
目录

| | | |
|-------|-----------------|----|
| 1 | 矩阵基本概念..... | 4 |
| 1.1 | 行列式..... | 4 |
| 1.1.1 | n 阶行列式..... | 4 |
| 1.1.2 | 行列式性质..... | 4 |
| 1.2 | 矩阵类型..... | 5 |
| 1.3 | 矩阵行列式..... | 5 |
| 1.4 | 逆矩阵..... | 5 |
| 1.5 | 矩阵的秩..... | 6 |
| 1.6 | 矩阵的迹..... | 6 |
| 1.7 | 特征值和特征向量..... | 6 |
| 1.8 | 相似矩阵..... | 7 |
| 2 | 欧式空间和酉空间..... | 8 |
| 2.1 | 共轭转置..... | 8 |
| 2.2 | 欧式空间与酉空间..... | 9 |
| 2.3 | 内积..... | 9 |
| 2.3.1 | 内积运算..... | 10 |
| 2.3.2 | 内积性质..... | 10 |
| 2.4 | 模(范数)..... | 11 |
| 2.5 | 正交..... | 11 |
| 2.6 | 酉(U)阵..... | 12 |
| 2.6.1 | U 阵定义..... | 12 |
| 2.6.2 | U 阵性质..... | 13 |
| 2.7 | Jordan 标准形..... | 13 |
| 2.7.1 | 概念..... | 13 |
| 2.7.2 | 基本求法..... | 14 |
| 2.7.3 | 幂零性质..... | 14 |
| 3 | 矩阵分解..... | 16 |
| 3.1 | QR 分解..... | 16 |
| 3.1.1 | 定理..... | 16 |
| 3.1.2 | Q、R 的求法..... | 16 |
| 3.2 | 镜面阵..... | 16 |
| 3.2.1 | 定义..... | 16 |

| | |
|------------------------------|----|
| 3.2.2 性质 | 17 |
| 3.2.3 扩展性质 | 17 |
| 3.2.4 应用 | 17 |
| 3.3 秩一/满秩分解 | 18 |
| 3.3.1 秩一分解 | 18 |
| 3.3.2 秩一方阵公式 | 19 |
| 3.3.3 满秩分解 | 19 |
| 3.3.4 满秩分解的求法 | 19 |
| 3.4 正规矩阵及 Schur 分解 | 20 |
| 3.4.1 Schur 引理 | 20 |
| 3.4.2 正规矩阵 | 21 |
| 3.4.3 定理与推论 | 23 |
| 3.5 厄米特(Hermite)分解 | 24 |
| 3.5.1 定理 | 24 |
| 3.5.2 应用 | 25 |
| 3.5.3 正定矩阵 | 25 |
| 3.5.4 Hermite 阵的性质 | 25 |
| 3.5.5 常见的 Hermite 阵 | 26 |
| 3.5.6 斜 Hermite 分解 | 26 |
| 3.6 特征值/酉阵 Q 的求解技巧 | 27 |
| 3.6.1 平移法则(求特征根) | 27 |
| 3.6.2 特殊矩阵分解酉阵 Q 的求解方法 | 28 |
| 3.6.3 换位公式(求特征根) | 29 |
| 3.7 奇异值分解 | 31 |
| 3.7.1 一些定理 | 31 |
| 3.7.2 正奇异值 | 32 |
| 3.7.3 简奇异值分解与奇异值分解 | 33 |
| 3.7.4 简奇异值分解方法 | 33 |
| 3.7.5 推论 | 35 |
| 3.8 谱分解 | 38 |
| 3.8.1 单纯阵 | 38 |
| 3.8.2 零化式 | 39 |
| 3.8.3 谱分解 | 41 |
| 3.8.4 谱分解的求法 | 42 |
| 3.8.5 谱分解的应用 | 45 |
| 4 广义逆矩阵 | 48 |
| 4.1 广义逆矩阵的定义 | 48 |
| 4.2 性质 | 49 |
| 4.3 广义逆求解方式 | 50 |

| | | |
|-------|------------------|----|
| 4.4 | 正规方程..... | 51 |
| 4.5 | 最小二乘解..... | 52 |
| 4.6 | 补充公式..... | 54 |
| 4.7 | A -定义与性质..... | 55 |
| 4.7.1 | A -定义..... | 55 |
| 4.7.2 | A -相关公式..... | 56 |
| 4.7.3 | 更多广逆..... | 57 |
| 5 | 矩阵分析..... | 58 |
| 5.1 | 向量范数..... | 58 |
| 5.1.1 | 向量范数..... | 58 |
| 5.1.2 | 常见向量范数..... | 58 |
| 5.2 | 矩阵范数..... | 59 |
| 5.2.1 | 定义..... | 59 |
| 5.2.2 | 常见的矩阵范数..... | 60 |
| 5.2.3 | 谱半径..... | 61 |
| 5.2.4 | 算子范数..... | 62 |
| 5.2.5 | 小范数定理..... | 64 |
| 5.3 | 矩阵级数..... | 66 |
| 5.4 | 特征根估计..... | 68 |
| 5.4.1 | 估计方法..... | 68 |
| 5.4.2 | 盖尔(Ger)圆盘..... | 68 |
| 5.5 | 矩阵函数..... | 70 |
| 5.5.1 | 收敛定理..... | 70 |
| 5.5.2 | 常见解析函数..... | 71 |
| 5.5.3 | 幂等公式..... | 72 |
| 5.5.4 | 分块公式..... | 72 |
| 5.5.5 | 根遗传公式..... | 74 |
| 5.5.6 | Euler 公式..... | 74 |
| 5.5.7 | 幂 0 公式..... | 75 |
| 5.5.8 | 矩阵函数求法总结..... | 75 |
| 5.6 | 矩阵函数应用..... | 77 |
| 5.6.1 | 基本定义与公式..... | 77 |
| 5.6.2 | 应用..... | 77 |
| 5.6.3 | 求解齐次线性微分方程组..... | 78 |
| 6 | 矩阵直积..... | 80 |
| 6.1 | 定义与性质..... | 80 |
| 6.1.1 | 定义..... | 80 |
| 6.1.2 | 基本性质..... | 80 |

1 矩阵基本概念

1.1 行列式

1.1.1 n 阶行列式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{(p_1 p_2 \dots p_n)} (-1)^{\tau(p_1 p_2 \dots p_n)} a_{1p_1} a_{2p_2} \cdots a_{np_n}$$

其中 $p_1 p_2 \dots p_n$ 指的是自然数 $1, 2, \dots, n$ 的一个排列， $\tau(p_1 p_2 \dots p_n)$ 为排列 $p_1 p_2 \dots p_n$ 的逆序数。

逆序数：一个自然数排列中，前面数大于后面数的组合叫做一个逆序，逆序的总和叫做逆序数。

1.1.2 行列式性质

- i. 行列式与其转置行列式相等
- ii. 行列式的任意两行(列)互换位置，行列式变号
- iii. 若行列式中某一行(列)所有元素有公因子 k ，则可把公因子提到行列式记号之外。即：

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ k * a_{i1} & k * a_{i2} & \cdots & k * a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = k * \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

1.2 矩阵类型

实数矩阵、复数矩阵、方阵、行/列矩阵(行/列向量)

系数矩阵：方程组的系数组成的矩阵

单位矩阵：主对角线为 1，其他全为 0 的方阵

对角矩阵：除主对角线以外全是 0 的矩阵，常写为 $\text{diag}(a_1, a_2, \dots, a_n)$

奇异矩阵： $\det(A) = 0$

1.3 矩阵行列式

N 阶矩阵 A 的行列式记作 $\det(A)$ ，行列式的相关定理：

- i. $\det(A) = \det(A^T)$
- ii. $\det(AB) = \det(A) * \det(B)$
- iii. 若 A 为一个 n 阶三角矩阵，则 $\det(A)$ 等于矩阵 A 对角元素的乘积
- iv. A 为 n 阶矩阵，若 A 的某行或某列全为 0，或者 A 有两行或两列相等，则 $\det(A) = 0$

1.4 逆矩阵

设 A 为 n 阶矩阵，在相同数域上存在 n 阶矩阵 B，使得 $AB = I$ (单位矩阵)，则 A、B 互为逆矩阵，非奇异矩阵。

A 的逆矩阵记作 A^{-1} :

- i. 矩阵 A 为可逆矩阵的充要条件为 $\det(A) \neq 0$
- ii. 矩阵 A 可逆, 则矩阵 A 的行/列向量线性无关, 矩阵 A 满秩
- iii. 矩阵 A 可逆, 则 $Ax=b$ 有唯一解(两边同乘以 A^{-1} , $x=A^{-1}b$)
- iv. 逆矩阵具有唯一性
- v. 伴随矩阵 $A^* = |A| * A^{-1}$

1.5 矩阵的秩

矩阵中线性无关的纵列的最大个数叫做列秩, 线性无关的行列的最大个数叫做行秩。

方阵的行秩和列秩相等, 简称为矩阵的秩, 记为 $\text{rank}(A)$ 或 $r(A)$ 或 $\text{rk}(A)$ 。

对于矩阵 A 而言, 其解空间为矩阵 X, 即 $AX=0$ 。则 $r(A)=n-\dim N(X)$ 。
即 A 的秩等于 A 的行数(或列数, 选较小的)-解空间的维度。

1.6 矩阵的迹

N 阶矩阵主对角线元素的和叫做矩阵的迹, 记作 $\text{tr}(A)$ 。

- i. 迹是 n 阶矩阵所有特征值的和
- ii. $\text{tr}(mA + nB) = m * \text{tr}(A) + n * \text{tr}(B)$
- iii. $\text{tr}(AB) = \text{tr}(BA)$

1.7 特征值和特征向量

设 A 是 n 阶方阵, 如果数 λ 和 n 维非零列向量 x 使关系式 $Ax=\lambda x$ 成立, 那么这样的数 λ 称为矩阵 A 特征值, 非零向量 x 称为 A 的对应于特征值 λ

的特征向量。

式 $Ax=\lambda x$ 也可写成 $(A-\lambda E)X=0$ 。这是 n 个未知数 n 个方程的齐次线性方程组，它有非零解的充分必要条件是系数行列式 $|A-\lambda E|=0$ 。

矩阵 A 的特征多项式： $P_A(\lambda) = |\lambda E - A| = \det(\lambda E - A)$

Hamilton-Cayley 定理：

矩阵 A 的特征多项式也是它的零化多项式，即对于矩阵 A 的特征多项式： $P_A(\lambda) = \det(\lambda I - A) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda^1 + a_0$

有： $P_A(A) = \det(AI - A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A^1 + a_0 = 0$

- i. 三角/对角矩阵对角元即为它的特征值
- ii. N 阶矩阵的迹等于其特征值之和
- iii. 相似矩阵特征值相同
- iv. $\lambda(AB) = \lambda(BA)$ (AB 为方阵)
- v. $\lambda(kA) = k * \lambda(A)$
- vi. $\lambda(A^k) = [\lambda(A)]^k$

证明：

设矩阵 A 的特征根为 λ ，则 $AX = \lambda X \rightarrow A^2 X = A\lambda X = \lambda AX = \lambda^2 X$ ，所以 λ^2 为矩阵 A^2 的特征根，同理可证 $\lambda(A^k) = \lambda^k$

1.8 相似矩阵

设 A, B 为 n 阶矩阵，如果有 n 阶可逆矩阵 P 存在，使得： $P^{-1}AP = B$ ，则称矩阵 A 和矩阵 B 相似，记作 $A \sim B$ 。

运算 $P^{-1}AP$ 称为相似变换， P 称为相似变换矩阵。

性质：

- i. $A \sim A$
- ii. 若 $A \sim B$ ，则 $B \sim A$
- iii. 若 $A \sim B$ ， $B \sim C$ ，则 $A \sim C$
- iv. 若 $A \sim B$ ，则 A, B 秩相同、行列式相同、迹相同，特征值和特征多项式相同
- v. 若 $A \sim B$ ，则 A, B 具有相同的可逆性

若矩阵相似于某个对角矩阵，则该矩阵可对角化。

2 欧式空间和酉空间

2.1 共轭转置

数的共轭转置：

实数 x 的共轭仍然是实数本身，复数 $z = a + bi$ 的共轭 $\bar{z} = a - bi$ 。

数 x 的共轭转置 $x^H = \bar{x}$ (复数) = x (实数)

向量的共轭转置：

向量 $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ 的共轭转置 $X^H = (\bar{x}_1 \quad \cdots \quad \bar{x}_n)$

$$X^H Y = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n$$

$$Y^H X = \bar{y}_1 x_1 + \cdots + \bar{y}_n x_n = \overline{\bar{x}_1 y_1 + \cdots + \bar{x}_n y_n} = \overline{X^H Y}$$

矩阵的共轭转置：

矩阵 $A = (X_1 \quad \dots \quad X_n)$, 其中 X_i 是列向量, 矩阵 A 的共轭转置

$$A^H = \begin{pmatrix} X_1^H \\ \dots \\ X_n^H \end{pmatrix}$$

共轭与转置的性质:

- i. $\overline{\overline{A}} = A$
- ii. $(AB)^T = B^T A^T$
- iii. $A^H = A^T$ (实数矩阵)

共轭转置的一些性质:

- i. $(A^H)^H = A$
- ii. $(kA)^H = \bar{k}A^H$
- iii. $(A + B)^H = A^H + B^H \quad (A - B)^H = A^H - B^H$
- iv. $(AB)^H = B^H A^H$

2.2 欧式空间与酉空间

欧式空间: 有限维的实数内积空间

$$R^n = \{ \text{全体实向量 } X = (x_1, x_2, \dots, x_n)^T \mid x_i \in R \wedge 1 \leq i \leq n \}$$

酉空间: 有限维的复数内积空间

$$C^n = \{ \text{全体复向量 } Z = (z_1, z_2, \dots, z_n)^T \mid z_i = a_i + ib \in C \wedge 1 \leq i \leq n \}$$

2.3 内积

在实数 R (复数 C) 的有限维线性空间 V 内, 若 $\forall X, Y \in V$, 有一种规则 (X, Y) 使之对应一个实数(复数), 则称该实数(复数)为 X, Y 的内积, 该规则满足

以下条件:

- i. 对称性: $(X, Y) = (Y, X)$ 实数 $= \overline{(Y, X)}$ 复数
- ii. 可加性: $(X + Y, Z) = (X, Z) + (Y, Z)$
- iii. 齐次性: $(kX, Y) = k(Y, X)$
- iv. 非负性: $(X, X) \geq 0$, 当且仅当 $X = \theta$ 时, $(X, X) = 0$

称 V 为实(复)内积空间。

2.3.1 内积运算

$$\text{令 } X = (x_1, x_2, \dots, x_n)^T, \quad Y = (y_1, y_2, \dots, y_n)^T$$

若 $X, Y \in R^n$,

$$(X, Y) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = Y^H X$$

若 $X, Y \in C^n$,

$$(X, Y) = x_1 \overline{y_1} + x_2 \overline{y_2} + \dots + x_n \overline{y_n} = Y^H X$$

$$\text{所以 } (X, Y) = Y^H X = \overline{X^H Y} = \overline{(Y, X)}$$

2.3.2 内积性质

上述四条: 对称、可加、齐次、非负

- i. $(X, Y) = (Y, X)$ 实数 $= \overline{(Y, X)}$ 复数
- ii. $(kX, Y) = Y^H (kX) = k(Y^H X) = k(Y, X)$
 $(X, kY) = (kY)^H X = \bar{k} Y^H X = \bar{k} (X, Y)$

2.4 模(范数)

定义：非负实数 $\sqrt{(X, X)}$ 称为 X 的长度(模或范数)，记为 $\|X\|$

定理：设 V 是欧氏空间，则：

- i. $\|kX\| = |k| * \|X\|$
- ii. $\|X + Y\|^2 + \|X - Y\|^2 = 2(\|X\|^2 + \|Y\|^2)$
- iii. $|(X, Y)| \leq \|X\| * \|Y\|$
- iv. $\|X + Y\| \leq \|X\| + \|Y\|$

复数向量的模：

复数的模： $|a + bi| = \sqrt{a^2 + b^2}$

复向量的模： $|(a_1 + b_1i, \dots, a_n + b_ni)| = \sqrt{\sum (a_j^2 + b_j^2)}$

2.5 正交

定义：欧氏空间中，若向量 X, Y 满足 $(X, Y) = 0$ ，则称 X 与 Y 正交，记作 $X \perp Y$ 。

定理：

- i. $X \perp Y \leftrightarrow \|X + Y\|^2 = \|X\|^2 + \|Y\|^2$

推广，若 X_1, X_2, \dots, X_n 两两正交，则

$$\left\| \sum_{i=1}^n X_i \right\|^2 = \sum_{i=1}^n \|X_i\|^2$$

- ii. 设 X_1, X_2, \dots, X_n 为 V 中非零向量的两两正交向量组，则 X_1, X_2, \dots, X_n 必线性无关。
- iii. 对于 n 维欧氏空间的任一基 X_1, X_2, \dots, X_n ，均可找到一组标准正交

基。

标准正交基的求法：

设正交基为 Y_1, Y_2, \dots, Y_n .

取 $Y_1 = X_1$ ，令 $Y_2 = X_2 + kY_1$ ，由于 Y_1, Y_2 正交，所以 $(Y_1, Y_2) = 0$

由 $(Y_1, Y_2) = (Y_2, Y_1) = (X_2 + kY_1, Y_1) = (X_2, Y_1) + k(Y_1, Y_1) = 0$ 可得：

$$k = -\frac{(X_2, Y_1)}{(Y_1, Y_1)}$$

令 $Y_3 = X_3 + k_1Y_1 + k_2Y_2$ ，由 $(Y_1, Y_3) = (Y_2, Y_3) = 0$ 可得：

$$k_2 = -\frac{(X_3, Y_2)}{(Y_2, Y_2)}, k_1 = -\frac{(X_3, Y_1)}{(Y_1, Y_1)}$$

因此， $Y_{m+1} = X_{m+1} + \sum_{i=1}^m k_i Y_i$ ，其中：

$$k_i = -\frac{(X_{m+1}, Y_i)}{(Y_i, Y_i)}$$

此时得到了一组正交基 Z_1, Z_2, \dots, Z_n ，将其单位化即可得到标准正交基 Z_1, Z_2, \dots, Z_n .

$$Z_i = \frac{1}{\|Y_i\|} Y_i$$

2.6 酉(U)阵

2.6.1 U 阵定义

1. 若 $A = A_{n \times p} = (\alpha_1, \dots, \alpha_p)$ ，各列正交， $\alpha_1 \perp \dots \perp \alpha_p$ ，则 $A^H A$ 为对角矩阵，称 A 为预备半 U 阵。

$$\begin{aligned}
A^H A &= \begin{pmatrix} \alpha_1^H \\ \vdots \\ \alpha_p^H \end{pmatrix} (\alpha_1 \quad \cdots \quad \alpha_p) = \begin{pmatrix} \alpha_1^H \alpha_1 & \alpha_1^H \alpha_2 & \cdots \\ \alpha_2^H \alpha_1 & \alpha_2^H \alpha_2 & \cdots \\ \vdots & \vdots & \cdots \end{pmatrix} \\
&= \begin{pmatrix} \overline{(\alpha_1, \alpha_1)} & \overline{(\alpha_1, \alpha_2)} & \cdots \\ \overline{(\alpha_2, \alpha_1)} & \overline{(\alpha_2, \alpha_2)} & \cdots \\ \vdots & \vdots & \cdots \end{pmatrix} \\
&= \begin{pmatrix} |\alpha_1|^2 & 0 & \cdots & 0 \\ 0 & |\alpha_2|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & |\alpha_p|^2 \end{pmatrix}
\end{aligned}$$

2. 若上述 A 中的 $\alpha_1, \dots, \alpha_p$ 都是单位长, 则 $A^H A = I_p$, 称 A 为半 U 阵。
3. 若上述 A 为方阵, 则称 A 为 U 阵。

2.6.2 U 阵性质

1. $A^H A = A A^H = I, A^H = A^{-1}$
2. $(AX, AY) = (X, Y)$
 $(AX, AY) = (AY)^H (AX) = Y^H A^H A X = Y^H I X = Y^H X = (X, Y)$
3. $\|AX\|^2 = (AX, AX) = (X, X) = \|X\|^2$
4. $X_1 \perp \cdots \perp X_n \rightarrow AX_1 \perp \cdots \perp AX_n$

2.7 Jordan 标准形

2.7.1 概念

定义:

每个 n 阶的复数矩阵(单纯阵)都相似于一个 Jordan 标准形矩阵, 该矩阵如果不考虑若当块的次序, 则 Jordan 标准形是唯一的。

若当块:

$$J_t(\lambda) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}_{t \times t}$$

特殊的

$$J_1(\lambda) = (\lambda)$$

$$J_t = J_t(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}_{t \times t}$$

若当标准形:

$$J = \begin{pmatrix} J_{t_1}(\lambda_1) & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & J_{t_k}(\lambda_k) \end{pmatrix}_{n \times n} \quad (t_1 + \dots + t_k = n)$$

2.7.2 基本求法

对于 n 阶矩阵 A , 其 Jordan 标准求法如下:

1. 对 $\lambda I - A$ 进行初等行变换, 转换成一个对角阵, 对角元为 1 或者 $a(t)$,

$$\text{其中 } a(t) = (\lambda - \lambda_1)^{t_1} \dots (\lambda - \lambda_j)^{t_j}$$

2. 记所以 $a(t)$ 中的元素为 $(\lambda - \lambda_1)^{t_1}, \dots, (\lambda - \lambda_k)^{t_k}$, 其中 $(t_1 + \dots + t_k = n)$

3. 则 $J = \begin{pmatrix} J_{t_1}(\lambda_1) & 0 & 0 \\ 0 & \dots & 0 \\ 0 & 0 & J_{t_k}(\lambda_k) \end{pmatrix}_{n \times n}$

2.7.3 幂零性质

若 n 阶矩阵 A 满足: $A^t = 0$, 则

1. 所有 Jordan 块的对角元都为 0

-
2. Jordan 块的最大阶数为 t ，且 $J_t(0)$ 是其中的一个 Jordan 块
 3. Jordan 块的个数等于矩阵 A 的零度(解空间维度 $= n - r(A)$)
 4. 所有 Jordan 块的阶数和为 n

若矩阵 A 的最小零化式为： $(A - \lambda)^t = 0$ ，则

1. $J_t(\lambda)$ 是其中的一个 Jordan 块

若矩阵 A 的行列式 $|\lambda I - A| = (\lambda - \lambda_1)^{t_1}, \dots, (\lambda - \lambda_j)^{t_j}$ ，则：

1. Jordan 块的对角元分别为 $\lambda_1, \dots, \lambda_j$

例题：

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

解： $A^2 = 0$ ，所以 $J_2(0)$ 是其最大阶数的 Jordan 块

因为 $r(A) = 2$ ，所以零度为 2，所以共有两 Jordan 块

所以 $J = \begin{pmatrix} J_2(0) & 0 \\ 0 & J_2(0) \end{pmatrix}$

2. 矩阵 $A = \begin{pmatrix} 0 & -4 & 0 \\ 1 & -4 & 0 \\ 1 & -2 & -2 \end{pmatrix}$

解： $\lambda(A) = \{-2, -2, -2\}$ $|\lambda I - A| = (\lambda + 2)^3$

所以 Jordan 的对角元都是 -2

又因为最小式 $m(x) = (x + 2)^2$ ，所以一个 Jordan 块为 $J_2(-2)$

所以 $J = \begin{pmatrix} J_2(-2) & 0 \\ 0 & J_1(-2) \end{pmatrix}$

3 矩阵分解

3.1 QR 分解

3.1.1 定理

设满秩矩阵 $A \in R^{n \times n}$, 则存在正交矩阵 Q 及正线上三角阵 R , 满足 $A = QR$, 且分解是唯一。

正交矩阵: $QQ^H = I$, Q 为 U 阵

正线上三角阵: 主对角线上元为正

由 $A = QR$ 可得, $Q^H A = Q^H QR = R$

3.1.2 Q 、 R 的求法

记 $A = (X_1, \dots, X_p)$, 由 3.5 节正交基的求法可以求得一组标准正交基 $(Z_1, \dots, Z_p) = Q$

$$Z_k = \frac{1}{\|Y_k\|} Y_k$$
$$Y_k = X_k - \sum_{i=1}^{k-1} \frac{(X_k, Y_i)}{(Y_i, Y_i)} Y_i$$

$$R = Q^H A$$

3.2 镜面阵

3.2.1 定义

矩阵 $A = I - \frac{2XX^H}{\|X\|^2}$, $X \in C^n$ 被称为镜面阵, 可以把矩阵 A 理解为一个平面, 向量 X 为这个平面的法向量。

3.2.2 性质

$$1. A = A^H = A^{-1} \quad A^2 = I$$

$$A^H = I - \frac{2}{\|X\|^2} (XX^H)^H = A$$

$$A^2 = I^2 - \frac{4}{\|X\|^2} XX^H + \frac{4}{\|X\|^4} (XX^H)^2$$

$$\text{因为 } (XX^H)^2 = XX^H XX^H = X(X, X)X^H = X\|X\|^2 X^H$$

$$\text{所以 } A^2 = I$$

$$\text{因为 } AA^{-1} = I = A^2, \text{ 所以 } A = A^{-1}$$

2. A 为 U 阵，各列正交+单位长+方阵

$$3. AX = -X$$

X 是镜面 A 的法向量，所以 AX 即为 X 关于镜面 A 的投影-X。

4. 若 $X \perp Y$ ，则 $AY = Y$

因为 $X \perp Y$ ，所以 Y 在镜面 A 上，所以 AY 即为镜面 A 上的向量 Y

关于镜面的投影，仍是 Y。

性质 3 和 4 都可以带入向量推导出来。

3.2.3 扩展性质

对于 $A = I - \frac{2XX^H}{\|X\|^2}$ ， $X \in C^n$ ，若 $X = \alpha - \beta$ ，则有以下性质：

$$1. A\alpha = \beta$$

3.2.4 应用

利用镜面阵对矩阵 A 进行 QR 分解。

$$\text{设矩阵 } A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix},$$

$$\text{记矩阵 } A_i = \begin{pmatrix} a_{i,i} & \cdots & a_{i,n} \\ \vdots & \ddots & \vdots \\ a_{n,i} & \cdots & a_{n,n} \end{pmatrix}, \text{ 即 } A_i \text{ 为矩阵 } A \text{ 去掉前 } i-1 \text{ 行和前 } i-1 \text{ 列}$$

之后的矩阵。

$$\text{对于矩阵 } A = (\alpha_1, \alpha_2, \dots, \alpha_n)$$

$$\text{取 } \alpha = \alpha_1, \beta = \begin{pmatrix} |\alpha| \\ 0 \\ \vdots \end{pmatrix}, P_1 = I - \frac{2XX^H}{\|X\|^2}, \text{ 其中 } X = \alpha - \beta$$

$$\text{此时: } P_1 \alpha_1 = \beta = \begin{pmatrix} |\alpha_1| \\ 0 \\ \vdots \end{pmatrix}$$

$$P_1 A = P_1 A_1 = \begin{pmatrix} |\alpha_1| & \cdots \\ 0 & A_2 \\ \vdots & \end{pmatrix}$$

对 A_2 也求出一个 P_2 , 此时

$$P'_2 P_1 A = \begin{pmatrix} 1 & 0 & \cdots \\ 0 & P_2 & \\ \vdots & & \end{pmatrix} P_1 A = \begin{pmatrix} |\alpha_1| & \cdots & \cdots \\ 0 & |\alpha'_2| & \cdots \\ \vdots & 0 & A_3 \end{pmatrix}$$

$$\text{这样 } P'_n \cdots P'_2 P_1 A = \begin{pmatrix} |\alpha_1| & \cdots & \cdots \\ 0 & \cdots & \cdots \\ \vdots & 0 & |\alpha'_n| \end{pmatrix} = R, \text{ 所以 } P'_n \cdots P'_2 P_1 = Q^H$$

3.3 秩一/满秩分解

3.3.1 秩一分解

设 $A = A_{m \times n}$, 秩 $r(A) = \text{rank}(A) = 1$, 即各列成倍数关系

$$\text{则 } A = \alpha \beta = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} (k_1 \quad \cdots \quad k_n)$$

3.3.2 秩一方阵公式

$$A = A_{n \times n}, \text{ 秩 } r(A) = \text{rank}(A) = 1$$

1. 秩一分解: $A = \alpha\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} (b_1 \cdots b_n)$
2. $\lambda(A) = \{\text{tr}(A), 0, \dots, 0\}$ 。A的唯一非零特征根就是A的迹(对角元素之和), 并且相应的特征向量为 α , 即 $A\alpha = \lambda_1\alpha$ 。
3. 对于 $b_1x_1 + \cdots + b_nx_n = 0$, 即 $\beta X = 0$, 恰有 $n-1$ 个无关解 $\{Y_1, \dots, Y_{n-1}\}$, 是矩阵 A 的 0 根特征向量: $AY_1 = 0Y_1, \dots, AY_{n-1} = 0Y_{n-1}$ 。

$$\text{证明: } AY_i = 0Y_i = 0 = (\alpha\beta)Y_i = \alpha(\beta Y_i)$$

$$\beta Y_i = 0$$

3.3.3 满秩分解

定理: 任一矩阵可分解为一个列满秩和行满秩矩阵的乘积。

记号 $A \in C_r^{m \times n}$, 表示 $r(A)=r$, 且 $A \in C^{m \times n}$

满秩分解定理可表示为: 设 $A \in C_r^{m \times n} (r > 0)$, 则存在 $F \in C_r^{m \times r}, G \in C_r^{r \times n}$, 使 $A = FG$ 。

3.3.4 满秩分解的求法

对于矩阵 $A_{m \times n} = (\alpha_1, \alpha_2, \dots, \alpha_n) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$, 对其进行初等行

变换之后可以得到矩阵 A 的标准形 \hat{A}_r , \hat{A}_r 只有前 r 行包含非零元, 其他行全为 0.

$$\hat{A}_r = \begin{pmatrix} & 1 & & 0 & \\ \cdots & 0 & \cdots & 1 & \cdots \\ & \vdots & & \vdots & \\ & 0 & & 0 & \\ & & \mathbf{0} & & \end{pmatrix}$$

并且 \hat{A}_r 的前 r 行的若干列可以组成一个单位阵 I , 即这些列只有一个非零元 1, 记这些列的列号为 $\{\beta_1, \dots, \beta_r\}$.

此时, \hat{A}_r 的前 r 行构成的新矩阵即为 $G_{r \times n}$, 原矩阵 A 中 $\{\beta_1, \dots, \beta_r\}$ 列构成的新矩阵就是 $F_{m \times r}$ 。

例题: $A = \begin{pmatrix} 0 & 2i & i & 0 & 4+2i & 1 \\ 0 & 0 & 0 & -3 & -6 & -3-3i \\ 0 & 2 & 1 & 1 & 4-4i & 1 \end{pmatrix}$

对其进行初等变换之后 $\hat{A}_r = \begin{pmatrix} 0 & \mathbf{1} & 0.5 & \mathbf{0} & 1-2i & -0.5i \\ 0 & \mathbf{0} & 0 & \mathbf{1} & 2 & 1+i \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

此时 $G_{r \times n} = \begin{pmatrix} 0 & 1 & 0.5 & 0 & 1-2i & -0.5i \\ 0 & 0 & 0 & 1 & 2 & 1+i \end{pmatrix}$

$F_{m \times r} = \begin{pmatrix} 2i & 0 \\ 0 & -3 \\ 2 & 1 \end{pmatrix}$

3.4 正规矩阵及 Schur 分解

3.4.1 Schur 引理

已知任意矩阵 $A \in C^{m \times n}$ 都相似与一上三角矩阵, 即存在矩阵 P 使得 $P^{-1}AP$ 为上三角矩阵。

由此可得定理: 对任意矩阵 $A \in C^{m \times n}$, 存在 U 阵 U , 使得:

$$U^H A U = \begin{pmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

即任意复方阵 A 酉相似与一上三角阵，且主对角元为 A 的特征值。

证明：由 QR 分解可知， $P = QR = UR$ (Q, U 都是酉阵)

$$\text{记 } P^{-1}AP = K, A = PKP^{-1}$$

$$U^H AU = U^H PKP^{-1}U = U^H URKR^{-1}U^{-1}U = RKR^{-1}$$

因为 K, R 都是上三角阵，所以 RKR^{-1} 为上三角阵

3.4.2 正规矩阵

设 $A \in C^{m \times n}$ ，若 A 满足 $A^H A = AA^H$ ，则称 A (方阵) 为正规矩阵 (规范阵)。

若 A 正规，则 A^H 也正规。

正规矩阵包括：

- 1) 厄米特阵： $A^H = A$ ，实对称矩阵： $A^T = A$
- 2) 斜厄米特阵： $A^H = -A$ ，实反对称矩阵： $A^T = -A$
- 3) 酉阵： $A^{-1} = A^H$ 实正交阵： $A^T = A^{-1}$
- 4) 复对角矩阵： $\begin{pmatrix} a_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_n \end{pmatrix}$

引理(平移法)： A 正规 $\leftrightarrow kA \pm cI$ 正规

证明：令 $B = kA - cI$

$$\begin{aligned} B^H B &= (kA - cI)^H (kA - cI) = (kA^H - \bar{c}I)(kA - cI) = k^2 A^H A - \\ &kcA^H - \bar{c}kA + c\bar{c}I = k^2 AA^H - kcA^H - \bar{c}kA + c\bar{c}I = (kA - cI)kA^H - \\ &(kA - cI)\bar{c} = BB^H \end{aligned}$$

Eg1 : $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ 正规?

解: $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + aI$
 $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ 为实反对称阵, 正规, 所以 A 正规

引理(酉相似): A 正规 $\leftrightarrow Q^H A Q$ 正规, Q 为任意酉阵

A 非正规 $\leftrightarrow Q^H A Q$ 非正规, Q 为任意酉阵

引理(分块上三角): 若 $A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$ 正规 $\leftrightarrow C$ 为 0 阵, 且 B 、 D 都正规(条件: B 、 D 为方阵)

证明: 迹公式: $\text{tr}(A^H A) = \text{tr}(A A^H) = \sum |a_{ij}|^2$

$\text{tr}(A^H A) = 0 \rightarrow A$ 为 0 阵

$$A A^H = \begin{pmatrix} B B^H + C C^H & C D^H \\ D C^H & D D^H \end{pmatrix} = A^H A = \begin{pmatrix} B^H B & B^H C \\ C^H B & C^H C + D^H D \end{pmatrix}$$

$$B B^H + C C^H = B^H B$$

$$D D^H = C^H C + D^H D$$

因为: $\text{tr}(B B^H) + \text{tr}(C C^H) = \text{tr}(B^H B)$ $\text{tr}(B B^H) = \text{tr}(B^H B)$

所以: $\text{tr}(C C^H) = 0$, C 为 0 阵, $C C^H$ 和 $C^H C$ 也为 0 阵

所以 $B B^H = B^H B$ $D D^H = D^H D$

因此, 可以得到以下推论:

- 1) 若上三角矩阵 A 正规, 则 A 为对角阵。
- 2) 若 A 为严格上三角(不是对角阵), 则 A 非正规。

证明:

记 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ & A_1 & & \\ & & & \end{pmatrix}$, 由引理可知, $a_{12} = \cdots = a_{1n} = 0$, A_1 正规。

对 A_i 进行递归, 可得 $a_{i,i+1} = \cdots = a_{i,n} = 0$, 所以 A 为对角阵。

3.4.3 定理与推论

定理: 设 $A \in C^{m \times n}$, 则 A 是正规矩阵当且仅当 A 酉相似于一个对角阵,

即: $A^H A = A A^H \leftrightarrow U^H A U = \text{diag}\{\lambda_1, \cdots, \lambda_n\}$

利用上节引理证明:

A 正规 $\rightarrow U^H A U = D$ 正规

D 正规, 且 D 为上三角 $\rightarrow D$ 为对角阵

直接证明:

充分性:

由 Schur 引理可知, $U^H A U = K$, $A = U K U^H$, K 为上三角阵

$$A^H A = U K^H U^H U K U^H = U K^H K U^H$$

$$A A^H = U K U^H U K^H U^H = U K K^H U^H$$

因为 $A^H A = A A^H$, 所以 $K^H K = K K^H$

记 $K = (r_{ij})_{n \times n}$, $r_{ij} = 0 (i > j)$, 由 $K^H K = K K^H$ 可得 $r_{ij} = 0 (i < j)$, 所

以 K 为对角阵

必要性:

$$U^H A U = \text{diag}\{\lambda_1, \cdots, \lambda_n\}, \quad U^H A^H U = \text{diag}\{\overline{\lambda_1}, \cdots, \overline{\lambda_n}\}$$

$$(U^H A U)(U^H A^H U) = \text{diag}\{|\lambda_1|^2, \cdots, |\lambda_n|^2\} = U^H A A^H U$$

$$(U^H A^H U)(U^H A U) = \text{diag}\{|\lambda_1|^2, \dots, |\lambda_n|^2\} = U^H A^H A U$$

$$\text{所以 } A^H A = A A^H$$

推论 1: A 为正规阵, 当且仅当 A 有 n 个特征向量构成 $C^{m \times n}$ 的一组标准正交基, 且 A 不同特征值得特征向量正交。

3.5 厄米特(Hermite)分解

3.5.1 定理

1. 若 $A \in C^{n \times n}$, A 为厄米特阵 ($A = A^H$), 则存在酉阵 Q, 使得:

$$Q^{-1} A Q = Q^H A Q = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \text{ (对角阵)}$$

且 $\lambda_1 \cdots \lambda_n$ 都为实数。

证明:

$$\text{由许尔(Schur)分解可知: } Q^H A Q = D = \begin{pmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

$$(Q^H A Q)^H = Q^H A^H Q = Q^H A Q \rightarrow D^H = D$$

所以 $* = 0$, $\lambda_i = \bar{\lambda}_i$, λ_i 为实数。

2. 若 $P^{-1} A P = D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$, 则 $P = (X_1, \dots, X_n)$ 中各列都是 A 的

特征向量, 且线性无关。

3. 若 $A \in C^{n \times n}$, A 为厄米特阵 ($A = A^H$), 则 A 恰有 n 个相互正交的特征向量。(有定理 2 可得)

3.5.2 应用

1. 若 $A = A^H$, $X \in C^n$, 则 $f(X) = X^H A X$ 的值都为实数。

$$f^H = (X^H A X)^H = X^H A^H X = X^H A X = f$$

因为 f 为一个数, 所以 f 为实数。

2. 若 $A = A^H$, 任取 $\lambda_i \in \lambda(A)$, $\lambda(A)$ 为 A 的特征值集合, 则

$$\lambda_i = \frac{X_i^H A X_i}{|X_i|^2} \quad (X_i \text{ 为 } \lambda_i \text{ 对应的非零特征向量})$$

证明:

$$A X_i = \lambda_i X_i \rightarrow X_i^H A X_i = \lambda_i X_i^H X_i = \lambda_i |X_i|^2$$

3. 由上述两点可得: **Hermite** 阵的特征值都为实数。

3.5.3 正定矩阵

定义: 对任意 $X \in C^n$, $f(X) = X^H A X \geq 0$, 则称 A 为半正定阵, 记 $A \geq 0$ 。

对任意非零 $X \in C^n$, $f(X) = X^H A X > 0$, 则称 A 为正定阵, 记 $A > 0$ 。

结合 $\lambda_i = \frac{X_i^H A X_i}{|X_i|^2}$ 可得:

$$A \geq 0, A = A^H \leftrightarrow \lambda_i \geq 0$$

$$A > 0, A = A^H \leftrightarrow \lambda_i > 0$$

3.5.4 Hermite 阵的性质

1. $A = A^H \rightarrow \lambda(A)$ 全为实数 (由 3.5.2(1)可证)
2. $A = A^H \rightarrow f(X) = X^H A X$ 的值全为实数
3. $A = A^H, A \geq 0 \leftrightarrow \lambda_i \geq 0$
4. $A = A^H, A > 0 \leftrightarrow \lambda_i > 0$

5. 若 $A \geq 0$, $A = A^H$, 则存在 $B \geq 0$ 使得, $B^2 = A$, B 叫做 A 的平方根,

记 $B = \sqrt{A}$, 可以写为 $A = (\sqrt{A})^2$

$$\text{证明: } Q^H A Q = D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

$$\sqrt{D} = \begin{pmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}$$

$$\text{取 } B = Q \sqrt{D} Q^H$$

$$B^2 = (Q \sqrt{D} Q^H)(Q \sqrt{D} Q^H) = Q D Q^H = A$$

6. 对于任意矩阵 $A \in C^{m \times n}$, $A^H A$ 和 $A A^H$ 是半正定的厄米特阵

3.5.5 常见的 Hermite 阵

1. 实对称矩阵: $A^H = A^T = A$
2. 任一方阵 A , $A + A^H$ 为 Hermite 阵

3.5.6 斜 Hermite 分解

若 $A \in C^{n \times n}$, A 为斜厄米特阵 ($A^H = -A$), 则存在酉阵 Q , 使得:

$$Q^{-1} A Q = Q^H A Q = \begin{pmatrix} b_1 i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & b_n i \end{pmatrix} \text{ (对角阵)}$$

$\lambda(A) = \{b_1 i, \dots, b_n i\}$ 为纯虚数。

证明:

$$A^H = -A \rightarrow \frac{A}{i} \text{ 为 Hermite 阵}$$

$$\text{所以: } Q^H \left(\frac{A}{i} \right) Q = D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

$$Q^H A Q = iD = \begin{pmatrix} \lambda_1 i & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n i \end{pmatrix}$$

3.6 特征值/酉阵 Q 的求解技巧

3.6.1 平移法则(求特征根)

思路：

将一个复杂矩阵转换成秩一矩阵 $A = \alpha\beta$ ，再利用秩一矩阵 $\lambda(A) = \{tr(A), 0, \dots, 0\}$ 特征向量 $= \alpha$ ，求复杂矩阵的特征值。

平移法则：

矩阵 $kA \pm cI$ 与矩阵 A 具有相同的特征向量 $\{X_i\}$ ，并且：

$$\lambda(kA \pm cI) = \{k\lambda_1 \pm c, \dots, k\lambda_n \pm c\}, \text{ 其中 } \lambda(A) = \{\lambda_1, \dots, \lambda_n\}$$

$$\text{即：} AX_i = \lambda_i X_i \leftrightarrow (kA \pm cI)X_i = (k\lambda_i \pm c)X_i$$

例题：

$$A = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix}, \text{ 求 } \lambda(A) \text{ 并给出一个特征向量。}$$

解：

$$A - I = \begin{pmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix} \text{ (秩一),}$$

$$\text{所以 } \lambda(A - I) = \{-4, 0, 0, 0\}, \text{ 特征向量 } \alpha = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

所以 $\lambda(A) = \lambda(A - I) + 1 = \{-3, 1, 1, 1\}$

$$A\alpha = -3\alpha$$

3.6.2 特殊矩阵分解酉阵 Q 的求解方法

适用范围：正规矩阵的 Schur 分解、Hermite 分解

$$Q^H A Q = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \text{ (对角阵)}$$

求法：

1. 求 A 的特征根
2. 求 A 的特征向量
3. 求 A 特征向量的标准正交基
4. 以这些标准正交基为列向量的矩阵即为 Q

例题：

$$\text{设 } A = \begin{pmatrix} -1 & -3 & 3 & -3 \\ -3 & -1 & -3 & 3 \\ 3 & -3 & -1 & -3 \\ -3 & 3 & -3 & -1 \end{pmatrix}, \text{ 求正交阵 } Q, \text{ 使得 } Q^H A Q \text{ 为对角阵。}$$

解：

$$(A + 4I) = \begin{pmatrix} 3 & -3 & 3 & -3 \\ -3 & 3 & -3 & 3 \\ 3 & -3 & 3 & -3 \\ -3 & 3 & -3 & 3 \end{pmatrix} \text{ (秩一)}$$

$$\lambda(A + 4I) = \{12, 0, 0, 0\} \rightarrow \lambda(A) = \{8, -4, -4, -4\}$$

$$\lambda = 8 \text{ 时, } \alpha_1 = (-1 \ 1 \ -1 \ 1)^T$$

$$\lambda = -4 \text{ 时, } \alpha_2 = (1 \ 1 \ 0 \ 0)^T, \alpha_3 = (1 \ 0 \ -1 \ 0)^T,$$

$$\alpha_4 = (1 \ 0 \ 0 \ 1)^T,$$

利用正交化法则：

$$Y_{m+1} = X_{m+1} + \sum_{i=1}^m k_i Y_i, \text{ 其中: } k_i = -\frac{(X_{m+1}, Y_i)}{(Y_i, Y_i)}$$

$$Y_1 = (-1, 1, -1, 1)$$

$$Y_2 = X_2 - \frac{(X_2, Y_1)}{(Y_1, Y_1)} Y_1 = (1, 1, 0, 0)$$

$$Y_3 = X_3 - \frac{(X_3, Y_1)}{(Y_1, Y_1)} Y_1 - \frac{(X_3, Y_2)}{(Y_2, Y_2)} Y_2 = (\frac{1}{2}, -\frac{1}{2}, -1, 0)$$

$$Y_4 = X_4 - \frac{(X_4, Y_1)}{(Y_1, Y_1)} Y_1 - \frac{(X_4, Y_2)}{(Y_2, Y_2)} Y_2 - \frac{(X_4, Y_3)}{(Y_3, Y_3)} Y_3 = (\frac{1}{3}, -\frac{1}{3}, \frac{1}{3}, 1)$$

$$Z_1 = (-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}), \quad Z_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0)$$

$$Z_3 = (\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, 0), \quad Z_4 = (\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{3}{2\sqrt{3}})$$

$$Q = \{Z_1, Z_2, Z_3, Z_4\}$$

3.6.3 换位公式(求特征根)

换位公式：

$A \in C^{n \times p}, B \in C^{p \times n}, AB \in C^{n \times n} (n \geq p)$, 则：

$$1) \quad |\lambda I_n - AB| = \lambda^{n-p} |\lambda I_p - BA|$$

2) AB 与 BA 的特征根只差 $n-p$ 个0根

$$\text{即: } \lambda(BA) = \{\lambda_1, \dots, \lambda_p\}$$

$$\lambda(AB) = \{\lambda_1, \dots, \lambda_p, 0, \dots, 0\}$$

所以 AA^H 与 $A^H A$ 特征根只差若干个0根

$$3) \quad \text{tr}(AB) = \text{tr}(BA) = \lambda_1 + \dots + \lambda_n$$

证明:

$$\begin{aligned} \text{构造矩阵 } M &= \begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}, N = \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}, P = \begin{pmatrix} I & A \\ 0 & I \end{pmatrix} \\ MP &= \begin{pmatrix} AB & ABA \\ B & BA \end{pmatrix} = PN \rightarrow P^{-1}MP = N \rightarrow M \sim N (\text{相似}) \\ \rightarrow |\lambda I - M| &= |\lambda I - N| \\ \rightarrow \begin{pmatrix} \lambda I - AB & 0 \\ -B & \lambda I - BA \end{pmatrix} &= \begin{pmatrix} \lambda I_n & 0 \\ -B & \lambda I - BA \end{pmatrix} \\ \rightarrow |\lambda I - AB| \lambda^p &= \lambda^n |(\lambda I - BA)| \\ \rightarrow |\lambda I_n - AB| &= \lambda^{n-p} |\lambda I_p - BA| \end{aligned}$$

例题:

Eg1. $A = \begin{pmatrix} 3 & 1 & 3 \\ 1 & 2 & 2 \\ -1 & 2 & 2 \end{pmatrix}$, 求 $\lambda(A)$ 和 $|\lambda I - A|$

解:

$$\begin{aligned} A - I &= \begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix} = BC = \begin{pmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \\ CB &= \begin{pmatrix} 1 & 3 \\ 0 & 3 \end{pmatrix} \\ \lambda(CB) &= \{1, 3\} \rightarrow \lambda(BC) = \{1, 3, 0\} \rightarrow \lambda(A) = \{2, 4, 1\} \\ |\lambda I - A| &= (\lambda - 2)(\lambda - 4)(\lambda - 1) \end{aligned}$$

Eg2. $A = \begin{pmatrix} -1 & i & 0 \\ -i & 0 & -i \\ 0 & i & -1 \end{pmatrix}$, 求 $\lambda(A)$ 和 $|\lambda I - A|$

解:

$$A + I = \begin{pmatrix} 0 & i & 0 \\ -i & 1 & -i \\ 0 & i & 0 \end{pmatrix} = BC = \begin{pmatrix} 0 & i \\ -i & 1 \\ 0 & i \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$CB + 1 = \begin{pmatrix} 1 & 2i \\ -i & 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -i \end{pmatrix} (1 \quad 2i) \text{(秩一)}$$

$$\lambda(CB + 1) = \{3, 0\} \rightarrow \lambda(CB) = \{2, -1\} \rightarrow \lambda(BC) = \{2, -1, 0\}$$

$$\rightarrow \lambda(A) = \{1, -2, -1\}$$

Eg3 用平移法求 $A = I - \frac{2XX^H}{\|X\|^2}$ 的 $\lambda(A)$, 其中 $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

解:

$$XX^H = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (\overline{x_1} \quad \cdots \quad \overline{x_n}) \text{(秩一)}$$

$$A - I = -\frac{2}{\|X\|^2} (XX^H) \text{(秩一)}$$

$$\lambda(XX^H) = \{\sum x_i^2, 0, \dots, 0\} = \{\|X\|^2, 0, \dots, 0\}$$

$$\lambda\left(-\frac{2}{\|X\|^2} (XX^H)\right) = -\frac{2}{\|X\|^2} \lambda(XX^H) = \{-2, 0, \dots, 0\}$$

$$\lambda(A - I) = \{-2, 0, \dots, 0\}$$

$$\lambda(A) = \{-1, 1, \dots, 1\}$$

3.7 奇异值分解

3.7.1 一些定理

对于矩阵 $A \in \mathbb{C}^{m \times n}$

1) $A^H A$ 和 $A A^H$ 是半正定的厄米特阵, 且具有相同的非零特征值

$X^H A^H A X = (AX, AX) \geq 0$, 所以 $A^H A$ 为半正定阵。

2) $\text{rank}(A^H A) = \text{rank}(A A^H) = \text{rank}(A)$

记 $A^H A$ 的解空间为 X ,

$$A^HAX = 0 \rightarrow X^HA^HAX = X^H0 = 0 \rightarrow (AX)^HAX = (AX, AX) = 0 \\ \rightarrow AX = 0$$

所以 A^HA 和 A 的解空间相等，所以 $\text{rank}(A^HA) = \text{rank}(AA^H) = \text{rank}(A)$

高阵有左侧逆

设 $B \in C^{m \times p}$ 为高阵， $r(B) = p$ ，则存在 $B_L B = I$ ， $B_L = (B^H B)^{-1} B^H$

证明：

$r(B) = p \rightarrow r(B^H B) = p, B^H B \in C^{p \times p} \rightarrow B^H B$ 满秩 $\rightarrow (B^H B)^{-1}$ 存在，
所以存在 $B_L = (B^H B)^{-1} B^H$ ，使得 $B_L B = (B^H B)^{-1} (B^H B) = I$

低阵有右侧逆

设 $C \in C^{p \times m}$ 为低阵， $r(C) = p$ ，则存在 $CC_R = I$ ， $C_R = C^H(CC^H)^{-1}$

用法：

1. 若 $BCX = 0$ ， B 为高阵，则 $CX = 0$

证： $BCX = 0 \rightarrow B_L(BCX) = 0 \rightarrow CX = 0$

2. 若 $BX = BY$ ， B 为高阵，则 $X = Y$

证： $BX = BY \rightarrow B_L(BX) = B_L(BY) \rightarrow IX = IY$

3.7.2 正奇异值

设 $A = A_{m \times n}$ ， $r(A) = r > 0$ ，则 A^HA 和 AA^H 恰有 r 个正根 $\lambda_1, \dots, \lambda_r > 0$ 。

$\lambda_1, \dots, \lambda_r$ 可重复

$$\lambda(AA^H) = \{\lambda_1, \dots, \lambda_r, 0, \dots, 0\}$$

$$\lambda(A^HA) = \{\lambda_1, \dots, \lambda_r, 0, \dots, 0\}$$

记 A 的正奇异值为: $S^+(A) = \{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}\}$

A 的全体奇异值为: $S(A) = \{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}, 0, \dots, 0\}$

3.7.3 简奇异值分解与奇异值分解

对于矩阵 $A \in C^{m \times n}$, $r(A)=r$, 则存在半酉阵 $P \in C^{m \times r}$ 和半酉阵 $Q \in C^{n \times r}$ ($PP^H = I, QQ^H = I$), 使得:

$$A = PS_rQ^H, \text{ 其中 } S_r = \text{diag}\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}\}$$

对于矩阵 $A \in C^{m \times n}$, $r(A)=r$, 则存在酉阵 $W \in C^{m \times m}$ 和酉阵 $V \in C^{n \times n}$, 使得 $A = W \begin{pmatrix} S_r & 0 \\ 0 & 0 \end{pmatrix}_{m \times n} V^H$, 其中 $S_r = \text{diag}\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}\}$ 。

3.7.4 简奇异值分解方法

解法一:

1. 求 A^HA 的特征根 $\lambda_1, \dots, \lambda_r > 0$ 与特征向量 X_1, \dots, X_r (相互正交)

$$S_r = \text{diag}\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}\}$$

2. 令 $P = \left(\frac{AX_1}{|AX_1|}, \dots, \frac{AX_r}{|AX_r|} \right)$, $Q = \left(\frac{X_1}{|X_1|}, \dots, \frac{X_r}{|X_r|} \right)$

3. 可得 $A = PS_rQ^H$

4. 在简奇异值分解 ($A = PS_rQ^H$) 的基础之上将 P 和 Q 扩展为酉阵, 其中 $W =$

$(P \ P_2)$, $V = (Q \ Q_2)$, 即可得奇异值分解

$$A = W \begin{pmatrix} S_r & 0 \\ 0 & 0 \end{pmatrix} V^H。$$

解法二：

$$A = PS_r Q^H, A^H = (PS_r Q^H)^H = Q S_r^H P^H = Q S_r P^H (S_r \text{ 为对角阵})$$

所以如果已知某个矩阵A的奇值分解，就可以求得矩阵 A^H 的奇值分解。

Eg1 : $A = \begin{pmatrix} 1 & 2i \\ i & 1 \\ i & 1 \end{pmatrix}$, 求正奇异值和简(正)SVD

解：

1. $A^H A = \begin{pmatrix} 3 & 0 \\ 0 & 6 \end{pmatrix}$ (对角) $\rightarrow \lambda_1 = 3, \lambda_2 = 6, X_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, X_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

正奇值为 $\{\sqrt{3}, \sqrt{6}\}$, $S_r = \text{diag}\{\sqrt{3}, \sqrt{6}\}$

2. $AX_1 = \begin{pmatrix} 1 \\ i \\ i \end{pmatrix}, AX_2 = \begin{pmatrix} 2i \\ 1 \\ 1 \end{pmatrix} \quad |AX_1| = \sqrt{3}, |AX_2| = \sqrt{6},$

$$P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2i}{\sqrt{6}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}, Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

3. $A = PS_r Q^H$

4. $W = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2i}{\sqrt{6}} & 0 \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \end{pmatrix}, Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A = P \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{6} \\ 0 & 0 \end{pmatrix} Q^H$

Eg2 : $A = \begin{pmatrix} 1 & -1 \\ 0 & 0 \\ 1 & -1 \end{pmatrix}$, 求正奇异值和简(正)SVD

解：

1. $A^H A = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$ (秩 1) \rightarrow 正根 $\lambda_1 = 4, X_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

正奇值为{2}, $S_r = \text{diag}\{2\}$

$$2. \quad AX_1 = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} \quad |AX_1| = 2\sqrt{2}$$

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, Q = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -1 \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$3. \quad A = PS_r Q^H$$

$$4. \quad W = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -1 & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad A = P \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} Q^H$$

Eg3: $B = \begin{pmatrix} 1 & -i & -i \\ -2i & 1 & 1 \end{pmatrix}$, 求正奇异值和简(正)SVD

解:

$$A = B^H = \begin{pmatrix} 1 & 2i \\ i & 1 \\ i & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2i}{\sqrt{6}} \\ i & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{6} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^H$$

$$B = \begin{pmatrix} 1 & -i & -i \\ -2i & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{6} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2i}{\sqrt{6}} \\ i & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}^H$$

3.7.5 推论

Eg1: 对于方阵A, 奇异值的乘积等于特征根的乘积的绝对值

$$\lambda(A) = \{\lambda_1, \dots, \lambda_n\} (\text{包含 } 0 \text{ 根})$$

$$\det(A) = \lambda_1 * \dots * \lambda_n$$

$$\lambda(A^H A) = \{t_1, \dots, t_n\}$$

$$S(A) = \{s_1, \dots, s_n\} = \{\sqrt{t_1}, \dots, \sqrt{t_n}\}$$

$$\det(A^H) = \det(\bar{A})^T = \det(\bar{A}) = \overline{\det(A)}$$

$$\rightarrow \det(A^H A) = \det(A^H) \det(A) = \overline{\det(A)} \det(A) = |\det(A)|^2$$

$$\rightarrow t_1 * \dots * t_n = (s_1 * \dots * s_n)^2 = (\lambda_1 * \dots * \lambda_n)^2$$

$$\rightarrow s_1 * \dots * s_n = |\lambda_1 * \dots * \lambda_n|$$

Eg2: 已知矩阵 A 的正奇异值分解为 $A = P\Delta Q^H$, 可以得到矩阵 $B = \begin{pmatrix} A \\ A \end{pmatrix}$ 和矩阵 $C = (A, A)$ 的正奇异值分解。

解:

$$B^H B = (A^H, A^H) \begin{pmatrix} A \\ A \end{pmatrix} = 2A^H A$$

$$\lambda(B^H B) = 2 \lambda(A^H A)$$

$$S^+(B) = \sqrt{2} S^+(A)$$

$$B = \begin{pmatrix} A \\ A \end{pmatrix} = \begin{pmatrix} P\Delta Q^H \\ P\Delta Q^H \end{pmatrix} = \begin{pmatrix} P \\ P \end{pmatrix} \Delta Q^H = \frac{1}{\sqrt{2}} \begin{pmatrix} P \\ P \end{pmatrix} (\sqrt{2}\Delta) Q^H$$

$$\text{令 } \bar{P} = \frac{1}{\sqrt{2}} \begin{pmatrix} P \\ P \end{pmatrix}, \bar{\Delta} = \sqrt{2}\Delta, \bar{Q} = Q$$

其中: $\bar{P}^H \bar{P} = \frac{1}{2} (P^H, P^H) \begin{pmatrix} P \\ P \end{pmatrix} = P^H P = I$, \bar{P} 是一个酉阵

$$B = \bar{P} \bar{\Delta} \bar{Q}^H$$

同样:

$$CC^H = (A, A) \begin{pmatrix} A^H \\ A^H \end{pmatrix} = 2AA^H$$

$$\lambda(CC^H) = 2\lambda(AA^H) = 2\lambda(A^HA)$$

$$S^+(C) = \sqrt{2}S^+(A)$$

$$C = (A, A) = (P\Delta Q^H, P\Delta Q^H) = P\Delta(Q^H, Q^H) = P(\sqrt{2}\Delta) \frac{1}{\sqrt{2}} \begin{pmatrix} Q \\ Q \end{pmatrix}^H$$

$$\text{令 } \bar{P} = P, \bar{\Delta} = \sqrt{2}\Delta, \bar{Q} = \frac{1}{\sqrt{2}} \begin{pmatrix} Q \\ Q \end{pmatrix}$$

$$C = \bar{P}\bar{\Delta}\bar{Q}^H$$

Eg3: 矩阵的极分解

对于任一方阵 $A = A_{n \times n}$, 有

$$A = WDV^H = W \begin{pmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_n \end{pmatrix} V^H, \quad A^H = VDW^H$$

$$AA^H = WD^2W^H = (WDW^H)^2$$

因为 WDW^H 为半正定矩阵, 所以 $\sqrt{AA^H} = WDW^H$

$$A = WDV^H = A = WDW^H W V^H = \sqrt{AA^H} U$$

其中 $U = WV^H$ 为酉阵

$A = \sqrt{AA^H} U = U \sqrt{A^H A}$ 叫做 A 的极分解

特别: $A = (z)$ (复数)

$$z = \sqrt{zz^H} e^{i\theta}$$

Eg4: $A = A_{m \times n} = (a_{ij})_{m \times n}, S^+(A) = \{s_1, \dots, s_r\}$

证明: $\text{tr}(AA^H) = \text{tr}(A^H A) = \sum |a_{i \cdot j}|^2 = s_1^2 + \dots + s_r^2$

解:

AA^H 为 Hermite 阵, 所以 $\text{tr}(AA^H) = \text{tr}(A^H A)$

记: $\lambda(AA^H) = \{\lambda_1, \dots, \lambda_r\}$ $\lambda_i = s_i^2$

$\text{tr}(AA^H) = \text{tr}(A^H A) = \lambda_1 + \dots + \lambda_r = s_1^2 + \dots + s_r^2$

矩阵 AA^H 的第 k 行对角元:

$$x_k = (a_{1k} \quad \dots \quad a_{mk}) \begin{pmatrix} a_{1k}^H \\ \vdots \\ a_{mk}^H \end{pmatrix} = \sum_{i=1}^m |a_{i*k}|^2$$

所以 $\text{tr}(AA^H) = x_1 + \dots + x_m = \sum |a_{i*j}|^2$

3.8 谱分解

3.8.1 单纯阵

定义:

$A = A_{n*n}$ 叫做单纯阵 $\leftrightarrow A \sim D = \text{diag}(\lambda_1, \dots, \lambda_n) \leftrightarrow P^{-1}AP = D$

注: $P^{-1}AP = D = \text{diag}(\lambda_1, \dots, \lambda_n) \leftrightarrow P = (X_1, \dots, X_n)$ 各列都是矩阵 A 的特征向量且线性无关, 即 $AX_i = \lambda_i X_i, 1 \leq i \leq n$

定理:

1. 方阵 A 为单纯阵 $\leftrightarrow A$ 有 n 个线性无关的特征向量。
2. 方阵 A 为单纯阵 $\leftrightarrow A$ 的每个 k 重根恰有 k 个线性无关的特征向量。
3. 若方阵 A 恰有 n 个互异的特征根, 则 A 为单纯阵。
4. 若方阵 A 的每个 k 重根恰有 k 个线性无关的特征向量, 则 A 为单纯阵。

判定:

1. 若方阵A恰有 n 个互异的特征根, 则A为单纯阵。
2. 对于方阵A的任一 $k(k>1)$ 重根 $\lambda_i \in \lambda(A)$,
若 $\text{rank}(A - \lambda_i I) = n - k$, 则A为单纯阵, 否则为非单。
3. 设方阵A有 k 个互异的特征根 $\lambda_1, \dots, \lambda_k (k \leq n)$,
若 $(A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_k I) = 0$, 则A为单纯阵, 否则为非单。

3.8.2 零化式

定义:

若存在方阵A和多项式 $f(x) = c_p x^p + \cdots + c_1 x^1 + c_0$,

使得: $f(A) = c_p A^p + \cdots + c_1 A^1 + c_0 I = 0$.

称 $f(x)$ 是A的一个”0 化式”, A叫做 $f(x)$ 的一个矩阵根。

注: 若 $f(x)$ 是A的一个零化式, 则对任意 $g(x) \rightarrow f(A)g(A) = 0 \rightarrow f(x)g(x)$ 也是A的一个零化式。

推理:

对于任一固定的方阵A, 可求出次数最低的零化式, 记为 $m_A(x)$ 。

Caylay 定理:

方阵 A 的特征多项式:

$$T(x) = |xI - A| = x^n + a_{n-1}x^{n-1} + \cdots + a_0$$

$$T(A) = A^n + a_{n-1}A^{n-1} + \dots + a_0I = 0$$

Eg1: $A = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 3 & 4 \\ 1 & 1 & 3 \end{pmatrix}$, 求 A 得极小式 $m(x) = ?$

解:

$$\lambda(A - I) = \{5, 0, 0\} \rightarrow \lambda(A) = \{6, 1, 1\}$$

$$\text{特征多项式: } T(x) = (x - 6)(x - 1)^2$$

$$\text{零化式(Caylay): } T(A) = (A - 6)(A - 1)^2 = 0$$

$$\text{因为: } (A - 6)(A - 1) = 0$$

$$\text{所以极小式: } m(x) = (A - 6)(A - 1)$$

应用: 若 $f(x)$ 无重根且 $f(A) = 0$, 则 A 为单纯阵。

Eg2: 已知 $A^2 - 3A + 2I = 0$, 则 A 为单。

解:

$$f(x) = x^2 - 3x + 2 = (x - 1)(x - 2) \text{ 无重根且 } f(A) = 0$$

所以 A 为单纯阵

Eg3: 单纯阵判定: $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

解:

$$\lambda(A - I) = \{1, 0, 0\} \rightarrow \lambda(A) = \{2, 1, 1\}$$

$$(A - 2)(A - 1) = 0 \text{ (第三条判定规则)}$$

所以 A 是单纯阵

3.8.3 谱分解

单纯矩阵: $P^{-1}AP = D = \text{diag}(\lambda_1, \dots, \lambda_n)$

$$D = \lambda_1 \text{diag}(1, 0, 0, \dots) + \lambda_2 \text{diag}(0, 1, 0, \dots) + \dots + \lambda_n \text{diag}(0, \dots, 0, 1)$$

$$D = \lambda_1 E_1 + \lambda_2 E_2 + \dots + \lambda_n E_n$$

有以下特点:

- 1) $E_1 + E_2 + \dots + E_n = I$
- 2) $E_i E_j = 0 (i \neq j)$
- 3) $E_i E_i = E_i$ (幂等阵)
- 4) $E_i^H = E_i$ (Hermit 阵)

$$P^{-1}AP = D \rightarrow A = PDP^{-1}$$

$$\begin{aligned} A &= P(\lambda_1 E_1 + \dots + \lambda_n E_n)P^{-1} = \lambda_1 P E_1 P^{-1} + \dots + \lambda_n P E_n P^{-1} \\ &= \lambda_1 F_1 + \dots + \lambda_n F_n (F_i = P E_i P^{-1}) \end{aligned}$$

$A = \lambda_1 F_1 + \dots + \lambda_n F_n$ 叫做A的谱分解

F_i 有以下特点:

- 1) $F_1 + F_2 + \dots + F_n = I$
- 2) $F_i F_j = 0 (i \neq j)$
- 3) $F_i F_i = F_i$ (幂等阵)
- 4) 若P为 Hermit 阵, 则 $F_i^H = F_i$, 否则 F_i^H 不一定等于 F_i 。

证明:

$$1) F_1 + F_2 + \dots + F_n = P(E_1 + \dots + E_n)P^{-1} = PIP^{-1} = I$$

$$2) F_i F_j = P E_i P^{-1} P E_j P^{-1} = P E_i E_j P^{-1} = P 0 P^{-1} = 0$$

$$3) F_i F_i = P E_i E_i P^{-1} = P E_i P^{-1} = F_i$$

单阵谱分解公式:

若 $A = A_{n \times n}$ 为单阵, 全体不同根为 t_1, t_2, \dots, t_k ,

则有: $A = t_1 G_1 + t_2 G_2 + \dots + t_k G_k (k \leq n)$

其中: $G_1 + \dots + G_k = I$, $G_i G_j = 0 (i \neq j)$, $G_i G_i = G_i$ (幂等阵)

G_1, \dots, G_k 叫谱阵。

证明:

$$A = \lambda_1 F_1 + \dots + \lambda_n F_n,$$

对于重根 $\lambda_i = \dots = \lambda_j$, $\lambda_i F_i + \dots + \lambda_j F_j = t(F_i + \dots + F_j) = tG$

所以 $A = t_1 G_1 + t_2 G_2 + \dots + t_k G_k$

同样:

$$1) G_1 + G_2 + \dots + G_k = F_1 + F_2 + \dots + F_n = I$$

$$2) G_i G_j = (F_i + \dots + F_j)(F_{i2} + \dots + F_{j2}) = F_i F_{i2} + \dots = 0$$

$$3) G_i G_i = (F_i + \dots + F_j)(F_i + \dots + F_j) = F_i^2 + \dots + F_j^2 + F_i F_{i+1} + \dots + F_j F_{j-1} = F_i^2 + \dots + F_j^2 = F_i + \dots + F_j = G_i$$

3.8.4 谱分解的求法

谱分解: $A = t_1 G_1 + t_2 G_2 + \dots + t_k G_k$

由上节性质可知: $G_i^p = G_i (p \in N^+)$

$$\begin{aligned}
 A^p &= (t_1 G_1 + t_2 G_2 + \cdots + t_k G_k)^p \\
 &= t_1^p G_1^p + \cdots + t_k^p G_k^p + (\text{一堆零项}) \\
 &= t_1^p G_1 + \cdots + t_k^p G_k
 \end{aligned}$$

$$A^0 = I = G_1 + \cdots + G_k$$

根据 $A^p = t_1^p G_1 + \cdots + t_k^p G_k$ 可得:

1. 对于任一多项式: $f(x) = c_0 + c_1 x + \cdots + c_p x^p$, 有 $f(A) = f(t_1)G_1 + \cdots + f(t_k)G_k$

$$\begin{aligned}
 \text{证 明 : } f(A) &= c_0 A^0 + c_1 A + \cdots + c_p A^p = (c_0 + c_1 t_1 + c_2 t_1^2 + \cdots + c_p t_1^p)G_1 + \cdots + \\
 &+ (c_0 + c_1 t_k + c_2 t_k^2 + \cdots + c_p t_k^p)G_k = f(t_1)G_1 + \cdots + f(t_k)G_k
 \end{aligned}$$

2. 由 1 可知, 取 $f(x) = (x - t_1) \cdots (x - t_k)$

$$f(A) = f(t_1)G_1 + \cdots + f(t_k)G_k = 0G_1 + \cdots + 0G_k = 0$$

由上述已知推论, 可取 $f(x) = \cancel{(x - t_1)} \cdots (x - t_k) = (x - t_2) \cdots (x - t_k)$

$$f(A) = \cancel{(A - t_1)} \cdots (A - t_k) = (A - t_2) \cdots (A - t_k)$$

$$f(A) = f(t_1)G_1 + \cdots + f(t_k)G_k = f(t_1)G_1 + 0 + \cdots + 0$$

$$G_1 = \frac{f(A)}{f(t_1)} = \frac{\cancel{(A - t_1)} \cdots (A - t_k)}{\cancel{(t_1 - t_1)} \cdots (t_1 - t_k)}$$

$$G_2 = \frac{(A - t_1)\cancel{(A - t_2)} \cdots (A - t_k)}{(t_2 - t_1)\cancel{(t_2 - t_2)} \cdots (t_2 - t_k)}$$

$$G_i = \frac{(A - t_1) \cdots \cancel{(A - t_i)} \cdots (A - t_k)}{(t_i - t_1) \cdots \cancel{(t_i - t_i)} \cdots (t_i - t_k)}$$

特殊，若单阵A只有两个不同特征根 $t_1 \neq t_2$ (可以是0根)

$$A = t_1 G_1 + t_2 G_2$$

$$G_1 = \frac{(A - t_2)(A - t_2)}{(t_1 - t_2)(t_1 - t_2)} = \frac{(A - t_2)}{(t_1 - t_2)}$$

$$G_2 = \frac{(A - t_1)(A - t_2)}{(t_2 - t_1)(t_2 - t_2)} = \frac{(A - t_1)}{(t_2 - t_1)}$$

$$G_1 + G_2 = I$$

Eg1 : $A = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix}$, 求 A^{100}

解: $\lambda(A) = \{4, 1\}$ $\lambda_1 = 4, \lambda_2 = 1$ A 为单阵

$$\text{所以 } A = \lambda_1 G_1 + \lambda_2 G_2 \rightarrow A^{100} = \lambda_1^{100} G_1 + \lambda_2^{100} G_2$$

$$G_1 = \frac{(A - \lambda_2)}{(\lambda_1 - \lambda_2)} = \frac{A - 1}{4 - 1} = \frac{1}{3} \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

$$G_2 = \frac{(A - \lambda_1)}{(\lambda_2 - \lambda_1)} = \frac{A - 4}{1 - 4} = \frac{1}{3} \begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix}$$

$$\text{所以 } A^{100} = 4^{100} G_1 + 1^{100} G_2$$

Eg2 : $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, 求 A^{100}

解: $\lambda(A) = \{2, 1, 1\} \rightarrow \lambda_1 = 2, \lambda_2 = 1 \rightarrow (A - \lambda_1)(A - \lambda_2) = 0 \rightarrow A$ 为单阵

$$\text{所以 } A^{100} = \lambda_1^{100} G_1 + \lambda_2^{100} G_2$$

$$G_1 = \frac{(A - \lambda_2)}{(\lambda_1 - \lambda_2)} = \frac{A - 1}{2 - 1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$G_2 = \frac{(A - \lambda_1)}{(\lambda_2 - \lambda_1)} = \frac{A - 2}{1 - 2} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

$$A^{100} = 2^{100}G_1 + 1^{100}G_2$$

3.8.5 谱分解的应用

平方谱公式:

若 $A \geq 0$ (半正定), 且有谱公式: $A = \lambda_1 G_1 + \lambda_2 G_2 + \cdots + \lambda_k G_k$, 则有平方

根公式: $\sqrt{A} = A^{\frac{1}{2}} = \sqrt{\lambda_1} G_1 + \cdots + \sqrt{\lambda_k} G_k$

$$(\sqrt{A})^2 = (\sqrt{\lambda_1} G_1 + \cdots + \sqrt{\lambda_k} G_k)^2 = \sqrt{\lambda_1}^2 G_1^2 + \cdots + \sqrt{\lambda_k}^2 G_k^2 + 0 = A$$

Eg1: $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} (A \geq 0) \quad \lambda(A) = \{3, 1\}$, 求 \sqrt{A}

解: $A = 3G_1 + 1G_2 \rightarrow \sqrt{A} = \sqrt{3}G_1 + \sqrt{1}G_2$

$$G_1 = \frac{(A - \lambda_2)}{(\lambda_1 - \lambda_2)} = \frac{A - 1}{3 - 1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$G_2 = \frac{(A - \lambda_1)}{(\lambda_2 - \lambda_1)} = \frac{A - 3}{1 - 3} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

逆谱公式:

若 A 为单阵, 全体不同特征根为 $\lambda_1, \cdots, \lambda_k$ (非 0),

有谱公式: $A = \lambda_1 G_1 + \cdots + \lambda_k G_k$

$$A^{-1} = \frac{1}{\lambda_1} G_1 + \cdots + \frac{1}{\lambda_k} G_k$$

引理: $AG_i = \lambda_i G_i (1 \leq i \leq k)$

$$AG_i = (\lambda_1 G_1 + \cdots + \lambda_k G_k) G_i = \lambda_1 G_1 G_i + \cdots + \lambda_i G_i G_i + \cdots +$$

$$\lambda_k G_k G_i = 0 + \lambda_i G_i + 0 = \lambda_i G_i$$

$$\begin{aligned}
 AA^{-1} &= A \left(\frac{1}{\lambda_1} G_1 + \cdots + \frac{1}{\lambda_k} G_k \right) = \frac{1}{\lambda_1} (AG_1) + \cdots + \frac{1}{\lambda_k} (AG_k) \\
 &= \frac{1}{\lambda_1} (\lambda_1 G_1) + \cdots + \frac{1}{\lambda_k} (\lambda_k G_k) = G_1 + \cdots + G_k = I
 \end{aligned}$$

特征向量相关推论：

1) 若 $AP = tP$ ，则 $P = (X_1, \cdots, X_n)$ 的各列都是 A 关于特征值 t 的特征向量。

原因： $AP = tP \rightarrow AX_i = tX_i$

2) 若 $(A - t)P = 0$ ，则 P 的各列都是 A 关于特征值 t 得特征向量。

a) 若 $A^2 = A \rightarrow (A - 1)A = A^2 - A = 0$ ，所以 A 的各列是 $(A - 1)$ 的特向量， $(A - 1)$ 的各列是 A 的特向量。

b) 若 $(A - \lambda_1)(A - \lambda_2) = 0$ ， $(A - \lambda_1)$ 是关于 λ_2 的特向量， $(A - \lambda_2)$ 是关于 λ_1 的特向量。

3) 谱公式： $A = \lambda_1 G_1 + \cdots + \lambda_k G_k$ 中， G_1, \cdots, G_k 的各列都是 A 的特向量。

4) 谱阵 G_1, \cdots, G_k 中，恰有 n 个无关的特向量。

证明：只要证 $\text{rank}(G_1, \cdots, G_k) = n$

$$(G_1, \cdots, G_k) \begin{pmatrix} I \\ \vdots \\ I \end{pmatrix} = G_1 + \cdots + G_k = I$$

根据秩的相关定理：

$$\text{rank}(AB) \leq \text{rank}(A) \quad \text{rank}(AB) \leq \text{rank}(B)$$

$$\text{rank}(A_{m \times n}) \leq m \quad \text{rank}(A_{m \times n}) \leq n$$

$$n = \text{rank}(I) \leq \text{rank}(G_1, \cdots, G_k)_{n \times nk} \leq n$$

$$\text{rank}(G_1, \cdots, G_k) = n$$

A与f(A)的”遗传公式”:

令 $A = A_{n \times n}$, 全体根 $\lambda(A) = \{\lambda_1, \dots, \lambda_n\}$ 则:

1) 对任一多项式 $f(x) = c_0 + c_1x^1 + \dots + c_px^p$ 的 $f(A)$, 全体根 $\lambda(f(A)) = \{f(\lambda_1), \dots, f(\lambda_n)\}$

2) 若A有特征向量 X_1, \dots, X_n , 则 $f(A)$ 也有特向量 X_1, \dots, X_n

a) 即 $f(A)X_i = f(\lambda_i)X_i$

证明: $AX = tX \rightarrow A^2X = A(tX) = tAX = t^2X \rightarrow A^pX = t^pX$

所以 $(c_0I + c_1A + \dots + c_pA^p)X = (c_0 + c_1t + \dots + c_pt^p)X$

因为 $AX_i = \lambda_iX_i$

所以 $f(A)X_i = f(\lambda_i)X_i$

$$(P^{-1}AP)^k = P^{-1}A^kP$$

$$f(P^{-1}AP) = c_0 + c_1(P^{-1}AP) + \dots + c_p(P^{-1}AP)^p = P^{-1}f(A)P$$

$$\text{记 } P^{-1}AP = B = \begin{pmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \text{ (上三角)}$$

$$f(P^{-1}AP) = P^{-1}f(A)P = f(B) = \begin{pmatrix} f(\lambda_1) & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(\lambda_n) \end{pmatrix} \text{ (上三角)}$$

4 广义逆矩阵

4.1 广义逆矩阵的定义

A^+ 定义：设 $A = A_{m \times n}$ ，若有 $X = X_{n \times m}$ ，满足 4 个条件：

$$1) AXA = A$$

$$2) XAX = X$$

$$3) (AX)^H = AX$$

$$4) (XA)^H = XA$$

称 X 为 A 的一个加号逆，记 $X = A^+$ 。

常见的 A^+ ：

$$1) (0_{m \times n})^+ = 0_{n \times m}$$

$$\text{特别：}(0_n)^+ = 0_n \quad 0^+ = 0$$

$$2) \text{可逆方阵} A \text{ 有 } A^+ = A^{-1}$$

$$3) \text{复数 } a \text{ 可作为 } 1 \text{ 阶阵： } a^+ = \begin{cases} \frac{1}{a}, & a \neq 0 \\ 0, & a = 0 \end{cases}$$

A^+ 的唯一性：

假设 X, Y 都适合 A^+ 条件：

$$\begin{aligned} Y &= YAY = Y(AXA)Y = Y(AX)^H(AY)^H = Y((AY)(AX))^H = Y((AYA)X)^H \\ &= Y(AX)^H = YAX = YA(XAX) = (YA)^H(XA)^HX \\ &= ((XA)(YA))^HX = (X(AYA))^HX = XAX = X \end{aligned}$$

4.2 性质

A^+ 高阵公式:

设 $B = B_{m \times r}$ 为高阵, $\text{rank}(B) = r$, 则: $B^+ = B_L = (B^H B)^{-1} B^H$, 且 $B^+ B = I$

低阵公式:

设 $C = C_{r \times m}$ 为低阵, $\text{rank}(C) = r$, 则: $C^+ = C_R = C^H (C C^H)^{-1}$, 且 $C C^+ = I$

1. 若 A 可逆, 则 $(A^{-1})^H = (A^H)^{-1}$
2. 若 $A^H = A$ 且可逆: $(A^{-1})^H = (A^H)^{-1} = A^{-1}$
3. 若 $B = \begin{pmatrix} A & 0 \end{pmatrix}$, 则 $B^+ = \begin{pmatrix} A^+ \\ 0 \end{pmatrix}$, 若 $B = \begin{pmatrix} A \\ 0 \end{pmatrix}$, 则 $B^+ = \begin{pmatrix} A^+ & 0 \end{pmatrix}$
4. 若 $A = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$, 则 $A^+ = \begin{pmatrix} B^+ & 0 \\ 0 & D^+ \end{pmatrix}$

$$B B^+ B = B \quad B^+ B B^+ = B^+ \quad (B B^+)^H = B B^+$$

高低分解公式:

若 $A = A_{m \times n} = BC$ 为高低分解, 则有 $A^+ = C^+ B^+$, 且 $B^+ B = I = C C^+$,

$$B^+ = (B^H B)^{-1} B^H, \quad C^+ = C^H (C C^H)^{-1}$$

可以证 $C^+ B^+$ 满足正号逆的公式。

例题:

$$1) A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \text{(高阵)}$$

$$\text{解: } A^+ = (A^H A)^{-1} A^H = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$2) A = \begin{pmatrix} 1 & i & 0 \\ 1 & 1 & 1 \end{pmatrix} (\text{低阵})$$

$$\text{解: } A^+ = A^H (AA^H)^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{-i}{3} \\ \frac{-i}{2} & \frac{1}{3} \\ 0 & \frac{1}{3} \end{pmatrix}$$

$$3) A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\text{解: } A = \begin{pmatrix} B & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$A^+ = \begin{pmatrix} B^+ \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

4.3 广义逆求解方式

1. A^+ 第一公式(高低分解):

若 $A = BC$ 为高低分解, 则 $A^+ = C^+B^+$, 其中 $B^+ = (B^HB)^{-1}B^H$, $C^+ = C^H(CC^H)^{-1}$ 。

2. A^+ 第二公式(奇异值分解):

若 $A = PS_rQ^H$ 为正 SVD 分解, 其中 $S_r = \text{diag}\{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_r}\} > 0$, P 、 Q 为半酉阵, 则 $A^+ = QS_r^{-1}P^H$ 。

3. A^+ 秩一第二公式:

若 $A = (a_{ij})_{m \times n} \in C^{m \times n}$, 且 $\text{rank}(A) = 1$, 则 $A^+ = \frac{1}{\sum |a_{ij}|^2} A^H$ 。

证明: 令 $A = PS_rQ^H$, 则 $A^+ = QS_r^{-1}P^H$, $A^H = QS_rP^H$

因为 $\text{rank}(A) = \text{rank}(A^HA) = 1$

所以 $\lambda(A^HA) = \{\text{tr}(A^HA), 0, 0, \dots\}$

$$\lambda_1 = \text{tr}(A^H A) = \sum |a_{ij}|^2 \quad S_r = (\sqrt{\lambda_1})$$

$$\text{所以 } A^+ = Q(\sqrt{\lambda_1})^{-1} P^H = \frac{1}{\lambda_1} Q(\sqrt{\lambda_1}) P^H = \frac{1}{\lambda_1} A^H$$

4. A^+ 第三公式(QR 分解):

若 $A = QR$, R 为上三角, Q 为酉阵, 则 $A^+ = R^{-1} Q^H$ 。

5. A^+ 谱公式:

若 A 为正规, 且有谱公式: $A = \lambda_1 G_1 + \cdots + \lambda_k G_k$, 则

$$A^+ = \lambda_1^+ G_1 + \cdots + \lambda_k^+ G_k$$

4.4 正规方程

定义:

任一方程 $AX = b$, 可产生方程 $A^H AX = A^H b$, 叫做正规方程。

任一矩阵 $A = A_{m \times n} \in C^{m \times n}$ 可产生两个子空间:

1. 令 $N(A) = \{X | AX = 0, X \in C^n\}$, 叫作 A 的核空间(解空间)。
2. 令 $R(A) = \{AX | X \in C^n\} \in C^m$, 叫作 A 的值域或像空间(A 的列空间)。

引理:

1. 正规方程 $A^H AX = A^H b$ 一定有解(相容), 且有特解 $X_0 = A^+ b$ 。

证明: $A^H AX_0 = A^H A A^+ b = A^H (A A^+)^H b = [(A A^+)^H A]^H b = A^H b$

2. $A^H AX = A^H b$ 有以下通解公式:

$$X = X_0 + (k_1 Y_1 + \cdots + k_{n-r} Y_{n-r}),$$

其中 $X_0 = A^+ b$, $Y = k_1 Y_1 + \cdots + k_{n-r} Y_{n-r}$, $Y_i \in N(A)$, $Y \in N(A)$

3. 若 $AX = b$ 有解, 则必有特解 $X_0 = A^+ b$, 有通解 $X = X_0 + (k_1 Y_1 + \cdots +$

$k_{n-r}Y_{n-r})$, Y_i 是 $AX = 0$ 的基本解。

证明：可设 X_1 为 $AX = b$ 的任一解： $AX_1 = b$

将 $X_0 = A^+b$ 代入公式：

左边 $= AA^+b = AA^+AX_1 = AX_1 = b =$ 右边

所以 X_0 是 $AX = b$ 的解

正交引理：

1. 若 $A = A_{m \times n} \in C^{m \times n}$, $b \in C^m$, $X_0 = A^+b \in C^n$, 则 $X_0 \perp N(A)$ (核空间),

即 $X_0 \perp X$, $X \in N(A) = \{X | AX = 0\}$

证明：任取 $X \in N(A)$, 有 $AX = 0$

$$\begin{aligned}(X, X_0) &= X_0^H X = (A^+b)^H X = (A^+AA^+b)^H X = (A^+b)^H (A^+A)^H X \\ &= (A^+b)^H A^+AX = (A^+b)^H A^+(AX) = 0\end{aligned}$$

2. 若 $A = A_{m \times n} \in C^{m \times n}$, $b \in C^m$, $X_0 = A^+b \in C^n$, 则 $(AX_0 - b) \perp R(A)$ (像

空间), 即 $(AX_0 - b) \perp AX$, $X \in C^n$

证明：

$$\begin{aligned}(AX_0 - b, AX) &= (AX)^H (AX_0 - b) = X^H A^H (AX_0 - b) = X^H (A^H AX_0 - \\ &A^H b) = X^H 0 \text{(定理 1: } A^H AX_0 = A^H b)\end{aligned}$$

4.5 最小二乘解

最小二乘解：

若 $AX = b$ 无解，称 $AX = b$ 不相容(矛盾方程), $|AX - b|^2 > 0$, 则最小值
 $\min\{|AX - b|^2\} = |AX_0 - b|^2$, $X_0 = A^+b$,

此时 X_0 是 $AX = b$ 的一个最小二乘解，其他最小二乘解也满足 $AX = AX_0 \Leftrightarrow A(X - X_0) = 0$ 。

全体的最小二乘解为 $X = X_0 + (k_1 Y_1 + \cdots + k_{n-r} Y_{n-r})$ ， Y_i 是 $AX = 0$ 的基本解。

证明：

$$AX - b = (AX - AX_0) + (AX_0 - b) = A(X - X_0) + (AX_0 - b)$$

因为 $AX \in R(A)$, $AX_0 \in R(A)$ ，所以 $A(X - X_0) \in R(A)$ ，

所以 $A(X - X_0) \perp (AX_0 - b)$

$$|AX - b|^2 = |A(X - X_0)|^2 + |AX_0 - b|^2 + 0 \geq |AX_0 - b|^2 (X = X_0)$$

求解： $AY = 0, (Y = X - X_0)$ ，

通解 $Y = t_1 Y_1 + t_2 Y_2 + \cdots + t_{n-r} Y_{n-r}$ ， $r = \text{rank}(A)$ ， Y 为 $AY = 0$ 的基本解，即 $Y_i \in N(A)$ 。

所以最小二乘解 $X - X_0 = t_1 Y_1 + t_2 Y_2 + \cdots + t_{n-r} Y_{n-r}$

$$X = X_0 + t_1 Y_1 + t_2 Y_2 + \cdots + t_{n-r} Y_{n-r}$$

例题：求 $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} X = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ 得最小二乘解。

解： $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ (秩 1), $b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$$X_0 = A^+ b = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{3}{4} \end{pmatrix}$$

设 $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ， $AY = 0 \rightarrow y_1 = -y_2$ ，所以 $Y = t \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

所以所有最小二乘解： $X = \begin{pmatrix} \frac{3}{4} \\ \frac{4}{3} \\ \frac{3}{4} \\ \frac{4}{4} \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

极小范数解：

若 $AX = b$ 有解，则必有特解 $X_0 = A^+b$ ，通解： $X = X_0 + (k_1Y_1 + \dots + k_{n-r}Y_{n-r})$ ， $X_0 \perp Y_i$ 。 X_0 是全体解中最小长度(范数)解，称为极小范数解，也叫最佳小二解。

$$\text{证明： } |X|^2 = |X_0|^2 + |Y|^2 \geq |X_0|^2$$

推论：

1. 若 $AX = b$ 有解，则 $AX = b$ 和 $A^HAX = A^Hb$ 同解， $N(A) = N(A^HA)$ ， $\text{rank}(A) = \text{rank}(A^HA)$
2. 若 $AX = b$ 无解， $A^HAX = A^Hb$ 的通解恰为 $AX = b$ 的全体最小二乘解。

4.6 补充公式

公式：

1. $A^+ = (A^HA)^+A^H$
2. $A^+ = A^H(AA^H)^+$

证明：

$$\text{令 } A = PS_rQ^H, A^H = QS_rP^H, A^+ = QS_r^{-1}P^H$$

$$A^HA = (QS_rP^H)(PS_rQ^H) = QS_r^2Q^H \text{ (正好是 } A^HA \text{ 的正SVD分解)}$$

$$(A^HA)^+ = Q(S_r^2)^{-1}Q^H$$

$$(A^H A)^+ A^H = \left(Q(S_r^2)^{-1} Q^H \right) (Q S_r P^H) = Q S_r^{-1} P^H = A^+$$

用法:

矩阵A不一定是正规矩阵，所以不一定可以进行谱分解，但 $A^H A$ 一定可以。

$$A^H A = \lambda_1 G_1 + \cdots + \lambda_k G_k$$

$$(A^H A)^+ = \lambda_1^+ G_1 + \cdots + \lambda_k^+ G_k$$

其他结论：矩阵方程 $AXB = D$

1) 若 $AXB = D$ 有解(相容)则必有特解 $X_0 = A^+ D B^+$ 与 $Y_0 = A^- D B^-$

证：设任一解为 X_1

$$A Y_0 B = A A^- D B^- B = A A^- (A X_1 B) B^- B = (A A^- A) X_1 (B B^- B) = A X_1 B$$

2) 若 $AXB = D$ 有解，则 $X_0 = A^+ D B^+$ 是最小范数解： $\|X\|^2 \geq \|X_0\|^2$

3) 若 $AXB = D$ 不相容，则 $X_0 = A^+ D B^+$ 是最佳小二解

4.7 A^- 定义与性质

4.7.1 A^- 定义

若 $A = A_{m \times n}$ 与 $X = X_{n \times m}$ ，适合条件 $AXA = A$ ，则称X是A的一个减号逆，记作 $X = A^-$ 或 $X = A^{(1)}$ ，可写： $AA^-A = A$ 。

$A^- \in A^{(1)} = \{X | AXA = A\}$ ($A^{(1)}$ 代表只符合广义逆四条性质中的第一条)

特别: $A^+ \in A^{(1)}$

注: A^- 不唯一

Eg: $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, $A^- = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 或 $A^- = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

$$a^- = \begin{cases} \text{全体数}x, & a = 0 \\ \frac{1}{a}, & a \neq 0 \end{cases} \quad (a, x \in C)$$

若 $A = A_{n \times n}$ (方阵) 可逆, 则 $A^{-1} = A^-$

性质:

- 1) 若 $AX = b$ 有解, 则必有特解 $X_0 = A^+b$ 与 $Y_0 = A^-b$
- 2) 若 $AXB = D$ 有解(相容)则必有特解 $X_0 = A^+DB^+$ 与 $Y_0 = A^-DB^-$

4.7.2 A^- 相关公式

公式一:

若 $A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m \times n}$ (标准形), 则全体 $A^- = \begin{pmatrix} I_r & C \\ D & F \end{pmatrix}_{n \times m}$, 其中子块 CDF 元素任意取值。

$$\text{Pf: } AA^-A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_r & C \\ D & F \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

其他标准形: $A = (I_m \ 0)_{m \times n}$ $A^- = \begin{pmatrix} I_m \\ C \end{pmatrix}_{n \times m}$

$$A = \begin{pmatrix} I_n \\ 0 \end{pmatrix}_{m \times n} \quad A^- = (I_n \ C)_{n \times m}$$

公式二:

若 $A = A_{m \times n}$ 与标准形 $B = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m \times n}$ 等价, 即存在 $PAQ = B$, PQ 可逆, 则全体 $A^- = Q \begin{pmatrix} I_r & C \\ D & F \end{pmatrix}_{n \times m} P$.

$$\text{Pf: } PAQ = B \rightarrow A = P^{-1}BQ^{-1}$$

$$AA^-A = (P^{-1}BQ^{-1})Q \begin{pmatrix} I_r & C \\ D & F \end{pmatrix}_{n*m} P(P^{-1}BQ^{-1}) =$$

$$P^{-1} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I_r & C \\ D & F \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} = P^{-1} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} = A$$

$$\text{令} \begin{pmatrix} A & I_m \\ I_n & 0 \end{pmatrix} (\text{行列变}) \rightarrow \begin{pmatrix} B & P \\ Q & 0 \end{pmatrix}, B = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

$$\text{全体} A^- = Q \begin{pmatrix} I_r & C \\ D & F \end{pmatrix}_{n*m} P$$

4.7.3 更多广逆

$A^{(1,3)}$ 定义：若 $A = A_{m*n}$ 与 $X = X_{n*m}$ ，适合条件：

$$1) \quad AXA = A$$

$$2) \quad (AX)^H = AX$$

记 $X = A^{(1,3)}$ (不唯一)

$A^{(1,4)}$ 定义：若 $A = A_{m*n}$ 与 $X = X_{n*m}$ ，适合条件：

$$1) \quad AXA = A$$

$$2) \quad (XA)^H = XA$$

记 $X = A^{(1,4)}$ (不唯一)

5 矩阵分析

5.1 向量范数

5.1.1 向量范数

任 $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in C^n$, 规定长度(范数)为:

$$\|X\| = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$$

记为 $|X|$ 。

性质:

- 1) 正性: $\|X\| > 0 (X \neq \vec{0})$, 且 $\|X\| = 0 \leftrightarrow X = \vec{0}$
- 2) 齐性: $\|kX\| = |k|\|X\|$, $k \in C$
- 3) 三角性: $\|X + Y\| \leq \|X\| + \|Y\|$

$$\|X - Y\| \geq ||\|X\| - \|Y\||$$

范数定义

V 是线性空间(在复数域), φ 是 V 上一个实值函数: $\varphi(X) = \|X\|$ 叫范数, 如果适合正性、齐性、三角性。

5.1.2 常见向量范数

空间 $V \in C^n$ 上常见范数 $\varphi(X)$

1. 取 $\varphi(X) = \|X\|_2 = \sqrt{(X, X)} = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$ 叫 2 范数, 又叫 F 范数(长度), 有公式: $\|X\| = \sqrt{\text{tr}(X^H X)} = \sqrt{\text{tr}(X X^H)}$
2. 取 $\varphi(X) = \|X\|_\infty = \max\{|x_1|, \cdots, |x_n|\}$, 叫最大值范数, 又叫“ ∞ 范数”
3. 取 $\varphi(X) = \|X\|_1 = |x_1| + \cdots + |x_n|$, 叫和范数

4. 取 $\varphi(X) = \|X\|_p = (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}}$, 叫 p 范数, $p \geq 1$.

取极限: $\lim_{p \rightarrow \infty} (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}} = \max\{|x_1|, \dots, |x_n|\} = \|X\|_\infty$

等价定理:

C^n 上任 2 种范数 $\|X\|_a$ 与 $\|X\|_b$, 适合: $k_1 \leq \frac{\|X\|_a}{\|X\|_b} \leq k_2, \forall X \in C^n$, 其中 $0 < k_1 < k_2$ 为固定正数, 称 $\|X\|_a$ 与 $\|X\|_b$ 等价, 记 $\|X\|_a \approx \|X\|_b$

5.2 矩阵范数

5.2.1 定义

方阵空间 $C^{n \times n}$ 上一函数 $\varphi(A) = \|A\|$ 叫一个方阵范数, 当且仅当满足以下性质:

- 1) 正性: $\varphi(A) = \|A\| > 0, A \neq 0$ 。且 $\varphi(0) = \|0\| = 0$
- 2) 齐性: $\varphi(kA) = |k|\varphi(A)$
- 3) 三角性 $\varphi(A + B) \leq \varphi(A) + \varphi(B)$

推论: $|||A| - |B||| \leq \|A - B\|$

- 4) 次乘性(相容性): $\varphi(AB) \leq \varphi(A)\varphi(B)$

只满足(1)~(3)的函数 $\varphi(A) = \|A\|$ 叫向量式范数

性质 4 的推广: $\|A^k\| \leq \|A\|^k$

对 $\|AB\| \leq \|A\|\|B\|$ 迭代 k 次: $\|A_1 A_2 \dots A_k\| \leq \|A_1\| \dots \|A_k\|$

令 $A_1 = A_2 = \dots = A_k$, $\|A^k\| \leq \|A\|^k$

5.2.2 常见的矩阵范数

$C^{n \times n}$ 上常见范数: $A = (a_{ij})_{n \times n}$

1. $\varphi(A) = \|A\|_1 = \max\{L_1, \dots, L_n\}$ 叫列范数, $L_j = |a_{1j}| + \dots + |a_{nj}|$

2. $\varphi(A) = \|A\|_\infty = \max\{R_1, \dots, R_n\}$ 叫行范数, $R_j = |a_{j1}| + \dots + |a_{jn}|$

3. $\varphi(A) = \|A\|_F = \sqrt{\sum |a_{ij}|^2} = \sqrt{\text{tr}(A^H A)}$ 叫 F 范数

4. $\varphi(A) = \|A\|_2 = \sqrt{\lambda_1}$ 最大奇值

$\lambda(A^H A) = \{\lambda_1, \dots, \lambda_n\}$ 并且 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$

5. $\varphi(A) = \|A\|_M = \sum |a_{ij}|$ 叫总和范数

6. $\varphi(A) = \|A\|_G = n * \max\{|a_{ij}|\}$

等价定理:

任 2 个方阵范数 $\|A\|_a, \|A\|_b$, $0 < k_1 \leq \frac{\|A\|_a}{\|A\|_b} \leq k_2$

常见向量范数之间的等价关系 $X = (x_1, \dots, x_n)^T$:

$$1) \quad 1 \leq \frac{\|X\|_1}{\|X\|_\infty} \leq n$$

$$2) \quad 1 \leq \frac{\|X\|_1}{\|X\|_2} \leq \sqrt{n}$$

Pf: (1) $\|X\|_1 = |x_1| + \dots + |x_n| \leq n * \max\{x_1, \dots, x_n\} = n\|X\|_\infty$

$$\|X\|_1 = |x_1| + \dots + |x_n| \geq \max\{x_1, \dots, x_n\} = \|X\|_\infty$$

(2) 由柯西不等式可得:

$$\|X\|_1 = |x_1| + \dots + |x_n| \leq \sqrt{1^2 + \dots + 1^2} \sqrt{|x_1|^2 + \dots + |x_n|^2} = \sqrt{n} \|X\|_2$$

$$\|X\|_1^2 = (|x_1| + \dots + |x_n|)^2 \geq |x_1|^2 + \dots + |x_n|^2 = \|X\|_2^2$$

$$\|X\|_1 \geq \|X\|_2$$

5.2.3 谱半径

谱半径 $\rho(A) = \max\{|\lambda_1|, \dots, |\lambda_n|\}$, $A = A_{n \times n}$, $\lambda(A) = \{\lambda_1, \dots, \lambda_n\}$

性质:

$$1) \quad \rho(A) \geq 0$$

$$2) \quad \text{齐性: } \rho(kA) = |k|\rho(A)$$

$$\lambda(kA) = \{k\lambda_1, \dots, k\lambda_n\}$$

$$3) \quad \rho(A^p) = \rho(A)^p$$

$$\lambda(A^p) = \{\lambda_1^p, \dots, \lambda_n^p\}$$

谱范数不等式: $\rho(A) \leq \|A\|$, 对一切方阵成立

证明:

$$\text{令 } |\lambda_1| \geq \dots \geq |\lambda_n|, \quad \rho(A) = |\lambda_1|$$

$$\text{取特征向量 } X: AX = \lambda_1 X$$

$$\text{令 } B = (X, \dots, X)$$

$$AB = (AX, \dots, AX) = (\lambda_1 X, \dots, \lambda_1 X) = \lambda_1 B$$

$$\|AB\| = \|\lambda_1 B\| = |\lambda_1| \|B\|$$

由方阵范数条件 4 可知:

$$|\lambda_1| \|B\| = \|AB\| \leq \|A\| \|B\|$$

$$|\lambda_1| \leq \|A\|$$

5.2.4 算子范数

定义：已知 $\|X\|_V$ 为 C^n 上给定的向量范数，固定一个方阵 A ， $\|AX\|_V$ 在 C^n 上连续，令 $\varphi(A) = \max\{\|AX\|_V\}$ （前提条件： $\|X\|_V = 1$ ）。

即当 $\|X\|_V = 1$ 时，存在 X_0 ， $\|X_0\|_V = 1$ ，且 $\varphi(A) = \|AX_0\|_V$

有以下性质：

$$\|AX\|_V \leq \varphi(A)\|X\|_V$$

且满足矩阵范数的 4 个性质。

这时 $\varphi(A)$ 是方阵范数，记为 $\varphi(A) = \|A\|_V$ ，称为 $\|X\|_V$ 导出的“算子范数”。

性质证明：

先证： $\|AX\|_V \leq \varphi(A)\|X\|_V$

$$\text{任 } Y = \frac{X}{\|X\|} \rightarrow \|Y\| = 1$$

$$\|AY\| \leq \max\{\|AX\|\} = \varphi(A)$$

$$\left\| A \frac{X}{\|X\|} \right\| = \frac{\|AX\|}{\|X\|} \leq \varphi(A)$$

$$\|AX\| \leq \varphi(A)\|X\|$$

证性质 3： $\varphi(A+B) = \max\{\|(A+B)X\|\} = \|(A+B)X_0\| = \|AX_0 + BX_0\| \leq \|AX_0\| + \|BX_0\| \leq \varphi(A) + \varphi(B)$ X_0 为 $\|(A+B)X\|$ 取最大值对应的 X

证性质 4：

$$\varphi(AB) = \max\{\|(AB)X\|\} = \|(AB)X_0\| = \|A(BX_0)\|$$

$$= \left\| A \frac{1}{\|BX_0\|} (BX_0) \right\| * \|BX_0\| \leq \varphi(A) * \|BX_0\| \leq \varphi(A)\varphi(B)$$

常见的算子范数:

1. $\|X\|_1$ 导出 $\varphi(A) = \|A\|_1$ (列)
2. $\|X\|_\infty$ 导出 $\varphi(A) = \|A\|_\infty$ (行)
3. $\|X\|_2$ 导出 $\varphi(A) = \max\{\|AX\|_2\} = \|A\|_2$ (最大奇值)

注: $\|AX\|_1 \leq \|A\|_1 \|X\|_1$

$$\|AX\|_\infty \leq \|A\|_\infty \|X\|_\infty$$

$$\|AX\|_2 \leq \|A\|_2 \|X\|_2$$

一般 $\|AX\|_v = \|A\|_v \|X\|_v$

证明: $\varphi(A) = \max\{\|AX\|_2\} = (\text{最大奇值})\sqrt{\lambda_1}$ ($\|X\|_2 = 1$), 记为 $\|A\|_2$

令长度 $|X| = \|X\|_2 = \sqrt{|x_1|^2 + \dots + |x_n|^2}$

即证 $|AX|^2$ (最大值) $= \lambda_1$, 条件: $|X|^2 = 1$

$$|AX|^2 = (AX)^H(AX) = X^H(A^H A)X \text{ (二次形)}$$

$$A^H A \geq 0 \text{ (半正定, Hermite 阵)}$$

由 Hermite 分解可知, 存在酉阵 Q :

$$Q^H(A^H A)Q = D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$

$$\text{全体奇值 } S(A) = \{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}\}$$

由酉阵保长性质可知: $|QX|^2 = |X|^2$

$$X^H(A^H A)X = Y^H Q^H(A^H A)QY = Y^H D Y \quad (X = QY)$$

$$Y^H D Y = (\overline{y_1} \quad \dots \quad \overline{y_n}) \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$X^H(A^H A)X = \lambda_1 |y_1|^2 + \cdots + \lambda_n |y_n|^2 \leq \lambda_1 |y_1|^2 + \cdots + \lambda_1 |y_n|^2 = \lambda_1 |Y|^2$$

$|Y|^2 = 1$ 时, $X^H(A^H A)X$ 取最大值 λ_1

所以 $|X| = \|X\|_2$ 时, $\|A\|_2 = \sqrt{\lambda_1}$

单位阵范数性质:

1. 对任一算子范数 $\|A\|$, 必有 $\|I_n\| = 1$

由定义 $\varphi(I_n) = \max\{\|I_n X\|\} = 1 (\|X\| = 1)$

推论: 若 $\|I_n\| > 1$, 则为非算子范数

2. 对于其它方阵范数, 必有 $\|I_n\| \geq 1$

谱范不等式: $1 = \rho(I_n) \leq \|I_n\|$

5.2.5 小范数定理

引理: 若 $\|A\|$ 为 $C^{n \times n}$ 上已知的方阵范数, P 为可逆阵,

令 $\varphi(A) = \|P^{-1}AP\|$ (或 $\varphi(A) = \|PAP^{-1}\|$)

则 $\varphi(A)$ 为 $C^{n \times n}$ 上一个新范数, 满足条件 1~4

证明条件 3: $\varphi(A+B) \leq \varphi(A) + \varphi(B)$

$$\varphi(A+B) = \|P^{-1}(A+B)P\| = \|P^{-1}AP + P^{-1}BP\| \leq \|P^{-1}AP\| +$$

$$\|P^{-1}BP\| = \varphi(A) + \varphi(B)$$

证明条件 4: $\varphi(AB) \leq \varphi(A)\varphi(B)$

$$\varphi(AB) = \|P^{-1}(AB)P\| = \|(P^{-1}AP)(P^{-1}BP)\| \leq \|P^{-1}AP\| \|P^{-1}BP\| =$$

$$\varphi(A)\varphi(B)$$

小范数定理:

固定一个方阵A与小正数 $\varepsilon > 0$, 则存在一个方阵范数 $\|A\|_\varepsilon$, 使得:

$$\|A\|_\varepsilon \leq \rho(A) + \varepsilon$$

证明:

$$\text{由 Jordan 形 可知: } P^{-1}AP = B = \begin{pmatrix} \lambda_1 & * & 0 & 0 \\ \vdots & \lambda_2 & * & \vdots \\ \vdots & 0 & \ddots & \\ 0 & \dots & & \lambda_n \end{pmatrix} \text{(双线上三角,*}$$

为 1 或 0)

$$\text{令 } D = \begin{pmatrix} \varepsilon^1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \varepsilon^n \end{pmatrix}, D^{-1} = \begin{pmatrix} \frac{1}{\varepsilon^1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\varepsilon^n} \end{pmatrix}$$

$$D^{-1}BD = \begin{pmatrix} \lambda_1 & * \varepsilon & 0 & 0 \\ \vdots & \lambda_2 & * \varepsilon & \vdots \\ \vdots & 0 & \ddots & \\ 0 & \dots & & \lambda_n \end{pmatrix}$$

$$\|D^{-1}BD\|_\infty \leq |\lambda_1| + \varepsilon$$

令 $W = PD$ (可逆)

$$\|A\|_\varepsilon = \|W^{-1}AW\|_\infty = \|D^{-1}BD\|_\infty \leq |\lambda_1| + \varepsilon = \rho(A) + \varepsilon$$

推论: 若 $\rho(A) < 1$, 则存在小范数, 使 $\|A\|_\varepsilon < 1$ 。

证明: 令 $\varepsilon = \frac{1-\rho(A)}{2}$

$$\text{则存在 } \|A\|_\varepsilon \leq \rho(A) + \varepsilon = \frac{1+\rho(A)}{2} < 1$$

条件数定义: $C(A) = \|A\|\|A^{-1}\|$ (A 可逆)

$$C(A) = \|A\|\|A^{-1}\| \geq \|AA^{-1}\| = \|I\| \geq 1$$

5.3 矩阵级数

设方阵序列 $A_1, A_2, \dots \in C^{n \times n}$, 记 $A_k = (a_{ij}^k)_{n \times n}$,

$$S_p = A_1 + \dots + A_p = \left(\sum_{k=1}^p a_{ij}^k \right)_{n \times n}$$

若矩阵级数: $\sum_{k=1}^{\infty} A_k = A_1 + \dots + A_k + \dots \xrightarrow{\text{收敛}} S$

则部分和: $S_p = A_1 + \dots + A_p \quad \lim_{p \rightarrow \infty} S_p = S = (s_{ij})_{n \times n}$

若每个级数 $\sum_{k=1}^{\infty} a_{ij}^k \xrightarrow{\text{收敛}} s_{ij}$, 则 $\sum_{k=1}^{\infty} A_k \xrightarrow{\text{收敛}} S$

注: 若 $\sum_{i=1}^{\infty} |a_i| < \infty$, 则 $\sum_{i=1}^{\infty} a_i$ 收敛

引理: 若 $\sum_{k=1}^{\infty} \|A_k\| = \|A_1\| + \dots + \|A_k\| + \dots$ 收敛, 则 $\sum_{k=1}^{\infty} A_k$ 收敛

即 “绝对收敛, 必收敛”

证明: 只要证 n^2 个 $\sum_{k=1}^{\infty} a_{ij}^k$ 都绝对收敛

取 M 范数 $\|A\|_M = \sum |a_{ij}|$

等价性: $0 < k_1 < \frac{\|A\|_M}{\|A\|} < k_2$

$$\rightarrow \sum_{k=1}^{\infty} \|A_k\|_M < k_2 \sum_{k=1}^{\infty} \|A_k\| \text{ (收敛)}$$

$\rightarrow \|A_1\|_M + \dots + \|A_k\|_M + \dots$ 收敛

$\rightarrow \sum (\sum_{k=1}^{\infty} |a_{ij}^k|)$ 收敛

\rightarrow 每个级数 $\sum_{k=1}^{\infty} |a_{ij}^k|$ 收敛

$\rightarrow \sum_{k=1}^{\infty} a_{ij}^k$ 收敛

$\rightarrow \sum (\sum_{k=1}^{\infty} a_{ij}^k)$ 收敛

$\rightarrow \sum_{k=1}^{\infty} A_k$ 收敛

牛曼公式:

1) $\rho(A) < 1$ 时, $\sum_{k=0}^{\infty} A^k = I + A + \cdots + A^n + \cdots = (I - A)^{-1}$ (收敛)

2) 某个 $\|A\| < 1$ 时, $\sum_{k=0}^{\infty} A^k = (I - A)^{-1}$

3) 若 $\rho(A) \geq 1$, 则 $\sum_{k=0}^{\infty} A^k$ 发散

证明: 若 $\rho(A) < 1$, 则 $\sum_{k=1}^{\infty} A^k$ 收敛

$\rho(A) < 1$, 则存在小范数 $\|A\|_{\varepsilon} = b < 1$

$\|A^k\|_{\varepsilon} \leq \|A\|_{\varepsilon}^k = b^k$

$\rightarrow \sum_{k=1}^{\infty} \|A^k\|_{\varepsilon} \leq \sum_{k=1}^{\infty} b^k$ ($0 < b < 1$) 收敛

则 $\sum_{k=1}^{\infty} A^k$ 收敛 (绝对收敛引理)

令 $\sum_{k=0}^{\infty} A^k = S$

$S(I - A) = S - SA$

$= (I + A + \cdots + A^n + \cdots) - (I + A + \cdots + A^n + \cdots)A = I$

所以 $S = (I - A)^{-1}$

注: 若 $\sum_{k=1}^{\infty} A^k$ 收敛, 则 $A^k \xrightarrow{\text{收敛}} 0$

证明: $\rho(A) \geq 1$, $\sum_{k=1}^{\infty} A^k$ 不收敛

反证法：若 $\sum_{k=1}^{\infty} A^k$ 收敛，则 $A^k \xrightarrow{\text{收敛}} 0$

$$\rightarrow \|A^k\|_M \xrightarrow{\text{收敛}} 0$$

$$\rightarrow \rho(A^k) \leq \|A^k\|_M$$

$$\rightarrow \rho(A)^k \xrightarrow{\text{收敛}} 0$$

$$\rightarrow \rho(A) < 1 (\text{矛盾})$$

5.4 特征根估计

5.4.1 估计方法

谱范不等式 $\rho(A) \leq \|A\|$ ：

$$1) \rho(A) = \max\{|\lambda_1|, \dots, |\lambda_n|\} \leq \|A\|_{\infty}$$

$$2) \rho(A) \leq \|A\|_1$$

$$3) \rho(A) \leq \|A\|_G$$

许尔估计：

$$|\lambda_1|^2 + \dots + |\lambda_n|^2 \leq \sum |a_{ij}|^2, \text{ 当且仅当 } A \text{ 正规时等号成立}$$

5.4.2 盖尔(Ger)圆盘

盖尔圆盘：以 a_{ii} 为圆心， R_i 为半径的圆

定义：方阵 $A = (a_{ij})_{n \times n}$ 恰有 n 个盖尔圆盘 $G_1 \sim G_n$ ，其中：

$$G_1: |z - a_{11}| \leq R_1 \quad \dots \quad G_n: |z - a_{nn}| \leq R_n$$

半径： $R_i = (\sum_{j=1}^n |a_{ij}|) - |a_{ii}|$ 即每行除对角元之外元素绝对值之和

盖尔定理：

-
1. 全体特征根 $\lambda(A)$ 被所有盖尔圆盘所覆盖，即：

$$\{\lambda_1, \dots, \lambda_n\} \subset G_1 \cup \dots \cup G_n$$

2. 若 p 个圆盘相交(或相切)形成一个连通分支 D_p ，且与其余 $n-p$ 个圆盘分离，则 D_p 中恰有 p 个根(含重复)。

特别，一个独立圆 D_1 中恰有 1 个根。

应用：若矩阵 A 的盖尔盘比较难求，可以求 $B = P^{-1}AP$ 的盖尔盘。因为相似矩阵特征根相同。

实方阵性质：

1. n 个中心 a_{11}, \dots, a_{nn} 都在实轴上。
2. 实方阵的虚根成双出现(共轭虚根)

$$\text{例如： } \lambda_1 = a + bi, \lambda_2 = a - bi$$

对角占优阵

若对角占优阵 A 适合条件：

$$\text{行对角占优： } |a_{ii}| > R_i$$

$$\text{列对角占优： } |a_{ii}| > R'_i (\text{非对角元列元素之和})$$

推论：

若 A 对角占优，则 $\forall G_i, \mathbf{0} \notin G_i$ ，即 $\mathbf{0} \notin \lambda(A) = \{\lambda_1, \dots, \lambda_n\}$ ， $|A| = \lambda_1 * \dots * \lambda_n \neq 0$ ， A 为可逆阵。

A 与转置 A^T :

1. 具有相同特征根: $\lambda(A) = \lambda(A^T)$
2. 特征式相同: $|\lambda I - A| = |\lambda I - A^T|$
3. A^T 的(Ger)行半径恰是 A 的(Ger)列半径

5.5 矩阵函数

5.5.1 收敛定理

令解析函数 $f(x) = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 x + \dots$ 收敛半径为 r , 或者

$$f(x) = \sum_{k=0}^{\infty} a_k (x - c) = a_0 + a_1 (x - c) + \dots$$

将 $x=A$ 代入, $f(x) = \sum_{k=0}^{\infty} c_k A^k$ 或 $f(x) = \sum_{k=0}^{\infty} a_k (A - c)$

收敛定理: $f(x) = \sum_{k=0}^{\infty} c_k x^k$ 收敛半径为 r , 且 $A \in C^{n \times n}$

1. $\rho(A) < r$ 时, $f(A)$ 收敛
2. $\rho(A) > r$ 时, $f(A)$ 发散
3. $\rho(A) = r$ 时, 待定

收敛半径求法: 记 $\lim_{n \rightarrow \infty} \left| \frac{c_{n+1}}{c_n} \right| = p$

- 1) $p \neq 0, r = \frac{1}{p}$
- 2) $p = 0, r = +\infty$
- 3) $p = +\infty, r = 0$

例: $f(x) = \sum_{k=0}^{\infty} x^k$, $p=1$, 收敛半径 $r=1$. $\rho(A) < 1$ 时, $f(A)$ 收敛

5.5.2 常见解析函数

3 个解析函数 $f(x) = e^x$, $f(x) = \sin x$, $f(x) = \cos x$, 收敛半径都是 $+\infty$

$$f(x) = e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

$$f(x) = \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

$$f(x) = \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^k \frac{x^{2k}}{(2k)!}$$

任方阵 A 代入解析函数:

$$f(A) = e^A = I + A + \frac{A^2}{2!} + \dots$$

$$f(A) = \sin A = A - \frac{A^3}{3!} + \frac{A^5}{5!} - \dots + (-1)^k \frac{A^{2k+1}}{(2k+1)!}$$

$$f(A) = \cos A = I - \frac{A^2}{2!} + \frac{A^4}{4!} - \dots + (-1)^k \frac{A^{2k}}{(2k)!}$$

$$\sin(-A) = -\sin A$$

$$\cos(-A) = \cos A$$

若 A 为单阵, 则对任一解析函数 $f(x) = \sum_{k=0}^{\infty} c_k x^k$, 必有

$$f(A) = f(\lambda_1)G_1 + \dots + f(\lambda_n)G_n$$

引入参数 t (实数或复数)

$$e^{tA} = I + tA + \frac{(tA)^2}{2!} + \dots$$

$$\sin(tA) = tA - \frac{(tA)^3}{3!} + \frac{(tA)^5}{5!} - \dots + (-1)^k \frac{(tA)^{2k+1}}{(2k+1)!}$$

$$\cos(tA) = I - \frac{(tA)^2}{2!} + \frac{(tA)^4}{4!} - \dots + (-1)^k \frac{(tA)^{2k}}{(2k)!}$$

$$\frac{de^{tA}}{dt} = 0 + A + \frac{2A^2t}{2!} + \dots = A \left(I + tA + \frac{(tA)^2}{2!} + \dots \right) = Ae^{tA}$$

$$\frac{d \sin(tA)}{dt} = A \cos(tA)$$

$$\frac{d \cos(tA)}{dt} = -A \sin(tA)$$

5.5.3 幂等公式

公式：若 $A^2 = A$ (可知 $A^k = A$)，则 $e^{tA} = I + (e^t - 1)A$

证明：

$$\begin{aligned} e^{tA} &= I + tA + \frac{(tA)^2}{2!} + \frac{(tA)^3}{3!} + \dots = I + tA + \frac{t^2A}{2!} + \frac{t^3A}{3!} + \dots \\ &= I + A \left(t + \frac{t^2}{2!} + \frac{t^3}{3!} + \dots \right) = I + (e^t - 1)A \end{aligned}$$

Eg: $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, 求 e^{tA}

$$A^2 = A$$

$$e^{tA} = I + (e^t - 1)A = \begin{pmatrix} e^t & e^t - 1 \\ 0 & 1 \end{pmatrix}$$

5.5.4 分块公式

公式：若 $A = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix}$ ，则 $A^k = \begin{pmatrix} B^k & 0 \\ 0 & D^k \end{pmatrix}$ ， $f(x)$ 为解析函数

$$f(A) = \begin{pmatrix} f(B) & 0 \\ 0 & f(D) \end{pmatrix}$$

$$\text{同样: } A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix} \rightarrow f(A) = \begin{pmatrix} f(A_1) & 0 & 0 \\ 0 & f(A_2) & 0 \\ 0 & 0 & f(A_3) \end{pmatrix}$$

$$\text{特别: } A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \rightarrow f(A) = \begin{pmatrix} f(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(\lambda_n) \end{pmatrix}$$

$$f(x) = e^x \rightarrow e^A = \begin{pmatrix} e^{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{\lambda_n} \end{pmatrix}$$

$$\text{Eg : } A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \text{求 } e^{A+B}$$

$$A^2 = A, B^2 = B$$

$$e^A = \begin{pmatrix} e & e-1 \\ 0 & 1 \end{pmatrix} \quad e^B = \begin{pmatrix} e & e-2 \\ 0 & 1 \end{pmatrix}$$

$$A+B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \text{ (对角阵)}$$

$$e^{A+B} = \begin{pmatrix} e^2 & 0 \\ 0 & 1 \end{pmatrix}$$

$$e^{A+B} \neq e^A e^B$$

定理: 若 $AB = BA$ (可交换), 则 $e^A e^B = e^{A+B} = e^{B+A} = e^B e^A$

注: 若 $AB = BA$, $(A+B)^2 = A^2 + 2AB + B^2$

$$(A+B)^p = A^p + pA^{p-1}B + \cdots + pAB^{p-1} + B^p \text{ (二项公式)}$$

由上述定理可知:

$$(-A)A = A(-A) \rightarrow e^A e^{-A} = e^{A+A} = e^0 = I$$

$$\text{同理 } e^{tA} e^{-tA} = I$$

可逆公式: e^A 必有逆, 且 $(e^A)^{-1} = e^{-A}$. $(e^{tA})^{-1} = e^{-tA}$

Eg: $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, 用定义求 e^{tA} , e^{tB}

$$A^1 = A, A^2 = -I, A^3 = -A, A^4 = I, A^5 = A$$

$$\begin{aligned} e^{tA} &= I + tA + \frac{t^2(-I)}{2!} + \frac{t^3(-A)}{3!} + \frac{t^4(I)}{4!} + \frac{t^5(A)}{5!} + \dots = \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} \dots\right) A + \\ &\left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} \dots\right) I = (\sin t)A + (\cos t)I = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \\ e^{tB} &= e^{-tA} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \end{aligned}$$

5.5.5 根遗传公式

公式: 若 $\lambda(A) = \{\lambda_1, \dots, \lambda_n\}$, 则 $f(A) = \sum c_k A_k$ 的全体根 $\lambda(f(A)) = \{f(\lambda_1), \dots, f(\lambda_n)\}$

行列式: $|f(A)| = \det[f(A)] = f(\lambda_1) * \dots * f(\lambda_n)$

Eg: $f(x) = e^x$, $\lambda(e^A) = \{e^{\lambda_1}, \dots, e^{\lambda_n}\}$

$$|e^A| = e^{\lambda_1} * \dots * e^{\lambda_n} = e^{\lambda_1 + \dots + \lambda_n} = e^{a_{11} + \dots + a_{nn}}$$

5.5.6 Euler 公式

$$1) e^{iA} = \cos A + i \sin A$$

$$2) e^{-iA} = \cos A - i \sin A \quad (\text{将(1)中 } A \text{ 换为 } -A)$$

$$3) \cos A = \frac{1}{2}(e^{iA} + e^{-iA})$$

$$4) \sin A = \frac{1}{2i}(e^{iA} - e^{-iA})$$

A 为常数(t)也成立

Eg1: $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, 求 e^{tA} , $(e^{tA})^{-1}$, $|e^{tA}|$

解:

$$|e^{tA}| = e^{tr(tA)} = e^{t+2t+t} = e^{4t}$$

$$\lambda(A) = \{2, 1, 1\} \text{ (平移法)}$$

$$(A - 1)(A - 2) = 0 \rightarrow A \text{ 为单阵} \rightarrow f(A) = f(\lambda_1)G_1 + f(\lambda_2)G_2$$

$$G_1 = \frac{A - \lambda_2}{\lambda_1 - \lambda_2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, G_2 = I - G_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

$$e^{tA} = e^{2t}G_1 + e^tG_2$$

$$(e^{tA})^{-1} = e^{-tA} = e^{-2t}G_1 + e^{-t}G_2$$

5.5.7 幂 0 公式

幂 0 阵: $A^k = 0 (k \geq 2)$ 且 $A^{k-1} \neq 0$

幂 0 公式: 若 $(A - a)^k = 0$, 则 $f(A) = f(a)I + f'(a)(A - a) + \frac{f''(a)}{2!}(A - a)^2 + \dots + \frac{f^{k-1}(a)}{(k-1)!}(A - a)^{k-1}$

Eg: $A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \end{pmatrix}$, 求 $e^{tA}, (e^{tA})^{-1}$

$$(A - 2)^2 = 0$$

$$f(A) = f(2)I + f'(2)(A - 2)$$

$$e^{tA} = e^{2t}I + te^{2t}(A - 2)$$

$$(e^{tA})^{-1} = e^{-tA}$$

5.5.8 矩阵函数求法总结

1. 幂等阵: $e^{tA} = I + (e^t - 1)A$

2. 分块阵: $A = \begin{pmatrix} B & 0 \\ 0 & D \end{pmatrix} \rightarrow f(A) = \begin{pmatrix} f(B) & 0 \\ 0 & f(D) \end{pmatrix}$

$$A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \rightarrow f(A) = \begin{pmatrix} f(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(\lambda_n) \end{pmatrix}$$

3. 单纯阵: $(A - \lambda_1) * \dots * (A - \lambda_k) = 0$

$$f(A) = f(\lambda_1)G_1 + \cdots + f(\lambda_k)G_k$$

4. 幂 0 阵 : $f(A) = f(a)I + f'(a)(A - a) + \frac{f''(a)}{2!}(A - a)^2 + \cdots + \frac{f^{k-1}(a)}{(k-1)!}(A - a)^{k-1}$

Eg: 已知矩阵 A 的最小式为 $(x - 2)^2(x - 1)$, 则可知 A 的广谱公式为:

$f(A) = f(1)P_1 + f(2)P_2 + f'(2)P$, 通过构造法求 P_1 、 P_2 、 P , 并计算 $\cos(2\pi A)$ 。

解:

$$\text{令 } f(x) = (x - 2)^2, f(1) = 1, f(2) = 0, f'(2) = 0$$

$$\text{所以 } P_1 = (A - 2I)^2$$

$$\text{令 } f(x) = \text{常数 } k, f(1) = f(2) = k, f'(2) = 0, k = kP_1 + kP_2$$

$$\text{所以 } I = P_1 + P_2, P_2 = -(A - I)(A - 3I)$$

$$\text{令 } f(x) = (x - 1)(x - 2), f(1) = f(2) = 0, f'(2) = 1$$

$$\text{所以 } P = (A - 1)(A - 2)$$

$$\cos(2\pi A) = \cos(2\pi)P_1 + \cos(4\pi)P_2 - 2\pi \sin(4\pi)P = P_1 + P_2 = I$$

5.6 矩阵函数应用

5.6.1 基本定义与公式

定义：

令 $A(t) = \left(a_{ij}(t) \right)_{m \times n}$ ，其中 $a_{ij}(t)$ 是关于 t 的函数。

求导： $\frac{dA(t)}{dt} = A'(t) = \frac{da_{ij}(t)}{dt} = a'_{ij}(t)$

积分： $\int_a^b A(t)dt = \left(\int_a^b a_{ij}(t)dt \right)$

常数矩阵导数 $(A)' = 0$

公式：

$$[A(t)B(t)]' = A'(t)B(t) + A(t)B'(t)$$

$$\frac{de^{tA}}{dt} = Ae^{tA}, \quad \frac{de^{-tA}}{dt} = -Ae^{-tA}$$

$$[e^{-tA}B(t)]' = -Ae^{-tA}B(t) + e^{-tA}B'(t) = e^{-tA}[B'(t) - AB(t)]$$

$$\int_a^b A(t)dt = A(b) - A(a)$$

$$\int_a^b (A(t) + B(t))dt = \int_a^b A(t)dt + \int_a^b B(t)dt$$

$$\int_a^b kA(t)dt = k \int_a^b A(t)dt$$

5.6.2 应用

Eg1: 已知 $e^{tA} = \frac{1}{3} \begin{pmatrix} e^{5t} + 2e^{2t} & 2e^{5t} - 2e^{2t} \\ e^{5t} - e^{2t} & 2e^{5t} + e^{2t} \end{pmatrix}$ ，用导数求矩阵 A

解：左边开导 $= Ae^{tA}$

$$\text{右边开导} = \frac{1}{3} \begin{pmatrix} 5e^{5t} + 4e^{2t} & 10e^{5t} - 4e^{2t} \\ 5e^{5t} - 2e^{2t} & 10e^{5t} + 2e^{2t} \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} e^{tA}$$

$$\text{所以 } A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$$

Eg2：已知 $\sin(At) = B(t)$ ，表示出矩阵 A

开导: $A \sin(At) = B'(t)$

令 $t=0$ 得: $A = B'(0)$

5.6.3 求解齐次线性微分方程组

$\frac{dX(t)}{dt} = AX(t)$ 有通解公式: $X(t) = e^{tA} * C$, C 为常数矩阵。

Pf:

$$\text{原式} \leftrightarrow e^{-tA} * \frac{dX(t)}{dt} = e^{-tA} * AX(t)$$

$$\leftrightarrow e^{-tA} \frac{dX(t)}{dt} - Ae^{-tA}X(t) = 0 \quad \text{为什么可以调换??}$$

$$\leftrightarrow [e^{-tA}X(t)]' = 0$$

$$\leftrightarrow e^{-tA}X(t) = C (\text{常数阵})$$

$$\leftrightarrow X(t) = e^{tA} * C$$

定理: $\begin{cases} \frac{dX(t)}{dt} = AX(t) \\ X(0) = C_0 \end{cases}$ 有唯一解 $X(t) = e^{tA} * C_0$

Eg: 求解 $\begin{cases} x'_1(t) + x_2(t) = 0 \\ x'_2(t) - x_1(t) = 0 \end{cases}$, 且 $X(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

解: 令 $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $X = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$

$$\frac{dX(t)}{dt} = \begin{pmatrix} x'_1(t) \\ x'_2(t) \end{pmatrix}, AX(t) = \begin{pmatrix} -x_2(t) \\ x_1(t) \end{pmatrix}$$

所以原式可写为: $\frac{dX(t)}{dt} = AX(t)$

所以唯一解为 $X(t) = e^{tA} * X(0) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos t - \sin t \\ \sin t + \cos t \end{pmatrix}$

e^{tA} 的值在之前的例题中求得

6 矩阵直积

6.1 定义与性质

6.1.1 定义

令 $A = (a_{ij})_{m \times n}$, $B = (b_{ij})_{p \times q}$, 矩阵的张量积(直积) $A \otimes B$ 定义如下:

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \cdots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} = (a_{ij}B)$$

6.1.2 基本性质

基本性质:

- 1) $A \otimes B \neq B \otimes A$
- 2) $A \otimes (B + C) = A \otimes B + A \otimes C$
- 3) $(A + B) \otimes C = A \otimes C + B \otimes C$
- 4) $I_n \otimes I_p = I_{np} = I_p \otimes I_n$

分块公式(右进法则):

- 1) $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \otimes F = \begin{pmatrix} A \otimes F & B \otimes F \\ C \otimes F & D \otimes F \end{pmatrix}$
- 2) $\begin{pmatrix} A \\ B \end{pmatrix} \otimes F = \begin{pmatrix} A \otimes F \\ B \otimes F \end{pmatrix}$
- 3) $(A \ B \ C) \otimes F = (A \otimes F \ B \otimes F \ C \otimes F)$

但 $F \otimes (A \ B) \neq (F \otimes A \ F \otimes B)$

转置公式:

$$(A \otimes B)^T = A^T \otimes B^T, \text{ 比较 } (AB)^T = B^T A^T$$

$$\overline{A \otimes B} = \bar{A} \otimes \bar{B} \quad (A \otimes B)^H = A^H \otimes B^H$$

吸收公式:

$$(A \otimes B)(C \otimes D) = (AC \otimes BD)$$

条件: AC 、 BD 有定义

$$\text{推广: } (A_1 \otimes B_1) \cdots (A_k \otimes B_k) = (A_1 \cdots A_k \otimes B_1 \cdots B_k)$$

$$\text{特别: } (A \otimes B)^k = A^k \otimes B^k$$

逆公式:

$$\text{若 } A、B \text{ 都可逆, 则 } (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

$$\text{证明: } (A \otimes B) * (A^{-1} \otimes B^{-1}) = (AA^{-1} \otimes BB^{-1}) = I$$

$$\text{推广: } (A \otimes B)^+ = A^+ \otimes B^+$$

Eg :

- 1) 若 A 、 B 都是 Hermite 阵, 则 $A \otimes B$ 为 Hermite 阵
- 2) 若 A 、 B 都是幂等阵, 则 $A \otimes B$ 为幂等阵
- 3) 若 A 、 B 都是酉阵, 则 $A \otimes B$ 为酉阵

$$\text{解: 若 } A = A^H, B = B^H, \text{ 则 } (A \otimes B)^H = A^H \otimes B^H = A \otimes B$$

$$\text{若 } A^k = A, B^k = B, \text{ 则 } (A \otimes B)^k = A^k \otimes B^k = A \otimes B$$

$$\text{若 } A^H = A^{-1}, B^H = B^{-1}, \text{ 则 } (A \otimes B)^H = A^H \otimes B^H = A^{-1} \otimes B^{-1} = (A \otimes B)^{-1}$$

6.1.3 扩展性质

上三角性质:

若 D_1 、 D_2 都是上三角阵, 则 $D_1 \otimes D_2$ 是上三角阵

D_1 对角元为 $\{a_1, \cdots, a_n\}$, D_2 对角元为 $\{b_1, \cdots, b_n\}$, 则 $D_1 \otimes D_2$ 对角元为

$$\{a_1 b_1, \dots, a_1 b_n, \dots, a_n b_1, \dots, a_n b_n\}$$

相似性质:

若 $A \sim D_1, B \sim D_2$, 则 $A \otimes B \sim D_1 \otimes D_2$

证明: $P^{-1}AP = D_1 \quad Q^{-1}BQ = D_2$

$$(P \otimes Q)^{-1}(A \otimes B)(P \otimes Q) = (P^{-1} \otimes Q^{-1})(A \otimes B)(P \otimes Q) = \\ (P^{-1}AP) \otimes (Q^{-1}BQ) = D_1 \otimes D_2$$

特征根定理:

$\lambda(A) = \{\lambda_1, \dots, \lambda_n\}, \lambda(B) = \{t_1, \dots, t_n\}$, 则 $\lambda(A \otimes B) = \{\text{全体 } \lambda_i t_j\}$

Eg: 证明: $(A \quad 0)^+ = \begin{pmatrix} A^+ \\ 0 \end{pmatrix}$

$$(A \quad 0) = (1 \quad 0) \otimes A$$

$$(A \quad 0)^+ = ((1 \quad 0) \otimes A)^+ = (1 \quad 0)^+ \otimes A^+ = \begin{pmatrix} A^+ \\ 0 \end{pmatrix}$$

同样: $(A \quad A \quad A)^+ = \frac{1}{3} \begin{pmatrix} A^+ \\ A^+ \\ A^+ \end{pmatrix}$

$$\begin{pmatrix} A \\ A \\ A \end{pmatrix}^+ = \frac{1}{3} (A^+ \quad A^+ \quad A^+)$$

Eg: $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ -1 & -1 & -1 & -1 \\ 0 & -1 & 0 & -1 \end{pmatrix}$, 求 A^+

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

$$A^+ = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}^+ \otimes \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^+ = \frac{1}{4} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \otimes \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$