維基百科
自由的百科全書

多元正态分布

维基百科，自由的百科全书

多变量正态分布亦称为**多变量高斯分布**。它是单维正态分布向多维的推广。它同矩阵正态分布有紧密的联系。

一般形式

N维随机向量 $X = [X_1, \dots, X_N]^T$如果服从多变量正态分布，必须满足下面的三个等價条件：

1. 任何线性组合 $Y = a_1 X_1 + \cdots + a_N X_N$服从正态分布。
2. 存在随机向量 $Z = [Z_1, \dots, Z_M]^T$(它的每个元素服从独立标准正态分布），向量 $\mu = [\mu_1, \dots, \mu_N]^T$及$N \times M$ 矩阵 A满足 $X = AZ + \mu$.
3. 存在μ和一个对称半正定阵 Σ满足 X的特征函数

$$\phi_X\left(u;\mu,\Sigma\right) = \mathrm{e}^{i\mu^T u - \frac{1}{2} u^T \Sigma u}$$

如果 Σ是非奇异的，那么该分布可以由以下的概率密度函数来描述：[1]

$$f_{\mathbf{x}}(x_1, \dots, x_k) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \mathrm{e}^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})},$$

注意这里的$|\Sigma|$表示协方差矩阵的行列式。

二元的情况

在二维非奇异的情况下（$k = \mathrm{rank}(\Sigma) = 2$），向量 $[X\ Y]'$ 的概率密度函数为：

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \mathrm{e}^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]}$$

其中 ρ 是 X 与 Y 之间的相关系数，$\sigma_X > 0$ 且 $\sigma_Y > 0$。在这种情况下，

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}.$$

参考文献

1. UIUC, Lecture 21. *The Multivariate Normal Distribution* (http://www.math.uiuc.edu/~r-ash/Stat/StatLec21-25.pdf) （页面存档备份 (https://web.archive.org/web/20160623194512/http://www.math.uiuc.edu/~r-ash/Stat/StatLec21-25.pdf)，存于互联网档案馆）, 21.5:"Finding the Density".