协方差矩阵的性质

davidc

2015

随机向量 $\mathbf{X} \in \mathbf{R}^n$ 与均值向量 $\mathbf{m}_x$ 的协方差矩阵定义如下：

$$\mathbf{C}_x = E\left[(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^T\right]$$

协方差矩阵 $\mathbf{C}_x$ 的第 $(i,j)^{\text{th}}$ 个元素由下式给出

$$C_{ij} = E\left[\left(X_i - m_i\right)\left(X_j - m_j\right)\right] = \sigma_{ij}$$

该协方差矩阵 $\mathbf{C}_x$ 的对角线项是随机向量 $\mathbf{X}$ 的分量的方差，即，

$$C_{ii} = E\left[\left(X_i - m_i\right)^2\right] = \sigma_i^2$$

由于对角线项都是正的，所以协方差矩阵的‘迹’是正的，即，

$$\text{Trace}\left(\mathbf{C}_x\right) = \sum_{i=1}^{n} C_{ii} > 0$$

协方差矩阵 $\mathbf{C}_x$ 是对称的，即 $\mathbf{C}_x = \mathbf{C}_x^T$，因为：

$$C_{ij} = \sigma_{ij} = \sigma_{ji} = C_{ji}$$

协方差矩阵 $\mathbf{C}_x$ 是半正定的，即对于 $\mathbf{a} \in \mathbf{R}^n$：

$$\begin{aligned} E\left\{\left[(\mathbf{X}-\mathbf{m})^T\mathbf{a}\right]^2\right\} = E\left\{\left[(\mathbf{X}-\mathbf{m})^T\mathbf{a}\right]^T\left[(\mathbf{X}-\mathbf{m})^T\mathbf{a}\right]\right\} \geq 0 \\ E\left[\mathbf{a}^T(\mathbf{X}-\mathbf{m})(\mathbf{X}-\mathbf{m})^T\mathbf{a}\right] \geq 0, \quad \mathbf{a} \in \mathbf{R}^n \\ \mathbf{a}^T\mathbf{C}_x\mathbf{a} \geq 0, \quad \mathbf{a} \in \mathbf{R}^n \end{aligned}$$

由于协方差矩阵 $\mathbf{C}_x$ 是对称的，即与通常的内积自伴，其特征值都是实的和正的，属于不同特征值的特征向量是正交的，即，

$$\mathbf{C}_x = \mathbf{V}\ \mathbf{V}^T = \sum_{i=1}^{n} \lambda_i \vec{v}_i \vec{v}_i^T$$

因此，协方差矩阵的行列式为正，即，

$$\det\left(\mathbf{C}_X\right) = \prod_{i=1}^{n} \lambda_i \geq 0$$

协方差矩阵的特征向量将随机向量变换成统计上不相关的随机变量，即变换成具有对角协方差矩阵的随机向量。协方差矩阵的瑞利系数 (Rayleigh coefficient) 上下以最大和最小特征值为界：

$$\lambda_{\min} \leq \tfrac{\mathbf{a}^T \mathbf{C}_x \mathbf{a}}{\mathbf{a}^T \mathbf{a}}, \quad \mathbf{a} \in \mathbf{R} \leq \lambda_{\max}$$