

数学期望(expectation)的含义: 均值(mean)

第四章、数字特征与特征函数

§4.1 数学期望

- 含义: 平均值.
- 时间平均: 大量观测值的算术平均.

X 的(独立的)观测值 a_1, a_2, \dots, a_n , 其平均值为

$$\frac{1}{n}(a_1 + \dots + a_n).$$

- 空间平均: 所有可能值的加权平均(总和).

假设 X 为离散型, 分布列为 $P(X = x_k) = p_k, \forall k$. 那么,

$$a_1 + \cdots + a_n = \sum_k x_k n_k,$$

$$n_k = |\{m : 1 \leq m \leq n, a_m = x_k\}|.$$

- 根据概率的频率含义:

$$\frac{n_k}{n} \approx p_k,$$

$$\text{因此, } \frac{1}{n}(a_1 + \cdots + a_n) \approx \sum_k x_k p_k.$$

离散型

- 若 $\sum_k |x_k|p_k < \infty$, 则称

$$\sum_k x_k p_k$$

为 X 的(数学)期望, 记为 EX . (期望存在, 定义4.1.1.)

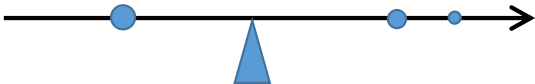
- 若 $\sum_k |x_k|p_k = \infty$, 则说 X 的期望不存在. 例4.1.5.

$$x_k = (-1)^k \frac{2^k}{k}, \quad p_k = \frac{1}{2^k}, \quad k \geq 1.$$

- 记

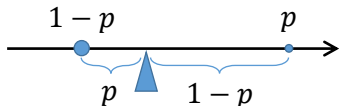
$$x^+ = x \vee 0, \quad x^- = (-x) \vee 0.$$

- 若 $\sum_k x_k^+ p_k$ 一个收敛而另一个发散, 也称 $EX = \sum_k x_k p_k$ 为 X 的期望. (期望不存在).
- 分布的数字特征: 若 $X \stackrel{d}{=} Y$, 则 $EX = EY$.
- 重心:



• 例4.1.1. Bernoulli 分布. $E1_A = P(A)$.

• 例4.1.3. 泊松分布.



(1) $X \geq 0$, EX 有意义.

(2) $x_k = k$:

$$kp_k = k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda p_{k-1}, \quad k \geq 1.$$

(3) 计算:

$$EX = \sum_{k=0}^{\infty} k \cdot p_k = \sum_{\ell=0}^{\infty} \lambda \frac{\lambda^{\ell}}{\ell!} e^{-\lambda} = \lambda.$$

- 习题四、7. X 取非负整数:

$$EX = \sum_{n=1}^{\infty} P(X \geq n) = \sum_{n=0}^{\infty} P(X > n).$$

- 证明:

$$EX = \sum_{k=1}^{\infty} k p_k = \sum_{k=1}^{\infty} \sum_{n=1}^k p_k = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} p_k.$$

- 例4.1.4. 几何分布. $P(X > n) = q^n, \forall n \geq 0$.

$$EX = \sum_{n=0}^{\infty} q^n = \frac{q^0}{1-q} = \frac{1}{p}.$$

连续型.

- 离散逼近: $x_0 < x_1 < \cdots < x_n$,

$$\sum_i \underbrace{x_i p(x_i)}_{\text{离散}} \Delta x_i \rightarrow \int \underbrace{x p(x)}_{\text{连续}} dx.$$

- 若 $\int |x|p(x)dx < \infty$, 则称

$$\int x p(x) dx$$

为 X 的(数学)期望, 记为 EX . (期望存在, 定义4.1.2.)

- $\int x^{\pm} p(x) dx$ 不同时为 ∞ 时,

$$EX = \int xp(x) dx.$$

- 例4.1.13. 柯西分布, EX 无意义.

$$p(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

- 若 $X \geq 0$, 则

$$EX = \int_0^{\infty} G(x) dx.$$

- 证明:

$$EX = \int_0^{\infty} xp(x) dx = -xG(x)|_0^{\infty} + \int_0^{\infty} G(x) dx,$$

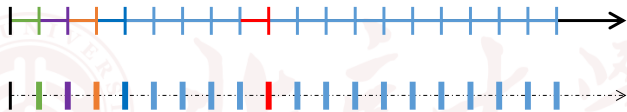
$$\text{其中, } aG(a) \leq \int_a^{\infty} yp(y) dy.$$

- 例4.1.12. $X \sim \text{Exp}(\lambda)$.

$$\int_0^{\infty} P(X > x) dx = \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda}.$$

一般情形.

- 先假设 $X \geq 0$. 微调位置: $Y = \frac{1}{n}[nX]$.



- 直观: $|X - Y| \leq \frac{1}{n}$, 因此 $|EX - EY| \leq \frac{1}{n}$.
- 计算 EY :

$$EY = \sum_{k=1}^{\infty} \frac{1}{n} P\left(Y \geq \frac{k}{n}\right) = \sum_{\ell=0}^{\infty} \frac{1}{n} P\left(X > \frac{\ell}{n}\right) \rightarrow \int_0^{\infty} G(x) dx.$$

- 或者微调位置: $\hat{Y} = Y - \frac{1}{n}$.



$$E\hat{Y} = \sum_{k=0}^{\infty} \frac{k}{n} \left(F\left(\frac{k+1}{n}\right) - F\left(\frac{k}{n}\right) \right) \rightarrow \int_0^{\infty} x dF(x).$$

- 一般的 X : $\int x dF(x)$.



- Lebesgue-Stieltjes 积分: $\Delta = \max_i \Delta x_i$.

$$\int x dF(x) := \lim_{\Delta \rightarrow 0} \sum_i x_i (F(x_{i+1}) - F(x_i)).$$

- 若 $\int |x| dF(x) < \infty$, 则称

$$\int x dF(x)$$

为 X 的期望, 记为 EX . (期望存在, 定义4.1.3.)

- 定义与离散型、连续型的一致.

- $X \geq 0$, 则

$$EX = \int_0^{\infty} P(X > x) dx.$$

若 EX^{\pm} 不全为 ∞ , 则 $EX := EX^+ - EX^-$. (期望有意义)

- 期望存在:

$$EX^{\pm} < \infty \quad \text{iff} \quad E|X| < \infty \quad \text{iff} \quad \int |x| dF(x).$$

- 若 X 有界(i.e., $P(|X| \leq M) = 1$), 则期望存在.
- 期望是分布的数字特征.

数学期望的性质

- $X \equiv c$, 则 $EX = c$.

- 单调性:

若 $X \geq Y$, 则 $EX \geq EY$.

- 线性:

$$E(aX) = aEX, \quad E(X + Y) = EX + EY.$$

- 若 $X \geq 0$ 且 $EX = 0$, 则 $X = 0$.

(1) $P(X > \frac{1}{n}) = 0$:

$$0 = EX \geq EX1_{\{X > \frac{1}{n}\}} \geq E\frac{1}{n} \cdot 1_{\{X > \frac{1}{n}\}} = \frac{1}{n}P\left(X > \frac{1}{n}\right).$$

(2) $P(X > 0) = \lim_n P\left(X > \frac{1}{n}\right) = 0$.

- 若 $X \geq 0$ 且 $EX < \infty$, 则

$$\lim_{x \rightarrow \infty} xG(x) = \lim_{x \rightarrow \infty} EX \cdot 1_{\{X > x\}} = 0.$$

(1) $xG(x)$:

$$\leq 2 \int_{x/2}^x G(y) dy \leq 2 \int_{x/2}^{\infty} G(y) dy \rightarrow 0.$$

(2) $EX \cdot 1_{\{X > x\}}$:

$$\begin{aligned} &= \int_0^{\infty} P(X > y) dy = \int_0^{\infty} P(X > x, X > y) dy \\ &= \int_0^x P(X > x) dy + \int_x^{\infty} P(X > y) dy. \end{aligned}$$

- $X \geq 0$, 则

$$\lim_{x \rightarrow \infty} EX \cdot 1_{\{X \leq x\}} \rightarrow EX.$$

(1) $EX < \infty$: $EX \cdot 1_{\{X > x\}} \rightarrow 0$.

(2) $EX = \infty$: $\forall M$,

$$\begin{aligned} EX \cdot 1_{\{X \leq x\}} &\geq \int_0^M P(y < X \leq x) dy \\ &\geq \int_0^M P(X > y) dy - MP(X > x) \\ &\rightarrow \int_0^M P(X > y) dy. \quad (x \rightarrow \infty). \end{aligned}$$

令 $M \rightarrow \infty$, 得

$$\lim_{x \rightarrow \infty} EX \cdot 1_{\{X \leq x\}} = \infty.$$

- 函数的期望: $Y = f(X)$, 则 $EY = \int f(x)dF_X(x)$. (定理4.1.1, 了解)
- 离散型:

$$Ef(X) = \sum_k f(x_k)p_k. \quad (4.1.18)$$

- 连续型:

$$Ef(X) = \int f(x)p(x)dx. \quad (4.1.20)$$

- 高维:

$$Ef(\vec{X}) = \sum_k f(\vec{x}_k)p_k, \quad Ef(\vec{X}) = \int f(\vec{x})p(\vec{x})d\vec{x}.$$

- 相互独立则

$$E(XY) = (EX)(EY).$$

- 例. X, Y 相互独立且是连续型, 则

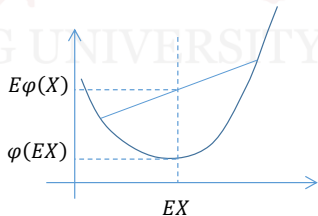
$$EXY = \iint xy p_X(x) p_Y(y) dx dy.$$

- Jensen不等式(4.3.7):

对任意凸函数 φ ,

$$E\varphi(X) \geq \varphi(EX).$$

例. $EX^2 \geq (EX)^2.$



例4.1.11. $X \sim N(\mu, \sigma^2)$, 则 $\mu = EX$.

- $Z \sim N(0, 1)$, 期望存在, $Z \stackrel{d}{=} -Z$, 则

$$EZ = E(-Z) = -EZ \Rightarrow EZ = 0.$$

- $Z := \frac{X-\mu}{\sigma} \sim N(0, 1)$.

$$EX = E(\mu + \sigma Z) = \mu.$$

- 一般地, 若 $X - \mu \stackrel{d}{=} \mu - X$ 且 $E|X| < \infty$, 则 $EX = \mu$.

例4.1.16. 随机数目的期望.

- $X = 1_{A_1} + \cdots + 1_{A_n}$, 则 $EX = \sum_i P(A_i)$.
- $X \sim B(n, p)$ 或 $X \sim H(N, M, n)$, $EX = np$, ($p = \frac{M}{N}$).

$$X = X_1 + \cdots + X_n, \quad X_i \sim B(1, p).$$

- 匹配问题(例1.5.6) n 封信, n 个信封, 平均装对1封.
- 例. 在有 N 个顶点的完全图中, 每条边独立以概率 p 涂红, $1 - p$ 涂蓝, 得到 X 个同色三角形.

$$EX = C_N^3(p^3 + (1 - p)^3).$$

条件期望

假设 $E|Y| < \infty$.

- (X, Y) 是离散/连续型:

$$\varphi(x) := \sum_j y_j P(Y = y_j | X = x), \quad \int y p_{Y|X}(y|x) dy.$$

- (\vec{X}, Y) 是离散/连续型, $E(Y|\vec{X}) = \varphi(\vec{X})$,

$$\varphi(\vec{x}) = \sum_j y_j P(Y = y_j | \vec{X} = \vec{x}), \quad \int y p_{Y|\vec{X}}(y|\vec{x}) dy.$$

- Y 关于 X 的条件期望: $E(Y|X) := \varphi(X)$.
- 条件期望的线性:

$$E(a(X) + b(X)Y|X) = a(X) + b(X)E(Y|X).$$

重期望公式: $EY = EE(Y|X) = E\varphi(X)$. (4.2.59)

- 离散型:

$$\begin{aligned} E\varphi(X) &= \sum_i \varphi(x_i)P(X = x_i) \\ &= \sum_{i,j} y_j P(X = x_i, Y = y_j) = EY. \end{aligned}$$

- 连续型:

$$E\varphi(X) = \int \varphi(x)p_X(x)dx = \iint yp_{X,Y}(x,y)dxdy = EY.$$

- $E\left(E(Y|\vec{X}, \vec{W})|\vec{W}\right) = E(Y|\vec{W})$.

$$\hat{E} = E(\cdot|\vec{W}), \quad E(Y|\vec{X}, \vec{W}) = \hat{E}(Y|\vec{X}).$$

习题二、43. 每个虫卵独立地以概率 p 孵化为幼虫.

虫卵数 X 的期望存在, $Y =$ 幼虫数. 求 EY .

- $\mathcal{L}(Y|X = n) = B(n, p)$:

$$E(Y|X = n) = np \Rightarrow E(Y|X) = Xp.$$

- $EY = E E(Y|X) = pEX$.
- $\xi = \xi_1, \xi_2, \dots$ i.i.d., X 与它们独立, 期望都存在.

$$Y = \xi_1 + \dots + \xi_X \Rightarrow EY = (EX) \cdot (E\xi).$$

- 证: $\mathcal{L}(Y|X = n) = \mathcal{L}(\xi_1 + \dots + \xi_n)$.

例(Polya 坛子). 最初有 b 个黑球, r 个红球. 每次取一个, 放回并放入 c 个同色球. $B_n =$ “第 n 个是黑球”, 则 $P(B_n) = \frac{b}{b+r}$.

• 记 $X_n =$ “ n 次后坛中黑球数”, $Y_n = \frac{X_n}{b+r+nc}$.

• 则

$$P(B_{n+1}|X_n = i) = \frac{i}{b+r+nc} =: y_i.$$

• 从而,

$$\begin{aligned} P(B_{n+1}) &= \sum_i P(X_n = i)P(B_{n+1}|X_n = i) \\ &= \sum_i P(Y_n = y_i)y_i = EY_n. \end{aligned}$$

- $X_{n+1} = X_n + c \cdot \underbrace{1_{B_{n+1}}}$:

$$E(X_{n+1}|X_n) = X_n + c \cdot Y_n = Y_n(b + r + (n + 1)c),$$

$$E(Y_{n+1}|X_n) = Y_n.$$

- 于是,

$$EY_{n+1} = EY_n = \cdots = EY_0 = \frac{b}{b+r}.$$

- $P(B_{n+1}) = EY_n = \frac{b}{b+r}, \forall n \geq 0.$

例. 假设 U_1, U_2, \dots i.i.d. $\sim U(0, 1)$. 记 $S_n = U_1 + \dots + U_n$.
求 EX , 其中

$$X := \inf\{n : S_n \geq 1\}.$$

- 令 $X_a := \inf\{n : S_n \geq a\}$. 记 $f(a) = EX_a$. 待求 $f(1)$.
- 分析 X_a . 记 $\hat{X}_b := \inf\{n - 1 : U_2 + \dots + U_n \geq b\}$, 则
$$X_a = 1_{\{U_1 \geq a\}} + 1_{\{U_1 < a\}}(1 + \hat{X}_{a-U_1}) = 1 + 1_{\{U_1 < a\}} \cdot \hat{X}_{a-U_1}.$$
- \hat{X}_b 与 U_1 相互独立, 且 $\hat{X}_b \stackrel{d}{=} X_b$. 因此,

$$E[1_{\{U_1 < a\}} \cdot \hat{X}_{a-U_1} | U_1 = x] = E(1_{\{x < a\}} \cdot \hat{X}_{a-x}) = 1_{\{x < a\}} \cdot f(a-x).$$

- $f(a) = 1 + \int_0^a f(y)dy, 0 < a \leq 1:$

$$E(1_{\{U_1 < a\}} \cdot f(a - U_1)) = \int_0^a f(a - x)dx.$$

- $f(a) = e^a:$

$$f'(a) = f(a) \Rightarrow (\ln f(a))' = 1 \Rightarrow f(a) = Ce^a.$$

因为 $f(0) = 1$, 所以 $C = 1$.

- $EX = f(1) = e.$

§4.2 方差、相关系数、矩

- 假设 $E|X| < \infty$. 若

$$E(X - EX)^2$$

有限, 则称它为 X 的方差(variance), 记为 $\text{var}(X)$ 或 $D(X)$.

- 称

$$\sigma_X := \sqrt{\text{var}(X)}$$

为 X 的标准差/均方差(standard deviation). (定义4.2.1)

- 矩(moment): $EX^k, E(X - EX)^k, Ee^{aX}$. (定义4.2.5)

- 计算公式: 假设 $E|X| < \infty$,

$$(X - EX)^2 = X^2 - 2(EX) \cdot X + (EX)^2,$$

$$\text{var}(X) = EX^2 - (EX)^2.$$

- 方差有限: $EX^2 < \infty$.
- 方差发散: $E|X| < \infty$, 但 $EX^2 = \infty$
 $\text{var}(X) = E(X - EX)^2 = \infty$.
- 方差是分布的数字特征: 若 $X \stackrel{d}{=} Y$ 则 $\text{var}(X) = \text{var}(Y)$.
- 含义: 权重的分散程度.

若 $\text{var}(X) = 0$, 则 $X = EX$.

- 线性变换:

$$\text{var}(a + bX) = E(bX - bEX)^2 = b^2 \text{var}(X).$$

- 标准化:

$$X^* = \frac{X - \mu}{\sigma}, \quad EX^* = 0, \quad \text{var}(X^*) = 1.$$

例4.2.1. & 例4.2.2.

• $X \sim B(1, p), X^2 = X,$

$$\text{var}(X) = p - p^2 = pq.$$

• $Y \stackrel{d}{=} 1_{A_1} + \cdots + 1_{A_n}.$

(1) $Y^2 = \sum_{i,j} 1_{A_i} \cdot 1_{A_j} = \sum_{i,j} 1_{A_i A_j},$

(2) $EY^2 = \sum_{i,j} P(A_i A_j).$

$$\text{var}(Y) = \sum_{i,j} P(A_i A_j) - \left(\sum_i P(A_i) \right)^2.$$

(3) 进一步, 假设 A_1, \dots, A_n 两两独立. 记 $p_i = P(A_i)$ 则

$$\text{var}(Y) = \sum_i p_i + \sum_{i \neq j} p_i p_j - \sum_{i,j} p_i p_j = \sum_i p_i (1 - p_i) = \sum_i \text{var}(1_{A_i}).$$

例4.2.3. $X \sim P(\lambda)$.

- $EX = \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} \lambda \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} = \lambda.$

- $\text{var}(X) = EX^2 - (EX)^2.$

(1) $EX(X-1)$:

$$\sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{\ell=0}^{\infty} \lambda^2 \frac{\lambda^{\ell}}{\ell!} e^{-\lambda} = \lambda^2.$$

(2) $\star\star = EX(X-1) + EX = \lambda^2 + \lambda,$

$$\text{var}(X) = \star\star - \lambda^2 = \lambda.$$

例4.2.4. $X \sim U(a, b)$.

- $U \sim U(0, 1)$:

$$EU^2 = \int_0^1 x^2 dx = \frac{1}{3}, \quad \text{var}(U) = \frac{1}{12}.$$

- $X \sim U(a, b)$,

$$U = \frac{X - a}{b - a} \sim U(0, 1).$$

- $\text{var}(X)$:

$$\text{var}(a + (b - a)U) = \frac{(b - a)^2}{12}.$$

例4.2.5. $X \sim N(\mu, \sigma^2)$.

● 标准正态分布. $Z := X^* = \frac{X-\mu}{\sigma} \sim N(0, 1)$.

● 二阶矩:

$$\begin{aligned} EZ^2 &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} x^2 e^{-\frac{x^2}{2}} dx = -\frac{2}{\sqrt{2\pi}} \int_0^{\infty} x de^{-\frac{x^2}{2}} \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{x^2}{2}} dx = 1. \end{aligned}$$

● $\text{var}(X) = \text{var}(\mu + \sigma Z) = \sigma^2 \text{var}(Z) = \sigma^2$.

Chebyshev's 不等式 (4.2.10):

$$P(|X - EX| \geq \varepsilon) \leq \frac{\text{var}(X)}{\varepsilon^2}, \quad \forall \varepsilon > 0.$$

证明技巧:

- $P(A) = E1_A$.
- 找 $Y \geq 1_A$, 满足下式, 于是 $P(A) \leq EY$.

$$Y \geq 0; \quad Y|_A \geq 1.$$

- 例. $Y = \frac{(X-EX)^2}{\varepsilon^2}$; $Y = \frac{|X-EX|^r}{\varepsilon^r}$. (Markov 不等式)(5.2.22)
- 例. 若 $X \geq 0$, 则 $P(X \geq C) \leq EX/C$.
- 例. $P(X \geq C) \leq Ee^{a(X-C)}$, 其中 $a > 0$.

协方差、相关系数

- 假设 $EX^2, EY^2 < \infty$. (定义4.2.3) 称

$$E(X - EX)(Y - EY).$$

为 X, Y 的协方差(covariance), 记为 $\text{cov}(X, Y)$ 或 $\sigma_{X, Y}$.

- 和的方差:

$$\begin{aligned}\text{var}(X + Y) &= E\left(\left((X + Y) - (EX + EY)\right)^2\right) \\ &= \text{var}(X) + \text{var}(Y) + 2E(X - EX)(Y - EY).\end{aligned}$$

- 相互独立则 $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$.

- 计算:

$$\begin{aligned} & (X - EX)(Y - EY) \\ &= XY - (EX) \cdot Y - X \cdot (EY) + (EX) \cdot (EY). \end{aligned}$$

- 于是,

$$\text{cov}(X, Y) = EXY - EXEY.$$

- 对称、双线性函数, $\tilde{X} = aX + c$, $\tilde{Y} = bY + d$:

$$\text{cov}(\tilde{X}, \tilde{Y}) = ab \cdot \text{cov}(X, Y).$$

Cauchy-Schwarz's 不等式 (定理4.2.1):

$$(EXY)^2 \leq EX^2 EY^2.$$

- 不妨设 $0 < EX^2, EY^2 < \infty$. 期望存在: $|XY| \leq \frac{1}{2}(X^2 + Y^2)$.
- 二次函数:

$$f(t) = E(tX + Y)^2 = (EX^2)t^2 + 2(EXY)t + EY^2 \geq 0.$$

- 等号成立 iff $f(t_0) = 0$ iff $Y = -t_0X$.

$$t_0 = -\frac{EXY}{EX^2}, \quad f(t_0) = \frac{EX^2 EY^2 - (EXY)^2}{EX^2}.$$

- Hölder's 不等式: $0 < k, \ell < \infty, \frac{1}{k} + \frac{1}{\ell} = 1$,

$$EXY \leq (E|X|^k)^{1/k} (E|Y|^\ell)^{1/\ell}.$$

- 假设 $0 < \sigma_X^2, \sigma_Y^2 < \infty$. 称

$$\rho_{X,Y} := \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}.$$

为 X, Y 的(线性)相关系数, 简记为 ρ . (定义4.2.3)

- 若 $\tilde{X} = aX + c, \tilde{Y} = bY + d$, 则

$$\rho_{\tilde{X}, \tilde{Y}} = \rho_{X,Y} \quad (ab > 0) \quad \text{或} \quad -\rho_{X,Y} \quad (ab < 0).$$

$$\rho_{X,Y} = \rho_{X^*, Y^*} = \text{cov}(X^*, Y^*) = EX^*Y^*.$$

- $|\rho| \leq 1$:

$$\rho = 1 \text{ iff } Y^* = X^* \text{ iff } Y = aX + b, a > 0;$$

$$\rho = -1 \text{ iff } Y^* = -X^* \text{ iff } Y = aX + b, a < 0.$$

$$\rho = EX^*Y^* = \langle X^*, Y^* \rangle = \cos \theta.$$

- 不、正、负相关:

$$\text{cov}(X, Y) = 0, \quad > 0, \quad < 0.$$

- 完全正、负相关 (定义4.2.4):

$$\rho = 1, \quad -1.$$

- 独立则线性不相关.
- 反之不然! 例:

$$X \sim N(0, 1), \quad Y = X^2.$$

例4.2.8. $U \sim U(0, 2\pi)$. $X = \cos U$, $Y = \cos(U + a)$.

- $Y = \cos \hat{U} \stackrel{d}{=} X$, $\hat{U} \sim U(0, 2\pi)$.
- 期望、方差: $EX = 0$, $EX^2 = \frac{1}{2}$.

$$EX^2 = \frac{1}{2\pi} \int_0^{2\pi} \cos^2 \theta d\theta = \frac{1}{2}$$

- 协方差、相关系数. $\rho_{X,Y} = \cos a$:

$$\text{cov}(X, Y) = EXY = \frac{1}{2\pi} \int_0^{2\pi} \cos \theta \cos(\theta + a) d\theta = \frac{1}{2} \cos a.$$

- $a = 0$: $Y = X$, $\rho = 1$, 完全正相关.
- $a = \pi$: $Y = -X$, $\rho = -1$, 完全负相关.
- $a = \frac{\pi}{2}$ 或 $\frac{3\pi}{2}$: $\rho = 0$, 不相关, 但是不独立.

例. 假设 A, B 是(可测)事件.

- A, B 不、正、负相关: $P(AB) =, >, < P(A)P(B)$.

$$\text{cov}(1_A, 1_B) = E1_A1_B - E1_A E1_B = P(AB) - P(A)P(B).$$

- A, B 不相关 iff A 与 B 相互独立. (性质4.2.5)
- 例. $A =$ “产品一是次品”, $B =$ “产品二是次品”.

放回抽样, 不相关(独立) vs 不放回抽样, 负相关.

- $|P(AB) - P(A)P(B)| \leq \frac{1}{4}$. (4.2.31), 习题一45.

$$\rho_{A,B} := \frac{P(AB) - P(A)P(B)}{\sqrt{P(A)(1 - P(A))} \cdot \sqrt{P(B)(1 - P(B))}}. \quad (4.2.30)$$

例. 二元正态. 设 (X, Y) 密度如下, 则 $\rho_{X, Y} = \rho$ (4.2.28).

$$p(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(u^2-2\rho uv+v^2)},$$

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

- 由定理3.2.1, $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$.

故 $U := X^*$, $V := Y^*$ 的密度为 $\hat{p}(u, v)$.

- $\rho_{X, Y} = EX^*Y^* = \rho$. (4.2.28)

$$\begin{aligned} &= \iint u \cdot v \cdot \hat{p}(u, v) du dv \\ &= \int \int u \cdot v \cdot \frac{1}{\sqrt{2\pi}\sqrt{2\pi}\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}((u-\rho v)^2+(1-\rho^2)v^2)} du dv \\ &= \int \rho v \cdot v \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv = \rho EZ^2 = \rho. \end{aligned}$$

- 假设 (X, Y) 服从二元正态分布. 则

X, Y 不相关($\rho = 0$) iff 相互独立($p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$).

- 例4.2.9. $X, Y \sim N(0, 1)$, $\rho_{X,Y} = 0$, 但不独立.

$$p(x, y) = p(x)p(y) + \frac{1}{2\pi}e^{-\pi^2}g(x)g(y)$$

$$p(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \quad g(x) = \cos x \cdot \mathbf{1}_{\{|x| < \pi\}}.$$

- $\vec{X} = (X_1, \dots, X_n)^T$ 的数字特征:

$$\text{期望: } E\vec{X} = (EX_1, \dots, EX_n)^T,$$

$$\text{协方差矩阵: } \Sigma = (\sigma_{ij})_{n \times n}, \quad \sigma_{ij} = \text{cov}(X_i, X_j). \quad (4.2.18).$$

- Σ 半正定: 对称,

$$\begin{aligned} \vec{x}^T \Sigma \vec{x} &= \sum_{i,j} \sigma_{ij} x_i x_j = \sum_{i,j} x_i x_j E(X_i - \mu_i)(X_j - \mu_j) \\ &= E \left(\sum_i x_i (X_i - \mu_i) \right)^2. \end{aligned}$$

- Σ 正定指:

$$\sum_i x_i (X_i - \mu_i) = 0, \quad (\text{a.s.}) \quad \Rightarrow \quad x_1 = \dots = x_n = 0.$$

- 因此, Σ 正定 iff $1, X_1, \dots, X_n$ 线性无关.

最优预测: 假设 $EY^2 < \infty$

最优预测1. 找 $\hat{y} \in \mathbb{R}$ 使得 $Q(\cdot)$ 达到最小,

$$Q(v) := E(Y - v)^2, \quad \forall v \in \mathbb{R}.$$

- 分解:

$$Y - v = (Y - \hat{y}) + (\hat{y} - v), \quad \forall v.$$

- 目标: 交叉项为0.

$$Q(v) = Q(\hat{y}) + (\hat{y} - v)^2 + 2\underbrace{(\hat{y} - v) \cdot E(Y - \hat{y})}_{=0}.$$

- 取 $\hat{y} = EY$ 即可. (4.2.8).
- 上述分解为正交分解, 且取 $v = 0$ 知

$$E(Y - \hat{y})^2 = EY^2 - E\hat{y}^2 = \text{var}(Y).$$

最优预测2. 找 $\hat{a}, \hat{b} \in \mathbb{R}$ 使得

$$Q(a, b) := E(Y - (a + bX))^2, \quad a, b \in \mathbb{R}$$

达到最小. 其中, $EX = 0, EX^2 = 1$.

- $\hat{a} = E(Y - \hat{b}X) = EY$: $Y - (a + bX) = (Y - bX) - a$.
- 令 $Y_0 = Y - EY$, 问题化为找 \hat{b} 使得 $f(\cdot)$ 达到最小.

$$f(b) = Q(\hat{a}, b) = E(Y_0 - bX)^2.$$

- (正交)分解及其交叉项.

$$Y_0 - bX = (Y_0 - \hat{b}X) \oplus (\hat{b} - b)X.$$

$$f(b) = f(\hat{b}) + (\hat{b} - b)^2 + \underbrace{2(\hat{b} - b) \cdot E(Y_0 - \hat{b}X)X}_{=0}.$$

- 取 $\hat{b} = EXY_0 = \text{cov}(X, Y)$ 即可.
- $E(Y - \hat{a} - \hat{b}X)^2 = EY_0^2 - E(\hat{b}X)^2 = (1 - \rho_{X,Y}^2) \sigma_Y^2$.

特别地, $\rho_{X,Y} = \pm 1$ 当且仅当 $Y = \hat{a} + \hat{b}X$.

最优预测3*. 找 $\hat{\psi}(\cdot)$ 使得 $Q(\cdot)$ 达到最小.

$$Q(\psi) = E(Y - \psi(X))^2, \quad \psi(\cdot) \text{ 使得 } E\psi(X)^2 < \infty.$$

- 正交分解:

$$Y - \psi(X) = (Y - \hat{\psi}(X)) \oplus (\hat{\psi}(X) - \psi(X)).$$

- 根据条件期望的线性与重期望公式,

$$E\star\star = E(E(\star\star|X)) = E\left(\star(E(Y|X) - \hat{\psi}(X))\right).$$

- 取 $\hat{\psi}(X) = E(Y|X) = \varphi(X)$ 即可. 进一步,

$$E(Y - \varphi(X))^2 = EY^2 - E\varphi(X)^2.$$

§4.4 母函数

注：在本节中，随机变量都取非负整数值。

- 设 $P(X = k) = p_k, k \geq 0$. 称

$$\sum_{k=0}^{\infty} p_k s^k, \quad s \in [-1, 1]$$

为 X 的母函数(generating function), 记为 $g_X(s)$ 或 $g(s)$.

- 表达式:

$$g(s) = E s^X = p_0 + p_1 s + p_2 s^2 + \cdots + p_k s^k + \cdots$$

- 母函数刻画分布: $X \stackrel{d}{=} Y$ iff $g_X = g_Y$.

$$g^{(k)}(0) = p_k \cdot k!, \quad \forall k \geq 0.$$

- $g(s) = p_0 + p_1s + p_2s^2 + \cdots + p_k s^k + \cdots = E s^X.$

- 矩: 对 $s \in (-1, 1)$,

$$g'(s) = p_1 + 2p_2s + \cdots + k p_k s^{k-1} + \cdots = EX s^{X-1},$$

$$g''(s) = 2p_2 + \cdots + k(k-1)p_k s^{k-2} + \cdots = EX(X-1)s^{X-2}.$$

$$g^{(\ell)}(s) = EX(X-1)\cdots(X-\ell+1)s^{X-\ell}.$$

- 例,

$$EX = g'(1) := \lim_{s \rightarrow 1^-} g'(s), \quad (EX < \infty \text{ or } EX = \infty \text{ 都成立})$$

$$EX^2 = g''(1) + g'(1).$$

- 例4.4.4. $X \sim G(p).$

$$g(s) = \sum_{k=1}^{\infty} q^{k-1} p s^k = \frac{ps}{1-qs}.$$

- 乘积: 若 X 与 Y 独立, 则 $g_{X+Y}(s) = g_X(s)g_Y(s)$.

$$E s^{X+Y} = E s^X s^Y = E s^X \cdot E s^Y.$$

- 例4.4.5. $X \sim B(n, p)$.

$$g_n(s) = (q + ps)^n.$$

- 例4.4.6. $X \sim P(\lambda)$.

$$g_\lambda(s) = \sum_{k=0}^{\infty} s^k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda(s-1)}.$$

- 复合: 设 $\xi = \xi_1, \xi_2, \dots$ i.i.d., 且它们与 W 独立.

令 $X = \xi_1 + \dots + \xi_W$, 则 $g_X = g_W \circ g_\xi$.

- (1) 条件期望:

$$E(s^X | W = k) = E(s^{\xi_1 + \dots + \xi_k} | W = k) = E_\star = g_\xi(s)^k.$$

- (2) 重期望公式:

$$E s^X = E(E(s^X | W)) = E(g_\xi(s)^W) = g_W(g_\xi(s)) = g_W \circ g_\xi(s).$$

- 期望:

$$EX = g'_X(1) = g'_W(g_\xi(1))g'_\xi(1) = EW \cdot E\xi.$$

$$EX = E(E(X|W)) = E(W \cdot E\xi) = EW \cdot E\xi.$$

- 复合泊松分布: 若 $W \sim P(\lambda)$, 则 $X = \xi_1 + \cdots + \xi_W$ 的母函数为

$$g_X(s) = \exp \{ \lambda (g_\xi(s) - 1) \}.$$

称 X 服从复合泊松分布.

- 例4.4.10. $W \sim P(\lambda)$, $\xi \sim B(1, p)$.

$$g_X(s) = \exp \{ \lambda (q + ps - 1) \} = e^{\lambda p(s-1)}.$$

- 凸组合: X, Y, ξ 相互独立, $\xi \sim B(1, p)$. 令

$$W = X \cdot 1_{\{\xi=1\}} + Y \cdot 1_{\{\xi=0\}},$$

则 $g_W = p \cdot g_X + (1 - p) \cdot g_Y$.

§4.5 特征函数

- 称

$$Ee^{itX} = E \cos(tX) + \sqrt{-1}E \sin(tX), \quad \forall t \in \mathbb{R}$$

为 X 的特征函数(characteristic function), 记为 $f_X(t)$ 或 $f(t)$.

- $x + iy = (x, y)$. $e^{itX} = (\cos(tX), \sin(tX))$.
- 基本性质1. $f(0) = 1$.
- $\|f(t)\| = \|Ee^{itX}\| \leq E\|e^{itX}\| = 1$.

$\varphi : x + iy \mapsto \|x + iy\| = \sqrt{x^2 + y^2}$ 是凸函数.

● 基本性质2. f 一致连续:

(1) $\|\cdot\|$ 凸:

$$\|f(t + \varepsilon) - f(t)\| \leq E\|e^{i(t+\varepsilon)X} - e^{itX}\| = E\|e^{i\varepsilon X} - 1\|.$$

(2) $\forall M > 0$,

$$EY = EY \cdot 1_{\{|X| \leq M\}} + EY \cdot 1_{\{|X| > M\}}.$$

(3) 取 M 使得 $P(|X| > M) < \frac{\delta}{4}$, 则

$$** \leq 2 \cdot P(|X| > M) < \frac{\delta}{2},$$

$$** \leq \max_{|x| \leq M} \|e^{i\varepsilon x} - 1\| \leq \varepsilon M < \frac{\delta}{2}, \quad (\text{取 } \varepsilon < \frac{\delta}{2M}).$$

● 基本性质3. f 半正定:

$\forall t_1, \dots, t_n \in \mathbb{R}$, 令 $a_{kj} = f(t_k - t_j)$, 则

$\mathbf{A} = (a_{kj})_{n \times n}$ 半正定,

i.e. (i) $\bar{\mathbf{A}}^T = \mathbf{A}$, (ii) $\sum_{k,j} a_{kj} \lambda_k \bar{\lambda}_j \geq 0, \forall \lambda_1, \dots, \lambda_n \in \mathbb{C}$.

● 验证(i): 记 $t = t_j - t_k$.

$$f(-t) = Ee^{i(-t)X} = E\overline{e^{itX}} = \overline{f(t)}. \quad (4.5.13)$$

● 验证(ii):

$$\sum_{k,j} a_{kj} \lambda_k \bar{\lambda}_j = \sum_{k,j} Ee^{i(t_k - t_j)X} \lambda_k \bar{\lambda}_j = E \left\| \sum_j \lambda_j e^{it_j X} \right\|^2 \geq 0.$$

- Bochner-Khinchine定理(定理5.2.9): 若 $f : \mathbb{R} \rightarrow \mathbb{C}$ 满足 $f(0) = 1$, 连续, 半正定, 则存在 X 使得 $f = f_X$.
- 逆转公式& 唯一性(定理4.5.1~ 4.5.2):

$$F(x) - F(y) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ity} - e^{-itx}}{it} f(t) dt, \quad \forall x, y \in C(F).$$

- (定理4.5.3) 假设 $\int \|f(t)\| dt < \infty$. 则相应分布函数 F 连续可导, 且

$$p(x) = F'(x) = \frac{1}{2\pi} \int e^{-itx} f(t) dt. \quad (4.5.25).$$

- 性质4. 乘积: X 与 Y 独立, 则 $f_{X+Y}(t) = f_X(t)f_Y(t)$.

(1) $X \sim B(n, p)$, $f_n(t) = (q + pe^{it})^n = (1 + p(e^{it} - 1))^n$.

(2) $X \sim P(\lambda)$, $f_\lambda(t) = e^{\lambda(e^{it}-1)}$.

- 凸组合: 设 X, Y, ξ 相互独立, $\xi \sim B(1, p)$.

令 $W = X\xi + Y(1 - \xi)$, 则 $f_W = pf_X + (1 - p)f_Y$.

- 性质5. 若 EX^k 存在, 则 $f^{(k)}(0) = i^k EX^k$, 且

$$f(t) = 1 + f'(0)t + \frac{f''(0)}{2!}t^2 + \cdots + \frac{f^{(k)}(0)}{k!}t^k + o(t^k).$$

- 性质6. $f_{aX+b}(t) = Ee^{itaX+itb} = e^{itb} f_X(at)$.

例4.5.5. $X \sim N(\mu, \sigma^2)$.

- $Z = X^* \sim N(0, 1)$:

$$\begin{aligned} f_Z(t) &= \frac{1}{\sqrt{2\pi}} \int e^{-\frac{x^2}{2} + itx} dx \\ &= \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}(x-it)^2 + \frac{1}{2}(it)^2} dx = e^{-\frac{t^2}{2}}. \end{aligned}$$

(严格计算见书)

- $X = \mu + \sigma Z$:

$$f_X(t) = Ee^{it(\mu + \sigma Z)} = e^{it\mu - \frac{1}{2}\sigma^2 t^2}.$$

随机向量的特征函数:

- 联合特征函数:

$$f_{\vec{X}}(\vec{t}) = Ee^{i\vec{t}\cdot\vec{X}} = Ee^{i(t_1X_1+\cdots+t_nX_n)}.$$

- 逆转公式: $f_{\vec{X}}(\vec{t}) \rightarrow F_{\vec{X}}(\vec{x})$.
- 若干性质(见书).
- 边缘: 例,

$$f_X(t) = f_{X,Y}(t, 0).$$

- X 与 Y 独立 iff $f_{X,Y}(t, s) = f_X(t)f_Y(s)$.
 - X 与 Y 独立 $\Rightarrow f_{X+Y}(t) = f_X(t)f_Y(t)$.
- 反之不然(习题四、50).

§4.6 多元正态分布

- $\vec{\mu} \in \mathbb{R}^n$ (列向量), Σ 为 $n \times n$ 的正定矩阵. 记

$$\vec{X} = (X_1, \dots, X_n)^T \sim N(\vec{\mu}, \Sigma).$$

- 联合密度:

$$p_{\vec{X}}(\vec{x}) = \frac{1}{\sqrt{2\pi}^n \sqrt{|\Sigma|}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu})^T \Sigma^{-1} (\vec{x} - \vec{\mu}) \right\}.$$

- n 维标准正态 $\vec{Z} = (Z_1, \dots, Z_n)^T \sim N(\vec{0}, \mathbf{I})$. 联合密度:

$$p_{\vec{Z}}(\vec{z}) = \frac{1}{\sqrt{2\pi}^n} \exp \left\{ -\frac{1}{2} \|\vec{z}\|^2 \right\}.$$

等价地, Z_1, \dots, Z_n i.i.d., $\sim N(0, 1)$.

- 设 $\vec{Z} \sim N(\vec{0}, \mathbf{I})$, \mathbf{A} 非退化, 则

$$\vec{Y} := \vec{\mu} + \mathbf{A}\vec{Z} \sim N(\vec{\mu}, \Sigma), \quad (\Sigma = \mathbf{A}\mathbf{A}^T.)$$

- 证: $p_{\vec{Y}}(\vec{y}) = p_{\vec{Z}}(\vec{z}) \left| \frac{\partial \vec{z}}{\partial \vec{y}} \right| = C \exp\{-\frac{1}{2}\|\vec{z}\|^2\}$.

$$\begin{aligned} \|\vec{z}\|^2 &= \|\mathbf{A}^{-1}(\vec{y} - \vec{\mu})\|^2 = (\vec{y} - \vec{\mu})^T \mathbf{A}^{-1,T} \mathbf{A}^{-1} (\vec{y} - \vec{\mu}) \\ &= (\vec{y} - \vec{\mu})^T (\mathbf{A}\mathbf{A}^T)^{-1} (\vec{y} - \vec{\mu}). \end{aligned}$$

- 设 $\vec{X} \sim N(\vec{\mu}, \Sigma)$. 取 $\mathbf{A} = \sqrt{\Sigma}$. 则

$$\vec{X} \stackrel{d}{=} \vec{Y} \Rightarrow \vec{V} := \mathbf{A}^{-1}(\vec{X} - \vec{\mu}) \stackrel{d}{=} \vec{Z}.$$

- 因此, $\exists \vec{V}$ 满足:

$$\vec{V} \sim N(\vec{0}, \mathbf{I}), \quad \vec{X} = \vec{\mu} + \mathbf{A}\vec{V}.$$

- 进一步, \vec{X} 的非退化线性变换都服从 n 维正态分布.

高斯分布与特征函数

- $N(\vec{0}, \mathbf{I})$:

$$f_{\vec{Z}}(\vec{t}) = Ee^{i\vec{t}\cdot\vec{Z}} = \prod_{k=1}^n Ee^{it_k Z_k} = \prod_{k=1}^n e^{-t_k^2/2} = e^{-\frac{1}{2}\|\vec{t}\|^2}.$$

- $N(\vec{\mu}, \Sigma)$: $\vec{X} = \vec{\mu} + \mathbf{A}\vec{Z}$, $\mathbf{A} = \sqrt{\Sigma}$, $\mathbf{A}\mathbf{A}^T = \Sigma$.

$$\begin{aligned} f_{\vec{X}}(\vec{t}) &= Ee^{i\vec{t}\cdot(\vec{\mu} + \mathbf{A}\vec{Z})} = e^{i\vec{t}\cdot\vec{\mu}} Ee^{i(\mathbf{A}^T\vec{t})\cdot\vec{Z}} \\ &= e^{i\vec{t}\cdot\vec{\mu}} e^{-\frac{1}{2}\|\mathbf{A}^T\vec{t}\|^2} = \exp\left\{i\vec{t}\cdot\vec{\mu} - \frac{1}{2}\vec{t}^T \Sigma \vec{t}\right\}. \end{aligned}$$

- 设 $\vec{\mu} \in \mathbb{R}^n$, $\Sigma_{n \times n}$ **半正定**. 若 \vec{X} 的联合特征函数为 \star , 则称 \vec{X} 服从高斯分布 $N(\vec{\mu}, \Sigma)$. 也称 \vec{X} 为一个高斯向量.
- $\vec{\mu} + \mathbf{A}\vec{Z} \sim N(\vec{\mu}, \Sigma)$, $\mathbf{A} = \sqrt{\Sigma}$. Σ 非退化 vs 退化.
- 高斯向量的任意线性变换仍服从高斯分布.

数字特征

- 假设 $\vec{X} \sim N(\vec{\mu}, \Sigma)$. (其中 Σ 半正定.)
- $\vec{X} \stackrel{d}{=} \vec{\mu} + \mathbf{A}\vec{Z} =: \vec{Y}$, 其中 $\mathbf{A} = (a_{ij})_{n \times n} = \sqrt{\Sigma}$.
- 期望: $E\vec{X} = \vec{\mu}$.

$$EX_i = EY_i = E(\mu_i + a_{i1}Z_1 + \cdots + a_{in}Z_n) = \mu_i.$$

- 协方差矩阵: $(\sigma_{ij}) = (\text{cov}(X_i, X_j)) = \Sigma$.

$$\begin{aligned}\text{cov}(X_i, X_j) &= E \sum_k a_{ik} \underline{Z_k} \sum_\ell a_{j\ell} \underline{Z_\ell} = \sum_{k,\ell} a_{ik} a_{j\ell} \times E \underline{Z_k Z_\ell} \\ &= \sum_k a_{ik} a_{jk} = \sigma_{ij}.\end{aligned}$$

定理 (定理4.6.6)

$\vec{X} \sim N(\vec{\mu}, \Sigma)$ iff $\forall a_1, \dots, a_n \in \mathbb{R}, Y := \sum_{k=1}^n a_k X_k \sim N(\mu, \sigma^2)$.

• \Rightarrow :

$$Ee^{itY} = Ee^{i(ta_1, \dots, ta_n) \cdot \vec{X}} = \exp \left\{ it \sum_{k=1}^n a_k \mu_k - \frac{1}{2} t^2 (\vec{a}^T \Sigma \vec{a}) \right\}.$$

• \Leftarrow : (1) 数字特征: 根据假设有 $EX_i^2 < \infty$. 令 $E\vec{X} = \vec{\mu}$; \vec{X} 的协方差阵为 Σ . 从而

$$\mu = EY = \sum_{k=1}^n a_k \mu_k, \quad \sigma^2 = \text{cov}(Y, Y) = \vec{a}^T \Sigma \vec{a}.$$

• (2) 特征函数:

$$\begin{aligned} Ee^{i\vec{t} \cdot \vec{X}} &= Ee^{i \cdot 1 \cdot \sum_{k=1}^n t_k X_k} = \exp \left\{ i\mu - \frac{1}{2} \sigma^2 \right\} \quad (\vec{a} = \vec{t}) \\ &= \exp \left\{ i\vec{t} \cdot \vec{\mu} - \frac{1}{2} \vec{t}^T \Sigma \vec{t} \right\}. \end{aligned}$$

- 改写:

$$\vec{X} = (Y_1, \dots, Y_r; W_{r+1}, \dots, W_n)^T,$$

$$\vec{\mu} = (\nu_1, \dots, \nu_r; w_{r+1}, \dots, w_n)^T.$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

- $\vec{Y} \sim N(\vec{\nu}, \Sigma_{11})$:

$$f_{\vec{Y}}(\vec{s}) = f_{\vec{X}}(\vec{s}; \vec{0}) = \exp \left\{ i \vec{s} \cdot \vec{\nu} - \frac{1}{2} \vec{s}^T \Sigma_{11} \vec{s} \right\}.$$

- Y_i 与 W_k 不相关, $\forall i \leq r < k$ (即 $\Sigma_{12} = 0$) $\Leftrightarrow \vec{Y}$ 与 \vec{W} 独立:

$$\Sigma_{12} = 0 \quad \Rightarrow \quad f_{\vec{X}}(\vec{s}; \vec{u}) = f_{\vec{Y}}(\vec{s}) f_{\vec{W}}(\vec{u}).$$

$\Sigma_{n \times n}$ 退化的情形. 假设 $\vec{X} \sim N(\vec{0}, \Sigma)$.

- 结论: 存在 $\vec{V} \sim N(\vec{0}, \mathbf{I}_{n \times n})$ 以及 $\mathbf{M}_{n \times n}$ 使得 $\vec{X} = \mathbf{M}\vec{V}$.
- 记 $\mathbf{A} = (a_{ij})_{n \times n} = \sqrt{\Sigma}$, 取 $\vec{Z} \sim N(\vec{0}, \mathbf{I}_{n \times n})$, 则

$$\vec{X} \stackrel{d}{=} \mathbf{A}\vec{Z} =: (Y_1, \dots, Y_r; W_{r+1}, \dots, W_n)^T.$$

- $r = \text{rank}(\mathbf{A}) \geq 1$. 令 $\vec{\alpha}_i = (a_{i1}, \dots, a_{in})$.
不妨设前 r 行 $\vec{\alpha}_1, \dots, \vec{\alpha}_r$ 线性无关, 且

$$\vec{\alpha}_k = b_{k1}\vec{\alpha}_1 + \dots + b_{kr}\vec{\alpha}_r, \quad k = r+1, \dots, n.$$

- $\vec{W} = \mathbf{B}_{(n-r) \times r} \vec{Y}$: $Y_i = \vec{\alpha}_i \cdot \vec{Z}, i \leq r,$

$$W_k = \vec{\alpha}_k \cdot \vec{Z} = b_{k1}Y_1 + \dots + b_{kr}Y_r, \quad k = r+1, \dots, n.$$

- $X_k \stackrel{\text{a.s.}}{=} b_{k1}X_1 + \cdots + b_{kr}X_r, \quad k = r+1, \cdots, n.$

$$X_k - (b_{k1}X_1 + \cdots + b_{kr}X_r) \stackrel{d}{=} W_k - (b_{k1}Y_1 + \cdots + b_{kr}Y_r) = 0.$$

- $(X_1, \cdots, X_r)^T \stackrel{d}{=} \vec{Y} = (Y_1, \cdots, Y_r)^T = \hat{\mathbf{A}}\vec{Z}$ 服从正态分布:

因为 $\vec{\alpha}_1, \cdots, \vec{\alpha}_r$ 线性无关, 所以

$$\hat{\Sigma}_{r \times r} = \hat{\mathbf{A}}\hat{\mathbf{A}}^T = \Sigma_{11} \quad \text{满秩}, \quad \hat{\mathbf{A}}_{r \times n} = (a_{ij})_{1 \leq i \leq r, 1 \leq j \leq n}.$$

- 利用 Σ 非退化情形的结果知,

$$(X_1, \cdots, X_r)^T = \mathbf{C}_{r \times r}\vec{V}, \quad \text{其中, } \vec{V} \sim N(\vec{0}, \mathbf{I}_{r \times r}).$$

- 结论: $\vec{V} = (V_1, \cdots, V_r; V_{r+1}, \cdots, V_n)^T \sim N(\vec{0}, \mathbf{I}_{n \times n})$

$$\vec{X} = \mathbf{M}\vec{V}, \quad \text{其中, } \mathbf{M} = \begin{pmatrix} \mathbf{C} & \vec{0} \\ \mathbf{BC} & \vec{0} \end{pmatrix}.$$

条件分布

- 假设 $\vec{X} = (Y_1, \dots, Y_r; W_{r+1}, \dots, W_n)^T \sim N(\vec{\mu}, \Sigma)$.

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

- 假设 \vec{Y} 服从正态分布, 即 Σ_{11} 非退化. 求 $\mathcal{L}(\vec{W} | \vec{Y} = \vec{y})$.
- 不妨假设 $E\vec{X} = 0$, 否则考虑 $\vec{X} - E\vec{X}$.
- 目标: 找 $\mathbf{B}_{(n-r) \times r}$ 使得 $\vec{V} = (\vec{W} - \mathbf{B}\vec{Y})$ 与 \vec{Y} 不相关, 即

$$\text{cov}(Y_i, V_k) = EY_i V_k = 0, \quad i \leq r < k.$$

- 解释: $(Y_1, \dots, Y_r; V_{r+1}, \dots, V_n)^T$ 服从高斯分布, 故**表明 \vec{Y} 与 \vec{V} 相互独立.

- 协方差: $i \leq r < k$,

$$\begin{aligned} \text{cov}(V_k, Y_i) &= E \left(W_k - \sum_{j \leq r} b_{kj} Y_j \right) Y_i \\ &= \sigma_{ki} - \sum_{j \leq r} b_{kj} \sigma_{ji} = (\boldsymbol{\Sigma}_{21} - \mathbf{B}\boldsymbol{\Sigma}_{11})_{ki}. \end{aligned}$$

- 取 $\mathbf{B} = \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}$ 即可. 于是 \vec{Y} 与 \vec{V} 独立, 且 $\vec{W} = \mathbf{B}\vec{Y} + \vec{V}$.
- $\vec{V} = \vec{W} - \mathbf{B}\vec{Y} \sim N(\vec{0}, \tilde{\boldsymbol{\Sigma}}_{22})$, 其中

$$EV_k V_\ell = E \left(W_k - \sum_{j \leq r} b_{kj} Y_j \right) V_\ell = EW_k \left(W_\ell - \sum_{j \leq r} b_{\ell j} Y_j \right)$$

$$\Rightarrow \tilde{\boldsymbol{\Sigma}}_{22} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}.$$

- 结论: 在 $\vec{Y} = \vec{y}$ 的条件下, $\vec{W} = \mathbf{B}\vec{y} + \vec{V}$, 故

$$\mathcal{L}(\vec{W} | \vec{Y} = \vec{y}) = N(\mathbf{B}\vec{y}, \tilde{\boldsymbol{\Sigma}}_{22}).$$