Но A^T – тоже ортогональна. Её строки – это столбцы матрицы A. Таким образом, столбцы матрицы A тоже нормированы и попарно ортогональны ■

Легко доказать и обратное утверждение, **если строки (столбцы) матрицы ортонормированы, то матрица будет ортогональной**. (Достаточно проделать обратный ход предыдущего доказательства).

Построение ортогональных матриц.

Докажем теорему, позволяющую дополнять ортонормированную систему векторов ещё одним вектором, не нарушая свойства ортонормированности.

Теорема. Пусть $X_1, \ldots, X_k$ – вещественные нормированные попарно ортогональные столбцы высоты n, и пусть $k < n$. Тогда существует нормированный столбец X_{k+1} ортогональный столбцам $X_1, \ldots, X_k$.

Доказательство. Пусть

$$
\begin{aligned}
&X_1 = (x_{11}, \ldots, x_{n1})^T \\
&\ldots\ldots\ldots\ldots \\
&X_k = (x_{1k}, \ldots, x_{nk})^T \\
&X_{k+1} = (z_1, \ldots, z_n)^T.
\end{aligned}
$$

Запишем требования ортогональности столбцов X_1 и X_{k+1}, X_2 и X_{k+1}, $\ldots$ X_k и X_{k+1} и нормированности столбца X_{k+1} в виде системы уравнений

$$
\begin{cases}
x_{11}z_1 + x_{21}z_2 + \ldots + x_{n1}z_n = 0, \\
x_{12}z_1 + x_{22}z_2 + \ldots + x_{n2}z_n = 0, \\
\ldots\ldots\ldots\ldots\ldots \\
x_{1k}z_1 + x_{2k}z_2 + \ldots + x_{nk}z_n = 0, \\
z_1^2 + z_2^2 + \ldots + z_n^2 = 1.
\end{cases}
$$

Первые k уравнений образуют линейную однородную систему, причем число уравнений k меньше числа неизвестных n. Поэтому система имеет нетривиальные решения. Пусть $z_1^*, z_2^*, \ldots, z_n^*$ – одно из них. Если ${z_1^*}^2 + {z_2^*}^2 + \ldots + {z_n^*}^2 = r^2$, то $\frac{1}{r}z_1^*, \frac{1}{r}z_2^*, \ldots, \frac{1}{r}z_n^*$ – решение ВСЕЙ системы (напомним, что умножая на константу решение однородной системы, мы получаем решение этой же системы. В данном случае новое решение к тому же ещё будет удовлетворять и последнему уравнению). Таким образом, $X_{k+1} = \left(\frac{1}{r}z_1^*, \frac{1}{r}z_2^*, \ldots, \frac{1}{r}z_n^*\right)^T$ ■

Следствие 1. Заметим, что при $k = n$ столбцы составляют ортогональную матрицу. Эта матрица невырожденная, а значит, система для определения X_{k+1} окажется несовместной (СА-

МОСТОЯТЕЛЬНО ОБЪЯСНИТЬ ПОЧЕМУ), так что более чем n попарно ортогональных нормированных столбцов не может существовать.

Следствие 2. Любую прямоугольную матрицу размерности $n \times m$ $(m < n)$, состоящую из попарно ортогональных нормированных столбцов (их количество меньше их высоты), можно дополнить до ортогональной матрицы (это будет уже квадратная матрица размерности $n \times n$). Действительно, к исходной матрице можно присоединять по одному новые столбцы до тех пор, пока не получим ортогональную матрицу. Алгоритм присоединения указан в доказательстве теоремы (решаем соответствующую систему). В частности, матрицу-столбец (её размерность $n \times 1$) можно дополнить до ортогональной матрицы, если этот столбец является нормированным. Последнее означает, что ЛЮБОЙ нормированный столбец может быть принят за первый столбец ортогональной матрицы.

Следствие 3. Теорема и два предыдущих следствия из неё остаются верны и для строк.

Ортогональное преобразование квадратичной формы к каноническому виду

С точки зрения геометрии линейное преобразование $X = PY$ с ортогональной матрицей P (ортогональное преобразование) представляет наибольший интерес. Дело в том, что одним из распространённых преобразований на плоскости (в пространстве) является преобразование поворота прямоугольной декартовой системы координат вокруг начала координат (Формулы поворота мы рассматривали в прошлом семестре. НАЙТИ ЭТИ ФОРМУЛЫ. Также см. пример ортогональной матрицы в самом начале файла). В результате получается, что если мы ищем ортогональное преобразование, то с геометрической точки зрения мы ищем преобразование, являющееся суперпозицией указанного поворота и изменением на противоположное направления одной из координатных осей.

Исследуем возможность ортогонального преобразования квадратичной формы к каноническому виду.

Теорема. Вещественная квадратичная форма может быть приведена к каноническому виду посредством преобразования переменных с ортогональной матрицей.

Доказательство. Применим метод математической индукции по числу n переменных. При $n = 1$ форма имеет вид $a_{11}x_1^2$ – она уже в каноническом виде. В качестве ортогонального преобразования здесь можно взять тождественное преобразование. Т.е. база индукции выполнена.

Допустим, что теорема уже доказана для форм от $n-1$ переменных. Докажем, что тогда теорема выполняется для формы от n переменных.

Пусть $f(x_1, \ldots, x_n) = X^T A X$, где $X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$ $A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$ – вещественная симметричная матрица. Пусть $X_1 = (p_{11}, p_{21}, \ldots, p_{n1})^T$ – какой-нибудь НОРМИРОВАННЫЙ собственный вектор матрицы A, соответствующий некоторому собственному λ_1. Т.е. имеет место равенство

$$AX_1 = \lambda_1 X_1,$$

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} p_{11} \\ \ldots \\ p_{n1} \end{pmatrix} = \lambda_1 \begin{pmatrix} p_{11} \\ \ldots \\ p_{n1} \end{pmatrix},$$

$$\begin{pmatrix} a_{11}p_{11} + a_{12}p_{21} + \ldots + a_{1n}p_{n1} \\ a_{21}p_{11} + a_{22}p_{21} + \ldots + a_{2n}p_{n1} \\ \ldots \\ a_{n1}p_{11} + a_{n2}p_{21} + \ldots + a_{nn}p_{n1} \end{pmatrix} = \begin{pmatrix} \lambda_1 p_{11} \\ \lambda_1 p_{21} \\ \ldots \\ \lambda_1 p_{n1} \end{pmatrix}. \tag{2}$$

Дополним этот вектор столбцами так, чтобы получилась ОРТОГОНАЛЬНАЯ матрица P – это всегда можно сделать (см. следствие 2 из предыдущей теоремы).

$$P = \begin{pmatrix} p_{11} & p_{12} & \ldots & p_{1n} \\ p_{21} & p_{22} & \ldots & p_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ p_{n1} & p_{n2} & \ldots & p_{nn} \end{pmatrix}.$$

Рассмотрим линейное преобразование $X = PY$ исходной квадратичной формы. Покажем, что преобразование с этой матрицей "улучшает" матрицу квадратичной формы. Матрица преобразованной формы есть $P^T AP$. Вычисляем

$$AP = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} p_{11} & \ldots & p_{1n} \\ \ldots & \ldots & \ldots \\ p_{n1} & \ldots & p_{nn} \end{pmatrix} = [\text{распишем только первый столбец}] =$$

$$= \begin{pmatrix} a_{11}p_{11} + a_{12}p_{21} + \ldots + a_{1n}p_{n1} & \ldots \\ a_{21}p_{11} + a_{22}p_{21} + \ldots + a_{2n}p_{n1} & \ldots \\ \ldots & \ldots \\ a_{n1}p_{11} + a_{n2}p_{21} + \ldots + a_{nn}p_{n1} & \ldots \end{pmatrix} = [\text{соотношение (2)}] = \begin{pmatrix} \lambda_1 p_{11} & \ldots \\ \lambda_1 p_{21} & \ldots \\ \ldots & \ldots \\ \lambda_1 p_{n1} & \ldots \end{pmatrix}.$$

Считаем далее

$$P^TAP = \begin{pmatrix} p_{11} & \dots & p_{n1} \\ \dots & \dots & \dots \\ p_{1n} & \dots & p_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 p_{11} & \dots \\ \lambda_1 p_{21} & \dots \\ \dots & \dots \\ \lambda_1 p_{n1} & \dots \end{pmatrix} = [\text{распишем только первый столбец}] =$$

$$= \begin{pmatrix} \lambda_1(p_{11}^2 + p_{21}^2 + \ldots + p_{n1}^2) & \dots \\ \lambda_1(p_{12}p_{11} + p_{22}p_{21} + \ldots + p_{n2}p_{n1}) & \dots \\ \dots & \dots \\ \lambda_1(p_{1n}p_{11} + p_{2n}p_{21} + \ldots + p_{nn}p_{n1}) & \dots \end{pmatrix} = [(*)] = \begin{pmatrix} \lambda_1 & \dots \\ 0 & \dots \\ \dots & \dots \\ 0 & \dots \end{pmatrix}$$

(*) – по предыдущей теореме столбцы ортогональной матрицы ортонормированы, следовательно $(p_{11}^2 + p_{21}^2 + \ldots + p_{n1}^2)$ – скалярный квадрат первого столбца матрицы P и эта сумма равна 1; $(p_{12}p_{11} + p_{22}p_{21} + \ldots + p_{n2}p_{n1})$ – скалярное произведение 1-го и 2-го столбца матрицы P и оно равно 0; и т.д. $(p_{1n}p_{11} + p_{2n}p_{21} + \ldots + p_{nn}p_{n1})$ – скалярное произведение 1-го и n-го столбца матрицы P и оно равно 0.

Матрица P^TAT – симметрична (это будет матрица формы в новых переменных Y), поэтому имеет вид (первая строчка повторяет первый столбец)

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & b_{n2} & \dots & b_{nn} \end{pmatrix},$$

где $B = \begin{pmatrix} b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots \\ b_{n2} & \dots & b_{nn} \end{pmatrix}$ – симметричная матрица.

Рассмотрим квадратичную форму от $n-1$ переменной с матрицей B. В силу индуктивного предположения, существует ортогональная матрица Q, такая что $Q^TBQ = \operatorname{diag}(\lambda_2, \ldots, \lambda_n)$. Положим $Q_1 = \begin{pmatrix} 1 & \mathbb{O} \\ \mathbb{O} & Q \end{pmatrix}$. Очевидно, Q_1 – ортогональна (так как её столбцы нормированы и попарно ортогональны).

Применим теперь к исходной квадратичной форме с матрицей A преобразование с матрицей PQ_1. Вычисляем матрицу формы после этого преобразования (вычисляем непосредственно, но вычисления лучше записывать в виде блочных матриц):

$$(PQ_1)^T A(PQ_1) = (Q_1^T P^T)A(PQ_1) = Q_1^T (P^T AP)Q_1 = Q_1^T \begin{pmatrix} \lambda_1 & \mathbb{O} \\ \mathbb{O} & B \end{pmatrix} Q_1 =$$
$$= \begin{pmatrix} 1 & \mathbb{O} \\ \mathbb{O} & Q^T \end{pmatrix} \begin{pmatrix} \lambda_1 & \mathbb{O} \\ \mathbb{O} & B \end{pmatrix} \begin{pmatrix} 1 & \mathbb{O} \\ \mathbb{O} & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & \mathbb{O} \\ \mathbb{O} & Q^T B \end{pmatrix} \begin{pmatrix} 1 & \mathbb{O} \\ \mathbb{O} & Q \end{pmatrix} =$$
$$= \begin{pmatrix} \lambda_1 & \mathbb{O} \\ \mathbb{O} & Q^T BQ \end{pmatrix} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Так как PQ_1 – ортогональная матрица (как произведение ортогональных – Свойство 3 ортогональных матриц), а матрица новой формы получилась диагональной, то теорема доказана ■

Коэффициенты канонического вида квадратичной формы и столбцы преобразующей ортогональной матрицы.

В предыдущем разделе мы доказали принципиальную возможность приведения квадратичной формы к каноническому виду ортогональным преобразованием, но не сказали, какова структура матрицы преобразования и чему равны коэффициенты канонического вида квадратичной формы. Хотя небольшой намёк был сделан: коэффициенты канонического вида связаны с собственными числами матрицы формы, а столбцы матрицы преобразования – собственные векторы. Это действительно так. Докажем соответствующие утверждения.

Итак, пусть P – ортогональная матрица преобразования формы $f(X) = X^T AX$ к каноническому виду, т.е. $P^T AP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Покажем, что матрицы A и $P^T AP$ имеют одинаковые характеристические многочлены. Характеристический многочлен матрицы $P^T AP$ равен (напомним, что у ортогональных матриц $P^T = P^{-1}$)

$$\det(P^T AP - \lambda E) = \det(P^T AP - \lambda(P^{-1}EP)) = \det(P^T AP - P^T(\lambda E)P) =$$
$$= \det\left(P^T(A - \lambda E)P\right) = \det P^T \det(A - \lambda E)\det P = \det P^T \det P \det(A - \lambda E) =$$
$$= \det(P^T P)\det(A - \lambda E) = \det(E)\det(A - \lambda E) = \det(A - \lambda E).$$

Значит, характеристический многочлен матрицы A исходной формы и матрицы формы $P^T AP$ в новых переменных совпадают. (Здесь в доказательстве важно, что P ортогональная матрица.)

Характеристический многочлен диагональной матрицы $P^T AP = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ равен

$$\begin{vmatrix} \lambda_1-\lambda & 0 & \dots & 0 \\ 0 & \lambda_2-\lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n-\lambda \end{vmatrix} = (\lambda_1-\lambda)(\lambda_2-\lambda)\dots(\lambda_n-\lambda).$$

Таким образом, каково бы ни было ортогональное преобразование квадратичной формы к каноническому виду, **коэффициенты этого канонического вида равны собственным значениям матрицы квадратичной формы, причем каждое собственное значение повторяется столько раз, какова его кратность как корня характеристического полинома**.

Теперь найдем, чему равны столбцы матрицы P (напомним $P^T = P^{-1}$).

Преобразуем равенство

$$P^T AP = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Домножим слева

$$P * | \quad P^T AP = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

$$AP = P\mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Если обозначить P_i $(i = 1 \dots n)$ – столбцы матрицы P, то последнее равенство можно записать в более удобном для нас виде – в виде соотношения между блочными матрицами

$$A(P_1, P_2, \dots, P_n) = (P_1, P_2, \dots, P_n)\mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

или, выполнив умножение

$$(AP_1, AP_2, \dots, AP_n) = (\lambda_1 P_1, \lambda_2 P_2, \dots, \lambda_n P_n).$$

Приравнивая одинаковые блоки в последнем равенстве, получим,

$$AP_i = \lambda_i P_i, \quad (i = 1 \dots n).$$

Итак, **столбцы преобразующей ортогональной матрицы являются СООТВЕТСТВУЮЩИМИ собственными векторами матрицы квадратичной формы**.

Казалось бы, мы ответили на вопрос, как искать матрицу ортогонального преобразования и чему равны коэффициенты канонического вида формы. Вроде бы достаточно найти собственные значения и соответствующие им собственные векторы. Но может получиться одна неприятность: мы знаем, что столбцы матрицы ортогонального преобразования должны быть ортогональны и нормированы. Нормировать мы их можем всегда, но вот ортогональными они вроде как быть не обязаны. Оказывается, что эта неприятность возникает, только если имеются кратные собственные значения. А именно, верна следующая теорема.

Теорема. Собственные векторы вещественной симметричной матрицы, соответствующие РАЗЛИЧНЫМ собственным значениям ортогональны.

Доказательство. Заметим сначала, что условие ортогональности двух столбцов

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \quad \text{и} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}$$

можно записать через операцию произведения матриц как $X^T Y = 0$ или $Y^T X = 0$.

Пусть A – вещественная симметричная матрица. X_1 и X_2 – ее собственные векторы, соответствующие различным её собственным значениям $\lambda_1 \neq \lambda_2$. Вычислим двумя способами число $a = X_2^T A X_1$ (то, что это матрица размерности 1×1, т.е. по сути просто число, убедиться самостоятельно). С одной стороны:

$$a = X_2^T A X_1 = [AX_1 = \lambda_1 X_1] = X_2^T \lambda_1 X_1 = \lambda_1 X_2^T X_1.$$

С другой стороны:

$$a = X_2^T A X_1 = [AX_2 = \lambda_2 X_2 \Rightarrow (AX_2)^T = (\lambda_2 X_2)^T \Rightarrow X_2^T A = \lambda_2 X_2^T] = \lambda_2 X_2^T X_1.$$

Вычитаем из одного результата другой:

$$0 = a - a = \lambda_1 X_2^T X_1 - \lambda_2 X_2^T X_1 = X_2^T X_1 (\lambda_1 - \lambda_2).$$

Так как $\lambda_1 \neq \lambda_2$, то $X_2^T X_1 = 0$, т.е. столбцы ортогональны ■

Все эти теоретические рассуждения позволяют сформулировать алгоритм приведения квадратичной формы ортогональным преобразованием к каноническому вид:

1) Записать матрицу A квадратичной формы.

2) Найти собственные числа A. С учётом кратности их должно быть ровно n (т.е. столько, каков порядок матрицы A).

3) Используя найденные собственные числа, записать канонический вид квадратичной формы: коэффициенты канонического вида равны собственным числам с учётом их кратности. Порядок следования собственных чисел выбирается произвольно (от этого будет зависеть порядок столбцов матрицы преобразования).

Например, если оказалось, что матрица A формы от 5 переменных имеет собственные числа -1, -1, 0, 2, 3, то каноническим видом будет

$$(-1)y_1^2 + (-1)y_2^2 + 0y_3^2 + 2y_4^2 + 3y_5^2 = -y_1^2 - y_2^2 + 2y_4^2 + 3y_5^2.$$

Можно было упорядочить собственные числа по-другому, например, 2, 3, -1, -1, 0, тогда канонический вид будет иной

$$2y_1^2 + 3y_2^2 - y_3^2 - y_4^2.$$

Оба этих ответа верные, но в каждом случае будет своя матрица преобразования.

Матрица преобразования состоит из нормированных ортогональных собственных столбцов, соответсвующих собственным числам. Подробный алгоритм их поиска будет рассмотрен на практике.

Одновременное приведение двух квадратичных форм к каноническому виду.

Очень часто нам приходится работать с несколькими объектами, находясь в одной системе. Так может быть и с квадратичными формами.

Пусть дана пара действительных квадратичных форм от n переменных

$$f(X) = X^T AX, \quad g(X) = X^T BX.$$

Зададимся вопросом, существует ли такое невырожденное линейное преобразование неизвестных $X = PY$, которое одновременно приводило бы обе эти формы к каноническому виду? В общем случае ответ будет отрицательным. Рассмотрим, например, пару форм

$$f(x_1, x_2) = x_1^2, \quad g(x_1, x_2) = x_1 x_2.$$

Пусть существует невырожденное линейное преобразование

$$x_1 = c_{11}y_1 + c_{12}y_2, \quad x_2 = c_{21}y_1 + c_{22}y_2 \tag{1}$$

приводящее обе эти формы к каноническому виду. Если это преобразование приводит форму f к каноническому виду, то

$$f(x_1, x_2) = x_1^2 = (c_{11}y_1 + c_{12}y_2)^2 = c_{11}^2 y_1^2 + 2c_{11}c_{12}y_1y_2 + c_{12}^2 y_2^2.$$

В каноническом виде нет смешанного произведения, поэтому один из коэффициентов c_{11} или c_{12} должен быть равен нулю. Меняя, если нужно, нумерацию неизвестных y_1, y_2 можно положить, что $c_{12} = 0$ и поэтому $c_{11} \neq 0$ (иначе преобразование будет вырожденным). Посмотрим, как преобразуется при этом преобразовании форма g:

$$g(x_1, x_2) = c_{11}y_1(c_{21}y_1 + c_{22}y_2) = c_{11}c_{21}y_1^2 + c_{11}c_{22}y_1y_2$$

Так как форма g также должна была перейти в канонический вид, то $c_{22} = 0$ ($c_{11} \neq 0$). Но тогда преобразование (1) – вырожденное, что неверно.

В случае, если одна из квадратичных форм является положительно определённой, то ситуация совсем иная. А именно, справедлива теорема:

Теорема. Если f и g – пара действительных квадратичных форм от n неизвестных, причём g – положительно определена, тогда существует невырожденное линейное преобразование, одновременно приводящее форму g к нормальному виду, а форму f – к каноническому.

Доказательство. Выполним сначала какое-нибудь преобразование

$$X = PY,$$

приводящее форму g к нормальному виду:

$$g = y_1^2 + y_2^2 + \ldots + y_n^2. \tag{2}$$

На форму f подействуем этим же преобразованием и получим некоторую форму

$$f = \varphi(y_1, \ldots, y_n).$$

Выполним теперь ортогональное преобразование

$$Y = QZ,$$

приводящее форму $\varphi(y_1, \ldots, y_n)$ к **каноническому виду**

$$f = \varphi(y_1, \ldots, y_n) = \lambda_1 z_1^2 + \ldots + \lambda_n z_n^2.$$

Так как матрица формы (2) единичная $\mathbb{E}$, то учитывая ортогональность матрицы Q, матрица формы в переменных z будет

$$Q^T \mathbb{E} Q = Q^T Q = \mathbb{E},$$

т.е. форма (2) останется в **нормальном виде**

$$g = y_1^2 + y_2^2 + \ldots + y_n^2 = z_1^2 + z_2^2 + \ldots + z_n^2.$$

Таким образом, преобразование

$$X = (PQ)Z$$

является искомым ■

Доказанная теорема является достаточным условием, но не является необходимым. Так например, формы

$$x_1^2 + x_2^2 - x_3^2, \quad x_1^2 - x_2^2 - x_3^2$$

уже обе имеют канонический вид, но среди них нет положительно определённых.

Линейные пространства

Множество L элементов $x, y, z, \ldots$ любой природы называют **линейным (векторным) пространством** над полем K, если выполнены три условия:

1. Задано сложение элементов L, т.е. закон, по которому любым элементам $x, y \in L$ ставится в соответствие элемент $z \in L$, называемый суммой элементов x и y и обозначаемый $z = x + y$;

2. Задано умножение элемента на число, т.е. закон, по которому любому элементу $x \in L$ и любому числу $\lambda \in K$ ставится в соответствие элемент $z \in L$, называемый произведением элемента x на λ и обозначаемый $z = \lambda x$;

3. Указанные законы (линейные операции) подчиняются следующим аксиомам линейного пространства:

   1) сложение коммутативно: $x + y = y + x$;

   2) сложение ассоциативно: $(x + y) + z = x + (y + z)$;

   3) существует такой элемент $\mathbb{O} \in L$, что $x + \mathbb{O} = x$ для любого $x \in L$;

   4) для каждого элемента $x \in L$ существует такой элемент $(-x) \in L$, что $x + (-x) = \mathbb{O}$;

   5) для любого $x \in L$: $1 \cdot x = x$ (1 – единица поля – нейтральный элемент относительно операции умножения, заданной в поле K);

   6) умножение на элементы поля K ассоциативно: $\lambda(\mu x) = (\lambda\mu)x$ $(\mu, \lambda \in K)$;

   7) выполнены свойства дистрибутивности: $(\lambda + \mu)x = \lambda x + \mu x$, $\lambda(x + y) = \lambda x + \lambda y$ $(\lambda, \mu \in K)$

Элементы линейного пространства принято называть **векторами**. Элемент $\mathbb{O}$, существование которого постулируется аксиомой 3), называют **нулевым вектором**, а элемент $(-x)$ – вектором, противоположным к вектору x.

Далее под полем K будем понимать $\mathbb{C}$ или $\mathbb{R}$. Там где какое либо утверждение будет справедливо только для одного из полей, это будет оговорено особо.

Пример 1. Примеры пространств (в каждом случае САМОСТОЯТЕЛЬНО проверить выполнение аксиом линейного пространства).

– множество V_3 (V_2) всех геометрических векторов в пространстве (на плоскости) с линейными операциями над векторами;

– множество матриц типа $m \times n$, элементами которых являются действительные числа, с линейными операциями над матрицами также удовлетворяет всем аксиомам линейного пространства;

– множество многочленов переменного x степени, не превышающей n, которые как функции можно складывать и умножать на действительные числа;

– множество функций, непрерывных на отрезке, с обычными операциями сложения функций и умножения функции на число. (При сложении непрерывных функций мы получаем непрерывную функцию, при умножении непрерывной функции на число также получаем непрерывную функцию. Поэтому сложение функций и умножение функции на число, не выводящее за пределы множества непрерывных на отрезке функций, можно рассматривать как операции линейного пространства. Легко убедиться, что для этих операций верны все аксиомы линейного пространства.)

Свойства линейного пространства

1. Любое линейное пространство имеет только один нулевой вектор

   Доказательство. Пусть есть два нулевых вектора $\mathbb{O}$ и $\mathbb{O}'$

$$\mathbb{O} = [\text{аксиома } 3)] = \mathbb{O} + \mathbb{O}' = [1)] = \mathbb{O}' + \mathbb{O} = [3)] = \mathbb{O}'.$$

2. Каждый вектор линейного пространства имеет только один противоположный вектор.

   Доказательство. Пусть есть два противоположных вектора $(-x)$ и $(-x)'$. Вычислим двумя способами выражение $(-x) + x + (-x)'$

$$(-x) + x + (-x)' = [2)] = ((-x) + x) + (-x)' = [1)] = (x + (-x)) + (-x)' = [4)] = \\ = \mathbb{O} + (-x)' = [1)] = (-x)' + \mathbb{O} = [3)] = (-x)'$$

$$(-x) + x + (-x)' = (-x) + (x + (-x)') = (-x) + \mathbb{O} = (-x).$$

   Значит, $(-x) = (-x)'$.

3. Если вектор $(-x)$ противоположен вектору x, то вектор x противоположен вектору $(-x)$.

   Доказательство. Если вектор $(-x)$ противоположен вектору x, то

$$\mathbb{O} = x + (-x) = [1)] = (-x) + x$$

   Из аксиомы 4) следует требуемое утверждение.

4. Для любых двух векторов a и b существует и притом единственный вектор x, такой что $a+x=b$.

   Доказательство. Покажем существование. Пусть $x=(-a)+b$. Тогда

$$a+((-a)+b)=(a+(-a))+b=\mathbb{O}+b=b.$$

   Значит, $a+x=b$. Единственность следует из рассуждений:

$$a+x=b \leftrightarrow (-a)+(a+x)=(-a)+b \leftrightarrow ((-a)+a)+x=(-a)+b \leftrightarrow$$

$$\mathbb{O}+x=(-a)+b \leftrightarrow x=(-a)+b.$$

   Полученный вектор совпал с нашим.

   Последнее свойство позволяет ввести еще одну операцию в линейном пространстве, которая является противоположной сложению. **Разностью** двух векторов $b-a$ называют вектор x, такой что $a+x=b$. Из доказательства свойства 4 вытекает, что

$$b-a=(-a)+b=b+(-a).$$

5. Произведение произвольного элемента линейного пространства на число 0 равно нулевому вектору: $0\cdot x=\mathbb{O}$.

   Доказательство.

$$0\cdot x=0\cdot x+\mathbb{O}=0\cdot x+x+(-x)=0\cdot x+1\cdot x+(-x)=$$
$$=(0+1)\cdot x+(-x)=x+(-x)=\mathbb{O}.$$

6. $(-x)=(-1)\cdot x$

   Доказательство. По свойству 2) противоположный – единственный. Проверим, что $(-1)\cdot x$ – противоположный к x:

$$x+(-1)\cdot x=1\cdot x+(-1)\cdot x=(1+(-1))x=0\cdot x=\mathbb{O}.$$

7. $\lambda\cdot\mathbb{O}=\mathbb{O}$.

   Доказательство. По свойству 5

$$\lambda\cdot\mathbb{O}=\lambda(0\cdot\mathbb{O})=(\lambda 0)\cdot\mathbb{O}=0\cdot\mathbb{O}=\mathbb{O}.$$

Линейная зависимость

(Повторим известные нам утверждения, но применительно к другим объектами – векторам линейного пространства.)

Пусть L – линейное пространство над полем K, $x_1, \ldots, x_n \in L$, $\alpha_1, \ldots, \alpha_n \in K$. Выражение

$$\alpha_1 x_1 + \ldots + \alpha_n x_n$$

называется **линейной комбинацией** векторов $x_1, \ldots, x_n \in L$, $\alpha_1, \ldots, \alpha_n$ – коэффициенты линейной комбинации. Линейная комбинации называется **тривиальной**, если $\alpha_1 = \ldots = \alpha_n = 0$, иначе эта комбинация называется **нетривиальной**.

Векторы $x_1, \ldots, x_n \in L$ называются **линейно зависимыми**, если существуют элементы поля $\alpha_1, \ldots, \alpha_n \in K$ не все равные нулю, такие что

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = \mathbb{O}.$$

В противном случае векторы называют **линейно независимыми**.

Пример. Многочлены $(x-2)^2$, $(x+2)^2$, x из множества $P_2(x)$ линейно зависимы. ($P_2(x)$ – множество многочленов, степень которых не выше двух. Легко доказать, что это линейное пространство относительно сложения многочленов и умножения на вещественные числа). Так как

$$1 \cdot (x-2)^2 - 1 \cdot (x+2)^2 + 8 \cdot x = \mathbb{O},$$

то существует ненулевой набор коэффициентов $(1, -1, 8)$, обращающий линейную комбинацию указанных векторов линейного пространства в ноль (в функцию тождественно равную нулю, а не в число ноль).

Пример. Система геометрических векторов $e_1 = (1, 0)$, $e_2 = (0, 1)$ пространства V_2 линейно независима, так как

$$\alpha_1 e_1 + \alpha_2 e_2 = \mathbb{O},$$

только если $\alpha_1 = \alpha_2 = 0$.

Свойства линейно зависимых векторов

Теорема 1. Для того чтобы система векторов $x_1, \ldots, x_n \in L$ была линейно зависима, необходимо и достаточно, чтобы один из векторов системы являлся линейной комбинацией остальных.

Доказательство. Пусть векторы линейно зависимы, т.е. существует набор не всех равных нулю $\alpha_1, \ldots, \alpha_n \in K$, таких, что линейная комбинация векторов обращается в ноль:

$$\alpha_1 x_1 + \alpha_2 x_2 \ldots + \alpha_n x_n = \mathbb{O}.$$

Пусть $\alpha_1 \neq 0$. Тогда выразим x_1

$$x_1 = -\frac{\alpha_2}{\alpha_1}x_2 - \ldots - \frac{\alpha_n}{\alpha_1}x_n.$$

Покажем теперь обратное. Если

$$x_1 = \alpha_2 x_2 + \ldots + \alpha_n x_n,$$

тогда перенося всё в одну сторону, получаем равенство, доказывающее линейную зависимость векторов

$$-1 \cdot x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = \mathbb{O} \quad \blacksquare$$

Пример. В линейном пространстве функций, непрерывных на $[a, b]$ система функций 1, $\sin^2 x$, $\cos 2x$ линейно зависима, так как

$$\cos 2x = 1 - 2\sin^2 x$$

Теорема 2. Если в системе векторов $x_1, \ldots, x_n \in L$ какие-то k векторов линейно зависимы $(0 < k < n)$, то и вся система векторов линено зависима.

Доказательство. Пусть первые k векторов линейно зависимы, тогда

$$\alpha_1 x_1 + \ldots + \alpha_k x_k = \mathbb{O},$$

где среди $\alpha_1, \ldots, \alpha_k$ не все равны нулю. Добавим в этом равенстве нулевые слагаемые

$$\alpha_1 x_1 + \ldots + \alpha_k x_k + 0 \cdot x_{k+1} + \ldots + 0 \cdot x_n = \mathbb{O}.$$

Исходя из определения линейной зависимости, делаем вывод, что все векторы $x_1, \ldots, x_n$ линейно зависимы $\blacksquare$

Следствие. Если система векторов линейно независима, то любая ее подсистема линейно независима.

(Доказательство от противного.)

Теорема 3. Если в системе векторов $x_1, \ldots, x_n \in L$ есть нулевой вектор, то вся система векторов линейно зависима.

Доказательство. Не ограничивая общности, считаем, что $x_1 = \mathbb{O}$. Тогда будет верным равенство

$$1 \cdot \mathbb{O} + 0 \cdot x_2 + \ldots + 0 \cdot x_n = \mathbb{O}.$$

Так как не все коэффициенты в этой линейной комбинации равны нулю, то все векторы $x_1, \ldots, x_n$ линейно зависимы $\blacksquare$

Базис линейного пространства.

Система векторов $b_1, \ldots, b_n \in L$ называется базисом пространства L, если выполнены условия

1) векторы $b_1, \ldots, b_n$ – линейно независимы;

2) для любого $x \in L$ существуют числа $x_1, \ldots x_n \in K$ (называемые **координатами**), такие что

$$x = x_1 b_1 + \ldots + x_n b_n.$$

Это равенство называют **разложением вектора** x **по базису** $b_1, \ldots, b_n$.

Теорема 1. Каждый элемент $x \in L$ может быть разложен по базису $b_1, \ldots, b_n \in L$ единственным образом.

Доказательство. Предположим, что для вектора x есть два разложения по одному и тому же базису.

$$x = x_1 b_1 + \ldots + x_n b_n, \quad x = x_1' b_1 + \ldots + x_n' b_n,$$

то вычитая эти равенства получаем

$$\mathbb{O} = (x_1 - x_1')b_1 + \ldots + (x_n - x_n')b_n.$$

Значит, в силу линейной независимости векторов базиса, коэффициенты в этом разложении равны нулю и $x_i = x_i'$ $(i = 1, \ldots, n)$. Следовательно, оба эти разложения по базису одинаковы ■

Пример. Рассмотрим множество $P_2(x)$ – пространство многочленов степени не выше 2. В качестве базиса возьмем 1, x, x^2. Эта совокупность независима, т.к. равенство

$$\alpha_1 \cdot 1 + \alpha_2 \cdot x + \alpha_3 \cdot x^2 = \mathbb{O}$$

есть равенство двух многочленов, а они равны, только если равны соответствующие коэффициенты, т.е $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Любой квадратичный многочлен может быть выражен через 1, x, x^2 (коэффициенты многочлена – коэффициенты разложения по базису):

$$ax^2 + bx + c = a \cdot x^2 + b \cdot x + c \cdot 1.$$

Теорема 2. При сложении любых двух векторов в линейном пространстве их координаты в одном и том же базисе складываются, а при умножении вектора на число его координаты умножаются на это число.

Доказательство следует непосредственно из разложений по базису. Рассмотрим два вектора, которые разложены по одному и тому же базису

$$x = x_1 b_1 + \ldots + x_n b_n, \quad y = y_1 b_1 + \ldots + y_n b_n.$$

Тогда (в преобразованиях используем аксиомы линейного пространства)

$$x + y = (x_1b_1 + \ldots + x_nb_n) + (y_1b_1 + \ldots + y_nb_n) = (x_1 + y_1)b_1 + \ldots + (x_n + y_n)b_n.$$

Значит, вектор $x+y$ имеет координаты $x_1+y_1, \ldots, x_n+y_n$. Аналогичные рассуждения проводим при умножении на число λ

$$\lambda x = \lambda(x_1b_1 + \ldots + x_nb_n) = (\lambda x_1)b_1 + \ldots + (\lambda x_n)b_n.$$

Значит, вектор λx имеет координаты $\lambda x_1, \ldots, \lambda x_n$ ■

Заметим, что по сути дела, если фиксировать порядок векторов в базисе, то линейную комбинацию, представляющую вектор, можно заменить упорядоченным набором коэффициентов и, тем самым, упростить запись. Более того, можно ввести матричный способ записи векторных соотношений. Базис удобно записывать как матрицу строку

$$B = (b_1, \ldots, b_n).$$

координаты разложения вектора по этому базису как матрицу-столбец

$$X = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$$

Тогда разложение x по базису можно записать

$$x = BX.$$

Размерность линейного пространства.

Максимальное количество линейно независимых векторов в данном линейном пространстве называют **размерностью** линейного пространства.

Если размерность линейного пространства L равна n, т.е. существует линейно независимая система из n векторов, а любая система векторов, содержащая $n+1$ вектор или более, линейно зависима, то говорят, что это линейное пространство n-мерно. Размерность такого линейного пространства обозначают $\dim L = n$.

Существуют линейные пространства, в которых можно выбрать линейно независимую систему, содержащую сколь угодно большое количество векторов. Такие линейные пространства называют **бесконечномерными**. В отличие от них, n-мерные линейные пространства называют **конечномерными**.

Пример. Линейное пространство $C[0,1]$ – функций, непрерывных на отрезке [0,1], является бесконечномерным, так как для любого натурального n система многочленов $1, x, x^2, \ldots, x^n$, являющихся элементами этого линейного пространства, линейно независима. Действительно, приравняем их линейную комбинацию к нулю

$$\alpha_0 \cdot 1 + \alpha_1 \cdot x + \ldots + \alpha_n \cdot x^n = \mathbb{O}.$$

Справа стоит нулевой многочлен, а многочлены равны тогда и только тогда, когда равны соответствующие коэффициенты, т.е. все коэффиценты α_0, … α_n равны нулю и многочлены $1, x, x^2, \ldots, x^n$ линейно независимы ■

Размерность линейного пространства тесно связана с количеством векторов, которое может иметь базис линейного пространства.

Теорема. Если линейное пространство L n-мерно, то любая линейно независимая система из n векторов является его базисом.

Доказательство. Пусть $b_1, \ldots, b_n$ – линейно независимые векторы. Так как L n-мерно, то для любого вектора $x \in L$ совокупность $b_1, \ldots, b_n, x$ – линейно зависима. По определению линейно зависимых векторов существуют $\alpha_1, \ldots, \alpha_{n+1}$ не все равные нулю, такие что

$$\alpha_1 b_1 + \ldots + \alpha_n b_n + \alpha_{n+1} x = \mathbb{O}.$$

Заметим, что $\alpha_{n+1} \neq 0$, иначе векторы $b_1, \ldots, b_n$ были бы линейно зависимы. Следовательно, x из этого равенства можно выразить:

$$x = -\frac{\alpha_1}{\alpha_{n+1}} b_1 - \ldots - \frac{\alpha_n}{\alpha_{n+1}} b_n.$$

Итак, $b_1, \ldots, b_n$ – линейно независимы и через них выражается любой вектор линейного пространства. Значит, по определению они являются базисом ■

Теорема. Если в линейном пространстве L существует базис из n векторов, то $\dim L = n$.

Доказательство. Пусть $b_1, \ldots, b_n$ – базис. Достаточно показать, что любая система из $n+1$ векторов $c_1, \ldots, c_{n+1}$ линейно зависима. Разложим эти векторы по базису

$$c_i = \alpha_{i1} b_1 + \ldots + \alpha_{in} b_n \quad (i = 1, \ldots, n+1)$$

Составим линейную комбинацию этих векторов и приравняем её к нулю

$$\lambda_1 c_1 + \ldots + \lambda_{n+1} c_{n+1} = \mathbb{O}.$$

Подставим в последнее равенство выражения для c_i

$$\lambda_1(\alpha_{11} b_1 + \ldots + \alpha_{1n} b_n) + \lambda_2(\alpha_{21} b_1 + \ldots + \alpha_{2n} b_n) + \ldots + \lambda_{n+1}(\alpha_{n+11} b_1 + \ldots + \alpha_{n+1n} b_n) = \mathbb{O}.$$

Перегруппируем по b_i

$$(\lambda_1\alpha_{11}+\lambda_2\alpha_{21}+\ldots+\lambda_{n+1}\alpha_{n+11})b_1+\ldots+(\lambda_1\alpha_{1n}+\lambda_2\alpha_{2n}+\ldots+\lambda_{n+1}\alpha_{n+1n})b_n=\mathbb{O}.$$

В силу линейной независимости, коэффициенты при b_i равны нулю. Это значит, что для поиска λ_i мы имеем линейную однородную систему n уравнений с n неизвестными

$$\begin{cases} \lambda_1\alpha_{11}+\lambda_2\alpha_{21}+\ldots+\lambda_{n+1}\alpha_{n+11}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ \lambda_1\alpha_{1n}+\lambda_2\alpha_{2n}+\ldots+\lambda_{n+1}\alpha_{n+1n}=0. \end{cases}$$

Так как число уравнений в системе (n) меньше, чем число неизвестных $(n+1)$, эта система имеет бесконечное множество решений, в частности ненулевое. Это означает, что векторы c_i линейно зависимы ■

Из последних двух теорем следует, что если в линейном пространстве есть базис длины n, то любой другой базис этого пространства имеет тоже n векторов. И в следующем разделе мы будем как раз рассматривать несколько базисов одного и того же линейного пространства.

Пример. Линейное пространство M_2 вещественных квадратных матриц второго порядка

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

рассматриваемое относительно операции сложения матриц и умножения их на вещественные числа, имеет базис

$$E_1=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ E_2=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ E_3=\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ E_4=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

(это доказать самостоятельно). Поэтому $\dim M_2=4$ ■