Московский Физико–Технический Институт

Кафедра высшей математики

Линейная алгебра
Весенний семестр

(Поток Чубарова Игоря Андреевича)

Над текстом работал:

Лазарь Влад

Version 1.6.5

Долгопрудный, 2026 г.

Вступление

Все пожелания, угрозы писать сюда.

Данный документ является не совсем "билетами" в традиционном их понимании для подготовки к экзаменам. Когда я взялся за вёрстку, мне хотелось, чтобы написанное тут было связкой с прекрасными семинарами Алексеева Василия Антоновича, где в крайне удачной форме изложены решения некоторых задач. На протяжении всего документа я постарался выдержать определённый стиль повествования — как можно меньше сухих выжимок теории из лекций. Также я хотел добавить больше физического смысла в некоторые темы. Что касается лекций моего года (в 2022 году их вёл Чубаров Игорь Андреевич) — некоторые доказательства я решил написать иначе. Советую читать и понимать это в течение семестра, иначе пользы будет примерно как и всегда, когда вы готовитесь к экзамену за пару дней.

Я постарался исключить все грубые ошибки. Однако, несмотря на бесчисленные перепроверки, переписывание документа и переосмысление того, как можно объяснить материал лучше, опечатки тут, конечно же, есть. Во всяком случае, по любой, даже самой незначительной, ошибке пишите мне. Здесь будет находиться последняя версия документа. *Перед прочтением не поленитесь скачать актуальную версию, так как изменения я вношу весьма активно.*

В качестве видеоматериалов хотел бы посоветовать лекции Чубарова Игоря Андреевича, Подлипской Ольги Геннадьевны (удалены с YouTube и LMS, загружены в ТГ с таймкодами), Ершова Андрея Владимировича, Кожевникова Павла Александровича, а также семинары Юденковой Марии Алексеевны.

Содержание

13. Диагональный и канонический вид квадратичных форм. Приведение квадратичных форм к каноническому виду с помощью элементарных преобразований. Приведение квадратичных форм к каноническому виду методом Лагранжа. Закон инерции квадратичных форм. Знакоопределённые квадратичные формы. Критерий Сильвестра

14. Аксиоматика евклидова пространства. Неравенство Коши–Буняковского. Неравенство треугольника. Матрица Грама и её свойства. Переход от одного ортонормированного базиса к другому. Ортогональные матрицы

15. Ортогональное дополнение подпространства. Ортогональное проектирование на подпространство. Ортогональные базисы. Процесс ортогонализации

16. Линейные преобразования евклидова пространства. Преобразование, сопряжённое данному. Матрица сопряжённого преобразования. Свойства сопряжённого преобразования

17. Самосопряжённые преобразования. Свойства их собственных векторов и собственных значений. Существование ортонормированного базиса из собственных векторов самосопряжённого преобразования

18. Свойства ортогональных преобразований. Свойства корней характеристического многочлена и собственных векторов ортогональных преобразований

19. Сингулярное разложение. Полярное разложение линейных преобразований евклидова пространства и матриц

20. Построение ортонормированного базиса, в котором квадратичная форма имеет диагональный вид. Одновременное приведение к диагональному виду пары квадратичных форм, одна из которых положительно определена

Билет № 1. Ранг матрицы. Теорема о ранге матрицы. Теорема о базисном миноре.

Определение [строчный ранг матрицы]. Пусть в матрице A существует линейно независимая (ЛНЗ) система из r строк, и нет линейно независимой системы из большего числа строк. Тогда мы будем говорить, что строчный ранг A равен r.

Аналогично определяется *столбцовый ранг матрицы.*

Определение [базисная подматрица]. В матрице A размера $m \times n$ подматрица порядка r называется базисной, если она невырождена, а все квадратные подматрицы большего порядка, если они существуют, вырождены. Столбцы и строки матрицы A, на пересечении которых стоит базисная подматрица, называются базисными столбцами и строками A.

Определение [ранг матрицы]. Рангом матрицы называется порядок базисной подматрицы или, иначе, самый большой порядок, для которого существуют невырожденные подматрицы. Ранг нулевой матрицы по определению считают нулём.

Отметим два очевидных свойства ранга:

- Ранг матрицы не меняется при транспонировании, так как при транспонировании матрицы все её подматрицы транспонируются, и при этом невырожденные подматрицы остаются невырожденными, а вырожденные — вырожденными.
- Если A' — подматрица матрицы A, то ранг A' не превосходит ранга A, так как любая невырожденная подматрица, входящая в A', входит и в A.

Определение [базисный минор]. Минор $M^{i_1,\ i_2,\ \dots,\ i_r}_{j_1,\ j_2,\ \dots,\ j_r}$ — детерминант базисной подматрицы. Верхние индексы $i_1,\ i_2,\ \dots,\ i_r$ — это номера строк, которые входят в этот минор. Аналогично нижние индексы $j_1,\ j_2,\ \dots,\ j_r$ — это номера столбцов, которые входят в этот минор.

Теорема о ранге матрицы. Ранг любой матрицы равен её строчному и столбцовому рангам.

Доказательство:

Пусть ранг матрицы A равен $r > 0$. Требуется доказать, что в матрице A имеется r столбцов (строк), образующих линейно *независимую* систему, и что всякие $r+1$ столбцов (строк) образуют линейно *зависимую* систему. Доказательство для строк и столбцов одно и то же, проведём его для столбцов.

Раз ранг матрицы равен r, то в ней имеется подматрица P с отличным от нуля детерминантом. Без ограничения общности можно предположить, что этот минор подматрицы P является угловым (ведь если он не угловой, то угловым мы его сделаем перестановкой строк и столбцов. Причём при элементарных преобразованиях ранг матрицы не изменится):

$$
\begin{pmatrix}
a_{1,\ 1} & \dots & a_{1,\ r} & a_{1,\ r+1} & \dots & a_{1,\ n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
a_{r,\ 1} & \dots & a_{r,\ r} & a_{r,\ r+1} & \dots & a_{r,\ n} \\
a_{r+1,\ 1} & \dots & a_{r+1,\ r} & a_{r+1,\ r+1} & \dots & a_{r+1,\ n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
a_{m,\ 1} & \dots & a_{m,\ r} & a_{m,\ r+1} & \dots & a_{m,\ n}
\end{pmatrix}, \text{ где } \det(P) = \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} \neq 0.
$$

Так как $\det(P) \neq 0$, то векторы $\boldsymbol{w}_1 = \{a_{11},\ \dots,\ a_{r1}\}$, $\dots$, $\boldsymbol{w}_r = \{a_{1r},\ \dots,\ a_{rr}\}$ линейно независимы, а тогда линейно независимы и векторы

$$
\boldsymbol{v}_1 = \{a_{11},\ \dots,\ a_{r1},\ \dots,\ a_{m1}\}, \quad \dots, \quad \boldsymbol{v}_r = \{a_{1r},\ \dots,\ a_{rr},\ \dots,\ a_{mr}\}.
$$

Итак, в матрице A ранга r имеется линейно независимая система, состоящая из r столбцов. Докажем тогда, что всякие $r+1$ столбцов матрицы A линейно зависимы. Предполагаем снова, что отличен от нуля левый верхний угловой минор порядка r матрицы A.

Вспомним, что среди векторов, являющихся линейными комбинациями данных r векторов, нельзя найти более r линейно независимых; поэтому достаточно доказать, что каждый столбец $\boldsymbol{v}_h = \{a_{1h},\ \dots,\ a_{rh},\ \dots,\ a_{mh}\}$, где $h = \overline{r+1,\ m}$, матрицы A является линейной комбинацией первых r столбцов.

Взяв любое $i \leqslant m$, построим определитель D_i, где последняя строка — это i-ая строка матрицы A.

$$
D_i = \begin{vmatrix} a_{11} & \dots & a_{1r} & a_{1h} \\ \vdots & \ddots & \vdots & \vdots \\ a_{r1} & \dots & a_{rr} & a_{rh} \\ a_{i1} & \dots & a_{ir} & a_{ih} \end{vmatrix}.
$$

При любом i этот детерминант равен 0. В самом деле, если $i \leqslant r$, то этот детерминант имеет две одинаковые строки (поскольку первые r строк уже включают i-ую строку). Если же $i \geqslant r$, то D_i есть детерминант некоторого минора $(r+1)$ порядка матрицы A, и он равен нулю, так как ранг матрицы A по предположению равен r. Итак, $D_i = 0$ при $i \leqslant m$.

Разложим детерминант D_i по элементам последней строки:

$$
A_1 = (-1)^{(r+1)+1} \begin{vmatrix} a_{12} & \dots & a_{1r} & \dots & a_{1h} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{r2} & \dots & a_{rr} & \dots & a_{rh} \end{vmatrix},
$$

и так далее ...

$$
A_k = (-1)^{(r+1)+k} \begin{vmatrix} a_{1,\ 1} & \dots & a_{1,\ k-1} & a_{1,\ k+1} & \dots & a_{1,\ r} & \dots & a_{1,\ h} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{r,\ 1} & \dots & a_{r,\ k-1} & a_{r,\ k+1} & \dots & a_{r,\ r} & \dots & a_{r,\ h} \end{vmatrix},
$$

и, наконец,

$$A_{r+1} = (-1)^{(r+1)+(r+1)} \begin{vmatrix} a_{11} & \dots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \dots & a_{rr} \end{vmatrix} = (-1)^{2(r+1)} \det(P) = \det(P) \neq 0.$$

Имеем: $0 = D_i = A_1 a_{i1} + A_2 a_{i2} + \ldots + A_r a_{ir} + A_{r+1} a_{ih}$, где $i = \overline{1,\, m}$.

Эти соотношения выражают равенство $A_1 \boldsymbol{v}_1 + A_2 \boldsymbol{v}_2 + \ldots + A_r \boldsymbol{v}_r + A_{r+1} \boldsymbol{v}_h = 0$, в котором заведомо коэффициент $A_{r+1} = \det(P)$ отличен от нуля, и которое поэтому можно разрешить относительно $\boldsymbol{v}_h$: $\boldsymbol{v}_h = \lambda_1 \boldsymbol{v}_1 + \ldots + \lambda_r \boldsymbol{v}_r$ при $\lambda_1 = -\frac{A_1}{A_{r+1}},\ \ldots,\ \lambda_r = -\frac{A_r}{A_{r+1}}$.

Итак, мы представили произвольный столбец $\boldsymbol{v}_h$ матрицы в виде линейной комбинации первых r столбцов этой матрицы, чем и закончим доказательство теоремы.

■

Теорема о базисном миноре. Если $M^{i_1,\ i_2,\ \ldots,\ i_r}_{j_1,\ j_2,\ \ldots,\ j_r}$ — базисный минор матрицы A, то строки $\boldsymbol{a}_{i_1},\ \ldots,\ \boldsymbol{a}_{i_r}$ и столбцы $\boldsymbol{a}^{\uparrow}_{j_1},\ \ldots,\ \boldsymbol{a}^{\uparrow}_{j_r}$ являются базисными в матрице A.

Доказательство:

Без ограничения общности будем считать, что матрица A имеет следующий вид (базисная матрица в левом верхнем углу):

$$\begin{pmatrix} a_{11} & \dots & a_{1r} & \dots & a_{1j} & \dots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\ a_{r1} & \dots & a_{rr} & \dots & a_{rj} & \dots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \\ a_{i1} & \dots & a_{ir} & \dots & a_{ij} & \dots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots \end{pmatrix}.$$

Докажем, что $\boldsymbol{a}_i = \lambda_1 \boldsymbol{a}_1 + \ldots + \lambda_r \boldsymbol{a}_r$.

Рассмотрим детерминант матрицы, состоящей из базисной матрицы и добавленными к ней i-ой строки и j-го столбца. Рассмотрим следующий минор, и разложим его по j-му столбцу (заметим, что мы буквально повторяем часть доказательства прошлой теоремы):

$$\begin{vmatrix} a_{11} & \dots & a_{1r} & a_{1j} \\ \vdots & \ddots & \vdots & \vdots \\ a_{r1} & \dots & a_{rr} & a_{rj} \\ a_{i1} & \dots & a_{ir} & a_{ij} \end{vmatrix} = a_{1j}(-1)^{1+r+1} \cdot M^j_1 + \ldots + a_{rj}(-1)^{2r+1} \cdot M^j_r + a_{ij}(-1)^{1+r+1+r} \cdot M^{1,\ \ldots,\ r}_{1,\ \ldots,\ r} = 0.$$

Значит, $a_{ij} = -\frac{1}{M}\left(\sum\limits_{k=1}^{r} a_{kj} A^j_k\right)$, где $A^j_i = (-1)^{i+j} M^j_i$ — алгебраическое дополнение элемента a_{ij}, и $M = M^{1,\ \ldots,\ r}_{1,\ \ldots,\ r}$ — базисный минор. Причём, A^j_k от j не зависит, то есть $a_{i*} = -\frac{1}{M}(\sum\limits_{k=1}^{r} a_{k*} A^*_k)$.

А ведь это как раз и означает, что i-ая строка — линейная комбинация $\boldsymbol{a}_1,\ \ldots, \boldsymbol{a}_r$.

■

Билет № 2. Системы линейных уравнений. Общее решение системы линейных уравнений. Фундаментальная система решений. Теорема Кронекера–Капелли. Теорема Фредгольма.

Определение [СЛУ]. Систему уравнений вида

$$\begin{cases} a_1^1 x_1 + a_2^1 x_2 + \ldots + a_n^1 x_n = b_1 \\ a_1^2 x_1 + a_2^2 x_2 + \ldots + a_n^2 x_n = b_2 \\ \qquad\qquad\vdots \\ a_1^m x_1 + a_2^m x_2 + \ldots + a_n^m x_n = b_m \end{cases}$$

будем называть системой линейных уравнений (СЛУ) с n неизвестными. Коэффициенты при $x_1, \ \ldots, \ x_n$ будем записывать в виде матрицы, называемой *матрицей системы*:

$$\begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \ldots & a_n^m \end{pmatrix} \equiv \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{n1} \\ a_{12} & a_{22} & \ldots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \ldots & a_{nm} \end{pmatrix}.$$

Числа, стоящие в правых частях уравнений системы, образуют столбец $\boldsymbol{b}$, называемый столбцом свободных членов. Если все элементы столбца свободных членов равны нулю, то система называется однородной (ОСЛУ). Причём однородная система всегда совместна (то есть имеет хотя бы одно решение), так как, очевидно, всегда есть как минимум тривиальное решение.

Определение [ФСР]. Фундаментальная система решений — базис в линейном пространстве решений однородной системы уравнений.

Определение [фундаментальная матрица]. Фундаментальная матрица — матрица, столбцы которой образуют ФСР.

Лемма о вычислении общего решения СЛУ. Пусть $A\boldsymbol{x} = \boldsymbol{b}$ и $A\boldsymbol{y} = \boldsymbol{0}$ — соответствующая ОСЛУ, тогда $\forall \boldsymbol{x}$ решение системы $A\boldsymbol{x} = \boldsymbol{b}$ представляется в виде $\boldsymbol{x} = \boldsymbol{x}_\text{ч} + \boldsymbol{y}$, где $\boldsymbol{x}_\text{ч}$ — частное решение неоднородной системы и $\boldsymbol{y}$ — произвольное решение ОСЛУ.

Доказательство [Необходимость $\Rightarrow$]:

Из того, что $\boldsymbol{x}_\text{ч}$ — частное решение $A\boldsymbol{x} = \boldsymbol{b}$, следует, что $A\boldsymbol{x}_\text{ч} = \boldsymbol{b}$. Тогда $A(\boldsymbol{x} - \boldsymbol{x}_\text{ч}) = \boldsymbol{0} \Rightarrow \boldsymbol{y} = \boldsymbol{x} - \boldsymbol{x}_\text{ч}$ — решение системы $A\boldsymbol{y} = \boldsymbol{0}$, или $\boldsymbol{x} = \boldsymbol{x}_\text{ч} + \boldsymbol{y}$.

Доказательство [Достаточность $\Leftarrow$]:

Пусть $\boldsymbol{y}$ — решение системы $A\boldsymbol{y} = \boldsymbol{0}$. Тогда $\boldsymbol{x} = \boldsymbol{x}_\text{ч} + \boldsymbol{y}$ — это есть решение $A\boldsymbol{x} = \boldsymbol{b}$, т. к. $A(\boldsymbol{x}_\text{ч} + \boldsymbol{y}) = A\boldsymbol{x}_\text{ч} + A\boldsymbol{y} = A\boldsymbol{x}_\text{ч} = \boldsymbol{b}$.

Итого: $\boxed{\boldsymbol{x}_\text{неодн} = \boldsymbol{x}_\text{ч} + \boldsymbol{y}_\text{одн}}$ ■

Общий вид ФСР.

Возьмём $\widetilde{A} = (A \mid \boldsymbol{b})$. Сведём матрицу $\widetilde{A}$ элементарными преобразованиями и перестановкой столбцов к следующему виду:

$$\left(\begin{array}{c|c|c} E_r & C & \boldsymbol{b'} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right), \text{ где } \boldsymbol{b'} \text{ — столбец } (b_1, \ \ldots, \ b_r)^T.$$

В матрице $\widetilde{A}$ после преобразований остаётся $\mathrm{rg}(\widetilde{A}) = \mathrm{rg}(A) = r$ ненулевых строк. Тогда запишем уравнения в следующем виде:

$$x_i + \sum_{j=r+1}^{n} c_{ij}x_j = b_i, \, i = \overline{1, \, r}.$$

Переменные $x_1, \ \ldots, \ x_r$ будем называть главными неизвестными, а остальные — свободными переменными.

$$\begin{cases} x_1 = b_1 - \sum\limits_{j=r+1}^{n} c_{1j}x_j \\ \qquad \vdots \\ x_r = b_r - \sum\limits_{j=r+1}^{n} c_{rj}x_j \end{cases}$$

Тогда распишем $\boldsymbol{x}$:

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 - \sum\limits_{j=r+1}^{n} c_{1j}x_j \\ \vdots \\ b_r - \sum\limits_{j=r+1}^{n} c_{rj}x_j \\ x_{r+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \sum_{j=r+1}^{n} x_j \begin{pmatrix} -c_{1j} \\ \vdots \\ -c_{rj} \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

где 1 стоит на $(j - r)$-ой позиции нижнего блока.

Матрица $\Phi = \begin{pmatrix} -\sum\limits_{j=r+1}^{n} c_{1j}x_j \\ \vdots \\ -\sum\limits_{j=r+1}^{n} c_{rj}x_j \\ E_{n-r} \end{pmatrix}$ — это матрица, столбцы которой образуют фундаментальную систему решений (ФСР). Будем обозначать Φ следующим образом:

$$\Phi = \left(\begin{array}{c|c|c} -c_{1,\ r+1} & \ldots & -c_{1n} \\ \vdots & \ddots & \vdots \\ -c_{r,\ r+1} & \ldots & -c_{rn} \\ \hline 1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 1 \end{array}\right) = \left(\begin{array}{c} -C \\ \hline E_{n-r} \end{array}\right).$$

А теперь докажем, что построенная система векторов Φ действительно является базисом пространства решений однородной системы $A\boldsymbol{x} = \mathbf{0}$, то есть:

1. Каждый столбец Φ является решением однородной системы.
2. Столбцы Φ линейно независимы.
3. Любое решение однородной системы есть линейная комбинации столбцов Φ.

Для проверки первого свойства воспользуемся блочным представлением матрицы коэффициентов A после элементарных преобразований (без столбца свободных членов). Её можно записать как $(E_r \mid C)$.

Умножим эту матрицу на произвольный столбец φ_j из ФСР. Это умножение удобно провести в блочной форме:

$$(E_r \mid C) \cdot \varphi_j = (E_r \mid C) \cdot \left(\begin{array}{c} -C_j \\ \hline e_j \end{array}\right) = E_r \cdot (-C_j) + C \cdot e_j = -C_j + C_j = \mathbf{0}.$$

А ведь это и значит, что каждый столбец Φ является решением однородной системы.

Линейная независимость столбцов матрицы Φ очевидно следует из её структуры: её нижний блок представляет собой единичную матрицу E_{n-r}. Ранги блочной матрицы и её нижнего блока связаны соотношением: $\mathrm{rg}(\Phi) \geqslant \mathrm{rg}(E_{n-r}) = n - r$. Так как в Φ ровно $n - r$ столбцов, её ранг равен $n - r$, значит, её столбцы линейно независимы.

Третье свойство следует из самого способа построения общего решения однородной системы. Любое решение $\boldsymbol{x}_0$ системы $A\boldsymbol{x} = \mathbf{0}$ получается из общего решения неоднородной системы при $\boldsymbol{b} = \mathbf{0}$ как было показано ранее, то есть представляется в виде линейной комбинации столбцов ФСР с какими-то коэффициентами.

Итак, столбцы матрицы Φ образуют базис (фундаментальную систему решений) пространства решений однородной системы уравнений $A\boldsymbol{x} = \mathbf{0}$. Размерность этого пространства равна $n - r$.

Теорема Кронекера–Капелли. СЛУ $A\boldsymbol{x} = \boldsymbol{b}$ совместна $\Leftrightarrow \operatorname{rg}(A) = \operatorname{rg}(A \mid \boldsymbol{b})$.

Доказательство [Необходимость $\Rightarrow$]:

Пусть существует решение $\boldsymbol{x}$, тогда $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{b}$. То есть столбец $\boldsymbol{b}$ есть линейная комбинация $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, а значит, дописывание $\boldsymbol{b}$ к матрице A не увеличит её ранг.

Доказательство [Достаточность $\Leftarrow$]:

Имеем $\operatorname{rg}(A) = \operatorname{rg}(A \mid \boldsymbol{b}) \Rightarrow$ у них одна и та же базисная подматрица. Значит, $\boldsymbol{b}$ может быть выражен как линейная комбинация столбцов базисной подматрицы. Итак, система $A\boldsymbol{x} = \boldsymbol{b}$ разрешима.

■

Теорема Фредгольма. СЛУ $A\boldsymbol{x} = \boldsymbol{b}$ совместна $\Leftrightarrow \forall \boldsymbol{y}$, где $\boldsymbol{y}$ — решение системы $A^T\boldsymbol{y} = \boldsymbol{0}$, выполняется $\boldsymbol{b}^T\boldsymbol{y} = 0$.

Доказательство [Необходимость $\Rightarrow$]:

Пусть $\exists \boldsymbol{x}_0 \colon A\boldsymbol{x}_0 = \boldsymbol{b}$ и $\boldsymbol{y}$ — некоторое решение системы $A^T\boldsymbol{y} = \boldsymbol{0} \Leftrightarrow \boldsymbol{y}^T A = \boldsymbol{0}$.
Домножим $\big[A\boldsymbol{x}_0 = \boldsymbol{b}\big]$ на $\boldsymbol{y}^T$ слева $\Rightarrow \big(\boldsymbol{y}^T A\big)\boldsymbol{x}_0 = \boldsymbol{y}^T\boldsymbol{b} = 0 = \boldsymbol{b}^T\boldsymbol{y}$.

Доказательство [Достаточность $\Leftarrow$]:

Предположим теперь, что система $A\boldsymbol{x} = \boldsymbol{b}$ несовместна. Это равносильно тому, что в упрощённом виде её расширенной матрицы $\big(A \mid \boldsymbol{b}\big)$ есть строка $(0 \ \ldots \ 0 \mid 1)$. Так как упрощённый вид получается из исходной матрицы элементарными преобразованиями строк, строка $(0 \ \ldots \ 0 \mid 1)$ является линейной комбинацией строк матрицы $\big(A \mid \boldsymbol{b}\big)$.

Обозначим коэффициенты этой линейной комбинации через $\boldsymbol{y}_0^T$. Тогда:

$$\boldsymbol{y}_0^T\big(A \mid \boldsymbol{b}\big) = (0 \ \ldots \ 0 \mid 1)$$

Это равносильно тому, что $\boldsymbol{y}_0^T A = \boldsymbol{0}$ и $\boldsymbol{y}_0^T\boldsymbol{b} = 1$. Из первого условия $\boldsymbol{y}_0^T A = \boldsymbol{0}$ следует, что $A^T\boldsymbol{y}_0 = \boldsymbol{0}$, то есть $\boldsymbol{y}_0$ — решение сопряжённой однородной системы $A^T\boldsymbol{y} = \boldsymbol{0}$.

Но из второго условия $\boldsymbol{y}_0^T\boldsymbol{b} = 1$ следует, что $\boldsymbol{b}^T\boldsymbol{y}_0 = 1 \neq 0$. По условию теоремы, для любого решения $\boldsymbol{y}$ системы $A^T\boldsymbol{y} = \boldsymbol{0}$ должно выполняться $\boldsymbol{b}^T\boldsymbol{y} = 0$, но мы нашли $\boldsymbol{y}_0$, для которого $A^T\boldsymbol{y}_0 = \boldsymbol{0}$, но $\boldsymbol{b}^T\boldsymbol{y}_0 = 1 \neq 0$ — это противоречие нашему изначальному предположению о несовместности системы $A\boldsymbol{x} = \boldsymbol{b}$.

■

Билет № 3. Аксиоматика линейного пространства. Линейная зависимость и линейная независимость систем элементов в линейном пространстве. Размерность и базис.

Немного о группах, кольцах и полях.

Группа, кольцо и поле — это алгебраические структуры, которые обобщают свойства операций над числами.

1. Группа — это множество G с бинарной операцией $\circ$ (например, сложение или умножение), удовлетворяющее следующим 4 аксиомам:

1. $\forall\, \boldsymbol{a},\, \boldsymbol{b} \in G\colon\; \boldsymbol{a} \circ \boldsymbol{b} \in G$ — замкнутость.
2. $(\boldsymbol{a} \circ \boldsymbol{b}) \circ \boldsymbol{c} = \boldsymbol{a} \circ (\boldsymbol{b} \circ \boldsymbol{c})\;\; (\forall\, \boldsymbol{a},\, \boldsymbol{b},\, \boldsymbol{c} \in G)$ — ассоциативность.
3. $\exists\, \boldsymbol{e} \in G\colon \forall \boldsymbol{a} \in G\colon \boldsymbol{a} \circ \boldsymbol{e} = \boldsymbol{e} \circ \boldsymbol{a} = \boldsymbol{a}$ — существование нейтрального элемента.
4. $\forall\, \boldsymbol{a} \in G : \exists\, \boldsymbol{a}^{-1} \in G\colon \boldsymbol{a} \circ \boldsymbol{a}^{-1} = \boldsymbol{e}$ — существование обратного элемента.

Причём операция $\circ$ не обязана быть коммутативной. Группа называется *абелевой*, если:

$$\forall\, \boldsymbol{a},\, \boldsymbol{b} \in G\colon \boldsymbol{a} \circ \boldsymbol{b} = \boldsymbol{b} \circ \boldsymbol{a}.$$

Приведём примеры групп:

1. Целые числа $\mathbb{Z}$ с операцией сложения "+". Обозначают как $\big(\mathbb{Z},\; +\big)$. Здесь нейтральным элементом является 0. А обратным, например, к 42 является -42.
2. Ненулевые рациональные числа $\mathbb{Q}\backslash\{0\}$ с умножением "$\cdot$". Обозначают как $\big(\mathbb{Q}^*,\; \cdot\big)$. Здесь нейтральным элементом является 1. А обратным к 42 является $\frac{1}{42}$.

2. Кольцо — это множество R, на котором заданы две операции (условно называемые сложением и умножением), удовлетворяющие следующим условиям:

1. Множество $\big(R,\; +\big)$ является абелевой группой.
2. $(\boldsymbol{a} \cdot \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot (\boldsymbol{b} \cdot \boldsymbol{c})\;\; (\forall\, \boldsymbol{a},\, \boldsymbol{b},\, \boldsymbol{c} \in R)$ — ассоциативность относительно *умножения.*
3. $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}\;\; (\forall\, \boldsymbol{a},\, \boldsymbol{b},\, \boldsymbol{c} \in R)$ — дистрибутивность–1.
4. $(\boldsymbol{a} + \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot \boldsymbol{c} + \boldsymbol{b} \cdot \boldsymbol{c}\;\; (\forall\, \boldsymbol{a},\, \boldsymbol{b},\, \boldsymbol{c} \in R)$ — дистрибутивность–2.

Причём в кольце не требуется существования обратного элемента по умножению (т. е. деления может не быть) и коммутативности умножения.

Приведём примеры колец:

1. Целые числа $\mathbb{Z}$ с обычными "+" и "$\cdot$".
2. Матрицы $n \times n$ с операциями сложения и умножения.

3. Поле — коммутативное кольцо с единицей, где каждый ненулевой элемент обратим относительно умножения:

1. Множество $\big(F,\ +\big)$ является абелевой группой.

2. Множество $\big(F\backslash\{0\},\ \cdot\big)$ является абелевой группой.

3. Дистрибутивность умножения относительно сложения.

Приведём примеры полей:

1. Рациональные числа $\mathbb{Q}$, вещественные $\mathbb{R}$, комплексные $\mathbb{C}$.

2. Конечные поля (поля Галуа), например, арифметика по простому модулю.

Это всё было для того, чтобы дать определение линейному пространству, которое строится над *полем.*

Полагаем, что L — произвольное множество, $\mathbb{F}$ — некоторое поле (например, поле $\mathbb{R}$). На L определена бинарная операция, которую мы будем называть сложением и обозначать "$+$", а также $\forall\lambda \in \mathbb{F}$ определена операция, которую мы будем называть умножением вектора на скаляр и обозначать $\lambda \cdot \boldsymbol{a}$ или $\lambda\boldsymbol{a}$, где $\boldsymbol{a} \in L$.

Определение [векторное (линейное) пространство]. L называется векторным (или линейным) пространством над полем $\mathbb{F}$, если выполнены следующие **8** свойств (аксиом векторного пространства):

1. $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$ $\ (\forall\,\boldsymbol{a},\,\boldsymbol{b},\,\boldsymbol{c}\in L)$ — ассоциативность сложения.

2. $\exists\,\mathbf{0}_L \in L\colon \forall\,\boldsymbol{a}\in L\colon \boldsymbol{a}+\mathbf{0}_L=\mathbf{0}_L+\boldsymbol{a}=\boldsymbol{a}$ — существование нулевого вектора.

3. $\forall\,\boldsymbol{a}\in L\colon \exists\,\boldsymbol{x}\in L\colon \boldsymbol{x}+\boldsymbol{a}=\boldsymbol{a}+\boldsymbol{x}=\mathbf{0}_L$ — существование противоположного вектора.

4. $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}\quad(\forall\,\boldsymbol{a},\,\boldsymbol{b}\in L)$ — коммутативность сложения.

5. $\lambda\cdot(\boldsymbol{a}+\boldsymbol{b})=\lambda\cdot\boldsymbol{a}+\lambda\cdot\boldsymbol{b}\ \ (\forall\,\boldsymbol{a},\,\boldsymbol{b}\in L;\ \forall\,\lambda\in\mathbb{F})$ — дистрибутивность–1.

6. $(\lambda+\mu)\cdot\boldsymbol{a}=\lambda\cdot\boldsymbol{a}+\mu\cdot\boldsymbol{a}\ \ (\forall\,\boldsymbol{a}\in L;\ \forall\ \lambda,\,\mu\in\mathbb{F})$ — дистрибутивность–2.

7. $1\cdot\boldsymbol{a}=\boldsymbol{a}\quad(\forall\,\boldsymbol{a}\in L;\ 1\in\mathbb{F})$ — нейтральность единичного скаляра.

8. $(\lambda\mu)\cdot\boldsymbol{a}=\lambda\cdot(\mu\cdot\boldsymbol{a})\ \ (\forall\,\boldsymbol{a}\in L;\ \forall\,\lambda,\,\mu\in\mathbb{F})$ — ассоциативность умножения на скаляры.

Определение [линейная комбинация]. Пусть $\boldsymbol{a}_1,\ \ldots,\boldsymbol{a}_k\in L;\ \lambda_1,\ \ldots,\lambda_k\in\mathbb{F}$.
Сумма $\sum\limits_{i=1}^{k}\lambda_i\boldsymbol{a}_i$ называется линейной комбинацией векторов $\boldsymbol{a}_1,\ \ldots,\boldsymbol{a}_k$ с коэффициентами $\lambda_1,\ \ldots,\lambda_k$.
Причём линейная комбинация $\sum\limits_{i=1}^{k}\lambda_i\boldsymbol{a}_i$ называется *тривиальной*, если все её коэффициенты равны 0 (очевидно, что тогда сама линейная комбинация будет равна $\mathbf{0}$).

Определение [линейная независимость]. Система векторов в L называется линейно независимой, если нулевой вектор раскладывается единственным образом по этой системе векторов как их тривиальная линейная комбинация.

Определение [линейная зависимость]. Если существует нетривиальная линейная комбинация векторов, равная нулевому вектору, то такая система векторов называется линейно зависимой.

Некоторые свойства векторов в системе.

1. Система из $k > 1$ векторов ЛЗ $\Leftrightarrow$ хотя бы один вектор этой системы — линейная комбинация остальных.

2. Если в системе есть $\mathbf{0}$, то она является ЛЗ.

3. Если подсистема некоторой системы векторов ЛЗ, то и вся система ЛЗ.

4. Всякая подсистема ЛНЗ системы является ЛНЗ системой.

5. Если система векторов является базисом пространства L, то разложение любого вектора из L по этому базису единственно.

Определение [ранг системы векторов]. Целое неотрицательное число r называется рангом непустой системы $\mathbb{A}$ векторов из L, если в системе $\mathbb{A}$ найдётся линейно независимая подсистема из r векторов, а любая подсистема из $r+1$ векторов является линейно зависимой.

Определение [бесконечный ранг]. Будем говорить, что система $\mathbb{A}$ имеет бесконечный ранг, если $\forall\, r \in \mathbb{N}$ в $\mathbb{A}$ найдётся линейно независимая подсистема из r векторов.

Примеры таких систем (вероятно, эти примеры не случайны и встретятся вам в дальнейшем...):

1. $\mathbb{A} = \{1,\, x,\, x^2,\, x^3,\, \dots\}$;

2. $\mathbb{A} = \{1,\ \sin(x),\ \cos(x),\ \sin(2x),\ \cos(2x),\ \dots\}$;

3. $\mathbb{A} = \{e^x,\, e^{2x},\, e^{3x},\, \dots\}$.

Если V — подпространство L, порождённое системой $\mathbb{A}$, то ранг системы $\mathbb{A}$ называется размерностью подпространства V и обозначается $\dim(V)$.

Пространство размерности k называют *k-мерным.* Если $\dim(V) < \infty$, то V называют конечномерным, иначе — бесконечномерным.

Определение [базис]. Базисом в L называется максимальная линейно независимая система векторов, по которой любой вектор из L раскладывается единственным образом. Если такая система конечна, пространство L называется конечномерным.

Билет № 4. Разложение по базису в линейном пространстве. Координатное представление элементов линейного пространства и операций с ними. Матрица перехода. Изменение координат при изменении базиса в линейном пространстве. Теорема об изоморфизме.

Пусть в линейном пространстве L выбраны базисы $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$. Тогда для любого $\boldsymbol{x} \in L$ можем записать его координатные представления:

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i = \boldsymbol{e}\boldsymbol{x}_{\boldsymbol{e}} \quad \text{и} \quad \boldsymbol{x}' = \sum_{i=1}^{n} x'_i \boldsymbol{e}'_i = \boldsymbol{e}'\boldsymbol{x}_{\boldsymbol{e}'}.$$

Операции сложения векторов и умножения на скаляр.

1. Координатный столбец суммы векторов равен сумме координатных столбцов:

$$\boldsymbol{x} + \boldsymbol{y} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i + \sum_{j=1}^{n} y_j \boldsymbol{e}_j = \sum_{i=1}^{n} (x_i + y_i)\boldsymbol{e}_i = \boldsymbol{e}(\boldsymbol{x} + \boldsymbol{y}).$$

2. Координатный столбец произведения вектора на скаляр равен произведению координатного столбца этого вектора на скаляр:

$$\alpha\boldsymbol{x} = \alpha \sum_{i=1}^{n} x_i \boldsymbol{e}_i = \sum_{i=1}^{n} (\alpha x_i)\boldsymbol{e}_i = \boldsymbol{e}(\alpha\boldsymbol{x}).$$

Матрица перехода.

Рассмотрим на примере трёхмерного пространства. Пусть есть два базиса в пространстве: "старый" $\boldsymbol{e} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ и "новый" $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$.

Координаты некоторого вектора $\boldsymbol{\vartheta}$ запишем в виде столбцов: $\boldsymbol{\vartheta}_{\boldsymbol{e}} = (x_1, x_2, x_3)^T$ в базисе $\boldsymbol{e}$ и $\boldsymbol{\vartheta}_{\boldsymbol{e}'} = (x'_1, x'_2, x'_3)^T$ в базисе $\boldsymbol{e}'$.

Пусть нам известно представление вектора $\boldsymbol{\vartheta}$ в базисе $\boldsymbol{e}'$:

$$\boldsymbol{\vartheta} = x'_1\boldsymbol{e}'_1 + x'_2\boldsymbol{e}'_2 + x'_3\boldsymbol{e}'_3 = \boldsymbol{e}'\boldsymbol{\vartheta}_{\boldsymbol{e}'}.$$

Также нам известно представление базиса $\boldsymbol{e}'$ в базисе $\boldsymbol{e}$:

$$\begin{cases} \boldsymbol{e}'_1 = a_{11}\cdot\boldsymbol{e}_1 + a_{12}\cdot\boldsymbol{e}_2 + a_{13}\cdot\boldsymbol{e}_3 \\ \boldsymbol{e}'_2 = a_{21}\cdot\boldsymbol{e}_1 + a_{22}\cdot\boldsymbol{e}_2 + a_{23}\cdot\boldsymbol{e}_3 \\ \boldsymbol{e}'_3 = a_{31}\cdot\boldsymbol{e}_1 + a_{32}\cdot\boldsymbol{e}_2 + a_{33}\cdot\boldsymbol{e}_3 \end{cases} \xRightarrow{\text{в матрично-векторном виде}} \quad \boldsymbol{e}' = \boldsymbol{e}\begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = \boldsymbol{e}S_{\boldsymbol{e}\to\boldsymbol{e}'},$$

где $S_{\boldsymbol{e}\to\boldsymbol{e}'}$ — матрица перехода от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$. Нас интересует, как в двух базисах $\boldsymbol{e}$ и $\boldsymbol{e}'$ связаны между собой представления вектора $\boldsymbol{\vartheta}$.

Разложим вектор $\boldsymbol{\vartheta}$ по базису $\boldsymbol{e}$:

$$\boldsymbol{\vartheta} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3 = \boldsymbol{e}\boldsymbol{\vartheta}_{\boldsymbol{e}}.$$

Получается, что один и тот же вектор можно представить по-разному:

$$\boldsymbol{\vartheta} = \boldsymbol{e}\boldsymbol{\vartheta}_{\boldsymbol{e}} = \boldsymbol{e}'\boldsymbol{\vartheta}_{\boldsymbol{e}'}.$$

Тогда, выразив $\boldsymbol{e}'$ через $\boldsymbol{e}$, получим:

$$\boldsymbol{e}\boldsymbol{\vartheta}_{\boldsymbol{e}} = \boldsymbol{e}'\boldsymbol{\vartheta}_{\boldsymbol{e}'} = \big(\boldsymbol{e}S_{\boldsymbol{e}\to\boldsymbol{e}'}\big)\boldsymbol{\vartheta}_{\boldsymbol{e}'} = \boldsymbol{e}\big(S_{\boldsymbol{e}\to\boldsymbol{e}'}\boldsymbol{\vartheta}_{\boldsymbol{e}'}\big).$$

Итак, в двух базисах компоненты векторов связаны следующим образом:

$$\boxed{\begin{cases} \boldsymbol{e}' = \boldsymbol{e}S_{\boldsymbol{e}\to\boldsymbol{e}'} \\ \boldsymbol{\vartheta}_{\boldsymbol{e}} = S_{\boldsymbol{e}\to\boldsymbol{e}'}\boldsymbol{\vartheta}_{\boldsymbol{e}'} \end{cases}}.$$

Отметим ещё одно свойство:

С одной стороны справедливо, что $\boldsymbol{e}' = \boldsymbol{e}S_{\boldsymbol{e}\to\boldsymbol{e}'}$, отсюда $\boldsymbol{e} = \boldsymbol{e}'\big(S_{\boldsymbol{e}\to\boldsymbol{e}'}\big)^{-1}$ (отметим, что обратная матрица обязательно существует, так как строки $S_{\boldsymbol{e}\to\boldsymbol{e}'}$ являются ЛНЗ, а значит детерминант $S_{\boldsymbol{e}\to\boldsymbol{e}'}$ не равен нулю). А с другой стороны $\boldsymbol{e} = \boldsymbol{e}'S_{\boldsymbol{e}'\to\boldsymbol{e}}$. Тогда в силу единственности разложения $\boldsymbol{e}$ по $\boldsymbol{e}'$ получаем:

$$\boxed{S_{\boldsymbol{e}'\to\boldsymbol{e}} = \big(S_{\boldsymbol{e}\to\boldsymbol{e}'}\big)^{-1}}.$$

Таким образом, матрица перехода $S_{\boldsymbol{e}\to\boldsymbol{e}'}$ всегда обратима.

Капля в море изоморфизма.

Анекдот [про изоморфизм].

70-е годы. Группа студентов матмеха Ленинградского универа стоит в очереди в кафе "Белочка" (возле метро "Василеостровская", не знаю, есть ли сейчас это кафе) и спорят между собой о некой математической функции. В воздухе мелькают слова: "Это изоморфизм!", "Да, нет же, это гомоморфизм!" и проч. Дама бальзаковского возраста делает замечание: "Молодые люди, здесь порядочные женщины стоят, а вы выражаетесь!".

Говорят, что между элементами двух множеств U и V установлено взаимно однозначное соответствие, если указано правило, которое каждому элементу $\boldsymbol{u} \in U$ сопоставляет единственный элемент $\boldsymbol{v} \in V$, причём каждый элемент $\boldsymbol{v} \in V$ оказывается сопоставленным единственному элементу $\boldsymbol{u} \in U$. Взаимно однозначное соответствие будем обозначать $U \leftrightarrow V$.

Два линейных пространства U и V называются изоморфными, если между их элементами можно установить такое взаимно однозначное соответствие, что выполняются условия:

1. Сумме векторов пространства U соответствует сумма соответствующих векторов пространства V.

2. Произведению числа на вектор пространства U соответствует произведение того же числа на соответствующий вектор пространства V

Другими словами, изоморфизм — это взаимно однозначное соответствие, сохраняющее линейные операции.

Некоторые свойства изоморфизма.

1. При изоморфизме линейных пространств U и V их нулевые элементы соответствуют друг другу ($\mathbf{0}_U \leftrightarrow \mathbf{0}_V$). Образом противоположного вектора $(-\boldsymbol{u})$ является вектор, противоположный образу вектора $\boldsymbol{u}$ (то есть $-\varphi(\boldsymbol{u})$).

2. Линейной комбинации векторов пространства U соответствует линейная комбинация соответствующих векторов пространства V с теми же коэффициентами. То есть если $\varphi\colon U \to V$ — изоморфизм, то для любых векторов $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in U$ и любых скаляров $\alpha_1, \ldots, \alpha_k$ справедливо:

$$\varphi(\alpha_1\boldsymbol{u}_1 + \ldots + \alpha_k\boldsymbol{u}_k) = \alpha_1\varphi(\boldsymbol{u}_1) + \ldots + \alpha_k\varphi(\boldsymbol{u}_k).$$

3. Линейно независимой (линейно зависимой) системе векторов пространства U соответствует линейно независимая (линейно зависимая) система векторов пространства V.

4. Если пространство U изоморфно пространству V, а V изоморфно пространству W, то пространства U и W также изоморфны.

5. Любое n-мерное вещественное пространство V изоморфно n-мерному арифметическому пространству $\mathbb{R}^n$, а n-мерное линейное комплексное пространство изоморфно $\mathbb{C}^n$.

Теорема об изоморфизме. Два конечномерных линейных пространства (над одним и тем же числовым полем) изоморфны тогда и только тогда, когда они имеют одну и ту же размерность.

Доказательство [Необходимость $\Rightarrow$]:

Если пространства изоморфны ($U \leftrightarrow V$), то базису $(\boldsymbol{u}_1, \boldsymbol{u}_2, \ \ldots, \boldsymbol{u}_n)$ пространства U соответствует линейно независимая система векторов $(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$ пространства V, которую в случае необходимости можно дополнить до базиса пространства V. Следовательно, $\dim(U) \leqslant \dim(V)$. Аналогично получаем противоположное неравенство $\dim(V) \leqslant \dim(U)$.

Таким образом, $\dim(U) = \dim(V)$.

Доказательство [Достаточность $\Leftarrow$]:

Пусть пространства U и V определены над полем $\mathbb{R}$ и $\dim(U) = \dim(V) = n$. Тогда, выбрав любые базисы в пространствах U и V, установим изоморфизмы $U \leftrightarrow \mathbb{R}^n$ и $V \leftrightarrow \mathbb{R}^n$, если U и V — вещественные пространства. А если пространства U и V определены над полем $\mathbb{C}$ комплексных чисел, то $U \leftrightarrow \mathbb{C}^n$ и $V \leftrightarrow \mathbb{C}^n$. В итоге пространства U и V изоморфны.

■

Билет № 5. Подпространства в линейном пространстве. Способы задания подпространств. Сумма и пересечение подпространств. Формула размерности суммы двух подпространств. Прямая сумма.

Определение [подпространство в L]. Непустое подмножество L' векторов линейного прострaвнства L называется линейным подпространством, если:

1. Сумма любых двух векторов из L' принадлежит L';
2. Произведение каждого вектора из L' на любое число также принадлежит L'.

В силу этого определения любая линейная комбинация векторов из L' лежит в L'. Также нулевой вектор должен принадлежать L' как произведение $0 \cdot \boldsymbol{x}$, где $\boldsymbol{x} \in L'$.

Например, множество многочленов степени не выше 3 является подпространством пространства всех многочленов.

Справедливость аксиом линейного пространства для L' прямо вытекает из их справедливости для L. Таким образом, подпространство L' является линейным пространством.

Задание подпространства с помощью линейной оболочки.

Определение [линейная оболочка]. Рассмотрим некоторое множество векторов P линейного пространства L. Множество U, состоящее из всевозможных линейных комбинаций векторов из P, мы будем называть линейной оболочкой множества P.

То есть если $\boldsymbol{a}_1, \ \ldots, \boldsymbol{a}_m \in L$, то линейная оболочка $U = \langle \boldsymbol{a}_1, \ \ldots, \boldsymbol{a}_m \rangle = \left\{ \sum\limits_{i=1}^{m} \lambda_i \boldsymbol{a}_i : \lambda_i \in \mathbb{F} \right\}$. И в этом есть принципиальное различие между базисом и линейной оболочкой. Базис в $\mathbb{R}^n$ состоит из n векторов, а вот линейная оболочка в $\mathbb{R}^n$ — из всех векторов в $\mathbb{R}^n$ (то есть из несчётного количества векторов при $n > 1$).

Предложение 1. Линейная оболочка U является подпространством.

1. $\sum\limits_{i=1}^{m} \lambda_i \boldsymbol{a}_i + \sum\limits_{i=1}^{m} \lambda'_i \boldsymbol{a}_i = \sum\limits_{i=1}^{m} (\lambda_i + \lambda'_i) \boldsymbol{a}_i \in U$.
2. $\alpha \sum\limits_{i=1}^{m} \lambda_i \boldsymbol{a}_i = \sum\limits_{i=1}^{m} (\lambda_i \cdot \alpha) \boldsymbol{a}_i = \sum\limits_{i=1}^{m} \lambda'_i \boldsymbol{a}_i \in U$.

Предложение 2. Пусть L' — подпространство n-мерного пространства L. Тогда $\dim(L') \leqslant n$. Если $\dim(L') = n$, то L' совпадает с L.

Предложение 3. Размерность линейной оболочки множества из m-векторов не больше m.

Задание подпространства с помощью однородной системы линейных уравнений.

Пусть дано n-мерное линейное пространство L, и пусть в нём зафиксирован базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$. M — линейное подпространство в L.

Будем говорить, что система линейных уравнений задаёт подпространство M, если этой системе удовлетворяют координаты всех векторов из M и не удовлетворяют координаты никаких других векторов.

Из свойств решений однородной системы линейных уравнений следует, что любая однородная линейная система уравнений ранга r с n переменными задаёт в n-мерном пространстве L (если в нём зафиксирован базис) $(n-r)$-мерное линейное подпространство.

Теорема. Пусть L — n-мерное линейное пространство с фиксированным базисом $\boldsymbol{e}$. Тогда любое его m-мерное подпространтсво $M \subset L$ можно задать как множество решений некоторой однородной системы линейных уравнений $A\boldsymbol{x} = \mathbf{0}$, где A — матрица размера $n \times (n-m)$ ранга $(n-m)$.

Доказательство:

Пусть в базисе $\boldsymbol{e}$ вектор $\boldsymbol{x} = \sum\limits_{i=1}^{n} x_i \boldsymbol{e}_i \in L$ имеет координаты $\{x_1, \ldots, x_n\}$. Тогда мы хотим узнать, при каких $\{x_1, \ldots, x_n\}$ этот вектор принадлежит подпространству $M = \langle \boldsymbol{c}_1, \ldots, \boldsymbol{c}_m \rangle$, где, вообще говоря, каждый $\boldsymbol{c}_j$ также выражается в базисе $\boldsymbol{e}$: $\boldsymbol{c}_j = \sum\limits_{i=1}^{n} c_{ij} \boldsymbol{e}_i$.

Значит, в координатах базиса $\boldsymbol{e}$, M — это множество линейных комбинаций векторов-столбцов:

$$C = \begin{pmatrix} c_{11} & \ldots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{n1} & \ldots & c_{nm} \end{pmatrix} \in \mathbb{F}^{n \times m}.$$

Мы считаем систему $\{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m\}$ линейно независимой $\Rightarrow$ по определению имеем:

$$\big[\boldsymbol{x} \in \langle \boldsymbol{c}_1, \ldots, \boldsymbol{c}_m \rangle\big] \;\Leftrightarrow\; \left[\exists\, \lambda_1, \ldots, \lambda_m \in \mathbb{F} : \boldsymbol{x} = \sum_{j=1}^{m} \lambda_j \boldsymbol{c}_j\right].$$

Это равносильно совместности СЛУ с расширенной матрицей.

$$\hat{C} = \left(\begin{array}{ccc|c} c_{11} & \ldots & c_{1m} & x_1 \\ \vdots & \ddots & \vdots & \vdots \\ c_{n1} & \ldots & c_{nm} & x_n \end{array}\right) \sim \left(\begin{array}{ccc|c} c_{1j_1} & \ldots & 0 & x'_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & c_{mj_m} & x'_m \\ \hline & \mathbf{0} & & a_{m+1,1} \cdot x'_1 + \ldots + a_{m+1,n} \cdot x'_n \\ & \vdots & & \vdots \\ & \mathbf{0} & & a_{n,1} \cdot x'_1 + \ldots + a_{n,n} \cdot x'_n \end{array}\right)$$

Итак, нам нужна совместность $\hat{C}$ (иначе говоря, не должно быть противоречий в $\hat{C}$). Тогда для получения ОСЛУ, описывающую оболочку, получаем простое условие:

$$\begin{cases} a_{m+1,1} \cdot x_1 + \ldots + a_{m+1,n} \cdot x_n = 0, \\ \qquad\qquad\vdots \\ a_{n,1} \cdot x_1 + \ldots + a_{n,n} \cdot x_n = 0. \end{cases} \quad \Rightarrow \quad A\boldsymbol{x}^{\uparrow} = \mathbf{0}.$$

Отметим, что $\dim\big(\langle \boldsymbol{c}_1, \ \ldots, \ \boldsymbol{c}_m \rangle\big) = m$ и $\dim(M) = n - \mathrm{rg}(A) = n - (n - m) = m$. В итоге получили ОСЛУ $A\boldsymbol{x}^{\uparrow} = \mathbf{0}$, описывающую оболочку $\langle \boldsymbol{c}_1, \ \ldots, \ \boldsymbol{c}_m \rangle$.

■

Сумма и пересечение подпространств

Определение [сумма подпространств]. Суммой подпространств L' и L'' будем называть линейную оболочку их объединения $L' \cup L''$.
Обозначение: $L' + L''$.

Подробнее определение означает, что вектор x из $L' + L''$ (и только такой) представим в виде $\boldsymbol{x} = \sum\limits_i \alpha_i \boldsymbol{p}_i + \sum\limits_j \beta_j \boldsymbol{q}_j$, где векторы $\boldsymbol{p}_i$ лежат в L', а $\boldsymbol{q}_j$ — в L''.

Достаточно очевидно, что $\dim(L' + L'') \leqslant \dim(L') + \dim(L'')$. Если $L' \subseteq L''$, то $L' + L'' = L''$. В частности, для любого подпространства $L' + L' = L'$.

Определение [пересечение подпространств]. Пересечением подпространств L' и L'' будем называть множество векторов, которые принадлежат обоим подпространствам.
Обозначение: $L' \cap L''$.

Пересечение есть подпространство. Действительно, нулевой вектор лежит во всех подпространствах и, следовательно, пересечение — непустое множество. Если векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ лежат в $L' \cap L''$, то они лежат как в L', так и в L''. Поэтому вектор $\boldsymbol{x} + \boldsymbol{y}$ и при любом α вектор $\alpha \cdot \boldsymbol{x}$ также лежат и в L', и в L'', а значит, и в $L' \cap L''$.

Если в конечномерном пространстве подпространства заданы СЛУ, то их пересечение задаётся системой уравнений, получаемой объединением систем, задающих подпространства.

В более общем случае суммой подпространств $L^1, \ \ldots, \ L^s$ называется линейная оболочка их объединения. Аналогично сумме двух подпространств можем получить для суммы s подпространств:

$$\dim(L^1 + \ldots + L^s) \leqslant \dim(L^1) + \ldots + \dim(L^s).$$

Формула Грассмана. Пусть V_1, V_2 — подпространства в L. $\dim(L) < \infty,\ \dim(V_1) < \infty,\ \dim(V_2) < \infty$. Тогда выполняется $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$.

Доказательство:

Пусть $\dim(V_1 \cap V_2) = s \geqslant 0,\ \dim(V_1) = k_1 \geqslant s,\ \dim(V_2) = k_2 \geqslant s$. Пусть $\boldsymbol{e} = \{\boldsymbol{e}_1,\ \ldots, \boldsymbol{e}_s\}$ — базис в $V_1 \cap V_2$.

Дополним базис $\boldsymbol{e}$ до базиса V_1 векторами $\boldsymbol{f}_1,\ \ldots, \boldsymbol{f}_{k_1-s}$ и до базиса V_2 векторами $\boldsymbol{g}_1,\ \ldots, \boldsymbol{g}_{k_2-s}$. Тогда $\{\boldsymbol{e}_1,\ \ldots, \boldsymbol{e}_s, \boldsymbol{f}_1,\ \ldots, \boldsymbol{f}_{k_1-s}\}$ — базис в V_1, и $\{\boldsymbol{e}_1,\ \ldots, \boldsymbol{e}_s, \boldsymbol{g}_1,\ \ldots, \boldsymbol{g}_{k_2-s}\}$ — базис в V_2. Базис в $(V_1 + V_2):\ \{\boldsymbol{e}_1,\ \ldots, \boldsymbol{e}_s, \boldsymbol{f}_1,\ \ldots, \boldsymbol{f}_{k_1-s}, \boldsymbol{g}_1,\ \ldots, \boldsymbol{g}_{k_2-s}\}$.

А значит, что $\dim(V_1 + V_2) = s + k_1 - s + k_2 - s = k_1 + k_2 - s$.

■

Определение [прямая сумма]. Сумма подпространств $L^1,\ \ldots,\ L^s$ называется прямой суммой, если ее размерность равна сумме размерностей этих подпространств, то есть имеет максимальное из возможных значений. Обозначение: чаще используется "$+$", но если требуется подчеркнуть, что сумма прямая, — то "$\oplus$".

Теорема. Для того, чтобы сумма M подпространств $L^1,\ \ldots,\ L^s$ была прямой суммой, необходимо и достаточно выполнение любого из следующих четырёх свойств:

1. Любая система из $m \leqslant s$ ненулевых векторов, принадлежащих различным подпространствам L^i $(i = \overline{1,\ s})$, линейно независима.
2. Каждый вектор $x \in M$ раскладывается в сумму $x_1,\ \ldots,\ x_s$, где $x_i \in L^i$ $(i = \overline{1,\ s})$, однозначно.
3. Пересечение каждого из подпространств L^i с суммой остальных есть нулевое подпространство.
4. Объединение базисов подпространств L^i $(i = \overline{1,\ s})$ является базисом в M.

Доказательство этой теоремы в силу простоты представляется как упражнение. Но если есть необходимость прочитать доказательство, то оно описано в книге Б. В. Беклемишева "Курс аналитической геометрии и линейной алгебры" в 13 издании на 239 странице.

Предложение 4. Для любого подпространства L' пространства L найдётся такое подпространство L'', что $L = L' \oplus L''$.

Доказательство:

Выберем базис $\{\boldsymbol{e}_1,\ \ldots,\ \boldsymbol{e}_k\}$ подпространства L' и дополним его до базиса пространства L векторами $\{\boldsymbol{e}_{k+1},\ \ldots,\ \boldsymbol{e}_n\}$. Тогда обозначим линейную оболочку $\{\boldsymbol{e}_{k+1},\ \ldots,\ \boldsymbol{e}_n\}$ через L''. Покажем, что в таком случае $L = L' \oplus L''$:

1. $L = L' + L''$. Любой вектор $\boldsymbol{x} \in L$ однозначно раскладывается по полному базису:

$$\boldsymbol{x} = (\alpha_1 \boldsymbol{e}_1 + \ldots + \alpha_k \boldsymbol{e}_k) + (\alpha_{k+1} \boldsymbol{e}_{k+1} + \ldots + \alpha_n \boldsymbol{e}_n).$$

Первая скобка — это вектор из L', вторая — вектор из L''. Значит, $\boldsymbol{x} \in L' + L''$, и поэтому $L \subseteq L' + L''$. Обратное включение $L' + L'' \subseteq L$ очевидно, так как оба подпространства содержатся в L. Значит, $L = L' + L''$.

2. $L' \cap L'' = \{\mathbf{0}\}$. Предположим, существует ненулевой вектор $\boldsymbol{y} \in L' \cap L''$. Тогда его можно разложить по базисам обоих подпространств:

$$\boldsymbol{y} = \beta_1 \boldsymbol{e}_1 + \ldots + \beta_k \boldsymbol{e}_k = \gamma_{k+1} \boldsymbol{e}_{k+1} + \ldots + \gamma_n \boldsymbol{e}_n.$$

Перенеся все слагаемые в одну сторону, получим:

$$\beta_1 \boldsymbol{e}_1 + \ldots + \beta_k \boldsymbol{e}_k - \gamma_{k+1} \boldsymbol{e}_{k+1} - \ldots - \gamma_n \boldsymbol{e}_n = \mathbf{0}.$$

Так как векторы $\{\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_n\}$ линейно независимы, все коэффициенты в этой линейной комбинации должны быть нулевыми. Следовательно, $\boldsymbol{y} = \mathbf{0}$. Что и означает, что пересечение тривиально. Значит, сумма прямая — $L = L' \oplus L''$.

■

Билет № 6. Линейные отображения и линейные преобразования линейного пространства. Ядро и множество значений. Ранг линейного отображения. Условия инъективности, сюръективности и биективности. Операции над линейными преобразованиями.

Небольшое уточнение: обратные преобразования описаны в 7 билете. Я подумал, что эту тему лучше раскрыть после матриц линейного отображения.

Пусть L_1 и L_2 — линейные пространства.

Определение [отображение]. Отображение $\varphi\colon L_1 \to L_2$ — закон, по которому каждому вектору из L_1 ставится в соответствие единственный вектор из L_2. Если отображение φ переводит элемент $\boldsymbol{x} \in X$ в элемент $\boldsymbol{y} \in Y$, то можно записать $\varphi(\boldsymbol{x}) = \boldsymbol{y}$. При этом $\boldsymbol{y}$ называется образом $\boldsymbol{x}$ под действием отображения φ, а $\boldsymbol{x}$ — прообразом $\boldsymbol{y}$.

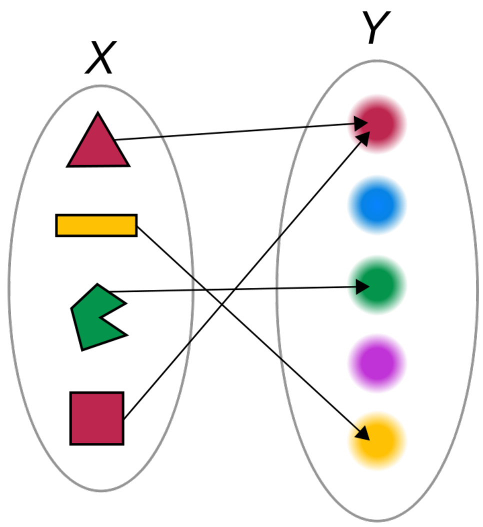

Рис. 1: Пример отображения. Каждому элементу из X соответствует один элемент из Y. Источник: Wikipedia.

Определение [лин. отобр.]. Отображение $\varphi\colon L_1 \to L_2$ линейно, если $\forall\ \boldsymbol{x}, \boldsymbol{y} \in L$ и $\forall \alpha \in \mathbb{R} \to$
1. $\varphi(\boldsymbol{x} + \boldsymbol{y}) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$;
2. $\varphi(\alpha \cdot \boldsymbol{x}) = \alpha \cdot \varphi(\boldsymbol{x})$.

Определение [лин. преобразование]. Линейное отображение называется линейным преобразованием, если пространства L_1 и L_2 совпадают.

Отметим свойства, которыми могут обладать произвольные отображения.

Определение [инъекция]. Отображение φ называется инъективным, если разные элементы отображаются в разные: $\boldsymbol{x}_1, \boldsymbol{x}_2 \in X$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2 \Rightarrow \varphi(\boldsymbol{x}_1) \neq \varphi(\boldsymbol{x}_2)$.

Определение [сюръекция]. Отображение φ называется сюръективным, если у любого элемента из Y есть хотя бы один прообраз: $\forall\ \boldsymbol{y} \in Y\ \exists\ \boldsymbol{x} \in X : \varphi(\boldsymbol{x}) = \boldsymbol{y}$.

Определение [биекция]. Отображение φ называется биективным, если у любого элемента из Y есть единственный прообраз: $\forall \boldsymbol{y} \in Y\ \exists!\ \boldsymbol{x} \in X : \varphi(\boldsymbol{x}) = \boldsymbol{y}$.

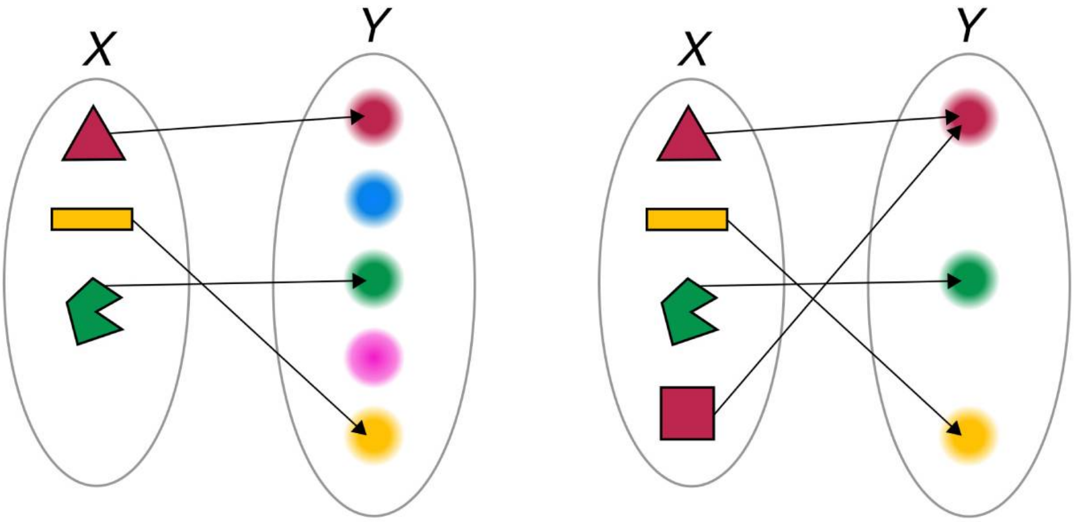

Рис. 2: Пример отображений (слева инъекция, справа сюръекция).
Источник: Методичка моего семинариста (Алексеева Василия Антоновича).

Определение [множество значений]. Множество, связанное с отображением φ — множество значений отображения $\mathrm{Im}(\varphi) \subseteq Y$, определяемое как совокупность всех элементов $\boldsymbol{y} \in Y$, в которые можно попасть под действием отображения: $\mathrm{Im}(\varphi) = \{\boldsymbol{y} \in Y \mid \exists\ \boldsymbol{x} \in X : \varphi(\boldsymbol{x}) = \boldsymbol{y}\}$. То есть сюръективность означает, что $\mathrm{Im}(\varphi) = Y$.

Определение [ядро]. Ядром отображения φ называется подмножество элементов $\mathrm{Ker}(\varphi) \subseteq X$, которые в результате действия φ отображаются в нулевой элемент пространства Y:
$\mathrm{Ker}(\varphi) = \{\boldsymbol{x} \in X \mid \varphi(\boldsymbol{x}) = \mathbf{0}_Y\}$.

Лемма. $\mathrm{Ker}(\varphi)$ является линейным пространством в X.

Доказательство:

Пусть $\boldsymbol{x}_1,\ \boldsymbol{x}_2 \in \mathrm{Ker}(\varphi)$. Значит, $\varphi(\boldsymbol{x}_1) = \mathbf{0}_Y$ и $\varphi(\boldsymbol{x}_2) = \mathbf{0}_Y$. В силу линейности φ имеем:

1. $\varphi(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \varphi(\boldsymbol{x}_1) + \varphi(\boldsymbol{x}_2) = \mathbf{0}_Y + \mathbf{0}_Y = \mathbf{0}_Y \Rightarrow \boldsymbol{x}_1 + \boldsymbol{x}_2 \in \mathrm{Ker}(\varphi)$.

Аналогично возьмём $\boldsymbol{x} \in \mathrm{Ker}(\varphi)$ и $\alpha \in \mathbb{R}$:

2. $\varphi(\alpha \cdot \boldsymbol{x}) = \alpha \cdot \varphi(x) = \alpha \cdot \mathbf{0}_Y = \mathbf{0}_Y \Rightarrow \alpha\boldsymbol{x} \in \mathrm{Ker}(\varphi)$.

■

Так как ядро φ является подпространством X, то в нём всегда как минимум есть нулевой вектор $\mathbf{0}_X$ пространства X.

Лемма [критерий инъективности].

Отображение $\varphi\colon\ L_1 \to L_2$ инъективно $\Leftrightarrow \operatorname{Ker}(\varphi) = \{\mathbf{0}_X\}$.

Доказательство [Необходимость $\Rightarrow$]:

Пусть отображение инъективно. Докажем, что ядро нулевое от противного. То есть пусть существует $\boldsymbol{x}' \in \operatorname{Ker}(\varphi)$, и при этом $\boldsymbol{x}' \neq \mathbf{0}$. В силу линейности φ имеем:
$\varphi(\boldsymbol{x} + \boldsymbol{x}') = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{x}') = \varphi(\boldsymbol{x})$.

Что мы получили: образы векторов $\boldsymbol{x}+\boldsymbol{x}'$ и $\boldsymbol{x}$ совпадают, а прообразы — нет, так как $\boldsymbol{x}' \neq \mathbf{0}_{L_1}$. Это противоречит инъективности, а значит и нашему предположению о том, что $\boldsymbol{x}' \neq \mathbf{0}_{L_1}$.

Доказательство [Достаточность $\Leftarrow$]:

Пусть ядро нулевое. Докажем, что отображение инъективно от противного. То есть $\exists\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in L_1\colon \varphi(\boldsymbol{x}_1) = \varphi(\boldsymbol{x}_2)$ и $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$. В силу линейности φ: $\varphi(\boldsymbol{x}_1) - \varphi(\boldsymbol{x}_2) = \varphi(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \mathbf{0}_{L_2}$. Значит $(\boldsymbol{x}_1 - \boldsymbol{x}_2) \in \operatorname{Ker}(\varphi)$. Так как $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, то в ядре нашёлся ненулевой элемент. Пришли к противоречию, а значит, что отображение инъективно.

■

Лемма. $\operatorname{Im}(\varphi)$ является линейным пространством в Y.

Доказательство:

Пусть $\boldsymbol{y}_1, \boldsymbol{y}_2 \in Y$. У каждого из них есть хотя бы один прообраз, то есть $\exists\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in X$ такие, что $\varphi(\boldsymbol{x}_1) = \boldsymbol{y}_1$, $\varphi(\boldsymbol{x}_2) = \boldsymbol{y}_2$. Покажем замкнутость множества значений в Y относительно сложения и умножения на скаляр ($\alpha \in \mathbb{R}$):

1. $\varphi(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \varphi(\boldsymbol{x}_1) + \varphi(\boldsymbol{x}_2) = \boldsymbol{y}_1 + \boldsymbol{y}_2 \Rightarrow \boldsymbol{y}_1 + \boldsymbol{y}_2 \in Y$;
2. $\varphi(\alpha \cdot \boldsymbol{x}) = \alpha \cdot \varphi(\boldsymbol{x}) = a \cdot \boldsymbol{y} \Rightarrow \alpha \cdot \boldsymbol{y} \in \operatorname{Im}(\varphi)$, где $\boldsymbol{x} \in X$, $\boldsymbol{y} \in Y$, $\varphi(\boldsymbol{x}) = \boldsymbol{y}$.

■

Так как множество значений является подпространством Y, то получим критерий сюръективности.

Лемма [критерий сюръективности].

Отображение $\varphi\colon\ L_1 \to L_2$ сюръективно $\Leftrightarrow \dim\bigl(\operatorname{Im}(\varphi)\bigr) = \dim(Y)$.

Без доказательства в силу очевидности критерия.

Операции над линейными отображениями. Рассмотрим множество всех линейных отображений из X в Y:

$$\mathfrak{F} = \{\varphi\colon X \to Y \mid \varphi - \text{линейное}\}$$

На $\mathfrak{F}$ введём операции сложения отображений и умножения отображения на число. Суммой $\varphi_1 + \varphi_2$ двух линейных отображений φ_1 и φ_2 будем считать отображение $X \to Y$, которое на произвольный $\boldsymbol{x} \in X$ действует как

$$(\varphi_1 + \varphi_2)(\boldsymbol{x}) \equiv \varphi_1(\boldsymbol{x}) + \varphi_2(\boldsymbol{x})$$

А за отображение $\alpha\varphi$, где $\alpha \in \mathbb{R}$ и $\varphi \in \mathfrak{F}$, будем считать отображение $X \to Y$, действующее на произвольный $\boldsymbol{x} \in X$ по правилу:

$$(\alpha\varphi)(\boldsymbol{x}) \equiv \alpha\varphi(\boldsymbol{x})$$

Достаточно легко убедиться в том, что линейные операции над линейными отображениями из $\mathfrak{F}$ дают линейные отображения из $\mathfrak{F}$:

1. Сумма линейных отображений:

$$\begin{aligned}(\varphi_1 + \varphi_2)(\boldsymbol{x}_1 + \boldsymbol{x}_2) &= \varphi_1(\boldsymbol{x}_1 + \boldsymbol{x}_2) + \varphi_2(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \varphi_1(\boldsymbol{x}_1) + \varphi_1(\boldsymbol{x}_2) + \varphi_2(\boldsymbol{x}_1) + \varphi_2(\boldsymbol{x}_2) \\ &= \big(\varphi_1(\boldsymbol{x}_1) + \varphi_2(\boldsymbol{x}_1)\big) + \big(\varphi_1(\boldsymbol{x}_2) + \varphi_2(\boldsymbol{x}_2)\big) = (\varphi_1 + \varphi_2)(\boldsymbol{x}_1) + (\varphi_1 + \varphi_2)(\boldsymbol{x}_2)\end{aligned}$$

2. Умножение отображения на число:

$$(\varphi_1 + \varphi_2)(\alpha\boldsymbol{x}) = \varphi_1(\alpha\boldsymbol{x}) + \varphi_2(\alpha\boldsymbol{x}) = \alpha\varphi_1(\boldsymbol{x}) + \alpha\varphi_2(\boldsymbol{x}) = \alpha\big(\varphi_1(\boldsymbol{x}) + \varphi_2(\boldsymbol{x})\big) = \alpha(\varphi_1 + \varphi_2)(\boldsymbol{x})$$

Более того, множество $\mathfrak{F}$ с введёнными операциями суммы и умножения на число образует линейное пространство, в чём можно убедиться, проверив выполнение **8** свойств.

Определение. Произведением линейных преобразований φ и ψ называется преобразование (обозначаемое через $\varphi\psi$), состоящее в последовательном выполнении сначала преобразования ψ, а затем преобразования φ.

По этому определению

$$(\varphi\psi)(\boldsymbol{x}) = \varphi\big(\psi(\boldsymbol{x})\big),$$

то есть сначала на вектор $\boldsymbol{x}$ действуют преобразованием ψ, а затем на полученный вектор $\psi(\boldsymbol{x})$ действуют преобразованием φ.

Преобразование $\varphi\psi$ линейно, так как:

$$\begin{aligned}(\varphi\psi)(\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2) = \varphi\big(\psi(\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2)\big) = \varphi\big(\lambda_1\psi(\boldsymbol{x}_1) + \lambda_2\psi(\boldsymbol{x}_2)\big) &= \lambda_1\varphi\big(\psi(\boldsymbol{x}_1)\big) + \lambda_2\varphi\big(\psi(\boldsymbol{x}_2)\big) \\ &= \lambda_1(\varphi\psi)(\boldsymbol{x}_1) + \lambda_2(\varphi\psi)(\boldsymbol{x}_2)\end{aligned}$$

И, вообще говоря, $\varphi\psi \neq \psi\varphi$!

Приведём пример, когда $\varphi\psi \neq \psi\varphi$. Рассмотрим два линейных преобразования в $\mathbb{R}^2$:

1. Преобразование φ — проекция на ось Ox:

$$\varphi(x,\, y) = (x,\, 0).$$

2. Преобразование ψ — поворот на 90° против часовой стрелки:

$$\psi(x,\, y) = (-y,\, x).$$

Теперь вычислим композиции $\varphi\psi$ и $\psi\varphi$ для произвольного вектора $\boldsymbol{v} = (x,\, y) \neq \mathbf{0}$:

1. $(\varphi\psi)(x,\, y) = \varphi\big(\psi(x,\, y)\big) = \varphi(-y,\, x) = (-y,\, 0)$.
2. $(\psi\varphi)(x,\, y) = \psi\big(\varphi(x,\, y)\big) = \psi(x,\, 0) = (0,\, x)$.

Отсюда можно заметить, что $\varphi\psi \neq \psi\varphi$.

Анекдот.
Проходит экзамен в виде теста — на каждый вопрос студенты должны отвечать "да" или "нет". Все старательно думают и что-то пишут, и только один студент выбирает ответ, подбрасывая монетку. Ну, думает про себя преподаватель, этот первым сдаст работу. Однако проходит некоторое время, все уже сдали свои работы, а этот студент всё сидит и подбрасывает монетку. Преподаватель не выдерживает и подходит к нему:
— Ну что, ты на все вопросы ответил?
— Да, — отвечает студент.
— А что же ты тогда делаешь?
— Проверяю.

Билет № 7. Матрицы линейного отображения и линейного преобразования для конечномерных пространств. Изменение матрицы линейного отображения (преобразования) при замене базисов. Операции над линейными преобразованиями в координатной (матричной) форме. Обратное преобразование.

Матрица линейного отображения.

Выберем базисы в пространствах X и Y: строки из базисных векторов $\boldsymbol{e} = \{\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_n\} \subset X$ и $\boldsymbol{f} = \{\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_m\} \subset Y$. Рассмотрим действие линейного отображения φ на $\boldsymbol{x} \in X$, столбец компонент которого в базисе $\boldsymbol{e}$ есть столбец $\boldsymbol{\xi} = \{x_1, \ \ldots, \ x_n\}^T$:

$$\varphi(\boldsymbol{x}) = \varphi(x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n) = x_1\varphi(\boldsymbol{e}_1) + \ldots + x_n\varphi(\boldsymbol{e}_n) = \underbrace{\big(\varphi(\boldsymbol{e}_1), \ \ldots, \ \varphi(\boldsymbol{e}_n)\big)}_{\substack{\text{строка}\\ \text{векторов}}} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{\substack{\text{столбец}\\ \text{координат}}}$$

Векторы $\varphi(\boldsymbol{e}_i) \in Y$ можно разложить по базису $\boldsymbol{f}$ (например, для $\varphi(\boldsymbol{e}_1)$):

$$\varphi(\boldsymbol{e}_1) = a_{11}\boldsymbol{f}_1 + \ldots + a_{m1}\boldsymbol{f}_m = (\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_m) \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}$$

Обозначим через $\boldsymbol{a}_i \in \mathbb{R}^m$ столбец $(a_{1i}, \ \ldots, \ a_{mi})^T$ координат вектора $\varphi(\boldsymbol{e}_i)$ в базисе $\boldsymbol{f}$. Тогда для образа вектора $\boldsymbol{x}$ имеем:

$$\varphi(\boldsymbol{x}) = \big(\varphi(\boldsymbol{e}_1), \ldots, \varphi(\boldsymbol{e}_n)\big) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m) \underbrace{(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)}_{A \in \mathbb{R}^{m \times n}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{f} A \boldsymbol{\xi}.$$

С другой стороны, вектор $\varphi(\boldsymbol{x}) \in Y$ раскладывается по базису $\boldsymbol{f}$ с некоторыми коэффициентами $\boldsymbol{\eta} = (y_1, \ \ldots, \ y_m)^T$:

$$\varphi(\boldsymbol{x}) = (\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_m) \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \boldsymbol{f}\boldsymbol{\eta}.$$

Получили два представления для вектора $\varphi(\boldsymbol{x})$:

$$\boldsymbol{f}\boldsymbol{\eta} = \boldsymbol{f} A \boldsymbol{\xi} \Rightarrow \boxed{\boldsymbol{\eta} = A\boldsymbol{\xi}}.$$

Матрицу A называют матрицей линейного отображения в паре базисов $\boldsymbol{e}$ и $\boldsymbol{f}$.

Изменение матрицы линейного отображения.

Наша задача — понять, как изменится матрица линейного отображения φ, когда мы перейдем к новой паре базисов $\boldsymbol{e}'$ и $\boldsymbol{f}'$ пространств X и Y.

Выберем в пространстве X новый базис $\boldsymbol{e}' = \{\boldsymbol{e}'_1,\ \ldots,\ \boldsymbol{e}'_n\}$. Нам известна матрица перехода $S \in \mathbb{R}^{n\times n}$ от старого базиса $\boldsymbol{e}$ к новому $\boldsymbol{e}'$: $\boldsymbol{e}' = \boldsymbol{e}S$. Также в пространстве Y выбран новый базис $\boldsymbol{f}' = \boldsymbol{f}P$, где $P \in \mathbb{R}^{m\times m}$ — матрица перехода от $\boldsymbol{f}$ к $\boldsymbol{f}'$.

При паре базисов $\boldsymbol{e}$ и $\boldsymbol{f}$ в прошлом разделе для произвольного $\boldsymbol{x} \in X$ получили:

$$\varphi(\boldsymbol{x}) = \boldsymbol{f}\boldsymbol{\eta} = \boldsymbol{f}A\boldsymbol{\xi}.$$

Аналогично в паре базисов $\boldsymbol{e}'$ и $\boldsymbol{f}'$ можем получить для того же $\boldsymbol{x}$:

$$\varphi(\boldsymbol{x}) = \boldsymbol{f}'\boldsymbol{\eta}' = \boldsymbol{f}'A'\boldsymbol{\xi}'.$$

Приравняем два представления одного и того же вектора:

$$\boldsymbol{f}A\boldsymbol{\xi} = \boldsymbol{f}'A'\boldsymbol{\xi}'.$$

Учитывая, что $\boldsymbol{f}' = \boldsymbol{f}P$ и $\boldsymbol{e}' = \boldsymbol{e}S \Rightarrow \boldsymbol{\xi} = S\boldsymbol{\xi}' \Rightarrow \boldsymbol{\xi}' = S^{-1}\boldsymbol{\xi}$, получим:

$$\boldsymbol{f}A\boldsymbol{\xi} = \boldsymbol{f}'A'\boldsymbol{\xi}' = (\boldsymbol{f}P)A'(S^{-1}\boldsymbol{\xi}) \to A\boldsymbol{\xi} = PA'S^{-1}\boldsymbol{\xi} \xrightarrow{\forall \boldsymbol{x}\in X} A = PA'S^{-1} \Rightarrow \boxed{A' = P^{-1}AS}.$$

В случае преобразования $\varphi\colon X \to X$:

$$\boxed{A' = S^{-1}AS}.$$

Итак, теперь нам интересно выяснить, как описываются операции над линейными преобразованиями в матричной форме. Пусть преобразования φ и ψ заданы матрицами A и B. Каковы матрицы преобразований $\varphi + \psi$, $\varphi\psi$, $\lambda\varphi$?

Перед ответом на этот вопрос сделаем, однако, очень важные замечания.

- Сумма преобразований $\varphi + \psi$ определена, только если φ и ψ действуют в одном и том же пространстве ($\varphi,\ \psi\colon L \to L$) и заданы в одном базисе $\{\boldsymbol{e}_1,\ \ldots,\ \boldsymbol{e}_n\}$.

- Композиция преобразований $\varphi\psi$ определена, если $\varphi\colon L_1 \to L_2$ и $\psi\colon L_2 \to L_3$, причём базис пространства L_2 должен быть одинаковым для матриц A (в базисах $\boldsymbol{e}$ и $\boldsymbol{f}$) и B (в базисах $\boldsymbol{f}$ и $\boldsymbol{g}$).

$$\text{Пусть } A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \text{ и } B = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix}$$

Теперь считаем $L_1 = L_2 = L_3$. Имеем:

$$\begin{cases} \varphi(\boldsymbol{e}_1) = a_{11}\boldsymbol{e}_1 + a_{12}\boldsymbol{e}_2 + \ldots + a_{1n}\boldsymbol{e}_n \\ \varphi(\boldsymbol{e}_2) = a_{21}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2 + \ldots + a_{2n}\boldsymbol{e}_n \\ \qquad\qquad\vdots \\ \varphi(\boldsymbol{e}_n) = a_{n1}\boldsymbol{e}_1 + a_{n2}\boldsymbol{e}_2 + \ldots + a_{nn}\boldsymbol{e}_n \end{cases} \qquad \begin{cases} \psi(\boldsymbol{e}_1) = b_{11}\boldsymbol{e}_1 + b_{12}\boldsymbol{e}_2 + \ldots + b_{1n}\boldsymbol{e}_n \\ \psi(\boldsymbol{e}_2) = b_{21}\boldsymbol{e}_1 + b_{22}\boldsymbol{e}_2 + \ldots + b_{2n}\boldsymbol{e}_n \\ \qquad\qquad\vdots \\ \psi(\boldsymbol{e}_n) = b_{n1}\boldsymbol{e}_1 + b_{n2}\boldsymbol{e}_2 + \ldots + b_{nn}\boldsymbol{e}_n \end{cases}$$

Чтобы найти матрицу преобразования $\varphi + \psi$, надо найти разложения векторов $(\varphi+\psi)(\boldsymbol{e}_1),\ \ldots,\ (\varphi+\psi)(\boldsymbol{e}_n)$ в базисе $\boldsymbol{e}$:

$$\begin{cases} (\varphi+\psi)(\boldsymbol{e}_1) = \varphi(\boldsymbol{e}_1) + \psi(\boldsymbol{e}_1) = (a_{11}+b_{11})\cdot\boldsymbol{e}_1 + (a_{21}+b_{21})\cdot\boldsymbol{e}_2 + \ldots + (a_{n1}+b_{n1})\cdot\boldsymbol{e}_n \\ (\varphi+\psi)(\boldsymbol{e}_2) = \varphi(\boldsymbol{e}_2) + \psi(\boldsymbol{e}_2) = (a_{12}+b_{12})\cdot\boldsymbol{e}_1 + (a_{22}+b_{22})\cdot\boldsymbol{e}_2 + \ldots + (a_{n2}+b_{n2})\cdot\boldsymbol{e}_n \\ \qquad\qquad\qquad\qquad\vdots \\ (\varphi+\psi)(\boldsymbol{e}_n) = \varphi(\boldsymbol{e}_n) + \psi(\boldsymbol{e}_n) = (a_{1n}+b_{1n})\cdot\boldsymbol{e}_1 + (a_{2n}+b_{2n})\cdot\boldsymbol{e}_2 + \ldots + (a_{nn}+b_{nn})\cdot\boldsymbol{e}_n \end{cases}$$

Теперь осознаём, что матрицей преобразования $\varphi + \psi$ в том же базисе $\boldsymbol{e}$ является матрица $A + B$, то есть сумме преобразований соответствует сумма их матриц.

Аналогично для получения матрицы преобразования $\varphi\psi$ находим разложения векторов $(\varphi\psi)(\boldsymbol{e}_1),\ \ldots,\ (\varphi\psi)(\boldsymbol{e}_n)$ в базисе $\boldsymbol{e}$:

$$\begin{aligned} (\varphi\psi)(\boldsymbol{e}_1) &= \varphi(\psi(\boldsymbol{e}_1)) = \varphi(b_{11}\boldsymbol{e}_1 + \ldots + b_{n1}\boldsymbol{e}_n) = b_{11}\varphi(\boldsymbol{e}_1) + \ldots + b_{n1}\varphi(\boldsymbol{e}_n) = \\ &= (a_{11}b_{11} + a_{12}b_{21} + \ldots + a_{1n}b_{n1})\cdot\boldsymbol{e}_1 + \ldots + (a_{n1}b_{11} + a_{n2}b_{21} + \ldots + a_{nn}b_{n1})\cdot\boldsymbol{e}_n \end{aligned}$$

$$\begin{aligned} (\varphi\psi)(\boldsymbol{e}_2) &= \varphi(\psi(\boldsymbol{e}_2)) = \varphi(b_{12}\boldsymbol{e}_1 + \ldots + b_{n2}\boldsymbol{e}_n) = b_{12}\varphi(\boldsymbol{e}_1) + \ldots + b_{n2}\varphi(\boldsymbol{e}_n) = \\ &= (a_{11}b_{12} + a_{12}b_{22} + \ldots + a_{1n}b_{n2})\cdot\boldsymbol{e}_1 + \ldots + (a_{n1}b_{12} + a_{n2}b_{22} + \ldots + a_{nn}b_{n2})\cdot\boldsymbol{e}_n \end{aligned}$$

$$\vdots$$

$$\begin{aligned} (\varphi\psi)(\boldsymbol{e}_n) &= \varphi(\psi(\boldsymbol{e}_n)) = \varphi(b_{1n}\boldsymbol{e}_1 + \ldots + b_{nn}\boldsymbol{e}_n) = b_{1n}\varphi(\boldsymbol{e}_1) + \ldots + b_{nn}\varphi(\boldsymbol{e}_n) = \\ &= (a_{11}b_{1n} + a_{12}b_{2n} + \ldots + a_{1n}b_{nn})\cdot\boldsymbol{e}_1 + \ldots + (a_{n1}b_{1n} + a_{n2}b_{2n} + \ldots + a_{nn}b_{nn})\cdot\boldsymbol{e}_n \end{aligned}$$

Замечаем, что элемент c_{ij} матрицы преобразования $\varphi\psi$ построен по закону:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}.$$

А это значит, что матрица преобразования $\varphi\psi$ равна произведению матриц A и B, то есть произведению преобразований соответствует произведение их матриц.

Таким образом, множество линейных преобразований пространства $\mathbb{R}^n$ над полем $\mathbb{F}$ относительно операций сложения и умножения изоморфно множеству квадратных матриц порядка n с элементами из поля $\mathbb{F}$. А так как указанное множество матриц образует кольцо, то это же можно сказать и о множестве линейных преобразований.

Обратное преобразование.

Теорема. Линейное преобразование φ пространства $\mathbb{R}^n$ взаимно однозначно тогда и только тогда, когда его матрица в каком-нибудь базисе невырождена.

Доказательство [Необходимость $\Rightarrow$]:

Пусть в некотором базисе преобразование φ имеет матрицу

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Тогда координаты вектора $\varphi(\boldsymbol{x}) = (\eta_1, \ \dots, \eta_n)$ следующим образом выражаются через координаты вектора $\boldsymbol{x} = (x_1, \ \dots, x_n)$:

$$\begin{cases} \eta_1 = a_{11}x_1 + \ldots + a_{1n}x_n \\ \qquad\vdots \\ \eta_n = a_{n1}x_1 + \ldots + a_{nn}x_n \end{cases} \tag{1}$$

Взаимная однозначность преобразования φ означает, что для любого набора чисел $\eta_1, \ \dots, \eta_n$ найдется ровно один набор чисел $x_1, \ \dots, x_n$, удовлетворяющих системе уравнений (1). Но система уравнений (1) имеет единственное решение относительно неизвестных $x_1, \ \dots, x_n$ тогда и только тогда, когда ее определитель отличен от нуля, то есть матрица A невырождена.

Доказательство [Достаточность $\Leftarrow$]:

Пусть матрица A линейного преобразования φ в некотором базисе невырождена (т.е. $\det(A) \neq 0$).

Рассмотрим систему уравнений $A\boldsymbol{x} = 0$. Так как $\det(A) \neq 0$, эта система имеет единственное решение $\boldsymbol{x} = \mathbf{0}$. Следовательно, ядро преобразования тривиально, и φ инъективно.

Так как A невырождена, то для любого $\boldsymbol{\eta}$ система $A\boldsymbol{x} = \boldsymbol{\eta}$ имеет единственное решение $\boldsymbol{x} = A^{-1}\boldsymbol{\eta}$. Значит, φ сюръективно.

Из инъективности и сюръективности следует, что φ является биекцией, то есть взаимно однозначным преобразованием. Более того, обратное преобразование φ^{-1} задаётся матрицей A^{-1}, которая существует благодаря невырожденности A.

■

Определение. Линейное преобразование φ пространства $\mathbb{R}^n$ называется обратимым (или невырожденным), если существует такое линейное преобразование ψ, что

$$\psi\varphi = \varphi\psi = \varepsilon, \tag{2}$$

где ε — тождественное преобразование.

Очевидно, что если некоторое преобразование ψ является решением для равенств (2), то это решение единственно, а само преобразование ψ линейно и невырождено. Такое преобразование называется обратным для φ и обозначается через φ^{-1}, так что

$$\varphi^{-1}\varphi = \varphi\varphi^{-1} = \varepsilon \tag{3}$$

Ясно, что φ обратимо тогда и только тогда, когда оно взаимно однозначно. При этом, поскольку кольцо линейных преобразований изоморфно кольцу матриц, то равенствам (3) будут соответствовать матричные равенства

$$BA = AB = E,$$

где B — матрица линейного преобразования φ^{-1} в том же базисе, что и A для φ. Отсюда как раз видно, что $B = A^{-1}$.

Утверждение [о размерности образа и ядра линейного отображения]:

$$\dim(\mathrm{Im}(\varphi)) + \dim(\mathrm{Ker}(\varphi)) = \dim(L).$$

Доказательство:

Пусть $\varphi : L \to P$, причём $\dim(L) = n$. Выберем в линейном пространстве L произвольный базис $\{\boldsymbol{e}_1,\ \ldots,\ \boldsymbol{e}_n\}$. Поскольку по определению $\mathrm{Im}(\varphi) = \{\boldsymbol{y} \mid \boldsymbol{y} = \varphi(\boldsymbol{x}),\ \boldsymbol{x} \in L\}$, то можно записать, что $\mathrm{Im}(\varphi) = L\{\varphi(\boldsymbol{e}_1),\ \ldots,\ \varphi(\boldsymbol{e}_n)\}$ — линейная оболочка, порождаемая совокупностью образов базисных векторов $\varphi(\boldsymbol{e}_1),\ \ldots,\ \varphi(\boldsymbol{e}_n)$, причём

$$\dim(\mathrm{Im}(\varphi)) = \mathrm{rg}(\varphi(\boldsymbol{e}_1),\ \ldots,\ \varphi(\boldsymbol{e}_n)) = \mathrm{rg}(A_\varphi) = r.$$

Рассмотрим ядро φ : $\mathrm{Ker}(\varphi) = \{\boldsymbol{x} \mid \varphi(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{x} \in L\}$. В выбранном базисе равенству $A_\varphi \boldsymbol{x} = \boldsymbol{0}$ соответствует однородная СЛУ: $A_\varphi \boldsymbol{\xi} = \boldsymbol{0}$, которая, как известно, имеет $n - \mathrm{rg}(A_\varphi) = n - r$ линейно независимых решений, образующих ФСР. Поскольку неизвестными данной системы являются координаты векторов, составляющих $\mathrm{Ker}(\varphi)$, то отсюда заключаем, что $\dim(\mathrm{Ker}(\varphi)) = n - r$. В результате получаем:

$$\dim(\mathrm{Im}(\varphi)) + \dim(\mathrm{Ker}(\varphi)) = r + (n - r) = n = \dim(L).$$

■

Билет № 8. Собственные векторы и собственные значения. Линейная независимость собственных векторов, принадлежащих различным собственным значениям. Инвариантные подпространства линейных преобразований. Ограничение преобразования на инвариантное подпространство. Собственные подпространства.

Немного про смысл происходящего.

Хорошее понимание собственных чисел и собственных векторов весьма принципиально в учёбе. Эти вещи будут преследовать вас как минимум 4 года, а может и больше.

Представьте себе какое-то преобразование, неважно какое — поворот, растяжение, отражение. Собственный вектор — вектор, который не поменял своего направления (в смысле коллинеарности) при этом преобразовании, а только сжался или растянулся. А собственное число как раз показывает коэффициент этого сжатия или растяжения.

Взяли вы, допустим, кусок ткани, начали её растягивать. Поверхность ткани при таком воздействии ведёт себя крайне интересно — по каким-то направлениям некоторые отрезки ткани как бы не поменяли своего направления, они только растянулись из-за внешней силы. То есть, если бы у нас были формальное описание поверхности этой ткани и матрица, отвечающая этой внейшней силе, мы бы вполне могли предсказать, по каким направлениям ткань только бы растянулась без изменения своего направления.

И вообще говоря, это используется везде, где удаётся свести физическую задачу к линейной алгебре. Физический смысл собственных векторов зависит от того, что это за задача, и каким образом она была сведена к линейной алгебре. Например:

- В физике колебаний собственные векторы задают моды колебаний, а собственные значения определяют частоты этих колебаний. Если вы правильно помните из школьной физики, резонанс в колебательной системе происходит на собственной частоте (как видите, смысл слова "собственной" сейчас для нас уже куда более глубокий, чем в 10 классе).

- Если вы хорошо учились в школе и успели заботать квантовую механику, то, конечно, знаете, что операторы (например, гамильтониан) описывают энергию системы. Собственные векторы — это квантовые состояния (например, орбитали электрона). Собственные значения — допустимые уровни энергии (например, $E_n = -13.6\ \frac{\text{эВ}}{n^2}$ в атоме водорода).

- В теории устойчивости важно понимать будет ли система (например, популяция хищников и жертв) стабильной или развалится. Матрица Якоби системы линейных уравнений имеет собственные значения λ. Если $\mathrm{Re}(\lambda) < 0$ — система устойчива. Если $\mathrm{Re}(\lambda) > 0$ — система неустойчива (например, экспоненциальный рост).

Собственные векторы и собственные значения.

Рассмотрим преобразование $\varphi\colon X \to X$. Нам было бы очень удобно работать в таком базисе $\boldsymbol{e}'$, в котором матрица преобразования имеет диагональный вид. Понятно, что такой базис не всегда получится найти.

Пусть найден базис $\boldsymbol{e}'$ такой, что матрица преобразования диагональна: $A' = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, где $\lambda_i \in \mathbb{R}$. В таком базисе действие преобразования φ устроено достаточно просто. Каждый базисный вектор просто умножается на некоторое число:

$$\varphi(\boldsymbol{e}'_i) = \lambda_i \boldsymbol{e}'_i, \quad \forall\, i = \overline{1, n}. \tag{1}$$

Значит, образ каждого базисного вектора $\varphi(\boldsymbol{e}'_i)$ коллинеарен этому же $\boldsymbol{e}'_i$. Теперь остаётся непонятным следующее — возможно ли будет найти векторы, обладающие свойством (1) для произвольного φ?

Определение [собственный вектор]. Ненулевой вектор $\boldsymbol{x} \in X$ называется собственным вектором преобразования $\varphi\colon X \to X$, соответствующим собственному значению $\lambda \in \mathbb{R}$, если $\varphi(\boldsymbol{x}) = \lambda \boldsymbol{x}$.

Выражение из определения выше перепишем в матричном виде (пусть преобразованию φ соответствует матрица A в некотором базисе:

$$A\boldsymbol{x} = \lambda \boldsymbol{x} \Leftrightarrow (A - \lambda E)\boldsymbol{x} = \boldsymbol{0}. \tag{2}$$

Уравнение (2) будет иметь нетривиальное решение в случае, когда определитель системы равен нулю:

$$\boxed{\det(A - \lambda E) = 0}. \tag{3}$$

Определение. Уравнение (3) — характеристическое уравнение для матрицы A.

Лемма. Собственные векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$, соответствующие различным собственным значениям $\lambda_1, \ldots, \lambda_m$, линейно независимы.

Доказательство:

Используем метод математической индукции по числу собственных значений m:

1. База индукции: $m = 1$. Так как собственный вектор $\boldsymbol{x} \neq \boldsymbol{0}$, то система $\{\boldsymbol{x}\}$ линейно независима.
2. Предположение индукции. Пусть для $(m-1)$ векторов утверждение верно.
3. Шаг индукции. Рассмотрим для m:

$$\alpha_1 \boldsymbol{x}_1 + \ldots + \alpha_{m-1}\boldsymbol{x}_{m-1} + \alpha_m \boldsymbol{x}_m = \boldsymbol{0} \mid \cdot \lambda_m \tag{4}$$

$$\alpha_1 \lambda_m \boldsymbol{x}_1 + \ldots + \alpha_{m-1}\lambda_m \boldsymbol{x}_{m-1} + \alpha_m \lambda_m \boldsymbol{x}_m = \boldsymbol{0} \tag{5}$$

Теперь подействуем φ на обе части (4):

$$\alpha_1\varphi(\boldsymbol{x}_1) + \ldots + \alpha_{m-1}\varphi(\boldsymbol{x}_{m-1}) + \alpha_m\varphi(\boldsymbol{x}_m) = \alpha_1\lambda_1\boldsymbol{x}_1 + \ldots + \alpha_{m-1}\lambda_{m-1}\boldsymbol{x}_{m-1} + \alpha_m\lambda_m\boldsymbol{x}_m = \mathbf{0} \quad (6)$$

Вычтем из выражения (5) выражение (6):

$$\alpha_1(\lambda_m - \lambda_1)\boldsymbol{x}_1 + \ldots + \alpha_{m-1}(\lambda_m - \lambda_{m-1})\boldsymbol{x}_{m-1} = \mathbf{0}$$

В силу того, что $\lambda_m - \lambda_i \neq 0$, $i = \overline{1, m-1}$, получаем

$$\alpha_1 = \ldots = \alpha_{m-1} = 0$$

Тогда из (4) выходит, что $\alpha_m\boldsymbol{x}_m = \mathbf{0} \Rightarrow \alpha_m = 0$, так как $\boldsymbol{x}_m \neq \mathbf{0}$.

Итак, мы получили, что только тривиальная комбинация даёт нулевой вектор. А это и значит, что собственные векторы $\boldsymbol{x}_1, \ \ldots, \boldsymbol{x}_m$, соответствующие различным собственным значениям $\lambda_1, \ \ldots, \lambda_m$, линейно независимы.

■

Инвариантные подпространства линейных преобразований.

Рассмотрим линейное пространство L и его линейное преобразование φ.

Определение [инвариантное под-во]. Подпространство $L' \subseteq L$ называется инвариантным относительно φ, если для каждого вектора $\boldsymbol{x}$ из L' образ $\varphi(\boldsymbol{x})$ лежит в L', или, что то же самое: $\varphi(L') \subseteq L'$.

Так, например, нулевое подпространство инвариантно относительно любого преобразования. Также пространство L, рассматриваемое как подпространство, является инвариантным относительно любого преобразования.

Приведём ещё один пример. В $\mathbb{R}^3$ выполняется некоторый поворот φ вокруг оси l, проходящей через точку O. Подпространствами, инвариантными относительно φ, будут:

1. Совокупность векторов, лежащих на оси l;
2. Совокупность векторов, лежащих в плоскости, проходящей через точку O и перпендикулярной оси l.

Преобразование φ каждому вектору из инвариантного подпространства L' сопоставляет вектор из L'. Этим определено преобразование подпространства φ, которое дальше мы будем называть ограничением φ на L' и обозначим φ'. Для векторов из L' по определению $\varphi'(\boldsymbol{x}) = \varphi(\boldsymbol{x})$, а для векторов, не принадлежащих L', преобразование φ' не определено.

Предложение 1. Если преобразования φ и ψ перестановочны (т. е. $\varphi\psi = \psi\varphi$), то ядро и множество значений одного из них инвариантны относительно другого.

Доказательство:
1. Если $\boldsymbol{x} \in \mathrm{Ker}(\varphi)$, то $\varphi(\boldsymbol{x}) = \mathbf{0}$, и потому $\psi(\varphi(\boldsymbol{x})) = \mathbf{0}$. Тогда $\varphi(\psi(\boldsymbol{x})) = \mathbf{0}$, а значит $\psi(\boldsymbol{x}) \in \mathrm{Ker}(\varphi)$.
2. Если $\boldsymbol{x} \in \mathrm{Im}(\varphi)$, то существует вектор $\boldsymbol{z}$ такой, что $\boldsymbol{x} = \varphi(\boldsymbol{z})$. Тогда $\psi(\boldsymbol{x}) = \psi(\varphi(\boldsymbol{z})) = \varphi(\psi(\boldsymbol{z}))$. Это означает, что $\psi(\boldsymbol{x}) \in \mathrm{Im}(\varphi)$.

■

Предложение 2. Матрица линейного преобразования является клеточно-диагональной тогда и только тогда, когда базис есть объединение базисов инвариантных подпространств.

Собственные подпространства

Мы найдем подпространство, инвариантное относительно заданного линейного преобразования φ, если найдем преобразование, перестановочное с φ и имеющее ненулевое ядро. Пусть A — матрица линейного преобразования φ. Перестановочны с A, прежде всего, многочлены от A и, в частности, простейшие из них — линейные. Многочлену от A можно придать вид $A - \lambda E$, где λ — некоторый коэффициент.

Определение. Если для числа λ подпространство $\mathrm{Ker}(A - \lambda E)$ — ненулевое, то λ называется собственным значением преобразования, а подпространство — собственным подпространством, соответствующим (или принадлежащим) собственному значению λ.

Важный частный случай: если преобразование A имеет ненулевое ядро, то это ядро — собственное подпространство, соответствующее собственному значению $\lambda = 0$. Ограничение A на этом инвариантном подпространстве — нулевое преобразование.

Предложение 3. Ограничение преобразования на собственном подпространстве является или нулевым преобразованием, или гомотетией: оно умножает каждый вектор этого подпространства на собственное значение.

Предложение 4. Собственные векторы и только они являются базисными векторами одномерных подпространств, инвариантных относительно φ.

Доказательство:
1. Пусть $\boldsymbol{x}$ — собственный вектор, а $\boldsymbol{y}$ принадлежит одномерному подпространству L' с базисом $\boldsymbol{x}$. Тогда $\boldsymbol{y} = \alpha\boldsymbol{x}$ и $\varphi(\boldsymbol{y}) = \alpha\varphi(\boldsymbol{x}) = \alpha\lambda\boldsymbol{x}$. Значит $\varphi(\boldsymbol{x})$ лежит в L'.
2. Пусть $\boldsymbol{x}$ — базис инвариантного подпространства L'. Тогда $\varphi(\boldsymbol{x})$ лежит в L' и раскладывается по базису: $\varphi(\boldsymbol{x}) = \lambda\boldsymbol{x}$. Так как $\boldsymbol{x} \neq \mathbf{0}$, то он собственный.

■

Предложение 5. В i-ом столбце матрицы линейного преобразования все элементы вне главной диагонали равны нулю тогда и только тогда, когда i-ый базисный вектор — собственный. В этом случае диагональный элемент столбца — собственное значение.

Из предложения 5 вытекает:

Предложение 6. Матрица преобразования A в базисе $\{\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_n\}$ диагональная тогда и только тогда, когда все базисные векторы — собственные. В этом случае диагональные элементы матрицы — соответствующие собственные значения.

Билет № 9. Характеристическое уравнение. Инвариантность характеристического многочлена. Выражение определителя и следа матрицы через корни характеристического много члена. Оценка размерности собственного подпространства. Условия диагонализуемости матрицы линейного преобразования.

Характеристический многочлен.

Определение. Пусть A — квадратная матрица порядка n с элементами a_{ij} из поля $\mathbb{F}$. Тогда многочлен

$$\Delta(\lambda) = |A - \lambda E| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix}$$

называется характеристическим многочленом матрицы A.

Уравнение $|A - \lambda E| = 0$ относительно λ называют характеристическим уравнением, а его корни — характеристическими числами матрицы A. Из определения определителя следует, что $\Delta(\lambda)$ есть многочлен от λ степени n, коэффициент старшего члена равен $(-1)^n$.

В вещественном пространстве в качестве множителей допускаются только вещественные числа, и собственные значения должны быть вещественными. В соответствии с этим имеет место:

Теорема. В комплексном пространстве все корни характеристического уравнения и только они являются собственными значениями. В вещественном пространстве то же справедливо для вещественных корней характеристического уравнения.

Пусть φ — линейное преобразование пространства $\mathbb{R}^n$. Выбирая различные базисы пространства $\mathbb{R}^n$, мы будем получать различные матрицы преобразования φ. Естественно возникает вопрос: зависит ли характеристический многочлен матрицы линейного преобразования от выбора базиса?

Теорема. Характеристический многочлен матрицы линейного преобразования не зависит от выбора базиса.

Доказательство:

Характеристический многочлен $\Delta(\lambda)$ матрицы A линейного преобразования φ есть определитель $|A - \lambda E|$. Как известно, в другом базисе матрица A' того же преобразования φ имеет вид:

$$A' = S^{-1}AS,$$

где S — матрица перехода к новому базису. В новом базисе характеристический многочлен есть определитель матрицы $A' - \lambda E$. Тогда имеем:

$$\det(A' - \lambda E) = \det(S^{-1}AS - \lambda S^{-1}S) = \det(S^{-1}(A - \lambda E)S).$$

Теперь заменим детерминант произведения произведением детерминантов сомножителей:

$$\det(A - \lambda E)\det(S^{-1})\det(S) = \det(A - \lambda E).$$

■

Утверждение [определитель и след матрицы через корни характеристичесго многочлена]. Пусть $\lambda_1, \ \dots, \lambda_n$ — корни характеристического многочлена $\Delta(\lambda) = \det(A - \lambda E)$. Тогда:

1. След матрицы равен сумме корней: $\mathrm{tr}(A) = \lambda_1 + \ldots + \lambda_n$.

2. Определитель матрицы равен произведению корней: $\det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n$.

Доказательство:
1) Сначала посмотрим, о чём идёт речь, на примере матрицы A размерами 2×2.
Пусть $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Её характеристический многочлен:

$$\Delta(\lambda) = \det(A - \lambda E) = \det\begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = (a - \lambda)(d - \lambda) - bc.$$

Раскроем скобки:

$$\Delta(\lambda) = \lambda^2 - (a + d)\lambda + (ad - bc) = (\lambda - \lambda_1)(\lambda - \lambda_2).$$

Корни этого многочлена $\lambda_1, \ \lambda_2$ — это собственные числа матрицы A. По теореме Виета для квадратного уравнения $\lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 = 0$ получаем:

$$\begin{cases} \lambda_1 + \lambda_2 = a + d = \mathrm{tr}(A) \\ \lambda_1 \cdot \lambda_2 = ad - bc = \det(A) \end{cases}$$

2) Рассмотрим в общем виде для произвольного n. Имеем:

$$\det(A - \lambda E) = (a_{11} - \lambda) \cdot \ldots \cdot (a_{nn} - \lambda) + \sigma(\lambda), \tag{1}$$

где под $\sigma(\lambda)$ подразумевается сумма всех членов со степенью $\lambda \leqslant n - 2$ (возникающих для недиагональных элементов матрицы A). Заметим, что можно переписать (1) в другом виде:

$$(-\lambda)^n + (-\lambda)^{n-1} \cdot (a_{11} + \ldots + a_{nn}) + \sigma'(\lambda) + \det(A). \tag{1*}$$

Здесь старший член λ^n возникает только при перемножении диагональных элементов и равен $(-\lambda)^n$. Член степени $n-1$ также определяется только диагональными элементами: при раскрытии определителя он получается, если взять $(-\lambda)$ из $n-1$ скобок и a_{ii} из одной, т.е. имеет вид $(-\lambda)^{n-1}\cdot(a_{11}+\ldots+a_{nn})$. Свободный член равен $\det(A)$, в чём можно убедиться с помощью подстановки $\lambda=0$. А также $\sigma'(\lambda)$ включает в себя все остальные слагаемые.

С другой стороны, по основной теореме алгебры, понятно, что:

$$\Delta(\lambda)=(\lambda_1-\lambda)\cdot\ldots\cdot(\lambda_n-\lambda)=(-\lambda)^n+(-\lambda)^{n-1}\cdot\sum_{i=1}^{n}\lambda_i+\ldots+\prod_{i=1}^{n}\lambda_i \tag{2}$$

Применяем всё могущество своей внимательности. Сравним (1*) и (2):

$$\begin{cases}\operatorname{tr}(A)=a_{11}+\ldots a_{nn}=\lambda_1+\ldots+\lambda_n\\ \det(A)=\lambda_1\cdot\ldots\cdot\lambda_n\end{cases}$$

■

Предложение 1. Если подпространство L' инвариантно относительно преобразования φ, то характеристический многочлен преобразования φ делится нацело на характеристический многочлен его ограничения на L'.

Оценка размерности собственного подпространства.

Теорема. Если собственное значение λ_0 преобразования φ есть корень характеристического многочлена кратности s, то размерность соответствующего собственного подпространства не превосходит s.

Доказательство:

Пусть корню соответствует собственное подпространство размерности k. Выберем там базис $\boldsymbol{e}_1,\ \ldots,\ \boldsymbol{e}_k$. Все векторы этого базиса — собственные с одним и тем же собственным значением λ_0. Поэтому согласно предложению 6 из 8 билета ограничение φ на L' имеет в этом базисе диагональную матрицу, все диагональные элементы которой равны λ_0. Характеристический многочлен этой матрицы есть $q(\lambda)=(\lambda_0-\lambda)^k$. По предложению 1 в этом случае характеристический многочлен $p(\lambda)$ преобразования делится на $(\lambda_0-\lambda)^k$. Следовательно, по определению кратности корня $k\leqslant s$.

■

Условия диагонализуемости матрицы линейного преобразования.

Пусть $\varphi(x)\colon X \to Y$ — линейное отображение, где X и Y — линейные пространства размерностей n и m соответственно. В пространствах X и Y выберем базисы: $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в пространстве X и $\boldsymbol{f} = \{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m\}$ в пространстве Y. В паре базисов $\boldsymbol{e}$ и $\boldsymbol{f}$ линейному отображению φ соответсвует матрица $A \in \mathbb{R}^{n\times m}$. При выборе другой пары базисов $\boldsymbol{e}' = \{\boldsymbol{e}_1', \ldots, \boldsymbol{e}_n'\}$ и $\boldsymbol{f}' = \{\boldsymbol{f}_1', \ldots, \boldsymbol{f}_m'\}$ матрица A', соответствующая линейному отображению φ, будет отличаться от матрицы A (за исключением возможных специальных совпадений). В наших интересах выбрать базисы $\boldsymbol{e}'$ и $\boldsymbol{f}'$ так, чтобы матрица A', как говорит мой семинарист, была "хорошего" для нас вида. Например, исследуем ситуацию, в которой матрица A' имеет диагональный вид.

Итак, рассмотрим сначала выбор пары базисов $\boldsymbol{e}'$ и $\boldsymbol{f}'$ качественно. Пусть $\boldsymbol{e}_1'$ — ненулевой вектор. Посмотрим на его образ $\varphi(\boldsymbol{e}_1')$. Если $\varphi(\boldsymbol{e}_1') = 0$, то первый столбец матрицы A' будет нулевым (что нас устраивает в плане диагональности). Если $\varphi(\boldsymbol{e}') \neq 0$, то положим $\boldsymbol{f}_1' \equiv \varphi(\boldsymbol{e}_1')$. В таком случае первый столбец A' будет первым столбцом единичной матрицы (что нас тоже устраивает).

Теперь выбираем вектор $\boldsymbol{e}_2'$, причем такой, чтобы система $\{\boldsymbol{e}_1', \boldsymbol{e}_2'\}$ была линейно независимой. Снова посмотрим на образ $\varphi(\boldsymbol{e}_2')$. Если $\varphi(\boldsymbol{e}_2') = 0$, то ничего дальше не делаем. А если $\varphi(\boldsymbol{e}_2') \neq 0$, то стоит задаться вопросом — как выбрать $\boldsymbol{f}_2'$ таким образом, чтобы система $\{\boldsymbol{f}_1', \boldsymbol{f}_2'\}$ была линейно независимой? Есть два варианта при ненулевом $\varphi(\boldsymbol{e}_2')$: либо он линейно независим с выбранным ранее $\boldsymbol{f}_1'$, либо линейно зависим с ним, то есть раскладывается по нему. В случае линейной независимости всё просто — просто берем $\boldsymbol{f}_2' \equiv \varphi(\boldsymbol{e}_2')$ (получаем во втором столбце A' второй столбец единичной матрицы). Если же $\varphi(\boldsymbol{e}_2')$ и $\boldsymbol{f}_1'$ линейно зависимы, то $\exists \alpha \in \mathbb{R}$ такая, что $\varphi(\boldsymbol{e}_2') = \alpha \boldsymbol{f}_1' \Rightarrow \varphi(\boldsymbol{e}_2') = \alpha\varphi(\boldsymbol{e}_1')$. Несложно заметить, что можно "поправить" вектор $\boldsymbol{e}_2'$ так, чтобы занулить его образ:

$$\boldsymbol{e}_2'' \equiv \boldsymbol{e}_2' - \alpha \boldsymbol{e}_1' \Rightarrow \varphi(\boldsymbol{e}_2'') = \varphi(\boldsymbol{e}_2' - \alpha\boldsymbol{e}_1') = \varphi(\boldsymbol{e}_2') - \alpha\varphi(\boldsymbol{e}_1') = 0.$$

Причём очевидно, что система $\{\boldsymbol{e}_1', \boldsymbol{e}_2''\}$ является линейно независимой. И вот в таком случае второй столбец матрицы A' будет нулевым.

Для следующих векторов алгоритм остаётся таким же.

В результате мы получили базис $\boldsymbol{e}' = \{\boldsymbol{e}_1', \ldots, \boldsymbol{e}_n'\}$ в X и, возможно, недостроенный базис $\boldsymbol{f}' = \{\boldsymbol{f}_1', \ldots, \boldsymbol{f}_r'\}$, $r \leqslant m$. Недостроенный — хотя бы потому, что не при любом исходе удавалось выбрать вектор в $\boldsymbol{f}'$ при выборе очередного вектора в $\boldsymbol{e}'$ (могло бы в принципе получиться и так, что ни одного $\boldsymbol{f}'$ вообще не было выбрано — если всегда оказывались нулевые столбцы в A'). Но линейно независимую систему $\{\boldsymbol{f}_1', \ldots, \boldsymbol{f}_r'\}$, если в ней недостаточно векторов, всегда можно достроить до базиса в Y, выбрав "как-нибудь" оставшиеся $m - r$ векторов.

Итого, нашли базисы $\boldsymbol{e}'$ и $\boldsymbol{f}'$ в X и Y такие, что матрица A' отображения $\varphi : X \to Y$ имеет диагональный вид.

Если все векторы базиса — собственные, то в нем матрица преобразования диагональная. Для произвольного линейного преобразования такого базиса может не существовать. Если он существует, то говорят, что матрица преобразования приводится к диагональному виду, а преобразование называют *диагонализуемым.*

Предложение 2. Преобразование φ пространства L диагонализуемо тогда и только тогда, когда L совпадает с суммой собственных подпространств φ.

Доказательство:

Если L совпадает с суммой собственных подпространств, то в L есть базис из собственных векторов, так как сумма собственных подпространств — прямая, и объединение их базисов — базис в L. Обратно, если есть базис из собственных векторов, то каждый вектор раскладывается по собственным векторам и потому принадлежит сумме собственных подпространств.

■

Предложение 3. Если преобразование n-мерного пространства L имеет n попарно различных собственных значений, то оно диагонализуемо.

Условие в предложении 3 не является необходимым. Вполне может случиться, что линейное преобразование имеет меньше, чем n, собственных значений, но всё же имеет базис из собственных векторов. Например, мы можем произвольно задать базис и взять преобразование, имеющее в этом базисе диагональную матрицу с несколькими равными элементами на диагонали. Если все элементы диагонали одинаковы (в частности, для тождественного и нулевого преобразований), то каждый ненулевой вектор будет собственным, и в каждом базисе матрица преобразования будет диагональной.

Билет № 10. Теорема Гамильтона–Кэли. Существование двумерного инвариантного подпространства, отвечающего комплексному корню характеристического многочлена линейного преобразования вещественного линейного пространства.

Теорема Гамильтона–Кэли. Если $f(t)$ — характеристический многочлен матрицы A, то $f(A)$ — нулевая матрица. Или иначе — всякая матрица является корнем своего характеристического многочлена.

Доказательство:

Обращу внимание на НЕправильное доказательство! Сразу же хочется сказать, что $f(A) = \det(A - AE) = \det(A - A) = 0$. *Но замечу, что в формулировке теоремы* $f(A)$ *— нулевая матрица, а не просто число 0. В этот момент на экзамене вам уже должно быть грустно от пересдачи..*

Пусть λ — не характеристическое число $\to$ матрица $A - \lambda E$ невырожденная $\to \exists$ матрица

$$(A - \lambda E)^{-1} = \frac{B(\lambda)}{\det(A - \lambda E)}. \tag{1}$$

Элементы матрицы B — элементы вида $b_{ij} = (-1)^{i+j} d_{ji}$, где d_{ji} — дополнительный минор (убираем j-ую строку и i-ый столбец, потом считаем детерминант такой матрицы).

$$A_{nn} = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

Допустим, мы выберем диагональный элемент $[a_{22} - \lambda]$. Тогда дополнительный минор d_{22} — многочлен степени $(n-1)$, где n — количество λ (так как убрали всего одну λ).

Если же выбрать недиагональный элемент, то как ни крути эту матрицу, будем убирать по две λ с диагонали. И дополнительный минор, например, d_{12} будет степени $(n-2)$. Тогда матрица $B(\lambda)$ представима в виде $B(\lambda) = B_0 + \lambda B_1 + \ldots + \lambda^{n-1} B_{n-1}$.

Пример, когда матрицу в таком виде имеет смысл представить:

$$\begin{pmatrix} \lambda^2 - 1 & \lambda + 1 & -\lambda \\ 1 & \lambda^2 + 2 & 3 \\ 0 & 0 & \lambda^2 + \lambda \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & 3 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \lambda + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \lambda^2.$$

Вернёмся обратно к доказательству теоремы.

Для определённости пусть $f(\lambda) = a_0 + a_1\lambda + \ldots + a_n\lambda^n$. Домножим (1) слева и справа на матрицу $A - \lambda E$ и на $\det(A - \lambda E)$:

$$(a_0 + a_1\lambda + \ldots + a_n\lambda^n) \cdot E = (B_0 + \ldots + \lambda^{n-1} B_{n-1}) \cdot (A - \lambda E).$$

Приравняем коэффициенты при одинаковых степенях λ с правой и левой частей. Домножим каждое равенство на A^i, где i — степень, соответствующая степени, рассматриваемого λ^i:

$$\begin{array}{lll}
\lambda^0\colon & a_0E = B_0A & \mid \cdot A^0 \\
\lambda^1\colon & a_1E = -B_0 + B_1A & \mid \cdot A^1 \\
\lambda^2\colon & a_2E = -B_1 + B_2A & \mid \cdot A^2 \\
 & \vdots & \\
\lambda^k\colon & a_kE = -B_{k-1} + B_kA & \mid \cdot A^k \\
 & \vdots & \\
\lambda^n\colon & a_nE = -B_{n-1} & \mid \cdot A^n
\end{array}$$

Сложим все левые и правые части равенств:

$$a_0E + a_1A + \ldots + a_kA^k + \ldots + a_nA^n = B_0A - B_0A + B_1A^2 - \ldots - B_{k-1}A^k + B_kA^{k+1} - \ldots - B_{n-1}A^n.$$

Как можно заметить — всё сократится. Получили нулевую матрицу.

$\blacksquare$

Добавлю, что отсюда мы сразу получаем следующее:
Если $f(\lambda)$ — характеристический многочлен преобразования φ, то $f(\varphi) = \mathbf{0}$.

Теорема. Любое линейное преобразование конечномерного действительного линейного пространства L обладает одномерным или двумерным инвариантным подпространством U.

Доказательство:

Пусть $\dim(L) = n$, и A — матрица преобразования φ в некотором базисе. Рассмотрим два случая.

1. Характеристический многочлен $\Delta(\lambda) = \det(A - \lambda E)$ имеет хотя бы один действительный корень $\lambda \in \mathbb{R}$. Тогда существует ненулевой собственный вектор $\boldsymbol{x} \in L$, для которого $\varphi(\boldsymbol{x}) = \lambda\boldsymbol{x}$.

   Итак, имеем $U = \langle \boldsymbol{x} \rangle$ — одномерное инвариантное подпространство.

2. Все корни $\Delta(\lambda)$ комплексные. Так как коэффициенты $\Delta(\lambda)$ вещественные, то корни идут парами $\lambda = \alpha + i\beta$ и $\overline{\lambda} = \alpha - i\beta$, где $\beta \neq 0$.

   Рассмотрим φ как линейное преобразование $\mathbb{C}^n$. Пусть $\boldsymbol{z} = (z_1, \ \ldots, z_n)^T \in \mathbb{C}$ — собственный вектор, соответствующий λ: $\varphi(\boldsymbol{z}) = A\boldsymbol{z} = \lambda\boldsymbol{z}$, $\boldsymbol{z} \neq \mathbf{0}$.

   Представим $\boldsymbol{z}$ в виде $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y}$, где $\boldsymbol{x} = (\mathrm{Re}(z_1)\ldots\mathrm{Re}(z_n))^T$ и $\boldsymbol{y} = (\mathrm{Im}(z_1)\ldots\mathrm{Im}(z_n))^T$ — вещественные векторы из $\mathbb{R}^n$. Подставим теперь в уравнение $A\boldsymbol{z} = \lambda\boldsymbol{z}$:

$$A(\boldsymbol{x} + i\boldsymbol{y}) = (\alpha + i\beta)(\boldsymbol{x} + i\boldsymbol{y}) = (\alpha\boldsymbol{x} - \beta\boldsymbol{y}) + i(\beta\boldsymbol{x} + \alpha\boldsymbol{y}).$$

Приравнивая вещественные и мнимые части, получаем:

$$A\boldsymbol{x} = \alpha\boldsymbol{x} - \beta\boldsymbol{y} \quad \text{и} \quad A\boldsymbol{y} = \beta\boldsymbol{x} + \alpha\boldsymbol{y}.$$

Причём, $\boldsymbol{x}$ и $\boldsymbol{y}$ линейно независимы над $\mathbb{R}^n$. Ведь если предположить, что $\boldsymbol{y} = k\boldsymbol{x}$, то $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y} = (1 + ik)\boldsymbol{x}$. И тогда из $A\boldsymbol{x} = \lambda\boldsymbol{z}$ следовало бы $A\boldsymbol{x} = \lambda\boldsymbol{x}$, то есть λ было бы вещественным — противоречие.

Итак, имеем $U = \langle\boldsymbol{x}, \boldsymbol{y}\rangle$ — двумерное инвариантное подпространство.

$\blacksquare$

Билет № 11. Линейные функции. Сопряжённое пространство (без доказательств). Биортогональный базис. Пространство, сопряжённое сопряжённому пространству (второе сопряжённое).

Линейные функции.

Определение [функция]. Если на линейном пространстве L задано правило, по которому $\forall \boldsymbol{x} \in L$ ставится в соответствие число из $\mathbb{F}$ ($\mathbb{R}$ или $\mathbb{C}$), то говорят, что на L задана функция f. На бесконечномерных пространствах их называют функционалами.

Определение [линейная функция]. f на L — линейная функция, если $\forall \boldsymbol{x}, \boldsymbol{y} \in L \;\&\; \forall \alpha \in \mathbb{F}:$
1. $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.
2. $f(\alpha\boldsymbol{x}) = \alpha f(\boldsymbol{x})$.

Примеры линейных (или не совсем) функций:
1. Аддитивность массы (масса сложного объекта почти всегда равна сумме масс составляющих его частей). То есть $f(m_1+m_2) = f(m_1) + f(m_2)$.
2. Функция, которая сопоставляет всем векторам ненулевое число, — не линейная. Так как не выполняется $f(\mathbf{0}) = 0$.
3. Функция, сопоставляющая каждому вектору его i-ую координату, — линейная.
4. Функция, сопоставляющая каждому вектору число 0, — тоже линейная.

Сопряжённое пространство.

Определение. Суммой линейных функций f и g называется функция h, значение которой для любого вектора $\boldsymbol{x}$ определяется равенством $h(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$. Произведением линейной функции f на число α называется функция g, значение которой на векторе $\boldsymbol{x}$ определяется как $g(\boldsymbol{x}) = \alpha f(\boldsymbol{x})$.

Пусть L — n-мерное линейное пространство с базисом $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Каждая линейная функция f на L однозначно задаётся своими значениями на базисных векторах: $\varphi_i = f(\boldsymbol{e}_i)$. Строка $\varphi = (\varphi_1, \ldots, \varphi_n)$ называется *строкой коэффициентов* функции f в базисе $\boldsymbol{e}$. Для любого вектора $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n$ имеем

$$f(\boldsymbol{x}) = \varphi_1 x_1 + \ldots + \varphi_n x_n.$$

Предложение 1. Пусть f и g — линейные функции, φ и ψ — их строки коэффициентов в некотором базисе $\boldsymbol{e}$. Тогда

1. Сумма $f+g$ — линейная функция, и её строка коэффициентов равна $\varphi + \psi$.

2. Для произвольного числа α произведение αf — линейная функция, и её строка коэффициентов есть $\alpha\varphi$.

Определение. Множество L^* всех линейных функций на n-мерном линейном пространстве L по отношению к введённым выше линейным операциям представляет собой n-мерное линейное пространство.

Существует взаимно однозначное отображение множества L^* на множество строк длины n, причём сумме функций соответствует сумма строк, а произведению функции на число — произведение её строки на это число. Поскольку аксиомы линейного пространства выполнены для операций со строками, они будут выполнены и для операций в L^*. Следовательно, L^* — линейное пространство, изоморфное пространству строк длины n.

Определение [сопряжённое пространство]. Линейное пространство L^* всех линейных функций на линейном пространстве L называется сопряжённым для L.

Предложение 2. $\dim(L^*) = \dim(L)$.

Доказательство:

$\forall \boldsymbol{x} \in L$, $\boldsymbol{x} = \sum\limits_{i=1}^{n} x_i \boldsymbol{e}_i$ рассмотрим n функций, которые дают значения i-ой координаты аргумента:

$$f_1(\boldsymbol{x}) = x_1, \ \ldots, \ f_i(\boldsymbol{x}) = x_i, \ \ldots, \ f_n(\boldsymbol{x}) = x_n.$$

То есть

$$f_i(\boldsymbol{e}_j) = \delta_{ij} \ = \ \begin{cases} 0, \ i \neq j \\ 1, \ i = j \end{cases}, \text{ где } (i, \ j = \overline{1, n}).$$

Докажем, что $f_1, \ \ldots, \ f_n$ — базис в L^*. Пусть $f = f_1\lambda_1 + \ldots + f_n\lambda_n = 0$. Имеем:

$$f(\boldsymbol{e}_1) = f_1(\boldsymbol{e}_1)\lambda_1 + f_2(\boldsymbol{e}_1)\lambda_2 + \ldots + f_n(\boldsymbol{e}_1)\lambda_n = 1 \cdot \lambda_1 + 0 \cdot \lambda_2 + \ldots + 0 \cdot \lambda_n = \lambda_1 = 0.$$

Аналогично $f(\boldsymbol{e}_2) = \lambda_2 = 0$. И так далее... $\Rightarrow$ $\{f_1, \ \ldots, \ f_n\}$ — ЛНЗ.

Рассмотрим $\forall f \colon L \to \mathbb{F}$ — линейную функцию:

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} a_i x_i = \sum_{i=1}^{n} a_i f_i(\boldsymbol{x}) = \sum_{i=1}^{n} (a_i f_i)(\boldsymbol{x}) = \left(\sum_{i=1}^{n} a_i f_i \right)(\boldsymbol{x})$$

Можем переобозначить $f_i = \boldsymbol{e}_i^*$. Тогда $\{\boldsymbol{e}_i^*\}$ — базис в L^*, он же — *биортогональный* к базису $\{\boldsymbol{e}_j\}$ в L.

■

Определение [биортогональный базис]. Базис $\{\boldsymbol{e}_1^*, \ \ldots, \ \boldsymbol{e}_n^*\}$ в L^*, определяемый формулой

$$\boldsymbol{e}_i^*(\boldsymbol{e}_j) = \delta_{ij} = \begin{cases} 0, \ i \neq j \\ 1, \ i = j \end{cases}, \text{ где } (i, \ j = \overline{1, n}).$$

называется биортогональным или взаимным базису $\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_n$ пространства L.

Элемент $\boldsymbol{f}$ пространства L^* со строкой коэффициентов $(\varphi_1, \ldots, \varphi_n)$, имеет разложение

$$\boldsymbol{f} = \varphi_1 \boldsymbol{e}_1^* + \ldots + \varphi_n \boldsymbol{e}_n^* \tag{1}$$

Введём столбец $\boldsymbol{e}^*$, составленный из функций $\boldsymbol{e}_i^*$. Теперь разложение (1) можно переписать в матричной форме:

$$\boldsymbol{f} = (\varphi_1, \ldots, \varphi_n) \cdot \begin{pmatrix} \boldsymbol{e}_1^* \\ \vdots \\ \boldsymbol{e}_n^* \end{pmatrix} = \varphi \boldsymbol{e}^*. \tag{2}$$

Таким образом, строка координат элемента $\boldsymbol{f} \in L^*$ во взаимном базисе $\boldsymbol{e}^*$ совпадает с его строкой коэффициентов в исходном базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства L. Если для пространства L^* придерживаться соглашения писать компоненты вектора в столбец, а базисные векторы в строку, то формулу (2) следовало бы написать $\boldsymbol{f} = \boldsymbol{e}^{*T} \varphi^T$.

Предложение 3. Матрицей перехода от базиса $\boldsymbol{e}^*$ к базису $\boldsymbol{e}'^*$ в пространстве L^* будет матрица $(S^{-1})^T$, то есть базисы связаны формулой:

$$\boldsymbol{e}'^{*T} = \boldsymbol{e}^{*T} (S^{-1})^T.$$

Пространство, сопряжённое сопряжённому пространству.

Пространство L^* тоже линейное пространство. То есть оно может иметь сопряжённое пространство L^{**}, элементы которого — линейные функции на L^*.

В силу того, что ранее было доказано предложение 2, мы можем утверждать, что если $\dim(L) = n$, то и $\dim(L^{**}) = n$.

Предложение 4. $\exists$ изоморфизм $L^{**} \cong L$, не зависящий от выбора базиса.

Вне билета хочу обратить внимание, что между E и E^{**} существует канонический изоморфизм, определённый соотношением:

$$x \in E \mapsto z \in E^{**},\ z(f) = f(x),\ \forall f \in E^*$$

Канонический изморфизм показывает, что пространства E и E^{**} играют симметричную роль.

Однако, для топологического линейного сопряжения, пространство, сопряжённое к сопряжённому, вообще говоря, с исходным не совпадает. Пространства, для которых $E^{**} = E$, называются рефлексивными — только для них, строго говоря, можно употреблять термин двойственное пространство.

Билет № 12. Билинейные и квадратичные функции (формы). Их координатное представление. Изменение матриц билинейных и квадратичных форм при изменении базиса.

Билинейные и квадратичные функции.

Определение [билинейная форма (БФ)]. Билинейной функцией или билинейной формой на линейном пространстве L называется отображение $b\colon L \times L \to \mathbb{F}$, линейное по каждому аргументу $(\boldsymbol{x},\, \boldsymbol{x}_1,\, \boldsymbol{x}_2,\, \boldsymbol{y},\, \boldsymbol{y}_1,\, \boldsymbol{y}_2 \in L;\ \alpha,\, \beta \in \mathbb{F})$:

1. $b(\boldsymbol{x}_1 + \boldsymbol{x}_2,\, \boldsymbol{y}) = b(\boldsymbol{x}_1,\, \boldsymbol{y}) + b(\boldsymbol{x}_2,\, \boldsymbol{y})$;
2. $b(\boldsymbol{x},\, \boldsymbol{y}_1 + \boldsymbol{y}_2) = b(\boldsymbol{x},\, \boldsymbol{y}_1) + b(\boldsymbol{x},\, \boldsymbol{y}_2)$;
3. $b(\alpha\boldsymbol{x},\, \boldsymbol{y}) = \alpha b(\boldsymbol{x},\, \boldsymbol{y})$;
4. $b(\boldsymbol{x},\, \beta\boldsymbol{y}) = \beta b(\boldsymbol{x},\, \boldsymbol{y})$.

Определение [симметричная БФ]. Билинейная форма называется симметричной, если

$$b(\boldsymbol{x},\, \boldsymbol{y}) = b(\boldsymbol{y},\, \boldsymbol{x}), \quad \forall \boldsymbol{x},\, \boldsymbol{y} \in L.$$

Примеры билинейных функций:

1. $f(\boldsymbol{x},\, \boldsymbol{y}) = x_1y_1 + \ldots + x_ny_n$ в $\mathbb{R}^n$.
2. Скалярное произведение в $\mathbb{R}^n$: $f(\boldsymbol{x},\, \boldsymbol{y}) = (\boldsymbol{x},\, \boldsymbol{y}) = |\boldsymbol{x}||\boldsymbol{y}|\cos(\alpha)$, где α — угол между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$.

Итак, раньше мы говорили о линейных функциях (у них был один аргумент), сейчас мы познакомились с понятием билинейной функции (у них уже два аргумента). Давайте попробуем подать в билинейную функцию один и тот же вектор (получаем снова функцию одного аргумента).

Определение [квадратичная форма (КФ)]. Пусть $b\colon L \times L \to F$ — симметричная билинейная функция. Тогда квадратичной функцией, порождённой функцией b, называется отображение $k\colon L \to \mathbb{F}$, которое определяется формулой: $k(\boldsymbol{x}) = b(\boldsymbol{x},\, \boldsymbol{x})$.

Примеры квадратичных функций:

1. Скалярное произведение $(\boldsymbol{x},\, \boldsymbol{y})$ в $\mathbb{R}^n$ — симметричная билинейная функция. Квадратичная форма $k(\boldsymbol{x}) = b(\boldsymbol{x},\, \boldsymbol{x}) = (\boldsymbol{x},\, \boldsymbol{x})$ сопоставляет вектору $\boldsymbol{x}$ квадрат его длины.
2. Функции $\boldsymbol{x}_1^2 - \boldsymbol{x}_1\boldsymbol{x}_2 + \boldsymbol{x}_3^2,\ \sqrt{2024}\cdot\boldsymbol{x}_1\boldsymbol{x}_2,\ i\boldsymbol{x}_1^2$ являются квадратичными формами.

Уже из приведённых выше примеров понятно, что квадратичная функция в общем случае линейной не является. В качестве примера нелинейности рассмотрим образ сумммы $\boldsymbol{x}_1+\boldsymbol{x}_2$ под действием произвольной квадратичной функции $k(\boldsymbol{x})$:

$$\begin{aligned}k(\boldsymbol{x}_1+\boldsymbol{x}_2) = b(\boldsymbol{x}_1+\boldsymbol{x}_2,\, \boldsymbol{x}_1+\boldsymbol{x}_2) = b(\boldsymbol{x}_1,\boldsymbol{x}_1)+b(\boldsymbol{x}_2,\boldsymbol{x}_2)+b(\boldsymbol{x}_1,\boldsymbol{x}_2)+b(\boldsymbol{x}_2,\boldsymbol{x}_1) = \\ = k(\boldsymbol{x}_1)+k(\boldsymbol{x}_2)+2b(\boldsymbol{x}_1,\boldsymbol{x}_2).\end{aligned}$$

Из последней формулы можно заметить, что билинейную функцию на паре векторов $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ можно выразить с помощью соответстующей ей квадратичной функции:

$$\boxed{b(\boldsymbol{x}_1,\, \boldsymbol{x}_2) = \frac{k(\boldsymbol{x}_1+\boldsymbol{x}_2)-k(\boldsymbol{x}_1)-k(\boldsymbol{x}_2)}{2}}.$$

То есть, как квадратичная форма порождается симметричной билинейной формой, так квадратичная форма *однозначно* восстановливает породившую её симметричную билинейную форму!

Из всех квадратичных функций на пространстве X выделяют несколько классов.

Определение. Квадратичная функция $k(\boldsymbol{x})$ называется

- положительно определённой, если $k(\boldsymbol{x}) > 0$, $\boldsymbol{x} \neq \boldsymbol{0}$;
- отрицательно определённой, если $k(\boldsymbol{x}) < 0$, $\boldsymbol{x} \neq \boldsymbol{0}$;
- положительно полуопределённой, если $k(\boldsymbol{x}) \geqslant 0$, $\forall \boldsymbol{x}$;
- отрицательно полуопределённой, если $k(\boldsymbol{x}) \leqslant 0$, $\forall \boldsymbol{x}$.

Возможно, вам пока не понятно, для чего вообще нужны билинейные и квадратичные формы. Но поверьте, это тот самый математический аппарат, который вам будет встречаться во всех следующих курсах МФТИ: кратные интегралы и теория поля, дифференциальные уравнения, гармонический анализ, аналитическая механика, теория вероятностей, вариационное исчисление, функциональный анализ, уравнения математической физики...

Я приведу один скромный пример применения квадратичной формы, и мы пойдём дальше. В 3 семестре вы будете заниматься нахождением условного экстремума функции многих переменных (то есть нахождением экстремума функции многих переменных при условиях связи — потому он и называется условным экстремумом). Квадратичная форма как раз и используется для изучения поведения функции вблизи критических точек. Она, по сути своей, выступает аналогом второй производной в двумерном случае, чтобы понять: критическая точка это точно максимум/минимум или всё же седловая точка.

Давайте, например, посмотрим на экстремумы функции $u(x,\, y) = x^4 + y^4 - 2x^2$. Первые производные u'_x и u'_y равны нулю в точках $(0,\, 0)$ и $(\pm 1,\, 0)$. Стационарные точки мы нашли, но точно ли они являются максимумами или минимумами? Давайте качественно попытаемся понять это из графика. По аналогии с функциями одной переменной понятно, что точки на "концах штанов" скорее всего являются минимумами, а вот точка $(0,\, 0)$ является седловой, так как в любой её окрестности есть точки как выше, так и ниже её. Это мы можем понять аналитически без графика с помощью анализа квадратичной формы или разложением в ряд Тейлора с остаточным членом в форме Пеано для функции многих переменных ровно так же, как делали это для функции одной переменной с помощью анализа первой и второй производных. Например, если квадратичная форма положительно определена, то рассматриваемый экстремум является локальным минимумом (смотрите, ровно так же как и в случае, когда вторая производная была больше нуля — рассматриваемый экстремум был локальным минимумом!).

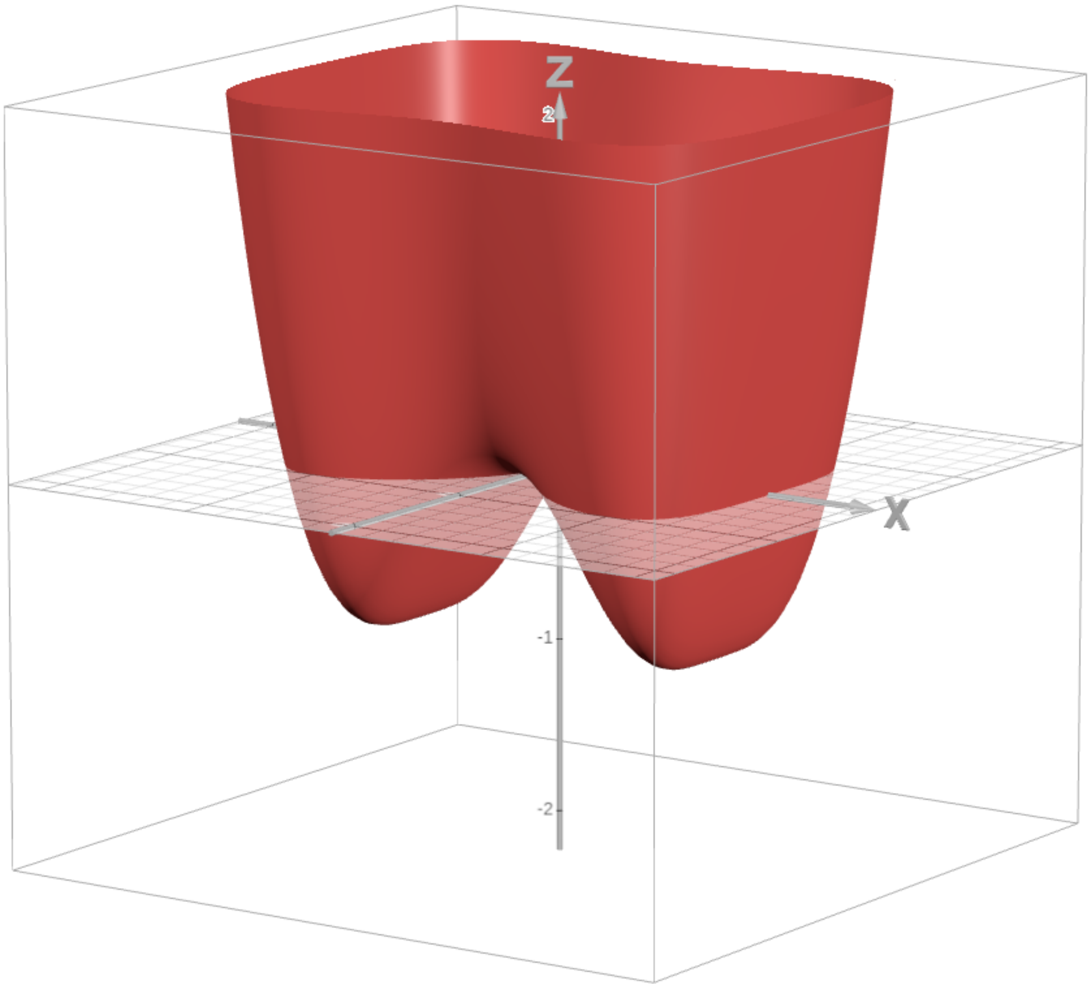

Рис. 3: График функции $u(x,\, y) = x^4 + y^4 - 2x^2$.

Координатное представление БФ и КФ.

Рассмотрим базис $\boldsymbol{e} = \{\boldsymbol{e}_1,\ \ldots,\ \boldsymbol{e}_n\}$. Любой вектор этого пространства можно представить как линейную комбинацию базисных векторов:

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n, \qquad \boldsymbol{y} = y_1\boldsymbol{e}_1 + \ldots + y_n\boldsymbol{e}_n.$$

Подставим это в билинейную функцию:

$$\begin{aligned} b(\boldsymbol{x}, \boldsymbol{y}) &= b(x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n,\, y_1\boldsymbol{e}_1 + \ldots + y_n\boldsymbol{e}_n) \\ &= x_1 b(\boldsymbol{e}_1,\, \boldsymbol{e}_1)y_1 + \ldots + x_1 b(\boldsymbol{e}_1,\, \boldsymbol{e}_n)y_n + \ldots + x_n b(\boldsymbol{e}_n,\, \boldsymbol{e}_1)y_1 + \ldots + x_n b(\boldsymbol{e}_n,\, \boldsymbol{e}_n)y_n \\ &= \sum_{i,\ j=1}^{n} x_i b(\boldsymbol{e}_i,\, \boldsymbol{e}_j)y_j. \end{aligned} \tag{1}$$

Если ввести столбцы $\boldsymbol{x} = (x_1,\ \ldots,\ x_n)^T$ и $\boldsymbol{y} = (y_1,\ \ldots,\ y_n)^T$ и матрицу

$$B = \begin{pmatrix} b(\boldsymbol{e}_1,\, \boldsymbol{e}_1) & \ldots & b(\boldsymbol{e}_1,\, \boldsymbol{e}_n) \\ \vdots & \ddots & \vdots \\ b(\boldsymbol{e}_n,\, \boldsymbol{e}_1) & \ldots & b(\boldsymbol{e}_n,\, \boldsymbol{e}_n) \end{pmatrix} \text{ — матрица билинейной функции,}$$

то выражение для (1) можно записать в следующем виде:

$$\boxed{b(\boldsymbol{x},\, \boldsymbol{y}) = \boldsymbol{x}^T B \boldsymbol{y}}. \tag{2}$$

Утверждение. Билинейная функция b симметрична тогда и только тогда, когда её матрица B в некотором базисе симметрична.

Доказательство [Необходимость $\Rightarrow$]:

Если b симметрична, то:

$$\begin{cases} b(\boldsymbol{x},\, \boldsymbol{y}) = b(\boldsymbol{y},\, \boldsymbol{x}) \\ \forall \boldsymbol{x},\, \boldsymbol{y} \end{cases} \Rightarrow \begin{cases} b(\boldsymbol{e}_i,\, \boldsymbol{e}_j) = b(\boldsymbol{e}_j,\, \boldsymbol{e}_i) \\ \forall i,\, j \end{cases} \Leftrightarrow B = B^T.$$

Доказательство [Достаточность $\Leftarrow$]:

Если матрица B симметрична, то

$$B = B^T \Leftrightarrow \begin{cases} b(\boldsymbol{e}_i,\, \boldsymbol{e}_j) = b(\boldsymbol{e}_j,\, \boldsymbol{e}_i) \\ \forall i,\, j \end{cases} \Rightarrow b(\boldsymbol{x}, \boldsymbol{y}) = \ldots + x_i b(\boldsymbol{e}_i, \boldsymbol{e}_j)y_j + \ldots = \ldots + y_j b(\boldsymbol{e}_j, \boldsymbol{e}_i)x_i + \ldots = b(\boldsymbol{y}, \boldsymbol{x}).$$

■

Для квадратичной функции $k(\,\cdot\,)$, построенной по симметричной билинейной $b(\cdot\,,\,\cdot)$ с матрицей B в базисе $\boldsymbol{e}$ получаем:

$$\boxed{k(\boldsymbol{x}) = \sum_{i,j=1}^{n} x_i b(\boldsymbol{e}_i,\, \boldsymbol{e}_j) x_j = \boldsymbol{x}^T B \boldsymbol{x}}.$$

Изменение матриц БФ и КФ при изменении базиса.

Лемма. Пусть квадратные симметричные матрицы A и B порядка n таковы, что для любого столбца $\boldsymbol{x}$ размером $n \times 1$ выполняется $\boldsymbol{x}^T A \boldsymbol{x} = \boldsymbol{x}^T B \boldsymbol{x}$. Тогда $A = B$.

Доказательство:

Рассмотрим матрицу $C = A - B = (c_{ij})$. Тогда $\boldsymbol{x}^T C \boldsymbol{x} = 0$ для любого столбца $\boldsymbol{x}$. Сначала в качестве $\boldsymbol{x}$ возьмём столбец $\boldsymbol{e}_i$, у которого на i-ом месте стоит единица, а остальные нули. Тогда $\boldsymbol{e}_i^T C \boldsymbol{e}_i = c_{ii} = 0$. Затем в качестве $\boldsymbol{x}$ возьмём столбец $\boldsymbol{e}_{ij}$, у которого на местах с номерами i и j стоят единицы, а остальные нули. Тогда $\boldsymbol{e}_{ij}^T C \boldsymbol{e}_{ij} = c_{ii} + 2c_{ij} + c_{jj} = 0$. Отсюда получаем $c_{ij} = 0$, а значит и $A = B$.

■

При замене базиса матрица билинейной функции, разумеется, меняется. Получим закон её изменения. Пусть в базисе $\boldsymbol{e}$ матрица билинейной функции b есть B. Пусть новый базис $\boldsymbol{e}'$ связан со старым матрицей перехода S, то есть $\boldsymbol{e}' = \boldsymbol{e}S$. Пусть $\boldsymbol{x}'$ и $\boldsymbol{y}'$ — координатные столбцы векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в базисе $\boldsymbol{e}'$. Тогда $\boldsymbol{x} = S\boldsymbol{x}'$ и $\boldsymbol{y} = S\boldsymbol{y}'$. По формуле (2) имеем

$$\boldsymbol{x}' B' \boldsymbol{y}' = b(\boldsymbol{x},\, \boldsymbol{y}) = (S\boldsymbol{x}')^T B (S\boldsymbol{y}') = \boldsymbol{x}'(S^T B S)\boldsymbol{y}'.$$

И это верно для любой пары $\boldsymbol{x}'$, $\boldsymbol{y}'$. Поэтому можно утверждать, что будут равны матрицы "посередине":

$$\boxed{B' = S^T B S}. \tag{3}$$

Билет № 13. Диагональный и канонический вид квадратичных форм. Приведение квадратичных форм к каноническому виду с помощью элементарных преобразований. Приведение квадратичных форм к каноническому виду методом Лагранжа. Закон инерции квадратичных форм. Знакоопределённые квадратичные формы. Критерий Сильвестра.

Определение [диагональная КФ]. Квадратичная форма вида $k(\boldsymbol{x}) = \sum\limits_{i=1}^{n} \alpha_i x_i^2$ называется диагональной. При этом КФ называется канонической, если $\alpha_i \in \{-1,\, 0,\, 1\}$.

Определение [ранг КФ]. Ранг квадратичной формы — число ненулевых коэффициентов в каноническом виде.

Лемма. Ранг матрицы квадратичной формы не зависит от базиса.

Доказательство:

Количество ненулевых элементов на диагонали матрицы B' формы $k(\,\cdot\,)$ в каноническом виде, очевидно, определяет ранг B'. Вот, что можно сказать о ранге матрицы B формы $k(\,\cdot\,)$ в произвольном базисе:

$$B' = S^T B S \Rightarrow \mathrm{rg}(B') = \mathrm{rg}(S^T B S) \xrightarrow[\text{ранга}]{\text{из свойств}} \mathrm{rg}(S^T B) = \mathrm{rg}(B).$$

Итого, ранг матрицы квадратичной формы сохраняется при переходе к новому базису. И в любом каноническом виде одной и той же квадратичной формы число ненулевых значений (± 1) на диагонали будет одинаковым и равным рангу матрицы квадратичной формы.

■

Приведение квадратичных форм к каноническому виду с помощью элементарных преобразований.

Теорема. Любая квадратичная форма может быть приведена к каноническому виду.

Доказательство:

Пусть K — матрица квадратичной формы. Будем делать элементарные преобразования строк и столбцов.

1) Основной случай ($\beta_{11} \neq 0$):

$$K = \begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \dots & \beta_{nn} \end{pmatrix}.$$

Вычитаем из каждой строки первую строку, умноженную на $\frac{\beta_{i1}}{\beta_{11}}$. Тогда получаем:

$$K \sim \left(\begin{array}{c|ccc} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \hline 0 & & & \\ \vdots & & C_1 & \\ 0 & & & \end{array}\right) \quad \big(C_1 \text{ — симметрическая матрица размера } (n-1)\times(n-1)\big).$$

Аналогично поступим и со столбцами. После преобразований:

$$K \sim \left(\begin{array}{c|ccc} \beta_{11} & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & C_1 & \\ 0 & & & \end{array}\right) \quad (C_1 \text{ — та же симметрическая матрица}).$$

2) Особый случай ($\beta_{11} = 0$):

- Если $\forall i : \beta_{i1} = 0$, тогда матрица уже имеет нужный вид.
- Если $\exists i : \beta_{i1} \neq 0$.
  Если при этом $\beta_{ii} \neq 0$, меняем местами первую и i-ую строки и первый с i-ым столбцы.

$$\begin{pmatrix} 0 & \dots & \beta_{1i} & \dots \\ \vdots & \ddots & \vdots & \ddots \\ \beta_{i1} & \dots & \beta_{ii} & \dots \\ \vdots & \ddots & \vdots & \ddots \end{pmatrix} \to \begin{pmatrix} \beta_{i1} & \dots & \beta_{ii} & \dots \\ \vdots & \ddots & \vdots & \ddots \\ 0 & \dots & \beta_{1i} & \dots \\ \vdots & \ddots & \vdots & \ddots \end{pmatrix} \to \begin{pmatrix} \beta_{ii} & \dots & \beta_{i1} & \dots \\ \vdots & \ddots & \vdots & \ddots \\ \beta_{1i} & \dots & 0 & \dots \\ \vdots & \ddots & \vdots & \ddots \end{pmatrix}.$$

  Мы пришли к основному случаю, то есть дальше проделываем алгоритм как тут.

  Если же $\beta_{ii} = 0$, то к первой строке прибавляем i-ую строку и к первому столбцу прибавляем i-ый столбец.

$$\begin{pmatrix} 0 & \dots & \beta_{1i} & \dots \\ \vdots & \ddots & \vdots & \ddots \\ \beta_{i1} & \dots & 0 & \dots \\ \vdots & \ddots & \vdots & \ddots \end{pmatrix} \to \begin{pmatrix} \beta_{i1} & \dots & \beta_{1i} & \dots \\ \vdots & \ddots & \vdots & \ddots \\ \beta_{i1} & \dots & 0 & \dots \\ \vdots & \ddots & \vdots & \ddots \end{pmatrix} \to \begin{pmatrix} 2\beta_{i1} & \dots & \beta_{1i} & \dots \\ \vdots & \ddots & \vdots & \ddots \\ \beta_{1i} & \dots & 0 & \dots \\ \vdots & \ddots & \vdots & \ddots \end{pmatrix}.$$

  И мы снова пришли к основному случаю.

3) Что делаем дальше? Мы привели матрицу квадратичной формы к следующему виду:

$$K \sim \left(\begin{array}{c|ccc} \varepsilon_1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & C_1 & \\ 0 & & & \end{array}\right).$$

Мы начинаем рекурсивно применять алгоритм, описанный в первых двух пунктах для симметрической матрицы в правом нижнем углу. В конце концов мы придём к виду

$$\begin{pmatrix} \varepsilon_1 & 0 & \dots & 0 \\ 0 & \varepsilon_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varepsilon_n \end{pmatrix} \xrightarrow{\text{делим } i\text{-ую строку и } i\text{-ый столбец на } \sqrt{|\varepsilon_i|}} \begin{pmatrix} \operatorname{sign}(\varepsilon_1) & 0 & \dots & 0 \\ 0 & \operatorname{sign}(\varepsilon_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \operatorname{sign}(\varepsilon_n) \end{pmatrix}.$$

Всё, привели к каноническому виду.

■

Приведение КФ к каноническому виду методом Лагранжа.

Метод Лагранжа — приведение квадратичной формы к каноническому виду методом выделения полных квадратов. И лучше всего показать этот метод на конкретном примере. Пусть задана квадратичная форма:

$$k(\boldsymbol{x}) = x_1^2 - x_1x_2 - 2x_1x_3 + \frac{x_2^2}{4} + x_3^2.$$

Способ 1: выделение квадратов.

В формуле $k(\boldsymbol{x})$ есть член x_1^2 , а также члены с x_1 в первой степени. Поэтому приведение к каноническому виду можно начать с того, чтоб собрать вместе все члены с x_1, дополнить, если надо, до квадрата, и сделать замену, выделив квадрат (так, чтобы члены с x_1 "ушли", а вместо этого появился бы только квадрат новой переменной):

$$k(\boldsymbol{x}) = \left(x_1^2 - x_1x_2 - 2x_1x_3 + \left(\frac{x_2}{2}\right)^2 + x_3^2 + 2\cdot\frac{x_2}{2}\cdot x_3\right) - \left(\frac{x_2}{2}\right)^2 - x_3^2 - 2\cdot\frac{x_2}{2}\cdot x_3 + \frac{x_2^2}{4} + x_3^2$$

$$k(\boldsymbol{x}) = \left(x_1 - \frac{x_2}{2} - x_3\right)^2 - x_2x_3.$$

Тогда делаем напрашивающуюся замену:

$$\begin{cases} x_1' = x_1 - \frac{1}{2}x_2 - x_3 \\ x_2' = x_2 \\ x_3' = x_3 \end{cases} \Rightarrow k(\boldsymbol{x}) = x_1'^2 - x_2'x_3'. \tag{1}$$

При приведении формы к диагональному (или каноническому) виду методом Лагранжа может возникнуть ситуация, когда квадрат сразу выделить и нельзя. Как в формуле выше: есть только смешанное произведение $x'_2 x'_3$, но ни $(x'_2)^2$, ни $(x'_3)^2$ нет. В таком случае можно сделать следующую замену ("трюк"):

$$\begin{cases} x'_1 = x''_1 \\ x'_2 = x''_2 - x''_3 \\ x'_3 = x''_2 + x''_3 \end{cases} \tag{2}$$

В результате вместо $x'_2 x'_3$ появится разность квадратов, и стандартный процесс выделения квадратов и замену переменных можно будет продолжить:

$$k(\boldsymbol{x}) = (x''_1)^2 - (x''_2 - x''_3)(x''_2 + x''_3) = (x''_1)^2 - (x''_2)^2 + (x''_3)^2.$$

Привели форму к каноническому виду. Теперь, по-хорошему, надо ещё найти соответствующую замену координат (переход от "старого" исходного базиса к "новому", где канонический вид). Из первой замены (1) получаем:

$$\begin{cases} x_1 = x'_1 + \dfrac{x'_2}{2} + x'_3 \\ x_2 = x'_2 \\ x_3 = x'_3 \end{cases} \quad \leftrightarrow \quad x = S_1 x', \quad S_1 = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

где S_1 — матрица, задающая переход от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$ (в котором координаты с одним "штрихом"). Далее, вторая замена (2):

$$\begin{cases} x'_1 = x''_1 \\ x'_2 = x''_2 - x''_3 \\ x'_3 = x''_2 + x''_3 \end{cases} \quad \leftrightarrow \quad x' = S_2 x'', \quad S_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix},$$

где S_2 — матрица перехода от базиса $\boldsymbol{e}'$ к базису $\boldsymbol{e}''$ (в котором как раз имеем канонический вид формы). Итоговая (комбинированная) замена:

$$\begin{cases} x_1 = x''_1 + \dfrac{3x''_2}{2} + \dfrac{x''_3}{2} \\ x_2 = x''_2 - x''_3 \\ x_3 = x''_2 + x'_3 \end{cases} \quad \leftrightarrow \quad x = S x'', \quad S = \begin{pmatrix} 1 & \frac{3}{2} & \frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Можно проверить, что $S = S_1 S_2$:

$$S_1 S_2 = \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{3}{2} & \frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix} = S.$$

Способ 2: элементарные преобразования матрицы.

При смене базиса матрица формы меняется по правилу (3):

$$B' = S^T BS.$$

Матрица S невырожденная, поэтому может быть представлена как произведение матриц, задающих, например, элементарные преобразования столбцов: $S = S_1 \cdot \ldots \cdot S_N$. Если подставить это в формулу выше:

$$B' = S_N^T \cdot \ldots \cdot S_1^T \cdot B \cdot S_1 \cdot \ldots \cdot S_N.$$

Получается, чтобы получить B', надо провести над B серию преобразований: столбцов (умножение на S_i справа) и строк (умножение на S_i^T слева), причём после преобразования столбцов должно сразу идти аналогичное преобразование строк (например, первая пара преобразований: $S_1^T BS_1$).

Для нахождения диагональной матрицы B' надо "приписать", например, справа к исходной матрице единичную, и выполнить над ней те же самые элементарные преобразования:

$$(B \mid E) \to (S_1^T BS_1 \mid S_1) \to \ldots \to (S_N^T \ldots S_1^T BS_1 \ldots S_N \mid \underbrace{S_1 \ldots S_N}_{S}).$$

Возвращаясь к форме из задачи, её матрица:

$$B = \begin{pmatrix} 1 & -\frac{1}{2} & -1 \\ -\frac{1}{2} & \frac{1}{4} & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Получим из неё диагональную:

$$\left(\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & -1 & 1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{4} & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & -\frac{1}{2} & 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & \frac{3}{2} & \frac{1}{2} \\ 0 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right),$$

где первый переход — это результат преобразования "ко второму столбцу прибавить половину первого, к третьему прибавить первый" (и то же самое со строками — но только для матрицы "слева"), а второй — это "ко второму столбцу прибавить третий, а из третьего вычесть второй" (и потом так же со строками).

Итого, матрица формы B' в каноническом виде и матрица перехода S к соответствующему базису (зависящая от выбора последовательности преобразований):

$$B' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & \frac{3}{2} & \frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Закон инерции квадратичных форм.

Определение [индекс инерции]. Пусть $L^{(+)}$ — подпространство максимальной размерности, где $k(\boldsymbol{x})$ положительно определена. Тогда $\dim(L^{(+)})$ — положительный индекс инерции. Аналогично для $L^{(-)}$.

Теорема [закон инерции КФ]. Число "+1" в каноническом виде квадратичной формы не зависит от канонического вида, к которому её привели, и равно положительному индексу инерции. Аналогично для отрицательного индекса инерции.

Доказательство:

Размерность пространства n. Без ограничения общности считаем, что мы привели квадратичную форму к каноническому виду ($r \leqslant n$):

$$k(\boldsymbol{x}) = x_1^2 + \ldots + x_j^2 - x_{j+1}^2 - \ldots - x_r^2.$$

Число "+1" равно j. Покажем, что j и равно положительному индексу инерции.
Рассмотрим $L_1 = \langle \boldsymbol{e}_1, \ \ldots, \boldsymbol{e}_j \rangle$ и $L_2 = \langle \boldsymbol{e}_{j+1}, \ \ldots, \boldsymbol{e}_n \rangle$. Причём, $\dim(L_1) = j$, $\dim(L_2) = n - j$.

Рассмотрим $\boldsymbol{x} \in L_1$:

$$k(\boldsymbol{x}) = x_1^2 + \ldots + x_j^2,$$

потому что, если $\boldsymbol{x} \in L_1$, то все его компоненты, начиная с $j+1$ равны 0. Если мы возьмем $\boldsymbol{x} \neq \boldsymbol{0}$, то хотя бы одна из компонент $\boldsymbol{x}$ не равна нулю. Значит, $k(\boldsymbol{x}) > 0$ на L_1.

Рассмотрим $\boldsymbol{x} \in L_2$:

$$k(\boldsymbol{x}) = -x_{j+1}^2 - \ldots - x_r^2,$$

тогда $k(\boldsymbol{x}) \leqslant 0$ на L_2, если даже $\boldsymbol{x} \neq \boldsymbol{0}$.

Покажем теперь, что j — максимально возможная размерность от противного.
Пусть $\exists L^{(+)}$: $k(\boldsymbol{x}) > 0$ на $L^{(+)}$ и $\dim(L^{(+)}) > j$. Мы знаем, что $L^{(+)}$ и L_2 — подпространства пространства L и при этом $\dim(L^{(+)}) + \dim(L_2) > n$. Это значит, что их пересечение ненулевое $\Rightarrow \exists \boldsymbol{z} \neq \boldsymbol{0} : \boldsymbol{z} \in L^{(+)} \cap L_2$.

Из того, что $\boldsymbol{z} \in L^{(+)}$ следует, что $k(\boldsymbol{z}) > 0$. А из того, что $\boldsymbol{z} \in L_2$ следует, что $k(\boldsymbol{z}) \leqslant 0$. Получили противоречие.

Значит, $\dim(L^{(+)}) = j$, и пространство L_1 — пространство максимальной размерности. Его размерность равна количеству "+1" в каком-то каноническом виде. А это значит, что в любом каноническом виде количество "+1", "–1" и "0" одинаковое.

■

Определение [сигнатура]. Сигнатура — разность между положительным и отрицательным индексами инерции.

Итого, мы можем выделить 4 параметра квадратичной формы: положительный индекс инерции, отрицательный индекс инерции, ранг, сигнатура. И чтобы однозначно охарактеризовать квадратичную форму, нам достаточно знать только 2 параметра из 4.

Повторим определение из билета № 12.

Определение. Квадратичная функция $k(\boldsymbol{x})$ называется

- положительно определённой, если $k(\boldsymbol{x}) > 0$, $\boldsymbol{x} \neq \boldsymbol{0}$;
- отрицательно определённой, если $k(\boldsymbol{x}) < 0$, $\boldsymbol{x} \neq \boldsymbol{0}$;
- положительно полуопределённой, если $k(\boldsymbol{x}) \geqslant 0$, $\forall \boldsymbol{x}$;
- отрицательно полуопределённой, если $k(\boldsymbol{x}) \leqslant 0$, $\forall \boldsymbol{x}$.

Теорема [критерий Сильвестра.] Квадратичная форма положительно определена $\Leftrightarrow$

$$\Leftrightarrow \text{угловые миноры } M_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} > 0 \quad \forall k = \overline{1,\, n}, \text{ где } K = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}.$$

Доказательство [Необходимость $\Rightarrow$]:

На диагонали матрицы квадратичной формы стоят значения квадратичной формы от базисных векторов. Так как квадратичная форма положительно определена, и базисные векторы ненулевые, то

$$k(\boldsymbol{e}_i) > 0 \ \forall i, \text{ где } \boldsymbol{e}_i \text{ — } i\text{-ый базисный вектор.}$$

Значит, диагональные элементы матрицы квадратичной формы в любом базисе положительны. Мы можем начать приводить матрицу квадратичной формы к каноническому виду с помощью элементарных преобразований строк и столбцов, причём особый случай точно не возникнет, на диагонали будут стоять ненулевые элементы.

При преобразованиях из каждой строки вычитаем линейную комбинацию строк, расположенных "выше", и из каждого столбца вычитаем только столбцы, расположенные "левее". То есть главные миноры меняться не будут. А так как в диагональном виде все диагональные элементы положительны, то все главные миноры в диагональном виде положительны, а значит и исходные миноры положительны.

Доказательство [Достаточность $\Leftarrow$]:

Пусть все главные миноры положительны. На k-ом шаге приведения к каноническому виду матрицы квадратичной формы мы получили следующий вид для B_k. Докажем по индукции, что особого случая не возникает, и $\varepsilon_i > 0 \quad \forall i = \overline{1,\, n}$.

$$
B_k = \left(\begin{array}{cccc|ccc}
\varepsilon_1 & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & \vdots & \ddots & \vdots \\
0 & \dots & 0 & \varepsilon_k & 0 & \dots & 0 \\
\hline
0 & \dots & \dots & 0 & & & \\
\vdots & \ddots & \ddots & \vdots & & C_k & \\
0 & \dots & \dots & 0 & & &
\end{array}\right).
$$

Используем метод математической индукции:

1. База индукции.

$$
B = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}, \text{ где } a_{11} > 0 \ \Rightarrow \varepsilon_1 > 0 \ \Rightarrow \text{особого случая не возникает.}
$$

2. Предположение индукции:

Пусть для k-ого шага включительно не возникает особых случаев, и $\varepsilon_k > 0$.

3. Шаг индукции.

$$
B_k = \left(\begin{array}{cccc|c|cc}
\varepsilon_1 & 0 & \dots & 0 & 0 & \dots & 0 \\
0 & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & \vdots & \ddots & \vdots \\
0 & \dots & 0 & \varepsilon_k & 0 & \dots & 0 \\
\hline
0 & \dots & \dots & 0 & \gamma & \dots & \vdots \\
\hline
\vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \dots & \dots & 0 & \dots & \dots & \cdot
\end{array}\right).
$$

Найдём M_{k+1}, например, по последней строке:

$$
M_{k+1} = (-1)^{(k+1)+(k+1)} \cdot \gamma \cdot M_k \Rightarrow \gamma = \frac{M_{k+1}}{M_k} > 0.
$$

Значит, особого случая не возникает, и $\varepsilon_{k+1} > 0$. В итоге получаем:

$$
B = \begin{pmatrix} \varepsilon_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \varepsilon_n \end{pmatrix},
$$

что и значит, что K положительно определена.

■

Следствие. Квадратичная форма $k(\boldsymbol{x})$ с матрицей K отрицательно определена $\Leftrightarrow$ знаки её ведущих главных миноров $M_1, \ \ldots, \ M_n$ чередуются, начиная с отрицательного:

$$\operatorname{sign}(M_s) = (-1)^s \quad \forall s = \overline{1, \, n},$$

то есть $\operatorname{sign}(M_1) = -1$, $\operatorname{sign}(M_2) = +1$, $\operatorname{sign}(M_3) = -1$, и так далее.

Доказательство:

Пусть $k(\boldsymbol{x})$ отрицательно определена. У этой квадратичной формы матрица K. Тогда $-k(\boldsymbol{x})$ положительно определена, и у этой квадратичной формы матрица $-K$.

Пусть M_s — s-ый главный минор квадратичной формы $k(\boldsymbol{x})$, и M'_s — s-ый главный минор квадратичной формы $-k(\boldsymbol{x})$. Причём, связаны они следующим образом:

$$M_s = (-1)^s M'_s, \text{ где } M'_s > 0 \quad \forall s = \overline{1, \, n}.$$

Значит, $\operatorname{sign}(M_s) = (-1)^s \quad \forall s = \overline{1, \, n}$.

$\blacksquare$

Билет № 14. Аксиоматика евклидова пространства. Неравенство Коши–Буняковского. Неравенство треугольника. Матрица Грама и её свойства. Переход от одного ортонормированного базиса к другому. Ортогональные матрицы.

Аксиоматика евклидова пространства.

Линейное пространство, введённое в предыдущих билетах, существенно отличается от множества векторов обычного геометрического пространства тем, что в линейном пространстве не определены понятия длины вектора и угла между векторами. Теперь мы опишем такие пространства, в которых эти понятия определены. В первом семестре, используя длину вектора и угол, мы определили скалярное произведение. Здесь удобнее поступить наоборот. Мы аксиоматически определим операцию скалярного умножения, а длину и угол определим с её помощью. Определение скалярного умножения для вещественных и для комплексных пространств формулируется различно.

Определение [евклидово пр-во]. Вещественное линейное пространство $\mathbb{E}$ называется евклидовым пространством, если каждой упорядоченной паре элементов $\boldsymbol{x}$ и $\boldsymbol{y}$ из $\mathbb{E}$ поставлено в соответствие вещественное число $(\boldsymbol{x},\, \boldsymbol{y})$, называемое скалярным произведением, так, что выполнены аксиомы:

1. $\forall \boldsymbol{x},\, \boldsymbol{y} \in \mathbb{E}\colon (\boldsymbol{x},\, \boldsymbol{y}) = (\boldsymbol{y},\, \boldsymbol{x})$;
2. $\forall \boldsymbol{x}_1,\, \boldsymbol{x}_2,\, \boldsymbol{y} \in \mathbb{E},\, \forall \alpha,\, \beta \in \mathbb{R}\colon (\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2,\, \boldsymbol{y}) = \alpha(\boldsymbol{x}_1,\, \boldsymbol{y}) + \beta(\boldsymbol{x}_2,\, \boldsymbol{y})$;
3. $\forall \boldsymbol{x} \in \mathbb{E},\, \boldsymbol{x} \neq \boldsymbol{0}\colon (\boldsymbol{x},\, \boldsymbol{x}) > 0$.

Кроме того, геометры традиционно требуют, чтобы $\mathbb{E}$ было конечномерным. Но для приложений в теории функций, где функциональные пространства бесконечномерные, большинство нижеследующих результатов остаются в силе. Данные аксиомы 1—4 в совокупности означают, что скалярное произведение есть билинейная (что следует из пунктов 2 и 3) и симметричная (1 пункт) функция, которая, кроме того, порождает положительно определенную квадратичную (4 пункт) функцию. Любая билинейная функция, обладающая данными свойствами, может быть принята за скалярное произведение.

Пример по мотивам задачи № 25.1.

Пусть $\boldsymbol{n}$ есть фиксированный ненулевой вектор в геометрическом пространстве. Можно ли принять за скалярное произведение функцию $(\boldsymbol{x},\, \boldsymbol{y})_1 \equiv (\boldsymbol{n},\, \boldsymbol{x},\, \boldsymbol{y})$?

Понятно, что нет, потому что, скалярный квадрат вектора в таком случае $(\boldsymbol{x},\, \boldsymbol{x})_1 = (\boldsymbol{n},\, \boldsymbol{x},\, \boldsymbol{x})$ равен нулю (объём параллелепипеда). Симметричность также "не работает". Однако линейность по первому аргументу есть.

А можно ли определить скалярное произведение как $(\boldsymbol{x}, \boldsymbol{y})_2 \equiv (\boldsymbol{x}+\boldsymbol{n},\, \boldsymbol{y}+\boldsymbol{n})$?

Снова нет, потому что, например, при $\boldsymbol{x} = -\boldsymbol{n}$ получим $(\boldsymbol{x},\, \boldsymbol{x})_2 = 0$. То есть нет положительной определённости. Линейности по первому аргументу также нет:

$$(\boldsymbol{x}_1+\boldsymbol{x}_2,\, \boldsymbol{y})_2 = \big((\boldsymbol{x}_1+\boldsymbol{x}_2)+\boldsymbol{n},\, \boldsymbol{y}+\boldsymbol{n}\big) = (\boldsymbol{x}_1,\, \boldsymbol{y}+\boldsymbol{n}) + (\boldsymbol{x}_2+\boldsymbol{n},\, \boldsymbol{y}+\boldsymbol{n})$$

(не хватает вектора $\boldsymbol{n}$ как слагаемого в первом аргументе у скобки справа, поэтому, очевидно, в общем случае нелинейно).

Рассмотрим следующую функцию — "кандидат" в скалярное: $(\boldsymbol{x},\, \boldsymbol{y})_3 \equiv (\boldsymbol{n},\, \boldsymbol{x})(\boldsymbol{n},\, \boldsymbol{y})$. Которая тоже не будет скалярным произведением, потому что при $\boldsymbol{x} \perp \boldsymbol{n}$ получается нуль:
$(\boldsymbol{x},\, \boldsymbol{x})_3 = (\boldsymbol{n},\, \boldsymbol{x})(\boldsymbol{n},\, \boldsymbol{x}) = 0$.

Функция же $(\boldsymbol{x},\, \boldsymbol{y})_4 \equiv |\boldsymbol{n}|(\boldsymbol{x},\, \boldsymbol{y})$ удовлетворяет всем свойствам скалярного.

А вариант $(\boldsymbol{x},\, \boldsymbol{y})_5 \equiv |\boldsymbol{x}||\boldsymbol{y}|$, отличающийся от "обычного скалярного" для векторов — направленных отрезков лишь тем, что нет косинуса угла. На самом деле вместе с этим лишается и свойства линейности. Например (при некоторых ненулевых $\boldsymbol{x}$ и $\boldsymbol{y}$):

$$(\boldsymbol{x}-\boldsymbol{x},\, \boldsymbol{y})_5 = |\boldsymbol{x}-\boldsymbol{x}||\boldsymbol{y}| = 0 \neq |\boldsymbol{x}||\boldsymbol{y}| + |-\boldsymbol{x}||\boldsymbol{y}| = (\boldsymbol{x},\, \boldsymbol{y})_5 + (-\boldsymbol{x},\, \boldsymbol{y})_5$$

Определение [длина вектора]. Пусть L — евклидово пространство со скалярным произведением $(\cdot\ ,\ \cdot)$. Длиной (нормой) вектора $\boldsymbol{x} \in L$ называется неотрицательное действительное число

$$|\boldsymbol{x}| = \sqrt{(\boldsymbol{x},\, \boldsymbol{x})}.$$

Определение [угол между векторами]. Пусть $\boldsymbol{x},\, \boldsymbol{y}$ — ненулевые векторы в евклидовом пространстве. Углом между векторами $\boldsymbol{x}$ и $\boldsymbol{y}$ называется единственное число $\varphi \in [0,\, \pi]$, для которого

$$\cos(\varphi) = \frac{(\boldsymbol{x},\, \boldsymbol{y})}{|\boldsymbol{x}||\boldsymbol{y}|}.$$

Неравенство Коши–Буняковского. Неравенство треугольника.

1. $\forall \boldsymbol{x},\, \boldsymbol{y} \in \mathbb{E} \;\Rightarrow\; |(\boldsymbol{x},\, \boldsymbol{y})| \leqslant |\boldsymbol{x}||\boldsymbol{y}|$. Причём, $|(\boldsymbol{x},\, \boldsymbol{y})| = |\boldsymbol{x}||\boldsymbol{y}|$, если $\boldsymbol{x},\, \boldsymbol{y}$ линейно зависимы.
2. $\forall \boldsymbol{x},\, \boldsymbol{y} \in \mathbb{E} \;\Rightarrow\; |\boldsymbol{x}+\boldsymbol{y}| \leqslant |\boldsymbol{x}| + |\boldsymbol{y}|$.

Доказательство:

1. Введём функцию $f(t) = (t\boldsymbol{x}+\boldsymbol{y},\, t\boldsymbol{x}+\boldsymbol{y}) = t^2(\boldsymbol{x},\, \boldsymbol{x}) + 2t(\boldsymbol{x},\, \boldsymbol{y}) + (\boldsymbol{y},\, \boldsymbol{y}) \geqslant 0$.
Тогда относительно t имеем: $D \leqslant 0 \Rightarrow \frac{1}{4}D = (\boldsymbol{x},\, \boldsymbol{y})^2 - (\boldsymbol{x},\, \boldsymbol{x})(\boldsymbol{y},\, \boldsymbol{y}) \leqslant 0$, таким образом $|(\boldsymbol{x},\, \boldsymbol{y})| \leqslant |\boldsymbol{x}||\boldsymbol{y}|$.
2. $(\boldsymbol{x}+\boldsymbol{y},\, \boldsymbol{x}+\boldsymbol{y}) = |\boldsymbol{x}|^2 + 2(\boldsymbol{x},\, \boldsymbol{y}) + |\boldsymbol{y}|^2 \leqslant |\boldsymbol{x}|^2 + 2|\boldsymbol{x}||\boldsymbol{y}| + |\boldsymbol{y}|^2 = (|\boldsymbol{x}|+|\boldsymbol{y}|)^2$, тогда $|\boldsymbol{x}+\boldsymbol{y}| \leqslant |\boldsymbol{x}|+|\boldsymbol{y}|$. ■

Матрица Грама и её свойства.

Все результаты, полученные ранее для симметричных билинейных и квадратичных функций, переносятся и на скалярное произведение. Так, скалярное тоже можно вычислять с помощью матрицы. Пусть в пространстве $\mathbb{E}$ выбран базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_n\}$. Тогда любой вектор $\boldsymbol{x} \in \mathbb{E}$ можно разложить по базису, а коэффициенты разложения собрать в координатный столбец $\boldsymbol{\xi}$:

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n = (\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \boldsymbol{e}\boldsymbol{\xi}.$$

Тогда, если при вычислении скалярного $(\boldsymbol{x}, \boldsymbol{y})$ подставить вместо векторов их разложения по базису:

$$(\boldsymbol{x}, \boldsymbol{y}) = (x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n, \, y_1\boldsymbol{e}_1 + \ldots + y_n\boldsymbol{e}_n) = \sum_{i, \ j=1}^{n} x_i(\boldsymbol{e}_i, \boldsymbol{e}_j)y_j = \boldsymbol{x}^T\Gamma\boldsymbol{y}.$$

Итак,

$$\boxed{(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T\Gamma\boldsymbol{y}}. \tag{1}$$

Матрица Γ билинейной функции $(\ \cdot\ ,\ \cdot\ )$ также может быть названа матрицей Грама системы векторов $\{\boldsymbol{e}_1, \ \ldots, \ \boldsymbol{e}_n\}$:

$$\Gamma = \begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & \ldots & (\boldsymbol{e}_1, \boldsymbol{e}_n) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{e}_n, \boldsymbol{e}_1) & \ldots & (\boldsymbol{e}_n, \boldsymbol{e}_n) \end{pmatrix}$$

Как матрица симметричной билинейной функции, матрица Γ, очевидно, симметрична: $\Gamma = \Gamma^T$. Помимо этого, так как Γ есть матрица положительно определённой квадратичной формы, то её определитель положителен $\det(\Gamma) > 0$. Более того, положительны все главные миноры: $\Delta_i > 0, \quad i = \overline{1, n}$ (критерий Сильвестра).

Переход от одного ортонормированного базиса к другому. Ортогональные матрицы.

Базис, в котором квадратичная форма имеет канонический вид, называется ортонормированным базисом. Так как она положительно определена, матрица Грама ортонормированного базиса — единичная: $(\boldsymbol{e}_i, \boldsymbol{e}_j) = 0$ при $i \neq j$ и $(\boldsymbol{e}_i, \boldsymbol{e}_i) = 1$ $(i, j = \overline{1, n})$. Это значит, что векторы ортонормированного базиса попарно ортогональны, а по длине равны единице.

Для ортонормированного базиса формула (1) имеет вид

$$(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T\boldsymbol{y}.$$

В n-мерном арифметическом пространстве $\mathbb{R}^n$ стандартный базис является ортонормированным относительно стандартного скалярного произведения. Число $\sqrt{(\boldsymbol{x}, \boldsymbol{x})} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}}$ называется евклидовой нормой столбца $\boldsymbol{x}$.

Определение [ортогональная матрица]. Матрица S называется ортогональной, если она невырожденная и $S^{-1} = S^T$.

Лемма. Ортогональные матрицы и только они могут быть матрицами перехода от одного ОНБ к другому.

Доказательство [Необходимость $\Rightarrow$]:

Пусть $\boldsymbol{e}$ — ОНБ. S — ортогональная матрица перехода от $\boldsymbol{e}$ к $\boldsymbol{e}'$. Тогда

$$\Gamma' = S^T \Gamma S = S^T E S = S^T S = E.$$

Доказательство [Достаточность $\Leftarrow$]:

Пусть $\boldsymbol{e}$ и $\boldsymbol{e}'$ — ОНБ. Тогда в этих базисах матрицы Грама $\Gamma = \Gamma' = E$. Мы уже знаем, что $\Gamma' = S^T \Gamma S$. Тогда имеем

$$E = S^T E S = S^T S \Rightarrow S^{-1} = S^T.$$

■

Лемма. Пусть S — ортогональная матрица. Тогда $\det(S) = \pm 1$.

Доказательство:

$$1 = \det(S^{-1}S) = \det(S^T S) = \det(S^T) \cdot \det(S) = \det(S) \cdot \det(S) = \det(S)^2 \Rightarrow \det(S) = \pm 1$$

■

Лемма. Матрица, обратная к ортогональной, является ортогональной.

Доказательство:

$$S^{-1} = S^T \Rightarrow (S^{-1})^T = (S^T)^T \Rightarrow S^{-1}(S^{-1})^T = S^{-1}(S^T)^T = S^{-1}S = E \Rightarrow S^{-1}(S^{-1})^T = E$$

■

Лемма. Произведение ортогональных матриц — ортогональная матрица.

Доказательство:

Пусть S, U — ортогональные матрицы. Распишем $(SU)^{-1}$:

$$(SU)^{-1} = U^{-1}S^{-1} = U^T S^T = (SU)^T \Rightarrow SU\text{— ортогональная матрица.}$$

■

Билет № 15. Ортогональное дополнение подпространства. Ортогональное проектирование на подпространство. Ортогональные базисы. Процесс ортогонализации.

Лемма. Пусть $\boldsymbol{x} \in \mathbb{E}$ и $\boldsymbol{x}$ ортогонален любому вектору из $\mathbb{E}$. Тогда $\boldsymbol{x} = \boldsymbol{0}$.

Доказательство:

$$\forall \boldsymbol{y} \in \mathbb{E} \to (\boldsymbol{x},\, \boldsymbol{y}) = 0 \Rightarrow (\boldsymbol{x},\, \boldsymbol{x}) = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}.$$

■

Лемма. Пусть $\boldsymbol{h}_1,\ \ldots,\ \boldsymbol{h}_n$ — ненулевые попарно ортогональные векторы в n-мерном евклидовом пространстве $\mathbb{E}$. Тогда $\boldsymbol{h}_1,\ \ldots,\ \boldsymbol{h}_n$ образуют базис, и разложение $\forall \boldsymbol{x} \in \mathbb{E}$ по этому базису задаётся формулой

$$\boldsymbol{x} = \sum_{i=1}^{n} \frac{(\boldsymbol{x}, \boldsymbol{h}_i)}{|\boldsymbol{h}_i|^2} \boldsymbol{h}_i.$$

Доказательство:

Рассмотрим матрицу

$$\begin{pmatrix} (\boldsymbol{h}_1,\, \boldsymbol{h}_1) & \ldots & (\boldsymbol{h}_1,\, \boldsymbol{h}_n) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{h}_n,\, \boldsymbol{h}_1) & \ldots & (\boldsymbol{h}_n,\, \boldsymbol{h}_n) \end{pmatrix} \xrightarrow{(\boldsymbol{h}_i,\ \boldsymbol{h}_j)\ =\ 0,\ i \neq j} \begin{pmatrix} (\boldsymbol{h}_1,\, \boldsymbol{h}_1) & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & (\boldsymbol{h}_n,\, \boldsymbol{h}_n) \end{pmatrix}.$$

Тогда определитель этой матрицы больше нуля, так как по условию $\boldsymbol{h}_1,\ \ldots,\ \boldsymbol{h}_n$ — ненулевые, а на диагоналях стоят скалярные квадраты. Определитель матрицы Грама системы векторов равен нулю тогда и только тогда, когда эти векторы линейно зависимы. А значит система векторов $\{\boldsymbol{h}_1,\ \ldots,\ \boldsymbol{h}_n\}$ — ЛНЗ.

Пусть $\boldsymbol{x} = \alpha_1 \boldsymbol{h}_1 + \ldots + \alpha_n \boldsymbol{h}_n$. Умножая это равенство скалярно на любой из $\boldsymbol{h}_i$, находим, что $\alpha_i = \frac{(\boldsymbol{x}, \boldsymbol{h}_i)}{|\boldsymbol{h}_i|^2}$. То есть разложение вектора по этому базису задается формулой

$$\boldsymbol{x} = \sum_{i=1}^{n} \frac{(\boldsymbol{x}, \boldsymbol{h}_i)}{|\boldsymbol{h}_i|^2} \boldsymbol{h}_i. \tag{1}$$

■

Определение [ортогональное дополнение]. Пусть $\mathbb{E}'$ — k-мерное подпространство в n-мерном евклидовом пространстве $\mathbb{E}$. Ортогональным дополнением подпространства $\mathbb{E}'$ называется множество всех векторов, ортогональных каждому вектору из $\mathbb{E}'$. Это множество обозначается $\mathbb{E}'^{\perp}$.

Лемма. Ортогональное дополнение k-мерного подпространства в n-мерном пространстве есть $(n-k)$-мерное линейное подпространство.

Доказательство:

Пусть $\boldsymbol{a}_1, \ \ldots, \boldsymbol{a}_k$ — базис в $\mathbb{E}'$. Пусть $\boldsymbol{y} \perp \boldsymbol{a}_i \ \forall i$. Покажем, что $\boldsymbol{y}$ ортогонален любому вектору из $\mathbb{E}'$. Возьмём $\boldsymbol{x} \in \mathbb{E}'$ и рассмотрим $(\boldsymbol{x}, \boldsymbol{y})$:

$$(\boldsymbol{x}, \boldsymbol{y}) = (\alpha_1 \boldsymbol{a}_1 + \ldots + \alpha_k \boldsymbol{a}_k, \boldsymbol{y}) = \alpha_1(\boldsymbol{a}_1, \boldsymbol{y}) + \ldots + \alpha_k(\boldsymbol{a}_k, \boldsymbol{y}) = 0.$$

В $\mathbb{E}$ выберем ОНБ, тогда координаты $\boldsymbol{a}_1$ в этом базисе будут $(\alpha_{11}, \ \ldots, \alpha_{1n}), \ \ldots,$ координаты $\boldsymbol{a}_k$ в этом базисе — $(\alpha_{k1}, \ \ldots, \alpha_{kn})$, а координаты $\boldsymbol{y}$ будут $(\eta_1, \ \ldots, \eta_n)$. Тогда

$$\begin{cases} (\boldsymbol{y}, \boldsymbol{a}_1) = \alpha_{11}\eta_1 + \ldots + \alpha_{1n}\eta_n \\ \qquad\qquad \vdots \\ (\boldsymbol{y}, \boldsymbol{a}_k) = \alpha_{k1}\eta_1 + \ldots + \alpha_{kn}\eta_n \end{cases} \Rightarrow \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{k1} & \ldots & \alpha_{kn} \end{pmatrix} \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Мы уже знаем, что система линейных уравнений задаёт некоторое подпространство. Причём размерность этого подпространства — (размерность пространства "минус" ранг матрицы). Ранг матрицы равен k, так как строки ЛНЗ. Значит подпространство размерности $(n-k)$.

■

Лемма. Для подпространств $L, L_1, L_2 \in \mathbb{E}$ справедливы равенства:
1) $(L^\perp)^\perp = L$; 2) $(L_1 + L_2)^\perp = L_1^\perp \cap L_2^\perp$; 3) $(L_1 \cap L_2)^\perp = L_1^\perp + L_2^\perp$ (без доказательства).

Доказательство:

1) $\boldsymbol{z} \in (L^\perp)^\perp \Leftrightarrow (\boldsymbol{z}, \boldsymbol{y}) = 0 \quad \forall \boldsymbol{y} \in L^\perp$, то есть $(L^\perp)^\perp \subseteq L$.
И при этом

$$\dim\big((L^\perp)^\perp\big) = n - \dim(L^\perp) = n - \big(n - \dim(L)\big) = \dim(L) \Rightarrow (L^\perp)^\perp = L.$$

2) Пусть $\boldsymbol{z} \in L_1^\perp \cap L_2^\perp$; $\quad \forall \boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 \in L_1 + L_2 \to (\boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{z}) = (\boldsymbol{x}_1, \boldsymbol{z}) + (\boldsymbol{x}_2, \boldsymbol{z}) = 0 \Rightarrow$
$\Rightarrow \boldsymbol{z} \in (L_1 + L_2)^\perp \Rightarrow L_1^\perp \cap L_2^\perp \subseteq (L_1 + L_2)^\perp$.
И при этом

$$\dim\big((L_1 + L_2)^\perp\big) = n - \dim(L_1 + L_2) = n - \dim(L_1) - \dim(L_2) + \dim(L_1 \cap L_2) =$$

$$= \big(n - \dim(L_1)\big) + \big(n - \dim(L_2)\big) - \big(n - \dim(L_1 \cap L_2)\big) = \dim(L_1^\perp) + \dim(L_2^\perp) - \dim\big((L_1 \cap L_2)^\perp\big) =$$

$$= \dim(L_1^\perp \cap L_2^\perp) \Rightarrow (L_1 + L_2)^\perp = L_1^\perp \cap L_2^\perp.$$

Ортогональные базисы. Процесс ортогонализации.

Теорема. Пусть $\mathbb{E}'$ — подпространство евклидова пространства, тогда $\mathbb{E} = \mathbb{E}' \oplus (\mathbb{E}')^{\perp}$.

Доказательство:

Вспомним критерий прямой суммы: сумма двух подпространств прямая, если сумма размерностей подпространств равна размерности пространства, и у подпространств нулевое пересечение.

Сумма размерностей подпространств: $\dim(\mathbb{E}') + \dim((\mathbb{E}')^{\perp}) = k + (n-k) = n$.

Также очевидно, что $\mathbb{E}' \cap (\mathbb{E}')^{\perp} = \mathbf{0}$, потому что только нулевой вектор может лежать и в $\mathbb{E}'$, и в $(\mathbb{E}')^{\perp}$.

■

Определение [ортогональные пространства]. Пространства $\mathbb{E}'$ и $\mathbb{E}''$ ортогональны, если $\mathbb{E}'' \subseteq (\mathbb{E}')^{\perp}$. Но тогда $\mathbb{E}' \subseteq (\mathbb{E}'')^{\perp}$.

Определение [ортогональная проекция]. Так как $\mathbb{E} = \mathbb{E}' \oplus (\mathbb{E}')^{\perp}$, каждый вектор $\boldsymbol{x} \in \mathbb{E}$ однозначно раскладывается в сумму векторов $\boldsymbol{x}_1 \in \mathbb{E}'$ и $\boldsymbol{x}_2 \in (\mathbb{E}')^{\perp}$. Вектор $\boldsymbol{x}_1$ называется ортогональной проекцией $\boldsymbol{x}$ на $\mathbb{E}'$. Аналогично $\boldsymbol{x}_2$ — ортогональная проекция $\boldsymbol{x}$ на $(\mathbb{E}')^{\perp}$.

Найдём ортогональную проекцию $\boldsymbol{x}$ на $\mathbb{E}'$ в предположении, что в $\mathbb{E}'$ задан некоторый *ортогональный!* базис $\boldsymbol{h}_1, \ \ldots, \ \boldsymbol{h}_k$. Дополним этот базис до ортогонального базиса в пространстве $\mathbb{E}$, присоединив к нему произвольный ортогональный базис $\boldsymbol{h}_{k+1}, \ \ldots, \ \boldsymbol{h}_n$ из $(\mathbb{E}')^{\perp}$. Так как сумма $\mathbb{E}$ и $(\mathbb{E}')^{\perp}$ — прямая, искомое разложение вектора $\boldsymbol{x}$ единственно, и мы, группируя слагаемые в формуле (1), получаем

$$\boldsymbol{x}_1 = \sum_{i=1}^{k} \frac{(\boldsymbol{x}, \boldsymbol{h}_i)}{|\boldsymbol{h}_i|^2} \boldsymbol{h}_i \tag{2}$$

Теорема. Пусть $\boldsymbol{x}_1$ — ортогональная проекция $\boldsymbol{x}$ на $\mathbb{E}' \in \mathbb{E}$, $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$. Тогда для всех $\boldsymbol{y} \in \mathbb{E}'$ таких, что $\boldsymbol{y} \neq \boldsymbol{x}_1$ выполняется:

$$|\boldsymbol{x}_2| = |\boldsymbol{x} - \boldsymbol{x}_1| < |\boldsymbol{x} - \boldsymbol{y}|.$$

Доказательство:

Пусть $\boldsymbol{z} = \boldsymbol{x}_1 - \boldsymbol{y} \in \mathbb{E}', \quad \boldsymbol{z} \neq \mathbf{0}$. Тогда

$$|\boldsymbol{x}-\boldsymbol{y}|^2 = |\boldsymbol{x}_1+\boldsymbol{x}_2-\boldsymbol{y}|^2 = |\boldsymbol{x}_2+\boldsymbol{z}|^2 = (\boldsymbol{x}_2+\boldsymbol{z}, \boldsymbol{x}_2+\boldsymbol{z}) = (\boldsymbol{x}_2, \boldsymbol{x}_2)+2(\boldsymbol{x}_2, \boldsymbol{z})+(\boldsymbol{z}, \boldsymbol{z}) = (\boldsymbol{x}_2, \boldsymbol{x}_2)+(\boldsymbol{z}, \boldsymbol{z})$$

$$|\boldsymbol{x} - \boldsymbol{y}|^2 = |\boldsymbol{x}_2|^2 + |\boldsymbol{z}|^2 > |\boldsymbol{x}_2|^2 \Rightarrow |\boldsymbol{x}_2| < |\boldsymbol{x} - \boldsymbol{y}|.$$

■

Формула (2) служит основой метода, позволяющего произвольный базис евклидова пространства преобразовать в ортогональный, а затем и в ортонормированный. Этот метод называется методом ортогонализации Грама–Шмидта.

Пусть в $\mathbb{E}$ задан некоторый базис $\{\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_n\}$. Положим $\boldsymbol{h}_1 = \boldsymbol{f}_1$. Затем из вектора $\boldsymbol{f}_2$ вычтем его ортогональную проекцию на линейную оболочку $\boldsymbol{h}_1$ и положим $\boldsymbol{h}_2$ равным полученной разности:

$$\boldsymbol{h}_2 = \boldsymbol{f}_2 - \frac{(\boldsymbol{f}_2, \, \boldsymbol{h}_1)}{|\boldsymbol{h_1}|^2}\boldsymbol{h}_1.$$

Отметим, что $\boldsymbol{h}_2$ раскладывается по $\boldsymbol{f}_1$ и $\boldsymbol{f}_2$, причём $\boldsymbol{h}_2 \neq \boldsymbol{0}$, так как в противном случае $\boldsymbol{f}_1$ и $\boldsymbol{f}_2$ были бы пропорциональны.

Будем продолжать таким же образом. Допустим, что построены попарно ортогональные ненулевые векторы $\boldsymbol{h}_1, \ \ldots, \boldsymbol{h}_k$, причём для любого $i \leqslant k$ вектор $\boldsymbol{h}_i$ раскладывается по $\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_i$. Положим

$$\boldsymbol{h}_{k+1} = \boldsymbol{f}_{k+1} - \sum_{i=1}^{k} \frac{(\boldsymbol{f}_{k+1}, \, \boldsymbol{h}_i)}{|\boldsymbol{h}_i|^2}\boldsymbol{h}_i.$$

Вектор $\boldsymbol{h}_{k+1}$ — проекция $\boldsymbol{f}_{k+1}$ на ортогональное дополнение линейной оболочки $\boldsymbol{h}_1, \ \ldots, \ \boldsymbol{h}_k$ и, потому, ортогонален всем $\boldsymbol{h}_i$ при $i \leqslant k$. Таким образом, векторы $\boldsymbol{h}_1, \ \ldots, \ \boldsymbol{h}_{k+1}$ попарно ортогональны.

Кроме того, $\boldsymbol{h}_{k+1}$ раскладывается по $\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_{k+1}$, так как для любого $i \leqslant k$ вектор $\boldsymbol{h}_i$ раскладывается по $\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_i$. Отсюда следует, что $\boldsymbol{h}_{k+1} \neq \boldsymbol{0}$, поскольку иначе векторы $\boldsymbol{f}_1, \ \ldots, \ \boldsymbol{f}_{k+1}$ оказались бы линейно зависимы.

После того, как будет преобразован последний вектор $\boldsymbol{f}_n$, мы получим ортогональную систему из n ненулевых векторов.

Итак, нами построен ортогональный базис $\boldsymbol{h}$. От него можно перейти к ортонормированному базису $\boldsymbol{e}$ из векторов $\boldsymbol{e}_i = \frac{\boldsymbol{h}_i}{|\boldsymbol{h}_i|}, \quad (i = \overline{1, \, n})$.

Билет № 16. Линейные преобразования евклидова пространства. Преобразование, сопряжённое данному. Матрица сопряжённого преобразования. Свойства сопряжённого преобразования.

Определение [сопряжённое преобразование]. Преобразование φ^* называется сопряжённым к φ, если $\forall \boldsymbol{x},\, \boldsymbol{y} \in \mathbb{E} \to \big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = \big(\boldsymbol{x},\, \varphi^*(\boldsymbol{y})\big)$.

Рассмотрим примеры:

1) $\varphi = \mathrm{id}$ (тождественное преобразование). Тогда:

$$\big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = \big(\boldsymbol{x},\, \varphi^*(\boldsymbol{y})\big) \Rightarrow \big(\boldsymbol{x},\, \boldsymbol{y} - \varphi^*(\boldsymbol{y})\big) = 0 \quad \forall \boldsymbol{x},\, \boldsymbol{y} \in \mathbb{E}.$$

Значит,

$$\boldsymbol{y} = \varphi^*(\boldsymbol{y})\ \forall \boldsymbol{y} \Rightarrow \varphi^* = \varphi = \mathrm{id}\,.$$

2) $\varphi = \mathbf{0}$ (нулевое преобразование).

$$\big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = \big(\boldsymbol{x},\, \varphi^*(\boldsymbol{y})\big) \Rightarrow 0 = (\mathbf{0},\, \boldsymbol{y}) = \big(\boldsymbol{x},\, \varphi^*(\boldsymbol{y})\big) \Rightarrow \varphi^*(\boldsymbol{y}) = \mathbf{0} \quad \forall \boldsymbol{y}$$

3) $\varphi(\boldsymbol{x}) = [\boldsymbol{a},\, \boldsymbol{x}], \quad \boldsymbol{a} \neq \mathbf{0}$.

$$\big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = \big([\boldsymbol{a},\, \boldsymbol{x}],\, \boldsymbol{y}\big) = (\boldsymbol{x},\, \boldsymbol{y},\, \boldsymbol{a}) = \big(\boldsymbol{x},\, [\boldsymbol{y},\, \boldsymbol{a}]\big).$$

Тогда

$$\big(\boldsymbol{x},\, [\boldsymbol{y},\, \boldsymbol{a}] - \varphi^*(\boldsymbol{y})\big) = 0 \Rightarrow \varphi^*(\boldsymbol{y}) = [\boldsymbol{y}{,}\boldsymbol{a}] = -[\boldsymbol{a},\, \boldsymbol{y}] = -\varphi(\boldsymbol{y}) \Rightarrow \varphi^* = -\varphi.$$

Нам повезло — во всех перечисленных примерах у преобразования φ было сопряжённое преобразование φ^*, причём оно было единственно. Но можно ли найти преобразование, у которого нет сопряжённого? А если сопряжённое преобразование нашлось, точно ли оно единственно?

Теорема. Для любого линейного преобразования φ в евклидовом пространстве $\mathbb{E}$ существует единственное сопряжённое преобразование φ^*.

Доказательство:

Сначала покажем вид матрицы преобразования φ^* , если это преобразование существует. Пусть A — матрица φ, A^* — матрица φ^*, ξ — координатный столбец $\boldsymbol{x}$, η —координатный столбец $\boldsymbol{y}$, Γ — матрица Грама. Распишем выражение $\big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = \big(\boldsymbol{x},\, \varphi^*(\boldsymbol{y})\big)$ (1):

$$(A\xi)^T\Gamma\eta = \xi^T\Gamma(A^*\eta) \Rightarrow \xi^T A^T\Gamma\eta = \xi^T\Gamma A^*\eta \quad \forall \xi,\, \eta \Rightarrow A^T\Gamma = \Gamma A^* \Rightarrow \boxed{A^* = \Gamma^{-1}A^T\Gamma}.$$

В ОНБ имеем $\Gamma = E \Rightarrow \boxed{A^* = A^T}$.

Теперь покажем, что преобразование φ^* всегда существует. Выберем ОНБ в $\mathbb{E}$, в качестве φ^* возьмём преобразование, заданное матрицей A^T, где A — матрица φ в этом базисе. Тогда, подставляя в (1), равенство выполняется $\Rightarrow$ φ^* сопряжено φ. То есть одно сопряжённое преобразование найдено.

Почему оно единственно? Если было бы другое преобразование, оно имело бы в этом ОНБ ту же самую матрицу. А если два преобразования в одном и том же базисе имеют одинаковые матрицы, эти преобразования совпадают. ■

Теорема. Множество значений линейного преобразования евклидового пространства $\mathbb{E}$ совпадает с ортогональным дополнением ядра сопряжённого преобразования.

$$\text{Или "по-человечески":} \quad \mathrm{Im}(\varphi) = (\mathrm{Ker}(\varphi^*))^{\perp}.$$

Доказательство:

Пусть $\boldsymbol{y} \in \big(\mathrm{Im}(\varphi)\big)^{\perp} \Leftrightarrow \forall \boldsymbol{x} \in \mathbb{E} \to \big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = 0 = \big(\boldsymbol{x},\, \varphi^*(\boldsymbol{y})\big) \Leftrightarrow \varphi^*(\boldsymbol{y}) = 0 \Leftrightarrow \boldsymbol{y} \in \mathrm{Ker}(\varphi^*)$.

А это и означает, что $\mathrm{Im}(\varphi) = \big(\mathrm{Ker}(\varphi^*)\big)^{\perp}$. По своей сути данная теорема — интерпретация теоремы Фредгольма.

Свойства сопряжённого преобразования.

1) $(\varphi^*)^* = \varphi$. $\big($В ОНБ, как отмечалось выше, $A^* = A^T$, кроме этого $(A^T)^T = A\big)$.
2) $(\varphi\psi)^* = \psi^*\varphi^*$. $\big($Так как $(AB)^T = B^T A^T\big)$.
3) Характеристические многочлены φ и φ^* совпадают. $\big($Так как совпадают ХМ для A и $A^T\big)$.
4) Все собственные значения и их кратности совпадают для φ и φ^*. $\big($Так как $(A - \lambda E)^T =$ $= (A^T - \lambda E)$, и детерминант не изменяется при транспонировании$\big)$.

Билет № 17. Самосопряжённые преобразования. Свойства их собственных векторов и собственных значений. Существование ортонормированного базиса из собственных векторов самосопряжённого преобразования.

Определение [самосопряжённое преобразование]. Преобразование φ^* называется самосопряжённым, если $\varphi^* = \varphi$.

Хотелось бы найти какие-то критерии самосопряжённости.

Теорема. φ — самосопряжённое преобразование $\Leftrightarrow$ матрица преобразования в ОНБ является симметрической.

Доказательство:

В ОНБ имеем $A^* = A^T$, а так как $A^* = A$, то $A = A^T$.

■

Когда мы работали с линейными преобразованиями, у нас возникали ситуации, когда вещественных собственных значений не было вовсе (например, поворот в плоскости). Что же будет с самосопряжёнными преобразованиями?

Теорема. Все собственные значения самосопряжённого преобразования вещественные.

Доказательство:

Докажем от противного. Пусть A — матрица преобразования в ОНБ (сразу смотрим на теорему выше — $A = A^* = A^T$). Предположим, что у A существует комплексное собственное значение λ. Тогда $\overline{\lambda}$ также является собственным значением (так как коэффициенты характеристического многочлена вещественные).

Пусть собственному значению λ соответствует комплексный собственный вектор $\boldsymbol{x}$, то есть:

$$(A - \lambda E)\boldsymbol{x} = \boldsymbol{0}.$$

Покажем, что $\overline{\lambda}$ соответствует $\overline{\boldsymbol{x}}$. Имеем:

$$(A - \lambda E)\boldsymbol{x} = \boldsymbol{0} \Rightarrow \overline{(A - \lambda E)\boldsymbol{x}} = \overline{\boldsymbol{0}} \Rightarrow (\overline{A} - \overline{\lambda} \cdot \overline{E})\overline{\boldsymbol{x}} = \overline{\boldsymbol{0}} \Rightarrow (A - \overline{\lambda}E)\overline{\boldsymbol{x}} = \boldsymbol{0}.$$

Значит $\overline{\boldsymbol{x}}$ — собственный вектор матрицы $A^T = A$, соответствующий $\overline{\lambda}$.

Рассмотрим теперь $\overline{\boldsymbol{x}}^T A\boldsymbol{x}$ (матрица размером 1×1):

$$\overline{\boldsymbol{x}}^T A\boldsymbol{x} = \lambda\overline{\boldsymbol{x}}^T\boldsymbol{x}.$$

Транспонируем $\overline{\boldsymbol{x}}^T A\boldsymbol{x}$ (но так как это скаляр, то транспонирование его не поменяет):

$$(\overline{\boldsymbol{x}}^T A\boldsymbol{x})^T = (\lambda\overline{\boldsymbol{x}}^T\boldsymbol{x})^T \Rightarrow \boldsymbol{x}^T A\overline{\boldsymbol{x}} = \lambda\boldsymbol{x}^T\overline{\boldsymbol{x}} \Rightarrow \overline{\lambda}\boldsymbol{x}^T\overline{\boldsymbol{x}} = \lambda\boldsymbol{x}^T\overline{\boldsymbol{x}} \Rightarrow (\overline{\lambda} - \lambda)\boldsymbol{x}^T\overline{\boldsymbol{x}} = 0.$$

Заметим, что $\boldsymbol{x}^T\overline{\boldsymbol{x}} = |\boldsymbol{x}|^2 \neq 0$. Значит $\overline{\lambda} = \lambda$. То есть λ — действительное собственное значение.

■

Теорема. Собственные векторы самосопряжённого преобразования, соответствующие различным собственным значениям, ортогональны.

Доказательство:

Пусть $\boldsymbol{x}$ — собственный вектор, соответствующий собственному значению λ, и $\boldsymbol{y}$ — собственный вектор, соответствующий собственному значению μ, причём $\lambda \neq \mu$.

$$\big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = \big(\boldsymbol{x},\, \varphi(\boldsymbol{y})\big) \Rightarrow \lambda(\boldsymbol{x},\, \boldsymbol{y}) = \mu(\boldsymbol{x},\, \boldsymbol{y}) \Rightarrow (\lambda - \mu)(\boldsymbol{x},\, \boldsymbol{y}) = 0 \Rightarrow (\boldsymbol{x},\, \boldsymbol{y}) = 0.$$

■

Лемма. Пусть $\mathbb{E}'$ инвариантно относительно самосопряжённого преобразования φ. Тогда $(\mathbb{E}')^{\perp}$ инвариантно относительно φ.

Доказательство:

Пусть $\boldsymbol{y} \in (\mathbb{E}')^{\perp}$. Также $\boldsymbol{x} \in \mathbb{E}' \Rightarrow \varphi(\boldsymbol{x}) \in \mathbb{E}'$, так как $\mathbb{E}'$ инвариантно относительно φ.

$$0 = \big(\varphi(\boldsymbol{x}),\, \boldsymbol{y}\big) = \big(\boldsymbol{x},\, \varphi(\boldsymbol{y})\big) \Rightarrow \varphi(\boldsymbol{y}) \in (\mathbb{E}')^{\perp}.$$

А это и значит, что $(\mathbb{E}')^{\perp}$ инвариантно относительно φ.

■

Теорема. Если существует ОНБ из собственных векторов преобразования φ, то φ — самосопряжённое.

Доказательство:

В ОНБ из собственных векторов матрица преобразования имеет диагональный вид. А значит, матрица преобразования в ОНБ симметрична. Тогда φ — самосопряжённое преобразование.

■

Теорема. Если φ — самосопряжённое преобразование, то существует ОНБ из собственных векторов φ.

Доказательство:

Пусть L — сумма всех собственных подпространств (т. е. инвариантное подпространство). Покажем, что L — инвариантное подпространство. Пусть $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ — собственные векторы, соответствующие разным СЗ, тогда:

$$\forall \boldsymbol{x} \in L \to \exists \alpha_1, \ldots, \alpha_k : \quad \boldsymbol{x} = \alpha_1 \boldsymbol{x}_1 + \ldots + \alpha_k \boldsymbol{x}_k.$$

Посмотрим также на преобазование φ:

$$\varphi(\boldsymbol{x}) = \alpha_1 \varphi(\boldsymbol{x}_1) + \ldots + \alpha_k \varphi(\boldsymbol{x}_k) = \alpha_1 \lambda_1 \boldsymbol{x}_1 + \ldots + \alpha_k \lambda_k \boldsymbol{x}_k.$$

Значит, L — инвариантное подпространство. А мы леммой выше уже показали, что если есть инвариантное подпространство относительно самосопряжённого преобразования, то верно, что ортогональное дополнение к этому подпространству будет также инвариантным подпространством. Итак, $L^\perp$ — инвариантное подпространство.

Также уже знаем, что $\mathbb{E} = L \oplus L^\perp$. Покажем, что $L^\perp$ — нулевое пространство. Пусть это не так, то есть $\exists \boldsymbol{x} \neq \boldsymbol{0}$, $\boldsymbol{x} \in L^\perp$. Рассмотрим ограничение преобразования на $L^\perp$. Если преобразование самосопряжённое, то и ограничение этого преобразования тоже самосопряжённое. А значит существует вещественное собственное значение, значит есть хотя бы один собственный вектор $\boldsymbol{x}$. Тогда $\boldsymbol{x} \in L \Rightarrow \boldsymbol{x} = \boldsymbol{0}$ — противоречие. Значит, $\mathbb{E} = L$.

В каждом собственном подпространстве берём ОНБ. Так как векторы из различных собственных подпространств будут попарно ортогональны, то объединение базисов — ОНБ в $\mathbb{E}$.

Причём, если λ — собственное значение самосопряжённого преобразованиями кратности k, ему соответствуют ровно k ЛНЗ собственных векторов. Это следует из того, что базис пространства — объединение бизисных собственных подпространств. Характеристический многочлен имеет $n = \dim(\mathbb{E})$ собственных значений с учётом кратности, и для линейного преобразования верно, что собственному значению кратности k соответствуют не более k собственных векторов.

■

Подытожим все полученные знания про самосопряжённые и линейные преобразования, как это в одной из последних лекций сделала Подлипская Ольга Геннадьевна.

Самосопряжённое преобразование	**Линейное преобразование**
Все СЗ вещественные.	СЗ могут быть как вещественные, так и комплексные.
Векторы, соответствующие различным СЗ, попарно ортогональны.	Векторы, соответствующие различным СЗ, линейно независимы.
Если есть корень характеристического многочлена кратности k, ему соответствуют ровно k СВ.	Если есть корень характеристического многочлена кратности k, ему соответствуют не более чем k СВ.
Существует ОНБ из СВ.	Может быть базис из СВ.

Таблица 1: Различия между линейными и самосопряжёнными преобразованиями.

Билет № 18. Свойства ортогональных преобразований. Свойства корней характеристического многочлена и собственных векторов ортогональных преобразований.

Определение [ортогональное преобразование]. Ортогональное преобразование евклидова пространства $\mathbb{E}$ — это такое преобразование φ, что $\forall\ \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{E} \to (\boldsymbol{x},\ \boldsymbol{y}) = \big(\varphi(\boldsymbol{x}), \varphi(\boldsymbol{y})\big)$.

Другими словами, ортогональное преобразование — это преобразование, сохраняющее скалярное произведение.

Теорема. Преобразование φ является ортогональным $\Leftrightarrow$ сопряжённое к нему является обратным к нему.

Доказательство [Необходимость $\Rightarrow$]:

$$\forall \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{E} \to (\boldsymbol{x},\ \boldsymbol{y}) = \big(\varphi(\boldsymbol{x}),\ \varphi(\boldsymbol{y})\big) = \Big(\boldsymbol{x},\ \varphi^*\big(\varphi(\boldsymbol{y})\big)\Big).$$

Перенесём всё в левую часть:

$$\Big(\boldsymbol{x},\ \boldsymbol{y} - \varphi^*\big(\varphi(\boldsymbol{y})\big)\Big) = 0 \Rightarrow \boldsymbol{y} = \varphi^*\big(\varphi(\boldsymbol{y})\big) = (\varphi^*\varphi)\boldsymbol{y}.$$

Тогда мы получаем, что преобразование $\varphi^*\varphi$ — тождественное. То есть

$$\varphi^*\varphi = \mathrm{id} \Rightarrow \varphi^{-1} = \varphi^*.$$

Доказательство [Достаточность $\Leftarrow$]:

По условию $\varphi^*\varphi = \mathrm{id}$. Рассмотрим $\big(\varphi(\boldsymbol{x}),\ \varphi(\boldsymbol{y})\big)$:

$$\big(\varphi(\boldsymbol{x}),\ \varphi(\boldsymbol{y})\big) = \Big(\boldsymbol{x},\ \varphi^*\big(\varphi(\boldsymbol{y})\big)\Big) = (\boldsymbol{x},\ \boldsymbol{y}) \quad \forall \boldsymbol{x},\ \boldsymbol{y} \in \mathbb{E} \Rightarrow \varphi \text{ — ортогональное.}$$

■

Лемма. Преобразование φ является ортогональным $\Leftrightarrow$ матрица преобразования в ОНБ является ортогональной.

Доказательство:

В ОНБ имеем $A^* = A^T$, и преобразование φ — ортогональное. Итого, $A^* = A^{-1} \Leftrightarrow A^{-1} = A^T$.

■

Лемма. Для любых ОНБ $\boldsymbol{e}$ и $\boldsymbol{f}$ существует единственное ортогональное преобразование, такое, что $\boldsymbol{f}_i = \varphi(\boldsymbol{e}_i)\ \forall i$.

Доказательство:

Рассмотрим преобразование, которое в базисе $\boldsymbol{e}$ имеет матрицу, столбцами которой являются координаты векторов из базиса $\boldsymbol{f}$. Это и есть матрица преобразования: в столбцах стоят координаты образов базисных векторов, то есть это преобразование такое, что $\boldsymbol{f}_i = \varphi(\boldsymbol{e}_i)\ \forall i$.

Матрица этого преобразования — это ещё и матрица перехода от $\boldsymbol{e}$ к $\boldsymbol{f}$. Тогда матрица перехода обязательно ортогональна.

Итак, матрица преобразования в базисе $\boldsymbol{e}$ ортогональна, причём $\boldsymbol{e}$ — ОНБ. Значит, из доказанной леммы следует, что φ — ортогональное преобразование.

■

Теорема. Все собственные значения ортогонального преобразования по модулю равны 1.

Доказательство:

1) *Если λ — вещественное.* Этому λ соответствует хотя бы один собственный вектор $\boldsymbol{x}$. Тогда:

$$(\boldsymbol{x},\, \boldsymbol{x}) = \big(\varphi(\boldsymbol{x}),\, \varphi(\boldsymbol{x})\big) = (\lambda\boldsymbol{x},\, \lambda\boldsymbol{x}) = \lambda^2(\boldsymbol{x},\, \boldsymbol{x}) \Rightarrow (1-\lambda^2)(\boldsymbol{x},\, \boldsymbol{x}) = 0 \Rightarrow \lambda^2 = 1 \Rightarrow |\lambda| = 1.$$

2) *Если λ — комплексное.* Значит, $\overline{\lambda}$ — тоже собственное значение. Этой паре собственных значений соответствует двумерное инвариантное подпространство $\mathbb{E}' = \operatorname{Ker}(\varphi^2 - p\cdot\varphi - q\cdot \mathrm{id})$, где $p = \lambda + \overline{\lambda}$ и $q = \lambda\overline{\lambda}$.

Рассмотрим ограничение φ на $\mathbb{E}'$. Это ограничение, очевидно, сохранило ортогональность. А значит, матрица этого преобразования в ОНБ является ортогональной.

У этой матрицы размер 2×2. Ортогональные матрицы 2×2 имеют один из следующих видов:

$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \quad\text{или}\quad \begin{pmatrix} -\cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Если определитель матрицы A' ограничения φ на $\mathbb{E}'$ равен -1, то матрица преобразования является симметрической, а тогда ограничение является самосопряжённым $\Rightarrow$ есть собственные векторы, определяемые комплексными собственными значениями, а их быть не может. Значит, этот случай не рассматриваем. Значит, $\det(A') = 1$.

Инвариантное подпространство $\mathbb{E}'$ является линейной оболочкой из $\operatorname{Ker}(\varphi^2 - p\varphi - q\cdot\mathrm{id})$ и его образа. То есть $\mathbb{E}' = \langle \boldsymbol{e},\, \varphi(\boldsymbol{e})\rangle$, где $\boldsymbol{e} \in \operatorname{Ker}(\varphi^2 - p\cdot\varphi - q\cdot\mathrm{id})$.

Рассмотрим матрицу ограничения φ на $\mathbb{E}'$ в базисе $\{\boldsymbol{e},\, \varphi(\boldsymbol{e})\}$. Образ первого базисного вектора $\varphi(\boldsymbol{e})$ в этом базисе он будет иметь координаты $(0,\ 1)$. Найдём образ второго базисного вектора, то есть образ вектора $\varphi(\boldsymbol{e})$:

$$\varphi\big(\varphi(\boldsymbol{e})\big) = p\varphi(\boldsymbol{e}) - q\boldsymbol{e}.$$

Итак, матрица преобразования будет иметь вид:

$$\begin{pmatrix} 0 & -q \\ 1 & p \end{pmatrix}.$$

Определитель матрицы преобразования не изменится при смене базиса. А ранее мы выяснили, что $\det(A') = 1$:

$$\begin{vmatrix} 0 & -q \\ 1 & p \end{vmatrix} = q = 1 = \lambda\overline{\lambda} \Rightarrow |\lambda| = 1.$$

■

Лемма. Собственные векторы ортогонального преобразования, отвечающие различным собственным значениям, ортогональны.

Доказательство:

Пусть собственные векторы $\boldsymbol{x}$, $\boldsymbol{y}$ отвечают различным собственным значениям λ и μ соответственно.

$$(\boldsymbol{x},\, \boldsymbol{y}) = \big(\varphi(\boldsymbol{x}),\, \varphi(\boldsymbol{y})\big) = (\lambda\boldsymbol{x},\, \mu\boldsymbol{y}) = \lambda\mu(\boldsymbol{x},\, \boldsymbol{y}) \Rightarrow (1 - \lambda\mu)(\boldsymbol{x},\, \boldsymbol{y}) = 0 \Rightarrow (\boldsymbol{x},\, \boldsymbol{y}) = 0.$$

■

Лемма. Если $\mathbb{E}'$ инвариантно относительно ортогонального преобразования φ, то $(\mathbb{E}')^{\perp}$ также инвариантно относительно φ.

Доказательство:

Пусть $\boldsymbol{x} \in \mathbb{E}'$. Так как $\mathbb{E}'$ инвариантно $\to \varphi(\mathbb{E}') \subseteq \mathbb{E}'$. Также φ ортогонально, значит размерность $\mathbb{E}'$ совпадает с размерностью образа: $\dim(\mathbb{E}') = \dim\big(\varphi(\mathbb{E}')\big)$. А такое может быть только если $\varphi(\mathbb{E}') = \mathbb{E}'$.

Пусть $\boldsymbol{y} \in (\mathbb{E}')^{\perp}$. Понятно, что $(\boldsymbol{x},\, \boldsymbol{y}) = \big(\varphi(\boldsymbol{x}),\, \varphi(\boldsymbol{y})\big) = 0$. Значит, $\varphi(\boldsymbol{y}) \in \big(\varphi(\mathbb{E}')\big)^{\perp} = (\mathbb{E}')^{\perp}$. Это выполнено $\forall \boldsymbol{y} \Rightarrow (\mathbb{E}')^{\perp}$ — инвариантное подпространство.

■

Билет № 19. Сингулярное разложение. Полярное разложение линейных преобразований евклидова пространства и матриц.

Сингулярное разложение.

Теперь мы будем предполагать, что преобразование φ невырожденно, то есть $\det(A_\varphi) \neq 0$, и выведем теорему, обобщающую теорему первого семестра о разложении аффинного преобразования в произведении ортогонального преобразования и двух сжатий к двум взаимно перпендикулярным прямым.

Теорема. Любое невырожденное линейное преобразование $\varphi : \mathbb{E} \to \mathbb{E}$ евклидова пространства $\mathbb{E}$ может быть представлено в виде произведения $\varphi = \kappa\gamma$, где κ — ортогональное преобразование, а γ — самосопряжённое преобразование с положительными собственными значениями. Такое разложение единственно.

Будем доказывать в основном матричную формулировку теоремы:

Для любой невырожденной вещественной матрицы A порядка n существуют единственные матрицы Q, C, такие что ($Q^T = Q^{-1}$, $C^T = C$, $\lambda_i > 0$):

$$A = QC.$$

Доказательство:

1) Покажем, прежде всего, что все собственные значения самосопряжённого преобразования $\varphi^*\varphi$ положительны. Пусть $\boldsymbol{f}$ — собственный вектор для $\varphi^*\varphi$.

$$(\varphi^*\varphi)(\boldsymbol{f}) = \lambda\boldsymbol{f} \Rightarrow ((\varphi^*\varphi)(\boldsymbol{f}),\, \boldsymbol{f}) = \lambda(\boldsymbol{f},\, \boldsymbol{f}) = (\varphi(\boldsymbol{f}),\, \varphi(\boldsymbol{f})) > 0.$$

Значит, $\lambda > 0$, поскольку $\varphi(\boldsymbol{f}) \neq 0$, так как φ невырожденно.

Теперь будем искать разложение матрицы $A = QC$. Транспонируем и перемножим:

$$A = QC \Rightarrow A^T = C^T Q^T = CQ^{-1} \Rightarrow A^T A = C^T (Q^{-1}Q) C = C^T C = C^2.$$

Матрица C^2 симметричная с положительными собственными значениями. Следовательно, для неё существует ортогональная матрица S:

$$S^{-1}C^2S = D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}, \text{ где } \lambda_i > 0,\, i = \overline{1,\, n} \Rightarrow C^2 = SDS^{-1} = SDS^T.$$

Рассмотрим матрицу $\Sigma = \sqrt{D} = \begin{pmatrix} \mu_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mu_n \end{pmatrix}$, где $\mu_i = \sqrt{\lambda_i} > 0$, $i = \overline{1,\, n}$. Тогда получим

$$C = S\Sigma S^{-1} = S\Sigma S^T.$$

Осталось найти ортогональную матрицу Q. Положим её просто равной $Q = AC^{-1}$. Проверим, что матрица Q ортогональна:

$$Q = AC^{-1}, \quad Q^T = (C^{-1})^T A^T = C^{-1}A^T \Rightarrow Q^T Q = (C^{-1}A^T)(AC^{-1}) = C^{-1}C^2C^{-1} = E.$$

2) Покажем единственность.

Допустим, $A = QC = \widetilde{Q}\widetilde{C}$, тогда $A^TA = C^2 = \widetilde{C}^2$.

По построению C, $\widetilde{C}$ имеют общий собственный базис и потому коммутируют, поэтому

$$C^2 - \widetilde{C}^2 = (C - \widetilde{C})(C + \widetilde{C}) = 0.$$

Так как C, $\widetilde{C}$ имеют положительные собственные значения в общем базисе, то у $C + \widetilde{C}$ положительные собственные значения, значит, $\exists \left(C + \widetilde{C}\right)^{-1}$ и потому $C = \widetilde{C} \Rightarrow Q = \widetilde{Q}$.

■

Запишем разложение $A = QC$ с учетом построения C:

$$A = QS\Sigma S^T = (QS)\Sigma S^T = U\Sigma V, \tag{1}$$

где U, V — ортогональные матрицы, Σ — диагональная матрица с положительными диагональными элементами.

Определение. Разложение (1) называется сингулярным разложением.

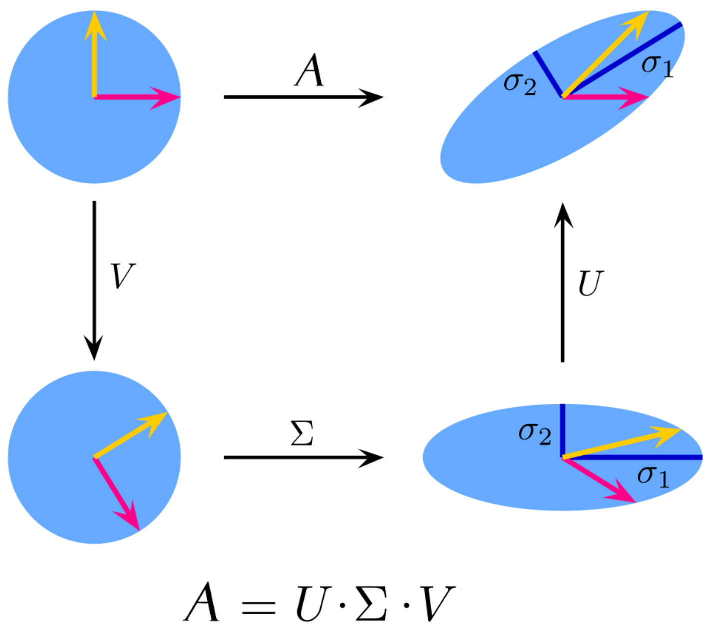

Рис. 4: Геометрический смысл сингулярного разложения в двумерном случае.
Источник: Wikipedia.

Пусть матрице A поставлен в соответствие линейный оператор. Сингулярное разложение можно переформулировать в геометрических терминах. Линейный оператор, отображающий элементы пространства $\mathbb{R}^n$ в себя, представим в виде последовательно выполняемых линейных операторов вращения и растяжения. Поэтому компоненты сингулярного разложения наглядно показывают геометрические изменения при отображении линейным оператором A множества векторов из векторного пространства в себя или в векторное пространство другой размерности.

Остаётся только один вопрос — где это может быть нужно? В первую очередь, SVD (singular value decomposition) используется в машинном обучении и обработке данных. Мы можем снижать размерности матриц и работать с ними по отдельности, что сильно ускоряет процесс. Если конкретнее — SVD используется в алгоритмах машинного обучения для решения задач классификации, регрессии и кластеризации; SVD используется для анализа текстовых данных, например, для тематического моделирования; SVD применяется для сжатия изображений, удаления шумов и реставрации повреждённых изображений...

Большое значение сингулярное разложение находит и в физике (решения уравнений Навье-Стокса, описывающих движение вязких жидкостей; разложения сигналов на компоненты), и во многих других областях.

Полярное разложение.

Далее мы будем предполагать, что матрица $A^{m\times n}$ квадратная ($m = n$). Вернемся к формуле (1) и перепишем её, вставив два сомножителя, произведение которых равно единичной матрице:

$$A = (UV^T)(V\Sigma V^T).$$

При такой расстановке скобок мы разложили A в произведение ортогональной матрицы $B = UV^T$ и симметричной матрицы $S = V\Sigma V^T$. Корни характеристического многочлена матрицы S неотрицательны — это сингулярные числа матрицы A. Мы показали, таким образом, что произвольную квадратную матрицу A можно разложить в произведение

$$A = BS. \tag{2}$$

ортогональной и симметричной матриц, причем характеристические числа симметричной матрицы неотрицательны.

Конечно, можно поступить несколько иначе — написать $A = (U\Sigma U^T)(UV^T)$. Это приводит нас к разложению $A = \tilde{S}B$, где ортогональный сомножитель расположен справа. Отметим, что в обоих случаях ортогональная матрица одна и та же: $B = UV^T$, симметричные матрицы различны, но их характеристические числа совпадают.

Определение. Разложение (2) квадратной матрицы в произведение ортогональной и симметричной матрицы с неотрицательными характеристическими числами (в каком бы то ни было порядке) называется полярным разложением.

Билет № 20. Построение ортонормированного базиса, в котором квадратичная форма имеет диагональный вид. Одновременное приведение к диагональному виду пары квадратичных форм, одна из которых положительно определена.

Билинейные формы в евклидовых пространствах.

Определение. φ — линейное преобразование, присоединённое к билинейное форме $b(\boldsymbol{x},\, \boldsymbol{y})$, если $\forall \boldsymbol{x},\, \boldsymbol{y} \in \mathbb{E} \to b(\boldsymbol{x},\, \boldsymbol{y}) = \big(\boldsymbol{x},\, \varphi(\boldsymbol{y})\big)$.

Теорема. У любой билинейной формы в евклидовом пространстве существует единственное присоединённое преобразование.

Доказательство:

Пусть B — матрица билинейной формы в некотором базисе, A — матрица присоединённого преобразования в этом же базисе, Γ — матрица Грама этого базиса. Пусть $\boldsymbol{\xi}$ — координатный столбец $\boldsymbol{x}$, $\boldsymbol{\eta}$ — координатный столбец $\boldsymbol{y}$. Тогда

$$b(\boldsymbol{x},\, \boldsymbol{y}) = \boldsymbol{\xi}^T B \boldsymbol{\eta} = \boldsymbol{\xi}^T \Gamma A \boldsymbol{\eta} \Rightarrow B = \Gamma A \Rightarrow \boxed{A = \Gamma^{-1} B}.$$

Если бы существовало другое присоединённое преобразование, то его матрица в том же базисе была бы такой же. То есть преобразования совпали бы. Значит, если присоединённое преобразование существует, то оно единственно.

Рассмотрим преобразование с матрицей $\Gamma^{-1}B$:

$$b(\boldsymbol{x},\, \boldsymbol{y}) = \boldsymbol{\xi}^T B \boldsymbol{\eta}, \quad \big(\boldsymbol{x},\, \varphi(\boldsymbol{y})\big) = \boldsymbol{\xi}^T \Gamma \Gamma^{-1} B \boldsymbol{\eta} = \boldsymbol{\xi}^T B \boldsymbol{\eta}.$$

Получили, что $b(\boldsymbol{x},\, \boldsymbol{y}) = \big(\boldsymbol{x},\, \varphi(\boldsymbol{y})\big)$. Значит, по определению, φ — присоединённое преобразование.

■

Лемма. Преобразование, присоединённое к симметрической билинейной форме, является самосопряжённым.

Доказательство:

В ОНБ матрица присоединённого преобразования совпадает с матрицей симметричной билинейной формы, то есть матрица этого преобразования является симмерической. А так как матрица в ОНБ симметрическая, то преобразование является самосопряжённым.

■

Определение. Преобразование, присоединённое к симметричной билинейной форме, называется преобразованием, присоединённым к соответствующей квадратичной форме.

Теорема. Для любой квадратичной формы существует ОНБ, в котором эта квадратичная форма имеет имеет диагональный вид.

Доказательство:

В ОНБ матрица квадратичной формы и матрица присоединённого преобразования совпадают. Преобразование присоединено к квадратичной форме, то есть к симметрической билинейной форме $\Rightarrow$ преобразование самосопряжённое. А тогда существует базис из собственных векторов, и в этом базисе матрица присоединённого преобразования имеет диагональный вид.

Таким образом, искомый базис — это ОНБ из собственных векторов присоединённого преобразования.

■

Теорема. Пусть L — конечномерное вещественное линейное пространство, а f и h — две квадратичные формы, заданные в некотором линейном пространстве. Если одна из них является положительно определённой, то существует базис, в котором обе квадратичные формы имеют диагональный вид.

Доказательство:

1) Пусть f положительно определена, тогда её матрица F (в любом базисе) симметрична и положительно определена $\Rightarrow$ можем определить на L скалярное произведение, связанное с f:

$$(\boldsymbol{x},\, \boldsymbol{y})_f = \boldsymbol{x}^T F \boldsymbol{y} \quad \forall \boldsymbol{x},\, \boldsymbol{y} \in L.$$

Отметим корректность этого скалярного произведения, так как F симметрична и $(\boldsymbol{x},\, \boldsymbol{x})_f > 0$ для всех $\boldsymbol{x} \neq \boldsymbol{0}$. Таким образом, пространство L становится евклидовым относительно $(\boldsymbol{x},\, \boldsymbol{y})_f$.

2) В евклидовом пространстве $(L,\, (\boldsymbol{x},\, \boldsymbol{y})_f)$ существует ортонормированный базис $\{\boldsymbol{e}_1,\, \ldots,\, \boldsymbol{e}_n\}$, то есть:

$$(\boldsymbol{e}_i,\, \boldsymbol{e}_j)_f = \delta_{ij} = \begin{cases} 0,\; i \neq j \\ 1,\; i = j \end{cases},\ \text{где } (i,\, j = \overline{1, n}).$$

В этом базисе матрица квадратичной формы f становится единичной, так как:

$$f(\boldsymbol{x}) = (\boldsymbol{x},\, \boldsymbol{x})_f = \sum_{i=1}^{n} x_i^2,$$

где $\boldsymbol{x} = (x_1,\, \ldots,\, x_n)^T$ — координаты вектора в базисе $\boldsymbol{e}$.

3) Матрица квадратичной формы h имеет диагональный вид в некотором базисе, ортонормированном относительно вспомогательного скалярного произведения. Причём матрица f в этом базисе — единичная.

■

Алгоритм приведения пары КФ к диагональному виду.

Пусть f — положительно определённая квадратичная форма, h — квадратичная форма.

Способ № 1:

1) Приводим к каноническому виду положительно определённую квадратичную форму f. Тем самым совершён переход к ОНБ $\boldsymbol{e}$ относительно вспомогательного скалярного произведения. Находим в этом базисе матрицу квадратичной формы h (пусть это будет матрица H').

2) В ОНБ $\boldsymbol{e}$ матрица H' — матрица присоединённого преобразования. Это присоединённое преобразования является самосопряжённым $\Rightarrow$ найдётся ОНБ из собственных векторов. Далее находим собственные векторы преобразования, присоединённого к h, а затем ортонормируем собственные векторы, вычисляя скалярные произведения по формуле $(\boldsymbol{x}, \boldsymbol{y}) = x_1y_1 + \ldots + x_ny_n$. Получили базис $\boldsymbol{e}'$.

В $\boldsymbol{e}'$ матрица преобразования, присоединённого к h — диагональная $\Rightarrow$ матрица квадратичной формы h диагональная, а матрица квадратичной формы f — единичная.

Способ № 2:

1) Найдём собственные значения и собственные векторы преобразования, присоединённого к h. Матрица присоединённого преобразования (относительно вспомогательного скалярного произведения) $F^{-1}H$. То есть необходимо решить:

$$\det(F^{-1}H - \lambda E) = \det(F^{-1}H - \lambda F^{-1}F) = 0.$$

Тогда

$$\det(F^{-1})\det(H - \lambda F) = 0 \Rightarrow \boxed{\det(H - \lambda F) = 0} \text{ — обобщ. хар. уравнение.}$$

В ортонормированном базисе из собственных векторов матрица f — единичная, матрица h имеет диагональный вид, а на диагонали стоят корни обобщённого характеристического уравнения.

И в целом, если сам базис не нужен, то алгоритм можно завершить, получив нужные диагональные виды матриц.

2) Чтобы найти базис, нужно решить систему

$$(F^{-1}H - \lambda E)\boldsymbol{x} = \boldsymbol{0}.$$

Эта система эквивалентна

$$(H - \lambda F)\boldsymbol{x} = \boldsymbol{0}.$$

После этого нужно нормировать собственные векторы, считая скалярные произведения по формуле $(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^T F \boldsymbol{y}$.