第五章　　Hermite矩阵与正定矩阵

5.1　Hermite矩阵与Hermite二次型

5.1.1　Hermite矩阵

- 例

$$\begin{bmatrix} 2 & 3-2\mathrm{i} \\ 3+2\mathrm{i} & -5 \end{bmatrix}, \quad \begin{bmatrix} -1 & 7 \\ 7 & 4 \end{bmatrix}$$

- **Hermite矩阵的性质**

(1) 如果 $\boldsymbol{A}$ 是Hermite矩阵，则对正整数 k，$\boldsymbol{A}^k$ 也是Hermite矩阵；

(2) 如果 $\boldsymbol{A}$ 是可逆Hermite矩阵，则 $\boldsymbol{A}^{-1}$ 也是Hermite矩阵；

(3) 如果 $\boldsymbol{A}$，$\boldsymbol{B}$ 是Hermite矩阵，则对实数 k，p，$k\boldsymbol{A}+p\boldsymbol{B}$ 也是Hermite矩阵；

(4) 若 $\boldsymbol{A}$，$\boldsymbol{B}$ 是Hermite矩阵，则 $\boldsymbol{AB}$ 是Hermite矩阵的充分必要条件是 $\boldsymbol{AB}=\boldsymbol{BA}$；

(5) $\boldsymbol{A}$ 是Hermite矩阵的充分必要条件是对任意方阵 $\boldsymbol{S}$，$\boldsymbol{S}^{\mathrm{H}}\boldsymbol{AS}$ 是Hermite矩阵。

- **定理 5.1.1** 设 $\boldsymbol{A} = (a_{jk}) \in \mathbb{C}^{n\times n}$，则 $\boldsymbol{A}$ 是Hermite矩阵的充分必要条件是对任意 $x \in \mathbb{C}^n$，$x^{\mathrm{H}}\boldsymbol{A}x$ 是实数。

- **定理 5.1.2** 设 $\boldsymbol{A}$ 为 n 阶Hermite矩阵，则
(1) $\boldsymbol{A}$ 的所有特征值全是实数；
(2) $\boldsymbol{A}$ 的不同特征值所对应的特征向量是互相正交的。
证明(2) 设 λ，μ 是 $\boldsymbol{A}$ 的两个不同特征值，相应的特征向量分别为 x，y，则

$$\boldsymbol{A}x = \lambda x \tag{1}$$

$$\boldsymbol{A}y = \mu y \tag{2}$$

由(1)

$$y^{\mathrm{H}}\boldsymbol{A}x = \lambda y^{\mathrm{H}}x$$

由(2),$\mu = \mu^{\mathrm{H}}$和$\boldsymbol{A} = \boldsymbol{A}^{\mathrm{H}}$

$$y^{\mathrm{H}}\boldsymbol{A}x = y^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}x = \mu y^{\mathrm{H}}x$$

于是 $(\lambda - \mu)y^{\mathrm{H}}x = 0$。由于 $\lambda \neq \mu$，故 x 与 y 正交。 $\square$

- **定理 5.1.3** 设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，则 $\boldsymbol{A}$ 是Hermite矩阵的充分必要条件是存在酉矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U} = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

  其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 均为实数。

- **定理 5.1.4** 设 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$，则 $\boldsymbol{A}$ 是实对称矩阵的充分必要条件是存在正交矩阵 $\boldsymbol{Q}$ 使得

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

  其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 均为实数。

5.1.2 矩阵的惯性

- 定义1.6.8　设$\boldsymbol{A},\boldsymbol{B} \in \mathbb{C}^{n\times n}$，如果存在$n$阶非奇异矩阵$\boldsymbol{P}$，使得

$$\boldsymbol{B} = \boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}$$

则称$\boldsymbol{A}$与$\boldsymbol{B}$是相合的。

- **定理 5.1.5**　设 $\boldsymbol{A}$ 是 n 阶Hermite矩阵，则 $\boldsymbol{A}$ 相合于矩阵

$$\boldsymbol{D}_0 = \begin{bmatrix} \boldsymbol{I}_s & 0 & 0 \\ 0 & -\boldsymbol{I}_{r-s} & 0 \\ 0 & 0 & \boldsymbol{0}_{n-r} \end{bmatrix}$$

其中 r =rank($\boldsymbol{A}$)，s 是 $\boldsymbol{A}$ 的正特征值(重特征值按重数计算)的个数。
证明　由定理5.1.3知，存在 n 阶酉矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{A} = \boldsymbol{U}\Lambda\boldsymbol{U}^{\mathrm{H}}$$

其中 Λ 是以 $\boldsymbol{A}$ 的特征值为对角元的对角矩阵。排列 $\boldsymbol{A}$ 的特征值使得

$$\boldsymbol{P}_1^{\mathrm{H}}\Lambda\boldsymbol{P}_1 = \mathrm{diag}(\lambda_1,\cdots,\lambda_s,\lambda_{s+1},\cdots,\lambda_r,0,\cdots,0) = \boldsymbol{D}$$

其中 $\boldsymbol{P}_1$ 是排列矩阵，$\lambda_i > 0(i = 1,\cdots,s), \lambda_j < 0(j = s+1,\cdots,r)$，令

$$\boldsymbol{D}_1 = \mathrm{diag}(\sqrt{\lambda_1},\cdots,\sqrt{\lambda_s},\sqrt{|\lambda_{s+1}|},\cdots,\sqrt{|\lambda_r|},1,\cdots,1)$$

显然，$\left(\boldsymbol{D}_1^{-1}\right)^{\mathrm{H}}\boldsymbol{D}\boldsymbol{D}_1^{-1} = \boldsymbol{D}_0$，于是$\left(\boldsymbol{D}_1^{-1}\right)^{\mathrm{H}}\boldsymbol{P}_1^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}\boldsymbol{P}_1\boldsymbol{D}_1^{-1} = \boldsymbol{D}_0$ 即 $\boldsymbol{A}$ 相合于 $\boldsymbol{D}_0$

- **定义**　式(5.1.3)中矩阵 $\boldsymbol{D}_0$ 称为 n 阶**Hermite矩阵 A 的相合标准形**。

- **定义**　设 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，$\pi(\boldsymbol{A}), \upsilon(\boldsymbol{A})$ 和 $\delta(\boldsymbol{A})$ 分别表示 $\boldsymbol{A}$ 的位于右半开平面、左半开平面和虚轴上的特征值(重特征值按重数计算)的个数。记

$$\mathrm{In}(\boldsymbol{A}) = \{\pi(\boldsymbol{A}), \upsilon(\boldsymbol{A}), \delta(\boldsymbol{A})\}$$

  则称In(A)为**矩阵 A 的惯性**。

- $\boldsymbol{A}$ 是 n 阶Hermite矩阵，则 $\pi(\boldsymbol{A}), \upsilon(\boldsymbol{A})$ 和 $\delta(\boldsymbol{A})$ 分别表示 $\boldsymbol{A}$ 的正、负和零特征值的个数(重特征值按重数计算)。因此 $\boldsymbol{A}$ 非奇异的充分必要条件为 $\delta(\boldsymbol{A}) = 0$，并且

$$\pi(\boldsymbol{A}) + \upsilon(\boldsymbol{A}) = \mathrm{rank}(\boldsymbol{A}) \tag{5.1.5}$$

- 定理5.1.6(惯性定理)　设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶Hermite矩阵，则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相合的充分必要条件是$\mathrm{In}(\boldsymbol{A}) = \mathrm{In}(\boldsymbol{B})$

5.1.3 Hermite二次型

- **定义** 由 n 个复变量 $x_1, \cdots, x_n \in \mathbb{C}$ 构成，系数为复数的二次齐式

$$f(x_1, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \overline{x}_i x_j$$

其中 $a_{ij} = \overline{a}_{ji}$，称为**Hermite二次型**。记

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

则 $\boldsymbol{A}$ 为**Hermite**矩阵。称 $\boldsymbol{A}$ 为**Hermite二次型的矩阵**，并且称 $\boldsymbol{A}$ 的秩为Hermite二次**型的秩**。于是，Hermite二次型可改写为

$$f(x) = x^{\mathrm{H}} \boldsymbol{A} x$$

其中 $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$。因此，一个Hermite二次型与一个Hermite矩阵相对应。

- 若作可逆线性变换 $x = \boldsymbol{P}y$，其中 $\boldsymbol{P}$ 为 n 阶可逆矩阵，$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$，则

$$f(x) = x^{\mathrm{H}}\boldsymbol{A}x = y^{\mathrm{H}}\boldsymbol{B}y$$

其中 $\boldsymbol{B} = \boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}$。显然 $\boldsymbol{B}$ 是Hermite矩阵，且与矩阵 $\boldsymbol{A}$ 相合。

- **定义**　Hermite二次型中最简单的一种是只包含平方项的二次型

$$\lambda_1\overline{y}_1y_1 + \lambda_2\overline{y}_2y_2 + \cdots + \lambda_n\overline{y}_ny_n$$

称为**Hermite二次型的标准形**。

- **定理 5.1.7**　对Hermite二次型 $f(x) = x^{\mathrm{H}}\boldsymbol{A}x$，存在酉线性变换 $x = \boldsymbol{U}y$ (其中 U 是酉矩阵)使得Hermite二次型 $f(x)$ 变成标准形

$$\lambda_1\overline{y}_1y_1 + \lambda_2\overline{y}_2y_2 + \cdots + \lambda_n\overline{y}_ny_n$$

其中 $\lambda_1, \cdots, \lambda_n$ 是Hermite矩阵 $\boldsymbol{A}$ 的特征值。

- **定理 5.1.8** 对Hermite二次型 $f(x)=x^{\mathrm{H}}\boldsymbol{A}x$，存在可逆线性变换 $x=\boldsymbol{P}y$ 使得Hermite二次型 $f(x)$ 化为

$$f(x)=x^{\mathrm{H}}\boldsymbol{A}x=\overline{y}_1y_1+\cdots+\overline{y}_sy_s-\overline{y}_{s+1}y_{s+1}-\cdots-\overline{y}_ry_r \tag{5.1.13}$$

其中 $r=\mathrm{rank}(\boldsymbol{A}), s=\pi(\boldsymbol{A})$。
式(5.1.13)称为Hermite二次型 $f(x)=x^{\mathrm{H}}\boldsymbol{A}x$ **的规范形**，其中 s 和 $r-s$ 分别称为Hermite二次型的**正惯性指数**和**负惯性指数**。

- 例 $f(x)=x_1^2+6x_1x_2+x_2^2$
作线性变换 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$，则

$$\begin{aligned} f(x) &= (y_1+y_2)^2+6(y_1+y_2)(y_2-y_1)+(y_2-y_1)^2 \\ &= y_1^2+2y_1y_2+y_2^2+6y_2^2-6y_1^2+y_2^2-2y_1y_2+y_1^2 \\ &= -4y_1^2+8y_2^2=-z_1^2+z_2^2 \end{aligned}$$

其中 $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}=\begin{bmatrix} \frac{1}{2} & \\ & \frac{1}{\sqrt{8}} \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$

- Hermite二次型 $f(x)=x^{\mathrm{H}}\boldsymbol{A}x$ 的正惯性指数 s 与秩 r 之间满足 $0\le s\le r\le n$，所以Hermite二次型可分为五种情况：

(1)若 $s=r=n$，则规范形为 $x^{\mathrm{H}}\boldsymbol{A}x=\sum\limits_{i=1}^{n}|y_i|^2$。若 $x\ne 0$，则 $y\ne 0$，$x^{\mathrm{H}}\boldsymbol{A}x>0$。

(2)若 $s=r$，则规范形为 $x^{\mathrm{H}}\boldsymbol{A}x=\sum\limits_{i=1}^{r}|y_i|^2$。对任意 $x\in\mathbb{C}^n$，都有 $x^{\mathrm{H}}\boldsymbol{A}x\ge 0$

(3)若 $s=0,r=n$，则规范形为 $x^{\mathrm{H}}\boldsymbol{A}x=-\sum\limits_{i=1}^{n}|y_i|^2$。若 $x\ne 0$，则 $y\ne 0$，$x^{\mathrm{H}}\boldsymbol{A}x<0$。

(4)若 $s=0$，则规范形为 $x^{\mathrm{H}}\boldsymbol{A}x=-\sum\limits_{i=1}^{r}|y_i|^2$。对任意 $x\in\mathbb{C}^n$，都有 $x^{\mathrm{H}}\boldsymbol{A}x\le 0$。

(5)若 $0<s<r$，则规范形为 $x^{\mathrm{H}}\boldsymbol{A}x=\sum\limits_{i=1}^{s}|y_i|^2-\sum\limits_{i=s+1}^{r}|y_i|^2$。对不同的 x，$x^{\mathrm{H}}\boldsymbol{A}x$ 之值可以大于0，小于0或等于0。

- **定义 5.1.1** 设 $f(x) = x^{\mathrm{H}}\boldsymbol{A}x$ 为Hermite二次型。
  (1)如果对任意 $x \in \mathbb{C}^n$ 且 $x \neq 0$，都有 $x^{\mathrm{H}}\boldsymbol{A}x > 0$，则称 $x^{\mathrm{H}}\boldsymbol{A}x$ 为正定的；
  (2)如果对任意 $x \in \mathbb{C}^n$，都有 $x^{\mathrm{H}}\boldsymbol{A}x \geq 0$，则称 $x^{\mathrm{H}}\boldsymbol{A}x$ 为半正定的；
  (3)如果对任意 $x \in \mathbb{C}^n$ 且 $x \neq 0$，都有 $x^{\mathrm{H}}\boldsymbol{A}x < 0$，则称 $x^{\mathrm{H}}\boldsymbol{A}x$ 为负定的；
  (4)如果对任意 $x \in \mathbb{C}^n$，都有 $x^{\mathrm{H}}\boldsymbol{A}x \leq 0$，则称 $x^{\mathrm{H}}\boldsymbol{A}x$ 为半负定的；
  (5)对不同的 $x \in \mathbb{C}^n$，$x^{\mathrm{H}}\boldsymbol{A}x$有时为正，有时为负，则称 $x^{\mathrm{H}}\boldsymbol{A}x$ 为不定的。

- **定理 5.1.9** 对Hermite二次型 $f(x) = x^{\mathrm{H}}\boldsymbol{A}x$，有
  (1) $x^{\mathrm{H}}\boldsymbol{A}x$ 正定的充分必要条件为 $s = r = n$；
  (2) $x^{\mathrm{H}}\boldsymbol{A}x$ 半正定的充分必要条件为 $s = r$；
  (3) $x^{\mathrm{H}}\boldsymbol{A}x$ 负定的充分必要条件为 $s = 0, r = n$；
  (4) $x^{\mathrm{H}}\boldsymbol{A}x$ 半负定的充分必要条件为 $s = 0$；
  (5) $x^{\mathrm{H}}\boldsymbol{A}x$ 不定的充分必要条件为 $0 < s < r$。

5.2 正定(半正定)矩阵

- **定义 5.2.1** 设 $\boldsymbol{A}$ 是 n 阶Hermite矩阵，如果对任意 $x \in \mathbb{C}^n$ 且 $x \neq 0$，都有 $x^{\mathrm{H}}\boldsymbol{A}x > 0$，则称 $\boldsymbol{A}$ 为**正定矩阵**，记作 $\boldsymbol{A} > 0$；如果对任意 $x \in \mathbb{C}^n$，都有 $x^{\mathrm{H}}\boldsymbol{A}x \geq 0$，则称 $\boldsymbol{A}$ 为**半正定矩阵**，记作 $\boldsymbol{A} \geq 0$。
- 正定（半正定）矩阵具有如下基本性质：
  (1) 单位矩阵 $\boldsymbol{I} > 0$；
  (2) 若 $\boldsymbol{A} > 0$，数 $k > 0$，则 $k\boldsymbol{A} > 0$；
  (3) 若 $\boldsymbol{A} > 0$，$\boldsymbol{B} > 0$，则 $\boldsymbol{A} + \boldsymbol{B} > 0$；
  (4) 若 $\boldsymbol{A} \geq 0$，$\boldsymbol{B} \geq 0$，则 $\boldsymbol{A} + \boldsymbol{B} \geq 0$。

- **定理 5.2.1** 设 $\boldsymbol{A}$ 是 n 阶Hermite矩阵，则下列命题等价：
(1) $\boldsymbol{A}$ 是正定矩阵；
(2) 对任意 n 阶可逆矩阵 $\boldsymbol{P}$，$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}$ 都是正定矩阵；
(3) $\boldsymbol{A}$ 的 n 个特征值均为正数；
(4) 存在 n 阶可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{I}$；
(5) 存在 n 阶可逆矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{A}=\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q}$；
(6) 存在 n 阶可逆Hermite矩阵 $\boldsymbol{S}$ 使得 $\boldsymbol{A}=\boldsymbol{S}^2$。
证明
$(1)\Rightarrow(6)$ 存在酉矩阵$\boldsymbol{U}$, 使得

$$\boldsymbol{A}=\boldsymbol{U}\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)\boldsymbol{U}^{\mathrm{H}}$$

其中，$\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $\boldsymbol{A}$ 的正特征值。令

$$\boldsymbol{S}=\boldsymbol{U}\mathrm{diag}(\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_n})\boldsymbol{U}^{\mathrm{H}}$$

则 $\boldsymbol{S}$ 是 n 阶可逆Hermite矩阵，并且 $\boldsymbol{A}=\boldsymbol{S}^2$。
$(6)\Rightarrow(1)$ 因为存在 n 阶可逆Hermite矩阵 $\boldsymbol{S}$ 使得 $\boldsymbol{A}=\boldsymbol{S}^2=\boldsymbol{S}^{\mathrm{H}}\boldsymbol{S}$，由(5)，即知 $\boldsymbol{A}$ 是正定矩阵。 $\square$

- **推论 5.2.1** 设 $\boldsymbol{A}$ 是 n 阶正定矩阵，其特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则
  (1) $\boldsymbol{A}^{-1}$ 是正定矩阵；
  (2)如果 $\boldsymbol{Q}$ 是任一 $n \times m$ 列满秩矩阵，则 $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q} > 0$；
  (3) $|\boldsymbol{A}| > 0$；
  (4) $\mathrm{tr}(\boldsymbol{A}) > \lambda_i (i = 1, 2, \cdots, n)$。
- **例 5.2.1** 设 n 阶Hermite矩阵 $\boldsymbol{A}$ 有如下分块

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^{\mathrm{H}} & \boldsymbol{A}_{22} \end{bmatrix}, \quad \boldsymbol{A}_{11} \in \mathbb{C}^{k \times k}$$

则 $\boldsymbol{A} > 0$ 的充分必要条件是 $\boldsymbol{A}_{11} > 0$ 和 $\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} > 0$。
证明 如果 $\boldsymbol{A}_{11}$ 非奇异，则

$$\begin{bmatrix} \boldsymbol{I}_k & 0 \\ -\boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}_{11}^{-1} & \boldsymbol{I}_{n-k} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^{\mathrm{H}} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_k & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \\ 0 & \boldsymbol{I}_{n-k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11} & 0 \\ 0 & \boldsymbol{A}_{22} - \boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \end{bmatrix}$$

即 $\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{12}^{\mathrm{H}} & \boldsymbol{A}_{22} \end{bmatrix}$ 相合于 $\begin{bmatrix} \boldsymbol{A}_{11} & 0 \\ 0 & \boldsymbol{A}_{22} - \boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \end{bmatrix}$。因此 $\boldsymbol{A} > 0$ 的充分必要条件是 $\boldsymbol{A}_{11} > 0$ 和 $\boldsymbol{A}_{22} - \boldsymbol{A}_{12}^{\mathrm{H}}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} > 0$。 □

- **定理 5.2.2** 设 $\boldsymbol{A}$ 是 n 阶Hermite矩阵，则下列命题等价：
  (1) $\boldsymbol{A}$ 是半正定矩阵；
  (2) 对任意 n 阶可逆矩阵 $\boldsymbol{P}$，$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}$ 是半正定矩阵；
  (3) $\boldsymbol{A}$ 的 n 个特征值均为非负数；
  (4) 存在 n 阶可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} \boldsymbol{I}_r & 0 \\ 0 & 0 \end{bmatrix}$，其中$r = \mathrm{rank}(\boldsymbol{A})$；
  (5) 存在秩为 r 的矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{A} = \boldsymbol{Q}^{\mathrm{H}}\boldsymbol{Q}$；
  (6) 存在 n 阶Hermite矩阵 $\boldsymbol{S}$ 使得 $\boldsymbol{A} = \boldsymbol{S}^2$。

- **推论 5.2.2** 设 $\boldsymbol{A}$ 是 n 阶半正定矩阵，其特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，则
  (1)如果 $\boldsymbol{Q}$ 是任一 $n \times m$ 矩阵，则 $\boldsymbol{Q}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{Q} \geq 0$；
  (2) $|\boldsymbol{A}| \geq 0$；
  (3) $\mathrm{tr}(\boldsymbol{A}) \geq \lambda_i (i = 1, 2, \cdots, n)$。

- 定理5.2.3 n 阶Hermite矩阵 $\boldsymbol{A}$ 正定的充分必要条件是 $\boldsymbol{A}$ 的顺序主子式均为正数，即

$$\Delta_k > 0, \quad k = 1, \cdots, n$$

- 定理5.2.4 n Hermite矩阵 $\boldsymbol{A}$ 正定的充分必要条件是 $\boldsymbol{A}$ 的所有主子式全大于零.

- 定理5.2.5 n 阶Hermite矩阵 $\boldsymbol{A}$ 半正定的充分必要条件是 $\boldsymbol{A}$ 的所有主子式均非负.

- 定理5.2.6　Hermite矩阵 $\boldsymbol{A}$ 正定的充分必要条件是存在非奇异下三角矩阵 $\boldsymbol{L}$ 使得

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{H}} \tag{5.2.3}$$

分解式(5.2.3)称为正定矩阵$\boldsymbol{A}$的Cholesky分解

证明　必要性.

若 $\boldsymbol{A}$ 是 n 阶正定矩阵， $\boldsymbol{A}$ 的顺序主子式全大于零. $\boldsymbol{A}$ 有唯一的LDU分解

$$\boldsymbol{A}=\boldsymbol{L}_1\boldsymbol{D}\boldsymbol{U}_1$$

其中 $\boldsymbol{L}_1,\boldsymbol{U}_1$ 分别为单位下三角矩阵和单位上三角矩阵，$\boldsymbol{D}=\mathrm{diag}(d_1,d_2,\cdots,d_n)$ 且 $d_i>0(i=1,2,\cdots,n)$.因为 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$,则

$$\boldsymbol{A}=\boldsymbol{L}_1\boldsymbol{D}\boldsymbol{U}_1=\boldsymbol{U}_1^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{L}_1^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}$$

由LDU分解的唯一性，有 $\boldsymbol{L}_1=\boldsymbol{U}_1^{\mathrm{H}}$.从而有$\boldsymbol{A}=\boldsymbol{L}_1\boldsymbol{D}\boldsymbol{L}_1^{\mathrm{H}}$，令

$$\boldsymbol{L}=\boldsymbol{L}_1\mathrm{diag}(\sqrt{d_1},\cdots,\sqrt{d_n})$$

则 $\boldsymbol{L}$ 是非奇异下三角矩阵，并且 $\boldsymbol{A}=\boldsymbol{L}\boldsymbol{L}^{\mathrm{H}}$.

- 定义5.2.2　设 $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$,如果存在复数 λ 和非零向量 $x \in \mathbb{C}^n$ 使得

$$\boldsymbol{A}x = \lambda \boldsymbol{B}x \tag{3}$$

  则称 λ 为广义特征值问题 $\boldsymbol{A}x = \lambda \boldsymbol{B}x$ 的特征值，非零向量 x 称为对应于特征值 λ 的特征向量.

- 如果 $\boldsymbol{B}$ 是 n 阶非奇异矩阵，则广义特征值问题可化为如下标准特征值问题

$$\boldsymbol{B}^{-1}\boldsymbol{A}x = \lambda x \tag{5.2.6}$$

- 定理5.2.7　设 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶Hermite矩阵，且 $\boldsymbol{B}>0$,则存在非奇异矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n),\boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}=\boldsymbol{I}$$

其中 $\lambda_1,\cdots,\lambda_n$ 是广义特征值问题的特征值.
证明　因为 $\boldsymbol{B}>0$，存在非奇异矩阵 $\boldsymbol{P}_1$ 使得 $\boldsymbol{P}_1^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}_1=\boldsymbol{I}$,而 $\boldsymbol{P}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}_1$ 仍为Hermite矩阵，则有酉矩阵 $\boldsymbol{U}$ 使

$$\boldsymbol{U}^{\mathrm{H}}(\boldsymbol{P}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}_1)\boldsymbol{U}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$$

令 $\boldsymbol{P}=\boldsymbol{P}_1\boldsymbol{U}$,则 $\boldsymbol{P}$ 非奇异，并且

$$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{P}_1^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}_1\boldsymbol{U}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}=\boldsymbol{I},\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{P}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}_1\boldsymbol{U}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$$

由上式得

$$\boldsymbol{P}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$$

即 $\boldsymbol{B}^{-1}\boldsymbol{A}$ 相似于 $\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$, 则 $\lambda_1,\cdots,\lambda_n$ 是矩阵 $\boldsymbol{B}^{-1}\boldsymbol{A}$ 的特征值，即是广义特征值问题的特征值.

5.3 矩阵不等式

- 定义5.3.1 设 $\boldsymbol{A},\boldsymbol{B}$ 都是 n 阶Hermite矩阵，如果 $\boldsymbol{A}-\boldsymbol{B}\geq 0$,则称 $\boldsymbol{A}$ 大于或等于 $\boldsymbol{B}$ (或称 $\boldsymbol{B}$ 小于或等于 $\boldsymbol{A}$), 记作 $\boldsymbol{A}\geq\boldsymbol{B}$ (或 $\boldsymbol{B}\leq\boldsymbol{A}$) ; 如果 $\boldsymbol{A}-\boldsymbol{B}>0$,则称 $\boldsymbol{A}$ 大于 $\boldsymbol{B}$ (或称 $\boldsymbol{B}$ 小于 $\boldsymbol{A}$),记作 $\boldsymbol{A}>\boldsymbol{B}$ (或 $\boldsymbol{B}<\boldsymbol{A}$).
- 任意两个实数总可以比较大小.但是任意两个 n 阶Hermite矩阵未必能”比较大小”,即并非 $\boldsymbol{A}\geq\boldsymbol{B}$ 或 $\boldsymbol{B}\geq\boldsymbol{A}$ 两者之中必有一个成立.例如对

$$\boldsymbol{A}=\begin{bmatrix}4&0\\0&5\end{bmatrix},\qquad \boldsymbol{B}=\begin{bmatrix}5&0\\0&4\end{bmatrix}$$

  $\boldsymbol{A}\geq\boldsymbol{B}$ 和 $\boldsymbol{B}\geq\boldsymbol{A}$ 均不成立.

- 对任意两个实数 a 和 b ,如果 $a\geq b$,而 $a\not>b$,则有 $a=b$. 但是对两个 $n(n\geq 2)$ 阶Hermite矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ ， 从 $\boldsymbol{A}\geq\boldsymbol{B}$ 和 $\boldsymbol{A}\not>\boldsymbol{B}$, 不能推出 $\boldsymbol{A}=\boldsymbol{B}$.例如，对

$$\boldsymbol{A}=\begin{bmatrix}1&0\\0&2\end{bmatrix},\quad ,\boldsymbol{B}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

  可见 $\boldsymbol{A}\geq\boldsymbol{B}$,且 $\boldsymbol{A}\not>\boldsymbol{B}$,但 $\boldsymbol{A}\neq\boldsymbol{B}$.

- 定理5.3.1　设 $\boldsymbol{A},\boldsymbol{A}_1,\boldsymbol{B},\boldsymbol{B}_1,\boldsymbol{C}$ 均为 n 阶Hermite矩阵，则
  (1) $\boldsymbol{A} \geq \boldsymbol{B}(\boldsymbol{A} > \boldsymbol{B})$ 的充分必要条件是 $-\boldsymbol{A} \leq -\boldsymbol{B}(-\boldsymbol{A} < -\boldsymbol{B})$;
  (2) $\boldsymbol{A} \geq \boldsymbol{B}(\boldsymbol{A} > \boldsymbol{B})$ 的充分必要条件是对任意 n 阶可逆矩阵 $\boldsymbol{P}$ 都有

$$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P} \geq \boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}(\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P} > \boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P})$$

  (3)若 $\boldsymbol{A} \geq \boldsymbol{B}(\boldsymbol{A} > \boldsymbol{B})$， k 为正数，则 $k\boldsymbol{A} \geq k\boldsymbol{B}(k\boldsymbol{A} > k\boldsymbol{B})$;
  (4)若 $\boldsymbol{A} \geq 0, -\boldsymbol{A} \geq 0$,则 $\boldsymbol{A} = 0$;
  (5)若 $\boldsymbol{A} \geq 0$, $\boldsymbol{B} \geq 0$,则 $\boldsymbol{A} + \boldsymbol{B} \geq 0$;
  (6)若 $\boldsymbol{A} \geq \boldsymbol{B}, \boldsymbol{B} \geq \boldsymbol{C}$,则 $\boldsymbol{A} \geq \boldsymbol{C}$;
  (7)若 $\boldsymbol{A} \geq \boldsymbol{B}, \boldsymbol{A}_1 \geq \boldsymbol{B}_1$,则 $\boldsymbol{A} + \boldsymbol{A}_1 \geq \boldsymbol{B} + \boldsymbol{B}_1$;
  (8)若 $\boldsymbol{A} \geq 0, \boldsymbol{B} > 0$,则 $\boldsymbol{A} + \boldsymbol{B} > 0$.
  (9)若 $\boldsymbol{A} \geq \boldsymbol{B}, \boldsymbol{B} > \boldsymbol{C}$， 则 $\boldsymbol{A} > \boldsymbol{C}$;
  (10)若 $\boldsymbol{A} > \boldsymbol{B}, \boldsymbol{P}$ 为 $n \times m$ 列满秩矩阵,则 $\boldsymbol{P}^{\mathrm{H}}A\boldsymbol{P} > \boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}$;
  (11)若 $\boldsymbol{A} \geq \boldsymbol{B}, \boldsymbol{P}$ 为 $n \times m$ 矩阵,则 $\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P} \geq \boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}$;
  (12)若 $\boldsymbol{A} > 0(\boldsymbol{A} \geq 0), \boldsymbol{C} > 0(\boldsymbol{C} \geq 0)$,且 $\boldsymbol{AC} = \boldsymbol{CA}$,则 $\boldsymbol{AC} > 0(\boldsymbol{AC} \geq 0)$.

- 定理5.3.2　设 $\boldsymbol{A}, \boldsymbol{B}$ 均为 n 阶Hermite矩阵,且 $\boldsymbol{A} \geq 0, \boldsymbol{B} > 0$,则
  (1) $\boldsymbol{B} \geq \boldsymbol{A}$ 的充分必要条件是 $\rho(\boldsymbol{A}\boldsymbol{B}^{-1}) \leq 1$;
  (2) $\boldsymbol{B} > \boldsymbol{A}$ 的充分必要条件是 $\rho(\boldsymbol{A}\boldsymbol{B}^{-1}) < 1$.
  证明　(1)由定理5.2.7知，有非奇异矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{I}, \boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$$

$\boldsymbol{B} \geq \boldsymbol{A}$ 等价于 $\boldsymbol{I} \geq \boldsymbol{\Lambda}$,而后者等价于 $\lambda_i \leq 1(i = 1, \cdots, n)$. 因为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是矩阵 $\boldsymbol{A}\boldsymbol{B}^{-1}$ 的全部特征值，所以 $\lambda_i \leq 1(i = 1, \cdots, n)$ 等价于 $\rho(\boldsymbol{A}\boldsymbol{B}^{-1}) \leq 1$.
类似可证明(2).

- 定理5.3.3　设 $\boldsymbol{A}$ 是 n 阶Hermite矩阵，则 $\lambda_{\min}(A)\boldsymbol{I} \leq \boldsymbol{A} \leq \lambda_{\max}(\boldsymbol{A})\boldsymbol{I}$, 这时 $\lambda_{\max}(\boldsymbol{A})$ 和 $\lambda_{\min}(\boldsymbol{A})$ 分别表示 $\boldsymbol{A}$ 的最大和最小特征值.

  证明　存在酉矩阵 $\boldsymbol{U}$ 使得

$$\boldsymbol{A} = \boldsymbol{U}\mathrm{diag}(\lambda_1, \cdots, \lambda_n)\boldsymbol{U}^{\mathrm{H}}$$

  其中 $\lambda_1 \geq \cdots \geq \lambda_n$ 是矩阵 $\boldsymbol{A}$ 的特征值，则

$$\begin{aligned}
&\lambda_{\max}(\boldsymbol{A}) = \lambda_1 \\
&\lambda_{\max}(\boldsymbol{A}) - \lambda_i \geq 0 \quad (i = 1, 2, \cdots, n) \\
&\lambda_{\min}(\boldsymbol{A}) = \lambda_n \\
&\lambda_i - \lambda_{\min}(\boldsymbol{A}) \geq 0 \quad (i = 1, 2, \cdots, n)
\end{aligned}$$

  并且

$$\begin{aligned}
\lambda_{\max}(\boldsymbol{A})\boldsymbol{I} - \boldsymbol{A} &= \boldsymbol{U}[\lambda_{\max}(\boldsymbol{A})\boldsymbol{I} - \mathrm{diag}(\lambda_1, \cdots, \lambda_n)]\boldsymbol{U}^{\mathrm{H}} \geq 0 \\
\boldsymbol{A} - \lambda_{\min}(\boldsymbol{A})\boldsymbol{I} &= \boldsymbol{U}[\mathrm{diag}(\lambda_1, \cdots, \lambda_n) - \lambda_{\min}(\boldsymbol{A})\boldsymbol{I}]\boldsymbol{U}^{\mathrm{H}} \geq 0
\end{aligned}$$

  即 $\lambda_{\max}(\boldsymbol{A})\boldsymbol{I} \geq \boldsymbol{A}, \boldsymbol{A} \geq \lambda_{\min}(\boldsymbol{A})\boldsymbol{I}$.

- 定理5.3.4　设 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶正定矩阵，则
(1)若 $\boldsymbol{A}\geq\boldsymbol{B}>0$,则 $\boldsymbol{B}^{-1}\geq\boldsymbol{A}^{-1}>0$;
(2)若 $\boldsymbol{A}>\boldsymbol{B}>0$,则 $\boldsymbol{B}^{-1}>\boldsymbol{A}^{-1}>0$.
证明　由定理5.2.7知，有非奇异矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{P}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{P}=\boldsymbol{I},\boldsymbol{P}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$$

由 $\mathbf{A}\geq\boldsymbol{B}$, 有 $\boldsymbol{\Lambda}\geq\boldsymbol{I}$,则 $\lambda_i\geq 1(i=1,\cdots,n)$.从而

$$0<\boldsymbol{\Lambda}^{-1}=\mathrm{diag}(\lambda_1^{-1},\cdots,\lambda_n^{-1})\leq\boldsymbol{I}$$

因为 $\boldsymbol{A}^{-1}=\boldsymbol{P}\boldsymbol{\Lambda}^{-1}\boldsymbol{P}^{\mathrm{H}},\boldsymbol{B}^{-1}=\boldsymbol{P}\boldsymbol{I}\boldsymbol{P}^{\mathrm{H}}$,便得 $\boldsymbol{B}^{-1}\geq\boldsymbol{A}^{-1}>0.$
将上面证明的 $\geq$ 号改为 $>$,可得 $\boldsymbol{B}^{-1}>\boldsymbol{A}^{-1}>0.$

- 定理5.3.5　设 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶正定矩阵，且 $\boldsymbol{AB}=\boldsymbol{BA}$,则
  (1)若 $\boldsymbol{A}\geq\boldsymbol{B}$,则 $\boldsymbol{A}^2\geq\boldsymbol{B}^2$;
  (2)若 $\boldsymbol{A}>\boldsymbol{B}$,则 $\boldsymbol{A}^2>\boldsymbol{B}^2$.
  证明　以(1)为例，(2)的证明类似.
  由 $\boldsymbol{AB}=\boldsymbol{BA}$,则

$$\boldsymbol{A}^2-\boldsymbol{B}^2=\boldsymbol{A}^2+\boldsymbol{AB}-\boldsymbol{BA}-\boldsymbol{B}^2=(\boldsymbol{A}-\boldsymbol{B})(\boldsymbol{A}+\boldsymbol{B})$$
$$\boldsymbol{A}^2-\boldsymbol{B}^2=\boldsymbol{A}^2+\boldsymbol{BA}-\boldsymbol{AB}-\boldsymbol{B}^2=(\boldsymbol{A}+\boldsymbol{B})(\boldsymbol{A}-\boldsymbol{B})$$

  因为 $\boldsymbol{A}-\boldsymbol{B}\geq 0,\boldsymbol{A}+\boldsymbol{B}\geq 0$,则由定理5.3.1(12)知 $\boldsymbol{A}^2-\boldsymbol{B}^2\geq 0$,即 $\boldsymbol{A}^2\geq\boldsymbol{B}^2$.

- 定理5.3.6　设 $\boldsymbol{A}$ 是 $m\times n$ 的行满秩矩阵， $\boldsymbol{B}$ 是 $n\times k$ 矩阵，则

$$\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B} \geq (\boldsymbol{AB})^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}(\boldsymbol{AB}) \tag{5.3.1}$$

其中等号成立的充分必要条件是存在一个 $m\times k$ 矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{B}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{C}$.
证明　因为

$$\begin{aligned}
0 &\leq [\boldsymbol{B}-\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}]^{\mathrm{H}}[\boldsymbol{B}-\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}]\\
&=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}-2\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}+\boldsymbol{B}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AA}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}\\
&=\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}-(\boldsymbol{AB})^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}
\end{aligned}$$

则得(5.3.1),并且上式右端为零矩阵的充分必要条件是

$$\boldsymbol{B}=\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}$$

如果这个条件满足，则令 $\boldsymbol{C}=(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}$,有 $\boldsymbol{B}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{C}$;反过来，如果存在一个 $m\times k$ 矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{B}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{C}$,则

$$\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AB}=\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{AA}^{\mathrm{H}})^{-1}\boldsymbol{AA}^{\mathrm{H}}\boldsymbol{C}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{C}=\boldsymbol{B}$$

Thank you!