Ортогональные проекции

Пусть $L \subset V$, $a \in V$, тогда $a = a_1 + a_2$, где $a_1 \in L$, $a_2 \in L^\perp$.

Опр: 1. Ортогональной проекцией $a \in V$ назовем $a_1 \in L$, обозначение $a_\parallel$.

Опр: 2. Ортогональной составляющей $a \in V$ назовем $a_2 \in L^\perp$, обозначение $a_\perp$.

Любой вектор можно представить в виде ортогональной проекции и ортогональной составляющей:

$$L \subset V,\ \forall a \in V,\ \exists!\, a = a_\parallel + a_\perp,\ a_\parallel \in L,\ a_\perp \in L^\perp$$

Расстояние от вектора до подпространства L

Лемма 1. Запишем один и тот же вектор двумя способами: $a \in V \Rightarrow a = a_\parallel + a_\perp = a_1 + a_2$, где $a_1 \in L$. Тогда верно следующее неравенство:

$$|a_\perp| \leq |a_2|$$

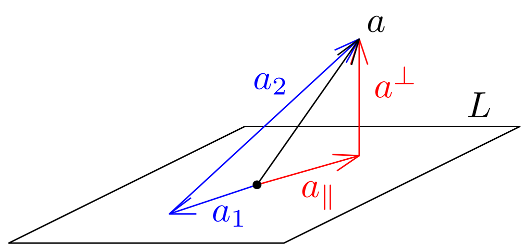

Рис. 1: Расстояние от a до L.

$\square$ Заметим, что $|a_\perp|^2 = |a - a_\parallel|^2 = (a - a_\parallel, a - a_\parallel)$, а также $\forall a \in L$, $(a, a_\perp) = 0$. Тогда:

$$a_1, a_\parallel \in L \Rightarrow (a_1, a_\perp) = (a_\parallel, a_\perp) = 0$$

Рассмотрим длину вектора a_2:

$$|a_2|^2 = |a - a_1|^2 = |a_\parallel + a_\perp - a_1|^2 = (a_\parallel + a_\perp - a_1, a_\parallel + a_\perp - a_1) =$$

$$= (a_\parallel, a_\parallel + a_\perp - a_1) + (a_\perp, a_\parallel + a_\perp - a_1) - (a_1, a_\parallel + a_\perp - a_1) =$$

$$= (a_\parallel, a_\parallel) + (a_\parallel, a_\perp) - (a_\parallel, a_1) + (a_\perp, a_\parallel) + (a_\perp, a_\perp) - (a_\perp, a_1) - (a_1, a_\parallel) - (a_1, a_\perp) + (a_1, a_1) =$$

$$= (a_\parallel, a_\parallel) + (a_\perp, a_\perp) + (a_1, a_1) - 2(a_\parallel, a_1) = |a_\perp|^2 + (a_\parallel, a_\parallel - a_1) + (a_1, a_1 - a_\parallel) =$$

$$= |a_\perp|^2 + (a_\parallel, a_\parallel - a_1) + (-a_1, a_\parallel - a_1) = |a_\perp|^2 + \underbrace{(a_\parallel - a_1, a_\parallel - a_1)}_{\geq 0} \geq |a_\perp|^2$$

где последнее неравенство верно в силу свойств скалярного произведения. $\blacksquare$

Опр: 3. Расстояние от $a \in V$ до L это длина ортогональной проекции $|a_\perp|$.

Лемма 2. Угол $(\widehat{a, a_\parallel})$ не больше, чем угол $(\widehat{a, a_1})$, при условии, что $a_1 \neq 0$. Или по другому:

$$(\widehat{a, a_\parallel}) \leq (\widehat{a, a_1}),\ a_1 \neq 0,\ a_1 \in L$$

где косинус угла между векторами мы понимаем в стандартном смысле: $\cos(\widehat{a, b}) = \dfrac{(a, b)}{|a| \cdot |b|}$.

☐ Пусть $a = a_1 + a_2 = b_1 + b_2$, где $a_1 \perp a_2$, $b_1 \perp b_2$. Тогда, если $|b_1| \geq |a_1|$, то:

$$(\widehat{a, a_1}) \geq (\widehat{a, b_1}) \Leftrightarrow \cos(\widehat{a, a_1}) \leq \cos(\widehat{a, b_1}),\ \sin(\widehat{a, a_1}) \geq \sin(\widehat{a, b_1})$$

Или по-другому это можно записать следующим образом:

$$\cos(\widehat{a, a_1}) = \frac{(a, a_1)}{|a|\cdot|a_1|} \leq \frac{(a, b_1)}{|a|\cdot|b_1|} = \cos(\widehat{a, b_1}) \Leftrightarrow \frac{(a_1, a_1)}{|a|\cdot|a_1|} = \frac{|a_1|}{|a|} \leq \frac{(b_1, b_1)}{|a|\cdot|b_1|} = \frac{|b_1|}{|a|} \Leftrightarrow |a_1| \leq |b_1|$$

Рассмотрим скалярное произведение вектора a с самим собой:

$$(a, a) = (a_1, a_1) + (a_1, a_2) + (a_2, a_1) + (a_2, a_2) = (a_1, a_1) + (a_2, a_2) \Leftrightarrow |a|^2 = |a_1|^2 + |a_2|^2 = |b_1|^2 + |b_2|^2$$

Следовательно, если $|b_1| \geq |a_1|$, то $|b_2| \leq |a_2|$ и $(\widehat{a, a_1}) \geq (\widehat{a, b_1})$.

Заменим теперь a_1 на $a'_1 = \lambda a_1$ и подберем λ таким образом, чтобы $\lambda a_1 \perp a'_2$, где $a'_2 = a - \lambda a_1$, $a = a'_1 + a'_2$. Таким образом:

$$\lambda a_1 \perp a'_2 \Leftrightarrow (\lambda a_1, a - \lambda a_1) = 0 \Rightarrow \lambda(a_1, a - \lambda a_1) = 0$$

Поскольку $a_1 \neq 0$ по условию и случай $\lambda = 0$ нас не интересует, то $\lambda \neq 0$ и верно следующее:

$$(a_1, a) - \lambda(a_1, a_1) = 0 \Rightarrow \lambda = \frac{(a_1, a)}{(a_1, a_1)},\ a_1 \neq 0 \Rightarrow (a_1, a_1) \neq 0$$

Рассмотрим угол между a и a_1 и угол между a и λa_1:

$$\lambda > 0 \Rightarrow (\widehat{a, a_1}) = (\widehat{a, a'_1}),\ \lambda < 0 \Rightarrow (\widehat{a, a_1}) = \pi - (\widehat{a, a'_1})$$

Знак λ определяется слагаемым (a_1, a): если $(\widehat{a, a_1}) \leq \frac{\pi}{2} \Rightarrow \lambda > 0$, если $(\widehat{a, a_1}) \geq \frac{\pi}{2} \Rightarrow \lambda < 0$. Это видно, если смотреть на формулу косинуса угла между векторами:

$$(\widehat{a, a_1}) \geq \frac{\pi}{2} \Rightarrow \cos(\widehat{a, a_1}) < 0 \Rightarrow \frac{(a, a_1)}{|a|\cdot|a_1|} < 0 \Rightarrow (a, a_1) = (a_1, a) < 0 \Rightarrow \lambda = \frac{(a_1, a)}{(a_1, a_1)} < 0$$

Заметим также, что угол между вектором и его проекцией всегда острый $\Rightarrow (\widehat{a, a_\parallel}) \leq \frac{\pi}{2}$, поэтому если $(\widehat{a, a_1}) \geq \frac{\pi}{2}$, то требуемое - очевидно. Рассмотрим случай $(\widehat{a, a_1}) \leq \frac{\pi}{2}$, $\lambda > 0$, тогда $(\widehat{a, a_1}) = (\widehat{a, a'_1})$. По предыдущей лемме: $|a_\perp| \leq |a'_2|$.

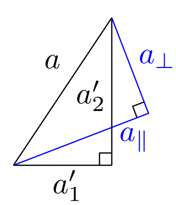

Рис. 2: Длина $a_\perp$ и a'_2.

Тогда верно следующее:

$$|a|^2 = |a_\perp|^2 + |a_\parallel|^2 = |a'_1|^2 + |a'_2|^2 \Rightarrow |a_\parallel| \geq |a'_1| \Rightarrow (\widehat{a, a_\parallel}) \leq (\widehat{a, a'_1})$$

поскольку напротив большей стороны лежит больший угол. Так как $\lambda > 0$, то:

$$(\widehat{a, a_\parallel}) \leq (\widehat{a, a'_1}) = (\widehat{a, a_1}) \leq \frac{\pi}{2}$$

■

Опр: 4. Углом между вектором и подпространством называется угол между вектором и его проекцией. При этом этот угол - наименьший, то есть $(\widehat{a, a_\parallel})$.

Матрица Грама

Пусть $a_1, \ldots, a_n$ - набор векторов.

Опр: 5. Матрицей Грама из набора векторов $a_1, \ldots, a_n$ называется матрица:

$$G = G(a_1, \ldots, a_n) = \begin{pmatrix} (a_1, a_1) & \ldots & (a_1, a_n) \\ \vdots & \ddots & \vdots \\ (a_n, a_1) & \ldots & (a_n, a_n) \end{pmatrix}$$

составленная из скалярных произведений (a_i, a_j).

Объем параллелепипедов

Опр: 6. Параллелепипед, построенный по векторам $a_1, \ldots, a_n$ есть следующее множество:

$$P(a_1, \ldots, a_n) = \{t_1 a_1 + \ldots + t_n a_n \colon 0 \leq t_i \leq 1\}$$

Например, рассмотрим параллелепипед построенный на векторах a_1, a_2:

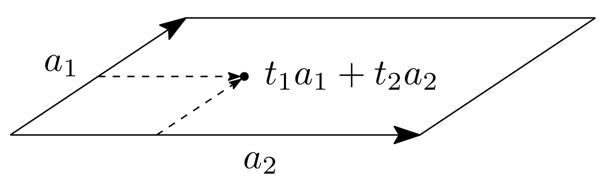

Рис. 3: Задание точек параллелепипеда $P(a_1, a_2)$.

Все точки на нём можно задать через $P(a_1, a_2) = \{t_1 a_1 + t_2 a_2 \colon 0 \leq t_i \leq 1\}$.

Пусть V_n это n-мерный объем $P(a_1, \ldots, a_n)$. Определеим V_n по индукции:

(1) $n = 1 \Rightarrow P(a_1) = \{t_1 a_1 \colon t_1 \in [0, 1]\} \Rightarrow V_1(P(a_1)) \coloneqq |a_1|$;

(2) $n = 2 \Rightarrow V_2(P(a_1, a_2)) \coloneqq |a_1| \cdot d(a_2, \langle a_1 \rangle)$, то есть произведение основания на высоту, где мы обозначили $d(a, L)$ - расстояние от вектора a до L;

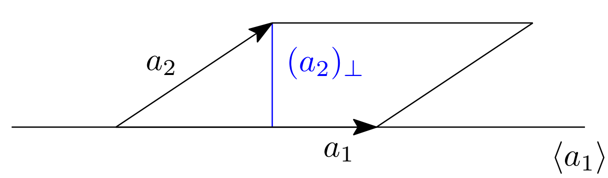

Рис. 4: Объем параллелепипеда $P(a_1, a_2)$.

(3) $n = 3 \Rightarrow V_3(P(a_1, a_2, a_3)) = V_2(P(a_1, a_2)) \cdot d(a_3, \langle a_1, a_2 \rangle)$, то есть произведение площади основания на высоту;

$\vdots$

(n) По индукции, пусть определено $V_{n-1}(P(a_1, \ldots, a_{n-1}))$, тогда:

$$V_n(P(a_1, \ldots, a_n)) = V_{n-1}(P(a_1, \ldots, a_{n-1})) \cdot d(a_n, \langle a_1, \ldots, a_{n-1} \rangle)$$

Пока это не очень хорошее определение, поскольку не ясно, зависит ли таким образом заданный объем от порядка $a_1, \ldots, a_n$. Рассмотрим следующую теорему.

Теорема 1. $V_n(P(a_1, \ldots, a_n))^2 = \det(G(a_1, \ldots, a_n))$.

□ Докажем по индукции:

База: $n = 1 \Rightarrow |a_1|^2 = (a_1, a_1) \Rightarrow$ утверждение верно.

Шаг: пусть теорема верна для $n-1 \Rightarrow$ выполним шаг индукции и рассмотрим матрицу Грама n-го порядка:

$$G = G(a_1, \ldots, a_n) = \begin{pmatrix} (a_1, a_1) & \ldots & (a_1, a_{n-1}) & (a_1, a_n) \\ \vdots & \ddots & \vdots & \vdots \\ (a_{n-1}, a_1) & \ldots & (a_{n-1}, a_{n-1}) & (a_{n-1}, a_n) \\ (a_n, a_1) & \ldots & (a_n, a_{n-1}) & (a_n, a_n) \end{pmatrix}$$

Запишем вектор a_n в виде суммы ортогональной проекции и ортогональной составляющей:

$$a_n = a_\parallel + a_\perp = \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1} + a_\perp$$

Подставим его в матрицу и рассмотрим изменения в последнем столбце. Поскольку $(a_i, a_\perp) = 0$, $\forall i < n$, то получим следующее:

$$(a_i, a_n) = (a_i, \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1} + a_\perp) = \lambda_1 (a_i, a_1) + \ldots + \lambda_{n-1}(a_i, a_{n-1}) + (a_i, a_\perp)$$

$$\begin{pmatrix} (a_1, a_n) \\ \vdots \\ (a_{n-1}, a_n) \\ (a_n, a_n) \end{pmatrix} = \begin{pmatrix} \lambda_1(a_1, a_1) & + & \ldots & + & \lambda_{n-1}(a_1, a_{n-1}) & + & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_1(a_{n-1}, a_1) & + & \ldots & + & \lambda_{n-1}(a_{n-1}, a_{n-1}) & + & 0 \\ \lambda_1(a_n, a_1) & + & \ldots & + & \lambda_{n-1}(a_n, a_{n-1}) & + & (a_n, a_\perp) \end{pmatrix} =$$

$$= \lambda_1 \cdot \begin{pmatrix} (a_1, a_1) \\ \vdots \\ (a_{n-1}, a_1) \\ (a_n, a_n) \end{pmatrix} + \ldots + \lambda_{n-1} \cdot \begin{pmatrix} (a_1, a_{n-1}) \\ \vdots \\ (a_{n-1}, a_{n-1}) \\ (a_n, a_{n-1}) \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ (a_n, a_\perp) \end{pmatrix}$$

таким образом, получили линейную комбинацию от 1-го до $n-1$-го столбца плюс еще один столбец. По свойству определителей выполнено следующее:

$$\begin{vmatrix} x_1 & \ldots & x_{n-1} & \lambda_1 x_1 + \ldots + \lambda_{n-1} x_{n-1} + x_n \end{vmatrix} = \begin{vmatrix} x_1 & \ldots & x_{n-1} & \lambda_1 x_1 \end{vmatrix} + \ldots + \begin{vmatrix} x_1 & \ldots & x_{n-1} & \lambda_{n-1} x_{n-1} \end{vmatrix} +$$

$$+ \begin{vmatrix} x_1 & \ldots & x_{n-1} & x_n \end{vmatrix} = 0 + \ldots + 0 + \begin{vmatrix} x_1 & \ldots & x_{n-1} & x_n \end{vmatrix} = \begin{vmatrix} x_1 & \ldots & x_{n-1} & x_n \end{vmatrix}$$

Поскольку $a_n = a_\parallel + a_\perp$, то $(a_n, a_\perp) = (a_\parallel + a_\perp, a_\perp) = |a_\perp|^2$. Рассмотрим определитель матрицы Грама:

$$|G| = |G(a_1, \ldots, a_n)| = \begin{vmatrix} (a_1, a_1) & \ldots & (a_1, a_{n-1}) & 0 \\ \vdots & \ddots & \vdots & \vdots \\ (a_{n-1}, a_1) & \ldots & (a_{n-1}, a_{n-1}) & 0 \\ (a_n, a_1) & \ldots & (a_n, a_{n-1}) & |a_\perp|^2 \end{vmatrix} = |a_\perp|^2 \cdot |G(a_1, \ldots, a_{n-1})|$$

где $|a_\perp|^2 = d(a_n, \langle a_1, \ldots, a_{n-1} \rangle)^2$ верно по определению, тогда:

$$|G(a_1, \ldots, a_n)| = |G(a_1, \ldots, a_{n-1})| \cdot d(a_n, \langle a_1, \ldots, a_{n-1} \rangle)^2$$

что дает равенство определителя матрицы Грама квадрату объема параллелепипеда. ■

Следствие 1. Определитель матрицы Грама неотрицателен: $\det G \geq 0$, $\forall a_1, \ldots, a_n$.

$\square$ Доказательство по теореме выше. $\blacksquare$

Следствие 2. $|G(a_1, \ldots, a_n)| = 0 \Leftrightarrow a_1, \ldots, a_n$ - линейно зависимы.

$\square$

($\Leftarrow$) Пусть $a_n = \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1}$, тогда будет верно:

$$\begin{vmatrix} (a_1, a_1) & \ldots & (a_1, a_{n-1}) & (a_1, \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1}) \\ \vdots & \ddots & \vdots & \vdots \\ (a_n, a_1) & \ldots & (a_n, a_{n-1}) & (a_n, \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1}) \end{vmatrix} = 0$$

($\Rightarrow$) $|G(a_1)|$, $|G(a_1, a_2)|, \ldots, |G(a_1, \ldots, a_k)|$, $|G(a_1, \ldots, a_k, a_{k+1})|, \ldots, |G(a_1, \ldots, a_n)| = 0 \Rightarrow$ если все равны нулю, то $G(a_1) = |a_1|^2 = 0 \Rightarrow a_1 = 0 \Rightarrow$ вектора линейно зависимы.

В том случае, если начиная с некоторого номера k, $|G(a_1, \ldots, a_k)| \neq 0$, но $|G(a_1, \ldots, a_k, a_{k+1})| = 0$, тогда:

$$V_k(P(a_1, \ldots, a_k)) \neq 0,\ V_{k+1}(P(a_1, \ldots, a_{k+1})) = 0 \Rightarrow$$
$$\Rightarrow V_{k+1}(P(a_1, \ldots, a_{k+1})) = V_k(P(a_1, \ldots, a_k)) \cdot d(a_{k+1}, \langle a_1, \ldots, a_k \rangle) = 0 \Rightarrow d(a_{k+1}, \langle a_1, \ldots, a_k \rangle) = 0 \Rightarrow$$
$$\Rightarrow a_{k+1} = (a_{k+1})_\parallel + (a_{k+1})_\perp = (a_{k+1})_\parallel$$

где последнее верно поскольку $|(a_{k+1})_\perp| = d(a_{k+1}, \langle a_1, \ldots, a_k \rangle) = 0 \Rightarrow a_{k+1} = (a_{k+1})_\parallel \in \langle a_1, \ldots, a_k \rangle \Rightarrow$ получили линейную зависимость векторов $a_1, \ldots, a_n$. $\blacksquare$

Следствие 3. При перестановке векторов $a_1, \ldots, a_n$ матрица Грама не меняется $\Rightarrow$ объем не зависит от порядка $a_1, \ldots, a_n$.

Пусть $e_1, \ldots, e_n, \tilde{e}_1, \ldots, \tilde{e}_n$ - два базиса, C - матрица перехода, $G = G(e_1, \ldots, e_n)$, $\widetilde{G} = G(\tilde{e}_1, \ldots, \tilde{e}_n)$. Обозначим следующим образом элементы этих матриц: $g_{ij} = (e_i, e_j)$, $\tilde{g}_{ij} = (\tilde{e}_i, \tilde{e}_j)$: $\tilde{e}_i = c_i^j e_j$ (суммирование по индексу j).

Пусть $\tilde{e}_i = c_i^k e_k$, $\tilde{e}_j = c_j^l e_l \Rightarrow$ подставим и получим следующее:

$$\tilde{g}_{ij} = (c_i^k e_k, c_j^l e_l) = \begin{cases} c_i^k c_j^l \cdot g_{kl}, & \mathbb{K} = \mathbb{R} \\ \overline{c_i^k} c_j^l \cdot g_{kl}, & \mathbb{K} = \mathbb{C} \end{cases}$$

или в матричном виде для случая $\mathbb{K} = \mathbb{R}$ получим такой результат:

$$\widetilde{G} = \begin{pmatrix} c_1^1 & \ldots & c_1^n \\ \vdots & \ddots & \vdots \\ c_n^1 & \ldots & c_n^n \end{pmatrix} \cdot \begin{pmatrix} g_{11} & \ldots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \ldots & g_{nn} \end{pmatrix} \cdot \begin{pmatrix} c_1^1 & \ldots & c_n^1 \\ \vdots & \ddots & \vdots \\ c_1^n & \ldots & c_n^n \end{pmatrix} = C^T G C$$

Следствие 4. Пусть G - матрица Грама над $\mathbb{R}$ для некоторого базиса, тогда существует невырожденная матрица C: $G = C^T C$.

$\square$ Возьмем первый базис произвольный, а второй ортонормированный, тогда:

$$G_1 = E,\ G = G_2 = C^T G_1 C = C^T E C = C^T C$$

$\blacksquare$

Метод наименьших квадратов

Рассмотрим следующую систему уравнений:

$$\begin{cases} a_1^1 x_1 + \ldots + a_1^n x_n & = & b_1 \\ \ddots & \vdots & \vdots \\ a_m^1 x_1 + \ldots + a_m^n x_n & = & b_m \end{cases}, \; m > n$$

Поскольку $m > n$, то скорее всего решений нет. Данная система эквивалентна следующей:

$$a_1 x_1 + \ldots + a_n x_n = b, \; a_1 = \begin{pmatrix} a_1^1 \\ \vdots \\ a_m^1 \end{pmatrix}, \ldots, a_n = \begin{pmatrix} a_1^n \\ \vdots \\ a_m^n \end{pmatrix}, \; b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Пусть $L = \langle a_1, \ldots, a_n \rangle$. Будет ли b принадлежать L? Может не принадлежать.

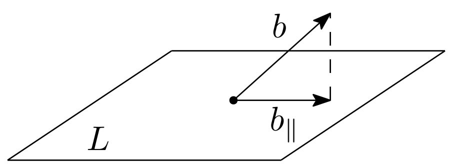

Рис. 5: Проекция вектора b на L.

Опр: 7. Псевдорешением называется набор $(x_1^\circ, \ldots x_n^\circ)$, который является решением системы:

$$a_1 x_1 + \ldots + a_n x_n = b_\parallel$$

Лемма 3. Верны следующие утверждения

(1) Псевдорешения всегда существуют;

(2) Если $a_1, \ldots, a_n$ линейно независимы, то псевдорешение единственно;

□ Очевидно, поскольку $a_1 x_1 + \ldots + a_n x_n$ - произвольный вектор в L. ■

Следствие 5. $|a_1 x_1 + \ldots + a_n x_n - b| \geq |a_1 x_1^\circ + \ldots + a_n x_n^\circ - b| = |b_\parallel - b|$, где $x_1^\circ, \ldots, x_n^\circ$ - псевдорешение.

Rm: 1. Другими словами, минимум $|a_1 x_1 + \ldots + a_n x_n - b|$ достигается на псевдорешении.