Евклидовы и унитарные пространства

Определение 1. Линейное пространство V над полем $\mathbb{R}$ называется *евклидовым*, если на нем определена функция $f : V \times V \to \mathbb{R}$ двух аргументов (обозначается $f(a,b) = (a,b)$ и называется скалярным произведением), удовлетворяющая следующим свойствам:

1) линейность по второму аргументу: $(a, b + \lambda c) = (a,b) + \lambda(a,c)$ для любых $a, b, c \in V$, $\lambda \in \mathbb{R}$;

$\lambda = 1$ $\quad b = 0$

2) симметричность: $(b,a) = (a,b)$ для любых $a, b \in V$;

3) положительная определенность: $(a,a) \geqslant 0$ для любого $a \in V$, причем, если $(a,a) = 0$, то $a = 0$.

Видно, что благодаря второму свойству эта функция также будет линейной и по первому аргументу, т.е. она *билинейна*.

Определение 2. Линейное пространство V над полем $\mathbb{C}$ называется *эрмитовым*, если на нем определена функция $f : V \times V \to \mathbb{C}$ двух аргументов (обозначается $f(a,b) = (a,b)$, называется скалярным произведением), удовлетворяющая следующим свойствам:

1) линейность по второму аргументу: $(a, b + \lambda c) = (a,b) + \lambda(a,c)$ для любых $a, b, c \in V$, $\lambda \in \mathbb{C}$;

$z = x + iy$ $\quad \bar{z} = x - iy$

2) эрмитовость: $(b,a) = \overline{(a,b)}$ для любых $a, b \in V$;

3) положительная определенность: $(a,a) \geqslant 0$, причем, если $(a,a) = 0$, то $a = 0$. Т.к. $(a,a) = \overline{(a,a)}$ (свойство 2), то число (a,a) вещественно, и неравенство $(a,a) \geqslant 0$ имеет смысл.

Используя второе свойство можно получить, что $(a + \lambda b, c) = (a,c) + \bar{\lambda}(b,c)$, т.е. она полу(анти)линейна по первому аргументу, такая функция называется полуторалинейной.

Пример:

Пусть $V = \mathbb{R}[t]$ — пространство многочленов над полем $\mathbb{R}$ (это один из немногих случаев, когда конечномерность пространства не играет существенной роли и не обязательно ограничиваться многочленами фиксированной степени), возьмем два произвольных вещественных числа a, b, $a < b$. Определим скалярное произведение двух многочленов $p(t), q(t)$ по следующей формуле: $(p(t), q(t)) = \int_a^b p(t)q(t)\,dt$. То, что выполнены первые два условия скалярного произведения, сразу вытекает из свойств интеграла, проверим положительную определенность. Действительно, $(p(t), p(t)) = \int_a^b p^2(t)\,dt \geqslant 0$, интеграл от неотрицательной функции неотрицателен, причем, если $\int_a^b p^2(t)\,dt = 0$, то $p(t) \equiv 0$. Аналогично, для пространства многочленов над полем $\mathbb{C}$ скалярное произведение можно задать формулой $(p(t), q(t)) = \int_a^b \overline{p(t)}q(t)\,dt$.

Определение 3. Длиной вектора a в евклидовом или эрмитовом пространстве называется число $|a| = \sqrt{(a,a)}$.

Это определение корректно, т.к. $(a,a) \geqslant 0$.

Лемма 4 (Неравенство Коши–Буняковского)**.** *Для любых двух векторов a, b евклидова или эрмитова пространства имеет место неравенство $|(a,b)| \leqslant |a| \cdot |b|$.*

Доказательство. Начнем с более простого — вещественного — случая. Рассмотрим скалярный квадрат $(a - \lambda b, a - \lambda b) \geqslant 0$, следовательно, для всех $\lambda \in \mathbb{R}$ имеем квадратичное неравенство $(a,a) - 2\lambda(a,b) + \lambda^2(b,b) \geqslant 0$, следовательно дискриминант этого квадратного трехчлена неположительный, т.е. $(a,b)^2 - (a,a)(b,b) \leqslant 0$. Перенося $(a,a)(b,b)$ в правую часть и извлекая корень, получаем искомое неравенство. $D/4$

Перейдем теперь к комплексному случаю. Возьмем произвольное $\lambda \in \mathbb{C}$. Тогда

$$(a - \lambda b, a - \lambda b) = (a,a) - \overline{\lambda}(b,a) - \lambda(a,b) + |\lambda|^2(b,b) \geqslant 0. \qquad \bar{\lambda}\cdot\lambda = |\lambda|^2$$

Т.к. (a,b) — комплексное число, то для некоторого угла φ выполнено равенство $(a,b) = |(a,b)|e^{i\varphi}$. Ограничимся только теми λ, для которых $\lambda e^{i\varphi} = \mu \in \mathbb{R}$, тогда наш скалярный квадрат можно переписать в виде

$$(a - \lambda b, a - \lambda b) = (a,a) - 2\mu|(a,b)| + \mu^2(b,b). \geqslant 0 \quad \forall \mu \in \mathbb{R}$$

Далее, действуя, как в вещественном случае, получаем нужный результат. $\square$

Лемма 5 (Неравенство треугольника)**.** *Для любых двух векторов a, b евклидова или эрмитова пространства имеет место неравенство $|a+b| \leqslant |a| + |b|$.*

Доказательство. Из неравенства Коши–Буняковского следует, что $(a,b) + (b,a) \leqslant 2|a| \cdot |b|$. Добавив к обеим частям неравенства $(a,a) + (b,b)$, получим

$$|a+b|^2 = (a+b, a+b) = (a,a) + (b,b) + (a,b) + (b,a) \leqslant (a,a) + (b,b) + 2|a| \cdot |b| = (|a| + |b|)^2,$$

откуда следует неравенство треугольника. $\square$

Процесс ортогонализации

Определение 6. Два вектора a, b называются ортогональными ($a \perp b$), если $(a, b) = 0$. Система векторов $e_1, \dots, e_n$ называется ортонормированной, если $(e_i, e_j) = \delta_{ij}$, т.е. если векторы попарно ортогональны и длина каждого из них равна 1.

Лемма 7. *Ортонормированная система векторов является линейно независимой.*

Доказательство. Пусть $\lambda_1 e_1 + \ldots + \lambda_n e_n = 0$. Умножим скалярно обе части этого равенства (слева) на вектор e_i: $(e_i, \lambda_1 e_1 + \ldots + \lambda_n e_n) = \lambda_1 (e_i, e_1) + \ldots + \lambda_n (e_i, e_n) = \lambda_i = 0$.
□

Пусть $e_1, \dots, e_j$ — некоторая линейно независимая система векторов. Обозначим через V_i линейную оболочку первых i векторов, $V_i = \langle e_1, \dots, e_i \rangle$, и получим расширяющуюся цепочку подпространств $V_1 \subset V_2 \subset \ldots \subset V_j$.

Лемма 8. *Существует такой набор попарно ортогональных векторов $a_1, \dots, a_j$, что для каждого номера i линейная оболочка $\langle a_1, \dots, a_i \rangle$ совпадает с V_i.*

Доказательство. (индукция по количеству векторов)
1) При $j = 1$ утверждение очевидно. $a_1 = e_1$
2) Пусть это утверждение выполнено для количества векторов, равного j, докажем его для $j + 1$. Т.к. утверждение верно для j векторов, то мы можем считать, что векторы $a_1, \dots, a_j$ с указанными свойствами уже построены. Поскольку, по предположению индукции, $\langle a_1, \dots, a_j \rangle = \langle e_1, \dots, e_j \rangle$, среди векторов $a_1, \dots, a_j$ не может быть нулевого вектора (иначе $\dim \langle a_1, \dots, a_j \rangle < j = \dim \langle e_1, \dots, e_j \rangle$).
Построим вектор a_{j+1} в виде

$$a_{j+1} = e_{j+1} + \lambda_1 a_1 + \ldots + \lambda_j a_j. \qquad (1)$$

Линейная оболочка векторов $a_1, \dots, a_{j+1}$ совпадает с $\langle e_1, \dots, e_{j+1} \rangle$ при любых λ_i, поэтому мы будем подбирать коэффициенты λ_i так, чтобы выполнялось условие $(a_{j+1}, a_i) = 0$ для всех $i = 1, \dots, j$. Рассмотрим скалярное произведение обеих частей (1) на a_i:

$$0 = (a_{j+1}, a_i) = (e_{j+1}, a_i) + \lambda_1 (a_1, a_i) + \ldots + \lambda_j (a_j, a_i).$$

j строк
$i = 1, \dots, j$

Поскольку $(a_k, a_i) = 0$ при $k \neq i$ по предположению индукции, то $0 = (e_{j+1}, a_i) + \lambda_i (a_i, a_i)$, следовательно $\lambda_i = -\frac{(e_{j+1}, a_i)}{(a_i, a_i)}$ (знаменатель отличен от нуля, т.к. $\cancel{a_j \neq 0}$). $a_i \neq 0$ □

Таким образом, чтобы получить вектор a_{j+1}, надо из вектора e_{j+1} вычесть его ортогональные проекции на векторы $a_1, \dots, a_j$. Этот метод ортогонализации называется методом Грама–Шмидта.
Пусть $e_1, \dots, e_n$ — базис в линейном пространстве V. Применим к нему процесс ортогонализации Грама-Шмидта и получим другой базис V, $a_1, \dots, a_n$. Это, действительно, базис, т.к. их количество равно размерности V, и, поскольку $\langle a_1, \dots, a_n \rangle = V$, они линейно независимы.

Лемма 9. *Матрица C перехода от базиса $e_1, \dots, e_n$ к базису $a_1, \dots, a_n$ верхнетреугольная, с единицами на диагонали.*

Доказательство. Поскольку $\langle e_1, \dots, e_j \rangle = \langle a_1, \dots, a_j \rangle$, для любых $\lambda_1, \dots, \lambda_j \in \mathbb{K}$ существуют такие $\mu^1, \dots, \mu^j \in \mathbb{K}$, что $\lambda_1 a_1 + \cdots + \lambda_j a_j = \mu^1 e_1 + \cdots + \mu^j e_j$. Тогда формула (1) примет вид

$$a_{j+1} = e_{j+1} + \mu^1 e_1 + \cdots + \mu^j e_j,$$

т.е.

$$\begin{aligned}
a_1 &= e_1;\\
a_2 &= e_2 + \mu_2^1 e_1;\\
a_3 &= e_3 + \mu_3^1 e_1 + \mu_3^2 e_2;\\
\cdots &\;\cdots\; \cdots\ \cdots\ \cdots\ \cdots\ \cdots\\
a_n &= e_n + \mu_n^1 e_1 + \cdots + \mu_n^{n-1} e_{n-1}.
\end{aligned}$$

Значит, матрица перехода имеет вид

$$C = \begin{pmatrix}
1 & \mu_2^1 & \mu_3^1 & \mu_4^1 & \vdots & \mu_n^1\\
0 & 1 & \mu_1^2 & \mu_2^2 & \vdots & \mu_n^2\\
0 & 0 & 1 & \mu_1^3 & \vdots & \mu_n^3\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$

$\square$

Ортогональные матрицы. QR-разложение

Определение 10. Базис $e_1, \dots, e_n$ евклидова или эрмитова пространства V называется *ортогональным*, если $(e_i, e_j) = 0$ при $i \neq j$. Он называется *ортонормированным*, если он ортогональный и $(e_i, e_i) = 1$, $i = 1, \dots, n$.

Если $e_1, \dots, e_n$ — ортогональный базис, то базис $\widetilde{e}_1, \dots, \widetilde{e}_n$, где $\widetilde{e}_i = \frac{1}{\sqrt{(e_i, e_i)}} e_i$, $i = 1, \dots, n$, — ортонормированный.

$= \frac{1}{|e_i|} e_i$

Определение 11. Матрица C с коэффициентами из $\mathbb{R}$ называется *ортогональной*, если $C^t C = E$. Матрица C с коэффициентами из $\mathbb{C}$ называется *унитарной*, если $\overline{C}^t C = E$ (здесь через $\overline{C}$ обозначает матрицу, полученную из матрицы C заменой всех элементов на их комплексно сопряженные).

Лемма 12. *Пусть $e_1 \dots, e_n$ — ортонормированный базис евклидова (эрмитова) пространства V, матрица C — матрица перехода к базису $\widetilde{e}_1, \dots, \widetilde{e}_n$. Базис $\widetilde{e}_1, \dots, \widetilde{e}_n$ ортонормированный тогда и только тогда, когда матрица C ортогональная (унитарная).*

Доказательство. Пусть $C = \begin{pmatrix} c_1^1 & \cdots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \cdots & c_n^n \end{pmatrix}$. Тогда $\widetilde{e}_i = c_i^j e_j$, $\widetilde{e}_k = c_k^l e_l$, поэтому $(\widetilde{e}_i, \widetilde{e}_k) = (c_i^j e_j, c_k^l e_l) = \overline{c}_i^j c_k^l (e_j, e_l) = \sum_{j=1}^n \overline{c}_i^j c_k^j$. Ортонормированность базиса $\widetilde{e}_1, \dots, \widetilde{e}_n$ равносильна тому, что $\sum_{j=1}^n \overline{c}_i^j c_k^j = 0$ при $i \neq k$, и $\sum_{j=1}^n \overline{c}_i^j c_i^j = 1$, $i = 1, \dots, n$. Но $\sum_{j=1}^n \overline{c}_i^j c_k^j$ — это элемент матрицы $\overline{C}^t C$ с индексами i и k.

$(\overline{C}^t C)_i^k$ $\square$

Таким образом, ортогональные (унитарные) матрицы — это матрицы перехода от ортонормированного базиса к ортонормированному.

Если рассматривать столбцы матрицы как координаты векторов (записанные в ортонормированном базисе), то эта матрица ортогональна (унитарна), если эти вектора-столбцы ортонормированы.

V мн. пр. во + , ·

(· , ·)

первый второй

вект. аргументы

матрицы,

функции,

числовые посл

многочлены

— — — —

$\mathbb{K} = \mathbb{R}$ или $\mathbb{C}$

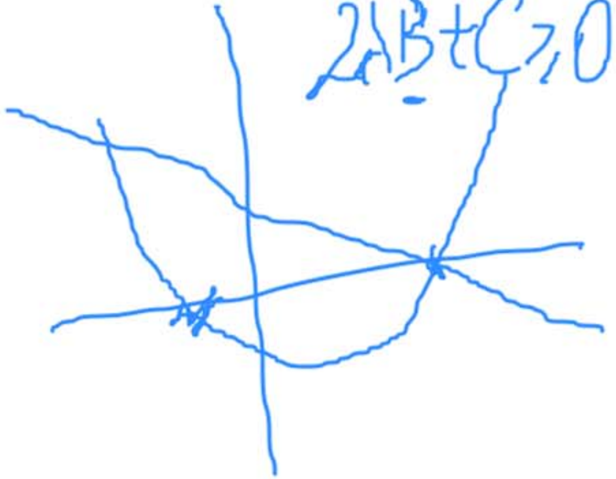
евклид. эрмит.

$(b,b)=0$

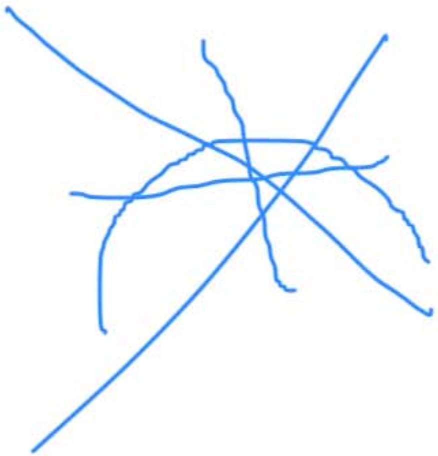
$2\lambda \underline{B} + C \geq 0$

$$\boxed{\lambda^2 \cdot A + 2\lambda \cdot B + C \geq 0 \text{ для всех } \lambda}$$

$A = (b,b)$, $B = (a,b)$, $C = (a,a)$



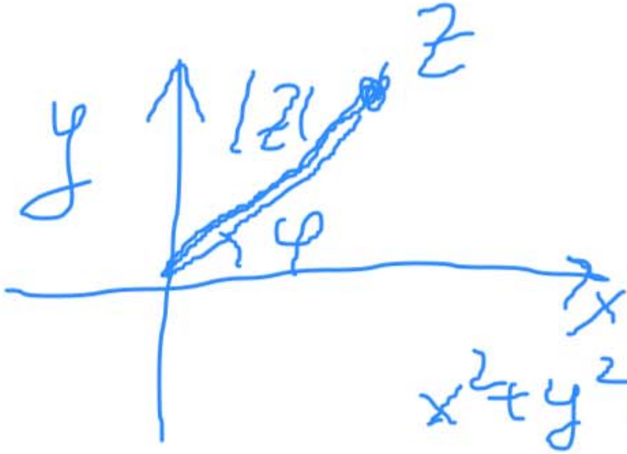
нет корней или единств. корень $\Longleftrightarrow D \leq 0$



$$\underbrace{(a,b)^2}_{=0} \leq (a,a)\underbrace{(b,b)}_{=0}$$

$z = |z|\cdot(\cos\varphi + i\sin\varphi) = |z|\cdot e^{i\varphi}$

φ – аргумент z



$x^2+y^2=|z|^2$

$\bar\lambda$ $(\lambda b, \lambda b) = \bar\lambda\lambda(b,b) = |\lambda|^2(b,b) \neq \lambda^2(b,b)$

$z = x+iy$

$|z| = \sqrt{x^2+y^2}$

$x = |z|\cos\varphi$

$y = |z|\sin\varphi$

$z \neq |z|$

$$(a,a)+\bar{\lambda}(b,a)+\underbrace{\lambda(a,b)}+|\lambda|^2(b,b) =$$

$$(b,a)=\overline{(a,b)}$$

$$(a,b)=|(a,b)|e^{i\varphi}$$

$$\overline{z_1\cdot z_2}=$$
$$=\overline{z_1}\cdot\overline{z_2}$$

$$=(a,a)+\underbrace{\overline{\lambda(a,b)}+\lambda\cdot(a,b)}+\ldots=$$

$$\lambda\cdot e^{i\varphi}=\mu$$

$$\overline{\lambda e^{i\varphi}}=\bar{\mu}$$

$$=(a,a)+\overline{\lambda|(a,b)|e^{i\varphi}}+\lambda|(a,b)|e^{i\varphi}+\ldots$$

$$=(a,a)+|(a,b)|\left(\overline{\lambda e^{i\varphi}}+\lambda\cdot e^{i\varphi}\right)+\ldots=(a,a)+2\mu\,|(a,b)|+\ldots$$

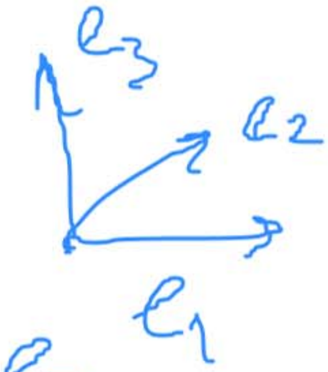
$$|(a,b)| \leq |(a,a)| \cdot |(b,b)|$$

$$|(a,b)| = |a| \cdot |b| \cdot |\cos\varphi| \leq$$
$$\leq |a| \cdot |b|$$

$$|(a,b)|^2 \leq (a,a) \cdot (b,b)$$

$$(p,q) = \int_a^b p(t)q(t)\,dt$$

$$\int_a^b p(t)q(t)\,dt^{\,2} \leq \int_a^b p^2(t)\,dt \cdot \int_a^b q^2(t)\,dt.$$



V_1 – лин. об. e_1 – прямая

V_2 – лин. об e_1 и e_2 – плоск.

V_3 – лин. об. e_1, e_2 и e_3

$\dim V_i = i$

$V_1 \subset V_2 \subset V_3 \subset \ldots$

$$\bar{C}^t = \begin{pmatrix} \bar{c}_1^1 & \cdots & \bar{c}_1^n \\ \cdots & \cdots & \cdots \\ \bar{c}_n^1 & \cdots & \bar{c}_n^n \end{pmatrix} \qquad C = \begin{pmatrix} c_1^1 & \cdots & c_n^1 \\ \cdots & \cdots & \cdots \\ c_1^n & \cdots & c_n^n \end{pmatrix}$$

$$(\bar{C}^t C)_1^1 = \bar{c}_1^1 c_1^1 + \cdots + \bar{c}_1^n c_1^n = 1$$

$$(\bar{C}^t C)_2^1 = \cdots$$