# 概率论与数理统计

## 第二十七讲

## 矩、协方差矩阵与多维正态分布

# 第27讲 矩、协方差矩阵与多维正态分布

**定义** 设$X,Y$为随机变量且各阶矩都存在，$k>0, l>0$.

(1) 称$E(X^k)$为$X$的$k$ **阶原点矩**（$k$**阶矩**）

(2) 称$E[(X-E(X))^k]$为$X$的$k$ **阶中心矩**

(3) 称$E(X^kY^l)$为$X,Y$ 的$k+l$ **阶混合矩**

(4) 称$E[(X-E(X))^k(Y-E(Y))^l]$为 $X,Y$ 的$k+l$ **阶混合中心矩**.

- 判断 $E(X)$ …… 1阶原点矩

  $D(X)$ …… 2阶中心矩

  $\mathrm{Cov}(X,Y)$ …… 2阶混合中心矩

- “矩”是来自于物理学中力矩的概念

## 结论

$$\because \ |X|^k \le 1 + |X|^{k+1}$$

$$\therefore \ E(|X|^k) \le 1 + E(|X|^{k+1})$$

这说明高阶矩存在,则低阶矩必存在.

# 第27讲 矩、协方差矩阵与多维正态分布

对于二维随机向量 $(X_1,X_2)$ ,记

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

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

$$c_{22}=E[(X_2-E(X_2))(X_1-E(X_1))]=\mathrm{Cov}(X_2,X_1)$$

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

$$C\triangleq\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

**定义** 称方阵 $C$ 为随机向量 $(X_1,X_2)$ 的**协方差矩阵**. (**协方差阵、协差阵**).

# 第27讲 矩、协方差矩阵与多维正态分布

对于$n$维随机向量 $\boldsymbol{X}=(X_1,X_2\cdots,X_n)$, 记

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

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

定义 称方阵 $C$为随机向量 $\boldsymbol{X}$ 的**协方差矩阵**.

**(协方差 阵、协差阵)**.

## 协方差矩阵的性质

(1) $C^T = C$, 即协方差阵为对称阵

(2) $C \geq 0$ , 即协方差阵为非负定阵

证 (1) 显然.

(2) 设 $\boldsymbol{t} = (t_1, t_2 \cdots, t_n)$为任意实数向量 , 则有

$$\boldsymbol{t} \cdot C \cdot \boldsymbol{t}^T = \sum_{j=1}^{n}\sum_{i=1}^{n} c_{ij} t_i t_j = \sum_{j=1}^{n}\sum_{i=1}^{n} E[(X_i - E(X_i))(X_j - E(X_j))] t_i t_j$$

$$= E[\sum_{j=1}^{n}\sum_{i=1}^{n} t_i (X_i - E(X_i)) \cdot t_j (X_j - E(X_j))]$$

$$= E[\sum_{i=1}^{n} t_i (X_i - E(X_i))]^2 \geq 0.$$

故 $\boldsymbol{t} \cdot C \cdot \boldsymbol{t}^T$ 为非负定二次型, 即 $C$ 为非负定阵.

## 二维正态随机向量的相关系数与协方差阵

设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，密度函数为

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\{-\frac{1}{2(1-\rho^2)}$$

$$\times[\boxed{\frac{(x-\mu_1)^2}{\sigma_1^2}}-2\rho\boxed{\frac{(x-\mu_1)}{\sigma_1}}\boxed{\frac{(y-\mu_2)}{\sigma_2}}+\boxed{\frac{(y-\mu_2)^2}{\sigma_2^2}}]\}$$

$$\mathrm{Cov}(X,Y)=E[(X-E(X))(Y-E(Y))]$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)\mathrm{d}x\mathrm{d}y$$

作变换 $\frac{x-\mu_1}{\sigma_1}=u,\ \frac{y-\mu_2}{\sigma_2}=v$

# 第27讲 矩、协方差矩阵与多维正态分布

令 $\dfrac{x-\mu_1}{\sigma_1}=u,\ \dfrac{y-\mu_2}{\sigma_2}=v$, 则有

$$\mathrm{Cov}(X,Y)=\frac{\sigma_1\sigma_2}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}uv\cdot\exp\{-\frac{u^2-2\rho uv+v^2}{2(1-\rho^2)}\}\mathrm{d}u\mathrm{d}v$$

$$=\frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}v\exp\{-\frac{v^2}{2(1-\rho^2)}\}\mathrm{d}v\int_{-\infty}^{\infty}u\cdot\exp\{-\boxed{\frac{u^2-2\rho uv}{2(1-\rho^2)}}\}\mathrm{d}u$$

令 $t=\dfrac{1}{\sqrt{1-\rho^2}}(u-\rho v)$ ,则有

$$\rho_{XY}=\frac{1}{2\pi}\int_{-\infty}^{\infty}v\cdot\exp\{-\frac{v^2}{2}\}\mathrm{d}v\int_{-\infty}^{\infty}(\boxed{\sqrt{1-\rho^2}\cdot t+}\rho v)\cdot\exp\{-\frac{t^2}{2}\}\mathrm{d}t$$

$$=\frac{\rho}{2\pi}\int_{-\infty}^{\infty}v^2\cdot\exp\{-\frac{v^2}{2}\}\mathrm{d}v\cdot\int_{-\infty}^{\infty}\exp\{-\frac{t^2}{2}\}\mathrm{d}t$$

$$=\frac{\rho}{2\pi}\sqrt{2\pi}\cdot\sqrt{2\pi}=\rho$$

# 第27讲 矩、协方差矩阵与多维正态分布

**定理** 设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 则

(1) $X,Y$ 的相关系数 $\rho_{XY}=\rho$

(2) $X,Y$ 的协方差阵为

$$C=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}$$

(3) $X,Y$相互独立 $\Leftrightarrow$ $\rho=0$

$\Leftrightarrow$ $\rho_{XY}=0$

$\Leftrightarrow$ $X,Y$互不相关

## 二维正态随机变量密度函数的矩阵表示法

设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，密度函数为

$$f(x,y)=\frac{1}{2\pi\boxed{\sigma_1\sigma_2\sqrt{1-\rho^2}}}\times$$

$$\exp\{-\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]\}$$

考虑协方差阵

$$C=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix},\quad \boxed{|C|=\sigma_1^2\sigma_2^2(1-\rho^2)}$$

## 二维正态随机变量密度函数的矩阵表示法

设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$，密度函数为

二次型

$$f(x,y)=\frac{1}{2\pi\quad|C|^{1/2}}\times$$

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

考虑协方差阵

$$C=\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix},|C|=\sigma_1^2\sigma_2^2(1-\rho^2)$$

利用伴随矩阵求 $C$ 的逆矩阵

$$C^{-1}=\frac{1}{|C|}\begin{bmatrix}c_{11}^* & c_{12}^*\\ c_{21}^* & c_{22}^*\end{bmatrix}=\frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{bmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{bmatrix}$$

## 二维正态随机变量密度函数的矩阵表示法

设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 密度函数为

$$f(x,y)=\frac{1}{2\pi\quad |C|^{1/2}}\times$$

二次型

$$\exp\{-\boxed{\frac{1}{2(1-\rho^2)}[\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}]}\}$$

记 $\boldsymbol{x}=(x,y),\boldsymbol{\mu}=(\mu_1,\mu_2)$, 则

$$\boxed{\text{二次型}}=\frac{1}{2|C|}[x-\mu_1,y-\mu_2]\begin{bmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{bmatrix}\begin{bmatrix}x-\mu_1\\ y-\mu_2\end{bmatrix}$$

$$=\frac{1}{2}(\boldsymbol{X}-\boldsymbol{\mu})C^{-1}(\boldsymbol{X}-\boldsymbol{\mu})^T$$

# 第27讲 矩、协方差矩阵与多维正态分布

- 二维正态随机变量密度函数的矩阵表示法

设 $(X,Y)\sim N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, 密度函数为

$$f(\boldsymbol{x})=\frac{1}{2\pi|C|^{1/2}}\exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})C^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^T\}$$

其中 $\boldsymbol{x}=(x,y),\boldsymbol{\mu}=(\mu_1,\mu_2),C$为协方差阵.

- 对比一维正态随机变量密度函数

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

- 问题 $n$维正态随机变量密度函数如何表示？

$$f(\boldsymbol{x})=\frac{1}{(2\pi)^{n/2}|C|^{1/2}}\exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})C^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^T\}\ ?$$

**定义** 设 $C$ 为 $n$ 阶正定对称阵, $\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_n)$ 为 $n$ 维已知向量. 记 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in\mathbf{R}^n$. 若 $n$ 维随机向量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 的密度函数为

$$f(\boldsymbol{x})=\frac{1}{(2\pi)^{n/2}|C|^{1/2}}\exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})C^{-1}(\boldsymbol{x}-\boldsymbol{\mu})^T\}$$

则称 $\boldsymbol{X}$ 服从 $n$ 维正态分布, 记为

$$\boldsymbol{X}=(X_1,X_2,\cdots,X_n)\sim N(\boldsymbol{\mu},C)$$

- 可以证明

(1) $f(\boldsymbol{x})>0$

(2) $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}f(x_1,x_2,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n=1$

(证略)

# 第27讲 矩、协方差矩阵与多维正态分布

**定理 ( $n$ 维正态分布的基本性质)**

设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)\sim N(\boldsymbol{\mu},C)$. 则

(1) $\mu_i=E(X_i)\ (i=1,2,\cdots,n)$

(2) $C=[c_{ij}]_{n\times n}$ 是 $(X_1,X_2,\cdots,X_n)$ 的协方差阵,且

$$D(X_i)=c_{ii}\quad(i=1,2,\cdots,n)$$

$$\mathrm{Cov}(X_i,X_j)=c_{ij}\quad(i,j=1,2,\cdots,n)$$

(3) $X_i\sim N(\mu_i,c_{ii})\quad(i=1,2,\cdots,n)$

(4) $X_1,X_2,\cdots,X_n$ 相互独立 $\Longleftrightarrow$ $X_1,X_2,\cdots,X_n$ 两两不相关

$\Longleftrightarrow$ $C=\mathrm{diag}(c_{11},c_{22},\cdots,c_{nn})$

(即 $C$ 为对角阵)

# 第27讲 矩、协方差矩阵与多维正态分布

**定理（$n$ 维正态分布的基本性质）**

设 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)\sim N(\boldsymbol{\mu},C)$. 则

(1) $\mu_i=E(X_i)\ (i=1,2,\cdots,n)$

(2) $C=[c_{ij}]_{n\times n}$ 是 $(X_1,X_2,\cdots,X_n)$ 的协方差阵,且

$$D(X_i)=c_{ii}\quad(i=1,2,\cdots,n)$$

$$\mathrm{Cov}(X_i,X_j)=c_{ij}\quad(i,j=1,2,\cdots,n)$$

(3) $X_i\sim N(\mu_i,c_{ii})\quad(i=1,2,\cdots,n)$

(4) $X_1,X_2,\cdots,X_n$ 相互独立 $\Longleftrightarrow$ $X_1,X_2,\cdots,X_n$ 两两不相关

$\Longleftrightarrow$ $C=\mathrm{diag}(c_{11},c_{22},\cdots,c_{nn})$

(5) 若 $X_1,X_2,\cdots,X_n$ 相互独立, 且各 $X_i\sim N(\mu_i,\sigma_i^2)$, 则

$$(X_1,X_2,\cdots,X_n)\sim N(\boldsymbol{\mu},C)$$

其中 $\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_n),C=\mathrm{diag}(\sigma_1^2,\sigma_2^2,\cdots,\sigma_n^2)$.

**定理（$n$ 维正态分布的性质）**

(1) $(X_1,X_2,\cdots,X_n)\sim N(\boldsymbol{\mu},C) \Longleftrightarrow X_1,X_2,\cdots,X_n$ 的任一非零线性组合 $l_1X_1+l_2X_2+\cdots+l_nX_n$ 服从一维正态分布.

(2) 正态随机向量的线性变换不变性：

若 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)\sim N(\boldsymbol{\mu},C)$，令

$$\begin{cases} Y_1=a_{11}X_1+a_{12}X_2+\cdots+a_{1n}X_n \\ Y_2=a_{21}X_1+a_{22}X_2+\cdots+a_{2n}X_n \\ \quad\vdots \\ Y_m=a_{m1}X_1+a_{m2}X_2+\cdots+a_{mn}X_n \end{cases}$$

则 $\boldsymbol{Y}=(Y_1,Y_2,\cdots,Y_m)$ 仍服从多维正态分布.

# 本讲结束，谢谢大家