第 1 章　绪　　论

1.1　引　　言

本章提纲挈领地引出了线性代数的几个重要概念：向量、向量空间、线性变换、线性方程组，旨在与初等数学知识进行衔接，适合自学，后续章节会对上述概念进一步深入探讨.

1.1.1　线性与非线性

什么是线性 (linear)？

若量与量之间成比例关系，我们通常称两者是线性关系 (或具有线性特性，也称为线性相关)，例如，若两个变量 x,y 之间服从

$$\frac{y}{x}=k \tag{1.1}$$

其中，$k\in\mathbb{R}$ 是常数，x,y 之间的比值是恒定的，则 x,y 之间是线性关系. 若 y 放大为 $2y$，由 x,y 的比例关系自然得出 x 变为了 $2x$. 当其中一个量发生改变时，我们可以由两者的线性关系直接得出另一个量. 这种规律的逻辑推理方式或思维模式即线性思维.

从几何的角度不难发现，式(1.1)可以写作 $y=kx$，表示二维平面上的一条直线 (过原点). 也就是说，两个量之间的关系为线性关系，则两者之间服从一次函数 (或方程)，几何上看是一条直线，如图 1.1 所示.

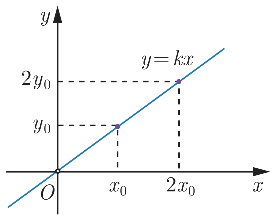

图 1.1　$y=kx$ 的图像

这种几何直观解释适用于低维条件 (两个变量之间) 下理解线性关系，并且不难得出：所谓非线性 (nonlinear)，即量与量之间不服从比例关系 (两个量的比值不是常数或两者不成比例)，从图像上看非线性呈现的是曲线，而不是直线.

为更好地拓展线性关系的概念内涵，研究多元函数的线性特征，数学家们将 $y=kx$ 的代数特性抽象出来，进而给出了更为一般的叙述来定义线性.

定义 1.1　线性

若函数 $f(x)$ 满足：

- 可加性：$f(x_1+x_2)=f(x_1)+f(x_2)$;
- 一次齐次性：$f(kx)=kf(x)$，k 是常数.

则称 $f(x)$ 为线性的.

可加性又称叠加性，即 x_1+x_2 作为自变量产生的效果 $f(x_1+x_2)$ 等于 x_1,x_2 独自产生的效果的叠加 $f(x_1)+f(x_2)$. 从另外一个角度看，对于 $\forall x_1$，当自变量增量为 x_2 时，函数增量均为 $f(x_2)$，即函数值的变化是均匀的.

一次齐次性也称为数乘性，即将 kx 代入 $f(x)$ 后，自变量 x 的幂次未发生改变，自变量放缩 k 倍，因变量 $f(x)$ 也进行相同比例的放缩，即函数值的变化是成比例的. 仅包含加法 (可加性) 与数乘 (齐次性) 两种的数学运算式称为线性运算.

例如，$y=3x$，$f(x,y,z)=x+2y+3z$ 都是线性函数，推广至 n 个自变量的线性函数，

$$y=a_1x_1+a_2x_2+\cdots+a_nx_n \quad (a_i\in\mathbb{R}, i=1,2,\cdots,n) \tag{1.2}$$

其中,我们称式(1.2)中等号右端 $a_1x_1+a_2x_2+\cdots+a_nx_n$ 为变量 $x_1,x_2,\cdots,x_n$ 的线性组合(或线性多项式)，$a_1,a_2,\cdots,a_n$ 是组合系数.

再如，$f(x,y)=xy$，$y=x^3$ 都不是线性函数，而是非线性函数，即不满足可加性和齐次性的函数.

那么,$y=3x+1$ 是不是线性函数?我们知道,$y=kx+b(k,b\in\mathbb{R},b\neq 0)$ 仍表示二维平面上的一条直线,这一数学结构是否为线性函数呢?不难验证,$kx+b$ 并不满足可加性和齐次性,因此严格来讲,在线性代数中,$f(x)=kx+b$ 不能算作线性函数. 然而，对 $y=kx+b$ 作恒等变形，

$$y-kx=b$$

其中，等号左端的数学结构 $y-kx$ 是线性的，是关于变量 x 和 y 的线性组合，所以，$y-kx=b$ 也可称作线性方程. 若 $b=0$，等式两边所有非零项所含未知变量的次数均为一次，故称作齐次线性方程；若 $b\neq 0$，则称作非齐次线性方程.

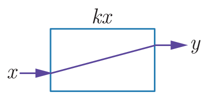

图 1.2 从系统视角看 $y=kx$

更进一步地，如果将 $y=kx$ 视为一个系统，如图 1.2 所示，该系统的功能是经处理后，将输入 x 转化成输出 y，即将 x 变换为 y，而这种变换是线性的，所以也称该变换为线性变换，其中 k 是变换 (比例) 系数.

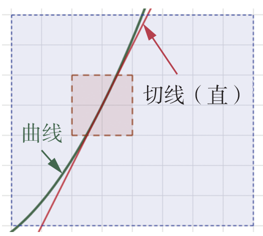

图 1.3 以直代曲

实际上，人们生活的现实世界本质上是非线性的，需要用非线性方法解决，如混沌、自组织、自适应、临界现象、“蝴蝶效应”等，为何还要研究和学习线性问题的解决方法呢？这是因为线性理论与方法相对完善，我们能够较好解决的问题中，线性问题占相当大的部分；另一个很重要的原因在于自然界中许多弱非线性现象可以在一定程度上近似为线性，或转化为线性问题进行解决，即“非线性问题线性化”. 如图 1.3 所示，在微积分中，“以直代曲”(光滑曲线在局部可以近似地用切线来代替) 就是局部线性化的体现.

1.1.2 线性代数的研究内容

所谓线性代数 (linear algebra)，就是揭示线性映射原理、分析处理线性关系的代数学分支，主要研究线性变换、有限维的线性方程组、行列式、矩阵、向量组的线性相关性、向量空间 (或线性空间)、特征值与特征向量、相似对角化、二次型等内容，是分析有限维线性空间及其线性变换、线性映射的基本理论.

1.1.3　线性代数的学习建议

为何要认真学习线性代数？一方面，线性代数理论完备、自洽，是训练抽象思维、逻辑推理、归纳演绎能力的最好素材；另一方面，线性代数是应用最广泛的数学分支之一，尤其是在大数据背景下，线性代数的基本理论、算法、数学符号和语言体系已经密切融入人工智能、数据挖掘、计算机图形学、密码学、军事运筹学、量子力学及其他数学分支领域的方方面面. 因此，熟练掌握线性代数的基本理论、方法和语言体系，是深入学习各领域专业知识的基石.

线性代数是关于数据的科学，在大数据时代，从海量数据中挖掘并获取有价值的信息，线性代数的地位更为凸显.

下面给出几条学习建议.

(1) 注重数学思维的培养. 在学习过程中，多问为什么，体会观察问题、抽象建模、探索求解、发现规律、猜想结论、总结论证的数学研究过程. 例如，在学习行列式这一主题时，围绕“行列式的概念是怎样形成的”“行列式有什么性质”“怎样计算行列式”“怎样通过降维将行列式按行展开”“如何应用行列式解线性方程组”等问题探究，综合提升自己分析问题、逻辑推理和精准计算等能力.

(2) 注重方法总结. 求解线性代数问题时，重复出现两次及以上频次的解题思路、策略和步骤，我们就可以称之为解题算法，要善于通过足量的习题训练，探寻并总结形成求解各型问题的算法.

(3) 注重多角度探究难点问题. 对较难理解的数学概念、定理及所研究的问题主线，要刻意并善于从不同角度、不同知识领域对其进行探究和解释 (数形结合就是一个很好的例子，低维情形下代数知识往往能够与几何知识联系起来)，并用自己的语言解释出来.

(4) 不拘泥于阅读一本书. 对相同的数学问题，不同教材的叙述逻辑和侧重点往往不同. 此外，随着学习的不断深入，线性代数的身影将在各专业领域中频繁出现，在不同的问题应用背景下会呈现出新的内涵.

本小节提到的线性相关、线性组合、线性变换等概念，会在后续章节深入讨论.

(5) 充分建立信心. 在诸多公式面前不必望而却步，沉下心来阅读，定能体会到形式化的数学符号背后朴素的哲理与逻辑.

下面就让我们步入线性代数的美妙世界，一同感悟其简约之美、纯粹之美、抽象之美、体系之美与实用之美！

1.2　向量和向量空间

1.2.1　向量的概念

物理学中有一类量，既有大小，又有方向，如位移、速度、加速度、力、力矩等，这一类量叫作向量 (或矢量).

定义 1.2　向量 (vector)

既有大小又有方向的量，称为向量或矢量.

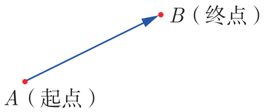

图 1.4　向量

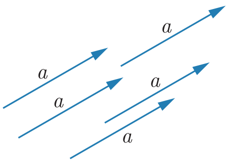

图 1.5　自由向量

常用一条有向线段来表示向量，线段的长度表示向量的大小，线段的方向表示向量的方向. 如图 1.4 所示，向量记为 $\overrightarrow{AB}$，$\vec{a}$ 或者 $\boldsymbol{a}$. 向量的长度记为 $\|\boldsymbol{a}\|$ 或 $|\boldsymbol{a}|$.

在物理学中，向量是可以在空间自由移动的量，只要保持大小和方向不变，向量就保持不变，称这种向量为自由向量，简称向量, 如图 1.5 所示.

在计算机等工程技术领域中，向量常被视为一个数据列表. 地球表面位置点 (西经 70°02′，北纬 36°53′) 的经纬度示意图如图 1.6 所示.

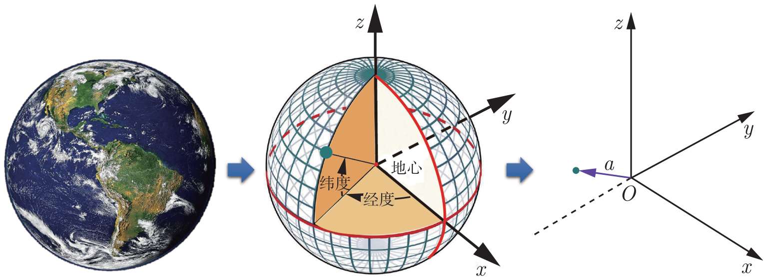

图 1.6　地球表面位置点的经纬度示意图

这个位置可以用下列数据列表来表示:

$$\begin{pmatrix} 70°02'\,\mathrm{W} \\ 36°53'\,\mathrm{N} \\ r \end{pmatrix} \text{ 或 } \begin{pmatrix} -70°02' \\ 36°53' \\ r \end{pmatrix}$$

其中，r 表示地球半径.

这里将地球视为理想球体，r 为常数.

在数学上，概括了上述两种观点，有向线段是向量的几何视角，有序数据列表是向量的代数视角，两者是向量这一数学概念在不同视角的表示. 图 1.6 中，数据列表代表的地球表面位置点与地心 (看作球心点) 之间唯一确定一个有向线段，即向量 $\boldsymbol{a}$，这实际上是利用三维向量对地球表面的位置点进行建模，数据列表的长度是 3，所以向量是三维的. 由此看来，有向线段 $\boldsymbol{a}$ 与有序数据列表两者之间是一一对应的.

☕ 数学上，通过对物理、计算机领域中的向量进行抽象化、符号化，借此研究和推广与之相关的数学结构和数学特性，为实际问题求解提供更为本质的视角.

本书主要研究欧几里得空间 (Euclidean space) 中的向量. 二维欧几里得空间记作 $\mathbb{R}^2$，是由全体二维向量 (具有两个分量的有序数组) 构成的集合，即

几何上，二维欧几里得空间习惯表示为 xOy 坐标平面.

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| \, x, y \in \mathbb{R} \right\}$$

其中，x, y 是向量的两个分量. 类似地，三维欧几里得空间记作 $\mathbb{R}^3$，是全体三维向量构成的集合，即

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, x, y, z \in \mathbb{R} \right\}$$

更一般地，n 维欧几里得空间记作 $\mathbb{R}^n$，是全体 n 维向量构成的集合.

定义 1.3 n 维向量 (n-dimensional vector)

n 维欧几里得空间

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| x_i \in \mathbb{R}, i = 1, 2, \cdots, n \right\}$$

中的元素称为n 维向量.

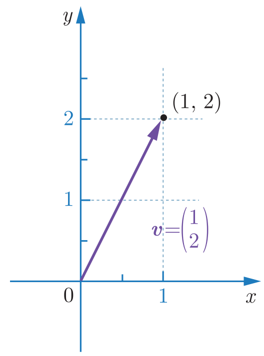

图 1.7 向量 $\boldsymbol{v}$

几何上，$\mathbb{R}^2$ 和 $\mathbb{R}^3$ 中的向量是从原点出发的有向线段，其终点坐标分量就是向量的分量. 例如，$\mathbb{R}^2$ 中的向量

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

是从原点 $(0, 0)$ 到终点 $(1, 2)$ 的有向线段，如图 1.7 所示. 向量的长度是有向线段的长度，$\boldsymbol{v}$ 的长度是 $\sqrt{1^2 + 2^2} = \sqrt{5}$. 如果向量进行平移运动，其方向与大小保持不变，例如，$(0, 0)$ 到 $(1, 2)$ 的有向线段与 $(2, 2)$ 到 $(3, 4)$ 的有向线段均表示向量 $\boldsymbol{v}$，如图 1.8 所示，两者大小 (长度) 相等、方向相同，即向量相等.

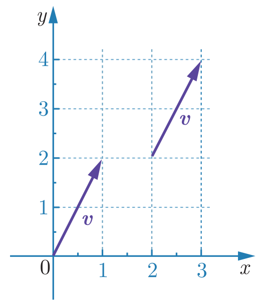

图 1.8 向量相等

所有分量均为 0 的向量称为零向量. 零向量的长度为 0，方向为任意方向，用 $\mathbf{0}$ 表示. 零向量 $\mathbf{0}$ 本质上是向量，与数字 0 不同.

长度为 1 的向量称为单位向量. 特别地，

$$\boldsymbol{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \boldsymbol{e_2} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \boldsymbol{e_n} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

称为单位坐标向量. 在 $\mathbb{R}^3$ 中单位坐标向量为

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{i}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \boldsymbol{j}, \quad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \boldsymbol{k}$$

1.2.2 向量的线性运算

向量的线性运算包括向量的加法和数乘.

数乘，即数与向量的乘法.

定义 1.4 向量的加法 (addition of vectors)

设向量 $\boldsymbol{a}, \boldsymbol{b}$ 是 $\mathbb{R}^n$ 中的两个向量，向量 $\boldsymbol{a}, \boldsymbol{b}$ 的和记作 $\boldsymbol{a} + \boldsymbol{b}$，且

$$\boldsymbol{a} + \boldsymbol{b} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

✍ 从定义中可以看出，对向量 $\boldsymbol{a}, \boldsymbol{b}$ 作加法运算，即两个向量的对应分量相加.

几何上，向量 $\boldsymbol{a},\boldsymbol{b}$ 加法遵循平行四边形法则或三角形法则. 例如，$\mathbb{R}^2$ 中，

$$\boldsymbol{a}=\begin{pmatrix}-1\\2\end{pmatrix},\quad \boldsymbol{b}=\begin{pmatrix}4\\2\end{pmatrix}$$

则

$$\boldsymbol{a}+\boldsymbol{b}=\begin{pmatrix}-1\\2\end{pmatrix}+\begin{pmatrix}4\\2\end{pmatrix}=\begin{pmatrix}3\\4\end{pmatrix}$$

其中，$\boldsymbol{a}+\boldsymbol{b}$ 是以原点为起点，以向量 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 为邻边的平行四边形对角线所在的向量，如图 1.9 和图 1.10 所示.

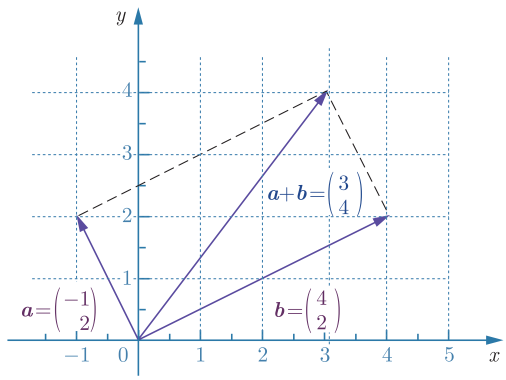

图 1.9　平行四边形法则

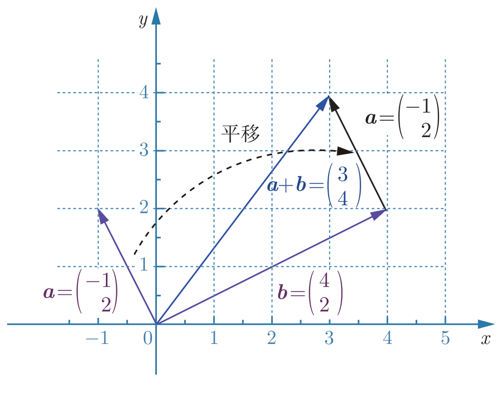

图 1.10　三角形法则

利用三角形法则有利于向量加法几何表示的推广. 如图 1.11 所示，$\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}+\cdots+\boldsymbol{z}$ 使向量平移并首尾相连，第一个向量的起点至最后一个向量终点之间的有向线段就是它们的和.

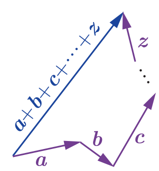

图 1.11　利用三角形法则表示多个向量相加

定义 1.5　向量的数乘 (scalar mutiplication of vectors)

实数 λ 与向量 $\boldsymbol{a}$ 的乘积记作 $\lambda\boldsymbol{a}$,

$$\lambda\boldsymbol{a}=\lambda\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}=\begin{pmatrix}\lambda a_1\\\lambda a_2\\\vdots\\\lambda a_n\end{pmatrix}$$

规定 $\lambda\boldsymbol{a}$ 是一个向量，它的大小为向量 $\boldsymbol{a}$ 的 $|\lambda|$ 倍，即

$$|\lambda\boldsymbol{a}|=|\lambda||\boldsymbol{a}|$$

当 $\lambda>0$ 时，$\lambda\boldsymbol{a}$ 与 $\boldsymbol{a}$ 同向；当 $\lambda<0$ 时，$\lambda\boldsymbol{a}$ 与 $\boldsymbol{a}$ 反向.

例如，设 $\boldsymbol{a}=\begin{pmatrix}1\\2\end{pmatrix}$，则 $\lambda\boldsymbol{a}=\lambda\begin{pmatrix}1\\2\end{pmatrix}$，当 λ 取 $-1,\dfrac{1}{2},\dfrac{3}{2},-2$ 时，相当于对向量 $\boldsymbol{a}$ 进行伸缩操作，$|\lambda|$ 即伸缩系数，如图 1.12 所示.

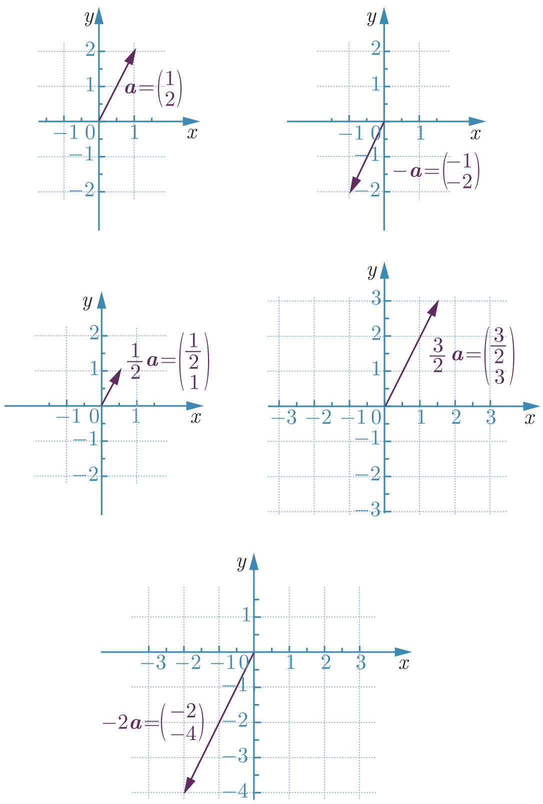

图 1.12　向量的数乘

特别地，当 $\lambda=-1$ 时，向量 $-\boldsymbol{a}$ 称为向量 $\boldsymbol{a}$ 的负向量.

借助负向量的定义，可以将向量的减法 $\boldsymbol{a}-\boldsymbol{b}$ 定义为

☕ 广义上，我们可以将向量减法视为向量加法的特殊情形而不再进行区分.

$$\boldsymbol{a}-\boldsymbol{b}=\boldsymbol{a}+(-\boldsymbol{b})$$

向量的线性运算满足下列 8 条运算规律.

设 $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ 都是 n 维向量，k,l 是常数，则有

(1) $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$
(2) $(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$
(3) $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$
(4) $\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$
(5) $1\cdot\boldsymbol{a}=\boldsymbol{a}$
(6) $k(l\boldsymbol{a})=(kl)\boldsymbol{a}$
(7) $k(\boldsymbol{a}+\boldsymbol{b})=k\boldsymbol{a}+k\boldsymbol{b}$
(8) $(k+l)\boldsymbol{a}=k\boldsymbol{a}+l\boldsymbol{a}$

1.2.3 向量的转置、内积及度量性质

向量包括列向量和行向量.

在线性代数中，不明确指出时，所谈的向量默认为列向量. 想要表达行向量时，可将其表示为列向量的转置(transposition). 例如，列向量

✍ 向量的转置运算是对向量的代数表示方式的改变，并不改变向量的大小和方向，但从代数上看，转置前后一般认为是两个不同的向量.

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

则行向量

$$(x_1,x_2,x_3)=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}^{\mathrm{T}}\overset{\text{记作}}{=\!=}\boldsymbol{x}^{\mathrm{T}}$$

向量的转置是可逆的，即 $\left(\boldsymbol{x}^{\mathrm{T}}\right)^{\mathrm{T}}=\boldsymbol{x}$.

✍ 在 $\mathbb{R}^2,\mathbb{R}^3$ 中，向量 $\boldsymbol{x}$ 的大小也称为模，记作 $|\boldsymbol{x}|$. 如图 1.13 所示，设向量 $\boldsymbol{x}=(x_1,x_2,x_3)^{\mathrm{T}}\in\mathbb{R}^3$，则 $\boldsymbol{x}$ 的模

$$|\boldsymbol{x}|=\sqrt{x_1^2+x_2^2+x_3^2}$$

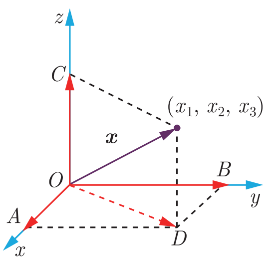

图 1.13 向量的模

为了节省书写空间，列向量 $\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$ 也常被写作 $(x_1,x_2,x_3)^{\mathrm{T}}$.

向量的度量性质主要有向量的长度、夹角、投影等.

n 维向量 $\boldsymbol{x}$ 的大小 (或长度) 称为 $\boldsymbol{x}$ 的范数(norm)，记作 $\|\boldsymbol{x}\|$.

定义 1.6 向量的范数

若向量

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in\mathbb{R}^n$$

则 $\|\boldsymbol{x}\|$ 称为 n 维向量 $\boldsymbol{x}$ 的范数，且

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \tag{1.3}$$

例 1.1　计算 n 维向量的范数：(1) 零向量 $\mathbf{0}$；(2) 单位坐标向量 $\boldsymbol{e}_1$.

解　(1) $\|\mathbf{0}\| = \sqrt{0^2 + 0^2 + \cdots + 0^2} = 0$;

(2) $\|\boldsymbol{e}_1\| = \sqrt{1^2 + 0^2 + \cdots + 0^2} = 1$.

单位向量的范数等于 1.

任意非零向量 $\boldsymbol{v}$，可运用下式单位化:

$$\boldsymbol{e}_v = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \tag{1.4}$$

向量的长度具有非负性，零向量 $\mathbf{0}$ 是唯一一个长度为 0 的向量，因此对任意向量 $\boldsymbol{v}$ 都有 $\|\boldsymbol{v}\| \geqslant 0$. 其中，$\|\boldsymbol{v}\| = 0$ 当且仅当 $\boldsymbol{v} = \mathbf{0}$. 另外，数乘 $\lambda\boldsymbol{v}$ 的范数为 $\boldsymbol{v}$ 范数的 $|\lambda|$ 倍，即 $\|\lambda\boldsymbol{v}\| = |\lambda|\,\|\boldsymbol{v}\|$ $(\lambda \in \mathbb{R})$. 这两条性质称为范数的非负性和齐次性.

范数的性质

(1)　非负性：当 $\boldsymbol{a} \neq 0$ 时，$\|\boldsymbol{a}\| > 0$；当 $\boldsymbol{a} = 0$ 时，$\|\boldsymbol{a}\| = 0$.

(2)　齐次性：　$\|\lambda\boldsymbol{a}\| = |\lambda|\|\boldsymbol{a}\|$.

定义 1.7　向量的内积 (inner product) 运算

设 $\mathbb{R}^n$ 中的两个 n 维向量

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

向量 $\boldsymbol{x}$ 与向量 $\boldsymbol{y}$ 的内积为

$$\begin{aligned} [\boldsymbol{x}, \boldsymbol{y}] &= \|\boldsymbol{x}\|\,\|\boldsymbol{y}\| \cos\theta \\ &= x_1y_1 + x_2y_2 + \cdots + x_ny_n \end{aligned} \tag{1.5}$$

其中，θ 为向量 $\boldsymbol{x}, \boldsymbol{y}$ 之间的夹角.

这种内积也称为欧几里得内积，与直角坐标系中的数量积等同看待、混同使用. 在常见教材中，内积也使用 $<\boldsymbol{x}, \boldsymbol{y}>$ 或 $(\boldsymbol{x}, \boldsymbol{y})$ 等记号表示.

内积的性质　设 $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ 为 n 维向量，λ 为实数.

(1) $[\boldsymbol{x}, \boldsymbol{y}] = [\boldsymbol{y}, \boldsymbol{x}]$

(2) $[\lambda\boldsymbol{x}, \boldsymbol{y}] = \lambda[\boldsymbol{y}, \boldsymbol{x}]$

(3) $[\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}] = [\boldsymbol{x}, \boldsymbol{z}] + [\boldsymbol{y}, \boldsymbol{z}]$

(4) 当 $\boldsymbol{x} = 0$ 时，$[\boldsymbol{x}, \boldsymbol{x}] = 0$；当 $\boldsymbol{x} \neq 0$ 时，$[\boldsymbol{x}, \boldsymbol{x}] > 0$.

在 $\mathbb{R}^2$ 中，向量

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

的数量积如图 1.14 所示，为

$$\begin{aligned} \boldsymbol{x} \cdot \boldsymbol{y} &= |\boldsymbol{x}||\boldsymbol{y}| \cos\theta \\ &= |\boldsymbol{x}|\mathrm{Prj}_{\boldsymbol{x}}\boldsymbol{y} \\ &= x_1x_2 + y_1y_2 \end{aligned}$$

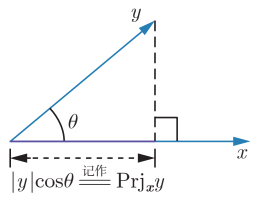

图 1.14　数量积

其中，$\mathrm{Prj}_{\boldsymbol{x}}\boldsymbol{y}$ 表示向量 $\boldsymbol{y}$ 在向量 $\boldsymbol{x}$ 上的投影大小 (是标量). 当 $\boldsymbol{x} \neq \mathbf{0}, \boldsymbol{y} \neq \mathbf{0}$ 时，向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的夹角为

$$\theta = \arccos \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{|\boldsymbol{x}||\boldsymbol{y}|}$$

特别地，当 $\boldsymbol{x} \perp \boldsymbol{y}$ 时，$\theta = \frac{\pi}{2}$，$\cos\theta = 0$，于是 $\boldsymbol{x} \cdot \boldsymbol{y} = |\boldsymbol{x}||\boldsymbol{y}| \cos\theta = 0$，即

$$\boldsymbol{x} \perp \boldsymbol{y} \Leftrightarrow \boldsymbol{x} \cdot \boldsymbol{y} = 0$$

⚒ 例 1.2 证明柯西-斯瓦茨 (Cauchy-Schwarz) 不等式:

$$[\boldsymbol{x},\boldsymbol{y}]^2 \leqslant [\boldsymbol{x},\boldsymbol{x}][\boldsymbol{y},\boldsymbol{y}]$$

证 $[\boldsymbol{x},\boldsymbol{y}]^2 = \|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2\cos\theta \leqslant \|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 = [\boldsymbol{x},\boldsymbol{x}][\boldsymbol{y},\boldsymbol{y}]$

当 $\theta=0$ 时，$\cos\theta=1$，向量 $\boldsymbol{y}$ 在 $\boldsymbol{x}$ 上的投影长度 $\mathrm{Prj}_{\boldsymbol{x}}\boldsymbol{y}=\|\boldsymbol{y}\|\cos\theta=\|\boldsymbol{y}\|$，此时，等号成立.

由向量内积的概念，当 $\boldsymbol{x}\neq\mathbf{0},\boldsymbol{y}\neq\mathbf{0}$ 时，n 维向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的夹角为

$$\theta=\arccos\frac{[\boldsymbol{x},\boldsymbol{y}]}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|}\quad(0\leqslant\theta<\pi)$$

在 $\mathbb{R}^2,\mathbb{R}^3$ 中，向量 $\boldsymbol{x}$ 与向量 $\boldsymbol{y}$ 正交等同于两者垂直，且垂直的充要条件为 $\boldsymbol{x}\cdot\boldsymbol{y}=0$.

特别地，当向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的夹角为 $\dfrac{\pi}{2}$ 时，称向量 $\boldsymbol{x},\boldsymbol{y}$ 正交(orthogonal)，两者互为正交向量. 显然，若非零向量 $\boldsymbol{x},\boldsymbol{y}$ 的内积为 0，即 $[\boldsymbol{x},\boldsymbol{y}]=\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cdot\cos\theta=0$，则 $\cos\theta=0$，于是 $\theta=\dfrac{\pi}{2}$，即向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交. 如果向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交，则 $\cos\theta=0$，即

$$\cos\theta=\frac{[\boldsymbol{x},\boldsymbol{y}]}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|}=0$$

所以，$[\boldsymbol{x},\boldsymbol{y}]=0$. 由此，总结得出：非零向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交的充要条件是

$$[\boldsymbol{x},\boldsymbol{y}]=x_1y_1+x_2y_2+\cdots+x_ny_n=0$$

✍ 所谓两两正交，是指在向量序列中任选两个不同的向量均正交.

下面简要介绍几个相关的重要概念，在第 4 章我们将着重讨论它们.

不难验证，单位坐标向量 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 两两正交.例如, $\mathbb{R}^3$ 中，单位坐标向量

$$\boldsymbol{i}=\begin{pmatrix}1\\0\\0\end{pmatrix},\quad \boldsymbol{j}=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad \boldsymbol{k}=\begin{pmatrix}0\\0\\1\end{pmatrix}$$

两两正交. 并且，对任意向量 $\boldsymbol{a}=(a_x,a_y,a_z)^{\mathrm{T}}\in\mathbb{R}^3$，有

$$\boldsymbol{a}=a_x\boldsymbol{i}+a_y\boldsymbol{j}+a_z\boldsymbol{k}$$

例如，$\boldsymbol{a}=(1,2,3)^{\mathrm{T}}=\boldsymbol{i}+2\boldsymbol{j}+3\boldsymbol{j}$, 故称向量 $\boldsymbol{a}$ 可由 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 线性表示，其中 $1,2,3$ 是线性表示的系数. 在线性代数中，单位坐标向量 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 也称为三维欧几里得空间 $\mathbb{R}^3$ 的一组基(basis)，$\mathbb{R}^3$ 中的任一向量 $\boldsymbol{a}=(a_x,a_y,a_z)^{\mathrm{T}}$ 均可由基 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 线性表示，且 (a_x,a_y,a_z) 为向量 $\boldsymbol{a}$ 在基 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 下的坐标.

通过线性组合的方式，可以由基构建出整个向量空间.

那么，是否只有 $\boldsymbol{i},\boldsymbol{j},\boldsymbol{k}$ 这种单位坐标向量才能成为 $\mathbb{R}^3$ 的一组基呢？不是！$\mathbb{R}^3$ 的基并不唯一，只要满足特定的条件就可以作为一组基，不同基之间可以互相变换，即基变换 (basis transformation)，读者可以带着兴趣提前查阅 4.3 节相关内容.

☕ 何谓空间？我们可以通俗地理解为空间是定义了结构(或元素之间关系) 的集合，即将抽象的运算系统用熟悉的几何空间来模拟.

1.2.4 向量空间中的线性变换

向量空间是一个由向量组成且满足特定的运算性质的数学结构. 引入向量空间的概念，便于从运算性质的角度描述和讨论向量的集合. 在线性代数中，向量空间一般是指欧几里得空间，如 $\mathbb{R}^n$.

定义 1.8　向量空间 (vector space)

设 V 为 n 维向量的集合，如果集合 V 非空，且集合 V 对向量的线性运算封闭，那么就称集合 V 为向量空间.

✍ 向量空间：
(1) 非空；
(2) 对线性运算封闭.
此外还需满足前文中提到的 8 条运算规律.

所谓对线性运算封闭，是指对 V 中任意两个向量进行加法和数乘运算，其运算结果仍在 V 中，即

(1) 若 $\boldsymbol{a}\in V,\boldsymbol{b}\in V$，则 $\boldsymbol{a}+\boldsymbol{b}\in V$；

(2) 若 $\boldsymbol{a}\in V,\lambda\in\mathbb{R}$，则 $\lambda\boldsymbol{a}\in V$.

三维向量的全体 $\mathbb{R}^3$ 是一个向量空间. 这是因为任意两个三维向量的和仍然是三维向量，数 λ 乘一个三维向量仍然是一个三维向量，即对向量的加法和数乘封闭，所以三维向量的全体是一个向量空间.

几何上，一个三维向量可以用空间的一个有向线段来表示，从而向量空间 $\mathbb{R}^3$ 可以看作以坐标原点 $\boldsymbol{O}$ 为起点的有向线段的全体. 由于以原点为起点的有向线段与终点一一对应，因此 $\mathbb{R}^3$ 也可看作点空间.

⚒ 例 1.3　判断集合 $V=\{\boldsymbol{x}=(1,x_2,x_3)^{\mathrm{T}}\,|\,x_2,x_3\in\mathbb{R}\}$ 是否为向量空间，为什么？

解　否. 若向量 $\boldsymbol{a}=\boldsymbol{b}=(1,2,3)^{\mathrm{T}}$，满足 $\boldsymbol{a},\boldsymbol{b}\in V$，而

$$\boldsymbol{a}-\boldsymbol{b}=(0,0,0)^{\mathrm{T}}\notin V$$

说明集合 V 对加法运算不封闭，故 V 不是向量空间.

✍ 向量空间必须包含零向量. 例 1.3表明，若集合中不包含零向量，则可直接判定其不是向量空间.

本例也可令 $\boldsymbol{a}=(1,2,3)^{\mathrm{T}}$, 而 $2\boldsymbol{a}=(2,4,6)^{\mathrm{T}}\notin V$, 说明 V 对数乘运算不封闭，故 V 不是向量空间.

⚒ 例 1.4　设 $\boldsymbol{a},\boldsymbol{b}$ 为两个二维向量，证明：集合

$$L=\{\boldsymbol{x}=\lambda\boldsymbol{a}+\mu\boldsymbol{b}|\lambda,\mu\in\mathbb{R}\}$$

是一个向量空间.

证　设

$$\boldsymbol{x_1}=\lambda_1\boldsymbol{a}+\mu_1\boldsymbol{b}\quad(\lambda_1,\mu_1\in\mathbb{R})$$

$$\boldsymbol{x_2}=\lambda_2\boldsymbol{a}+\mu_2\boldsymbol{b}\quad(\lambda_2,\mu_2\in\mathbb{R})$$

则

$$\boldsymbol{x_1}+\boldsymbol{x_2}=(\lambda_1+\lambda_2)\boldsymbol{a}+(\mu_1+\mu_2)\boldsymbol{b}\in L$$

$$k\boldsymbol{x_1}=(k\lambda_1)\boldsymbol{a}+(k\mu_1)\boldsymbol{b}\in L$$

所以，集合 L 是一个向量空间.

本例中，$L=\{\boldsymbol{x}=\lambda\boldsymbol{a}+\mu\boldsymbol{b}|\lambda,\mu\in\mathbb{R}\}$ 表明 L 中任一向量均可表示为向量 $\boldsymbol{a},\boldsymbol{b}$ 的线性组合，因此，向量空间 L 称为由向量 $\boldsymbol{a},\boldsymbol{b}$ 生成的向量空间或张成的向量空间，可记为 $\mathrm{Span}\{\boldsymbol{a},\boldsymbol{b}\}$. 若向量 $\boldsymbol{a},\boldsymbol{b}$ 共线，则 L 表示与 $\boldsymbol{a},\boldsymbol{b}$ 共线的所有向量；若向量 $\boldsymbol{a},\boldsymbol{b}$ 不共线，则 L 表示 $\boldsymbol{a},\boldsymbol{b}$ 所在的整个平面空间.

☕ 定义 1.9 中，向量 $\boldsymbol{y}$ 是一个 m 维向量，向量 $\boldsymbol{x}$ 是一个 n 维向量，该线性变换的功能是建立 $\mathbb{R}^n$ 空间中向量 $\boldsymbol{x}$ 与 $\mathbb{R}^m$ 空间中向量 $\boldsymbol{y}$ 之间的映射. 当 $m<n$ 时，线性变换使向量的维数降低了，这种线性变换称为降维线性变换. 降维线性变换可以理解为一种"投影"或"压缩"，这种变换往往是不可逆的、退化的.

定义 1.9　线性变换 (linear transformation)

设 n 维向量 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ 与 m 维向量 $\boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_m\end{pmatrix}$ 存在如下的线性映射：

$$\begin{cases}y_1=a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\y_2=a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\\vdots\\y_m=a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n\end{cases}$$

这个映射就称为向量 $\boldsymbol{x}$ 到向量 $\boldsymbol{y}$ 的线性变换，记为 $\boldsymbol{y}=\mathcal{T}(\boldsymbol{x})$.

线性变换可以看作一种函数关系，即任意给定 n 维向量 $\boldsymbol{x}$，通过线性变换函数 $\mathcal{T}$ 可以得到向量 $\boldsymbol{y}$，即 $\boldsymbol{y}=\mathcal{T}(\boldsymbol{x})$.

线性变换也可以理解为一种运动形式，即从向量 $\boldsymbol{x}$ 至向量 $\boldsymbol{y}$ 的变换运动. 几何上，该运动对向量 $\boldsymbol{x}$ 进行伸缩或者旋转操作，转化成向量 $\boldsymbol{y}$. 同一个线性变换对向量空间中所有的向量进行同比例的伸缩和旋转 (功能相同). 假设有无穷多向量，若它们的矢端端点构成一条直线，在同一个线性变换下，变换后的向量的所有矢端端点仍然构成一条直线. 线性变换把直线仍然变换为直线，这是把这种变换称为"线性"的原因.

向量可用起点在原点的有向线段表示，有向线段带箭头的端点称为向量的矢端.

线性变换的几何解释

设 $\mathbb{R}^2$ 中有线性变换

$$\mathcal{T}:\begin{cases}y_1=-x_2\\y_2=x_1\end{cases}$$

对于向量 $\boldsymbol{x}=(2,1)^{\mathrm{T}}$，由上述线性变换可以得到变换后的向量

$$\boldsymbol{y}=\begin{pmatrix}-1\\2\end{pmatrix}$$

从图 1.15 可以看到，线性变换 $\mathcal{T}$ 使向量 $\boldsymbol{x}$ 逆时针旋转 90° 变换为向量 $\boldsymbol{y}$，并且变换过程中保持向量的长度不变.

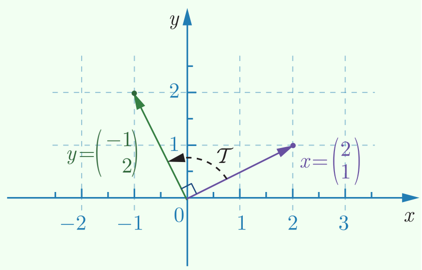

图 1.15　线性变换 $\mathcal{T}$ 作用于向量

下面换一个角度分析. 上述线性变换 $\mathcal{T}$ 还可以写成

$$\begin{cases}y_1=0x_1-x_2\\y_2=x_1+0x_2\end{cases}\tag{1.6}$$

可以看到，线性变换 $\mathcal{T}$ 由系数列向量

$$\boldsymbol{i}'=\begin{pmatrix}0\\1\end{pmatrix}\quad 与\quad \boldsymbol{j}'=\begin{pmatrix}-1\\0\end{pmatrix}$$

唯一确定. 任意给定向量 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$，由式(1.6)可求得变换后的向量 $\boldsymbol{y}=(y_1,y_2)^{\mathrm{T}}$. 事实上，线性变换 $\mathcal{T}$ 可以写成向量形式：

$$\boldsymbol{y}=\mathcal{T}(\boldsymbol{x})=x_1\begin{pmatrix}0\\1\end{pmatrix}+x_2\begin{pmatrix}-1\\0\end{pmatrix}=x_1\boldsymbol{i}'+x_2\boldsymbol{j}' \tag{1.7}$$

其中，x_1,x_2 是向量 $\boldsymbol{y}$ 在基 $\boldsymbol{i}',\boldsymbol{j}'$ 下的坐标. 我们知道，x_1,x_2 也是向量 $\boldsymbol{x}$ 在基 $\boldsymbol{i},\boldsymbol{j}$ 下的坐标，即

$$\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\end{pmatrix}=x_1\begin{pmatrix}1\\0\end{pmatrix}+x_2\begin{pmatrix}0\\1\end{pmatrix}=x_1\boldsymbol{i}+x_2\boldsymbol{j} \tag{1.8}$$

比较式(1.7)和式(1.8)不难发现，在线性变换 $\mathcal{T}$ 的作用下，相当于对基 $\boldsymbol{i},\boldsymbol{j}$ 所决定的整个坐标系逆时针旋转 $90°$，向量 $\boldsymbol{y}$ 在基 $\boldsymbol{i}',\boldsymbol{j}'$ 下的坐标与向量 $\boldsymbol{x}$ 在基 $\boldsymbol{i},\boldsymbol{j}$ 下的坐标保持不变.

例如，当给定向量 $\boldsymbol{x}=(2,1)^{\mathrm{T}}$ 时，变换后的向量

$$\boldsymbol{y}=2\begin{pmatrix}0\\1\end{pmatrix}+1\cdot\begin{pmatrix}-1\\0\end{pmatrix}=\begin{pmatrix}2\times 0+1\times(-1)\\2\times 1+1\times 0\end{pmatrix}=\begin{pmatrix}-1\\2\end{pmatrix}$$

线性变换 $\mathcal{T}$ 下的基向量如图 1.16 所示.

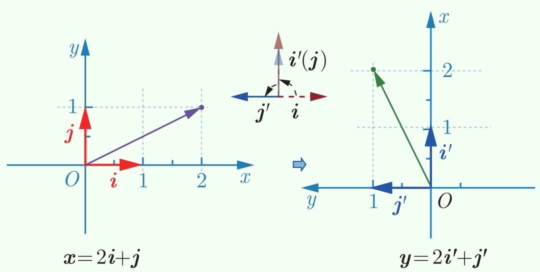

图 1.16　线性变换下的基向量

x 轴上的 $\boldsymbol{i}=(1,0)^{\mathrm{T}}$ 变换为 $\boldsymbol{i}'=(0,1)^{\mathrm{T}}$ 表明 x 轴逆时针旋转了 $90°$，y 轴上的 $\boldsymbol{j}=(0,1)^{\mathrm{T}}$ 变换为 $\boldsymbol{j}'=(-1,0)^{\mathrm{T}}$，表明 y 轴逆时针旋转了 $90°$，向量 $\boldsymbol{x}$ 跟随整个坐标系逆时针旋转了 $90°$ 到达向量 $\boldsymbol{y}$ 所在位置，由于向量和坐标系同时旋转，即整个坐标平面逆时针旋转了 $90°$，相对位置没有改变，所以在新基下的坐标和旧基下的坐标相同.

例 1.5　设有线性变换

$$\mathcal{T}:\begin{cases}y_1=x_1+6x_2+3x_3\\y_2=3x_1+3x_2+x_3\\y_3=5x_1+2x_2\end{cases}$$

当 $\boldsymbol{x}=(3,2,1)^{\mathrm{T}}$ 时，求线性变换后的向量 $\boldsymbol{y}$.

解　向量

$$\begin{aligned}\boldsymbol{y}&=3\begin{pmatrix}1\\3\\5\end{pmatrix}+2\begin{pmatrix}6\\3\\2\end{pmatrix}+\begin{pmatrix}3\\1\\0\end{pmatrix}\\&=\begin{pmatrix}3\times 1+2\times 6+3\\3\times 3+2\times 3+1\\3\times 5+2\times 2+0\end{pmatrix}=\begin{pmatrix}18\\16\\19\end{pmatrix}\end{aligned}$$

例 1.6　设有线性变换

$$\mathcal{T}:\begin{cases} y_1 = x_1 \\ y_2 = 0 \end{cases}$$

当 $\boldsymbol{x}=(1,2)^{\mathrm{T}}$ 时，求线性变换后的向量 $\boldsymbol{y}$.

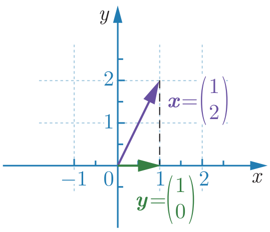

图 1.17　投影线性变换

通俗地看，该线性变换将 $\mathbb{R}^2$ 空间“压缩”为 $\mathbb{R}$，是不可逆的、退化的变换.

解　方法一：把 $\boldsymbol{x}=(1,2)^{\mathrm{T}}$ 代入得

$$\boldsymbol{y}=\begin{pmatrix}1\\0\end{pmatrix}$$

表明该线性变换将向量 $\boldsymbol{x}$ 向 x 轴投影，投影向量就是向量 $\boldsymbol{y}$，如图 1.17 所示.

方法二：线性变换变形为

$$\begin{cases} y_1 = \ \ x_1 + 0x_2 \\ y_2 = 0x_1 + 0x_2 \end{cases}$$

实际上，线性变换 $\mathcal{T}$ 将基 $\boldsymbol{i}=(1,0)^{\mathrm{T}},\boldsymbol{j}=(0,1)^{\mathrm{T}}$ 变换为新基 $\boldsymbol{i}'=(1,0)^{\mathrm{T}},\boldsymbol{j}'=(0,0)^{\mathrm{T}}$，即线性变换的结果是将基向量向 x 轴投影，所以，向量 $\boldsymbol{y}$ 就是向量 $\boldsymbol{x}$ 的投影向量，且根据 $\boldsymbol{y}$ 在新基 $\boldsymbol{i}',\boldsymbol{j}'$ 下的坐标与 $\boldsymbol{x}$ 在旧基 $\boldsymbol{i},\boldsymbol{j}$ 下的坐标相同，$\boldsymbol{y}=\boldsymbol{i}'+2\boldsymbol{j}'$，所以

$$\boldsymbol{y}=\begin{pmatrix}1\\0\end{pmatrix}$$

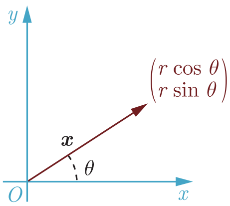

图 1.18　初始向量

例 1.7　说明旋转线性变换

$$\mathcal{T}:\begin{cases} y_1 = \cos\varphi x_1 - \sin\varphi x_2 \\ y_2 = \sin\varphi x_1 + \cos\varphi x_2 \end{cases}$$

的几何意义.

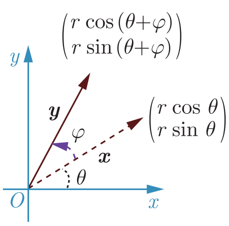

图 1.19　向量逆时针旋转 φ 角

解　方法一：设 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}=(r\cos\theta,r\sin\theta)^{\mathrm{T}}$，如图 1.18 所示. 代入线性变换得

$$\begin{aligned}\boldsymbol{y}&=\begin{pmatrix}y_1\\y_2\end{pmatrix}\\&=\begin{pmatrix}r\cos\theta\cos\varphi-r\sin\theta\sin\varphi\\r\cos\theta\sin\varphi+r\sin\theta\cos\varphi\end{pmatrix}\\&=\begin{pmatrix}r\cos(\theta+\varphi)\\r\sin(\theta+\varphi)\end{pmatrix}\end{aligned}$$

线性变换 $\mathcal{T}$ 将向量 $\boldsymbol{x}$ 逆时针旋转了 φ 角，如图 1.19 所示.

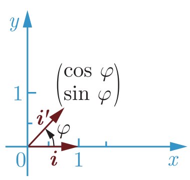

图 1.20　基向量 $\boldsymbol{i}$ 旋转 φ 角

方法二：如图 1.20 和图 1.21 所示，线性变换 $\mathcal{T}$ 将基 $\boldsymbol{i}=(1,0)^{\mathrm{T}},\boldsymbol{j}=(0,1)^{\mathrm{T}}$ 变换为新基

$$\boldsymbol{i}'=\begin{pmatrix}\cos\varphi\\\sin\varphi\end{pmatrix},\quad \boldsymbol{j}'=\begin{pmatrix}-\sin\varphi\\\cos\varphi\end{pmatrix}$$

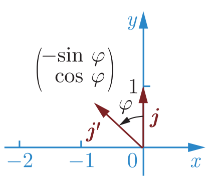

图 1.21　基向量 $\boldsymbol{j}$ 旋转 φ 角

基向量 $\boldsymbol{i},\boldsymbol{j}$ 均逆时针旋转了 φ 角，可理解为基向量的旋转携带着整个坐标平面逆时针旋转 φ 角，所以线性变换将向量 $\boldsymbol{x}$ 逆时针旋转了 φ 角.

1.3　线性方程组

1.3.1　线性方程组的定义

定义 1.10　n 元非齐次线性方程组

设有 n 个未知数 m 个方程的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{1.9}$$

其中，a_{ij} 是第 i 个方程的第 j 个未知数的系数，b_i 是第 i 个方程的常数项，$i = 1, 2, \cdots, m; j = 1, 2, \cdots, n$. 当常数项 $b_1, b_2, \cdots, b_m$ 不全为 0 时，线性方程组 (1.9) 叫作 n 元非齐次线性方程组.

当 $b_1, b_2, \cdots, b_m$ 全为 0 时，式 (1.9) 成为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \quad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \tag{1.10}$$

式 (1.10) 称为n 元齐次线性方程组. n 元线性方程组 (1.9) 和 (1.10) 统称线性方程组.

1.3.2　线性方程组的解

线性方程组若有解，称它相容；若无解，则称它不相容.

对 n 元齐次线性方程组 (1.10)，$x_1 = x_2 = \cdots = x_n = 0$ 必定是它的解，这个解叫作 n 元齐次线性方程组 (1.10) 的零解. 齐次线性方程组一定有零解，所以必相容. 如果存在不全为 0 的一组数 $(x_1, x_2, \cdots, x_n)$ 也满足方程组 (1.10)，则 $(x_1, x_2, \cdots, x_n)$ 叫作齐次线性方程组 (1.10) 的非零解. 一个齐次线性方程组不一定有非零解，需要具体问题具体分析.

因此，针对线性方程组，主要研究三个基本问题.

(1) 线性方程组是否有解；
(2) 在有解时，解是否唯一；
(3) 解不唯一时，如何求出 (表示出) 这些解.

假设有三个非齐次线性方程组

$$\text{(i)}\begin{cases} x_1 + x_2 = 2 \\ x_1 - x_2 = 2 \end{cases} \qquad \text{(ii)}\begin{cases} x_1 + x_2 = 2 \\ x_1 + x_2 = 1 \end{cases} \qquad \text{(iii)}\begin{cases} x_1 + x_2 = 2 \\ -x_1 - x_2 = -2 \end{cases}$$

从几何上看，方程组中每一个方程表示平面中的一条直线，两条直线的交点 (x_1, x_2) 为线性方程组的解.

如图 1.22 所示，方程组 (i) 中两条直线交于一点 $(2,0)$，因此，$(2,0)$ 是线性方程组 (i) 的解；方程组 (ii) 是不相容的，无解，对应的两条直线平行；方程组 (iii) 中两个方程表示同一条直线，即两条直线重合，直线上任意一点都是方程组 (iii) 的解，所以方程组 (iii) 有无穷多解. 一般地，二元一次线性方程组表示的两条直线有三种几何关系：两条直线交于一点；两条直线平行；两条直线重合. 因此，线性方程组的解的情况有三种：唯一解、无解、无穷多解.

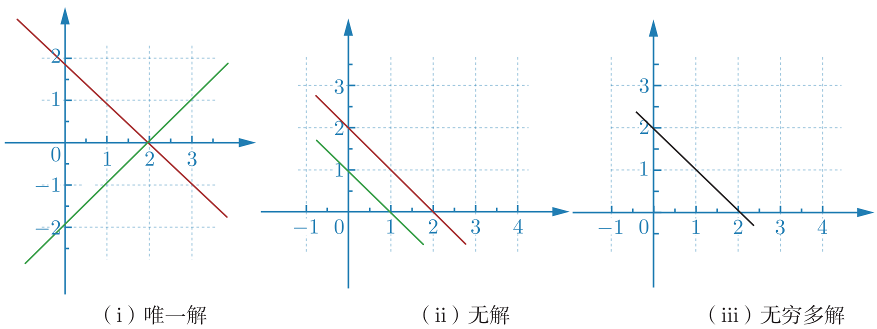

图 1.22　方程组 (i)、(ii)、(iii) 的解的几何解释

线性方程组的解的另一种几何解释

线性方程组 (1.9) 本质上就是一个由系数列向量

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

共同决定的线性变换 $\mathcal{T}$. 该线性变换将向量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 变换为向量 $\boldsymbol{b} = (b_1, b_2, \cdots, b_m)^{\mathrm{T}}$. 因此，求线性方程组的解，本质上就是已知线性变换 $\mathcal{T}$ 和变换后的向量 $\boldsymbol{b}$，求初始向量 $\boldsymbol{x}$，如图 1.23 所示.

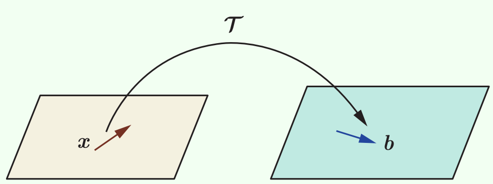

图 1.23　线性方程组的解的几何解释

1.3.3　消元法

线性方程组的基本解法是消元法，又称高斯消元法.

⚒ 例 1.8　使用消元法求解二元一次线性方程组：

$$\begin{cases} 3x-2y=0 \\ -x+y=2 \end{cases}$$

解　交换两个方程的位置，得到与原方程组同解的方程组

$$\begin{cases} -x+y=2 & (1) \\ 3x-2y=0 & (2) \end{cases}$$

将方程 (1) 左右两端同时乘以 2，得到同解方程组

$$\begin{cases} -2x+2y=4 & (3) \\ 3x-2y=0 \end{cases}$$

将方程 (3) 与方程 (2) 相加得

$$x=4$$

将 $x=4$ 代入方程 (1) 得 $y=6$，所以方程组的解为

$$\begin{cases} x=4 \\ y=6 \end{cases}$$

关于线性方程组的研究，最早出现在我国《九章算术》(公元 1 世纪左右) 的第八章“方程”中，采用分离系数法表示线性方程组 (相当于现在的矩阵)、采用直除法求解方程组 (与矩阵的初等变换一致) 的完整解法，这是世界数学史上一项重大的成就，在隋唐时期传入朝鲜、日本.

从例 1.8 可以看出，消元法解线性方程组包含三种类型运算：

(1) 交换两个方程的位置;
(2) 一个方程左右两端乘上一个常数;
(3) 将一个方程的常数倍加到另外一个方程上.

在 3.6 节学习矩阵的三种初等行变换时，你会发现其与方程的三种运算是一一对应的关系，从而揭示了使用矩阵的初等行变换解线性方程组的方法本质上就是消元法.

⚒ 例 1.9　研究线性方程组：

$$\begin{cases} 3x-2y=0 & (1) \\ -x+y=2 & (2) \\ 2x-y=2 & (3) \end{cases}$$

解　方程 (3) = 方程 (1)+ 方程 (2)，即方程 (3) 能提供的关于解的信息完全可由方程 (1) 和方程 (2) 替代，故本例中的方程组与例 1.8 方程组同解，此时称线性方程组等价. 称方程 (3) 可由方程 (1) 和方程 (2) 线性表示，或方程 (1)、方程 (2)、方程 (3) 线性相关.

另外，方程 (1) 和方程 (2) 不能相互线性表示，称方程 (1) 和方程 (2) 线性无关. 方程组中线性无关的方程个数的最大值 r 称为方程组的秩，这 r 个线性无关的方程构成的方程组称为最大无关方程组，本例中 $r=2$. 最大无关方程组与原方程组同解.

消元法解线性方程组是一种重要的方法，但当未知数的个数增加时，解题过程会迅速变得非常冗长而繁杂，因而需要更加强有力的数学工具来解决这一问题，后续我们会看到，围绕线性方程组求解这一问题主线，线性代数提供了行列式、矩阵、向量组等理论方法，很好地解决了这个问题.

习　题　1

1. 设向量 $\boldsymbol{a}=(-3,4,-2,4)$，计算 $\|\boldsymbol{a}\|$.

2. k 为何值时向量 $\boldsymbol{\alpha}=(1,-2,2,-1)$ 与 $\boldsymbol{\beta}=(1,1,k,3)$ 正交？

3. 设向量 $\boldsymbol{x}=(1,a,b)$ 与向量 $\boldsymbol{\alpha}=(2,2,2),\boldsymbol{\beta}=(3,1,3)$ 都正交，求 a,b 的值.

4. 设向量 $\boldsymbol{\alpha}_1=(-1,1,2,-1),\boldsymbol{\alpha}_2=(0,3,8,-2),\boldsymbol{\alpha}_3=(3,1,2,2)$，计算 $\|3\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3\|$.

5. 设 $\boldsymbol{a},\boldsymbol{b}$ 都是 n 维单位列向量，求 $\boldsymbol{a}+\boldsymbol{b}$ 与 $\boldsymbol{a}-\boldsymbol{b}$ 的内积 $[\boldsymbol{a}+\boldsymbol{b},\boldsymbol{a}-\boldsymbol{b}]$.

6. 设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2\in\mathbb{R}^n$，$\|\boldsymbol{\alpha}_1\|=\|\boldsymbol{\alpha}_2\|=1$，且内积 $[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2]=\dfrac{1}{4}$，求 $\|\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2\|$.

7. 当 k 为何值时，向量 $\boldsymbol{\beta}=(1,k,5)^{\mathrm{T}}$ 能由向量 $\boldsymbol{\alpha}_1=(1,2,3)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,1,1)^{\mathrm{T}}$ 线性表示？

8. 已知向量 $\boldsymbol{a},\boldsymbol{b}$，两者之间夹角为 θ，如图 1.24 所示，求向量 $\boldsymbol{b}$ 在向量 $\boldsymbol{a}$ 上的投影向量 $\boldsymbol{\xi}$.

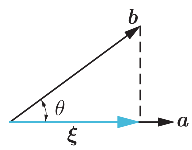

图 1.24　题 8 图