Учебник по Эконометрике

Лекция 1: Векторные Пространства

Джон Стачурски

Лекции: Акшай Шенкер
Перевел: Алексей Кедо

27 сентября 2020 г.

Обзор

Линейная алгебра является основой математики и, в частности, эконометрики:

- проведение базовых вычислений с данными
- решение линейных уравнений используя данные
- продвинутые операции, такие как квадратичная минимизация

В центре внимания данной главы:

1. векторные пространства: линейные операции, нормы, линейные подпространства, линейная независимость, базисы и т.д.
2. теорема ортогональной проекции

Векторное пространство

Символ $\mathbb{R}^N$ показывает набор любых векторов длины N, или N векторов

N-вектор $\mathbf{x}$ – это список из N действительных чисел:

$$\mathbf{x} = (x_1, \ldots, x_N) \text{ ,где } \quad x_n \in \mathbb{R} \text{ для любого } n$$

Также мы можем записать $\mathbf{x}$ вертикально, вот так:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

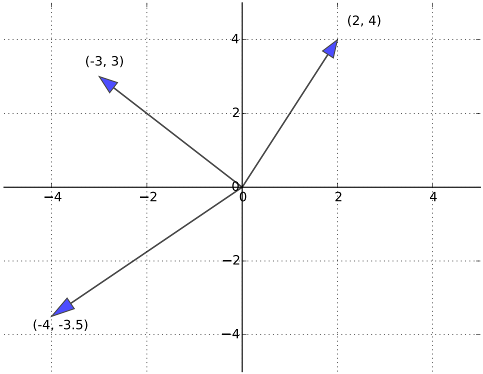


Рис.: Три вектора в $\mathbb{R}^2$

Вектор из единиц будет обозначен $\mathbf{1}$

$$\mathbf{1} := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Вектор из нулей будет обозначен $\mathbf{0}$

$$\mathbf{0} := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Линейные операции

Две базовые алгебраические операции:

1. Сложение векторов
2. Умножение на скаляр

1. **Сумма** $\mathbf{x} \in \mathbb{R}^N$ и $\mathbf{y} \in \mathbb{R}^N$ обозначается

$$\mathbf{x} + \mathbf{y} =: \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_N + y_N \end{pmatrix}$$

Пример 1:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \\ 6 \\ 8 \end{pmatrix} := \begin{pmatrix} 3 \\ 6 \\ 9 \\ 12 \end{pmatrix}$$

Пример 2:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} := \begin{pmatrix} 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}$$

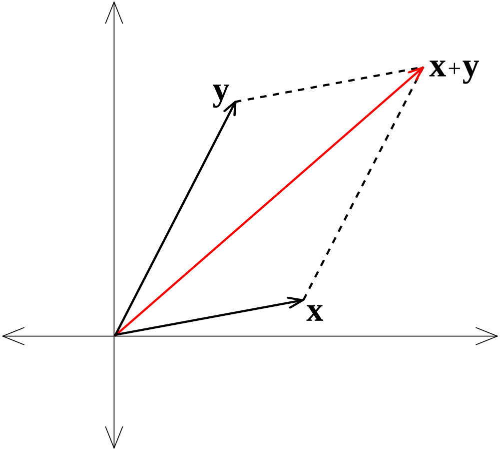


Рис.: Сложение векторов

2. **Умножение на скаляр** $\alpha \in \mathbb{R}$ и $\mathbf{x} \in \mathbb{R}^N$ обозначается

$$\alpha \mathbf{x} = \alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} := \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_N \end{pmatrix}$$

Пример 1:

$$0.5\begin{pmatrix}1\\2\\3\\4\end{pmatrix} := \begin{pmatrix}0.5\\1.0\\1.5\\2.0\end{pmatrix}$$

Пример 2:

$$-1\begin{pmatrix}1\\2\\3\\4\end{pmatrix} := \begin{pmatrix}-1\\-2\\-3\\-4\end{pmatrix}$$

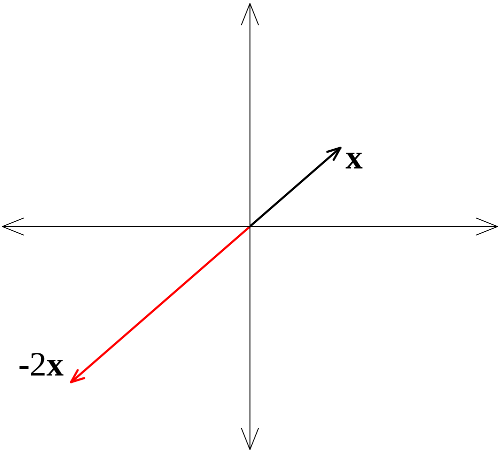


Рис.: Уножение на скаляр

Вычитание выполняется поэлементно, аналогично умножению

$$\mathbf{x} - \mathbf{y} := \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \\ \vdots \\ x_N - y_N \end{pmatrix}$$

Определение можно дать в терминах сложения и умножения на скаляр

$$\mathbf{x} - \mathbf{y} := \mathbf{x} + (-1)\mathbf{y}$$

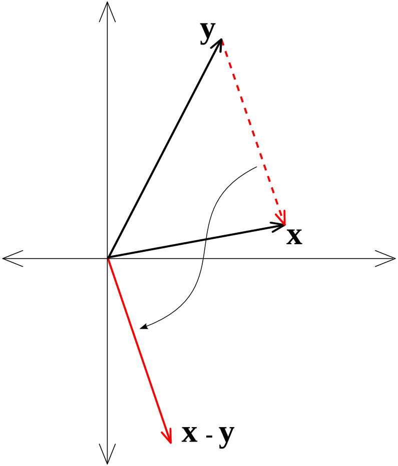


Рис.: Разница векторов

Скалярное произведение

Скалярное произведение двух векторов $\mathbf{x}$ и $\mathbf{y}$ в $\mathbb{R}^N$ обозначается $\langle \mathbf{x}, \mathbf{y} \rangle$, и является суммой произведения их элементов:

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{n=1}^{N} x_n y_n$$

Факт. (??)

Для любых $\alpha, \beta \in \mathbb{R}$ и любых $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, верны следующие утверждения:

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$,
2. $\langle \alpha \mathbf{x}, \beta \mathbf{y} \rangle = \alpha\beta \langle \mathbf{x}, \mathbf{y} \rangle$, и
3. $\langle \mathbf{x}, \alpha \mathbf{y} + \beta \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle + \beta \langle \mathbf{x}, \mathbf{z} \rangle$.

Свойства можно легко проверить с помощью определений умножения на скаляр и скалярного произведения.

Для 2., например, возьмите любые $\alpha, \beta \in \mathbb{R}$ и любые $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$:

$$\langle \alpha \mathbf{x}, \beta \mathbf{y} \rangle = \sum_{n=1}^{N} \alpha x_n \beta y_n = \alpha\beta \sum_{n=1}^{N} x_n y_n = \alpha\beta \langle \mathbf{x}, \mathbf{y} \rangle$$

Нормы и расстояние

Норма (Эвклида) $\mathbf{x} \in \mathbb{R}^N$ обозначается

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

Интерпретация:

- $\|\mathbf{x}\|$ показывает “длину” вектора $\mathbf{x}$
- $\|\mathbf{x} - \mathbf{y}\|$ показывает расстояние между векторами $\mathbf{x}$ и $\mathbf{y}$

Факт. (??) Для любого $\alpha \in \mathbb{R}$ и любых $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$, верны следующие утверждения:

1. $\|\mathbf{x}\| \geq 0$ и $\|\mathbf{x}\| = 0$ тогда и только тогда, когда $\mathbf{x} = \mathbf{0}$
2. $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (неравенство треугольника)
4. $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|$ (неравенство Коши — Буняковского)

Свойства 1. и 2. легко доказать (упражнение)

Свойство 4. рассматривается в ЕТ упражнении ??

Покажем доказательство свойства 3. с помощью свойств скалярного произведения (факт ??)

$$
\begin{aligned}
\|\mathbf{x}+\mathbf{y}\|^2 &= \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\
&\leq \langle \mathbf{x}, \mathbf{x} \rangle + 2|\langle \mathbf{x}, \mathbf{y} \rangle| + \langle \mathbf{y}, \mathbf{y} \rangle
\end{aligned}
$$

Применим неравенство Коши — Буняковского

$$\|\mathbf{x}+\mathbf{y}\|^2 \leq (\|\mathbf{x}\|+\|\mathbf{y}\|)^2$$

Убираем квадраты и получаем неравенство треугольника

Линейная комбинация векторов $\mathbf{x}_1, \ldots, \mathbf{x}_K$ in $\mathbb{R}^N$ — это вектор

$$\mathbf{y} = \sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_K \mathbf{x}_K$$

где $\alpha_1, \ldots, \alpha_K$ скаляры

Пример.

$$0.5 \begin{pmatrix} 6.0 \\ 2.0 \\ 8.0 \end{pmatrix} + 3.0 \begin{pmatrix} 0 \\ 1.0 \\ -1.0 \end{pmatrix} = \begin{pmatrix} 3.0 \\ 4.0 \\ 1.0 \end{pmatrix}$$

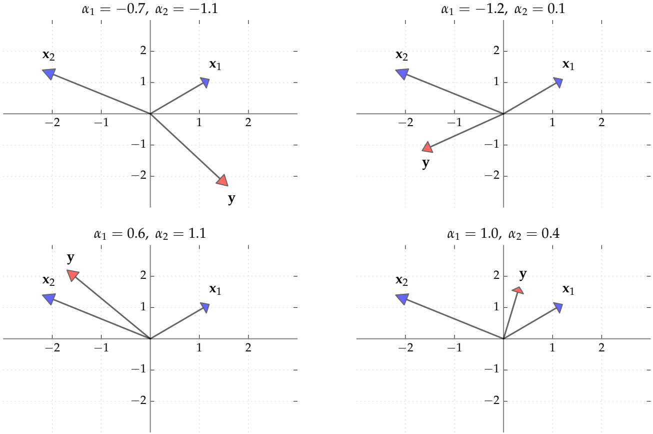


Рис.: Линейные комбинации $\mathbf{x}_1, \mathbf{x}_2$

Линейная оболочка

Пусть $X \subset \mathbb{R}^N$ некое непустое множество

Множество всех возможных линейных комбинаций X называют **линейной оболочкой** X, обозначается $\operatorname{span}(X)$

Для конечного $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\}$ линейную оболочку можно записать так

$$\operatorname{span}(X) := \left\{ \text{все } \sum_{k=1}^{K} \alpha_k \mathbf{x}_k, \text{ где } (\alpha_1, \ldots, \alpha_K) \in \mathbb{R}^K \right\}$$

Пример. Четыре вектора, обозначенные $\mathbf{y}$ на предыдущем рисунке, лежат в линейной оболочке $X = \{\mathbf{x}_1, \mathbf{x}_2\}$

Может ли *любой* вектор в $\mathbb{R}^2$ быть создан линейной комбинацией $\mathbf{x}_1, \mathbf{x}_2$?

Ответ да. Мы докажем это в §**??**

Пример. Пусть $X = \{\mathbf{1}\} \subset \mathbb{R}^2$, где $\mathbf{1} := (1, 1)$

Линейная оболочка X состоит из всех векторов вида

$$\alpha\mathbf{1} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} \quad , \text{ где } \quad \alpha \in \mathbb{R}$$

Представляет собой прямую на плоскости, которая проходит через

- вектор $\mathbf{1}$ (при $\alpha = 1$)
- начало координат $\mathbf{0}$ (при $\alpha = 0$)

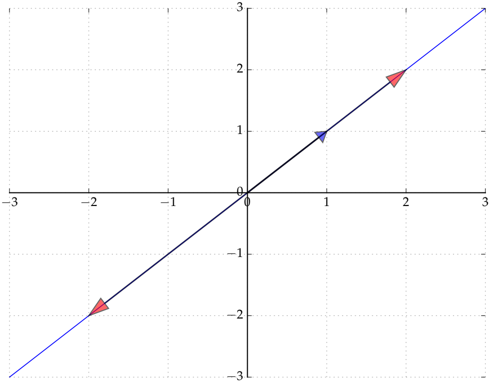


Рис.: Линейная оболочка $\mathbf{1} := (1,1)$ в $\mathbb{R}^2$

Пример. Набор канонически базисных векторов $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ линейно независим в $\mathbb{R}^N$

Доказательство. Пусть $\alpha_1, \ldots, \alpha_N$ коэффициенты, такие что $\sum_{k=1}^N \alpha_k \mathbf{e}_k = \mathbf{0}$

Эквивалентно,

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix} = \sum_{k=1}^N \alpha_k \mathbf{e}_k = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

В частности, $\alpha_k = 0$ для всех k

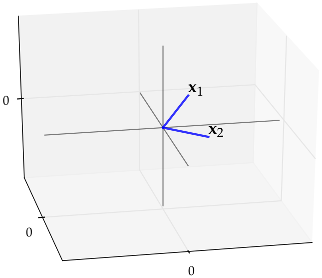
Пример. Пусть $\mathbf{x}_1 = (3, 4, 2)$ и $\mathbf{x}_2 = (3, -4, 0.4)$

По определению, линейная оболочка — все возможные вектора в виде

$$\mathbf{y} = \alpha \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} + \beta \begin{pmatrix} 3 \\ -4 \\ 0.4 \end{pmatrix} \quad \text{, где} \quad \alpha, \beta \in \mathbb{R}$$

Это плоскость, проходящая через

- вектор $\mathbf{x}_1$
- вектор $\mathbf{x}_2$
- начало координат $\mathbf{0}$



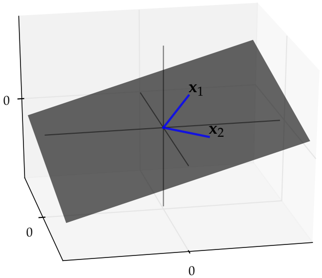


Рис.: Линейная оболочка $\mathbf{x}_1, \mathbf{x}_2$

Пример. Рассмотрим векторы $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\} \subset \mathbb{R}^N$, где

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \mathbf{e}_N := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

$\mathbf{e}_n$ состоит из нулей, кроме n-ого элемента равного 1

Вектора $\mathbf{e}_1, \ldots, \mathbf{e}_N$ назвают **канонически базисными векторами** of $\mathbb{R}^N$

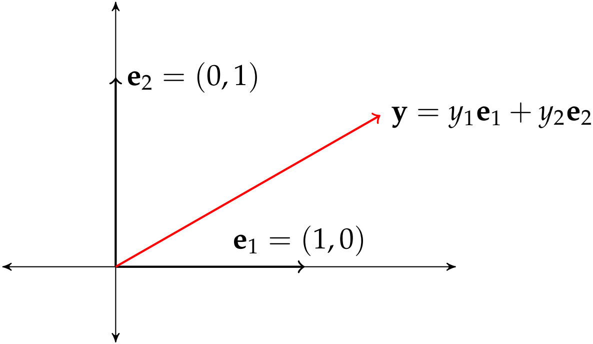


Рис.: Базисные векторы в $\mathbb{R}^2$

Пример. (прод.)

Линейная оболочка $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ эквивалентна всему $\mathbb{R}^N$

Доказательство для $N = 2$:

Пусть $\mathbf{y} \in \mathbb{R}^2$, тогда

$$\mathbf{y} := \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y_2 \end{pmatrix}$$

$$= y_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$$

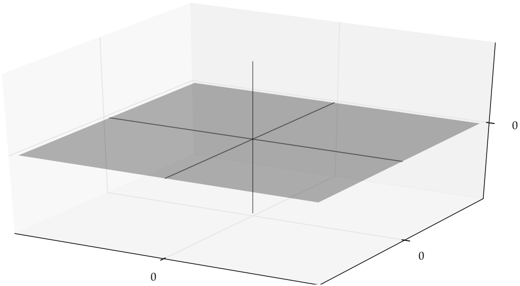
Таким образом, $\mathbf{y} \in \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\}$

Так как $\mathbf{y}$ произвольный, мы показали, что $\operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\} = \mathbb{R}^2$

Пример. Рассмотрим множество

$$P := \{(x_1, x_2, 0) \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R}\}$$

Графически, $P =$ плоскость в $\mathbb{R}^3$с координатой высоты $= 0$



Пример. (прод.)

Если $\mathbf{e}_1 = (1,0,0)$ и $\mathbf{e}_2 = (0,1,0)$, тогда $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} = P$

Чтобы подтвердить утверждение, возьмем $\mathbf{x} = (x_1, x_2, 0)$, любой элемент P. Можно записать $\mathbf{x}$ в следующем виде

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2$$

Другими словами, $P \subset \mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\}$

И наоборот, у нас имеется $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} \subset P$ (почему?)

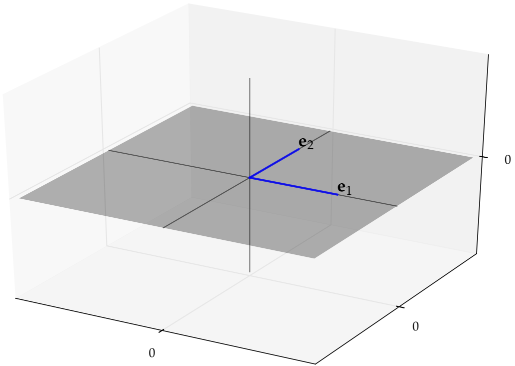


Рис.: $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} = P$

Факт. (??) Если X и Y непустые подмножества $\mathbb{R}^N$ и $X \subset Y$, тода $\operatorname{span}(X) \subset \operatorname{span}(Y)$

Доказательство. Возьмем любой непустой $X \subset Y \subset \mathbb{R}^N$

Пусть $\mathbf{z} \in \operatorname{span}(X)$, тогда имеется

$$\mathbf{z} = \sum_{k=1}^{K} \alpha_k \mathbf{x}_k \text{ для некоторых } \mathbf{x}_1, \ldots, \mathbf{x}_K \in X,\ \alpha_1, \ldots, \alpha_K \in \mathbb{R}$$

Доказательство. (прод.) Так как $X \subset Y$, каждый $\mathbf{x}_k$ также находится в Y, получается

$$\mathbf{z} = \sum_{k=1}^{K} \alpha_k \mathbf{x}_k \text{ для некоторых } \mathbf{x}_1, \ldots, \mathbf{x}_K \in Y,\ \alpha_1, \ldots, \alpha_K \in \mathbb{R}$$

Значит, $\mathbf{z} \in \operatorname{span}(Y)$

Линейная независимость

Важные вопросы:

- Когда матрица обратима?
- Когда аргументы регрессии страдают от коллинеарности?
- Когда система линейных уравнений имеет решение?

Все эти вопросы тесно связаны с линейной независимостью

Непустое множество векторов $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$ называется **линейно независимым**, если

$$\sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \mathbf{0} \implies \alpha_1 = \cdots = \alpha_K = 0$$

Пример. Рассмотрим $\mathbf{x}_1 = (1.2, 1.1)$ и $\mathbf{x}_2 = (-2.2, 1.4)$

Пусть α_1 и α_2 скаляры, такие что

$$\alpha_1 \begin{pmatrix} 1.2 \\ 1.1 \end{pmatrix} + \alpha_2 \begin{pmatrix} -2.2 \\ 1.4 \end{pmatrix} = \mathbf{0}$$

Это является линейной системой из двух уравнений α_1 и α_2

Единственное решение $\alpha_1 = \alpha_2 = 0$

$\{\mathbf{x}_1, \mathbf{x}_2\}$ линейно независимы

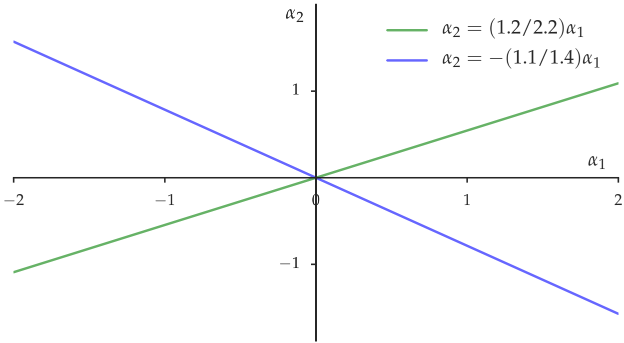


Рис.: Единственное решение $\alpha_1 = \alpha_2 = 0$

Пример. Базисные векторы $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ линейно независимы в $\mathbb{R}^N$

Проверим это. Пусть $\alpha_1, \ldots, \alpha_N$ коэффициенты, такие что $\sum_{k=1}^N \alpha_k \mathbf{e}_k = \mathbf{0}$, тогда

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix} = \sum_{k=1}^N \alpha_k \mathbf{e}_k = \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

То есть $\alpha_k = 0$ для всех k

Таким образом, $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ линейно независимы

Теорема. (??) Пусть $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$. Для $K > 1$, следующие утверждения эквивалентны:

1. X линейно независим
2. X_0 — подходящее подмножество $X \implies \operatorname{span} X_0$ — подходящее подмножество $\operatorname{span} X$
3. Ни один из векторов X не может быть записан как линейная комбинация оставшихся

Доказательство есть в упражнениях. См. ET упр. **??** и решение

Пример. Если убрать любой базисный вектор, линейная оболочка уменьшится

Рассмотрим случай с $N = 2$

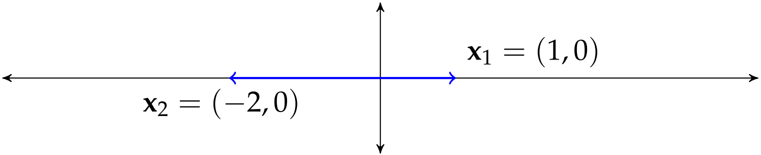
Мы знаем, что $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\} = \mathbb{R}^2$

- Если убрать любой элемент из $\mathrm{span}\{\mathbf{e}_1, \mathbf{e}_2\}$, линейная оболочка превратится в прямую.

Тем не менее, пусть $\mathbf{x}_1 = (1, 0)$ и $\mathbf{x}_2 = (-2, 0)$

Векторы не являются линейно независимыми, так как $\mathbf{x}_2 = -2\mathbf{x}_1$

- Если убрать любой из векторов, линейная оболочка не изменится — останется горизонтальной прямой
- имеется $\mathbf{x}_2 = -2\mathbf{x}_1$, значит любой вектор может быть записан как линейная комбинация оставшегося



Факт. (**??**) Если $X := \{\mathbf{x}_1, \dots, \mathbf{x}_K\}$ линейно независимы, тогда

1. любое подмножество X линейно независимо,
2. X не содержит $\mathbf{0}$, и
3. $X \cup \{\mathbf{x}\}$ линейно независимы для всех $\mathbf{x} \in \mathbb{R}^N$, таких что $\mathbf{x} \notin \operatorname{span} X$.

Доказательство показано в упражнении (упр. **??** в ET)

Линейная независимость и единственность

Линейная независимость - ключевое условие для того, чтобы решение системы линейных уравнений существовало *и* было единственным

Теорема. (??) Пусть $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\}$ некоторое множество векторов в $\mathbb{R}^N$. Следующие утверждения эквивалентны:

1. X линейно независим
2. Для каждого $\mathbf{y} \in \mathbb{R}^N$ существует не более одного множества скаляров $\alpha_1, \ldots, \alpha_K$, такого что

$$\mathbf{y} = \alpha_1 \mathbf{x}_1 + \cdots + \alpha_K \mathbf{x}_K \tag{1}$$

Доказательство. (1. $\implies$ 2.)

Пусть X линейно независим, возьмем любой $\mathbf{y}$

Предположим противоположное — (1) выполняется для нескольких наборов скаляров, получается

$$\exists\ \alpha_1, \ldots, \alpha_K \text{ и } \beta_1, \ldots, \beta_K \ \text{ s.t. } \ \mathbf{y} = \sum_{k=1}^{K} \alpha_k \mathbf{x}_k = \sum_{k=1}^{K} \beta_k \mathbf{x}_k$$

$$\therefore \quad \sum_{k=1}^{K} (\alpha_k - \beta_k) \mathbf{x}_k = \mathbf{0}$$

$$\therefore \quad \alpha_k = \beta_k \quad \text{для всех} \quad k$$

Доказательство. (2. $\implies$ 1.)

Если 2. выполняется, то существует не более одного набора скаляров, такого что

$$\mathbf{0} = \sum_{k=1}^{K} \alpha_k \mathbf{x}_k$$

Так как при $\alpha_1 = \cdots = \alpha_k = 0$ это равенство выполняется, больше не существует скаляров, при которых $\mathbf{0} = \sum_{k=1}^{K} \alpha_k \mathbf{x}_k$

Значит, X линейно независим по определению

Линейные подпространства

Непустое подмножество S множества $\mathbb{R}^N$ называется **линейным подпространством** (или просто **подпространством**) множества $\mathbb{R}^N$, если

$$\mathbf{x}, \mathbf{y} \in S \text{ и } \alpha, \beta \in \mathbb{R} \implies \alpha\mathbf{x} + \beta\mathbf{y} \in S$$

Другими словами, $S \subset \mathbb{R}^N$ 'заперт' с векторным сложением и умножением на скаляр

Пример. Если X — некое непустое подмножество $\mathbb{R}^N$, тогда $\operatorname{span} X$ — линейное подпространство $\mathbb{R}^N$

Пример. $\mathbb{R}^N$ — линейное подпространство $\mathbb{R}^N$

Пример. Возьмем любой вектор $\mathbf{a} \in \mathbb{R}^N$, множество $A := \{\mathbf{x} \in \mathbb{R}^N : \langle \mathbf{a}, \mathbf{x} \rangle = 0\}$ является линейным подпространством $\mathbb{R}^N$

Чтобы показать это, пусть $\mathbf{x}, \mathbf{y} \in A$, $\alpha, \beta \in \mathbb{R}$ и $\mathbf{z} := \alpha\mathbf{x} + \beta\mathbf{y} \in A$

Получается

$$\langle \mathbf{a}, \mathbf{z} \rangle = \langle \mathbf{a}, \alpha\mathbf{x} + \beta\mathbf{y} \rangle = \alpha \langle \mathbf{a}, \mathbf{x} \rangle + \beta \langle \mathbf{a}, \mathbf{y} \rangle = 0 + 0 = 0$$

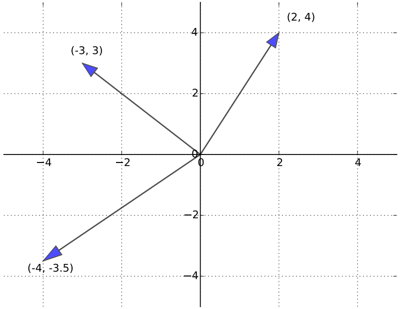
Значит, $\mathbf{z} \in A$

Факт. (??) Если S — линейное подпространство $\mathbb{R}^N$, тогда

1. $\mathbf{0} \in S$
2. $X \subset S \implies \operatorname{span} X \subset S$, и
3. $\operatorname{span} S = S$

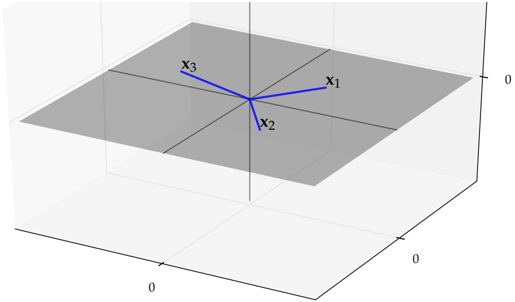
Теорема. (??) Пусть S — линейное подпространство $\mathbb{R}^N$. Если S охватывает K векторов, тогда любое линейно независимое подмножество S имеет не более K векторов

Пример. Вспомним базисные векторы $\{\mathbf{e}_1, \mathbf{e}_2\}$, охватывающие $\mathbb{R}^2$. Из Теоремы **??** следует, что три вектора ниже линейно зависимы



Пример. Плоскость $P := \{(x_1, x_2, 0) \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R}\}$ из примера **??** в ЕТ может быть образована двумя векторами

По теореме **??**, три вектора на рисунке ниже линейно независимы



Базисы и Размерность

Теорема. (??) Пусть $X := \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ — некие N векторов в $\mathbb{R}^N$. Следующие утверждения эквивалентны:

1. $\operatorname{span} X = \mathbb{R}^N$
2. X линейно независим

Для доказательства смотрите **??** в ET

Пусть S — линейное подпространство $\mathbb{R}^N$ и $B \subset S$

Множество B называется **базисом** S, если

1. B охватывает S и
2. B линейно независимо

Из Теоремы **??**, когда B является базисом S, каждая точка S имеет ровно представление как линейной комбинации элементов B

Из Теоремы **??**, любые N линейно независимых векторов в $\mathbb{R}^N$ формируют базис в $\mathbb{R}^N$

Пример. Вспомним плоскость из примера выше

$$P := \{(x_1, x_2, 0) \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R}\}$$

Мы показали, что $\operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\} = P$ для

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Более того, $\{\mathbf{e}_1, \mathbf{e}_2\}$ линейно независимы (почему?)

Значит, $\{\mathbf{e}_1, \mathbf{e}_2\}$ являются базисом для P

Теорема. (**??**) Если S — линейное подпространство $\mathbb{R}^N$ отличное от $\{\mathbf{0}\}$, тогда

1. S имеет по меньшей мере один базис и
2. каждый базис S имеет одинаковое количество элементов.

Если S — линейное подпространство $\mathbb{R}^N$, то обычное число, определенное в Теореме **??** называется **размерностью** S, и обозначается $\dim S$

Пример. Для $P := \{(x_1, x_2, 0) \in \mathbb{R}^3 : x_1, x_2 \in \mathbb{R}\}$, $\dim P = 2$, потому что

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

базис имеет два элемента

Пример. Прямая $\{\alpha\mathbf{x} \in \mathbb{R}^N : \alpha \in \mathbb{R}\}$, проходящая через начало координат, имеет размерность 1

Говорят, что в $\mathbb{R}^N$ одноэлементное подпространство $\{\mathbf{0}\}$ имеет нулевую размерность

Возьмем множество из K векторов, насколько большой будет его линейная оболочка с точки зрения размерности?

Теорема. (??) Если $X := \{\mathbf{x}_1, \ldots, \mathbf{x}_K\} \subset \mathbb{R}^N$, то

1. $\dim \operatorname{span} X \leq K$ и
2. $\dim \operatorname{span} X = K$ тогда и только тогда, когда X линейно независим

Для доказательства смотрите упражнение **??** в ET

Факт. (??) Следующие утверждения верны:

1. Пусть S и S' являются линейными подпространствами $\mathbb{R}^N$ размерности K. Если $S \subset S'$, то $S = S'$
2. Если S — линейное подпространство $\mathbb{R}^N$ размерности M и $M < N$, то $S \neq \mathbb{R}^N$

Линейные отображения

Функция $T\colon \mathbb{R}^K \to \mathbb{R}^N$ называется **линейной**, если

$$T(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha T\mathbf{x} + \beta T\mathbf{y} \qquad \forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^K,\ \forall\, \alpha, \beta \in \mathbb{R}$$

Примечание:

- Линейные функции обычно записываются с большой буквы
- Обычно, когда это удобно, аргументы в скобках опускаются

Пример. Функция $T\colon \mathbb{R} \to \mathbb{R}$, определяемая как $Tx = 2x$, линейна

Чтобы увидеть это, возьмем любые α, β, x, y в $\mathbb{R}$, тогда

$$T(\alpha x + \beta y) = 2(\alpha x + \beta y) = \alpha 2x + \beta 2y = \alpha Tx + \beta Ty$$

Пример. Функция $f\colon \mathbb{R} \to \mathbb{R}$, определяемая как $f(x) = x^2$, нелинейна

Чтобы увидеть это, возьмем $\alpha = \beta = x = y = 1$, получается

$$f(\alpha x + \beta y) = f(2) = 4$$

Тем не менее, $\alpha f(x) + \beta f(y) = 1 + 1 = 2$

Примечание: Неправильно думать о линейных функциях как о тех, чей график является прямой

Пример. Функция $f : \mathbb{R} \to \mathbb{R}$, определяемая $f(x) = 1 + 2x$, нелинейна

Возьмем $\alpha = \beta = x = y = 1$. Имеется

$$f(\alpha x + \beta y) = f(2) = 5$$

Тем не менее, $\alpha f(x) + \beta f(y) = 3 + 3 = 6$

Такой вид функции называется **афинной** функцией

По определению, если T линейна, то изменение порядка в

$$T[\sum_{k=1}^{K} \alpha_k \mathbf{x}_k] = \sum_{k=1}^{K} \alpha_k T\mathbf{x}_k$$

будет действительным пока $K = 2$

Индукция расширяет это до произвольных K

Факт. (??) Если $T\colon \mathbb{R}^K \to \mathbb{R}^N$ является линейным отображением, то

$$\mathrm{rng}(T) = \mathrm{span}(V) \quad \text{, где} \quad V := \{T\mathbf{e}_1, \ldots, T\mathbf{e}_K\}$$

где $\mathbf{e}_k$ является k-ым базисным вектором в $\mathbb{R}^K$

Доказательство. Любой $\mathbf{x} \in \mathbb{R}^K$ может быть выражен как $\sum_{k=1}^K \alpha_k \mathbf{e}_k$. Значит, $\mathrm{rng}(T)$ является множеством всех точек следующей формы

$$T\mathbf{x} = T\left[\sum_{k=1}^K \alpha_k \mathbf{e}_k\right] = \sum_{k=1}^K \alpha_k T\mathbf{e}_k$$

так как $\alpha_1, \ldots, \alpha_K$ меняется для всех комбинаций. Это совпадает с определением $\mathrm{span}(V)$

Ядром линейного отображения $T\colon \mathbb{R}^K \to \mathbb{R}^N$ называют

$$\ker(T) := \{\mathbf{x} \in \mathbb{R}^K : T\mathbf{x} = \mathbf{0}\}$$

Факт. (??) Если $T\colon \mathbb{R}^K \to \mathbb{R}^N$ линейное отображение, то

$$\operatorname{rng} T = \operatorname{span} V, \quad \text{где } V := \{T\mathbf{e}_1, \ldots, T\mathbf{e}_K\}$$

Доказательства простые (выполните в качестве упражнения)

Линейная независимость и биекция

Много научных и практических задач являются задачами 'обратными'

- мы наблюдаем результаты, но не их причины
- как можно работать в обратном порядке, от результатов к причинам?

Примеры

- какие предпочтения потребителей привели к наблюдаемому рыночному поведению?
- какие ожидания привели к данному сдвигу обменных курсов?

В общем, обратную задачу можно выразить как

outcome

model

$$F(x) = y$$

what x led to outcome y?

- имеет ли эта задача решение?
- является ли оно едиственным?

Ответы зависят от того, является ли F сюръекцией, инъекцией и т.д.

Лучший вариант — биекция

Но возникают и другие ситуации

Теорема. (**??**) Если T — линейная функция из $\mathbb{R}^N$ в $\mathbb{R}^N$, то все следующие утверждения эквивалентны:

1. T является биекцией.
2. T является сюръекцией.
3. T является инъекцией.
4. $\ker T = \{\mathbf{0}\}$.
5. $V := \{T\mathbf{e}_1, \ldots, T\mathbf{e}_N\}$ линейно независимо.
6. $V := \{T\mathbf{e}_1, \ldots, T\mathbf{e}_N\}$ формирует базис $\mathbb{R}^N$.

Смотрите упражнение **??** в ЕТ для доказательства

Если любое из этих условий выполняется, то T называют **несингулярной**. В ином случае T назвают **сингулярной**

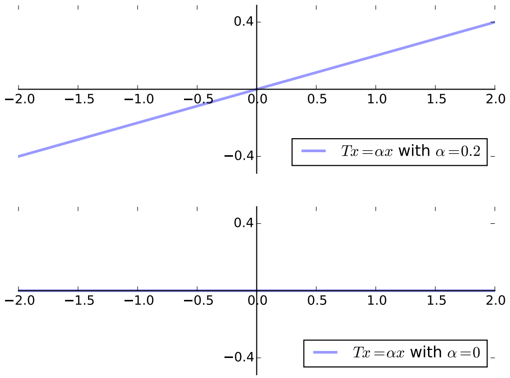


Рис.: Случай с $N = 1$, несингулярная и сингулярная функции

Если T несингулярна, то, будучи биекцией, она должна иметь обратную функцию T^{-1}, которая так же является биекцией (факт **??** на странице **??**)

Факт. (??) Если $T\colon \mathbb{R}^N \to \mathbb{R}^N$ несингулярна, то и T^{-1} несингулярна.

Для доказательства, смотрите упр. **??**

Отображения при различных размерностях

Помните, что результаты выше применимы к отображениям из $\mathbb{R}^N$ в $\mathbb{R}^N$

Все меняется, когда мы смотрим на линейные отображения при различных размерностях

Общие правили для линейных отображений:

- отображения из меньших в большие размерности не могут быть сюръекцией
- отображения из больших в меньшие размерности не могут быть инъекцией

Ни один из случаев не может быть биекцией

Теорема. (??) Для линейного отображения T из $\mathbb{R}^K \to \mathbb{R}^N$, следующие утверждения верны:

1. Если $K < N$, то T не сюръекция.
2. Если $K > N$, то T не инъекция.

Доказательство. (часть 1)

Пусть $K < N$ и отображение $T \colon \mathbb{R}^K \to \mathbb{R}^N$ линейное

Пусть $V := \{T\mathbf{e}_1, \ldots, T\mathbf{e}_K\}$, имеется

$$\dim(\operatorname{rng}(T)) = \dim(\operatorname{span}(V)) \leq K < N$$

$$\therefore \quad \operatorname{rng}(T) \neq \mathbb{R}^N$$

Значит, T не сюръекция

Доказательство. (часть 2)

Предположим обратное, что T является инъекцией

Пусть $\alpha_1, \dots, \alpha_K$ набор векторов, такой что

$$\alpha_1 T\mathbf{e}_1 + \cdots + \alpha_K T\mathbf{e}_K = \mathbf{0}$$

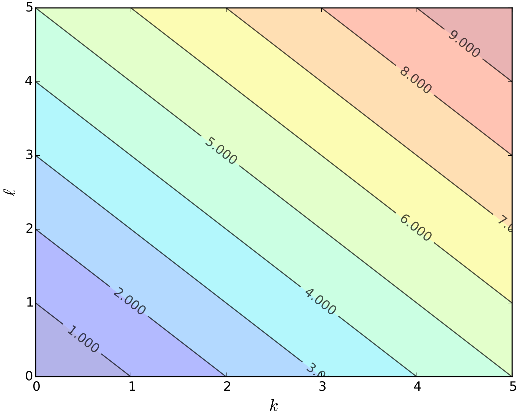
$$\therefore \quad T(\alpha_1 \mathbf{e}_1 + \cdots + \alpha_K \mathbf{e}_K) = \mathbf{0} \qquad \text{(по линейности)}$$

$$\therefore \quad \alpha_1 \mathbf{e}_1 + \cdots + \alpha_K \mathbf{e}_K = \mathbf{0} \qquad \text{(так как } \ker(T) = \{\mathbf{0}\}\text{)}$$

$$\therefore \quad \alpha_1 = \cdots = \alpha_K = 0 \qquad \text{(по независимости } \{\mathbf{e}_1, \dots \mathbf{e}_K\}\text{)}$$

Мы показали, что $\{T\mathbf{e}_1, \dots, T\mathbf{e}_K\}$ линейно независимы

Но тогда $\mathbb{R}^N$ содержит линейно независимое множество с $K > N$ векторами — противоречие



Пример. Функция издержек $c(k, \ell) = rk + w\ell$ не может быть инъекцией

Ортогональные Векторы и проекции

Ключевой концепцией курса является ортогональность – не только векторов, но и случайных величин

Пусть $\mathbf{x}$ и $\mathbf{z}$ векторы в $\mathbb{R}^N$

Если $\langle \mathbf{x}, \mathbf{z} \rangle = 0$, то мы называем $\mathbf{x}$ и $\mathbf{z}$ **ортогональными**

Записывается $\mathbf{x} \perp \mathbf{z}$

В $\mathbb{R}^2$ ортогональный значит перпендикулярный

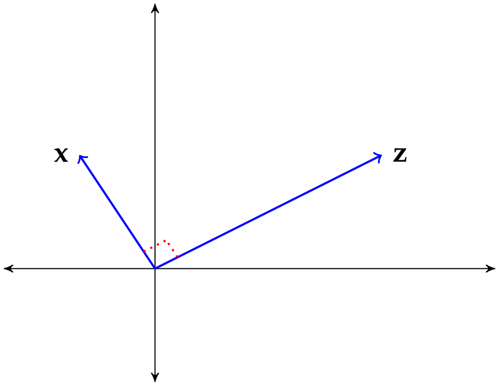


Рис.: $\mathbf{x} \perp \mathbf{z}$

Пусть S линейное подпространство

Говорят, что **$\mathbf{x}$ ортогонален S**, если $\mathbf{x} \perp \mathbf{z}$ для всех $\mathbf{z} \in S$

Записывается $\mathbf{x} \perp S$

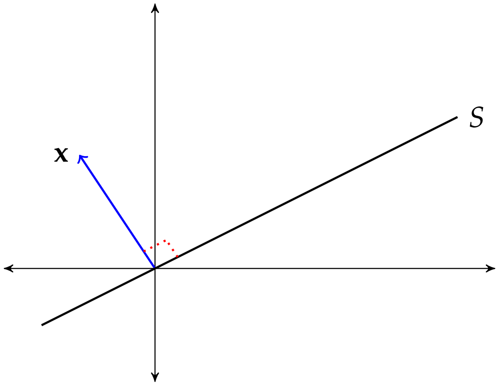


Рис.: $\mathbf{x} \perp S$

Факт. (??) (Теорема Пифагора)

Если $\{\mathbf{z}_1, \ldots, \mathbf{z}_K\}$ является ортогональным множеством, то

$$\|\mathbf{z}_1 + \cdots + \mathbf{z}_K\|^2 = \|\mathbf{z}_1\|^2 + \cdots + \|\mathbf{z}_K\|^2$$

Докажите в качестве упражнения

Факт. (??) Если $O \subset \mathbb{R}^N$ является ортогональным множеством и $\mathbf{0} \notin O$, то O линейно независимо

Ортогональной множество $O \subset \mathbb{R}^N$ называется **ортонормальное множество** Если $\|\mathbf{u}\| = 1$ для всех $\mathbf{u} \in O$

Ортонормальное множество, охватывающее линейное подпространство S в $\mathbb{R}^N$ называют **ортонормальным базисом** S

- примером ортонормального базиса для всего в $\mathbb{R}^N$ является канонический базис $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$

Факт. (??) Если $\{\mathbf{u}_1, \ldots, \mathbf{u}_K\}$ является ортонормальным множеством и $\mathbf{x} \in \operatorname{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_K\}$, то

$$\mathbf{x} = \sum_{k=1}^{K} \langle \mathbf{x}, \mathbf{u}_k \rangle \mathbf{u}_k$$

Возьмем $S \subset \mathbb{R}^N$, **ортогональное дополнение** S будет

$$S^\perp := \{\mathbf{x} \in \mathbb{R}^N : \mathbf{x} \perp S\}$$

Факт. (??) Для любого непустого $S \subset \mathbb{R}^N$, пространство $S^\perp$ является линейным подпространством $\mathbb{R}^N$

Доказательство. Если $\mathbf{x}, \mathbf{y} \in S^\perp$ и $\alpha, \beta \in \mathbb{R}$, то $\alpha\mathbf{x} + \beta\mathbf{y} \in S^\perp$, так как для любых $\mathbf{z} \in S$

$$\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle = \alpha \times 0 + \beta \times 0 = 0$$

Факт. (??) Для $S \subset \mathbb{R}^N$, выполняется $S \cap S^\perp = \{\mathbf{0}\}$

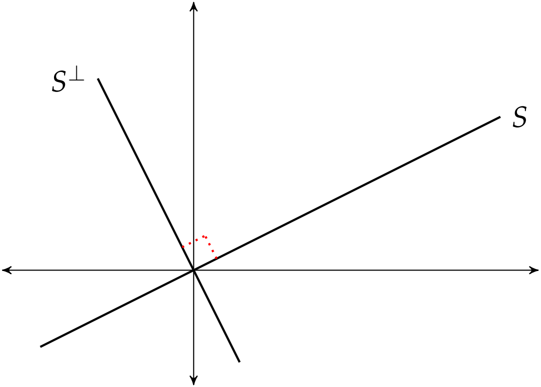


Рис.: Ортогональное дополнение S в $\mathbb{R}^2$

Теорема ортогональной проекции

Задача:

При данном $\mathbf{y} \in \mathbb{R}^N$ и подпространстве S, найти ближайший элемент S к $\mathbf{y}$

Формально: Решаем

$$\hat{\mathbf{y}} := \operatorname*{argmin}_{\mathbf{z} \in S} \|\mathbf{y} - \mathbf{z}\| \tag{2}$$

Существование, единственность решения не очевидны

Теорема ортогональной проекции: $\hat{\mathbf{y}}$ всегда существует, причем в единственном числе

Также дает полезную характеристику

Теорема. (??) [Теорема ортогональной проекции I]

Пусть $\mathbf{y} \in \mathbb{R}^N$ и S является непустым линейным подпространством в $\mathbb{R}^N$.

Следующие утверждения верны:

1. Задача оптимизации (2) имеет ровно одно решение
2. $\hat{\mathbf{y}} \in \mathbb{R}^N$ является решением (2) тогда и только тогда, когда $\hat{\mathbf{y}} \in S$ и $\mathbf{y} - \hat{\mathbf{y}} \perp S$

Единственное решение $\hat{\mathbf{y}}$ называется **ортогональной проекцией y на S**

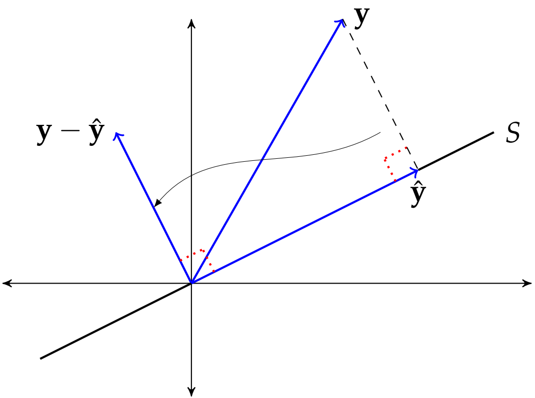


Рис.: Ортогональная проекция

Доказательство. (достаточности 2.)

Пусть $\mathbf{y} \in \mathbb{R}^N$ и S является линейным подпространством в $\mathbb{R}^N$

Пусть $\hat{\mathbf{y}}$ — вектор в S, удовлетворяющий условию $\mathbf{y} - \hat{\mathbf{y}} \perp S$

Пусть $\mathbf{z}$ — некая точка в S. Получается

$$\|\mathbf{y} - \mathbf{z}\|^2 = \|(\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{z})\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{z}\|^2$$

Второе равенство следует из $\mathbf{y} - \hat{\mathbf{y}} \perp S$ и Теоремы Пифагора

Так как $\mathbf{z}$ был произвольной точкой в S, получается
$\|\mathbf{y} - \mathbf{z}\| \geq \|\mathbf{y} - \hat{\mathbf{y}}\|$ для всех $\mathbf{z} \in S$

Пример. Пусть $\mathbf{y} \in \mathbb{R}^N$ и $\mathbf{1} \in \mathbb{R}^N$ является вектором из единиц

Пусть S является множеством постоянных векторов в $\mathbb{R}^N$ — S является линейной оболочкой $\{\mathbf{1}\}$

Ортогональная проекция $\mathbf{y}$ на S — это $\hat{\mathbf{y}} := \bar{y}\mathbf{1}$, где $\bar{y} := \frac{1}{N}\sum_{n=1}^{N} y_n$

Ясно, что $\hat{\mathbf{y}} \in S$

Чтобы показать, что $\mathbf{y} - \hat{\mathbf{y}}$ ортогонален к S, нужно проверить $\langle \mathbf{y} - \hat{\mathbf{y}}, \mathbf{1} \rangle = 0$ (смотрите упр. **??** на странице **??**). Это верно, так как

$$\langle \mathbf{y} - \hat{\mathbf{y}}, \mathbf{1} \rangle = \langle \mathbf{y}, \mathbf{1} \rangle - \langle \hat{\mathbf{y}}, \mathbf{1} \rangle = \sum_{n=1}^{N} y_n - \bar{y} \langle \mathbf{1}, \mathbf{1} \rangle = 0$$

Если зафиксировать подпространство S, мы получим функциональную связь

$$\mathbf{y} \;\mapsto\; \text{его ортогональная проекция } \hat{\mathbf{y}} \in S$$

Это четко определенная функция из $\mathbb{R}^N$ в $\mathbb{R}^N$

Функция обычно обозначается как $\mathbf{P}$

- $\mathbf{P}(\mathbf{y})$ или $\mathbf{P}\mathbf{y}$ представляет $\hat{\mathbf{y}}$

$\mathbf{P}$ называется **ортогональным проекционным отображением на S**, записывается как

$$\mathbf{P} = \operatorname{proj} S$$

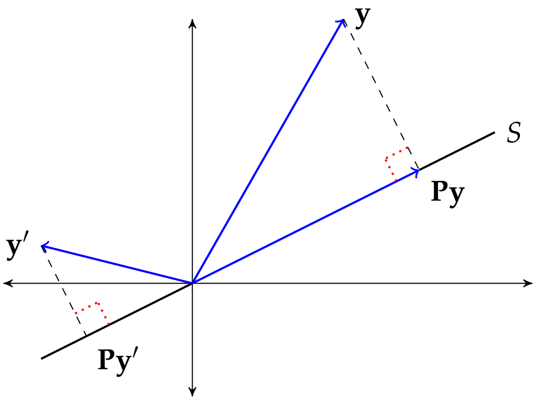


Рис.: Ортогональная проекция под $\mathbf{P}$

Теорема. (??) [Теорема ортогональной проекции II]

Пусть S является неким линейным подпространством $\mathbb{R}^N$, и $\mathbf{P} = \operatorname{proj} S$. Следующие утверждения верны:

1. $\mathbf{P}$ — линейная функция

Более того, для любого $\mathbf{y} \in \mathbb{R}^N$, соблюдается

2. $\mathbf{Py} \in S$,
3. $\mathbf{y} - \mathbf{Py} \perp S$,
4. $\|\mathbf{y}\|^2 = \|\mathbf{Py}\|^2 + \|\mathbf{y} - \mathbf{Py}\|^2$,
5. $\|\mathbf{Py}\| \leq \|\mathbf{y}\|$,
6. $\mathbf{Py} = \mathbf{y}$ тогда и только тогда, когда $\mathbf{y} \in S$, и
7. $\mathbf{Py} = \mathbf{0}$ тогда и только тогда, когда $\mathbf{y} \in S^{\perp}$.

Для доказательства смотрите страницу ?? и упражнение ??

Ниже приводится основополагающий результат

Факт. (??) Если $\{\mathbf{u}_1, \ldots, \mathbf{u}_K\}$ является ортонормальным базисом для S, то для каждого $\mathbf{y} \in \mathbb{R}^N$,

$$\mathbf{P}\mathbf{y} = \sum_{k=1}^{K} \langle \mathbf{y}, \mathbf{u}_k \rangle \, \mathbf{u}_k \tag{3}$$

Доказательство. Для начала, правая сторона (3) находится в S, так как это линейная комбинация векторов, охватывающих S

Далее, мы знаем, что $\mathbf{y} - \mathbf{P}\mathbf{y} \perp S$ тогда и только тогда, когда $\mathbf{y} - \mathbf{P}\mathbf{y} \perp \mathbf{u}_j$ для каждого $\mathbf{u}_j$ из множества базисных векторов (упражнение упр. **??**)

Дл любого $\mathbf{y} - \mathbf{P}\mathbf{y} \perp \mathbf{u}_j$, выполняется слудующее

$$\langle \mathbf{y} - \mathbf{P}\mathbf{y}, \mathbf{u}_j \rangle = \langle \mathbf{y}, \mathbf{u}_j \rangle - \sum_{k=1}^{K} \langle \mathbf{y}, \mathbf{u}_k \rangle \langle \mathbf{u}_k, \mathbf{u}_j \rangle$$

$$= \langle \mathbf{y}, \mathbf{u}_j \rangle - \langle \mathbf{y}, \mathbf{u}_j \rangle = 0$$

Это подтверждает, что $\mathbf{y} - \mathbf{P}\mathbf{y} \perp S$

Факт. (??) Пусть S_i является линейным подпространством $\mathbb{R}^N$ для $i = 1, 2$ и $\mathbf{P}_i = \operatorname{proj} S_i$. Если $S_1 \subset S_2$, то

$$\mathbf{P}_1\mathbf{P}_2\mathbf{y} = \mathbf{P}_2\mathbf{P}_1\mathbf{y} = \mathbf{P}_1\mathbf{y} \quad \text{для всех } \mathbf{y} \in \mathbb{R}^N$$

Остаточная проекция

Спроецируем $\mathbf{y}$ на S, где S является линейным подпространством $\mathbb{R}^N$

- Ближайшая точка к $\mathbf{y}$ на S — это $\hat{\mathbf{y}} := \mathbf{P}\mathbf{y}$, здесь $\mathbf{P} = \operatorname{proj} S$
- Если $\mathbf{y}$ не находится в S, ошибка $\mathbf{y} - \mathbf{P}\mathbf{y}$ существует

Введем оператор $\mathbf{M}$, который берет $\mathbf{y} \in \mathbb{R}^N$ и возвращает остаток

$$\mathbf{M} := \mathbf{I} - \mathbf{P} \tag{4}$$

где $\mathbf{I}$ является тождественным отображением $\mathbb{R}^N$

Для любых $\mathbf{y}$ выполняется $\mathbf{My} = \mathbf{Iy} - \mathbf{Py} = \mathbf{y} - \mathbf{Py}$

В регрессионном анализе $\mathbf{M}$ проявляется как матрица, называемая 'аннигилятором'

Мы говорим о $\mathbf{M}$ как об остаточной проекции

Пример. Вспомним, что проекция $\mathbf{y} \in \mathbb{R}^N$ на $\mathrm{span}\{\mathbf{1}\}$ — это $\bar{y}\mathbf{1}$

Остаточная проекция $\mathbf{M}_c\mathbf{y} := \mathbf{y} - \bar{y}\mathbf{1}$

- вектор ошибок, полученный, когда элементы вектора предсказываются его выборочным средним

Факт. (??) Пусть S является подпространством $\mathbb{R}^N$, $\mathbf{P} = \operatorname{proj} S$, и $\mathbf{M}$ является остаточной проекцией, определенной в (4). Верны следующие утверждения:

1. $\mathbf{M} = \operatorname{proj} S^{\perp}$
2. $\mathbf{y} = \mathbf{Py} + \mathbf{My}$ для любых $\mathbf{y} \in \mathbb{R}^N$
3. $\mathbf{Py} \perp \mathbf{My}$ для любых $\mathbf{y} \in \mathbb{R}^N$
4. $\mathbf{My} = \mathbf{0}$ тогда и только тогда, когда $\mathbf{y} \in S$
5. $\mathbf{P} \circ \mathbf{M} = \mathbf{M} \circ \mathbf{P} = \mathbf{0}$

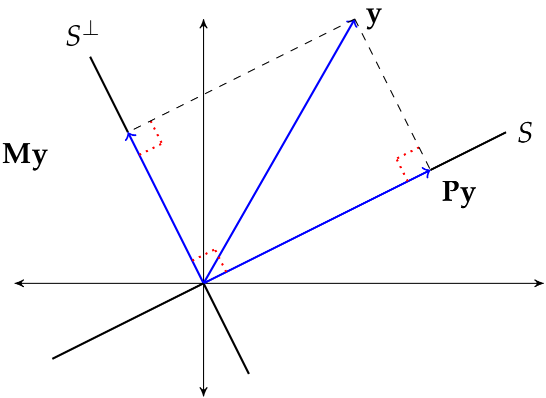


Рис.: Остаточная проекция

Если S_1 и S_2 — два подпространства $\mathbb{R}^N$ и $S_1 \subset S_2$, то $S_2^\perp \subset S_1^\perp$

Результат факта **??** обратный для $\mathbf{M}$

Факт. (**??**) Пусть S_1 и S_2 — два подпространства $\mathbb{R}^N$ и $\mathbf{y} \in \mathbb{R}^N$. Пусть $\mathbf{M}_1$ и $\mathbf{M}_2$ являются проекциями на $S_1^\perp$ и $S_2^\perp$ соответственно. Если $S_1 \subset S_2$, то

$$\mathbf{M}_1\mathbf{M}_2\mathbf{y} = \mathbf{M}_2\mathbf{M}_1\mathbf{y} = \mathbf{M}_2\mathbf{y}$$

Ортогонализация Грама-Шмидта

Вспомним, что любое ортогональное множество $\mathbb{R}^N$ не содержащее $\mathbf{0}$ линейно независимо – факт **??**

Вот (важное) частично обратное этому утверждению

Теорема. (**??**) Для каждого линейно независимого множества $\{\mathbf{b}_1, \ldots, \mathbf{b}_K\} \subset \mathbb{R}^N$, существует ортонормальное множество $\{\mathbf{u}_1, \ldots, \mathbf{u}_K\}$, такое что

$$\operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_k\} = \operatorname{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} \quad \text{для} \quad k = 1, \ldots, K$$

Формальные доказательства решаются как упражнения **??** to **??**

Доказательство дает важный алгоритм построения ортонормированного множества $\{\mathbf{u}_1, \ldots, \mathbf{u}_K\}$

Первый шаг — построить ортогональные множества $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ с линейной оболочкой, идентичной $\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ для каджого k

Построение $\{\mathbf{v}_1, \ldots, \mathbf{v}_K\}$ использует **ортогонализацию Грама-Шмидта**:

Для каждого $k = 1, \ldots, K$, пусть

1. $B_k := \operatorname{span}\{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$,
2. $\mathbf{P}_k := \operatorname{proj} B_k$ и $\mathbf{M}_k := \operatorname{proj} B_k^{\perp}$,
3. $\mathbf{v}_k := \mathbf{M}_{k-1}\mathbf{b}_k$, где $\mathbf{M}_0$ является тождественным отображением, и
4. $V_k := \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

В шаге 3. мы сопоставляем каждый последующий элемент $\mathbf{b}_k$ в подпространство, ортогональное подпространству, созданному с помощью $\mathbf{b}_1, \ldots, \mathbf{b}_{k-1}$

В завершение, определим $\mathbf{u}_k$ с помощью $\mathbf{u}_k := \mathbf{v}_k / \|\mathbf{v}_k\|$

Множество векторов $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ является ортонормальным с линейной оболочкой, равной V_k