

4.2. 特征值与特征向量

上一节我们知道线性变换在不同基下的矩阵相似. 在计算矩阵幂等问题时, 可以将矩阵 $\mathbf{A}$ 的问题转化为其相似矩阵的问题.

与方阵 $\mathbf{A}$ 相似的矩阵中, 哪个矩阵 “最简单”?

“最简单” 的矩阵莫过于对角矩阵, 如果 $\mathbf{A}$ 相似于对角矩阵, 那么我们称 $\mathbf{A}$ 可对角化(diagonalizable).

并不是所有的矩阵都可对角化! 例如:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

不可对角化! (可待定系数法验证)

假设 n 阶矩阵 $\mathbf{A} \in M_n(\mathbb{R})$ 可对角化,

即存在可逆矩阵 $\mathbf{P} \in M_n(\mathbb{R})$ 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ 为对角矩阵.

$\mathbf{P}$ 可逆, 其列向量 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_n$ 线性无关, 组成 $\mathbb{R}^n$ 的一组基.

假设 $\mathbf{D} = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n]$, 那么 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$, 即 $\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D}$, 也就意味着

$$(\mathbf{A}\boldsymbol{\eta}_1, \mathbf{A}\boldsymbol{\eta}_2, \dots, \mathbf{A}\boldsymbol{\eta}_n) = (\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

这等价于

$$\mathbf{A}\boldsymbol{\eta}_i = \lambda_i \boldsymbol{\eta}_i, i = 1, 2, \dots, n.$$

矩阵 $\mathbf{A}$ 可对角化当且仅当存在 $\mathbb{R}^n$ 的一组基 $\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \dots, \boldsymbol{\eta}_n$, 以及 $\mathbb{R}$ 中 n 个数 $\lambda_1, \lambda_2, \dots, \lambda_n$, 使得

$$\mathbf{A}\boldsymbol{\eta}_i = \lambda_i\boldsymbol{\eta}_i, i = 1, 2, \dots, n$$

即, 这组基中的向量在 $\mathbf{A}$ 的作用下只是变为自身的数量倍.

定义

假设 σ 为 $\mathbb{R}$ 向量空间 V 上的线性变换, 如果 V 中的非零向量 $\boldsymbol{\alpha}$ 满足

$$\sigma(\boldsymbol{\alpha}) = \lambda\boldsymbol{\alpha},$$

其中 $\lambda \in \mathbb{R}$. 那么 $\boldsymbol{\alpha}$ 称为 σ 的特征向量(eigenvector), 而 λ 称为特征向量 $\boldsymbol{\alpha}$ 对应的特征值(eigenvalue).

定理 (4.5)

$\mathbf{A} \in M_n(\mathbb{R})$ 可对角化当且仅当 $\mathbb{R}^n$ 中存在 n 个线性无关的 $\mathbf{A}$ 的特征向量.

怎样才能找出 $\mathbf{A}$ 的所有特征值和特征向量呢?

假设 $\boldsymbol{\eta} \in \mathbb{R}^n$ 为 $\mathbf{A}$ 的特征向量, 对应的特征值为 λ , 则

$$\mathbf{A}\boldsymbol{\eta} = \lambda\boldsymbol{\eta}, \quad \text{即 } (\lambda\mathbf{I} - \mathbf{A})\boldsymbol{\eta} = \mathbf{0}.$$

因此, $\boldsymbol{\eta}$ 为齐次线性方程组

$$(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$$

的非零解, 这个方程组有非零解的充要条件是 $|\lambda\mathbf{I} - \mathbf{A}| = 0$.

命题 (4.6)

对矩阵 $\mathbf{A} \in M_n(\mathbb{R})$, $\lambda \in \mathbb{R}$ 是 $\mathbf{A}$ 的特征值当且仅当 $|\lambda \mathbf{I} - \mathbf{A}| = 0$.

这个命题把 $\mathbf{A}$ 的特征值限定在了一个比较小的范围内. 考虑矩阵

$$t\mathbf{I} - \mathbf{A} = \begin{pmatrix} t - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & t - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & t - a_{nn} \end{pmatrix},$$

将 $t\mathbf{I} - \mathbf{A}$ 看作元素为关于 t 的多项式的矩阵, 并求其行列式.

$$f_{\mathbf{A}}(t) := \det(t\mathbf{I} - \mathbf{A}),$$

称为 $\mathbf{A}$ 的**特征多项式**(characteristic polynomial). 根据命题 (4.6), $\lambda \in \mathbb{R}$ 是 $\mathbf{A}$ 的特征值当且仅当 λ 是 $\mathbf{A}$ 的特征多项式 $f_{\mathbf{A}}(t)$ 的根. 由此可见 $\mathbf{A}$ 的特征值最多有 n 个.

对于 $\lambda \in \mathbb{R}$, 我们定义

$$V_\lambda := \{\alpha \in \mathbb{R}^n \mid A\alpha = \lambda\alpha\}.$$

当 λ 为 A 的特征值时, 即 $V_\lambda \neq \{\mathbf{0}\}$, 称 V_λ 为 A 的特征值 λ 对应的特征子空间(eigenspace).

例

找出下面矩阵的所有特征值和相应的特征向量.

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

解: 先算出 $\mathbf{A}$ 的特征多项式

$$f_{\mathbf{A}}(t) = \begin{vmatrix} t-1 & -2 & 0 \\ 0 & t-1 & 0 \\ 0 & 0 & t-2 \end{vmatrix} = (t-1)^2(t-2).$$

因此, $\mathbf{A}$ 的特征值是 $\lambda_1 = 1$, $\lambda_2 = 2$.

解 $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ 可以得到特征子空间 $V_1 = \langle \mathbf{e}_1 \rangle$.

解 $(2\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ 可得 $V_2 = \langle \mathbf{e}_3 \rangle$.

所有特征向量都在 V_1 或者 V_2 中, 任意 3 个特征向量线性相关.
因此, $\mathbf{A}$ 没有 3 个线性无关的特征向量, 不能对角化!

由行列式的定义, n 阶矩阵的行列式可以写成 $n!$ 项的和, 不计符号的情况下, 其中每项都是矩阵中互不同行且不同列的元素乘积. 而 $t\mathbf{I} - \mathbf{A}$ 的行列式中, 除了

$$(t - a_{11})(t - a_{22}) \cdots (t - a_{nn})$$

这一项外, 其他项中至少有两个元素不是对角线上的元素 (行列式中, 不可能有一项有 $n - 1$ 个对角线元素). 因此

$$f_{\mathbf{A}}(t) = (t - a_{11})(t - a_{22}) \cdots (t - a_{nn}) + g(t),$$

其中 $g(t)$ 是次数不超过 $n - 2$ 的多项式. 这样, 我们发现 $f_{\mathbf{A}}(t)$ 的前两项为

$$t^n - (a_{11} + a_{22} + \cdots + a_{nn})t^{n-1}.$$

有趣的是, t^{n-1} 系数是 A 对角线元素之和的相反数. 为了方便, 我们记

$$\text{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{nn},$$

称为 $\mathbf{A}$ 的迹(trace).

另外, $f_{\mathbf{A}}(0) = |-\mathbf{A}| = (-1)^n |\mathbf{A}|$, 即 $f_{\mathbf{A}}(t)$ 的常数项为 $(-1)^n |\mathbf{A}|$.

命题 (4.7)

两个相似的矩阵具有相同的特征多项式.

证明. 假设 $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, 则

$$f_{\mathbf{B}}(t) = |t\mathbf{I} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}| = |\mathbf{P}^{-1}(t\mathbf{I} - \mathbf{A})\mathbf{P}| = |\mathbf{P}^{-1}| \cdot |t\mathbf{I} - \mathbf{A}| \cdot |\mathbf{P}| = f_{\mathbf{A}}(t).$$

推论 (4.8)

两个相似的矩阵具有相同的特征值、迹、行列式和秩.

如果两个矩阵的迹、行列式或者秩不同, 那么它们不可能相似.