第四节　矩、协方差矩阵

- 原点矩　中心矩
- 协方差矩阵
- n 元正态分布的概率密度
- 小结　布置作业

一、 原点矩　中心矩

定义　设X是随机变量，若 $E(X^k), k=1,2,\cdots$ 存在，称它为**X的k阶原点矩**，简称 **k阶矩**

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

- 均值 **E(X)** 是 **X** 的一阶原点矩
- 方差 **D(X)** 是 **X** 的二阶中心矩

【应用】矩估计法；三阶中心距用来衡量**RV**的分布是否有偏，四阶中心距用来衡量**RV**的分布在均值附近的陡峭程度

设 X 和 Y 是随机变量，若

$$E(X^kY^l) \quad k,l=1,2,\ldots$$ 存在，

称它为 X 和 Y 的 $k+l$ 阶混合（原点）矩.

若 $E\{[X-E(X)]^k[Y-E(Y)]^l\}$ 存在，

称它为X 和 Y 的 $k+l$ 阶混合中心矩.

- 协方差$Cov(X,Y)$是X和Y的二阶混合中心矩

二、协方差矩阵

将二维随机变量（X_1,X_2）的四个二阶中心矩

$$c_{11}=E\{[X_1-E(X_1)]^2\}$$

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

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

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

排成矩阵的形式: $\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$

对称矩阵

称此矩阵为（X_1,X_2）的协方差矩阵.

例　若 $D(X)=1, D(Y)=4, \rho_{XY}=-1/4$,

求$(X+Y, 3X-Y)$的协方差矩阵.

解: $\mathrm{Cov}(X,Y)=\rho_{XY}\sqrt{D(X)}\sqrt{D(Y)}=-\dfrac{1}{2}$

$$D(X+Y)=D(X)+D(Y)+2\mathrm{Cov}(X,Y)=4$$

$$D(3X-Y)=9D(X)+D(Y)-6\mathrm{Cov}(X,Y)=16$$

$$\begin{aligned}&\mathrm{Cov}(X+Y,3X-Y)\\&=E[X+Y-E(X)-E(Y)]\,[3X-Y-3E(X)+E(Y)]\\&=E\{([X-E(X)]+[Y-E(Y)])\\&\qquad(3[X-E(X)]-[Y-E(Y)])\}\\&=3D(X)+2\mathrm{Cov}(X,Y)-D(Y)=-2\end{aligned}$$

（$X+Y, 3X-Y$)协方差矩阵$C=\begin{pmatrix}4 & -2\\ -2 & 16\end{pmatrix}$

类似定义n 维随机变量$(X_1,X_2,\dots,X_n)$ 的协方差矩阵.

若 $c_{ij}=Cov(X_i,X_j)=E\{[X_i-E(X_i)][X_j-E(X_j)]\}$

$(i,j=1,2,\dots,n)$

都存在, 称 矩阵

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

对称矩阵

半正定矩阵

为$(X_1,X_2,\dots,X_n)$ 的协方差矩阵

【应用】用来表示多维随机变量的概率密度，可通过研究协方差矩阵来研究多维随机变量。

以二维正态分布变量 (X_1, X_2) 的概率密度函数为例.

由于 $f(x_1,x_2)=\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}$

$$\boldsymbol{f}(\boldsymbol{x}_1,\boldsymbol{x}_2)=\frac{\mathbf{1}}{\mathbf{2}\pi|\mathbf{C}|}\exp\{-\frac{\mathbf{1}}{\mathbf{2}}(\boldsymbol{X}-\mu)^{T}\boldsymbol{C}^{-1}(\boldsymbol{X}-\mu)\}$$

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

上式花括号内的式子可写成如下矩阵形式:

引入矩阵 $X=\begin{pmatrix}x_1\\x_2\end{pmatrix}$, $\mu=\begin{pmatrix}\mu_1\\\mu_2\end{pmatrix}$ 及 $(\boldsymbol{X}_1,\boldsymbol{X}_2)$ 的协方差矩阵

$$C=\begin{pmatrix}c_{11}&c_{12}\\c_{21}&c_{22}\end{pmatrix}=\begin{pmatrix}\sigma_1^2&\rho\sigma_1\sigma_2\\\rho\sigma_1\sigma_2&\sigma_2^2\end{pmatrix},$$ 其逆矩阵 $$\mathrm{C}^{-1}=\frac{\boldsymbol{C}^*}{|\boldsymbol{C}|}=\frac{1}{|\boldsymbol{C}|}\begin{pmatrix}\boldsymbol{A}_{11}&\boldsymbol{A}_{21}\\\boldsymbol{A}_{12}&\boldsymbol{A}_{22}\end{pmatrix}$$

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

$$(X-\mu)^{\mathrm{T}}C^{-1}(X-\mu)=\frac{1}{|C|}(x_1-\mu_1,x_2-\mu_2)\begin{pmatrix}\sigma_2^2&-\rho\sigma_1\sigma_2\\-\rho\sigma_1\sigma_2&\sigma_1^2\end{pmatrix}\begin{pmatrix}x_1-\mu_1\\x_2-\mu_2\end{pmatrix}$$

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

n 维正态分布的概率密度函数

设 $X^{\mathrm{T}}=(X_1,X_2,\ldots,X_n)$ 是一个n维随机向量，若它的概率密度为

$$f(x_1,x_2,\ \ldots,x_n)=\frac{1}{(2\pi)^{n/2}\,|C|^{1/2}}\exp\{-\frac{1}{2}(X-\mu)^T C^{-1}(X-\mu)\}$$

则称 X 服从 n 维正态分布.

其中，$X=(x_1,x_2,\cdots,x_n)^T$,

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

$|C|$是它的行列式，C^{-1} 表示C的逆矩阵，

三、n元正态分布变量的性质

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

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

2. n 维随机变量 $(X_1, X_2, \cdots, X_n)$服从 n 维正态分布的充要条件是 $X_1, X_2, \cdots, X_n$ 的任意的线性组合

$l_1X_1 + l_2X_2 + \cdots + l_nX_n$ 服从一维正态分布

(其中 $l_1, l_2, \cdots, l_n$ 不全为零).

3. 正态变量的线性变换不变性.

若 $X=(X_1, X_2, \ldots, X_n)$ 服从n维正态分布，

$Y_1,Y_2,\ldots,Y_k$是X_j（$j=1,2,\ldots,n$）的线性函数，

则 $(Y_1,Y_2,\ldots, Y_k)$ 也服从多元正态分布.

4. 设$(X_1,X_2,\ldots,X_n)$服从n维正态分布，则

“$X_1,X_2,\ldots,X_n$相互独立”

等价于

“$X_1,X_2,\ldots,X_n$两两不相关”

例　设随机变量X和Y相互独立且$X\sim N(1,2)$, $Y\sim N(0,1)$. 试求$Z=2X-Y+3$的概率密度.

解: $X\sim N(1,2), Y\sim N(0,1)$，且 X 与 Y 独立,

故X 和Y 的联合分布为正态分布，X 和Y 的任意线性组合是正态分布.

即　$Z\sim N(E(Z), D(Z))$

$E(Z)=E(2X-Y+3)=2E(X)-E(Y)+3=2+3=5$

$D(Z)=D(2X-Y+3)=4D(X)+D(Y)=8+1=9$

$Z\sim N(5, 3^2)$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

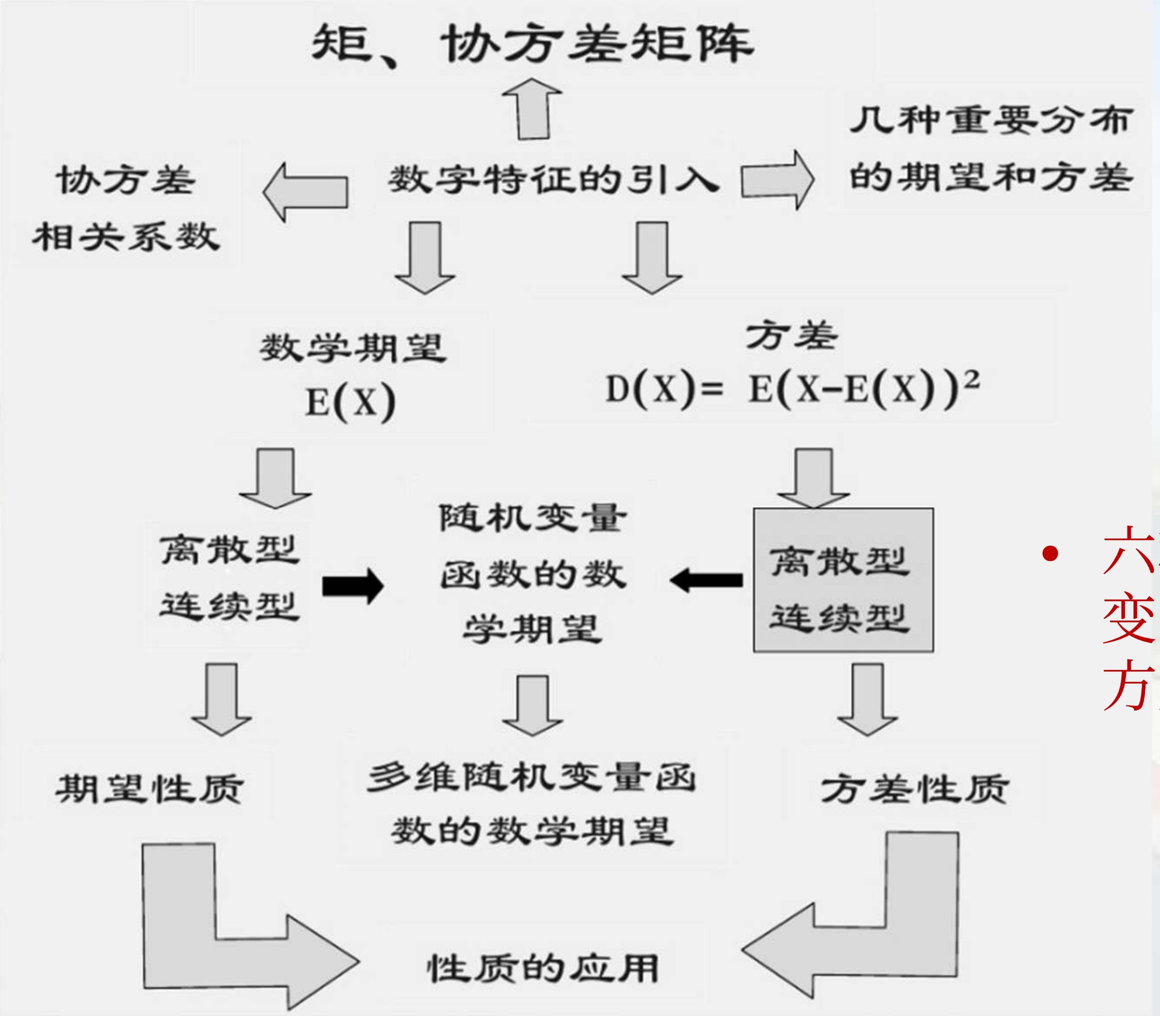
故 Z 的概率密度是

$$f_Z(z)=\frac{1}{3\sqrt{2\pi}}e^{-\frac{(z-5)^2}{18}}, \quad -\infty<z<\infty$$

四、小结

在这一节中我们学习了随机变量的原点矩和中心矩以及协方差矩阵.

一般地,高维随机变量的分布是不知道的,或者太复杂,以至于在数学上不易处理,因此在实际中协方差矩阵就显得重要了.



- 六种常用随机变量的期望与方差