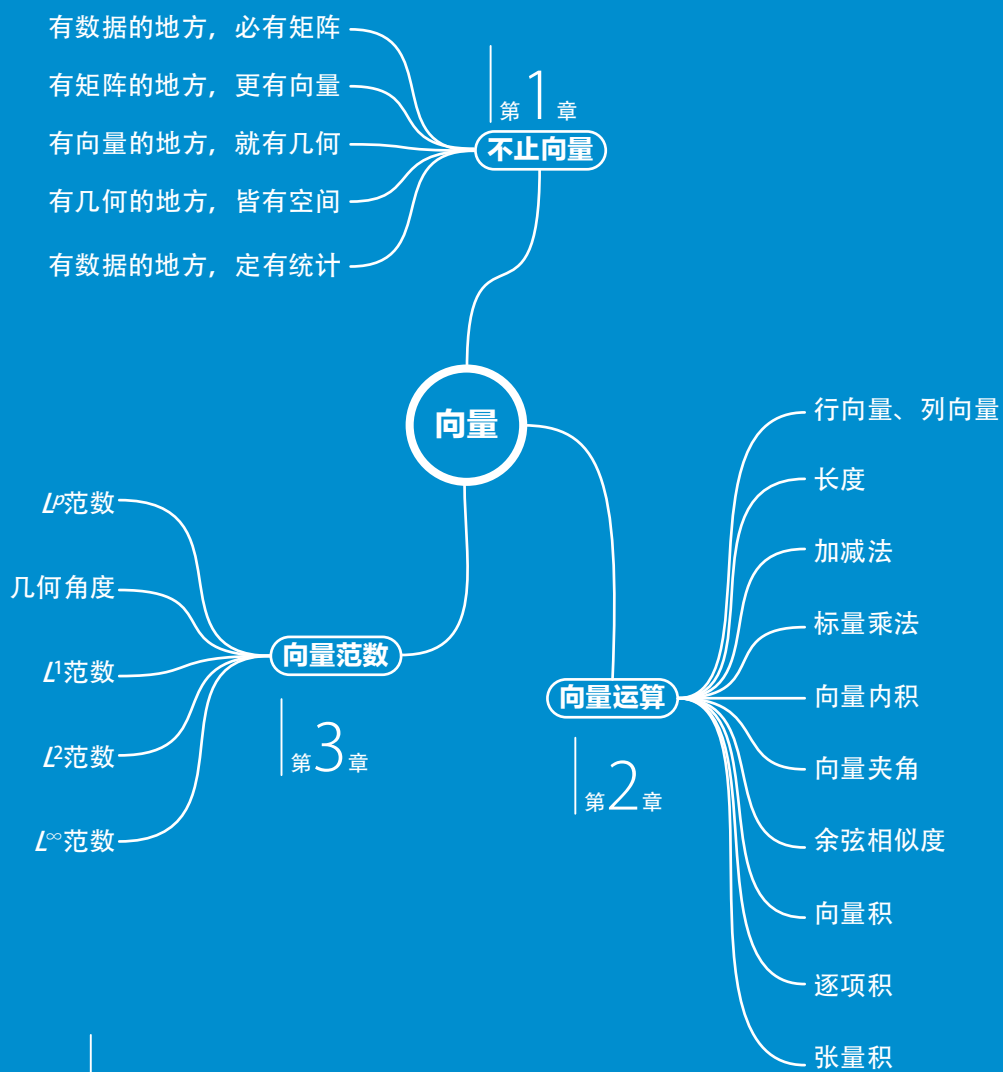


A series of horizontal white lines of varying lengths, creating a sense of motion or a musical staff, positioned behind the large number 01.

01

Section 01

向 量



01

Vector and More

不止向量

一个有关向量的故事，从鸢尾花数据讲起



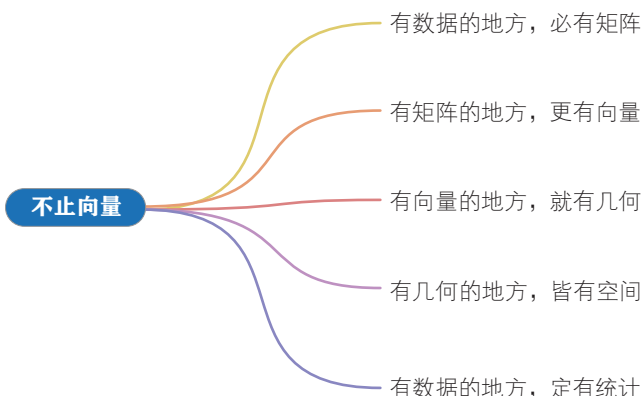
科学的每一次巨大进步，都源于颠覆性的大胆想象。

Every great advance in science has issued from a new audacity of imagination.

—— 约翰·杜威 (John Dewey) | 美国著名哲学家、教育家、心理学家 | 1859 — 1952



- ◀ `sklearn.datasets.load_iris()` 加载鸢尾花数据
- ◀ `seaborn.heatmap()` 绘制热图



1.1 有数据的地方，必有矩阵

本章主角虽然是**向量** (vector)，但是这个有关向量的故事要先从**矩阵** (matrix) 讲起。

简单来说，矩阵是由若干行或若干列元素排列得到的**数组** (array)。矩阵内的元素可以是实数、虚数、符号，甚至是代数式。

从数据角度来看，矩阵就是表格！

鸢尾花数据集

数据科学、机器学习算法和模型都是“数据驱动”。没有数据，任何的算法都无从谈起，数据是各种算法的绝对核心。优质数据本身就极具价值，甚至不需要借助任何模型；反之，则是**垃圾进，垃圾出** (Garbage in, garbage out, GIGO)。

本书使用频率最高的数据是鸢尾花卉数据集。数据集的全称为**安德森鸢尾花卉数据集** (Anderson's Iris data set)，是植物学家**埃德加·安德森** (Edgar Anderson) 在加拿大魁北克加斯帕半岛上采集的鸢尾花样本数据。图1.1所示为鸢尾花数据集部分数据。

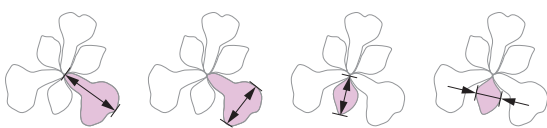
图1.1给出的这些样本都归类于鸢尾属下的三个亚属，分别是**山鸢尾** (setosa)、**变色鸢尾** (versicolor) 和**弗吉尼亚鸢尾** (virginica)。每一类鸢尾花收集了50条样本记录，共计150条。



注意：本书用大写、粗体、斜体字母代表矩阵，如 $\mathbf{X}$ 、 $\mathbf{A}$ 、 Σ 、 $\mathbf{A}$ 。特别地，本书用 $\mathbf{X}$ 代表样本数据矩阵，用 Σ 代表方差协方差矩阵 (variance covariance matrix)。本书用小写、粗体、斜体字母代表向量，如 $\mathbf{x}$ 、 $\mathbf{x}_1$ 、 $\mathbf{x}^{(1)}$ 、 $\mathbf{v}$ 。

鸢尾花的四个特征被用作样本的定量分析，它们分别是**花萼长度** (sepal length)、**花萼宽度** (sepal width)、**花瓣长度** (petal length) 和**花瓣宽度** (petal width)。

如图1.2所示，本书常用**热图** (heatmap) 可视化矩阵。不考虑鸢尾花分类标签，鸢尾花数据矩阵 $\mathbf{X}$ 有150行、4列，因此 $\mathbf{X}$ 也常记做 $\mathbf{X}_{150 \times 4}$ 。



Index	Sepal length X_1	Sepal width X_2	Petal length X_3	Petal width X_4	Species C
1	5.1	3.5	1.4	0.2	Setosa C_1
2	4.9	3	1.4	0.2	
3	4.7	3.2	1.3	0.2	
...	...	...	...	...	
49	5.3	3.7	1.5	0.2	
50	5	3.3	1.4	0.2	Versicolor C_2
51	7	3.2	4.7	1.4	
52	6.4	3.2	4.5	1.5	
53	6.9	3.1	4.9	1.5	
...	...	...	...	...	
99	5.1	2.5	3	1.1	Virginica C_3
100	5.7	2.8	4.1	1.3	
101	6.3	3.3	6	2.5	
102	5.8	2.7	5.1	1.9	
103	7.1	3	5.9	2.1	
...	...	...	...	...	
149	6.2	3.4	5.4	2.3	
150	5.9	3	5.1	1.8	



图1.1 鸢尾花数据，数值数据 (单位: cm)

行向量、列向量

前文提到，矩阵可以视作由一系列行向量、列向量构造而成。

反方向来看，矩阵切丝、切片可以得到行向量、列向量。如图1.2所示， $\mathbf{X}$ 任一行向量 ($\mathbf{x}^{(1)}$ 、 $\mathbf{x}^{(2)}$ 、 $\dots$ 、 $\mathbf{x}^{(150)}$) 代表一朵鸢尾花样本花萼长度、花萼宽度、花瓣长度和花瓣宽度测量结果；而 $\mathbf{X}$ 某一列向量 ($\mathbf{x}_1$ 、 $\mathbf{x}_2$ 、 $\mathbf{x}_3$ 、 $\mathbf{x}_4$) 为鸢尾花某个特征的样本数据。

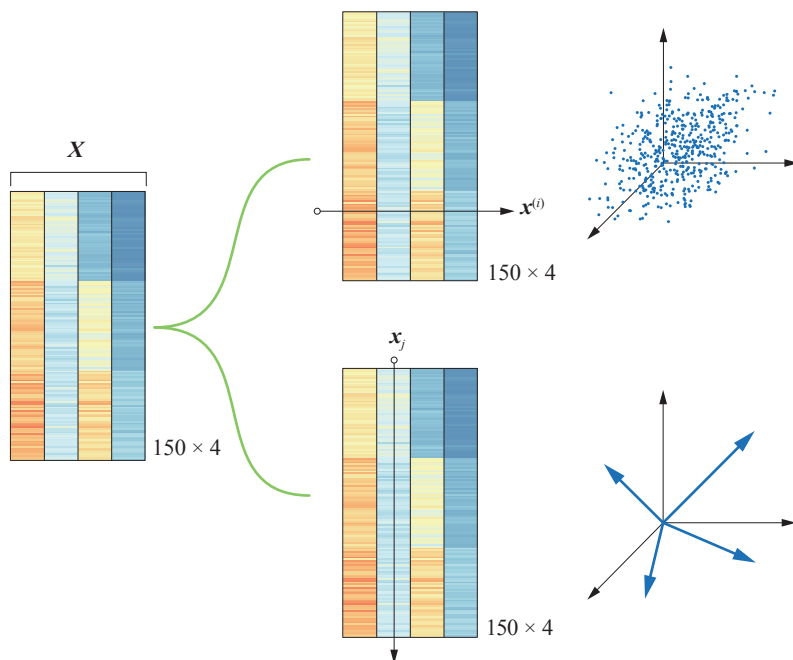


图1.2 矩阵可以分割成一系列行向量或列向量

图片

数据矩阵其实无处不在。

再举个例子，大家日常随手拍摄的照片实际上就是数据矩阵。图1.3所示为作者拍摄的一张鸢尾花照片。把这张照片做黑白处理后，它变成了形状为 2990×2714 的矩阵，即2990行、2714列。



鸢尾花书《数据有道》将采用主成分分析 (Principal Component Analysis, PCA) 继续深入分析图1.3这幅鸢尾花黑白照片。

图1.3所示照片显然不是矢量图。不断放大后，我们会发现照片的局部变得越来越模糊。继续放大，我们发现这张照片竟然是由一系列灰度热图构成的。再进一步，提取其中图片的4个像素点，也就是矩阵的4个元素，我们可以得到一个 2×2 实数矩阵。

对于大部分机器学习应用，如识别人脸、判断障碍物等，并不需要输入彩色照片，黑白照片的数据矩阵含有的信息就足够用。

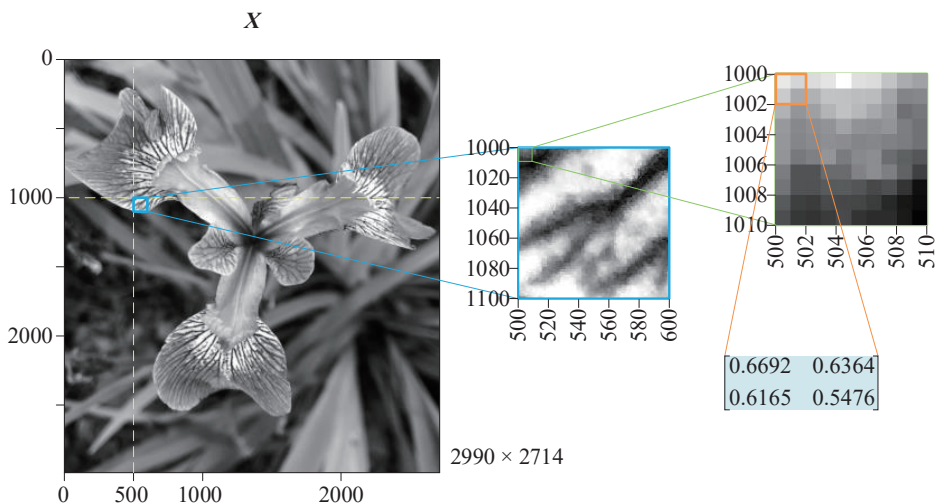


图1.3 照片也是数据矩阵

1.2 有矩阵的地方，更有向量

行向量

首先，矩阵 $\mathbf{X}$ 可以看作是由一系列行向量 (row vector) 上下叠加而成的。

如图1.4所示，矩阵 $\mathbf{X}$ 的第 i 行可以写成行向量 $\mathbf{x}^{(i)}$ 。上标圆括号中的 i 代表序号，对于鸢尾花数据集， $i = 1 \sim 150$ 。

举个例子， $\mathbf{X}$ 的第1行行向量记做 $\mathbf{x}^{(1)}$ ，具体为

$$\mathbf{x}^{(1)} = [5.1 \quad 3.5 \quad 1.4 \quad 0.2]_{1 \times 4} \quad (1.1)$$

行向量 $\mathbf{x}^{(1)}$ 代表鸢尾花数据集编号为1的样本。行向量 $\mathbf{x}^{(1)}$ 的四个元素依次代表花萼长度 (sepal length)、花萼宽度 (sepal width)、花瓣长度 (petal length) 和花瓣宽度 (petal width)。长、宽的单位均为厘米 (cm)。

行向量 $\mathbf{x}^{(i)}$ 也可以看做1行、4列的矩阵，即形状为 1×4 。

虽然Python是**基于0编号** (zero-based indexing)，但是本书对矩阵行、列编号时，还是延续线性代数传统，采用**基于1编号** (one-based indexing)。鸢尾花书《编程不难》专门介绍过两种不同的编号方式。

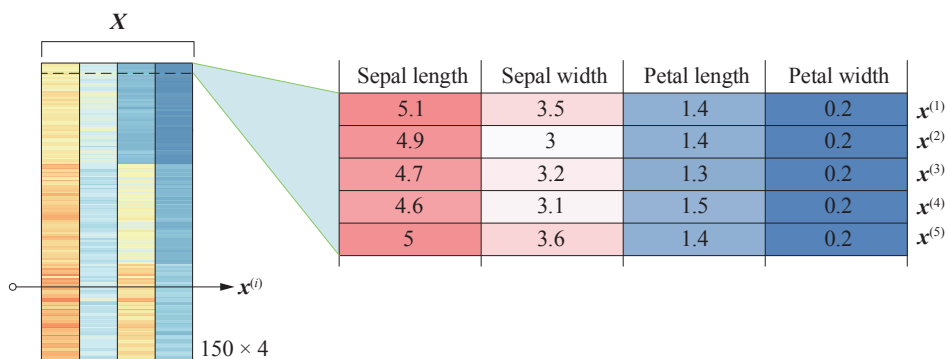


图1.4 鸢尾花数据，行向量代表样本数据点

列向量

矩阵 $\mathbf{X}$ 也可以看做是由一系列**列向量** (column vector) 左右排列而成的。

如图1.2所示，矩阵 $\mathbf{X}$ 的第 j 列可以写成列向量 $\mathbf{x}_j$ ，下标 j 代表列序号。对于鸢尾花数据集，若不考虑分类标签，则 $j = 1 \sim 4$ 。

比如， $\mathbf{X}$ 的第1列向量记做 $\mathbf{x}_1$ ，具体为

$$\mathbf{x}_1 = \begin{bmatrix} 5.1 \\ 4.9 \\ \vdots \\ 5.9 \end{bmatrix}_{150 \times 1} \quad (1.2)$$

列向量 $\mathbf{x}_1$ 代表鸢尾花150个样本数据花萼长度数值。列向量 $\mathbf{x}_1$ 可以看做150行、1列的矩阵，即形状为 150×1 。整个数据矩阵 $\mathbf{X}$ 可以写成四个列向量，即 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4]$ 。

再次强调：为了区分数据矩阵中的行向量和列向量，在编号时，本书中行向量采用上标加圆括号，如 $\mathbf{x}^{(i)}$ ；而列向量编号采用下标，如 $\mathbf{x}_i$ 。

此外，大家熟悉的**三原色光模式** (RGB color mode) 中的每种颜色实际上也可以写成列向量，如图1.5所示的7种颜色。在本书第7章中，我们将用RGB解释向量空间等概念。

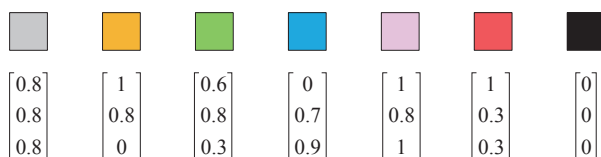


图 1.5 7种颜色对应的RGB颜色向量

大家可能会问，元素数量均为150的 $\mathbf{x}_1$ 、 $\mathbf{x}_2$ 、 $\mathbf{x}_3$ 、 $\mathbf{x}_4$ 这四个向量到底意味着什么？有没有什么办法可视化这四个列向量？怎么量化它们之间的关系？答案会在本书第12章揭晓。

当然，我们不要被向量、矩阵这些名词吓到。矩阵就是一个表格，而这个表格可以划分成若干行、若干列，它们分别叫行向量、列向量。

1.3 有向量的地方，就有几何

数据云、投影

取出鸢尾花前两个特征，即花萼长度和花萼宽度所对应的数据，把它们以坐标的形式画在平面直角坐标系（记做 $\mathbb{R}^2$ ）中，我们便得到平面散点图。如图1.6所示。这幅散点图好比样本“数据云”。

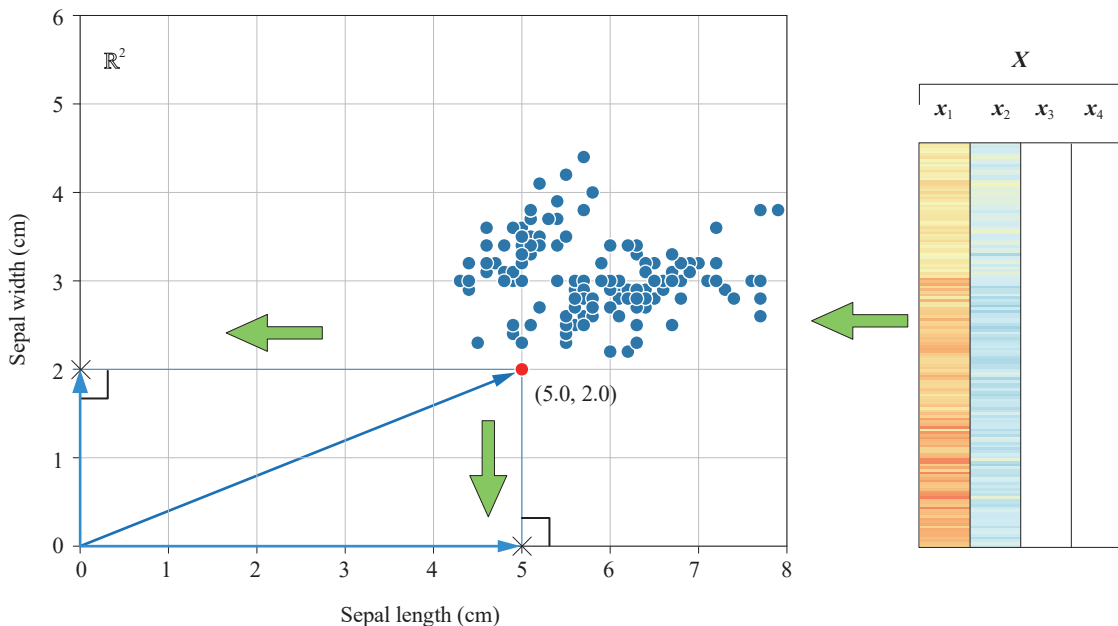


图1.6 鸢尾花前两个特征数据散点图

图1.6中数据点 $(5.0, 2.0)$ 可以写成行向量 $[5.0, 2.0]$ 。 $(5.0, 2.0)$ 是序号为61的样本点，对应的行向量可以写成 $\mathbf{x}^{(61)}$ 。

从几何视角来看， $[5.0, 2.0]$ 在横轴的**正交投影** (orthogonal projection) 结果为5.0，代表该点的横坐标为5.0。 $[5.0, 2.0]$ 在纵轴的正交投影结果为2.0，代表其纵坐标为2.0。

正交 (orthogonality) 是线性代数的概念，是垂直的推广。正交投影很好理解，即原数据点和投影点连线垂直于投影点所在直线或平面。打个比方，头顶正上方阳光将物体的影子投影在地面，而阳光光线垂直于地面。如无特别强调，本书的投影均指正交投影。

从集合视角来看， $(5.0, 2.0)$ 属于平面 $\mathbb{R}^2$ ，即 $(5.0, 2.0) \in \mathbb{R}^2$ 。图1.6中整团数据云都属于 $\mathbb{R}^2$ 。再者，如图1.6所示，从向量角度来看，行向量 $[5.0, 2.0]$ 在横轴上投影的向量为 $[5.0, 0]$ ，在纵轴上投影的向量为 $[0, 2.0]$ 。而 $[5.0, 0]$ 和 $[0, 2.0]$ 两个向量合成就是 $[5.0, 2.0] = [5.0, 0] + [0, 2.0]$ 。

再进一步，将图1.6整团数据云全部正交投影到横轴，得到图1.7。图1.7中 $\times$ 代表的数据实际上就是鸢尾花数据集第一列的花萼长度数据。图1.7中的横轴相当于一个一维空间，即数轴 $\mathbb{R}$ 。

我们也可以把整团数据云全部投影在纵轴，得到图1.8。图中的 $\times$ 是鸢尾花数据第二列的花萼宽度数据。

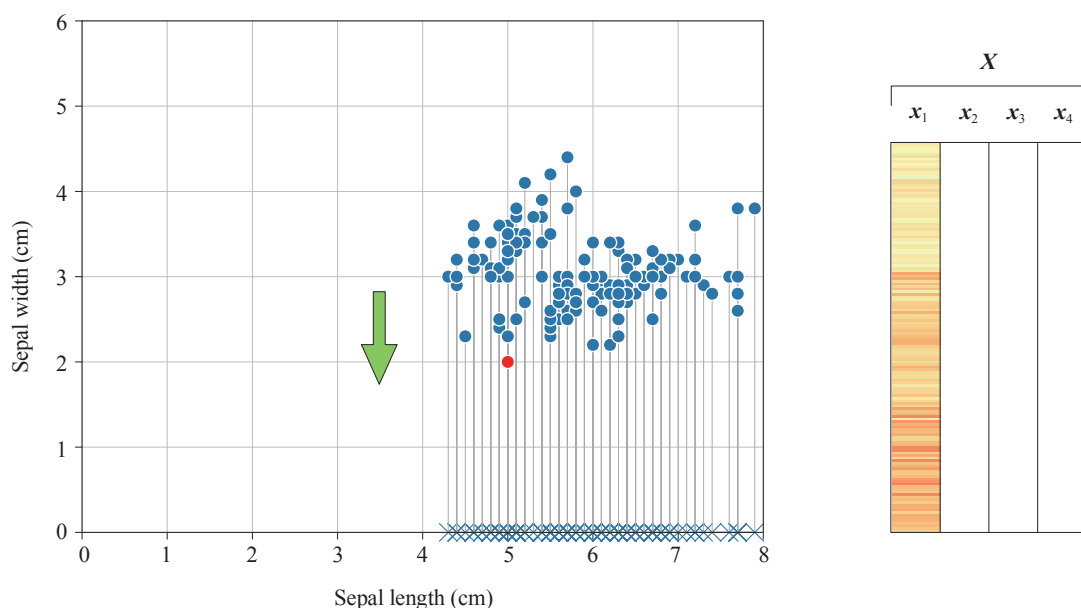


图1.7 二维散点正交投影到横轴

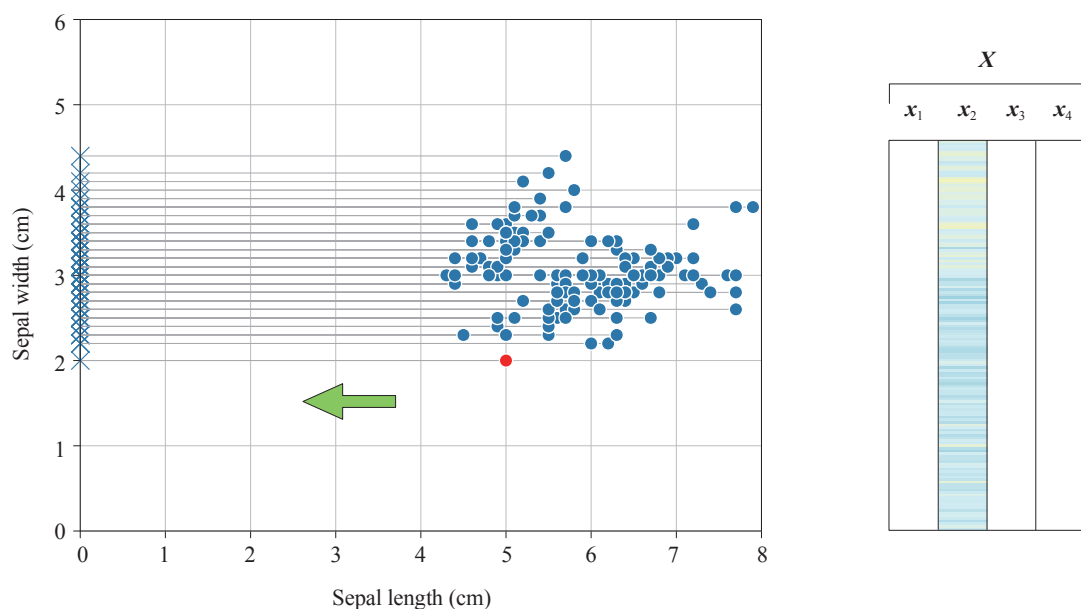


图1.8 二维散点正交投影到纵轴

投影到一条过原点的斜线

你可能会问，是否可以将图1.7中所有点投影在一条斜线上？

答案是肯定的。

如图1.9所示，鸢尾花数据投影到一条斜线上，这条斜线通过原点，与横轴夹角为 15° 。观察图1.9，我们已经发现投影点似乎是 $\mathbf{x}_1$ 与 $\mathbf{x}_2$ 的某种组合。也就是说， $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 分别贡献 $v_1\mathbf{x}_1$ 和 $v_2\mathbf{x}_2$ ，两种成分的合成 $v_1\mathbf{x}_1 + v_2\mathbf{x}_2$ 就是投影点坐标。 $v_1\mathbf{x}_1 + v_2\mathbf{x}_2$ 也叫**线性组合** (linear combination)。

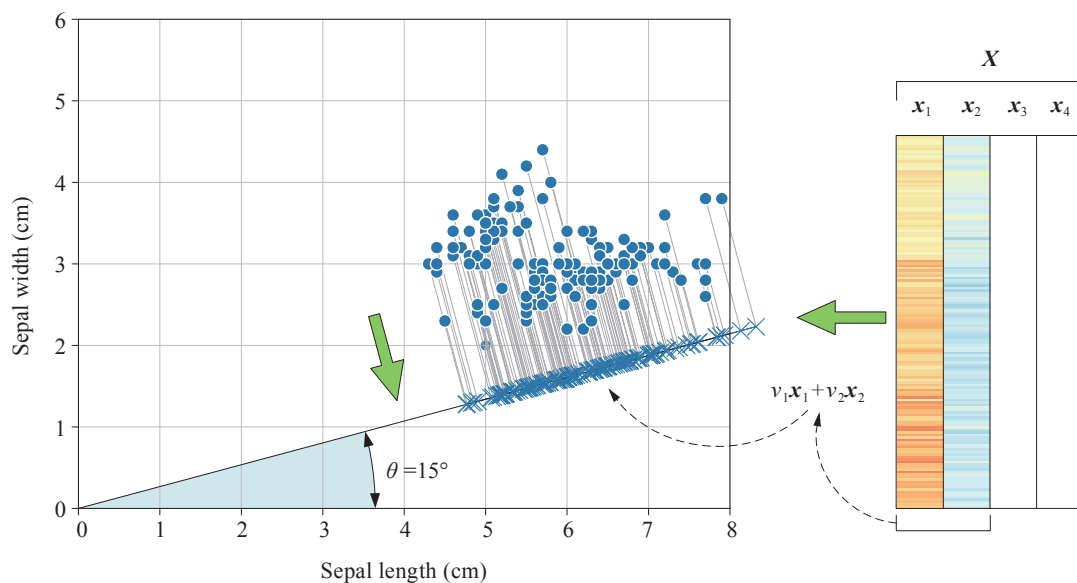


图1.9 二维散点正交投影到一条斜线



大家可能会问，怎么计算图1.9中的投影点坐标呢？这种几何变换有何用途？这是本书第9、10章要探究的问题。

三维散点图、成对特征散点图

取出鸢尾花前三个特征（花萼长度、花萼宽度、花瓣长度）对应的数据，并在三维空间 $\mathbb{R}^3$ 绘制散点图，得到图1.10所示的散点图。而图1.6相当于图1.10在水平面（浅蓝色背景）的正交投影结果。

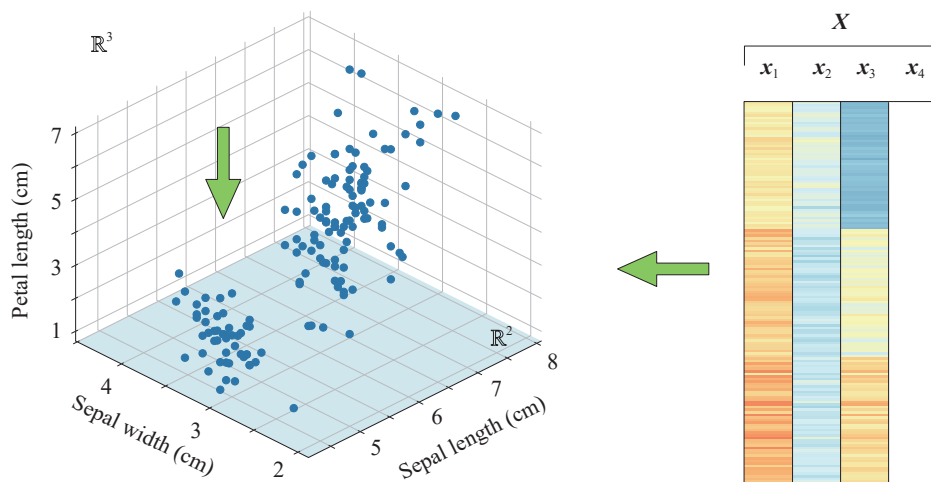


图1.10 鸢尾花前三个特征数据散点图

回顾鸢尾花书《数学要素》一册介绍过的成对特征散点图，具体如图1.11所示。成对特征散点图不仅可视化鸢尾花的四个特征（花萼长度、花萼宽度、花瓣长度和花瓣宽度），而且通过散点颜色还可以展示鸢尾花的三个类别（山鸢尾、变色鸢尾、弗吉尼亚鸢尾）。图1.11中的每一幅散点图相当于四维空间数据在不同平面上的投影结果。

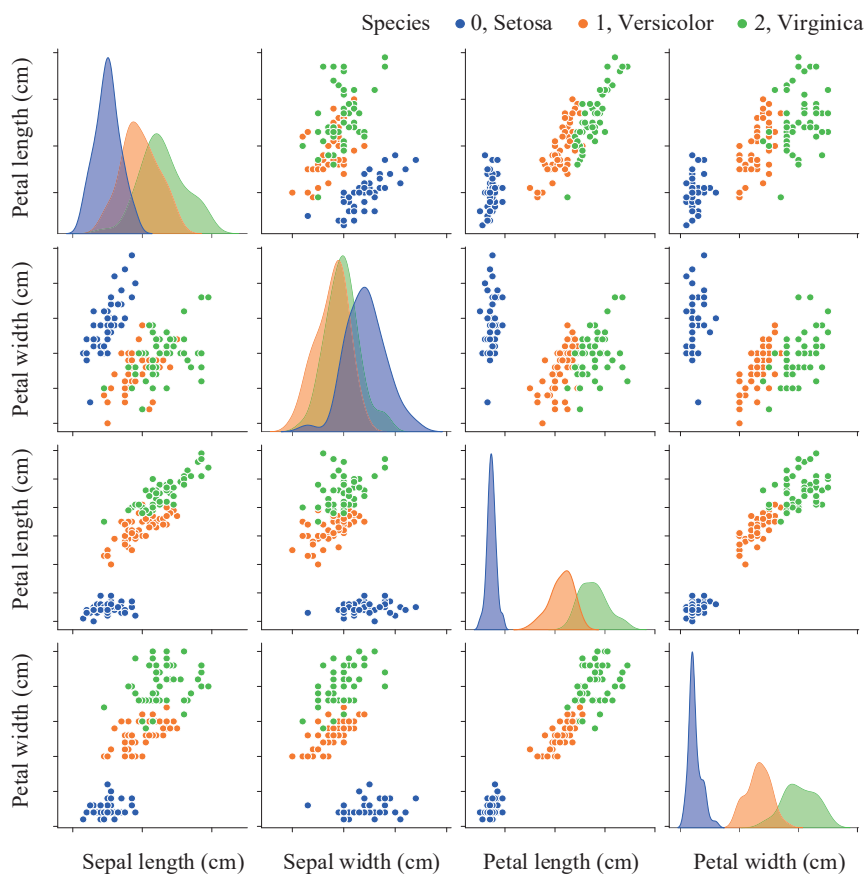


图1.11 鸢尾花数据成对特征散点图 (考虑分类标签, 图片来自鸢尾花书《数学要素》一册)

统计视角：移动向量起点

如图1.12所示, 本节前文行向量的起点都是原点, 即零向量 $\mathbf{0}$ 。而平面 $\mathbb{R}^2$ 这个二维空间则“装下”了这150个行向量。

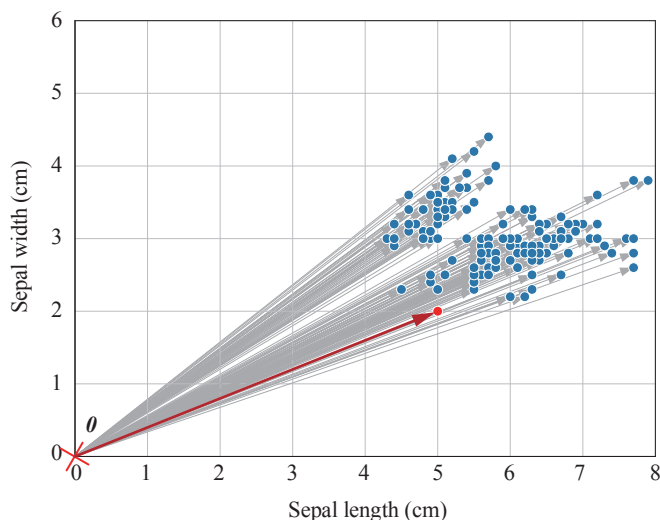


图1.12 向量起点为原点

但是，统计视角下，向量的起点移动到了数据**质心** (centroid)。所谓数据质心就是数据每一特征均值构成的向量。

这一点也不难理解，大家回想一下，我们在计算方差、均方差、协方差、相关性系数等统计度量时，都会去均值。从向量角度来看，这相当于移动了向量起点。

如图1.13所示，将向量的起点移动到质心后，向量的长度、绝对角度 (如与坐标系横轴夹角)、相对角度 (向量两两之间的夹角) 都发生了显著变化。

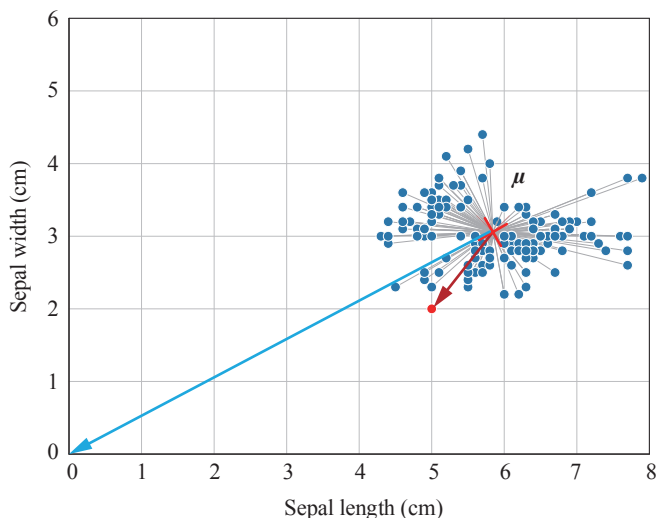


图1.13 向量起点为质心

将图1.13整团数据云质心平移到原点，这个过程就是去均值过程，结果如图1.14所示。数据矩阵 $\mathbf{X}$ 去均值化得到的数据矩阵记做 $\mathbf{X}_c$ ，显然 $\mathbf{X}_c$ 的质心位于原点 (0, 0)。去均值并不影响数据的单位，图1.14横轴、纵轴的单位都是厘米。

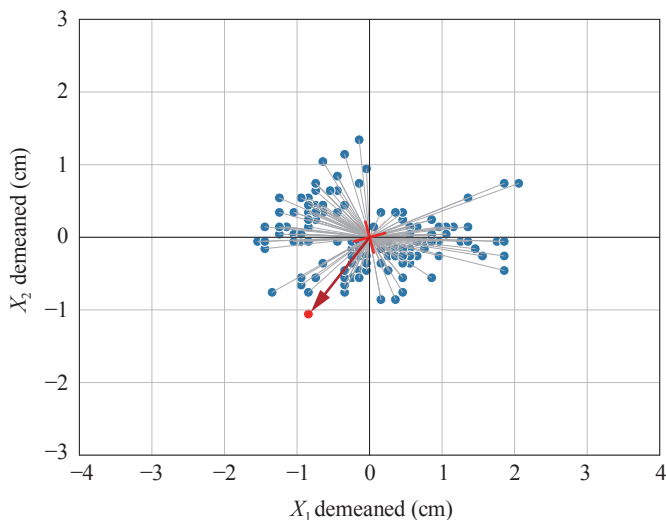


图1.14 数据去均值化



观察图1.11，我们发现，如果考虑数据标签的话，每一类标签样本数据都有自己质心，叫做分类质心，这是本书第22章要讨论的话题。此外，本书最后三章的“数据三部曲”会把数据、矩阵、向量、矩阵分解、空间、优化、统计等板块联结起来。

1.4 有几何的地方，皆有空间

从线性方程组说起

从代数视角来看，**矩阵乘法** (matrix multiplication) 代表**线性映射** (linear mapping)。比如，在 $A_{m \times n} \mathbf{x}_{n \times 1} = \mathbf{b}_{m \times 1}$ 中，矩阵 $A_{m \times n}$ 扮演的角色就是完成 $\mathbf{x} \rightarrow \mathbf{b}$ 的线性映射。列向量 $\mathbf{x}_{n \times 1}$ 在 $\mathbb{R}^n$ 中，列向量 $\mathbf{b}_{m \times 1}$ 在 $\mathbb{R}^m$ 中。

$A_{m \times n} \mathbf{x}_{n \times 1} = \mathbf{b}_{m \times 1}$ 也叫做**线性方程组** (system of linear equations)。在鸢尾花书《数学要素》“鸡兔同笼三部曲”中，我们用线性方程组解决过鸡兔同笼问题。下面我们简单回顾一下。

《孙子算经》这样引出鸡兔同笼问题：“今有雉兔同笼，上有三十五头，下有九十四足，问雉兔各几何？”

将这个问题写成线性方程组为

$$\begin{cases} 1 \cdot x_1 + 1 \cdot x_2 = 35 \\ 2 \cdot x_1 + 4 \cdot x_2 = 94 \end{cases} \Rightarrow \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_x = \underbrace{\begin{bmatrix} 35 \\ 94 \end{bmatrix}}_b \quad (1.3)$$

即

$$A\mathbf{x} = \mathbf{b} \quad (1.4)$$

未知变量构成的列向量 $\mathbf{x}$ 可以利用下式求解，即

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 35 \\ 94 \end{bmatrix} = \begin{bmatrix} 2 & -0.5 \\ -1 & 0.5 \end{bmatrix} \begin{bmatrix} 35 \\ 94 \end{bmatrix} = \begin{bmatrix} 23 \\ 12 \end{bmatrix} \quad (1.5)$$

式中：逆矩阵 A^{-1} 完成 $\mathbf{b} \rightarrow \mathbf{x}$ 的线性映射。

这里用到了矩阵乘法 (matrix multiplication)、矩阵逆 (matrix inverse) 相关知识。本书第4、5、6三章将介绍矩阵相关运算，居于核心的运算当然是矩阵乘法。

几何视角

从几何视角来看，式 (1.3) 中矩阵 A 完成的是**线性变换** (linear transformation)。如图1.15所示，矩阵 A 把方方正正的方格变成平行四边形网格，对应的计算为

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}}_A \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\mathbf{e}_1} = \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{\mathbf{a}_1}, \quad \underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix}}_A \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\mathbf{e}_2} = \underbrace{\begin{bmatrix} 1 \\ 4 \end{bmatrix}}_{\mathbf{a}_2} \quad (1.6)$$

而式 (1.6) 的结果恰好是矩阵 $A = [\mathbf{a}_1, \mathbf{a}_2]$ 的两个列向量 $\mathbf{a}_1$ 和 $\mathbf{a}_2$ 。

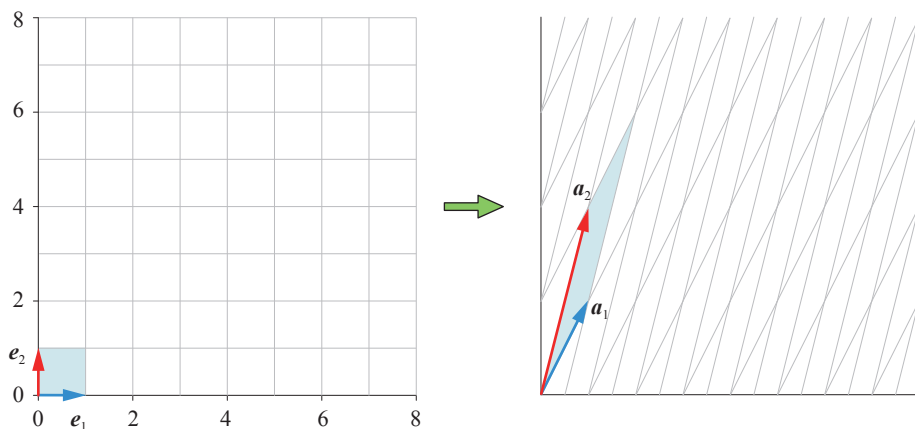


图1.15 矩阵 A 完成的线性变换

观察图1.15中的左图，整个直角坐标系整个方方正正的网格由 $[e_1, e_2]$ 张成，就好比 e_1 和 e_2 是撑起这个二维空间的“骨架”。再看图1.15中的右图， $[a_1, a_2]$ 同样张成了整个直角坐标系，不同的是网格为平行四边形。 $[e_1, e_2]$ 和 $[a_1, a_2]$ 都叫做空间 $\mathbb{R}^2$ 的**基底** (base)。

将 A 写成 $[a_1, a_2]$ ，展开得到

$$\begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 a_1 + x_2 a_2 = b \quad (1.7)$$



本书将在第7章专门讲解基底、线性组合等向量空间概念。

式 (1.7) 代表基底 $[a_1, a_2]$ 中两个基底向量的线性组合。

从正圆到旋转椭圆

圆锥曲线，特别是椭圆，在鸢尾花书中扮演重要角色，这一切都源于多元高斯分布概率密度函数，而线性变换和椭圆又有着千丝万缕的联系。

如图1.16所示，同样利用矩阵 A ，我们可以把一个单位圆转化为旋转椭圆。图1.16中，任意向量 x 起点为原点，终点落在单位圆上，经过 A 的线性变换变成 $y = Ax$ 。

图1.16旋转椭圆的半长轴长度约为4.67，半短轴长度约为0.43，半短轴和横轴夹角约为 -16.85° 。要获得这些椭圆信息，我们需要一个线性代数利器——**特征值分解** (eigen decomposition)。

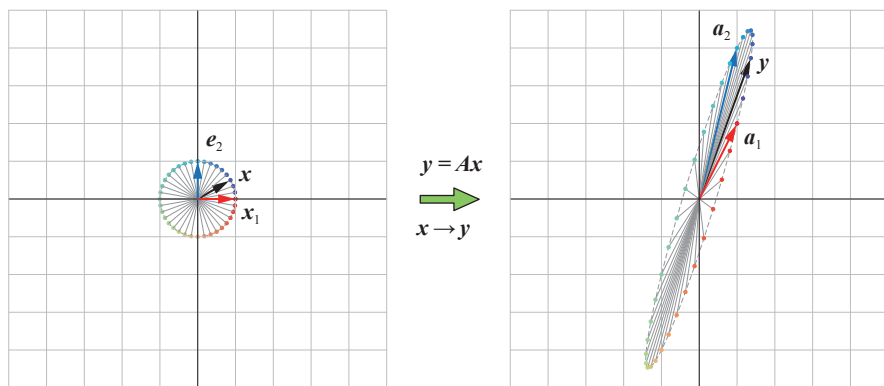


图1.16 矩阵 A 将单位圆转化为旋转椭圆

特征值分解

相信读者对特征值分解并不陌生。如图1.17所示，我们在鸢尾花书《数学要素》鸡兔同笼三部曲的“鸡兔互变”中简单介绍过特征值分解，大家如果忘记了，建议回顾一下。

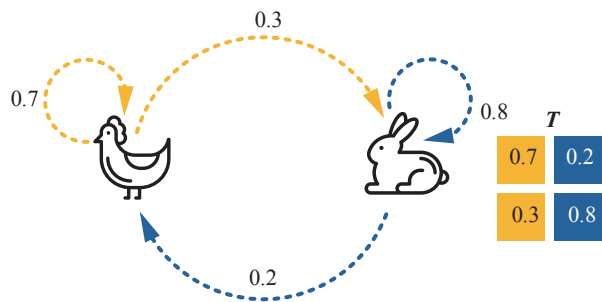


图1.17 鸡兔同笼三部曲中“鸡兔互变”(图片来自鸢尾花书《数学要素》第25章)

剧透一下，鸢尾花数据矩阵 X 本身并不能完成特征值分解。但是图1.18中的格拉姆矩阵 $G = X^T X$ 可以完成特征值分解，分解过程如图1.18所示。请大家特别注意图1.18中的矩阵 V 。正如图1.15右图中的 $A = [a_1, a_2]$ 张成了一个平面，矩阵 $V = [v_1, v_2, v_3, v_4]$ 则张成了一个四维空间 $\mathbb{R}^4$ ！

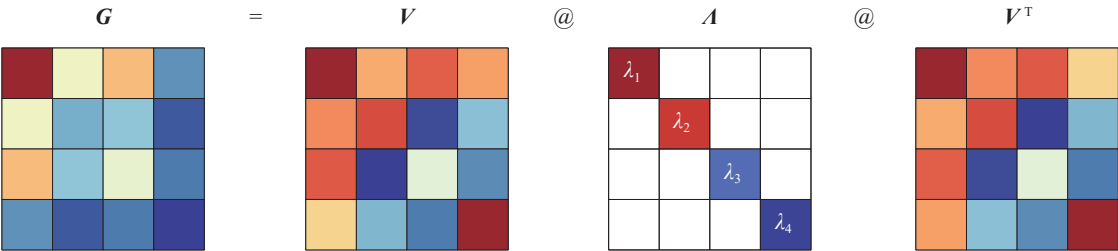


图1.18 矩阵 X 的格拉姆矩阵的特征值分解

本书第13、14章专门探讨特征值分解。此外，本书将在第20、21章利用线性代数工具分析圆锥曲线和二次曲面。

奇异值分解

在**矩阵分解** (matrix decomposition) 这个工具库中，最全能的工具叫**奇异值分解** (Singular Value Decomposition, SVD)。因为不管形状如何，任何实数矩阵都可以完成奇异值分解。

图1.19所示为对鸢尾花数据矩阵的SVD分解，其中的 U 和 V 都各自张成不同的空间。

本书第15、16章专门讲解奇异值分解，第23章则利用SVD分解引出四个空间。

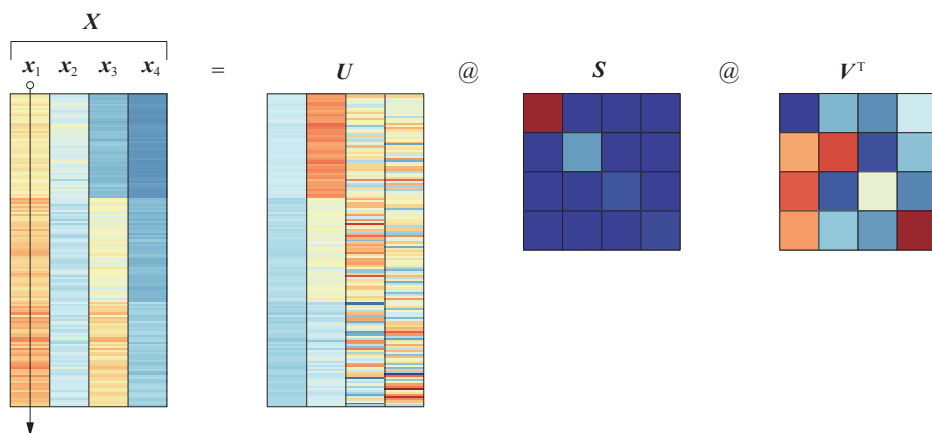


图1.19 对矩阵 X 进行SVD分解

1.5 有数据的地方，定有统计

前文提到，图1.20所示鸢尾花数据每一列代表鸢尾花的一个特征，如花萼长度（第1列，列向量 $\mathbf{x}_1$ ）、花萼宽度（第2列，列向量 $\mathbf{x}_2$ ）、花瓣长度（第3列，列向量 $\mathbf{x}_3$ ）和花瓣宽度（第4列，列向量 $\mathbf{x}_4$ ）。这些列向量可以看成是 X_1 、 X_2 、 X_3 、 X_4 四个随机变量的样本值集合。

从统计视角来看，我们可以计算样本数据各个特征的均值（ μ_j ）和不同特征上样本数据的均方差（ σ_j ）。图1.20中四幅子图中的曲线代表各个特征样本数据的**概率密度估计**（probability density estimation）曲线。有必要的话，我们还可以在图中标出 μ_j 、 $\mu_j \pm \sigma_j$ 、 $\mu_j \pm 2\sigma_j$ 对应的位置。

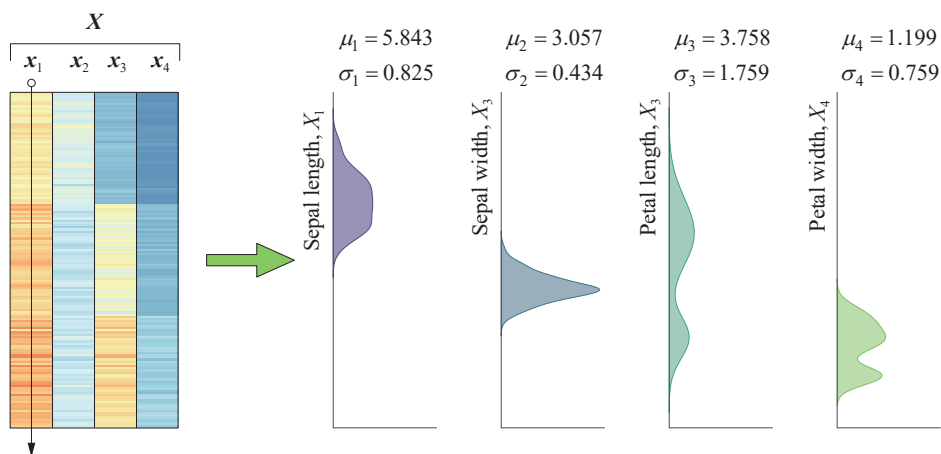


图1.20 鸢尾花数据每个特征的基本统计描述

实际应用时，我们还会对原始数据进行处理，常见的操作有**去均值**（demean）、**标准化**（standardization）等。

多个特征之间的关系，我们可以采用**格拉姆矩阵**（Gram matrix）、**协方差矩阵**（covariance matrix）、**相关性系数矩阵**（correlation matrix）等矩阵来描述。

图1.21所示为本书后续要用到的鸢尾花数据矩阵 $\mathbf{X}$ 衍生得到的几种矩阵。注意，图1.2和图1.21中矩阵 $\mathbf{X}$ 的热图采用不同的色谱值。

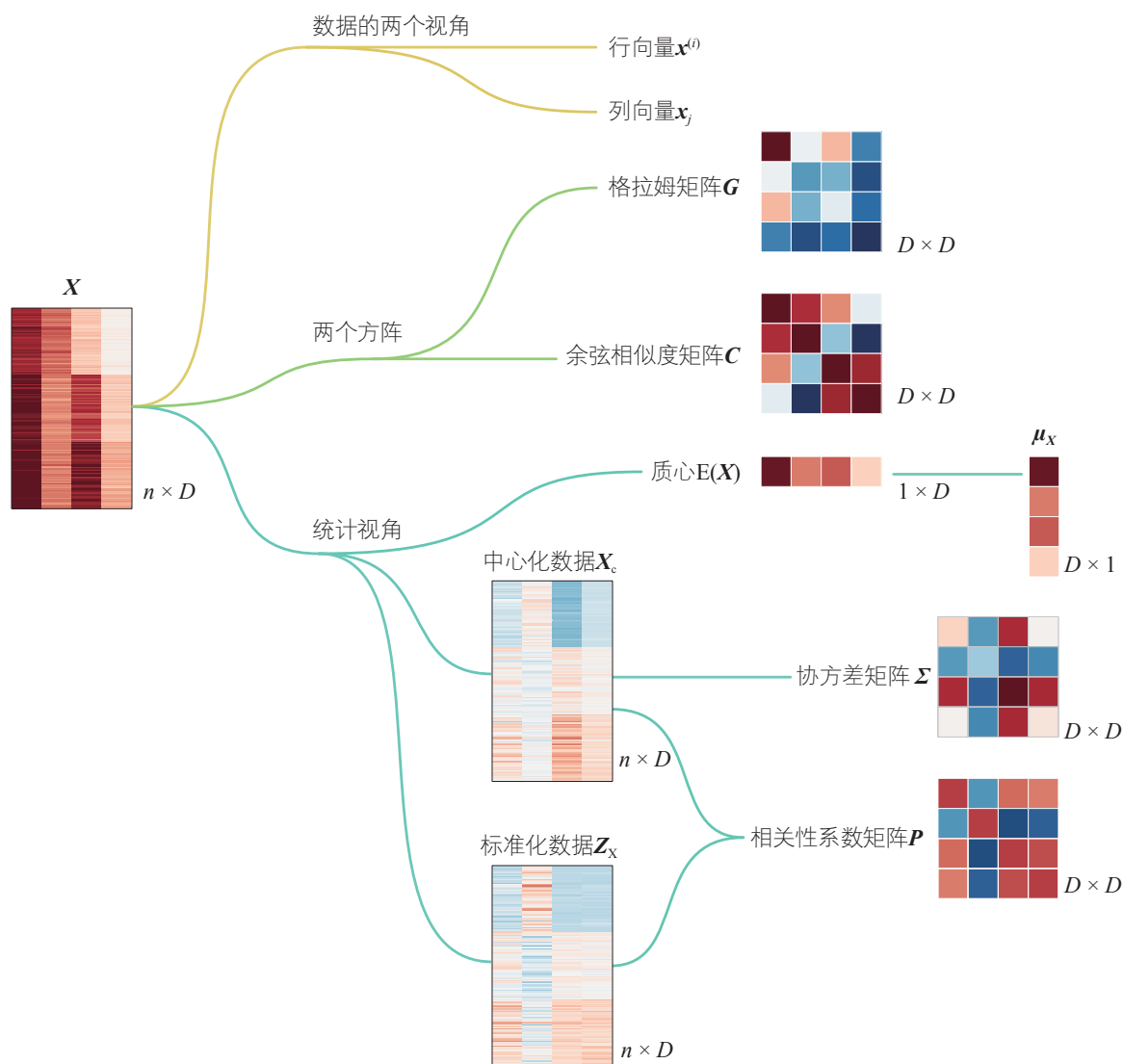


图1.21 鸢尾花数据衍生得到的几个矩阵 (图片来自本书第24章)

本书第22章将介绍如何获得图1.21所示的矩阵，本书第24章将探讨图1.21中主要矩阵和各种矩阵分解 (matrix decomposition) 之间的有趣关系。

本章只配套一个代码文件 `Streamlit_Bk4_Ch1_01.py`。这段代码中，我们用 `Streamlit` 和 `Plotly` 分别绘制了鸢尾花数据集的热图、平面散点图、三维散点图、成对特征散点图。这四幅图都是可交互图像。



本章以向量为主线，回顾了《数学要素》“鸡兔同笼三部曲”的主要内容，预告了本书的核心内容。目前不需要大家理解本章提到的所有术语，只希望大家记住以下几句话：

有数据的地方，必有矩阵！

有矩阵的地方，更有向量！

有向量的地方，就有几何！

有几何的地方，皆有空间！

有数据的地方，定有统计！



对线性代数概念感到困惑的读者，推荐大家看看3Blue1Brown制作的视频。很多视频网站上都可以找到译制视频。如下为3Blue1Brown线性代数部分网页入口：

◀ <https://www.3blue1brown.com/topics/linear-algebra>