

向量和线性代数

CHAPTER OBJECTIVES

在这一章我们将把视角从简单的数投向一种更复杂的数学结构，完成以下目标：

1. 了解向量和其基本的计算
2. 了解向量几何上的意义
3. 学会简单的线性代数，并尝试应用

1.1 二维向量

■ 向量与向量运算

DEFINITION 1

二维向量

一个有序实数对称为一个**二维向量**，我们用大写字母来表示向量

$$\mathbf{U} = (u_1, u_2)$$

数 u_1, u_2 称为向量的**分量**，二维向量的全体称为**二维空间**，记作 \mathbb{R}^2

DEFINITION 2

向量的代数运算

- 用一个实数 c **乘** 向量 $\mathbf{U} = (u_1, u_2)$ 得到的结果记为 $c\mathbf{U}$ ，定义为用 c 乘上 \mathbf{U} 的各个分量

$$c\mathbf{U} = (cu_1, cu_2)$$

- 向量 \mathbf{U} 和向量 \mathbf{V} 的**和** $\mathbf{U} + \mathbf{V}$ ，定义为两个向量对应分量相加得到的向量：

$$\mathbf{U} + \mathbf{V} = (u_1 + v_1, u_2 + v_2)$$

向量的数乘和相加满足通常的代数性质：

PROPERTIES

- 交换律: $UV = V + U$
- 结合律: $(U + V) + W = U + (V + W)$
- 分配律: $(a + b)U = aU + bV$
- 分配律: $c(U + V) = cU + cV$
- 加法逆元: $U + (-U) = 0$

注意：向量的另一种表示方法，是包裹在中括号之中，例如向量 (u_1, u_2) 可以被表示为：

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad \begin{bmatrix} u_1 & u_2 \end{bmatrix}^T$$

下面我们开始讨论向量及向量运算在几何上的意义：

■ 向量的几何意义

我们注意到任意一个向量 $U = \begin{bmatrix} x \\ y \end{bmatrix}$ 都可以视为：

在这里，我们可以将原本的 U 写作：

$$x \begin{bmatrix} 1 & 0 \end{bmatrix}^T + y \begin{bmatrix} 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$U = (x, y) = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

我们可以将 $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$ 和 $\begin{bmatrix} 0 & 1 \end{bmatrix}^T$ 分别看作是 x 轴和 y 轴的单位向量，那么可以将原本的向量 U 视为 $x - y$ 平面上的一个点。

通过将向量视为平面中的点，运算具有下述几何解释：

- 数乘：**对于非零向量 U 与实数 c ，点 cU 位于通过原点与 U 的直线上。若 $c > 0$ ，方向相同；若 $c < 0$ ，方向相反。
- 加法：**若 $0, U, V$ 不在一条直线上，则 $0, U, U + V, V$ 构成一个平行四边形的四个顶点。
- 线段表达：**设 $0 \leq c \leq 1$ ，则点 $V + cU$ 位于从 V 到 $V + U$ 的线段上。

■ 线性组合与线性无关

线性组合

向量 U 与 V 的一个**线性组合**是一个形如如下结构的向量：

$$aU + bV$$

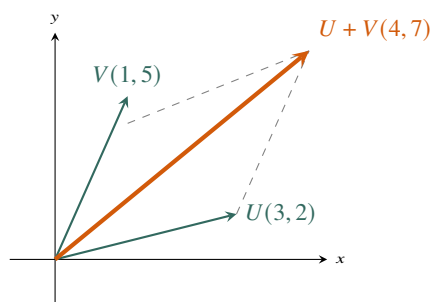


图 1.1: 向量加法的平行四边形法则

DEFINITION 3

DEFINITION 3

其中 a 与 b 是实数。

EXAMPLE 1 线性组合实例

证明 $U = (5, 3)$ 是 $(1, 1)$ 与 $(-1, 1)$ 的一个线性组合。

解 我们需要寻找 a, b 使得 $a(1, 1) + b(-1, 1) = (5, 3)$ 。这产生方程组：

$$\begin{cases} a - b = 5 \\ a + b = 3 \end{cases}$$

解得 $a = 4, b = -1$ 。因此 $U = 4(1, 1) - (1, -1)$ 。

DEFINITION 4

物理意义：

如果两个向量线性相关，说明它们“共线”，即其中一个向量的信息可以由另一个完全替代。

线性无关 (Linear Independence)

称向量 U 与 V **线性无关**，如果它们的某个线性组合 $aU + bV = \mathbf{0}$ 仅在 $a = 0, b = 0$ 时成立。

若存在不全为零的系数使得组合为零，则称它们为 **线性相关**。

THEOREM 1

线性无关的性质

给定位平面中两个线性无关的向量 C 和 D ，则 \mathbb{R}^2 中的每一个向量 U 都可以 **唯一**地写成 C 和 D 的线性组合：

$$U = aC + bD$$

线性函数

DEFINITION 5

线性函数

从 \mathbb{R}^2 到实数集 \mathbb{R} 的一个函数 $\ell : U \mapsto \ell(U)$ 称为线性的，如果：

1. $\ell(cU) = c\ell(U)$
2. $\ell(U + V) = \ell(U) + \ell(V)$

THEOREM 2

线性函数的结构

THEOREM 2

从 \mathbb{R}^2 到实数集 \mathbb{R} 的每一个线性函数 ℓ 仅当它具有如下形式：

$$\ell(x, y) = px + qy$$

其中 p, q 是实数。

■ 平面的基向量和矩阵

在前文我们提到，对于向量 $\mathbf{v} = (x, y)$ ，我们可以将其视为：

$$\begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

其中， $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ 是一个简写符号，用来指明向量的分量所对应的基向量，第一列对应第一个分量的**基向量**，第二列对应第二个分量的**基向量**。

DEFINITION 6

基向量 (Basis Vectors)

在二维平面 \mathbb{R}^2 中，一组 **基向量** 是指两个线性无关的向量 \mathbf{v}_1 和 \mathbf{v}_2 。

- **标准基**：通常我们取 $\mathbf{v}_1 = (1, 0)$ 和 $\mathbf{v}_2 = (0, 1)$ 。在物理中，它们常被表示为单位矢量 \hat{i} 和 \hat{j} 。
- **唯一表示**：平面上的任何向量 \mathbf{v} 都可以唯一地表示为基向量的线性组合： $\mathbf{v} = x_1 \mathbf{v}_1 + y_2 \mathbf{v}_2$ 。

物理意义：

基向量 就像是测量和描述物理世界的“坐标轴”和“基本单位”。任何物理量（力、速度）都可以沿着这些基本方向进行分解。

矩阵的几何意义：基向量的集合

当我们写出一个 2×2 矩阵时，它的每一列都具有深刻的几何意义。

PROPERTIES

矩阵的物理/几何意义

对于一个 2×2 矩阵 $\mathbf{A} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ ：

1. **第一列** $\begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$ ：表示第一个分量（通常是 x 坐标）所对应的 **基向量**。
2. **第二列** $\begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$ ：表示第二个分量（通常是 y 坐标）所对应的 **基向量**。

因此，一个矩阵可以被视为一个基向量的集合。

重要观察：

矩阵乘法 M 本质上是将向量 的坐标 (倍数) 重新分配给矩阵 M 中定义的新基向量。

EXAMPLE 2 非标准基下的向量表示

假设我们选择一组新的基向量: $v_1 = (1, 1)$ 和 $v_2 = (-1, 1)$ 。那么, 由它们构成的矩阵是 $M' = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ 。如果一个点在“新坐标系”下的坐标是 $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, 它所代表的实际向量是什么?

解 根据矩阵列的定义, 矩阵的第一列是新的 x 轴方向的基向量, 第二列是新的 y 轴方向的基向量。将“新坐标”与“基向量矩阵”相乘:

$$actual = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2-1 \\ 2+1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

物理意义： 矩阵在这里充当了“坐标转换器”，它将一组 **抽象坐标** 转化为 **实际的物理向量**。

注意事项：坐标与向量的区分

在这一框架下, 我们需要严格区分两个概念:

- **坐标：** 一个列向量 $\begin{bmatrix} x \\ y \end{bmatrix}$, 它仅仅是一组数值系数 (倍数)。
- **矩阵：** 一个 2×2 矩阵, 它存储了基向量 (方向和大小) 的信息。

只有当 **坐标** 与 **矩阵** (基向量的集合) 相乘时, 它才具有了确定的 **物理意义**, 表示空间中的一个实际向量。

利用矩阵我们可以快速处理基变量的变换, 从而处理坐标变换:

EXAMPLE 3 向量的旋转变换

已知向量 $\begin{bmatrix} x \\ y \end{bmatrix}$ 。若将该向量绕原点逆时针旋转 θ 角得到新向量', 求' 在标准基下的坐标。

解 根据“矩阵是基向量集合”的原理, 旋转后的向量 U' 可以看作是: 原来的坐标份额 (x, y) 作用在了**旋转后的新基向量**上。

我们将标准基 \hat{i}, \hat{j} 统一旋转 θ 角, 得到一组新基向量:

$$\hat{i}^R = \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}, \quad \hat{j}^R = \begin{bmatrix} -\sin \theta \\ \cos \theta \end{bmatrix}$$

新向量' 依然保持着 x 份的第一基向量和 y 份的第二基向量, 只不过基向量变了:

$$U' = x\hat{i}^R + y\hat{j}^R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

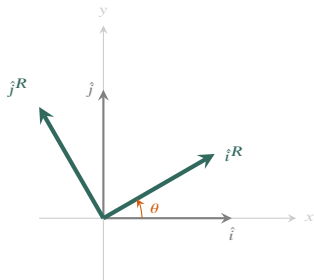


图 1.2: 基向量的旋转

展开矩阵乘法，得到旋转后的坐标：

$$\mathbf{U}' = \begin{bmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \end{bmatrix}$$

物理意义：旋转矩阵的每一列，本质上就是旋转后的基向量在原坐标系中的“投影”。

■

在上面的例题中，我们将对向量的变换改为对基向量的变换，我们可以将线性变换在应用场景中我们常常会遇到对向量的多次变换，我们可以利用基向量变换的思想引入矩阵乘法：

EXAMPLE 4

对向量 \mathbf{U} 进行 \mathbf{A}, \mathbf{B} 先后两次线性变换，
