线性代数

向量、矩阵与线性变换

XYZW 丛书

目录

前言 **1**

怎样读这本书 1

记号 1

与丛书其他各本的关系 2

第 1 章 向量 **3**

1.1 几何直觉 3

1.2 向量空间 3

1.3 线性组合 3

1.4 长度与内积 4

第 2 章 矩阵 **5**

2.1 线性组合的记账方式 5

2.2 矩阵乘法 5

2.3 结合律 5

2.4 常见的矩阵 6

前言

线性代数研究「可以相加、可以伸缩」的对象，以及保持这两种运算的映射。它是几何、微积分、概率统计乃至机器学习共用的语言。

怎样读这本书

每一章都按同样的顺序展开：先给一幅图，再给定义，最后给例子。第一次读时可以跳过证明，但不要跳过例子。

先画图

遇到新定义，先在平面 $\mathbb{R}^2$ 里把它画出来。大多数命题在二维里看懂了，推广到 $\mathbb{R}^n$ 只是记号问题。

记号

记号	含义
$\mathbb{R}^n$	n 维实向量空间
u, v, w	向量
A, B	矩阵
$\lVert v \rVert$	向量 v 的长度

本书把向量写成列的形式[1]，例如

$$v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

[1] 有些教材写成行向量，两种写法只差一次转置。

与丛书其他各本的关系

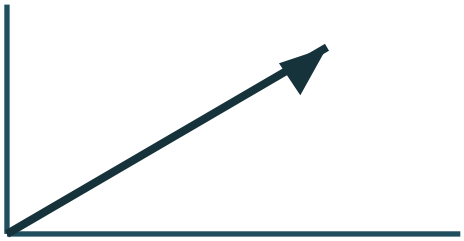
极限的严格定义见《微积分》的 极限。期望的线性性（见 随机变量 › 期望的线性性）本质上说的是：期望是一个线性映射。

准备好了就从向量 开始。

第 1 章　向量

1.1　几何直觉

在平面上，向量是一支带方向的箭头。两支箭头首尾相接就是相加，把箭头拉长、缩短或反向就是数乘。



1.2　向量空间

向量空间

设 V 是一个集合，配有加法 $V \times V \to V$ 与数乘 $\mathbb{R} \times V \to V$。若对任意 $u, v, w \in V$ 与 $a, b \in \mathbb{R}$，交换律、结合律、零元、负元与分配律都成立，则称 V 是 $\mathbb{R}$ 上的**向量空间**。

最常见的例子是 $\mathbb{R}^n$：分量逐个相加，逐个乘以同一个数。

1.3　线性组合

线性组合

给定向量 $v_1, \dots, v_k$ 与数 $a_1, \dots, a_k$，向量

$$w = a_1 v_1 + a_2 v_2 + \cdots + a_k v_k$$

称为 $v_1, \dots, v_k$ 的一个**线性组合**。

全体线性组合构成的集合叫作这些向量的**张成**，记作 $\operatorname{span}(v_1, \dots, v_k)$。

1.4　长度与内积

$\mathbb{R}^n$ 中向量的长度由勾股定理给出：

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

两个向量的内积 $\langle v, w \rangle = \sum_{i=1}^{n} v_i w_i$ 满足 $\langle v, v \rangle = \|v\|^2$。

用代码验算

```python
import math

def norm(v: list[float]) -> float:
    return math.sqrt(sum(x * x for x in v))

print(norm([3.0, 4.0]))  # 5.0
```

下一章用矩阵把线性组合写得更紧凑，见矩阵 › 矩阵乘法。

第 2 章　矩阵

2.1　线性组合的记账方式

把向量 $v_1, \dots, v_n \in \mathbb{R}^m$ 按列排好，就得到 $m \times n$ 矩阵 $A = (v_1 \ \cdots \ v_n)$。于是线性组合可以写成矩阵与向量的乘积:

$$Ax = x_1 v_1 + \cdots + x_n v_n.$$

这正是线性组合 的另一种写法。

2.2　矩阵乘法

矩阵乘法

设 A 是 $m \times n$ 矩阵，B 是 $n \times p$ 矩阵，乘积 $C = AB$ 是 $m \times p$ 矩阵，其元素为

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

例如

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 11 \\ 4 \end{pmatrix}.$$

顺序很重要

一般来说 $AB \neq BA$。即使 AB 有定义，BA 也可能根本没有定义。

2.3　结合律

结合律

只要乘积有定义，就有 $(AB)C = A(BC)$。

证明

比较两边第 (i, j) 个元素：

$$((AB)C)_{ij} = \sum_l \Big(\sum_k a_{ik} b_{kl} \Big) c_{lj} = \sum_k a_{ik} \Big(\sum_l b_{kl} c_{lj} \Big) = (A(BC))_{ij}.$$

中间一步只交换了有限和的次序。

2.4　常见的矩阵

名称	条件	例子
单位矩阵 I	$AI = IA = A$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
对角矩阵	非对角元都是 0	$\mathrm{diag}(2, 3)$
对称矩阵	$A^\top = A$	$\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$