上海师范大学 Shanghai Normal University

《线性代数》

13-对称矩阵和正定矩阵 (Symmetric Matrices and Positive Definite Matrices)

杨启哲

上海师范大学信机学院计算机系

2025 年 5 月 11 日

定义 1 **[特征值和特征向量].**

令 A 是一个 $n\times n$ 的矩阵，$\lambda\in\mathbb{C}$，$\boldsymbol{x}\in\mathbb{C}^n\setminus\{0\}$。如果:

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

则称 λ 是 A 的特征值 (Eigenvalue)，$\boldsymbol{x}$ 是 λ 对应的特征向量 (Eigenvector)。

定义 2 **[特征多项式 (Characteristic Polynomial)].**

令 A 是 $n\times n$ 的矩阵，则称 $f(\lambda)=\det(\lambda E-A)$ 是 A 的特征多项式 (Characteristic Polynomial).

计算特征值的方法

1. 计算 A 的特征多项式，即 $\lambda E - A$ 的行列式 $\det(\lambda E - A)$.
2. 计算 $\det(\lambda E - A)$ 的根，我们一共会得到 n 个特征值（可能重复）。这使得 $\lambda E - A$ 变成一个奇异矩阵 (singular).
3. 对每个 λ，通过解方程 $(\lambda E - A)\boldsymbol{x} = 0$ 来获取 λ 对应的特征向量 $\boldsymbol{x}$.

定理 3 **[Fundamental Theorem of Algebra].**

对于任一的 $a_1, \cdots, a_n \in \mathbb{C}$，我们有

$$x^n + a_1 x^{n-1} + \cdots a_{n-1} x + a_n = (x - b_1)(x - b_2) \cdots (x - b_n)$$

这里 $b_1, \cdots b_n \in \mathbb{C}$，即任一单变元的 n 次复系数多项式恰好有 n 个复数根（可重复）。

推论 4.

令 A 是一个 $n \times n$ 的矩阵，则在计算重根的意义下，A 恰好有 n 个 $\mathbb{C}$ 中的特征值。

定义 5 **[Diagonalization].**

一个方阵 A 是可对角化的 (Diagonalizable)，如果其存在一个可逆矩阵 X 和对角矩阵 Λ 使得:

$$A = X\Lambda X^{-1}, \text{或者等价的 } \Lambda = X^{-1}AX$$

推论 6.

若 $A = X\Lambda X^{-1}$，其中 $X = \text{diag}(\lambda_1, \cdots, \lambda_n)$，则对于任意的 $k \geqslant 1$ 我们有:

$$A^k = X\Lambda^k X^{-1} = X \begin{bmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{bmatrix} X^{-1}$$

定理 7.

$n \times n$ 的矩阵可对角化当且仅当 A 有n 个线性无关的特征向量。

定理 8.

令 $\mathbf{x}_1, \cdots, \mathbf{x}_k$ 是 A 的 k 个特征向量，并且其对应的特征值两两不相同。则 $\mathbf{x}_1, \cdots, \mathbf{x}_k$ 是线性无关的。

推论 9.

如果 $n \times n$ 的矩阵 A 有 n 个不同的特征值，则 A 是可对角化的。

主要内容

- 对称矩阵

- 正定矩阵

对称矩阵

我们称一个方阵是对称 (Symmetric)的，如果其满足:

$$A = A^{\mathrm{T}}$$

例 10.

下述矩阵都是对称的:

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 6 \\ 3 & 5 & 1 & 7 \\ 4 & 6 & 7 & 8 \end{bmatrix}$$

对称矩阵的对角化?

让我们思考一下如果一个对称矩阵 S 可以对角化会发生什么? 假设其可以对角化为:

$$S = X\Lambda X^{-1}$$

那我们有:

$$S^T = (X\Lambda X^{-1})^T = (X^T)^{-1}\Lambda^T X^T = (X^T)^{-1}\Lambda X^T = S = X\Lambda X^{-1}$$

一个理想的状况是 $X^T = X^{-1}$, 即$X^TX = E$。事实上也正是如此, 我们将证明对于对称矩阵:

1. 特征值是实数。
2. 不同特征值的特征向量是正交的。

特征值是实数!

我们将首先证明对于对称矩阵 S，其特征值是实数。

定理 11.

所有实对称矩阵的特征值都是实数。

我们还可以证明一个更强的版本:

定理 12.

所有实对称矩阵的特征向量是实数，并且每个特征值都有一个对应的实特征向量。

复数的一些复习

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:

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

引理 13.

给定复数 $x, y \in \mathbb{C}$，我们有:

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

定义 14 **[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：

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

引理 15.

令 A 是一个 $m \times n$ 的复矩阵：

1. 对任意的 $\lambda \in \mathbb{C}$，我们有：$\overline{\lambda A} = \bar{\lambda}\bar{A}$.
2. 对任意的复向量 $\boldsymbol{x} \in \mathbb{C}^n$, 我们有：$\overline{A\boldsymbol{x}} = \bar{A}\bar{\boldsymbol{x}}$.

引理 16.

令 $\boldsymbol{x} \in \mathbb{C}^n$，如果 $\bar{\boldsymbol{x}}^T\boldsymbol{x} = 0$, 则 $\boldsymbol{x} = \boldsymbol{0}$.

特征值是实数的证明

证明：[定理12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。我们有：

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

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

特征向量是实向量的证明

证明：[定理12的证明续] 假设 $\mathbf{x}_j = a_j + b_j i$，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。

定理 17.

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： 由假设我们有：

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

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。

由上述定理我们可以直接得到:

推论 18.

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。

而事实上，我们并不需要“n 个两两不同的特征值”这个条件。

任何一个实对称矩阵都可以对角化 (II)

上海师范大学 Shanghai Normal University

定理 19.

令 S 是一个 $n \times n$ 的实对称矩阵，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。显然这是一个当且仅当的关系，因为另一个方向是显然成立的。

定理19的证明 (I)

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：

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

定理19的证明 (II)

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 $P^TP = E$。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。

定理19的证明 (III)

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：

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

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

$$\Lambda = Q^{T}SQ$$

并且有归纳假设，Λ 上对角线的元素都是 S 的特征值。 $\square$

对称矩阵的谱分解

给定实对称矩阵 S，由上述定理，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。

我们讨论了 $n \times n$ 的实对称矩阵 S。

- 其所有的特征值都是实数，并且都有着对应的实特征向量。
- 不同特征值对应的特征向量都是正交的。
- 任何一个实对称矩阵都可以被对角化，并且其谱分解可以写成 n 个秩为 1 的矩阵之和，即：

$$S = Q^T \Lambda Q = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

正定矩阵

在很多应用中，所有特征值是正的对称矩阵起到了非常大的作用。

定义 20 **[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 $\lambda > 0$。

例 21.

$$\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

定理 22.

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 $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$，都有：

$$\mathbf{x}^T S \mathbf{x} > 0$$

$\mathbf{x}^T S \mathbf{x}$ 是什么？

令 $\mathbf{x} = (x_1, \cdots, x_n)$，$S = \left[s_{ij}\right]_{n \times n}$，则我们有：

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

即 $\mathbf{x}^T A \mathbf{x}$ 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)。

一般地，对于 $x_1, \cdots, x_n$ 的一个二次齐次函数：

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

我们总可以将其转换成如下的矩阵形式：

$$f(x_1, \cdots, x_n) = \boldsymbol{x}^T S\boldsymbol{x}, \text{ 其中：} \boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \ S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{12} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ s_{1n} & s_{2n} & \cdots & s_{nn} \end{bmatrix}$$

这意味着我们总可以用一个实对称矩阵来表示一个关于 $x_1, \cdots, x_n$ 的一个二次齐次函数。特别的，正定矩阵就是指的是那些恒大于 0 的二次型。

定理22的一个应用

我们先来看一个简单的应用。

定理 23.

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

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

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

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

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

由于 $\operatorname{rank}(A) = n$，从而：

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

我们先证 (⇒) 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:

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

从而:

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

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

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

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

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

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

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

与假设矛盾。 $\square$

正定性的定义

通过上述的讨论，我们可以把正定的概念扩充到任何一个实矩阵上去。

定义 24.

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

我们需要注意的是，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：

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

二次型的转化

通过上述的证明，可以看到，对参数作一些变换可以使得对应的二次函数变得简单：

- $\boldsymbol{x}^{\mathrm{T}} S \boldsymbol{x} = s_{11} x_1^2 + \cdots + s_{nn} x_n^2 + 2 \sum_{1 \leqslant i \neq j \leqslant n} s_{ij} x_i x_j$
- 令 $\boldsymbol{x} = Q^{\mathrm{T}} \boldsymbol{y}$ 后我们有：

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

我们将后者这种只含平方项的二次型为标准二次型，特别的，如果其系数进一步化简为 $1, 1, \cdots, 1, -1, \cdots, -1$，我们称其为规范二次型。

上海师范大学 Shanghai Normal University

定理22的证明已经给了我们一种转化二次型的方法，即找对应矩阵的谱分解。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

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \Lambda = P^TSP = \begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix}$$

二次型的转化-对角化 (II)

从而令:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

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:

$$\begin{cases} y_1 = \frac{1}{\sqrt{2}} z_1 \\ y_2 = z_2 \\ y_3 = z_3 \end{cases}$$

即有:

$$f = -z_1^2 + z_2^2 + z_3^2$$

二次型的转化-配方法 (I)

上海师范大学 Shanghai Normal University

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:

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

二次型的转化-配方法 (II)

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

我们令

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

则我们有:

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

需要注意的是，该方法并不一定满足系数矩阵是正交的。

二次型的转化-配方法 (II)

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。

可以看到，通过一些可逆的系数变换，我们可以将一个二次型进行转换，即，令

$$f = x^T A x$$

则可以通过一个可逆逆矩阵 C, 使得 $x = Cy$，并且:

$$f = y^T C^T A C y = y^T B y$$

我们称满足这样一个关系的矩阵 A, B 是合同的。

定义 25.

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。

正定性和行列式的关系

现在我们来看一下正定性和行列式的关系。

定理 26.

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：

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

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：

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

行列式大于 0 的对称矩阵不一定正定

上述定理的逆命题并不一定正确。比如考察下列矩阵:

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

其行列式 $=4>0$，但是我们有:

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

定义 27 [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).

定理 28.

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 $\Delta_k > 0$.

直观理解

直观上来讲，当每个顺序主子式都大于 0 时，我们有:

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

这里可以不妨假设 $\Delta_0 = 1$.

定理28的证明 (I)

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 $\mathbf{x} = (x_1, \cdots, x_k) \in \mathbb{R}^k \setminus \{0\}$，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 $\Delta_k = \det(S_k) > 0$.

定理28的证明 (II)

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:

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

上海师范大学 Shanghai Normal University

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:

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

定理28的证明 (V)

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:

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

$\square$

我们讨论了矩阵一个性质，正定性。

正定矩阵的性质

令 S 是一个对称矩阵，

1. S 是正定的。
2. S 的特征值都大于 0.
3. S 的顺序主子式都大于 0。
4. 存在一个矩阵 A，使得 $S = A^{T}A$.

我们还讨论了二次型的概念，以及变换成标准二次型的方法。

- 利用对称矩阵的谱分解。
- 配方法。