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 $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$，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 $\boldsymbol{A}$ 为对称矩阵，则满足 $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$。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。

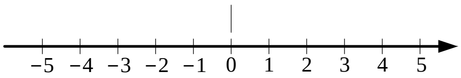

图 1. 实数数轴

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)，$\mathrm{i}^2 = -1$。把实数、虚数放在一起，我们便得到复数 (complex number)，比如 $a + b\mathrm{i}$ (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。

实对称矩阵的谱分解

给定一个 $n \times n$ 实对称矩阵 $\boldsymbol{A}$，其特征值分解可以写成

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}} \tag{1}$$

其中，$\boldsymbol{V}$ 为正交矩阵，满足 $\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}} = \boldsymbol{V}^{\mathrm{T}}\boldsymbol{V} = \boldsymbol{I}$。

$\boldsymbol{V}$ 的列向量为方阵 $\boldsymbol{A}$ 的特征向量。$\boldsymbol{\Lambda}$ 为对角方阵，$\boldsymbol{\Lambda}$ 的主对角元素 $\boldsymbol{A}$ 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。

上一节的例子

接着上一节的例子，2×2 实对称矩阵 $\boldsymbol{A}$ 的特征值分解 (谱分解)：

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}} = \underbrace{\begin{bmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{\boldsymbol{V}} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}}_{\boldsymbol{\Lambda}} \underbrace{\begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}}_{\boldsymbol{V}^{\mathrm{T}}} = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix} \tag{2}$$

将 $\boldsymbol{A}$ 作用于二维列向量 $\boldsymbol{x}$，对应矩阵乘法

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y} \tag{3}$$

$\boldsymbol{A}$ 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

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x} = \boldsymbol{y} \tag{4}$$

图 2 展示 $\boldsymbol{A}$ 对 $\boldsymbol{e}_1$、$\boldsymbol{e}_2$ 两个单位向量的几何变换。

$\boldsymbol{A}$ 对 $\boldsymbol{e}_1$ 几何操作结果为

$$\boldsymbol{A}\boldsymbol{e}_1 = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5/4 \\ -3/4 \end{bmatrix} \tag{5}$$

$\boldsymbol{A}$ 对 $\boldsymbol{e}_2$ 几何操作结果为

$$\boldsymbol{A}\boldsymbol{e}_2 = \begin{bmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 5/4 \end{bmatrix} \tag{6}$$

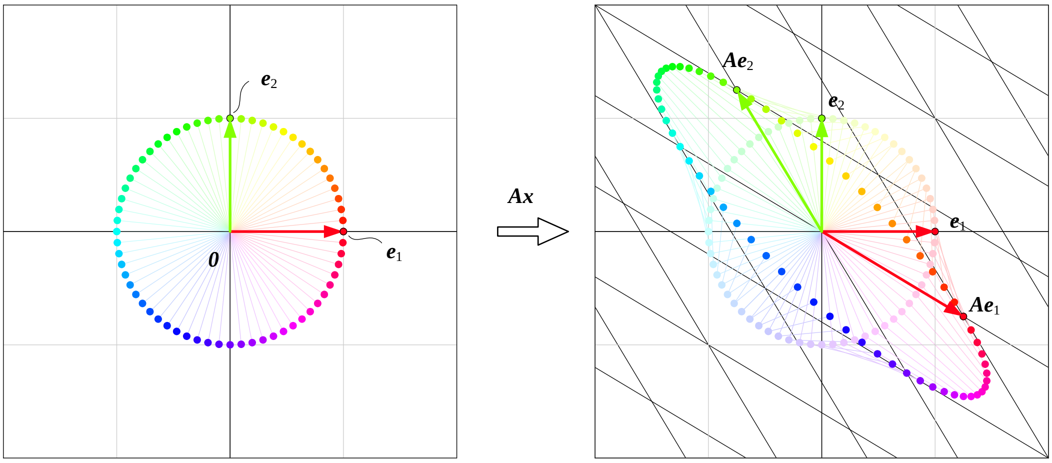

图 2. 矩阵 A 对 e_1、e_2 的线性变换

请大家写出图 3 对应的矩阵乘法。

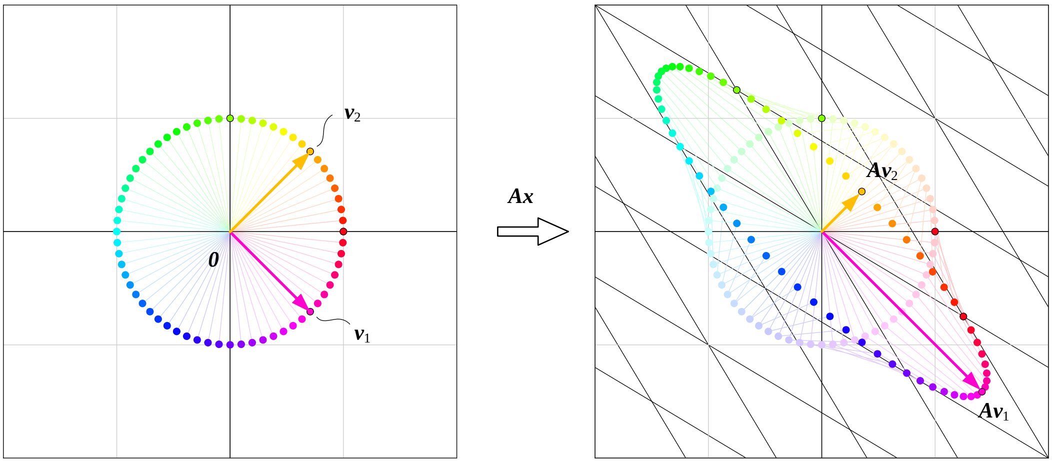

图 3. 矩阵 A 对 v_1、v_2 的线性变换

下面让看看 $V\Lambda V^{\mathrm{T}}$ 分步作用的效果。

e_1、e_2：旋转 → 缩放 → 旋转

$V\Lambda V^{\mathrm{T}}$ 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。

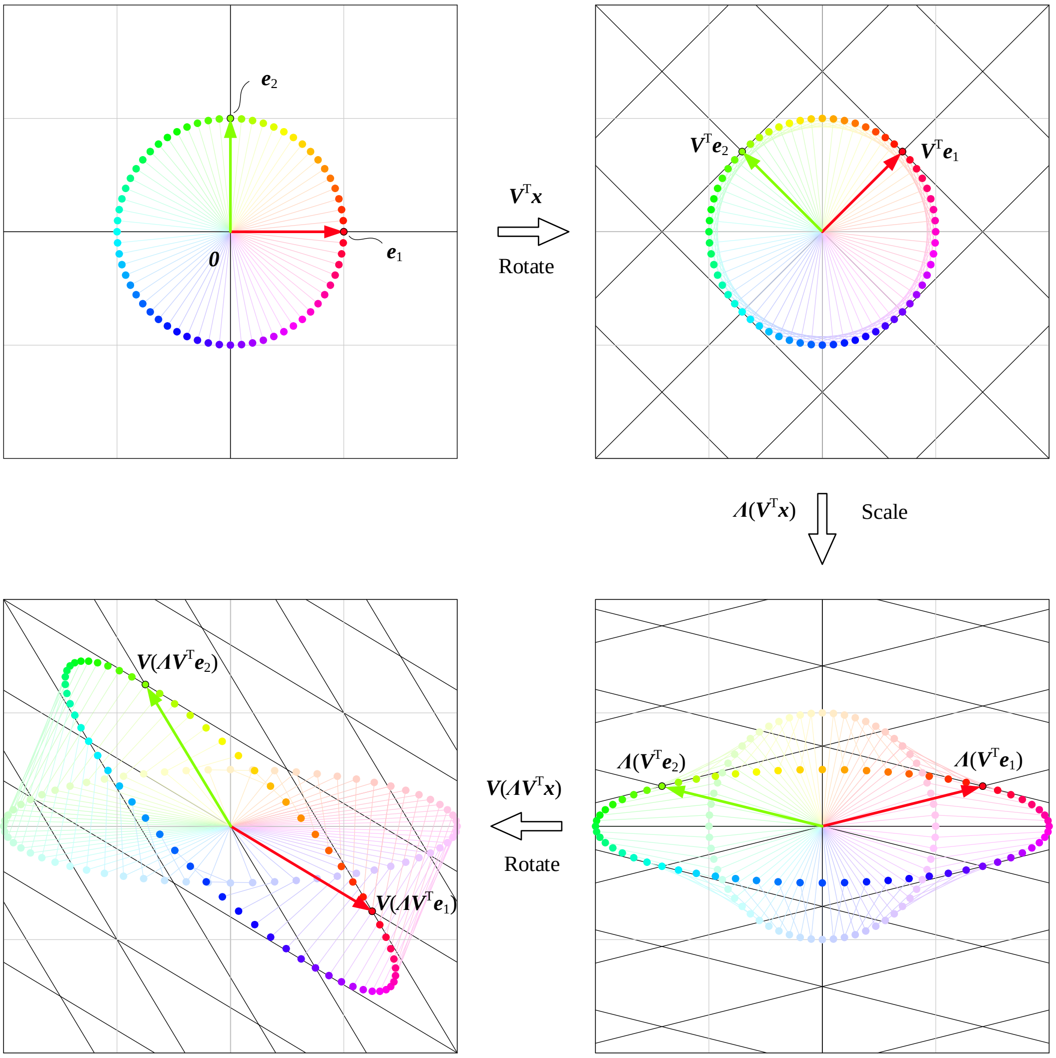

图 4. 矩阵 $\boldsymbol{V\Lambda V}^{\mathrm{T}}$ 对 $\boldsymbol{e}_1$、$\boldsymbol{e}_2$ 的线性变换

先用旋转矩阵 $\boldsymbol{V}^{\mathrm{T}}$ 进行旋转

$$\boldsymbol{V}^{\mathrm{T}} = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \tag{7}$$

⚠ 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，$\boldsymbol{V}^{\mathrm{T}}$ 对 $\boldsymbol{e}_1$ 的旋转结果为

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_1 = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \tag{8}$$

$\boldsymbol{V}^{\mathrm{T}}$ 对 $\boldsymbol{e}_2$ 的旋转结果为

$$\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_2 = \begin{bmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \tag{9}$$

容易计算得到的 $\boldsymbol{V}^{\mathrm{T}}$ 行列式为 1；显然，旋转不改变面积。

(2) 中缩放矩阵 $\boldsymbol{\Lambda}$ 为

$$\boldsymbol{\Lambda}=\begin{bmatrix}2 & 0\\ 0 & 1/2\end{bmatrix} \tag{10}$$

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 $\boldsymbol{A}$ 的行列式，等于它所有特征值的乘积。

$\boldsymbol{\Lambda}$ 对 $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_1$ 的缩放结果为

$$\boldsymbol{\Lambda}\left(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_1\right)=\begin{bmatrix}2 & 0\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}\sqrt{2}/2\\ \sqrt{2}/2\end{bmatrix}=\begin{bmatrix}\sqrt{2}\\ \sqrt{2}/4\end{bmatrix} \tag{11}$$

$\boldsymbol{\Lambda}$ 对 $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_2$ 的缩放结果为

$$\boldsymbol{\Lambda}\left(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_2\right)=\begin{bmatrix}2 & 0\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}-\sqrt{2}/2\\ \sqrt{2}/2\end{bmatrix}=\begin{bmatrix}-\sqrt{2}\\ \sqrt{2}/4\end{bmatrix} \tag{12}$$

最后，用如下 $\boldsymbol{V}$ 旋转

$$\boldsymbol{V}=\begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\\ -\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix} \tag{13}$$

$\boldsymbol{V}$ 对 $\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_1$ 的旋转结果为

$$\boldsymbol{V}\left(\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_1\right)=\begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\\ -\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}\begin{bmatrix}\sqrt{2}\\ \sqrt{2}/4\end{bmatrix}=\begin{bmatrix}5/4\\ -3/4\end{bmatrix} \tag{14}$$

$\boldsymbol{V}$ 对 $\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_2$ 的旋转结果为

$$\boldsymbol{V}\left(\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{e}_2\right)=\begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\\ -\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}\begin{bmatrix}-\sqrt{2}\\ \sqrt{2}/4\end{bmatrix}=\begin{bmatrix}-3/4\\ 5/4\end{bmatrix} \tag{15}$$

$\boldsymbol{v}_1$、$\boldsymbol{v}_2$：旋转 → 缩放 → 旋转

下面，让我们再看 $\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}$ 对特征向量 $\boldsymbol{v}_1$、$\boldsymbol{v}_2$ 的作用。

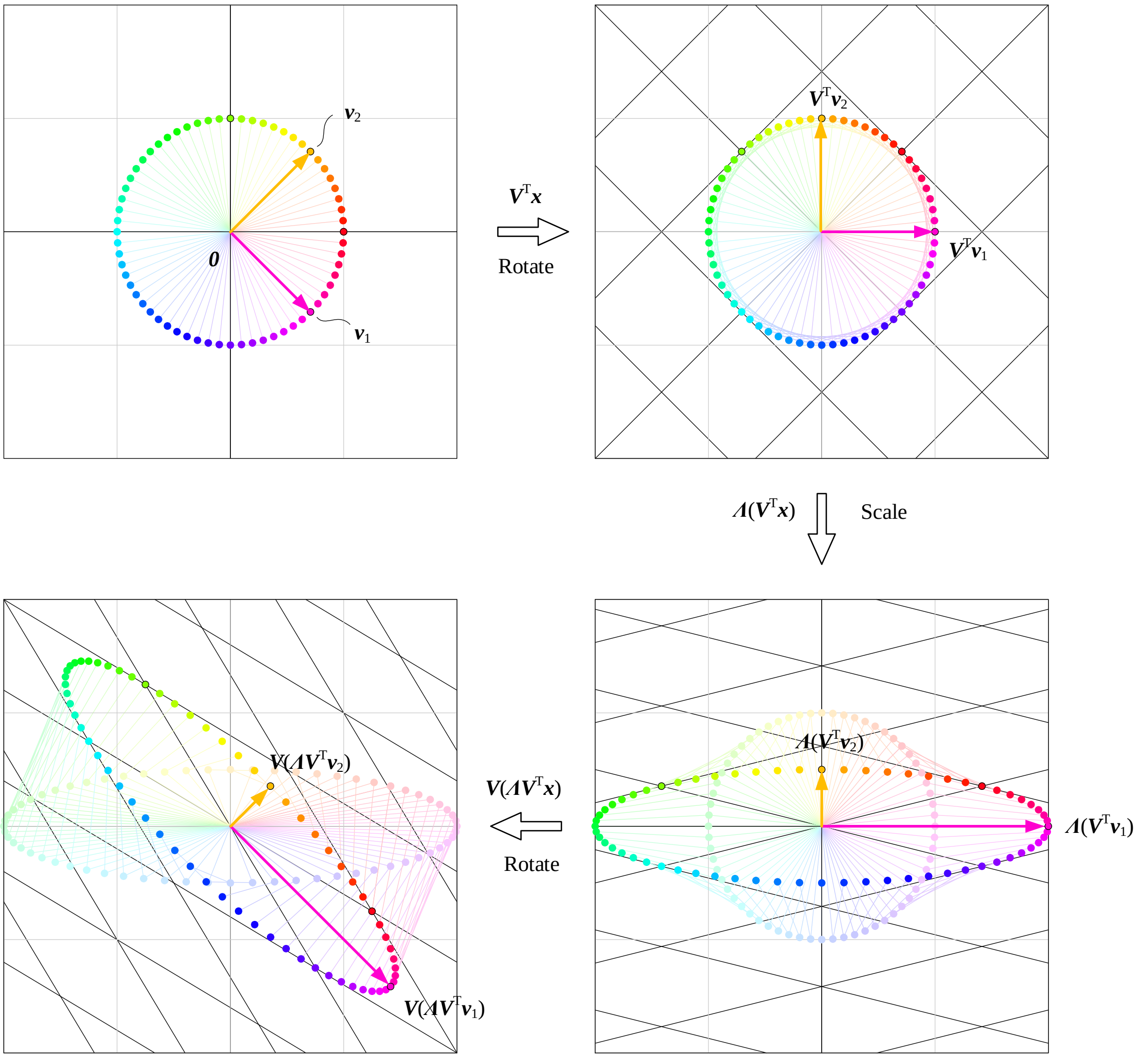

图 5. 矩阵 $V\Lambda V^{\mathrm{T}}$ 对 v_1、v_2 的线性变换

比如，V^{T} 对 v_1 的旋转结果为

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

V^{T} 对 v_2 的旋转结果为

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

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

$$\begin{aligned} V^{\mathrm{T}}v_1 &= \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{\mathrm{T}} v_1 = \begin{bmatrix} v_1^{\mathrm{T}} \\ v_2^{\mathrm{T}} \end{bmatrix} v_1 = \begin{bmatrix} v_1^{\mathrm{T}}v_1 \\ v_2^{\mathrm{T}}v_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ V^{\mathrm{T}}v_2 &= \begin{bmatrix} v_1 & v_2 \end{bmatrix}^{\mathrm{T}} v_2 = \begin{bmatrix} v_1^{\mathrm{T}} \\ v_2^{\mathrm{T}} \end{bmatrix} v_2 = \begin{bmatrix} v_1^{\mathrm{T}}v_2 \\ v_2^{\mathrm{T}}v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \tag{18}$$

下面让我们看 Λ 的缩放效果。

$\boldsymbol{\Lambda}$ 对 $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_1$ 的缩放结果为

$$\boldsymbol{\Lambda}\left(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_1\right)=\begin{bmatrix}2 & 0\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=\begin{bmatrix}2\\ 0\end{bmatrix} \tag{19}$$

$\boldsymbol{\Lambda}$ 对 $\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_2$ 的缩放结果为

$$\boldsymbol{\Lambda}\left(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_2\right)=\begin{bmatrix}2 & 0\\ 0 & 1/2\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}0\\ 1/2\end{bmatrix} \tag{20}$$

如图 5 所示，单位圆变成了正椭圆。

最后，用 $\boldsymbol{V}$ 旋转。$\boldsymbol{V}$ 对 $\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_1$ 的旋转结果为

$$\boldsymbol{V}\left(\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_1\right)=\begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\\ -\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}\begin{bmatrix}2\\ 0\end{bmatrix}=\begin{bmatrix}\sqrt{2}\\ -\sqrt{2}\end{bmatrix}=2\times\begin{bmatrix}\sqrt{2}/2\\ -\sqrt{2}/2\end{bmatrix}=2\boldsymbol{v}_1 \tag{21}$$

$\boldsymbol{V}$ 对 $\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_2$ 的缩放结果为

$$\boldsymbol{V}\left(\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{v}_2\right)=\begin{bmatrix}\sqrt{2}/2 & \sqrt{2}/2\\ -\sqrt{2}/2 & \sqrt{2}/2\end{bmatrix}\begin{bmatrix}0\\ 1/2\end{bmatrix}=\begin{bmatrix}\sqrt{2}/4\\ \sqrt{2}/4\end{bmatrix}=0.5\times\begin{bmatrix}\sqrt{2}/2\\ \sqrt{2}/2\end{bmatrix}=0.5\boldsymbol{v}_2 \tag{22}$$

正椭圆旋转之后得到旋转椭圆。

[$\boldsymbol{v}_1$, $\boldsymbol{v}_2$] 中看椭圆

正交矩阵 $\boldsymbol{V}$ 的列向量构成一个标准正交基，图 6 所示为在 [$\boldsymbol{v}_1$, $\boldsymbol{v}_2$] 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。

本章最后要介绍的主成分分析 (Principal Component Analysis, PCA) 将用到这一点。

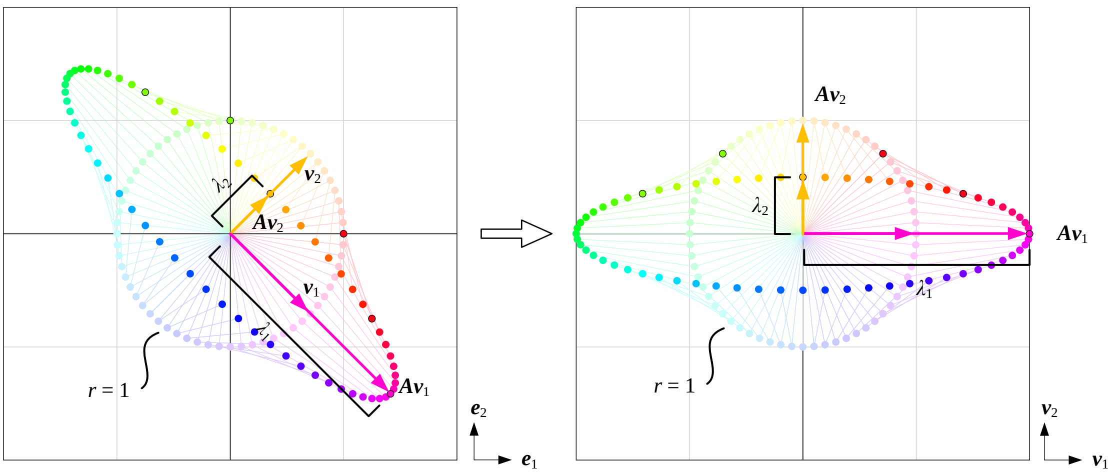

图 6. 分别在 [$\boldsymbol{e}_1$, $\boldsymbol{e}_2$]、[$\boldsymbol{v}_1$, $\boldsymbol{v}_2$] 中观察椭圆

这个椭圆一定有解析式，那么这个解析式和矩阵 $\boldsymbol{A}$ 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

$$PF_1 + PF_2 = 2a \tag{23}$$

其中，$a > c > 0$；c 是焦点到椭圆中心的距离。

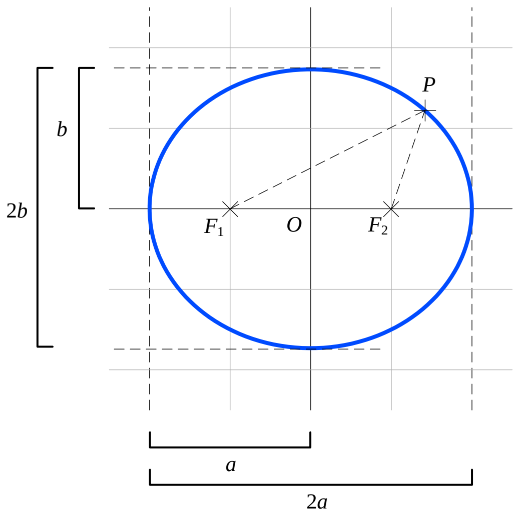

图 7．中心位于原点，焦点位于横轴的椭圆

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

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \tag{24}$$

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

$$\frac{x_1^2}{\lambda_1^2}+\frac{x_2^2}{\lambda_1^2}=1 \tag{25}$$

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

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

据此，(25) 也可以写成矩阵乘法形式，即

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

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

$$\boldsymbol{z}=\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{28}$$

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

$$\|\boldsymbol{z}\|^2=\boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}=1 \tag{29}$$

展开得到

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}^{\mathrm{T}}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}=z_1^2+z_2^2=1 \tag{30}$$

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

$$\boldsymbol{A}\boldsymbol{z}=\boldsymbol{x} \tag{31}$$

由于矩阵 $\boldsymbol{A}$ 可逆，因此

$$\boldsymbol{z}=\boldsymbol{A}^{-1}\boldsymbol{x} \tag{32}$$

把 (32) 带入 (29) 得到

$$\left(\boldsymbol{A}^{-1}\boldsymbol{x}\right)^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{x}=1 \tag{33}$$

整理

$$\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{A}^{-1}\right)^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{x}=1 \tag{34}$$

矩阵 $\boldsymbol{A}$ 的逆为

$$A^{-1}=\begin{bmatrix}5/4 & -3/4\\ -3/4 & 5/4\end{bmatrix}^{-1}=\begin{bmatrix}5/4 & 3/4\\ 3/4 & 5/4\end{bmatrix} \tag{35}$$

带入 (34) 整理得到旋转椭圆的解析式

$$2.125x_1^2+3.75x_1x_2+2.125x_2^2=1 \tag{36}$$

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 $\boldsymbol{A}=\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}$，$\boldsymbol{A}$ 的逆可以写成

$$\boldsymbol{A}^{-1}=\left(\boldsymbol{V}^{\mathrm{T}}\right)^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{V}^{-1}=\boldsymbol{V}\boldsymbol{\Lambda}^{-1}\boldsymbol{V}^{\mathrm{T}} \tag{37}$$

(37) 带入 (34) 得到

$$\boldsymbol{x}^{\mathrm{T}}\left(\boldsymbol{V}\boldsymbol{\Lambda}^{-1}\boldsymbol{V}^{\mathrm{T}}\right)^{\mathrm{T}}\boldsymbol{V}\boldsymbol{\Lambda}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x}=1 \tag{38}$$

整理

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{\Lambda}^{-1}\underbrace{\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}}_{\boldsymbol{I}}\boldsymbol{\Lambda}^{-1}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x}=1 \tag{39}$$

得到

$$\left(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x}\right)^{\mathrm{T}}\boldsymbol{\Lambda}^{-2}\left(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{x}\right)=1 \tag{40}$$

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

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}^{-2}\boldsymbol{y}=1 \tag{41}$$

即

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

整理得到

$$\frac{y_1^2}{\lambda_1^2}+\frac{y_2^2}{\lambda_2^2}=1 \tag{43}$$

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，这是本书后续要讨论的话题。

更一般的谱分解

理解了 2 × 2 实对称矩阵 $\boldsymbol{A}$ 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 $\boldsymbol{A}$，谱分解后的结果为 $\boldsymbol{A}=\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}$ 。$\boldsymbol{V}$ 为正交矩阵，满足 $\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}=\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}=\boldsymbol{I}$。$\boldsymbol{\Lambda}$ 为对角方阵，主对角线上为特征值。图中，$\boldsymbol{\Lambda}$ 主对角线元素从左上到右下角由大到小排列。

几何角度来看，$\boldsymbol{V}$ 在 $\mathbb{R}^n$ 完成旋转，$\boldsymbol{\Lambda}$ 在 $\mathbb{R}^n$ 完成缩放。

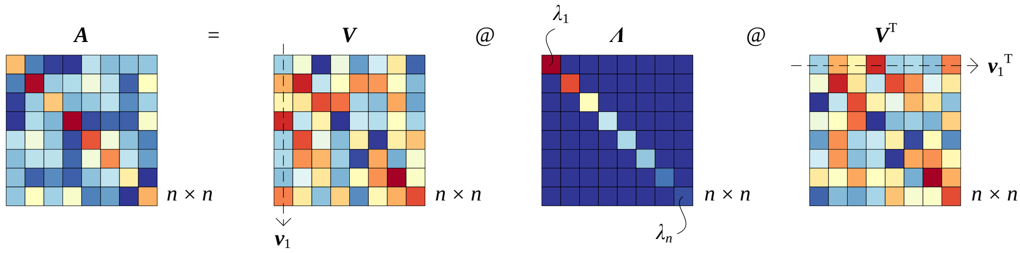

图 8. $n \times n$ 实对称矩阵 $\boldsymbol{A}$ 谱分解

矩阵乘法第二视角

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

$$
\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1^{\mathrm{T}} \\ \boldsymbol{v}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{v}_n^{\mathrm{T}} \end{bmatrix} \\
&= \lambda_1 \boldsymbol{v}_1 \boldsymbol{v}_1^{\mathrm{T}} + \lambda_2 \boldsymbol{v}_2 \boldsymbol{v}_2^{\mathrm{T}} + \cdots + \lambda_n \boldsymbol{v}_n \boldsymbol{v}_n^{\mathrm{T}}
\end{aligned}
\tag{44}
$$

这样我们把矩阵 $\boldsymbol{A}$ 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，$\lambda_i \boldsymbol{v}_i \boldsymbol{v}_i^{\mathrm{T}}$ 就是秩一矩阵；否则，如果 λ_i 为 0，$\lambda_i \boldsymbol{v}_i \boldsymbol{v}_i^{\mathrm{T}}$ 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。

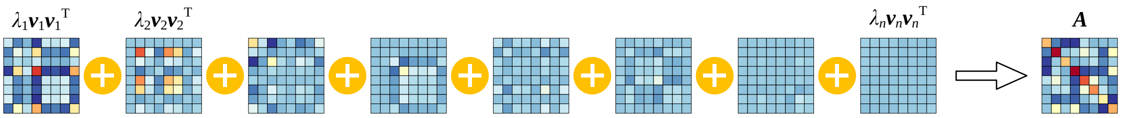

图 9. n 个矩阵的叠加得到 $\boldsymbol{A}$

对角化

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

$$
\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{\Lambda} \tag{45}
$$

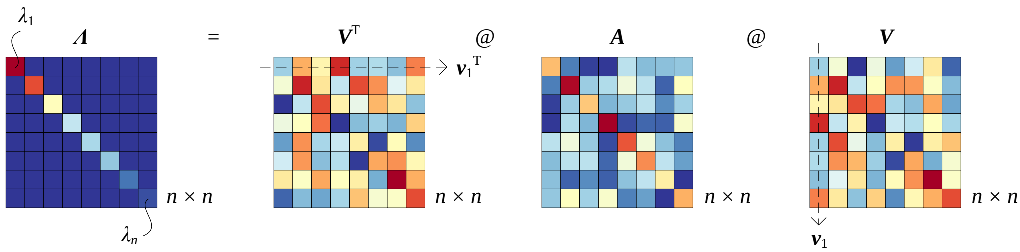

图 10. $n \times n$ 实对称矩阵 $\boldsymbol{A}$ 对角化

用矩阵乘法第一视角展开 (45) 得到

$$
\begin{aligned}
\boldsymbol{V}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{V} &= \begin{bmatrix} \boldsymbol{v}_1^{\mathrm{T}} \\ \boldsymbol{v}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{v}_n^{\mathrm{T}} \end{bmatrix} \boldsymbol{A} \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_2 & \cdots & \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_n \\ \boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_2 & \cdots & \boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}
\end{aligned} \tag{46}
$$

让我们先看 (46) 主对角线矩阵乘法。

以 $\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1$ 为例，由于 $\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1$，且 $\boldsymbol{v}_1$ 为单位向量，$\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1$ 可以写成

$$
\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_1 = \lambda_1\left\|\boldsymbol{v}_1\right\|_2^2 = \lambda_1 \tag{47}
$$

上式是瑞利商 (Rayleigh quotient) 的特殊形式，用来度量向量 $\boldsymbol{v}_1$ 在 $\boldsymbol{A}$ 作用下的缩放效果。

给定实数对称矩阵 $\boldsymbol{A}$，$\boldsymbol{A}$ 的瑞利商定义为：

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

本书后续将专门介绍瑞利商。

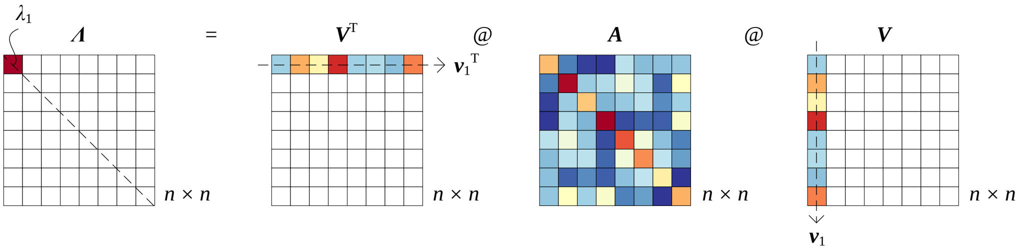

图 11. $n \times n$ 实对称矩阵 $\boldsymbol{A}$ 对角化，主对角线元素

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

$$
\boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1\boldsymbol{v}_2^{\mathrm{T}}\boldsymbol{v}_1 = 0 \tag{49}
$$

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

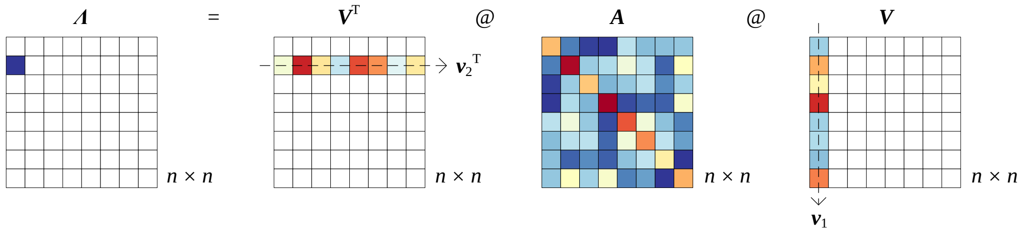

图 12． $n \times n$ 实对称矩阵 $\boldsymbol{A}$ 对角化，非主对角线元素

此外，由于 $\boldsymbol{A}$ 为实对称矩阵，(49) 还可以转置得到

$$\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_2 = \lambda_1\boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_2 = 0 \tag{50}$$

逆矩阵

只有所有特征值均非 0，实对称矩阵 $\boldsymbol{A}$ 才有逆，即

$$\boldsymbol{A}^{-1} = \left(\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\mathrm{T}}\right)^{-1} = \left(\boldsymbol{V}^{\mathrm{T}}\right)^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{V}^{-1} = \boldsymbol{V}\boldsymbol{\Lambda}^{-1}\boldsymbol{V}^{\mathrm{T}} = \boldsymbol{V}\begin{bmatrix} 1/\lambda_1 & & & \\ & 1/\lambda_2 & & \\ & & \ddots & \\ & & & 1/\lambda_n \end{bmatrix}\boldsymbol{V}^{\mathrm{T}} \tag{51}$$

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

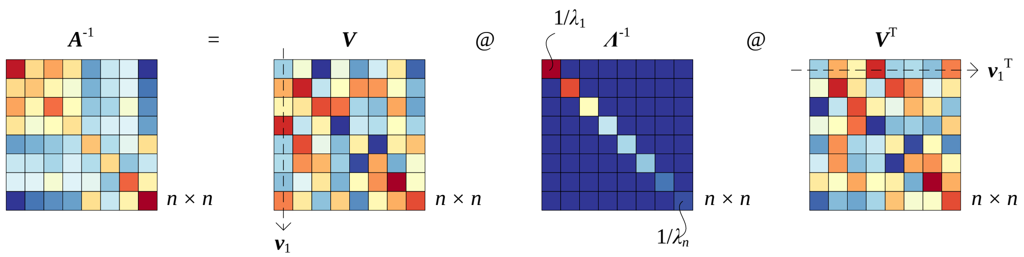

图 13． $n \times n$ 实对称矩阵 $\boldsymbol{A}^{-1}$ 谱分解

前文中，大家最常见到的对称矩阵是格拉姆矩阵。本章后续将专门讲解格拉姆矩阵。

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。