \times n$.

7.1.1. Умножение матриц. Пусть элементы K-модулей $\mathcal{A}$ и $\mathcal{B}$ можно билинейно перемножать со значениями в K-модуле $\mathcal{C}$, т. е. имеется отображение $\mathcal{A} \times \mathcal{B} \to \mathcal{C}$, $(a, b) \to ab$, линейное по каждому аргументу в том смысле, что

$$(x_1a_1 + x_2a_2)(y_1b_1 + y_2b_2) = x_1y_1a_1b_1 + x_1y_2a_1b_2 + x_2y_1a_2b_1 + x_2y_2a_2b_2$$

для всех $a_1, a_2 \in \mathcal{A}$, $b_1, b_2 \in \mathcal{B}$, $x_1, x_2, y_1, y_2 \in K$. Тогда для всех $m, s, n \in \mathbb{N}$ имеется *умножение матриц* $\mathrm{Mat}_{m\times s}(\mathcal{A}) \times \mathrm{Mat}_{s\times n}(\mathcal{B}) \to \mathrm{Mat}_{m\times n}(\mathcal{C})$, $(A, B) \mapsto AB$. Иными словами, произведение двух матриц определено, когда ширина левой матрицы равна с высоте правой, при этом матрица-произведение имеет столько же строк, сколько в левом сомножителе, и столько же столбцов, сколько в правом. При $m = n = 1$ результатом умножения строки ширины s на столбец высоты s является матрица размера 1×1, т. е. один элемент. Он определяется так:

$$(a_1, a_2, \dots, a_s)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{pmatrix} \overset{\text{def}}{=} a_1b_1 + a_2b_2 + \dots + a_sb_s = \sum_{k=1}^{s} a_kb_k\,. \tag{7-1}$$

Для произвольных m и n элемент c_{ij} матрицы $C = AB$ равен произведению i-й строки из A на j-й столбец из B, посчитанному по формуле (7-1):

$$c_{ij} = \left(a_{i1},\ a_{i2},\ \dots\ ,\ a_{is}\right) \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj} \end{pmatrix} = \sum_{k=1}^{s} a_{ik}b_{kj}\,. \tag{7-2}$$

Иначе можно сказать, что в j-том столбце матрицы AB стоит линейная комбинация s столбцов матрицы A с коэффициентами из j-го столбца матрицы B. Это описание получается, если подставить в формулу (7-1) в качестве элементов b_i — числа из j-го столбца матрицы B, а в качестве элементов a_j — столбцы матрицы A, интерпретируемые как векторы координатного модуля K^m с координатами, выписанными в столбик.

Упражнение 7.1. Удостоверьтесь, что это описание согласуется с формулой (7-2).

Например, для того, чтобы превратить матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \tag{7-3}$$

в матрицу из четырёх столбцов, равных, соответственно, сумме 1-го столбца матрицы A со 2-м, умноженным на λ, сумме 1-го и 3-го столбцов матрицы A, сумме 3-го столбца матрицы A со 2-м, умноженным на μ, и сумме всех трёх столбцов матрицы A, умноженных на их номера, надо умножить матрицу A справа на матрицу

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

Упражнение 7.2. Проверьте это прямым вычислением по формуле (7-2).

Симметричным образом, если в формуле (7-1) взять в в качестве элементов a_j числа из i-й строки матрицы A, а в качестве b_i — строки матрицы B, являющиеся векторами координатного модуля K^n с координатами, выписанными в строчку, то можно сказать, что i-й строкой матрицы AB является линейная комбинация строк матрицы B с коэффициентами, стоящими в i-й строке матрицы A. Например, если в той же матрице (7-3) 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 7.3. Проверьте это прямым вычислением по формуле (7-2).

Обратите внимание, что предыдущие два описания произведения AB получаются друг из друга заменой слова «столбец» на слово «строка» и наоборот с одновременной перестановкой букв A и B местами. Матрица $C^t = (c^t_{ij})$ размера $n\times m$, по строкам которой записаны столбцы $m\times n$ матрицы $C = (c_{ij})$, называется *транспонированной* к матрице C. Её элементы $c^t_{ij} = c_{ji}$ получаются отражением элементов матрицы C относительно биссектрисы левого верхнего угла матрицы.

Упражнение 7.4. Убедитесь, что для матриц с элементами из любого коммутативного кольца K выполняется равенство $(AB)^t = B^t A^t$, т. е. транспонирование обращает порядок сомножителей в произведениях матриц, элементы которых коммутируют друг с другом.

Упражнение 7.5. Убедитесь, что при наличии билинейного умножения K-модулей $\mathcal{A} \times \mathcal{B} \to \mathcal{C}$ произведение матриц $\mathrm{Mat}_{m\times s}(\mathcal{A}) \times \mathrm{Mat}_{s\times n}(\mathcal{B}) \to \mathrm{Mat}_{m\times n}(\mathcal{C})$ тоже билинейно, т. е.

$$(\lambda_1 A_1 + \lambda_2 A_2)(\mu_1 B_1 + \mu_2 B_2) = \lambda_1\mu_1 A_1 B_1 + \lambda_1\mu_2 A_1 B_2 + \lambda_2\mu_1 A_2 B_1 + \lambda_2\mu_2 A_2 B_2$$

для любых $A_1, A_2 \in \mathrm{Mat}_{m\times s}(\mathcal{A})$, $B_1, B_2 \in \mathrm{Mat}_{m\times s}(\mathcal{B})$, $\lambda_1, \lambda_2, \mu_1, \mu_2 \in K$.

Упражнение 7.6. Пусть заданы билинейные умножения K-модулей $\mathcal{A} \times \mathcal{B} \to \mathcal{P}$, $\mathcal{B} \times \mathcal{C} \to \mathcal{Q}$, $\mathcal{P} \times \mathcal{C} \to \mathcal{R}$ и $\mathcal{A} \times \mathcal{Q} \to \mathcal{R}$. Убедитесь, что если они ассоциативны, т. е. $(ab)c = a(bc)$ для всех $a \in \mathcal{A}$, $b \in \mathcal{B}$, $c \in \mathcal{C}$, то и произведение матриц

$$\mathrm{Mat}_{m\times s}(\mathcal{A}) \times \mathrm{Mat}_{s\times t}(\mathcal{B}) \times \mathrm{Mat}_{t\times n}(\mathcal{C}) \to \mathrm{Mat}_{m\times n}(\mathcal{R})$$

ассоциативно, т. е. $(AB)C = A(BC)$ для всех $A \in \mathrm{Mat}_{m\times s}(\mathcal{A})$, $B \in \mathrm{Mat}_{s\times t}(\mathcal{B})$, $C \in \mathrm{Mat}_{t\times n}(\mathcal{R})$.

Предостережение 7.1. Умножение матриц не коммутативно. Например, в $\mathrm{Mat}_{2x2}(\mathbb{Z})$

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

7.1.2. Матрицы перехода. Пусть в некотором K-модуле M заданы два набора векторов:

$$\boldsymbol{u} = (u_1, \dots, u_n) \quad \text{и} \quad \boldsymbol{w} = (w_1, \dots, w_m),$$

причём первый из них содержится в линейной оболочке второго, т. е. каждый вектор u_j имеет вид $u_j = w_1 \cdot c_{1j} + w_2 \cdot c_{2j} + \dots + w_m \cdot c_{mj}$, где $c_{ij} \in K$. Эти n равенств удобно собираются в одну матричную формулу $\boldsymbol{u} = \boldsymbol{w} \cdot C_{\boldsymbol{wu}}$, где $\boldsymbol{u} = (u_1, \dots, u_n)$ и $\boldsymbol{w} = (w_1, \dots, w_m)$ суть матрицы-строки с элементами из M, а матрица $C_{\boldsymbol{wu}} = \left(c_{ij}\right)$ получается подстановкой в матрицу $\boldsymbol{u}$ вместо каждого из векторов u_j столбца коэффициентов его линейного выражения через векторы w_i. Матрица $C_{\boldsymbol{wu}}$ называется *матрицей перехода* от векторов $\boldsymbol{u}$ к векторам $\boldsymbol{w}$. Название объясняется тем, что если имеется набор векторов $\boldsymbol{v} = (v_1, \dots, v_k)$, линейно выражающихся через векторы $\boldsymbol{u}$ по формулам $\boldsymbol{v} = \boldsymbol{u}C_{\boldsymbol{uv}}$, то выражение векторов $\boldsymbol{v}$ через векторы $\boldsymbol{w}$ задаётся матрицей

$$C_{\boldsymbol{wv}} = C_{\boldsymbol{wu}}C_{\boldsymbol{uv}}, \tag{7-4}$$

которая возникает при подстановке $\boldsymbol{u} = \boldsymbol{w}C_{\boldsymbol{wu}}$ в разложение $\boldsymbol{v} = \boldsymbol{u}C_{\boldsymbol{uv}}$. Таким образом, если записывать линейные выражения $v = u_1x_1 + \dots + u_nx_n = w_1y_1 + \dots + w_my_m$ произвольного вектора $v \in \mathrm{span}(u_1, \dots, u_n)$ через векторы $\boldsymbol{u}$ и $\boldsymbol{w}$ в виде $v = \boldsymbol{u}x = \boldsymbol{w}y$, где $x = (x_1, \dots, x_n)^t$ и $y = (y_1, \dots, y_m)^t$ суть столбцы коэффициентов, то эти столбцы будут связаны соотношением

$$y = C_{\boldsymbol{wu}}x.$$

Отметим, что когда набор векторов $\boldsymbol{w} = (w_1, \dots, w_m)$ линейно зависим, у каждого вектора v из их линейной оболочки имеется много *разных* линейных выражений через векторы w_j. Поэтому обозначение $C_{\boldsymbol{wv}}$ в этой ситуации не корректно в том смысле, что элементы матрицы $C_{\boldsymbol{wv}}$ определяются наборами векторов $\boldsymbol{w}$ и $\boldsymbol{v}$ не однозначно. Тем не менее, равенство (7-4) вполне осмысленно и означает, что имея какие-нибудь линейные выражения $C_{\boldsymbol{wu}}$ и $C_{\boldsymbol{uv}}$ векторов $\boldsymbol{u}$ через $\boldsymbol{w}$ и векторов $\boldsymbol{v}$ через $\boldsymbol{u}$, мы можем явно предъявить одно из линейных выражений $C_{\boldsymbol{wv}}$ векторов $\boldsymbol{v}$ через векторы $\boldsymbol{w}$, перемножив матрицы $C_{\boldsymbol{wu}}$ и $C_{\boldsymbol{uv}}$.

Если же набор векторов $\boldsymbol{e} = (e_1, \dots, e_n)$ является базисом, то матрица перехода $C_{\boldsymbol{ew}}$, выражающая произвольный набор векторов $\boldsymbol{w} = (w_1, \dots, w_m)$ через базис $\boldsymbol{e}$ однозначно определяется по наборам $\boldsymbol{e}$ и $\boldsymbol{w}$, т. е. два набора векторов $\boldsymbol{u}$, $\boldsymbol{w}$ совпадают если и только если выполняется равенство $C_{\boldsymbol{eu}} = C_{\boldsymbol{ew}}$.

7.1.3. Обратимые матрицы. В этом разделе мы рассматриваем квадратные $n \times n$ матрицы с элементами из коммутативного кольца K с единицей. Матрица

$$E = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_{n\times n}(K)\,,$$

по диагонали которой стоят единицы, а в остальных местах — нули, называется *единичной*.

Упражнение 7.7. Убедитесь, что $AE = A$ и $EA = A$ всякий раз, когда такие произведения определены.

Матрица $C \in \mathrm{Mat}_{n\times n}(K)$ называется *обратимой*, если существуют такие матрицы A и B, что $AC = E = CB$. В этом случае матрицы A и B автоматически равны друг другу, так как

$$A = AE = A(CB) = (AC)B = EB = B\,.$$

Это вычисление заодно показывает, что матрица $C^{-1} \stackrel{\text{def}}{=} A = B$ однозначно определяется по C свойством $C^{-1}C = CC^{-1} = E$. Матрица C^{-1}, если существует, называется *обратной* к C.

Упражнение 7.8. Докажите, что обратимость матрицы $C \in \mathrm{Mat}_{n\times n}(K)$ равносильна обратимости транспонированной к ней матрицы[1] C^t.

Предложение 7.1
Набор векторов $\boldsymbol{u} = (u_1, \dots, u_n)$ свободного K-модуля с базисом $\boldsymbol{e} = (e_1, \dots, e_n)$ является базисом если и только если матрица перехода[2] $C_{\boldsymbol{eu}} \in \mathrm{Mat}_{n\times n}(K)$ обратима, и тогда $C_{\boldsymbol{eu}}^{-1} = C_{\boldsymbol{ue}}$.

Доказательство. Пусть векторы $\boldsymbol{u}$ образуют базис. Так как каждый набор векторов имеет единственное выражение через базис, $C_{\boldsymbol{ue}}C_{\boldsymbol{eu}} = C_{\boldsymbol{uu}} = E$ и $C_{\boldsymbol{eu}}C_{\boldsymbol{ue}} = C_{\boldsymbol{ee}} = E$ по формуле (7-4). Тем самым, $C_{\boldsymbol{ue}} = C_{\boldsymbol{eu}}^{-1}$. Наоборот, если матрица $C_{\boldsymbol{eu}}$ обратима, то умножая обе части равенства $\boldsymbol{u} = \boldsymbol{e}C_{\boldsymbol{eu}}$ справа на $C_{\boldsymbol{eu}}^{-1}$, мы получаем линейное выражение $\boldsymbol{e} = \boldsymbol{u}C_{\boldsymbol{eu}}^{-1}$ базиса $\boldsymbol{e}$ через векторы $\boldsymbol{u}$ и заключаем, что последние линейно порождают весь модуль. Если существует линейное соотношение $\boldsymbol{u}x = 0$, где $x \in K^n$ — столбец коэффициентов, то $\boldsymbol{e}C_{\boldsymbol{eu}}x = 0$, откуда $C_{\boldsymbol{eu}}x = 0$. Умножая обе части слева на $C_{\boldsymbol{eu}}^{-1}$, получаем $x = 0$. Тем самым, векторы $\boldsymbol{u}$ линейно независимы и образуют базис по лем. 6.1 на стр. 85. □

Пример 7.1 (замена координат при смене базиса)
Пусть набор векторов $\boldsymbol{w} = (w_1, \dots, w_m)$ выражается через базис $\boldsymbol{e} = (e_1, \dots, e_n)$ как $\boldsymbol{w} = \boldsymbol{e}C_{\boldsymbol{ew}}$. Если $\boldsymbol{u} = \boldsymbol{e}C_{\boldsymbol{eu}}$ — другой базис, то выражение векторов $\boldsymbol{w}$ через базис $\boldsymbol{u}$ имеет вид $\boldsymbol{w} = \boldsymbol{e}C_{\boldsymbol{ew}} = \boldsymbol{u}C_{\boldsymbol{eu}}^{-1}C_{\boldsymbol{ew}}$, т. е. $C_{\boldsymbol{uw}} = C_{\boldsymbol{eu}}^{-1}C_{\boldsymbol{ew}}$. В частности, если вектор $v = \boldsymbol{e}x$ имеет в базисе $\boldsymbol{e}$ столбец координат x, то в базисе $\boldsymbol{u} = \boldsymbol{e}C_{\boldsymbol{eu}}$ он имеет столбец координат $y = C_{\boldsymbol{eu}}^{-1}x$

Следствие 7.1
Следующие условия на квадратную матрицу $A \in \mathrm{Mat}_{n\times n}(\Bbbk)$ с элементами из поля $\Bbbk$ эквивалентны:

1) матрица A обратима

[1] См. упр. 7.4 на стр. 94.
[2] См. n° 7.1.2 на стр. 95.

2) столбцы матрицы A линейно независимы

3) столбцы матрицы A линейно порождают координатное пространство $\Bbbk^n$,

и то же самое верно с заменой столбцов на строки.

Доказательство. Обозначим через $a_1, \dots, a_n$ столбцы матрицы A, воспринимаемые как векторы координатного пространства $\Bbbk^n$. Матрица A является матрицей перехода от этих векторов к стандартному базису пространства $\Bbbk^n$. По предл. 7.1 обратимость матрицы A равносильна тому, что векторы a_i образуют в $\Bbbk^n$ базис, что по сл. 6.3 на стр. 88 равносильно каждому из условий (2), (3). Самое последнее утверждение вытекает из упр. 7.8 на стр. 96. □

Пример 7.2 (обратимые 2×2-матрицы над коммутативным кольцом)
Возводя матрицу

$$C = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_{2\times 2}(K)$$

в квадрат, получим

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

откуда

$$(a + d) \cdot C - C^2 = \begin{pmatrix} (ad - bc) & 0 \\ 0 & (ad - bc) \end{pmatrix} = (ad - bc) \cdot E\,. \tag{7-5}$$

Число $\det C \stackrel{\text{def}}{=} ad - bc$ называется *определителем* 2×2-матрицы C.

Упражнение 7.9. Докажите для любых $A, B \in \mathrm{Mat}_{2x2}(K)$ равенство $\det(AB) = \det(A) \cdot \det(B)$.

Из упражнения вытекает, что определитель любой обратимой матрицы обратим в K, поскольку вычисляя определитель обеих частей матричного равенства $C \cdot C^{-1} = E$, получаем

$$\det(C) \cdot \det(C^{-1}) = \det E = 1\,.$$

С другой стороны, если $\det C = ad - bc$ обратим в K, то равенство (7-5) переписывается как

$$C \cdot \big((a + d)E - C\big) \cdot (ad - bc)^{-1} = E\,.$$

Тем самым, матрица C обратима и

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\det C}\big((a + d)E - C\big) = (ad - bc)^{-1} \begin{pmatrix} a & -b \\ -c & d \end{pmatrix}. \tag{7-6}$$

Итак, 2×2 матрица обратима если и только если обратим её определитель.

7.1.4. Матрицы линейных отображений. Пусть K-модули N и M линейно порождаются наборами векторов $\boldsymbol{u} = (u_1, \dots, u_n)$ и $\boldsymbol{w} = (w_1, \dots, w_m)$ соответственно. Всякое K-линейное отображение $F : N \to M$ однозначно задаётся набором своих значений

$$F(\boldsymbol{u}) \stackrel{\text{def}}{=} \big(F(u_1), \dots, F(u_n)\big)\,, \tag{7-7}$$

на порождающих векторах и действует на произвольный вектор $v = \boldsymbol{u}x$, где $x \in K^n$ — столбец коэффициентов линейного выражения вектора v через образующие $\boldsymbol{u}$, по правилу

$$F(\boldsymbol{u}x) = F\Big(\sum_{i=1}^{n} u_i x_i\Big) = \sum_{i=1}^{n} F(u_i) \cdot x_i = F(\boldsymbol{u})x\,. \tag{7-8}$$

Матрица перехода от векторов (7-7) к образующим $\boldsymbol{w}$ модуля M обозначается

$$F_{\boldsymbol{wu}} = C_{\boldsymbol{w},F(\boldsymbol{u})} \in \mathrm{Mat}_{m\times n}(K)$$

и называется *матрицей отображения* F в образующих $\boldsymbol{w}$ и $\boldsymbol{u}$. В её j-м стоят коэффициенты линейного выражения вектора $F(u_j)$ через векторы $\boldsymbol{w}$. Согласно (7-8) произвольный вектор $\boldsymbol{u}x$ со столбцом коэффициентов x переводится отображением F в вектор $\boldsymbol{w}F_{\boldsymbol{wu}}x$ со столбцом коэффициентов $F_{\boldsymbol{wu}}x$. Из (7-8) также вытекает, что для любого набора векторов $\boldsymbol{v} = (v_1, \dots , v_k)$ в N, любой матрицы $A \in \mathrm{Mat}_{\ell\times k}(K)$ и любого K-линейного отображения $F : N \to M$ выполняется равенство $F(\boldsymbol{v}A) = F(\boldsymbol{v})A$. Если K-модуль L порождается векторами $\boldsymbol{v} = (v_1, \dots , v_\ell)$ и K-линейные отображения $F : N \to L$ и $G : L \to M$ имеют матрицы $F_{\boldsymbol{vu}}$ и $G_{\boldsymbol{wv}}$, соответственно, в образующих $\boldsymbol{v}$, $\boldsymbol{u}$ и в образующих $\boldsymbol{w}$, $\boldsymbol{v}$, то композиция $H = GF : N \to M$ имеет в образующих $\boldsymbol{w}$, $\boldsymbol{u}$ матрицу $H_{\boldsymbol{wu}} = G_{\boldsymbol{wv}}F_{\boldsymbol{vu}}$, ибо $H(\boldsymbol{u}) = G\big(F(\boldsymbol{u})\big) = G\big(\boldsymbol{v}F_{\boldsymbol{vu}}\big) = G(\boldsymbol{v})F_{\boldsymbol{vu}} = \boldsymbol{w}G_{\boldsymbol{wv}}F_{\boldsymbol{vu}}$.

Отметим, что когда образующие $\boldsymbol{w}$ линейно зависимы, то как и в n° 7.1.2, матрица $F_{\boldsymbol{wu}}$ линейного отображения F определяется образующими $\boldsymbol{w}$ и $\boldsymbol{u}$ не однозначно, так как набор векторов $F(\boldsymbol{u})$ имеет много разных линейных выражений через векторы $\boldsymbol{w}$. Предыдущие формулы означают при этом, что если задано какое-то выражение $v = \boldsymbol{u}x$ вектора v через образующие $\boldsymbol{u}$, то столбец коэффициентов $y = F_{\boldsymbol{wu}}x$ даёт одно из возможных линейных выражений $F(v) = \boldsymbol{w}y$ вектора $F(v)$ через образующие $\boldsymbol{w}$, и что получить одну из возможных матриц для композиции отображений можно перемножив какие-нибудь из матриц этих отображений в том же порядке, в каком берётся композиция.

Также важно понимать, что когда образующие $\boldsymbol{u}$ линейно зависимы, матрица $F_{\boldsymbol{wu}}$ не может быть произвольной: для любого линейного соотношения $\boldsymbol{u}x = 0$ между векторами $\boldsymbol{u}$ в N в модуле M должно выполняться соотношение $\boldsymbol{w}F_{\boldsymbol{wu}}x = 0$. Иными словами, если модули $M = K^n / R_M$ и $M = K^m / R_N$ заданы при помощи образующих и соотношений, как в прим. 6.12 на стр. 86, то матрица $A \in \mathrm{Mat}_{m\times n}(K)$ тогда и только тогда является матрицей некоторого линейного отображения $F : N \to M$, когда для любого столбца $x \in R_N$ столбец $Ax \in R_M$. Это матричная переформулировка предл. 6.3 на стр. 84 в обозначениях из прим. 6.12.

Если же модули N и M оба свободны и наборы векторов $\boldsymbol{u}$ и $\boldsymbol{w}$ являются их базисами, то сопоставление K-линейному отображению $F : N \to M$ его матрицы $F_{\boldsymbol{wu}}$ в этих базисах задаёт K-линейный изоморфизм $\mathrm{Hom}_K(N, M) ⥲ \mathrm{Mat}_{m\times n}(K)$, $F \mapsto F_{\boldsymbol{wu}}$.

Упражнение 7.10. Убедитесь, что сопоставление отображению его матрицы линейно.

В частности, для свободных K-модулей N и M конечного ранга модуль $\mathrm{Hom}_K(N, M)$ тоже свободен и $\mathrm{rk}\,\mathrm{Hom}_K(N, M) = \mathrm{rk}\, N \cdot \mathrm{rk}\, M$.

Пример 7.3 (замена матрицы линейного отображения при смене базиса)

Если K-линейный гомоморфизм свободных модулей $F : N \to M$ имеет в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ матрицу $F_{\boldsymbol{wu}}$, то он переводит векторы $\boldsymbol{e} = \boldsymbol{u}C_{\boldsymbol{ue}}$ любого другого базиса $\boldsymbol{e}$ в N в векторы

$$F(\boldsymbol{e}) = F(\boldsymbol{u}C_{\boldsymbol{ue}}) = F(\boldsymbol{u})C_{\boldsymbol{ue}} = \boldsymbol{w}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}}\,.$$

Если выбрать в M другой базис $\boldsymbol{f}$, через который исходный базис $\boldsymbol{w}$ выражается по формуле $\boldsymbol{w} = \boldsymbol{f}C_{\boldsymbol{fw}}$, и подставить это выражение в предыдущую формулу вместо $\boldsymbol{w}$, мы получим, что $F(\boldsymbol{e}) = \boldsymbol{f}C_{\boldsymbol{fw}}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}}$. Таким образом, матрица $F_{\boldsymbol{fe}}$ отображения F в базисах $\boldsymbol{e}$ и $\boldsymbol{f}$ выражается через матрицу $F_{\boldsymbol{wu}}$ того же отображения в базисах $\boldsymbol{u}$ и $\boldsymbol{w}$ по формулам

$$F_{\boldsymbol{fe}} = C_{\boldsymbol{fw}}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}} = C_{\boldsymbol{wf}}^{-1}F_{\boldsymbol{wu}}C_{\boldsymbol{ue}} = C_{\boldsymbol{fw}}F_{\boldsymbol{wu}}C_{\boldsymbol{eu}}^{-1}\,. \tag{7-9}$$

Пример 7.4 (матрица линейного эндоморфизма)
Линейный эндоморфизм $F : M \to M$ модуля M, порождённого векторами $\boldsymbol{w} = (w_1, \dots, w_m)$, обычно принято записывать квадратной матрицей $F_{\boldsymbol{w}} \stackrel{\text{def}}{=} F_{\boldsymbol{ww}}$, в j-м столбце которой стоят коэффициенты линейного выражения вектора $F(w_j)$ через тот же самый набор образующих $\boldsymbol{w}$. Эта матрица называется *матрицей эндоморфизма* F в образующих $\boldsymbol{w}$. Если векторы $\boldsymbol{w}$ составляют базис модуля M, то при переходе к другому базису $\boldsymbol{e} = \boldsymbol{u}C_{\boldsymbol{ue}}$ матрица эндоморфизма поменяется по формулам (7-9):

$$F_{\boldsymbol{e}} = C_{\boldsymbol{ew}}F_{\boldsymbol{w}}C_{\boldsymbol{we}} = C_{\boldsymbol{we}}^{-1}F_{\boldsymbol{w}}C_{\boldsymbol{we}} = C_{\boldsymbol{ew}}F_{\boldsymbol{w}}C_{\boldsymbol{ew}}^{-1}\,. \tag{7-10}$$

7.1.5. Ранг матрицы над полем. В этом разделе мы рассматриваем матрицы с элементами из произвольного поля $\Bbbk$. Размерность линейной оболочки столбцов матрицы $A \in \mathrm{Mat}_{m\times n}(\Bbbk)$ в координатном векторном пространстве $\Bbbk^m$ называется *рангом* матрицы A и обозначается $\operatorname{rk} A$. Каждая матрица A задаёт линейное отображение $F_A : \Bbbk^n \to \Bbbk^m$, $x \mapsto Ax$, которое переводит координатный столбец $x \in \Bbbk^n$ в координатный столбец $F_A(x) = Ax \in \Bbbk^m$. В стандартных базисах[1] $\boldsymbol{e}$ и $\boldsymbol{f}$ координатных пространств $\Bbbk^n$ и $\Bbbk^m$ матрица $F_{\boldsymbol{fe}}$ оператора F_A совпадает матрицей A. Поэтому линейная оболочка столбцов матрицы A представляет собою образ оператора F_A. Тем самым, $\operatorname{rk} A = \dim \operatorname{im} F_A$.

Лемма 7.1
Ранг матрицы не меняется при умножении на обратимые матрицы слева или справа.

Доказательство. Если матрицы $D \in \mathrm{Mat}_{n\times n}(\Bbbk)$ и $C \in \mathrm{Mat}_{m\times m}(\Bbbk)$ обратимы, то матрица CAD является матрицей описанного выше оператора $F_A : \Bbbk^n \to \Bbbk^m$, $x \mapsto Ax$, но не в стандартных базисах $\boldsymbol{e}$ и $\boldsymbol{f}$ координатных пространств $\Bbbk^n$ и $\Bbbk^m$, а в новых базисах $\boldsymbol{u} = \boldsymbol{e}D$ и $\boldsymbol{w} = \boldsymbol{f}C^{-1}$. В самом деле, $F_{\boldsymbol{wu}} = C_{\boldsymbol{wf}}F_{\boldsymbol{fe}}C_{\boldsymbol{eu}} = CAD$ по формуле (7-9). Тем самым, размерность образа линейного оператора F_A равна размерности линейной оболочки столбцов матрицы CAD. □

Следствие 7.2
Размерность линейной оболочки строк произвольной матрицы A тоже не меняется при умножении матрицы A слева или справа на любые обратимые матрицы.

Доказательство. Применим лем. 7.1 к транспонированной матрице A^t. □

Теорема 7.1 (теорема о ранге матрицы)
Для любой матрицы $A \in \mathrm{Mat}_{m\times n}(\Bbbk)$ выполняется равенство $\operatorname{rk} A = \operatorname{rk} A^t$. Иными словами, линейная оболочка строк матрицы A в координатном пространстве $\Bbbk^n$ и линейная оболочка столбцов матрицы A в координатном пространстве $\Bbbk^m$ имеют равные размерности.

[1] См. прим. 6.12 на стр. 86 и в частности формулу (6-10).

Доказательство. Рассмотрим задаваемое матрицей A линейное отображение

$$F_A : \Bbbk^n \to \Bbbk^m , \quad x \mapsto Ax ,$$

выберем в $\Bbbk^n$ базис $\boldsymbol{u} = (u_1, \dots , u_r, u_{r+1}, \dots , u_n)$ так, чтобы его векторы $u_{r+1}, \dots , u_n$ составили базис в $\ker F_A$. В доказательстве предл. 6.6 на стр. 90 мы видели, что векторы $w_j = F_A(u_j)$, где $1 \leqslant j \leqslant r$, образуют в этом случае базис в $\operatorname{im} F_A$, так что $r = \dim \operatorname{im} F_A = \operatorname{rk} A$. Дополним эти векторы w_j до базиса $\boldsymbol{w} = (w_1, \dots , w_m)$ всего пространства $\Bbbk^m$. Матрица $F_{\boldsymbol{wu}} = \big(f_{ij}\big)$ оператора F_A в базисах $\boldsymbol{w}$ и $\boldsymbol{u}$ пространств $\Bbbk^m$ и $\Bbbk^n$ имеет $f_{ii} = 1$ при $1 \leqslant i \leqslant r$ и нули во всех остальных местах. В частности, линейная оболочка её строк в координатном пространстве $\Bbbk^n$ и линейная оболочка её столбцов в координатном пространстве $\Bbbk^m$ имеют одну и ту же размерность r. Согласно прим. 7.3 матрица $F_{\boldsymbol{wu}} = C_{\boldsymbol{wf}} A C_{\boldsymbol{eu}}$ получается из матрицы $A = F_{\boldsymbol{fe}}$ оператора F_A в стандартных базисах $\boldsymbol{e}$ и $\boldsymbol{f}$ пространств $\Bbbk^n$ и $\Bbbk^m$ умножением слева и справа на обратимые матрицы переходов $C_{\boldsymbol{wf}}$ и $C_{\boldsymbol{eu}}$. Согласно лем. 7.1 и сл. 7.2 такое умножение не меняет размерностей линейных оболочек строк и столбцов матрицы A. □

7.2. Ассоциативные алгебры над полем. Векторное пространство A над полем $\Bbbk$ называется *алгеброй* над $\Bbbk$ или $\Bbbk$-*алгеброй*, если на нём имеется такое умножение $A \times A \to A$, что произведение ab линейно по a при фиксированном b, и линейно по b при фиксированном a, т. е.

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

для всех $a_1, a_2, b_1, b_2 \in A$ и $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \Bbbk$. Алгебра A называется *ассоциативной*, если

$$\forall\, a, b, c \in A \quad (ab)c = a(bc) ,$$

и *коммутативной* — если $ab = ba$ для всех $a, b \in A$. Ассоциативная алгебра, в которой есть нейтральный элемент по отношению к умножению, т. е. такой $e \in A$, что $ea = ae = a$ для всех $a \in A$, называется *алгеброй с единицей*.

Упражнение 7.11. 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 $\Bbbk[x_1, \dots , x_n]$ и прочие коммутативные $\Bbbk$-алгебры в смысле прим. 5.5 на стр. 70.

7.2.1. Алгебра матриц. Модельными примерами некоммутативных ассоциативных алгебр с единицами являются алгебры квадратных матриц $\operatorname{Mat}_n(\Bbbk) \overset{\text{def}}{=} \operatorname{Mat}_{n\times n}(\Bbbk)$ и алгебры

$$\operatorname{End}_{\Bbbk}(V) \overset{\text{def}}{=} \operatorname{Hom}_{\Bbbk}(V, V)$$

линейных эндоморфизмов $V \to V$ векторных пространств V над полем $\Bbbk$. Если $\dim V = n$, то каждый базис $\boldsymbol{e} = (e_1, \dots , e_n)$ задаёт линейный изоморфизм $V \simeq \Bbbk^n$, переводящий вектор $v = \boldsymbol{e}x \in V$ в столбец $x \in \Bbbk^n$ его координат в базисе $\boldsymbol{e}$, а также изоморфизм алгебр[1]

$$\operatorname{End}_{\Bbbk}(V) \simeq \operatorname{Mat}_n(\Bbbk) , \quad F \mapsto F_{\boldsymbol{e}} , \tag{7-11}$$

переводящий линейное отображение $F : V \to V$ в его матрицу[2] $F_{\boldsymbol{e}}$ в базисе $\boldsymbol{e}$. Эти два изоморфизма согласованы в том смысле, что отображение F переводит вектор v со столбцом координат x в вектор Fv со столбцом координат $F_{\boldsymbol{e}}x$.

[1]Т. е. перестановочный с умножением в алгебре $\Bbbk$-линейный изоморфизм векторных пространств.

[2]См. n° 7.1.4 выше, в частности прим. 7.4 на стр. 99.

Стандартный базис матричной алгебры составляют матрицы E_{ij}, единственным ненулевым элементом которых является единица, стоящая в i-й строке и j-м столбце. Произвольная матрица $A = (a_{ij})$ линейно выражается через них по формуле $A = \sum_{i,j} a_{ij}E_{ij}$. Прообразами базисных матриц E_{ij} при изоморфизме (7-11) являются линейные операторы $E_{ij} : V \to V$, которые мы будем обозначать теми же буквами и которые действуют на базисные векторы e_k пространства V по правилам

$$E_{ij}(e_k) = \begin{cases} e_i & \text{при } k = j \\ 0 & \text{при } k \neq j. \end{cases}$$

Отсюда немедленно получается таблица умножения базисных элементов E_{ij}:

$$E_{ik}E_{\ell j} = \begin{cases} E_{ij} & \text{при } k = \ell \\ 0 & \text{при } k \neq \ell, \end{cases} \tag{7-12}$$

из которой лишний раз видно, что алгебра некоммутативна (скажем, $E_{12}E_{21} \neq E_{21}E_{12}$).

Упражнение 7.12. Составьте таблицу *коммутаторов* $[E_{ik}, E_{\ell j}] \overset{\text{def}}{=} E_{ik}E_{\ell j} - E_{\ell j}E_{ik}$.

Пример 7.5
Вычислим A^{2020} для матрицы $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Поскольку $A = E + E_{12}$ и матрицы E и E_{12} коммутируют, вычислить $(E + E_{12})^{2020}$ можно по обычной формулой бинома[1]. А так как $E_{12}^n = 0$ при $n > 1$, мы получаем

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{2020} = (E + E_{12})^{2020} = E + 2020\, E_{12} = \begin{pmatrix} 1 & 2020 \\ 0 & 1 \end{pmatrix}.$$

Упражнение 7.13. Покажите, что $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ при всех $n \in \mathbb{Z}$.

7.2.2. Обратимые элементы. Элемент a алгебры A с единицей $e \in A$ называется *обратимым*, если существует такой элемент $a^{-1} \in A$, 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 $a'a = aa'' = e$, ибо они автоматически совпадут друг с другом[2]: $a' = a'e = a'(aa'') = (a'a)a'' = ea'' = a''$. Эта выкладка заодно показывает, что обратный к a элемент a^{-1} однозначно определяется по a равенствами $aa^{-1} = a^{-1}a = e$.

7.2.3. Алгебраические и трансцендентные элементы. Каждый ненулевой элемент a любой ассоциативной $\Bbbk$-алгебры A с единицей задаёт ненулевой гомоморфизм вычисления

$$\mathrm{ev}_a : \Bbbk[t] \to A, \quad f(x) \mapsto f(a), \tag{7-13}$$

переводящий многочлен $f(x) = f_0 + f_1x + \dots + f_mx^m$ в элемент $f(a) = f_0e + f_1a + \dots + f_ma^m \in A$ — результат подстановки в f значения[3] $x = a$. Если ядро $\ker \mathrm{ev}_a = 0$, элемент a называется *трансцендентным* над $\Bbbk$. В этом случае гомоморфизм (7-13) инъективен, и все целые неотрицательные степени элемента a линейно независимы над $\Bbbk$. В частности, алгебра A бесконечномерна как векторное пространство над $\Bbbk$.

[1] См. формулу (1-9) на стр. 8.
[2] Ср. с n° 7.1.3 на стр. 96.
[3] При этом мы считаем, что $f_0 = f_0x^0$ и $a^0 \overset{\text{def}}{=} e$.

Если ядро гомоморфизма (7-13) ненулевое, элемент a называется *алгебраическим* над $\Bbbk$. В этом случае $\ker \mathrm{ev}_a$ является ненулевым собственным главным идеалом[1] в $\Bbbk[x]$. Порождающий его многочлен со старшим коэффициентом 1 обозначается $\mu_a(x)$ и называется *минимальным многочленом* элемента a. Он однозначно характеризуется как приведённый многочлен наименьшей степени, для которого $\mu_a(a) = 0$, и делит все многочлены, аннулирующие элемент a. Если алгебра A конечномерна как векторное пространство над $\Bbbk$, то все её элементы алгебраичны над $\Bbbk$. В частности, любая квадратная матрица конечного размера и любой линейный эндоморфизм конечномерного векторного пространства удовлетворяют некоторому полиномиальному уравнению.

Пример 7.6 (аннулирующий многочлен матрицы)
Поскольку $\dim_{\Bbbk} \mathrm{Mat}_n(\Bbbk) = n^2$, матрицы A^k, где $0 \leqslant k \leqslant n^2$, линейно зависимы над $\Bbbk$ для любой матрицы $A \in \mathrm{Mat}_n(\Bbbk)$. Это означает, что каждая $n \times n$ матрица удовлетворяет нетривиальному полиномиальному уравнению степени не выше n^2. Вскоре мы увидим[2], что эта априорная оценка степени сильно завышена, и степень минимального многочлена любой $n \times n$ матрицы в действительности не превышает n. Для матриц размера 2×2 это видно из прим. 7.2 на стр. 97: полученная там формула (7-5) утверждает, что матрица

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

удовлетворяет квадратному уравнению $x^2 - (a+b)\,x + (ad - bc) = 0$.

7.2.4. Нильпотентные элементы. Элемент a алгебры A называется *нильпотентным*, если $a^n = 0$ для некоторого $n \in \mathbb{N}$. Каждый нильпотентный элемент a корректно задаёт аналогичный (7-13) гомоморфизм вычисления $\mathrm{ev}_a : \Bbbk[\![x]\!] \to A$, $f(x) \mapsto f(a)$, подставляющий элемент a вместо переменной x в формальные степенные ряды. В частности, для такого элемента a определены элементы e^a, $\ln(1+a)$ и $(1+a)^s$ с произвольным $s \in \Bbbk$, которые удовлетворяют в алгебре A всем алгебраическим соотношениям, что имеются между рядами e^x, $\ln(1+x)$ и $(1+x)^s$ в кольце $\Bbbk[\![x]\!]$. Например, элемент $b = (1+a)^{1/2} \in A$ имеет $b^2 = 1 + a$.

Упражнение 7.14. Предъявите такую рациональную 3×3 матрицу B, что $B^3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

7.3. Некоммутативные кольца. Абелева группа R с операцией умножения $R \times R \to R$ называется *кольцом*, если это умножение ассоциативно, т. е. $f(gh) = (fg)h$ для всех $f, g, h \in R$, двусторонне дистрибутивно по отношению к сложению, т. е. $f(g+h) = fg + fh$ и $(f+g)h = fh + gh$ для всех $f, g, h \in R$, и существует такой элемент $1 \in R$, что $1 \cdot f = f \cdot 1 = f$ для всех $f \in R$. Элемент 1 называется *единицей* кольца R.

Упражнение 7.15. Покажите, что $0 \cdot f = 0$ для всех $f \in R$ и что единица единственна.

Алгебры $\mathrm{Mat}_n(\Bbbk)$ и $\mathrm{End}_{\Bbbk}(V)$ являются примерами некоммутативных колец. Первый из этих примеров допускает значительное обобщение, а именно, квадратные $n \times n$ матрицы с элементами из любого кольца R тоже образуют кольцо $\mathrm{Mat}_n(R)$ относительно операций сложения и

[1] Напомню, что $\Bbbk[x]$ является кольцом главных идеалов, см. n° 5.3 на стр. 70.

[2] См. ?? на стр. ??.

умножения матриц, определённых в самом начале этого параграфа[1]. А именно, сумма $S = F+G$ и произведение $P = FG$ матриц $F = \left(f_{ij}\right)$ и $G = \left(g_{ij}\right)$ имеют матричными элементами

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

Упражнение 7.16. Убедитесь, что умножение в $\mathrm{Mat}_n(R)$ ассоциативно и дистрибутивно по отношению к сложению, а матрицы E_{ij}, единственным ненулевым элементом которых является единица, стоящая в i-й строке и j-м столбце, перемножаются по форм. (7-12) на стр. 101, и единичная матрица $E = \sum E_{ii}$ является единицей кольца $\mathrm{Mat}_n(R)$.

Вычисления с матрицами, элементы которых лежат в некоммутативном кольце, требуют большей осторожности, чем вычисления с матрицами, элементы которых можно переставлять друг с другом в произведениях. Например, ключевая формула (7-5) из прим. 7.2 на стр. 97 перестаёт быть верной для матриц над некоммутативным кольцом, как и полученные в прим. 7.2 критерий обратимости и формула для обратной матрицы к матрице 2×2.

Упражнение 7.17. Убедитесь в этом.

Пример 7.7 (примеры обратимых матриц 2×2)
Покажем, что матрица

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

с элементами из произвольного[2] кольца R обратима если и только если обратимы её диагональные элементы a и d. 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

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

[1]См. n° 7.1 на стр. 93.

[2]В том числе некоммутативного.

Упражнение 7.18. Покажите, что матрицы $\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$ и $\begin{pmatrix} 0 & b \\ c & d \end{pmatrix}$ обратимы если и только если обратимы оба элемента c, b, и в этом случае

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

Пример 7.8 (обратимость унитреугольных матриц)
Диагональ, идущая из левого верхнего угла квадратной матрицы в правый нижний, называется *главной*. Если все стоящие под (соотв. над) главной диагональю элементы нулевые, матрица называется *верхней* (соотв. *нижней*) *треугольной*.

Упражнение 7.19. Проверьте, что верхние и нижние треугольные матрицы являются подкольцами[1] в кольце $\mathrm{Mat}_n(R)$ для любого кольца R.

Треугольные матрицы с единицами на главной диагонали называются *унитреугольными*. Покажем, что каждая верхняя унитреугольная матрица $A = (a_{ij})$ обратима в кольце $\mathrm{Mat}_n(R)$ для любого кольца R, и обратная к A матрица $B = A^{-1}$ тоже верхняя унитреугольная с наддиагональными элементами

$$\begin{aligned} b_{ij} &= \sum_{s=0}^{j-i-1} (-1)^{s+1} \sum_{i<\nu_1<\ldots<\nu_s<j} a_{i\nu_1} a_{\nu_1\nu_2} a_{\nu_2\nu_3} \,\ldots\, a_{\nu_{s-1}\nu_s} a_{\nu_s j} = \\ &= -a_{ij} + \sum_{i<k<j} a_{ik}a_{kj} - \sum_{i<k<\ell<j} a_{ik}a_{k\ell}a_{\ell j} + \sum_{i<k<\ell<m<j} a_{ik}a_{k\ell}a_{\ell m}a_{mj} - \cdots . \end{aligned} \tag{7-14}$$

Для этого запишем матрицу A в виде линейной комбинации матриц E_{ij} из упр. 7.16 выше[2]

$$A = E + \sum_{i<j} a_{ij}E_{ij} = E + N \,,$$

где матрица $N = \sum_{i<j} a_{ij}E_{ij}$ представляет собою наддиагональную часть матрицы A. В силу форм. (7-12) на стр. 101 и упр. 7.16, коэффициент при E_{ij} в матрице N^k равен[3] нулю при $j-i < k$, а при $j - i \geqslant k$ представляет собою сумму всевозможных произведений

$$\underbrace{a_{i\nu_1} \cdot a_{\nu_1\nu_2} \cdot \,\cdots\, \cdot a_{\nu_{k-2}\nu_{k-1}} \cdot a_{\nu_{k-1}j}}_{k\text{ сомножителей}} \,, \quad \text{где} \quad i < \nu_1 < \nu_2 < \ldots \nu_{k-1} < j \,.$$

В частности $N^k = 0$ при всех k, больших размера матрицы A. Полагая $x = E$, $y = N$ в равенстве[4]

$$(x+y)(x^{m-1} - x^{m-2}y + \cdots + (-1)^{m-2}xy^{m-2} + (-1)^{m-1}y^{m-1}) = x^m - y^m \,,$$

[1] Т. е. замкнуты относительно сложения и умножения.

[2] См. также форм. (7-12) на стр. 101.

[3] Продуктивно представлять себе E_{ij} как стрелку, ведущую из числа j в число i на числовой прямой. Произведение k сомножителей E_{ij} отлично от нуля если и только если конец каждой стрелки совпадает с началом предыдущей, и в этом случае такое произведение равно сумме всех перемножаемых стрелок, рассматриваемых как целочисленные векторы на числовой прямой. Таким образом, каждое ненулевое произведение k стрелок имеет длину как минимум k, а разложения элемента E_{ij} в произведение k таких элементов находятся в биекции со всевозможными способами пройти из j в i за k шагов.

[4] Поскольку матрицы E и N коммутируют друг с другом, в результате этой подстановки мы получим верное матричное равенство.

при достаточно большом m мы получим матричное равенство $A(E - N + N^2 - \cdots) = E$, откуда

$$A^{-1} = E - N + N^2 - N^3 + \dots\ ,$$

что и утверждалось.

Упр. 7.4. 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. 7.6. Пусть $AB = P$, $BC = Q$, тогда (i,j)-е элементы произведений PC и AQ равны друг другу: $\sum_k p_{ik}c_{kj} = \sum_k \sum_\ell (a_{i\ell}b_{\ell k})c_{kj} = \sum_{k\ell} a_{i\ell}(b_{\ell k}c_{kj}) = \sum_\ell a_{i\ell} \sum_k b_{\ell k}c_{kj} = \sum_\ell a_{i\ell}q_{\ell j}$.

Упр. 7.8. Поскольку $(AB)^t = B^tA^t$, матрица B обратна матрице A если и только если матрица B^t обратна матрице A^t.

Упр. 7.9. Прямое вычисление:

$$\begin{aligned}(a_{11}b_{11} + a_{12}b_{21})(a_{21}b_{12} + a_{22}b_{22}) - (a_{11}b_{12} + a_{12}b_{22})(a_{21}b_{11} + a_{22}b_{21}) = \\ = (a_{11}a_{22} - a_{12}a_{21})(b_{11}b_{22} - b_{12}b_{21})\,.\end{aligned}$$

Упр. 7.11. Первое доказывается выкладкой $0 \cdot a = (b + (-1) \cdot b)a = ba + (-1)ba = 0$, второе — выкладкой $e' = e' \cdot e'' = e''$.

Упр. 7.12. Ответ:

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

Упр. 7.15. См. указания к упр. 7.11

Упр. 7.18. Можно воспользоваться тем, что

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