Магистратура ВШЭ. Лекции

Чепуркин К.М.

0.1. Векторные пространства

Базовым объектом линейной алгебры является векторное пространство.

Определение (Векторное пространство)**.** Векторным пространством над полем K называется множество V вместе с отображениями $+\colon V \times V \to V$ и $\cdot\colon K \times V \to V$, удовлетворяющее свойствам:
1) $\forall u, v, w \in V\ u + (v + w) = (u + v) + w$
2) $\forall u, v \in V\ u + v = v + u$
3) Существует нулевой элемент $0 \in V$, такой что $\forall v \in V\ 0 + v = v + 0 = v$. Такой элемент обязан быть единственным в силу предыдущих свойств.
4) $\forall u \in V\ \exists(-u) \in V$, такой что $u+(-u) = 0 = (-u)+u$. Такой элемент единственен и называется противоположенным к элементу u.
5) $\forall v \in V$ верно, что $1 \cdot v = v$
6) $\forall v \in V$, $\forall \lambda, \mu \in K$ верно, что $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$.
7) $\forall u, v \in V$, $\forall \lambda \in K$ верно, что $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$.
8) $\forall v \in V\ \forall \lambda, \mu \in K$ верно, что $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$.

Примеры:
0) Само поле K вместе со сложением и умножением.
1) Пространство столбцов K^n. Умножение и сложение покомпонентное.
2) Обобщая. Пространство матриц $M_{m\times n}(K)$.
3) Пусть X – множество. Рассмотрим множество всех функций из K в X , то есть K^X. Это векторное пространство над полем K с поточечным сложением и умножением.
4) Рассмотрим множество непрерывных вещественнозначных функций на отрезке $[0, 1]$. Это векторное пространство над $\mathbb{R}$.
5) Рассмотрим множество последовательностей над полем K, удовлетворяющих заданному линейному рекуррентному соотношению $a_k x_{n+k} + \cdots + a_0 x_n = 0$. Это векторное пространство над K.
6) Рассмотрим множества всех многочленов $K[x_1, \ldots, x_n]$, всех рациональных функций $K(x_1, \ldots, x_n)$, рядов $K[[x_1, \ldots, x_n]]$. Это векторные пространства относительно естественного сложения и умножения на элементы K.
7) $\mathbb{C}$ является векторным пространством над $\mathbb{R}$

Однако обычно векторные пространства возникают в следующей ситуации.

Определение (Подпространство)**.** Пусть V – векторное пространство над полем K. Подмножество $U \subseteq V$ называется подпространством V, если
1) U – подгруппа V.
2) $\forall \lambda \in K$, $\forall u \in U$ верно, что $\lambda u \in U$.
По другому говоря, операции на V можно сузить на U, с тем, чтобы U стало векторным пространством относительно этих операций.

Примеры:
1) Множество решений уравнения $Ax = 0$, где матрица $A \in M_{m\times n}(K)$ образует подпространство в K^n. Это один из самых интересных для нас примеров.
2) Рассмотрим множество непрерывных на отрезке $[0, 1]$ функций, принимающих значение 0 в точках $0, \frac{1}{2}, 1$. Это подпространство в $C[0, 1]$.
3) Рассмотрим множество многочленов степени не выше n от одной переменной

$$K[x]_{\le n} = \{f \in K[x] \mid \deg f \le n\}.$$

Это подпространство в $K[x]$.
4) Рассмотрим множество правильных дробей $\frac{f}{g} \in K(x)$. Это подпространство в $K(x)$.

Замечание. Подпространства наследует с объемлющего пространства структуру и потому сами являются примерами векторных пространств.

Как и в жизни элементы пространства обычно не интересны сами по себе, а только во взаимоотношении с окружающими их сосед.

Определение. (Линейная комбинация) Линейной комбинацией векторов $v_1, \ldots, v_n$ с коэффициентами $\lambda_1, \ldots, \lambda_n$, называется выражение

$$\lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Будем говорить, что элемент $v \in V$ представим в виде линейной комбинации векторов $v_1, \ldots, v_n$, если

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Если хотя бы один из элементов $\lambda_1, \ldots, \lambda_n$ не равен 0, то говорят, что линейная комбинация нетривиальна.

Лемма 1. Рассмотрим набор $v_1, \dots, v_n \in V$. Пусть $w_1 = \mu_{11}v_1 + \dots + \mu_{1n}v_n, \dots, w_m = \mu_{m1}v_1 + \dots + \mu_{mn}v_n$. Рассмотрим набор $\lambda_1, \dots, \lambda_m$. Тогда вектор $w = \sum \lambda_i w_i$ является линейной комбинацией набора v_i.

Определение (Линейная зависимость)**.** Набор векторов $v_1, \dots, v_n$ называется линейно зависимым, если 0 является нетривиальной линейной комбинацией $v_1, \dots, v_n$, то есть существуют $\lambda_1, \dots, \lambda_n \in K$ не все равные 0, что

$$0 = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Определение (Линейная независимость)**.** Набор векторов $v_1, \dots, v_n$ называется линейно независимым, если он не является линейно зависимым, то есть если $\lambda_1, \dots, \lambda_n \in K$ такие, что

$$0 = \lambda_1 v_1 + \dots + \lambda_n v_n, \text{ то } \lambda_1 = \dots = \lambda_n = 0.$$

Примеры:
0) Набор из одного нуля линейно зависим.
1) Пусть v_1 и v_2 два вектора из V. Они линейно зависимы тогда и только тогда, когда они пропорциональны.
2) Рассмотрим пространство K^n и набор столбцов

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

Это линейно независимая система векторов.
3) Аналогично в пространстве матриц $M_{m\times n}(K)$ имеется набор матриц e_{ij} вида

$$e_{ij} = \begin{array}{cc} & \begin{array}{cccc} & j & & \end{array} \\ \begin{array}{c} \\ \\ i \\ \\ \end{array} & \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & 1 & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix} \end{array}$$

4) Все мономы x^α в кольце $K[x_1, \dots, x_n]$ линейно независимы
5) Набор $\frac{1}{x-\lambda}$ линейно независим.
6) Любой поднабор линейно независимого набора линейно независим.

Теорема 1 (О линейной зависимости линейных комбинаций.)**.** Пусть $u_1, \dots, u_m$ и $v_1, \dots, v_n$ два набора векторов. При этом все вектора u_i являются линейными комбинациями v_j, то есть $u_i = \sum_{j=1}^n \lambda_{ij} v_j$. Тогда, если $m > n$, то u_i обязательно линейно зависимы.

Доказательство. Индукция по n. $n = 1$. Все вектора u_i кратны v_1 и пропорциональны друг другу.

Будем поступать как в методе Гаусса. Запишем u_i в виде $u_i = \sum_{j=1}^n \lambda_{ij} v_j$. Если для некоторого индекса j все $\lambda_{ij} = 0$, то можно воспользоваться предположением индукции.

Рассмотрим вектор u_i, что $\lambda_{in} \neq 0$. Тогда перейдём к набору

$$u'_s = u_s - \frac{\lambda_{sn}}{\lambda_{in}} u_i = \sum_{j=1}^{n-1} \mu_{sj} v_j, \quad s \neq i.$$

Это набор из $m-1$ элемента, которые суть линейные комбинации $v_1, \dots, v_{n-1}$. Этот набор линейно зависим по индукционному предположению, то есть существуют ν_s не все равные нулю, что

$$0 = \sum \nu_s u'_s = -\left(\sum \nu_s \frac{\lambda_{sn}}{\lambda_{in}}\right) u_i + \sum \nu_s u_s.$$

Заметим, что не все коэффициенты при u_s равны нулю. Таким образом мы получили нетривиальную линейную зависимость. □

Определение. Пусть $v_1, \dots, v_k \in V$. Тогда линейная оболочка $\langle v_1, \dots, v_k \rangle$ называется множество всех линейных комбинаций $v_1, \dots, v_n$.

$$\langle X \rangle = \{v \in V \mid \lambda_1, \dots, \lambda_k \in K, \text{ что } v = \lambda_1 v_1 + \dots + \lambda_k v_k\}$$

Определение. Будем говорить, что набор $v_\alpha \in V$ $\alpha \in I$ является порождающим для V, если $\langle \{v_\alpha\}_{\alpha \in I} \rangle = V$. Иными словами, для любого $v \in V$ существуют $\lambda_1, \dots, \lambda_n$, что $v = \sum \lambda_i v_i$.

Определение (Базис)**.** Набор векторов v_α, $\alpha \in I$ называется базисом пространства V, если он является порождающей и линейно независимой системой векторов в V.

Примеры:
1) Набор векторов $e_i \in K^n$ является базисом. Этот базис называют стандартным.
2) Набор e_{ij} является базисом $M_{m\times n}(K)$.
3) Мономы x^α являются базисом $K[x_1,\ldots,x_n]$.
4) Элементы $1, x, \ldots, x^n, \ldots$ и $\frac{1}{(x-\lambda)^n}$ по всем $\lambda \in \mathbb{C}$ являются базисом в $\mathbb{C}(x)$.
5) $1, i$ – базис $\mathbb{C}$ над $\mathbb{R}$.

Лемма 2 (Переформулировки)**.** Пусть $e_1,\ldots,e_n$ – набор элементов пространства V. Тогда следующие свойства эквивалентны:
1) $e_1,\ldots,e_n$ – базис V.
2) $e_1,\ldots,e_n$ – минимальный по включению среди порождающих V наборов векторов.
3) Для любого $v \in V$ существуют единственные $\lambda_1,\ldots,\lambda_n \in K$, что $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$. Такие элементы $\lambda_1,\ldots,\lambda_n$ называются координатами вектора v в базисе e. 4) $e_1,\ldots,e_n$ – максимальный по включению набор среди линейно независимых наборов векторов из V.

Доказательство. 1) в 2). Пусть набор не минимален. Тогда в нём есть лишний вектор. Например, v_1. Выкинем его. Но тогда $v_1 = \sum_{i\ge 2}\lambda_i v_i$. Это даёт нетривиальную линейную зависимость.

2) в 3). Существование разложения следует из порождаемости. Пусть $\sum \lambda_i v_i = \sum \mu_i v_i$. Тогда либо все $\mu_i = \lambda_i$, либо, скажем, $\mu_1 - \lambda_1 \neq 0$. Тогда $v_1 = \frac{1}{\mu_1-\lambda_1}\sum_{i\ge 2}(\mu_i - \lambda_i)v_i$. Тогда из системы можно выкинуть v_1 и всё равно будет порождающая система.

3) в 4). Единственность разложения для 0 означает линейную независимость. Покажем максимальность. Пусть система не максимальна. Тогда к ней можно добавить вектор $v \neq 0$ и она останется независимой. Но такого не может быть, так как $v = \sum \lambda_i v_i$, что очевидно даёт линейную зависимость.

4) в 1). Пусть система не порождает V. Тогда добавим к ней вектор $v \in V \setminus \langle v_i\rangle$. Если система зависима, то есть $\lambda v + \sum \lambda_i v_i = 0$, то либо $\lambda = 0$ и тогда получаем, что v_i зависимы, что не так, либо $v = -\lambda^{-1}\sum \lambda_i v_i$, что противоречит определению v. Таким образом, система $v, v_1, \ldots, v_n$ независима, что противоречит максимальности. □

Определение. Пространство V называется конечномерным, если оно порождено конечной системой векторов $e_1,\ldots,e_n$.

Теорема 2 (Теорема о дополнении до базиса)**.** Пусть V – конечно порождённое векторное пространство. Тогда любой набор линейно независимых векторов $u_1,\ldots,u_m$ можно дополнить до базиса при помощи элементов заданного конечного порождающего множества $v_1,\ldots,v_n$.

Доказательство. Будем добавлять элементы из порождающего множества к линейно независимой системе до тех пор, пока полученный набор будет оставаться линейно независимым. Индукция по числу элементов из порождающего множества не лежащих в линейной оболочке $\langle u_i\rangle$. База. Если все элементы задействованы, то

$$V = \langle v_1,\ldots,v_n\rangle \subseteq \langle u_i\rangle \subseteq V.$$

Следовательно, система u_i является порождающей, и является базисом.

Рассмотрим порождающий набор $v_1,\ldots,v_n$. Пусть элемент v_i не лежит в $\langle u_i\rangle$. Рассмотрим новую систему $\{u_i\}\cup\{v_i\}$. Эта система линейно независима и, очевидно, в её линейной оболочке лежит больше векторов из порождающей системы. Тогда её можно дополнить до базиса. □

Следствие 1. В любом конечномерном пространстве есть конечный базис.

Доказательство. Пустое множество линейно независимо. Дополним его до базиса при помощи конечной порождающей системы. □

Теорема 3 (Теорема о равномощности базисов)**.** Пусть V – конечно порождённое пространство. Тогда любые два базиса V конечны и равномощны.

Доказательство. Можно считать, что один базис $e_1,\ldots,e_n$ конечен. Рассмотрим другой базис f_α. Тогда по теореме о линейной зависимости линейных комбинаций число независимых элементов в f_α не более n. Пусть их $m \le n$. Тогда заметим, что e_j выражаются через f_i. Тогда $n \le m$. □

Следствие 2. Размер любого базиса не больше размера любой порождающей системы и не меньше размера любой независимой системы векторов.

Определение (Размерность)**.** Пусть V – конечномерное векторное пространство. Размерностью V называется количество элементов в каком-то базисе V.

Замечание. В дальнейшем будем считать, что все пространства, кроме тех, что встречаются в примерах – конечномерны.

Теорема 4. Любое подпространство U конечномерного пространства V конечномерно и $\dim U \leq \dim V$. Равенство достигается только в случае $U = V$.

Доказательство. Пусть в пространстве U нет конечного базиса. Тогда рассмотрим линейно независимую систему $u_1, \ldots, u_m$ в U самого большого размера. Такая есть, так как линейно независимая система в U лежит в V и, следовательно, ограничена размерностью V. Если $\langle u_1, \ldots, u_m \rangle = U$, то теорема доказана. Иначе рассмотрим вектор $u_{m+1} \in U$, что $u_{m+1} \notin \langle u_1, \ldots, u_m \rangle$. Новая система $u_1, \ldots, u_m, u_{m+1}$ линейно независима и больше предыдущей. Противоречие. Значит в U есть конечный базис. Этот базис есть линейно независимая система внутри V и, как мы уже отмечали, имеет размер меньший, чем у базиса V и равный, только при $U = V$ (иначе можно нетривиально дополнить). □

Лемма 3. Пусть размерность V равна n, а $f_1, \ldots, f_n \in V$ набор векторов с одним из условий:
1) $f_1, \ldots, f_n$ – линейно независимая система
2) $f_1, \ldots, f_n$ – порождает V.
Тогда $f_1, \ldots, f_n$ – базис V.

Примеры:
1) Размерность пространства $M_{m \times n}(K)$ равна mn. Базисом является e_{ij}. Такой базис называется стандартным
2) Размерность пространства $K[x]_{\leq n} = n + 1$. Базис $1, x, \ldots, x^n$.
3) Размерность $\mathbb{C}$ над $\mathbb{R}$ равна 2.
4) Размерность $K[x]/p(x) = \deg p(x)$.

Как определить по системе векторов $v_1, \ldots, v_k$ из V, что она линейно независима? Пусть заранее известен $e_1, \ldots, e_n$ базис V. Разложим $v_j = \sum a_{ij} e_i$. Тогда сумма $0 = \sum \lambda_j v_j = \sum\sum \lambda_j a_{ij} e_i$ обнуляется тогда и только тогда, когда λ_i решение системы уравнений $A\lambda = 0$. Итого, вектора линейно независимы тогда и только тогда, когда у этой системы нет нетривиальных решений и зависимы иначе.

Как определить, что система порождающая? Для этого необходимо и достаточно получить любой базисный элемент e_i, то есть добиться существования линейной комбинации $\sum \lambda_i v_i - e_j = 0$ по всем j. Это эквивалентно тому что система уравнений с расширенной матрицей $(A|e_i)$ имеет решение (в последнем случае e_i – это столбец из стандартного базиса).

0.2. Сумма и пересечение подпространств

Определение (Сумма и пересечение подпространств).

Лемма 4 (Формула Грассмана). Пусть U и V подпространства некоторого W тогда имеет место соотношение

$$\dim U + \dim V = \dim U \cap V + \dim(U + V).$$

Доказательство. Выберем базис $e_1, \ldots, e_k$ в $U \cap V$ и дополним его до базисов U и V при помощи наборов $u_1, \ldots, u_m$ и $v_1, \ldots, v_l$ соответственно.

Я утверждаю, что базисом $U + V$ тогда является набор $e_1, \ldots, e_k, u_1, \ldots, u_m, v_1, \ldots v_l$. Во-первых это порождающая система. Во-вторых заметим, что если имеет место линейная зависимость

$$\sum \lambda_i e_i + \sum \mu_j u_j + \sum \nu_s v_s = 0,$$

то это соотношение можно представить в виде

$$\sum \lambda_i e_i + \sum \mu_j u_j = -\sum \nu_s v_s$$

Справа написан вектор из U, а слева – из V. Таким образом, и правая и левая часть лежит в пересечении. Но нетривиальная линейная комбинация $v_1, \ldots, v_l$ не может лежать в пересечении, так как иначе набор $e_1, \ldots, e_k, v_1, \ldots, v_l$ был бы линейно зависим.

□

Определение (Прямая сумма подпространств).

Замечание. Для любого подпространства $U \leq W$ существует пространство V, что $W = U \oplus V$.

Теорема 5. Пусть W – векторное пространство, а U_i его подпространства. Тогда следующие условия эквивалентны
1) W раскладывается в прямую сумму U_i
2) $U_i \cap \sum_{j \neq i} U_j = \{0\}$ и $\dim U_1 + \cdots + \dim U_n = \dim W$

0.3. Линейные отображения

Определение (Линейное отображение). Пусть U, V – два векторных пространства над полем K. Отображение $L\colon U \to V$ называется линейным, если
1) $\forall a, b \in U$ верно, что $L(a+b) = L(a) + L(b)$.
2) $\forall a \in U$, $\lambda \in K$ верно, что $L(\lambda a) = \lambda L(a)$.

Определение. Линейное отображение называется моно-, эпи-, изоморфизмом, если оно инъективно, сюръективно, биективно.

Примеры:
1) Пусть задана матрица $A \in M_{m\times n}(K)$. Тогда отображение $K^n \to K^m$, заданное как $x \to Ax$ линейно.
2) Пусть задан многочлен $p(x) \in K[x]$. Тогда отображение $K[x] \to K[x]$ $f(x) \to p(x)f(x)$ линейно.
3) Пусть $p(x) \in K[x]$ – многочлен. Тогда отображение $K[x] \to K[x]$, заданное как $f(x) \to f(p(x))$ линейно.
4) Пусть $g(x,y)$ непрерывная функция из $[0,1]^2 \to \mathbb{R}$. Тогда отображение $f(x) \to \int_0^1 f(y)g(x,y)dy$ линейно как отображение $C([0,1]) \to C([0,1])$.

Итак, благодаря примерам, мы поняли, что отображение $K^n \to K^m$, заданное с помощью матрицы A это частный случай линейного отображения. Таким образом, решение системы линейных уравнений можно понимать как нахождение прообраза элемента при линейном отображении. В дальнейшем мы увидим, что общий случай можно свести и разными способами к указанному частному случаю. Но прежде посмотрим, на общие качественные свойства линейных отображений.

Определение (Ядро).

Лемма 5 (Базовые свойства). 0) Пусть V – векторное пространство, а $\lambda \in K$. Тогда отображение $x \to \lambda x$ является линейным отображением.
1) Пусть L_1 и L_2 – два линейных отображения $V \to W$. Тогда $L_1 + L_2$ – их поточечная сумма – тоже линейное отображение.
2) Пусть $\mu \in K$ и $L\colon V \to W$ – линейное отображение. Тогда μL – тоже линейное отображение.
3) Пусть $L_1\colon V_1 \to V_2$, а $L_2\colon V_2 \to V_3$. Тогда $L_1 \circ L_2$ – линейное отображение $V_1 \to V_3$.
4) Пусть $L\colon V \to W$ линейное отображение. Тогда обратное к нему линейно.
5) Для композиции линейных отображений выполнена дистрибутивность. А именно, если $L_1, L_2\colon V \to W$, а $L_3\colon W \to U$, а $\lambda_1, \lambda_2 \in K$, то

$$(\lambda_1 L_1 + \lambda_2 L_2) \circ L_3 = \lambda_1(L_1 \circ L_3) + \lambda_2(L_2 \circ L_3).$$

6) Так же, если $L_1\colon V \to W$, а $L_2, L_3\colon W \to U$, и $\lambda_1, \lambda_2 \in K$, то

$$L_1 \circ (\lambda_1 L_2 + \lambda_2 L_3) = \lambda_1(L_1 \circ L_2) + \lambda_2(L_1 \circ L_3).$$

7) Пусть $L\colon V \to W$. Тогда $\operatorname{Ker} L$ и $\operatorname{Im} L$ – это подпространства V и W соответственно.
8) Линейное отображение инъективно тогда и только тогда, когда $\operatorname{Ker} L = \{0\}$. 9) Если линейное отображение обратимо, то обратное к нему тоже линейно.

Доказательство. □

Следствие 3. Множество всех отображений $\operatorname{Hom}(U, V)$ является векторным пространством.

Доказательство. Необходимо дополнительно проверить, что сложение ассоциативно, коммутативно, любой элемент имеет противоположенный и что умножение на скаляр ассоциативно. Эти свойства верны потому что они верны в V. □

Основная теорема, которая говорит про устройство линейных отображений, следующая:

Теорема 6. Пусть V_1, V_2 — векторные пространства над полем K. Пусть $e_1, \ldots, e_n$ - базис V_1, а $f_1, \ldots, f_n$ — набор каких-то векторов из V_2. Тогда существует единственное отображение $L\colon V_1 \to V_2$, что $L(e_i) = f_i$. Более того, если вектор $v \in V_1$ раскладывается по базису $v = \sum \lambda_i e_i$, то $L(v) = \sum \lambda_i f_i$.

Доказательство. Очевидно, что отображение должно быть задано такой формулой и это показывает единственность. Осталось показать, что оно линейно. Действительно, если два вектора $v = \sum \lambda_i e_i$ и $u = \sum \mu_i e_i$, то $u + v = \sum(\lambda_i + \mu_i)e_i$ есть разложение для суммы. Осталось посчитать $L(u+v)$ и раскрыть скобки. □

Следствие 4. Если L переводит некоторый базис V в базис W, то L обратимо. Обратно, если линейное отображение $L\colon V \to W$ является изоморфизмом, то L переводит любой базис V в базис W.

Доказательство. Пусть $e_1, \dots, e_n$ – базис V и $L(e_1), \dots, L(e_n)$ – базис W. Тогда есть единственное линейное отображение $T\colon W \to V$, которое переводит $L(e_i)$ в e_i. Заметим, что композиция $U \circ L$ переводит $e_i \to e_i$. Тогда по единственности такая композиция тождественна. Аналогично композиция в другом порядке тождественна. Тогда T – обратное отображение к L. В другую сторону. Пусть $e_1, \dots, e_n$ – базис V. Покажем, что $L(e_1), \dots, L(e_n)$ – базис W. Пусть есть линейная зависимость

$$0 = \lambda_1 L(e_1) + \dots + \lambda_n L(e_n).$$

Тогда

$$0 = L^{-1}(0) = L^{-1}(L(\lambda_1 e_1 + \dots + \lambda_n e_n)) = \lambda_1 e_1 + \dots + \lambda_n e_n.$$

Тогда все λ_i равны 0. Покажем, что $L(e_i)$ порождают W. Рассмотрим вектор $w \in W$. Тогда

$$w = L(v) = L\left(\sum \lambda_i e_i\right) = \sum \lambda_i L(e_i).$$

□

Это позволяет нам установить следующее

Следствие 5. Два конечномерных пространства V и W изоморфны тогда и только тогда, когда $\dim V = \dim W$. В частности $V \cong K^{\dim V}$. Более того, задание базиса равносильно заданию изоморфизма $V \to K^{\dim V}$.

Доказательство. Рассмотрим в качестве W пространство K^n и стандартный базис $e_1, \dots, e_n$. Тогда базису $f_1, \dots, f_n$ в V соответствует линейное отображение переводящее $f_i \to e_i$. Это отображение берёт вектор $v \in V$, раскладывает его как $v = \sum_{i=1}^n \lambda_i f_i$ и переводит его в столбец координат λ_i. Обратно, если есть изоморфизм $L\colon V \to K^n$, то можно рассмотреть обратное отображение L^{-1}. Тогда $L^{-1}(e_i)$ – это базис V. □

Определение. Изоморфизм между пространством V и пространством K^n называется линейной системой координат на пространстве V. Мы знаем что любая линейная система координат происходит из некоторого базиса. Отображение, сопоставляющее вектору v его i-ую координату, называется i-ой координатной функцией.

Далее я всегда буду предполагать, что все пространства конечномерны. Можно ли как-то попроще разобраться с нашими модельными примерами конечномерных пространств – пространствами K^n?

Следствие 6. Все линейные отображения $L\colon K^n \to K^m$ имеют вид $L(x) = Ax$, где A — матрица $m \times n$

Доказательство. Если задано линейное отображение L, то по нему определяется матрица A составленная из столбцов $L(e_i)$, где e_i – это стандартный базис K^n. Тогда отображения $x \to L(x)$ и $x \to Ax$ оба переводят $e_i \to L(e_i)$, то есть совпадают на базисе и, значит, совпадают везде. □

Определение. Рангом линейного отображения $L\colon V \to W$ называется размерность его образа $\operatorname{rank} L = \dim \operatorname{Im} L$. Рангом матрицы $A \in M_{m\times n}(K)$ называется ранг соответствующего ей линейного отображения $K^n \to K^m$. Иными словами ранг матрицы это размерность пространства, порождённого столбцами этой матрицы.

Теорема 7. Пусть $L\colon V \to W$ — линейное отображение между конечномерными пространствами. Тогда существует $e_1, \dots, e_n$ – базис V, такой что $e_1, \dots, e_k$ – базис $\operatorname{Ker} L$, а $L(e_{k+1}), \dots, L(e_n)$ – базис $\operatorname{Im} L$.

Доказательство. Выберем $e_1, \dots, e_k$ – базис $\operatorname{Ker} L$ и дополним его элементами $e_{k+1}, \dots, e_n$ до базиса всего V. Я утверждаю, что $L(e_{k+1}, \dots, L(e_n))$ являются базисом образа. Действительно, образ L порождён $L(e_1), \dots, L(e_n)$, но первые k элементов этого набора нули. Поэтому их можно исключить, что означает, что $L(e_{k+1}), \dots, L(e_n)$ порождают образ L.

Пусть сумма $\sum_{i=k+1}^n \lambda_i L(e_i) = 0$. Тогда элемент $\sum_{i=k+1}^n \lambda_i e_i$ лежит в ядре L. Отсюда получаем, что имеет место равенство

$$\sum_{i=k+1}^{n} \lambda_i e_i = \sum_{i=1}^{k} \mu_i e_i$$

которое приводит к нулевой линейной комбинации всех e_i. Но e_i – базис V и поэтому все коэффициенты, в частности, λ_i равны нулю. □

Следствие 7. Пусть $L\colon V \to W$ — линейное отображение между конечномерными пространствами. Тогда

$$\dim V = \dim \operatorname{Ker} L + \dim \operatorname{Im} L = \dim \operatorname{Ker} L + \operatorname{rank} L.$$

Доказательство. В обозначениях предыдущей теоремы размерность ядра равна k, размерность образа равна $n - k$, что в сумме и даёт n. □

Следствие 8 (Принцип Дирихле для линейных отображений)**.** Пусть V и W два пространства размерности n и $L\colon V \to W$ – линейное отображение между ними. Тогда L – сюръективно тогда и только тогда, когда L – инъективно.

Следствие 9. Квадратная $A \in M_n(K)$ обратима тогда и только тогда, когда $\operatorname{rank} A = n$. Обратимые матрицы так же называются невырожденными.

Доказательство. Следует из принципа Дирихле для линейных отображений. □

Итак, у нас есть модельная задача: пусть $A\colon K^n \to K^m$ линейное отображение и элемент $y \in K^m$. Необходимо описать прообраз элемента y, то есть решить систему линейных уравнений $Ax = y$. Про эту задачу мы понимаем всё. Наша текущая задача – научиться сводить общую задачу к модельной. Однако тут возникает нюанс – это можно сделать разными способами. Точнее:

Определение. Пусть V_1, V_2 - векторные пространства над полем K с базисами $e_1, \ldots, e_n$ и $f_1, \ldots, f_m$ соответственно. Пусть $L\colon V_1 \to V_2$ - линейное отображение. Составим матрицу A, такую, что её i-ый столбец состоит из координат $L(e_i)$ в базисе f. Такая матрица называется матрицей линейного отображения L в базисах e и f.

Теорема 8. Пусть $L\colon V_1 \to V_2$ – линейное отображение, а e и f – базисы V_1 и V_2. Тогда матрица отображения L – это единственная такая матрица A, что $Ax = y$, где x – координаты вектора $v \in V_1$, а y – координаты его образа в V_2.

Лемма 6. Пусть $L_1, L_2\colon U \to V$ и e – базис U, а f – базис V. Тогда матрица $L_1 + L_2$ в базисах e и f есть сумма матриц L_1 и L_2. Аналогично про домножение на скаляр.

Доказательство. Пусть $L_1 e_j = \sum A_{ij} f_i$, а $L_2 = \sum B_{ij} f_i$. Тогда $L_1(e_j) + L_2(e_j) = \sum (A_{ij} + B_{ij}) f_i$, что и доказывает требуемое. □

Следствие 10. Пусть пространства U и V имеют размерности n и m. Допустим выбраны базисы $e_1, \ldots, e_m$ и $f_1, \ldots, f_m$ пространств U и V соответственно. Тогда сопоставление линейному отображению $L \in \operatorname{Hom}(U, V)$ его матрицы в указанных базисах есть изоморфизм

$$\operatorname{Hom}(U, V) \xrightarrow{\sim} M_{m\times n}(K).$$

Пусть есть две матрицы $A \in M_{n\times m}(K)$ и $B \in M_{m\times l}(K)$. Какая матрица соответствует композиции линейных отображений, построенных по A и B?

Теорема 9. Пусть U, V, W пространства с базисами $e_1, \ldots, e_n$, $f_1, \ldots, f_m$, $g_1, \ldots, g_l$. Пусть $L\colon U \to V$ и $T\colon V \to W$ два линейных отображения. Пусть A и B матрицы T и L в парах базисов v,w и u, v. Тогда матрица $T \circ L$ в базисах u, w равна произведению AB.

Доказательство. ;Пусть $u \in U$, а x – его столбец координат. По ассоциативности произведения матриц имеем

$$A(Bx) = (AB)x$$

Таким образом, AB и есть матрица произведения. □

Какая матрица соответствует изоморфизму? Та, которая осуществляет изоморфизм пространств K^n и K^m.

Определение. Матрица $A \in M_{m\times n}$ называется обратимой, если существует $B \in M_{n\times m}$, что $AB = E_m$, $BA = E_n$. Матрица B называется обратной к A и обозначается A^{-1}.

Следствие 11. Если матрица $A \in M_{n\times m}$ обратима, то $n = m$. Обратная матрица единственна.

Доказательство. Обратимая матрица осуществляет изоморфизм пространства K^n и K^m. Но если пространства изоморфны, то их размерности равны, то есть $n = m$. □

Следствие 12. Если квадратная матрица A обратима с одной стороны, то она обратима.

Доказательство. Следует из принципа Дирихле для линейных отображений. □

Настала пора ответить на следующий вопрос: что происходит с матрицей линейного отображения, если мы заменим базисы пространств?

Определение. Рассмотрим пространство V размерности n и два базиса $e_1, \ldots, e_n$ – старый и $f_1, \ldots, f_n$ – новый. Рассмотрим матрицу отображения Id_V из базиса f в базис e. Эта матрица называется перехода из базиса e в базис f.

Это отображение устроено следующим образом: оно берёт столбец x и сопоставляет ему вектор $v = \sum x_i e_i$, у которого ровно такие координаты, а затем считает его координаты в новом базисе. То есть матрица C – это способ пересчитать координаты вектора из старого базиса в новый. Точнее

Замечание. Если есть вектор v и его координаты в базисе $e_1, \ldots, e_n$ это столбец x, а координаты в базисе $f_1, \ldots, f_n$ – это столбец y, то верно соотношение

$$y = Cx.$$

Замечание. Имеют место следующие соотношения $C_f^e = (C_e^f)^{-1}$ и $C_f^g C_e^f = C_e^g$. В частности, матрица замены базиса всегда обратима.

Теперь можно разобраться с ситуацией про замену базиса.

Теорема 10. Пусть даны пространства V, W и линейное отображение $L\colon V \to W$. Рассмотрим в пространстве V два базиса $e_1, \dots, e_n$ (старый) и $e'_1, \dots, e'_n$ (новый). Аналогично в W — $f_1, \dots, f_m$ (старый) и $f'_1, \dots, f'_m$ (новый). В указанных предположениях матрицу A' линейного отображения L в базисах e' и f' можно выразить через матрицу A и матрицы замены C и D следующим образом

$$A' = DAC^{-1}.$$

Доказательство. $L = Id_W \circ L \circ Id_V$. Тогда $[L]_e^f = [Id_W]_{f'}^f [L]_{e'}^{f'} [Id_V]_e^{e'}$, что и требовалось ☐

Следствие 13. Для любой матрицы $A \in M_{m\times n}$ ранга r существуют обратимые матрицы C и D, что CAD^{-1} имеет вид

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Следствие 14. Для любой матрицы $A \in M_{m\times n}(K)$ ранга r существуют матрицы $B \in M_{m\times r}$ и $C \in M_{r\times n}(K)$, что $A = BC$.

0.4. Проекция на подпространство

Следствие 15. Для любого подпространства $U \le K^n$ размерности $n - d$ существует линейное отображение $L\colon K^n \to K^d$, такое, что $U = \operatorname{Ker} L$. Иными словами, такое подпространство задаётся d уравнениями.

Доказательство. Рассмотрим базис U, $f_1, \dots, f_k$, дополним его до базиса K^n. Построим отображение L, такое что $L(f_i) = e_i$ при $i \le k$ и $L(f_i) = 0$ иначе. ☐

Сделаем ещё несколько уточнений для случая двух подпространств:

Лемма 7. Пространство W раскладывается в прямую сумму U и V тогда и только тогда, когда существует оператор $L\colon W \to W$, что $L^2 = L$ и $U = \operatorname{Im} L$, а $V = \operatorname{Ker} L$. Более того, такой оператор единственен.

Определение. Оператор, удовлетворяющий тождеству $L^2 = L$ называется проектором. Любой проектор однозначно определяет пару U и V. Проектор, соответствующий такой паре подпространств, называется проекцией на U вдоль V.

0.5. Свойства ранга

Наша ближайшая задача – классифицировать все линейные отображения. Для этого мы посмотрим на их численную характеристику – ранг. Ранг линейного отображения отвечает за его обратимость и за размерность множества решений уравнения $Lx = y$. Исследуем свойства ранга.

Лемма 8. Пусть $S\colon V \to W$, $T\colon W \to U$, тогда

$$\operatorname{rank} T \circ S \le \min(\operatorname{rank} T, \operatorname{rank} S).$$

Доказательство. Для начала покажем, что если есть $L\colon V_1 \to V_2$ и $V' \le V_1$, то $\dim L(V') \le \dim V'$. Действительно

$$\dim V' = \dim \operatorname{Im} L|_{V'} + \dim \operatorname{Ker} L|_{V'} \ge \dim L(V').$$

Теперь $\operatorname{Im} T\circ S \subseteq \operatorname{Im} T$, поэтому $\operatorname{rank} T\circ S \le \operatorname{rank} T$. Далее $\operatorname{rank} T\circ S = \dim \operatorname{Im} T\circ S = \dim T(S(V)) \le \dim S(V) = \operatorname{rank} S$. Что и требовалось. ☐

Лемма 9. Пусть $S, T\colon V \to W$ – два линейных отображения. Тогда

$$\operatorname{rank}(T + S) \le \operatorname{rank} T + \operatorname{rank} S.$$

Доказательство. Для начала покажем, что, если $V_1, V_2 \le V$, то $\dim V_1 + V_2 \le \dim V_1 + \dim V_2$. Действительно, если $e_1, \dots, e_k$ базис V_1, и $f_1, \dots, f_l$ базис V_2, то $e_1, \dots, e_k, f_1, \dots, f_l$ порождает $V_1 + V_2$. Осталось заметить, что размер базиса меньше размера любой порождающей системы. Теперь заметим, что $\operatorname{Im}(S + T) \le \operatorname{Im} S + \operatorname{Im} T$. Тогда $\operatorname{rank}(S + T) \le \dim(\operatorname{Im} S + \operatorname{Im} T) \le \operatorname{rank} S + \operatorname{rank} T$. ☐

Лемма 10 (Ранг не меняется при подкрутке на изоморфизмы)**.** Пусть $L\colon U_1 \to V_1$ – линейное отображение, а $G\colon U_1 \to U_2$ и $H\colon V_1 \to V_2$ изоморфизмы. Тогда $\operatorname{rank}(H \circ L \circ G^{-1}) = \operatorname{rank} L$.

Доказательство. Для начала покажем, что, если некоторое линейное отображение $T\colon V \to W$ обратимо, то $\dim T(U) = \dim U$ для любого $U \le V$. Действительно, если T обратимо, то оно инъективно и его сужение на U тоже инъективно. Тогда $\dim U = \dim T(U) + \dim \operatorname{Ker} T|_U = \dim T(U)$. Заметим, что если G обратимо, то G^{-1} тоже и, следовательно, сюръективно. Тогда $G^{-1}(X) = U$. Теперь $\operatorname{rank} L = \dim L(U) = \dim H(L(U)) = \dim H(L(G^{-1}(X)) = \operatorname{rank} H \circ L \circ G^{-1}$. $\square$

Напомню, что ранг матрицы – это ранг соответствующего линейного отображения. Иными словами, это размерность пространства, порождённого столбцами матрицы.

Однако, в методе Гаусса удобно смотреть на строчки, чем на столбцы. Это замечание приводит нас к определению строчного ранга.

Определение. Пусть $A \in M_{m\times n}(K)$. Определим $\operatorname{rank}_{row}$ – строчный ранг матрицы A как размерность пространства, порождённого её строками.

Замечание. Очевидно равенство $\operatorname{rank}_{row} A = \operatorname{rank} A^\top$.

Поэтому исследуем свойства операции транспонирования матрицы.

Лемма 11. Пусть A, B – матрицы. Тогда
0) ${A^\top}^\top = A$.
1) $(A + \lambda B)^\top = A^\top + \lambda B^\top$.
2) $(AB)^\top = B^\top A^\top$.
3) $(A^\top)^{-1} = (A^{-1})^\top$.

Доказательство. Покажем второй пункт. Имеем

$$(AB)^\top_{ij} = (AB)_{ji} = \sum_k A_{jk} B_{ki} = \sum_k B^\top_{ik} A^\top_{kj} = (B^\top A^\top)_{ij}.$$

Теперь покажем пункт 3. Пусть $AA^{-1} = E_n$. Тогда $E_n = E_n^\top = (AA^{-1})^\top = (A^{-1})^\top A^\top$. Аналогично из равенства $A^{-1}A = E_n$ следует $E_n = A^\top (A^{-1})^\top$, что завершает доказательство. $\square$

Замечание. Заметим, что, вообще говоря, проверять второе условие не обязательно. Действительно, предположим, что $BA = E_n$, где A, B квадратные матрицы размера n. Покажем, что тогда $B = A^{-1}$. Заметим, что матрица A имеет нулевое ядро так как в противном случае матрица E_n имела бы нетривиальное ядро. Тогда матрица A обратима по принципу Дирихле, потому что обратимо соответствующее линейное отображение. Тогда $B = BAA^{-1} = E_n A^{-1} = A^{-1}$.

Теорема 11. Пусть $A \in M_{m\times n}(K)$. Тогда ранг по строчкам и ранг по столбцам совпадают.

Доказательство. Представим матрицу A в виде произведения

$$A = C \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} D,$$

где $C \in \mathrm{GL}_m(K)$, а $D \in \mathrm{GL}_n(K)$, а r – ранг матрицы. Тогда

$$A^\top = D^\top \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} C^\top.$$

Так как ранг не меняется при домножении на обратимые матрицы, то $\operatorname{rank} A^\top = \operatorname{rank} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} = r = \operatorname{rank} A$. $\square$

Есть и другое определение – минорный ранг.

Определение. Пусть $A \in M_{m\times n}(K)$ и даны два множества индексов $I = \{i_1 < \dots < i_k\} \subseteq \overline{1,m}$, $J = \{j_1 < \dots < j_l\} \subseteq \overline{1,n}$. Подматрицей $A_{I,J} \in M_{k\times l}$ определим так, чтобы ${A_{I,J}}_{s,t} = a_{i_s, j_t}$.

Определение (Минорный ранг)**.** Пусть $A \in M_{m\times n}(K)$. Минорный ранг матрицы A – это наибольшее такое k, что существует невырожденная квадратная подматрица $A_{I,J}$ размера k внутри A.

Теорема 12. Ранг матрицы равен её минорному рангу.

Доказательство. Пусть A – матрица ранга r. Тогда в матрице A есть r линейно независимых столбцов $v_1, \dots, v_r$. Составим из них матрицу A'. Ранг A' равен r. Теперь, так как строчный и столбцовый ранги совпадают, то в A' есть r линейно независимых строк. Они дают квадратную невырожденную подматрицу в A' и следовательно в A. $\square$

Сейчас перед нами стоит вопрос, как вычислить ранг матрицы. При решении системы линейных уравнений нам пригодились элементарные преобразования. Покажем, что они не влияют на ранг матрицы.

Для этого покажем, что элементарное преобразование можно реализовать домножив исходную матрицу на некоторую обратимую матрицу. Точнее, рассмотрим матрицу

$$E_{ij}(\lambda) = E_n + \lambda e_{ij} = \begin{array}{cc} & \begin{matrix} & j & & \end{matrix} \\ \begin{matrix} \\ \\ i \\ \\ \end{matrix} & \begin{pmatrix} 1 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & \lambda & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix} \end{array}.$$

Домножение на матрицу $E_{ij}(\lambda)$ приводит к тому, что матрица

$$A' = E_{ij}(\lambda)A.$$

получается из матрицы A прибавлением j-ой строки к i-ой с коэффициентом λ. Действительно, домножение матрицы A на матрицу e_{ij} выделяет из матрицы A j-ую строку и записывает её на позицию i.

Определение. Матрица $E_{ij}(\lambda)$ называется матрицей элементарного преобразований первого типа.

Что соответствует замене местами двух строк матрицы A? Для этого надо поставить i-ую строчку в позицию j и наоборот, а остальные строки оставить на месте. Итого имеем матрицу

$$T_{(ij)} = \begin{array}{cc} & \begin{matrix} & & j & i & & \end{matrix} \\ \begin{matrix} \\ \\ j \\ i \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 0 & 1 & & \\ & & 1 & 0 & & \\ & & & & \ddots & \\ 0 & & & & & 1 \end{pmatrix} \end{array} = E_n - e_{ii} - e_{jj} + e_{ij} + e_{ji}.$$

Определение. Матрица $T_{(ij)}$ называется матрицей элементарного преобразований второго типа или матрицей транспозиции.

А что вообще происходит при перестановке строк? Каждой перестановке $\sigma \in S_n$ однозначно соответствует матрица T_σ заданная правилом

$$(T_\sigma)_{ij} = \begin{cases} 1, & \text{если } \sigma(j) = i \\ 0, & \text{иначе} \end{cases}.$$

Определение. Матрица T_σ называется матрицей перестановки σ.

Замечание. Заметим, что произведение $T_\sigma T_\tau = T_{\sigma\tau}$. Это даёт гомоморфизм из группы перестановок в группу $\mathrm{GL}_n(K)$. Любая матрица, у которой в каждом столбце ровно один единичный элемент, а все остальные 0 является матрицей перестановки. Можно написать аналогичное условие для строк.

Наконец, преобразованию третьего типа соответствует матрица

$$D_i(\lambda) = \begin{array}{cc} & \begin{matrix} & & i & & \end{matrix} \\ \begin{matrix} \\ \\ i \\ \\ \\ \end{matrix} & \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} \end{array}.$$

Определение. Матрица $D_i(\lambda)$, $\lambda \in K^*$ называется матрицей элементарного преобразований третьего типа.

Замечание. Заметим, что $E_{ij}^{-1}(\lambda) = E_{ij}(-\lambda)$ и $D_i(\lambda) = D_i(\lambda^{-1})$, а $T_{ij}^{-1} = T_{ij}$.

Прежде чем пойти дальше сформулируем давно напрашивающиеся определения про преобразования столбцов, которые нам пригодятся в дальнейшем.

Определение. Определим элементарные преобразования столбцов матрицы $A \in M_{n\times m}(K)$ следующим образом:
1) Преобразование первого типа соответствует домножению

$$A \to AE_{ij}(\lambda),$$

которое приводит к тому, что i-ый столбец прибавляется к j-ому с коэффициентом λ.
2) Преобразование столбцов второго типа – это перестановка i-го и j-го столбца местами, что соответствует домножению на $T_{(ij)}$

$$A \to AT_{(ij)}.$$

3) Преобразование третьего типа – это домножение на $D_i(\lambda)$ – умножение i-го столбца на $\lambda \in K^*$.

Теорема 13. Элементарные преобразования строк и столбцов не меняют ранг матрицы.

Доказательство. Преобразования строк и столбцов соответствуют домножениям на обратимые матрицы и поэтому не меняют ранг. □

Композиция элементарных преобразований, это конечно не элементарное преобразование, но можно кое-что заметить

Определение. Пусть $A \in M_{n\times m}(K)$. Тогда матрица A называется верхнетреугольной, если $\forall i < n, j < m$ и $i < j$, то $A_{ij} = 0$. Аналогично определяются нижнетреугольные матрицы. Верхнетреугольная матрица называется верхней унитреугольной, если на диагонали стоят единицы. Аналогично для нижнетреугольных.

Замечание. Матрица $E_{ij}(\lambda)$ является верхнеунитреугольной, если $i < j$ и нижней унитреугольной иначе. Произведение верхнетреугольных матриц – снова верхнетреугольная. И т.д.

Переформулируем и дополним теорему, которую мы доказали, когда обсуждали метод Гаусса на языке матриц.

Теорема 14 (LU разложение)**.** Пусть $A \in M_{m\times n}(K)$. Тогда существует набор матриц элементарных преобразований $S_1, \dots, S_k \in M_{m\times m}(K)$, что матрица $S_1 \dots S_k A$ имеет ступенчатый вид. Существует набор матриц элементарных преобразований первого типа $S_1, \dots, S_k$ с $i < j$ (то есть они все нижние унитреугольные) и матрица перестановки $P \in M_m(K)$, что $S_1 \dots S_k PA$ является верхнетреугольной матрицей. Существуют две обратимые матрицы $L \in M_m(K)$ – нижняя унитреугольная и $U \in M_n(K)$ – верхнетреугольная и матрицы перестановок $P \in M_m(K)$, $Q \in M_n(K)$, что $LPAQU$ есть матрица вида

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Если матрица A обратима, то её можно представить в виде PLU, где P – матрица перестановки, L – нижняя унитреугольная, U – верхняя треугольная матрицы. В частности обратимая матрица есть произведение матриц элементарных преобразований.

Приведём пример использования свойства ранга. Для этого посмотрим на пример с поисковой системой. Пусть G граф на n вершинах. Тогда рассмотрим квадратную матрицу $P(G)$ размера n

$$P(G)_{ij} = \begin{cases} \frac{1}{d_j^{out}}, & \text{если есть ребро } j \to i \\ 1, & \text{если } i = j \text{ и из вершины } i \text{ нет исходящих рёбер} \\ 0, & \text{иначе} \end{cases}.$$

Эта матрицу еще называется матрицей случайного блуждания на графе G. Смысл состоит в том, что она моделирует следующую ситуацию – если мы с вероятностями $w_1, \dots, w_n$ находимся в вершинах графа и на следующем шаге хотим из j-ой вершины перейти в i-ую по ребру графа, при этом проход по каждому ребру равновероятен, то после такого действия новый набор вероятностей будет иметь вид $P(G)w$. Матрица P обладает тем свойством, что сумма элементов в каждом столбце равна 1.

Теперь систему для нахождения «важностей» вершин графа можно переписать в виде $E_n w = P(G)w$ или

$$(E_n - P(G))w = 0.$$

Нам хочется показать, что у этой системы есть нетривиальное решение, то есть, что ранг матрицы $E_n - P(G)$ строго меньше n. Для этого можно показать, что ранг

$$E_n - P(G)^\top$$

меньше n, а это легко сделать – достаточно заметить, что матрица $E_n - P(G)^\top$ имеет столбец $(1, 1, \dots, 1)^\top$ в своём ядре. Заметьте – это не помогает найти решение исходной системы, а лишь доказывает его существование.

0.6. Определитель

Есть ли какая-нибудь численная характеристика, которая позволяет сказать, что матрица $A \in M_n(K)$ обратима? Напомню, что обратимость матрицы A равносильна линейной независимости её столбцов. Если посмотреть на картинку над $\mathbb{R}$, то видно, что вектора линейно независимы тогда и только тогда, когда объём параллелепипеда на них натянутого отличен от нуля.

Определение. Пусть V – векторное пространство размерности n над $\mathbb{R}$, тогда для набора $v_1, \dots, v_n \in V$ определим параллелепипед

$$D(v_1, \dots, v_n) = \left\{ \sum_{i=1}^{n} \lambda_i v_i \mid \text{ где } \lambda_i \in [0,1] \right\}.$$

Обозначим для краткости $\mathrm{Vol}(v_1, \dots, v_n) = \mathrm{Vol}(D(v_1, \dots, v_n))$. Более того, работая с набором векторов из пространство K^n я буду отождествлять его с матрицей $M_n(K)$.

Попробуем понять есть ли возможность как-то аксиоматизировать понятие объёма параллелепипеда, так, чтобы оно работало над произвольным полем и для произвольного пространства. А заодно узнаем как его посчитать.

Отображение $\mathrm{Vol}\colon M_n(K) \to K$ типа объёма естественно удовлетворяет следующим свойствам...
0) Объём единичного кубика это $\mathrm{Vol}(E) = 1$.
1) При растяжении одного вектора объём меняется пропорционально $\mathrm{Vol}(\dots, \lambda v, \dots) = |\lambda| \mathrm{Vol}(\dots, v, \dots)$.
2) Исходя из принципа Кавальери $\mathrm{Vol}(\dots, v, \dots, u, \dots) = \mathrm{Vol}(\dots, v, \dots, u + \lambda v, \dots)$.
3) Если в наборе есть два одинаковых вектора, то $\mathrm{Vol}(\dots, v, \dots, v, \dots) = 0$.
Для общего пространства V условие типа 0) не имеет смысла, так как нет возможности выбрать какой-то канонический базис в V.

Безусловно, модуль числа в свойстве 1) нет возможности определить над произвольным полем. Да и вообще ни о какой положительности, как для объёма, речи нет. Вместо свойства 1) хочется потребовать
1') $\omega(\dots, \lambda v, \dots) = \lambda \omega(\dots, v, \dots)$.
Заметим, что если мы находимся над $\mathbb{R}$ и для отображения ω выполнены свойства 0), 1'), 2), то для отображения $|\omega|$ выполнены свойства 0),1),2).

Наконец, с алгебраической точки зрения свойство 2) не самое удобное. Вместо него мы рассмотрим свойство
2') $\omega(\dots, u + v, \dots) = \omega(\dots, u, \dots) + \omega(\dots, v, \dots)$.

Определение. Отображение $\omega\colon V^l \to K$ называется полилинейным, если для любого $1 \leq i \leq l$

$$\omega(v_1, \dots, v_i + \lambda u_i, \dots, v_l) = \omega(v_1, \dots, v_i, \dots, v_l) + \lambda \omega(v_1, \dots, u_i, \dots, v_l).$$

Здесь одновременно зашифрованы свойства типа 1') и 2').

Лемма 12. Для полилинейного отображения $\omega\colon V^l \to K$ выполнено, что если $e_1, \dots, e_n$ базис V, то

$$\omega(v_1, \dots, v_l) = \sum_{i_1, \dots, i_n} \omega(e_{i_1}, \dots, e_{i_l}) \prod_{j=1}^{l} a_{i_j, j}$$

где $a_{i,j}$ – это i-ая координата v_j.

Определение. Полилинейное отображение $\mu\colon V^l \to K$ называется антисимметричным или кососимметричным, если

$$\omega(v_1, \dots, v, \dots, v, \dots, v_l) = 0.$$

и симметричным, если

$$\omega(v_1, \dots, v_i, \dots, v_j, \dots, v_l) = \omega(v_1, \dots, v_j, \dots, v_i, \dots, v_l)$$

Теперь надо разобраться с естественностью рассматриваемых понятий и как-то отметить их свойства.

Лемма 13. Пусть V – пространство размерности n. Для кососимметричного полилинейного отображения $\mu\colon V^l \to K$ выполнено, что
1) $\omega(\dots, u, \dots, v, \dots) = -\omega(\dots, v, \dots, u, \dots)$.
2) В случае, если $\mathrm{char}\, K \neq 2$, то из свойства, описанного в пункте 1), следует кососимметричность.
3) Для любой перестановки $\sigma \in S_l$ выполнено, что $\omega(v_{\sigma(1)}, \dots, v_{\sigma(l)}) = \mathrm{sgn}(\sigma) \omega(v_1, \dots, v_n)$.
4) $\omega(\dots, v, \dots, u, \dots) = \omega(\dots, v, \dots, u + \lambda v, \dots)$.
5) Пусть $l = n$. Для набора векторов $v_1, \dots, v_n$ и базиса $e_1, \dots, e_n$ выполнено, что

$$\mu(v_1, \dots, v_n) = \omega(e_1, \dots, e_n) \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j), j} = \omega(e_1, \dots, e_n) \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i, \sigma(i)},$$

где $a_{i,j}$ – это i-ая координата j-ого вектора.

Доказательство. Докажем первое утверждение. Распишем

$$
\begin{aligned}
0 &= \omega(\dots, u+v, \dots, u+v, \dots) = \\
&= \omega(\dots, u, \dots, u, \dots) + \omega(\dots, v, \dots, u, \dots) + \omega(\dots, u, \dots, v, \dots) + \omega(\dots, v, \dots, v, \dots) = \\
&= \omega(\dots, v, \dots, u, \dots) + \omega(\dots, u, \dots, v, \dots),
\end{aligned}
$$

что и доказывает утверждение.
Теперь, если $\operatorname{char} K \neq 2$, то переставляя одинаковые столбцы получаем

$$\omega(\dots, v, \dots, v, \dots) = -\omega(\dots, v, \dots, v, \dots).$$

То есть $2\omega(\dots, v, \dots, v, \dots) = 0$. Осталось поделить на 2.
По определению знака перестановки

$$\operatorname{sgn}(\sigma) = (-1)^k, \text{ где } k \text{ – число транспозиций в разложении } \sigma$$

откуда получаем, что применить σ это тоже самое, что применить k транспозиций, то есть изменить знак k раз, что и требовалось.
Докажем 4).

$$\omega(\dots, v, \dots, u+\lambda v, \dots) = \omega(\dots, v, \dots, u, \dots) + \lambda\omega(\dots, v, \dots, v, \dots) = \omega(\dots, v, \dots, u, \dots).$$

Для доказательства свойства 5) воспользуемся общим выражением для полилинейной формы.

$$\omega(v_1, \dots, v_n) = \sum_{i_1, \dots, i_n} \omega(e_{i_1}, \dots, e_{i_n}) \prod_{j=1}^{n} a_{i_j, j}.$$

Если два индекса совпадают, то $\omega(e_{i_1}, \dots, e_{i_n}) = 0$, а вместе с ним и всё слагаемое. Остаются только наборы с разными i_α которые однозначно соответствуют перестановкам $\sigma(k) = i_k$. Теперь заметим, что $\omega(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \operatorname{sgn}(\sigma)\omega(e_1, \dots, e_n)$, что доказывает первое равенство.
Теперь

$$\prod_{j=1}^{n} a_{\sigma(j), j} = \prod_{i=1}^{n} a_{i, \sigma^{-1}(i)}.$$

Если вместо σ поставить сумму по σ^{-1}, то с одной стороны сумма не поменяется, а с другой стороны по тождеству выше, можно будет перекинуть σ на другой индекс. □

Определение. Пусть $n = \dim V$. Антисимметричная полилинейная форма $\omega\colon V^n \to K$ называется формой объёма на V. Если такая форма не равна 0, то будем говорить, что она невырождена.

До этого момента мы говорили про объекты которых, возможно, просто не существует. Настала пора предъявить для них конструкцию и доказать её единственность.

Определение. Определителем det называется отображение $\det\colon M_n(K) \to K$, такое, что

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{1 \le i \le n} a_{i\sigma(i)}.$$

Определение. Пусть $e_1, \dots, e_n$ базис пространства V. Определим отображение $\operatorname{Vol}_e\colon V^n \to K$, такое, что

$$\operatorname{Vol}_e(v_1, \dots, v_n) = \det(e(v_1), \dots, e(v_n)),$$

где $e\colon V \to K^n$ – отображение сопоставления координат.

Теорема 15. Следующие свойства верны:
1) Определитель является формой объёма на K^n, такой, что $\det E = 1$.
2) Если V – пространство размерности n, то любая форма объёма на V имеет вид

$$\omega = \omega(e_1, \dots, e_n) \operatorname{Vol}_e.$$

В частности, если есть два базиса e и f, то $\operatorname{Vol}_f = \det(C_e^f) \operatorname{Vol}_e$.
3) Пространство форм объёма одномерно.
4) Для любой невырожденной формы объёма ω верно утверждение: $\omega(v_1, \dots, v_n) = 0$ тогда и только тогда, когда $v_1, \dots, v_n$ линейно зависимы.

Доказательство. Определитель является полилинейной функцией столбцов, так как каждое слагаемое содержит только одну координату из каждого столбца матрицы A. Осталось показать, что если в матрице A пара столбцов одинакова, то её определитель равен нулю. Пусть равны столбцы k и l. Рассмотрим транспозицию $\tau = (kl)$. Тогда все перестановки разбиваются на классы A_n и $A_n\tau$. Выпишем теперь сумму

$$\sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} a_{\sigma(j)j} = \sum_{\sigma\in A_n} \prod_{j=1}^{n} a_{\sigma(j)j} - \sum_{\sigma\tau\in A_n\tau} \prod_{j=1}^{n} a_{\sigma(\tau(j))j}.$$

Заметим, что $a_{\sigma(\tau(j))j} = a_{\sigma(j)j}$, если $j \neq k, l$. Но при $j = k$, благодаря тому что $\tau(k) = l$ и равенству столбцов матрицы A, имеем $a_{\sigma(\tau(k))k} = a_{\sigma(l)k} = a_{\sigma(l)l}$. Аналогично $a_{\sigma(\tau(l))l} = a_{\sigma(k)k}$, что означает, что произведения для одной и той же σ в правой и левой части суммы одинаковы.

Далее, используя предыдущую лемму, получаем $\omega(v_1, \dots, v_n) = \omega(e_1, \dots, e_n) \operatorname{Vol}_e(v_1, \dots, v_n)$. Осталось связать Vol_e и Vol_f для разных базисов. Для этого надо вычислить $\operatorname{Vol}_f(e_1, \dots, e_n)$. Для этого надо построить матрицу из координат столбцов e_i в базисе f и посчитать её определитель. Но это матрица C_e^f. Итого $\operatorname{Vol}_f = \det(C_e^f) \operatorname{Vol}_e$.
Все формы пропорциональны Vol_e для некоторого базиса e. Пусть теперь $v_1, \dots, v_n$ набор векторов. Тогда, если $v_1, \dots, v_n$ линейно зависимы, то существует нетривиальная линейная комбинация $\sum \lambda_i v_i = 0$. Пусть $\lambda_1 \neq 0$, тогда можно считать, что $\lambda_1 = 1$. Тогда, прибавляя к первому аргументу остальные с соответствующими коэффициентами, получаем набор с нулевым первым вектором. $\omega(0, \dots)$ с одной стороны 0, а с другой равен определителю исходной матрицы. Обратно, если $v_1, \dots, v_n$ независимы, то $\omega = \omega(v_1, \dots, v_n) \operatorname{Vol}_v$. Так как $\omega \neq 0$, то коэффициент $\omega(v_1, \dots, v_n) \neq 0$. □

Лемма 14. Для определителей квадратных матриц верны следующие свойства:
0) $\det(A) = \det(A^T)$.
1) Определитель не меняется при элементарных преобразованиях первого типа для строк и столбцов. При смене строк местами меняется знак определителя. При домножении строки на λ, определитель домножается на λ.
2) $\det(AB) = \det(A)\det(B)$.
3) $\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det(A)\det(C)$.
4) Определитель верхнетреугольной или нижнетреугольной матрицы равен произведению диагональных элементов.
5) $\det(A^{-1}) = (\det A)^{-1}$.
6) $\det\colon \operatorname{GL}(V) \to K^*$ является гомоморфизмом групп.

Доказательство. Свойство 0) мы уже доказали. Из него сразу же следует свойство 1). Докажем 2).
Заметим, что форма $B \to \det AB$ линейна по столбцам B. Тогда это форма объёма и следовательно она пропорциональна форме $B \to \det B$. Для того, чтобы найти коэффициент пропорциональности, подставим $B = E_n$. Имеем $\det AE_n = \det A = c \det E_n = c$. Откуда для любого B получаем

$$\det AB = \det A \det B.$$

Покажем 3). Сначала посчитаем

$$\det \begin{pmatrix} E_n & B \\ 0 & E_m \end{pmatrix}.$$

Заметим, что с помощью элементарных преобразований 1-го типа из неё можно сделать E_{n+m}. Тогда определитель равен 1. Теперь заметим, что форма

$$A \to \det \begin{pmatrix} A & B \\ 0 & E_m \end{pmatrix}$$

форма объёма на K^n. Как и раньше подставив $A = E_n$ получаем что коэффициент пропорциональности с $A \to \det A$ равен 1.
Теперь отображение

$$C \to \det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

есть форма объёма, по строчкам C. Подставляя $B = E_m$ находим коэффициент пропорциональности $\det A$, что и даёт

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \det C.$$

Применяя предыдущее условие несколько раз приходим к пункту 4).
Беря равенство $AA^{-1} = E_n$ и считая определитель получаем $\det A \det A^{-1} = E_n$, что завершает доказательство 5). Пункт 6) следует из 2). □

С вычислительной точки зрения формула для определителя бесполезна – в ней $n!$ слагаемых. Но она даёт некоторые важные следствия. Например, то, что определитель есть однородный многочлен степени n с целыми коэффициентами от элементов матрицы (следовательно, понятие определителя можно ввести над любым коммутативным кольцом по этой формуле).

Примеры:
Приведём пример, вычислив определитель Вандермонда.

$$\det\begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix} = \prod_{i>j}(\lambda_i - \lambda_j).$$

Мы стартовали с понятия объёма параллелепипеда в $\mathbb{R}^n$. Отображение $A \to |\det A|$ удовлетворяет всем предполагаемым свойствам объёма. Может ли тем не менее объём параллелепипеда считаться по-другому?
Нет. Так как мы знаем, что модуль определителя и объём одинаково изменяются при элементарных преобразованиях. Точнее:

Лемма 15. Пусть дано отображение $\mathrm{Vol}\colon M_n(\mathbb{R}) \to \mathbb{R}$, такое что
1) $\mathrm{Vol}(E_n) = 1$
2) $\mathrm{Vol}(\dots, u + \lambda v, \dots, v, \dots) = \mathrm{Vol}(\dots, u, \dots, v, \dots)$
3) $\mathrm{Vol}(\dots, \lambda v, \dots) = |\lambda| \mathrm{Vol}(\dots, v, \dots)$
Тогда $\mathrm{Vol}(A) = |\det A|$.

Доказательство. Покажем, что эти свойства однозначно позволяют вычислить $\mathrm{Vol}(A)$. Так как отображение $A \to |\det A|$ тоже им удовлетворяет, то это гарантирует их совпадение. Действительно, если матрицы A вырождена, то элементарными преобразованиями первого типа можно сделать так, чтобы один столбец A стал нулевым. По третьему свойству

$$\mathrm{Vol}(v_1, \dots, 0, \dots, v_n) = \mathrm{Vol}(v_1, \dots, 0 \cdot 0, \dots, v_n) = 0 \cdot \mathrm{Vol}(v_1, \dots, 0, \dots, v_n) = 0.$$

Если же A невырождена, то элементарными преобразованиями первого типа её можно привести к диагональному виду не меняя определителя. Теперь вынося по свойству 3) диагональные элементы приводим матрицу A к единичному виду. Осталось воспользоваться свойством 1). □

Теперь дадим важное в дальнейшем определение:

Определение. Пусть V – пространство. Тогда линейное отображение $L\colon V \to V$ называется оператором на пространстве V. Пусть $e_1, \dots, e_n$ базис V. Тогда матрицей оператора L в базисе e называется матрица L_e^e (базисы с разных концов взяты одинаковыми).

Это определение удобно, потому что при фиксированном базисе композиции операторов $L_1 \circ L_2$ соответствует произведение их матриц $A_1 A_2$ в базисе e. Если бы базисы были выбраны не согласовано, этого бы не было. В частности оператору L^n соответствует матрица A^n.

Определение. Пусть $L\colon V \to V$ – линейный оператор. Тогда определим $\det L = \det A$, где A – матрица оператора в каком-то базисе.

Замечание. Линейный оператор обратим тогда и только тогда, когда $\det L \neq 0$.

Это определение требует с одной стороны требует проверки корректности, а с другой пояснения. Корректность получается из соображения, что при смене базиса $e \to f$ матрица A заменяется на $C_e^f A (C_e^f)^{-1}$. Но

$$\det C_e^f A (C_e^f)^{-1} = \det C_e^f \det A (\det C_e^f)^{-1} = \det A.$$

Теперь проинтерпретируем полученное определение. Рассмотрим какую-нибудь невырожденную форму объёма ω на пространстве V. Рассмотрим следующее отношение

$$\frac{\omega(L(e_1), \dots, L(e_n))}{\omega(e_1, \dots, e_n)}.$$

Если L не обратим, то это 0. Если обратим, то $L(e_1), \dots, L(e_n)$ базис и

$$\omega(L(e_1), \dots, L(e_n)) = \omega(e_1, \dots, e_n) \det L_e^e.$$

Откуда всё выражение просто равно $\det L_e^e = \det L$. В частности не зависит ни от выбора формы объёма, ни от выбора базиса. Это означает, что определитель оператора есть коэффициент изменения объёма параллелепипеда (вне зависимости от формы объёма и выбора параллелепипеда).

Возвращаясь к вещественным числам, заметим, что понятие определителя выросло из понятия объёма и немного его переросло. А именно, для набора элементов из пространства $\mathbb{R}^n$ можно посчитать не только объём, но ещё и некоторый знак. О смысле этого знака и пойдёт сейчас речь.

Определение. Будем говорить, что два базиса пространства V над $\mathbb{R}$ одинаково ориентированы, если матрица перехода между ними имеет положительный определитель.

Замечание. Отношение одинаковой ориентированности есть отношение эквивалентности.

Определение. Выбор одного из классов эквивалентности базисов вещественного векторного пространства V называется заданием ориентации.

Определение. Пусть V – векторное пространство над $\mathbb{R}$. Будем говорить, что линейный оператор $L\colon V \to V$ сохраняет ориентацию, если $\det L > 0$ и не сохраняет, если $\det L < 0$.

Лемма 16. Сохраняющее ориентацию отображение переводит одинаково ориентированные базисы в одинаково ориентированные.

Доказательство. Пусть e – базис. Если матрица L это A, то A – это матрица перехода от Le к e, что и доказывает утверждение. □

Примеры:
1) Симметрия относительно прямой в $\mathbb{R}^2$ меняет ориентацию.
2) Поворот сохраняет ориентацию.
3) Центральная симметрия (домножение на -1) меняет или нет ориентацию в зависимости от размерности пространства.

Определение. Определим группу операторов $\mathrm{SL}(V) = \{L\colon V \to V \mid \det L = 1\}$. Если V – вещественное векторное пространство, то это операторы, которые сохраняют понятие объёма и выбор ориентации пространства.

0.7. Линейные операторы

Рассмотрим последовательность $x_{n+2} = x_{n+1} + x_n$, $x_0 = a, x_1 = b$. Как посчитать $x_{1000} =?$. Для того, чтобы воспользоваться рекуррентой, надо сделать 1000 операций. Можно ли меньше? С одной стороны вы знаете ответ – надо найти характеристический многочлен и посчитать его корни, а потом свести всё к вычислению геометрической прогрессии. Однако в этом случае для получения точного ответа придётся возиться с иррациональными корнями. Попробуем сделать по другому. Заменим наше соотношение системой

$$\begin{cases} x_{n+1} = x_n + y_n \\ y_{n+1} = x_n \end{cases}.$$

Перепишем её в следующем виде

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = A \begin{pmatrix} x_n \\ y_n \end{pmatrix}, \text{ где } A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Тогда x_{1000} это первая координата столбца $A^{1000}\begin{pmatrix} b \\ a \end{pmatrix}$. Итого достаточно просто возвести матрицу в 1000 степень. Заметим, что это частный случай общей задачи: имеется последовательность x_n из K^m удовлетворяющая соотношению $x_{n+1} = Ax_n$. Требуется найти какой-то её член. Например – дан набор состояний $s_1, \ldots, s_k$ и даны вероятности перехода между состояниями a_{ij} за один шаг. Вопрос: что произойдёт с системой после n шагов? Итак, вопрос можно ли как-то в общем виде упростить возведение матрицы в степень? Замечу, что это вопрос напрямую связан со свойством произведения матриц, или, более инвариантно, операторов на пространстве V. Ближайшей нашей темой будет разбирательство с новой структурой – пространством с оператором.

Определение. Два пространства с V_1 и V_2 с операторами L_1 и L_2 эквивалентны, если существует изоморфизм векторных пространств $C\colon V_1 \to V_2$, что $L_2 = C \circ L_1 \circ C^{-1}$.

Замечание. Две матрицы $A, B \in M_n(K)$ эквивалентны как операторы на K^n, тогда и только тогда, когда существует матрица $C \in \mathrm{GL}_n(K)$, что $A = CBC^{-1}$. Часто такие матрицы называются подобными. Эквивалентные операторы имеют подобные матрицы при любом выборе базиса.

Замечание. Два оператора эквивалентны тогда и только тогда, когда в некоторых базисах их матрицы одинаковы.

Доказательство. Пусть $L_2 = C \circ L_1 \circ C^{-1}$. Пусть $e_1, \ldots, e_n$ базис V_1, а $C(e_1), \ldots, C(e_n)$ базис V_2. В указанных базисах матрица C единичная, что даёт равенство матриц L_1 и L_2. □

Определение. Пусть V – пространство с оператором L. Пусть $U \leq V$. Тогда U называется инвариантным подпространством, если $L(U) \leq U$.

Замечание. Это условие позволяет сузить оператор L с V на U.

Посмотрим на простейший случай, когда инвариантное пространство одномерно. Пусть оно порождено вектором v. Тогда условие инвариантности перепишется как $L(v) = \lambda v$ для некоторого $\lambda \in K$.

Определение. Пусть V – пространство с оператором L. Тогда вектор $0 \neq v \in V$ называется собственным вектором с собственным числом λ относительно оператора L, если $Lv = \lambda v$.

Примеры:
1) Рассмотрим пространство последовательностей и оператор сдвига $S(x)_n = x_{n+1}$. Тогда собственный вектор – это геометрическая прогрессия.
2) Рассмотрим тот же контекст. Тогда несложно увидеть, что подпространство последовательностей, удовлетворяющих линейному рекуррентному соотношению с постоянными коэффициентами является инвариантным относительно оператора сдвига.
3) Рассмотрим алгебру $K[x]/p(x)q(x)$. Тогда подпространство многочленов делящихся на $p(x)$ является инвариантным относительно оператора $f(x) \to xf(x)$ домножения на x.
4) Подпространство многочленов степени меньшей или равной n инвариантно относительно оператора дифференцирования.
5) Если $p(x)$ многочлен, то $\operatorname{Ker} p(L)$ инвариантно относительно L.
6) Пусть $v \in V$. Тогда $V' = \langle v, Lv, L^2v, \ldots\rangle$ является инвариантным пространством, порождённым v. Такое пространство называется циклическим.

Лемма 17. Пусть $U \leq V$ – подпространство, а $L\colon V \to V$ – линейный оператор. Тогда U инвариантно относительно L тогда и только тогда, когда в базисе $e_1, \ldots, e_k, e_{k+1}, \ldots, e_n$, где $e_1, \ldots, e_k$ базис U матрица оператора имеет блочно диагональный вид

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

Доказательство. Образ $L(e_i)$ при $i \leq k$ лежит в U и раскладывается по базису U, то есть по первым k векторам. □

Замечание. У нас снова всплыли блочные матрицы и на этот раз нам необходимо обсудить как перемножаются матрицы в таком виде. Общая формулировка выглядит так. Если есть две матрицы

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \text{ и } B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

которые перемножаемы, плюс ко всему размеры A_{ik} и B_{kj} согласованы, то тогда

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

То есть матрицы перемножаются по блокам.

Как найти инвариантные подпространства? Например, собственные векторы.

Определение. Определим характеристический многочлен оператора L как $\chi_L(t) = \det(tE_n - A)$, где A – матрица L в некотором базисе.

Лемма 18. Характеристический многочлен корректно определён.

Доказательство. Пусть A матрица оператора L в базисе e, A' – в базисе f, а C матрица перехода из $e \to f$. Тогда $A' = CAC^{-1}$. Рассмотрим $\chi_A(t)$, как элемент $K(t)$. Тогда C – тоже матрица над $K(t)$ и

$$\det(tE - A') = \det(C(tE - A)C^{-1}) = \det(C)\det(tE - A)\det C^{-1} = \det(tE - A).$$

Раз эти выражения равны как элементы $K(t)$, то и как элементы $K[t]$. □

Лемма 19. Элемент $\lambda \in K$ является собственным числом оператора L тогда и только тогда, когда λ корень $\chi_L(t)$.

Доказательство. λ собственное число тогда и только тогда, когда есть ненулевой v, что $Lv = \lambda v$ тогда и только тогда, когда $(\lambda \operatorname{Id} - L)v = 0$ тогда и только тогда, когда матрица этого оператора вырождена тогда и только тогда, когда $\det(\lambda E - A) = 0$. □

Коэффициенты многочлена $\chi_L(t)$ являются инвариантами оператора L. Давайте посмотрим на них чуть внимательнее.

Замечание. $\det L = (-1)^n \chi_L(0)$

Определение. Пусть A – матрица размера n, тогда

$$\operatorname{Tr} A = \sum_{i=1}^{n} a_{ii}.$$

Замечание. След матрицы A это $-a_{n-1}$, где $\chi_A(t) = \sum a_i t^i$. В частности, можно определить след оператора L как след его матрицы.

Имеют место следующие свойства следа

Лемма 20. След обладает следующими свойствами:
1) $\operatorname{Tr} CAC^{-1} = \operatorname{Tr} A$ для обратимой матрицы C.
2) След равен сумме собственных чисел с учётом кратностей, как корней характеристического многочлена (над любым полем, где характеристический многочлен раскладывается на линейные множители).
3) $\operatorname{Tr} AB = \operatorname{Tr} BA$ для квадратных A и B.
4) $\operatorname{Tr} A = \operatorname{Tr} A^\top$.
5) $\operatorname{Tr}(A + \lambda B) = \operatorname{Tr} A + \lambda \operatorname{Tr} B$.

Доказательство. 1) следует из инвариантности. 2) Многочлен $\chi(t) = \prod(t - \lambda_i)$. Осталось раскрыть скобки. 3) Если расписать, то равенство эквивалентно

$$\sum_{i,k} a_{ik} b_{ki} = \sum_{j,l} b_{jl} a_{lj}$$

Которое верно, особенно если взять $k = j$ и $i = l$. 4),5) ясно. □

Замечание. Аналогично определитель равен произведению собственных чисел с учётом кратностей.

Определение. Кратность собственного числа λ оператора L как корня $\chi_L(t)$ называется его алгебраической кратностью.

Определение. Оператор называется диагонализуемым, если в некотором базисе V его матрица диагональна.

Лемма 21. Матрица оператора L в базисе $v_1, \ldots, v_n$ диагональна тогда и только тогда, когда все v_i – собственные вектора L. В этом случае на диагонали матрицы стоят собственные числа оператора L.

Доказательство. Если $Lv_i = \lambda_i v_i$, то в i столбце λ_i стоит на диагонали, а остальное – 0. □

Лемма 22. Пусть $v_1, \ldots, v_n$ собственные вектора L. Если при этом для каждого с.ч. λ набор всех векторов $v_{i_1} \ldots, v_{i_k}$ с одинаковыми $\lambda = \lambda_{i_1} = \cdots = \lambda_{i_k}$ линейно независим, то тогда и все $v_1, \ldots, v_n$ линейно независимы.

Доказательство. Пусть есть линейная комбинация

$$\sum c_i v_i = 0,$$

состоящая из минимального числа векторов (пусть они нумеруются от 1 до k). Тогда так же

$$0 = L\left(\sum c_i v_i\right) = \sum c_i \lambda_i v_i$$

Умножая на λ_1 и вычитая получаем

$$0 = \sum_{i=2}^{k} c_i(\lambda_1 - \lambda_i) v_i$$

Если в последней сумме есть хоть одно ненулевое слагаемое, то приходим к противоречию. Обратное бывает только если $\lambda_i = \lambda_1$ для всех i. Но в этом случае мы можем воспользоваться независимостью из условия леммы. □

Определение. Алгебраической кратностью собственного числа λ называется его кратность как корня характеристического многочлена. Геометрической кратностью λ называется размерность $\operatorname{Ker} \lambda \operatorname{Id} - L$.

Теорема 16. Пусть K – алгебраически замкнутое поле. Тогда оператор L диагонализуем в том и только том случае, когда для всякого собственного числа λ его алгебраическая кратность равна его геометрической кратности. Более того, алгебраическая кратность собственного числа не меньше его геометрической кратности.

Доказательство. Пусть L диагонализуем, то $\dim \operatorname{Ker} \lambda \operatorname{Id} - L$ совпадает с количеством λ на диагонали, что равно алгебраической кратности собственного числа.

Обратно. Сумма алгебраических кратностей k_i собственных чисел равна

$$\sum k_i = \deg \chi_L(t) = n = \dim V.$$

По условию для различных корней λ_i, можно выбрать k_i линейно независимых векторов $v_{i1}, \ldots, v_{ik_i}$. Если объединить эти наборы, то по предыдущей лемме они будут независимы. Их ровно n штук. Следовательно это базис. Если есть k_i линейно независимых собственных векторов, то их можно дополнить до базиса. У полученной матрицы оператора будет блочный вид, в первом блоке которого будет стоять матрица $\lambda_i E_{k_i}$. Следовательно характеристический многочлен делится на $(t - \lambda_i)^{k_i}$. □

Следствие 16 (Критерий диагонализуемости)**.** Пусть K – алгебраически замкнутое поле. Если характеристический многочлен не имеет кратных корней, то оператор L диагонализуем.

Доказательство. В этом случае алгебраическая кратность наименьшая возможная. □

0.8. Жорданова форма

В этом разделе я буду предполагать, что поле K алгебраически замкнуто. Ровно те же результаты можно получить, предполагая, что поле K содержит корни характеристического многочлена рассматриваемого оператора. Как я уже обмолвился в этом разделе мы получим полную классификацию операторов над алгебраически замкнутым полем K, то есть для каждого класса эквивалентности мы построим некоторую каноническую модель и научимся сравнивать оператор с его моделью.

Лемма 23. Пусть L – оператор на пространстве V, а многочлен $g(t) = p(t)q(t)$ аннулятор L, причём $(p(t), q(t)) = 1$. Тогда подпространство V раскладывается в прямую сумму

$$V = \operatorname{Ker} p(L) \oplus \operatorname{Ker} q(L).$$

Доказательство. Рассмотрим линейное разложение $1 = a(t)p(t) + b(t)q(t)$. Тогда любой вектор v представим в виде

$$v = a(L)p(L)v + b(L)q(L)v.$$

Тогда $q(L)a(L)p(L)v = 0 = p(L)b(L)q(L)v$. Следовательно, $V = \operatorname{Ker} p(L) + \operatorname{Ker} q(L)$. Покажем, что ядра пересекаются по нулю. Пусть $v \in \operatorname{Ker} p(L) \cap \operatorname{Ker} q(L)$. Тогда $v = a(L)p(L)v + b(L)q(L)v = 0$. □

Определение. Пусть $p(x)$ – неприводимый многочлен над $K[x]$. Пространство V с оператором L называется p-примарным, если существует $\alpha \in \mathbb{N}$, что p^α аннулирует L.

Определение. Пусть V – пространство с оператором L, а V' – инвариантное подпространство. Тогда определим оператор $\overline{L}$ на V/V' следующим образом:

$$\overline{L}(\overline{v}) = \overline{L(v)}.$$

Замечание. Если $p(x)$ – многочлен, а $v \in V$, то $p(\overline{L})\overline{v} = \overline{p(L)v}$.

Замечание. Если $e_1, \ldots, e_n$ базис V и $\langle e_1, \ldots, e_k \rangle$ инвариантное пространство относительно L. Если матрица L в этом базисе имеет вид

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

то C – это матрица $\overline{L}$ в базисе $\overline{e_{k+1}}, \ldots, \overline{e_n}$. И, следовательно,

$$\chi_L(t) = \chi_{L|_{V'}}(t) \cdot \chi_{\overline{L}}(t).$$

Теорема 17 (Теорема Гамильтона-Кэли)**.** Пусть L – оператор на V. Тогда $\chi_L(L) = 0$.

Доказательство. Докажем по индукции. Случай $\dim V = 1$ ясен. Шаг. Так как K алгебраически замкнуто, то у характеристического многочлена есть корень λ_1 и собственный вектор e_1. Рассмотрим фактор $V/\langle e_1 \rangle$. Для него теорема выполнена. Заметим, что

$$\chi_L(t) = (t - \lambda_1)\chi_{\overline{L}}(t).$$

Пусть $v \in V$. Тогда $\chi_{\overline{L}}(L)v = ce_1$ так как в факторе этот элемент равен 0. Но тогда

$$\chi_L(L)v = (L - \lambda_1 E)\chi_{\overline{L}}(L)v = (L - \lambda_1 E)ce_1 = 0$$

□

Определение. Матрица $k \times k$

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

называется жордановой клеткой размера k с собственным числом λ.

Замечание. Заметим, что для того, чтобы в базисе $e_1, \dots, e_n$ матрица оператора L была жордановой клеткой необходимо и достаточно, чтобы $(L - \lambda E)e_i = e_{i-1}$ для $i \geq 2$ и $(L - \lambda E)e_1 = 0$. В частности оператор $L - \lambda E$ должен быть нильпотентным.

Теорема 18. Пусть $L\colon V \to V$ — оператор на конечномерном пространстве над алгебраически замкнутым полем K. Тогда существует базис $e_1, \dots, e_n$ в котором матрица L имеет вид

$$A = \begin{pmatrix} J_{k_1}(\lambda_1) & & & \\ & J_{k_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{k_s}(\lambda_s) \end{pmatrix}.$$

Более того, такая матрица единственна с точностью до перестановки блоков. Эта матрица называется матрицей оператора в форме Жордана. Базис, в котором матрица оператора имеет такой вид называется жордановым базисом.

Доказательство. Сначала докажем единственность. Прежде всего если нам дана матрица в жордановой форме, то мы легко можем вычислить её характеристический многочлен. Он равен $\prod(t - \lambda_i)$ где λ_i – все числа на диагонали, откуда сразу становится ясно, что λ_i – собственные числа L. Более того, алгебраическая кратность k собственного числа λ равна сумме размеров клеток с этим собственным числом.

Пусть сами размеры клеток для собственного числа λ заданы как набор чисел h_i. Имеем $\sum h_i = k$. Итак, набору клеток соответствует разбиение числа k в сумму некоторого числа слагаемых. Удобно упорядочить эти слагаемые по возрастанию $h_i \geq h_{i+1}$.

Можно ли как-то лучше визуализировать себе структуру жордановой формы? Каждому разбиению числа на слагаемые однозначно соответствует картинка

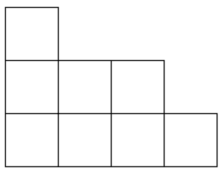

Рис. 1. $8 = 3 + 2 + 2 + 1$, соответствует одной клетке размера 3, двум клеткам размера 2, одной клетке размера 1

А именно, сопоставим каждому h_i столбик высоты h_i. Такие (обычно, правда, перевёрнутые) картинки называются диаграммами Юнга.

Как же восстановить эту картинку зная оператор L и собственное число λ?

Рассмотрим оператор $N = L - \lambda E$. Понятно, что сколько клеток с λ было у L в указанном базисе, столько же клеток с с.ч. 0 будет и у N и они будут такого же размера. Заметим, что жордановы клетки с собственным числом 0 являются нильпотентными матрицами.

Вспомним, что каждой клетке соответствует базис $v_i, Nv_i, \dots, N^{h_i}v_i$. Заметим, что эти вектора образуют базис пространства $\operatorname{Ker} N^k$. Действительно, N^k обнуляется на векторах v_i и их образах. С другой стороны N^k обратим на дополнительном слагаемом.

Пририсуем эти вектора к нашей картинке следующим образом – поместим v_i наверху соответствующего клетке столбика, Nv_i на ступень ниже и т.д. Таким образом заполним все ячейки диаграммы. При действии оператора N на диаграмму происходит следующее – все вектора съезжают на единицу вниз, кроме самых нижних, которые переходят в 0. Итого, количество ячеек в диаграмме Юнга для собственного числа λ оператора L на высоте не более s равно $\dim \operatorname{Ker}(L - \lambda E)^s$. Это позволяет однозначно восстановить разбиение числа и, следовательно, конфигурацию клеток, если мы знаем числа $\dim \operatorname{Ker}(L - \lambda E)^s$.

Точнее, число ячеек в строке уровня s равно $\dim \operatorname{Ker}(L - \lambda E)^s - \dim \operatorname{Ker}(L - \lambda E)^{s-1}$.

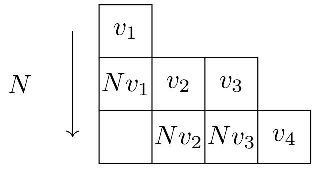

Рис. 2. $8 = 3 + 2 + 2 + 1$, расставляем базисные вектора

0.9. Евклидовы пространства

Определение. Пусть V – векторное пространство над K. Отображение $h\colon V \times V \to K$ называется билинейной формой, если
1) $\forall \lambda \in K\ \forall u, v, w \in V$ верно, что $h(u + \lambda v, w) = h(u, w) + \lambda h(v, w)$,
2) и по второй координате: $h(w, u + \lambda v) = h(w, u) + \lambda h(w, v)$.

Примеры:
1) Рассмотрим пространство K^n и определим на нём билинейную форму $h(u, v) = \sum_{i=1}^{n} x_i y_i$.
2) Рассмотрим пространство многочленов $\mathbb{R}[x]$ или пространство непрерывных функций на отрезке $[a, b]$, а так же какую-нибудь непрерывную функцию $w(x)$ и зададим билинейную форму $h(f, g) = \int_a^b f(x)g(x)w(x)dx$.
3) Рассмотрим пространство один раз непрерывно дифференцируемых функций на отрезке $C^1[a, b]$ и введём на нём билинейную форму по правилу $h(f, g) = \int_a^b f'(x)g(x)dx$.
4) Рассмотрим пространство матриц $M_n(K)$ и введём на нём билинейную форму $h(A, B) = \mathrm{Tr}(AB)$.
5) Рассмотрим конечномерную алгебру A над полем K. Тогда любой элемент $a \in A$ задаёт линейное отображение $L_a\colon A \to A$, переводящее $x \to ax$. У этого линейного отображения есть след. Для простоты обозначим его как $\mathrm{Tr}\, a$. Тогда отображение $A \times A \to K$ заданное по правилу $\mathrm{Tr}_{A/K}(u, v) = \mathrm{Tr}(uv)$ является билинейной формой. Замечу, что конструкция из предыдущего пункта не является частным случаем этой, а отличается на константу.

Как обычно, поведение объекта линейной алгебры определяется его взаимодействием с каким-либо базисом.

Определение. Пусть $e_1, \ldots, e_n$ базис V, а h – билинейная форма на V. Тогда матрица A составленная из элементов $h(e_i, e_j)$ называется матрицей билинейной формы.

Лемма 24. Пусть V – пространство с базисом $e_1, \ldots, e_n$. Тогда имеет место взаимооднозначное соответствие между билинейными формами h на V и матрицами $A \in M_n(K)$. В частности, если вектор v имеет столбец координат x, а вектор u – столбец y, то значение $h(u, v)$ можно найти по формуле $y^\top Ax$.

Лемма 25. Пусть V – пространство с билинейной формой h и базисом $e_1, \ldots, e_n$. Пусть матрица h в этом базисе – это A. Если выбрать другой базис f с матрицей перехода C, то в новом базисе матрица A будет иметь вид

$$A' = C^\top AC.$$

Наибольший интерес среди билинейных форм вызывают формы со специальными свойствами:

Определение. Билинейная форма h называется симметричной, если $h(u, v) = h(v, u)$. Форма h называется кососимметричной, если $h(u, v) = -h(v, u)$.

Лемма 26. Форма h симметрична тогда и только тогда, когда её матрица симметрична, то есть $A^\top = A$ и кососимметрична, если $A^\top = -A$.

Перейдём к основному определению этого параграфа.

Определение. Векторное пространство V над $\mathbb{R}$ вместе с заданной на нём положительно определённой симметричной билинейной формой $\langle \cdot , \cdot \rangle$ называется евклидовым пространством. Форма $\langle \cdot , \cdot \rangle$ называется скалярным произведением.

Определение. Определим норму на евклидовом пространстве как $\|v\| = \sqrt{\langle v, v \rangle}$. Норма задаёт расстояние по правилу $\rho(u, v) = \|u - v\|$. Будем говорить, что два вектора ортогональны, если $\langle u, v \rangle = 0$.

Лемма 27 (Неравенство Коши-Буняковского)**.** В евклидовом пространстве выполнено неравенство

$$\langle u, v \rangle \leq \|u\|\|v\|.$$

Доказательство. Квадратный трёхчлен $\langle u, u \rangle + 2\lambda \langle u, v \rangle + \lambda^2 \langle v, v \rangle = \langle u + \lambda v, u + \lambda v \rangle \geq 0$ всегда положителен. Значит у него нет корней, то есть дискриминант отрицателен. То есть

$$4 \langle u, v \rangle^2 \leq 4||u||^2||v||^2.$$

□

Лемма 28. Введённая норма действительно является нормой.

Доказательство. Необходимо показать неравенство $||u+v|| \leq ||u|| + ||v||$. Оно эквивалентно

$$||u+v||^2 \leq ||u||^2 + ||v||^2 + 2||u||||v||$$

Расписывая левую часть получаем эквивалентное

$$||u||^2 + ||v||^2 + 2\langle u, v\rangle \leq ||u||^2 + ||v||^2 + 2||u||||v||.$$

Сокращая справа и слева приходим к уже известному неравенству. $\square$

Лемма 29. Пусть V – евклидово пространство. Тогда для всякого подпространства U имеет место ортогональное разложение $V = U \oplus U^{\perp}$. Если есть такое разложение, то оператор проекции на U называется ортогональной проекцией.

Доказательство. Положительно определённая форма невырождена. Ограничение положительно определённой формы на любое подпространство положительно определено. $\square$

Так же полезно будет проинтерпретировать геометрическое понятие угла. Прежде всего разберёмся с углом между двумя векторами.

Определение. Пусть $x, y \neq 0$ два вектора в V. Если V – евклидово, то углом между ними называется такое число $0 \leq \varphi \leq \pi$, что

$$\cos\varphi = \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

Угол между векторами равен $\frac{\pi}{2}$ тогда и только тогда, когда $\langle x, y\rangle = 0$, то есть когда векторы ортогональны.

0.10. Ортогонализация Грама-Шмидта

В случае евклидовых пространств ограничение положительно определённой формы на любое подпространство невырождено и, таким образом, нахождение базиса, в котором форма имеет канонический вид несколько облегчается. Это приводит к тому, что мы можем уточнить сам результат и немного изменить алгоритм нахождения подходящего базиса.

Итак пусть дан набор векторов $e_1, \ldots, e_n$ евклидового пространства V. Ортогонализацией набора $e_1, \ldots, e_n$ называется новый набор векторов $f_1, \ldots, f_n$ такой, что
1) $f_i \perp f_j$, если $i \neq j$
2) $\|f_i\| = 1$.
3) $\forall 1 \leq k \leq n$ $\langle e_1, \ldots, e_k\rangle = \langle f_1, \ldots, f_k\rangle$

Определение. Набор векторов со свойством 2) называется нормированным. со свойствами 1),2) – ортонормированным. Если этот набор базис, то он называется ортогональным или ортонормированным базисом соответственно.

Замечание. Заметим, что если мы нашли набор ненулевых векторов со свойствами 1) и 3), то несложно сделать из него нормированный набор, взяв вектора $\frac{f_i}{\|f_i\|}$.

Теорема 19. Пусть V – евклидово пространство. Задача ортогонализации разрешима для линейно независимого набора векторов из V.

Доказательство. Перейдём к решению задачи добиваясь только условий 1) и 3). Будем последовательно искать вектора f_i в виде $f_i = e_i + \lambda_1 f_1 + \cdots + \lambda_{i-1} f_{i-1}$. Этот подход приводит к ответу

$$f_i = e_i - \sum_{j<i} \frac{\langle f_j, e_i\rangle}{\langle f_j, f_j\rangle} f_j.$$

Так как вектора линейно независимы, то $f_i \neq 0$. Это означает, что можно поделить на его норму и добиться нормированности. $\square$

Процесс ортогонализации позволяет строить ортогональный базис для различных подпространств.

Следствие 17. В евклидовом и унитарном пространстве любой ортонормированный набор векторов можно дополнить до ортонормированного базиса.

Следствие 18. Если U подпространство в унитарном пространстве V, то $U^{\perp}$ есть подпространство V и $V = U \oplus U^{\perp}$.

Доказательство. Пусть $e_1, \dots, e_k$ – ортонормированный базис U, а $e_{k+1}, \dots, e_n$ – его дополнение до ортонормированного базиса всего пространства. По определению $e_{k+1}, \dots, e_n \in U^\perp$, откуда $\dim U^\perp \geq n-k$. Покажем, что $U \cap U^\perp = \{0\}$. Это одновременно покажет нам, что $\dim U^\perp = n - k$, так как иначе из формулы Грассмана будет следовать, что пересечение ненулевое. Если вектор лежит в $U \cap U^\perp$, то $\langle x, x \rangle = 0$, что противоречит положительной определённости. □

Лемма 30 (Теорема Пифагора)**.** Пусть $x = \sum c_i e_i \in V$, где e_i это ортонормированный базис V. Тогда $\|x\|^2 = \sum |c_i|^2$.

Доказательство. Раскрыть скобки в выражении $\langle x, x \rangle$. □

Лемма 31. (Нахождение координат в ортогональном базисе) Пусть набор e_i — ортогональный базис V. Тогда для любого $x \in V$

$$x = \sum \frac{\langle e_i, x \rangle}{\langle e_i, e_i \rangle} e_i \text{ и следовательно } \|x\|^2 = \sum \frac{|\langle x, e_i \rangle|^2}{\langle e_i, e_i \rangle}.$$

В случае нормированного базиса формула упрощается – исчезает знаменатель.

Доказательство. Пусть $x = \sum c_i e_i$. Тогда $\langle e_i, x \rangle = c_i \langle e_i, e_i \rangle$, что и требовалось. □

Следствие 19. Пусть $e_1, \dots, e_k$ — ортогональный базис U – подпространства V. Тогда для проекции x на U имеет место формула:

$$pr_U x = \sum \frac{\langle x, e_i \rangle}{\langle e_i, e_i \rangle} e_i.$$

Так же, справедливо, что

$$||pr_{U^\perp} x||^2 + ||pr_U x||^2 = ||x||^2.$$

Доказательство. Рассмотрим базис $e_1, \dots, e_k$ и дополним его векторами $e_{k+1}, \dots, e_n$ до ортогонального базиса всего V. Тогда

$$x = \sum_{i=1}^{k} \frac{\langle e_i, x \rangle}{\langle e_i, e_i \rangle} e_i + \sum_{i=k+1}^{n} \frac{\langle e_i, x \rangle}{\langle e_i, e_i \rangle} e_i$$

Первая часть суммы лежит в U, а вторая в ортогональном дополнении. По единственности такого разложения получаем требуемое. Осталось применить теорему Пифагора. □

Для чего может пригодиться понятие ортогональной проекции? Прежде всего для определения расстояния. Расстояния между подмножествами.

Определение. Пусть A и B подмножества метрического пространства. Тогда расстоянием $\rho(A, B)$ положим равным

$$\rho(A, B) = \inf_{\substack{x \in A \\ y \in B}} \rho(x, y).$$

Наша задача — научиться считать расстояние между аффинными подпространствами A_1 и A_2 в евклидовом пространстве. Попробуем это сделать. Представим $A_1 = L_1 + x$ и $A_2 = L_2 + y$.

Разберём сначала случай расстояния от точки до линейного подпространства.

Теорема 20. Пусть $U \leq V$ подпространство, $x \in V$. Тогда расстояние $\rho(x, U)$ достигается на проекции $pr_U(x)$ и равно $\|x - pr_U(x)\| = \|pr_{U^\perp}(x)\|$.

Доказательство. Рассмотрим $u \in U$, и представим $x = y + z$, где $pr_U x = y \in U$, $pr_{U^\perp} x = z \in U^\perp$. Тогда

$$\rho(x, u)^2 = ||x - u||^2 = ||y - u||^2 + ||z||^2 \geq ||z||^2 = ||pr_{U^\perp} x||^2.$$

С другой стороны равенство достигается при $u = y = pr_U x$. □

Замечание. Если размерность $U^\perp$ мала, то может быть легче найти проекцию на $U^\perp$, найдя ортогональный базис $U^\perp$.

Всё это немедленно приводит к решению общей задачи.

Следствие 20. Пусть $A_1 = L_1 + x$ и $A_2 = L_2 + y$ – аффинные подпространства. Тогда $\rho(A_1, A_2) = \rho(y - x, L_1 + L_2)$. То есть задача сводится к ранее разобранной.

Доказательство.

$$\inf_{\substack{u+x \in L_1 + x \\ v+y \in L_2 + y}} ||u + x - v - y|| = \inf_{u - v \in L_1 + L_2} ||u - v - (y - x)|| = \inf_{u \in L_1 + L_2} ||u - (y - x)||.$$

□

Мы с вами ввели понятие объёма параллелепипеда с помощью определителя и показали, что оно однозначно определено с точностью до множителя, который фиксируется некоторой нормировкой. Вопрос – даёт ли структура евклидового пространства возможность зафиксировать эту нормировку? Ответ да, с точностью до знака, то есть, до ориентации пространства.

Определение. Пусть V – евклидово или унитарное пространство. Тогда будем говорить, что форма объёма согласована с метрикой, если объём параллелепипеда, натянутого на некоторый ортогональный базис, равен 1.

Определение. Пусть $e_1, \dots, e_k$ набор векторов V. Тогда матрицей Грамма называется матрица

$$G(e_1, \dots, e_k)_{ij} = \langle e_i, e_j \rangle .$$

Матрица Грамма отличается от матрицы скалярного произведения как би(полутора)линейной формы, только тем, что даёт некоторый ответ в случае зависимого набора векторов.

Лемма 32. Форма объёма принимающая значение 1 на ортогональном базисе принимает значения ± 1 на всех остальных ортонормированных базисах и более того, имеет место формула

$$(\mathrm{Vol}(v_1, \dots, v_n))^2 = \det G(v_1, \dots, v_n)$$

Доказательство. Пусть $e_1, \dots, e_n$ ортогональный базис, в котором форма объёма принимает значение 1. Тогда $\mathrm{Vol} = \mathrm{Vol}_e$. Обозначим за A матрицу $e(v_1), \dots, e(v_n)$. Тогда

$$\mathrm{Vol}(v_1, \dots, v_n)^2 = \det^2(A) = \det A^\top \det A = \det A^\top A = \det G(v_1, \dots, v_n)$$

В частности, если v_i ортонормирован, то $G(v_1, \dots, v_n) = E_n$ и $\mathrm{Vol}(v_1, \dots, v_n)^2 = 1$. □

Замечание. Определитель матрицы Грамма обнуляется тогда и только тогда, когда вектора v_i линейно зависимы.

Следствие 21. Форма объёма, согласованная с метрикой на евклидовом пространстве единственна с точностью до знака. А понятие объёма параллелепипеда – единственно.

0.11. Ортогональные и унитарные операторы

Основным отличием структуры евклидового и унитарного пространства от просто векторного пространства является понятие расстояние и поэтому некоторое время мы посвятим преобразованиям, это расстояние сохраняющим.

Определение. Пусть X и Y – метрические пространства. Тогда отображение $f\colon X \to Y$ называется изометрическим вложением, если для всех $x_1, x_2 \in X$ $\rho(x_1, x_2) = \rho(f(x_1), f(x_2))$. Если к тому же f – биекция, то f называют изометрией.

Замечание. Изометрическое вложение всегда инъективно, так как $f(x) = f(y)$ тогда и только тогда, когда $\rho(f(x), f(y)) = 0$ что происходит только если $\rho(x, y) = 0$, то есть $x = y$.

Определение. Пусть V — евклидово пространство. Ортогональным оператором на V называется такой линейный оператор $L\colon V \to V$, что $||Lx|| = ||x||$.

Теорема 21. Пусть $L\colon V \to V$ – оператор на евклидовом пространстве. Следующие условия эквивалентны:
1) L – ортогональный оператор.
2) $\langle Lx, Ly \rangle = \langle x, y \rangle$ для всех $x, y \in V$.
3) L переводит любой ортонормированный базис в ортонормированный базис.
4) В любом ортонормированном базисе A – матрица L – удовлетворяет условию $A^\top A = E_n$.
5) L переводит некоторый ортонормированный базис в ортонормированный базис.
6) В некотором ортонормированном базисе A – матрица L – удовлетворяет условию $A^\top A = E_n$.

Доказательство. Переход $1 \to 2$ верен, так как из равенства квадратичных форм следует равенство билинейных форм. Переход $2 \to 3$ ясен.

Покажем, что $3 \leftrightarrow 4$ и $5 \leftrightarrow 6$. Для этого покажем, что если фиксировать ортонормированный базис $e_1, \dots, e_n$, то $L(e_i)$ ортонормирован тогда и только тогда, когда A – матрица L – удовлетворяет условию $A^\top A = E_n$. Для этого необходимо и достаточно заметить, что

$$\langle Le_i, Le_j \rangle = (A^\top A)_{ij}.$$

Покажем это равенство. Имеем, что v_j – j-ый столбец A – составлен из координат Le_j. i-ая строка $A^\top$ равна тогда $v_i^\top$. Тогда

$$\langle Le_i, Le_j \rangle = v_i^\top E_n v_j = v_i^\top v_j = (A^\top A)_{ij}.$$

Во втором равенстве E_n играет роль матрицы Грама ортонормированного базиса $e_1, \dots, e_n$. Итак $3 \leftrightarrow 4$ и $5 \leftrightarrow 6$. Из 3) очевидно следует 5).

Покажем 6 → 1. Пусть $e_1, \ldots, e_n$ ортонормированный базис, в котором имеет место свойство $A^\top A = E_n$. Тогда

$$||Lx||^2 = e(x)^\top A^\top A e(x) = e(x)^\top e(x) = ||x||^2.$$

Если отображение L сохраняет норму, то оно сохраняет и расстояние, следовательно, получившееся отображение изометрия из V в себя. Оно является автоморфизмом V. С другой стороны, в определении мы видим равенство двух квадратичных форм $\|Lx\|^2 = \|x\|^2$, значит равны соответствующие симметрические билинейные формы $\langle Lx, Ly\rangle = \langle x, y\rangle$. Расписав это равенство в ортонормированном базисе получаем для B – матрицы ортогонального оператора L соотношение $B^\top B = E$. □

Следствие 22. В частности, ортогональный оператор L сохраняет углы между векторами.

Следствие 23. Пусть $e_1, \ldots, e_n$ — ортонормированный базис V. Линейный оператор L, который в базисе e_i имеет матрицу, составленную из столбцов $v_1, \ldots, v_n$, является ортогональным тогда и только тогда, когда $v_1, \ldots, v_n$ — ортонормированный базис $\mathbb{R}^n$.

Доказательство. Пусть B – матрица L – составлена из столбцов v_i в ортогональном базисе e. Тогда соотношение $B^\top B = E$ эквивалентно ортогональности и нормированности v_i. □

Замечание. Множество всех ортогональных матриц размера n будем будем обозначать $\mathrm{O}_n(\mathbb{R})$. Заметим, что определитель ортогональной матрицы либо плюс, либо минус единица. Определим множество $\mathrm{SO}_n(\mathbb{R}) \leq \mathrm{O}_n(\mathbb{R})$ – специальную ортогональную группу состоящую из матриц

$$\mathrm{SO}_n(\mathbb{R}) = \{A \in \mathrm{O}_n(\mathbb{R}) \mid \det A = 1\}.$$

сохраняющих ориентацию. Её ещё называют группой вращений.

0.12. Сопряжённые операторы

Геометрия евклидовых и унитарных пространств позволяет сводить определённые вопросы про билинейные формы к вопросам про операторы и наоборот. А именно всякому оператору L соответствует билинейная (полуторалинейная) форма $\langle x, Ly\rangle$. Или ещё – форма $\langle Lx, y\rangle$ тоже билинейная (полуторалинейная). Посмотрим на эти соответствия поближе.

Лемма 33. Пусть L – линейный оператор на V, тогда существует единственный оператор L^*, что

$$\langle u, Lv\rangle = \langle L^* u, v\rangle.$$

Точнее, если $e_1, \ldots, e_n$ ортонормированный базис и A – это матрица L, то матрица L^* будет равна $\overline{A}^\top$.

Доказательство. Распишем равенство $\langle u, Lv\rangle = \langle L^* u, v\rangle$. в ортогональной системе координат. Получим

$$x^\top A y = (Bx)^\top y = x^\top B^\top y,$$

что равносильно равенству матриц $B^\top = A$ или $B = A^\top$. □

Определение. Пусть L — оператор на евклидовом пространстве V. Сопряжённым оператором к L называется единственный такой оператор $L^* \colon V \to V$, что $\langle L^* x, y\rangle = \langle x, Ly\rangle$ для всех $x, y \in V$.

Следствие 24. Сопряжённый оператор к оператору L существует и единственен. Более того, если задан ортонормированный базис $e_1, \ldots, e_n$ и A – матрица L, то матрица L^* есть $\overline{A}^\top$.

Лемма 34 (Общие свойства)**.** $(L+T)^* = L^* + T^*$
$(LT)^* = T^* L^*$
$(\lambda L)^* = \overline{\lambda} L^*$.
$(L^{-1})^* = (L^*)^{-1}$.
$L^{**} = L$.

Доказательство. Фиксируем ортонормированный базис. Тогда все свойства следуют из свойств транспонирования и сопряжения ($\overline{AB} = \overline{A}\,\overline{B}$). Однако их можно показать и из определения сопряжённого оператора. Например,

$$\langle x, ABy\rangle = \langle A^* x, By\rangle = \langle B^* A^* x, y\rangle,$$

откуда видно, что $(AB)^* = B^* A^*$. □

Определение. Пусть L оператор на евклидовом или унитарном пространстве V называется самосопряжённым, если $L^* = L$.

Определение. Пусть L оператор на евклидовом или унитарном пространстве V называется кососимметричным (косоэрмитовым), если $L^* = -L$.

Замечание. Пусть $e_1, \dots e_n$ ортонормированный базис, тогда оператор L самосопряжён тогда и только тогда, когда его матрица в этом базисе удовлетворяет условию $L^\top = L$.

Доказательство. Равенство операторов равносильно равенству их матриц в ортонормированном базисе. □

Примеры:
0) Любая симметричная матрица $A = A^\top$ задаёт самосопряжённый оператор на $\mathbb{R}^n$ относительно стандартного скалярного произведения.
1) Условие ортогональности оператора можно переписать в виде $L^*L = 1$ или, что эквивалентно, $L^* = L^{-1}$. Таким образом, сопряжённый оператор к ортогональному – это обратный оператор.
2) Рассмотрим оператор дифференцирования на пространстве $C^\infty_{per=1}(\mathbb{R})$ – бесконечно дифференцируемых функций, с периодом 1. Тогда сопряжённый оператор к оператору дифференцирования $f \to \frac{df}{dx}$ есть оператор $-\frac{df}{dx}$.
3) Пусть $\mathbb{R}[x]$ пространство многочленов и $g(x,y) \in \mathbb{R}[x,y]$ – многочлен. Тогда сопряжённый к оператору $f \to \int_a^b f(y)g(x,y)dy$ это оператор $f \to \int_a^b f(x)g(x,y)dx$.

0.13. Спектральные теоремы

Как всегда при обсуждении линейных операторов разумно задать вопрос про их собственные числа. Наша цель – разобраться со спектральными операторами.

Теорема 22. Оператор L на унитарном пространстве V нормален тогда и только тогда, когда существует ортонормированный базис $e_1, \dots, e_n$ в котором матрица L диагональна.

Доказательство. Доказательство идёт индукцией по размерности. Если оператор нормален, то у L и L^* есть общий собственный вектор v_1, так как они коммутируют.

Возьмём к нему ортогональное дополнение $U = \langle v_1 \rangle^\perp$. Это будет инвариантное подпространство для L и L^*. При этом ограничение L^* на U – это сопряжённый к $L|_U$. По индукции это даёт ортонормированный базис для L на U, а вместе с v_1 базис из собственных векторов на всём V. □

Прежде чем перейти к вещественному случаю покажем несколько лемм.

Лемма 35. Пусть L нормальный оператор на унитарном. Тогда существует комплексный или многочлен $p(x)$, что $L^* = p(L)$.

Доказательство. Рассмотрим унитарный случай. Тогда существует ортонормированный базис в котором матрица L диагональна, а на диагонали стоят числа $\lambda_1, \dots, \lambda_n$. Тогда матрица L^* тоже диагональна, но на диагонали стоят числа $\overline{\lambda_i}$. Найдём комплексный многочлен $p(x)$, что $p(\lambda_i) = \overline{\lambda_i}$. Тогда $p(L) = L^*$. □

Следствие 25. Пусть $U \leq V$ подпространство унитарного пространства V на котором действует нормальный оператор L. Тогда если U инвариантно относительно L, то U инвариантно относительно L^* и, следовательно, $U^\perp$ инвариантно относительно L и U инвариантно относительно L^*.

Доказательство. Если U инвариантно относительно L, то U инвариантно относительно $p(L) = L^*$. Тогда по лемме **??** $U^\perp$ инвариантно относительно $L^{**} = L$ и по той же лемме $U^{\perp\perp}$ инвариантно относительно L^*. □

Следствие 26. Пусть L – нормальный оператор на евклидовом пространстве V и $U \leq V$ инвариантно относительно L. Тогда $U^\perp$ инвариантно относительно L.

Доказательство. Пусть $u_1, \dots, u_k$ ортонормированный базис U. Дополним его до ортонормированного базиса V векторами $u_{k+1}, \dots u_n$. Пусть A матрица L в этом базисе. Тогда $AA^\top = A^\top A$ и A задаёт нормальный оператор $\mathbb{C}^n \to \mathbb{C}^n$. Так как подпространство $\langle e_1, \dots, e_k \rangle$ инвариантно относительно A, то и $e_{k+1}, \dots e_n$ инвариантно относительно A, то есть элементы $a_{ij} = 0$ при $j > k$ и $i \leq k$. Но это значит, что $\langle e_{k+1}, \dots, e_n \rangle$ инвариантно относительно A как вещественное пространство, то есть $\langle u_{k+1}, \dots, u_n \rangle$ инвариантно относительно L, что и требовалось. □

Замечание. Ограничение нормального оператора на инвариантное подпространство нормально.

Лемма 36. Пусть A вещественная квадратная матрица размера n. Тогда если $v \in \mathbb{C}^n$ собственный вектор A с собственным числом λ, то $\overline{v}$ собственный вектор A с собственным числом $\overline{\lambda}$.

Доказательство. Действительно
$$\overline{\lambda}\overline{x} = \overline{\lambda x} = \overline{Ax} = \overline{A}\overline{x} = A\overline{x}.$$
□

Теорема 23. Оператор L на евклидовом пространстве V нормален тогда и только тогда, когда существует ортонормированный базис в котором его матрица A блочно-диагональная, при этом блоки имеют или размер 1×1 или 2×2 вида
$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$
Такое представление единственно с точностью до порядка блоков.

Доказательство. Прежде всего покажем единственность блочной структуры. Пусть e_i – ортонормированный базис в котором L имеет матрицу в указанном виде. Тогда блочное представление матрицы даёт возможность посчитать характеристический многочлен. Клетка 1×1 даёт множитель $t-\lambda$, то есть соответствует вещественному собственному числу. Клетке 2×2 соответствует множитель в характеристическом многочлене $t^2 - 2at + a^2 + b^2$. Этот многочлен раскладывается на множители
$$t^2 - 2at + a^2 + b^2 = (t - a - bi)(t - a + bi).$$
Если $b \neq 0$, то корни этого многочлена не вещественны, а если равно 0, то мы попадаем в уже разобранную диагональную ситуацию. Тогда количество клеток вида
$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$
с $b \neq 0$ равно кратности корня $a + bi$ у характеристического многочлена.

Перейдём к доказательству существования. Индукция по размерности V. База индукции при $\dim V = 1$ тривиальна. Покажем переход. Рассмотрим два случая. Первый состоит в том, что у L есть собственный вектор $e_1 \in V$. Тогда перейдём к ортогональному дополнению $\langle e_1\rangle^{\perp}$ и воспользуемся индукционным предположением.

Во втором случае рассмотрим какой-то ортонормированный базис e пространства V. Тогда A – матрица L – вещественна и удовлетворяет соотношению $\overline{A}^{\top} A = A\overline{A}^{\top}$, и, следовательно, задаёт нормальный оператор $\mathbb{C}^n \to \mathbb{C}^n$. Пусть $v = e_1 + ie_2$ и $\overline{v} = e_1 - ie_2$ собственные вектора для собственных чисел $\lambda = a + bi$ и $\overline{\lambda} = a - bi$. Тогда
$$\langle e_1, e_2\rangle = \frac{1}{4}\left\langle v + \overline{v}, -i(v - \overline{v})\right\rangle = \frac{1}{4}(||v||^2 - ||\overline{v}||^2) = 0.$$
Покажем, что норма e_i равна $\frac{1}{\sqrt{2}}$.
$$||e_1||^2 = \langle e_1, e_1\rangle = \frac{1}{4}\left\langle v + \overline{v}, v + \overline{v}\right\rangle = \frac{1}{4}(||v||^2 + ||\overline{v}||^2) = \frac{1}{2}.$$

Итого, вектора $\sqrt{2}e_1, \sqrt{2}e_2$ вещественны независимы ортогональны и нормированы над $\mathbb{C}$ и следовательно над $\mathbb{R}$. В базисе e_1, e_2 матрица A действует как
$$e_1 \to \frac{1}{2}(\lambda v + \overline{\lambda}\overline{v}) = ae_1 - be_2$$
$$e_2 \to \frac{1}{2i}(\lambda v - \overline{\lambda}\overline{v}) = \frac{1}{2i}(2ibe_1 + 2iae_2)$$
Разумеется, она действует так же и в базисе $\sqrt{2}e_1, \sqrt{2}e_2$. Откуда получаем, что пространство $\langle e_1, e_2\rangle$ инвариантно относительно A над $\mathbb{R}$, и в базисе $\sqrt{2}e_1, \sqrt{2}e_2$ матрица A действует как
$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$
Осталось применить индукционное предположение к ортогональному дополнению $\langle e_1, e_2\rangle$. □

Теперь можно легко получить характеризацию самосопряжённых, унитарных, вещественных ортогональных, и кососимметрических операторов.

Теорема 24. Пусть L – оператор в евклидовом (унитарном) пространстве V. Тогда L – самосопряжённый тогда и только тогда, когда существует ортонормированный базис V состоящий из собственных векторов оператора L и все собственные числа L – вещественны.

Доказательство. Пусть L оператор на унитарном пространстве. Возьмём ортонормированный базис из его собственных векторов и распишем условие самосопряжённости. Оно означает, что $\overline{A}^{\top} = A$. Но A диагональна и на диагонали стоят собственные числа. Итого на них получается уравнение $\overline{\lambda} = \lambda$, что гарантирует их вещественность.

Пусть теперь L оператор на евклидовом пространстве. Тогда есть ортонормированный базис, в котором матрица L составлена из блоков 1×1 или 2×2 вида
$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$
Но матрица L в ортогональном базисе должна быть симметрична. Отсюда $b = -b$, то есть $b = 0$, что и требовалось. □

Задача 1. Докажите спектральную теорему в случае самосопряжённого оператора напрямую.

Теорема 25. Оператор L – унитарный тогда и только тогда, когда его собственные числа по модулю равны 1 и существует ортонормированный базис из собственных векторов.

Доказательство. Рассмотрим ортонормированный базис из собственных векторов e. Матрица A оператора L в этом базисе диагональна и удовлетворяет соотношению $\overline{A}^\top A = E_n$. Для собственных чисел A это означает, что

$$|\lambda|^2 = \overline{\lambda}\lambda = 1.$$

□

Теорема 26. Оператор L на евклидовом пространстве V ортогональный тогда и только тогда, когда существует ортонормированный базис в котором матрица A блочно-диагональная, при этом блоки имеют размер 1 и состоят из ± 1 или имеют размера 2 и имеют вид

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Доказательство. По вещественной версии спектральной теоремы для самосопряжённых операторов есть ортонормированный базис в котором матрица L состоит из блоков 1×1 или 2×2 вида

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Но матрица L в ортогональном базисе должна быть ортогональной и, следовательно, вещественной и унитарной. Откуда получаем, что её собственные числа по модулю равны 1. Если эти числа вещественные, то они равны ± 1. Если же они не вещественные, то имеют вид $a + bi = \cos\varphi + i\sin\varphi$, что даёт необходимый вид для a и b.

Обратно, оператор L ортогонален так как его матрица ортогональна в ортонормированном базисе по условию теоремы. □

Задача 2. Покажите аналогичные теоремы для кососимметричных и косоэрмитовых операторов.

Следствие 27. Пусть L – ортогональный оператор на $\mathbb{R}^3$. Если $\det L = -1$, то в некотором ортогональном базисе матрица L имеет вид

$$\begin{pmatrix} -1 & & \\ & \cos\varphi & \sin\varphi \\ & -\sin\varphi & \cos\varphi \end{pmatrix}$$

Если же, $\det L = 1$, то в некотором ортогональном базисе матрица L имеет вид

$$\begin{pmatrix} 1 & & \\ & \cos\varphi & \sin\varphi \\ & -\sin\varphi & \cos\varphi \end{pmatrix}.$$

То есть любой элемент $\mathrm{SO}_3(\mathbb{R})$ есть вращение относительно некоторой прямой.

О матрицах из $\mathrm{SO}_n(\mathbb{R})$ мы поговорим позже, а пока приведём пример преобразований не сохраняющих ориентацию. Самый геометрически наглядный пример – это отражение относительно гиперплоскости H. Чтобы получить линейное, а не аффинное преобразование потребуем, чтобы H проходила через 0, то есть была линейным подпространством. Тогда все точки из H задаются тем, что они ортогональны некоторому вектору v (любая гиперплоскость есть ядро любого линейного функционала, а любой линейный функционал имеет вид $\langle \_, v\rangle$ для некоторого v из-за невырожденности скалярного произведения). Не умаляя общности потребуем, чтобы вектор v был нормирован. Тогда отражение относительно H действует следующим образом:

$$u \to u - 2\langle u, v\rangle v.$$

Это, очевидно, линейное преобразование. Покажем, что оно ортогонально. Для этого проще всего найти ортонормированный базис из собственных векторов и проверить, что все с.ч. равны ± 1. Первый кандидат – это v, который есть собственный вектор с собственным числом -1. Остальные вектора – это $e_1, \dots, e_{n-1}$ – ортонормированный базис H. Это вектора с собственным числом 1.

Ещё один класс примеров ортогональных операторов на $\mathbb{R}^n$ даётся матрицами перестановок

$$(P_\sigma)_{ij} = \begin{cases} 1, & \text{если } i = \sigma(j). \\ 0, & \text{иначе} \end{cases}.$$

0.14. Задачи на максимизацию

Теперь обратимся к вопросам, связанным с вещественными самосопряжёнными операторами. Для этого заметим, что с каждым самосопряжённым оператором L на евклидовом пространстве можно связать билинейную симметричную форму $\langle x, Ly\rangle$ или, что эквивалентно, квадратичную форму $\langle x, Lx\rangle$. Безусловно по квадратичной форме можно обратно восстановить оператор.

Рассмотрим один из вопросов, связанных с такой конструкцией, а именно, рассмотрим задачу о нахождении нормы линейного оператора $L\colon U \to V$ между двумя евклидовыми пространствами. Для того, чтобы найти норму необходимо найти

$$\max_{x\neq 0}\sqrt{\frac{\langle Lx, Lx\rangle}{\|x\|^2}} = \sqrt{\max_{x\neq 0}\frac{\langle L^*Lx, x\rangle}{\|x\|^2}} = \sqrt{\max_{\|x\|=1}\langle L^*Lx, x\rangle}.$$

Таким образом, нахождение нормы оператора свелось к задаче максимизации квадратичной формы на единичной сфере. Заметим, что максимум действительно достигается благодаря компактности сферы.

Оказывается, что довольно легко найти максимум или минимум квадратичной формы на сфере.

Теорема 27. Пусть V – евклидово пространство, A – самосопряжённый оператор на V, а $q(x) = \langle x, Ax\rangle$ – соответствующая квадратичная форма. Тогда

$$\max_{x\in V}\frac{q(x)}{||x||^2} = \max_{\substack{x\in V\\ ||x||=1}} q(x) = \lambda_1,$$

где λ_1 - наибольшее собственное число оператора A и достигается на собственном векторе v_1, соответствующему λ_1. Аналогично минимум равен минимальному собственному числу A.

Доказательство. Пусть $v = \sum c_i e_i$, причём $1 = ||v||^2 = \sum c_i^2$. Тогда $\langle Av, v\rangle = \sum c_i^2\lambda_i$, что меньше $\langle Ae_1, e_1\rangle = \lambda_1 = \sum \lambda_1 c_i^2$. □

Эта теорема, кроме, собственно, решения задачи, даёт геометрическую характеризацию первого собственного числа. Вопрос: можно ли аналогично охарактеризовать другие собственные числа? Ответ получается не таким простым, но, тем не менее, полезным.

Теорема 28 (Куранта-Фишера)**.** Пусть $q(x) = \langle x, Ax\rangle$. Тогда k-ое по убыванию собственное число λ_k для A есть

$$\lambda_k = \max_{\dim L = k}\min_{\substack{x\in L\\ ||x||=1}} q(x) = \min_{\dim L = n-k+1}\max_{\substack{x\in L\\ ||x||=1}} q(x).$$

Причем максимум достигается на инвариантном подпространстве, содержащем собственные вектора для $\lambda_1, \ldots, \lambda_k$.

Доказательство. Пусть U — подпространство на котором достигается максимум, причём допустим, что максимум больше λ_k. Тогда рассмотрим подпространство $W = \langle v_k, \ldots, v_n\rangle$, где v_i — собственный вектор соответствующий i-ому по убыванию собственному числу. Заметим, что $U \cap W = \{0\}$, так как на W форма принимает значения меньше или равные λ_k, а на U – строго большие. Однако $\dim W = n - k + 1$. Приходим к противоречию с подсчётом размерности пересечения. □

Введём определение:

Определение. Пусть q – квадратичная форма на евклидовом пространстве. Рассмотрим ортонормированный базис u_i пространства V. Тогда положим

$$\operatorname{Tr} q = \sum q(u_i).$$

Если в базисе u_i форме $q(x) = x^\top Ax$ соответствует симметричная матрица A, то $\operatorname{Tr} q(x) = \operatorname{Tr}(A)$.

Замечание. Определение не зависит от выбора ортонормированного базиса. Действительно, если замена координат ортогональна, то матрица q в новой системе координат имеет вид $C^\top AC = C^{-1}AC$. Осталось заметить, что след последней матрицы очевидно равен следу A.

Замечание. Если форме q соответствует самосопряжённый оператор A, то $\operatorname{Tr} q = \sum \lambda_i$, где λ_i – собственные числа A.

Следствие 28. Пусть U некоторое подпространство, а $q(x) = x^\top Ax$. Пусть собственные числа A — это λ_i, а собственные числа оператора, соответствующего $q(x)|_U$ – это μ_i упорядоченные по убыванию. Тогда

$$\lambda_{i+n-m} \le \mu_i \le \lambda_i.$$

Доказательство. Заметим, что
$$\mu_i = \max_{\substack{L\le U\\ \dim L=i}} \min_{\substack{x\in L\\ ||x||=1}} q(x),$$
что очевидно меньше, чем
$$\max_{\substack{L\le V\\ \dim L=i}} \min_{\substack{x\in L\\ ||x||=1}} q(x) = \lambda_i.$$
Неравенство в другую сторону получается из второго описания в теореме Куранта-Фишера. □

Следствие 29. Пусть $q(x) = x^\top Ax$, U — некоторое подпространство, $\dim U = k$. Пусть собственные числа A — это λ_i. Тогда
$$\operatorname{Tr} q|_U \le \sum_{i=1}^{k} \lambda_i = \operatorname{Tr} q|_{V_k},$$
где V_k подпространство натянутое на первые k собственных векторов q.

Доказательство. Нам известны неравенства на собственные числа μ_i, ограничения $q|_U$. А именно, $\mu_i \le \lambda_i$. Но тогда и
$$\sum_{i=1}^{k} \mu_i \le \sum_{i=1}^{k} \lambda_i.$$
□

0.15. Метод главных компонент

Рассмотрим следующую задачу: имеется массив данных — набор векторов $x_1, \ldots, x_s \in V$, где V – это евклидово пространство размерности n. Подразумевается, что точек очень много. Предположим, что при идеальных измерениях между координатами этих векторов есть линейные зависимости. Все отклонения от этой погрешности вызваны небольшими погрешностями. Задача состоит в том, чтобы восстановить линейную зависимость.

Переформулируем задачу геометрически. Пусть наши вектора удовлетворяют линейным условиям $Ax_i = b$ для некоторой матрицы A ранга $n-k$. Это значит, что они лежат в аффинном подпространстве $\operatorname{Ker} A + a_0$ размерности k. Если же нам даны вектора с погрешностями, то задачу таким образом можно поставить в виде: найти аффинное подпространство размерности k наиболее близкое к данным точкам $x_1, \ldots, x_s$. Понятие «наиболее близкое» требует конкретизации. Вообще говоря, тут есть выбор. Мы будем считать, что подходящее пространство $W = L + a_0$ должно давать минимум следующего выражения:
$$\sqrt{\sum_{i=1}^{s} \rho(x_i, W)^2} \to \min.$$
Совершенно ясно, что корень квадратный тут для красоты, и минимизировать нужно $\sum_{i=1}^{s} \rho(x_i, W)^2$.

Прежде всего установим, что какой бы «минимайзер» $W = L + a_0$ мы не нашли, в W всегда будет лежать среднее $\frac{1}{s}\sum_{i=1}^{s} x_i$ и, следовательно, в качестве a_0 всегда можно взять среднее.

Действительно, выпишем условие $\sum_{i=1}^{s} \rho(x_i - a_0, L)^2 = \sum ||pr_{L^\perp}(x_i - a_0)||^2$ минимально. Продифференцируем по координатам a_0. Получим
$$\sum -2pr_{L^\perp} x_i + 2pr_{L^\perp} a_0 = 0$$
Это условие означает, что проекции a_0 и среднего $\frac{1}{s}\sum x_i$ на подпространство $L^\perp$ совпадают. То есть эти две величины отличаются на элемент L. Тогда среднее лежит в W.

Итак, вычтя из всех x_i их среднее можно считать $a_0 = 0$, а все точки x_i удовлетворяют равенству $\sum_{i=1}^{s} x_i = 0$. Замечу, что это равенство нигде в дальнейшем не будет использовано. Благодаря такой замене мы свели задачу к поиску подпространства L размерности k, которое минимизирует
$$\sum_{i=1}^{s} \rho(x_i, L)^2 = \sum_{i=1}^{s} ||pr_{L^\perp}(x_i)||^2.$$

Воспользуемся тем, что $||x||^2 = ||pr_L x||^2 + ||pr_{L^\perp} x||^2$ или $||x||^2 - ||pr_L x||^2 = ||pr_{L^\perp} x||^2$. Тогда получаем, что
$$\sum_{i=1}^{s} ||pr_{L^\perp}(x_i)||^2 = \sum_{i=1}^{s} ||x_i||^2 - ||pr_L x_i||^2$$
должно быть минимально. Тогда сумма $\sum_i ||pr_L x_i||^2$ должна быть максимальна.

Для того чтобы посчитать проекцию выберем в L ортонормированный базис $u_1, \ldots, u_k$. Тогда перепишем

$$\sum_{i=1}^{s} ||pr_L x_i||^2 = \sum_{i=1}^{s}\sum_{j=1}^{k} \langle x_i, u_j\rangle^2 = \sum_{j=1}^{k}\sum_{i=1}^{s} \langle x_i, u_j\rangle^2 .$$

Внутренняя сумма теперь есть значение некоторой квадратичной формы на векторе u_j. Разберёмся какой. Рассмотрим матрицу X размера $s \times n$, чьи строки это вектора x_i $i \in \overline{1, s}$. Тогда вектор $d = Xu_j$ состоит из скалярных произведений $\langle v_i, u_j\rangle$. Тогда выражение $\langle d, d\rangle = (u_j X^\top)Xu_j = \sum_{i=1}^{s} \langle x_i, u_j\rangle^2$, то есть как раз тому, что участвует в нашей сумме. Рассмотрим симметричную матрицу $A = X^\top X$ и обозначим соответствующую ей форму за q. Тогда нам надо минимизировать выражение

$$\sum_{j=1}^{k} q(u_j).$$

Таким образом мы ищем максимум $\operatorname{Tr} q_L$ по всем подпространствам L размерности k, где форма q соответствует матрице $X^\top X$. Сформулируем теперь ответ.

Теорема 29. Пусть есть набор векторов $x_1, \ldots, x_s \in V$. Определим симметричную, положительно определённую матрицу A, как $A = X^\top X$, где строчки матрицы X – это вектора x_i. Тогда минимум по всем аффинным подпространствам V размерности k выражения $\sum_{i=1}^{s} \rho(x_i, W)^2$ достигается при $a_0 = \frac{1}{s}\sum x_i$ и $L = \langle v_1, \ldots, v_k\rangle$, где v_i — собственные вектора A, причём соответствующие собственные числа упорядочены по убыванию.

При вычислении с помощью метода главных компонент часто используют несколько другие конструкции. В частности, в методе главных компонент возникает матрица $X^\top X$ и её собственные числа. Про них поговорим подробнее.

Определение (Сингулярные значения)**.** Пусть A – линейное отображение $A\colon U \to V$ между евклидовыми пространствами. Тогда сингулярными значениями A называются числа $\sigma_i = \sqrt{d_i}$, где $d_i > 0$ – положительные собственные числа оператора $A^*A\colon V \to V$. Если же говорить на языке матриц, то для матрицы X её сингулярными значениями будут корни из собственных чисел $X^\top X$.

На самом деле мы не обсуждали определение сопряжённого линейного отображения, а только сопряжённого оператора. Напишу немного об этом.

Определение. Пусть A – линейное отображение $A\colon U \to V$ между евклидовыми пространствами. Тогда сопряжённым отображением к A, называется такое линейное отображение A^*, что $\langle A^*x, y\rangle = \langle x, Ay\rangle$ для всех $x \in V$ и $y \in U$.

Теорема 30. Сопряжённое линейное отображение единственно. Более того, если в U и V выбрать ортонормированные базисы, то матрица сопряжённого отображения в этих базисах будет равна транспонированной матрице исходного.

Доказательство. Достаточно доказать последнюю часть, чтобы показать единственность и существование. Выберем ортонормированные базисы в U и V – u_j и v_i. Обозначим матрицу A в этом базисе за X, а кандидата на A^* за Y. Тогда для равенства из определения сопряжённости необходимо и достаточно, его выполнения на базисных. Иными словами необходимо и достаточно, чтобы $\langle X^*e_i, e_j\rangle = \langle e_i, Xe_j\rangle$. Но первая часть даёт X^*_{ji}, а вторая – X_{ij}. Итого необходимо и достаточно, чтобы $X^* = X^\top$. □

Теперь обсудим важную конструкцию, проясняющую геометрический смысл сингулярных значений.

Теорема 31 (SVD разложение)**.** Пусть A – линейное отображение $A\colon U \to V$ между евклидовыми пространствами. Тогда существуют такие ортонормированный базисы U и V, что матрица A имеет вид

$$\begin{pmatrix} \sigma_1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \sigma_r & 0 \\ 0 & \ldots & 0 & 0 \end{pmatrix},$$

где r – ранг A, числа $\Sigma = \sigma_1, \ldots, \sigma_r$ её сингулярные значения. На языке матриц это означает, что для любой матрицы $X \in M_{m\times n}$ существуют матрицы L – размера m и R – размера n, что

$$X = L\Sigma R,$$

с теми же условиями на r и σ_i.

Доказательство. Рассмотрим оператор $B = A^*A$. Тогда существуют ортонормированный базис $e_1, \dots, e_n$ в котором оператор B диагонален, с неотрицательными числами на диагонали $d_1 \geq \dots \geq d_n \geq 0$. Имеем $d_i = \sigma_i^2$ для единственного положительного σ_i. Посмотрим на вектора $Ae_i \in U$. Они ортогональны. Действительно

$$\langle Ae_i, Ae_j \rangle = \langle A^*Ae_i, e_j \rangle = \langle d_i e_i, e_j \rangle,$$

что равно нулю, если $i \neq j$. В случае $i = j$ получаем $||e_i||^2 = d_i$. Возьмём

$$f_i = \frac{Ae_i}{\sqrt{d_i}}$$

и дополним этот набор до ортонормированного базиса пространства U. Итого имеем $e_1, \dots, e_n$ ортонормированный базис U и $f_1, \dots, f_m$ – ортонормированный базис V. Посмотрим матрицу A в этих базисах. По определению $Ae_i = \sqrt{d_i} f_i$. Это и даёт требуемый вид матрице оператора A Напоследок осталось решить вопрос, как выглядит матрица R. В нашей конструкции матрица R есть матрица замены из стандартного базиса в базис из собственных векторов e_i матрицы $X^\top X$. Если за C обозначить матрицу из столбцов e_i, то $R = C^{-1}$, но C ортогональна и поэтому можно написать $R = C^T$, то есть строки R – собственные вектора $X^\top X$. Часто эти вектора называют правыми сингулярными векторами X.

□

Наличие такого разложения означает, что для всякого линейного отображения можно так выбрать декартову систему координат, что в этой системе координат это отображение будет выглядеть как растяжение вдоль каких-то осей.

Задача 3. Получите аналогичное описание для L. Покажите так же, что ничего кроме Σ в аналогичном разложении получится не может.

SVD-разложение используется в практическом решении задачи из метода главных компонент и позволяет сразу найти не только пространство, но и проекцию начальных точек на него. Формализуется это так: рассмотрим матрицу X, чьи строки равны $x_i^\top$. Тогда если все её строки заменить на их проекции на оптимальное подпространство L, то получится матрица ранга k или меньше. Эта матрица будет ближайшей к исходной в смысле вот такой вот матричной нормы, называемой, нормой Фробениуса

$$||X||_F = \sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{\operatorname{Tr} X^\top X}.$$

Таким образом нахождение проекций точек можно переформулировать, как нахождение ближайшей к матрице X матрице ранга меньше или равного k. Это легко сделать, зная SVD-разложение

Теорема 32. Пусть $X \in M_{m\times n}(\mathbb{R})$. И SVD-разложение X имеет вид $X = L\Sigma R$, где на диагонали Σ стоят $\sigma_1, \dots, \sigma_r$ и нули. Тогда наилучшим приближением ранга k в смысле нормы Фробениуса к матрице X будет матрица $X^{(k)} = L\Sigma^{(k)}R$, где на диагонали $\Sigma^{(k)}$ стоят $\sigma_1, \dots, \sigma_k$ и нули.

Доказательство. Для того, чтобы найти матрицу $X^{(k)}$ необходимо спроецировать строки X на подпространство $L = \langle v_1^\top, \dots, v_k^\top \rangle$, где v_i ортонормированный базис из собственных векторов $X^\top X$. Вспомним, что строки R есть $v_i^\top$. Для того, чтобы спроецировать одну строку a на пространство $V^{(k)}$ необходимо вычислить сумму $\sum_{i=1}^k (av_i)v_i^\top$. Применив это целиком к матрице $X = L\Sigma R$ получим

$$X^{(k)} = \sum_{i=1}^k X v_i v_i^\top = L\Sigma \left(\sum_{i=1}^k R v_i v_i^\top \right).$$

Вычислим последнюю сумму. Эта сумма считает проекции строк R на L. Но первые k строк лежат в L, а остальные ортогональны L. Итого имеем

$$R^{(k)} = \sum_{i=1}^k R v_i v_i^\top = \begin{pmatrix} v_1^\top \\ \vdots \\ v_k^\top \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Осталось заметить, что $\Sigma^{(k)}R = \Sigma^{(k)}R^{(k)} = \Sigma R^{(k)}$. □

SVD-разложение используется в разных задачах, в том числе и для сжатия изображений. Для простоты рассмотрим случай квадратного $n\times n$ чёрно белого изображения. Сделаем из него вещественную матрицу X размера $n\times n$ и найдём SVD-разложение $L\Sigma K$. Тогда приближение $X^{(k)}$ задаётся $L\Sigma^{(k)}R$. Однако, как мы уже заметили, вместо матрицы R можно взять матрицу $R^{(k)}$. Аналогично вместо L можно взять $L^{(k)}$ – выкинув из L последние $n-k$ столбцов. Для хранения матрицы $\Sigma^{(k)}$ нужно k параметров, для матриц $L^{(k)}$ и $R^{(k)}$ по kn параметров. Итого нужно $2kn+k$ параметров. Однако чтобы не хранить отдельно Σ её можно домножить на L и хранить $L\Sigma$. В таком случае необходимо $2kn$ параметров. При $k<\frac{n}{2}$ это даёт эффект сжатия.

Однако, это не предел. Посмотрим, сколько параметров нужно, чтобы задать X – матрицу ранга k. Пусть главный минор размера k матрицы X не ноль (априори мы знаем, что какой-то минор такого размера не ноль). Тогда для j-ой строки матрицы, начиная с номера $j\geq k+1$ есть набор чисел $a_{1,j},\dots,a_{k,j}$, которые есть коэффициенты в линейной комбинации дающей из первых строк j-ую. Аналогично для столбцов. Такой набор данных задаётся $k^2+2k(n-k)=2kn-k^2$ параметрами. Осталось заметить, что всегда $2kn-k^2\leq n^2$ так как $0\leq n^2-2kn+k^2=(n-k)^2$. Если невырожденным оказался не главный минор, то дополнительно нужно задать $2k$ дискретных параметров задающий номера строк и столбцов невырожденного минора.

0.16. Метод Лагранжа, критерий Сильвестра

В этом разделе речь пойдёт про симметрические формы. В начале мы дадим для них альтернативное описание.

Определение. Квадратичная форма – это отображение $q\colon V\to K$, такое, что в некоторой линейной системе координат это отображение есть однородный многочлен степени 2, то есть имеет вид $\sum_{i\leq j} b_{ij}x_ix_j$. Матрицей квадратичной формы в указанной системе координат называется матрица

$$a_{ij}=\begin{cases} b_{ii}, & \text{если } i=j,\\ \frac{b_{ij}}{2}, & \text{если } i\neq j.\end{cases}.$$

Если вектор v имеет столбец координат x, то $q(v)=x^\top Ax$.

Замечание. Матрица A единственная симметричная матрица, что $q(v)=x^\top Ax$. Действительно, единственный способ решить уравнения $b_{ij}=a_{ij}+a_{ji}$ и $a_{ij}=a_{ji}$ есть $a_{ij}=a_{ji}=\frac{b_{ij}}{2}$. Таким образом, при выборе базиса возникает взаимооднозначное соответствие

$$\text{Симметричные билинейные формы} \leftrightarrow \text{Симметричные матрицы} \leftrightarrow \text{Квадратичные формы}$$

Покажем, что соответствие симметричных билинейных и квадратичных форм не зависит от базиса.

Лемма 37. Пусть h – симметричная билинейная форма на V. Тогда $q(v)=h(v,v)$ – это квадратичная форма. При этом, в любой системе координат матрица q есть A – матрица h.

Доказательство. $q(v)=h(v,v)=x^\top Ax$. Матрица A – симметричная и, следовательно, есть матрица q.

Замечание. Пусть q – квадратичная форма. Тогда форма $h(u,v)=\frac{q(u+v)-q(u)-q(v)}{2}$ – симметричная билинейная. Эта конструкция обратна к конструкции из предыдущего факта. В этом случае форма h называется поляризацией квадратичной формы q.

Определение. Квадратичная форма невырождена, если соответствующая ей симметричная билинейная форма невырождена.

Определение. Пусть h – симметричная билинейная форма на V, тогда система векторов $e_1,\dots,e_k$ называется ортогональной, если $h(e_i,e_j)=0$, при $i\neq j$. Если указанная система векторов является базисом, то такой базис называют ортогональным.

Замечание. Матрица билинейной формы в ортогональном базисе имеет диагональный вид, а выражение для квадратичной формы даёт сумму квадратов с коэффициентами $\sum\lambda_ix_i^2$.

Определение. Будем говорить, что симметрические билинейные или квадратичные формы эквивалентны, если в некоторых базисах они имеют одинаковые матрицы.

Вопрос: всегда ли можно найти ортогональный базис и насколько форма матрицы зависит от выбора ортогонального базиса?

На первый вопрос ответ положительный.

Теорема 33. Пусть V – пространство с симметричной билинейной формой h. Тогда в V существует ортогональный относительно h базис.

Доказательство. Пусть h – не ноль. Тогда существует вектор e_1, что $h(e_1, e_1) = q(e_1) \neq 0$, потому что форма q не нулевая. Действительно, если $h(e_1, e_1) = 0$ для какого-то $e_1 \neq 0$, то есть e_2, что $h(e_1, e_2) \neq 0$. Тогда либо $h(e_2, e_2) \neq 0$, либо $h(e_1+e_2, e_1+e_2) = 2h(e_1, e_2) \neq 0$. Теперь $h|_{\langle e_1 \rangle}$ невырождена и следовательно $V = \langle e_1 \rangle \oplus \langle e_1 \rangle^\perp$. Далее индукция. □

Обсудим алгоритм который стоит за этим доказательством

Нам удобнее всего будет работать с формой q и представлять её в виде однородного многочлена второй степени. После чего описанный алгоритм можно будет условно назвать выделением полного квадрата.

Пусть форма $q(x)$ в координатах имеет вид

$$q(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \dots + 2a_{1n}x_1x_n + q'(x_2, \dots, x_n).$$

Первый случай Предположим, что $a_{11} \neq 0$. Тогда представим $q(x)$ в виде, выделив полный квадрат

$$q(x) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \dots + \frac{a_{1n}}{a_{11}}x_n\right)^2 - \frac{a_{12}^2}{a_{11}}x_2^2 - 2\frac{a_{12}a_{13}}{a_{11}}x_2x_3 - \dots - \frac{a_{1n}^2}{a_{11}}x_n^2 + q'(x_2, \dots, x_n).$$

Новые переменные выглядят следующим образом:

$$\begin{aligned} y_1 &= x_1 + \frac{a_{12}}{a_{11}}x_2 + \dots + \frac{a_{1n}}{a_{11}}x_n, \\ y_2 &= x_2, \\ &\vdots \\ y_n &= x_n. \end{aligned}$$

Или в подходящую сторону

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

Видно, что кроме формы q' возникает ещё поправка, которая содержит слагаемые $\lambda x_i x_j$ $i, j \geq 2$. Таким образом мы обнулили a_{1j} и сделали первый вектор новой системы координат ортогональным остальным, как и в доказательстве. Заметим так же, что указанное преобразование над матрицей эквивалентно одновременному применению одинаковых элементарных преобразований строк и столбцов.

Второй случай. $a_{11} = 0$. Если $a_{ii} \neq 0$, то меняем первую и i-ую координаты местами и продолжаем как раньше.

Третий случай. Все $a_{ii} = 0$. Пусть $a_{12} \neq 0$. Тогда сделаем замену $x_1 = y_1 + y_2$, $x_2 = y_1 - y_2$, $y_i = x_i$, $i \geq 3$. Получим $2a_{12}$ при y_1^2 и $-2a_{12}$ при y_2^2. Теперь находимся в ситуации первого случая.

Четвёртый случай. Все $a_{ii} = 0$ и все $a_{1i} = 0$. Тогда форма не зависит от первой переменной и можно смело переходить к следующей переменной.

Определение. Пусть A – матрица. Числа $d_i = \det A_i$, где A_i – подматрица A составленная из элементов первых i строк и столбцов называются главными минорами матрицы A. Будем считать $d_0 = 1$.

Теперь можно доказать формулу

Теорема 34 (Теорема Якоби)**.** Пусть V – векторное пространство, q – квадратичная форма, A – её матрица в некотором базисе $e_1, \dots, e_n$. Пусть главные миноры d_i не равны 0. Тогда матрица A – невырожденная и может быть приведена к диагональному виду с числами $\frac{d_i}{d_{i-1}}$ на диагонали.

Доказательство. Заметим, что если мы всё время пользуемся первым случаем из алгоритма, то домножение слева на матрицу перехода $C^\top$ будет эквивалентно прибавлению строки ко всем остальным строкам матрицы A с коэффициентами $-\frac{a_{1i}}{a_{11}}$ (аналогично при домножении на C справа происходит преобразование столбцов). Заметим, что при таком преобразовании главные миноры матрицы вообще не меняются (внутри каждого минора происходят преобразования первого типа, которые не меняют определитель). Покажем, что только первый случай реализуется.

Пусть мы доказали это для шага i. На шаге $i+1$ первые i столбцов и строк содержат только ненулевые диагональные элементы $a_{11}, \dots, a_{ii}$. Тогда $d_{i+1} = a_{11} \dots a_{i+1i+1}$. Так как d_{i+1} не поменялось и, следовательно не равно нулю, то и $a_{i+1i+1} \neq 0$. Следовательно реализуется первый случай.

Теперь посмотрим, что происходит, после приведения матрицы к диагональному виду. Заметим, что для всех i $d_i = a_{11} \dots a_{ii}$. Тогда $a_{ii} = \frac{d_i}{d_{i-1}}$, что и требовалось. □

Насколько диагональный вид для квадратичной формы единственен? Сразу же можно отметить очевидный инвариант – а именно $\operatorname{rank} A$, который не меняется при замене базиса и в диагональном виде равен числу не нулей на диагонали.

Определение. Сигнатурой формы над $\mathbb{R}$ называется пара чисел (k, l) – число плюсов и число минусов в каноническом виде. Заметим, что сумма $l + k = \operatorname{rank} q$.

Следствие 30. Пусть q – квадратичная форма на вещественном векторном пространстве V. Тогда существует линейная система координат в которой форма имеет вид

$$q(x) = x_1^2 + \dots + x_k^2 - x_{k+1}^2 - \dots - x_{k+l}^2.$$

Такой вид квадратичной формы будем называть каноническим.

Определение. Квадратичная форма называется положительно определённой, если $\forall v \neq 0\ q(v) > 0$. Симметричная билинейная форма называется положительно определённой, если соответствующая форма $q(v) = h(v, v)$ положительно определена. Симметричная матрица называется положительно определённой, если соответствующая форма положительно определена.

Теорема 35. Сигнатура формы q не зависит от способа приведения формы к каноническому виду. Точнее – число k равно размерности наибольшего подпространства, ограничение формы на которое положительно определено.

Доказательство. Рассмотрим базис $e_1, \dots, e_k, e_{k+1}, \dots, e_{l+k}, \dots, e_n$, что матрица формы q диагональна, и первые k её диагональных компонент положительны, следующие l отрицательны, а остальные 0. Пусть U подпространство $\dim U \geq k + 1$, что $q|_U > 0$. Тогда исходя из подсчёта размерности $U \cap \langle e_{k+1}, \dots, e_n \rangle \neq \{0\}$. Но это приводит к противоречию, так как $q(v)$ для $0 \neq v \in U \cap \langle e_{k+1}, \dots, e_n \rangle$ выполнено, что $q(v) > 0$ и $q(v) \leq 0$ одновременно. $\square$

Следствие 31. Пусть q – форма на вещественном пространстве V размерности n. Тогда канонический вид q однозначно определяется n и её сигнатурой.

Можно ли как-то ещё найти сигнатуру не приводя форму к диагональному виду, а воспользовавшись другими знаниями? Ответ: да, можно. А именно

Следствие 32 (Критерий Сильвестра)**.** Пусть V – векторное пространство над $\mathbb{R}$, q – квадратичная форма, A – её матрица в некотором базисе $e_1, \dots, e_n$. Пусть главные миноры d_i матрицы A все не равны 0. Тогда число перемен знака в последовательности $1 = d_0, d_1, \dots, d_n$ равно числу отрицательных квадратов в каноническом виде.

Доказательство. По теореме Якоби существует система координат в которой форма имеет диагональную матрицу с числами $\lambda_i = \frac{d_i}{d_{i-1}}$ на диагонали. Тогда последовательность δ_i меняет знак тогда и только тогда, когда $\lambda_i < 0$. $\square$

Поговорим теперь про частный случай положительно определённых форм

Лемма 38. Положительно определённая форма всегда невырождена.

Доказательство. $h(x, x) > 0$ и поэтому не равно 0. $\square$

Теорема 36. Пусть дана форма q на вещественном пространстве V и её матрица A в некотором базисе. Следующие условия эквивалентны:
1) Форма q положительно определена.
2) Матрица A представима в виде $A = C^\top C$ для некоторой невырожденной матрицы C.
3) Главные миноры матрицы A положительны.

Доказательство. Если форма $q > 0$, то она невырождена и имеет сигнатуру $(n, 0)$. Тогда есть обратимая матрица C, что $E_n = C^\top A C$. Тогда $A = {C^{-1}}^\top C^{-1}$. Обратно, пусть $A = C^\top C$, тогда если $x^\top A x = (Cx)^\top C x \geq 0$. Более того, это выражение равно нулю только если $Cx = 0$. Но такое возможно, только если $x = 0$.

Пусть теперь $d_i > 0$, тогда $q > 0$ по критерию Сильвестра. Обратно, если $q > 0$, то $q|_{\langle e_1, \dots, e_k \rangle} > 0$. Но определитель матрицы этой формы и есть главный минор порядка k. Отсюда $d_k > 0$ как определитель произведения $\det C^\top C = \det^2 C$. $\square$

Замечание. Матрицу C можно выбрать верхнетреугольной.

Критерий Сильвестра пригоден в случае, если имеется задача с параметром. Иначе, вычисление определителя практически эквивалентно приведению к диагональному виду.

Покажем, как можно применять понятие положительной определённости. Точнее, того, что положительная определённость гарантирует невырожденность.

Задача: Рассмотрим множество $\{1, \dots, n\}$. Сколько может быть различных подмножеств $C_1, \dots, C_m$, таких, что $|C_i \cap C_j| = t$ одинаково для всех $i \neq j$?

Мы покажем, что есть ограничение $m \leq n$. Прежде всего, это надо сделать в случае, если $|C_i| = t$ для некоторого i. В этом случае $C_i \subseteq C_j$ для всех j. Тогда $C'_j = C_j \setminus C_i$ лежат в множестве $\{1, \dots, n\} \setminus C_i$, которое состоит из $n - t$ элементов. Множества C'_j не пусты, но имеют пустое пересечение друг с другом. Тогда их меньше чем $n - t$ штук, откуда – каждое множество должно содержать уникальный элемент. Итого

$$m - 1 \leq n - t \leq n - 1 \text{ или, по-другому, } m \leq n.$$

Теперь покажем, что в ситуации $d_i = |C_i| > t$ выполнено то же неравенство. Для этого составим матрицу инцидентности

$$B_{ij} = \begin{cases} 1, & \text{если } i \in C_j. \\ 0, & \text{иначе} \end{cases}.$$

Теперь матрица $B^\top B$ есть квадратная симметричная матрица размера m. Я утверждаю, что $B^\top B$ положительно определена. Для этого найдём её явно.

$$B^\top B = \begin{pmatrix} d_1 & t & t \\ t & \ddots & t \\ t & t & d_m \end{pmatrix} = \begin{pmatrix} d_1 - t & & 0 \\ & \ddots & \\ 0 & & d_m - t \end{pmatrix} + \begin{pmatrix} t & \dots & t \\ \vdots & \ddots & \vdots \\ t & \dots & t \end{pmatrix}.$$

Тогда при $x \neq 0$ имеем

$$x^\top B^\top B x = \sum (d_i - t) x_i^2 + t(\sum x_i)^2 > 0.$$

Теперь

$$m = \operatorname{rank} B^\top B \leq \operatorname{rank} B \leq n.$$