1. Expectations and Covariance matrices of random vectors

Definition 1 (随机向量的期望与协方差) 对 $\forall\ S \in \mathbb{R}^p,\ T \in \mathbb{R}^q$，记：

$$E(S) = [E(S_{[1]}),\ E(S_{[2]}),\ \cdots,\ E(S_{[p]})]^T$$

为向量 S 的期望 (expectation)

$$\mathrm{cov}(S,T) = E\{(S-E(S))(T-E(T))^T\} \in \mathbb{R}^{p\times q}$$

为向量 S 和 T 的协方差矩阵，特别地，记

$$\mathrm{cov}(T) = \mathrm{cov}(T,T) = E(TT^T) - E(T)\cdot E(T)^T$$

为 T 的协方差矩阵 (covariance matrix)

Proposition 1 (期望、协方差性质)

$C_1 \in \mathbb{R}^{c_1\times p}$ $\quad C_2 \in \mathbb{R}^{c_2\times q}$ $\quad S \in \mathbb{R}^p$ $\quad T \in \mathbb{R}^q$

(a) $E(C_1 S) = C_1 E(S)$

(b) 若 $p = q$ 则 $E(S+T) = E(S) + E(T)$

(c) $\mathrm{cov}(S,T) = [\mathrm{cov}(T,S)]^T$

(d) $\mathrm{cov}(T)$ 是半正定的 (positive semi-definite, PSD)

(e) $\mathrm{cov}(C_1 S, C_2 T) = C_1\cdot \mathrm{cov}(S,T)\cdot C_2^T \in \mathbb{R}^{c_1\times c_2}$

(f) 若 $c_1 = c_2$ 则

$$\mathrm{cov}(C_1\cdot S + C_2\cdot T) = \mathrm{cov}(C_1 S) + \mathrm{cov}(C_2 T) + \mathrm{cov}(C_1 S, C_2 T) + \mathrm{cov}(C_2 T, C_1 S)$$

2. Multivariate Normal Distributions (MNDs)

Definition 2 (特征函数 characteristic function)

$$\varphi_X(t) = E\{\exp(it^TX)\} \quad \forall t \in \mathbb{R}^n \quad i^2 = -1$$

如果随机向量 X 满足如下特征函数：

$$\varphi_X(t) = E\{\exp(it^TX)\} = \exp(i\mu^Tt - \frac{1}{2} \cdot t^T \Sigma t)$$

其中 $t \in \mathbb{R}^n$ $X \in \mathbb{R}^n$ $i^2 = -1$。则称 X 服从 n维正态分布 (n-dimentional normal distribution)，期望为 $\mu \in \mathbb{R}$，协方差矩阵为 $\Sigma \in \mathbb{R}^{n\times n}$，记作 $X \sim N(\mu, \Sigma)$

Remark：Σ 为半正定 (PSD) 即可

Theorem 1 (X 服从正态 $\Leftrightarrow$ X 所有线性组合服从正态) 下面二者等价：

(a) 随机向量 X 服从多元正态分布 (n维正态分布)

(b) 对 $\forall$ $v \in \mathbb{R}^n$，随机变量 $v^TX \in \mathbb{R}$ 均服从正态分布

$$(a) \Leftrightarrow (b)$$

Proposition 2 (密度函数) 随机向量 $X \in \mathbb{R}^n$ 服从 $X \sim N(\mu, \Sigma)$

则有 $E(X) = \mu$ $\quad \mathrm{cov}(X) = \Sigma$，且当 Σ 正定时，X 的概率密度函数为：

$$f(x) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \cdot \exp\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\} \quad x \in \mathbb{R}^n$$

概率密度函数 $p(x) = \frac{1}{(2\pi)^n} \cdot \int_{\mathbb{R}^n} \varphi_X(t) \cdot \exp(-it^TX)\, dt$

Remark 正态随机向量的分布完全取决于 μ 和 Σ

Theorem 2 (closedness under linear operations) 已知 $X \sim N(\mu, \Sigma)$ 对 $\forall$ 确定的 $A \in \mathbb{R}^{s\times n}$ 和 $b \in \mathbb{R}^{n}$，均有

$$AX + b \sim N(A\mu + b, A\Sigma A^T)$$

下面讨论独立性、边际分布、条件分布

Proposition 3 (独立性 independence) 随机向量 X_1 和 X_2，构造 $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ 及任意 $t = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$ 则有

$$X_1 \perp X_2 \iff \varphi_X(t) = \varphi_{X_1}(t_1)\,\varphi_{X_2}(t_2)$$

Remark：如果 X_1, X_2 形状相同，则有若 $X_1 \perp X_2$，则有

$$\varphi_{X_1+X_2}(t_1) = \varphi_{X_1}(t_1)\,\varphi_{X_2}(t_2)$$

反之不成立

Theorem 3 (边际分布 marginal distribution) 已知 X 是 X_1, X_2 的联合分布 (joint distribution) 服从如下联合正态分布：

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$$

则有 X_1, X_2 的边际分布：

$$X_1 \sim N(\mu_1, \Sigma_{11}) \qquad X_2 \sim N(\mu_2, \Sigma_{22})$$

更进一步，如果 $\Sigma_{12} = \Sigma_{21}^T = 0$，则有 $X_1 \perp X_2$

Remark： 边际分布 $X_1 \sim N(\mu_1, \Sigma_{11})$ $X_2 \sim N(\mu_2, \Sigma_{22})$

且 $cov(X_1, X_2) = 0$

无法推出二者独立

Theorem 4 (条件分布 conditional distribution) 已知 X 是 X_1, X_2 的联合分布 (joint distribution) 服从如下联合正态分布：

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

且 Σ_{22} 可逆，则

$$X_1 | X_2 \sim N(\mu_1 + \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot (X_2 - \mu_2), \quad \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21})$$

即给定 X_2 时 X_1 的分布.

Remark 如上的假定可视为 X_1 (response)，X_2 (covariate) 的线性模型

$$X_1 = a_0 + b_0 X_2 + \varepsilon$$

其中： $a_0 = \mu_1 - \Sigma_{12} \Sigma_{22}^{-1} \mu_2$

$b_0 = \Sigma_{12} \Sigma_{22}^{-1}$

随机误差 $\varepsilon \sim N(0, \Sigma_{11} - \Sigma_{12} \cdot \Sigma_{22}^{-1} \cdot \Sigma_{21})$

3. Quadratic forms of random vectors

Definition 3 (二次型) 对 $\forall A \in \mathbb{R}^{n\times n}$，标量 X^TAX 称为矩阵 A 的二次型 ($X \in \mathbb{R}^n$)

Lemma 1 (迹的交换) 对于 $\forall C_1 \in \mathbb{R}^{p\times q}$ 和 $C_2 \in \mathbb{R}^{q\times p}$ 均有

$$tr(C_1C_2) = tr(C_2C_1)$$

Remark: 只能交换 2 个矩阵，$tr(ABC) \neq tr(CBA)$

Theorem 5 (二次型期望) 已知 $A \in \mathbb{R}^{n\times n}$，随机向量 $X \in \mathbb{R}^n$，已知 $E(X) = \mu$ $\quad cov(X) = \Sigma$ 则有

$$E(X^TAX) = \mu^TA\mu + tr(A\Sigma)$$

下面，讨论正态分布的二次型相关

Theorem 6 $A \in \mathbb{R}^{n\times n}$ 是实对称矩阵，$B \in \mathbb{R}^{p\times n}$ 且 $BA = 0$，如果随机向量 $X \sim N(\mu, \sigma^2 \cdot I)$ $\sigma \in \mathbb{R}^+$。则有 $BX \perp X^TAX$

Theorem 7 $A, B \in \mathbb{R}^{n\times n}$ 均为实对称阵，满足 $AB = 0$，如果随机向量 $X \sim N(\mu, \sigma^2 \cdot I)$ $\sigma \in \mathbb{R}^+$。则有 $X^TAX \perp X^TBX$

4. Chi-Squre distribution (χ^2)

Definition 4 若随机变量 $X_1, X_2, \cdots, X_n \sim N(0,1)$ 且相互独立。记

$Y = X_1^2 + X_2^2 + \cdots + X_n^2$，服从自由度为 n 的 χ^2 分布 (chi-square)

$$Y \sim \chi^2(n)$$

Proposition 4 ☆

(a) 正态随机向量的二次型：

$X \in \mathbb{R}^n$ 且 $X \sim N(\mu, \Sigma)$ 且 Σ 可逆，则有

$$(X-\mu)^T \cdot \Sigma^{-1} \cdot (X-\mu) \sim \chi^2(n)$$

(b) χ^2 分布的期望和方差：

若 $Y \sim \chi^2(n)$ 则

$$E(Y) = n \qquad var(Y) = 2n$$

(c) χ^2 分布的独立可加性：

若 $Y_1 \sim \chi^2(n_1)$ $Y_2 \sim \chi^2(n_2)$ 且 $Y_1 \perp Y_2$ 则有

$$Y_1 + Y_2 \sim \chi^2(n_1+n_2)$$

下面，讨论正态随机向量二次型 & 幂等矩阵 (idempotent)

Lemma 2 (幂等矩阵的特征值) 若矩阵 C 满足 $CC = C^2 = C$

则它的特征值一定为 0 或 1

Theorem 8 已知 $X \sim N(\mu, I)$，有一实对称阵 $A \in R^{n\times n}$ 满足

$A^2 = A$ $\mu^T A \mu = 0$，记 $r = rank(A) > 0$，则有

$$X^T A X \sim \chi^2(r)$$

If $X \sim N(\mu, \sigma^2 I)$, $A \in R^{n\times n}$ satisfying $A^2 = A$, $\mu^T A\mu = 0$, then

$$\frac{X^T A X}{\sigma^2} \sim \chi^2(r), \text{ where } r = rank(A)$$

If $A \in R^{n\times n}$, and denote $\mu = E(X)$, $\Sigma = cov(X)$, then we have

$$E(X^T A X) = \mu^T A \mu + tr(A\Sigma)$$