

Ch05: 多维随机变量及其数字特征

Operations on Multi-dimensional Random Variables – 2

November 19, 2024

引言

已知二维随机向量 (X, Y) 的概率分布, 求随机变量 (X, Y) 的函数 $Z = g(X, Y)$ 的概率分布, 分离散和连续两种情况讨论.

离散型随机变量函数的分布

已知离散型随机向量 (X, Y) 的联合分布列, 求随机变量 $Z = g(X, Y)$ 的分布列:

- 根据 X, Y 的各种取值, 计算随机变量 Z 的取值;

$\begin{array}{c} \text{Y} \\ \diagdown \\ \text{X} \end{array}$	y_1	y_2	$\dots$	y_n
x_1	$X = x_1, Y = y_1$	$X = x_1, Y = y_2$	$\dots$	$X = x_1, Y = y_n$
x_2	$X = x_2, Y = y_1$	$X = x_2, Y = y_2$	$\dots$	$X = x_2, Y = y_n$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_m	$X = x_m, Y = y_1$	$X = x_m, Y = y_2$	$\dots$	$X = x_m, Y = y_n$

- 对相同的 Z 值合并, 对应的概率相加.

离散型随机变量函数的分布——重要性质

- < 和函数 > $Z = X + Y$ 的分布列为 $P(Z = X + Y = k) = \sum_{i=1}^k a_i b_{k-i}$.
- 二项分布之和是二项分布, 即 $Z = X + Y \sim B(n_1 + n_2, p)$.
- 泊松分布之和是泊松分布, 即 $Z = X + Y \sim P(\lambda_1 + \lambda_2)$.

连续型随机向量函数的分布

设二维连续型随机向量 (X, Y) 的联合概率密度为 $f(x, y)$, 求随机变量 $Z = g(X, Y)$ 的概率密度:

- 先计算分布函数 (积分区域)

$$F_Z(z) = P(Z \leq z) = P(g(x, y) \leq z) = \int \int_{g(x, y) \leq z} f(x, y) \, dx \, dy$$

- 对分布函数 $F_Z(z)$ 求导得到密度函数

$$f_Z(z) = F'_Z(z)$$

连续型随机变量函数的分布——重要性质

- < 求和基本公式 > $Z = X + Y$ 的密度函数为

$$f_Z(z) = \int_{-\infty}^{+\infty} f(x, z-x) dx = \int_{-\infty}^{+\infty} f(z-y, y) dy$$

- < 和函数 + 独立性 > $Z = X + Y$ 的密度函数为

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x)f_Y(z-x) dx = \int_{-\infty}^{+\infty} f_X(z-y)f_Y(y) dy$$

- 均匀分布: 例 0.108.

- 指数分布: 例 0.109.

- 正态分布之和是正态分布, 即 $X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

- < 乘/除法 > $Z = XY$ 和 $Z = Y/X$ 的概率密度分别为

$$f_{XY}(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f(x, \frac{z}{x}) dx \quad f_{Y/X}(z) = \int_{-\infty}^{+\infty} |x| f(x, xz) dx$$

- < 最大值/最小值 > $Y = \max(X_1, \dots, X_n)$ 和 $Z = \min(X_1, \dots, X_n)$ 的分布函数分别为

$$\begin{cases} F_Y(y) = F_{X_1}(y) F_{X_2}(y) \dots F_{X_n}(y) \\ F_Z(z) = 1 - [1 - F_{X_1}(z)] [1 - F_{X_2}(z)] \dots [1 - F_{X_n}(z)] \end{cases}$$

- < 复合函数 > $(u, v) = (u(x, y), v(x, y))$ **OR** $(x, y) = (x(u, v), y(u, v))$, 则有

$$f_{UV}(u, v) = f_{XY}(x(u, v), y(u, v))|J|,$$

其中, J 为变换的雅可比行列式, 即

$$J = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| = \left| \frac{\partial(u, v)}{\partial(x, y)} \right|^{-1} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}^{-1}.$$

例题补充：例 0.114

例 0.114 设 X_1, X_2, X_3 相互独立, 并且有相同的概率分布

$$P(X_i = 1) = p, \quad P(X_i = 0) = 1 - p, \quad i = 1, 2, 3$$

考虑变量

$$Y_1 = \begin{cases} 1, & \text{若 } X_1 + X_2 \text{ 为奇数,} \\ 0, & \text{若 } X_1 + X_2 \text{ 为偶数,} \end{cases}$$
$$Y_2 = \begin{cases} 1, & \text{若 } X_2 + X_3 \text{ 为奇数,} \\ 0, & \text{若 } X_2 + X_3 \text{ 为偶数,} \end{cases}$$

求乘积变量 $Y_1 Y_2$ 的概率分布.

解答：例 0.114

题目：如上所述.

解答：

• $Y_1, Y_2 \in \{0, 1\} \Rightarrow Y_1 Y_2 \in \{0, 1\}$, 原问题转化为“求 $P(Y_1 Y_2 = 1)$ 和 $P(Y_1 Y_2 = 0)$ ”.

• For the case of $Y_1 Y_2 = 1$,

$$\begin{aligned} P(Y_1 Y_2 = 1) &= P(Y_1 = 1, Y_2 = 1) \\ &= P(\text{若 } X_1 + X_2 \text{ 为奇数, 若 } X_2 + X_3 \text{ 为奇数}) \\ &= P(\{X_1 = 0, X_2 = 1, X_3 = 0\} \cup \{X_1 = 1, X_2 = 0, X_3 = 1\}) \quad \text{事件不相容} \\ &= P(\{X_1 = 0, X_2 = 1, X_3 = 0\}) + P(\{X_1 = 1, X_2 = 0, X_3 = 1\}) \\ &= P(X_1 = 0)P(X_2 = 1)P(X_3 = 0) + P(X_1 = 1)P(X_2 = 0)P(X_3 = 1) \\ &= p(1 - p)^2 + p^2(1 - p) . \end{aligned}$$

• For the case of $Y_1 Y_2 = 0$,

$$P(Y_1 Y_2 = 1) = 1 - P(Y_1 Y_2 = 0) = 1 - p(1 - p) .$$

例题补充：例 0.115

例 0.115 设系统 L 由两个相互独立的子系统 L_1 和 L_2 连接而成, 其连接的方式分别为 (1) 串联; (2) 并联. 设 L_1 和 L_2 的寿命分别为 X 和 Y , 已知它们的密度函数分别为

$$f_X(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

$$f_Y(y) = \begin{cases} \beta e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

其中 $\alpha > 0, \beta > 0$. 试分别就以上两种连接方式求系统 L 的寿命 Z 的密度函数.

解答：例 0.115

题目：如上所述.

解答：建模要点 – 串联和并联各代表何种运算？

- L_1 和 L_2 串联时, L 的寿命 $Z = \min(X, Y)$, 而 X 和 Y 的分布函数为

$$F_X(x) = \begin{cases} 1 - e^{-\alpha x}, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad F_Y(y) = \begin{cases} 1 - e^{-\beta y}, & y > 0 \\ 0, & y \leq 0 \end{cases}$$

由此 Z 的分布函数为

$$F_Z(z) = 1 - (1 - F_X(x))(1 - F_Y(y)) = \begin{cases} 1 - e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

于是 Z 的密度函数为

$$f_Z(z) = \begin{cases} (\alpha + \beta)e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

- L_1 和 L_2 并联时, L 的寿命 $Z = \max(X, Y)$, 由此 Z 的分布函数为

$$F_Z(z) = F_X(x)F_Y(y) = \begin{cases} (1 - e^{-\alpha z})(1 - e^{-\beta z}), & z > 0 \\ 0, & z \leq 0 \end{cases}$$

于是 Z 的密度函数为

$$f_Z(z) = \begin{cases} \alpha e^{-\alpha z} + \beta e^{-\beta z} + (\alpha + \beta)e^{-(\alpha+\beta)z}, & z > 0 \\ 0, & z \leq 0 \end{cases}$$

例题补充：例 0.116

例 0.116 假设一电路装有 3 个同种电气元件, 其工作状态相互独立, 且无故障工作时间都服从参数为 $\lambda > 0$ 的指数分布 $e(\lambda)$. 当 3 个元件都无故障时, 电路正常工作, 否则整个电路都不能正常工作. 试求电路正常工作的时间 T 的概率分布.

解答：例 0.116

题目：如上所述.

解答：

- 解法一，以 $X_i (i = 1, 2, 3)$ 表示第 i 个电气元件无故障工作的时间，则 X_1, X_2, X_3 相互独立且同分布，其分布函数为

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

设 $G(t)$ 是 T 的分布函数, $T = \min\{X_1, X_2, X_3\}$. 当 $t \leq 0$ 时, $G(t) = 0$. 当 $t > 0$, 有

$$\begin{aligned} G(t) &= P\{T \leq t\} = 1 - \{T > t\} = 1 - P\{X_1 > t, X_2 > t, X_3 > t\} \\ &= 1 - P\{X_1 > t\}P\{X_2 > t\}P\{X_3 > t\} = 1 - [1 - F(t)]^3 \\ &= 1 - e^{-3\lambda x} \end{aligned}$$

得知 T 服从参数为 3λ 的指数分布:

$$G(t) = \begin{cases} 1 - e^{-3\lambda t}, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

- 解法二, 本题也可以直接利用公式计算, 因为 X_1, X_2, X_3 相互独立且同分布, 而 $T = \min(X_1, X_2, X_3)$, 故

$$G(t) = 1 - [1 - F(t)]^3 = \begin{cases} 1 - e^{-3\lambda t}, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

例题补充：例 0.117

例 0.117 假设一设备开机后无故障工作时间 X 服从指数分布, 平均无故障工作的时间 $\mathbb{E}(X)$ 为 5 小时. 设备定时开机. 出现故障时自动关机, 而在无故障的情况下工作 2 小时便关机. 试求该设备每次开机无故障工作的时间 Y 的分布函数 $F(y)$.

解答：例 0.117

题目：如上所述.

解答：

- 设 X 的分布参数为 λ . 由于 $\mathbb{E}(X) = \frac{1}{\lambda} = 5$, 可见 $\lambda = \frac{1}{5}$, 显然

$$Y = \min(X, 2)$$

对于 $y < 0$, $F(y) = 0$. 对于 $y \geq 2$, $F(y) = 1$. 设 $0 \leq y < 2$, 有

$$F(y) = P\{Y \leq y\} = P\{\min(X, 2) \leq y\}$$

于是 Y 的分布函数为

$$F(y) = \begin{cases} 0, & y < 0 \\ 1 - e^{-\frac{y}{5}}, & 0 \leq y < 2 \\ 1, & y \geq 2 \end{cases}$$

例题补充：例 0.118

例 0.118 设某班车起点站上客人数 X 服从参数为 λ ($\lambda > 0$) 的泊松分布, 每位乘客在中途下车的概率为 p ($0 < p < 1$), 且中途下车与否相互独立. 以 Y 表示在中途下车的人数. 求:

- (1) 在发车时有 n 个乘客的条件下, 中途有 m 人下车的概率;
- (2) 二维随机变量 (X, Y) 的概率分布.

解答：例 0.118

题目：如上所述.

解答：这是一个关于条件分布的例题.

- 对于第 1 小题, 有

$$P\{Y = m \mid X = n\} = C_n^m p^m (1 - p)^{n-m}, \quad 0 \leq m \leq n, n = 0, 1, 2, \dots$$

- 对于第 2 小题, 有

$$P\{X = n, Y = m\} = P\{Y = m \mid X = n\}P\{X = n\} = C_n^m p^m (1 - p)^{n-m} \cdot \frac{e^{-\lambda}}{n!} \lambda^n$$

其中, $0 \leq m \leq n, n = 0, 1, 2, \dots$.

例题补充：例 0.119

例 0.119 [作业加强题] 设某种型号的电子元件的寿命 (以小时计) 近似地服从 $\mathcal{N}(160, 20^2)$, 随机地选择 4 只. 求其中没有一只寿命小于 180 小时的概率.

解答：例 0.119

题目：如上所述.

解答：

- 随机取 4 只, 记其寿命分别为: X_1, X_2, X_3, X_4 . 这四个变量独立同分布, 即

$$X_i \sim \mathcal{N}(160, 20^2) \quad i = 1, 2, 3, 4$$

- 记 $X = \min\{X_1, X_2, X_3, X_4\}$.
- 所求问题可以转化为 $P(X \geq 180)$, 则有

$$\begin{aligned} P(X \geq 180) &= [1 - F(180)]^4 \\ &= \left[1 - \Phi\left(\frac{180 - 160}{20}\right)\right]^4 \\ &= (1 - 0.8413)^4 \\ &\approx 0.000634 \end{aligned}$$

多维随机变量的分布函数和密度函数

将二维随机向量及分布推广到多维随机向量, 二维与多维随机变量没有本质性的区别, 只是相关概念和结论的扩展.

定义 0.49 设 $(X_1, X_2, \dots, X_n)$ 为 n 维随机向量, 对任意实数 $x_1, x_2, \dots, x_n$, 函数

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

称为 n 维随机向量 $(X_1, X_2, \dots, X_n)$ 的分布函数, 或随机变量 $X_1, X_2, \dots, X_n$ 的联合分布函数. 若存在可积函数 $f(x_1, x_2, \dots, x_n)$, 使得对任意实数 $x_1, x_2, \dots, x_n$ 有

$$F(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(u_1, u_2, \dots, u_n) \, du_1 \, du_2 \cdots du_n,$$

则称 $(X_1, X_2, \dots, X_n)$ 为连续型随机向量, 称 $f(x_1, x_2, \dots, x_n)$ 为 n 维联合密度函数.

n 维联合密度函数的性质

- 非负性. 对任意实数 $x_1, x_2, \dots, x_n$ 有 $f(x_1, x_2, \dots, x_n) \geq 0$;
- 规范性.

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(u_1, u_2, \dots, u_n) \, du_1 \, du_2 \cdots du_n = 1$$

- 连续性. 若 $f(x_1, x_2, \dots, x_n)$ 在点 $(x_1, x_2, \dots, x_n)$ 处连续, 则有

$$\frac{\partial F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n} = f(x_1, x_2, \dots, x_n)$$

- 有界区域可积. 设 G 是 n 维空间的一片区域, 则有

$$P((X_1, X_2, \dots, X_n) \in G) = \int \cdots \int_G f(u_1, u_2, \dots, u_n) \, du_1 \, du_2 \cdots du_n$$

多维随机变量的边缘分布函数和边缘密度函数

定义 0.50 n 维随机向量 $(X_1, X_2, \dots, X_n)$ 中任意 k ($k \leq n$) 个分量所构成的随机向量, 它的分布函数和密度函数被称为 k 维边缘分布函数和 k 维边缘密度函数.

例如, 随机向量 $(X_1, X_2, \dots, X_n)$ 前 k 维随机向量的边缘分布函数和边缘密度函数分布为

$$\begin{aligned} F_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) &= P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_k \leq x_k) \\ &= \lim_{\substack{x_{k+1} \rightarrow +\infty \\ x_n \rightarrow +\infty}} F(x_1, x_2, \dots, x_k) \end{aligned}$$

$$f_{X_1, X_2, \dots, X_k}(x_1, x_2, \dots, x_k) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(u_1, \dots, u_k, u_{k+1}, \dots, u_n) \, du_{k+1} \cdots du_n$$

多维随机变量的独立性

将二维随机向量及分布推广到多维随机向量, 二维与多维随机变量没有本质性的区别, 只是相关概念和结论的扩展.

定义 0.51 若随机向量 $(X_1, X_2, \dots, X_n)$ 的联合分布函数 $F(x_1, x_2, \dots, x_n)$ 满足

$$F(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2) \dots F_{X_n}(x_n)$$

则称 $X_1, X_2, \dots, X_n$ 相互独立. 若随机向量 $X = (X_1, X_2, \dots, X_m)$ 和 $Y = (Y_1, Y_2, \dots, Y_n)$ 的联合分布函数 $F(x_1, \dots, x_m, y_1, \dots, y_n)$ 满足

$$F(x_1, \dots, x_m, y_1, \dots, y_n) = F_X(x_1, \dots, x_m)F_Y(y_1, \dots, y_n)$$

则称随机向量 X 和 Y 相互独立.

多维正态分布

多维随机向量中最重要的常用是多维正态分布.

定义 0.52 给定一个向量 $\boldsymbol{\mu} \in \mathbb{R}^n$ 和正定矩阵 $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$, 对任意实数向量 $\boldsymbol{x} = (x_1, x_2, \dots, x_n)^\top$, 若随机向量 $X = (X_1, X_2, \dots, X_n)$ 的密度函数为

$$f(\boldsymbol{x}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1/2} \exp \left(-(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) / 2 \right)$$

则称随机向量 X 服从参数为 $\boldsymbol{\mu}$ 和 $\boldsymbol{\Sigma}$ 的**多维正态分布 (multivariate normal distribution)**, 记为 $X = (X_1, X_2, \dots, X_n) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. 特别地, 当 $n = 2$ 时, 二维随机变量 $(X, Y) \sim \mathcal{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ 可以写成矩阵形式

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \quad \text{和} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

多维正态分布

当 $n = 2$ 时, 三维随机变量 $(X_1, X_2, X_3) \sim \mathcal{N}(\mu_1, \mu_2, \mu_3, \sigma_1^2, \sigma_2^2, \sigma_3^2, \boldsymbol{\rho})$ 可以写成矩阵形式

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}$$

和

$$\boldsymbol{\Sigma} = \begin{pmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \text{Cov}(X_1, X_3) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \text{Cov}(X_2, X_3) \\ \text{Cov}(X_3, X_1) & \text{Cov}(X_3, X_2) & \text{Var}(X_3) \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_2 & \rho_{23}\sigma_1\sigma_2 & \sigma_3^2 \end{pmatrix}$$

协方差矩阵是一个对称的矩阵, 因此是有实特征值的, 进而是非半正定的。

多维标准正态分布及标准化

回顾: 一维随机变量 $X \sim \mathcal{N}(\mu, \sigma^2)$, 则

$$Y = (X - \mu)/\sigma \sim \mathcal{N}(0, 1) .$$

即正态变量都可以通过一个线性变换(标准化)化成标准状态变量.

定义 0.53 当 $\mu = \mathbf{0}_n$ (全为零的 n 维向量), 以及 $\Sigma = \mathbf{I}_n$ ($n \times n$ 单位阵) 时, 正态分布 $\mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$ 被称为 n 维标准正态分布.

定理 0.28 设 n 维随机向量 $X = (X_1, X_2, \dots, X_n) \sim \mathcal{N}(\mu, \Sigma)$, 以及正定矩阵 Σ 的特征值分解 $\Sigma = \mathbf{U}^\top \Lambda \mathbf{U}$, 则随机向量

$$Y = \Lambda^{-1/2} \mathbf{U} (X - \mu) \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n) .$$

Proof of Theorem 0.28

证明 根据 $Y = \Lambda^{-1/2}U(X - \mu)$ 可得 $X = U^T\Lambda^{1/2}Y + \mu$, 已知 X 的概率密度函数为

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

根据 n 维随机变量函数 (定理 5.10 的多维情况) 的概率密度公式有

$$f_Y(y) = f_X \left(U^T \Lambda^{1/2} y + \mu \right) \left| U^T \Lambda^{1/2} \right| ,$$

其中 $y = (y_1, y_2, \dots, y_n)^T$. 根据特征值分解 $\Sigma = U^T \Lambda U$ 有

$$\left| U^T \Lambda^{1/2} \right| = |\Sigma|^{1/2} ,$$

以及将 $x = U^T \Lambda^{1/2} y + \mu$ 代入有

$$(x - \mu)^T \Sigma^{-1} (x - \mu) = y^T y .$$

由此可得随机向量 $Y = \Lambda^{-1/2}U(X - \mu)$ 的密度函数为

$$f_Y(y) = (2\pi)^{-n/2} \exp \left(-y^T y / 2 \right) ,$$

Proof of Theorem 0.28

- 随机向量 $X = (X_1, X_2, \dots, X_n)$ 的密度函数为

$$f(\mathbf{x}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left(-(\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) / 2 \right)$$

- 根据

$$Y = \Lambda^{-1/2} \mathbf{U} (X - \boldsymbol{\mu}) = \Sigma^{-1/2} (X - \boldsymbol{\mu})$$

则有

$$\mathbf{y}^\top \mathbf{y} = (\mathbf{x} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

- 带入, 则有

$$f_Y(\mathbf{y}) = (2\pi)^{-n/2} \exp \left(-\mathbf{y}^\top \mathbf{y} / 2 \right)$$

多维正态分布的可加性

定理 **0.29** [线性性] 随机向量 $X = (X_1, X_2, \dots, X_n) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 则有

$$Y = \mathbf{A}X + \mathbf{b} \sim \mathcal{N}(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top)$$

其中 $|\mathbf{A}| \neq 0$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ 和 $\mathbf{b} \in \mathbb{R}^{n \times 1}$.

多维正态分布其他性质 – 该证明留作思考题

定理 0.30 设随机向量 $X = (X_1, X_2, \dots, X_n)^\top$ 和 $Y = (Y_1, Y_2, \dots, Y_m)^\top$, 以及

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix} \right),$$

则有

- 随机向量 X 和 Y 的边缘分布为 $X \sim \mathcal{N}(\mu_x, \Sigma_{xx})$ 和 $Y \sim \mathcal{N}(\mu_y, \Sigma_{yy})$
- 随机向量 X 和 Y 相互独立的充要条件是 $\Sigma_{xy} = \mathbf{0}_{m \times n}$ (元素全为零的 $m \times n$ 矩阵)
- 在 $X = x$ 的条件下, 随机向量 Y 的分布

$$Y \mid X = x \sim \mathcal{N}(\mu_y + \Sigma_{yx}\Sigma_{xx}^{-1}(x - \mu_x), \Sigma_{yy} - \Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy})$$

- 在 $Y = y$ 的条件下, 随机向量 X 的分布

$$X \mid Y = y \sim \mathcal{N}(\mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y), \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})$$