线性系统理论补充 2：有关对称矩阵的一个结论

2024 年 11 月 21 日

定理：特征值与二次型取值的关系（Rayleigh-Ritz 不等式）

设 $A = A^{\mathrm{T}} \in \mathbf{R}^{n\times n}$，并设 $\lambda_{\max}$ 和 $\lambda_{\min}$ 分别是 A 的最大特征值和最小特征值。则有

$$\lambda_{\max} = \max_{x\in\mathbf{R}^n, x\neq 0} \frac{x^{\mathrm{T}}Ax}{\|x\|_2^2} \tag{1}$$

$$\lambda_{\min} = \min_{x\in\mathbf{R}^n, x\neq 0} \frac{x^{\mathrm{T}}Ax}{\|x\|_2^2} \tag{2}$$

证. 由于 A 为实对称矩阵，所以 A 可以正交相似对角化，即存在 R 满足 $R^{\mathrm{T}}R = I$，以及一对角矩阵 $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$（各对角线元素即为 A 的特征值，其中可能有重复），使得

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

因此

$$x^{\mathrm{T}}Ax = x^{\mathrm{T}}R^{\mathrm{T}}\Lambda Rx = (Rx)^{\mathrm{T}}\Lambda Rx$$

又注意到

$$\|x\|_2^2 = x^{\mathrm{T}}x = x^{\mathrm{T}}Ix = x^{\mathrm{T}}R^{\mathrm{T}}Rx = (Rx)^{\mathrm{T}}Rx = \|Rx\|_2^2$$

令 $y = Rx = (y_1, \dots, y_n)^{\mathrm{T}}$，则

$$\frac{x^{\mathrm{T}}Ax}{\|x\|_2^2} = \frac{y^{\mathrm{T}}\Lambda y}{\|y\|_2^2} = \frac{\sum_{i=1}^n \lambda_i y_i^2}{\sum_{i=1}^n y_i^2}$$

将上式中各 λ_i 都以 $\lambda_{\max}$ 代之，则

$$\frac{x^{\mathrm{T}}Ax}{\|x\|_2^2} \leq \frac{\lambda_{\max}\sum_{i=1}^n y_i^2}{\sum_{i=1}^n y_i^2} = \lambda_{\max}$$

并注意到取 x 为对应于最大特征值的特征向量时，上式可取得等号，即证明了 (1)；类似可证明 (2)。

注 1. 结论中的这种比值称为**瑞利商（Rayleigh quotient）**。

注 2. 很容易将该结论推广到 n 阶 Hermite 矩阵上（即 $A = A^{\mathrm{H}} \in \mathbf{C}^{n\times n}$），只要将结论中的所有转置换成共轭转置，再将证明中的正交相似对角化换成

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

即可（其中 U 为幺正矩阵）。