主成份分析的线性代数基础

王宁宁

广州医科大学

线性代数基础

Cauchy-Schwarz 不等式

Cauchy-Schwarz 不等式

设 $\boldsymbol{b}$ 和 $\boldsymbol{d}$ 是两个任意 $p \times 1$ 向量，有：

$$(\boldsymbol{b}'\boldsymbol{d}) \leq (\boldsymbol{b}'\boldsymbol{b})(\boldsymbol{d}'\boldsymbol{d})$$

则当且仅当对某个常数 c 有 $\boldsymbol{b} = c\boldsymbol{d}$ 时，等号成立。

证明

若 $\boldsymbol{b}$ 和 $\boldsymbol{d}$ 有一个为零，则结果显然成立。

$$
\begin{aligned}
0 &\le (\boldsymbol{b} - x\boldsymbol{d})'(\boldsymbol{b} - x\boldsymbol{d}) \\
&= \boldsymbol{b}'\boldsymbol{b} - x\boldsymbol{d}'\boldsymbol{b} - \boldsymbol{b}'(x\boldsymbol{d}) + x^2\boldsymbol{d}'\boldsymbol{d} \\
&= \boldsymbol{b}'\boldsymbol{b} - 2x\boldsymbol{b}'\boldsymbol{d} + x^2\boldsymbol{d}'\boldsymbol{d} \qquad \left(\text{加减}\frac{(\boldsymbol{b}'\boldsymbol{d})^2}{\boldsymbol{d}'\boldsymbol{d}}\right) \\
&= \boldsymbol{b}'\boldsymbol{b} - \frac{(\boldsymbol{b}'\boldsymbol{d})^2}{\boldsymbol{d}'\boldsymbol{d}} + \frac{(\boldsymbol{b}'\boldsymbol{d})^2}{\boldsymbol{d}'\boldsymbol{d}} - 2x\boldsymbol{b}'\boldsymbol{d} + x^2\boldsymbol{d}'\boldsymbol{d} \\
&= \boldsymbol{b}'\boldsymbol{b} - \frac{(\boldsymbol{b}'\boldsymbol{d})^2}{\boldsymbol{d}'\boldsymbol{d}} + (\boldsymbol{d}'\boldsymbol{d})\left(x - \frac{\boldsymbol{b}'\boldsymbol{d}}{\boldsymbol{d}'\boldsymbol{d}}\right)^2
\end{aligned}
$$

取 $x = \frac{\boldsymbol{b}'\boldsymbol{d}}{\boldsymbol{d}'\boldsymbol{d}}$ 得证。等号显然当对某个常数 c 有 $\boldsymbol{b} = c\boldsymbol{d}$ 时成立

矩阵的谱分解

矩阵的谱分解

设 $\boldsymbol{B}_{p\times p}$ 是任意一个正定矩阵，有：

$$\boldsymbol{B} = \lambda_1 \boldsymbol{e}_1 \boldsymbol{e}_1' + \lambda_2 \boldsymbol{e}_2 \boldsymbol{e}_2' + \cdots + \lambda_p \boldsymbol{e}_p \boldsymbol{e}_p'$$

其中 $\lambda_i, \boldsymbol{e}_i$，是 $\boldsymbol{B}$ 的第 i 个特征根及相应的特征向量。

矩阵的谱分解

矩阵的谱分解

设 $\boldsymbol{B}_{p\times p}$ 是任意一个正定矩阵，有：

$$\boldsymbol{B}=\lambda_1\boldsymbol{e}_1\boldsymbol{e}_1{}'+\lambda_2\boldsymbol{e}_2\boldsymbol{e}_2{}'+\cdots+\lambda_p\boldsymbol{e}_p\boldsymbol{e}_p{}'$$

其中 $\lambda_i,\boldsymbol{e}_i$，是 $\boldsymbol{B}$ 的第 i 个特征根及相应的特征向量。

推论

$$\boldsymbol{B}^{\frac{1}{2}}=\sqrt{\lambda_1}\boldsymbol{e}_1\boldsymbol{e}_1{}'+\sqrt{\lambda_2}\boldsymbol{e}_2\boldsymbol{e}_2{}'+\cdots+\sqrt{\lambda_p}\boldsymbol{e}_p\boldsymbol{e}_p{}'$$

$$\boldsymbol{B}^{-\frac{1}{2}}=\frac{1}{\sqrt{\lambda_1}}\boldsymbol{e}_1\boldsymbol{e}_1{}'+\frac{1}{\sqrt{\lambda_2}}\boldsymbol{e}_2\boldsymbol{e}_2{}'+\cdots+\frac{1}{\sqrt{\lambda_p}}\boldsymbol{e}_p\boldsymbol{e}_p{}'$$

广义 Cauchy-Schwarz 不等式

广义 Cauchy-Schwarz 不等式

设 $\boldsymbol{b}$ 和 $\boldsymbol{d}$ 是两个任意 $p \times 1$ 向量，$\boldsymbol{B}_{p\times p}$ 是正定矩阵，有：

$$(\boldsymbol{b}'\boldsymbol{d}) \leq (\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})(\boldsymbol{d}'\boldsymbol{B}^{-1}\boldsymbol{d})$$

则当且仅当对某个常数 c 有 $\boldsymbol{b} = c\boldsymbol{B}^{-1}\boldsymbol{d}$ 时，等号成立。

广义 Cauchy-Schwarz 不等式

广义 Cauchy-Schwarz 不等式

设 $\boldsymbol{b}$ 和 $\boldsymbol{d}$ 是两个任意 $p \times 1$ 向量，$\boldsymbol{B}_{p\times p}$ 是正定矩阵，有：

$$(\boldsymbol{b}'\boldsymbol{d}) \leq (\boldsymbol{b}'\boldsymbol{B}\boldsymbol{b})(\boldsymbol{d}'\boldsymbol{B}^{-1}\boldsymbol{d})$$

则当且仅当对某个常数 c 有 $\boldsymbol{b} = c\boldsymbol{B}^{-1}\boldsymbol{d}$ 时，等号成立。

证明

只需注意到：

$$\boldsymbol{b}'\boldsymbol{d} = \boldsymbol{b}'\boldsymbol{I}\boldsymbol{d} = \boldsymbol{b}'\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{d} = (\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{b})'(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{d})$$

用 $\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{b}, \boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{d}$ 替换 $\boldsymbol{b}, \boldsymbol{d}$ 利用 Cauchy-Schwarz 不等式即得。

二次型极大化

二次型极大化

设 $\boldsymbol{B}_{p\times p}$ 是任意一个正定矩阵，$\boldsymbol{d}_{p\times 1}$ 是任意给定的向量，对任意 $\boldsymbol{x}_{p\times 1}$ 有：

$$\max \frac{(\boldsymbol{x}'\boldsymbol{d})^2}{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}} \leq \boldsymbol{b}'\boldsymbol{B}^{-1}\boldsymbol{d}$$

则当且仅当对某个常数 c 有 $\boldsymbol{x} = c\boldsymbol{B}^{-1}\boldsymbol{d}$ 时，等号成立。

二次型极大化

设 $\boldsymbol{B}_{p\times p}$ 是任意一个正定矩阵，$\boldsymbol{d}_{p\times 1}$ 是任意给定的向量，对任意 $\boldsymbol{x}_{p\times 1}$ 有：

$$\max \frac{(\boldsymbol{x}'\boldsymbol{d})^2}{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}} \leq \boldsymbol{b}'\boldsymbol{B}^{-1}\boldsymbol{d}$$

则当且仅当对某个常数 c 有 $\boldsymbol{x} = c\boldsymbol{B}^{-1}\boldsymbol{d}$ 时，等号成立。

证明

利用广义 Cauchy-Schwarz 不等式有:

$$(\boldsymbol{x}'\boldsymbol{d})^2 \leq (\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x})(\boldsymbol{b}'\boldsymbol{B}^{-1}\boldsymbol{d})$$

利用 $\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}$ 除以上式两边即得结论，而且当 $\boldsymbol{x} = c\boldsymbol{B}^{-1}\boldsymbol{d}$ 时，达到上界。

单位球面上二次型极大化

单位球面上二次型极大化

设 $\boldsymbol{B}_{p\times p}$ 是任意一个正定矩阵，其特征值为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$，相应的特征向量为 $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_p$，对任意单位向量 $\boldsymbol{x}_{p\times 1}$ 有：

$$\max_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = \lambda_1 \quad \text{当 } \boldsymbol{x} = \boldsymbol{e}_1 \text{ 时达到}$$

$$\min_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = \lambda_p \quad \text{当 } \boldsymbol{x} = \boldsymbol{e}_p \text{ 时达到}$$

特别的：

$$\max_{\boldsymbol{x}\perp e_1,\cdots,e_k} \frac{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = \lambda_{k+1} \quad \text{当 } \boldsymbol{x} = \boldsymbol{e}_{k+1} \text{ 时达到}$$

证明

只证第一式：设 $\boldsymbol{B}^{\frac{1}{2}} = \boldsymbol{P}\Lambda\boldsymbol{P}', \boldsymbol{y} = \boldsymbol{P}'\boldsymbol{x}$，其中 $\boldsymbol{P}$ 是正交阵

$$\max_{\boldsymbol{x}\neq 0} \frac{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} = \frac{\boldsymbol{x}'\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{P}\boldsymbol{P}'\boldsymbol{x}} = \frac{\boldsymbol{x}'\boldsymbol{P}\Lambda\boldsymbol{P}'\boldsymbol{P}\Lambda\boldsymbol{P}'\boldsymbol{x}}{\boldsymbol{y}'\boldsymbol{y}}$$

$$= \frac{\boldsymbol{y}'\Lambda\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{y}} = \frac{\sum_{i=1}^{p} \lambda_i y_i^2}{\sum_{i=1}^{p} y_i^2} \leq \lambda_1 \frac{\sum_{i=1}^{p} y_i^2}{\sum_{i=1}^{p} y_i^2} = \lambda_1$$

当 $\boldsymbol{x} = \boldsymbol{e}_1$ 时：

$$\frac{\boldsymbol{e}_1'\boldsymbol{B}\boldsymbol{e}_1}{\boldsymbol{e}_1'\boldsymbol{e}_1} = \boldsymbol{e}_1'\boldsymbol{B}\boldsymbol{e}_1 = \lambda_1$$