Теорема 13. *Любая обратимая матрица C допускает единственное представление в виде произведения $C = QR$, где Q — ортогональная (унитарная) матрица, а R — верхнетреугольная, с положительными числами на диагонали.*

Доказательство. Пусть $e_1, \ldots, e_n$ — ортонормированный базис. Обратимость C позволяет ее рассматривать как матрицу перехода к некоторому другому базису $\widetilde{e}_1, \ldots, \widetilde{e}_n$. Применим к базису $\widetilde{e}_1, \ldots, \widetilde{e}_n$ процесс ортогонализации Грама–Шмидта, в результате которого получим ортогональный базис $\widetilde{a}_1, \ldots, \widetilde{a}_n$. Нормализуем его, и получим уже ортонормированный базис $a_1, \ldots, a_n$. Поскольку матрица перехода от базиса $\widetilde{e}_1, \ldots, \widetilde{e}_n$ к базису $\widetilde{a}_1, \ldots, \widetilde{a}_n$ верхнетреугольная с единицами на диагонали, матрица перехода от базиса $\widetilde{e}_1, \ldots, \widetilde{e}_n$ к базису $a_1, \ldots, a_n$ тоже верхнетреугольная, но на диагонали могут быть произвольные положительные числа (т.к. мы делили вектора $\widetilde{a}_i$ на их длины)). Обозначим эту матрицу через T.

Итак, $(\widetilde{e}_1, \ldots, \widetilde{e}_n) = (e_1, \ldots, e_n)C$, $(a_1, \ldots, a_n) = (\widetilde{e}_1, \ldots, \widetilde{e}_n)T$. Отсюда $(a_1, \ldots, a_n) = (e_1, \ldots, e_n)CT$. Поскольку базис $a_1, \ldots, a_n$ ортонормированный, матрица CT, служащая матрицей перехода от одного ортонормированного базиса к другому ортонормированному, является ортогональной (унитарной). Обозначим ее Q. Тогда $CT = Q$. Обозначим также $R = T^{-1}$. Тогда $C = QR$. Матрица, обратная к верхнетреугольной, также верхнетреугольна, и на ее диагонали также положительные числа.

Мы доказали существование требуемого представления в виде произведения. Остается доказать единственность. Пусть $C = Q_1R_1 = Q_2R_2$, где Q_1, Q_2 ортогональные (унитарные), а R_1, R_2 верхнетреугольные с положительными числами на диагонали. Тогда $Q_2^{-1}Q_1 = R_2R_1^{-1}$. Простая проверка показывает, что как ортогональные (унитарные) матрицы, так и верхнетреугольные с положительными числами на диагонали, образуют группу, т.е. обратные элементы и произведения принадлежат этой же группе. Таким образом, ортогональная матрица $Q = Q_2^{-1}Q_1$ равна верхнетреугольной с положительными числами на диагонали $R = R_2R_1^{-1}$. Остается проверить, что единственная матрица, которая является одновременно ортогональной (унитарной) и верхнетреугольной с положительными числами на диагонали, является единичной матрицей. Рассмторим такую матрицу $Q = \begin{pmatrix} q_1^1 & q_2^1 & q_3^1 & \cdots & q_n^1 \\ 0 & q_2^2 & q_3^2 & \cdots & q_n^2 \\ 0 & 0 & q_3^3 & \cdots & q_n^3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & q_n^n \end{pmatrix}$. Ее ортогональность (унитарность) означает, что вектор-столбцы образуют ортонормированный базис. Единичная длина первого столбца означает, что $q_1^1 = 1$ (напомним, что на диагонали должны быть положительные числа). Ортогональность первого и второго столбцов дает $q_2^1 = 0$. Ортогональность третьего столбца с первым и вторым дает $q_3^1 = q_3^2 = 0$ и т.д.

□

Ортогональное дополнение

Определение 14. Пусть $V \subset W$ — подпространство евклидова или эрмитова пространства. *Ортогональным дополнением* $V^{\perp}$ к V в W называется множество, состоящее из векторов, ортогональных всем векторам из V, т.е. $V^{\perp} = \{w \in W : (v, w) = 0 \ \ \forall v \in V\}$.

Очевидным образом проверяется, что $V^{\perp}$ является подпространством, а не просто подмножеством.

Лемма 15. $W = V \oplus V^{\perp}$.

Доказательство. Пусть $e_1, \ldots, e_k$ — базис в V, дополним его до базиса всего пространства W векторами $e_{k+1}, \ldots, e_n$. Применив процесс ортогонализации Грама–Шмидта, получим ортогональный базис $a_1, \ldots, a_k, a_{k+1}, \ldots, a_n$ в W, причем его первая часть будет базисом в V, т.к. $\langle a_1, \ldots, a_k \rangle = \langle e_1, \ldots, e_k \rangle = V$. Покажем, что векторы $a_{k+1}, \ldots, a_n$ образуют базис в $V^{\perp}$. Пусть $v \in V^{\perp}$, $v = \nu_1 a_1 + \ldots + \nu_k a_k + \nu_{k+1} a_{k+1} + \ldots + \nu_n a_n$ — разложение вектора v по базису пространства W. Коэффициенты $\nu_1, \ldots, \nu_k$ должны быть нулевыми, так как иначе вектор v не был бы ортогонален всем векторам $a_1, \ldots, a_k$. Верно и обратное: если первые k координат какого-то вектора в базисе $a_1, \ldots, a_n$ равны нулю, то этот вектор принадлежит $V^{\perp}$. Следовательно, произвольный вектор $w \in W$ может быть представлен в виде суммы двух слагаемых — из V и из $V^{\perp}$: $w = \underbrace{\lambda_1 a_1 + \ldots + \lambda_k a_k}_{\in V} + \underbrace{\lambda_{k+1} a_{k+1} + \ldots + \lambda_n a_n}_{\in V^{\perp}}$,

т.е. $W = V + V^{\perp}$.

Докажем, что эта сумма прямая. Возьмем произвольный вектор $v \in V \cap V^{\perp}$. Т.к. $v \in V^{\perp}$, то $(v, w) = 0$ для любого вектора $w \in V$. Поскольку $v \in V$, мы можем в качестве w взять сам вектор v, тогда $(v, v) = 0$, значит, $v = 0$. Следовательно, пересечение состоит только из нулевого вектора, и сумма — прямая. □

Из разложения в прямую сумму $W = V \oplus V^{\perp}$ следует, что любой вектор $a \in W$ можно единственным способом представить в виде $a = a_0 + a_{\perp}$, где $a_0 \in V$, $a_{\perp} \in V^{\perp}$. Вектор a_0 называется (ортогональной) проекцией вектора a на подпространство V, а вектор $a_{\perp}$ называется ортогональной составляющей вектора a.

Пусть имеется другое разложение вектора a: $a = a_1 + a_2$, где $a_1 \in V$, (а на a_2 дополнительных условий нет), тогда имеет место

Утверждение 16. $|a_2| \geqslant |a_{\perp}|$.

Доказательство. Обозначим $a_0 - a_1 = b \in V$, тогда $a = a_1 + a_2 = a_0 - (a_0 - a_1) + a_2 = a_0 + (a_2 - b)$, следовательно, $a_2 - b = a_{\perp}$, т.е. $a_2 = a_{\perp} + b$. Тогда $(a_2, a_2) = (a_{\perp}, a_{\perp}) + 2\underbrace{(a_{\perp}, b)}_{=0} + (b, b) \geqslant (a_{\perp}, a_{\perp})$, откуда получаем $|a_2| \geqslant |a_{\perp}|$. □

$(a_2, a_2) = |a_2|^2$
$(a_{\perp}, a_{\perp}) = |a_{\perp}|^2$

Определение 17. Углом между двумя ненулевыми векторами в евклидовом пространстве называется величина $(\widehat{a, b}) := \arccos \frac{(a,b)}{|a| \cdot |b|}$.

Видно, что это определение корректно, т.к. $-1 \leqslant \frac{(a,b)}{|a| \cdot |b|} \leqslant 1$, и не противоречит здравому смыслу, т.е. угол равен нулю тогда и только тогда, когда вектора коллинеарны, и угол — прямой тогда и только тогда, когда вектора ортогональны.

В многомерном случае геометрия аналогична обычной геометрии. Так, например, в прямоугольных треугольниках с одинаковой гипотенузой, чем больше угол, тем больше противолежащий катет.

Утверждение 18. *Пусть* $c = a_1 + b_1 = a_2 + b_2$, $a_1 \perp b_1$, $a_2 \perp b_2$ *и* $|b_1| < |b_2|$. *Тогда* $(\widehat{a_1, c}) < (\widehat{a_2, c})$.

Доказательство. Поскольку $(a_i, c) = (a_i, a_i + b_i) = |a_i|^2$, $i = 1, 2$, то $\cos(\widehat{a_i, c}) = \frac{|a_i|}{|c|}$, а из теоремы Пифагора $\sin(\widehat{a_i, c}) = \frac{|b_i|}{|c|}$, что и доказывает утверждение.

Утверждение 19. *(обозначения те же, что и ранее)* $(\widehat{a, a_0}) \leqslant (\widehat{a, a_1})$.

Доказательство. Найдем такое число λ, чтобы вектор λa_1 был бы ортогонален вектору $b = a - \lambda a_1$:

$a = a_0 + a_\perp = \lambda a_1 + b$

$$(\lambda a_1, a - \lambda a_1) = 0 \Leftrightarrow (a_1, a - \lambda a_1) = 0 \Leftrightarrow (a_1, a) - \lambda \underbrace{(a_1, a_1)}_{\neq 0} = 0 \Leftrightarrow \lambda = \frac{(a_1, a)}{(a_1, a_1)}.$$

Если $(a_1, a) \leqslant 0$, то угол $(\widehat{a, a_1}) \geqslant \pi/2$ и утверждение очевидно. Если $(a_1, a) > 0$, то $\lambda > 0$ и угол между a и a_1 равен углу между a и λa_1. Применив предыдущую лемму, получаем требуемое утверждение. $\square$

Определение 20. Расстоянием $d(a, V)$ от вектора a до подпространства V называется наименьшее из всех возможных длин векторов, соединяющих векторы (точки) подпространства V с данным вектором, т.е. $d(a, V) := \min_{a_1 \in V} |a - a_1|$. Углом $(\widehat{a, V})$ между вектором a и подпространством V называется наименьший из всех углов между вектором a и произвольным вектором $a_1 \in V$, т.е. $(\widehat{a, V}) := \min_{a_1 \in V} (\widehat{a, a_1})$.

Очевидно, что $d(a, V) = |a_\perp|$ — расстояние от вектора до подпространства равно длине ортогональной составляющей при проекции вектора на подпространство, а $(\widehat{a, V}) = (\widehat{a, a_0})$ — угол между вектором и подпространством равен углу между вектором и его проекцией на данное подпространство.

Определение 21. Аффинное пространство $(\mathcal{A}, V, +)$ называется евклидовым аффинным пространством, если линейное пространство V является евклидовым.

Это позволяет говорить о расстояниях между точками и между точкой и аффинным подпространством. Если $A, B \in \mathcal{A}$, то эти две точки определяют вектор $v \in V$, $B = A + v$, и расстоянием $|AB|$ естественно назвать длину вектора v. Если $L \subset V$ — линейное подпространство, и если аффинное подпространство $\mathcal{B}$ задано как $\mathcal{B} = \{A + v : v \in L\}$, то расстоянием от точки B до аффинного подпространства $\mathcal{B}$ естественно назвать расстояние между вектором v (определенным равенством $B = A + v$) и линейным подпространством L.

Метод наименьших квадратов

Допустим, что мы исследуем какое-нибудь природное явление и хотим описать его линейной формулой, т.е. мы предполагаем, что какая-то величина b линейно зависит от других — $a_1, \ldots, a_n$, и хотим получить эту зависимость $b = a_1x^1 + \ldots + a_nx^n$, т.е. узнать неизвестные коэффициенты $x^1, \ldots, x^n$. Мы делаем m измерений (для точности берем $m > n$) и решаем систему уравнений $\begin{cases} a_1^1x^1 + \ldots + a_n^1x^n & = & b^1 \\ \ldots \quad \ldots \quad \ldots & & \\ a_1^mx^1 + \ldots + a_n^mx^n & = & b^m \end{cases}$. Вообще говоря, эта переопределенная система не имеет решения. Поэтому нам надо найти наиболее приближенное решение $x^1, \ldots, x^n$ в том смысле, что отклонение значений b^j от $c^j = a_1^jx^1 + \ldots + a_n^jx^n$ будет наименьшим. В качестве отклонения удобно рассмотреть корень из суммы квадратов отклонений координат $\sqrt{(b^1 - c^1)^2 + \ldots + (b^m - c^m)^2} = |b - c|$, что равно длине вектора $b - c$. Будем искать такое псевдо-решение.

Рассмотрим векторы

$$a_1 = \begin{pmatrix} a_1^1 \\ \vdots \\ a_1^m \end{pmatrix}, \quad \ldots, \quad a_n = \begin{pmatrix} a_n^1 \\ \vdots \\ a_n^m \end{pmatrix}, \quad b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix}.$$

Пусть $V = \langle a_1, \ldots, a_n \rangle$. Обычно m намного больше n, и векторы $a_1, \ldots, a_n$ линейно независимы. Если они все-таки линейно зависимы, следует отбросить какое-то их количество, чтобы оставшиеся образовали базис подпространства V. Будем считать, что это уже сделано, и векторы $a_1, \ldots, a_n$ линейно независимы.

Спроектируем вектор b на подпространство V. Получим разложение $b = b_0 + b_\perp$, и, как мы доказали ранее, $|b_\perp| = |b - b_0|$ будет наименьшей длиной векторов, соединяющих b с V, т.е. b_0 определяет искомое максимально приближенное псевдо-решение. Чтобы найти его, разложим вектор b_0 по базису подпространства V, $b_0 = x^1a_1 + \ldots + x^na_n$. Тогда, взяв скалярные произведения с векторами базиса, получаем следующие равенства:

$$\begin{cases} (a_1, x^1a_1 + \ldots + x^na_n) & = & (a_1, b) \\ \ldots \quad \ldots \quad \ldots \quad \ldots & & \\ (a_n, x^1a_1 + \ldots + x^na_n) & = & (a_n, b), \end{cases}$$

т.е. надо решить систему уравнений (уже квадратную):

$$\begin{cases} (a_1, a_1)x^1 + \ldots + (a_1, a_n)x^n & = & (a_1, b) \\ \ldots \quad \ldots \quad \ldots & & \\ (a_n, a_1)x^1 + \ldots + (a_n, a_n)x^n & = & (a_n, b). \end{cases}$$

Решив ее, найдем координаты $x^1, \ldots, x^n$ вектора b_0, которые и есть наше псевдо-решение.

Матрица этой системы уравнений

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

называется *матрицей Грама*. Далее мы убедимся, что ее определитель отличен от нуля, когда векторы $a_1, \ldots, a_n$ линейно независимы.

Параллелепипеды и матрица Грама

Определение 22. Пусть $a_1, \ldots, a_n$ — система векторов в векторном пространстве V. *Параллелепипедом*, натянутым на векторы $a_1, \ldots, a_n$ называется множество векторов (точек) $\Pi(a_1, \ldots, a_n) = \{x \in V : x = x^1 a_1 + \ldots + x^n a_n, \quad 0 \leqslant x^1, \ldots, x^n \leqslant 1\}$.

Определение 23. Определим *n-мерный объем* Vol_n параллелепипеда $\Pi(a_1, \ldots, a_n)$ индуктивно:

1) одномерный объем $\mathrm{Vol}_1 \Pi(a_1) := |a_1|$ — это длина вектора;

2)
$$\mathrm{Vol}_n \Pi(a_1, \ldots, a_n) := \mathrm{Vol}_{n-1} \Pi(a_1, \ldots, a_{n-1}) \cdot d(a_n, \langle a_1, \ldots, a_{n-1} \rangle).$$

Очевидно, что объем есть неотрицательная величина. Корректность этого определения, т.е. независимость объема от порядка векторов при индуктивном переходе, вытекает из следующей теоремы.

Теорема 24. $(\mathrm{Vol}_n \Pi(a_1, \ldots, a_n))^2 = \det G(a_1, \ldots, a_n)$.

Доказательство. (по индукции)

1) При $n = 1$, очевидно, $|a_1|^2 = (a_1, a_1)$.

2) Пусть утверждение верно для размерности $n-1$, докажем его для размерности n. Спроектируем вектор a_n на линейную оболочку векторов $a_1, \ldots, a_{n-1}$: $a_n = (a_n)_0 + (a_n)_\perp = \lambda_1 a_1 + \ldots + \lambda_{n-1} a_{n-1} + b$, где $b = (a_n)_\perp$ (т.е. b ортогонален векторам $a_1, \ldots, a_{n-1}$) и $|b| = d(a_n, \langle a_1, \ldots, a_{n-1} \rangle)$. Имеем:

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

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

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

$$= \underbrace{\lambda_1 \det\begin{pmatrix} (a_1,a_1) & \ldots & (a_1,a_{n-1}) & (a_1,a_1) \\ \vdots & \ddots & \vdots & \vdots \\ (a_n,a_1) & \ldots & (a_n,a_{n-1}) & (a_n,a_1) \end{pmatrix}}_{=0} + \ldots$$

$$\ldots + \underbrace{\lambda_{n-1} \det\begin{pmatrix} (a_1,a_1) & \ldots & (a_1,a_{n-1}) & (a_1,a_{n-1}) \\ \vdots & \ddots & \vdots & \vdots \\ (a_n,a_1) & \ldots & (a_n,a_{n-1}) & (a_n,a_{n-1}) \end{pmatrix}}_{=0} +$$

=0

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

$$= \det\begin{pmatrix} (a_1,a_1) & \ldots & (a_1,a_{n-1}) & \underbrace{(a_1,b)}_{=0} \\ \vdots & \ddots & \vdots & \vdots \\ (a_{n-1},a_1) & \ldots & (a_{n-1},a_{n-1}) & \underbrace{(a_{n-1},b)}_{=0} \\ (a_n,a_1) & \ldots & (a_n,a_{n-1}) & \underbrace{(a_n,b)}_{=(b,b)} \end{pmatrix} =$$

разл. по посл. столбцу

$$= \det\begin{pmatrix} (a_1,a_1) & \ldots & (a_1,a_{n-1}) \\ \vdots & \ddots & \vdots \\ (a_{n-1},a_1) & \ldots & (a_{n-1},a_{n-1}) \end{pmatrix} \cdot (b,b) = \det G(a_1,\ldots,a_{n-1}) \cdot |b|^2 =$$

$$= (\mathrm{Vol}_{n-1}\,\Pi(a_1,\ldots,a_{n-1}))^2 \cdot |b|^2 = (\mathrm{Vol}_n\,\Pi(a_1,\ldots,a_n))^2.$$

$\begin{pmatrix} 2 & 1 \\ 0 & \frac{1}{5} \end{pmatrix}$ не ортог

$\begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{pmatrix}$

$$\bar{C}^t C = E \Rightarrow C^{-1} = \bar{C}^t.$$

$$\begin{pmatrix} & \\ 0 & \end{pmatrix}^{-1} = \begin{pmatrix} & \\ 0 & \end{pmatrix}$$

$a, b \in V^{\perp}$ $\overset{?}{\Longrightarrow}$ $a+b \in V^{\perp}$

$\lambda a \in V^{\perp}$

$a \perp v \quad \forall v \in V$

$b \perp v \quad \forall v \in V$

$(a, v) = 0 \quad \forall v \in V$

$(b, v) = 0 \quad \forall v \in V$

$(\underline{a+b}, v) = 0 \Rightarrow (a, v) + (b, v) = 0.$

$v \in V^{\perp}$ $\quad a_1, \ldots, a_k$ – базис V

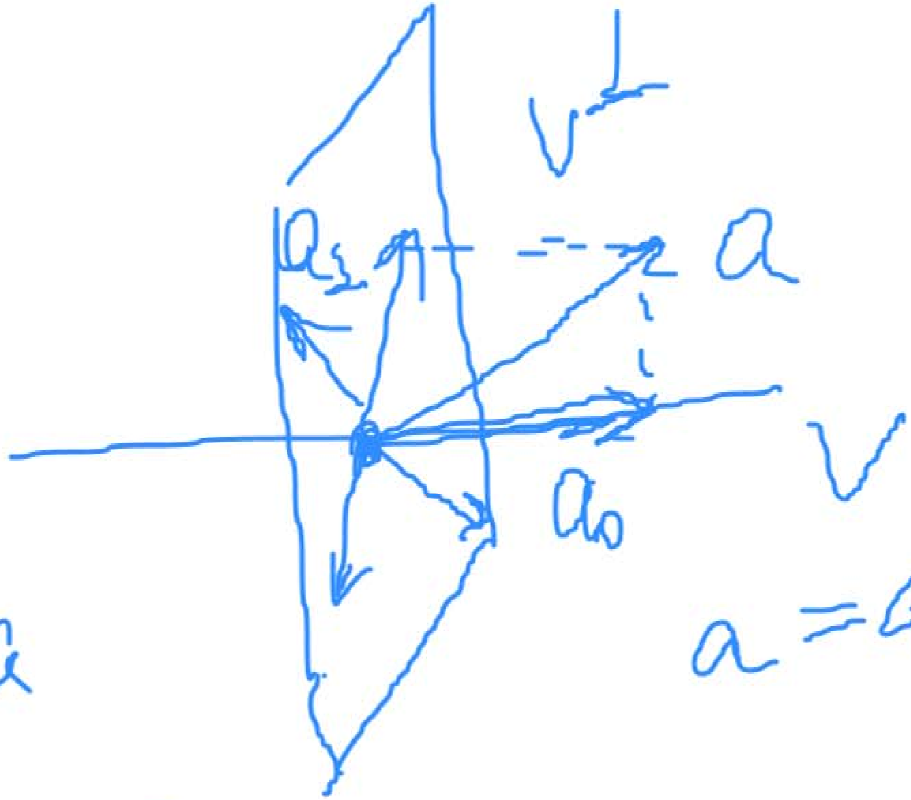
$v \perp$ любому вект. из V

$\Rightarrow (v, a_1) = 0, \ldots, (v, a_k) = 0$

$v = \boxed{\nu_1 a_1 + \ldots + \nu_k a_k} + \nu_{k+1} a_{k+1} + \ldots$

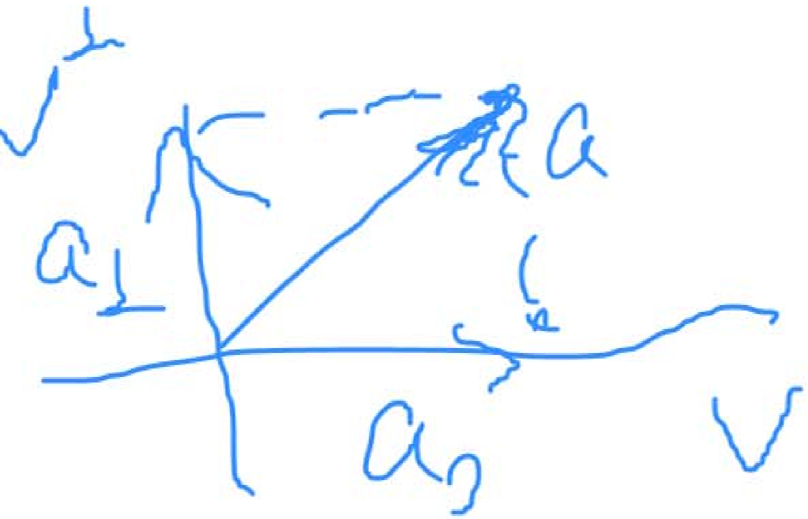
$\nu_1 (a_1, a_1) = 0 \Rightarrow \nu_1 = 0$

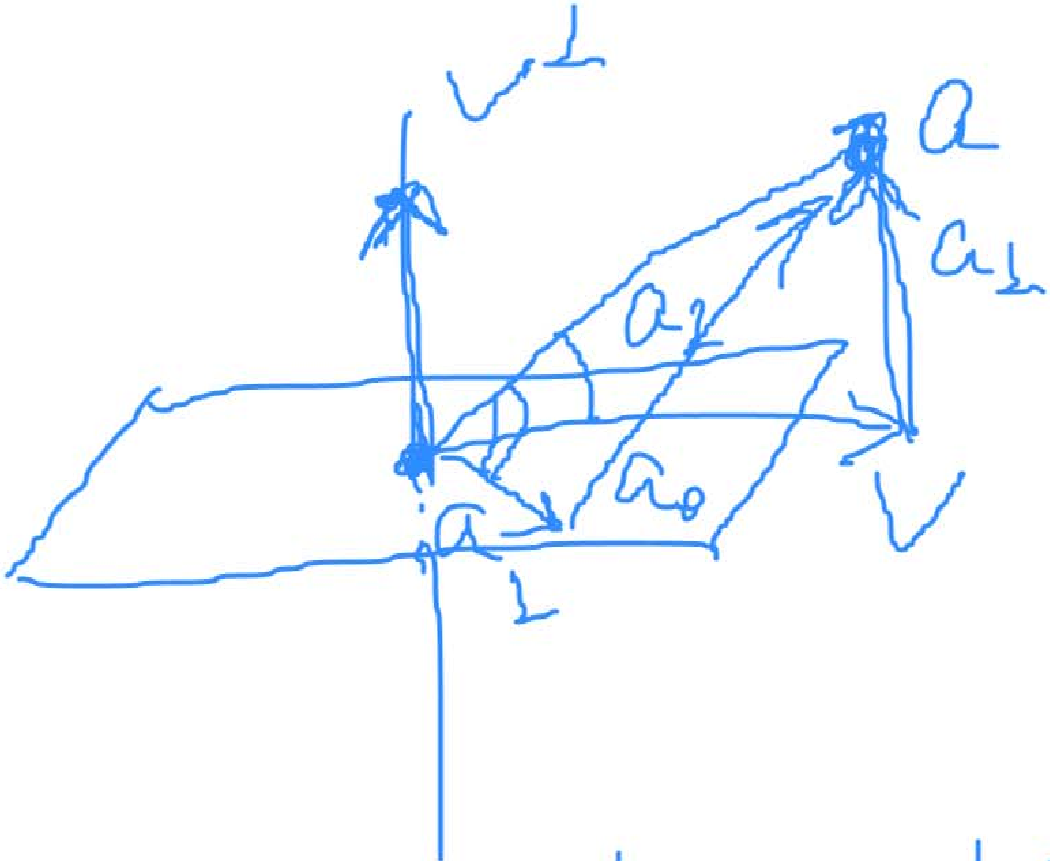
$W = \mathbb{R}^3$



$(V^\perp)^\perp = V$

$a = a_0 + a_\perp$





$a_1 \in V$

$a_2 = a - a$

$a = a_1 + a_2$

$\in V$

$|a_2| \geq |a_{\perp}|$

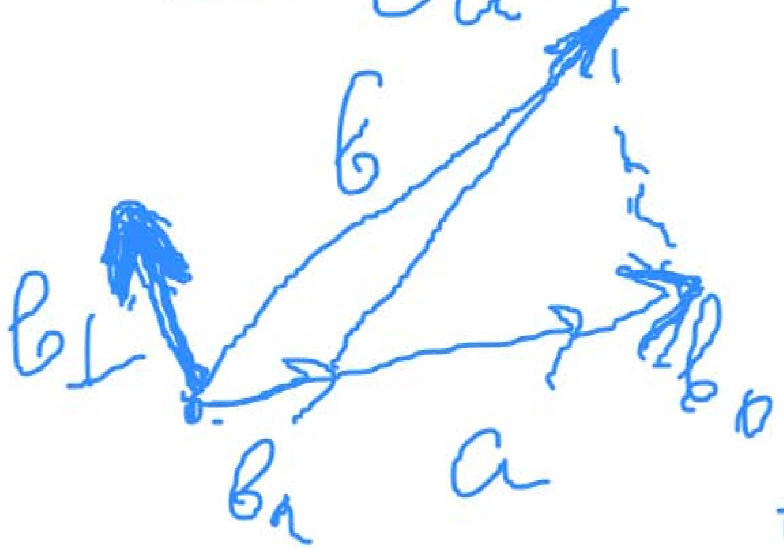
a_1

$$\boxed{b = ax}$$

$$\begin{cases} b^1 = a^1 x \\ b^2 = a^2 x \end{cases}$$

$n = 1$

$$a = \begin{pmatrix} a^1 \\ a^2 \end{pmatrix} \qquad b = \begin{pmatrix} b^1 \\ b^2 \end{pmatrix}$$



$\boxed{b_0 = a \cdot x}$ псевдореше

$b - b_0$ - ми

$b = b_1 + b_2$

$b - b_1$ из всевозм

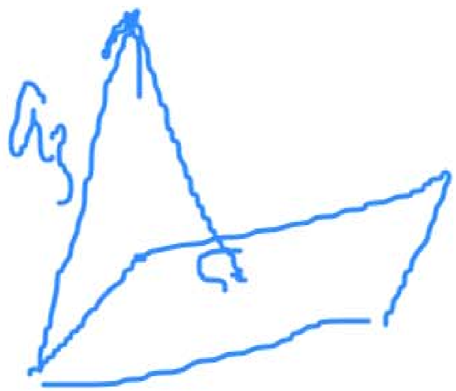
$b_1 \in \langle a \rangle$

$x = x^1 a_1 + x^2 a_2$

$0 \le x^1 \le 1$

$0 \le x^2 \le 1$

$n = 1$



$V_3 \ell_2 \Pi(a_1, a_2) \cdot d(a_3, \langle a_1, a_2 \rangle)$