# *Multivariate Statistical Analysis*

# 多元统计分析

2026 年 3 月 26 日

# 已学知识点 (Recap)

## 第 4 章 多元分布

### 4.1 分布函数与密度函数

- 累积分布函数 (cdf) 的定义为 $F(x) = \mathbb{P}(X < x)$.
- 如果概率密度函数 (pdf) 存在，则 $F(x) = \int_{-\infty}^{x} f(u)\,\mathrm{d}u$.
- 概率密度函数 (pdf) 满足 $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$.
- 将随机向量 $X$ 分块为 $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$，设 $X_1$ 与 $X_2$ 的联合 (即 $X$ 的) 分布函数为 $F$，则 $X_1$ 的边际分布函数为 $F_{X_1}(x_1) = \mathbb{P}(X_1 < x_1)$，$X_1$ 的边际概率密度函数为 $f_{X_1}(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\,\mathrm{d}x_2$. 对联合概率密度函数求导亦可得到相同的边际概率密度函数.
- 给定 $X_1 = x_1$ 时 $X_2$ 条件概率密度函数为 $f\left(x_2 \middle| x_1\right) = \dfrac{f(x_1, x_2)}{f_{X_1}(x_1)}$.

# 已学知识点 (Recap)

## 第 4 章 多元分布

### 4.1 分布函数与密度函数

- 两个随机向量 $X_1$ 与 $X_2$ 独立，当且仅当 $f(\boldsymbol{x}_1,\boldsymbol{x}_2)=f_{X_1}(\boldsymbol{x}_1)\cdot f_{X_2}(\boldsymbol{x}_2)$，也等价于 $f\left(\boldsymbol{x}_2\middle|\boldsymbol{x}_1\right)=f_{X_2}(\boldsymbol{x}_2)$，$f\left(\boldsymbol{x}_1\middle|\boldsymbol{x}_2\right)=f_{X_1}(\boldsymbol{x}_1)$.
- 不同的联合概率密度函数可以有相同的边际概率密度函数.

# 已学知识点 (Recap)

## 第 4 章 多元分布

### 4.2 矩与特征函数

- 随机向量 $\boldsymbol{X}$ 的期望为 $\boldsymbol{\mu} = \int \boldsymbol{x} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$，协方差矩阵为 $\Sigma = \mathrm{Var}(\boldsymbol{X}) = \mathbb{E}\left[\left(\boldsymbol{X}-\boldsymbol{\mu}\right)\left(\boldsymbol{X}-\boldsymbol{\mu}\right)^{\mathrm{T}}\right]$，将其记为 $\boldsymbol{X} \sim \left(\boldsymbol{\mu}, \Sigma\right)$.
- 求期望是线性运算，即 $\mathbb{E}\left(\alpha\boldsymbol{X}+\beta\boldsymbol{Y}\right) = \alpha\mathbb{E}(\boldsymbol{X}) + \beta\mathbb{E}(\boldsymbol{Y})$. 如果 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 相互独立，则 $\mathbb{E}\left(\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}\right) = \mathbb{E}(\boldsymbol{X})\,\mathbb{E}\left(\boldsymbol{Y}^{\mathrm{T}}\right)$.
- 两个随机向量 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的协方差矩阵为
$$\Sigma_{\boldsymbol{XY}} = \mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y}) = \mathbb{E}\left[(\boldsymbol{X}-\mathbb{E}(\boldsymbol{X}))(\boldsymbol{Y}-\mathbb{E}(\boldsymbol{Y}))^{\mathrm{T}}\right] = \mathbb{E}\left(\boldsymbol{X}\boldsymbol{Y}^{\mathrm{T}}\right) - \mathbb{E}(\boldsymbol{X})\,\mathbb{E}\left(\boldsymbol{Y}^{\mathrm{T}}\right).$$
如果 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 独立，则 $\mathrm{Cov}(\boldsymbol{X},\boldsymbol{Y}) = 0$.
- 条件期望 $\mathbb{E}\left(X_2 \,\middle|\, X_1\right)$ 是用 $X_1$ 的一个函数对 $X_2$ 在均方误差 (MSE) 意义下的最佳近似.

# *Multivariate Distributions*

# 多 元 分 布

# 矩与特征函数 (Moments and Characteristic Functions)

- 特征函数
  - 设随机向量 $\boldsymbol{X} \in \mathbb{R}^p$ 的概率密度函数为 $f(\boldsymbol{x})$，定义 $\boldsymbol{X}$ 的特征函数为

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) \triangleq \mathrm{E}\left(\mathrm{e}^{\mathbf{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{X}}\right) = \int \mathrm{e}^{\mathbf{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{x}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\,, \quad \boldsymbol{t} \in \mathbb{R}^p$$

虚数单位 $\mathbf{i}^2 = -1$

  - 性质：
    - $\varphi_{\boldsymbol{X}}(\mathbf{0}) = 1$
    - $\left|\varphi_{\boldsymbol{X}}(\boldsymbol{t})\right| \le 1$

$$\int_{-\infty}^{\infty} \left|\varphi(\boldsymbol{x})\right| \mathrm{d}\boldsymbol{x} < \infty$$

    - 如果 $\varphi$ 绝对可积，则 $f(\boldsymbol{x}) = \dfrac{1}{(2\pi)^p}\displaystyle\int_{-\infty}^{\infty} \mathrm{e}^{-\mathbf{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{x}}\,\varphi_{\boldsymbol{X}}(\boldsymbol{t})\,\mathrm{d}\boldsymbol{t}$.
    - $$\begin{cases} \varphi_{X_1}(t_1) & = \varphi_{\boldsymbol{X}}(t_1, 0, \ldots, 0) \\ \quad\vdots & \\ \varphi_{X_k}(t_k) & = \varphi_{\boldsymbol{X}}(0, \ldots, 0, t_k, 0, \ldots 0) \\ \quad\vdots & \\ \varphi_{X_p}(t_p) & = \varphi_{\boldsymbol{X}}(0, \ldots, 0, t_p) \end{cases}$$

# 矩与特征函数 (Moments and Characteristic Functions)

- 特征函数
  - 设随机向量 $\boldsymbol{X} \in \mathbb{R}^p$ 的概率密度函数为 $f(\boldsymbol{x})$，定义 $\boldsymbol{X}$ 的特征函数为

虚数单位 $\mathbf{i}^2 = -1$

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) \triangleq \mathrm{E}\left(\mathrm{e}^{\mathbf{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{X}}\right) = \int \mathrm{e}^{\mathbf{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{x}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\,, \quad \boldsymbol{t} \in \mathbb{R}^p$$

  - 性质：
    - 如果 $X_1, X_2, \ldots, X_p$ 相互独立，则对 $\boldsymbol{t} = \left(t_1, t_2, \ldots, t_p\right)^{\mathrm{T}}$，我们有

$$\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \varphi_{X_1}\left(t_1\right) \cdot \varphi_{X_2}\left(t_2\right) \cdot \cdots \cdot \varphi_{X_p}\left(t_p\right)$$

    - 如果 $X_1, X_2, \ldots, X_p$ 相互独立，则对 $t \in \mathbb{R}$，我们有

$$\varphi_{X_1+X_2+\cdots+X_p}(t) = \varphi_{X_1}(t) \cdot \varphi_{X_2}(t) \cdot \cdots \cdot \varphi_{X_p}(t)$$

    - 由特征函数可以得到任意阶的交叉乘积矩 (cross-product moments)： $\forall j_k \geq 0$，$k = 1, 2, \ldots, p$ 以及 $\boldsymbol{t} = \left(t_1, t_2, \ldots, t_p\right)^{\mathrm{T}}$，我们有

$$\mathrm{E}\left(X_1^{j_1} \cdot X_2^{j_2} \cdot \cdots \cdot X_p^{j_p}\right) = \frac{1}{\mathbf{i}^{j_1+j_2+\cdots+j_p}}\left[\frac{\partial\, \varphi_{\boldsymbol{X}}(\boldsymbol{t})}{\partial\, t_1^{j_1}\, \partial\, t_2^{j_2} \cdots \partial\, t_p^{j_p}}\right]_{\boldsymbol{t}=\boldsymbol{0}}$$

# 矩与特征函数 (Moments and Characteristic Functions)

- 例：考虑概率密度函数 $f(x_1, x_2) = \begin{cases} \dfrac{1}{2}x_1 + \dfrac{3}{2}x_2\,, & 0 \le x_1,\, x_2 \le 1 \\ 0\,, & \text{其它} \end{cases}$

$$\varphi_X(\boldsymbol{t}) = \int \mathrm{e}^{\mathbf{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{x}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \qquad \boldsymbol{t} = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathrm{e}^{\mathbf{i}u} = \cos u + \mathbf{i}\sin u$$

$$= \int_0^1\int_0^1 \left[\cos\left(t_1x_1 + t_2x_2\right) + \mathbf{i}\sin\left(t_1x_1 + t_2x_2\right)\right]\left(\frac{1}{2}x_1 + \frac{3}{2}x_3\right)\mathrm{d}x_1\,\mathrm{d}x_2$$

$$= \frac{0.5\mathrm{e}^{\mathbf{i}t_1}\left(3\mathbf{i}t_1 - 3\mathbf{i}\mathrm{e}^{\mathbf{i}t_2}t_1 + \mathbf{i}t_2 - \mathbf{i}\mathrm{e}^{\mathbf{i}t_2}t_2 + t_1t_2 - 4\mathrm{e}^{\mathbf{i}t_2}t_1t_2\right)}{t_1^2t_2^2} - \frac{0.5\left(3\mathbf{i}t_1 - 3\mathbf{i}\mathrm{e}^{\mathbf{i}t_2}t_1 + \mathbf{i}t_2 - \mathbf{i}\mathrm{e}^{\mathbf{i}t_2}t_2 - 3\mathrm{e}^{\mathbf{i}t_2}t_1t_2\right)}{t_1^2t_2^2}$$

# 矩与特征函数 (Moments and Characteristic Functions)

- 例：设 $X \in \mathbb{R}^1$ 服从标准正态分布 $f_X(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$$
\begin{aligned}
\varphi_X(t) &= \int e^{\mathbf{i}tx} f(x)\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\mathbf{i}tx}\, e^{-\frac{x^2}{2}}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(x^2 - 2\mathbf{i}tx + \mathbf{i}^2t^2\right)}\, e^{\frac{1}{2}\mathbf{i}^2t^2}\,\mathrm{d}x \\
&= e^{-\frac{t^2}{2}} \int_{-\infty}^{\infty} \boxed{\frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mathbf{i}t)^2}{2}}}\,\mathrm{d}x \\
&= e^{-\frac{t^2}{2}}
\end{aligned}
$$

$\mathbf{i}^2 = -1$

$N\left(\mathbf{i}t, 1\right)$ 的概率密度函数

# 矩与特征函数 (Moments and Characteristic Functions)

常用分布的特征函数

| 分布 | 密度函数 | 特征函数 |
|---|---|---|
| 均匀分布 | $f(x) = \dfrac{1}{b-a} \cdot \mathrm{I}\left(x \in [a, b]\right)$ | $\varphi_X(t) = \dfrac{\mathrm{e}^{\mathbf{i}bt} - \mathrm{e}^{\mathbf{i}at}}{(b-a)\mathbf{i}t}$ |
| $N_1\left(\mu, \sigma^2\right)$ | $f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right]$ | $\varphi_X(t) = \mathrm{e}^{\mathbf{i}\mu t - \frac{\sigma^2 t^2}{2}}$ |
| $\chi_n^2$ | $f(x) = \mathrm{I}(x > 0) \cdot \dfrac{x^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}}{\Gamma\left(\frac{n}{2}\right)\ 2^{\frac{n}{2}}}$ | $\varphi_X(t) = (1 - 2\mathbf{i}t)^{-\frac{n}{2}}$ |
| $N_p(\boldsymbol{\mu}, \Sigma)$ | $f(x) = \left\lvert 2\pi\Sigma \right\rvert^{-\frac{1}{2}} \exp\left[-\dfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma(\boldsymbol{x}-\boldsymbol{\mu})\right]$ | $\varphi_X(\boldsymbol{t}) = \mathrm{e}^{\mathbf{i}\boldsymbol{t}^{\mathrm{T}}\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{t}^{\mathrm{T}}\Sigma\boldsymbol{t}}$ |

**定理 4.4 (Cramer-Wold)** $\boldsymbol{X} \in \mathbb{R}^p$ 的分布由形如 $\boldsymbol{t}^{\mathrm{T}}\boldsymbol{X}$ 的所有 (一维) 分布的集合完全确定，其中 $\boldsymbol{t} \in \mathbb{R}^p$.

线性组合 $\sum_{k=1}^{p} t_k X_k$

# 矩与特征函数 (Moments and Characteristic Functions)

- 累积函数 (cumulant functions)
  - 矩 $m_k = \int \boldsymbol{x}^k f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ 通常有助于我们描述分布的特征.
  - $d = 1$ 维的正态分布可以完全由其标准正态密度函数 $f(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{x^2}{2}}$ 以及矩

    $\underbrace{\mu = m_1}_{\mathrm{E}(X)}$ 和 $\sigma^2 = \underbrace{m_2}_{\mathrm{E}\left(X^2\right)} - m_1^2$ 所描述.

# 矩与特征函数 (Moments and Characteristic Functions)

- 累积函数 (cumulant functions)
  - 假设 $X$ 是一个 (1 维) 随机变量，其密度函数为 $f$，且具有有限的 $k$ 阶矩.
  - 则其特征函数 $\varphi_X(t) = \mathrm{E}\left(\mathrm{e}^{\mathbf{i}tX}\right)$ 有如下各阶导数

$$\frac{1}{\mathbf{i}^j}\left[\frac{\partial^j \log\left\{\varphi_X(t)\right\}}{\partial t^j}\right]_{t=0} \triangleq \kappa_j, \qquad j = 1, 2, \ \ldots, k$$

称其为累积量 (cumulant) 或半不变量 (semi-invariant)

  - 累积量 $\kappa_j$ $\left(j > 1\right)$ 不随平移变换 $X \mapsto X + a$ 而改变.

# 矩与特征函数 (Moments and Characteristic Functions)

- 累积函数 (cumulant functions)

  - 累积量与前 $k$ 阶矩 $m_1, m_2, \ldots, m_k$ 有如下关系

$$\kappa_k = (-1)^{k-1}\begin{vmatrix} m_1 & 1 & 0 & \cdots & 0 \\ m_2 & \binom{1}{0} m_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_k & \binom{k-1}{0} m_{k-1} & \binom{k-1}{1} m_{k-2} & \cdots & \binom{k-1}{k-2} m_1 \end{vmatrix}$$

$$= (-1)^{k-1}\begin{vmatrix} m_1 & 1 & 0 & \cdots & 0 \\ m_2 & \mathrm{C}_1^0 \cdot m_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_k & \mathrm{C}_{k-1}^0 \cdot m_{k-1} & \mathrm{C}_{k-1}^1 \cdot m_{k-2} & \cdots & \mathrm{C}_{k-1}^{k-2} \cdot m_1 \end{vmatrix}$$

# 矩与特征函数 (Moments and Characteristic Functions)

- 例：当 $k=1$ 时，

$$\kappa_k = (-1)^{k-1}\begin{vmatrix} m_1 & 1 & 0 & \cdots & 0 \\ m_2 & C_1^0\cdot m_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_k & C_{k-1}^0\cdot m_{k-1} & C_{k-1}^1\cdot m_{k-2} & \cdots & C_{k-1}^{k-2}\cdot m_1 \end{vmatrix}$$

$$\kappa_1 = m_1$$

  - 当 $k=2$ 时，

$$\kappa_2 = -\begin{vmatrix} m_1 & 1 \\ m_2 & C_1^0\cdot m_1 \end{vmatrix} = m_2 - m_1^2$$

  - 当 $k=3$ 时，

$$\kappa_3 = \begin{vmatrix} m_1 & 1 & 0 \\ m_2 & C_1^0\cdot m_1 & 1 \\ m_3 & C_2^0\cdot m_2 & C_2^1\cdot m_1 \end{vmatrix} = \begin{vmatrix} m_1 & 1 & 0 \\ m_2 & m_1 & 1 \\ m_3 & m_2 & 2m_1 \end{vmatrix} = m_3 - 3m_1m_2 + 2m_1^3$$

  - 类似地，当 $k=4$ 时

$$\kappa_4 = m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4$$

# 矩与特征函数 (Moments and Characteristic Functions)

- 例：当 $k=1$ 时，

$$\kappa_1 = m_1$$

$$\kappa_k = (-1)^{k-1}\begin{vmatrix} m_1 & 1 & 0 & \cdots & 0 \\ m_2 & C_1^0 \cdot m_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_k & C_{k-1}^0 \cdot m_{k-1} & C_{k-1}^1 \cdot m_{k-2} & \cdots & C_{k-1}^{k-2} \cdot m_1 \end{vmatrix}$$

  - 也可以由累积量计算矩：

$$m_1 = \kappa_1$$
$$m_2 = \kappa_2 + \kappa_1^2$$
$$m_3 = \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3$$
$$m_4 = \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4$$

# 矩与特征函数 (Moments and Characteristic Functions)

- 例：当 $k=1$ 时，

$$\kappa_1 = m_1 \qquad \kappa_k = (-1)^{k-1}\begin{vmatrix} m_1 & 1 & 0 & \cdots & 0 \\ m_2 & \mathrm{C}_1^0\cdot m_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_k & \mathrm{C}_{k-1}^0\cdot m_{k-1} & \mathrm{C}_{k-1}^1\cdot m_{k-2} & \cdots & \mathrm{C}_{k-1}^{k-2}\cdot m_1 \end{vmatrix}$$

  - 累积量和中心矩 $\mu_k = \mathrm{E}(X-\mu)^k$ 的关系如下，其中 $\mu = m_1$

$$\kappa_2 = \mu_2$$
$$\kappa_3 = \mu_3$$
$$\kappa_4 = \mu_4 - 3\mu_2^2$$

# 矩与特征函数 (Moments and Characteristic Functions)

- 例：当 $k = 1$ 时，

$$\kappa_k = (-1)^{k-1} \begin{vmatrix} m_1 & 1 & 0 & \cdots & 0 \\ m_2 & \mathrm{C}_1^0 \cdot m_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_k & \mathrm{C}_{k-1}^0 \cdot m_{k-1} & \mathrm{C}_{k-1}^1 \cdot m_{k-2} & \cdots & \mathrm{C}_{k-1}^{k-2} \cdot m_1 \end{vmatrix}$$

$$\kappa_1 = m_1$$

- 偏度 (skewness) 和峰度 (kurtosis)

$$\gamma_3 = \frac{\mathrm{E}(X-\mu)^3}{\sigma^3} = \frac{\kappa_3}{\kappa_2^{3/2}}, \quad \gamma_4 = \frac{\mathrm{E}(X-\mu)^4}{\sigma^4} = \frac{\mu_4}{\kappa_2^2} = \frac{\kappa_4 + 3\mu_2^2}{\kappa_2^2} = \frac{\kappa_4 + 3\kappa_2^2}{\kappa_2^2} = \frac{\kappa_4}{\kappa_2^2} + 3$$

- 对正态分布而言：$\gamma_3 = 0$，$\gamma_4 = 3$.

# 变换 (transformations)

- 问题：随机变量 $X$ 的概率密度为 $f_X(x)$. 求 $Y = 3X$ 的概率密度.

$$f_Y(y) = \frac{1}{3} f_X\left(\frac{y}{3}\right) \qquad f_X(x) \xRightarrow{Y=aX+b} f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

- 问题：随机向量 $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ 的概率密度为 $f_{\boldsymbol{X}}(\boldsymbol{x})$. 求

$$\boldsymbol{Y} = \begin{pmatrix} 3X_1 \\ X_1 - 4X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & -4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \mathscr{A}\boldsymbol{X}$$

的概率密度.

# 变换 (transformations)

- 问题：对于一对一的变换 $u: \mathbb{R}^p \to \mathbb{R}^p$，当 $\boldsymbol{X} = u(\boldsymbol{Y})$ 时，求 $\boldsymbol{Y}$ 的概率密度.

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} u_1\left(Y_1, Y_2, \ldots, Y_p\right) \\ u_2\left(Y_1, Y_2, \ldots, Y_p\right) \\ \vdots \\ u_p\left(Y_1, Y_2, \ldots, Y_p\right) \end{pmatrix}$$

- 变换 $u$ 的 Jacobian 矩阵定义为

$$\mathcal{J} = \left(\frac{\partial x_i}{\partial y_j}\right)_{p\times p} = \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_p} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_p}{\partial y_1} & \frac{\partial x_p}{\partial y_2} & \cdots & \frac{\partial x_p}{\partial y_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial u_1}{\partial y_1} & \frac{\partial u_1}{\partial y_2} & \cdots & \frac{\partial u_1}{\partial y_p} \\ \frac{\partial u_2}{\partial y_1} & \frac{\partial u_2}{\partial y_2} & \cdots & \frac{\partial u_2}{\partial y_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial u_p}{\partial y_1} & \frac{\partial u_p}{\partial y_2} & \cdots & \frac{\partial u_p}{\partial y_p} \end{pmatrix}$$

- 用 $\text{abs}\left(\left|\mathcal{J}\right|\right)$ 表示 Jacobian 矩阵行列式的绝对值，则 $\boldsymbol{Y}$ 的概率密度为

$$f_Y(\boldsymbol{y}) = \text{abs}\left(\left|\mathcal{J}\right|\right) \cdot f_X\left[u(\boldsymbol{y})\right]$$

# 变换 (transformations)

- 对线性变换 $\boldsymbol{Y} = \mathscr{A}\boldsymbol{X} + \boldsymbol{b}$，其中 $\mathscr{A}$ 为非奇异阵.

$$\Longrightarrow \quad \boldsymbol{X} = \mathscr{A}^{-1}\left(\boldsymbol{Y} - \boldsymbol{b}\right)$$

$$\Longrightarrow \quad \mathscr{J} = \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{Y}} = \mathscr{A}^{-1}$$

$$f_{Y}(\boldsymbol{y}) = \text{abs}\left(\left|\mathscr{A}\right|^{-1}\right) \cdot f_{X}\left[\mathscr{A}^{-1}\left(\boldsymbol{y} - \boldsymbol{b}\right)\right]$$

# 变换 (transformations)

- 例： $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \in \mathbb{R}^2,\ f_{\boldsymbol{X}}(\boldsymbol{x}) = f_{\boldsymbol{X}}(x_1, x_2),\ \boldsymbol{Y} = \mathscr{A}\boldsymbol{X} + \boldsymbol{b} = \begin{pmatrix} X_1 + X_2 \\ X_1 - X_2 \end{pmatrix}$

$$\Longrightarrow \quad \mathscr{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Longrightarrow \quad |\mathscr{A}| = -2, \quad \mathrm{abs}\left(|\mathscr{A}|^{-1}\right) = \frac{1}{2}, \quad \mathscr{A}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

$$\Longrightarrow \quad f_{\boldsymbol{Y}}(\boldsymbol{y}) = \mathrm{abs}\left(|\mathscr{A}|^{-1}\right) \cdot f_{\boldsymbol{X}}\left[\mathscr{A}^{-1}(\boldsymbol{y} - \boldsymbol{b})\right]$$

$$= \frac{1}{2} f_{\boldsymbol{X}}\left[\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right]$$

$$= \frac{1}{2} f_{\boldsymbol{X}}\left[\frac{1}{2}(y_1 + y_2), \frac{1}{2}(y_1 - y_2)\right]$$

## 变换 (transformations)

- 例：设随机变量 $X \in \mathbb{R}^1$ 的概率密度函数为 $f_X(x)$. 令 $Y = \mathrm{e}^X$.

$$x = u(y) = \log(y) \quad \Longrightarrow \quad \mathscr{J} = \frac{dx}{dy} = \frac{1}{y}$$

$$\Longrightarrow \quad f_Y(y) = \mathrm{abs}\left(\left|\mathscr{J}\right|\right) \cdot f_X\left[u(y)\right]$$

$$= \frac{1}{y} f_X\left[\log(y)\right]$$

# 多元正态分布

- 多元正态分布 $\boldsymbol{X} \sim N_p\left(\boldsymbol{\mu}, \Sigma\right)$:

$x \in \mathbb{R}^p$ $\quad$ $x \in \mathbb{R}^p$

$$f_X(\boldsymbol{x}) = \left|2\pi\Sigma\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^{\mathrm{T}}\Sigma^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)\right\}$$

- 多元标准正态分布 $\boldsymbol{Y} \sim N_p\left(\boldsymbol{0}, \mathscr{I}_p\right)$:

$$\mathscr{I}_p = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}_{p\times p}$$

$$f_Y(\boldsymbol{y}) = (2\pi)^{-p/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}\right)$$

$$= (2\pi)^{-p/2}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y_1^2}{2}} \cdot (2\pi)^{-p/2}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y_2^2}{2}} \cdot \cdots \cdot (2\pi)^{-p/2}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y_p^2}{2}}$$

$Y_1 \sim N(0,\ 1)$ $\quad$ $Y_2 \sim N(0,\ 1)$ $\quad$ $Y_p \sim N(0,\ 1)$

# 多元正态分布

**定理 4.5** 设 $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$，$\boldsymbol{Y} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu})$ (Mahalanobis 变换). 则

$$\boldsymbol{Y} \sim N_p\left(\boldsymbol{0}, \mathscr{I}_p\right),$$

即，$\boldsymbol{Y}$ 的元素 $Y_j \in \mathbb{R}$ 是相互独立、服从一元标准正态分布 $N(0, 1)$ 的随机变量.

**证明** $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma) \implies f_X(\boldsymbol{x}) = \left|2\pi\Sigma\right|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}} \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right\}$

$$\boldsymbol{Y} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu}) \implies \boldsymbol{X} = \Sigma^{1/2}\boldsymbol{Y} + \boldsymbol{\mu}, \qquad \mathscr{J} = \frac{\partial \boldsymbol{X}}{\partial \boldsymbol{Y}} = \Sigma^{1/2}$$

$$\boldsymbol{X} = u(\boldsymbol{Y}) \implies f_Y(\boldsymbol{y}) = \mathrm{abs}\left(\left|\mathscr{J}\right|\right) \cdot f_X\left[u(\boldsymbol{y})\right]$$

$$f_Y(\boldsymbol{y}) = \mathrm{abs}\left(\left|\Sigma^{1/2}\right|\right) \cdot \left|2\pi\Sigma\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(\Sigma^{1/2}\boldsymbol{y}\right)^{\mathrm{T}} \Sigma^{-1} \left(\Sigma^{1/2}\boldsymbol{y}\right)\right\}$$

$$= (2\pi)^{-p/2} \cdot \left|\Sigma^{1/2}\right| \cdot \left|\Sigma\right|^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\left(\Sigma^{1/2}\right)^{\mathrm{T}} \Sigma^{-1}\Sigma^{1/2}\boldsymbol{y}\right\} = (2\pi)^{-p/2} \exp\left\{-\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}\right\}$$

# 多元正态分布

- 如何由多元标准正态分布 $N_p\left(\mathbf{0}\ ,\ \mathscr{I}_p\right)$ 构造多元正态分布 $N_p\left(\boldsymbol{\mu}\ ,\ \Sigma\right)$ ?

$$\boldsymbol{Y}=\Sigma^{1/2}\boldsymbol{X}+\boldsymbol{\mu} \quad \Longleftarrow \quad \boldsymbol{X}\sim N_p\left(\mathbf{0},\ \mathscr{I}_p\right)$$

$$\boldsymbol{Y}=\mathscr{A}\boldsymbol{X}+\boldsymbol{b} \quad \Longrightarrow \quad f_Y(\boldsymbol{y})=\mathrm{abs}\left(\left|\mathscr{A}\right|^{-1}\right)\cdot f_X\left[\mathscr{A}^{-1}\left(\boldsymbol{y}-\boldsymbol{b}\right)\right]$$

$$f_X(\boldsymbol{x})=(2\pi)^{-p/2}\exp\left(-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\right)$$

$$\begin{aligned}
f_Y(\boldsymbol{y}) &= \mathrm{abs}\left(\left|\Sigma^{1/2}\right|^{-1}\right)\cdot f_X\left[\Sigma^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu})\right] \\
&= \left(\left|\Sigma\right|^{-1/2}\right)\cdot\left[(2\pi)^p\right]^{-1/2}\cdot\exp\left(-\frac{1}{2}\left\{\left[\Sigma^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu})\right]\right\}^{\mathrm{T}}\left[\Sigma^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu})\right]\right) \\
&= \left|2\pi\Sigma\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\boldsymbol{y}-\boldsymbol{\mu}\right)^{\mathrm{T}}\Sigma^{-1}\left(\boldsymbol{y}-\boldsymbol{\mu}\right)\right\}
\end{aligned}$$

$$\Longrightarrow \quad \boldsymbol{Y}=\Sigma^{1/2}\boldsymbol{X}+\boldsymbol{\mu}\sim N_p(\boldsymbol{\mu},\ \Sigma)$$

**定理 4.6** 设 $\boldsymbol{X}\sim N_p\left(\boldsymbol{\mu},\ \Sigma\right)$，$\mathscr{A}_{p\times p}$ 是非奇异阵，$\boldsymbol{c}\in\mathbb{R}^p$，则 $\boldsymbol{Y}=\mathscr{A}\boldsymbol{X}+\boldsymbol{c}$ 也服从 $p$ 元正态分布，即

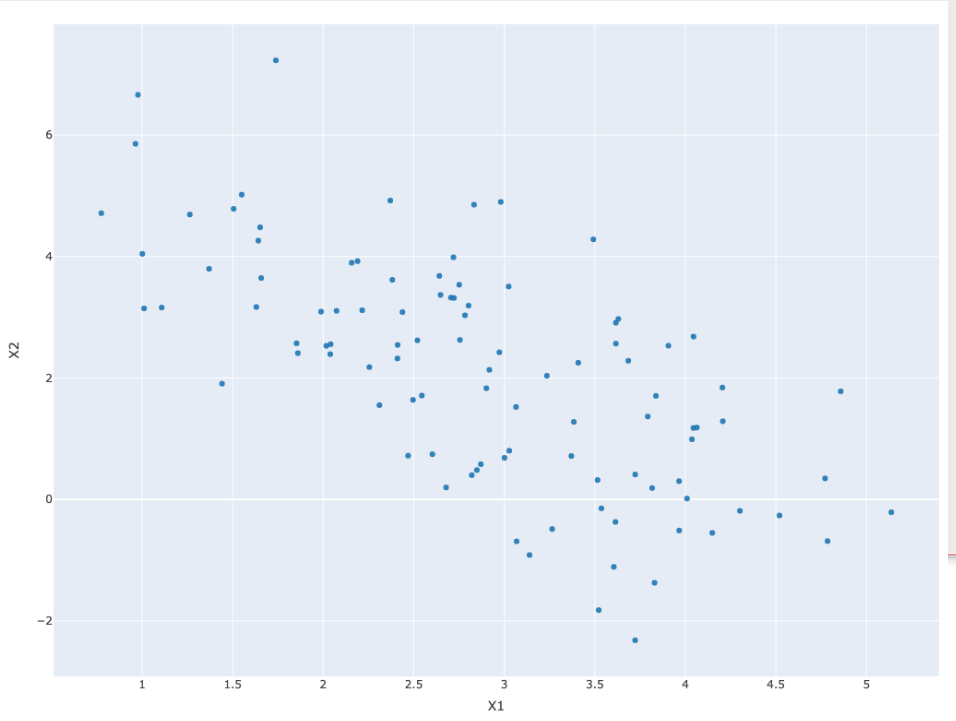
$$\boldsymbol{Y}\sim N_p\left(\mathscr{A}\boldsymbol{\mu}+\boldsymbol{c},\ \mathscr{A}\Sigma\mathscr{A}^{\mathrm{T}}\right)$$

# 多元正态分布

- 分布 $N_p(\boldsymbol{\mu}, \Sigma)$ 的几何特征
  - $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2(\boldsymbol{\mu}, \Sigma)$，其中 $\boldsymbol{\mu} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$，$\Sigma = \begin{pmatrix} 1 & -1.5 \\ -1.5 & 4 \end{pmatrix}$.

```
library(ggplot2)
library(plotly)
library(MASS)
mu = c(3, 2)
Sigma = matrix(c(1, -1.5, -1.5, 4), nrow = 2, byrow = TRUE)
X = mvrnorm(n = 100, mu, Sigma)
colnames(X) = c("X1", "X2")
X = as.data.frame(X)

fig5 = plot_ly(X, x = ~X1, y = ~X2) %>%
  add_markers(alpha = 0.9) %>%
  layout(plot_bgcolor='#e5ecf6',
        xaxis = list(zerolinecolor = '#ffff', zerolinewidth = 2, gridcolor = 'ffff'),
        yaxis = list(zerolinecolor = '#ffff', zerolinewidth = 2, gridcolor = 'ffff'))
fig5
```

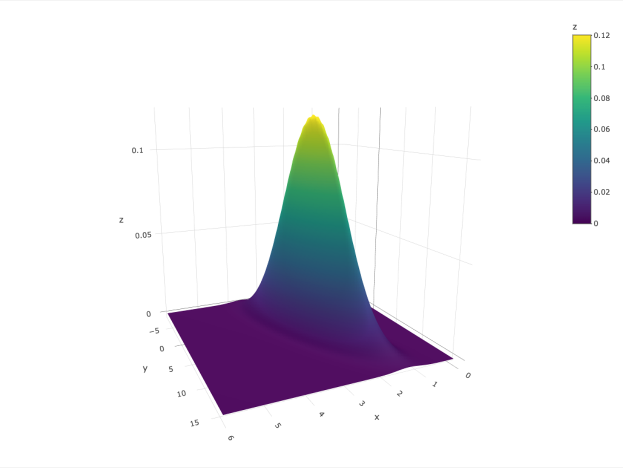


# 多元正态分布

- 分布 $N_p(\boldsymbol{\mu}, \Sigma)$ 的几何特征
  - $\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2(\boldsymbol{\mu}, \Sigma)$，其中 $\boldsymbol{\mu} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$，$\Sigma = \begin{pmatrix} 1 & -1.5 \\ -1.5 & 4 \end{pmatrix}$.

```
## ---------- 联合密度函数曲面 ----------
n = 80
mu1 = 3
mu2 = 2
s1 = 1
s2 = 2
rho = -1.5/2
x = seq(0, 6, length = n) * s1
y = seq(-4 , 8, length = n) * s2
f = function(x, y){
  (2 * pi * s1 * s2 * sqrt(1-rho^2))^-1 * exp(-0.5 * (1 - rho^2)^-1 *
                 ((x-mu1)^2/s1^2 - 2 * rho * (x - mu1) * ( y - mu2) / (s1 * s2) + (y - mu2)^2/s2^2))
}
z = outer(x , y, f)

fig6 = plot_ly() %>%
  add_surface(x = ~x, y = ~y, z = ~z)
fig6
```



# 多元正态分布

- 分布 $N_p(\boldsymbol{\mu}, \Sigma)$ 的几何特征

$$f_X(\boldsymbol{x}) = \left|2\pi\Sigma\right|^{-1/2} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}$$

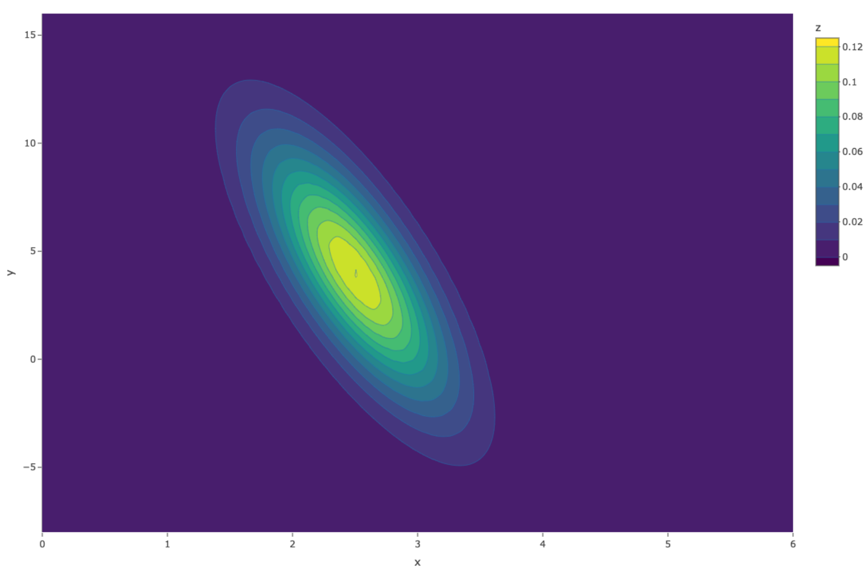
椭圆

$$(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = d^2$$

常数

$$f_X(\boldsymbol{x}) = \left|2\pi\Sigma\right|^{-1/2} \exp\left\{-\frac{d^2}{2}\right\}$$

```
## ---------- 联合密度函数曲面的等高线 ----------
fig7 = plot_ly(x = ~x, y = ~y, z = ~z) %>%
  add_contour()
fig7
```



# 多元正态分布

- 分布 $N_p\left(\boldsymbol{\mu}, \Sigma\right)$ 的几何特征

**定理 4.7** 设 $\boldsymbol{X} \sim N_p\left(\boldsymbol{\mu}, \Sigma\right)$，则二次型变量 $U = \left(\boldsymbol{X}-\boldsymbol{\mu}\right)^{\mathrm{T}} \Sigma^{-1} \left(\boldsymbol{X}-\boldsymbol{\mu}\right)$ 服从 $\chi_p^2$ 分布.

$p = 1$

$$U = \frac{(X-\mu)^2}{\sigma^2} = \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi_1^2$$

**定理 4.5** 设 $\boldsymbol{X} \sim N_p\left(\boldsymbol{\mu}, \Sigma\right)$，$\boldsymbol{Y} = \Sigma^{-1/2}\left(\boldsymbol{X}-\boldsymbol{\mu}\right)$ (Mahalanobis 变换). 则

$$\boldsymbol{Y} \sim N_p\left(\boldsymbol{0}, \mathscr{I}_p\right),$$

即，$\boldsymbol{Y}$ 的元素 $Y_j \in \mathbb{R}$ 是相互独立、服从一元标准正态分布 $N(0, 1)$ 的随机变量.

$p > 1$

$$U = \left(\boldsymbol{X}-\boldsymbol{\mu}\right)^{\mathrm{T}} \Sigma^{-1/2}\Sigma^{-1/2} \left(\boldsymbol{X}-\boldsymbol{\mu}\right)$$

$$= \left[\Sigma^{-1/2}\left(\boldsymbol{X}-\boldsymbol{\mu}\right)\right]^{\mathrm{T}} \left[\Sigma^{-1/2}\left(\boldsymbol{X}-\boldsymbol{\mu}\right)\right]$$

$$\triangleq \boldsymbol{Y} \sim N_p\left(\boldsymbol{0}, \mathscr{I}_p\right)$$

$$= \sum_{k=1}^{p} Y_k^2 \sim \chi_p^2$$

$$Y_k \overset{\text{i.i.d.}}{\sim} N(0, 1)$$