

作者	生姜 DrGinger
脚本	生姜 DrGinger
视频	崔崔 CuiCui
开源学习资源	https://github.com/Visualize-ML
平台	https://www.youtube.com/@DrGinger_Jiang https://space.bilibili.com/3546865719052873 https://space.bilibili.com/513194466

11.2 谱分解



本节你将掌握的核心技能：

- ▶ 实对称矩阵：满足 $A = A^T$ ，且元素均为实数。
- ▶ 实对称矩阵谱分解结果可表示为“旋转 → 缩放 → 旋转”。
- ▶ 谱分解找到一个新坐标系将椭圆摆正。
- ▶ 单位圆 → 旋转椭圆，反映出特征值的几何含义。
- ▶ 只有所有特征值非零时方阵才可逆。

上一节提到，实数对称矩阵的特征值分解叫做谱分解。简单来说，实对称矩阵是个对称矩阵，并且所有元素均为实数。本书前文提过，如果矩阵 A 为对称矩阵，则满足 $A = A^T$ 。几何角度来看，对称矩阵沿主对角线镜像对称。

⚠ 注意，不特殊强调的话，本书默认矩阵的元素都是实数。

让我们先回顾实数这个概念。

实数、虚数、复数

大家应该对实数这个概念不陌生。实数 (real number) 包括有理数 (可以表示为两个整数之比，比如无限循环小数 $1/3$)、无理数 (不能表示为两个整数之比，比如无限不循环小数 π)。

如图 1 所示，每一个实数都对应数轴上的一个点；反之，数轴上的每一个点也都对应着一个实数。也就是说，有理数 $1/3$ 、无理数 π 在数轴上都有它们各自的位置。

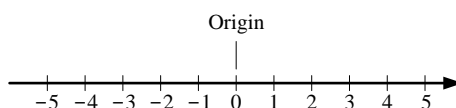


图 1. 实数数轴

与实数相对的就是虚数 (imaginary number)，虚数的定义来源于对负数开平方的扩展。一般定义 i 为虚数单位 (imaginary unit)， $i^2 = -1$ 。

把实数、虚数放在一起，我们便得到复数 (complex number)，比如 $a + bi$ (其中， a 、 b 均为实数)。

⚠ 注意，我们之所以提到实数、虚数、复数这些概念是因为，大家很快就会看到，特征值分解的特征值、特征向量中会出现复数。但是，复数矩阵相关的运算不是本书涉及的话题。

实对称矩阵的谱分解

给定一个 $n \times n$ 实对称矩阵 A ，其特征值分解可以写成

$$A = V\Lambda V^T \quad (1)$$

其中， V 为正交矩阵，满足 $VV^T = V^TV = I$ 。

V 的列向量为方阵 A 的特征向量。 Λ 为对角方阵， Λ 的主对角元素 Λ 的特征值。

我们管实对称矩阵的特征值分解叫做谱分解 (spectral decomposition)。大家很快就看到“谱”字的来由。

⚠ 注意，本书中提到的对称矩阵，都默认为实对称矩阵。

上一节的例子

接着上一节的例子， 2×2 实对称矩阵 A 的特征值分解 (谱分解)：

$$A = V\Lambda V^T = \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_V \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}}_\Lambda \underbrace{\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{V^T} = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} \quad (2)$$

将 A 作用于二维列向量 x ，对应矩阵乘法

$$Ax = y \quad (3)$$

A 可以看作一个复合几何变换，根据 (2) 把它拆解为三个连续动作

$$Ax = V\Lambda V^T x = y \quad (4)$$

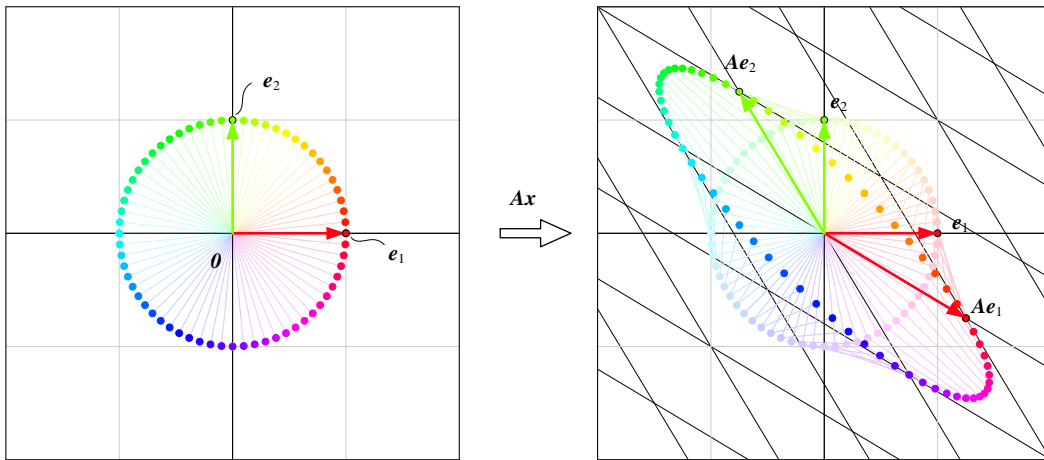
图 2 展示 A 对 e_1 、 e_2 两个单位向量的几何变换。

A 对 e_1 几何操作结果为

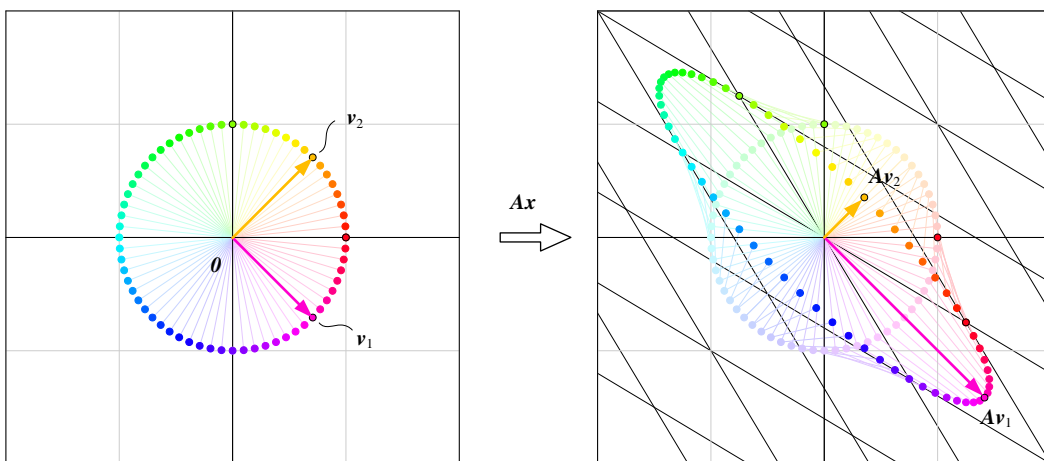
$$Ae_1 = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5/4 \\ -3/4 \end{bmatrix} \quad (5)$$

A 对 e_2 几何操作结果为

$$Ae_2 = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 5/4 \end{bmatrix} \quad (6)$$

图 2. 矩阵 A 对 e_1 、 e_2 的线性变换

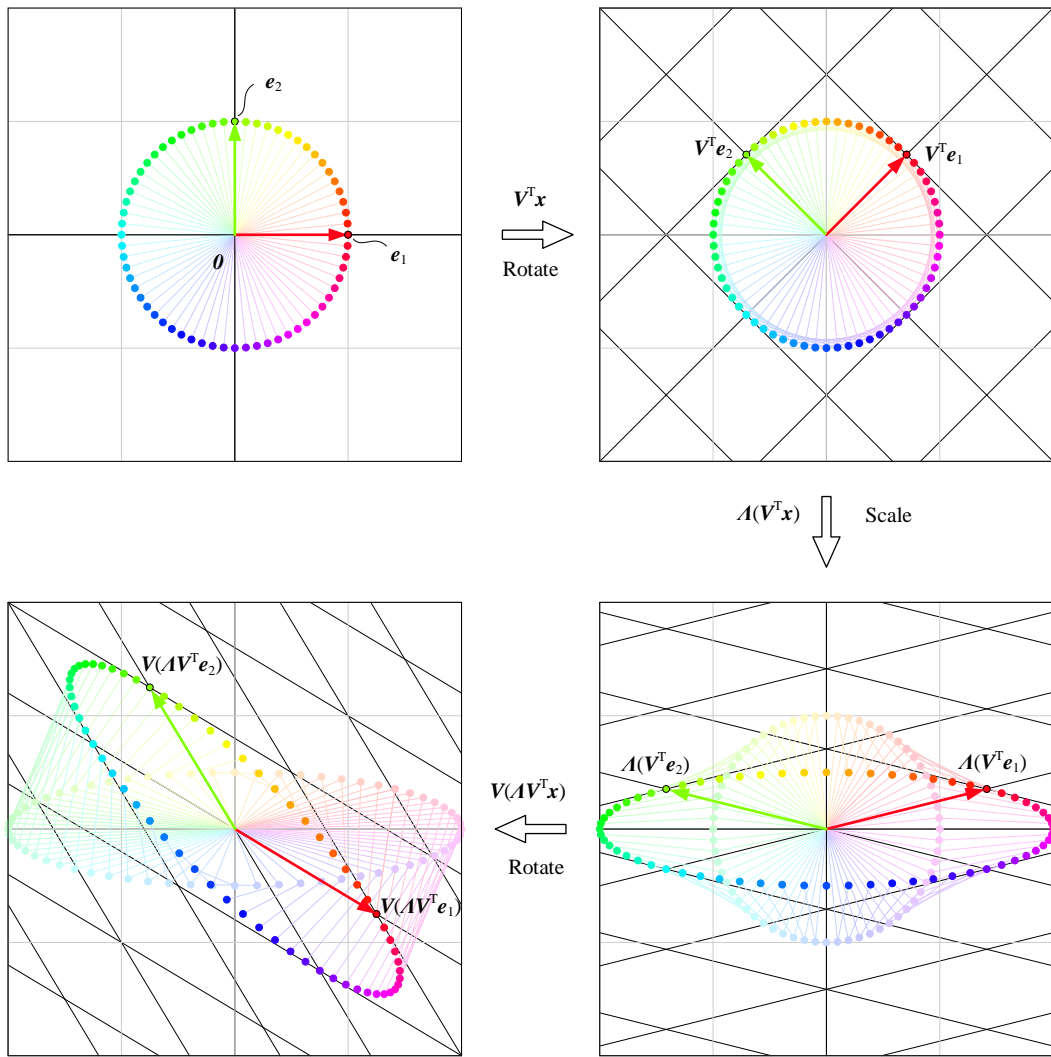
? 请大家写出图 3 对应的矩阵乘法。

图 3. 矩阵 A 对 v_1 、 v_2 的线性变换

下面让看看 $VAIV^T$ 分步作用的效果。

e_1 、 e_2 : 旋转 → 缩放 → 旋转

$VAIV^T$ 对 e_1 、 e_2 分步几何操作具体如图 4 所示。下面让我们逐步分析。

图 4. 矩阵 VAV^T 对 e_1 、 e_2 的线性变换

先用旋转矩阵 V^T 进行旋转

$$V^T = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \quad (7)$$

⚠ 本书前文提过，单位矩阵、旋转矩阵、置换矩阵、镜像矩阵，以及它们的任意顺序组合，都是正交矩阵。为了方便讨论，本章中，我们把正交矩阵的几何操作简化为旋转。

比如， V^T 对 e_1 的旋转结果为

$$V^T e_1 = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \quad (8)$$

V^T 对 e_2 的旋转结果为

$$V^T e_2 = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \quad (9)$$

容易计算得到的 V^T 行列式为 1；显然，旋转不改变面积。

(2) 中缩放矩阵 A 为

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \quad (10)$$

这个对角方阵的几何作用是对横轴放大 2 倍，纵轴缩小为 1/2。

巧合的是，这个对角方阵的行列式也是 1。这意味着图 4 中椭圆的面积和单位圆相同。方阵 A 的行列式，等于它所有特征值的乘积。

A 对 $V^T e_1$ 的缩放结果为

$$A(V^T e_1) = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ \sqrt{2}/4 \end{bmatrix} \quad (11)$$

A 对 $V^T e_2$ 的缩放结果为

$$A(V^T e_2) = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} -\sqrt{2} \\ \sqrt{2}/4 \end{bmatrix} \quad (12)$$

最后，用如下 V 旋转

$$V = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \quad (13)$$

V 对 $AV^T e_1$ 的旋转结果为

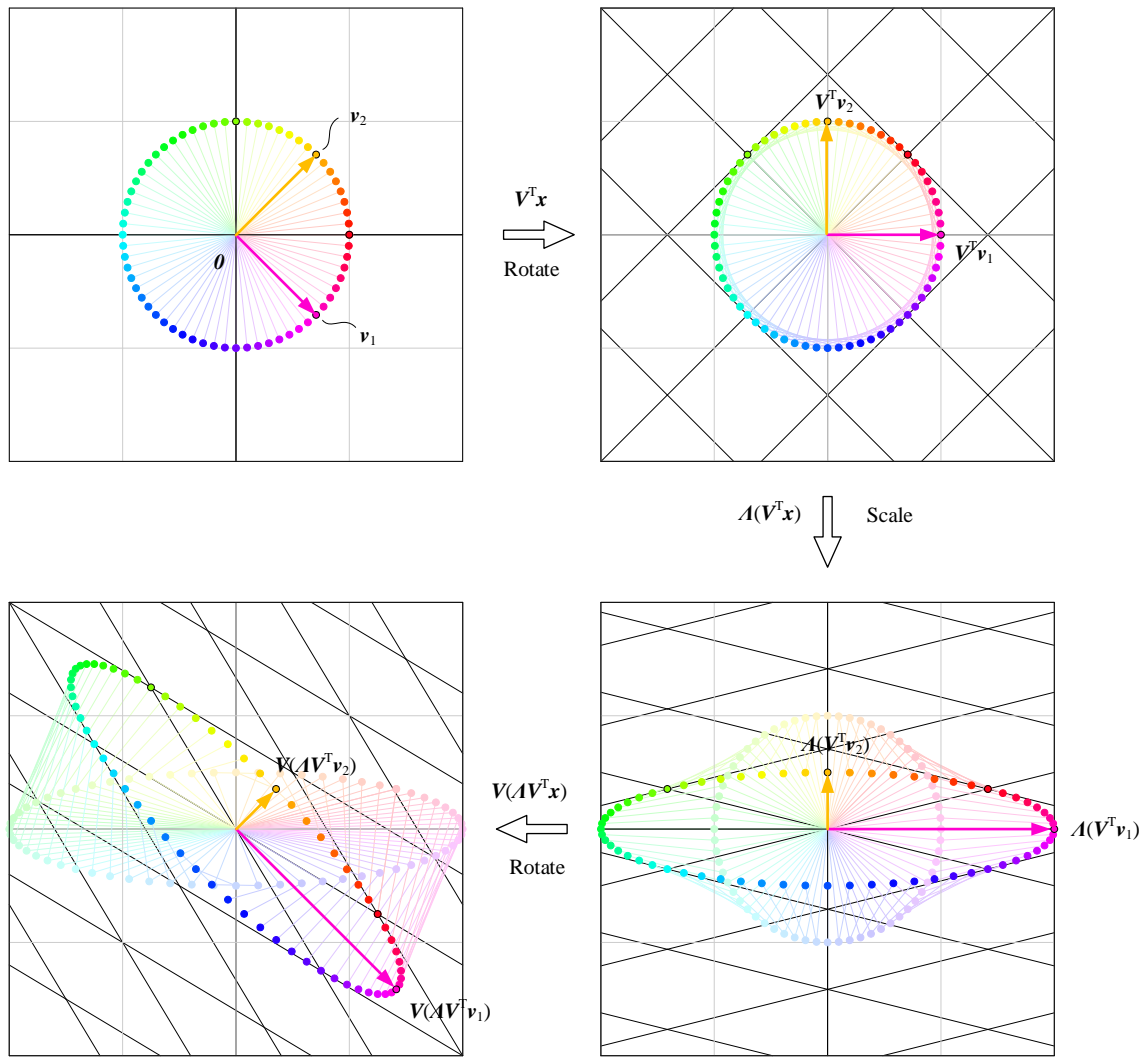
$$V(AV^T e_1) = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ \sqrt{2}/4 \end{bmatrix} = \begin{bmatrix} 5/4 \\ -3/4 \end{bmatrix} \quad (14)$$

V 对 $AV^T e_2$ 的旋转结果为

$$V(AV^T e_2) = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} -\sqrt{2} \\ \sqrt{2}/4 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 5/4 \end{bmatrix} \quad (15)$$

v_1, v_2 : 旋转 → 缩放 → 旋转

下面，让我们再看 VAV^T 对特征向量 v_1, v_2 的作用。

图 5. 矩阵 VAV^T 对 v_1 、 v_2 的线性变换

比如, V^T 对 v_1 的旋转结果为

$$V^T v_1 = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (16)$$

V^T 对 v_2 的旋转结果为

$$V^T v_2 = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (17)$$

我们竟然得到了标准正交基的基底向量 e_1 、 e_2 ? 这显然不是巧合!

把正交矩阵 V 展开写成列向量形式, (16)、(17) 可以写成

$$\begin{aligned} V^T v_1 &= [v_1 \ v_2]^T v_1 = \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} v_1 = \begin{bmatrix} v_1^T v_1 \\ v_2^T v_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ V^T v_2 &= [v_1 \ v_2]^T v_2 = \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix} v_2 = \begin{bmatrix} v_1^T v_2 \\ v_2^T v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \quad (18)$$

下面让我们看 A 的缩放效果。

本 PDF 文件为作者草稿, 发布目的为方便读者在移动终端学习, 终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有, 请勿商用, 引用请注明出处。

代码及 PDF 文件下载: <https://github.com/Visualize-ML>

本书配套微视频均发布在 B 站——生姜 DrGinger: <https://space.bilibili.com/513194466>

欢迎大家批评指教, 本书专属邮箱: jiang.visualize.ml@gmail.com

A 对 $V^T v_1$ 的缩放结果为

$$A(V^T v_1) = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad (19)$$

A 对 $V^T v_2$ 的缩放结果为

$$A(V^T v_2) = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} \quad (20)$$

如图 5 所示，单位圆变成了正椭圆。

最后，用 V 旋转。 V 对 $AV^T v_1$ 的旋转结果为

$$V(AV^T v_1) = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ -\sqrt{2} \end{bmatrix} = 2 \times \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix} = 2v_1 \quad (21)$$

V 对 $AV^T v_2$ 的缩放结果为

$$V(AV^T v_2) = \begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/4 \\ \sqrt{2}/4 \end{bmatrix} = 0.5 \times \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = 0.5v_2 \quad (22)$$

正椭圆旋转之后得到旋转椭圆。

$[v_1, v_2]$ 中看椭圆

正交矩阵 V 的列向量构成一个标准正交基，图 6 所示为在 $[v_1, v_2]$ 中观察椭圆。容易发现，几何角度来看，谱分解相当于帮我们找到一个坐标系将椭圆摆正！

λ_1 、 λ_2 为椭圆的半轴长度。



下一章要介绍的主成分分析 (Principal Component Analysis, PCA) 将用到这一点。

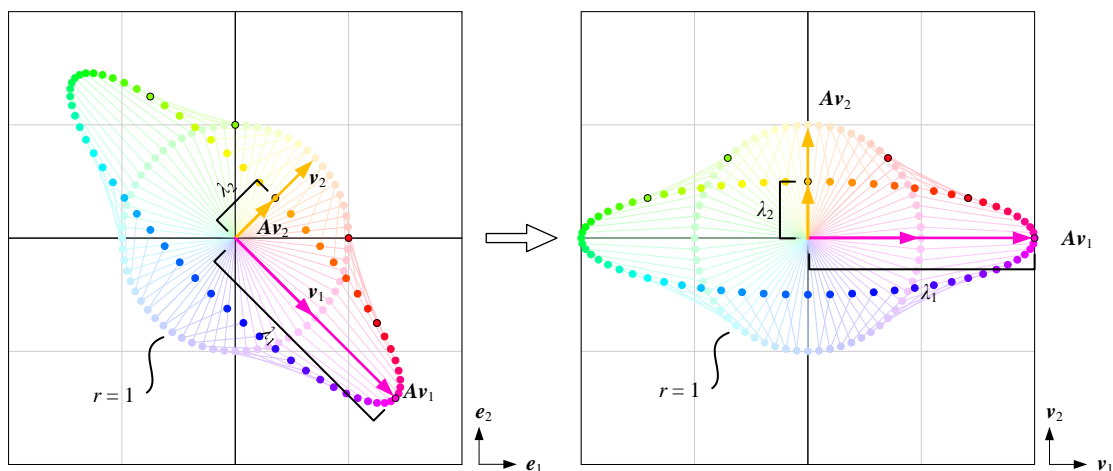


图 6. 分别在 $[e_1, e_2]$ 、 $[v_1, v_2]$ 中观察椭圆

这个椭圆一定有解析式，那么这个解析式和矩阵 A 有什么关系。如何求椭圆的半长轴、半短轴长度？下面就让我们回答这些问题。

椭圆

椭圆 (ellipse) 是几何中的一个非常重要的曲线，属于圆锥曲线 (conic section) 的一种。

简单来说，椭圆是平面内到两个定点的距离之和为定值的点的轨迹。这两个定点叫做焦点 (foci)，通常记作 F_1 、 F_2 。

如图 7 所示，椭圆上任意一点 P 满足

$$PF_1 + PF_2 = 2a \quad (23)$$

其中， $a > c > 0$ ； c 是焦点到椭圆中心的距离。

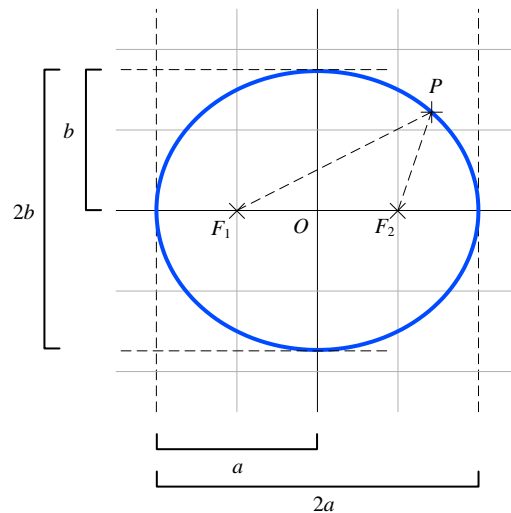


图 7. 中心位于原点，焦点位于横轴的椭圆

图 7 中这个以原点为中心，焦点在 x_1 轴 (横轴) 上的椭圆对应的方程为

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \quad (24)$$

其中， $a > b > 0$ 。

如图 7 所示，长轴 (major axis) 穿过两个焦点，长度为 $2a$ 。长度 a 是半长轴 (semi-major axis)；

短轴 (minor axis) 垂直于长轴，长度为 $2b$ 。长度 b 是半短轴 (semi-minor axis)。

我们管长轴、短轴和横、纵轴平行的椭圆叫做正椭圆；当长轴、短轴绕椭圆中心旋转，得到的是旋转椭圆。

回头再看图 6 右图，我们发现把椭圆摆正之后，椭圆半长轴长度为 λ_1 、半短轴的长度为 λ_2 ；椭圆的解析式为

$$\frac{x_1^2}{\lambda_1^2} + \frac{x_2^2}{\lambda_1^2} = 1 \quad (25)$$

写成矩阵乘法形式

我们可以很容易把 (24) 写成矩阵乘法形式

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1/a^2 & 0 \\ 0 & 1/b^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1 \quad (26)$$

据此，(25) 也可以写成矩阵乘法形式，即

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_2^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1 \quad (27)$$

我们如何一步步地从单位圆得到这个椭圆的呢？

单位圆 → 椭圆

假设单位圆上点对应单位向量

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \quad (28)$$

这个单位向量长度（模、大小、 L^2 范数、欧几里得范数）为 1，对应等式

$$\|\mathbf{z}\|^2 = \mathbf{z}^T \mathbf{z} = 1 \quad (29)$$

展开得到

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}^T \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = z_1^2 + z_2^2 = 1 \quad (30)$$

矩阵 \mathbf{A} 作用在 \mathbf{z} 后，得到 \mathbf{x} ，对应矩阵乘法

$$\mathbf{A}\mathbf{z} = \mathbf{x} \quad (31)$$

由于矩阵 \mathbf{A} 可逆，因此

$$\mathbf{z} = \mathbf{A}^{-1}\mathbf{x} \quad (32)$$

把 (32) 带入 (29) 得到

$$(\mathbf{A}^{-1}\mathbf{x})^T \mathbf{A}^{-1}\mathbf{x} = 1 \quad (33)$$

整理

$$\mathbf{x}^T (\mathbf{A}^{-1})^T \mathbf{A}^{-1}\mathbf{x} = 1 \quad (34)$$

矩阵 \mathbf{A} 的逆为

$$\mathbf{A}^{-1} = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix}^{-1} = \begin{bmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{bmatrix} \quad (35)$$

带入 (34) 整理得到旋转椭圆的解析式

$$2.125x_1^2 + 3.75x_1x_2 + 2.125x_2^2 = 1 \quad (36)$$

要从这个式子直接得到椭圆的半长轴、半短轴的长度没那么容易。

让我们换个角度，由于 $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ ， \mathbf{A} 的逆可以写成

$$\mathbf{A}^{-1} = (\mathbf{V}^T)^{-1} \mathbf{\Lambda}^{-1} \mathbf{V}^{-1} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T \quad (37)$$

(37) 带入 (34) 得到

$$\mathbf{x}^T (\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T)^T \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T \mathbf{x} = 1 \quad (38)$$

整理

$$\mathbf{x}^T \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T \mathbf{x} = 1 \quad (39)$$

得到

$$(\mathbf{V}^T \mathbf{x})^T \mathbf{\Lambda}^{-2} (\mathbf{V}^T \mathbf{x}) = 1 \quad (40)$$

令 $\mathbf{y} = \mathbf{V}^T \mathbf{x}$ ，上式可以写成

$$\mathbf{y}^T \mathbf{\Lambda}^{-2} \mathbf{y} = 1 \quad (41)$$

即

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_2^2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 1 \quad (42)$$

整理得到

$$\frac{y_1^2}{\lambda_1^2} + \frac{y_2^2}{\lambda_2^2} = 1 \quad (43)$$

这便是大家想要的正椭圆形式。



此外，除了椭圆，圆锥曲线还有双曲线、抛物线等等，它们又和线性代数又怎样的关系，这是本书第 13 章要讨论的话题。

更一般的谱分解

理解了 2×2 实对称矩阵 \mathbf{A} 谱分解之后，让我们聊聊更一般的谱分解。

如图 8 所示， $n \times n$ 实对称矩阵 \mathbf{A} ，谱分解后的结果为 $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ 。 \mathbf{V} 为正交矩阵，满足 $\mathbf{V}\mathbf{V}^T = \mathbf{V}^T\mathbf{V} = \mathbf{I}$ 。

$\mathbf{\Lambda}$ 为对角方阵，主对角线上为特征值。

图 8 中， A 主对角线元素从左上到右下角由大到小排列。

为了方便讨论，几何角度来看， V 在 \mathbb{R}^n 完成旋转， A 在 \mathbb{R}^n 完成缩放。

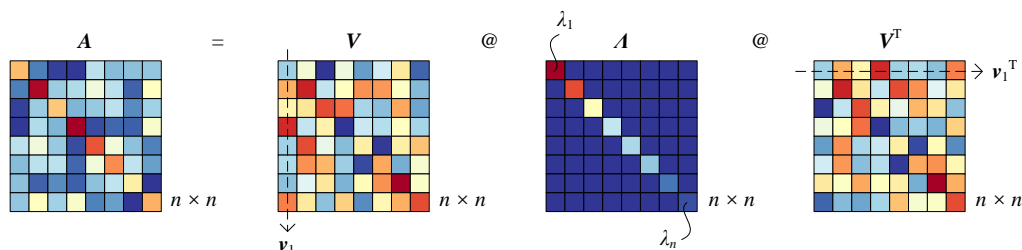


图 8. $n \times n$ 实对称矩阵 A 谱分解

矩阵乘法第二视角

用矩阵乘法第二视角——外积视角——把 $A = V\Lambda V^T$ 写成

$$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \quad (44)$$

$$= \lambda_1 v_1 v_1^T + \lambda_2 v_2 v_2^T + \cdots + \lambda_n v_n v_n^T$$

这样我们把矩阵 A 写成 n 个矩阵的叠加，如图 9。特征值越大，子图的颜色越“鲜艳”。只要 λ_i 不为 0， $\lambda_i v_i v_i^T$ 就是秩一矩阵；否则，如果 λ_i 为 0， $\lambda_i v_i v_i^T$ 就是零矩阵。

将一个实对称矩阵进行谱分解，就是将它“拆解”图 9 中的子图方阵。这些方阵就像是白光通过棱镜被分解时，展现出一系列有序排列的颜色带，这些不同颜色光代表不同波长成分。

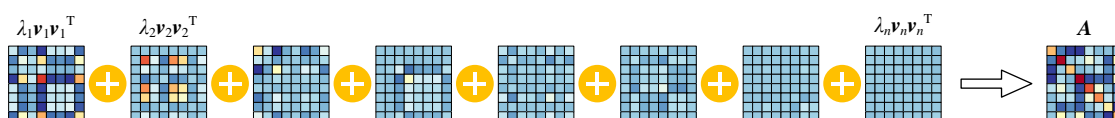
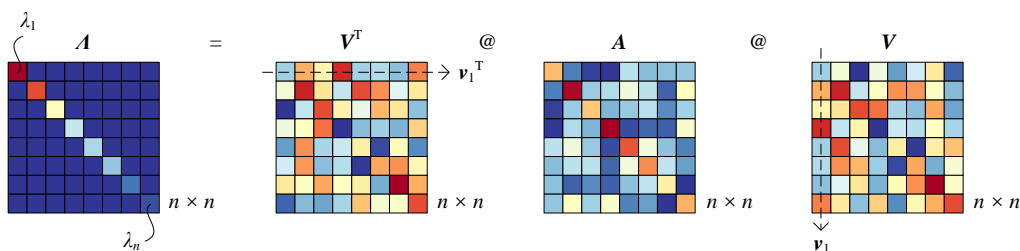


图 9. n 个矩阵的叠加得到 A

对角化

显然对角方阵 V 可逆，如图 10 所示，实对称矩阵 A 的对角化则对应矩阵乘法

$$V^T A V = \Lambda \quad (45)$$

图 10. $n \times n$ 实对称矩阵 A 对角化

用矩阵乘法第一视角展开 (45) 得到

$$\begin{aligned}
 V^T A V &= \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} A \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} \\
 &= \begin{bmatrix} v_1^T A v_1 & v_1^T A v_2 & \cdots & v_1^T A v_n \\ v_2^T A v_1 & v_2^T A v_2 & \cdots & v_2^T A v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n^T A v_1 & v_n^T A v_2 & \cdots & v_n^T A v_n \end{bmatrix} \\
 &= \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}
 \end{aligned} \tag{46}$$

让我们先看 (46) 主对角线矩阵乘法。

以 $v_1^T A v_1 = \lambda_1$ 为例，由于 $A v_1 = \lambda_1 v_1$ ，且 v_1 为单位向量， $v_1^T A v_1 = \lambda_1$ 可以写成

$$v_1^T A v_1 = \lambda_1 v_1^T v_1 = \lambda_1 \|v_1\|_2^2 = \lambda_1 \tag{47}$$

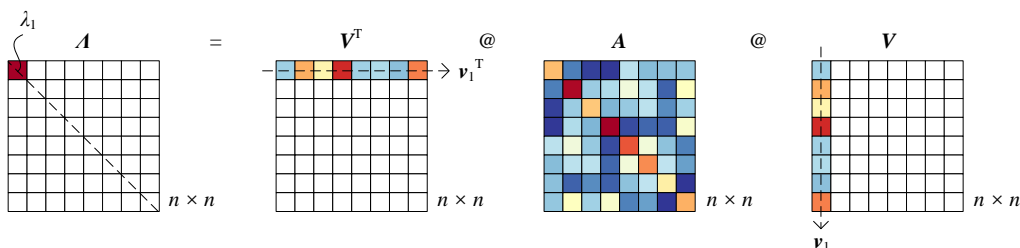
上式是瑞利商 (Rayleigh quotient) 的特殊形式，用来度量向量 v_1 在 A 作用下的缩放效果。

给定实数对称矩阵 A ， A 的瑞利商定义为：

$$R(x) = \frac{x^T A x}{x^T x} \tag{48}$$



本书第 13 章、第 3 节将专门介绍瑞利商。

图 11. $n \times n$ 实对称矩阵 A 对角化，主对角线元素

下面再看 (46) 非主对角线乘法。以 $\mathbf{v}_2^T \mathbf{A} \mathbf{v}_1 = 0$ 为例，由于 \mathbf{V} 的列向量两两正交，这个式子可以写成

$$\mathbf{v}_2^T \mathbf{A} \mathbf{v}_1 = \lambda_1 \mathbf{v}_2^T \mathbf{v}_1 = 0 \quad (49)$$

上式告诉我们，在 \mathbf{A} 的特定线性转换中，特征向量 \mathbf{v}_1 、 \mathbf{v}_2 依然保持了正交性。

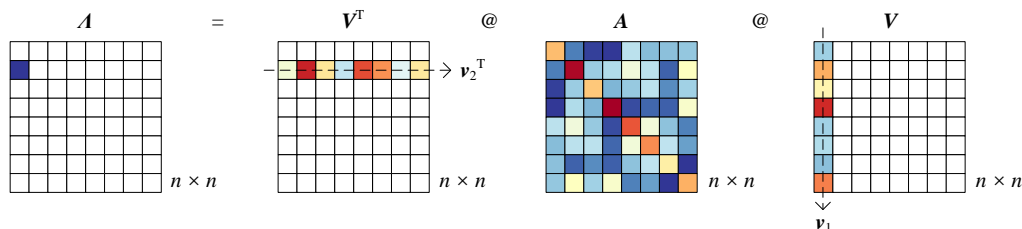


图 12. $n \times n$ 实对称矩阵 \mathbf{A} 对角化，非主对角线元素

此外，由于 \mathbf{A} 为实对称矩阵，(49) 还可以转置得到

$$\mathbf{v}_1^T \mathbf{A} \mathbf{v}_2 = \lambda_1 \mathbf{v}_1^T \mathbf{v}_2 = 0 \quad (50)$$

逆矩阵

只有所有特征值均非 0，实对称矩阵 \mathbf{A} 才有逆，即

$$\mathbf{A}^{-1} = (\mathbf{V} \mathbf{A} \mathbf{V}^T)^{-1} = (\mathbf{V}^T)^{-1} \mathbf{A}^{-1} \mathbf{V}^{-1} = \mathbf{V} \mathbf{A}^{-1} \mathbf{V}^T = \mathbf{V} \begin{bmatrix} 1/\lambda_1 & & & \\ & 1/\lambda_2 & & \\ & & \ddots & \\ & & & 1/\lambda_n \end{bmatrix} \mathbf{V}^T \quad (51)$$

上式相当于逆矩阵 \mathbf{A}^{-1} 的谱分解，具体如图 13 所示。

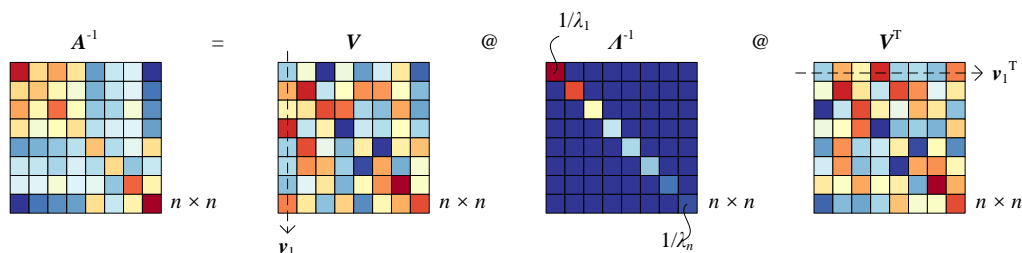


图 13. $n \times n$ 实对称矩阵 \mathbf{A}^{-1} 谱分解



前文中，大家最常见到的对称矩阵是格拉姆矩阵。本章第 4 节将专门讲解格拉姆矩阵。

提取元素

本节最后让我们展开聊聊 (46) 中一组有趣矩阵乘法。

本 PDF 文件为作者草稿，发布目的为方便读者在移动终端学习，终稿内容以清华大学出版社纸质出版物为准。

版权归清华大学出版社所有，请勿商用，引用请注明出处。

代码及 PDF 文件下载：<https://github.com/Visualize-ML>

本书配套微视频均发布在 B 站——生姜 DrGinger：<https://space.bilibili.com/513194466>

欢迎大家批评指教，本书专属邮箱：jiang.visualize.ml@gmail.com

比如，如下矩阵乘法相当于提取单位矩阵 \mathbf{I} 的第 1 行、第 1 列元素

$$\mathbf{e}_1^T \mathbf{I} \mathbf{e}_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 1 \quad (52)$$

再如，下面这个矩阵乘法相当于提取单位矩阵 \mathbf{I} 的第 2 行、第 1 列元素

$$\mathbf{e}_2^T \mathbf{I} \mathbf{e}_1 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0 \quad (53)$$

❓ 请大家思考如何提取单位矩阵 \mathbf{I} 的剩余其他元素。

如果实对称矩阵 \mathbf{A} 可以谱分解得到 $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ ，如下矩阵乘法相当于提取对角方阵 $\mathbf{\Lambda}$ (不是 \mathbf{A}) 第 1 行、第 1 列元素

$$\underset{\mathbf{e}_1^T}{\mathbf{v}_1^T} \mathbf{A} \underset{\mathbf{e}_1}{\mathbf{v}_1} = \underset{\mathbf{e}_1^T}{\mathbf{v}_1^T} \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \underset{\mathbf{e}_1}{\mathbf{v}_1} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \quad (54)$$

相当于在 \mathbf{V} 列向量构成的规范正交基中，我们把实对称矩阵 \mathbf{A} “摆正”得到对角方阵 $\mathbf{\Lambda}$ 。

再如，下面这个矩阵乘法相当于提取对角方阵 $\mathbf{\Lambda}$ 的第 2 行、第 1 列元素

$$\underset{\mathbf{e}_2^T}{\mathbf{v}_2^T} \mathbf{A} \underset{\mathbf{e}_1}{\mathbf{v}_1} = \underset{\mathbf{e}_2^T}{\mathbf{v}_2^T} \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T \underset{\mathbf{e}_1}{\mathbf{v}_1} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0 \quad (55)$$


❓ 请大家思考如何提取对角方阵 $\mathbf{\Lambda}$ 的剩余其他元素。



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 请对如下实对称矩阵谱分解，并从几何角度解释矩阵分解结果。

► $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$


$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Q2. 请修改 LA_11_01_01.ipynb 可视化 **Q1** 中的谱分解结果。