第四次习题课讲义

矩阵分类的巅峰—相似

王向阳

2023 年 12 月 17 日

如果你想了解一个东西, 却不知道它的精妙所在, 那将会是一种遗憾.

目录

1 线性变换

1.1 线性变换退化到哪里去 — Rank-Nullity 定理

定义 1. 设 $\mathscr{A}: U \to V$ 是线性映射. 集合

$$\mathrm{Im}(\mathscr{A}) = \{\mathscr{A}(\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in U\}$$

称为 U 在 $\mathscr{A}$ 下的像, 或者说是值域, 也记为 $\mathscr{A}(U)$. 集合

$$\mathrm{Ker}(\mathscr{A}) = \{\boldsymbol{\alpha} \in U : \mathscr{A}(\boldsymbol{\alpha}) = \mathbf{0} \in V\}$$

称为 $\mathscr{A}$ 的核, 也记为 $\mathscr{A}^{-1}(\mathbf{0})$.

容易根据子空间的定义说明:

定理 1. $\mathrm{Im}(\mathscr{A}), \mathrm{Ker}(\mathscr{A})$ *分别为 V 和 U 的子空间.*

定义 2. 设 $\mathscr{A}: U \to V$ 是线性映射. $\dim(\mathrm{Im}(\mathscr{A}))$ 称为 $\mathscr{A}$ 的秩, 记作 $\rho(\mathscr{A})$, $\dim(\mathrm{Ker}(\mathscr{A}))$ 称为 $\mathscr{A}$ 的零度 (nullity), 记作 $\nu(\mathscr{A})$.

定理 2 (Rank-Nullity 定理)**.** *设 $\mathscr{A}: U \to V$ 是线性映射, 则*

$$\dim U = \rho(\mathscr{A}) + \nu(\mathscr{A}).$$

证明. 设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_\nu$ 是 $\mathrm{Ker}(\mathscr{A})$ 的一组基, 设 $\dim(U) = n$, 将其扩充成 U 的一组基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$.

设 $\mathscr{A}\boldsymbol{\alpha}_{\nu+i} = \boldsymbol{\beta}_i, i = 1, 2, \ldots, n-\nu$. 我们说明其线性无关.

若其线性相关, 则存在不全为零的 $k_1, k_2, \ldots, k_{n-\nu} \in F$, 使得

$$k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2 + \cdots + k_{n-\nu}\boldsymbol{\beta}_{n-\nu} = 0.$$

从而 $k_1\boldsymbol{\alpha}_{\nu+1} + k_2\boldsymbol{\alpha}_{\nu+2} + \cdots + k_{n-\nu}\boldsymbol{\alpha}_n \in \mathrm{Ker}(\mathscr{A})$. 所以可以用 $\mathrm{Ker}(\mathscr{A})$ 中的基线性表示. 所以存在 $l_1, l_2, \ldots, l_\nu \in F$, 使得

$$k_1\boldsymbol{\alpha}_{\nu+1} + k_2\boldsymbol{\alpha}_{\nu+2} + \cdots + k_{n-\nu}\boldsymbol{\alpha}_n = l_1\boldsymbol{\alpha}_1 + l_2\boldsymbol{\alpha}_2 + \cdots + l_\nu\boldsymbol{\alpha}_\nu.$$

由于这些向量线性无关, 所以系数只能为 0. 矛盾!

又因为 $\mathrm{Im}(\mathscr{A})$ 都可以用 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_{n-\nu}$ 线性表示. 所以 $\dim(\mathrm{Im}(\mathscr{A})) = n - \nu$. 即 $n = \rho(\mathscr{A}) + \nu(\mathscr{A})$. $\square$

1.2 线性变换变到哪里去 — 基制定的 "规则"

一个线性变换可以由一组基的像确定. 所以选取一组好的基, 能够更简单的表现这个线性变换.

例 1. 设 A 是数域 F 上 n 阶方阵. *证明: 存在数域 F 上 n 阶可逆方阵 P, 使得 PA 是幂等方阵* (即 $(PA)^2 = PA$).

证明. 设 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n\}$ 是数域 F 上 n 维线性空间 V 的基. 由

$$\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n)A$$

可以确定 V 的一个线性变换 $\mathscr{A}$. 设 $\operatorname{rank}(A) = r$. 从而 $\dim(\operatorname{Ker}(\mathscr{A})) = n - r$.

设 $\boldsymbol{\beta}_{r+1}, \boldsymbol{\beta}_{r+2}, \dots, \boldsymbol{\beta}_n$ 是 $\operatorname{Ker}(\mathscr{A})$ 的一组基, 我们将其扩充为 V 的一组基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n$. 可以得到 $\mathscr{A}(\boldsymbol{\beta}_1), \mathscr{A}(\boldsymbol{\beta}_2), \dots, \mathscr{A}(\boldsymbol{\beta}_r)$ 是 $\operatorname{Im}(\mathscr{A})$ 的一组基, 设其为 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots, \boldsymbol{\gamma}_r$. 再把 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots, \boldsymbol{\gamma}_r$ 扩充为 V 的一组基 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots, \boldsymbol{\gamma}_n$.

由

$$\mathscr{B}(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots, \boldsymbol{\gamma}_n) = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n).$$

可以确定一个可逆线性变换 $\mathscr{B} : V \to V$. 容易验证:

$$(\mathscr{B}\mathscr{A})^2 = \mathscr{B}\mathscr{A}.$$

记

$$\mathscr{B}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n)P,$$

可以知道 P 是可逆的. 所以

$$\mathscr{B}\mathscr{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n)PA.$$

因为 $(\mathscr{B}\mathscr{A})^2 = \mathscr{B}\mathscr{A}$. 可以得到 $(PA)^2 = PA$.

□

2 相似下的不变量 — 特征值

设 n 阶方阵 A 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$.

特征值能够将矩阵的各个参数联系起来.

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

由于相似的两个矩阵, 具有相同的特征值. 所以上面这些量也是相似下的不变量.

设 $A = (a_{ij})$, 根据韦达定理, 我们还能得到 (虽然我们说其它系数很难算, 但是不是不能算):

因为

$$\det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{vmatrix}.$$

设其为 $f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_n$. 则

$$\begin{aligned} a_i &= (-1)^i \sum_{1 \le k_1 < k_2 < \cdots < k_i \le n} \lambda_{k_1}\lambda_{k_2}\dots\lambda_{k_i} \\ &= (-1)^i \sum_{1 \le k_1 < k_2 < \cdots < k_i \le n} \det(A\begin{pmatrix} k_1 & k_2 & \dots & k_i \\ k_1 & k_2 & \dots & k_i \end{pmatrix}). \end{aligned}$$

即
$$\sum_{1\le k_1<k_2<\cdots<k_i\le n}\lambda_{k_1}\lambda_{k_2}\ldots\lambda_{k_i}=\sum_{1\le k_1<k_2<\cdots<k_i\le n}\det(A\begin{pmatrix}k_1 & k_2 & \ldots & k_i\\ k_1 & k_2 & \ldots & k_i\end{pmatrix}).$$

一个重要公式: 若 A,B 分别是 $m\times n, n\times m$ 矩阵. 则
$$\det(\lambda I_m-AB)=\lambda^{m-n}\det(\lambda I_n-BA).$$
由此容易知道当 A,B 都是 n 阶方阵时, AB 和 BA 具有完全相同的特征值.

一些基本结论:

- 上 (下) 三角阵的特征值是全体对角元;
- 准对角矩阵的特征值是各个块的特征值的并;
- A 和 A^T 具有相同的特征值, 但未必有相同的特征向量;
- A 是可逆矩阵当且仅当它的特征值都不为零;
- 属于 A 同一特征值的两个特征向量的非零和仍是 A 的这个特征值的特征向量 (构成子空间), 不同特征值的特征向量的和一定不是 A 的特征向量.

2.1 两个得到特征值的基本方法

- $A\boldsymbol{x}=\lambda\boldsymbol{x}(\boldsymbol{x}\neq 0)$;
- 特征多项式 $\det(\lambda I-A)$ 的根.

例 2. *设 A 是一个 n 阶方阵, A 的各行各列恰好有一个非零元素, 且为 1 或 -1. 证明: A 的特征值都是单位根.*

证明. A 的任意两列, 是相互正交的. 于是有 $A^TA=I$, 即 A 是一个正交矩阵.

设 λ 是 A 的任意一个特征值, $A\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}(\boldsymbol{\alpha}\neq\mathbf{0})$. 两边取共轭转置, 再右乘 $A\boldsymbol{\alpha}$.
$$\bar{\boldsymbol{\alpha}}^T\boldsymbol{\alpha}=\bar{\boldsymbol{\alpha}}^TA^TA\boldsymbol{\alpha}=|\lambda|^2\bar{\boldsymbol{\alpha}}^T\boldsymbol{\alpha}.$$
于是 $|\lambda|=1$, 即 A 的特征值都是单位根. □

例 3. *如果 $A^2=aI$. 则 A 的特征值只能为 $\pm\sqrt{a}$.*

例 4. *如果 $A^2=aA$, 则 A 的特征值只能为 a 和 0.*

例 5 (马尔科夫矩阵)**.** *设 n 阶方阵 $A=(a_{ij})$ 满足条件:*
(1) $0\le a_{ij}\le 1,\forall i,j=1,2,\ldots,n$;
(2) $\sum_{i=1}^n a_{ij}=1(j=1,2,\ldots,n)$.
求证:
(1) $\lambda=1$ 是 A 的一个特征值;
(2) 对于其它的特征值 λ', 都有 $|\lambda'|\le 1$.

证明. (1) 观察到 $I-A$ 每列之和为 0. 因此行列式为 0, 所以 1 是 A 的一个特征值.

(2) 反证. 设 A 有个特征值 λ' 满足 $|\lambda'| > 1$, 我们知道 λ' 也是 A^T 的特征值. 设 $(b_1, b_2, \ldots, b_n)^T$ 是 A^T 属于 λ' 的特征向量, 即

$$A^T \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \lambda' \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

设 $|b_i| = \max\{|b_1|, |b_2|, \ldots, |b_n|\}$. 所以

$$(a_{1i}, a_{2i}, \ldots, a_{ni}) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \lambda' b_i.$$

而左式取模 $\leq \sum_{j=1}^{n} a_{ji}|b_i| = |b_i|$. 而右边取模为 $|\lambda'||b_i| > |b_i|$, 矛盾. □

这样的模小于等于 1 的矩阵, 我们一般称之为收敛矩阵. 我们知道, 如果一个向量 $\boldsymbol{\alpha}$ 能被这种矩阵的一组特征向量线性表示, 那么 $A^n\boldsymbol{\alpha}$ 当 n 趋向于无穷时不会跑到很远的地方. 若其可以对角化, 则对于任意向量都将会如此. 若其能表示的特征向量组对应模为 1 的特征值只有 1 , 则其极限存在.

例 6. *设 A 是 n 阶方阵, 则 A 是数量矩阵当且仅当任意 n 维非零向量都是 A 的特征向量.*

证明. 设 $A = aI$, 则 $A\boldsymbol{x} = a\boldsymbol{x}, \forall\ \boldsymbol{x} \in F^n$.

反之, 因为

$$A(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

所以 $A = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. 根据不同特征值的特征向量和一定不是 A 的特征向量, 所以 $\lambda_1 = \lambda_2 = \cdots = \lambda_n$, 从而 A 是数量矩阵. □

例 7. *设 A, B 均为 n 阶方阵. 如果 $r(A) + r(B) < n$, 证明 A 和 B 有公共的特征值和特征向量.*

证明. 由 $r(A) + r(B) < n$ 可知, A 和 B 均不可逆, 所以 0 是 A 和 B 的公共特征值.

因为 $r\begin{pmatrix} A \\ B \end{pmatrix} \leq r(A) + r(B) < n$. 所以 $\begin{pmatrix} A \\ B \end{pmatrix}\boldsymbol{x} = \boldsymbol{0}$ 有非零解. 且有 $A\boldsymbol{x} = \boldsymbol{0}, B\boldsymbol{x} = \boldsymbol{0}$. 这个非零解就是 A, B 公共属于 0 的特征向量. □

2.2 特征值 — 求行列式和迹的利器

例 8. *已知三阶方阵 A 满足 $\det(A - I) = \det(A - 2I) = \det(A + I) = \lambda$. 当 $\lambda = 0$ 时, 求 $\det(A + 3I)$. 当 $\lambda = 2$ 时, 求 $\det(A + 3I)$.*

解. (1) 即 $1, 2, -1$ 是 A 的特征值, 从而 $A + 3I$ 的特征值为 $4, 5, 2$. 所以 $\det(A + 3I) = 40$.

(2) 设 $\det(A + \lambda I) = \lambda^3 + a_0\lambda^2 + a_1\lambda + a_2$. 于是

$$a_0 - a_1 + a_2 = 3,$$
$$4a_0 - 2a_1 + a_2 = 10,$$

$$a_0 + a_1 + a_2 = 1.$$

或者由 $\lambda^3+a_0\lambda^2+a_1\lambda+a_2-2$ 的根有 $-1,-2,1$. 因此 $\lambda^3+a_0\lambda^2+a_1\lambda+a_2-2=(\lambda+1)(\lambda+2)(\lambda-1)$. 所以 $\det(A+\lambda I)=(\lambda+1)(\lambda+2)(\lambda-1)+2$.

解得 $a_0=2, a_1=-1, a_2=0$. 于是 $\det(A+3I)=27+9\times 2+3\times(-1)=42$. □

例 9. 已知 n 阶方阵

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

其中 $\sum_{i=1}^n a_i=1, \sum_{i=1}^n a_i^2=n$. 求 A 的全部特征值; 求 A 的行列式和迹.

解. (1) 观察到 $I+A=B^TB$, 其中 $B=\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ 1 & 1 & \dots & 1 \end{pmatrix}$.

所以

$$\begin{aligned}\det(\lambda I-A)&=\det((\lambda+1)I-B^TB)\\&=(\lambda+1)^{n-2}\det((\lambda+1)I-BB^T)\\&=(\lambda+1)^{n-2}\det(\begin{pmatrix}\lambda+1 & 0\\ 0 & \lambda+1\end{pmatrix}-\begin{pmatrix} n & 1\\ 1 & n\end{pmatrix})\\&=(\lambda+1)^{n-2}(\lambda+2-n)(\lambda-n).\end{aligned}$$

因此 A 的全部特征值是 $\lambda_1=-1(n-2\text{重}), \lambda_2=n, \lambda_3=n-2$.

(2) 可以直接算, 但这里根据特征值也可以得到行列式为 $(-1)^{n-2}n(n-2)$, 迹为 n.

□

设 A 的特征值为 $\lambda_1,\lambda_2,\dots,\lambda_n$. 我们很容易根据特征多项式算出 $A+\mu I$ 的所有特征值. 即

$$\det(\lambda I-(A+\mu I))=\det((\lambda-\mu)I-A)=(\lambda-\mu-\lambda_1)(\lambda-\mu-\lambda_2)\cdots(\lambda-\mu-\lambda_n).$$

因此 $A+\mu I$ 全部特征值 $\lambda_1+\mu,\lambda_2+\mu,\dots,\lambda_n+\mu$. 那么对于 A 的高次多项式该如何处理呢?

例 10. 设 A 的特征值为 $\lambda_1,\lambda_2,\dots,\lambda_n$. 求 A^2 的所有特征值.

解. 因为

$$\begin{aligned}\det(\lambda I-A^2)&=\det(-(\sqrt{\lambda}I-A)(-\sqrt{\lambda}I-A))\\&=(-1)^n\det(\sqrt{\lambda}I-A)\det(-\sqrt{\lambda}I-A)\\&=(\sqrt{\lambda}-\lambda_1)(\sqrt{\lambda}-\lambda_2)\dots(\sqrt{\lambda}-\lambda_n)\cdot\\&\quad(\sqrt{\lambda}+\lambda_1)(\sqrt{\lambda}+\lambda_1)\cdots(\sqrt{\lambda}+\lambda_n)\\&=(\lambda-\lambda_1^2)(\lambda-\lambda_2^2)\cdots(\lambda-\lambda_n^2).\end{aligned}$$

所以 A^2 的所有特征值为 $\lambda_1^2,\lambda_2^2,\dots,\lambda_n^2$. □

例 11. 证明: 设 n 阶方阵 A 的全部特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$, 则对任意的多项式 f. 矩阵 $f(A)$ 的特征值为 $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$.

一个不难验证的事实是, 如果 $\lambda, \boldsymbol{x}$ 的 A 的特征值和特征向量, 那么 $f(\lambda), \boldsymbol{x}$ 是 $f(A)$ 的特征值和特征向量. 我们还要保证特征值的重数不会变化, 而且不会出现新的特征值.

证明. 我们先证明

对任意的多项式 $g(x)$, $|g(A)| = g(\lambda_1)g(\lambda_2)\cdots g(\lambda_n)$.

不妨设 $g(x)$ 是首 1 的多项式. 设

$$g(x) = (x-\mu_1)(x-\mu_2)\cdots(x-\mu_n).$$

则

$$g(A) = (A-\mu_1 I)(A-\mu_2 I)\cdots(A-\mu_n I).$$

再考虑多项式

$$g(\lambda_i) = (\lambda_i-\mu_1)(\lambda_i-\mu_2)\cdots(\lambda_i-\mu_n).$$

而

$$\det(\lambda I - A) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

从而

$$\det(A-\mu_i I) = (\lambda_1-\mu_i)(\lambda_2-\mu_i)\cdots(\lambda_n-\mu_i).$$

所以

$$|g(A)| = \prod_{i,j=1}^{n}(\lambda_i-\mu_j) = \prod_{i=1}^{n} g(\lambda_i).$$

即证.

下面考虑 $\det(\lambda I - f(A))$. 设 $g(x) = \lambda - f(x)$.

根据上面的式子可以知道

$$|g(A)| = (\lambda - f(\lambda_1))(\lambda - f(\lambda_2))\cdots(\lambda - f(\lambda_n)).$$

所以全部特征值为 $f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)$. □

例 12. 设 n 阶循环矩阵 C 为

$$\begin{pmatrix} c_1 & c_2 & \ldots & c_n \\ c_n & c_1 & \ldots & c_{n-1} \\ \vdots & \vdots & & \vdots \\ c_2 & c_3 & \ldots & c_1 \end{pmatrix}$$

求 C 的所有特征值及相应的特征向量; 求 $\det(C)$.

解. 设 $A = \begin{pmatrix} \mathbf{0} & I_{n-1} \\ 1 & \mathbf{0} \end{pmatrix}$, 则 $C = c_1 I + c_2 A + c_3 A^2 + \cdots + c_n A^{n-1}$. 设 $f(x) = c_1 + c_2 x + \cdots + c_n x^{n-1}$.

由于 $\det(\lambda I - A) = \lambda^n - 1 = \prod_{i=0}^{n-1}(\lambda - \omega_k)$. 其中 $\omega_k = \mathrm{e}^{\frac{k}{n}2\pi\mathrm{i}}(k = 0, 1, \ldots, n-1)$ 为 n 次单位根.

所以 A 的特征值为 $\omega_0, \omega_1, \ldots, \omega_{n-1}$, 对应特征向量为 $\boldsymbol{x}_k = (1, \omega_k, \ldots, \omega_k^{n-1})^T (k = 0, 1, \ldots, n-1)$. 且互不相同.

因此 C 的特征值为 $f(\omega_k), k = 0, 1, \ldots, n-1$, 对应的特征向量分别为 $\boldsymbol{x}_k, k = 0, 1, \ldots, n-1$.

进而 $\det(C) = f(1)f(\omega_1)\cdots f(\omega_{n-1})$. □

3 矩阵的可对角化

我们说 A 与 B 相似, 是指存在可逆矩阵 P, 使得 $B = P^{-1}AP$. 其是一种等价关系, 记作 $A \sim B$. 一些基本结论:

- 两个相似的方阵有相同的特征值, 反之, 不成立, 即有相同的两个特征值组的方阵未必相似 (几何重数可能不一致, 但实际上即使几何重数一致, 也可能不相似);
- n 阶方阵可对角化, 即相似于对角矩阵 $\Leftrightarrow$ A 有 n 个线性无关的特征向量.

$$T^{-1}AT = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),$$

  其中 $T = \begin{pmatrix} T_1 & T_2 & \ldots & T_n \end{pmatrix}$. 这个式子成立当且仅当 $AT_i = \lambda_i T_i, i = 1, 2, \ldots, n$.
- 任意两个特征值不同的方阵一定可以对角化 (每个特征值的几何重数为 1, 等于代数重数 1).

例 13. 证明两个实数域上的的矩阵 A 与 B 在复数域上相似, 则它们在实数域上亦相似.

证明. 因为 A 与 B 在复数域上相似, 即存在可逆矩阵 $T = P + \mathrm{i}Q$, 使得 $T^{-1}AT = B$. 其中 P 和 Q 为实数域上的矩阵.

因为 $AT = TB$, 所以

$$AP = PB, AQ = QB.$$

因为 T 可逆, 所以可以考虑 $P + \lambda Q$. 因为

$$\det(P + \lambda Q)$$

是关于 λ 的多项式, 因为 $P + iQ$ 可逆, 所以其不为零多项式. 由于其根只有有限多个, 从而存在某个实数 λ', $\lambda = \lambda'$ 时其不为 0, 于是 $P + \lambda' Q$ 可逆, 且有

$$A(P + \lambda' Q) = (P + \lambda' Q)B.$$

其中 $P + \lambda' Q$ 为实数域上的可逆矩阵. 所以 A 和 B 在实数域上也相似. □

例 14. 如果 A 和 B 相似, C 和 D 相似, $\begin{pmatrix} O & A \\ C & O \end{pmatrix}$ 与 $\begin{pmatrix} O & B \\ D & O \end{pmatrix}$ 是否相似?

解. 不一定. 主准对角线是对的, 副准对角线不一定.

比如 $\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \sim \begin{pmatrix}1 & 0\\ 1 & -1\end{pmatrix}$. 而

$$A = \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & -1\\ 1 & 0 & 0\end{pmatrix}, B = \begin{pmatrix}0 & 1 & 0\\ 0 & 1 & -1\\ 1 & 0 & 0\end{pmatrix}$$

不相似, 因为特征多项式分别为

$$\det(\lambda I - A) = \lambda^3 + 1 \neq \det(\lambda I - B) = \lambda^3 - \lambda^2 + 1.$$

因此两者不相似. □

3.1 可对角化的本质 — 特征向量的伸缩

矩阵能不能对角化, 主要是特征向量在起作用. 如果方阵 A 有 n 个线性无关的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$(线性无关是为了构成一组基), 则 $\mathscr{A} : \boldsymbol{x} \mapsto A\boldsymbol{x}$ 在自然基下的矩阵为 A, 在特征向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$ 下的矩阵就是对角矩阵, 对角元为相应的特征值. 而特征值的大小主要作用是影响特征向量的伸缩的大小, 但由于线性变换的特性, 伸缩不一致的特征向量组一定线性无关.

反之, 如果一个矩阵可以对角化, 那么一定可以找到 n 个线性无关的特征向量.

- 矩阵不同特征值的代数重数之和为 n;
- 每个特征值的几何重数小于等于代数重数;

从而, 我们知道, 如果想要有 n 个线性无关的特征向量, 即几何重数的和达到 n, 必有代数重数等于几何重数.

例 15. 方阵 $A = \begin{pmatrix}1 & 1 & 0\\ 0 & 2 & 1\\ 0 & 0 & 3\end{pmatrix}$ 的相似标准形为 $\underline{\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{pmatrix}}$.

解. 特征值互不相同, 可以对角化, 且为对角元素. □

例 16 (秩 1 矩阵)**.** 设 $A = \begin{pmatrix}a_1b_1 & a_1b_2 & \ldots & a_1b_n\\ a_2b_1 & a_2b_1 & \ldots & a_2b_n\\ \vdots & \vdots & & \vdots\\ a_nb_1 & a_nb_2 & \ldots & a_nb_n\end{pmatrix}$. 求 A 的特征值. 并讨论 A 是否可以对角化.

解. (1) 当 A 存在非零行时, $A\boldsymbol{x} = \mathbf{0}$ 的系数矩阵的秩为 1, 所以解空间的维数为 $n-1$. 所以 $n-1$ 个属于 0 特征向量, 所以 0 的代数重数至少为 $n-1$. 又因为 $\operatorname{tr}(A)$ 是 A 所有的特征值的和, 所以剩下一个特征值为 $\operatorname{tr}(A)$. 因此 A 的全部特征值为 $0, 0, \ldots, 0, \sum_{i=1}^n a_ib_i$. A 为零矩阵的时候的特征值显然也是这个形式.

(2) 当 $\operatorname{tr}(A) \neq 0$ 时, 特征值 0 的代数重数和几何重数都为 $n-1$, 又特征值 $\operatorname{tr}(A) \neq 0$, 其对应的特征向量与属于 0 的 $n-1$ 个线性无关的向量可以构成一组基. 以这组基为列向量构成的方阵 T, 可以使

得 A 对角化，即

$$T^{-1}AT = \begin{pmatrix} \mathrm{tr}(A) & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}.$$

若 $\mathrm{tr}(A) = 0, A \neq 0$, 因为特征值 0 的代数重数为 n, 而几何重数为 $n-1$, 因此不能对角化. $A = 0$ 本身就是对角矩阵.

或者根据

$$\begin{aligned} \det(\lambda I - A) &= \det(\lambda I - \begin{pmatrix} a_1 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix}) \\ &= \lambda^{n-1}(\lambda - \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix} \begin{pmatrix} a_1 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}) \\ &= \lambda^{n-1}(\lambda - \sum_{i=1}^{n} a_i b_i). \end{aligned}$$

由此也可以得到 A 的特征值. □

根据这些特征值, 我们知道 A 的特征多项式为 $f(\lambda) = \lambda^n - \mathrm{tr}(A)\lambda^{n-1}$.

例 17. 说明

- $\begin{pmatrix} 1 & 2 & \dots & n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$;
- 设 A 可逆, 则 $\begin{pmatrix} A & O \\ B & O \end{pmatrix} \sim \begin{pmatrix} A & O \\ O & O \end{pmatrix}$.

解. (1) 设 $T = \begin{pmatrix} 1 & -2 & \dots & -n \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$, 则 $T^{-1}AT = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$.

(2) 设 $T = \begin{pmatrix} I & O \\ BA^{-1} & I \end{pmatrix}$, 则 $T^{-1} \begin{pmatrix} A & O \\ B & O \end{pmatrix} T = \begin{pmatrix} A & O \\ O & O \end{pmatrix}$. □

例 18. 如果 $A^2 = A$, 则 A 相似于它的相似于它的相抵标准形 $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 $r = r(A)$.

证明. 这里用相抵标准形的证法.

因为 $r(A)=r$, 所以存在 n 阶可逆矩阵 P 和 Q, 使得

$$A=P\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}Q.$$

考察 QAQ^{-1},

$$QAQ^{-1}=QP\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}.$$

因为 $A^2=A$(我们可以由此看到, 相似于幂等的矩阵也是幂等矩阵), 所以

$$(QAQ^{-1})^2=QAQ^{-1}.$$

即

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

因为 QP 可逆, 设 $R=QP$, 则有

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

可以将 R 对应分块为 $\begin{pmatrix}R_1 & R_2\\ R_3 & R_4\end{pmatrix}$. 则可以得到

$$R_1=I_r.$$

因此

$$QAQ^{-1}=\begin{pmatrix}I_r & R_2\\ R_3 & R_4\end{pmatrix}\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}=\begin{pmatrix}I_r & O\\ R_3 & O\end{pmatrix}.$$

因此根据上例结果可以知道 $A\sim\begin{pmatrix}I_r & O\\ R_3 & O\end{pmatrix}\sim\begin{pmatrix}I_r & O\\ O & O\end{pmatrix}$. □

或者利用本章的方法, 得到 $r(A)+r(I-A)=n$, 得到 0 和 1 的几何重数和为 n, 从而可以对角化. 对角元素只能 1 或 0.

例 19. 令 $A=\begin{pmatrix}1 & \frac{1}{2} & 0\\ \frac{1}{2} & 1 & -\frac{1}{2}\\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\end{pmatrix}$, $\boldsymbol{\alpha}=(3,-1,2)^T$, 求 $\lim_{n\to\infty}A^n\boldsymbol{\alpha}$.

解. 因为 $\det(\lambda I-A)=(\lambda-1)^2(\lambda-\frac{1}{2})$. 所以 A 的特征值为 1 和 $\frac{1}{2}$.

特征值 λ	线性方程组 $(\lambda I - A)\boldsymbol{x} = \mathbf{0}$	属于 λ 特征向量
$\lambda_1 = \lambda_2 = 1$	$\begin{pmatrix} 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	$c_1(1,0,1)^T(c_1 \neq 0)$
$\lambda_3 = \frac{1}{2}$	$\begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	$c_2(1,-1,0)^T(c_2 \neq 0)$

因为特征值 1 的几何重数为 1, 而代数重数为 2. 因此 A 不能对角化.

但是注意到 $\boldsymbol{\alpha} = 2\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. 因此 $A^n\boldsymbol{\alpha} = 2A^n\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + A^n\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

因此

$$\lim_{n\to\infty} A^n\boldsymbol{\alpha} = \lim_{n\to\infty}(2\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{2^n}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}) = 2\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = (2,0,2)^T.$$

□

3.2 可对角化的天敌 — 幂零的非零矩阵

例 20. 设 T 是 n 维线性空间 V 的线性变换, $n > 1, \boldsymbol{\alpha} \in V$. 设 $T^n(\boldsymbol{\alpha}) = \mathbf{0}$, 但是 $T^{n-1}\boldsymbol{\alpha} \neq \mathbf{0}$.

证明: 向量组 $\boldsymbol{\alpha}, T\boldsymbol{\alpha}, \ldots, T^{n-1}\boldsymbol{\alpha}$ 线性无关. 并且 T 不能对角化.

解. 设

$$k_1\boldsymbol{\alpha} + k_2T\boldsymbol{\alpha} + \cdots + k_nT^{n-1}\boldsymbol{\alpha} = \mathbf{0}.$$

其中 $k_1, k_2, \ldots, k_n \in F$. 对两边同时用 T 作用 $(n-1)$ 次得到

$$k_1T^{n-1}\boldsymbol{\alpha} = \mathbf{0}.$$

由于 $T^{n-1}\boldsymbol{\alpha} \neq \mathbf{0}$, 所以 $k_1 = 0$. 所以

$$k_2T\boldsymbol{\alpha} + \cdots + k_nT^{n-1}\boldsymbol{\alpha} = \mathbf{0}.$$

对两边同时用 T 作用 $(n-2)$ 次得到

$$k_2T^{n-1}\boldsymbol{\alpha} = \mathbf{0}.$$

所以 $k_2 = 0$. 重复上述过程就可以得到 $k_1 = k_2 = \cdots = k_n = 0$. 所以 $\boldsymbol{\alpha}, T\boldsymbol{\alpha}, \ldots, T^{n-1}\boldsymbol{\alpha}$ 线性无关.

因为 V 是 n 维线性空间, 所以 $\boldsymbol{\alpha}, T\boldsymbol{\alpha}, \ldots, T^{n-1}\boldsymbol{\alpha}$ 是 V 的一组基. 所以假设 $T(\boldsymbol{\beta}) = \lambda\boldsymbol{\beta}(\boldsymbol{\beta} \neq 0)$.

由于 $\boldsymbol{\beta}$ 可以用 $\boldsymbol{\alpha}, T\boldsymbol{\alpha}, \ldots, T^{n-1}\boldsymbol{\alpha}$ 线性表示, 所以 $T^n\boldsymbol{\beta} = \mathbf{0} = \lambda^n\boldsymbol{\beta}$. 因此 $\lambda = 0$. 所以, T 的特征值只能全为 0, 若其在某组基下的矩阵为对角矩阵. 则 T 将这组基映成 0, 从而只能为零变换. 与 $T^{n-1}\boldsymbol{\alpha} \neq 0$ 矛盾. □

例 21. 设 $A = (a_{ij})$ 是数域 F 上的 n 阶上三角形矩阵, 证明:

(1) 若 $a_{11}, a_{22}, \ldots, a_{nn}$ 两两不等, 则 A 可对角化;

(2) 若 $a_{11} = a_{22} = \cdots = a_{nn}$, 且至少存在一个 $a_{kl} \neq 0 (k < l)$, 则 A 不可对角化.

证明. (1) 因为

$$\det(\lambda I - A) = (\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn}).$$

所以 A 的特征值为 $a_{11}, a_{22}, \ldots, a_{nn}$, 互不相同, 因此 A 可对角化.

(2) 若 $a_{11} = a_{22} = \cdots = a_{nn}$, 则 A 的全部特征值为 $a_1 1$, 而

$$r(a_{11} I - A) < n - 1.$$

于是 A 不可对角化. □

$A + \lambda I$ 与 A 的特征值差个 λ, 具有完全相同的特征向量. 因此 $A + \lambda I$ 可对角化当且仅当 A 可对角化. 根据严格上三角形矩阵是幂零矩阵, 因此不可对角化, 所以你可以由此得到相同的结论.

4 方阵与多项式

4.1 线性代数最优雅的定理 — Hamilton-Caley 定理

哈密顿-凯莱定理. 设 A 是数域 F 上的 n 阶矩阵,

$$f(\lambda) = |\lambda I - A| = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n$$

是 A 的特征多项式, 则

$$f(A) = A^n + a_1 A^{n-1} + \cdots + a_n I = O.$$

回顾前面秩为 1 的方阵, 我们就知道有 $A^2 - \mathrm{tr}(A) A = O$, 即 $A^2 = \mathrm{tr}(A) A$.

在证明这个定理之前, 我们先证明几个例题.

例 22 (特征向量扩充的艺术)**.** 复数域上任一方阵必相似于上三角方阵.

证明. 数学归纳法.

$n = 1$ 结果显然成立. 假设结论对 $n - 1$ 阶矩阵成立, 设 A 是 n 阶矩阵.

设 λ 是 n 阶方阵 A 的特征值, 则存在非零向量使得 $A\boldsymbol{\alpha}_1 = \lambda \boldsymbol{\alpha}_1$. 我们把 $\boldsymbol{\alpha}_1$ 扩充成 $\mathbb{C}^n$ 中的一组基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n$. 则有

$$A(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n) \begin{pmatrix} \lambda & * \\ \mathbf{0} & A_1 \end{pmatrix},$$

其中 A_1 是一个 $n - 1$ 阶复方阵. 设 $T = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_n)$, 则 T 可逆, 且有

$$T^{-1} A T = \begin{pmatrix} \lambda & * \\ \mathbf{0} & A_1 \end{pmatrix}.$$

由假设归纳可知, 存在 $n - 1$ 阶可逆矩阵 P, 使得 $P^{-1} A_1 P$ 是一个上三角形方阵. 令 $C = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & P \end{pmatrix}$.

则

$$(TC)^{-1}ATC = C^{-1}T^{-1}ATC = \begin{pmatrix} \lambda & * \\ \mathbf{0} & P^{-1}A_1P \end{pmatrix},$$

这是一个上三角形方阵. 证毕. □

例 23. 设 A 是一个上三角形方阵, $f(x)$ 是 A 的特征多项式, 则 $f(A) = O$.

证明. 设上三角形的方阵

$$A = \begin{pmatrix} \lambda_1 & a_{12} & \dots & a_{1n} \\ & \lambda_2 & \dots & a_{2n} \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix},$$

其主对角线上的元素正好是它的全部特征值. 则

$$f(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n).$$

于是设 $A_1 = A - \lambda_1 I, A_2 = A - \lambda_2 I, \dots, A_n = A - \lambda_n I$. 因为其都是关于 A 的多项式, 所以都可以交换. 因为

$$A_1 A_2 \dots A_i \boldsymbol{e}_i = \mathbf{0}.$$

所以根据交换性, $f(A)(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n) = f(A) = O$. □

现在我们可以证明凯莱-哈密顿定理: 存在可逆矩阵 T, 使得 $T^{-1}AT = B$ 是个上三角形方阵. 其与 A 有相同的特征多项式，且 $f(B) = \mathrm{O}$. 从而 $f(A) = Tf(B)T^{-1} = O$.

上三角形方阵一些性质:

- 上三角形方阵相加, 相乘, 取逆, 取伴随都是上三角形矩阵;
- 上三角形方阵的特征值, 是它的全部对角元素;
- 上三角形方阵相加, 相乘得到的方阵的对角元素是对应对角元素相加, 相乘.

根据上三角方阵的这些性质, 你很容易通过化成上三角形的方式得到逆矩阵, 伴随矩阵, 甚至前面说的多项式的特征值.

4.2 高山流水 — Hamilton-Caley 定理的应用

根据哈密顿-凯莱定理, 我们很容易得到可逆方阵 A 的逆可以表示成关于 A 的多项式, 进而很容易得到任意矩阵的伴随矩阵也可以表示成关于它的多项式.

例 24. 设 $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$.

(1) 证明: $A^n = -A^{n-2} + A^2 + I (n \geq 3)$;

(2) 计算 A^{103} 和 A^{102}.

解. A 的特征多项式为

$$f(\lambda)=|\lambda I-A|=\lambda^3-\lambda^2+\lambda-1.$$

由哈密顿-凯莱定理可以知道

$$A^3-A^2+A-I=O,$$

即 $A^3+A=A^2+I$.

(1) $A^m+A^{m-2}=A^{m-2}(A^2+I)=A^{m-2}(A^3+A)=A^{m+1}+A^{m-1}(m\geq 3)$.

所以

$$A^n+A^{n-2}=A^{n-1}+A^{n-3}=\cdots=A^2+I.$$

即有 $A^n=-A^{n-2}+A^2+I$.

(2) $A^{103}=-A^{101}+A^2+I=A^{99}-A^2-I+A^2+I=A^{99}=\cdots=A^3=-A+A^2+I=\begin{pmatrix}1&0&1\\0&0&-1\\0&1&0\end{pmatrix}$.

$A^{102}=A^2=\begin{pmatrix}1&1&1\\0&-1&0\\0&0&-1\end{pmatrix}$. □

例 25. 求 A^{500}, 其中

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

解. A 的特征多项式为 $f(\lambda)=|\lambda I-A|=\lambda^4-\lambda^3$. 所以 $A^4-A^3=O$, 即 $A^4=A^3$.

从而 $A^{500}=A^3=\begin{pmatrix}1&0&0&0\\-1&0&0&0\\1&0&0&0\\2&0&0&0\end{pmatrix}$. □

5 综合应用

5.1 求方阵的高次幂

例 26. 设实数数列 $\{a_n\}$ 满足递推关系式

$$a_{k+1}=\frac{1}{2}(a_{k+1}+a_k)(k=0,1,2,\ldots)$$

以及初始条件 $a_0=0,a_1=\frac{1}{2}$, 求这个数列的通项公式, 并且求出 $\lim_{n\to\infty}a_n$.

解. 由数列的递推关系式得到

$$\begin{pmatrix}a_{k+2}\\a_{k+1}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\1&0\end{pmatrix}\begin{pmatrix}a_{k+1}\\a_k\end{pmatrix}.$$

设 $A=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\1&0\end{pmatrix}$, 则

$$|\lambda I-A|=(\lambda-1)(\lambda+\frac{1}{2}).$$

特征值 λ	线性方程组 $(\lambda I-A)\boldsymbol{x}=\mathbf{0}$	属于 λ 特征向量
$\lambda_1=1$	$\begin{pmatrix}\frac{1}{2}&-\frac{1}{2}\\-1&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$	$c_1(1,1)^T(c_1\neq 0)$
$\lambda_2=-\frac{1}{2}$	$\begin{pmatrix}-1&-\frac{1}{2}\\-1&-\frac{1}{2}\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$	$c_2(1,-2)^T(c_2\neq 0)$

由此, 可以知道, 设 $T=\begin{pmatrix}1&1\\1&-2\end{pmatrix}$, 则有 $T^{-1}AT=\begin{pmatrix}1&0\\0&-\frac{1}{2}\end{pmatrix}$.

从而

$$\begin{pmatrix}a_n\\a_{n-1}\end{pmatrix}=A^{n-1}\begin{pmatrix}\frac{1}{2}\\0\end{pmatrix}=T\begin{pmatrix}1&0\\0&(-\frac{1}{2})^{n-1}\end{pmatrix}T^{-1}\begin{pmatrix}\frac{1}{2}\\0\end{pmatrix}=-\frac{1}{3}\begin{pmatrix}-1+(-\frac{1}{2})^{n}\\-1+(-\frac{1}{2})^{n-1}\end{pmatrix}.$$

因此 $a_n=-\frac{1}{3}(-1+(-\frac{1}{2})^n)$. 当 $n\to\infty$ 时, $a_n\to\frac{1}{3}$. □

5.2 交换的 A,B

例 27. *设 n 阶方阵 A 有 n 个互不相同的特征值. 证明: 如果 n 阶方阵 B 满足 $AB=BA$, 那么 B 相似于对角矩阵, 并且 B 可以表示为 A 的多项式.*

证明. 因为 A 有 n 个互不相同的特征值, 设为 $\lambda_1,\lambda_2,\ldots,\lambda_n$. 因此可以对角化, 即存在 n 阶可逆矩阵 T, 使得 $T^{-1}AT=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

又

$$T^{-1}ATT^{-1}BT=T^{-1}ABT=T^{-1}BAT=T^{-1}BTT^{-1}AT.$$

设 $D=T^{-1}AT$, 则对角矩阵 D 与 $T^{-1}BT$ 可以交换. 由于对角矩阵 D 的对角元素互不相同, 因此 $T^{-1}BT$ 只能为对角矩阵, 设为 $\mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_n)$.

于是存在 $n-1$ 次多项式 f, 使得 $f(\lambda_i)=\mu_i(1\leq i\leq n)$. 从而 $T^{-1}BT=T^{-1}f(A)T$, 即 $B=f(A)$. □

例 28. *设 A,B 为 n 阶复方阵, 如果 $AB=BA$, 则 A,B 一定有公共的特征向量.*

解. 设 $V_\lambda(A)$ 为 A 属于 λ 的特征子空间. 设 $\boldsymbol{x}\in V_A(\lambda)$, 则 $AB\boldsymbol{x}=BA\boldsymbol{x}=\lambda B\boldsymbol{x}$. 所以 $B\boldsymbol{x}\in V_\lambda(A)$. 于是 B 在子空间 $V_\lambda(A)$ 一定有一个特征向量, 即为公共特征向量. □

例 29. *设 A,B 是复数域 $\mathbb{C}$ 上的 n 阶方阵, 且 $AB=BA$. 求证: 存在一个 n 阶可逆矩阵 P, 使得 $P^{-1}AP$ 与 $P^{-1}BP$ 都是上三角形矩阵.*

解. 用数学归纳法. $n=1$ 显然成立. 假设对 $n-1$ 阶复矩阵成立, 对于 n 阶复矩阵 A,B, 由于 $AB=BA$, 所以 A,B 有公共的特征向量, 设为 $\boldsymbol{x}$.

设

$$Ax = \lambda x, Bx = \mu x.$$

将 $\boldsymbol{x}$ 扩充为 $\mathbb{C}^n$ 的一组基 $\boldsymbol{x}, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$. 设 T 以这组基为列向量, 则有

$$AT = T \begin{pmatrix} \lambda & * \\ \mathbf{0} & A_1 \end{pmatrix}, BT = T \begin{pmatrix} \mu & * \\ \mathbf{0} & B_1 \end{pmatrix}$$

根据 $AB = BA$ 可以得到 $A_1B_1 = B_1A_1$, 根据归纳假设, 存在 $n-1$ 阶可逆矩阵 T_1, 使得 $T_1^{-1}A_1T_1$, $T_1^{-1}B_1T_1$ 同时为上三角形矩阵. 设 $S = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & T_1 \end{pmatrix}$.

则

$$(TS)^{-1}A(TS), (TS)^{-1}B(TS).$$

均为上三角形矩阵. 即 $P = TS$ 即可, 容易看到其确实是一个可逆矩阵. □

例 30. *设 A, B 是 n 阶矩阵, 且 $AB = BA$. 证明: 如果 A, B 都可对角化, 那么存在可逆矩阵 P, 使得 $P^{-1}AP, P^{-1}BP$ 同时为对角矩阵.*

解. 因为 A 可对角化, 从而存在 n 阶可逆矩阵 T, 使得

$$T^{-1}AT = \begin{pmatrix} \lambda_1 I_{m_1} & & & \\ & \lambda_2 I_{m_2} & & \\ & & \ddots & \\ & & & \lambda_s I_{m_s} \end{pmatrix},$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_s$ 为 A 互不相同的特征值, $m_1 + m_2 + \ldots + m_s = n$. 由 $AB = BA$ 可以得到

$$T^{-1}AT, T^{-1}BT$$

也可以交换. 可以对 $T^{-1}BT$ 对应分块为 $(B_{ij})_{s\times s}$. 由于可以交换, 可以得到 $B_{ij} = O(i \neq j)$. 所以

$$T^{-1}BT = \mathrm{diag}(B_{11}, B_{22}, \ldots, B_{ss}).$$

因为 B 可对角化, 所以 $T^{-1}BT$ 可以对角化, 因此每个对角块可以对角化. 即存在 T_i 使得 $T_i^{-1}B_{ii}T_i$ 是一个对角矩阵. 设

$$P = T\mathrm{diag}(T_1, T_2, \ldots, T_s),$$

就有 $P^{-1}AP, P^{-1}BP$ 同时为对角矩阵. □

6 写出三阶矩阵的特征向量 — 瞪眼法

本节仅为个人小结, 仅供参考. 不希望大家在不必要的计算中花费太多时间, 甚至出现计算错误. 本节用大写字母 C 表示一个常数.

首先我们要知道一个基本的事情是, 与非零向量 (a, b) 正交的一个向量是 $(b, -a)$, 即交换取反.

在开始之前, 我们先看一下各个秩的三阶矩阵的样子, 由于求特征向量时系数矩阵的行列式是为 0 的, 所以我们主要关注不可逆矩阵的样子.

三阶矩阵相抵下的分类

- 矩阵秩为 0, 只有零矩阵 O;
- 矩阵秩为 1, 每一行是非零行的常数倍;
- 矩阵秩为 2, 行列式为 0 但不符合上面两个;
- 矩阵秩为 3, 行列式不为 0.

也就是, 在不可逆非零矩阵中, 如果每一行是非零行的常数倍, 秩就是 1, 否则就为 2.

如果 $\lambda I - A$ 的秩为 1(观察到每一行都是非零行的倍数), 我们只需提取其中一行, 其解空间的维数为 2. 对应有下面两种情况.

秩为 1 的矩阵	
$(a,b,c)(a\neq 0)$	$C_1(b,-a,0)^T + C_2(c,0,-a)^T(C_1, C_2$ 不同时为零)
$(0,b,c)$	$C_1(1,0,0)^T + C_2(0,c,-b)^T(C_1, C_2$ 不同时为零)

如果 $\lambda I - A$ 的秩为 2(由于其行列式为 0 且自身非零矩阵, 秩不为 1 肯定就是 2), 我们提取两个最简单的线性无关的行即可, 可以通过剩下的行检验解的正确性.

秩为 2 的矩阵		
$\begin{pmatrix} a_1 & k_1b_1 & k_1b_2 \\ a_2 & k_2b_1 & k_2b_2 \end{pmatrix}$	只有两列成比例	$C(0,b_2,-b_1)^T(C\neq 0)$
$\begin{pmatrix} a & 0 & 0 \\ b & c & d \end{pmatrix}$	只有两列成比例的特殊情况	$C(0,d,-c)^T(C\neq 0)$
$\begin{pmatrix} a & b & 0 \\ c & d & 0 \end{pmatrix}$	相同位置都是 0	$C(0,0,1)^T(C\neq 0)$
$\begin{pmatrix} a & b & 0 \\ c & 0 & b \end{pmatrix}$	不同位置的 0, 相同的数	$C(b,-a,-c)^T(C\neq 0)$
$\begin{pmatrix} a & b & c \\ 0 & d & e \end{pmatrix}$	有一个为 0	$C(\text{待解},e,-d)^T(C\neq 0)$

一般的可以通过一次初等行变换, 得到一个 0 即可.

一个特别的形式:

$$\begin{pmatrix} -(n-1)a & a & \dots & a \\ a & -(n-1)a & \dots & a \\ \vdots & \vdots & & \vdots \\ a & a & \dots & -(n-1)a \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

$n\geq 2, a\neq 0$ 时, 设系数矩阵为 A. 通解为 $C(1,1,\dots,1)^T(C\neq 0)$.

同时, 我们可以做一下初等变换

$$D_i(\lambda^{-1})AD_i(\lambda)\boldsymbol{x} = \boldsymbol{0}$$

的通解为 $C(1,1,\dots,\lambda^{-1},\dots,1)(C\neq 0)$.

$$S_{ij}AS_{ij}\boldsymbol{x} = \boldsymbol{0}$$

的通解仍为 $C(1,1,\dots,1)(C\neq 0)$. 你能看出来它们是什么样子吗?