- 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 $p(\boldsymbol{x})$，则其均值向量定义为：

$$\boldsymbol{m} = \mathrm{E}(\boldsymbol{x}) = \int_x x p(\boldsymbol{x}) d\boldsymbol{x}$$

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

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

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

其中 N 为样本的数目，$\boldsymbol{x}^j$ 表示第 j 个样本。

协方差矩阵为：

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

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 $\boldsymbol{x}$ 和 $\boldsymbol{m}$ 的第 l 和 k 个分量。

协方差矩阵写成向量形式为：

$\boldsymbol{C} = E\{(\boldsymbol{x}-\boldsymbol{m})(\boldsymbol{x}-\boldsymbol{m})^T\} = E\{\boldsymbol{x}\boldsymbol{x}^T\} - \boldsymbol{m}\boldsymbol{m}^T$（注：用乘法分配率展开 $E\{\boldsymbol{x}\boldsymbol{m}^T\} = E\{\boldsymbol{x}\}\boldsymbol{m}^T = \boldsymbol{m}\boldsymbol{m}^T, E\{\boldsymbol{m}\boldsymbol{x}^T\} = \boldsymbol{m}E\{\boldsymbol{x}^T\} = \boldsymbol{m}\boldsymbol{m}^T$）

协方差矩阵的估计量（当 N>>1 时）为：

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

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