

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

11.3 非对称方阵 EVD



本节你将掌握的核心技能：

- ▶ 非对称矩阵的 EVD， V 不再是正交矩阵。
- ▶ V 与其逆矩阵 (如果存在的话) 的变换作用互为抵消，两者行列式乘积为 1。
- ▶ 本节的 V 拆解剪切与缩放操作。
- ▶ 本节的非对称矩阵 EVD 分解结果可视作“缩放 → 剪切 → 缩放 → 剪切 → 缩放”。
- ▶ 等比例缩放矩阵：所有方向都是特征方向。
- ▶ 不等比例缩放的几何意义：每个方向有不同的拉伸因子。
- ▶ 旋转矩阵：出现复数时通常无法实现实数对角化。
- ▶ 特征值为 0 意味着线性变换中存在降维。
- ▶ 负特征值意味着镜像。
- ▶ 存在特征值 0 和 1，方阵仍可能可对角化。

上一节中，我们看到实对称矩阵的特征值分解结果对应“旋转 → 缩放 → 旋转”；但是，很多非对称矩阵的特征值分解结果对应的几何操作就没那么显而易见了。本节用一个例子展示非对称方阵特征值分解 (EVD) 的几何操作。当然这也是很多情况中的一个，本节最后还会展示更多场景。

非对称矩阵的 EVD

给定如下非对称方阵 A

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

方阵 A 的特征值分解为

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

$$\mathbf{A} = \mathbf{V}\mathbf{A}\mathbf{V}^{-1} = \underbrace{\begin{bmatrix} 1 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 \end{bmatrix}}_{\mathbf{V}} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & \sqrt{2} \end{bmatrix}}_{\mathbf{V}^{-1}} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad (2)$$

上式中， \mathbf{V} 的列向量还是单位向量，但是不正交。

⚠ 注意，(2) 最右侧是 \mathbf{V} 的逆矩阵。由于 \mathbf{V} 不是正交矩阵， \mathbf{V} 的逆矩阵不能写成 \mathbf{V}^T 。

计算 \mathbf{V} 的格拉姆矩阵 $\mathbf{V}^T @ \mathbf{V}$

$$\mathbf{V}^T @ \mathbf{V} = \begin{bmatrix} 1 & 0 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 1 & -\sqrt{2}/2 \\ -\sqrt{2}/2 & 1 \end{bmatrix} \quad (3)$$

将 \mathbf{V} 写成 $[\mathbf{v}_1, \mathbf{v}_2]$ ，上述矩阵乘法可以展开写成

$$\mathbf{V}^T @ \mathbf{V} = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^T \mathbf{v}_1 & \mathbf{v}_1^T \mathbf{v}_2 \\ \mathbf{v}_2^T \mathbf{v}_1 & \mathbf{v}_2^T \mathbf{v}_2 \end{bmatrix} \quad (4)$$

比较 (3) 和 (4)，我们发现对角线元素为 1，意味着 \mathbf{v}_1 、 \mathbf{v}_2 为单位向量；显然 \mathbf{v}_1 、 \mathbf{v}_2 并非正交，本例中两者夹角为 135 度。

而 \mathbf{V} 的逆矩阵为

$$\mathbf{V}^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & \sqrt{2} \end{bmatrix} \quad (5)$$

我们发现， \mathbf{V}^{-1} 的列向量并非都是单位向量。

可对角化

值得“庆幸”的是，(2) 中 \mathbf{V} 的列向量线性无关，方阵 \mathbf{V} 满秩；也就是说， \mathbf{V} 可逆。

这意味着 \mathbf{A} 可对角化，即

$$\mathbf{A} = \mathbf{V}^{-1} \mathbf{A} \mathbf{V} = \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & \sqrt{2} \end{bmatrix}}_{\mathbf{V}^{-1}} \underbrace{\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} 1 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 \end{bmatrix}}_{\mathbf{V}} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad (6)$$

当一个矩阵被成功对角化后，许多与矩阵相关的计算问题都会显著简化。

例如，行列式的计算就不再需要繁复的展开式，而是可以直接通过对角方阵的主对角线上元素的乘积得到。比如，矩阵 \mathbf{A} 的行列式值为其特征值乘积：

$$\det(\mathbf{A}) = \prod_{j=1}^D \lambda_j \quad (7)$$

除了行列式之外，对角化还极大地方便了矩阵的幂运算。

如果 \mathbf{A} 可对角化，矩阵 \mathbf{A} 的平方可以写成：

$$A^2 = VAV^{-1}VAV^{-1} = VA^2V^{-1} = V \begin{bmatrix} (\lambda_1)^2 & & \\ & (\lambda_2)^2 & \\ & & \ddots \\ & & & (\lambda_D)^2 \end{bmatrix} V^{-1} \quad (8)$$

类似地， A 的 n 次幂可以写成：

$$A^n = VAV^{-1} \dots VAV^{-1} = VA^nV^{-1} = V \begin{bmatrix} (\lambda_1)^n & & \\ & (\lambda_2)^n & \\ & & \ddots \\ & & & (\lambda_D)^n \end{bmatrix} V^{-1} \quad (9)$$

下面还是用几何视角帮助我们理解 A 以及 A 的特征值分解对应的几何操作。

非对称矩阵对 e_1 、 e_2 作用

先看矩阵 A 对单位向量的几何操作。

如图 1 所示，非对称矩阵 A 对 e_1 作用对应矩阵乘法 Ae_1

$$Ae_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \quad (10)$$

非对称矩阵 A 对 e_2 作用对应矩阵乘法 Ae_2

$$Ae_2 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (11)$$

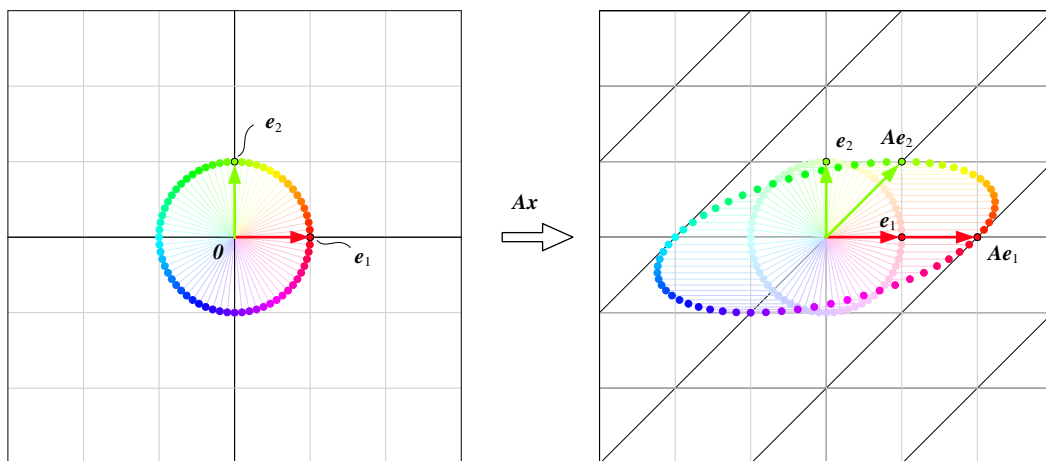
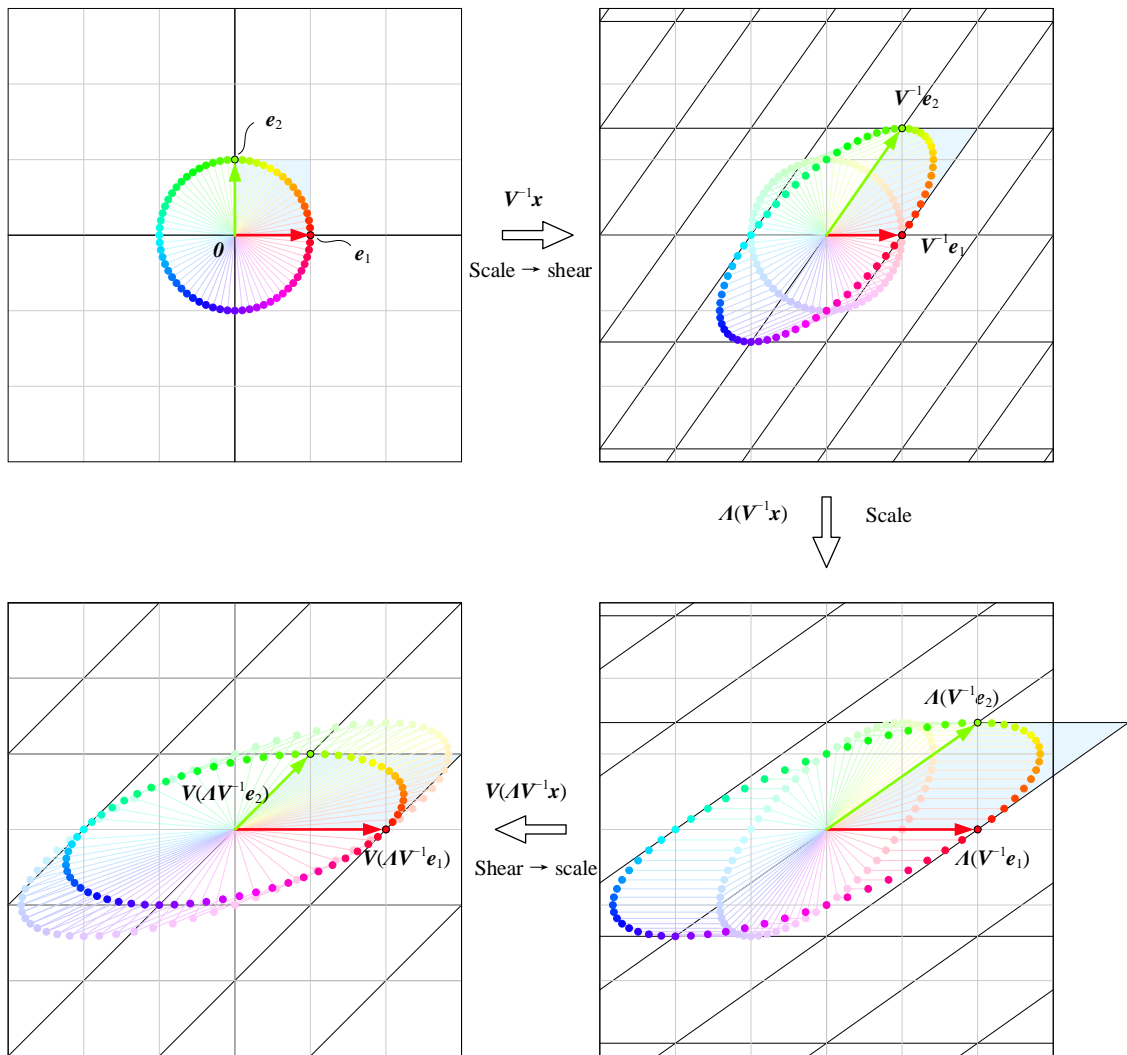


图 1. 矩阵 A 对 e_1 、 e_2 的线性变换

图 2 所示为矩阵连乘 VAV^{-1} 对 e_1 、 e_2 的分步线性变换。

对角方阵 A 的作用还是缩放，这一点和谱分解中的对角方阵作用完全一致。

我们看到 V^{-1} 和 V 对图形的作用似乎是“剪切 + 缩放”，下面让我展开分析这一点。

图 2. 矩阵连乘 VAV^{-1} 对 e_1, e_2 的线性变换

拆解 V

仔细观察 V , V 可以拆解成两个矩阵的乘积

$$V = \begin{bmatrix} 1 & -\sqrt{2}/2 \\ 0 & \sqrt{2}/2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{2}/2 \end{bmatrix}}_{\text{Scale, } S} @ \underbrace{\begin{bmatrix} 1 & -\sqrt{2}/2 \\ 0 & 1 \end{bmatrix}}_{\text{Shear, } K} = S @ K \quad (12)$$

K 是沿横轴剪切

$$K = \begin{bmatrix} 1 & -\sqrt{2}/2 \\ 0 & 1 \end{bmatrix} \quad (13)$$

S 是缩放矩阵

$$S = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{2}/2 \end{bmatrix} \quad (14)$$

准确来说, 上式是沿纵轴缩放。

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>

本书配套微课视频均发布在 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

(12) 中 $Vx = SKx$ 对列向量 x 几何操作的先后顺序是，先剪切 (K)，再缩放 (S)。

V^{-1} 对 e_1 、 e_2 作用

这样的话，我们可以把 V^{-1} 写成

$$V^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & \sqrt{2} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \sqrt{2}/2 \\ 0 & 1 \end{bmatrix}}_{\text{Shear, } K^{-1}} @ \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{2} \end{bmatrix}}_{\text{Scale, } S^{-1}} = K^{-1}S^{-1} \quad (15)$$

如图 3 所示， V^{-1} 的作用相当于先用 S^{-1} 缩放，再用 K^{-1} 剪切。

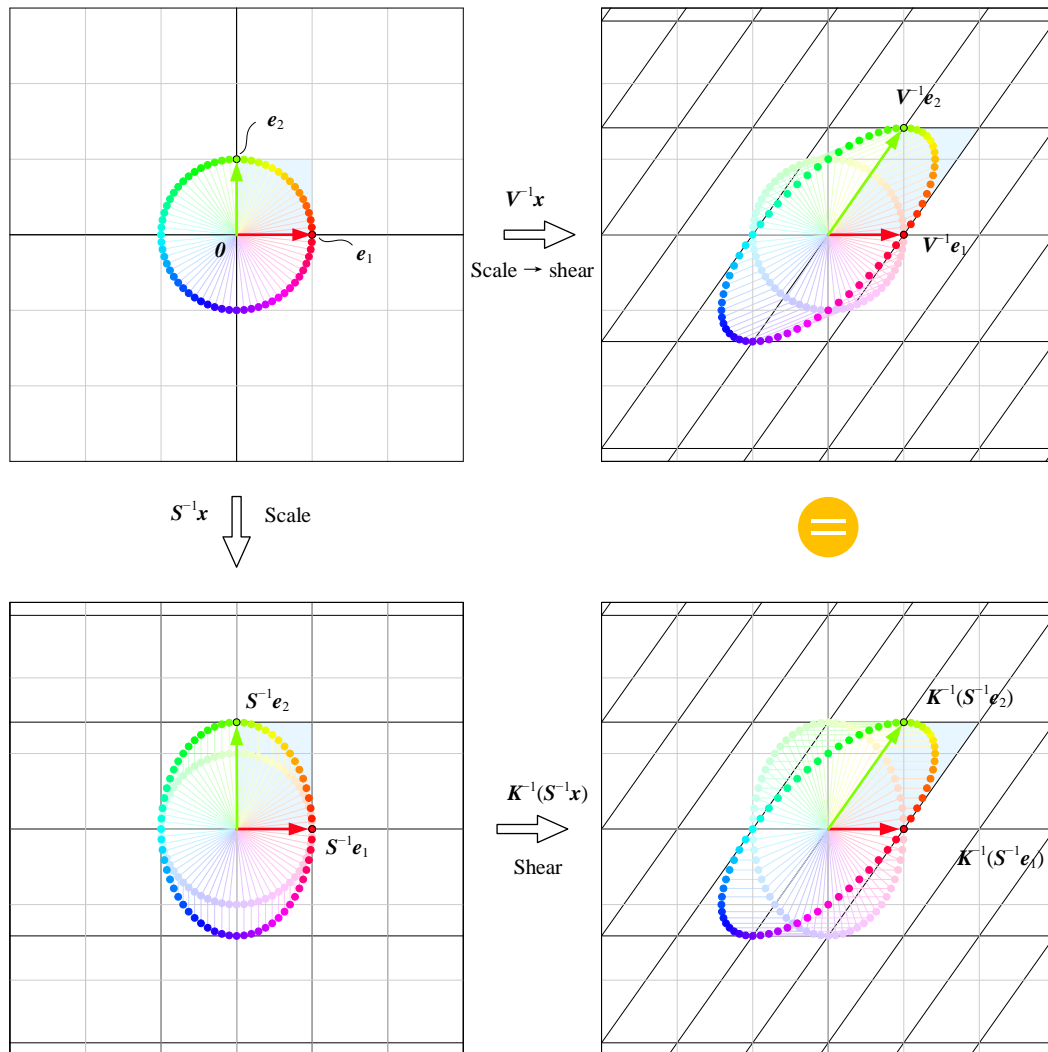


图 3. 矩阵 V^{-1} 对 e_1 、 e_2 的线性变换，拆解成先缩放 (S^{-1})、再剪切 (K^{-1})

V 逆矩阵 V^{-1} 的行列式为

$$\det(V^{-1}) = \det\left(\begin{bmatrix} 1 & 1 \\ 0 & \sqrt{2} \end{bmatrix}\right) = \sqrt{2} \quad (16)$$

这意味着图 3 中平行四边形网格面积放大 $\sqrt{2}$ 倍。显然，剪切矩阵 \mathbf{K}^{-1} 不会让图形面积发生变化，缩放矩阵 \mathbf{S}^{-1} 才是导致图形面积变化的几何操作。

\mathbf{V} 的作用

如图 4 所示， \mathbf{V} 的作用则和 \mathbf{V}^{-1} 相反，先剪切 (\mathbf{K})、再缩放 (\mathbf{S})。

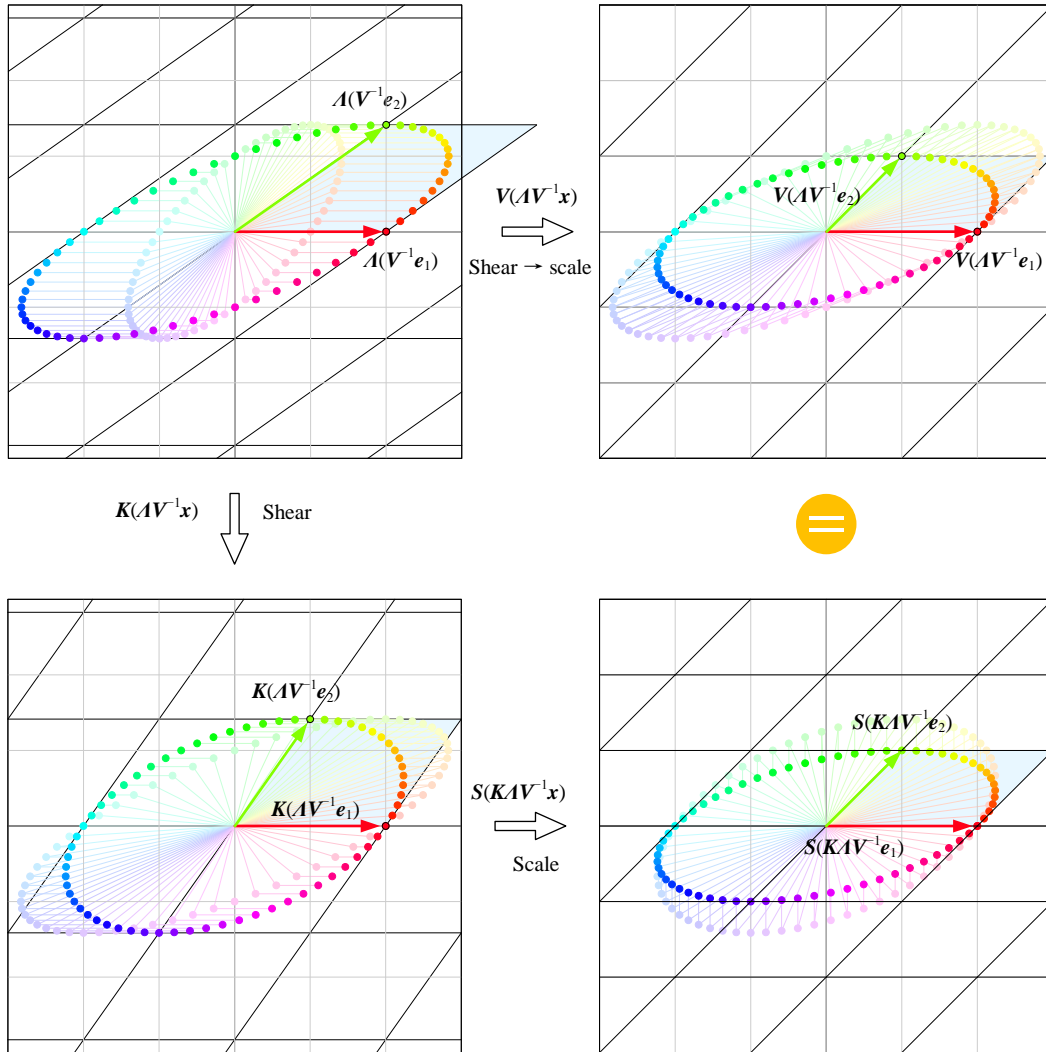


图 4. 矩阵 \mathbf{V} 的线性变换，拆解成先剪切 (\mathbf{K})、再缩放 (\mathbf{S})

矩阵 \mathbf{V} 的行列式为

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

这意味着图形面积缩小为原来的 $\sqrt{2}/2$ 。

(17)、(16) 这两个行列式的乘积为 1。本书前文提过，原矩阵与逆矩阵的行列式的乘积恒等于 1。

这意味着， \mathbf{V}^{-1} 、 \mathbf{V} 对于面积的影响相互抵消，即

$$\det(\mathbf{V}^{-1}) \det(\mathbf{V}) = 1 \quad (18)$$

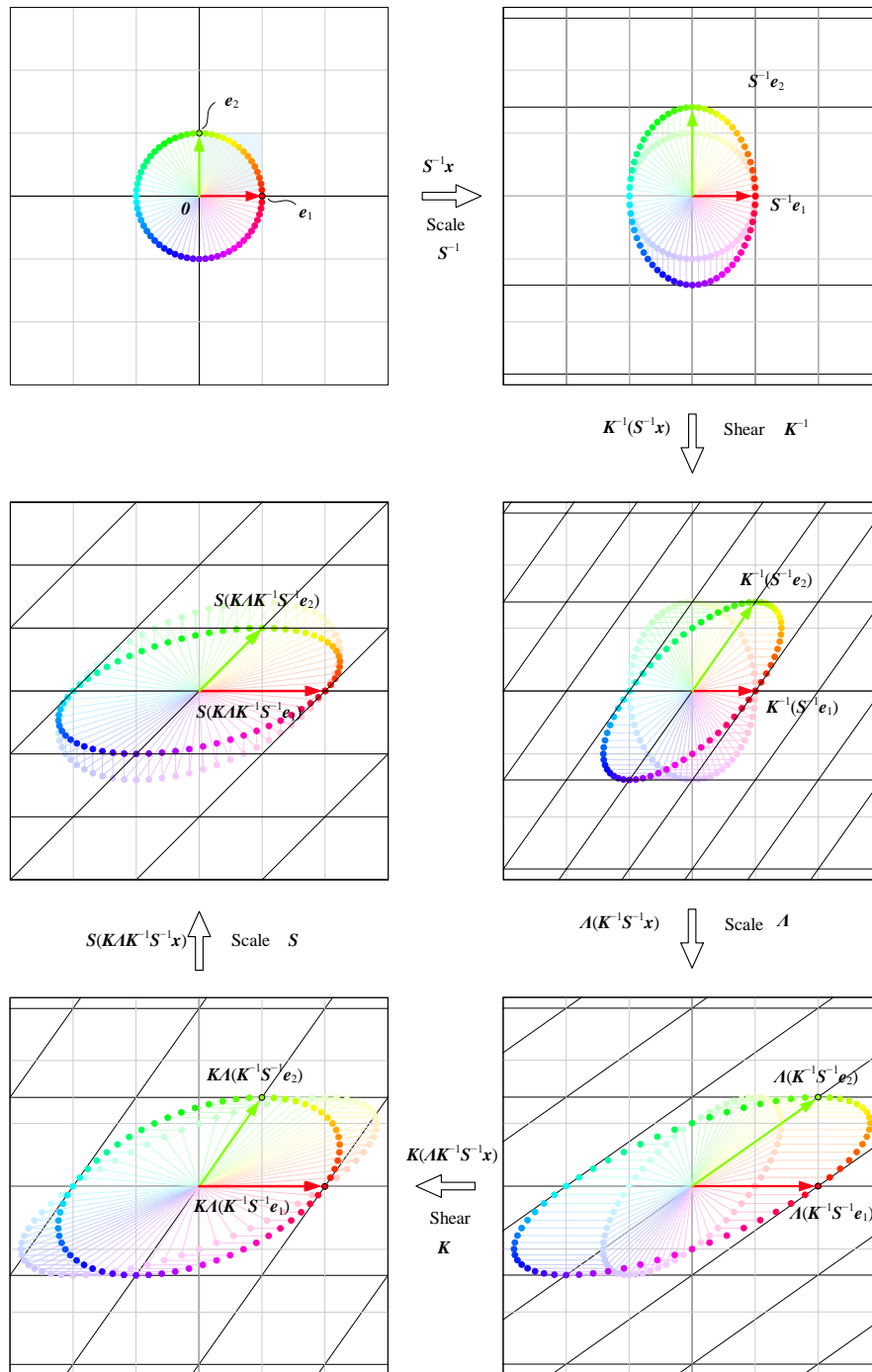
A 线性变换导致面积变化来自 $\det(A)$ ，即

$$\det(A) = \det(A) \quad (19)$$

缩放 → 剪切 → 缩放 → 剪切 → 缩放

也就是说矩阵 A 的特征值分解可以拆解成这样一系列动作“缩放 (S^{-1}) → 剪切 (K^{-1}) → 缩放 (A) → 剪切 (K) → 缩放 (S)”，具体如图 5 所示。

$$A = VAV^{-1} = \underbrace{SKA}_{\substack{V \\ V^{-1}}} \underbrace{K^{-1}S^{-1}}_{V^{-1}} \quad (20)$$

图 5. “缩放 → 剪切 → 缩放 → 剪切 → 缩放”对 e_1 、 e_2 的线性变换

非对称矩阵对 v_1 、 v_2 作用

下面让我们再看非对称矩阵 A 对特征向量 v_1 、 v_2 的几何操作。

如图 6 所示，非对称矩阵 A 对特征向量 v_1 作用对应矩阵乘法 Av_1

$$Av_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 2 \times_{\lambda_1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (21)$$

v_1

非对称矩阵 A 对特征向量 v_2 作用对应矩阵乘法 Av_2

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

$$Av_2 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \underbrace{1}_{\lambda_2} \times \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}_{v_2} \quad (22)$$

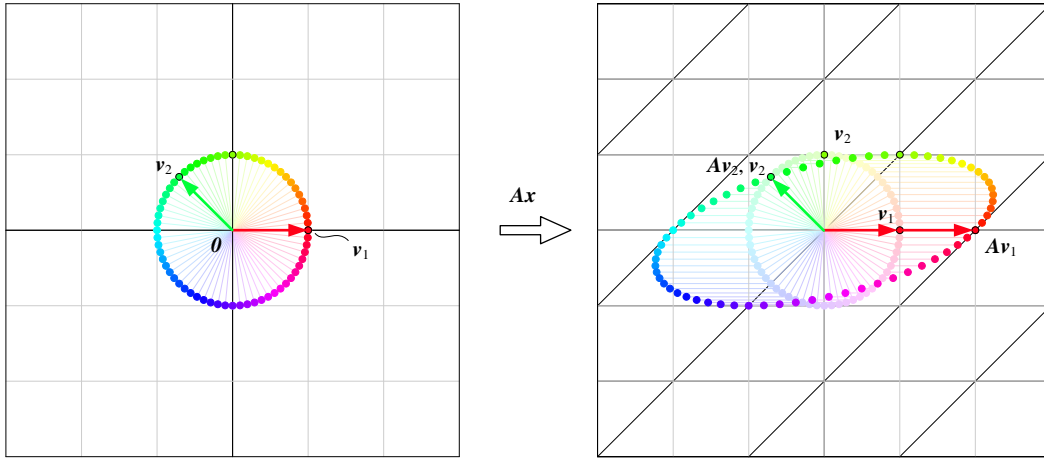
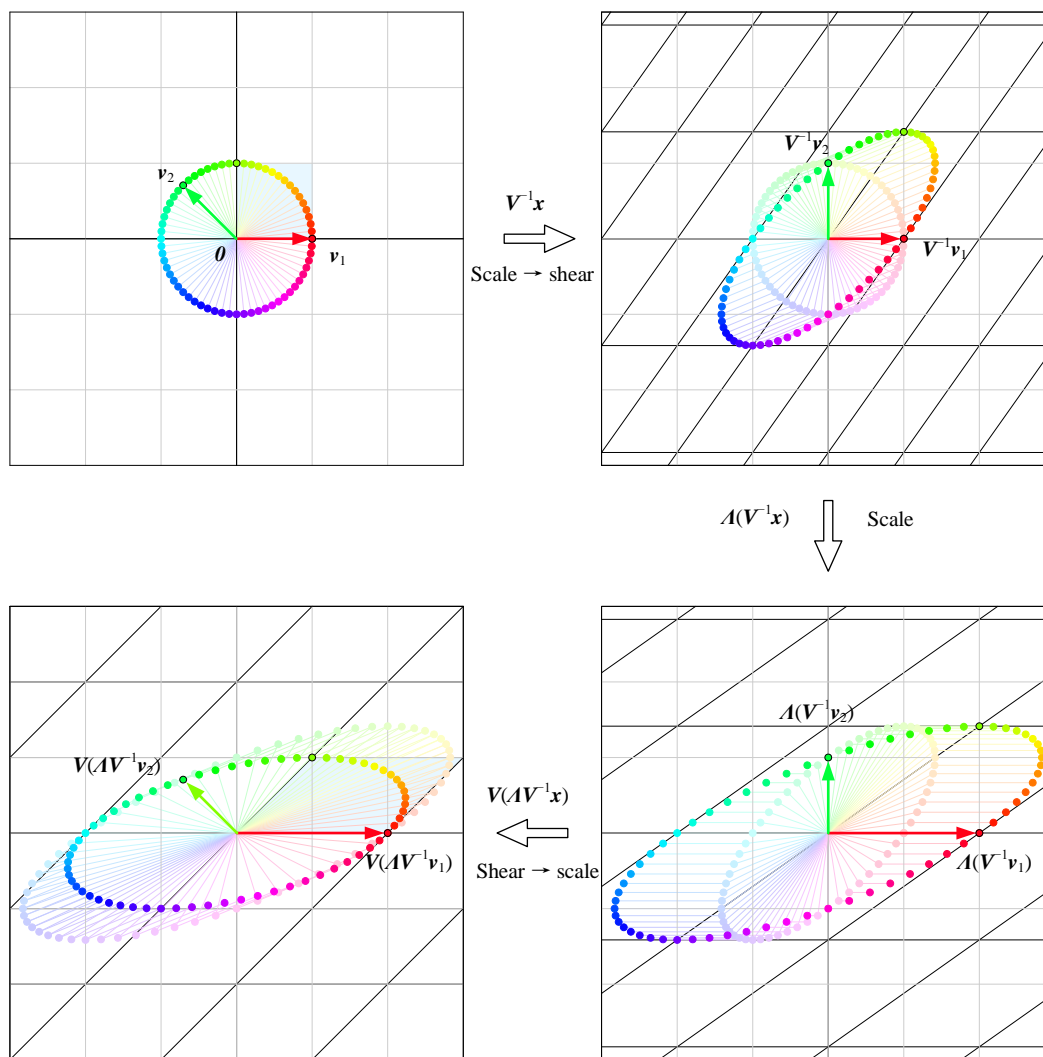
图 6. 矩阵 A 对 v_1 、 v_2 的线性变换

图 7 所示为矩阵连乘 VAV^{-1} 对 v_1 、 v_2 的分步线性变换。

? 请大家自己分析图 7 的分步操作。

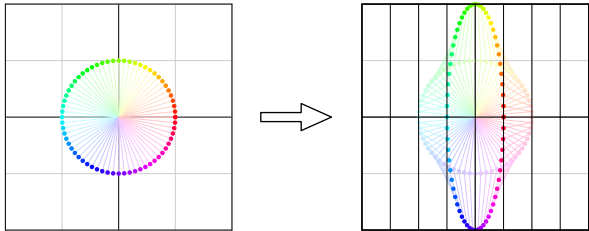
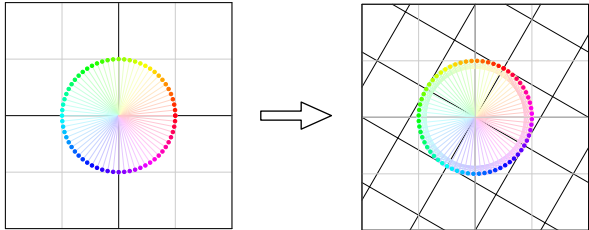
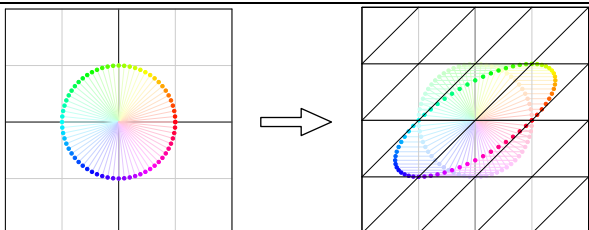
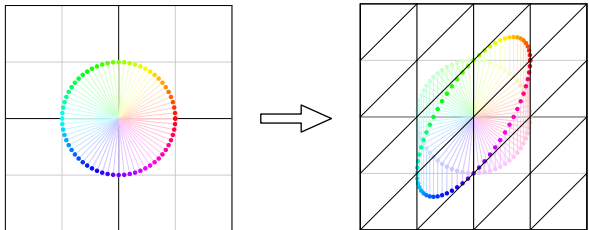
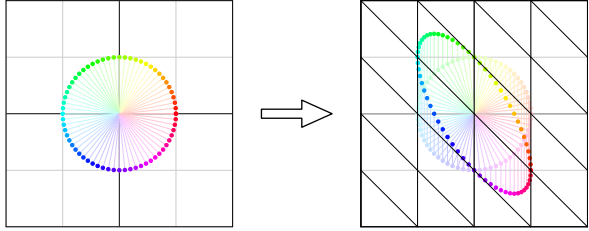
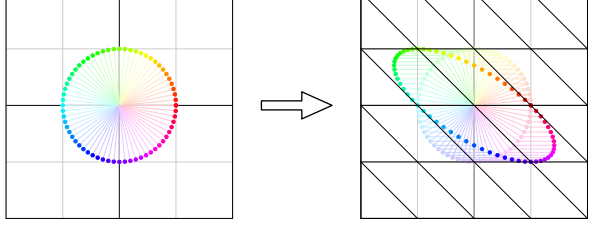
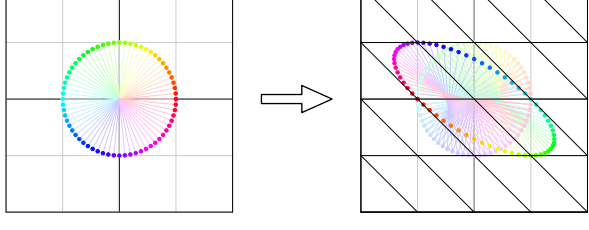
图 7. 矩阵连乘 VAV^{-1} 对 v_1 、 v_2 的线性变换

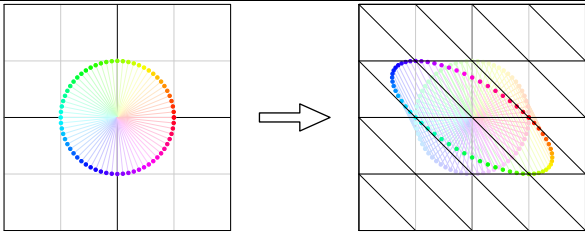
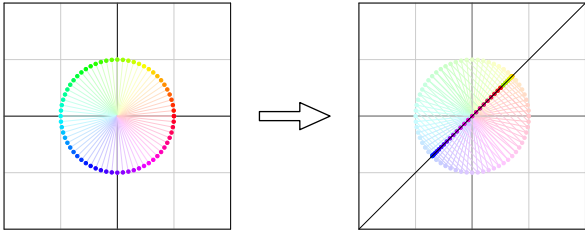
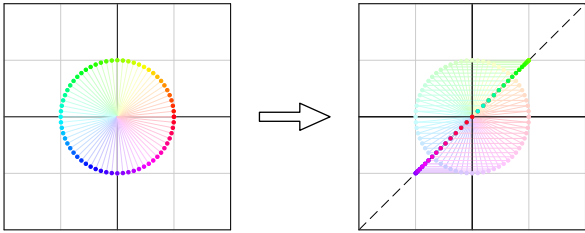
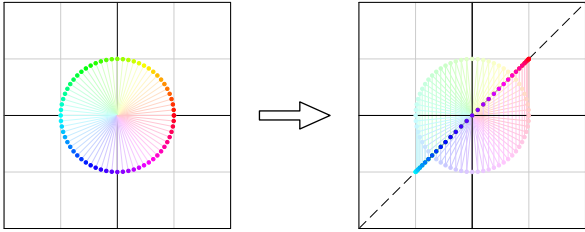
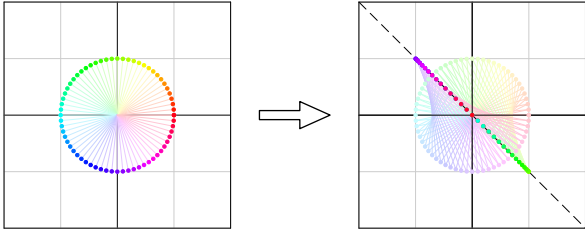
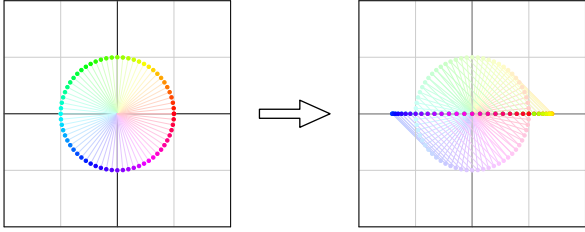
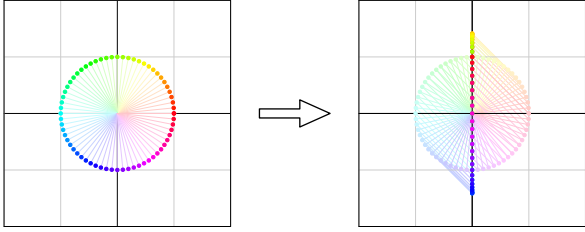
更多特征值分解

以上讲述的仅仅是非对称方阵特征值分解几何视角的一个案例。特征值分解还会遇到很多有趣的结果，请大家逐一分析表 1。

表 1. 更多特征值分解

矩阵 A	特征值, 特征向量	几何变换
$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ <p>可以对角化 (已经是对角方阵)</p>	$\lambda_{1,2} = 2$ $v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 注意: 任何非零向量都可以作为特征向量	

$\begin{bmatrix} 1/2 & 0 \\ 0 & 2 \end{bmatrix}$ <p>可以对角化 (已经是对角方阵)</p>	$\lambda_1 = \frac{1}{2}, \lambda_2 = 2$ $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	
$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ <p>如果 $\sin \theta$ 为 0, 特征值为实数。矩阵可以对角化, 特征值为实数正负 1, 有两个线性无关的特征向量</p>	$\lambda_{1,2} = \cos \theta \pm i \sin \theta$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ +i \end{bmatrix}$	
$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ <p>不可以对角化</p>	$\lambda_{1,2} = 1$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$	
$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ <p>不可以对角化</p>	$\lambda_{1,2} = 1$ $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$	
$\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$ <p>不可以对角化</p>	$\lambda_{1,2} = 1$ $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	
$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$ <p>不可以对角化</p>	$\lambda_{1,2} = 1$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$	
$\begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}$ <p>不可以对角化</p>	$\lambda_{1,2} = -1$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$	

$\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$ 可以对角化	$\lambda_1 = 1, \lambda_2 = -1$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\sqrt{5}/5 \\ 2\sqrt{5}/5 \end{bmatrix}$	
$\begin{bmatrix} 1 & 1/2 \\ 2 & 2 \end{bmatrix}$ 可以对角化	$\lambda_1 = 0, \lambda_2 = 1$ $\mathbf{v}_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	
$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ 可以对角化	$\lambda_1 = 0, \lambda_2 = 1$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	
$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ 可以对角化	$\lambda_1 = 0, \lambda_2 = 1$ $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	
$\begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}$ 可以对角化	$\lambda_1 = 0, \lambda_2 = -1$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	
$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ 可以对角化	$\lambda_1 = 1, \lambda_2 = 0$ $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	
$\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ 可以对角化	$\lambda_1 = 1, \lambda_2 = 0$ $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$	

$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 可以对角化	$\lambda_1 = 2, \lambda_2 = 0$ $\mathbf{v}_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	
$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ 可以对角化	$\lambda_1 = 0, \lambda_2 = -2$ $\mathbf{v}_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	
$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 可以对角化	$\lambda_1 = 1, \lambda_2 = -1$ $\mathbf{v}_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$	



请大家用 DeepSeek/ChatGPT 等工具完成本节如下习题。

Q1. 请大家把表 1 所有方阵全部手算、Python 编程完成特征值分解，并逐个分析结果。分析表 1 每个案例的时候，请大家格外注意以下几方面

- ▶ 方阵是否可逆 (行列式是否不为 0)；
- ▶ 特征方程是什么；
- ▶ 特征值的特点，是否为 0？是否为负数？是否出现复数？
- ▶ 特征向量是否线性相关；如果线性无关，则 \mathbf{V} 可逆 (方阵可对角化)。