上海师范大学
Shanghai Normal University

《线性代数》

12-特征值与特征向量 (Eigenvalues and Eigenvectors)

杨启哲

上海师范大学信机学院计算机系

2024 年 5 月 17 日

复习-行列式和矩阵乘法、转置矩阵的关系

定理 1 **[Transpose].**

令 A 是一个 $n \times n$ 的矩阵，我们有：

$$\det(A) = \det(A^{\mathsf{T}})$$

定理 2 **[Product of Determinants].**

令 A 是一个 $n \times n$ 的矩阵，B 是一个 $n \times n$ 的矩阵，我们有：

$$\det(AB) = \det(A)\det(B)$$

这也说明，对矩阵的初等行变换对行列式的影响跟初等列变换是完全相同的

复习-行列式的定义

定义 3.

令 A 是一个 $n \times n$ 的矩阵，我们定义：

$$\det(A) = \sum_{\sigma \in \mathrm{Perm}(n)} (-1)^{\tau(\sigma)} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

定理 4.

$\det(A)$ 满足下列性质：

- $D(e_1, \cdots, e_n) = 1$
- $D(a_1, \cdots, a_i, \cdots, a_j, \cdots a_n) = -D(a_1, \cdots, a_j, \cdots, a_i, \cdots a_n)$
- $D(a_1, \cdots, ca_i + da_i', \cdots, a_n) = cD(a_1, \cdots, a_i, \cdots, a_n) + dD(a_1, \cdots, a_i', \cdots, a_n)$

上海师范大学 Shanghai Normal University

定理 5 **[Uniqueness].**

$\det(A)$ 是唯一一个 $\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n\text{个}} \to \mathbb{R}^n$ 的函数满足下述三个性质:

- $D(e_1, \cdots, e_n) = 1$
- $D(a_1, \cdots, a_i, \cdots, a_j, \cdots a_n) = -D(a_1, \cdots, a_j, \cdots, a_i, \cdots a_n)$
- $D(a_1, \cdots, ca_i + da_i', \cdots, a_n) = cD(a_1, \cdots, a_i, \cdots, a_n) + dD(a_1, \cdots, a_i', \cdots, a_n)$

令 A 是一个 $n \times n$ 的矩阵，并且 $n \geqslant 2$。对于任意的 $i, j \in [n]$，我们定义：

M_{ij} 是将 A 第 i 行第 j 列的元素删去后得到的 $n-1 \times n-1$ 的矩阵。

定义 6 **[Cofactor].**

下列数：

$$C_{ij} = (-1)^{i+j} \det(M_{ij})$$

称之为代数余子式 (Cofactor)，特别的 $\det(M_{ij})$ 称之为余子式。

Shanghai Normal University

定理 7.

1. $\det(\begin{bmatrix} a \end{bmatrix}) = a.$
2. 对于任意的 $n \geqslant 2$，并且 A 是一个 $n \times n$ 的矩阵:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

我们有:

$$\det(A) = a_{11}C_{11} + \cdots + a_{1n}C_{1n}$$

上海师范大学 Shanghai Normal University

定理 8.

1. $\det(\begin{bmatrix} a \end{bmatrix}) = a$.
2. 对于任意的 $n \geqslant 2$，并且 A 是一个 $n \times n$ 的矩阵:

$$\begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots & \ddots & \vdots \\ a_{11} & \cdots & a_{ni} & \cdots & a_{nn} \end{bmatrix}$$

我们有:

$$\det(A) = a_{1i}C_{1i} + \cdots + a_{1i}C_{1i}$$

上海师范大学 Shanghai Normal University

定理 9.

1. $\det(\begin{bmatrix} a \end{bmatrix}) = a.$
2. 对于任意的 $n \geqslant 2$，并且 A 是一个 $n \times n$ 的矩阵:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

我们有:

$$\det(A) = a_{i1}C_{i1} + \cdots + a_{in}C_{in}$$

方程组求解

对于任意的 $i \in [n]$，令 A_i 是将 A 的第 i 列替换成 b 后的矩阵，则方程组 $Ax = b$ 的解可以表示为:

$$x = \begin{bmatrix} \frac{\det(A_1)}{\det(A)} \\ \vdots \\ \frac{\det(A_n)}{\det(A)} \end{bmatrix}$$

求逆矩阵

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix},$$

其中 $\begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix}$ 称作伴随矩阵 A^*。

复习-关于行列式的总结

- 行列式跟首元的关系:

$$\det(A) = \det(U) = p_1 p_2 \cdots p_n \text{ 或者 } \det(A) = -\det(U) = -p_1 p_2 \cdots p_n$$

- 行列式的正式定义，The Big Formula:

$$\det(A) = \sum_{\sigma \in Perm(n)} (-1)^{\tau(\sigma)} a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

- 代数余子式，行列式按行 (列) 展开:

$$\det(A) = a_{1i} C_{1i} + \cdots + a_{1i} C_{1i}$$

- 一个应用：$Cramer's\ Rule$，用行列式解方程，求矩阵的逆，伴随矩阵。

主要内容

Shanghai Normal University

- 特征值介绍 (Introduction to Eigenvalues)
- 对角化矩阵 (Diagonalizing a Matrix)

特征值介绍 (Introduction to Eigenvalues)

我们知道斐波那契数列：$0, 1, 1, 2, 3, 5, 8, 13, 21, \cdots$,，即我们可以有如下的表示：

$$F_0 = 0,\ F_1 = 1,\ F_{k+2} = F_{k+1} + F_k$$

我们设法求这个递推式，我们可以重写这个式子：

$$F_{k+2} + aF_{k+1} = c(F_{k+1} + aF_k)$$

其中：

$$ac = 1$$

$$c - a = 1$$

解上述方程得：

$$\begin{cases} a & = \frac{-1+\sqrt{5}}{2} \\ c & = \frac{1+\sqrt{5}}{2} \end{cases} \quad \text{或} \quad \begin{cases} a & = \frac{-1-\sqrt{5}}{2} \\ c & = \frac{1-\sqrt{5}}{2} \end{cases}$$

即我们有:

$$F_{k+2}+\frac{-1+\sqrt{5}}{2}F_{k+1}=\frac{1+\sqrt{5}}{2}(F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k)$$

$$F_{k+2}+\frac{-1-\sqrt{5}}{2}F_{k+1}=\frac{1-\sqrt{5}}{2}(F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k)$$

从而我们有:

$$F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k=(\frac{1+\sqrt{5}}{2})^k(F_1+\frac{-1+\sqrt{5}}{2}F_0)$$

$$F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k=(\frac{1-\sqrt{5}}{2})^k(F_1+\frac{-1-\sqrt{5}}{2}F_0)$$

两式相减:

$$F_k=\frac{1}{\sqrt{5}}((\frac{1+\sqrt{5}}{2})^k-(\frac{1-\sqrt{5}}{2})^k)$$

斐波那契数列的矩阵形式 (I)

现在我们用矩阵得过程来探索一下。由:

$$F_{k+1} = F_k + F_{k-1}$$

我们有:

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$$

从而我们有:

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix} = \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right)^2 \begin{bmatrix} F_{k-1} \\ F_{k-2} \end{bmatrix} = \cdots = \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}$$

另一方面:

$$F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k=\frac{1+\sqrt{5}}{2}(F_k+\frac{-1+\sqrt{5}}{2}F_{k-1})$$
$$F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k=\frac{1-\sqrt{5}}{2}(F_k+\frac{-1-\sqrt{5}}{2}F_{k-1})$$

我们有:

$$\begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix}=\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} F_k+\frac{-1+\sqrt{5}}{2}F_{k-1} \\ F_k+\frac{-1-\sqrt{5}}{2}F_{k-1} \end{bmatrix}$$

从而通过类似的讨论我们可以得到:

$$\begin{aligned}\begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix}&=\left(\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}\right)^k\begin{bmatrix} F_1+\frac{-1+\sqrt{5}}{2}F_0 \\ F_1+\frac{-1-\sqrt{5}}{2}F_0 \end{bmatrix}\\ &=\begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{bmatrix}\begin{bmatrix} F_1+\frac{-1+\sqrt{5}}{2}F_0 \\ F_1+\frac{-1-\sqrt{5}}{2}F_0 \end{bmatrix}\end{aligned}$$

而注意到:

$$\begin{bmatrix} F_1+\frac{-1+\sqrt{5}}{2}F_0 \\ F_1+\frac{-1-\sqrt{5}}{2}F_0 \end{bmatrix}=\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} F_1 \\ F_0 \end{bmatrix}, \quad \begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix}=\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}$$

从而我们有:

$$\begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix}=\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}=\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$$

$$\begin{bmatrix} F_{k+1}+\frac{-1+\sqrt{5}}{2}F_k \\ F_{k+1}+\frac{-1-\sqrt{5}}{2}F_k \end{bmatrix}=\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} F_k+\frac{-1+\sqrt{5}}{2}F_{k-1} \\ F_k+\frac{-1-\sqrt{5}}{2}F_{k-1} \end{bmatrix}=\begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}\begin{bmatrix} F_k \\ F_{k-1} \end{bmatrix}$$

$A = X^{-1}\Lambda X(\mathrm{I})$

上海师范大学 Shanghai Normal University

我们得到了:

$$\begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}$$

令:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & \frac{-1+\sqrt{5}}{2} \\ 1 & \frac{-1-\sqrt{5}}{2} \end{bmatrix}$$

我们有:

$$XA = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} X, \quad \text{即:} \quad A = X^{-1} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} X$$

$A = X^{-1}\Lambda X$(II)

注意到此时:

$$X^{-1} = \begin{bmatrix} \frac{5+\sqrt{5}}{10} & \frac{5-\sqrt{5}}{10} \\ \frac{\sqrt{5}}{5} & -\frac{\sqrt{5}}{5} \end{bmatrix}$$

令 X^{-1} 的列向量是 x_1, x_2。我们有:

$$AX^{-1} = \begin{bmatrix} Ax_1 & Ax_2 \end{bmatrix} = X^{-1}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 & \lambda x_2 \end{bmatrix}$$

也就是:

$$Ax_1 = \lambda_1 x_1, \quad Ax_2 = \lambda_1 x_2$$

我们称 λ_1, λ_2 是 A 的特征值 (Eigenvalue),而 x_1, x_2 是相应的特征向量 (Eigenvector)

$A = X^{-1}\Lambda X$ 的另一个用处:

$$
\begin{aligned}
\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} &= \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\
&= \left(X^{-1} \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} X \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\
&= \left(X^{-1} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{bmatrix} X \right) \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}
\end{aligned}
$$

我们来尝试给出特征值和特征向量的定义：

定义 10 [特征值和特征向量，尝试定义].

令 A 是一个 $n \times n$ 的矩阵，λ 是一个实数，x 是一个非零的 n 维向量。如果：

$$Ax = \lambda x$$

则称 λ 是 A 的特征值 (Eigenvalue)，x 是 λ 对应的特征向量 (Eigenvector)。

让我们关注一下斐波那契的递推矩阵:

$$A=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}$$

我们有:

$$\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\frac{5+\sqrt{5}}{10}\\ \frac{\sqrt{5}}{5}\end{bmatrix}=\frac{1+\sqrt{5}}{2}\begin{bmatrix}\frac{5+\sqrt{5}}{10}\\ \frac{\sqrt{5}}{5}\end{bmatrix},\quad \begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}\frac{5-\sqrt{5}}{10}\\ -\frac{\sqrt{5}}{5}\end{bmatrix}=\frac{1-\sqrt{5}}{2}\begin{bmatrix}\frac{5-\sqrt{5}}{10}\\ -\frac{\sqrt{5}}{5}\end{bmatrix}$$

从而对于矩阵 A:

- 其有两个特征值: $\lambda_1=\frac{1+\sqrt{5}}{2}$, $\lambda_2=\frac{1-\sqrt{5}}{2}$, 对应的特征向量为:
  $x_1=\begin{bmatrix}\frac{5+\sqrt{5}}{10}\\ \frac{\sqrt{5}}{5}\end{bmatrix}$, $x_2=\begin{bmatrix}\frac{5-\sqrt{5}}{10}\\ -\frac{\sqrt{5}}{5}\end{bmatrix}$.
- 回顾之前的计算:

$$AX^{-1}=A\begin{bmatrix}x_1 & x_2\end{bmatrix}=\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}=X^{-1}\Lambda$$

特征值和特征向量的几何解释 (I)

令 A 是一个 $n \times n$ 的矩阵，将其看作一个函数$f_A : \mathbb{R}^n \to \mathbb{R}^n$，即:

$$f_A(x) = Ax$$

则当 x 是 A 的特征向量，即 $Ax = \lambda x$ 时我们有:

$$f_A(x) = \lambda x$$

这意味着f_A 不改变 x 的方向。

考察矩阵:

$$A = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix}$$

我们有:

$$A \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \quad A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

从而 $1, \frac{1}{2}$ 是其两个特征值, $\begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ 是相应的特征向量。注意到这两个向量是线性无关的，意味着任何一个向量都可以表示成其线性组合，比如:

$$\begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} + 0.2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

从而我们有:

$$A^{100} \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix} = (1)^{100} \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} + 0.2(\frac{1}{2})^{100} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \approx \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$$

$\lambda = 0$ 的情况

现在我们来讨论 $\lambda = 0$ 的情况，假设 $n \times n$ 的矩阵 A 存在特征值 0，即存在一个非零的 $x \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ 满足:

$$Ax = 0x = \mathbf{0}$$

显然这说明 $\dim(N(A)) \geqslant 1$.

引理 11.

下述是等价的:

1. A 具有特征值 0。
2. $N(A) \neq Z (= \{\mathbf{0}\})$。
3. $\operatorname{rank}(A) < n$，即 A 是奇异的 (singular).
4. $\det(A) = 0$.

对角矩阵的特征值

现在我们来考察对角矩阵:

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

显然我们有:

$$A e_i = \lambda_i e_i, \quad \text{这里} e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

从而 Λ 的特征值为$\lambda_1, \cdots, \lambda_n$，每个$\lambda_i$对应的特征向量为$e_i$。

投影矩阵的特征值

现在我们再来回顾投影矩阵，令 A 是一个 $m \times n$ 且 $\mathrm{rank}(A) = n$ 的矩阵，则投影至 $C(A)$ 的投影矩阵 P 可以表示为：

$$P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}$$

注意到 Px 是 x 到 $C(A)$ 上的投影，所以 $Px = \lambda x$ 只有两种情况：

1. $x \in C(A)$，此时 $Px = 1x = x$
2. $x \in C(A)^{\perp}$ 即 $x \in N(A^{\mathsf{T}})$，此时 $Px = 0x = \mathbf{0}$

回顾特征值，实际上就是：

$$Ax = \lambda x \quad \Longrightarrow \quad (A - \lambda I)x = \mathbf{0} \text{ 有非零解}$$

定理 12.

令 A 是 $n \times n$ 的矩阵，λ 是 A 的特征值当且仅当：

$$\det(A - \lambda I) = 0$$

证明.

$$\begin{aligned}
\lambda\text{是 } A \text{ 的特征值} &\Longleftrightarrow Ax = \lambda x \text{ 有非零解} \\
&\Longleftrightarrow (A - \lambda I)x = \mathbf{0} \text{ 有非零解} \\
&\Longleftrightarrow \operatorname{rank}(A - \lambda I) < n \\
&\Longleftrightarrow \det(A - \lambda I) = 0
\end{aligned}$$

$\square$

我们将 λ 看成一个变量，则对于 $n \times n$ 的矩阵 A:

$$f(\lambda) = \det(A - \lambda I_n)$$

是一个关于 λ 的 n 次多项式。(为什么？)

定义 13 **[特征多项式 (Characteristic Polynomial)].**

令 A 是 $n \times n$ 的矩阵，则称 $f(\lambda) = \det(A - \lambda I_n)$ 是 A 的特征多项式 (Characteristic Polynomial).

一些例子 (I)

我们首先来看下斐波那契矩阵:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

其特征多项式为:

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - \lambda - 1$$

该多项式的两个零点为:

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}$$

对角矩阵:

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

的特征多项式为:

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

消元法会改变特征值!

上海师范大学 Shanghai Normal University

我们来考察一下消元法:

$$A=\begin{bmatrix}1&3\\2&6\end{bmatrix}\implies U=\begin{bmatrix}1&3\\0&0\end{bmatrix}$$

我们有:

- $f_A(\lambda)=(1-\lambda)(6-\lambda)-6=\lambda(\lambda-7)\implies\lambda_1=0,\ \lambda_2=7.$
- $f_U(\lambda)=(1-\lambda)(0-\lambda)-0=\lambda(\lambda-1)\implies\lambda_1=0,\ \lambda_2=1.$

迹 (Trace)

我们定义 A 的主对角线上元素之和称为 A 的迹 (Trace),即:

$$\text{trace}=a_{11}+a_{22}+\cdots+a_{nn}$$

则我们有:

- $\lambda_1+\lambda_2+\cdots+\lambda_n=\text{trace}=a_{11}+a_{22}+\cdots+a_{nn}$。
- $\det(A)=\lambda_1\lambda_2\cdots\lambda_n$

最后我们来关注一下旋转矩阵:

$$Q=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}=\begin{bmatrix}\cos\frac{\pi}{2} & -\sin\frac{\pi}{2}\\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2}\end{bmatrix}$$

其特征多项式为:

$$\det(A-\lambda I)=\begin{vmatrix}-\lambda & -1\\ 1 & -\lambda\end{vmatrix}=\lambda^2+1$$

其两个零点为:

$$\lambda_1=i,\ \lambda_2=-i$$

特征值和特征向量是复数!

Q 有两个复数特征值:

$$\lambda_1 = i,\ \lambda_2 = -i$$

其对应的特征向量为:

$$Q\begin{bmatrix} 1 \\ i \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix}$$

$$Q\begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} i \\ 1 \end{bmatrix}$$

也是在$\mathbb{C}^2$上的向量!

定理 14 **[Fundamental Theorem of Algebra].**

对于任一的 $a_1, \cdots, a_n \in \mathbb{C}$，我们有

$$x^n + a_1 x^{n-1} + \cdots a_{n-1} x + a_n = (x - b_1)(x - b_2) \cdots (x - b_n)$$

这里 $b_1, \cdots b_n \in \mathbb{C}$，即任一单变元的 n 次复系数多项式恰好有 n 个复数根（可重复）。

特征值和特征向量的定义

定义 15 **[特征值和特征向量].**

令 A 是一个 $n \times n$ 的矩阵，$\lambda \in \mathbb{C}$，$x \in \mathbb{C}^n \setminus \{\mathbf{0}\}$。如果：

$$Ax = \lambda x$$

则称 λ 是 A 的特征值 (Eigenvalue)，x 是 λ 对应的特征向量 (Eigenvector)。

推论 16.

令 A 是一个 $n \times n$ 的矩阵，则在计算重根的意义下，A 恰好有 n 个 $\mathbb{C}$ 中的特征值。

例 17.

考察矩阵

$$\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

其特征多项式为 $f(\lambda) = \lambda^2(\lambda^2 + 1)$，注意到 $f(\lambda) = 0$ 一共有 4 个解:

$$\lambda_1 = \lambda_2 = 0, \ \lambda_3 = i, \ \lambda_4 = -i$$

这里 0 是一个重根，在计算重复次数以后可以得到其一共有 4 个特征值。

在一般意义下:

- 矩阵可以有复数。
- 向量空间是可以通过复数扩充的，即数乘允许 $c \in \mathbb{C}$ 进行运算。

但是在我们这堂课中，为了保持简单:

- 我们只考虑实数的矩阵。
- 但是我们依旧会讨论到复数的特征值和特征向量。

计算特征值和特征向量的总结

我们可以如下计算特征值和特征向量:

1. 计算 A 的特征多项式，即 $A-\lambda I$ 的行列式 $\det(A-\lambda I)$.
2. 计算 $\det(A-\lambda I)$ 的根，我们一共会得到 n 个特征值（可能重复）。这使得 $A-\lambda I$ 变成一个奇异矩阵 (singular).
3. 对每个 λ，通过解方程 $(A-\lambda I)x=\mathbf{0}$ 来获取 λ 对应的特征向量 x.

对角化矩阵 (Diagonalizing a Matrix)

$$
\begin{aligned}
\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} &= \left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\
&= \left(X' \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix} (X')^{-1} \right)^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} \\
&= \left(X' \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{bmatrix} (X')^{-1} \right) \begin{bmatrix} F_1 \\ F_0 \end{bmatrix}
\end{aligned}
$$

中间的对角矩阵 Λ 极大的简化了我们对 $\left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right)^k$ 的计算，我们把这过程称为对角化(Diagonalization)。

定义 18 **[Diagonalization].**

一个方阵 A 是可对角化的 (Diagonalizable)，如果其存在一个可逆矩阵 X 和对角矩阵 Λ 使得:

$$A = X\Lambda X^{-1}, \text{ 或者等价的 } \Lambda = X^{-1}AX$$

推论 19.

若 $A = X\Lambda X^{-1}$，其中 $X = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$，则对于任意的 $k \geqslant 1$ 我们有:

$$A^k = X\Lambda^k X^{-1} = X \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix} X^{-1}$$

我们也称这是矩阵的谱分解。

对角化的一个直观理解

$$AX = X\Lambda$$

1. X 的每一列都应该是 A 的特征向量。
2. X 是可逆的，所以 A 需要有 n 个线性无关的特征向量。

定理 20.

$n \times n$ 的矩阵可对角化当且仅当 A 有n 个线性无关的特征向量。

定理20的证明 (I)

假设 A 有 n 个特征值 $\lambda_1,\cdots,\lambda_n$(可以重复)，对应的特征向量 $x_1,\cdots,x_n$ 是线性无关的:

$$Ax_1=\lambda_1x_1,\cdots,Ax_n=\lambda_nx_n$$

定义下列矩阵:

$$\Lambda=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)=\begin{bmatrix}\lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k\end{bmatrix},\quad X=\begin{bmatrix}x_1,\cdots,x_n\end{bmatrix}$$

则我们有:

- $AX=A\begin{bmatrix}x_1,\cdots,x_n\end{bmatrix}=\begin{bmatrix}Ax_1,\cdots,Ax_n\end{bmatrix}=\begin{bmatrix}\lambda_1x_1,\cdots,\lambda_nx_n\end{bmatrix}=\begin{bmatrix}x_1,\cdots,x_n\end{bmatrix}\Lambda=X\Lambda.$
- $\mathrm{rank}(X)=n$，从而 X 可逆，即:

$$A=X\Lambda X^{-1}$$

定理20的证明 (II)

另一方面，如果:

$$A = X\Lambda X^{-1}$$

定义下列矩阵:

$$\Lambda = \mathrm{diag}(\lambda_1, \cdots, \lambda_n) = \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix}, \quad X = \begin{bmatrix} x_1, \cdots, x_n \end{bmatrix}$$

我们自然容易验证:

- 对于每个 x_i，有 $Ax_i = \lambda_i x_i$。
- $x_1, \cdots, x_n$ 是线性无关的。

$\square$

考察矩阵:

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

其特征多项式为 $f_A(\lambda) = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3)$，分别计算其对应的特征向量为:

- $\lambda = 1$ 时解方程 $(A - I)x = \mathbf{0}$ 得一个其特征向量为 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
- $\lambda = 3$ 时解方程 $(A - 3I)x = \mathbf{0}$ 得一个其特征向量为 $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

从而我们有:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$A^k = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 + 3^k & 1 - 3^k \\ 1 - 3^k & 1 + 3^k \end{bmatrix}$$

一些例子 (II)

下列矩阵:

$$A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

不可对角化。

- A 的特征值为 $\lambda = 0$，只有一个特征向量。
- B 的特征值为 $\lambda = 0$，只有一个特征向量。
- C 的特征值为 $\lambda = 1$，只有一个特征向量。

可逆和对角化没有关系!

- 矩阵是否可逆取决于是否有零特征值 ($\lambda = 0$)。
- 矩阵是否可对角化取决于是否有 n 个特征向量。

定理 21.

令 $x_1, \cdots, x_k$ 是 A 的 k 个特征向量，并且其对应的特征值两两不相同。则 $x_1, \cdots, x_k$ 是线性无关的。

推论 22.

如果 $n \times n$ 的矩阵 A 有 n 个不同的特征值，则 A 是可对角化的。

定理的直观理解

$$Ax = \lambda x = c_1 \lambda x_1 + \cdots c_k \lambda x_k$$

$$A(c_1 x_1 + \cdots c_k x_k) = c_1 \lambda_1 x_1 + \cdots c_k \lambda_k x_k$$

当 λ 互不相等的时候，上述两个式子不可能相等。

定理21的证明 (I)

上海师范大学 Shanghai Normal University

我们对 k 做归纳法，$k=1$ 的时候是显然的。

假设 $k \geqslant 2$，并且:

$$c_1 x_1 + \cdots + c_k x_k = \mathbf{0}$$

我们需要证明 $c_1 = c_2 = \cdots = c_k = 0$，反设结论不成立，即存在 $c_i \neq 0$，即:

$$x_i = \sum_{j \in [k] \setminus \{i\}} -\frac{c_j}{c_i} x_j$$

从而我们有:

$$\lambda_i x_i = A x_i = \sum_{j \in [k] \setminus \{i\}} -\frac{c_j}{c_i} A x_j = \sum_{j \in [k] \setminus \{i\}} -\frac{c_j}{c_i} \lambda_j x_j$$

定理21的证明 (II)

$$\lambda_i x_i = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}\lambda_j x_j$$

注意到，由归纳假设：

$$x_1,\cdots,x_{i-1},x_{i+1}\cdots x_k \text{ 是线性无关的}$$

- 如果 $\lambda_i = 0$，我们有 $\frac{c_j}{c_i}\lambda_j = 0$，从而 $\frac{c_j}{c_i} = 0$，$x_i = \mathbf{0}$，矛盾。
- 如果 $\lambda_i \neq 0$，我们有：

$$x_i = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}x_j \text{ 和 } x_i = \sum_{j\in[k]\setminus\{i\}} -\frac{c_j}{c_i}\frac{\lambda_j}{\lambda_i}x_j$$

注意到 $\lambda_j \neq \lambda_i$，所以存在不全为 0 的 $d_1,\cdots,d_{i-1},d_{i+1},\cdots,d_k$ 满足：

$$d_1x_1+\cdots+d_{i-1}x_{i-1}+d_{i-1}x_{i-1}+\cdots+d_kx_k = \mathbf{0}$$

与 $x_1,\cdots,x_{i-1},x_{i+1}\cdots x_k$ 是线性无关的矛盾。

让我们再来考虑一下对角化的形式：

$$A = X\Lambda X^{-1}$$

当特征向量组成的矩阵 X 改变的时候，我们得到了无数个不同的 A。

- 但这些矩阵 A 的特征值都是相同的。

更一般的，对于

$$A = BCB^{-1}$$

即使 C 不可以对角化，我们也可以得到跟 C 具有相同特征值的一大类矩阵。

我们称这样一个关系叫作相似 (similar)。

定义 23 **[相似矩阵（Similar Matrix].**

给定矩阵 A, B，如果存在可逆矩阵 P，使得：

$$A = PBP^{-1}$$

则称矩阵 A 和 B 是相似的。

引理 24.

相似矩阵 A 和 B 的特征值相同。

证明. 假设 $Bx = \lambda x$，则我们有：

$$A(Px) = PBP^{-1}(Px) = PBx = \lambda Px$$

$\square$

相似矩阵的一个例子，以及直观理解

考察如下两个矩阵：

$$Q = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix},\ Q' = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix}$$

Q 和 Q′ 是相似的，考虑如下两组基：

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\ \text{和} f_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ f_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

给定 $u = (5, 4) = 5e_1 + 4e_2$ 和 $v = (5, 9) = 5f_1 + 4f_2$，我们有：

- $Qu = Q\begin{bmatrix} e_1 & e_2 \end{bmatrix}\begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 2e_1 & 3e_2 \end{bmatrix}\begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} e_1 & e_2 \end{bmatrix}\begin{bmatrix} 10 \\ 12 \end{bmatrix}$
- $Q'v = Q'\begin{bmatrix} f_1 & f_2 \end{bmatrix}\begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 2f_1 & 3f_2 \end{bmatrix}\begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} f_1 & f_2 \end{bmatrix}\begin{bmatrix} 10 \\ 12 \end{bmatrix}$

两个相似矩阵本质上是在不同基下的相同线性变换。

阶段总结

- 特征值和特征向量。
- 求特征值和特征向量的方法。
- 对角化矩阵。
- 相似矩阵的概念。