

● 均值和协方差矩阵的估计量定义

设模式的类概率密度函数为 $p(\mathbf{x})$ ，则其均值向量定义为：

$$\mathbf{m} = E(\mathbf{x}) = \int_{\mathbf{x}} \mathbf{x} p(\mathbf{x}) d\mathbf{x}$$

其中，样本 \mathbf{x} 和均值向量 \mathbf{m} 为 n 维向量，即 $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $\mathbf{m} = (m_1, m_2, \dots, m_n)^T$ 。

若以样本的平均值作为均值向量的近似值，则均值估计量 $\hat{\mathbf{m}}$ 为：

$$\hat{\mathbf{m}} = \frac{1}{N} \sum_{j=1}^N \mathbf{x}^j$$

其中 N 为样本的数目， \mathbf{x}^j 表示第 j 个样本。

协方差矩阵为：

$$\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} & \text{L} & c_{1n} \\ c_{21} & c_{22} & \text{L} & c_{2n} \\ \text{M} & \text{M} & & \text{M} \\ c_{n1} & c_{n2} & \text{L} & c_{nn} \end{pmatrix}$$

其每个元素 c_{lk} 定义为：

$$\begin{aligned} c_{lk} &= E\{(x_l - m_l)(x_k - m_k)\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_l - m_l)(x_k - m_k) p(x_l, x_k) dx_l dx_k \end{aligned}$$

其中， x_l 、 x_k 和 m_l 、 m_k 分别为 \mathbf{x} 和 \mathbf{m} 的第 l 和 k 个分量。

协方差矩阵写成向量形式为：

$$\mathbf{C} = E\{(\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})^T\} = E\{\mathbf{x}\mathbf{x}^T\} - \mathbf{m}\mathbf{m}^T \quad (\text{注：用乘法分}$$

配律展开 $E\{\mathbf{x}\mathbf{m}^T\} = E\{\mathbf{x}\}\mathbf{m}^T = \mathbf{m}\mathbf{m}^T, E\{\mathbf{m}\mathbf{x}^T\} = \mathbf{m}E\{\mathbf{x}^T\} = \mathbf{m}\mathbf{m}^T)$

协方差矩阵的估计量（当 $N \gg 1$ 时）为：

$$\hat{\mathbf{C}} \approx \frac{1}{N} \sum_{j=1}^N (\mathbf{x}^j - \hat{\mathbf{m}})(\mathbf{x}^j - \hat{\mathbf{m}})^T$$

这里，样本模式总体为 $\{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^k, \dots, \mathbf{x}^N\}$ 。因为计算估计量时没有真实的均值向量 \mathbf{m} 可用，只能用均值向量的估计量 $\hat{\mathbf{m}}$ 来代替，**会存在偏差**。