FDU 高等线性代数 5. 正定矩阵

本文根据邵美悦老师授课内容整理而成，并参考了以下教材：

- Matrix Analysis (R. Horn & C. Johnson) Chapter $4, 7$
- 矩阵分析 (R. Horn & C. Johnson) 第 $4, 7$ 章

欢迎批评指正!

5.1 相合变换

5.1.1 二次型

考虑多元函数 $f:\mathbb{R}^n\rightarrow\mathbb{R}$ 在 $x=x_0$ 处的二阶 Taylor 展开式：

$$f(x)=f(x_0)+\nabla f(x_0)^{\mathrm{T}}(x-x_0)+\frac{1}{2}(x-x_0)^{\mathrm{T}}\nabla^2 f(x_0)(x-x_0)+o(\|x-x_0\|^2)$$

Taylor 展开式提供了在某点 $x=x_0$ 处附近近似多元函数的方法.
在处理一般的近似求解问题时，通常展开到二阶项就足够了.
二阶项 $(x-x_0)^{\mathrm{T}}\nabla^2 f(x_0)(x-x_0)$ 就是一个二次型，其中 $\nabla^2 f(x_0)$ 是函数 f 在 x_0 处的 Hesse 矩阵.

给定实方阵 $A\in\mathbb{R}^{n\times n}$，
我们称 $f(x)=x^{\mathrm{T}}Ax$ $(\forall\ x\in\mathbb{R}^n)$ 为一个**实二次型** (Real Quadratic Form)
注意到 $x^{\mathrm{T}}Ax=x^{\mathrm{T}}(\frac{A+A^{\mathrm{T}}}{2})x$ $(\forall\ x\in\mathbb{R}^n)$，故我们通常默认 A 为实对称阵 (即满足 $A^{\mathrm{T}}=A$)
事实上，如果我们不限制 A 为对称阵，则系数矩阵 A 将不唯一.
值得注意的是，用上三角阵 (或下三角阵) 来表示二次型也是可行的，
因为上三角阵可以包含对称阵的所有信息.
可是一旦对 x 做线性变换 $x=Cy$ (其中 $C\in\mathbb{R}^{n\times n}$ 为非奇异阵)，则二次型变为
$x^{\mathrm{T}}Ax=y^{\mathrm{T}}(C^{\mathrm{T}}AC)y$，
其中 $C^{\mathrm{T}}AC$ 不一定能保持 A 的上三角结构，但一定能保持 A 的对称性.

对应地，给定 Hermite 阵 $A\in\mathbb{C}^{n\times n}$，
我们称 $f(x)=x^{\mathrm{H}}Ax$ $(\forall\ x\in\mathbb{C}^n)$ 为一个 **Hermite 型** (Hermitian Form)
它是一个 $\mathbb{C}^n\rightarrow\mathbb{R}$ 的实值函数.
对 x 做线性变换 $x=Cy$ (其中 $C\in\mathbb{C}^{n\times n}$ 为非奇异阵)，则二次型变为 $x^{\mathrm{H}}Ax=y^{\mathrm{H}}(C^{\mathrm{H}}AC)y$，
其中 $C^{\mathrm{H}}AC$ 一定能保持 A 的 Hermite 性.

二次型最早用来研究二次曲线，例如椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 对应的实二次型为：

$$\begin{bmatrix}x & y\end{bmatrix}\begin{bmatrix}\frac{1}{a^2} & \\ & \frac{1}{b^2}\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=0$$

对于非齐次的二次曲线问题 $x^{\mathrm{T}}Ax+2b^{\mathrm{T}}x+a=0$，对应的实二次型为：

$$\begin{bmatrix}x^{\mathrm{T}} & 1\end{bmatrix}\begin{bmatrix}A & b\\ b^{\mathrm{T}} & a\end{bmatrix}\begin{bmatrix}x\\ 1\end{bmatrix}=0$$

通过二次型，可以把齐次二次曲线和非齐次二次曲线放在一起研究.

5.1.2 相合关系

给定复方阵 $A, B \in \mathbb{C}^{n\times n}$

- 若存在一个非奇异阵 $C \in \mathbb{C}^{n\times n}$ 使得 $B = C^{\mathrm{T}}AC$，则我们称 A, B 是**相合的** (congruent)
- 若存在一个非奇异阵 $C \in \mathbb{C}^{n\times n}$ 使得 $B = C^{\mathrm{H}}AC$，则我们称 A, B 是**共轭相合的** (conjugate congruent)

(Matrix Analysis 定理 $4.5.5$)
我们容易验证以下事实:

- ① (共轭) 相合的矩阵具有相同的秩.
- ② 若 A 是对称阵，则 $C^{\mathrm{T}}AC$ 一定是对称阵 (即使 C 奇异，结论也成立)
  若 A 是 Hermite 阵，则 $C^{\mathrm{H}}AC$ 一定是 Hermite 阵 (即使 C 奇异，结论也成立)
  通常来说我们感兴趣的是保持矩阵类型不变的相合，即对称阵之间的相合，以及 Hermite 阵之间的共轭相合.
- ③ (共轭) 相合是等价关系，即满足自反性、对称性和传递性.

任意给定 Hermite 阵 $A \in \mathbb{C}^{n\times n}$，其特征值一定是实数.
我们定义 A 的**惯性指数** (inertia) 是有序的三元组:

$$\mathrm{Inertia}(A) = (n_+(A), n_-(A), n_0(A))$$

其中 $n_+(A)$ 是 A 正特征值的个数，$n_-(A)$ 是 A 负特征值的个数，$n_0(A)$ 是 A 零特征值的个数.
它们满足:

$$\begin{aligned}n_+(A) + n_-(A) + n_0(A) &= n \\ \mathrm{rank}(A) &= n_+(A) + n_-(A)\end{aligned}$$

我们定义 Hermite 阵 A 的**惯性矩阵**为 $I(A) = I_{n_+(A)} \oplus I_{n_-(A)} \oplus I_{n_0(A)}$

(Matrix Analysis 定理 $4.5.7$)
每一个 Hermite 矩阵 $A \in \mathbb{C}^{n\times n}$ 都共轭相合于它的惯性矩阵 $I(A) = I_{n_+(A)} \oplus I_{n_-(A)} \oplus I_{n_0(A)}$

- **证明:**
  任意给定 Hermite 阵 $A \in \mathbb{C}^{n\times n}$，它作为一个正规矩阵一定可以酉对角化 (即谱分解)
  即存在酉矩阵 $U \in \mathbb{C}^{n\times n}$ 使得 $\Lambda := U^{\mathrm{H}}AU$ 为对角阵
  设 Λ 的排列先是正特征值，接着是负特征值，最后是零特征值，则它可表示为:

  $$\Lambda = \mathrm{diag}\{\underbrace{\lambda_1, \ldots, \lambda_{n_+(A)}}_{n_+(A)}, \underbrace{\mu_1, \ldots, \mu_{n_-(A)}}_{n_-(A)}, \underbrace{0, \ldots, 0}_{n_0(A)}\}$$

  其中 $\lambda_1, \ldots, \lambda_{n_+(A)} > 0$ 而 $\mu_1, \ldots, \mu_{n_-(A)} < 0$
  我们将其分解为:

  $$\begin{aligned}&\Lambda = DI(A)D \\ &\text{where } \begin{cases} I(A) = I_{n_+(A)} \oplus I_{n_-(A)} \oplus I_{n_0(A)} \\ D = \mathrm{diag}\{\underbrace{\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_{n_+(A)}}}_{n_+(A)}, \underbrace{\sqrt{-\mu_1}, \ldots, \sqrt{\mu_{n_-(A)}}}_{n_-(A)}, \underbrace{1, \ldots, 1}_{n_0(A)}\} \end{cases}\end{aligned}$$

  记 $C = UD^{-1}$ (显然它是非奇异阵)，则我们有:

  $$C^{\mathrm{H}}AC = D^{-1}U^{\mathrm{H}}AUD^{-1} = D^{-1}\Lambda D^{-1} = I(A)$$

  命题得证.

(Sylvester 惯性定理, Matrix Analysis 定理 $4.5.8$**)**
Hermite 阵 $A,B\in\mathbb{C}^{n\times n}$ 是共轭相合的，当且仅当它们具有相同的惯性指数.

- 上述定理表明任意 Hermite 阵都只能与唯一的形如惯性矩阵的矩阵共轭相合.
  因此我们称 Hermite 阵 A 的惯性矩阵 $I(A)=I_{n_+(A)}\oplus I_{n_-(A)}\oplus I_{n_0(A)}$ 为**共轭相合标准型**.
- **证明:**
  充分性显然，下证必要性:
  若 Hermite 阵 $A,B\in\mathbb{C}^{n\times n}$ 是共轭相合的，则存在非奇异阵 $C\in\mathbb{C}^{n\times n}$ 使得 $A=C^{\mathrm{H}}BC$
  注意到共轭相合的矩阵具有相同的秩，故零惯性指数 $n_0(A)=n_0(B)$
  要证明 A,B 具有相同的惯性指数，只需证明正惯性指数 $n_+(A)=n_+(B)$ 即可.

  设 $v_1,\ldots,v_{n_+(A)}$ 是与 A 的正特征值 $\lambda_1,\ldots,\lambda_{n_+(A)}$ 相伴的特征向量 (自然都是非零向量，且相互正交)
  定义子空间 $S_+(A):=\operatorname{span}\{v_1,\ldots,v_{n_+}(A)\}$
  对于 $S_+(A)$ 中的任意非零向量 $x=\alpha_1v_1+\cdots+\alpha_{n_+(A)}v_{n_+(A)}$ (其中 $\alpha_1,\ldots,\alpha_{n_+(A)}\in\mathbb{C}$ 不全为零)
  我们都有:

$$
\begin{aligned}
x^{\mathrm{H}}Ax&=(\alpha_1v_1+\cdots+\alpha_{n_+(A)}v_{n_+(A)})^{\mathrm{H}}A(\alpha_1v_1+\cdots+\alpha_{n_+(A)}v_{n_+(A)})\\
&=|\alpha_1|^2v_1^{\mathrm{H}}Av_1+\cdots+|\alpha_{n_+(A)}|^2v_{n_+(A)}^{\mathrm{H}}Av_{n_+(A)}\\
&=|\alpha_1|^2\lambda_1\|v_1\|_2^2+\cdots+|\alpha_{n_+(A)}|^2\lambda_{n_+(A)}\|v_{n_+(A)}\|_2^2\\
&>0
\end{aligned}
$$

  对于任意非零的 $x\neq 0_n\in S_+(A)$，记 $y=Cx$，则我们有:

$$
\begin{aligned}
y^{\mathrm{H}}By&=(Cx)^{\mathrm{H}}B(Cx)\\
&=x^{\mathrm{H}}C^{\mathrm{H}}BCx\quad(\text{note that }A=C^{\mathrm{H}}BC)\\
&=x^{\mathrm{H}}Ax\quad(\text{note that }x\neq 0_n\in S_+(A))\\
&>0
\end{aligned}
$$

  因此我们有 $n_+(B)\geq\dim(\{y=Cx:x\in S_+(A)\})=\dim(S_+(A))=n_+(A)$ 成立.

  > **Courant-Fischer 定理的推论: (Matrix Analysis 推论** $4.2.12$**)**
  > 给定 Hermite 阵 $A\in\mathbb{C}^{n\times n}$，特征值按非减的次序排列: $\lambda_{\min}=\lambda_1\leq\cdots\leq\lambda_n=\lambda_{\max}$
  > 设 S 是 $\mathbb{C}^n$ 的一个给定的 k 维子空间
  >
  > - ① 若对于任意单位向量 $x\in S$ 都有 $x^{\mathrm{H}}Ax\leq 0$，则 A 至少有 k 个非正的特征值.
  >   若对于任意单位向量 $x\in S$ 都有 $x^{\mathrm{H}}Ax<0$，则 A 至少有 k 个负的特征值.
  > - ② 若对于任意单位向量 $x\in S$ 都有 $x^{\mathrm{H}}Ax\geq 0$，则 A 至少有 k 个非负的特征值.
  >   若对于任意单位向量 $x\in S$ 都有 $x^{\mathrm{H}}Ax>0$，则 A 至少有 k 个正的特征值.

  如果在上述推理过程中将 A,B 的角色互换，那么我们就得到 $n_+(A)\geq n_+(B)$
  于是我们有 $n_+(A)=n_+(B)$ 成立，进而有 $n_-(A)=n_-(B)$ 成立.
  综上所述，A,B 具有相同的惯性指数.
- **一点说明:**
  在已经说明了零惯性指数 $n_0(A)=n_0(B)$ 的基础之上，
  关于正惯性指数 $n_+(A)=n_+(B)$ 的证明还可利用**反证法**:

  设 Hermite 阵 $A,B\in\mathbb{C}^{n\times n}$ 的惯性矩阵为:

$$
\begin{aligned}
I(A)&=I_{n_+(A)}\oplus I_{n_-(A)}\oplus I_{n_0(A)}\\
I(B)&=I_{n_+(B)}\oplus I_{n_-(B)}\oplus I_{n_0(B)}
\end{aligned}
$$

  根据 **Matrix Analysis 定理** $4.5.7$ 可知 A 共轭相合于 $I(A)$，而 B 共轭相合于 $I(B)$

若 A, B 是共轭相合的，则根据共轭相合关系的传递性可知 $I(A)$ 与 $I(B)$ 也是共轭相合的.
于是存在非奇异阵 $C \in \mathbb{C}^{n\times n}$ 使得 $C^{\mathrm{H}} I(A) C = I(B)$

(反证法) 假设 $n_+(A) \neq n_+(B)$ (不妨假设 $n_+(A) > n_+(B)$，对应地有 $n_-(A) < n_-(B)$)
我们定义:

$$
\begin{aligned}
V_1 &:= \mathrm{span}\{C^{-1}e_1, \dots, C^{-1}e_{n_+(A)}\} \\
V_2 &:= \mathrm{span}\{e_{n_+(B)+1}, \dots, e_n\}
\end{aligned}
$$

则我们有:

$$
\begin{aligned}
\dim(V_1 \cap V_2) &= \dim(V_1) + \dim(V_2) - \dim(V_1 + V_2) \\
&= n_+(A) + (n - n_+(B)) - n \\
&= n_+(A) - n_+(B) \quad (\text{note that } n_+(A) > n_+(B)) \\
&> 0
\end{aligned}
$$

因此存在 $x \neq 0_n \in V_1 \cap V_2$，
即存在 $x \neq 0_n$ 使得 $\begin{cases} (Cx)^{\mathrm{H}} I(A)(Cx) = x^{\mathrm{H}}(C^{\mathrm{H}} I(A) C)x = x^{\mathrm{H}} I(B) x > 0 \\ x^{\mathrm{H}} I(B) x \leq 0 \end{cases}$
从而得出矛盾.
因此 $n_+(A) = n_+(B)$ 成立.

共轭相合变换虽然不改变 Hermite 阵特征值的符号，但却可以改变特征值的模长.
(Ostrowski 定理, Matrix Analysis 定理 $4.5.9$)
给定 Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 和非奇异阵 $C \in \mathbb{C}^{n\times n}$
设 $A, C^{\mathrm{H}}AC$ 的特征值都按非减的次序排列，
则对于任意 $k = 1, 2, \dots, n$，都存在正实数
$\alpha_k \in [\sigma_{\min}^2(C), \sigma_{\max}^2(C)] = [\lambda_{\min}(C^{\mathrm{H}}C), \lambda_{\max}(C^{\mathrm{H}}C)]$，
使得 $\lambda_k(C^{\mathrm{H}}AC) = \alpha_k \cdot \lambda_k(A)$ 成立.

5.2 正定矩阵

5.2.1 基本性质

给定 Hermite 阵 $A \in \mathbb{C}^{n\times n}$

- 若对于任意 $x \in \mathbb{C}^n$，都有 $x^{\mathrm{H}}Ax \geq 0$，则我们称 A 是**半正定的** (positive semidefinite)，记为 $A \succeq 0$
- 若对于任意非零向量 $x \in \mathbb{C}^n \backslash \{0_n\}$，都有 $x^{\mathrm{H}}Ax > 0$，则我们称 A 是**正定的** (positive definite)，记为 $A \succ 0$
- 若 $-A$ 半正定，则我们称 A 是**半负定的** (negative semidefinite)，记为 $A \preceq 0$
- 若 $-A$ 正定，则我们称 A 是**负定的** (negative definite)，记为 $A \prec 0$
- 若 A 既不是半正定的，也不是半负定的，则我们称 A 是**不定的** (indefinite)

Hermite (半) 正定阵具有如下性质:

- **(Matrix Analysis 结论 $7.1.6$ & $7.1.7$)**
若 Hermite 阵 A 是半正定的，则 $x^{\mathrm{H}}Ax = 0$ 当且仅当 $Ax = 0_n$
由此可知半正定阵是正定的，当且仅当它是非奇异的.
- **(Matrix Analysis 结论 $7.1.3$)**
任意给定半正定阵 $A_1, \dots, A_m \in \mathbb{C}^{n\times n}$ 和非负实数 $\alpha_1, \dots, \alpha_m$，则 $\sum_{k=1}^m \alpha_k A_k$ 一定是半正定的.
特殊地，如果存在某个 $k \in \{1, \dots, m\}$ 使得 A_k 是正定的且 $\alpha_k > 0$，则 $\sum_{k=1}^m \alpha_k A_k$ 是正定

的.

- **(Matrix Analysis 结论** $7.1.4$**)**
  正定阵 (半正定阵) 的所有特征值都是正实数 (非负实数).
- **(Matrix Analysis 结论** $7.1.2$ **&** $7.1.5$ **&** $7.1.10$**)**
  若 Hermite 阵 A 是 (半) 正定的，则其所有的主子阵都是 (半) 正定的.
  且 $\text{tr}\,(A)$ 和 $\det\,(A)$ 以及 A 的所有主子式都是正实数 (非负实数)

  特殊地，正定阵 (半正定阵) 的所有主对角元都是正实数 (非负实数).
  此外，若半正定阵 $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ 的对角元 $a_{kk} = 0$，则 A 的第 k 行和第 k 列均为零.
- **(Matrix Analysis 结论** $7.1.9$**)**
  若 Hermite 阵 A 是半正定的，则存在正定矩阵序列 $\{A_k\}$ 逐元素收敛到 A.
- 相合变换保持半正定性不变:
  (Matrix Analysis 结论 $7.1.8$**)**
  任意给定 Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 和 $C \in \mathbb{C}^{n\times m}$
  - 若 A 半正定，则 $C^{\text{H}}AC$ 也是半正定的，且满足 $\begin{cases} \text{Ker}(C^{\text{H}}AC) = \text{Ker}(AC) \\ \text{rank}(C^{\text{H}}AC) = \text{rank}(AC) \end{cases}$
  - 若 A 正定，则 $C^{\text{H}}AC$ 至少是半正定的，且满足
    $\begin{cases} \text{Ker}(C^{\text{H}}AC) = \text{Ker}(AC) = \text{Ker}(C) \\ \text{rank}(C^{\text{H}}AC) = \text{rank}(AC) = \text{rank}(C) \end{cases}$
    因此 $C^{\text{H}}AC \in \mathbb{C}^{m\times m}$ 是正定的，当且仅当 $\text{rank}(C) = m$

5.2.2 特征刻画

第一种刻画:
(Matrix Analysis 定理 $7.2.1$**)**
Hermite 阵是正定的 (半正定的)，当且仅当其所有特征值都是正实数 (非负实数)

- **(Matrix Analysis 推论** $7.2.2$**)**
  若 Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 是半正定的，则其正整数次幂 $A^k\ (k \in \mathbb{Z}_+)$ 也都是半正定的.
- **(Matrix Analysis 推论** $7.2.3$**)**
  若 Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 严格对角占优且所有对角元均为正实数，则 A 是正定的.
- **(Matrix Analysis 推论** $7.2.4$**)**
  若 Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 是半正定的，则其特征多项式的非零系数的符号是严格交错的.
  具体来说，设 $p_A(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_{n-m}t^{n-m}$
  其中 $a_{n-m} \neq 0$ 且 $a_{k+1}a_k < 0\ (\forall\, k = n-m, \ldots, n-1)$

第二种刻画:
(Sylvester 判别法, Matrix Analysis 定理 $7.2.5$**)**
设 $A \in \mathbb{C}^{n\times n}$ 是 Hermite 阵.

- A 是半正定的当且仅当其所有主子式 (注意不是顺序主子式) 都是非负实数.
- A 是正定的当且仅当其所有顺序主子式 (或逆序主子式) 都是正实数.

5.2.3 k 次根

对于任意正整数 $k \in \mathbb{Z}_+$，非负实数都有唯一的非负 k 次根.
Hermite 半正定阵有一个对应的性质.
(Matrix Analysis 定理 $7.2.6$**)**
设 Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 是半正定的，记 $r = \text{rank}(A)$，设 $k \in \mathbb{Z}_+$
则我们有:

- ① 存在唯一的 Hermite 半正定阵 B 使得 $B^k = A$ (因此我们可将其记为 $A^{\frac{1}{k}}$)
  具体来说，设 A 的谱分解为 $A = U\Lambda U^{\mathrm{H}}$，则 $A^{\frac{1}{k}} := U\Lambda^{\frac{1}{k}}U^{\mathrm{H}}$
  其中 $U \in \mathbb{C}^{n\times n}$ 为酉矩阵，$\Lambda = \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$ 的对角元为 A 的特征值
  而 $\Lambda^{\frac{1}{k}}$ 的定义为 $\Lambda^{\frac{1}{k}} := \operatorname{diag}\{\sqrt[k]{\lambda_1}, \ldots, \sqrt[k]{\lambda_n}\}$
- ② 存在实系数多项式 $p(\cdot)$ 使得 $A^{\frac{1}{k}} = p(A)$
  因此 $A^{\frac{1}{k}}$ 和任意与 A 可交换的矩阵都可交换.
- ③ $\operatorname{Range}(A^{\frac{1}{k}}) = \operatorname{Range}(A)$ (因此 $\operatorname{rank}(A^{\frac{1}{k}}) = \operatorname{rank}(A)$)
- ④ 若 A 退化为实对称阵，则 $A^{\frac{1}{k}}$ 也退化为实对称阵.

(Matrix Analysis 定理 $7.2.7$)
设 $A \in \mathbb{C}^{n\times n}$ 是 Hermite 阵，则我们有:

- ① A 是半正定的，当且仅当存在 $B \in \mathbb{C}^{m\times n}$ 使得 $A = B^{\mathrm{H}}B$
- ② 若半正定阵 $A = B^{\mathrm{H}}B$ (其中 $B \in \mathbb{C}^{m\times n}$)，则 $Ax = 0_n$ 当且仅当 $Bx = 0_m$，因此我们有:

$$\begin{cases} \operatorname{Ker}(A) = \operatorname{Ker}(B) \\ \operatorname{rank}(A) = \operatorname{rank}(B) \end{cases}$$

  特殊地，$A = B^{\mathrm{H}}B$ 是正定的当且仅当 B 是列满秩的.
  (另一种表述: Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 是正定的，当且仅当其相合于单位矩阵 I_n)

半正定阵的分解式 $A = B^{\mathrm{H}}B$ 可由多种方法得到，例如 Cholesky 分解.
(半正定阵的 Cholesky 分解, Matrix Analysis 推论 $7.2.9$)
Hermite 阵 $A \in \mathbb{C}^{n\times n}$ 是正定的 (半正定的)，
当且仅当存在一个对角元均为正实数 (非负实数) 的下三角阵 $L \in \mathbb{C}^{n\times n}$ 使得 $A = LL^{\mathrm{H}}$ (称为 Cholesky 分解)
特殊地，若 A 是正定的，则 Cholesky 分解是唯一的.
若 A 退化为实对称阵，则 L 可取为实矩阵.

- **证明:**
  设 $A^{\frac{1}{2}}$ 的 QR 分解为 $A^{\frac{1}{2}} = QR$
  记 $L := R^{\mathrm{H}}$，则我们有:

$$\begin{aligned} A &= A^{\frac{1}{2}}A^{\frac{1}{2}} \\ &= (A^{\frac{1}{2}})^{\mathrm{H}}A^{\frac{1}{2}} \\ &= (QR)^{\mathrm{H}}(QR) \\ &= R^{\mathrm{H}}Q^{\mathrm{H}}QR \\ &= LL^{\mathrm{H}} \end{aligned}$$

  正定矩阵 Cholesky 分解的唯一性可由满秩方阵的 QR 分解的唯一性得到，也可以通过反证法得到.
  假设存在两个对角元均为正实数的下三角阵 L_1, L_2 使得正定阵 $A = L_1L_1^{\mathrm{H}} = L_2L_2^{\mathrm{H}}$，
  则我们有 $L_2^{-1}L_1 = L_2^{\mathrm{H}}L_1^{-\mathrm{H}}$.
  注意到等式左端是两个下三角阵的乘积，因此仍是下三角阵.
  而等式右端是两个上三角阵的乘积，因此仍是上三角阵.
  对比可知矩阵乘积一定是一个对角阵，因此只考虑对角元即可:

$$\begin{gathered} \frac{L_1(i,i)}{L_2(i,i)} = \frac{L_2(i,i)}{L_1(i,i)} \quad (i = 1, \ldots, n) \\ \Updownarrow \\ L_1(i,i) = L_2(i,i) \quad (i = 1, \ldots, n) \end{gathered}$$

  这表明 L_1 和 L_2 的对角元完全相同，且 $L_2^{-1}L_1$ 的对角元均为 1.
  注意到:

$$L_2^{-1}L_1L_1^{\mathrm{H}}L_2^{-\mathrm{H}} = L_2^{-1}L_1(L_2^{-1}L_1)^{\mathrm{H}} = I_n$$

故 $L_2^{-1}L_1$ 为酉矩阵.
考虑到 $L_2^{-1}L_1$ 同时还是单位下三角阵，因此其非对角元均为零，即 $L_2^{-1}L_1 = I_n$.
(事实上，若酉矩阵 U 是下三角的，则它必为对角矩阵，且所有对角元的模长都为 1)
这说明 $L_1 = L_2$.

5.2.4 Schur 乘积定理

两个 Hermite 阵 $A, B \in \mathbb{C}^{n\times n}$ 的乘积 AB 是 Hermite 的，当且仅当它们可交换 (即 $AB = BA$)
两个 Hermite 阵 $A, B \in \mathbb{C}^{n\times n}$ 的 Hadamard 乘积 (即 Schur 乘积) $A \odot B$ 一定是 Hermite 的.
我们不仅要问，两个 Hermite 半正定阵的 Hadamard 乘积是否仍是 Hermite 半正定阵?

在解决上述问题之前，我们首先证明一个引理:
与 Hadamard 乘积相关联的双线性型可表示为矩阵乘积的迹.
(Matrix Analysis 引理 $7.5.2$)
对于任意 $A, B \in \mathbb{C}^{n\times n}$ 和 $x, y \in \mathbb{C}^n$，我们都有:

$$x^{\mathrm{H}}(A \odot B)y = \mathrm{tr}\left(\mathrm{diag}(\bar{x})A \cdot \mathrm{diag}(y)B^{\mathrm{T}}\right)$$

- **证明:**
  注意到:
  $$\begin{aligned}\mathrm{diag}(\bar{x})A &= [\bar{x}_i a_{ij}]_{i,j=1}^n \\ \mathrm{diag}(y)B^{\mathrm{T}} &= [y_i b_{ji}]_{i,j=1}^n\end{aligned}$$
  于是我们有:
  $$\begin{aligned}\mathrm{tr}\left(\mathrm{diag}(\bar{x})A \cdot \mathrm{diag}(y)B^{\mathrm{T}}\right) &= \sum_{i=1}^n \sum_{j=1}^n [\mathrm{diag}(\bar{x})A]_{i,j}[\mathrm{diag}(y)B^{\mathrm{T}}]_{j,i} \\ &= \sum_{i=1}^n \sum_{j=1}^n (\bar{x}_i a_{ij})(y_j b_{ij}) \\ &= \sum_{i=1}^n \sum_{j=1}^n \bar{x}_i (a_{ij} b_{ij}) y_j \\ &= x^{\mathrm{H}}(A \odot B)y\end{aligned}$$

(Schur 乘积定理, Matrix Analysis 定理 $7.5.3$)
设 Hermite 阵 $A, B \in \mathbb{C}^{n\times n}$ 是半正定的，则我们有:

- ① $A \odot B$ 半正定
- ② 若 A 正定且 B 的每个主对角元都是正实数，则 $A \odot B$ 正定.
  特殊地，若 A, B 正定，则 $A \odot B$ 正定.

证明:

- ① 假设 Hermite 阵 A, B 是半正定的.
  对于任意 $x \in \mathbb{C}^n$，记 $C := \mathrm{diag}(x)(B^{\mathrm{T}})^{\frac{1}{2}} = \mathrm{diag}(x)\bar{B}^{\frac{1}{2}}$
  则根据 **Matrix Analysis 引理** $7.5.2$ 我们有:

$$
\begin{aligned}
x^{\mathrm{H}}(A \odot B)x &= \operatorname{tr}\left(\operatorname{diag}(\bar{x})A \cdot \operatorname{diag}(x)B^{\mathrm{T}}\right) \\
&= \operatorname{tr}\left(\operatorname{diag}(\bar{x})A \cdot \operatorname{diag}(x)\bar{B}\right) \\
&= \operatorname{tr}\left(\bar{B}^{\frac{1}{2}}\operatorname{diag}(\bar{x})A \cdot \operatorname{diag}(x)\bar{B}^{\frac{1}{2}}\right) \\
&= \operatorname{tr}\left(C^{\mathrm{H}}AC\right)
\end{aligned}
$$

注意到合同变换保持 Hermite 半正定性，可知 $C^{\mathrm{H}}AC$ 也为 Hermite 半正定阵.
因此其对角元均为非负实数，于是有 $x^{\mathrm{H}}(A \odot B)x = \operatorname{tr}\left(C^{\mathrm{H}}AC\right) \geq 0$ 成立.
根据 $x \in \mathbb{C}^n$ 的任意性可知 $A \odot B$ Hermite 半正定.

- ② 假设 Hermite 阵 A 是正定的，Hermite 阵 B 是半正定的且每个主对角元均为正实数.
  设 A 的最小特征值为 $\alpha > 0$，B 的最小对角元为 $\beta > 0$
  根据 ① 可知 $(A - \alpha I_n) \odot B$ 是半正定的.
  因此对于任意 $x \in \mathbb{C}^n$ 都有:

  $$
  0 \leq x^{\mathrm{H}}((A - \alpha I_n) \odot B)x = x^{\mathrm{H}}(A \odot B)x - \alpha x^{\mathrm{H}}(I_n \odot B)x
  $$

  于是对于任意 $x \in \mathbb{C}^n \backslash \{0_n\}$ 我们都有:

  $$
  \begin{aligned}
  x^{\mathrm{H}}(A \odot B)x &\geq \alpha x^{\mathrm{H}}(I_n \odot B)x \quad (\text{note that } x^{\mathrm{H}}(A \odot B)y = \operatorname{tr}\left(\operatorname{diag}(\bar{x})A \cdot \operatorname{diag}(y)B^{\mathrm{T}}\right)) \\
  &= \alpha \operatorname{tr}\left(\operatorname{diag}(\bar{x})I_n \cdot \operatorname{diag}(x)B^{\mathrm{T}}\right) \\
  &= \alpha \sum_{i=1}^{n} b_{ii}|x_i|^2 \\
  &\geq \alpha\beta \sum_{i=1}^{n} |x_i|^2 \\
  &= \alpha\beta\|x\|_2^2 \\
  &> 0
  \end{aligned}
  $$

  因此 $A \odot B$ 是正定的.

(Moutard 定理, Matrix Analysis 定理 $7.5.4$)
Hermite 阵 $A = [a_{ij}] \in \mathbb{C}^{n \times n}$ 半正定，
当且仅当对于每个半正定阵 $B = [b_{ij}] \in \mathbb{C}^{n \times n}$ 都有 $\operatorname{tr}\left(AB^{\mathrm{T}}\right) = \sum_{i,j=1}^{n} a_{ij}b_{ij} \geq 0$

证明:

- **必要性:**
  设 A, B 为半正定阵.
  根据 Schur 乘积定理可知 $A \odot B$ 也是半正定阵.
  于是我们有 $\operatorname{tr}\left(AB^{\mathrm{T}}\right) = \operatorname{tr}\left(\operatorname{diag}(1_n)A \cdot \operatorname{diag}(1_n)B^{\mathrm{T}}\right) = 1_n^{\mathrm{H}}(A \odot B)1_n \geq 0$
- **充分性:**
  设对于每个半正定阵 $B = [b_{ij}] \in \mathbb{C}^{n \times n}$ 都有 $\operatorname{tr}\left(AB^{\mathrm{T}}\right) = \sum_{i,j=1}^{n} a_{ij}b_{ij} \geq 0$
  任意给定 $x \in \mathbb{C}^n$，取 $B = \bar{x}\bar{x}^{\mathrm{H}} = \bar{x}x^{\mathrm{T}}$ 即得到:

  $$
  x^{\mathrm{H}}Ax = \sum_{i,j=1}^{n} a_{ij}\bar{x}_i x_j = \operatorname{tr}\left(AB^{\mathrm{T}}\right) \geq 0
  $$

(Matrix Analysis 定理 $7.5.9$)
设 $A = [a_{ij}] \in \mathbb{C}^{n \times n}$ 半正定，则我们有:

- ① 对于任意 $k \in \mathbb{Z}_+$，$A^{(k)} := [a_{ij}^k]$ 都是半正定的.
  它们是正定的，如果 A 是正定的.

- ② 设 $f(z) := \sum_{k=0}^{\infty} a_k z^k$ 是具有非负系数且收敛半径 $R > 0$ 的解析函数.
  若 $|a_{ij}| < R\ (\forall\, i, j \in \{1, \ldots, n\})$，则 $F := [f(a_{ij})] \in \mathbb{C}^{n\times n}$ 是半正定的.
  若 A 正定且某个对角元 $a_{ii} > 0$ **(存疑, 书上有误)**，则 $F := [f(a_{ij})] \in \mathbb{C}^{n\times n}$ 是正定的.
- ③ Hadamard 指数矩阵 $E := [e^{a_{ij}}] \in \mathbb{C}^{n\times n}$ 是正定的，当且仅当 A 没有两行是相同的.

5.2.5 同时对角化

一种同时对角化的方法:
(Matrix Analysis 定理 $7.6.1$)
设 $A, B \in \mathbb{C}^{n\times n}$ 是 Hermite 阵.

- ① 若 A 正定，则存在非奇异阵 $S \in \mathbb{C}^{n\times n}$ 使得 $\begin{cases} A = SS^{\mathrm{H}} \\ B = S^{-H} D S^{-1} \end{cases}$

  其中对角阵 D 与 B 拥有相同的惯性指数.
  (因此若 B 半正定，则 D 的对角元均为非负实数; 若 B 正定，则 D 的对角元均为正实数)
- ② 若 A, B 半正定，且 $\operatorname{rank}(A) = r$，则存在非奇异阵 $S \in \mathbb{C}^{n\times n}$ 使得
  $\begin{cases} A = S(I_r \oplus 0_{(n-r)\times(n-r)})S^{\mathrm{H}} \\ B = S^{-H} D S^{-1} \end{cases}$

  其中对角阵 D 与 B 拥有相同的惯性指数，因此 D 的对角元均为非负实数且
  $\operatorname{rank}(D) = \operatorname{rank}(B)$

(Matrix Analysis 推论 $7.6.2$)
设 $A, B \in \mathbb{C}^{n\times n}$ 是 Hermite 阵.

- ① 若 A 正定，则 AB 可对角化且与 B 拥有相同的惯性指数.
  (因此若 B 半正定，则 AB 的特征值均为非负实数; 若 B 正定，则 AB 的特征值均为正实数)
- ② 若 A, B 半正定，则 AB 可对角化，且具有非负的特征值.

不过一个半正定阵与一个 Hermite 阵的乘积不一定可对角化:

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

下面的定理表明上述例子是典型的:
两个 Hermite 阵的乘积永远是拟可对角化的，且其 Jordan 标准型中的任意 Jordan 块都是幂零的.

(Matrix Analysis 定理 $7.6.3$)
若 $A, B \in \mathbb{C}^{n\times n}$ 是 Hermite 阵，且 A 是半正定的奇异矩阵，则 AB 相似于 $\Lambda \oplus N$
其中 Λ 是实对角阵，而 $N = J_2(0) \oplus \cdots \oplus J_2(0)$ 是 2 阶幂零 Jordan 块的直和.
(直和项 Λ 和 N 中的任意一个都有可能不出现)

另一种同时对角化的方法:
(Matrix Analysis 定理 $7.6.4$)
设 $A, B \in \mathbb{C}^{n\times n}$ 是 Hermite 阵.

- ① 若 A 正定，则存在非奇异阵 $S \in \mathbb{C}^{n\times n}$ 使得 $\begin{cases} A = SS^{\mathrm{H}} \\ B = S\Lambda S^{\mathrm{H}} \end{cases}$

其中对角阵 Λ 与 B 拥有相同的惯性指数.
(因此若 B 半正定，则 Λ 的对角元均为非负实数; 若 B 正定，则 Λ 的对角元均为正实数)
此外，Λ 的对角元即为 $A^{-1}B$ 的特征值.

- ② 若 A, B 半正定，且 $\text{rank}(A) = r$，则存在非奇异阵 $S \in \mathbb{C}^{n\times n}$ 使得
$\begin{cases} A = S(I_r \oplus 0_{(n-r)\times(n-r)})S^{\text{H}} \\ B = S\Lambda S^{\text{H}} \end{cases}$
其中对角阵 Λ 与 B 拥有相同的惯性指数，因此 Λ 的对角元均为非负实数且
$\text{rank}(\Lambda) = \text{rank}(B)$

上述结论的一个变形:
(Matrix Analysis 定理 $7.6.5$**)**
若 $A \in \mathbb{C}^{n\times n}$ 正定，而 $B \in \mathbb{C}^{n\times n}$ 是复对称的，则存在非奇异阵 $S \in \mathbb{C}^{n\times n}$ 使得 $\begin{cases} A = SS^{\text{H}} \\ B = S\Lambda S^{\text{T}} \end{cases}$
其中 Λ 是非负的对角矩阵.
Λ^2 的主对角元是可对角化矩阵 $A^{-1}B\bar{A}^{-1}\bar{B}$ 的特征值.

上述定理有以下应用:
(Matrix Analysis 定理 $7.6.6$**)**
函数 $f(A) = \log(\det(A))$ 是 Hermite 正定阵构成的集合 $\{A \in \mathbb{C}^{n\times n} : A \succeq 0\}$ 上的严格凹函数.

(Matrix Analysis 推论 $7.6.8$**)**
若 $A, B \in \mathbb{C}^{n\times n}$ Hermite 正定，则对于任意 $0 < \alpha < 1$ 都有:

$$\det(\alpha A + (1-\alpha)B) \geq (\det(A))^{\alpha}(\det(B))^{1-\alpha}$$

上式当且仅当 $A = B$ 时取等.

(Matrix Analysis 定理 $7.6.10$**)**
函数 $f(A) = \text{tr}\left(A^{-1}\right)$ 是 Hermite 正定阵构成的集合 $\{A \in \mathbb{C}^{n\times n} : A \succeq 0\}$ 上的严格凸函数.

The End