第 1 章 线性代数的重点

本书第 1 章是对应用线性代数的严肃介绍。如果读者的知识背景或线性代数基础不是很强，请不要匆忙地读完本章。本章首先用矩阵 $\boldsymbol{A}$ 的列说明乘法 $\boldsymbol{Ax}$ 和 $\boldsymbol{AB}$ 。这可能看起来只是形式上的，但实际上是十分基本的。

本章研究五个基本问题：

$$\boxed{\boldsymbol{Ax}=\boldsymbol{b}, \quad \boldsymbol{Ax}=\lambda\boldsymbol{x}, \quad \boldsymbol{Av}=\sigma\boldsymbol{u}, \quad \text{Minimize } \|\boldsymbol{Ax}\|^2/\|\boldsymbol{x}\|^2, \quad \text{分解矩阵}\boldsymbol{A}}$$

每个问题看起来都像一个普通的计算问题：

$$\text{寻找 } \boldsymbol{x}, \quad \text{寻找 } \boldsymbol{x} \text{ 与 } \lambda, \quad \text{寻找 } \boldsymbol{v}\text{、}\boldsymbol{u} \text{ 和 } \sigma, \quad \text{分解 } \boldsymbol{A}=\text{列}\times\text{行}$$

你将看到我们的目标是理解这些问题，这甚至比求解更为重要。首先想知道 $\boldsymbol{Ax}=\boldsymbol{b}$ 是否有解 $\boldsymbol{x}$。“向量 $\boldsymbol{b}$ 是否在 $\boldsymbol{A}$ 的列空间中?”“空间” 这个看上去平淡的词其实具有丰富的内涵，采用 “空间” 这个词是十分有效的。

特征值方程 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$ 则大为不同。这个方程中没有向量 $\boldsymbol{b}$，只有矩阵 $\boldsymbol{A}$ 的信息。想要得到特征向量的方向，以便 $\boldsymbol{Ax}$ 与 $\boldsymbol{x}$ 保持相同的方向。于是沿着这条线，与 $\boldsymbol{A}$ 相关的所有复杂性都消失了。向量 $\boldsymbol{A}^2\boldsymbol{x}$ 只是简单的 $\lambda^2\boldsymbol{x}$。矩阵 $\mathrm{e}^{\boldsymbol{A}t}$（这来自微分方程）只是将 $\boldsymbol{x}$ 乘以 $\mathrm{e}^{\lambda t}$。当知道每个 $\boldsymbol{x}$ 和 λ 时，就可以求解线性问题。

方程 $\boldsymbol{Av}=\sigma\boldsymbol{u}$ 类似，但也有所不同。有两个向量 $\boldsymbol{v}$ 和 $\boldsymbol{u}$。矩阵 $\boldsymbol{A}$ 可能是矩形的（非方形），并且填满了数据（非稀疏矩阵）。那么这个数据矩阵的哪一部分是重要的？运用奇异值分解（Singular Value Decomposition，SVD）可以找到它最简单的成分 $\sigma\boldsymbol{uv}^{\mathrm{T}}$。这些是矩阵（列 $\boldsymbol{u}$ 乘以行 $\boldsymbol{v}^{\mathrm{T}}$）。每个矩阵都是由这些正交的成分矩阵组成的。这样，**数据科学在 SVD 中遇到了线性代数**。

找到这些 $\sigma\boldsymbol{uv}^{\mathrm{T}}$ 成分矩阵就是主成分分析（Principal Component Analysis，PCA）的目标。

最小化和分解表达了基本的应用问题，它们导致了奇异向量 $\boldsymbol{v}$ 和 $\boldsymbol{u}$。计算最小二乘中的最佳向量 $\widehat{\boldsymbol{x}}$ 与 PCA 中的主成分 $\boldsymbol{v}_1$ 是拟合数据的代数问题。我们不会提供代码（那属于线上的工作），而是解释思路及想法。

当你理解清楚了列空间、零空间以及特征向量和奇异向量后，就可以应对各类应用，如最小二乘、傅里叶变换、统计中的 LASSO 以及使用神经网络进行深度学习中的随机梯度下降。

1.1 使用 $\boldsymbol{A}$ 的列向量实现 $\boldsymbol{Ax}$ 的相乘

我们希望读者已经有了一些有关线性代数的知识。这是一个十分迷人的学科——对更多的人而言线性代数可能比微积分还要有用（这是我们不会大声声张的观点）。但是，即使是传统的

线性代数课程，也错过了一些基本和重要的结论。第 1 章介绍矩阵向量的乘积 $\boldsymbol{Ax}$，以及矩阵的列空间和秩。

我们经常用例子来清楚地阐明观点。

例 1 用 $\boldsymbol{A}$ 的三行将 $\boldsymbol{A}$ 乘以 $\boldsymbol{x}$，然后用两列乘以 $\boldsymbol{x}$。

通过行：

$$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 \\ 2x_1 + 4x_2 \\ 3x_1 + 7x_2 \end{bmatrix} \quad (\text{行与}\boldsymbol{x} = (x_1, x_2)\text{的内积})$$

通过列：

$$\begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} \quad (\text{列}\boldsymbol{a}_1\text{与}\boldsymbol{a}_2\text{的组合})$$

由上看到两种方法可以得到相同的结果。第一种方法（一次一行）产生三个内积。由于用了点符号，这些运算也称为“点乘”：

$$\text{行} \cdot \text{列} = (2,\ 3) \cdot (x_1, x_2) = 2x_1 + 3x_2 \tag{1}$$

这是得到 $\boldsymbol{Ax}$ 三个独立分量的方法。这种做法是一个计算过程，而不是为了理解。因为这是一个低级运算过程。从理解这个更高的层次来看，可以使用向量法。

向量法将 $\boldsymbol{Ax}$ 视为 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 的“线性组合”，这是线性代数的基本运算。$\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 的线性组合包括两个步骤：

(1) 将两列 $\boldsymbol{a}_1$、$\boldsymbol{a}_2$ 分别和“数”x_1、x_2 相乘。

(2) 将这两个向量相加得到 $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 = \boldsymbol{Ax}$。

因此 $\boldsymbol{Ax}$ 是 $\boldsymbol{A}$ 的各个列的线性组合，这是根本性的。

这种想法将人们引向 $\boldsymbol{A}$ 的列空间的概念。关键想法是取 $\boldsymbol{A}$ 中各列的**所有组合**。所有实数 x_1、x_2 都是允许的——这个空间包含形如 $\boldsymbol{Ax}$ 的向量，其中 $\boldsymbol{x}$ 取所有向量。这样，得到了无限多的输出向量 $\boldsymbol{Ax}$，可以从几何上看到这些输出。

在这个例子中，每个 $\boldsymbol{Ax}$ 都是三维空间中的向量。该三维空间称为 $\mathbf{R}^3$（此处 $\mathbf{R}$ 表示实数，有三个复数分量的向量则位于 $\mathbf{C}^3$ 空间中）。只考虑实向量，有下面这个关键问题：

所有组合$\boldsymbol{Ax} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$产生整个三维空间中的哪些部分

答案是这些向量组成一个平面。这个平面包含沿 $\boldsymbol{a}_1 = (2, 2, 3)$ 方向的整条直线，因为每个 $x_1\boldsymbol{a}_1$ 向量都包括在内。这个平面还包含了沿 $\boldsymbol{a}_2$ 方向上的所有 $x_2\boldsymbol{a}_2$ 向量所在的那条直线。进一步而言，这个平面包括了在一条直线上的任何向量与在另一条直线上的任何向量之和。**这两个向量的求和填充了包含这两条直线的无限大平面**。但它并未填充满整个三维空间 $\mathbf{R}^3$。

定义 **列向量的组合构成了 $\boldsymbol{A}$ 的列空间。**

这里的列空间是一个平面，该平面包括零点 $(0, 0, 0)$，这是当 $x_1 = x_2 = 0$ 时产生的。平面包括 $(5, 6, 10) = \boldsymbol{a}_1 + \boldsymbol{a}_2$ 和 $(-1, -2, -4) = \boldsymbol{a}_1 - \boldsymbol{a}_2$。每个组合 $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2$ 在列空间中。但这

个列空间一定（概率是 1）不包括由 MATLAB 函数 rand(3,1) 产生的随机数。那么哪些点在这个平面上呢？

$$\boxed{\boldsymbol{b}=(b_1,b_2,b_3)\text{属于}\boldsymbol{A}\text{的列空间当且仅当}\boldsymbol{Ax}=\boldsymbol{b}\text{有解}\ \boldsymbol{x}=(x_1,x_2)}$$

只有看清这个事实，才能理解列空间 $\mathbf{C}(\boldsymbol{A})$：解 $\boldsymbol{x}$ 体现了如何将等式右侧的 $\boldsymbol{b}$ 表示为列向量的线性组合 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2$。对某些 $\boldsymbol{b}$，这是不可能的，因为它们不在列空间中。

例 2 $\boldsymbol{b}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 不属于 $\mathbf{C}(\boldsymbol{A})$。相应地，有 $\boldsymbol{Ax}=\begin{bmatrix}2x_1+3x_2\\2x_1+4x_2\\3x_1+7x_2\end{bmatrix}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ 无解。第一个方程和第二个方程导致 $x_1=\dfrac{1}{2}$，$x_2=0$。但这时

$$3\left(\frac{1}{2}\right)+7(0)=1.5\neq 1$$

第三个方程就不满足。这意味着，$\boldsymbol{b}=(1,1,1)$ 不属于这个列空间——即由 $\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 组成的平面。

例 3 求矩阵 $\boldsymbol{A}_2=\begin{bmatrix}2&3&5\\2&4&6\\3&7&10\end{bmatrix}$ 与 $\boldsymbol{A}_3=\begin{bmatrix}2&3&1\\2&4&1\\3&7&1\end{bmatrix}$ 的列空间。

解：$\boldsymbol{A}_2$ 的列空间与之前一样是同一个平面。新的列 $(5,6,10)$ 是第 1 列与第 2 列的和。因此第 3 列 $\boldsymbol{a}_3$ 已经在这个平面中，并没有添加任何新的信息。添加这个“相关”的列，得到的列空间还是原来的平面。

$\boldsymbol{A}_3$ 的列空间是整个三维空间 $\mathbf{R}^3$。例 2 展示了新的第 3 列 $(1,1,1)$ 不在平面 $\mathbf{C}(\boldsymbol{A})$ 中。列空间 $\mathbf{C}(\boldsymbol{A}_3)$ 变大。但是，在一个平面与整个三维空间之间并没有空当。xy 平面与第三个不在这个平面上的向量 (x_3,y_3,z_3)（这意味着 $z_3\neq 0$）的组合得到在 $\mathbf{R}^3$ 中的每个向量。

下面是 $\mathbf{R}^3$ 中所有可能的列空间的总列表，维数分别为零维、一维、二维和三维：

$\mathbf{R}^3$的子空间	**零向量**$(0,0,0)$ 本身
	所有向量 $x_1\boldsymbol{a}_1$ 的集合是一条**直线**
	所有向量 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2$ 的集合是一个**平面**
	所有向量 $x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+x_3\boldsymbol{a}_3$ 的集合是**整个$\mathbf{R}^3$** 空间

在上面总列表中，要求向量 $\boldsymbol{a}_1$、$\boldsymbol{a}_2$、$\boldsymbol{a}_3$ 是“**独立**”的。零向量的唯一组合是 $0\boldsymbol{a}_1+0\boldsymbol{a}_2+0\boldsymbol{a}_3$。因此，$\boldsymbol{a}_1$ 给出一条直线，$\boldsymbol{a}_1$ 和 $\boldsymbol{a}_2$ 给出一个平面，$\boldsymbol{a}_1$、$\boldsymbol{a}_2$ 和 $\boldsymbol{a}_3$ 给出 $\mathbf{R}^3$ 中的每个 $\boldsymbol{b}$。零向量在每个子空间中都存在。在线性代数的语言中：

(1) $\mathbf{R}^3$ 中的三个独立的列构成一个**可逆矩阵**：$\boldsymbol{AA}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}$。

(2) $\boldsymbol{Ax}=\mathbf{0}$ 要求 $\boldsymbol{x}=(0,0,0)$；然后 $\boldsymbol{Ax}=\boldsymbol{b}$，只有 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}$ 这个唯一的解。

由此给定一个 $n\times n$ 可逆矩阵的列的图，这些列的组合填充了此矩阵的列空间：整个 $\mathbf{R}^n$。我们需要将这样的想法和语言进一步拓展。

$\boldsymbol{A}$ 的独立列与秩

在介绍了以上内容后，本节并未结束，将找到 $\boldsymbol{A}$ 的列空间的一组**基**，并将 $\boldsymbol{A}$ **分解**成 $\boldsymbol{C}$ 乘以 $\boldsymbol{R}$，而且可以证明线性代数中的**第一个大定理**，还可以看到矩阵的秩和子空间的维数。

所有这一切来自对**独立**性的理解。我们的目标是构建一个矩阵 $\boldsymbol{C}$，其所有的列直接来自 $\boldsymbol{A}$，但不包括前几列任意组合而得的列。矩阵 $\boldsymbol{C}$ 的列（希望尽可能多）将是"独立的"。下面从 $\boldsymbol{A}$ 的 n 列中得到矩阵 $\boldsymbol{C}$ 的自然构造：

若 $\boldsymbol{A}$ 的第 1 列不全为零，则将其放入矩阵 $\boldsymbol{C}$ 中。

若 $\boldsymbol{A}$ 的第 2 列不是第 1 列的倍数，则将其放入 $\boldsymbol{C}$ 中。

若 $\boldsymbol{A}$ 的第 3 列不是第 1 列和第 2 列的组合，则将其放入 $\boldsymbol{C}$ 中。

……

最终 $\boldsymbol{C}$ 会有 r 列 $(r \leqslant n)$。

它们将成为 $\boldsymbol{A}$ 的列空间的"基"。

$\boldsymbol{A}$ 中剩余的列是 $\boldsymbol{C}$ 中这些基列的组合。

子空间的**基**是一整套独立的向量：**空间中的所有向量都是基向量的组合**。下面的例子将说明其要点。

例 4 若 $\boldsymbol{A}=\begin{bmatrix}1&3&8\\1&2&6\\0&1&2\end{bmatrix}$，则 $\boldsymbol{C}=\begin{bmatrix}1&3\\1&2\\0&1\end{bmatrix}$ $\begin{pmatrix}n=\boldsymbol{A}\text{的列数}=3\\ r=\boldsymbol{C}\text{的列数}=2\end{pmatrix}$

$\boldsymbol{A}$ 的第 3 列等于 $2(\boldsymbol{A}\text{的第1列})+2(\boldsymbol{A}\text{的第2列})$，因此不将其留在 $\boldsymbol{C}$ 中作为基向量。

例 5 若 $\boldsymbol{A}=\begin{bmatrix}1&2&3\\0&4&5\\0&0&6\end{bmatrix}$，则 $\boldsymbol{C}=\boldsymbol{A}$ $\begin{pmatrix}n=\boldsymbol{A}\text{的列数}=3\\ r=\boldsymbol{C}\text{的列数}=3\end{pmatrix}$

矩阵 $\boldsymbol{A}$ 是可逆的。它的列空间就是整个 $\mathbf{R}^3$。保留所有 3 列。

例 6 若 $\boldsymbol{A}=\begin{bmatrix}1&2&5\\1&2&5\\1&2&5\end{bmatrix}$，则 $\boldsymbol{C}=\begin{bmatrix}1\\1\\1\end{bmatrix}$ $\begin{pmatrix}n=\boldsymbol{A}\text{的列数}=3\\ r=\boldsymbol{C}\text{的列数}=1\end{pmatrix}$

数 r 是 $\boldsymbol{A}$ 的"**秩**"，同时也是 $\boldsymbol{C}$ 的秩。矩阵的秩是独立的列的数目。也可以从 $\boldsymbol{A}$ 的最后一列开始，从右数到左，这不会改变最终计数 r。这样会有不同的基，但是基向量的数目总是相同的。这个数 r 是 $\boldsymbol{A}$ 和 $\boldsymbol{C}$ 的列空间（它们是相同的空间）的"维数"。

矩阵的秩是其列空间的维数

矩阵 $\boldsymbol{C}$ 通过第三个矩阵 $\boldsymbol{R}$ 与 $\boldsymbol{A}$ 关联：$\boldsymbol{A}=\boldsymbol{CR}$。它们的形状是 $(m\times n)=(m\times r)\,(r\times n)$。能够由例 4 展示这个"$\boldsymbol{A}$ 的分解"：

$$\boxed{\boldsymbol{A}=\begin{bmatrix}1&3&8\\1&2&6\\0&1&2\end{bmatrix}=\begin{bmatrix}1&3\\1&2\\0&1\end{bmatrix}\begin{bmatrix}1&0&2\\0&1&2\end{bmatrix}=\boldsymbol{CR}} \tag{2}$$

当 $\boldsymbol{C}$ 乘以 $\boldsymbol{R}$ 的第 1 列 $\begin{bmatrix}1\\0\end{bmatrix}$ 时，就产生了 $\boldsymbol{C}$ 和 $\boldsymbol{A}$ 的第 1 列。

当 $\boldsymbol{C}$ 乘以 $\boldsymbol{R}$ 的第 2 列 $\begin{bmatrix}0\\1\end{bmatrix}$ 时，就得到 $\boldsymbol{C}$ 和 $\boldsymbol{A}$ 的第 2 列。

当 $\boldsymbol{C}$ 乘以 $\boldsymbol{R}$ 的第 3 列 $\begin{bmatrix}2\\2\end{bmatrix}$ 时，就得到 $2\,(\boldsymbol{C}\text{的第 1列}) + 2\,(\boldsymbol{C}\text{的第 2列})$。

这个结果与 $\boldsymbol{A}$ 的 3 列结果一致。将正确的数放入 $\boldsymbol{R}$ 中，$\boldsymbol{C}$ 中的列组合就产生了 $\boldsymbol{A}$ 中的各列，然后 $\boldsymbol{A}=\boldsymbol{CR}$ 将这些信息存储为矩阵的乘积。实际上，$\boldsymbol{R}$ 是线性代数中的一个著名的矩阵：

$$\boldsymbol{R} = \text{rref}(\boldsymbol{A}) = \boldsymbol{A}\textbf{的行约化阶梯型}\ (\text{没有全为零的行})$$

例 5 中 $\boldsymbol{C}=\boldsymbol{A}$，$\boldsymbol{R}=\boldsymbol{I}$（单位矩阵）。例 6 中，$\boldsymbol{C}$ 只有 1 列，因此在 $\boldsymbol{R}$ 中只有 1 行：

$$\boldsymbol{A} = \begin{bmatrix}1&2&5\\1&2&5\\1&2&5\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1&2&5\end{bmatrix} = \boldsymbol{CR} \quad \left(\begin{array}{c}\text{所有这些矩阵的秩 } r=1\\ \textbf{列秩} = \textbf{行秩}\end{array}\right)$$

$$\boxed{\text{独立列的列数等于独立行的行数}}$$

这个秩定理适用于每个矩阵。在线性代数中，总是处理列和行。m 行与 n 列包含同样的数 a_{ij}，但是代表不同的向量。

这个定理由 $\boldsymbol{A}=\boldsymbol{CR}$ 所证明。从另一个角度来看：通过行而不是列。矩阵 $\boldsymbol{R}$ 有 r 行。**用$\boldsymbol{C}$来乘，并取这些行的组合**。因为 $\boldsymbol{A}=\boldsymbol{CR}$，从 $\boldsymbol{R}$ 的 r 行得到 $\boldsymbol{A}$ 的每一行，这些 r 行是互相独立的，这样它们是 $\boldsymbol{A}$ 的**行空间的基**。$\boldsymbol{A}$ 的列空间与行空间的维数都为 r，其 r 个基向量是 $\boldsymbol{C}$ 的列与 $\boldsymbol{R}$ 的行。

为什么 $\boldsymbol{R}$ 有独立的行？再看一下例 4。

$$\boldsymbol{A} = \begin{bmatrix}1&3&8\\1&2&6\\0&1&2\end{bmatrix} = \begin{bmatrix}1&3\\1&2\\0&1\end{bmatrix}\underset{\substack{\uparrow\ \ \uparrow\\ 1\text{ 与 }0}}{\begin{bmatrix}1&0&2\\0&1&2\end{bmatrix}}\begin{matrix}\leftarrow\\ \leftarrow\end{matrix}\ \boldsymbol{R}\text{的独立行}$$

正是在 $\boldsymbol{R}$ 中的 1 与 0，可以看出：没有一行是其他行的线性组合。

数据科学中的一个重要分解是 $\boldsymbol{A}$ 的 “SVD”（奇异值分解）——第一个因子 $\boldsymbol{C}$ 有 r 个正交的列，而第二个因子 $\boldsymbol{R}$ 有 r 个正交的行。

习题 1.1

1. $\mathbf{R}^4$ 中三个非零向量的组合为零向量，并以 $\boldsymbol{Ax}=\mathbf{0}$ 形式表示。$\boldsymbol{A}$、$\boldsymbol{x}$ 和 $\mathbf{0}$ 的形状是怎样的？举例给出。

2. 假设 $\boldsymbol{A}$ 的一个列组合等于这些列的另一个不同的组合，记为 $\boldsymbol{Ax}=\boldsymbol{Ay}$。再找出 $\boldsymbol{A}$ 的两个列组合，使得它们等于零向量（在矩阵语言中，即找到 $\boldsymbol{Az}=\mathbf{0}$ 的两个解）。
3. (练习下标的用法) 向量 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 在 m 维空间 $\mathbf{R}^m$ 中，且组合 $c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2+\cdots+c_n\boldsymbol{a}_n$ 是零向量，这是向量层面的表述。
   (1) 试用矩阵语言写出上面的表述。用这些 $\boldsymbol{a}$ 向量作为矩阵 $\boldsymbol{A}$ 的列，并采用列向量 $\boldsymbol{c}=(c_1,c_2,\cdots,c_n)$。
   (2) 将上面的表述在标量层面写出。采用下标与 Σ 求和形式将数加起来。列向量 $\boldsymbol{a}_j$ 的分量是 $a_{1j},a_{2j},\cdots,a_{mj}$。
4. 假设 $\boldsymbol{A}$ 是 3×3 矩阵 $\mathbf{ones}(3,3)$，它的每个矩阵元素都是 1。求两个独立向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$，使其满足 $\boldsymbol{Ax}=\mathbf{0}$，$\boldsymbol{Ay}=\mathbf{0}$。将第一个方程 $\boldsymbol{Ax}=\mathbf{0}$ 写为 $\boldsymbol{A}$ 的列向量的组合。能求出第 3 个满足 $\boldsymbol{Az}=\mathbf{0}$ 的独立向量吗？
5. $\boldsymbol{v}=(1,1,0)$ 和 $\boldsymbol{w}=(0,1,1)$ 的线性组合填充了 $\mathbf{R}^3$ 中的一个平面。
   (1) 求一个垂直于 $\boldsymbol{v}$ 和 $\boldsymbol{w}$ 的向量 $\boldsymbol{z}$，则 $\boldsymbol{z}$ 垂直于平面中的任意向量 $c\boldsymbol{v}+d\boldsymbol{w}$：$(c\boldsymbol{v}+d\boldsymbol{w})^{\mathrm{T}}\boldsymbol{z}=c\boldsymbol{v}^{\mathrm{T}}\boldsymbol{z}+d\boldsymbol{w}^{\mathrm{T}}\boldsymbol{z}=0+0$。
   (2) 对一个不在这个平面上的向量 $\boldsymbol{u}$，验证 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{z}\neq 0$。
6. 若一个平行四边形的三个顶点的坐标是 $(1,1)$, $(4,2)$ 和 $(1,3)$，则第四个顶点的所有三种可能是什么？画出其中两个。
7. 描述 $\boldsymbol{A}=[\boldsymbol{v}\quad \boldsymbol{w}\quad \boldsymbol{v}+2\boldsymbol{w}]$ 的列空间。描述 $\boldsymbol{A}$ 的零空间，即所有满足 $\boldsymbol{Ax}=\mathbf{0}$ 的向量 $\boldsymbol{x}=(x_1,x_2,x_3)$ 的集合。将此平面（$\boldsymbol{A}$ 的列空间）与直线（$\boldsymbol{A}$ 的零空间）的“维数”加在一起：

$$\text{列空间的维数} + \text{零空间的维数} = \text{列数}$$

8. $\boldsymbol{A}=\boldsymbol{CR}$ 表示将 $\boldsymbol{A}$ 的列向量写为列空间的基的线性组合，其中基向量作为列向量构成矩阵 $\boldsymbol{C}$，系数构成矩阵 $\boldsymbol{R}$。若 3×3 矩阵 $\boldsymbol{A}_{ij}=j^2$，写出 $\boldsymbol{A}$、$\boldsymbol{C}$ 与 $\boldsymbol{R}$。
9. 假设一个 $m\times n$ 矩阵的列空间是 $\mathbf{R}^3$。m 代表什么？n 代表什么？秩 r 代表什么？
10. 求包含有 $\boldsymbol{A}_1$ 与 $\boldsymbol{A}_2$ 独立列的矩阵 $\boldsymbol{C}_1$ 和 $\boldsymbol{C}_2$：

$$\boldsymbol{A}_1=\begin{bmatrix}1 & 3 & -2\\ 3 & 9 & -6\\ 2 & 6 & -4\end{bmatrix}\qquad \boldsymbol{A}_2=\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{bmatrix}$$

11. 将习题 10 的矩阵分解为 $\boldsymbol{A}=\boldsymbol{CR}$ 的形式。矩阵 $\boldsymbol{R}$ 包含的数与 $\boldsymbol{C}$ 的列相乘就可以恢复 $\boldsymbol{A}$ 的列。这是看待矩阵乘法的一种方式，即**$\boldsymbol{C}$ 乘以$\boldsymbol{R}$ 的每一列**。
12. 给出 $\boldsymbol{A}_1$ 和 $\boldsymbol{A}_2$ 的列空间的基。求：这些列空间的维数（独立向量的数目），$\boldsymbol{A}_1$、$\boldsymbol{A}_2$ 的秩，$\boldsymbol{A}_1$、$\boldsymbol{A}_2$ 独立的行。
13. 构造一个秩为 2 的 4×4 矩阵。$\boldsymbol{C}$ 和 $\boldsymbol{R}$ 的形状是怎样的？
14. 假设两矩阵 $\boldsymbol{A}$、$\boldsymbol{B}$ 有相同的列空间。
   (1) 说明它们的行空间可以是不同的。
   (2) 说明矩阵 $\boldsymbol{C}$（即基列矩阵）可以是不同的。
   (3) 哪个数对矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是一样的？

15. 设 $\boldsymbol{A}=\boldsymbol{C}\boldsymbol{R}$，那么 $\boldsymbol{A}$ 的第一行是 $\boldsymbol{R}$ 各行的组合。哪个矩阵的哪部分存有这个组合的系数（用来乘 $\boldsymbol{R}$ 的行产生 $\boldsymbol{A}$ 的第一行的数）?
16. $\boldsymbol{R}$ 的行是 $\boldsymbol{A}$ 的行空间的一组基。这句话含义是什么？
17. 对下列分块矩阵，求 $\boldsymbol{A}=\boldsymbol{C}\boldsymbol{R}$。求它们的秩。

$$\boldsymbol{A}_1=\begin{bmatrix}\text{zeros} & \text{ones}\\ \text{ones} & \text{ones}\end{bmatrix}_{4\times 4},\quad \boldsymbol{A}_2=\begin{bmatrix}\boldsymbol{A}_1\\ \boldsymbol{A}_1\end{bmatrix}_{8\times 4},\quad \boldsymbol{A}_3=\begin{bmatrix}\boldsymbol{A}_1 & \boldsymbol{A}_1\\ \boldsymbol{A}_1 & \boldsymbol{A}_1\end{bmatrix}_{8\times 8}$$

18. 若 $\boldsymbol{A}=\boldsymbol{C}\boldsymbol{R}$，求矩阵 $\begin{bmatrix}\boldsymbol{0} & \boldsymbol{A}\\ \boldsymbol{0} & \boldsymbol{A}\end{bmatrix}$ 的 $\boldsymbol{C}\boldsymbol{R}$ 分解。
19. “消元法” 从第 i 行中减去第 j 行的 ℓ_{ij} 倍：这是“行运算”。给出步骤将例 4 中矩阵 $\boldsymbol{A}$ 化简为 $\boldsymbol{R}$（此处这个行阶梯形 $\boldsymbol{R}$ 有一行全为零）。其秩不会发生变化。

$$\boldsymbol{A}=\begin{bmatrix}1 & 3 & 8\\ 1 & 2 & 6\\ 0 & 1 & 2\end{bmatrix}\longrightarrow \boldsymbol{R}=\begin{bmatrix}1 & 0 & 2\\ 0 & 1 & 2\\ 0 & 0 & 0\end{bmatrix}=\mathbf{rref}(\boldsymbol{A})$$

新的开始与一本新的书

1.1 节启发我开始用一种全新的方式讲授线性代数，我意识到借助分量是小整数的矩阵，关键的概念能够在课程开始阶段给出。

下面解释这些概念，因为它们引导读者直接进入线性代数。这也促使我编写了另一本新的教科书，名为《适合每个人的线性代数》(*Linear Algebra for Everyone*)，读者可以访问 **math.mit.edu/everyone** 了解更多的信息。这些关键概念是:

独立的列　　列空间　　矩阵的秩　　矩阵乘积 $\boldsymbol{A}=\boldsymbol{C}\boldsymbol{R}$

$\boldsymbol{A}=\boldsymbol{C}\boldsymbol{R}$ 的例:

$$\begin{bmatrix}1 & 4 & 7\\ 2 & 5 & 8\\ 3 & 6 & 9\end{bmatrix}=\begin{bmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{bmatrix}\begin{bmatrix}1 & 0 & -1\\ 0 & 1 & 2\end{bmatrix}\quad\left(\begin{array}{l}\boldsymbol{C}\text{有 } \boldsymbol{A} \text{ 中的 2 个独立的列}\\ \boldsymbol{R}\text{有 2 个包含 } \boldsymbol{I} \text{ 的独立的行}\\ \boldsymbol{A}\text{的第 3 列} = -1\times(\text{第 1 列})+2\times(\text{第 2 列})\end{array}\right)$$

所有这三个矩阵的秩 $r=2$。$\boldsymbol{C}$ 表示 $\boldsymbol{A}$ 的列空间的一组基。$\boldsymbol{R}$ 表示 $\boldsymbol{A}$ 的行空间的一组基。这两个空间都是三维空间中的平面。$\boldsymbol{A}=\boldsymbol{C}\boldsymbol{R}$ 引出一个有三个因子紧密相连的形式，其中列与行对称地出现:

$$\boldsymbol{A}=\begin{bmatrix}1 & 4 & 7\\ 2 & 5 & 8\\ 3 & 6 & 9\end{bmatrix}=\begin{bmatrix}1 & 4\\ 2 & 5\\ 3 & 6\end{bmatrix}\begin{bmatrix}1 & 4\\ 2 & 5\end{bmatrix}^{-1}\begin{bmatrix}1 & 4 & 7\\ 2 & 5 & 8\end{bmatrix}\quad\left(\begin{array}{l}\text{来自 } \boldsymbol{A} \text{ 的 2 列在 } \boldsymbol{C} \text{ 中}\\ \boldsymbol{R}\text{来自 } \boldsymbol{A} \text{ 的 2 行形成了新的 } \boldsymbol{R}\\ 2\times 2 \text{ 的（行与列）重叠部分求逆}\end{array}\right)$$

这是第一次遇到逆矩阵，因此在这里加以解释：$\boldsymbol{A}\times\boldsymbol{A}^{-1}=$ 单位矩阵$\boldsymbol{I}$。

2×2 矩阵的逆
若 $ad=bc$，则逆不存在
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

这个新思路是直接处理矩阵乘积 $\boldsymbol{CR}$。它引出 $\boldsymbol{C}$ 的列的组合。这对向量来说是本质的一步：取它们的组合。当观察这两列的所有组合时，就在填充一个平面了。如果 $\boldsymbol{A}$ 的第三列不在这个平面上，$\boldsymbol{A}$ 就是可逆的：矩阵的秩是 3。可以阅读《适合每个人的线性代数》的 1.3 节和 1.4 节。

$\boldsymbol{A}=\boldsymbol{CR}$ 之后是线性代数的五大分解——从用 $\boldsymbol{A}=\boldsymbol{LU}$ 求解 n 个方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 开始。这些就是线性代数的组织原则。它们包括方阵的特征值与所有矩阵的奇异值，深入有效地反映出矩阵的信息，其中 $\boldsymbol{C}$ 直接从 $\boldsymbol{A}$ 中取得列。

当 $\boldsymbol{A}$ 中的数字有明确的意义（它们可能都是正的，或可能看到许多零），在 $\boldsymbol{C}$ 中保留这些性质不是一件坏事。

$$\text{推荐：验证} \quad \begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix}^{-1}\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

故 $\boldsymbol{A}$ 的上述两种分解的结果是一致的。

推荐：构造一个秩为 1 的 3×3 矩阵 $\boldsymbol{A}$，将其分解为 $\boldsymbol{A}=\boldsymbol{CR}$。

1.2 矩阵与矩阵相乘：$\boldsymbol{AB}$

内积（行乘列）得到 $\boldsymbol{AB}=\boldsymbol{C}$ 中的每个元素。

由 $\boldsymbol{A}$ 的第 2 行
$\boldsymbol{B}$ 的第 3 列
得到 $\boldsymbol{C}$ 中的 $\boldsymbol{c_{23}}$
$$\begin{bmatrix} \cdot & \cdot & \cdot \\ a_{21} & a_{22} & a_{23} \\ \cdot & \cdot & \cdot \end{bmatrix}\begin{bmatrix} \cdot & \cdot & b_{13} \\ \cdot & \cdot & b_{23} \\ \cdot & \cdot & b_{33} \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & c_{23} \\ \cdot & \cdot & \cdot \end{bmatrix} \tag{1}$$

点积 $c_{23}=(\boldsymbol{A}\text{的第 2 行})\cdot(\boldsymbol{B}\text{的第 3 列})$ 是各个 a 乘以相应 b 的和：

$$c_{23}=a_{21}b_{13}+a_{22}b_{23}+a_{23}b_{33}=\sum_{k=1}^{3}a_{2k}b_{k3}, \quad c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj} \tag{2}$$

这就是通常计算 $\boldsymbol{AB}=\boldsymbol{C}$ 中每个元素的方法。

得到乘积 $\boldsymbol{AB}$ 的另一种方法是将 **$\boldsymbol{A}$ 的列乘以 $\boldsymbol{B}$ 的行**。可以用具体的数说明两个主要的点：一列 $\boldsymbol{u}$ 乘以一行 $\boldsymbol{v}^{\mathrm{T}}$ 就得到一个矩阵。首先集中在 $\boldsymbol{AB}$ 那部分。这个矩阵 $\boldsymbol{uv}^{\mathrm{T}}$ 特别简单：

$$\text{“外积”} \quad \boldsymbol{uv}^{\mathrm{T}} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}\begin{bmatrix} 3 & 4 & 6 \end{bmatrix} = \begin{bmatrix} 6 & 8 & 12 \\ 6 & 8 & 12 \\ 3 & 4 & 6 \end{bmatrix} = \text{“秩为 1 的矩阵”}$$

一个 $m\times1$ 矩阵（列向量 $\boldsymbol{u}$）乘以一个 $1\times p$ 矩阵（行向量 $\boldsymbol{v}^{\mathrm{T}}$）得到一个 $m\times p$ 矩阵。注意这

个秩为 1 的矩阵 $\boldsymbol{uv}^{\mathrm{T}}$ 有何特别之处：

$$\boldsymbol{uv}^{\mathrm{T}}\text{的所有列是}\boldsymbol{u}=\begin{bmatrix}2\\2\\1\end{bmatrix}\text{的倍数，而所有行是}\boldsymbol{v}^{\mathrm{T}}=\begin{bmatrix}3&4&6\end{bmatrix}\text{的倍数。}$$

$\boldsymbol{uv}^{\mathrm{T}}$ 的列空间是一维的：它是沿 $\boldsymbol{u}$ 的方向的直线。列空间的维数（独立列的数目）是**矩阵的秩**（一个关键的数）。**所有非零矩阵 $\boldsymbol{uv}^{\mathrm{T}}$ 的秩都是 1**。它们是每个矩阵的完美构建基块。

也要注意到 $\boldsymbol{uv}^{\mathrm{T}}$ **的行空间是沿着 $\boldsymbol{v}$ 的直线**。根据定义，任何矩阵 $\boldsymbol{A}$ 的行空间是其转置矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 的列空间，记为 $\mathbf{C}(\boldsymbol{A}^{\mathrm{T}})$。这样，只要处理列向量就可以了。在这个例子中，将 $\boldsymbol{uv}^{\mathrm{T}}$ 转置（**行列互换**）得到矩阵 $\boldsymbol{uv}^{\mathrm{T}}$：

$$(\boldsymbol{uv}^{\mathrm{T}})^{\mathrm{T}}=\begin{bmatrix}\mathbf{6}&8&12\\\mathbf{6}&8&12\\\mathbf{3}&4&6\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}\mathbf{6}&\mathbf{6}&\mathbf{3}\\8&8&4\\12&12&6\end{bmatrix}=\begin{bmatrix}3\\4\\6\end{bmatrix}\begin{bmatrix}2&2&1\end{bmatrix}=\boldsymbol{vu}^{\mathrm{T}}$$

可以看到线性代数中的第一个大定理的最清晰的例子：

$$\boxed{\text{行秩}=\text{列秩}，\quad r\text{个独立的列}\Leftrightarrow r\text{个独立的行}}$$

一个非零矩阵 $\boldsymbol{uv}^{\mathrm{T}}$ 有一个独立的列与一个独立的行。所有的列都是 $\boldsymbol{u}$ 的倍数，而所有的行都是 $\boldsymbol{v}^{\mathrm{T}}$ 的倍数。对于这个矩阵，其秩 $r=1$。

AB = 秩为 1 的矩阵之和

用 $\boldsymbol{A}$ 的列乘以 $\boldsymbol{B}$ 的行，得到 $\boldsymbol{AB}$ 的完整乘积。令 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$ 为 $\boldsymbol{A}$ 的 n 列，则 $\boldsymbol{B}$ 必定有 n 行 $\boldsymbol{b}_1^*,\boldsymbol{b}_2^*,\cdots,\boldsymbol{b}_n^*$。这样矩阵 $\boldsymbol{A}$ 就能乘以 $\boldsymbol{B}$ 了。**它们的乘积 $\boldsymbol{AB}$ 是列向量 $\boldsymbol{a}_k$ 乘以行向量 $\boldsymbol{b}_k^*$ 的和**：

矩阵的列行相乘

$$\boxed{\boldsymbol{AB}=\begin{bmatrix}\\ \boldsymbol{a}_1&\cdots&\boldsymbol{a}_n\\ \\ \end{bmatrix}\begin{bmatrix}\boldsymbol{b}_1^*\\\vdots\\\boldsymbol{b}_n^*\end{bmatrix}=\underset{\text{秩为 1 的矩阵之和}}{\boldsymbol{a}_1\boldsymbol{b}_1^*+\boldsymbol{a}_2\boldsymbol{b}_2^*+\cdots+\boldsymbol{a}_n\boldsymbol{b}_n^*}}\tag{3}$$

下面是一个 2×2 的例子，用以显示 $n=2$ 项（列乘以行）以及它们的和 $\boldsymbol{AB}$：

$$\begin{bmatrix}1&0\\3&1\end{bmatrix}\begin{bmatrix}2&4\\0&5\end{bmatrix}=\begin{bmatrix}\mathbf{1}\\\mathbf{3}\end{bmatrix}\begin{bmatrix}\mathbf{2}&\mathbf{4}\end{bmatrix}+\begin{bmatrix}\mathbf{0}\\\mathbf{1}\end{bmatrix}\begin{bmatrix}\mathbf{0}&\mathbf{5}\end{bmatrix}=\begin{bmatrix}2&4\\6&12\end{bmatrix}+\begin{bmatrix}0&0\\0&5\end{bmatrix}=\begin{bmatrix}2&4\\6&17\end{bmatrix}\tag{4}$$

数与数相乘的次数：相乘 4 次得到 2、4、6、12，再相乘 4 次得到 0、0、0、5，总共相乘次数 $2^3=8$。当 $\boldsymbol{A}$、$\boldsymbol{B}$ 是 $n\times n$ 矩阵时，总是有 n^3 次乘法运算。而当 $\boldsymbol{AB}$ 为 $(m\times n)\times(n\times p)$ 矩阵时，则有 mnp 次相乘：n 个秩为 1 的矩阵，每个这样的矩阵都是 $m\times p$ 矩阵。

计数与通常的内积方法是一样的。$\boldsymbol{A}$ 的行乘以 $\boldsymbol{B}$ 的列需要 n 次乘法运算。对 $\boldsymbol{AB}$ 的每个数字计算：当 $\boldsymbol{AB}$ 是 $m \times p$ 矩阵时，需 mp 个点积。当将 $m \times n$ 矩阵乘以 $n \times p$ 矩阵时，相乘次数又是 mnp。

行乘以列	mp 个内积，	每个有 n 次相乘	mnp
列乘以行	n 个外积，	每个有 mp 次相乘	mnp

由上可以看出，它们其实是完全相同的乘法 $a_{ik}\,b_{kj}$，只是顺序不同。下面是 $\boldsymbol{C}=\boldsymbol{AB}$ 中的每个元素 c_{ij} 通过式 (3) 的外积与通过式 (2) 的内积是相同的代数证明：

$$\boldsymbol{a}_k\boldsymbol{b}_k^* \text{ 的 } i,j \text{ 元素是 } a_{ik}b_{kj}，\quad \text{相加得到 } c_{ij}=\sum_{k=1}^{n} a_{ik}\,b_{kj}=\text{第 } \boldsymbol{i} \text{ 行} \cdot \text{第 } \boldsymbol{j} \text{ 列}$$

从列乘以行得到的内在信息

之所以外积的方法在数据科学中如此重要，是因为我们正在寻找矩阵 $\boldsymbol{A}$ 的重要部分。通常不喜欢 $\boldsymbol{A}$ 中的最大数（尽管那也可能是重要的），而是希望得到的是 $\boldsymbol{A}$ 中的主要部分，**而这些部分就是秩为 1 的矩阵$\boldsymbol{uv}^{\mathrm{T}}$**。应用线性代数中的一个主题就是：

$$\text{将 } \boldsymbol{A} \text{ 分解为 } \boldsymbol{CR}，\quad \text{看 } \boldsymbol{A}=\boldsymbol{CR} \text{ 中的 } \boldsymbol{c}_k\boldsymbol{r}_k^* \text{ 部分}$$

将 $\boldsymbol{A}$ 分解为 $\boldsymbol{CR}$ 是矩阵相乘 $\boldsymbol{CR}=\boldsymbol{A}$ 的逆过程。分解更费时，特别是当其中的一些部分涉及特征值或奇异值时。但是，那些数包含矩阵 $\boldsymbol{A}$ 的内在信息，只有在做分解后才能看到。

下面是五个重要的矩阵分解，原始的乘积矩阵（通常用 $\boldsymbol{A}$）及其因子用标准字母表示。本书将解释所有这五个分解。

$$\boxed{\boldsymbol{A}=\boldsymbol{LU}，\quad \boldsymbol{A}=\boldsymbol{QR}，\quad \boldsymbol{S}=\boldsymbol{Q\Lambda Q}^{\mathrm{T}}，\quad \boldsymbol{A}=\boldsymbol{X\Lambda X}^{-1}，\quad \boldsymbol{A}=\boldsymbol{U\Sigma V}^{\mathrm{T}}}$$

这里只简单地列出每个分解的关键词和性质。

(1) $\boldsymbol{A}=\boldsymbol{LU}$ 来自消元法。通过行的组合可由 $\boldsymbol{A}$ 得到 $\boldsymbol{U}$，由 $\boldsymbol{U}$ 回到 $\boldsymbol{A}$。矩阵 $\boldsymbol{L}$ 为下三角矩阵，而 $\boldsymbol{U}$ 是上三角矩阵，参见式(4)。

(2) $\boldsymbol{A}=\boldsymbol{QR}$ 来自将列 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n$**正交化**，就如 “Gram-Schmidt” 方法那样。$\boldsymbol{Q}$ 由正交列组成 ($\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}=\boldsymbol{I}$)，而 $\boldsymbol{R}$ 是上三角矩阵。

(3) $\boldsymbol{S}=\boldsymbol{Q\Lambda Q}^{\mathrm{T}}$ 来自一个对称矩阵 $\boldsymbol{S}=\boldsymbol{S}^{\mathrm{T}}$ 的**特征值**$\lambda_1,\lambda_2,\cdots,\lambda_n$。特征值组成 $\boldsymbol{\Lambda}$ 的对角元。正交的特征向量是矩阵 $\boldsymbol{Q}$ 的列。

(4) $\boldsymbol{A}=\boldsymbol{X\Lambda X}^{-1}$ 是具有 n 个独立的特征向量的 $n \times n$ 矩阵 $\boldsymbol{A}$ 的**对角化**。$\boldsymbol{A}$ 的特征值在 $\boldsymbol{\Lambda}$ 的对角线上。$\boldsymbol{A}$ 的特征向量为 $\boldsymbol{X}$ 的列。

(5) $\boldsymbol{A}=\boldsymbol{U\Sigma V}^{\mathrm{T}}$ 是任意矩阵 $\boldsymbol{A}$（不论是否是方块矩阵）的**奇异值分解**。**奇异值**$\sigma_1,\sigma_2,\cdots,\sigma_r$ 在 $\boldsymbol{\Sigma}$ 中。单位正交的**奇异向量**在 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 中。

采用分解式 (3) 来说明想法。这个特别的 $\boldsymbol{Q\Lambda Q}^{\mathrm{T}}$ 分解由一个对称矩阵 $\boldsymbol{S}$ 开始。该矩阵具有正交的单位特征向量：$\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_n$。这些互相垂直的特征向量（其点积为 0）构成 $\boldsymbol{Q}$ 的列。$\boldsymbol{S}$ 和 $\boldsymbol{Q}$ 是线性代数的 “国王” 和 “皇后”。