

Глава IV. Векторные пространства и системы линейных уравнений

1. СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Системы алгебраических уравнений. Пусть k – поле. Алгебраическим уравнением над k с неизвестными x_1, \dots, x_n называется выражение вида

$$f(x_1, \dots, x_n) = 0,$$

где $f(x_1, \dots, x_n)$ – многочлен из $k[x_1, \dots, x_n]$. Отметим, что в уравнении знак равенства не имеет обычного содержательного смысла: он не означает, что слева и справа от него стоит один и тот же многочлен. Набор из нескольких алгебраических уравнений над одним и тем же полем с одними и теми же неизвестными называется системой алгебраических уравнений.

Пусть $f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)$ – многочлены с коэффициентами из поля k ; решением системы алгебраических уравнений

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0, \\ &\dots, \\ f_m(x_1, \dots, x_n) &= 0. \end{aligned}$$

называется набор (a_1, \dots, a_n) элементов из поля k , таких что

$$\begin{aligned} f_1(a_1, \dots, a_n) &= 0, \\ &\dots, \\ f_m(a_1, \dots, a_n) &= 0. \end{aligned}$$

Здесь знаки равенства уже имеют обычный содержательный смысл: для любого i значение $f_i(a_1, \dots, a_n)$ многочлена $f_m(x_1, \dots, x_n)$ принадлежит полю k , и требуется, чтобы этот элемент поля k равнялся 0.

Две системы алгебраических уравнений над одним и тем же полем с одними и теми же неизвестными называются равносильными, если всякое решение первой системы является решением второй системы, а всякое решение второй системы является решением первой системы. Следующее предложение дает достаточный признак равносильности двух систем алгебраических уравнений.

Предложение 1. Пусть

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0, & g_1(x_1, \dots, x_n) &= 0, \\ &\dots, & &\dots, \\ f_m(x_1, \dots, x_n) &= 0, & g_p(x_1, \dots, x_n) &= 0 \end{aligned}$$

две системы алгебраических уравнений над полем k . Если существует такое натуральное число $N \geq 1$, что N -е степени всех многочленов f_1, \dots, f_m принадлежат идеалу кольца $k[x_1, \dots, x_n]$, порожденному многочленами g_1, \dots, g_p , и наоборот, N -е степени всех многочленов g_1, \dots, g_p принадлежат идеалу кольца $k[x_1, \dots, x_n]$, порожденному многочленами f_1, \dots, f_m , то эти системы равносильны.

Доказательство. Пусть (a_1, \dots, a_n) – решение второй системы. Для любого i , $1 \leq i \leq m$, многочлен f_i^N принадлежит идеалу, порожденному g_1, \dots, g_p ; поэтому существуют такие многочлены $h_1, \dots, h_p \in k[x_1, \dots, x_n]$, что

$$(f_i(x_1, \dots, x_n))^N = h_1(x_1, \dots, x_n)g_1(x_1, \dots, x_n) + \dots + h_p(x_1, \dots, x_n)g_p(x_1, \dots, x_n).$$

Тогда

$$\begin{aligned} (f_i(a_1, \dots, a_n))^N &= h_1(a_1, \dots, a_n)g_1(a_1, \dots, a_n) + \dots + h_p(a_1, \dots, a_n)g_p(a_1, \dots, a_n) = \\ &= h_1(a_1, \dots, a_n) \cdot 0 + \dots + h_p(a_1, \dots, a_n) \cdot 0 = 0, \end{aligned}$$

Замечание. Если поле k алгебраически замкнуто, то это достаточное условие равносильности двух систем алгебраических уравнений является и необходимым. Доказательство этого утверждения, известного как "Теорема Гильберта о нулях", гораздо сложнее.

(2) замена одного из уравнений системы $f_i(x_1, \dots, x_n) = 0$ на уравнение $f_i(x_1, \dots, x_n) + h(x_1, \dots, x_n)f_j(x_1, \dots, x_n) = 0$, где $f_j(x_1, \dots, x_n) = 0$ – другое (не то же самое!) уравнение системы, а $h(x_1, \dots, x_n) \in k[x_1, \dots, x_n]$.

Системы линейных уравнений. Система уравнений, левые части которых являются многочленами полной степени не больше 1, называется системой линейных уравнений. Поскольку всякий многочлен из $k[x_1, \dots, x_n]$ полной степени не больше 1 имеет вид $a_1x_1 + \dots + a_nx_n - b$, системами линейных уравнений над полем k являются системы

$$a_{m,1}x_1 + \cdots + a_{m,n}x_n - b_m = 0,$$

$$a_{m1}x_1 + \cdots + a_{mn}x_n = b_m.$$
$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}.$$
$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

$$A \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Теорема Крамера. Пусть

– система n линейных уравнений с n неизвестными, и пусть A – матрица этой системы. Если матрица A обратима, то система имеет единственное решение

Доказательство. Для доказательства существования решения достаточно проверить, что столбец, указанный в формулировке теоремы, действительно представляет собой решение системы. Но это непосредственно следует из ассоциативности умножения матриц:

Теперь докажем единственность этого решения. Пусть столбец $(a_1, \dots, a_n)^T$ — произвольное решение нашей системы; тогда

т.е. любое решение системы совпадает с тем, которое указано в формулировке теоремы.

(1) умножение обеих частей одного из уравнений на элемент $a \neq 0$ из поля k ;

(2) прибавление к обеим частям одного из уравнений соответствующих частей другого уравнения, умноженных на элемент $a \in k$.

Добавим сюда еще одно преобразование, которое не меняет саму систему уравнений, но меняет матрицу и расширенную матрицу системы, которые зависят от того, в каком порядке записаны уравнения:

(3) перемена местами двух уравнений системы.

Таким образом, элементарные преобразования – это преобразования, переводящие систему

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

в систему

$$DA \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = D \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

где D – произвольная элементарная матрица. Отсюда сразу видно, что при элементарном преобразовании над системой линейных уравнений ее матрица и ее расширенная матрица претерпевают элементарные преобразования над строками, описываемые теми же самыми элементарными матрицами. Но мы знаем, что элементарными преобразованиями над строками расширенной матрицы системы ее можно сделать ступенчатой матрицей. Поэтому при помощи элементарных преобразований (которые, напомним, не меняют множество решений системы) любую систему линейных уравнений можно свести к равносильной ей системе вида

$$\begin{aligned} x_{j_1} + a'_{1,j_1+1}x_{j_1+1} + \dots + 0 \cdot x_{j_2} + \dots + 0 \cdot x_{j_r} + \dots + a'_{1n}x_n &= b'_1, \\ x_{j_2} + \dots + 0 \cdot x_{j_r} + \dots + a'_{2n}x_n &= b'_2, \\ &\dots\dots\dots \\ x_{j_r} + \dots + a'_{rn}x_n &= b'_r, \\ 0 &= 0, \\ &\dots\dots\dots \\ 0 &= 0 \end{aligned}$$

(здесь $1 \leq j_1 < j_2 < \dots < j_r \leq n$), или же вида

$$\begin{aligned} x_{j_1} + a'_{1,j_1+1}x_{j_1+1} + \dots + 0 \cdot x_{j_2} + \dots + 0 \cdot x_{j_r} + \dots + a'_{1n}x_n &= 0, \\ x_{j_2} + \dots + 0 \cdot x_{j_r} + \dots + a'_{2n}x_n &= 0, \\ &\dots\dots\dots \\ x_{j_r} + \dots + a'_{rn}x_n &= 0, \\ 0 &= 1, \\ &\dots\dots\dots \\ 0 &= 0. \end{aligned}$$

Во втором случае система не имеет решений, так как ни при каких значениях неизвестных значение многочлена 0 не окажется равным 1. В первом случае, наоборот, решения есть; более того, всем неизвестным x_j , кроме $x_{j_1}, x_{j_2}, \dots, x_{j_r}$, можно придать произвольные значения $\alpha_j \in k$, и тогда значения α_{j_s} неизвестных x_{j_s} определятся из системы уже однозначно:

$$\alpha_{j_s} = b'_s - \sum_{\substack{j > j_s \\ j \neq j_{s+1}, \dots, j_r}} a'_{sj} \alpha_j.$$

Если $r = n$, то "свободных" неизвестных не будет, и у системы будет лишь одно решение.

Итак, система линейных уравнений может вообще не иметь решений, или иметь единственное решение (как в случае, описываемом теоремой Крамера), или иметь много решений, зависящих от $n - r$ параметров, принимающих произвольные значения (здесь n – число неизвестных, а r – число ненулевых уравнений ступенчатой системы, равносильной исходной системе). Для того, чтобы разобраться в том, от чего зависит характер множества решений системы, нам придется воспользоваться некоторыми основополагающими фактами линейной алгебры, изложению которых посвящены следующие параграфы.

2. ВЕКТОРНЫЕ ПРОСТРАНСТВА

Определение векторного пространства. Пусть k – поле, а V – множество, на котором определено сложение, а также умножение на элементы из поля k . Таким образом, каждому элементам $u, v \in V$ сопоставляется их сумма $u + v \in V$, а каждому элементам $v \in V$, $a \in k$ – произведение $av \in V$. Множество V называется векторным (или линейным) пространством над полем k , если выполняются следующие свойства:

- (1) $u + (v + w) = (u + v) + w$ для любых $u, v, w \in V$;
- (2) $u + v = v + u$ для любых $u, v \in V$;
- (3) существует такой элемент $\theta \in V$, что $v + \theta = v$ для любого $v \in V$;
- (4) для всякого $v \in V$ существует такой элемент $-v \in V$, что $v + (-v) = \theta$;
- (5) $a(u + v) = au + av$ для любых $a \in k$, $u, v \in V$;
- (6) $(a + b)v = av + bv$ для любых $a, b \in k$, $v \in V$;
- (7) $(ab)v = a(bv)$ для любых $a, b \in k$, $v \in V$;
- (8) $1 \cdot v = v$ для любого $v \in V$ (здесь 1 означает единицу поля k).

Элементы векторного пространства называются векторами. Первые 4 свойства показывают, что относительно сложения векторное пространство является абелевой группой. Вектор θ из аксиомы 3 называется нулевым вектором, а вектор $-v$ из аксиомы 4 – противоположным к v вектором. Мы уже упоминали, что во всякой абелевой группе нулевой и противоположный элементы определены однозначно; поскольку доказательства тривиальны и не занимают много места, и поскольку мы часто будем использовать еще два столь же тривиальных утверждения (о первом из них мы уже говорили), приведем здесь эти доказательства.

Предложение 1. Пусть V – векторное пространство над полем k .

- (1) Если θ, θ_1 – такие векторы из V , что $v + \theta = v + \theta_1 = v$ для любого вектора $v \in V$, то $\theta = \theta_1$.
- (2) Если $u_1, u_2, v \in V$ таковы, что $u_1 + v = u_2 + v = \theta$, то $u_1 = u_2$.
- (3) $a\theta = \theta$ для любого $a \in k$.
- (4) $0 \cdot v = \theta$ для любого вектора $v \in V$.
- (5) $(-1) \cdot v = -v$ для любого вектора $v \in V$.

Доказательство. (1) Взяв в качестве v сначала вектор θ , а затем вектор θ_1 , получим: $\theta = \theta + \theta_1$, $\theta_1 = \theta_1 + \theta$, откуда и следует, что $\theta = \theta_1$.

(2) $u_1 = u_1 + \theta = u_1 + (v + u_2) = (u_1 + v) + u_2 = \theta + u_2 = u_2$.

(3) Пользуясь дистрибутивностью умножения по векторному сомножителю (аксиома 5), получаем:

$$a\theta = a\theta + \theta = a\theta + (a\theta + (-a\theta)) = (a\theta + a\theta) + (-a\theta) = a(\theta + \theta) + (-a\theta) = a\theta + (-a\theta) = \theta.$$

(4) Пользуясь аксиомой 6, получаем:

$$\begin{aligned} 0 \cdot v &= 0 \cdot v + \theta = 0 \cdot v + (0 \cdot v + (-0 \cdot v)) = (0 \cdot v + 0 \cdot v) + (-0 \cdot v) = \\ &= (0 + 0)v + (-0 \cdot v) = 0 \cdot v + (-0 \cdot v) = \theta. \end{aligned}$$

(5) По аксиомам (8) и (6) и по уже доказанному свойству (4), имеем:

$$\begin{aligned}(-1)v &= (-1)v + \theta = (-1)v + (v + (-v)) = ((-1)v + 1 \cdot v) + (-v) = \\ &= (-1 + 1)v + (-v) = 0 \cdot v + (-v) = \theta + (-v) = -v.\end{aligned}$$

Примеры векторных пространств. Почти, все что изучается в математике, происходит в различных векторных пространствах. Приведем лишь несколько примеров.

1. Изучающиеся в геометрии векторы на плоскости и в пространстве составляют векторные пространства над полем вещественных чисел \mathbb{R} . Мы будем называть их "геометрическими" векторными пространствами.

2. В поле вещественных чисел \mathbb{R} определено сложение и умножение на рациональные числа; из аксиом поля, в частности, вытекает, что в \mathbb{R} выполняются все требования, которым должно удовлетворять векторное пространство над полем рациональных чисел \mathbb{Q} . Совершенно аналогично, любое поле и даже тело может рассматриваться как векторное пространство над содержащимся в нем полем. Например, поле комплексных чисел \mathbb{C} и тело кватернионов \mathbb{H} являются векторными пространствами над каждым из полей \mathbb{Q} , \mathbb{R} , \mathbb{C} .

3. Множество всех непрерывных на отрезке $[0, 1]$ функций – векторное пространство над \mathbb{R} .

4. Пусть k – поле; тогда кольцо многочленов $k[x]$ – векторное пространство над полем k .

5. Пусть k – поле, а m, n – натуральные числа; тогда множество $k^{m \times n}$ всех матриц с компонентами из k является векторным пространством над полем k (и первые 8 свойств действий над матрицами как раз и являются аксиомами векторного пространства).

Из пространств $k^{m \times n}$ особенно важным является пространство столбцов $k^{m \times 1}$; оно называется арифметическим векторным пространством размерности m и обозначается k^m . Слова "размерности m " здесь не имеют никакого содержательного смысла (по крайней мере, пока), и их следует рассматривать как часть очень длинного имени пространства k^m . В дальнейшем, когда мы дадим определение размерности пространства, мы докажем, что размерность "арифметического векторного пространства размерности m " равна m .

Линейные комбинации. Пусть V – векторное пространство над полем k . Линейной комбинацией векторов $u_1, \dots, u_n \in V$ с коэффициентами $\alpha_1, \dots, \alpha_n \in k$ называется вектор $\alpha_1 u_1 + \dots + \alpha_n u_n \in V$. Удобно считать, что линейной комбинацией пустого набора векторов является нулевой вектор. Такое определение естественно, потому что линейная комбинация n векторов является суммой n слагаемых, так что линейная комбинация пустого набора должна быть суммой 0 слагаемых, а такую сумму естественно считать нулевым вектором. Таким образом, нулевой вектор является линейной комбинацией векторов из любого конечного набора векторов.

Нам часто будет удобно пользоваться следующим простым утверждением, которое можно было бы назвать "транзитивностью линейных комбинаций".

Лемма 1. Пусть v_1, \dots, v_m – какие-то векторы из векторного пространства V , и пусть u_1, \dots, u_n – линейные комбинации векторов из S (не обязательно различные). Тогда всякая линейная комбинация векторов u_1, \dots, u_n является и линейной комбинацией векторов v_1, \dots, v_m .

Доказательство. Пусть $w = \alpha_1 u_1 + \dots + \alpha_n u_n$, где $\alpha_1, \dots, \alpha_n \in k$. Каждый из векторов u_i , $1 \leq i \leq n$, является линейной комбинацией векторов v_1, \dots, v_m и потому представляется в виде $u_i = \beta_{i1} v_1 + \dots + \beta_{im} v_m$, где $\beta_{ij} \in k$ для всех j ,

$1 \leq j \leq m$. Поэтому вектор

$$v = \sum_{i=1}^n \alpha_i \sum_{j=1}^m \beta_{ij} v_j = \sum_{j=1}^m \left(\sum_{i=1}^n \alpha_i \beta_{ij} \right) v_j$$

представляет собой линейную комбинацию векторов v_1, \dots, v_m .

Подпространства. Порождающие множества. Пусть V – векторное пространство над полем k . Непустое подмножество U пространства V называется подпространством V , если для любых векторов $u, u' \in U$ и для любого $\alpha \in k$ векторы $u + u'$, αu принадлежат U . Если u – любой элемент из непустого множества U , то векторы $\theta = 0 \cdot u$, $-u = (-1)u$ принадлежит U . Таким образом, любое подпространство содержит нулевой вектор θ и вместе с каждым вектором u содержит также и противоположный вектор $-u$.

Любое подпространство U пространства V само является пространством над полем k . Действительно, при сложении векторов из U и при умножении вектора из U на элемент $\alpha \in k$ мы снова получаем векторы из U . Мы уже отметили, что в U выполняются аксиомы 3 и 4 определения векторного пространства; остальные аксиомы представляют собой равенства, которые выполняются во всем пространстве V , и тем более в его части U .

Часто бывает полезно немного ослабить или, наоборот, усилить требования, определяющие подпространство.

Предложение 2. Пусть V – векторное пространство над полем k , и пусть U – его непустое подмножество. Следующие условия равносильны.

- (1) U – подпространство V .
- (2) Для любых $u, u' \in U$ и любого $\alpha \in k$ вектор $\alpha u + u'$ принадлежит U .
- (3) Любая линейная комбинация векторов из U принадлежит U .

Доказательство. (1) \Rightarrow (3). Если U – подпространство V , то для любых векторов $u_1, \dots, u_n \in U$ и любых $\alpha_1, \dots, \alpha_n \in k$ векторы $\alpha_1 u_1, \dots, \alpha_n u_n$ принадлежат U , а тогда и их сумма $\alpha_1 u_1 + \dots + \alpha_n u_n$ принадлежит U .

(3) \Rightarrow (2). Если любая линейная комбинация векторов из U принадлежит U , то, в частности, и линейная комбинация $\alpha u + u' = \alpha u + 1 \cdot u'$ векторов $u, u' \in U$ принадлежит множеству U .

(2) \Rightarrow (1). Пусть $u \in U$; тогда по (2) элемент $\theta = (-1)u + u$ принадлежит U . Далее, опять по (2) элемент $\alpha u = \alpha u + \theta$ принадлежит U . Наконец, полагая в (2) $\alpha = 1$, получаем, что $u + u' = 1 \cdot u + u'$ принадлежит U для любых $u, u' \in U$. Итак, если выполнено условие (2), то U – подпространство V .

Пусть u_1, \dots, u_n – элементы векторного пространства V над полем k ; их линейной оболочкой называется множество всех их линейных комбинаций векторов из S . Мы будем обозначать линейную оболочку векторов $u_1, \dots, u_n \in V$ через $\langle u_1, \dots, u_n \rangle$. В частности, $\langle \emptyset \rangle = \{\theta\}$. Заметим, что для любого i , $1 \leq i \leq n$ вектор $u_i = 1 \cdot u_i$ является линейной комбинацией векторов u_1, \dots, u_n и потому $u_i \in \langle u_1, \dots, u_n \rangle$.

Предложение 3. Линейная оболочка любых векторов u_1, \dots, u_n из векторного пространства V является подпространством V .

Доказательство. Пусть $u, u' \in \langle u_1, \dots, u_n \rangle$, $\alpha \in k$; тогда линейная комбинация $\alpha u + u'$ линейных комбинаций u, u' векторов u_1, \dots, u_n сама является, по транзитивности линейных комбинаций, линейной комбинацией векторов u_1, \dots, u_n и потому принадлежит $\langle u_1, \dots, u_n \rangle$. По предложению 2 отсюда следует, что $\langle u_1, \dots, u_n \rangle$ – подпространство V .

Мы говорим, что векторы u_1, \dots, u_n порождают V , если $\langle u_1, \dots, u_n \rangle = V$. Отметим, что если к семейству векторов, порождающих V , добавить еще несколько векторов, то это расширенное семейство векторов тоже будет порождать V . Если в пространстве V найдется конечный набор векторов, который порождает это пространство, то пространство V называется конечно порожденным. В частности, линейная оболочка $\langle v_1, \dots, v_m \rangle$ любых векторов $v_1, \dots, v_m \in V$ порождается этими векторами и потому является конечно порожденным пространством (хотя само пространство V может и не быть конечно порожденным).

Линейно зависимость и независимость. Пусть V – векторное пространство над полем k . Векторы $v_1, \dots, v_n \in V$ называются линейно независимыми, если ни один из векторов v_i , $1 \leq i \leq n$, не является линейной комбинацией остальных векторов $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ из этого семейства. Векторы $v_1, \dots, v_n \in V$ называются линейно зависимыми, если они не являются линейно независимыми.

Заметим, что по этому определению пустое семейство векторов линейно независимо: ни один из векторов, принадлежащих ему, не является линейной комбинацией остальных векторов, потому что принадлежащих пустому множеству векторов попросту нет. Семейство, состоящее из единственного вектора $v \in V$, линейно независимо тогда и только тогда, когда этот вектор не представляет собой линейную комбинацию пустого множества векторов, т.е. когда $v \neq 0$. Если $v_1 = v_2$, то векторы v_1, v_2 линейно зависимы, поскольку $v_2 = 1 \cdot v_1$ – линейная комбинация вектора v_1 .

Заметим еще, что если векторы v_1, \dots, v_n линейно независимы, то любая часть этого семейства векторов тоже линейно независима; точнее говоря, это означает, что если $1 \leq i_1 < i_2 < \dots < i_m \leq n$, то векторы $v_{i_1}, v_{i_2}, \dots, v_{i_m}$ линейно независимы. В частности, если среди векторов v_1, \dots, v_n есть нулевой вектор или два одинаковых вектора, то векторы v_1, \dots, v_n линейно зависимы.

Легко понять, что в пространстве "геометрических" векторов два вектора линейно независимы тогда и только тогда, когда они не коллинеарны, а три вектора линейно независимы тогда и только тогда, когда они не компланарны.

Линейная комбинация $\alpha_1 v_1 + \dots + \alpha_n v_n$ векторов $v_1, \dots, v_n \in V$ называется тривиальной, если все коэффициенты $\alpha_1, \dots, \alpha_n$ равны 0. Тривиальной мы будем считать и линейную комбинацию пустого подмножества множества векторов. Ясно, что всякая тривиальная комбинация векторов равна нулевому вектору θ . Оказывается, обратное утверждение характеризует линейно независимые семейства векторов.

Предложение 4. Пусть V – векторное пространство над полем k . Векторы $v_1, \dots, v_n \in V$ линейно независимы тогда и только тогда, когда из того, что их линейная комбинация равна нулевому вектору θ , следует, что эта линейная комбинация тривиальна.

Доказательство. Мы будем доказывать следующее утверждение, очевидно, равносильное утверждению предложения: векторы $v_1, \dots, v_n \in V$ линейно зависимы тогда и только тогда, когда существует их нетривиальная линейная комбинация, равная нулевому вектору.

Если векторы $v_1, \dots, v_n \in V$ линейно зависимы, то найдется такое i , $1 \leq i \leq n$, что вектор v_i является линейной комбинацией векторов $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$. Это значит, что существуют элементы $\alpha_1, \dots, \alpha_{i-1}, \alpha_{i+1}, \dots, \alpha_n \in k$, такие что

$$v_i = \alpha_1 v_1 + \dots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \dots + \alpha_n v_n.$$

Тогда $\alpha_1 v_1 + \dots + \alpha_{i-1} v_{i-1} + (-1)v_i + \alpha_{i+1} v_{i+1} + \dots + \alpha_n v_n = \theta$, и эта линейная комбинация не тривиальна, потому что коэффициент при v_i равен $-1 \neq 0$.

Обратно, пусть $\alpha_1 v_1 + \dots + \alpha_i v_i + \dots + \alpha_n v_n$ – нетривиальная линейная комбинация векторов $v_1, \dots, v_n \in V$, равная нулевому вектору. Тогда среди коэффициентов этой комбинации есть хотя бы один ненулевой коэффициент α_i , и вектор

$$v_i = (-\alpha_1/\alpha_i)v_1 + \dots + (-\alpha_{i-1}/\alpha_i)v_{i-1} + (-\alpha_{i+1}/\alpha_i)v_{i+1} + \dots + (-\alpha_n/\alpha_i)v_n$$

оказывается линейной комбинацией векторов $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$.

Следующее простое утверждение оказывается очень полезным.

Лемма 2. Пусть V – векторное пространство над полем k , и пусть v_1, \dots, v_n, w – векторы из V . Если векторы v_1, \dots, v_n линейно независимы, а векторы v_1, \dots, v_n, w линейно зависимы, то вектор w является линейной комбинацией векторов v_1, \dots, v_n .

Доказательство. По определению линейно зависимого множества, какой-то из векторов v_1, \dots, v_n, w является линейной комбинацией остальных векторов; лемма утверждает, что в качестве такого вектора возможно взять именно добавленный вектор w . Действительно, поскольку векторы v_1, \dots, v_n, w линейно зависимы, существуют элементы $\alpha_1, \dots, \alpha_n, \beta \in k$, не все равные 0, такие что

$$\alpha_1 v_1 + \dots + \alpha_n v_n + \beta w = \theta.$$

Если $\beta = 0$, то ненулевым будет один из коэффициентов $\alpha_1, \dots, \alpha_n$; при этом будет

$$\alpha_1 v_1 + \dots + \alpha_n v_n = \alpha_1 v_1 + \dots + \alpha_n v_n + 0 \cdot w = \alpha_1 v_1 + \dots + \alpha_n v_n + \beta w = \theta,$$

а из этого по предложению 4 следует, что векторы v_1, \dots, v_n в противоречие с условием леммы линейно зависимы. Следовательно, $\beta \neq 0$, и

$$w = (-\alpha_1/\beta)v_1 + \dots + (-\alpha_n/\beta)v_n \in \langle v_1, \dots, v_n \rangle.$$

Базисы. Пусть V – векторное пространство над полем k . Мы говорим, что линейно независимые векторы $v_1, \dots, v_n \in V$ составляют максимальное линейно независимое семейство векторов пространства V , если для любого вектора $w \in V$ векторы v_1, \dots, v_n, w уже будут линейно зависимы. Если векторы $v_1, \dots, v_n \in V$ порождают пространство V , а при выбрасывании из этого набора любого вектора v_i , $1 \leq i \leq n$ оставшиеся векторы $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ уже не порождают V , то мы говорим, что векторы v_1, \dots, v_n составляют минимальное порождающее пространство V семейство.

Теорема 1. Пусть V – векторное пространство над полем k , и пусть $v_1, \dots, v_n \in V$. Следующие условия равносильны:

- (1) векторы v_1, \dots, v_n линейно независимы и порождают пространство V ;
- (2) векторы v_1, \dots, v_n составляют максимальное линейно независимое семейство векторов пространства V ;
- (3) векторы v_1, \dots, v_n составляют минимальное порождающее пространство V семейство.

Доказательство. (1) \Rightarrow (2). Пусть векторы v_1, \dots, v_n линейно независимы и порождают пространство V , и пусть w – любой вектор из V . Поскольку векторы v_1, \dots, v_n порождают V , вектор w является их линейной комбинацией, и потому векторы v_1, \dots, v_n, w линейно зависимы. Следовательно, векторы v_1, \dots, v_n составляют максимальное линейно независимое семейство векторов пространства V .

(2) \Rightarrow (3). Пусть векторы v_1, \dots, v_n составляют максимальное линейно независимое семейство векторов пространства V . Тогда для любого вектора $w \in V$ векторы v_1, \dots, v_n, w линейно зависимы, и по лемме 2 именно добавленный к линейно независимым векторам v_1, \dots, v_n вектор w является линейной комбинацией векторов v_1, \dots, v_n . Итак, векторы v_1, \dots, v_n порождают пространство V .

Напротив, при выбрасывании из этого семейства любого вектора v_i , $1 \leq i \leq n$, оставшиеся векторы $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ не порождают V , потому что векторы v_1, \dots, v_n линейно независимы, и уже вектор v_i не может быть линейной комбинацией векторов $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$. Таким образом, векторы v_1, \dots, v_n составляют минимальное порождающее пространство V семейства.

(3) \Rightarrow (1). Пусть векторы v_1, \dots, v_n составляют минимальное порождающее пространство V семейства; покажем, что они линейно независимы. Если бы это было не так, то нашелся бы вектор v_i , $1 \leq i \leq n$, который был бы линейной комбинацией остальных векторов $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$; но тогда любой вектор из V , который является линейной комбинацией векторов $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ и вектора v_i , был бы по транзитивности линейных комбинаций (лемма 1) линейной комбинацией векторов $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$, которые, таким образом, порождали бы пространство V , а это противоречит тому, что векторы v_1, \dots, v_n составляют минимальное порождающее пространство V семейства.

Если векторы v_1, \dots, v_n удовлетворяют любому из требований (1), (2), (3) только что доказанной теоремы (а значит, и всем этим требованиям), то мы говорим, что v_1, \dots, v_n – базис V . Отметим, что для базиса важен порядок, в котором расположены составляющие его векторы: базисы v_1, v_2, \dots, v_n и v_2, v_1, \dots, v_n считаются различными базисами.

Существование базиса. Все теоремы о существовании базисов, обладающих теми или иными свойствами, вытекают из следующего утверждения, усиливающего часть утверждений теоремы 1.

Лемма 3. Пусть векторы v_1, \dots, v_N порождают пространство V . Если векторы v_{i_1}, \dots, v_{i_n} , где $1 \leq i_1, \dots, i_n \leq N$, линейно независимы, а для любого j , $1 \leq j \leq N$, векторы $v_{i_1}, \dots, v_{i_n}, v_j$ линейно зависимы, то $v_{i_1}, v_{i_2}, \dots, v_{i_n}$ – базис пространства V .

Доказательство. По лемме 2 все векторы v_j ($1 \leq j \leq N$), являются линейными комбинациями векторов $v_{i_1}, v_{i_2}, \dots, v_{i_n}$. Поскольку векторы v_1, \dots, v_N порождают пространство V , любой вектор из V является их линейной комбинацией, а значит, по транзитивности линейных комбинаций, и линейной комбинацией векторов $v_{i_1}, v_{i_2}, \dots, v_{i_n}$. Таким образом, векторы $v_{i_1}, v_{i_2}, \dots, v_{i_n}$ не только линейно независимы, но еще и порождают пространство V , а это и значит, что $v_{i_1}, v_{i_2}, \dots, v_{i_n}$ – базис V .

Теорема 2. Пусть V – конечно порожденное векторное пространство над полем k . Тогда:

- (1) в пространстве V есть хотя бы один базис;
- (2) если система векторов v_1, \dots, v_N порождает пространство V , то из нее можно выбрать подсистему, являющуюся базисом пространства V ;
- (3) Если векторы u_1, \dots, u_r линейно независимы, то существует базис пространства V , содержащий все векторы u_1, \dots, u_r .

Доказательство. Все три утверждения доказываются одинаково. Поскольку V – конечно порожденное векторное пространство, найдутся векторы v_1, \dots, v_N , которые порождают пространство V . Пусть $1 \leq i_1, \dots, i_s \leq N$; при $s > N$ среди векторов $u_1, \dots, u_r, v_{i_1}, \dots, v_{i_s}$ будут одинаковые, и потому эти векторы линейно зависимы. Напротив, при $s = 0$ мы получаем семейство векторов u_1, \dots, u_r , которые линейно независимы. Пусть s , $0 \leq s \leq N$ – самое большое из чисел, обладающих свойством: существуют такие i_1, \dots, i_s , $1 \leq i_1, \dots, i_s \leq N$, что векторы $u_1, \dots, u_r, v_{i_1}, \dots, v_{i_s}$ линейно независимы. Если w – любой из векторов $u_1, \dots, u_r, v_1, \dots, v_N$, то векторы $u_1, \dots, u_r, v_{i_1}, \dots, v_{i_s}, w$ линейно зависимы: если $w = u_j$ для некоторого j , $1 \leq j \leq r$, то среди векторов $u_1, \dots, u_r, v_{i_1}, \dots, v_{i_s}, w$

Теорема о линейной зависимости линейных комбинаций. Мы покажем далее, что все базисы пространства состоят из одинакового числа элементов. Доказательство основано на следующем важном факте, называемом теоремой о линейной зависимости линейных комбинаций.

Замечание. В теореме ничего не говорится о случае, когда $m \leq n$: несмотря на то, что количество линейных комбинаций мало, они могут быть линейно зависимыми.

Поскольку векторы u_1, \dots, u_m являются линейными комбинациями векторов v_1, \dots, v_n , существуют такие элементы $\alpha_{ij} \in k$ ($1 \leq i \leq m$, $1 \leq j \leq n$), что

Если $\alpha_{11} = \alpha_{21} = \dots = \alpha_{m1} = 0$, то векторы u_1, \dots, u_m являются на самом деле линейными комбинациями $n - 1$ векторов v_2, \dots, v_n ; поэтому по предположению индукции $m > n > n - 1$ векторов u_1, \dots, u_m линейно зависимы. Пусть среди коэффициентов α_{i1} есть ненулевые; меняя, если надо, нумерацию векторов u_1, \dots, u_m мы можем, не умаляя общности, считать, что $\alpha_{11} \neq 0$. Тогда $m - 1$ векторов $w_i = u_i - (\alpha_{i1}/\alpha_{11})u_1$ ($2 \leq i \leq m$) являются линейными комбинациями $n - 1 < m - 1$ векторов v_2, \dots, v_n , и по предположению индукции они линейно зависимы. Следовательно, один из этих векторов w_j представляет собой линейную комбинацию векторов $w_2, \dots, w_{j-1}, w_{j+1}, \dots, w_m$. Но каждый из перечисленных векторов является линейной комбинацией векторов $u_1, u_2, \dots, u_{j-1}, u_{j+1}, \dots, u_m$, поэтому по транзитивности линейных комбинаций вектор w_j , а вместе с ним и вектор $u_j = w_j + (\alpha_{j1}/\alpha_{11})u_1$, представляет собой линейную комбинацию векторов $u_1, u_2, \dots, u_{j-1}, u_{j+1}, \dots, u_m$. А это значит, что векторы u_1, \dots, u_m линейно зависимы.

Предложение 5. Пусть $u_1, \dots, u_m; v_1, \dots, v_n$ – базисы пространства V . Тогда $m = n$.

Доказательство. Поскольку векторы u_1, \dots, u_m составляют базис пространства V , они порождают V . Векторы v_1, \dots, v_n , как и все векторы из пространства V , являются их линейными комбинациями; если бы было $n > m$, то по теореме о линейной зависимости линейных комбинаций векторы v_1, \dots, v_n были бы линейно зависимы, а это противоречит тому, что они составляют базис. Следовательно, $n \leq m$; точно так же доказывается, что $m \leq n$.

Пусть теперь пространство V конечно порождено; тогда в нем есть конечные базисы, и любые два базиса состоят из одинакового количества элементов. Число элементов в базисе конечно порожденного пространства V называется размерностью V и обозначается $\dim V$. Иногда одно и то же множество рассматривается как векторное пространство над разными полями; в таких случаях размерность V как векторного пространства над полем k обозначают через $\dim_k V$.

Всякое пространство конечной размерности n порождается своим базисом, состоящим из n элементов; поэтому оно конечно порождено. Обратно, у всякого конечно порожденного пространства размерность конечна. Поэтому конечно порожденное пространство – это то же самое, что пространство конечной размерности; такие пространства называются конечномерными.

Из теоремы 2 сразу следуют следующие утверждения.

Предложение 6. Пусть V – конечномерное пространство над полем k .

(1) Если векторы v_1, \dots, v_m порождают пространство V , то $m \geq \dim V$; если при этом $m = \dim V$, то v_1, \dots, v_m – базис V .

(2) Если векторы $v_1, \dots, v_m \in V$ линейно независимы, то $m \leq \dim V$; если при этом $m = \dim V$, то v_1, \dots, v_m – базис V .

Доказательство. (1). По теореме 2 из порождающей V системы векторов v_1, \dots, v_m можно выделить подсистему, являющуюся базисом и потому состоящую из $\dim V$ элементов. Следовательно, $m \geq \dim V$, а если $m = \dim V$, то выделенный базис совпадает со всей системой векторов v_1, \dots, v_m .

(2). По теореме 2 к линейно независимой системе векторов $v_1, \dots, v_m \in V$ можно добавить несколько векторов так, чтобы получился базис V ; поскольку любой базис состоит из $\dim V$ элементов, число элементов m исходной системы векторов не больше $\dim V$. Если же линейно $m = \dim V$, то количество добавленных векторов равно 0, и потому уже система v_1, \dots, v_m является базисом пространства.

Следствие. Пусть V – конечномерное пространство над полем k и пусть U – его подпространство. Тогда $\dim U \leq \dim V$. Если при этом $\dim U = \dim V$, то $U = V$.

Доказательство. Всякий базис u_1, \dots, u_m пространства U составляет линейно независимую систему векторов пространства V , поэтому $m = \dim U \leq \dim V$. Если при этом $m = \dim U = \dim V$, то u_1, \dots, u_m – базис V , и значит, всякий вектор из V является линейной комбинацией векторов u_1, \dots, u_m и потому принадлежит U .

Размерность арифметического векторного пространства. Напомним, что арифметическим векторным пространством размерности n называется пространство столбцов k^n .

Предложение 7. Пусть k – поле. Размерность пространства k^n равна n .

Доказательство. Обозначим через $e_i \in k^n$ ($1 \leq i \leq n$) столбец, у которого i -й элемент равен 1, а все остальные элементы равны 0:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Покажем, что e_1, e_2, \dots, e_n – базис k^n ; из этого и будет следовать, что $\dim k^n = n$. Действительно, для любых коэффициентов $\alpha_1, \alpha_2, \dots, \alpha_n \in k$ мы имеем: $\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n =$

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + \alpha_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \alpha_2 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Отсюда следует, во-первых, что любой столбец $(\alpha_1, \alpha_2, \dots, \alpha_n)^T \in k^n$ является линейной комбинацией $\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$ векторов e_1, e_2, \dots, e_n , а во-вторых, что если эта комбинация равна нулевому вектору, то $(\alpha_1, \alpha_2, \dots, \alpha_n)^T = (0, 0, \dots, 0)^T$, а потому $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$, так что векторы e_1, e_2, \dots, e_n линейно независимы.

Базис e_1, e_2, \dots, e_n пространства k^n , построенный в приведенном доказательстве, называется каноническим базисом пространства k^n .

Координаты вектора. Пусть V – конечномерное пространство над полем k и пусть v_1, \dots, v_n – его базис. Тогда любой вектор $v \in V$ представляется в виде линейной комбинации векторов базиса:

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n, \quad \alpha_1, \dots, \alpha_n \in k.$$

Если $v = \beta_1 v_1 + \dots + \beta_n v_n$ – другое представление того же вектора в виде линейной комбинации векторов базиса, то $(\alpha_1 - \beta_1)v_1 + \dots + (\alpha_n - \beta_n)v_n = \theta$, и, поскольку векторы v_1, \dots, v_n линейно независимы, отсюда следует, что $\alpha_1 - \beta_1 = \dots = \alpha_n - \beta_n = 0$. Таким образом, коэффициенты линейной комбинации векторов базиса, равной вектору v , определены однозначно; они называются координатами вектора v в базисе v_1, \dots, v_n . Чаще всего координаты вектора $\alpha_1, \dots, \alpha_n$ записывают в виде столбца $(\alpha_1, \alpha_2, \dots, \alpha_n)^T$, который называется столбцом координат вектора v в базисе v_1, \dots, v_n . Любой вектор может быть представлен в виде произведения двух матриц: строки из базисных элементов и столбца координат:

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n = v_1 \alpha_1 + \dots + v_n \alpha_n = (v_1, \dots, v_n) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Отметим, что действия над векторами и действия над столбцами координат согласованы: столбец координат суммы векторов в некотором базисе равен сумме столбцов координат слагаемых в том же базисе, а столбец координат произведения элемента $\alpha \in k$ на вектор равен произведению α на столбец координат вектора (в том же базисе). Действительно, пусть $(\alpha_1, \dots, \alpha_n)^T, (\beta_1, \dots, \beta_n)^T$ столбцы координат векторов $u, v \in V$ в базисе v_1, \dots, v_n ; тогда

$$u + v = (\alpha_1 v_1 + \dots + \alpha_n v_n) + (\beta_1 v_1 + \dots + \beta_n v_n) = (\alpha_1 + \beta_1)v_1 + \dots + (\alpha_n + \beta_n)v_n, \\ \alpha u = \alpha(\alpha_1 v_1 + \dots + \alpha_n v_n) = (\alpha \alpha_1)v_1 + \dots + (\alpha \alpha_n)v_n,$$

т.е. столбцы координат векторов $u+v, \alpha u$ в том же базисе равны $(\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)^T$ и $(\alpha \alpha_1, \dots, \alpha \alpha_n)$ соответственно.

Предложение 8. Пусть v_1, \dots, v_n – линейно независимые элементы из некоторого векторного пространства над полем k , и пусть $A, B \in k^{n \times m}$ – две матрицы с компонентами из k . Если $(v_1, \dots, v_n)A = (v_1, \dots, v_n)B$, то $A = B$.

$$v_1 A_{1j} + \dots + v_n A_{nj} = v_1 B_{1j} + \dots + v_n B_{nj}.$$

Предложение 9. Пусть векторы v_1, \dots, v_n порождают векторное пространство V над полем k , и пусть u_1, \dots, u_m – любые векторы из V . Тогда существует матрица $C \in k^{n \times m}$, такая что

Доказательство. Поскольку векторы v_1, \dots, v_n порождают пространство V , векторы $u_1, \dots, u_m \in V$ – их линейные комбинации, а значит, существуют элементы $\alpha_{ij} \in k$ ($1 \leq i \leq n$, $1 \leq j \leq m$), такие что

Эту систему равенств можно записать в виде одного матричного равенства

где через $C \in k^{n \times m}$ обозначена матрица

Обратим внимание на то, что элементы в матрице C расположены не там, где они были в предшествующей системе равенств, а на симметричных относительно диагонали позициях.

Теорема 4. Пусть u_1, \dots, u_n – базис векторного пространства над полем k , и пусть v_1, \dots, v_m – векторы из V , связанные с базисом соотношением

где $C \in k^{n \times m}$ – матрица с компонентами из k . Для того, чтобы векторы v_1, \dots, v_m тоже составляли базис пространства V , необходимо и достаточно, чтобы матрица C была обратимой матрицей (откуда, в частности, следует, что $m = n$).

Доказательство. Необходимость. Пусть v_1, \dots, v_m – базис V ; тогда векторы исходного базиса являются линейными комбинациями векторов v_1, \dots, v_m , и потому существует такая матрица $D \in k^{m \times n}$, что

$$(u_1, \dots, u_n) = (v_1, \dots, v_m)D.$$

Мы получаем теперь:

$$\begin{aligned}(u_1, \dots, u_n)E_n &= (u_1, \dots, u_n) = (v_1, \dots, v_m)D = \\ &= ((u_1, \dots, u_n)C)D = (u_1, \dots, u_n)(CD).\end{aligned}$$

Таким образом, произведения строки из линейно независимых векторов (u_1, \dots, u_n) на матрицы $E_n, CD \in k^{n \times n}$ совпадают; следовательно, $CD = E_n$. Точно так же показываем, что $DC = E_m$. Равенства $CD = E_n, DC = E_m$ означают, что D – обратная к C матрица, а значит, матрица C обратима.

Достаточность. Заметим, что размерность пространства V равна количеству n элементов в базисе u_1, \dots, u_n . Пусть матрица C обратима; тогда она necessarily является квадратной матрицей, т.е. $m = n = \dim V$. Домножая обе части соотношения

$$(v_1, \dots, v_n) = (u_1, \dots, u_n)C$$

на матрицу C^{-1} , получим, что $(u_1, \dots, u_n) = (v_1, \dots, v_n)C^{-1}$, откуда видно, что все векторы базиса u_1, \dots, u_n являются линейными комбинациями векторов v_1, \dots, v_n . Поскольку всякий вектор из V – линейная комбинация базисных векторов u_1, \dots, u_n , по транзитивности линейных комбинаций отсюда следует, что всякий вектор из V – линейная комбинация векторов v_1, \dots, v_n . Поэтому векторы v_1, \dots, v_n порождают пространство V ; но их количество n равно размерности пространства V , и поэтому они составляют базис V .

Пусть $u_1, \dots, u_n, v_1, \dots, v_n$ – два базиса векторного пространства V над полем k ; обратимая матрица $C \in k_n$, такая что $(v_1, \dots, v_n) = (u_1, \dots, u_n)C$ (она существует по предложению 9 и единственна по предложению 8) называется матрицей перехода от базиса u_1, \dots, u_n к базису v_1, \dots, v_n .

Предложение 10. Пусть $u_1, \dots, u_n, u'_1, \dots, u'_n$ – два базиса векторного пространства V над полем k , и пусть C – матрица перехода от первого базиса ко второму. Пусть, далее, $(\alpha_1, \dots, \alpha_n)^T, (\alpha'_1, \dots, \alpha'_n)^T$ – столбцы координат некоторого вектора $v \in V$ в первом и втором базисах. Тогда

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = C \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix}.$$

Доказательство. Обозначим для краткости столбцы координат $(\alpha_1, \dots, \alpha_n)^T, (\alpha'_1, \dots, \alpha'_n)^T$ через X, X' . Из определений столбца координат и матрицы перехода получаем соотношение

$$v = (u_1, \dots, u_n)X = (u'_1, \dots, u'_n)X', \quad (u'_1, \dots, u'_n) = (u_1, \dots, u_n)C;$$

поэтому

$$(u_1, \dots, u_n)X = (u'_1, \dots, u'_n)X' = ((u_1, \dots, u_n)C)X' = (u_1, \dots, u_n)(CX'),$$

откуда следует, что $X = CX'$, потому что векторы u_1, \dots, u_n линейно независимы.

Предложение 11. Пусть $u_1, \dots, u_n; v_1, \dots, v_n; w_1, \dots, w_n$ – три базиса векторного пространства V над полем k , и пусть C – матрица перехода от базиса u_1, \dots, u_n к базису v_1, \dots, v_n , а D – матрица перехода от базиса v_1, \dots, v_n к базису w_1, \dots, w_n . Тогда матрица перехода от базиса u_1, \dots, u_n к базису w_1, \dots, w_n равна CD , а матрица перехода от базиса v_1, \dots, v_n к базису u_1, \dots, u_n равна C^{-1} .

Доказательство. По определению матрицы перехода

$$(v_1, \dots, v_n) = (u_1, \dots, u_n)C, \quad (w_1, \dots, w_n) = (v_1, \dots, v_n)D,$$

откуда следует, что

$$(w_1, \dots, w_n) = (v_1, \dots, v_n)D = ((u_1, \dots, u_n)C)D = (u_1, \dots, u_n)(CD),$$

а это значит, что CD – матрица перехода от базиса u_1, \dots, u_n к базису w_1, \dots, w_n . Точно так же равенство

$$(u_1, \dots, u_n) = (u_1, \dots, u_n)(CC^{-1}) = ((u_1, \dots, u_n)C)C^{-1} = (v_1, \dots, v_n)C^{-1}$$

показывает, что C^{-1} – матрица перехода от базиса v_1, \dots, v_n к базису u_1, \dots, u_n .

3. РАНГ МАТРИЦЫ

Определение ранга матрицы. Пусть $A \in k^{m \times n}$ – матрица с компонентами из поля k , состоящая из m строк и n столбцов. Тогда все столбцы матрицы A принадлежат арифметическому векторному пространству k^m . Линейная оболочка столбцов является подпространством m -мерного пространства k^m и потому конечномерна; ее размерность называется рангом матрицы A и обозначается $\text{rank } A$. Ясно, что ранг матрицы $A \in k^{m \times n}$ не больше количества n столбцов, порождающих линейную оболочку, и не больше количества m строк матрицы, потому что размерность линейной оболочки, являющейся подпространством пространства k^m , не больше $\dim k^m = m$.

Ранг произведения матриц. Следующее почти тривиальное утверждение оказывается часто полезным.

Предложение 1. Пусть A, B – две матрицы с компонентами из поля k , состоящие из одинакового числа строк. Если существует такая матрица X с компонентами из k , что $B = AX$, то все столбцы матрицы B являются линейными комбинациями столбцов матрицы A . Обратно, если все столбцы матрицы B являются линейными комбинациями столбцов матрицы A , то существует такая матрица X с компонентами из k , что $B = AX$.

Доказательство. Пусть

$$\mathbf{a}_1 = \begin{pmatrix} A_{11} \\ \vdots \\ A_{m1} \end{pmatrix}, \dots, \mathbf{a}_n = \begin{pmatrix} A_{1n} \\ \vdots \\ A_{mn} \end{pmatrix}; \quad \mathbf{b}_1 = \begin{pmatrix} B_{11} \\ \vdots \\ B_{m1} \end{pmatrix}, \dots, \mathbf{b}_p = \begin{pmatrix} B_{1p} \\ \vdots \\ B_{mp} \end{pmatrix}$$

– столбцы матриц A, B , и пусть $X \in k^{n \times p}$ – такая матрица, что $B = AX$, т.е.

$$(\mathbf{b}_1 \quad \dots \quad \mathbf{b}_p) = (\mathbf{a}_1 \quad \dots \quad \mathbf{a}_n) X.$$

Поскольку матрицы, разбитые на блоки, можно перемножать, как если бы блоки были элементами матриц, мы получаем, что для любого j , $1 \leq j \leq p$ столбец $\mathbf{b}_j = \mathbf{a}_1 X_{1j} + \dots + \mathbf{a}_n X_{nj} = X_{1j} \mathbf{a}_1 + \dots + X_{nj} \mathbf{a}_n$ является линейной комбинацией столбцов $\mathbf{a}_1, \dots, \mathbf{a}_n$.

Обратно, пусть все столбцы \mathbf{b}_j являются линейными комбинациями столбцов $\mathbf{a}_1, \dots, \mathbf{a}_n$. Тогда существуют такие элементы $\alpha_{ij} \in k$, $1 \leq i \leq n$, $1 \leq j \leq p$, что $\mathbf{b}_j = \alpha_{1j} \mathbf{a}_1 + \dots + \alpha_{nj} \mathbf{a}_n$ для всех j , а это значит, что

$$B = (\mathbf{b}_1 \quad \dots \quad \mathbf{b}_p) = (\mathbf{a}_1 \quad \dots \quad \mathbf{a}_n) \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1p} \\ \dots & \dots & \dots \\ \alpha_{n1} & \dots & \alpha_{np} \end{pmatrix} = A \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1p} \\ \dots & \dots & \dots \\ \alpha_{n1} & \dots & \alpha_{np} \end{pmatrix}.$$

Следствие. Пусть k – поле, и пусть $A \in k^{m \times n}$, $B \in k^{n \times p}$. Тогда $\text{rank } AB \leq \text{rank } A$, $\text{rank } AB \leq \text{rank } B$.

Доказательство. По предыдущему предложению все столбцы матрицы AB принадлежат линейной оболочке столбцов матрицы A , и поэтому линейная оболочка столбцов матрицы AB является подпространством линейной оболочки столбцов матрицы A . Поэтому ранг матрицы AB , равный размерности линейной оболочки ее столбцов, не больше размерности линейной оболочки столбцов матрицы A , которая равна рангу A .

Чуть сложнее доказывается второе утверждение. Пусть $r = \text{rank } B$ и пусть $B' \in k^{p \times r}$ – матрица, столбцы которой составляют базис линейной оболочки столбцов матрицы B . Все столбцы матрицы B являются линейными комбинациями столбцов матрицы B' , и по предложению 1 существует такая матрица $X \in k^{r \times p}$, что $B = B'X$. Тогда по уже доказанному первому утверждению следствия $\text{rank } AB = \text{rank } A(B'X) = \text{rank } (AB')X \leq \text{rank } AB'$. Но матрица AB' состоит из r столбцов, поэтому ее ранг не может быть больше r , и мы получаем:

$$\text{rank } AB \leq \text{rank } AB' \leq r = \text{rank } B.$$

Инвариантность ранга при элементарных преобразованиях. Одно из основных свойств ранга матрицы – его инвариантность при элементарных преобразованиях.

Теорема 1. Ранг матрицы не меняется при умножении матрицы слева или справа на обратимую матрицу. В частности, ранг матрицы не меняется при элементарных преобразованиях над строками или над столбцами матрицы.

(Напомним, что элементарные преобразования над матрицей – это умножение матрицы слева или справа на элементарные матрицы, которые, как мы видели раньше, обратимы).

Доказательство. Пусть A матрица с компонентами из поля k , и пусть $B = AX$, $C = YA$, где X, Y – обратимые матрицы с компонентами из k . По только что доказанному следствию предложения 1

$$\text{rank } B = \text{rank } AX \leq \text{rank } A, \quad \text{rank } C = \text{rank } YA \leq \text{rank } A.$$

С другой стороны, $A = AE = A(XX^{-1}) = (AX)X^{-1} = BX^{-1}$, и точно так же $A = Y^{-1}C$; по тому же следствию

$$\text{rank } A = \text{rank } BX^{-1} \leq \text{rank } B, \quad \text{rank } A = \text{rank } Y^{-1}C \leq \text{rank } C.$$

Ранг ступенчатой матрицы. Как мы знаем, при помощи элементарных преобразований над строчками матрица может быть сделана ступенчатой. Элементарные преобразования не меняют ранг, поэтому если мы будем уметь вычислять ранг ступенчатых матриц, мы сможем найти ранг любой матрицы. Но ранг ступенчатой матрицы найти очень просто.

Теорема 2. Ранг ступенчатой матрицы с компонентами из поля равен числу ненулевых строк этой матрицы (или, что то же, числу ступенек).

Доказательство. Напомним, что матрица $A \in k^{m \times n}$ называется ступенчатой, если существуют такие числа r, j_1, \dots, j_r , что $0 \leq r \leq m$, $1 \leq j_1 < \dots < j_r \leq n$ и выполняются условия:

- (1) $A_{1j_1} = \dots = A_{rj_r} = 1$;
- (2) если $s \leq r$ и $i \neq s$, то $A_{ij_s} = 0$;
- (3) если $i > r$, то $A_{ij} = 0$ для всех j , $1 \leq j \leq n$;
- (4) если $i \leq r$, $j < j_i$, то $A_{ij} = 0$.

Из условий (1) и (2) мы видим, что в j_s -м столбце матрицы A s -й элемент равен 1, а все остальные элементы равны 0. Таким образом, j_s -й столбец матрицы A совпадает с вектором e_s из канонического базиса $\{e_1, \dots, e_m\}$ пространства

столбцов k^m . Векторы e_1, \dots, e_r линейно независимы; в то же время все столбцы матрицы A – их линейные комбинации, поскольку по условию (3) ненулевые элементы есть только в первых r строках матрицы A . Таким образом, векторы e_1, \dots, e_r составляют базис линейной оболочки столбцов матрицы A , и потому размерность этой линейной оболочки равна r . Но по определению ранга именно эта размерность и есть ранг матрицы A .

Следствие. Ранг матрицы

$$\begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix},$$

где E_r – единичная матрица порядка $r \geq 0$, а $\mathbf{0}^{r \times t}$, $\mathbf{0}^{s \times r}$, $\mathbf{0}^{s \times t}$ – нулевые матрицы, равен r .

Это утверждение – очевидное следствие теоремы 2: указанная в его формулировке матрица является ступенчатой.

Ранг транспонированной матрицы. В определении ранга строки и столбцы матрицы играли разные роли. Мы покажем, что на самом деле они равноправны по отношению к рангу.

Теорема 3. Пусть A – матрица с компонентами из некоторого поля k . Тогда $\text{rank } A^T = \text{rank } A$.

Доказательство. Мы видели, что существуют такие обратимые матрицы X, Y , такие что

$$XAY = \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix},$$

где E_r – единичная матрица порядка $r \geq 0$, а $\mathbf{0}^{r \times t}$, $\mathbf{0}^{s \times r}$. Тогда

$$Y^T A^T X^T = (XAY)^T = \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix}^T = \begin{pmatrix} E_r & \mathbf{0}^{r \times s} \\ \mathbf{0}^{t \times r} & \mathbf{0}^{t \times s} \end{pmatrix},$$

причем матрицы X^T , Y^T тоже обратимы. По теореме 1 и следствию теоремы 2 мы получаем:

$$\begin{aligned} \text{rank } A &= \text{rank } XAY = \text{rank} \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{s \times r} & \mathbf{0}^{s \times t} \end{pmatrix} = r = \\ &= \text{rank} \begin{pmatrix} E_r & \mathbf{0}^{r \times s} \\ \mathbf{0}^{t \times r} & \mathbf{0}^{t \times s} \end{pmatrix} = \text{rank } Y^T A^T X^T = \text{rank } A^T. \end{aligned}$$

Поскольку столбцы матрицы A^T являются строками матрицы A , мы получаем следующее

Следствие. Размерность линейной оболочки строк матрицы равна ее рангу.

Таким образом, ранг матрицы можно было бы определять не только как размерность линейной оболочки столбцов матрицы, но и как размерность линейной оболочки ее строк.

Невырожденные матрицы. Пусть k – поле. Квадратная матрица порядка n с компонентами из k называется невырожденной, если ее ранг равен n . Есть много утверждений, равносильных указанному свойству.

Теорема 4. Пусть k – поле, и пусть $A \in k_n = k^{n \times n}$ – квадратная матрица с компонентами из k . Следующие условия равносильны:

- (1) матрица A невырождена;
- (2) размерность линейной оболочки столбцов матрицы A равна n ;
- (3) размерность линейной оболочки строк матрицы A равна n ;
- (4) столбцы матрицы A линейно независимы;

- (5) строки матрицы A линейно независимы;
- (6) столбцы матрицы A порождают пространство столбцов k^n ;
- (7) строки матрицы A порождают пространство строк $k^{1 \times n}$;
- (8) матрица A обратима;
- (9) матрица A представляется в виде произведения элементарных матриц.

Доказательство. (1) \Leftrightarrow (2), (1) \Leftrightarrow (3) – просто перефразировки определений: матрица невырождена, когда ее ранг, т.е. размерность линейной оболочки ее столбцов (строк), равен n .

(2) \Rightarrow (6). Линейная оболочка столбцов матрицы A представляет собой подпространство пространства столбцов k^n . Если размерность этого подпространства совпадает с размерностью n всего пространства k^n , то подпространство совпадает со всем пространством, т.е. k^n является линейной оболочкой столбцов матрицы A , а это и значит, что столбцы матрицы A порождают пространство столбцов k^n .

(6) \Rightarrow (4). Если n столбцов матрицы A порождают пространство k^n , размерность которого равна n , то они линейно независимы.

(4) \Rightarrow (2). Если n столбцов матрицы A линейно независимы, то размерность их линейной оболочки не меньше n . С другой стороны, линейная оболочка столбцов матрицы содержится в k^n , и потому ее размерность не больше, чем $\dim k^n = n$.

Точно так же доказывается, что (3) \Rightarrow (7) \Rightarrow (5) \Rightarrow (3).

(1) \Rightarrow (8). Существуют такие обратимые матрицы X, Y , что $XAY = B$, где матрица B имеет вид

$$B = \begin{pmatrix} E_r & \mathbf{0}^{r \times t} \\ \mathbf{0}^{t \times r} & \mathbf{0}^{t \times t} \end{pmatrix}$$

(здесь $t = n - r$). Поскольку при умножении на обратимые матрицы ранг не меняется, мы получаем, что $\text{rank } A = \text{rank } B = r$; если A – невырожденная матрица, то $r = \text{rank } A = n$, а тогда получается, что B – единичная матрица. Итак, $XAY = E$, откуда следует, что матрица

$$A = (X^{-1}X)A(Y^{-1}Y) = X^{-1}(XAY)Y^{-1} = X^{-1}EY^{-1} = X^{-1}Y^{-1}$$

представляется в виде произведения обратимых матриц X^{-1}, Y^{-1} , а потому и сама обратима.

(8) \Rightarrow (1). Если матрица A обратима, то она получается умножением единичной матрицы E на обратимую матрицу A . Поскольку при умножении на обратимую матрицу ранг матрицы не меняется, мы получаем, что $\text{rank } A = \text{rank } E = n$, т.е. матрица A невырождена.

(8) \Leftrightarrow (9) – было доказано ранее.

В следующей главе мы добавим к этим девяти характеристикам невырожденных матриц еще одну – в терминах определителя матрицы.

Ранг как максимальный порядок невырожденных подматриц. Пусть $A \in k^{m \times n}$ – некоторая матрица, и пусть $1 \leq i_1 < \dots < i_p \leq m$, $1 \leq j_1 < \dots < j_q \leq n$. Подматрицей матрицы A , составленной из строк с номерами i_1, \dots, i_p и столбцов с номерами j_1, \dots, j_q мы называем матрицу

$$A_{i_1, \dots, i_p}^{j_1, \dots, j_q} = \begin{pmatrix} A_{i_1 j_1} & \dots & A_{i_1 j_q} \\ \dots & \dots & \dots \\ A_{i_p j_1} & \dots & A_{i_p j_q} \end{pmatrix}.$$

Теорема 5. Пусть A – матрица с компонентами из поля k , ранг которой равен r . Тогда все квадратные подматрицы матрицы A , порядок которых больше r , вырождены, но в матрице A существует невырожденная квадратная подматрица порядка r .

$$(*) \quad \begin{array}{rcl} & a_{11} & x_1 + a_{12} & x_2 + \cdots + a_{1n}x_n = b_1, \\ & a_{21} & x_1 + a_{21} & x_2 + \cdots + a_{2n}x_n = b_2, \\ & & \dots\dots\dots & \dots\dots\dots \\ & a_{m1} & x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{array}$$

С системой (*) связаны две матрицы

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix},$$

которые называются соответственно матрицей системы и расширенной матрицей системы.

Следующая теорема указывает необходимое и достаточное условие разрешимости системы (*).

Теорема 1 (теорема Кронекера-Капелли). Система (*) имеет хотя бы одно решение тогда и только тогда, когда ранг матрицы системы равен рангу ее расширенной матрицы. Если система имеет решение, то это решение единственно тогда и только тогда, когда ранг матрицы системы равен числу неизвестных.

Доказательство. Пусть

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \dots, \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}; \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

– столбцы матрицы системы A и столбец свободных членов. Линейная оболочка $\langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle$ столбцов матрицы системы является подпространством линейной оболочки $\langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b} \rangle$ столбцов расширенной матрицы. Если у системы есть решение $\alpha_1, \alpha_2, \dots, \alpha_n$, то, очевидно, $\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \dots + \alpha_n \mathbf{a}_n = \mathbf{b}$, т.е. $\mathbf{b} \in \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle$ и потому $\langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle = \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b} \rangle$; но тогда

$$\text{rank } A = \dim \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle = \dim \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b} \rangle = \text{rank } B.$$

Обратно, если $\text{rank } A = \text{rank } B$, то

$$\dim \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle = \text{rank } A = \text{rank } B = \dim \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b} \rangle,$$

т.е. размерность подпространства $\langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle$ равна размерности содержащего его пространства $\langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b} \rangle$, а потому подпространство совпадает со всем пространством. В частности, столбец $\mathbf{b} \in \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b} \rangle = \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle$ является линейной комбинацией столбцов $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$, и потому существуют такие $\alpha_1, \alpha_2, \dots, \alpha_n \in k$, что $\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \dots + \alpha_n \mathbf{a}_n = \mathbf{b}$. Но это соотношение означает, что $\alpha_1, \alpha_2, \dots, \alpha_n$ – решение системы (*).

Пусть $\alpha_1, \alpha_2, \dots, \alpha_n, \beta_1, \beta_2, \dots, \beta_n$ – два различных решения системы; тогда хотя бы один из элементов $\alpha_1 - \beta_1, \alpha_2 - \beta_2, \dots, \alpha_n - \beta_n$ отличен от 0, но

$$(\alpha_1 - \beta_1) \mathbf{a}_1 + \dots + (\alpha_n - \beta_n) \mathbf{a}_n = (\alpha_1 \mathbf{a}_1 + \dots + \alpha_n \mathbf{a}_n) - (\beta_1 \mathbf{a}_1 + \dots + \beta_n \mathbf{a}_n) = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

а это значит, что столбцы $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ линейно зависимы, и потому размерность их линейной оболочки меньше, чем число столбцов, т.е.

$$\text{rank } A = \dim \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle < n.$$

Обратно, пусть $\text{rank } A \neq n$; поскольку ранг матрицы не больше количества n ее столбцов, из этого предположения следует, что $\text{rank } A < n$. Тогда

$$\dim \langle \mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \rangle = \text{rank } A < n,$$

и потому столбцы $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ линейно зависимы, а значит, существуют элементы $\gamma_1, \gamma_2, \dots, \gamma_n$, не все равные 0 и такие, что $\gamma_1 \mathbf{a}_1 + \gamma_2 \mathbf{a}_2 + \dots + \gamma_n \mathbf{a}_n = \mathbf{0}$. Если система (*) имеет решение $\alpha_1, \alpha_2, \dots, \alpha_n$, то $\alpha_1 \mathbf{a}_1 + \dots + \alpha_n \mathbf{a}_n = \mathbf{b}$, и потому

$$(\alpha_1 + \gamma_1) \mathbf{a}_1 + \dots + (\alpha_n + \gamma_n) \mathbf{a}_n = (\alpha_1 \mathbf{a}_1 + \dots + \alpha_n \mathbf{a}_n) + (\gamma_1 \mathbf{a}_1 + \dots + \gamma_n \mathbf{a}_n) = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Доказательство. Пусть A – матрица однородной системы m линейных уравнений с n неизвестными. Ранг матрицы A не больше, чем число ее строк, т.е. число m уравнений системы. Если $m < n$, то $\text{rang } A \leq m < n$, и по теореме 2 система имеет нетривиальные решения.

Следствие 2. *Однородная система n линейных уравнений с n неизвестными имеет нетривиальные решения тогда и только тогда, когда матрица этой системы вырождена.*

Доказательство. Пусть A – матрица однородной системы n линейных уравнений с n неизвестными. Это квадратная матрица порядка n . По теореме 2 система имеет нетривиальные решения тогда и только тогда, когда $\text{rang } A < n$, т.е. когда матрица A вырождена.

Связь между множествами решений неоднородной и однородной систем. Мы уже использовали для систем линейных уравнений матричную запись

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Теперь мы еще больше сократим ее, введя обозначения для столбцов неизвестных и свободных членов

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix};$$

теперь система линейных уравнений записывается совсем коротко:

$$AX = \mathbf{b}.$$

Решениями этой системы являются столбцы $X_0 \in k^n$, такие что $AX_0 = \mathbf{b}$.

Наряду с неоднородной системой линейных уравнений $AX = \mathbf{b}$ мы будем рассматривать однородную систему с той же матрицей $AX = \theta$. Установим, как связаны множества решений этих систем.

Теорема 3. *Пусть $\mathfrak{X}, \mathfrak{Y} \subseteq k^n$ – множества решений неоднородной системы $AX = \mathbf{b}$ и однородной системы с той же матрицей $AX = \theta$, и пусть $X_0 \in \mathfrak{X}$ – какое-то решение неоднородной системы. Если $Y_0 \in \mathfrak{Y}$ – любое решение однородной системы, то $X_0 + Y_0$ – тоже решение неоднородной системы. Для всякого решения $X_1 \in \mathfrak{X}$ неоднородной системы найдется такое решение однородной системы $Y_0 \in \mathfrak{Y}$, что $X_1 = X_0 + Y_0$.*

Замечание. *Если мы обозначим через $X_0 + \mathfrak{Y}$ множество всех столбцов вида $X_0 + Y_0$, где $Y_0 \in \mathfrak{Y}$, то утверждение теоремы примет вид: $\mathfrak{X} = X_0 + \mathfrak{Y}$.*

Доказательство. Если $X_0 \in \mathfrak{X}$, $Y_0 \in \mathfrak{Y}$, то $AX_0 = \mathbf{b}$, $AY_0 = \theta$, и потому

$$A(X_0 + Y_0) = AX_0 + AY_0 = \mathbf{b} + \theta = \mathbf{b},$$

а это значит, что $X_0 + Y_0 \in \mathfrak{X}$. пусть теперь $X_1 \in \mathfrak{X}$ – еще одно решение неоднородной системы; это значит, что $AX_1 = \mathbf{b}$. Положим $Y_0 = X_1 - X_0$; тогда $X_1 = X_0 + Y_0$ и, кроме того,

$$AY_0 = A(X_1 - X_0) = AX_1 - AX_0 = \mathbf{b} - \mathbf{b} = \theta,$$

т.е. Y_0 – решение однородной системы, и $Y_0 \in \mathfrak{Y}$.

Подробнее о строении множества решений однородной системы линейных уравнений мы поговорим в следующем параграфе.

5. ЛИНЕЙНЫЕ ОТОБРАЖЕНИЯ

Определение и примеры линейных отображений. Пусть U и V – два векторных пространства над одним и тем же полем k . Отображение $\mathcal{A}: U \rightarrow V$ называется линейным, если для любых векторов $u, u' \in U$ и любого $\alpha \in k$ выполняются соотношения $\mathcal{A}(u + u') = \mathcal{A}(u) + \mathcal{A}(u')$, $\mathcal{A}(\alpha u) = \alpha \mathcal{A}(u)$.

Часто бывает ситуация, в которой U и V могут рассматриваться как пространства над разными полями. Если \mathcal{A} является линейным отображением U в V , рассматриваемыми как пространства над полем k , то мы говорим, что \mathcal{A} – k -линейное отображение, или линейное отображение над полем k .

Из определения линейного отображения сразу следуют "пустяковые теоремы" о линейных отображениях.

Предложение 1. Пусть U, V – векторные пространства над полем k , и пусть \mathcal{A} – линейное отображение пространства U в пространство V . Тогда:

- (1) $\mathcal{A}(\theta_U) = \theta_V$;
- (2) $\mathcal{A}(-u) = -\mathcal{A}(u)$ для любого $u \in U$;
- (3) $\mathcal{A}(u_1 + \dots + u_n) = \mathcal{A}(u_1) + \dots + \mathcal{A}(u_n)$ для любых векторов $u_1, \dots, u_n \in U$;
- (4) $\mathcal{A}(\alpha_1 u_1 + \dots + \alpha_n u_n) = \alpha_1 \mathcal{A}(u_1) + \dots + \alpha_n \mathcal{A}(u_n)$ для любых векторов $u_1, \dots, u_n \in U$ и любых $\alpha_1, \dots, \alpha_n \in k$.

Доказательство. (1) $\mathcal{A}(\theta_U) = \mathcal{A}(\theta_U + \theta_U) = \mathcal{A}(\theta_U) + \mathcal{A}(\theta_U)$, и потому

$$\begin{aligned} \theta_V &= -\mathcal{A}(\theta_U) + \mathcal{A}(\theta_U) = -\mathcal{A}(\theta_U) + (\mathcal{A}(\theta_U) + \mathcal{A}(\theta_U)) = \\ &= (-\mathcal{A}(\theta_U) + \mathcal{A}(\theta_U)) + \mathcal{A}(\theta_U) = \theta_V + \mathcal{A}(\theta_U) = \mathcal{A}(\theta_U). \end{aligned}$$

$$(2) \mathcal{A}(-u) = \mathcal{A}((-1)u) = (-1)\mathcal{A}(u) = -\mathcal{A}(u).$$

(3) Индукция по n .

(4) По определению линейного отображения $\mathcal{A}(\alpha_i u_i) = \alpha_i \mathcal{A}(u_i)$ для всех i , $1 \leq i \leq n$, а тогда по свойству (3)

$$\mathcal{A}(\alpha_1 u_1 + \dots + \alpha_n u_n) = \alpha_1 \mathcal{A}(u_1) + \dots + \alpha_n \mathcal{A}(u_n).$$

Как и при определении подпространства, полезно дать несколько более слабую форму определения линейного отображения.

Предложение 2. Пусть U, V – векторные пространства над полем k , и пусть \mathcal{A} – отображение пространства U в пространство V . Для того, чтобы отображение \mathcal{A} было линейным, необходимо и достаточно, чтобы для любых $u, u' \in U$ и любого $\alpha \in k$ выполнялось соотношение $\mathcal{A}(\alpha u + u') = \alpha \mathcal{A}(u) + \mathcal{A}(u')$.

Доказательство. Необходимость очевидна: если \mathcal{A} – линейное отображение, то $\mathcal{A}(\alpha u + u') = \mathcal{A}(\alpha u) + \mathcal{A}(u') = \alpha \mathcal{A}(u) + \mathcal{A}(u')$. Докажем достаточность. Полагая в соотношении $\mathcal{A}(\alpha u + u') = \alpha \mathcal{A}(u) + \mathcal{A}(u')$ сначала $\alpha = 1$, а затем $u' = \theta$, получим:

$$\mathcal{A}(u + u') = \mathcal{A}(1 \cdot u + u') = 1 \cdot \mathcal{A}(u) + \mathcal{A}(u') = \mathcal{A}(u) + \mathcal{A}(u'),$$

$$\mathcal{A}(\alpha u) = \mathcal{A}(\alpha u + \theta) = \alpha \mathcal{A}(u) + \mathcal{A}(\theta) = \alpha \mathcal{A}(u) + \theta = \alpha \mathcal{A}(u),$$

а это и значит, что отображение \mathcal{A} линейно.

В дальнейшем мы не будем обычно подчеркивать функциональный характер линейного отображения и рассматривать его действие на вектор как умножение; поэтому обычно мы будем писать $\mathcal{A}u$ вместо $\mathcal{A}(u)$.

Почти вся математика занимается исследованием линейных отображений. Приведем лишь несколько примеров.

1. Проекция векторов в трехмерном геометрическом пространстве на какую-то плоскость или прямую – линейное отображение.

2. Поворот всех векторов плоскости в положительном направлении на один и тот же угол, а также замена вектора на симметричный с ним относительно

некоторой прямой – линейные отображения пространства векторов на плоскости в себя.

3. Векторное умножение трехмерных геометрических векторов на некоторый фиксированный вектор – линейные отображения пространства трехмерных векторов в себя, а скалярное умножение на некоторый фиксированный вектор – линейное отображение пространства трехмерных геометрических векторов в поле вещественных чисел.

4. Пусть $C^1(0, 1)$ – пространство всех непрерывно дифференцируемых на отрезке $(0, 1)$ функций. Взятие производной является линейным отображением этого пространства в пространство всех непрерывных на $(0, 1)$ функций. Умножение на любую функцию из $C^1(0, 1)$ является линейным отображением пространства $C^1(0, 1)$ в себя.

5. Пусть k – поле; умножение на любой многочлен $f(x) \in k[x]$ и взятие производной являются линейными отображениями пространства многочленов $k[x]$ в себя (над полем k). Эти примеры – алгебраические аналоги примеров из пункта 4.

6. Взятие сопряженного для комплексного числа или для кватерниона представляет собой \mathbb{R} -линейное отображение пространства \mathbb{C} (соответственно, \mathbb{H}) в себя.

7. Транспонирование матриц является линейным отображением из пространства $k^{m \times n}$ в пространство $k^{n \times m}$.

8. Пусть k – поле, и пусть $A \in k^{m \times n}$ – матрица с компонентами из k , состоящая из m строк и n столбцов. Определим отображение $\mathcal{A} : k^n \rightarrow k^m$, положив $\mathcal{A}(X) = AX$ для любого столбца $X \in k^n$. Из свойств действий над матрицами следует, что это отображение линейно.

В заключение определим линейные отображения, которые существуют для любых пространств.

Пусть U, V – пространства над полем k . Отображение, которое ставит в соответствие любому вектору из U нулевой вектор пространства V , является линейным отображением; оно называется нулевым линейным отображением и обозначается через \mathcal{O} . Точнее говоря, речь идет не об одном линейном отображении: для каждой пары пространств есть свое нулевое отображение.

Для любого пространства V его тождественное отображение на себя является линейным; оно обозначается \mathcal{E} (или \mathcal{E}_V , если нужно явно обозначить пространство, на котором действует это отображение) и называется единичным отображением.

Ядро и образ линейного отображения. Пусть U и V – два векторных пространства над полем k , и пусть $\mathcal{A} : U \rightarrow V$ – линейное отображение. Ядром \mathcal{A} называется множество всех векторов $u \in U$, таких что $\mathcal{A}u = \theta$; ядро отображения \mathcal{A} обозначается через $\text{Ker } \mathcal{A}$. Пусть теперь U_0 – любое подмножество U ; обозначим через $\mathcal{A}U_0$ множество всех тех векторов $v \in V$, для которых есть такой элемент $u_0 \in U_0$ (прообраз v), что $\mathcal{A}u_0 = v$. Таким образом, $\mathcal{A}U_0 = \{\mathcal{A}u \mid u \in U_0\}$. В частности, множество $\mathcal{A}U$ называется образом линейного отображения \mathcal{A} и обозначается $\text{Im } \mathcal{A}$.

Предложение 3. Пусть U и V – два векторных пространства над полем k , и пусть $\mathcal{A} : U \rightarrow V$ – линейное отображение. Тогда его ядро $\text{Ker } \mathcal{A}$ является подпространством пространства U . Для любого подпространства U_0 пространства U множество $\mathcal{A}U_0$ – подпространство пространства V ; в частности, подпространством V является образ $\text{Im } \mathcal{A} = \mathcal{A}U$.

Доказательство. Пусть $u, u' \in \text{Ker } \mathcal{A}$; это значит, что $\mathcal{A}u = \mathcal{A}u' = \theta$. Если α – любой элемент поля k , то $\mathcal{A}(\alpha u + u') = \alpha \mathcal{A}u + \mathcal{A}u' = \theta$, т.е. $\alpha u + u' \in \text{Ker } \mathcal{A}$. По признаку подпространства это означает, что $\text{Ker } \mathcal{A}$ – подпространство пространства U .

Пусть теперь $v, v' \in \mathcal{A}U_0$; тогда существуют такие элементы $u, u' \in U_0$, что $\mathcal{A}u = v$, $\mathcal{A}u' = v'$. Если α – любой элемент поля k , то $\alpha u + u' \in U_0$, потому что U_0 – подпространство пространства U , и

$$\alpha v + v' = \alpha \mathcal{A}u + \mathcal{A}u' = \mathcal{A}(\alpha u + u') \in \mathcal{A}U_0;$$

по признаку подпространства это означает, что $\mathcal{A}U_0$ – подпространство пространства V .

Теорема 1. Пусть $\mathcal{A} : U \rightarrow V$ – линейное отображение векторных пространств над полем k , и пусть пространство U конечномерно. Тогда ядро и образ \mathcal{A} являются конечномерными пространствами, и

$$\dim U = \dim \text{Ker } \mathcal{A} + \dim \text{Im } \mathcal{A}.$$

Доказательство. Ядро $\text{Ker } \mathcal{A}$ отображения \mathcal{A} является подпространством конечномерного пространства U и потому оно тоже конечномерно. Пусть u_1, \dots, u_m – какой-то базис ядра $\text{Ker } \mathcal{A}$; тогда u_1, \dots, u_m – линейно независимые векторы пространства U , и их можно включить в некоторый базис $u_1, \dots, u_m, u_{m+1}, \dots, u_n$ пространства U . Все будет доказано, если мы удостоверимся в том, что векторы $\mathcal{A}u_{m+1}, \dots, \mathcal{A}u_n$ составляют базис образа $\text{Im } \mathcal{A}$; действительно, тогда

$$\dim \text{Ker } \mathcal{A} = m, \quad \dim \text{Im } \mathcal{A} = n - m, \quad \dim U = n,$$

и соотношение $\dim U = \dim \text{Ker } \mathcal{A} + \dim \text{Im } \mathcal{A}$ выполняется.

Если $v \in \text{Im } \mathcal{A}$, то существует вектор $u \in U$, такой что $\mathcal{A}u = v$; поскольку $u_1, \dots, u_m, u_{m+1}, \dots, u_n$ – базис пространства U , вектор u представляется в виде $u = \alpha_1 u_1 + \dots + \alpha_m u_m + \alpha_{m+1} u_{m+1} + \dots + \alpha_n u_n$, и потому

$$\begin{aligned} v = \mathcal{A}u &= \mathcal{A}(\alpha_1 u_1 + \dots + \alpha_m u_m + \alpha_{m+1} u_{m+1} + \dots + \alpha_n u_n) = \\ &= \alpha_1 \mathcal{A}u_1 + \dots + \alpha_m \mathcal{A}u_m + \alpha_{m+1} \mathcal{A}u_{m+1} + \dots + \alpha_n \mathcal{A}u_n = \\ &= \alpha_1 \theta + \dots + \alpha_m \theta + \alpha_{m+1} \mathcal{A}u_{m+1} + \dots + \alpha_n \mathcal{A}u_n = \alpha_{m+1} \mathcal{A}u_{m+1} + \dots + \alpha_n \mathcal{A}u_n. \end{aligned}$$

Таким образом, векторы $\mathcal{A}u_{m+1}, \dots, \mathcal{A}u_n$ порождают пространство $\text{Im } \mathcal{A}$; покажем, что они линейно независимы. Пусть $\alpha_{m+1} \mathcal{A}u_{m+1} + \dots + \alpha_n \mathcal{A}u_n = \theta$, где все коэффициенты $\alpha_{m+1}, \dots, \alpha_n$ принадлежат полю k ; тогда

$$\mathcal{A}(\alpha_{m+1} u_{m+1} + \dots + \alpha_n u_n) = \theta,$$

и потому вектор $\alpha_{m+1} u_{m+1} + \dots + \alpha_n u_n$ принадлежит ядру $\text{Ker } \mathcal{A}$ отображения \mathcal{A} . Следовательно, этот вектор является линейной комбинацией векторов u_1, \dots, u_m , составляющих базис пространства $\text{Ker } \mathcal{A}$, т.е.

$$\alpha_{m+1} u_{m+1} + \dots + \alpha_n u_n = \alpha_1 u_1 + \dots + \alpha_m u_m$$

для некоторых $\alpha_1, \dots, \alpha_m \in k$; таким образом, линейная комбинация

$$(-\alpha_1)u_1 + \dots + (-\alpha_m)u_m + \alpha_{m+1}u_{m+1} + \dots + \alpha_n u_n$$

равна нулевому вектору, а потому все ее коэффициенты равны 0. В частности, $\alpha_{m+1} = \dots = \alpha_n = 0$. Итак, всякая линейная комбинация векторов $\mathcal{A}u_{m+1}, \dots, \mathcal{A}u_n$, равная нулевому вектору θ , является тривиальной линейной комбинацией, а это значит, что векторы $\mathcal{A}u_{m+1}, \dots, \mathcal{A}u_n$ линейно независимы.

Структура множества решений однородной системы линейных уравнений. Мы сейчас применим теорему 1 к линейным отображениям из примера 8.

Лемма 1. Пусть k – поле, $A \in k^{m \times n}$ и пусть $\mathcal{A} : k^n \rightarrow k^m$ – линейное отображение, определенное формулой: $\mathcal{A}(X_0) = AX_0$ для любого столбца $X_0 \in k^n$. Тогда $\text{Ker } \mathcal{A}$ совпадает с множеством решений однородной системы линейных уравнений $AX = \theta$, а $\text{Im } \mathcal{A}$ совпадает с линейной оболочкой столбцов матрицы A .

Доказательство. Оба утверждения очевидны. Столбец $X_0 \in k^n$ тогда и только тогда принадлежит ядру отображения \mathcal{A} , когда $\theta = \mathcal{A}(X_0) = AX_0$, т.е. когда столбец X_0 представляет собой решение однородной системы линейных уравнений $AX = \theta$. Далее, пусть $\mathbf{a}_1, \dots, \mathbf{a}_n$ – столбцы матрицы A ; тогда образ отображения \mathcal{A} состоит из всех векторов вида

$$\mathcal{A} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = (\mathbf{a}_1 \quad \dots \quad \mathbf{a}_n) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1 \mathbf{a}_1 + \dots + \alpha_n \mathbf{a}_n,$$

где $\alpha_1, \dots, \alpha_n \in k$, т.е. этот образ совпадает с множеством всех линейных комбинаций столбцов матрицы A .

Теорема 2. Пусть k – поле. Множество решений однородной системы m линейных уравнения с n неизвестными $AX = \theta$, где $A \in k^{m \times n}$ – матрица системы, является подпространством пространства k^n , и размерность пространства решений этой однородной системы равна $n - r$.

Доказательство. Пусть \mathcal{A} – линейное отображение из леммы 1. Тогда множество решений \mathfrak{Y} системы совпадает с $\text{Ker } \mathcal{A}$ и потому является подпространством пространства k^n . Поскольку $\text{Im } \mathcal{A}$ – линейная оболочка столбцов матрицы A , размерность пространства $\text{Im } \mathcal{A}$ равна рангу матрицы A . По теореме 1

$$\dim \mathfrak{Y} = \dim \text{Ker } \mathcal{A} = \dim k^n - \dim \text{Im } \mathcal{A} = n - r.$$

Действия над линейными отображениями. Хотя основной целью, для которой мы стали рассматривать здесь линейные отображения, было доказательство теоремы 2, мы приведем здесь еще некоторые элементарные факты о линейных отображениях.

Пусть U, V – векторные пространства над одним и тем же полем k ; суммой двух отображений $\mathcal{A}, \mathcal{B} : U \rightarrow V$ а произведением элемента $\alpha \in k$ на отображение \mathcal{A} называются отображение $\mathcal{A} + \mathcal{B} : U \rightarrow V$, заданные формулами:

$$(\mathcal{A} + \mathcal{B})u = \mathcal{A}u + \mathcal{B}u, \quad (\alpha \mathcal{A})u = \alpha(\mathcal{A}u) \quad \text{для всех } u \in U.$$

Пусть теперь W – еще одно пространство над полем k ; произведением отображений $\mathcal{A} : V \rightarrow W$, $\mathcal{B} : U \rightarrow V$ называется отображение $\mathcal{AB} : U \rightarrow W$, заданное формулой

$$(\mathcal{AB})u = \mathcal{A}(\mathcal{B}u) \quad \text{для всех } u \in U.$$

Мы видим, что, как и для матриц, сумма и произведение отображений не всегда определены: сумма определена, если "начала" и "концы" слагаемых одинаковы, а произведение – если "начало" первого сомножителя совпадает с "концом" второго.

Предложение 4. Сумма, произведение линейных отображений (если они определены) и произведение элемента из основного поля на линейное отображение – снова линейные отображения.

Доказательство. Пусть \mathcal{A}, \mathcal{B} – линейные отображения, $\alpha \in k$, и пусть U – то пространство, из которого действует отображение \mathcal{B} . Для любых $u, u' \in U$ и любого $\beta \in k$ получаем, пользуясь линейностью отображений и определениями действий над ними (если в первом случае определена сумма $\mathcal{A} + \mathcal{B}$, а в третьем – произведение $\mathcal{A}\mathcal{B}$):

$$\begin{aligned} (\mathcal{A} + \mathcal{B})(\beta u + u') &= \mathcal{A}(\beta u + u') + \mathcal{B}(\beta u + u') = \beta \mathcal{A}u + \mathcal{A}u' + \beta \mathcal{B}u + \mathcal{B}u' = \\ &= \beta(\mathcal{A}u + \mathcal{B}u) + (\mathcal{A}u' + \mathcal{B}u') = \beta(\mathcal{A} + \mathcal{B})u + (\mathcal{A} + \mathcal{B})u'; \\ (\alpha \mathcal{A})(\beta u + u') &= \alpha(\mathcal{A}(\beta u + u')) = \alpha(\beta \mathcal{A}u + \mathcal{A}u') = \beta(\alpha \mathcal{A}u) + \alpha \mathcal{A}u' = \\ &= \beta(\alpha \mathcal{A})u + (\alpha \mathcal{A})u'; \\ (\mathcal{A}\mathcal{B})(\beta u + u') &= \mathcal{A}(\mathcal{B}(\beta u + u')) = \mathcal{A}(\beta \mathcal{B}u + \mathcal{B}u') = \beta \mathcal{A}(\mathcal{B}u) + \mathcal{A}(\mathcal{B}u') = \\ &= \beta(\mathcal{A}\mathcal{B})u + (\mathcal{A}\mathcal{B})u'. \end{aligned}$$

Эти соотношения и означают, что все рассматриваемые отображения линейны.

Действия над отображениями обладают теми же свойствами, что и действия над матрицами, только некоторые из них выглядят естественнее и доказываются проще, чем соответствующие свойства действий над матрицами. Приведем эти свойства. Всюду через $\mathcal{A}, \mathcal{B}, \mathcal{C}$ обозначены линейные отображения, а через α, β – элементы поля k . Как и для свойств действий над матрицами, многие из этих свойств надо читать так: если определена одна из частей равенства, то определена и другая, и они равны.

- (1) $(\mathcal{A} + \mathcal{B}) + \mathcal{C} = \mathcal{A} + (\mathcal{B} + \mathcal{C})$;
- (2) $\mathcal{A} + \mathcal{B} = \mathcal{B} + \mathcal{A}$;
- (3) $\mathcal{A} + \mathcal{O} = \mathcal{A}$;
- (4) $\mathcal{A} + (-1)\mathcal{A} = \mathcal{O}$;
- (5) $\alpha(\mathcal{A} + \mathcal{B}) = \alpha\mathcal{A} + \alpha\mathcal{B}$;
- (6) $(\alpha + \beta)\mathcal{A} = \alpha\mathcal{A} + \beta\mathcal{A}$;
- (7) $1 \cdot \mathcal{A} = \mathcal{A}$;
- (8) $(\alpha\beta)\mathcal{A} = \alpha(\beta\mathcal{A})$;
- (9) $(\mathcal{A}\mathcal{B})\mathcal{C} = \mathcal{A}(\mathcal{B}\mathcal{C})$;
- (10) $\mathcal{A}(\mathcal{B} + \mathcal{C}) = \mathcal{A}\mathcal{B} + \mathcal{A}\mathcal{C}$;
- (11) $(\mathcal{A} + \mathcal{B})\mathcal{C} = \mathcal{A}\mathcal{C} + \mathcal{B}\mathcal{C}$;
- (12) $\mathcal{E}\mathcal{A} = \mathcal{A}, \mathcal{A}\mathcal{E} = \mathcal{A}$;
- (13) $\alpha(\mathcal{A}\mathcal{B}) = (\alpha\mathcal{A})\mathcal{B} = \mathcal{A}(\alpha\mathcal{B})$.

Покажем, например, что умножение линейных отображений ассоциативно. Пусть $\mathcal{A} : W \rightarrow P, \mathcal{B} : V \rightarrow W, \mathcal{C} : U \rightarrow V$ – линейные отображения; тогда для любого $u \in U$

$$((\mathcal{A}\mathcal{B})\mathcal{C})u = (\mathcal{A}\mathcal{B})(\mathcal{C}u) = \mathcal{A}(\mathcal{B}(\mathcal{C}u)) = \mathcal{A}(\mathcal{B}\mathcal{C})u = (\mathcal{A}(\mathcal{B}\mathcal{C}))u,$$

т.е. отображения $(\mathcal{A}\mathcal{B})\mathcal{C}$ и $\mathcal{A}(\mathcal{B}\mathcal{C})$ совпадают.

Изоморфизм векторных пространств. Пусть U, V – векторные пространства над полем k . Всякое биективное линейное отображение $\mathcal{A} : U \rightarrow V$ называется изоморфизмом пространства U на пространство V ; если изоморфизм U на V существует, то говорят, что пространства U и V изоморфны.

Пусть $\mathcal{A} : U \rightarrow V$ – изоморфизм векторных пространств над полем k ; тогда отображение \mathcal{A} биективно, и поэтому существует обратное отображение $\mathcal{B} : V \rightarrow U$, т.е. такое отображение, что для всех $u \in U, v \in V$ выполняются соотношения $\mathcal{A}(\mathcal{B}v) = v, \mathcal{B}(\mathcal{A}u) = u$. Как известно, обратное отображение \mathcal{B} тоже биективно; покажем, что оно еще и линейно, т.е. тоже является изоморфизмом векторных пространств. Действительно, для любых векторов $v, v' \in V$ и любого $\alpha \in k$ мы

имеем:

$$\mathcal{A}(\mathcal{B}(\alpha v + v')) = \alpha v + v' = \alpha \mathcal{A}(\mathcal{B}v) + \mathcal{A}(\mathcal{B}v') = \mathcal{A}(\alpha \mathcal{B}v + \mathcal{B}v').$$

Поскольку \mathcal{A} – инъективное отображение, отсюда следует, что $\mathcal{B}(\alpha v + v') = \alpha \mathcal{B}v + \mathcal{B}v'$ для любых $v, v' \in V$, $\alpha \in k$, а это и означает, что отображение \mathcal{B} линейно.

Изоморфные пространства одинаковы по своей структуре, хотя природа их элементов может быть различна (пример – пространство "геометрических" векторов и \mathbb{R}^3). В частности, если $\mathcal{A} : U \rightarrow V$ – изоморфизм векторных пространств над полем k , и U_0 – конечномерное подпространство пространства U , то пространство $\mathcal{A}U_0 \subseteq V$ тоже конечномерно и его размерность равна размерности U_0 . Действительно, пусть $\mathcal{A}_0 : U_0 \rightarrow V$ – ограничение на U_0 отображения \mathcal{A} ; очевидно, что отображение \mathcal{A}_0 линейно, $\text{Ker } \mathcal{A}_0 = \text{Ker } \mathcal{A} = \{\theta\}$, $\text{Im } \mathcal{A}_0 = \mathcal{A}_0 U_0 = \mathcal{A}U_0$, и потому

$$\dim \mathcal{A}U_0 = \dim \text{Im } \mathcal{A}_0 = \dim U_0 - \dim \text{Ker } \mathcal{A}_0 = \dim U_0.$$

Предложение 5. Пусть $\mathcal{A} : U \rightarrow V$ – изоморфизм векторных пространств над полем k .

- (1) Если векторы $u_1, \dots, u_m \in U$ порождают подпространство U_0 пространства U , то векторы $\mathcal{A}u_1, \dots, \mathcal{A}u_m$ порождают подпространство $\mathcal{A}U_0$ пространства V .
- (2) Если векторы $u_1, \dots, u_m \in U$ линейно независимы, то и векторы $\mathcal{A}u_1, \dots, \mathcal{A}u_m \in V$ линейно независимы.
- (3) Если U_0 – конечномерное подпространство пространства U , то $\mathcal{A}U_0$ – конечномерное подпространство пространства V , и $\dim \mathcal{A}U_0 = \dim U_0$.
- (4) Если пространство U конечномерно, то пространство V конечномерно, и $\dim U = \dim V$.

Доказательство. (1) Пусть U_0 порождается векторами u_1, \dots, u_m ; тогда $U_0 = \{\alpha_1 u_1 + \dots + \alpha_n u_n \mid \alpha_1, \dots, \alpha_n \in k\}$, и потому

$$\begin{aligned} \mathcal{A}U_0 &= \{\mathcal{A}u \mid u \in U_0\} = \{\mathcal{A}(\alpha_1 u_1 + \dots + \alpha_n u_n) \mid \alpha_1, \dots, \alpha_n \in k\} = \\ &= \{\alpha_1 \mathcal{A}u_1 + \dots + \alpha_n \mathcal{A}u_n \mid \alpha_1, \dots, \alpha_n \in k\}, \end{aligned}$$

т.е. $\mathcal{A}U_0$ совпадает с множеством линейных комбинаций векторов $\mathcal{A}u_1, \dots, \mathcal{A}u_m$.

(2) Если векторы $\mathcal{A}u_1, \dots, \mathcal{A}u_m$ линейно зависимы, то существуют такие $\alpha_1, \dots, \alpha_n \in k$, не все равные 0, что $\alpha_1 \mathcal{A}u_1 + \dots + \alpha_n \mathcal{A}u_n = \theta$. Тогда

$$\mathcal{A}(\alpha_1 u_1 + \dots + \alpha_n u_n) = \alpha_1 \mathcal{A}u_1 + \dots + \alpha_n \mathcal{A}u_n = \theta = \mathcal{A}(\theta),$$

и, поскольку \mathcal{A} – инъективное отображение, $\alpha_1 u_1 + \dots + \alpha_n u_n = \theta$; но это означает, что векторы u_1, \dots, u_m линейно зависимы.

(3) Если u_1, \dots, u_m – базис U_0 , то векторы $\mathcal{A}u_1, \dots, \mathcal{A}u_m$ порождают подпространство $\mathcal{A}U_0$ по свойству (1) и линейно независимы по свойству (2); таким образом, $\mathcal{A}u_1, \dots, \mathcal{A}u_m$ – базис $\mathcal{A}U_0$, и потому $\dim \mathcal{A}U_0 = m = \dim U_0$.

(4) Поскольку отображение \mathcal{A} сюръективно, $\mathcal{A}U = V$; по свойству (3) тогда $\dim V = \dim \mathcal{A}U = \dim U$.

Укажем важнейший пример изоморфизма векторных пространств. Пусть V – конечномерное векторное пространство над полем k , и пусть v_1, \dots, v_n – его базис. Тогда отображение Φ_{v_1, \dots, v_n} , сопоставляющее каждому столбцу $X = (\alpha_1, \dots, \alpha_n)^T$ вектор

$$(v_1, \dots, v_n)X = (v_1, \dots, v_n) \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1 v_1 + \dots + \alpha_n v_n,$$

как и обратное к нему отображение Ψ_{v_1, \dots, v_n} , сопоставляющее каждому вектору из V столбец его координат, являются изоморфизмами векторных пространств.

Матрица линейного отображения. Пусть U, V – конечномерные векторные пространства над полем k , и пусть $\mathcal{A} : U \rightarrow V$ – линейное отображение. Выберем базисы u_1, \dots, u_n и v_1, \dots, v_m пространств U и V ; векторы $\mathcal{A}u_1, \dots, \mathcal{A}u_n$ принадлежат пространству V и потому являются линейными комбинациями элементов базиса v_1, \dots, v_m . Таким образом, существуют (очевидно, единственные) элементы $A_{ij} \in k$, такие что

[illegible]

Матрица

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}$$

называется матрицей линейного отображения \mathcal{A} в базисах u_1, \dots, u_n и v_1, \dots, v_m . Обратим внимание на то, что элементы в матрице A расположены не там, где они были в предшествующей системе равенств, а на симметричных относительно диагонали позициях.

Определение матрицы линейного отображения удобно записать в матричной форме: A – это такая матрица с компонентами из поля k , что

$$\mathcal{A}(u_1, \dots, u_n) = (\mathcal{A}u_1, \dots, \mathcal{A}u_n) = (v_1, \dots, v_m)A$$

(поскольку векторы v_1, \dots, v_m линейно независимы, матрица A определяется этим соотношением однозначно).

Предложение 6. Пусть u_1, \dots, u_n и v_1, \dots, v_m – базисы векторных пространств U и V над полем k , и пусть $A: U \rightarrow V$ – линейное отображение, а A – матрица отображения A в выбранных базисах. Далее, пусть $u \in U$, а $X = (x_1, \dots, x_n)^T$, $Y = (y_1, \dots, y_m)^T$ – столбцы координат векторов $u \in U$, $Au \in V$ в тех же базисах. Тогда $Y = AX$.

Доказательство. По определениям столбцов координат и матрицы линейного отображения мы имеем следующие соотношения:

$$u = (u_1, \dots, u_n)X, \quad \mathcal{A}u = (v_1, \dots, v_m)Y, \quad \mathcal{A}(u_1, \dots, u_n) = (v_1, \dots, v_m)A;$$

пользуясь ассоциативностью умножения матриц, получаем отсюда:

$$\begin{aligned}(v_1, \dots, v_m)Y &= \mathcal{A}u = \mathcal{A}((u_1, \dots, u_n)X) = (\mathcal{A}(u_1, \dots, u_n))X = \\ &= ((v_1, \dots, v_m)A)X = (v_1, \dots, v_m)(AX).\end{aligned}$$

Из того, что произведения строки линейно независимых векторов (v_1, \dots, v_m) на столбцы $Y, AX \in k^m$ совпадают, следует, что $Y = AX$.

Предложение 7. Пусть u_1, \dots, u_n и v_1, \dots, v_m – базисы векторных пространств U и V над полем k , и пусть $A : U \rightarrow V$ – линейное отображение, а A – матрица отображения A в выбранных базисах. Обозначим через $A' : k^n \rightarrow k^m$ отображение, сопоставляющее каждому столбцу $X \in k^n$ столбец $AX \in k^m$. Тогда отображения $\Phi_{v_1, \dots, v_m} A'$ и $A \Phi_{u_1, \dots, u_n}$ пространства k^n в пространство V совпадают.

Доказательство. Пусть $X \in k^n$ – столбец координат вектора $u \in U$ в базисе u_1, \dots, u_n , $Y \in k^m$ – столбец координат вектора $Au \in V$ в базисе v_1, \dots, v_m . Из определения изоморфизмов Φ_{u_1, \dots, u_n} , Φ_{v_1, \dots, v_m} следует, что

$$u = \Phi_{u_1, \dots, u_n} X, \quad Au = \Phi_{v_1, \dots, v_m} Y,$$

а из предыдущего предложения – что $Y = AX$. Таким образом, для любого столбца $X \in k^n$

$$A \Phi_{u_1, \dots, u_n} X = Au = \Phi_{v_1, \dots, v_m} Y = \Phi_{v_1, \dots, v_m} AX = \Phi_{v_1, \dots, v_m} A' X.$$

Доказанное утверждение удобно проиллюстрировать следующей диаграммой, на которой для краткости изоморфизмы Φ_{u_1, \dots, u_n} и Φ_{v_1, \dots, v_m} обозначены через $\Phi_{[u]}$ и $\Phi_{[v]}$:

$$(*) \quad \begin{array}{ccc} k^n & \xrightarrow{A'} & k^m \\ \Phi_{[u]} \downarrow & & \downarrow \Phi_{[v]} \\ U & \xrightarrow{A} & V. \end{array}$$

Двигаясь по стрелкам, можно двумя способами попасть из k^n в V ; по доказанному предложению, оба способа дают одно и то же отображение. Такие диаграммы называются коммутативными, а само утверждение можно кратко сформулировать так: диаграмма $(*)$ коммутативна.

Следствие. Пусть $A : U \rightarrow V$ – линейное отображение одного конечномерного пространства в другое, а A – матрица отображения A в каких-то базисах пространств U, V . Тогда $\text{rank } A = \dim \text{Im } A$.

Доказательство. В обозначениях предложения 7 покажем, что $\text{Im } A = \Phi_{[v]} \text{Im } A'$; поскольку, как мы уже знаем, $\dim \text{Im } A' = \text{rank } A$, и поскольку $\Phi_{[v]}$ – изоморфизм, отсюда будет следовать, что $\dim \text{Im } A = \dim \text{Im } A' = \text{rank } A$. Но соотношение $\text{Im } A = \Phi_{[v]} \text{Im } A'$ сразу следует из предложения 7; действительно, отображение $\Phi_{[u]}$ сюръективно, и потому

$$\begin{aligned} \text{Im } A &= \{Au \mid u \in U\} = \{A \Phi_{[u]} X \mid X \in k^n\} = \\ &= \{\Phi_{[v]} A' X \mid X \in k^n\} = \{\Phi_{[v]} Y \mid Y \in \text{Im } A'\}. \end{aligned}$$

Предложение 8. Пусть $A : U \rightarrow V$ – линейное отображение векторных пространств над полем k , и пусть $u_1, \dots, u_n, u'_1, \dots, u'_n$ – два базиса пространства U , а $v_1, \dots, v_m, v'_1, \dots, v'_m$ – два базиса пространства V . Далее, пусть C – матрица перехода от базиса u_1, \dots, u_n к базису u'_1, \dots, u'_n , D – матрица перехода от базиса v_1, \dots, v_m к базису v'_1, \dots, v'_m , A – матрица отображения A в базисах $u_1, \dots, u_n, v_1, \dots, v_m$, A' – матрица отображения A в базисах $u'_1, \dots, u'_n, v'_1, \dots, v'_m$. Тогда $A' = D^{-1}AC$.

Доказательство. По определениям матрицы перехода от одного базиса к другому и матрицы линейного отображения мы имеем следующие соотношения:

$$\begin{aligned} (u'_1, \dots, u'_n) &= (u_1, \dots, u_n)C, & (v'_1, \dots, v'_m) &= (v_1, \dots, v_m)D, \\ A(u_1, \dots, u_n) &= (v_1, \dots, v_m)A, & A(u'_1, \dots, u'_n) &= (v'_1, \dots, v'_m)A'; \end{aligned}$$

пользуясь ассоциативностью умножения матриц, получаем отсюда:

$$\begin{aligned}\mathcal{A}(u'_1, \dots, u'_n) &= \mathcal{A}((u_1, \dots, u_n)C) = (\mathcal{A}(u_1, \dots, u_n))C = \\ &= ((v_1, \dots, v_m)A)C = (v_1, \dots, v_m)(AC).\end{aligned}$$

С другой стороны,

$$\mathcal{A}(u'_1, \dots, u'_n) = (v'_1, \dots, v'_m)A' = ((v_1, \dots, v_m)D)A' = (v_1, \dots, v_m)(DA').$$

Из того, что произведения строки линейно независимых векторов (v_1, \dots, v_m) на матрицы $AC, DA' \in k^{m \times n}$ совпадают, следует, что $AC = DA'$, т.е.

$$A' = (D^{-1}D)A' = D^{-1}(DA') = D^{-1}AC.$$

Предложение 9. Пусть $\mathcal{A}, \mathcal{B} : U \rightarrow V$ – два линейных отображения векторного пространства U в векторное пространство V над тем же полем k , и пусть $u_1, \dots, u_n, v_1, \dots, v_m$ – базисы пространств U и V . Пусть, далее, A, B – матрицы отображений \mathcal{A}, \mathcal{B} в выбранных базисах. Тогда матрица суммы $\mathcal{A} + \mathcal{B}$ в тех же базисах равна $A + B$.

Доказательство. По определению матрицы линейного отображения мы имеем следующие соотношения:

$$\mathcal{A}(u_1, \dots, u_n) = (v_1, \dots, v_m)A, \quad \mathcal{B}(u_1, \dots, u_n) = (v_1, \dots, v_m)B.$$

Вспоминая определение суммы отображений, получаем:

$$\begin{aligned}(\mathcal{A} + \mathcal{B})(u_1, \dots, u_n) &= ((\mathcal{A} + \mathcal{B})u_1, \dots, (\mathcal{A} + \mathcal{B})u_n) = (\mathcal{A}u_1 + \mathcal{B}u_1, \dots, \mathcal{A}u_n + \mathcal{B}u_n) = \\ &= (\mathcal{A}u_1, \dots, \mathcal{A}u_n) + (\mathcal{B}u_1, \dots, \mathcal{B}u_n) = \\ &= (v_1, \dots, v_m)A + (v_1, \dots, v_m)B = (v_1, \dots, v_m)(A + B),\end{aligned}$$

а это и значит, что $A + B$ – матрица линейного отображения $\mathcal{A} + \mathcal{B}$.

Предложение 10. Пусть \mathcal{A} – линейное отображение векторного пространства U в векторное пространство V над тем же полем k , и пусть $u_1, \dots, u_n, v_1, \dots, v_m$ – базисы пространств U и V . Пусть, далее, A – матрица отображения \mathcal{A} в выбранных базисах. Тогда для любого $a \in k$ матрица отображения $a\mathcal{A}$ в тех же базисах равна aA .

Доказательство аналогично доказательству предыдущего предложения.

Предложение 11. Пусть U, V, W – векторные пространства над одним и тем же полем k , и пусть $\mathcal{A} : V \rightarrow W, \mathcal{B} : U \rightarrow V$ – линейные отображения. Пусть, далее, $u_1, \dots, u_n, v_1, \dots, v_m, w_1, \dots, w_p$ – базисы пространств U, V, W . Если A – матрица отображения \mathcal{A} в базисах $v_1, \dots, v_m, w_1, \dots, w_p$, а B – матрица отображения \mathcal{B} в базисах $u_1, \dots, u_n, v_1, \dots, v_m$, то матрица произведения отображений \mathcal{A}, \mathcal{B} в базисах $u_1, \dots, u_n, w_1, \dots, w_p$ равна произведению AB матриц сомножителей.

Доказательство. По определению матрицы линейного отображения мы имеем следующие соотношения:

$$\mathcal{A}(v_1, \dots, v_m) = (w_1, \dots, w_p)A, \quad \mathcal{B}(u_1, \dots, u_n) = (v_1, \dots, v_m)B.$$

Вспоминая определение произведения отображений, получаем:

$$\begin{aligned}(\mathcal{A}\mathcal{B})(u_1, \dots, u_n) &= \mathcal{A}(\mathcal{B}(u_1, \dots, u_n)) = \mathcal{A}((v_1, \dots, v_m)B) = (\mathcal{A}(v_1, \dots, v_m))B = \\ &= ((w_1, \dots, w_p)A)B = (w_1, \dots, w_p)(AB),\end{aligned}$$

а это и значит, что AB – матрица линейного отображения $\mathcal{A}\mathcal{B}$.

Доказанные предложения показывают, что казавшиеся искусственными определения действий над матрицами на самом деле являются реализациями очень естественных действий над отображениями.