线性代数 (XIII)

Lecturer: 杨启哲 Last modified: 2025 年 6 月 1 日

1 对称矩阵

一个方阵是对称 (Symmetric)的，如果其满足：

$$A = A^T$$

下列矩阵很显然都是对称的：

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix},$$

考虑将其对角化，我们有：

$$\begin{aligned}
\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} &= X_1 \Lambda_1 X_1^{-1} = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \\
\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} &= X_2 \Lambda_2 X_2^{-1} = \begin{bmatrix} \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & -\frac{5}{\sqrt{30}} \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{30}} & -\frac{5}{\sqrt{30}} \end{bmatrix}
\end{aligned}$$

仔细观察可以发现，其对应的特征矩阵 X 是正交的，即：

$$X_i^{-1} = X_i^T \quad (i = 1, 2)$$

在本节中我们将证明，这并不是巧合，而是对称矩阵的一个重要性质。事实上我们将说明：

1. 实对称矩阵的特征值都是实数。
2. 实对称矩阵不同特征值的特征向量是正交的，且其一定可以对角化。

实对称矩阵的特征值 我们将证明如下定理：

定理 1.1

实对称矩阵 S 的特征值是实数，特别的，其每个特征值都有一个对应的实特征向量。

为了完成这个定理的证明，我们需要使用一些复数上的知识：令 $x \in \mathbb{C}$，则我们有存在 $a, b \in \mathbb{R}$, 使得：

$$x = a + bi$$

我们定义 x 的共轭复数 (complex conjugate) 为:

$$\bar{x} = a - bi$$

共轭复数有如下的性质:

引理 1.2

给定复数 $x, y \in \mathbb{C}$，我们有:

1. 如果 $\bar{x}x = 0$，则 $x = 0$.
2. $\overline{x+y} = \bar{x} + \bar{y}$.
3. $\overline{xy} = \bar{x}\bar{y}$.

由共轭复数可以定义在复矩阵中的**共轭矩阵**:

定义 1.3 (共轭矩阵 (Conjugate Matrix))

对于一个矩阵 A，我们定义其共轭矩阵 $\bar{A}$ 为:

$$\bar{A}(i,j) = \overline{A(i,j)}.$$

为了证明定理1.1，还需要共轭矩阵上的如下性质:

引理 1.4

令 $A \in \mathbb{C}^{m\times n}$ 是一个 $m \times n$ 的复矩阵，$\mathbf{x} \in \mathbb{C}^n$ 是任一复向量，则:

1. 如果 $\bar{\mathbf{x}}^T\mathbf{x} = 0$, 则 $\mathbf{x} = \mathbf{0}$.
2. 对任意的 $\lambda \in \mathbb{C}$，我们有: $\overline{\lambda A} = \bar{\lambda}\bar{A}$.
3. $\overline{A\mathbf{x}} = \bar{A}\bar{\mathbf{x}}$.

现在可以给出定理1.1的证明了。

证明: [定理1.1的证明] 令 S 是 $n \times n$ 的实对称矩阵，$\lambda \in \mathbb{C}$ 是 S 的特征值，$\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ 是 λ 对应的特征向量，则:

$$\begin{aligned}
&S\mathbf{x} = \lambda\mathbf{x} \\
\Longrightarrow &\bar{S}\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}} \\
\Longrightarrow &\bar{\mathbf{x}}^T\bar{S}^T = \bar{\lambda}\bar{\mathbf{x}}^T \\
\Longrightarrow &\bar{\mathbf{x}}^T S = \bar{\lambda}\bar{\mathbf{x}}^T \qquad (S \text{ 是实对称的，从而 } \bar{S}^T = S) \\
\Longrightarrow &\bar{\mathbf{x}}^T S\mathbf{x} = \bar{\lambda}\bar{\mathbf{x}}^T\mathbf{x} \\
\Longrightarrow &\lambda\bar{\mathbf{x}}^T\mathbf{x} = \bar{\lambda}\bar{\mathbf{x}}^T\mathbf{x} \\
\Longrightarrow &(\lambda - \bar{\lambda})\bar{\mathbf{x}}^T\mathbf{x} = 0
\end{aligned}$$

显然由于 $\mathbf{x} \neq \mathbf{0}$，从而 $\bar{\mathbf{x}}^T\mathbf{x} \neq 0$，因此 $\lambda - \bar{\lambda} = 0$，即 λ 是实数。

下证每个特征值都有一个对应的实特征向量。假设 $x_j = a_j + b_j i$，即：

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + i \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \mathbf{a} + i\mathbf{b}$$

由于 $S\mathbf{x} = \lambda \mathbf{x}$，我们有：

$$S\mathbf{x} = S(\mathbf{a} + i\mathbf{b}) = \lambda(\mathbf{a} + i\mathbf{b})$$

由于 S 是实对称矩阵，λ 是实数，我们有：

$$S\mathbf{a} = \lambda \mathbf{a},\ S\mathbf{b} = \lambda \mathbf{b}$$

由于 $\mathbf{x} \neq 0$，$\mathbf{a}$ 和 $\mathbf{b}$ 至少有一个是 S 的特征向量。 □

实对称矩阵的对角化 现在来考察其特征向量，下述定理阐明了不同特征值对应的特征向量不仅是线性无关的，更是正交的：

定理 1.5

对于一个实对称矩阵 S，如果 λ_1 和 λ_2 是 S 的两个不同的特征值，$\mathbf{x}_1$ 和 $\mathbf{x}_2$ 是 λ_1 和 λ_2 对应的特征向量，则 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 是正交的。

证明： 由假设有：

$$S\mathbf{x}_1 = \lambda_1 \mathbf{x}_1,\ S\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$$

从而：

$$\begin{aligned} \lambda_1(\mathbf{x}_1 \cdot \mathbf{x}_2) &= (\lambda_1 \mathbf{x}_1) \cdot \mathbf{x}_2 = (S\mathbf{x}_1) \cdot \mathbf{x}_2 = (S\mathbf{x}_1)^T \mathbf{x}_2 = \mathbf{x}_1^T S^T \mathbf{x}_2 \\ &= \mathbf{x}_1^T S \mathbf{x}_2 = \mathbf{x}_1^T (S\mathbf{x}_2) = \mathbf{x}_1^T (\lambda_2 \mathbf{x}_2) = \mathbf{x}_1 \cdot (\lambda_2 \mathbf{x}_2) = \lambda_2(\mathbf{x}_1 \cdot \mathbf{x}_2) \end{aligned}$$

由于 $\lambda_1 \neq \lambda_2$，从而 $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$，即 $\mathbf{x}_1$ 和 $\mathbf{x}_2$ 是正交的。 □

定理1.5可以直接推出如下结论：

推论 1.6

令 S 是一个 $n \times n$ 的实对称矩阵，如果 S 存在n 个两两不同的特征值，则 S 可以对角化。更精确的说，存在一个正交矩阵 Q 和对角矩阵 Λ 使得：

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = Q^T S Q$$

这里 $\lambda_1, \cdots, \lambda_n$ 都是 S 的特征值。

证明： 令 S 的特征值为 $\lambda_1, \ldots, \lambda_n$，由定理1.1和定理1.5，$S$ 存在如下的 n 个特征向量 $\mathbf{q}_1, \ldots, \mathbf{q}_n$，满足：

1. 对每个 $i=1,\ldots,n$，$Sq_i=\lambda_i q_i$.
2. 对于任意的 $i\neq j$，q_i 和 q_j 是正交的，即 $q_i\cdot q_j=0$.
3. q_i 的模长为 1，即 $\|q_i\|=1$.

令 $Q=\begin{bmatrix} q_1 & q_2 & \cdots & q_n\end{bmatrix}$，则 Q 是一个正交矩阵，从而:

$$Q^TSQ=Q^TS\begin{bmatrix} q_1 & q_2 & \cdots & q_n\end{bmatrix}=\begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T\end{bmatrix}\begin{bmatrix} \lambda_1 q_1 & \lambda_2 q_2 & \cdots & \lambda_n q_n\end{bmatrix}=\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}.$$

□

事实上，我们并不需要“n 个两两不同的特征值”这个条件。

定理 1.7

令 S 是一个 $n\times n$ 的实矩阵，则 S 是对称的当且仅当存在一个正交矩阵 Q 和对角矩阵 Λ 使得:

$$\Lambda=\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}=Q^TSQ$$

这里 $\lambda_1,\cdots,\lambda_n$ 都是 S 的特征值。

证明: $(\Leftarrow)$ 的方向是显然的，注意到:

$$S^T=\left(Q\Lambda Q^T\right)^T=Q\Lambda Q^T=S$$

下证 $(\Rightarrow)$ 方向。令 S 是一个 $n\times n$ 的实对称矩阵，我们对 n 作归纳。

BASE: $n=1$ 时是显然的。

INDUCTION: 现在令 $n\geqslant 2$。令 S 的一个特征值为 $\lambda_1\in\mathbb{R}$，其对应的一个特征向量为 $x_1\in\mathbb{R}^n\setminus\{0\}$。我们不妨可以假设 $\|x_1\|=1$，否则我们可以令:

$$x_1\leftarrow\frac{x_1}{\|x_1\|}$$

通过 Gram–Schmidt 正交化，我们可以从 x_1 扩展出一组标准正交基:

$$x_1,x_2,\cdots,x_n$$

满足:

$$x_i\cdot x_j=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

定义矩阵 P:

$$P = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

不难验证 P 是正交矩阵，即 $P^TP = I$。并且我们有:

$$p^T x_i = e_i$$

从而:

$$\begin{aligned} P^TSP &= P^TS \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \\ &= \begin{bmatrix} P^TSx_1 & P^TSx_2 & \cdots & P^TSx_n \end{bmatrix} \\ &= \begin{bmatrix} P^T\lambda_1 x_1 & P^TSx_2 & \cdots & P^TSx_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 P^T x_1 & P^TSx_2 & \cdots & P^TSx_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 e_1 & P^TSx_2 & \cdots & P^TSx_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & a^T \\ 0 & B \end{bmatrix} \end{aligned}$$

这里 $a \in \mathbb{R}^{n-1}$ 是一个列向量，B 是一个 $(n-1) \times (n-1)$ 的矩阵。注意到:

$$\begin{bmatrix} \lambda_1 & 0^T \\ a & B^T \end{bmatrix} = \left(\begin{bmatrix} \lambda_1 & a^T \\ 0 & B \end{bmatrix} \right)^T = (P^TSP)^T = P^TS^TP = P^TSP = \begin{bmatrix} \lambda_1 & a^T \\ 0 & B \end{bmatrix}$$

从而我们有 $a = 0$，并且 $B = B^T$. 从而 B 是一个 $(n-1) \times (n-1)$ 实对称矩阵。由归纳假设，我们可以对 B 进行对角化，即存在 $(n-1) \times (n-1)$ 的正交矩阵 P' 和对角矩阵 Λ' 使得:

$$\Lambda' = (P')^T BP'$$

从而我们有:

$$P^TSP = \begin{bmatrix} \lambda_1 & 0^T \\ 0 & B \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0^T \\ 0 & (P')^T BP' \end{bmatrix} = \begin{bmatrix} 1 & 0^T \\ 0 & P' \end{bmatrix} \begin{bmatrix} \lambda_1 & 0^T \\ 0 & \Lambda' \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (P')^T \end{bmatrix}$$

令

$$M = \begin{bmatrix} 1 & 0^T \\ 0 & P' \end{bmatrix}$$

显然 M 是一个正交矩阵，并且我们有:

$$(PM)^TSPM = M^TP^TSPM = \begin{bmatrix} \lambda_1 & 0^T \\ 0 & \Lambda' \end{bmatrix}$$

从而令 $Q = PM, \Lambda = \begin{bmatrix} \lambda_1 & 0^T \\ 0 & \Lambda' \end{bmatrix}$，我们有:

$$\Lambda = Q^TSQ$$

并且有归纳假设，Λ 上对角线的元素都是 S 的特征值。 $\square$

实对称矩阵的谱分解 这一节的最后再进一步思考下实对称矩阵 S 的对角化，由定理1.7存在正交矩阵 Q 和对角矩阵 Λ 使得：

$$S = Q\Lambda Q^T$$

令 $Q = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix}$，$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$，我们有：

$$S = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

即：实对称矩阵的谱分解 (Spectral Decomposition)将其分解成了n 个秩为 1的矩阵之和。

例 1.8

考察矩阵 $S = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$，其特征多项式为：

$$f(\lambda) = \begin{vmatrix} \lambda - 1 & -2 & -1 \\ -2 & \lambda - 4 & -2 \\ -1 & -2 & \lambda - 1 \end{vmatrix} = \lambda^2(\lambda - 6)$$

从而 S 的特征值为 $\lambda_1 = 6, \lambda_2 = 0, \lambda_3 = 0$，对应的单位特征向量为：

$$\mathbf{q}_1 = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

由定理1.7，我们有：

$$S = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T + \lambda_3 \mathbf{q}_3 \mathbf{q}_3^T = 6 \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

这意味着一个 $n \times n$ 的实对称矩阵可以用 n 个特征值和对应的 n 个正交的特征向量来表示。这也是**奇异值分解 (SVD)** 的一个特例。

2 正定矩阵

这一节讨论的是所有特征值都大于 0 的矩阵。我们给出如下的定义：

定义 2.1 (Positive Definite Matrix)

一个对称矩阵 S 被称为是正定矩阵，如果其所有的特征值 λ 都满足 $\lambda > 0$。

如下的矩阵:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, B = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}, C = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

1. A 的特征值为 $\lambda_1 = 1, \lambda_2 = 4$.
2. B 的特征值为 $\lambda_1 = 7, \lambda_2 = 1$.
3. C 的特征值为 $\lambda_1 = 7, \lambda_2 = 1, \lambda_3 = 2$.

从而 A, B, C 都是正定矩阵。

下述定理给出了正定矩阵的另一个重要刻画:

定理 2.2

令 S 是一个实对称矩阵，则 S 是正定的，当且仅当对于所有的 $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$，都有:

$$\boldsymbol{x}^T S \boldsymbol{x} > 0.$$

二次型 在证明定理2.2之前，我们先搞清楚 $\boldsymbol{x}^T S \boldsymbol{x}$ 是什么。令 $\boldsymbol{x} = (x_1, \cdots, x_n)$, $S = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{bmatrix}$,

则我们有:

$$\boldsymbol{x}^T S \boldsymbol{x} = s_{11}x_1^2 + \cdots + s_{nn}x_n^2 + 2 \sum_{1 \leqslant i \neq j \leqslant n} s_{ij}x_i x_j,$$

即 $\boldsymbol{x}^T A \boldsymbol{x}$ 是关于 $x_1, \cdots, x_n$ 的一个二次齐次函数，我们称其为二次型 (Quadratic Forms):

定义 2.3 (Quadratic Forms)

含有 n 个变量 $x_1, \cdots, x_n$ 的二次齐次函数

$$f(x_1, \cdots, x_n) = s_{11}x_1^2 + \cdots + s_{nn}x_n^2 + \sum_{1 \leqslant i < j \leqslant n} a_{ij}x_i x_j$$

被称作二次型 (Quadratic Forms)，特别的，令 $a_{ij} = 2s_{ij} = 2s_{ji}$，则上述式子中的每一项 $a_{ij}x_i x_j$ 可以被改写为:

$$a_{ij}x_i x_j = s_{ij}x_i x_j + s_{ji}x_i x_j$$

再令 $S = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{n1} & \cdots & s_{nn} \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$，从而 f 可以有如下的矩阵表示:

$$f(x_1, \cdots, x_n) = \boldsymbol{x}^T S \boldsymbol{x}$$

对称矩阵 S 的表达是唯一的，所以称 S 就是 f 的二次型矩阵 (Quadratic Forms Matrix)。

注 2.4

由上述定义可知，$n \times n$ 的**对称矩阵**和 n 个变量的二次型形成了一一对应的关系，所以我们经常回混用这两个概念，如我们称一个二次型 f 的秩表示的其对应的矩阵 S 的秩。

证明：[定理2.2的证明] 先证 ($\Rightarrow$) 方向。由于 S 是一个实对称矩阵，从而存在一个正交矩阵 Q 满足：

$$\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = Q^T S Q$$

这里 $\lambda_1, \cdots, \lambda_n$ 是 S 的特征值，并且由假设：$\lambda_i > 0$.

定义：

$$\mathbf{x} = Q^T \mathbf{y}, \text{ 即：} \mathbf{y} = Q\mathbf{x}$$

从而：

$$\mathbf{x}^T S \mathbf{x} = (Q\mathbf{y})^T S (Q\mathbf{y}) = \mathbf{y}^T Q^T S Q \mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

注意到 Q 是可逆的，从而对于任意 $\mathbf{x} \neq \mathbf{0}$，我们有 $\mathbf{y} = Q\mathbf{x} \neq \mathbf{0}$，即：

$$\mathbf{x}^T S \mathbf{x} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 > 0$$

另一方面，反设存在 $\lambda_i \leqslant 0$，注意到：

$$\mathbf{x}^T S \mathbf{x} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = \lambda_i y_i^2 + \sum_{j \neq i} \lambda_j y_j^2$$

从而令 $y_i = 1$，$y_j = 0 (j \neq 0)$，由矩阵的可逆性我们有 $\mathbf{x} = Q\mathbf{y} \neq \mathbf{0}$，即：

$$\mathbf{x}^T S \mathbf{x} = \lambda_i y_i^2 + \sum_{j \neq i} \lambda_j y_j^2 = \lambda_j \leqslant 0$$

与假设矛盾。 □

定理2.2给出了正定矩阵的一个刻画，对应的二次型**恒大于** 0，自然也可以将其推广到任意的实矩阵上：

定义 2.5

给定一个 $n \times n$ 的实矩阵A，如果对任意向量 $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$：

1. $\mathbf{x}^T A \mathbf{x} > 0$，则称 A 是正定的。
2. $\mathbf{x}^T A \mathbf{x} \geqslant 0$，则称 A 是半正定的。
3. $\mathbf{x}^T A \mathbf{x} < 0$，则称 A 是负定的。

4. $\boldsymbol{x}^{\mathrm{T}} A \boldsymbol{x} \leqslant 0$，则称 A 是半负定的。
5. 若不满足以上任何一种条件，则称矩阵 A 是不定的。

注 2.6

定义2.5给出了正定矩阵在实矩阵上的一个定义，但在本课程中，**后续讨论的还是实对称矩阵的正定性**，其原因有二：

1. 如果不加对称这一条件，矩阵和二次型便没有了一一对应的关系。
2. 对于非对称矩阵来说，即使其所有的特征值大于 0，它也不一定是正定的。考虑下列的例子：

$$\begin{bmatrix} 1 & 100 \\ 0 & 2 \end{bmatrix}.$$

不难验证，其特征值为 $1, 2$。但是我们有：

$$\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 100 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 1 + 2 - 100 = -97 < 0.$$

下述定理给出了正定矩阵的一个充分性条件：

定理 2.7

令 A 是 $m \times n$ 的矩阵，并且 $\mathrm{rank}(A) = n$. 则 $S = A^{\mathrm{T}} A$ 是对称且正定的。

证明： 令 $\boldsymbol{x} \in \mathbb{R}^n$，则我们有：

$$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}} A^{\mathrm{T}} A \boldsymbol{x} = (A\boldsymbol{x})^{\mathrm{T}} (A\boldsymbol{x}) = A\boldsymbol{x} \cdot A\boldsymbol{x} = \|A\boldsymbol{x}\|^2$$

从而我们有：$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} \geqslant 0$ 并且：

$$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = 0 \iff \|A\boldsymbol{x}\| = 0 \implies A\boldsymbol{x} = \boldsymbol{0}$$

由于 $\mathrm{rank}(A) = n$，从而：

$$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = 0 \implies A\boldsymbol{x} = \boldsymbol{0} \iff \boldsymbol{x} = \boldsymbol{0}$$

□

二次型的转化　在定理 2.2 的证明中可以看到，通过对参数做可逆的线性变换，可以简化对应的二次型，比如在定理证明中：

1. 对应的二次型为：$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = s_{11} x_1^2 + \cdots + s_{nn} x_n^2 + 2\sum_{1 \leqslant i \neq j \leqslant n} s_{ij} x_i x_j$。
2. 如果令 $\boldsymbol{x} = Q^{\mathrm{T}} \boldsymbol{y}$，则二次型可表示为：

$$\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2.$$

正式的来说，称一个二次型 f:

1. **标准二次型**，如果其只包含平方项，即：

$$f = s_{11}x_1^2 + \cdots + s_{nn}x_n^2.$$

2. **规范二次型**，如果其只包含系数为 1 或 -1 的平方项，即：

$$f = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2.$$

一般的，通过一个可逆的线性变换可以将一个二次型进行转换，即，令

$$f = \boldsymbol{x}^T A \boldsymbol{x}$$

则可以通过一个可逆逆矩阵 C, 使得 $\boldsymbol{x} = C\boldsymbol{y}$，并且：

$$f = \boldsymbol{y}^T C^T A C \boldsymbol{y} = \boldsymbol{y}^T B \boldsymbol{y}$$

我们称满足这样一个关系的矩阵 A, B 是合同的。

定义 2.8

令 A, B 是 $n \times n$ 的矩阵，若存在可逆矩阵 C 使得：

$$B = C^T A C$$

则称 A 和 B 是合同的。

注意到无论是规范二次型还是标准二次型，其对应的矩阵 Λ 都是一个对角矩阵，从而由定理1.7，寻找一个二次型的标准型或规范型的问题其实就是计算实对称矩阵的谱分解，总结如下：

1. 计算实对称矩阵 S 的谱分解，即寻找一个正交矩阵 Q 和对角矩阵 Λ 使得：

$$S = Q \Lambda Q^T.$$

2. 令 $\boldsymbol{y} = Q^T \boldsymbol{y}$，则：

$$\boldsymbol{x}^T S \boldsymbol{x} = \boldsymbol{x}^T Q \Lambda Q^T \boldsymbol{x} = \left(Q^T \boldsymbol{x}\right)^T \Lambda \left(Q^T \boldsymbol{x}\right) = \boldsymbol{y}^T \Lambda \boldsymbol{y}.$$

   得到了对应的标准型。

3. 再将 $\boldsymbol{y}$ 作归一化，即可将标准型进一步转化成规范型。

下面通过一个例子来进行说明。以下列二次函数为例：

$$f(x_1, x_2, x_3) = -2x_1x_2 + 2x_1x_3 + 2x_2x_3$$

其对应的对称矩阵为：

$$S = \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

将其进行谱分解，则存在正交矩阵 P 和对角矩阵 Λ 使得

$$\begin{bmatrix}-2&0&0\\0&1&0\\0&0&1\end{bmatrix}=\Lambda=P^TSP=\begin{bmatrix}-\frac{1}{\sqrt3}&-\frac{1}{\sqrt3}&\frac{1}{\sqrt3}\\-\frac{1}{\sqrt2}&\frac{1}{\sqrt2}&0\\\frac{1}{\sqrt6}&\frac{1}{\sqrt6}&\frac{2}{\sqrt6}\end{bmatrix}\begin{bmatrix}0&-1&1\\-1&0&1\\1&1&0\end{bmatrix}\begin{bmatrix}-\frac{1}{\sqrt3}&-\frac{1}{\sqrt2}&\frac{1}{\sqrt6}\\-\frac{1}{\sqrt3}&\frac{1}{\sqrt2}&\frac{1}{\sqrt6}\\\frac{1}{\sqrt3}&0&\frac{2}{\sqrt6}\end{bmatrix}$$

从而令:

$$\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}-\frac{1}{\sqrt3}&-\frac{1}{\sqrt2}&\frac{1}{\sqrt6}\\-\frac{1}{\sqrt3}&\frac{1}{\sqrt2}&\frac{1}{\sqrt6}\\\frac{1}{\sqrt3}&0&\frac{2}{\sqrt6}\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

我们有:

$$f=-2y_1^2+y_2^2+y_3^2$$

如果要进一步变成规范型，我们只需要在令:

$$\begin{cases}y_1=\frac{1}{\sqrt2}z_1\\y_2=z_2\\y_3=z_3\end{cases}$$

即有:

$$f=-z_1^2+z_2^2+z_3^2.‘’$$

另一个二次型的转化方法-配方法 我们再介绍另一个方法，配方法。它不依赖于矩阵的运算。比如考虑:

$$f=x_1^2+2x_2^2+5x_3^2+2x_1x_2+2x_1x_3+6x_2x_3$$

我们考虑先将所有 x_1 的项集中起来:

$$x_1^2+2x_1x_2+2x_1x_3$$

对其配方可得:

$$x_1^2+2x_1x_2+2x_1x_3=(x_1+x_2+x_3)^2-x_2^2-x_3^2-2x_2x_3$$

从而:

$$f=(x_1+x_2+x_3)^2+x_2^2+4x_3^2+4x_2x_3=(x_1+x_2+x_3)^2+(x_2+x_3)^2$$

$$f=(x_1+x_2+x_3)^2+x_2^2+4x_3^2+4x_2x_3=(x_1+x_2+x_3)^2+(x_2+2x_3)^2$$

我们令

$$\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}=\begin{bmatrix}1&1&1\\0&1&2\\0&0&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix},\text{即:}\quad\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1&-1&1\\0&1&-2\\0&0&1\end{bmatrix}\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}$$

则我们有:

$$f=\boldsymbol{x}^T\begin{bmatrix}1&1&1\\1&2&3\\1&3&5\end{bmatrix}\boldsymbol{x}=\boldsymbol{y}^T\begin{bmatrix}1&0&0\\-1&1&0\\1&-2&1\end{bmatrix}\begin{bmatrix}1&1&1\\1&2&3\\1&3&5\end{bmatrix}\begin{bmatrix}1&-1&1\\0&1&-2\\0&0&1\end{bmatrix}\boldsymbol{y}=\boldsymbol{y}^T\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\boldsymbol{y}$$

需要注意的是，<u>**该方法并不一定满足系数矩阵是正交的**</u>。

如果只有类似 x_1x_2 的项，比如：

$$f = 2x_1x_2 + 2x_12x_3 - 6x_1x_3$$

我们也是可以用配方法的，只要注意到：

$$4x_1x_2 = (x_1 + x_2)^2 - (x_1 - x_2)^2$$

从而进行

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

代入再按之前配方的思路化简即可。

正定性与行列式的关系 这一节的最后，我们再给出正定性的另一个刻画，即用行列式来描述，下述定理给出了一个必要的条件：

定理 2.9

如果 S 是正定矩阵，则 $\det(S) > 0$.

证明： 由 S 是实对称的，从而存在正交矩阵 Q 使得：

$$\begin{bmatrix} \lambda_1 \cdots & 0 & \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = Q^TSQ$$

这里 $\lambda_1, \cdots, \lambda_n$ 是 S 的特征值。从而：

$$\begin{vmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{vmatrix} = \det(Q^T)\det(S)\det(Q)$$

即：

$$\det(S) = \det(Q^T)\det(S)\det(Q) = \lambda_1 \cdots \lambda_n > 0$$

□

定理2.9的逆命题并不一定正确。考察下列矩阵：

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix}$$

其行列式 $= 4 > 0$，但是我们有：

$$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = -1 < 0$$

事实上，为了充分的利用行列式来刻画正定性，我们需要引入一系列新的行列式，称之为顺序主子式 (Leading Principal Minor)：

定义 2.10（顺序主子式 (Leading Principal Minors)）

给定一个 $n \times n$ 的矩阵：

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

则对任意的 $k \in [n]$。定义下列行列式：

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}$$

是 A 的一个 k 阶顺序主子式 (Leading Principal Minors).

例 2.11

令 $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$，则：

$$\Delta_1 = 1,\ \Delta_2 = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = 1 \cdot 5 - 2 \cdot 4 = -3,\ \Delta_3 = \det(A) = 0.$$

直观来讲，当顺序主子式都 > 0 时，其对应的特征值满足：

$$\lambda_k = \frac{\Delta_k}{\Delta_{k-1}}$$

这里可以不妨假设 $\Delta_0 = 1$，从而蕴含了所有特征值都 > 0 的条件，即对应的矩阵是正定的，下面对其进行严格的证明：

定理 2.12

给定一个 $n \times n$ 的对称矩阵 S，S 是正定的当且仅当所有的顺序主子式 $\Delta_k > 0$.

证明： 令

$$S_k = \begin{bmatrix} s_{11} & \cdots & s_{1k} \\ \vdots & \ddots & \vdots \\ s_{k1} & \cdots & s_{kk} \end{bmatrix}$$

则我们有:$\Delta_k = \det(S_k)$. 我们先证 $(\Leftarrow)$ 方向,即说明每个 S_k 都是正定的。令 $\boldsymbol{x} = (x_1, \cdots, x_k) \in \mathbb{R}^k \backslash \{\boldsymbol{0}\}$, 注意到:

$$\begin{bmatrix} x_1 & \cdots & x_k \end{bmatrix} S_k \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_k & 0 & \cdots & 0 \end{bmatrix} S \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} > 0$$

从而 S_k 正定，因此 $\Delta_k = \det(S_k) > 0$.

另一个方向我们使用归纳法。

BASE: $n = 1$ 时是显然的。

INDUCTION: 现在考虑 $n \geqslant 2$ 的时候，我们有对于任意的 $k \in [n]$:

$$\Delta_k = \begin{vmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ s_{21} & s_{22} & \cdots & s_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{k1} & s_{k2} & \cdots & s_{kk} \end{vmatrix} = \begin{vmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ 0 & s_{22} - \frac{s_{12}s_{21}}{s_{11}} & \cdots & s_{2k} - \frac{s_{1k}s_{21}}{s_{11}} \\ 0 & \vdots & \ddots & \vdots \\ 0 & s_{k2} - \frac{s_{1k}s_{12}}{s_{11}} & \cdots & s_{kk} - \frac{s_{1k}s_{k1}}{s_{11}} \end{vmatrix}$$

对任意的 $2 \leqslant i, j \leqslant k$ 定义:

$$t_{ij} = s_{ij} - \frac{s_{1i}s_{1j}}{s_{11}} (= s_{ij} - \frac{s_{1i}s_{j1}}{s_{11}})$$

从而我们有:

$$\Delta_k = \begin{vmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ 0 & t_{22} & \cdots & t_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & t_{k2} & \cdots & t_{kk} \end{vmatrix} = s_{11} \begin{vmatrix} t_{22} & \cdots & t_{2k} \\ \vdots & \ddots & \vdots \\ t_{k2} & \cdots & t_{kk} \end{vmatrix}$$

注意到 $s_{11} = \Delta_1 > 0$，$\Delta_k > 0$，从而我们有:

$$\begin{vmatrix} t_{22} & \cdots & t_{2k} \\ \vdots & \ddots & \vdots \\ t_{k2} & \cdots & t_{kk} \end{vmatrix} > 0$$

对任意的 $2 \leqslant k \leqslant n$ 是成立的。定义 $(n-1) \times (n-1)$ 的矩阵 T:

$$T = \begin{bmatrix} t_{22} & \cdots & t_{2k} \\ \vdots & \ddots & \vdots \\ t_{k2} & \cdots & t_{kk} \end{bmatrix}$$

显然 T 是一个实对称矩阵，并且其所有的顺序主子式都是大于 0 的，从而有归纳假设 T 是正定的。从

而:

$$\begin{aligned}\mathbf{x}^{\mathrm{T}}S\mathbf{x} &= \sum_{i\in[n]}\sum_{j\in[n]} s_{ij}x_ix_j \\ &= \frac{(s_{11}x_1+\cdots+s_{1n}x_n)^2}{s_{11}} + \sum_{i=2}^{n}\sum_{j=2}^{n} t_{ij}x_ix_j \\ &= \frac{(s_{11}x_1+\cdots+s_{1n}x_n)^2}{s_{11}} + \begin{bmatrix} x_2 & \cdots & x_n \end{bmatrix} T \begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix}\end{aligned}$$

最后我们分两种情况讨论:

1. $\begin{bmatrix} x_2 & \cdots & x_n \end{bmatrix}^{\mathrm{T}} \neq \mathbf{0}$, 则由 T 是正定的, 我们有:

$$\mathbf{x}^{\mathrm{T}}S\mathbf{x} \geqslant \begin{bmatrix} x_2 & \cdots & x_n \end{bmatrix} T \begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix} > 0$$

2. $\begin{bmatrix} x_2 & \cdots & x_n \end{bmatrix}^{\mathrm{T}} = \mathbf{0}$, 这意味着 $x_1 \neq 0$, 注意到 $s_{11} > 0$, 我们有:

$$\mathbf{x}^{\mathrm{T}}S\mathbf{x} = \frac{(s_{11}x_1)^2}{s_{11}} = s_{11}x_1^2 > 0$$

□

定理2.12还有相应的判断矩阵是否负定的版本:

定理 2.13

令 S 是 $n \times n$ 的对称矩阵, 则 S 是负定的当且仅当所有的奇数阶顺序主子式 < 0, 偶数阶顺序主子式 > 0.