1.3 矩阵

1 $A=\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}$ 是一个 3×2 矩阵：$m=3$ 行且 $n=2$ 列。

2 $A\boldsymbol{x}=\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$ 是一个**列组合** $A\boldsymbol{x}=x_1\begin{bmatrix}1\\3\\5\end{bmatrix}+x_2\begin{bmatrix}2\\4\\6\end{bmatrix}$。

3 $A\boldsymbol{x}$ 的三个分量是 A 的三个行与向量 $\boldsymbol{x}$ 的点积：

$$\textbf{一次一行}\qquad A=\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}\begin{bmatrix}7\\8\end{bmatrix}=\begin{bmatrix}1\cdot7+2\cdot8\\3\cdot7+4\cdot8\\5\cdot7+6\cdot8\end{bmatrix}=\begin{bmatrix}23\\53\\83\end{bmatrix}$$

4 方程的矩阵形式 $A\boldsymbol{x}=\boldsymbol{b}$：$\begin{bmatrix}2&5\\3&7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}b_1\\b_2\end{bmatrix}$ 取代 $\begin{matrix}2x_1+5x_2=b_1\\3x_1+7x_2=b_2\end{matrix}$。

5 $A\boldsymbol{x}=\boldsymbol{b}$ 的解可写作 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$。但有的矩阵不给予 A^{-1}。

本节从三个向量 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ 开始。我将利用矩阵来组合它们。

$$\textbf{三个向量}\qquad \boldsymbol{u}=\begin{bmatrix}1\\-1\\0\end{bmatrix}\qquad \boldsymbol{v}=\begin{bmatrix}0\\1\\-1\end{bmatrix}\qquad \boldsymbol{w}=\begin{bmatrix}0\\0\\1\end{bmatrix}。$$

它们在三维空间中的线性组合为 $x_1\boldsymbol{u}+x_2\boldsymbol{v}+x_3\boldsymbol{w}$：

$$\textbf{向量的组合}\qquad x_1\begin{bmatrix}1\\-1\\0\end{bmatrix}+x_2\begin{bmatrix}0\\1\\-1\end{bmatrix}+x_3\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}x_1\\x_2-x_1\\x_3-x_2\end{bmatrix}. \tag{1}$$

现有一件重要的事：利用矩阵来重写这个组合。向量 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ 成为矩阵 A 的列。这个矩阵“**乘以**”向量 (x_1,x_2,x_3)：

$$\textbf{矩阵乘以向量 列的组合}\qquad A\boldsymbol{x}=\begin{bmatrix}1&0&0\\-1&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}x_1\\x_2-x_1\\x_3-x_2\end{bmatrix}. \tag{2}$$

数 x_1,x_2,x_3 均为向量 $\boldsymbol{x}$ 的分量。矩阵 A 乘以向量 $\boldsymbol{x}$ 与方程（1）中三个列的组合 $x_1\boldsymbol{u}+x_2\boldsymbol{v}+x_3\boldsymbol{w}$ **相同**。

由于此改写造成了在视角上的重要改变，所以这不仅仅是 $A\boldsymbol{x}$ 的定义。起初，数 x_1, x_2, x_3 乘上了向量。现在矩阵乘上了这些数。**矩阵 A 作用于向量 $\boldsymbol{x}$**。$A\boldsymbol{x}$ 的结果是 A **列的组合 $\boldsymbol{b}$**。

为了理解该运算行为，我将 $A\boldsymbol{x}$ 的分量写为 b_1, b_2, b_3：

$$A\boldsymbol{x} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2 - x_1} \\ \boldsymbol{x_3 - x_2} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \boldsymbol{b}. \tag{3}$$

其输入是 $\boldsymbol{x}$，输出是 $\boldsymbol{b} = A\boldsymbol{x}$。因为 $\boldsymbol{b}$ 包含输入向量 $\boldsymbol{x}$ 的差分，所以这个 A 是一个“**差分矩阵**”。顶部的差是 $x_1 - x_0 = x_1 - 0$。以下是个展示 $\boldsymbol{x} = (1,4,9)$ 差分的例子：$\boldsymbol{x}$ 含平方数，$\boldsymbol{b}$ 含奇数。

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} = \text{平方数} \qquad A\boldsymbol{x} = \begin{bmatrix} 1-0 \\ 4-1 \\ 9-4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \boldsymbol{b}. \tag{4}$$

这种模式会延续到 4×4 差分矩阵。下一个方阵应当是 $x_4 = 16$。下一个差分应该是 $x_4 - x_3 = 16 - 9 = 7$（下一个奇数）。矩阵一下求出了所有的差 $1, 3, 5, 7$。**重要注记：一次乘以一行**。你可能已经学过 $A\boldsymbol{x}$ 相乘，一个矩阵乘以一个向量。或许有不同的阐述，利用行而不是列。通常的方法是取每行与 $\boldsymbol{x}$ 的点积：

$A\boldsymbol{x}$ 也是与行的点积

$$A\boldsymbol{x} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} (1,0,0)\cdot(x_1,x_2,x_3) \\ (-1,1,0)\cdot(x_1,x_2,x_3) \\ (0,-1,1)\cdot(x_1,x_2,x_3) \end{bmatrix}. \tag{5}$$

这些点积与我们在方程（3）中写的 x_1，$x_2 - x_1$，$x_3 - x_2$ 相同。新方法是*一次使用 $A\boldsymbol{x}$ 的一列*。线性组合是线性代数的关键，$A\boldsymbol{x}$ 的结果是 A **列**的线性组合。

对于数，你可以按行将 $A\boldsymbol{x}$ 相乘。对于字母，按列是好的方法。第 2 章将重复这些矩阵乘法法则，并解释其中思想。

线性方程

另一个视角改变至关重要。到目前为止，数 x_1, x_2, x_3 已知。右手边的 $\boldsymbol{b}$ 未知。我们通过将 A 与 $\boldsymbol{x}$ 相乘来求出差分向量。**现在我们设想 $\boldsymbol{b}$ 为已知然后我们找 $\boldsymbol{x}$**。

旧问题：计算线性组合 $x_1\boldsymbol{u} + x_2\boldsymbol{v} + x_3\boldsymbol{w}$ 来求 $\boldsymbol{b}$。

新问题：$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 的哪种组合产生一个特定向量 $\boldsymbol{b}$？

这是*逆问题*——求能得出期望输出 $\boldsymbol{b} = A\boldsymbol{x}$ 的输入 $\boldsymbol{x}$。你以前见过这个问题，它作为一个关于 x_1, x_2, x_3 的线性方程组。方程右手边是 b_1, b_2, b_3。我现在要解方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 来求 x_1, x_2, x_3：

$$\begin{array}{l} \text{方程} \\ A\boldsymbol{x} = \boldsymbol{b} \end{array} \quad \begin{aligned} x_1 \quad &= b_1 \\ -x_1 + x_2 &= b_2 \\ -x_2 + x_3 &= b_3 \end{aligned} \qquad \begin{array}{l} \text{解} \\ \boldsymbol{x} = A^{-1}\boldsymbol{b} \end{array} \quad \begin{aligned} x_1 &= b_1 \\ x_2 &= b_1 + b_2 \\ x_3 &= b_1 + b_2 + b_3 \end{aligned} \tag{6}$$

允许我马上承认——大多数线性方程组都不易解出。在此例中，第一个方程决定了 $x_1 = b_1$。然后第二个方程得出 $x_2 = b_1 + b_2$。*因为 A 是一个三角矩阵，所以该方程可按顺序解出（从顶到底）*。

看一下右边 b_1, b_2, b_3 的两个特定选择 $0, 0, 0$ 和 $1, 3, 5$：

$$\boldsymbol{b}=\begin{bmatrix}0\\0\\0\end{bmatrix} \text{ 得 } \boldsymbol{x}=\begin{bmatrix}0\\0\\0\end{bmatrix} \qquad \boldsymbol{b}=\begin{bmatrix}1\\3\\5\end{bmatrix} \text{ 得 } \boldsymbol{x}=\boldsymbol{b}=\begin{bmatrix}1\\1+3\\1+3+5\end{bmatrix}=\boldsymbol{b}=\begin{bmatrix}1\\4\\9\end{bmatrix}。$$

第一个解（全零）比它看上去要重要。用句话说：*若输出为* $\boldsymbol{b}=\mathbf{0}$，*则输入必为* $\boldsymbol{x}=\mathbf{0}$。这句陈述对这个矩阵 A 来说是对的。它并不对所有矩阵是对的。我们的第二个例子将演示（对一个不同的矩阵 C）当 $C\neq 0$ 且 $\boldsymbol{x}\neq\mathbf{0}$ 时我们如何使 $C\boldsymbol{x}=\mathbf{0}$。

这个矩阵 A 是“**可逆的**”。我们可从 $\boldsymbol{b}$ 重新得出 $\boldsymbol{x}$。我们将 $\boldsymbol{x}$ 写作 $A^{-1}\boldsymbol{b}$。

可逆矩阵

让我复述一遍方程（6）的解 $\boldsymbol{x}$。将会出现一个和矩阵！

$$A\boldsymbol{x}=\boldsymbol{b}\text{的解为}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}b_1\\b_1+b_2\\b_1+b_2+b_3\end{bmatrix}=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}. \tag{7}$$

若 x 的差是 b，则 b 的和是 x。这对于奇数 $\boldsymbol{b}=(1,3,5)$ 与平方数 $\boldsymbol{x}=(1,4,9)$ 成立。它对所有向量都成立。方程（7）中的和矩阵是差分矩阵 A 的逆 A^{-1}。

例：$\boldsymbol{x}=(1,2,3)$ 的差分为 $\boldsymbol{b}=(1,1,1,)$。于是 $\boldsymbol{b}=A\boldsymbol{x}$ 且 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$：

$$A\boldsymbol{x}=\begin{bmatrix}1&0&0\\-1&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}\mathbf{1}\\\mathbf{2}\\\mathbf{3}\end{bmatrix}=\begin{bmatrix}\mathbf{1}\\\mathbf{1}\\\mathbf{1}\end{bmatrix} \qquad A^{-1}\boldsymbol{b}=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}\mathbf{1}\\\mathbf{1}\\\mathbf{1}\end{bmatrix}=\begin{bmatrix}\mathbf{1}\\\mathbf{2}\\\mathbf{3}\end{bmatrix}$$

解向量 $\boldsymbol{x}=(x_1,x_2,x_3)$ 的方程（7）告诉我们两个重要事实：

1. 对每个 $\boldsymbol{b}$，$A\boldsymbol{x}=b$ 都有唯一解。 **2.** 矩阵 A^{-1} 得出 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$。

下一章会问到其它方程 $A\boldsymbol{x}=\boldsymbol{b}$。有解么？怎么求解？

微积分注记。 让我将这些特殊矩阵与微积分联系起来。向量 $\boldsymbol{x}$ 转为函数 $x(t)$。差分 $A\boldsymbol{x}$ 变成导数 $dx/dt=b(t)$。在反方向上，加和 $A^{-1}\boldsymbol{b}$ 变成 $b(t)$ 的*积分*。**对差分加和就像对导数积分一样。**

微积分的基本定理表示：**积分是微分的逆过程。**

$$\boldsymbol{Ax=b} \text{ 与 } \boldsymbol{x=A^{-1}b} \qquad \frac{dx}{dt}=b \text{ 及 } x(t)=\int_0^t b\,dt \tag{8}$$

平方数 $0,1,4,9$ 的差分是奇数 $1,3,5$。$x(t)=t^2$ 的导数是 $2t$。一个理想的类推是要在 $t=1,2,3$ 时刻产生偶数 $b=2,4,6$。然而差分与导数不一致，并且我们的矩阵 A 要得出 $2t$ 而不是 $2t-1$：

$$\textbf{后向差分} \qquad x(t)-x(t-1)=t^2-(t-1)^2=t^2-(t^2-2t+1)=2t-1. \tag{9}$$

随后的学业中将证明“前向差分”得出 $2t+1$。其最佳选择（微积分课程中不常见）是运用 $x(t+1)-x(t-1)$ 的“中心差分”。将那个 Δt 除以从 $t-1$ 到 $t+1$ 的距离——2：

$$\boldsymbol{x(t)=t^2} \textbf{ 的中心差分} \qquad \frac{(t+1)^2-(t-1)^2}{2}=2t \quad \text{一点不差} \tag{10}$$

差分矩阵是重要的。中心差分是最佳的。我们的第二个例子是*不可逆的*。

本例除了将 $\boldsymbol{w}$ 改为新向量 $\boldsymbol{w}^*$ 外，保持相同的列 $\boldsymbol{u}$ 和 $\boldsymbol{v}$：

第二个例子 $\boldsymbol{u}=\begin{bmatrix}1\\-1\\0\end{bmatrix}$ $\boldsymbol{v}=\begin{bmatrix}0\\1\\-1\end{bmatrix}$ $\boldsymbol{w}^*=\begin{bmatrix}\mathbf{-1}\\0\\1\end{bmatrix}$。如今 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}^*$ 导出了一个**循环差分矩阵** C：

$$\textbf{循环差分}\qquad C\boldsymbol{x}=\begin{bmatrix}1&0&\mathbf{-1}\\-1&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}x_1-x_3\\x_2-x_1\\x_3-x_2\end{bmatrix}=\boldsymbol{b}. \tag{11}$$

这个矩阵 C 不是三角阵。当我们给出 $\boldsymbol{b}$ 时，并不易解出 $\boldsymbol{x}$。事实上三个方程要么有**无穷多解**（有时）要么**无解**（通常），因此求出 $C\boldsymbol{x}=\boldsymbol{b}$ 的解挺重要的：

$$\begin{array}{l}C\boldsymbol{x}=0\\ \text{无穷多解}\end{array}\qquad \begin{bmatrix}x_1-x_3\\x_2-x_1\\x_3-x_2\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\ \text{的解为所有向量}\ \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}c\\c\\c\end{bmatrix}. \tag{12}$$

当我们循环计算时，每个像 $\boldsymbol{x}=(3,3,3)$ 的常数向量都具有零差分。不定常数 c 就像我们往积分中添加的 $+C$ 一样。循环差分的第一个分量是循环回到 x_1-x_3，而不是从 $x_0=0$ 开始。

$C\boldsymbol{x}=\boldsymbol{b}$ 的最大可能性是根本就**无解**：

$$C\boldsymbol{x}=\boldsymbol{b}\qquad \begin{bmatrix}x_1-x_3\\x_2-x_1\\x_3-x_2\end{bmatrix}=\begin{bmatrix}1\\3\\5\end{bmatrix}\qquad \begin{array}{l}\text{左边加起来为 }0\\ \text{右边加起来为 }9\\ x_1,x_2,x_3\ \text{无解}\end{array} \tag{13}$$

从几何上看下这个例子。没有 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}^*$ 的组合会产生向量 $\boldsymbol{b}=(1,3,5)$。它们的组合不会占满整个三维空间。因为方程左边 x_1-x_3,x_2-x_1,x_3-x_2 总是加起来为 0，所以右边必须为 $b_1+b_2+b_3=0$ 才能容许 $C\boldsymbol{x}=\boldsymbol{b}$ 有解。换句话说：

所有线性组合 $x_1\boldsymbol{u}+x_2\boldsymbol{v}+x_3\boldsymbol{w}^*$ 均位于由 $b_1+b_2+b_3=0$ 给定的平面上。

本节课突然将代数和几何联系起来。线性组合可以充满整个空间，或仅仅一个平面。我们需要一张图来展示 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$（第一个例子）与 $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}^*$（都在同一平面）间至关重要的差别。

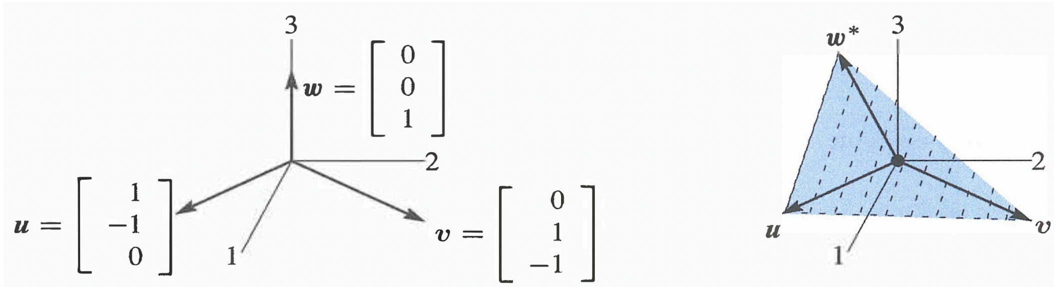

图 1.10：无关向量 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$。相关向量 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}^*$ 在一平面内。

线性无关与线性相关

图 1.10 展示了那些列向量，先是矩阵 A 的然后是 C 的。在两个图中，头两列 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 都是相同的。如果我们只考虑这两个向量的组合，我们将会得到一个二维平面。**关键问题是第三条向量是否在此平面上**：

无关　$\boldsymbol{w}$ 不在 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 张成的平面内。

相关　$\boldsymbol{w}^*$ 在 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 张成的平面内。

最重要的一点是新向量 $\boldsymbol{w}^*$ 是 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的线性组合：

$$\boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}^*=0 \qquad \boldsymbol{w}^*=\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}=-\boldsymbol{u}-\boldsymbol{v}. \tag{14}$$

三个向量 $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 的全部分量加和为 0。因此它们的全部分量将会使 $b_1+b_2+b_3=0$（正如我们在前文所看到的，通过将三个方程相加）。这是关于 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 所有组合的平面方程。由于 $\boldsymbol{w}^*$ 已经在该平面上，因此凭借囊括进 $\boldsymbol{w}^*$ 我们也*无法得到新向量*。

最初的 $\boldsymbol{w}=(0,0,1)$ 不在那平面上：$0+0+1 \neq 0$。$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 的组合占满整个三维空间。我们对此早有知晓，是因为方程（6）中的解 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$ 给出了构造任意 $\boldsymbol{b}$ 的正确组合。

带有第三列 $\boldsymbol{w}$ 与 $\boldsymbol{w}^*$ 的两个矩阵 A 与 C，让我论及两个线性代数关键词：（线性）无关与相关。课程的前半部分将进一步深入这些思想——如果你能较早的从这两个例子中理解它们的话那我很高兴：

$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ 是**无关的**。除了 $0\boldsymbol{u}+0\boldsymbol{v}+0\boldsymbol{w}=\mathbf{0}$ 外没有组合可得 $\boldsymbol{b}=\mathbf{0}$。

$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}^*$ 是**相关的**。其它组合如 $\boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}^*$，得出 $\boldsymbol{b}=\mathbf{0}$。

在三维中你可以将这些描绘出来。三个向量位于一个平面或并非如此。第 2 章在 n 维空间中拿 n 个向量。*无关与相关*是关键点。向量成为 $n \times n$ 矩阵的列：

列无关：$A\boldsymbol{x}=\mathbf{0}$ 有唯一解。A 是**可逆矩阵**。

列相关：$C\boldsymbol{x}=\mathbf{0}$ 有多个解。C 是**奇异矩阵**。

最终我们将在 m 维空间中拿 n 个向量。带这 n 个列的矩阵 A 现在是*矩形*（$m \times n$）。弄懂 $A\boldsymbol{x}=\boldsymbol{b}$ 是第 3 章的问题。

■ 复习关键点 ■

1. **矩阵乘以向量：**$A\boldsymbol{x} = A$ **的列组合**。
2. 当 A 是可逆矩阵时，$A\boldsymbol{x} = \boldsymbol{b}$ 的解是 $\boldsymbol{x} = A^{-1}\boldsymbol{b}$。
3. 循环差分矩阵 C 没有逆。它的三个列位于同一平面。这些相关列的加和为零向量。$C\boldsymbol{x} = \mathbf{0}$ 有多个解。
4. 本节展望关键思想，并未做完整解释。