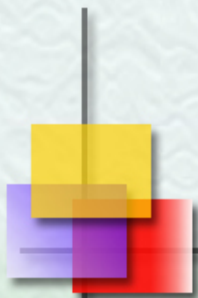


第四节 矩、协方差矩阵

一、基本概念

二、 n 维正态变量的性质

三、小结



一、基本概念

1. 定义

设 X 和 Y 是随机变量, 若 $E(X^k)$, $k = 1, 2, \dots$ 存在, 称它为 X 的 k 阶原点矩, 简称 k 阶矩.

若 $E\{[X - E(X)]^k\}$, $k = 2, 3, \dots$ 存在, 称它为 X 的 k 阶中心矩.

若 $E(X^k Y^l)$, $k, l = 1, 2, \dots$ 存在, 称它为 X 和 Y 的 $k + l$ 阶混合矩.

若 $E\{[X - E(X)]^k [Y - E(Y)]^l\}$, $k, l = 1, 2, \dots$ 存在, 称它为 X 和 Y 的 $k + l$ 阶混合中心矩.



2. 说明

- (1) 以上数字特征都是随机变量函数的数学期望;
- (2) 随机变量 X 的数学期望 $E(X)$ 是 X 的一阶原点矩, 方差为二阶中心矩, 协方差 $\text{Cov}(X, Y)$ 是 X 与 Y 的二阶混合中心矩;
- (3) 在实际应用中, 高于 4 阶的矩很少使用.

三阶中心矩 $E\{[X - E(X)]^3\}$ 主要用来衡量随机变量的分布是否有偏.

四阶中心矩 $E\{[X - E(X)]^4\}$ 主要用来衡量随机变量的分布在均值附近的陡峭程度如何.



3. 协方差矩阵

设 n 维随机变量 (X_1, X_2, \dots, X_n) 的二阶混合中心矩

$$c_{ij} = \text{Cov}(X_i, X_j) = E\{[X_i - E(X_i)][X_j - E(X_j)]\}$$
$$i, j = 1, 2, \dots, n$$

都存在, 则称矩阵

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

为 n 维随机变量的协方差矩阵.



例如 二维随机变量 (X_1, X_2) 的协方差矩阵为

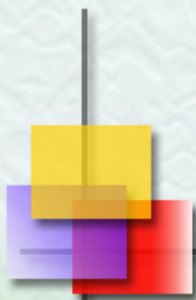
$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$

其中 $c_{11} = E\{[X_1 - E(X_1)]^2\},$

$$c_{12} = E\{[X_1 - E(X_1)][X_2 - E(X_2)]\},$$

$$c_{21} = E\{[X_2 - E(X_2)][X_1 - E(X_1)]\},$$

$$c_{22} = E\{[X_2 - E(X_2)]^2\}.$$



由于 $c_{ij} = c_{ji}$ ($i, j = 1, 2, \dots, n$), 所以协方差矩阵为对称的非负定矩阵.

协方差矩阵的应用

协方差矩阵可用来表示多维随机变量的概率密度, 从而可通过协方差矩阵达到对多维随机变量的研究



以二维随机变量 (X_1, X_2) 为例.

由于

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}.$$

引入矩阵 $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$.

及 (X_1, X_2) 的协方差矩阵 $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$,



$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix},$$

由此可得

$$\begin{aligned} C^{-1} &= \frac{1}{\det C} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \\ &= \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}. \end{aligned}$$



由于

$$\begin{aligned}
 & (X - \mu)^T C^{-1} (X - \mu) \\
 &= \frac{1}{\det C} (x_1 - \mu_1, x_2 - \mu_2) \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix} \\
 &= \frac{1}{1 - \rho^2} \left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right].
 \end{aligned}$$



于是 (X_1, X_2) 的概率密度可写成

$$f(x_1, x_2)$$

$$= \frac{1}{(2\pi)^{2/2} (\det C)^{1/2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\}.$$



推广

n 维随机变量 (X_1, X_2, \dots, X_n) 的概率密度可表示为 $f(x_1, x_2, \dots, x_n)$

$$= \frac{1}{(2\pi)^{n/2} (\det C)^{1/2}} \exp \left\{ -\frac{1}{2} (X - \mu)^T C^{-1} (X - \mu) \right\}.$$

其中 $X = (x_1, x_2, \dots, x_n)^T$,

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_n) \end{pmatrix}, \quad C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}.$$



二、 n 维正态变量的性质

1. n 维随机变量 (X_1, X_2, \dots, X_n) 的每一个分量 $X_i, i = 1, 2, \dots, n$ 都是正态变量;

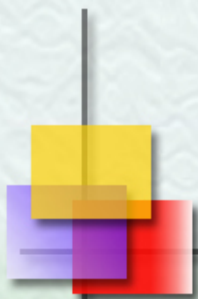
反之, 若 X_1, X_2, \dots, X_n 都是正态变量, 且相互独立, 则 (X_1, X_2, \dots, X_n) 是 n 维正态变量.

2. n 维随机变量 (X_1, X_2, \dots, X_n) 服从 n 维正态分布的充要条件是 X_1, X_2, \dots, X_n 的任意的线性组合 $l_1 X_1 + l_2 X_2 + \dots + l_n X_n$ 服从一维正态分布 (其中 l_1, l_2, \dots, l_n 不全为零).



3.若 (X_1, X_2, \dots, X_n) 服从 n 维正态分布, 设 Y_1, \dots, Y_k 是 $X_j (j = 1, 2, \dots, n)$ 的线性函数, 则 (Y_1, Y_2, \dots, Y_k) 也服从多维正态分布. **线性变换不变性**

4.设 (X_1, \dots, X_n) 服从 n 维正态分布, 则“ X_1, X_2, \dots, X_n 相互独立”与“ X_1, X_2, \dots, X_n 两两不相关”是等价的.



三、小结

1. 矩是随机变量的数字特征.

随机变量 X 的

{ 数学期望 $E(X)$ 是 X 的一阶原点矩;
方差 $D(X)$ 是 X 的二阶中心矩;
协方差 $\text{Cov}(X, Y)$ 是 X 与 Y 的二阶混合中心矩.

2. 正态变量是最重要的随机变量, 其性质一定要熟练掌握.

