0.1 特征值与特征向量

定义 0.1 (线性变换的特征值和特征向量)

设 φ 是数域 $\mathbb{K}$ 上线性空间 V 上的线性变换, 若 $\lambda_0 \in \mathbb{K}, x \in V$ 且 $x \neq 0$, 使

$$\varphi(x) = \lambda_0 x,$$

则称 λ_0 是线性变换 φ 的一个**特征值**, 向量 x 称为 φ 关于特征值 λ_0 的**特征向量**.

笔记 显然 φ 的关于特征值 λ_0 的全体特征向量加上零向量构成 V 的子空间.

定义 0.2 (线性变换的特征子空间)

设 λ_0 是线性空间 V 上的线性变换 φ 的特征值, 令

$$V_{\lambda_0} = \{\alpha \in V \mid \varphi(\alpha) = \lambda_0 \alpha\} = \{\alpha \in V \mid \alpha\text{是}\varphi\text{的属于}\lambda_0\text{的特征向量}\} \cup \{0\},$$

则显然 V_{λ_0} 是 V 的子空间, 称为 φ 的属于特征值 λ_0 的**特征子空间**.

笔记 显然 V_{λ_0} 是 φ 的不变子空间.

定义 0.3 (矩阵的特征值和特征向量)

设 A 是数域 $\mathbb{F}$ 上的 n 阶方阵, 若存在 $\lambda_0 \in \mathbb{F}$ 及 n 维非零列向量 α, 使

$$A\alpha = \lambda_0 \alpha$$

式成立, 则称 λ_0 为矩阵 A 的一个**特征值**,α 为 A 关于特征值 λ_0 的**特征向量**.

定义 0.4 (矩阵的特征子空间)

设 λ_0 是 $\mathbb{F}$ 上的 n 阶矩阵 A 的特征值, 令

$$V_{\lambda_0} = \{x \in \mathbb{F}^n \mid Ax = \lambda_0 x\} = \{x \in \mathbb{F}^n \mid x\text{是}A\text{的属于}\lambda_0\text{的特征向量}\} \cup \{0\},$$

则 V_{λ_0} 是线性方程组 $(\lambda_0 I_n - A)x = 0$ 的解空间, 从而是 $\mathbb{F}^n$ 的子空间, 称为 A 的属于特征值 λ_0 的**特征子空间**。

定义 0.5 (特征多项式)

设 A 是 n 阶方阵, 称 $|\lambda I_n - A|$ 为 A 的**特征多项式**.

定理 0.1 (相似矩阵有相同的特征多项式与特征值)

若 B 与 A 相似, 则 B 与 A 具有相同的特征多项式, 从而具有相同的特征值 (计重数).

证明 设 $B = P^{-1}AP$, 其中 P 是可逆矩阵, 则

$$|\lambda I_n - B| = |P^{-1}(\lambda I_n - A)P| = |P^{-1}||\lambda I_n - A||P| = |\lambda I_n - A|.$$

因此相似矩阵必有相同的特征多项式, 从而必有相同的特征值 (不计重数).

定理 0.2 (特征值的和与积)

矩阵 A 的 n 个特征值的和与积分别为

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(A),$$
$$\lambda_1 \lambda_2 \cdots \lambda_n = |A|.$$

证明 设

$$
\begin{aligned}
|\lambda I_n - A| &= \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \\
&= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n).
\end{aligned}
$$

由 Vieta 定理知 $\lambda_1 + \lambda_2 + \cdots + \lambda_n = -a_1$,$\lambda_1\lambda_2\cdots\lambda_n = (-1)^n a_n$. 由例题**??**可知 $a_1 = -(a_{11} + a_{22} + \cdots + a_{nn}) = -\mathrm{tr}(A)$,$a_n = (-1)^n|A|$. 因此矩阵 A 的 n 个特征值的和与积分别为

$$
\begin{aligned}
\lambda_1 + \lambda_2 + \cdots + \lambda_n &= \mathrm{tr}(A), \\
\lambda_1\lambda_2\cdots\lambda_n &= |A|.
\end{aligned}
$$

定义 0.6 (特征多项式)

设 φ 是线性空间 V 上的线性变换,φ 在 V 的某组基下的表示矩阵为 A, 由相似矩阵有相同特征值知 $|\lambda I_n - A|$ 与基或表示矩阵的选取无关, 称 $|\lambda I_n - A|$ 为 φ 的**特征多项式**, 记为 $|\lambda I_V - \varphi|$.

定理 0.3 (复方阵必相似于上三角阵)

任何复方阵必相似于一个上三角阵, 并且对角元素都是其特征值.

注 一般数域 $\mathbb{K}$ 上的矩阵未必相似于上三角阵.

证明 设 A 是 n 阶复方阵, 现对 n 用数学归纳法. 当 $n = 1$ 时结论显然成立. 假设对 $n - 1$ 阶矩阵结论成立, 现对 n 阶矩阵 A 来证明. 设 λ_1 是 A 的一个特征值, 则存在非零列向量 α_1, 使

$$
A\alpha_1 = \lambda_1\alpha_1.
$$

将 α_1 作为 C_n 的一个基向量, 并扩展为 C_n 的一组基 $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$. 将这些基向量按照列分块方式拼成矩阵 $P = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, 则 P 为 n 阶非异阵, 且

$$
\begin{aligned}
AP = A(\alpha_1, \alpha_2, \cdots, \alpha_n) &= (A\alpha_1, A\alpha_2, \cdots, A\alpha_n) \\
&= (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix},
\end{aligned}
$$

其中 A_1 是一个 $n - 1$ 阶方阵. 注意到 $P = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 非异, 上式即为

$$
P^{-1}AP = \begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}.
$$

因为 A_1 是一个 $n - 1$ 阶方阵, 所以由归纳假设可知, 存在 $n - 1$ 阶非异阵 Q, 使 $Q^{-1}A_1Q$ 是一个上三角阵. 令

$$
R = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix},
$$

则 R 是 n 阶非异阵, 且

$$
\begin{aligned}
R^{-1}P^{-1}APR &= \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}^{-1}\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} \\
&= \begin{pmatrix} 1 & O \\ O & Q^{-1} \end{pmatrix}\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 & * \\ O & Q^{-1}A_1Q \end{pmatrix}.
\end{aligned}
$$

这是一个上三角阵, 它与 A 相似, 并且对角元素都是其特征值.

推论 0.1

若数域 $\mathbb{K}$ 上的 n 阶方阵 A 的特征值全在 $\mathbb{K}$ 中, 则存在 $\mathbb{K}$ 上的非异阵 P, 使 $P^{-1}AP$ 是一个上三角阵.

证明 由复方阵必相似于上三角阵的证明类似可得.

0.1.1 直接利用定义计算和证明

例题 0.1 设 V 是 n 阶矩阵全体组成的线性空间, φ 是 V 上的线性变换: $\varphi(X) = AX$, 其中 A 是一个 n 阶矩阵。求证: φ 和 A 具有相同的特征值 (重数可能不同)。

证明 设 λ_0 是 A 的特征值, x_0 是对应的特征向量, 即 $Ax_0 = \lambda_0 x_0$。令 $X = (x_0, 0, \ldots, 0)$, 则 $\varphi(X) = AX = \lambda_0 X$ 且 $X \neq 0$, 因此 λ_0 也是 φ 的特征值.

反之, 设 λ_0 是 φ 的特征值, X 是对应的特征向量, 即 $\varphi(X) = AX = \lambda_0 X$。令 $X = (x_1, x_2, \ldots, x_n)$ 为列分块, 设第 i 个列向量 $x_i \neq 0$, 则 $Ax_i = \lambda_0 x_i$, 因此 λ_0 也是 A 的特征值.

例题 0.2 设 λ_1, λ_2 是矩阵 A 的两个不同的特征值, α_1, α_2 分别是 λ_1, λ_2 的特征向量, 求证: $\alpha_1 + \alpha_2$ 必不是 A 的特征向量。

证明 用反证法, 设 $A(\alpha_1 + \alpha_2) = \mu(\alpha_1 + \alpha_2)$, 又

$$A(\alpha_1 + \alpha_2) = A\alpha_1 + A\alpha_2 = \lambda_1\alpha_1 + \lambda_2\alpha_2,$$

于是 $(\lambda_1 - \mu)\alpha_1 + (\lambda_2 - \mu)\alpha_2 = 0$。由于属于不同特征值的特征向量线性无关, 故有 $\lambda_1 = \mu, \lambda_2 = \mu$, 从而 $\lambda_1 = \lambda_2$, 引出矛盾.

命题 0.1

设 φ 是线性空间 V 上的线性变换, V 有一个直和分解:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m,$$

其中 V_i 都是 φ-不变子空间。

(1) 设 φ 限制在 V_i 上的特征多项式为 $f_i(\lambda)$, 求证: φ 的特征多项式

$$f(\lambda) = f_1(\lambda) f_2(\lambda) \cdots f_m(\lambda).$$

(2) 设 λ_0 是 φ 的特征值, $V_0 = \{v \in V \mid \varphi(v) = \lambda_0 v\}$ 为特征子空间, $V_{i,0} = V_i \cap V_0 = \{v \in V_i \mid \varphi(v) = \lambda_0 v\}$, 求证:

$$V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}.$$

证明

(1) 取 V_i 的一组基, 将它们拼成 V 的一组基。记 A_i 是 φ 在 V_i 上的限制在 V_i 所取基下的表示矩阵, 则由定理**??**可知 φ 在 V 的这组基下的表示矩阵为分块对角矩阵 $A = \mathrm{diag}(A_1, A_2, \cdots, A_m)$, 于是

$$f(\lambda) = |\lambda I_n - A| = |\lambda I - A_1||\lambda I - A_2| \cdots |\lambda I - A_m|,$$

即 $f(\lambda) = f_1(\lambda) f_2(\lambda) \cdots f_m(\lambda)$.

(2) 任取 $\alpha \in V_0$, 设 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m$, 其中 $\alpha_i \in V_i$, 则

$$\varphi(\alpha_1) + \varphi(\alpha_2) + \cdots + \varphi(\alpha_m) = \varphi(\alpha) = \lambda_0\alpha = \lambda_0\alpha_1 + \lambda_0\alpha_2 + \cdots + \lambda_0\alpha_m.$$

注意到 $\varphi(\alpha_i) \in V_i, \varphi(\alpha) = \lambda_0\alpha \in V$, 故由直和的等价条件 (5) 可得 $\varphi(\alpha_i) = \lambda_0\alpha_i$, 即 $\alpha_i \in V_{i,0}$, 从而 $V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}$。注意到 $V_{i,0} \subseteq V_i$, 故

$$V_{i,0} \cap (V_{1,0} + \cdots + V_{i-1,0}) \subseteq V_i \cap (V_1 + \cdots + V_{i-1}) = \{0\}, \quad 2 \leq i \leq m,$$

于是由直和的等价条件 (2) 可知上述为直和。

推论 0.2

对分块对角矩阵 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_m\}$ 的任一特征值 λ_0, 其代数重数等于每个分块的代数重数之和, 其几何重数等于每个分块的几何重数之和.

证明 将命题 0.1的条件和结论代数化之后, 即可得到结论.

命题 0.2 (特征向量的延拓)

设 n 阶分块对角矩阵 $A=\text{diag}\{A_1,A_2,\cdots,A_m\}$, 其中 A_i 是 n_i 阶矩阵。

(1) 任取 A_i 的特征值 λ_i 及其特征向量 $x_i\in\mathbb{C}^{n_i}$, 求证: 可在 x_i 的上下添加适当多的零, 得到非零向量 $\tilde{x}_i\in\mathbb{C}^n$, 使得 $A\tilde{x}_i=\lambda_i\tilde{x}_i$, 即 $\tilde{x}_i$ 是 A 关于特征值 λ_i 的特征向量, 称为 x_i 的**延拓**.

(2) 任取 A 的特征值 λ_0, 并设 λ_0 是 $A_{i_1},\cdots,A_{i_r}$ 的特征值, 但不是其他 $A_j\ (1\le j\le m, j\neq i_1,\cdots,i_r)$ 的特征值, 求证: A 关于特征值 λ_0 的特征子空间的一组基可取为 $A_{i_1}(1\le k\le r)$ 关于特征值 λ_0 的特征子空间的一组基的延拓的并集.

证明

(1) 令 $\tilde{x}_i=(0,\cdots,0,x_i,0,\cdots,0)'$, 即 $\tilde{x}_i$ 的第 i 块为 x_i, 其余块均为 0, 显然 $\tilde{x}_i\neq 0$。容易验证 $A\tilde{x}_i=\lambda_i\tilde{x}_i$, 故结论成立。

(2) 由命题 0.1(2)以及直和的等价条件 (5) 即得.

例题 0.3 设 A 是 n 阶整数矩阵, p,q 为互素的整数且 $q>1$。求证：矩阵方程 $Ax=\frac{p}{q}x$ 必无非零解。

证明 用反证法。设上述矩阵方程有非零解, 则 $\frac{p}{q}$ 为 A 的特征值, 即为特征多项式 $f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$ 的根。由于 A 是整数矩阵, 故 $f(\lambda)$ 为整数系数多项式。由整数系数多项式有有理根的必要条件可知 $q\mid 1$, 从而 $q=\pm1$, 于是 $q\mid p$, 这与 p,q 互素矛盾。

例题 0.4 求下列 n 阶矩阵的特征值:

$$A=\begin{pmatrix}0&a&\cdots&a&a\\ b&0&\cdots&a&a\\ \vdots&\vdots&\ddots&\vdots&\vdots\\ b&b&\cdots&0&a\\ b&b&\cdots&b&0\end{pmatrix}.$$

解 若 $a=0$ 或 $b=0$, 则 A 是主对角元全为零的下三角或上三角矩阵, 故 A 的特征值全为零. 设 $a\neq0,b\neq0$, 则由命题**??**可知: 若 $a\neq b$, 则 $|\lambda I_n-A|=\frac{a(\lambda+b)^n-b(\lambda+a)^n}{a-b}$. 设 $\frac{b}{a}$ 的 n 次方根为 $\omega_i(1\le i\le n)$, 则

$$|\lambda I_n-A|=\frac{a(\lambda+b)^n-b(\lambda+a)^n}{a-b}=0\Rightarrow\left(\frac{\lambda+b}{\lambda+a}\right)^n=\frac{b}{a}$$
$$\Rightarrow\frac{\lambda+b}{\lambda+a}=w_i\ (1\le i\le n)\Rightarrow\lambda=\frac{a\omega_i-b}{1-\omega_i}(1\le i\le n).$$

从而 A 的特征值为 $\frac{a\omega_i-b}{1-\omega_i}(1\le i\le n)$.

若 $a=b$, 则 $|\lambda I_n-A|=(\lambda-(n-1)a)(\lambda+a)^{n-1}$, 从而 A 的特征值为 $(n-1)a$ (1 重), $-a$ (n - 1 重).

0.1.2 正向利用矩阵的多项式

定义 0.7 (矩阵多项式)

若 A 是一个 n 阶矩阵,$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$ 是一个多项式, 记

$$f(A)=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0I_n.$$

命题 0.3

设 n 阶矩阵 A 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, $f(x)$ 是一个多项式, 则 $f(A)$ 的全部特征值为 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$.

注 这个命题告诉我们: 如果能够将一个复杂矩阵写成一个简单矩阵的多项式, 那么就可以由简单矩阵的特征值得到复杂矩阵的特征值.

证明 因为任一 n 阶矩阵均复相似于上三角阵, 可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

因为上三角阵的和、数乘及乘方仍是上三角阵, 经计算不难得到

$$P^{-1}f(A)P=f(P^{-1}AP)=\begin{pmatrix}f(\lambda_1) & * & \cdots & * \\ 0 & f(\lambda_2) & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n)\end{pmatrix}.$$

因此 $f(A)$ 的全部特征值为 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$.

例题 0.5 设 n 阶矩阵 A 的全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 求 $2n$ 阶矩阵

$$\begin{pmatrix}A^2 & A \\ A^2 & A\end{pmatrix}$$

的全体特征值。

证明 由命题??可知

$$\left|\lambda I_{2n}-\begin{pmatrix}A^2 & A \\ A^2 & A\end{pmatrix}\right|=\left|\begin{pmatrix}\lambda I_n-A & -A^2 \\ -A^2 & \lambda I_n-A\end{pmatrix}\right|=|\lambda I_n-A-A^2||\lambda I_n-A+A^2|.$$

由命题 0.3可知 $A+A^2$ 的全体特征值为 $\lambda_i+\lambda_i^2(1\le i\le n)$, $A-A^2$ 的全体特征值为 $\lambda_i-\lambda_i^2(1\le i\le n)$, 因此所求矩阵的全体特征值为

$$\lambda_1+\lambda_1^2,\lambda_1-\lambda_1^2,\lambda_2+\lambda_2^2,\lambda_2-\lambda_2^2,\cdots,\lambda_n+\lambda_n^2,\lambda_n-\lambda_n^2.$$

命题 0.4 (循环矩阵的特征值)

求下列循环矩阵的特征值:

$$A=\begin{pmatrix}a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1\end{pmatrix}.$$

解 设 $J=\begin{pmatrix}O & I_{n-1} \\ 1 & O\end{pmatrix}$, $f(x)=a_1+a_2x+a_3x^2+\cdots+a_nx^{n-1}$, 则由循环矩阵的性质 2 可知 $A=f(J)$。经简单计算可得

$$|\lambda I_n-J|=\begin{vmatrix}\lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ -1 & 0 & 0 & \cdots & 0 & \lambda\end{vmatrix}\xlongequal{\text{按第一列展开}}\lambda\begin{vmatrix}\lambda & -1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \\ 0 & 0 & \cdots & 0 & \lambda\end{vmatrix}+(-1)^{n+1}(-1)\begin{vmatrix}-1 & 0 & \cdots & 0 & 0 \\ \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & \cdots & \lambda & -1\end{vmatrix}$$

$$=\lambda^n+(-1)^{n+2}(-1)^{n-1}=\lambda^n-1,$$

于是 J 的特征值为

$$\omega_k=\cos\frac{2k\pi}{n}+i\sin\frac{2k\pi}{n},\quad 0\le k\le n-1$$

因此 A 的特征值为 $f(1),f(\omega_1),\cdots,f(\omega_{n-1})$。

定义 0.8 (友矩阵)

$$A=\begin{pmatrix}0&0&0&\cdots&0&-a_0\\1&0&0&\cdots&0&-a_1\\0&1&0&\cdots&0&-a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&-a_{n-2}\\0&0&0&\cdots&1&-a_{n-1}\end{pmatrix}$$

称为多项式 $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 的 **友矩阵**.

♣

命题 0.5 (友矩阵的特征多项式及特征值)

设首一多项式 $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, $f(x)$ 的友矩阵

$$C=\begin{pmatrix}0&0&\cdots&0&-a_0\\1&0&\cdots&0&-a_1\\0&1&\cdots&0&-a_2\\\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&1&-a_{n-1}\end{pmatrix}.$$

(1) 求证：矩阵 C 的特征多项式就是 $f(\lambda)$。

(2) 设 $f(x)$ 的根为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，$g(x)$ 为任一多项式，求以 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$ 为根的 n 次多项式。♠

证明

(1)

$$|xE-A|=\begin{vmatrix}x&0&0&\cdots&0&a_0\\-1&x&0&\cdots&0&a_1\\0&-1&x&\cdots&0&a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&x&a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}\xlongequal[i=n,n-1,\cdots,2]{xr_i+r_{i-1}}\begin{vmatrix}0&0&0&\cdots&0&x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\\-1&0&0&\cdots&0&x^{n-1}+a_{n-1}x^{n-2}+\cdots+a_2x+a_1\\0&-1&0&\cdots&0&x^{n-2}+a_{n-1}x^{n-3}+\cdots+a_3x+a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&x^2+a_{n-1}x+a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}$$

$$\xlongequal{\text{按第一行展开}}\left(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a\right)(-1)^{n+1}\begin{vmatrix}-1&0&0&\cdots&0\\0&-1&0&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&0\\0&0&0&\cdots&-1\end{vmatrix}$$

$$=\left(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a\right)(-1)^{n+1}(-1)^{n-1}$$

$$=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a=f(x).$$

(2) 由假设及 (1) 的结论可知 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 C 的全体特征值，故由命题 0.3可知 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$ 是 $g(C)$ 的全体特征值，从而 $h(x)=|xI_n-g(C)|$ 即为所求的多项式。

0.1.3 反向利用矩阵的多项式

命题 0.6

设 n 阶矩阵 A 适合一个多项式 $g(x)$, 即 $g(A)=O$, 则 A 的任一特征值 λ_0 也必适合 $g(x)$, 即 $g(\lambda_0)=0$.

证明 设 α 是 A 关于特征值 λ_0 的特征向量, 经简单计算得

$$g(\lambda_0)\alpha=g(A)\alpha=0.$$

而 $\alpha\neq 0$, 因此 $g(\lambda_0)=0$.

命题 0.7 (幂零矩阵关于特征值的充要条件)

求证：n 阶矩阵 A 为幂零矩阵的充要条件是 A 的特征值全为零。

证明 若 A 为幂零矩阵, 即存在正整数 k, 使得 $A^k=O$, 则由命题 0.6可知 A 的任一特征值 λ_0 也适合 x^k, 于是 $\lambda_0=0$。

反之, 证法一:若 A 的特征值全为零, 则存在可逆矩阵 P, 使得 $P^{-1}AP=B$ 为上三角矩阵且主对角元素全为零。由上三角阵性质 (1) 可知 $B^n=O$, 于是 $A^n=(PBP^{-1})^n=PB^nP^{-1}=O$, 即 A 为幂零矩阵。

证法二:也可以利用 Cayley-Hamilton 定理来证明, 由于 A 的特征值全为零, 故其特征多项式为 λ^n, 从而 $A^n=O$.

例题 0.6 设 V 是数域 $\mathbb{F}$ 上的 n 阶方阵全体构成的线性空间, n 阶方阵

$$P=\begin{pmatrix}0&\cdots&0&1\\0&\cdots&1&0\\\vdots&&\vdots&\vdots\\1&0&\cdots&0\end{pmatrix},$$

V 上的线性变换 η 定义为 $\eta(X)=PX'P$. 试求 η 的全体特征值及其特征向量.

笔记 任意 n 阶矩阵 A 左乘 P 相当于行倒排, 右乘 P 矩阵相当于列倒排.

解 由 $P=P', P^2=I_n$ 容易验证 $\eta^2(X)=P(PX'P)P=X$, 即 $\eta^2=I_V$, 于是 η 的特征值也适合多项式 x^2-1, 从而特征值只能是 ±1.

设 $\eta(X_0)=PX_0'P=\pm X_0$, 这等价于 $(PX_0')P=\pm PX_0$, 即 PX_0 为对称矩阵或反对称矩阵.

令 $PX_0=E_{ii}, E_{ij}+E_{ji}$ (对称矩阵空间的基向量), 容易证明 η 关于特征值 1 的线性无关的特征向量为 $X_0=PE_{ii}\ (1\le i\le n), P(E_{ij}+E_{ji})\ (1\le i<j\le n)$.

令 $PX_0=E_{ij}-E_{ji}$ (反对称矩阵空间的基向量), 容易证明 η 关于特征值 -1 的线性无关的特征向量为 $X_0=P(E_{ij}-E_{ji})\ (1\le i<j\le n)$. 注意到这些特征向量恰好构成 V 的一组基, 故 η 的特征值为 1 $\frac{n(n+1)}{2}$ 重, -1 $\frac{n(n-1)}{2}$ 重.

例题 0.7 设 n 阶方阵 A 的每行每列只有一个元素非零, 并且那些非零元素为 1 或 -1, 证明: A 的特征值都是单位根.

证明 设 S 为由每行每列只有一个元素非零, 并且那些非零元素为 1 或 -1 的所有 n 阶方阵构成的集合, 由排列组合可得 $\overline{\overline{S}}=2^n n!$, 即 S 是一个有限集合. 注意到矩阵 $M\in S$ 当且仅当 $M=P_1P_2\cdots P_r$, 其中 P_k 是初等矩阵 P_{ij} 或 $P_i(-1)$, 因此对任意的 $M,N\in S$, $MN\in S$. 特别地, 由 $A\in S$ 可知 $A^k\in S\ (k\ge1)$, 即 $\{A,A^2,A^3,\cdots\}\subseteq S$, 于是存在正整数 $k>l$, 使得 $A^k=A^l$. 注意到 $|A|=\pm1$, 故 A 可逆, 于是 $A^{k-l}=I_n$, 从而 A 的特征值适合多项式 $x^{k-l}-1$, 即为单位根.

例题 0.8 设 A 是 n 阶实方阵, 又 I_n-A 的特征值的模长都小于 1, 求证: $0<|A|<2^n$.

证明 设 A 的特征值为 $\lambda_1,\cdots,\lambda_n$, 则 I_n-A 的特征值为 $1-\lambda_1,\cdots,1-\lambda_n$. 由假设 $|1-\lambda_i|<1$, 若 λ_i 是实数, 则 $0<\lambda_i<2$; 若 λ_i 是虚数, 则 $\overline{\lambda_i}$ 也是 A 的特征值, 此时 $1-\overline{\lambda_i}$ 也是 I_n-A 的特征值. 从而 $|1-\lambda_i|<1, |1-\overline{\lambda_i}|<1$, 于是

$$|1-\lambda_i^2|=|(1-\lambda_i)(1-\overline{\lambda_i})|=|1-\lambda_i||1-\overline{\lambda_i}|<1.$$

因此 $0<\lambda_i^2<2$, 故此时 $0<\lambda_i<\sqrt{2}$.

综上, 无论 λ_i 是实数还是虚数, 都有 $0<|\lambda_i|<2$. 由于 $|A|$ 等于所有特征值之积, 故 $0<|A|<2^n$.

命题 0.8 (逆矩阵的特征值)

设 n 阶矩阵 A 是可逆矩阵, 且 A 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则 A^{-1} 的全部特征值为 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$。

证明 首先注意到 A 是可逆矩阵, $\lambda_1\lambda_2\cdots\lambda_n=|A|\neq 0$, 因此每个 $\lambda_i\neq 0$ (事实上, A 可逆的充分必要条件是它的特征值全不为零)。由复方阵必相似于上三角阵可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

因为上三角矩阵的逆矩阵仍然是上三角矩阵, 经过计算不难得到

$$P^{-1}A^{-1}P=(P^{-1}AP)^{-1}=\begin{pmatrix}\lambda_1^{-1} & * & \cdots & * \\ 0 & \lambda_2^{-1} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{-1}\end{pmatrix}.$$

因此 A^{-1} 的全部特征值为 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$.

命题 0.9 (伴随矩阵的特征值)

设 n 阶矩阵 A 的全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，求证：$A^*$ 的全体特征值为

$$\prod_{i\neq 1}\lambda_i,\prod_{i\neq 2}\lambda_i,\cdots,\prod_{i\neq n}\lambda_i.$$

证明 因为任一 n 阶矩阵均复相似于上三角矩阵，故可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

注意到上三角矩阵的伴随矩阵仍是上三角矩阵，经计算可得

$$P^{-1}A^*P=P^*A^*(P^{-1})^*=(P^{-1}AP)^*=\begin{pmatrix}\prod_{i\neq 1}\lambda_i & * & \cdots & * \\ 0 & \prod_{i\neq 2}\lambda_i & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \prod_{i\neq n}\lambda_i\end{pmatrix}.$$

因此 A^* 的全部特征值为

$$\prod_{i\neq 1}\lambda_i,\prod_{i\neq 2}\lambda_i,\cdots,\prod_{i\neq n}\lambda_i.$$

0.1.4 特征值的降价公式

定理 0.4 (特征值的降价公式)

设 A 是 $m\times n$ 矩阵，B 是 $n\times m$ 矩阵，且 $m\geq n$。求证：

$$|\lambda I_m - AB| = \lambda^{m-n}|\lambda I_n - BA|.$$

特别地，若 A,B 都是 n 阶矩阵，则 AB 与 BA 有相同的特征多项式.

笔记 本质上就是打洞原理.

证明 证法一 (打洞原理): 当 $\lambda\neq 0$ 时, 考虑下列分块矩阵:

$$\begin{pmatrix} \lambda I_m & A \\ B & I_n \end{pmatrix},$$

因为 $\lambda I_m, I_n$ 都是可逆矩阵, 故由行列式的降阶公式可得

$$|I_n|\cdot|\lambda I_m - A(I_n)^{-1}B| = |\lambda I_m|\cdot|I_n - B(\lambda I_m)^{-1}A|,$$

即有

$$|\lambda I_m - AB| = \lambda^{m-n}|\lambda I_n - BA|$$

成立.

当 $\lambda = 0$ 时, 若 $m>n$, 则 $r(AB)\leq \min\{r(A),r(B)\}\leq \min\{m,n\} = n < m$, 故 $|-AB| = 0$, 结论成立；若 $m=n$, 则 $|-AB| = (-1)^n|A||B| = |-BA|$, 结论也成立。

事实上, $\lambda = 0$ 的情形也可以用 Cauchy-Binet 公式来处理, 还可以通过摄动法由 $\lambda\neq 0$ 的情形来得到.

证法二 (相抵标准型): 设 A 的秩等于 r, 则存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q, 使得

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

令

$$Q^{-1}BP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

其中 B_{11} 是 $r\times r$ 矩阵, 则

$$PABP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix},\quad Q^{-1}BAQ = \begin{pmatrix} B_{11} & O \\ B_{21} & O \end{pmatrix}.$$

因此

$$|\lambda I_m - AB| = \left|\begin{pmatrix} \lambda I_r - B_{11} & -B_{12} \\ O & \lambda I_{m-r} \end{pmatrix}\right| = \lambda^{m-r}|\lambda I_r - B_{11}|,$$

同理

$$|\lambda I_n - BA| = \left|\begin{pmatrix} \lambda I_r - B_{11} & O \\ -B_{21} & \lambda I_{n-r} \end{pmatrix}\right| = \lambda^{n-r}|\lambda I_r - B_{11}|.$$

比较上面两个式子即可得出结论。

证法三 (摄动法): 先证明 $m=n$ 的情形。若 A 可逆, 则 $BA = A^{-1}(AB)A$, 即 AB 和 BA 相似, 因此它们的特征多项式相等。对于一般的方阵 A, 可取到一列有理数 $t_k\to 0$, 使得 $t_kI_n + A$ 是可逆矩阵。由可逆情形的证明可得

$$|\lambda I_n - (t_kI_n + A)B| = |\lambda I_n - B(t_kI_n + A)|.$$

注意到上述两边的行列式都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k\to 0$, 即有 $|\lambda I_n - AB| = |\lambda I_n - BA|$ 成立。

再证明 $m > n$ 的情形。令

$$C = \begin{pmatrix} A & O \end{pmatrix}, \quad D = \begin{pmatrix} B \\ O \end{pmatrix},$$

其中 C, D 均为 $m \times m$ 分块矩阵, 则

$$CD = AB, \quad DC = \begin{pmatrix} BA & O \\ O & O \end{pmatrix}.$$

因此由方阵的情形可得

$$|\lambda I_m - AB| = |\lambda I_m - CD| = |\lambda I_m - DC| = \lambda^{m-n} |\lambda I_n - BA|.$$

例题 0.9 设 α 是 n 维实列向量且 $\alpha'\alpha = 1$，试求矩阵 $I_n - 2\alpha\alpha'$ 的特征值。

解 设 $A = I_n - 2\alpha\alpha'$，则由特征值的降价公式可得

$$|\lambda I_n - A| = |(\lambda - 1)I_n + 2\alpha\alpha'| = (\lambda - 1)^{n-1}(\lambda - 1 + 2\alpha'\alpha) = (\lambda - 1)^{n-1}(\lambda + 1).$$

因此, 矩阵 A 的特征值为 1(n-1) 重,-1(1 重).

例题 0.10 设 A 为 n 阶方阵, α, β 为 n 维列向量, 试求矩阵 $A\alpha\beta'$ 的特征值。

解 设 $B = A\alpha\beta'$, 则由特征值的降价公式可得

$$|\lambda I_n - B| = |I_n - (A\alpha)\beta'| = \lambda^{n-1}(\lambda - \beta' A\alpha).$$

若 $\beta' A\alpha \neq 0$, 则 B 的特征值为 0 $(n-1$ 重$)$, $\beta' A\alpha$ $(1$ 重$)$; 若 $\beta' A\alpha = 0$, 则 B 的特征值为 0 $(n$ 重$)$.

例题 0.11 设 a_i $(1 \le i \le n)$ 都是实数, 且 $a_1 + a_2 + \cdots + a_n = 0$, 试求下列矩阵的特征值:

$$A = \begin{pmatrix} a_1^2 & a_1a_2 + 1 & \cdots & a_1a_n + 1 \\ a_2a_1 + 1 & a_2^2 & \cdots & a_2a_n + 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_na_1 + 1 & a_na_2 + 1 & \cdots & a_n^2 \end{pmatrix}.$$

解 矩阵 A 可以分解为 $A = -I_n + BC$, 其中

$$B = \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}, \quad C = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

由特征值的降价公式得

$$\begin{aligned} |\lambda I_n - A| &= |(\lambda + 1)I_n - BC| \\ &= (\lambda + 1)^{n-2} |(\lambda + 1)I_2 - CB|. \end{aligned}$$

注意到 $a_1 + a_2 + \cdots + a_n = 0$, 故有

$$CB = \begin{pmatrix} a_1^2 + a_2^2 + \cdots + a_n^2 & 0 \\ 0 & n \end{pmatrix}.$$

因此 A 的特征值为 -1 $(n-2$ 重$)$, $n-1$, $a_1^2 + a_2^2 + \cdots + a_n^2 - 1$.

例题 0.12 设 A, B, C 分别是 $m \times m, n \times n, m \times n$ 矩阵, 满足: $AC = CB$, $\mathrm{r}(C) = r$. 求证: A 和 B 至少有 r 个相同的特征值。

注 不妨设 $C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的原因: 假设当 $C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 时, 结论已经成立, 则对于一般的满足条件的矩阵 C, 由条件我们有

$$r(C) = r, \quad AC = BC.$$

由 $r(C)=r$ 可知, 存在可逆矩阵 P,Q, 使得

$$PCQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

从而对 $AC=BC$ 两边同时左乘 P, 右乘 Q 得到

$$(PAP^{-1})(PCQ)=(PCQ)(Q^{-1}BQ).$$

于是由假设可知 PAP^{-1} 和 $Q^{-1}BQ$ 都至少有 r 个相同的特征值. 又因为相似矩阵有相同的特征值, 所以 A,B 也至少有 r 个相同的特征值. 故不妨设成立.

证明 设 P 为 m 阶非异阵, Q 为 n 阶非异阵, 使得

$$PCQ=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

注意到问题的条件和结论在相抵变换 $C\mapsto PCQ$, $A\mapsto PAP^{-1}$, $B\mapsto Q^{-1}BQ$ 下保持不变, 故不妨从一开始就假设 $C=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是相抵标准型. 设

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},\quad B=\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

为对应的分块, 则

$$AC=\begin{pmatrix} A_{11} & O \\ A_{21} & O \end{pmatrix},\quad CB=\begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}.$$

由 $AC=CB$ 可得 $A_{11}=B_{11}, A_{21}=0, B_{12}=0$. 于是

$$|\lambda I_m-A|=|\lambda I_r-A_{11}|\cdot|\lambda I_{m-r}-A_{22}|,$$
$$|\lambda I_n-B|=|\lambda I_r-B_{11}|\cdot|\lambda I_{n-r}-B_{22}|.$$

从而 A,B 至少有 r 个相同的特征值 (即 $A_{11}=B_{11}$ 的特征值).

0.1.5 特征值与特征多项式系数的关系

命题 0.10 (特征值与特征多项式系数的关系)

设 n 阶矩阵 A 的特征多项式为

$$f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n.$$

求证: a_r 等于 $(-1)^r$ 乘以 A 的所有 r 阶主子式之和, 即

$$a_r=(-1)^r\sum_{1\le i_1<i_2<\cdots<i_r\le n}A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix},\quad 1\le r\le n.$$

进一步, 若设 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则

$$\sum_{1\le i_1<i_2<\cdots<i_r\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}=\sum_{1\le i_1<i_2<\cdots<i_r\le n}A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix},\quad 1\le r\le n.$$

注 上述结论中最常用的是 $r=1$ 和 $r=n$ 的情形:

$$\lambda_1+\lambda_2+\cdots+\lambda_n=\mathrm{tr}(A),\quad \lambda_1\lambda_2\cdots\lambda_n=|A|.$$

特别地, **A 是非异阵的充要条件是 A 的特征值全不为零.** 因此, 特征值的计算是判断矩阵是否非异阵的重要依据.

证明 第一种结论是推论**??**.

由 Vieta 定理可得

$$\sum_{1\le i_1<i_2<\cdots<i_r\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}=(-1)^r a_r=(-1)^r(-1)^r\sum_{1\le i_1<i_2<\cdots<i_r\le n}A\begin{pmatrix} i_1 & i_2 & \cdots & i_r\\ i_1 & i_2 & \cdots & i_r\end{pmatrix}$$
$$=\sum_{1\le i_1<i_2<\cdots<i_r\le n}A\begin{pmatrix} i_1 & i_2 & \cdots & i_r\\ i_1 & i_2 & \cdots & i_r\end{pmatrix},1\le r\le n.$$

因此第二种结论也成立.

例题 0.13 设 n 阶方阵 A 满足

$$A^2-A-3I_n=O,$$

求证：$A-2I_n$ 是非奇异阵。

笔记 用特征值判断矩阵非异性.

证明 用反证法。设 $A-2I_n$ 为奇异阵, 则 2 是 A 的特征值。注意到 A 适合

$$f(x)=x^2-x-3,$$

但特征值 2 却不适合 $f(x)$, 这与命题 0.6矛盾。

例题 0.14

证明

例题 0.15

证明

例题 0.16

证明

例题 0.17

证明

例题 0.18

证明