Wishart 分布

1. 引入与动机

在统计学中，我们经常需要处理样本协方差矩阵的分布。

- 在一元（单变量）情况下，如果样本 $x_1, \ldots, x_n$ 来自正态分布 $N(\mu, \sigma^2)$，那么样本方差 $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ 的一个重要性质是 $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$（卡方分布，自由度为 $n-1$)。
- Wishart 分布可以看作是卡方分布在多元（多变量）情况下的推广。它描述的是**样本协方差矩阵（或与之成比例的样本离差阵）的抽样分布**，当原始数据来自多元正态分布时。

2. Wishart 分布的定义

有两种常见的定义方式，一种是基于零均值正态向量的和，另一种是基于样本离差阵。

定义 1: 基于独立同分布的零均值正态向量 (中心Wishart分布)

定义 设 $Z_1, Z_2, \ldots, Z_k$ 是 k 个独立的 p 维随机列向量，并且每一个 Z_a 都服从多元正态分布 $N_p(\mathbf{0}, \Sigma)$，其中 $\mathbf{0}$ 是 p 维零均值向量，Σ 是一个 $p \times p$ 的正定协方差矩阵 (称为参数矩阵或尺度矩阵)。

则随机矩阵 W 定义为这些向量外积的和：

$$W = \sum_{a=1}^{k} Z_a Z_a'$$

我们称随机矩阵 W 服从自由度为 k、参数矩阵为 Σ 的 **Wishart 分布**，记作：

$$W \sim W_p(k, \Sigma)$$

- W 是一个 $p \times p$ 的对称半正定矩阵。
- p 是随机向量的维度 (也是Wishart矩阵的维度)。
- k 是自由度 (degrees of freedom)，表示参与求和的独立向量的个数。

定义 2（称为来源可能更合适）：作为样本离差阵的分布 (更常见于应用)

定义 设 $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ 是 n 个独立的 p 维随机列向量，并且每一个 $X_{(a)}$ 都服从多元正态分布 $N_p(\mu, \Sigma)$。
令 $\bar{X} = \frac{1}{n}\sum_{a=1}^{n} X_{(a)}$ 为样本均值向量。
则样本离差阵 (Sum of Squares and Cross-Products matrix, SSCP) S 定义为:

$$S = \sum_{a=1}^{n}(X_{(a)} - \bar{X})(X_{(a)} - \bar{X})'$$

那么，S 服从自由度为 $n-1$、参数矩阵为 Σ 的 Wishart 分布:

$$S \sim W_p(n-1, \Sigma)$$

这里的自由度是 $n-1$ 而不是 n，因为从数据中估计了均值向量 μ (通过 $\bar{X}$)。

3. Wishart分布的概率密度函数 (PDF)

当自由度 $k \geq p$ 且参数矩阵 Σ 是正定的时候，Wishart分布 $W_p(k, \Sigma)$ 的随机矩阵 W 是几乎必然正定的，并且存在概率密度函数。
对于一个 $p \times p$ 的对称正定矩阵 w，其PDF为:

$$f(w; k, \Sigma) = \frac{|w|^{(k-p-1)/2} \exp\left(-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}w)\right)}{2^{kp/2}|\Sigma|^{k/2}\Gamma_p\left(\frac{k}{2}\right)}$$

其中:

- $|w|$ 是矩阵 w 的行列式。
- $\mathrm{tr}(\Sigma^{-1}w)$ 是矩阵 $\Sigma^{-1}w$ 的迹 (主对角线元素之和)。
- $\Gamma_p(\cdot)$ 是多元伽马函数 (multivariate Gamma function)，定义为:
  $\Gamma_p(x) = \pi^{p(p-1)/4}\prod_{j=1}^{p}\Gamma\left(x - \frac{j-1}{2}\right)$
  其中 $\Gamma(\cdot)$ 是一元伽马函数。

如果 w 不是对称正定矩阵，则 $f(w; k, \Sigma) = 0$。

4. 与卡方分布的关系

Wishart 分布是卡方分布的直接推广:

- 当 $p = 1$ (一维情况) 时，随机向量 Z_a 就是一个标量 $Z_a \sim N(0, \sigma^2)$。
  参数矩阵 Σ 就是一个标量 σ^2。
  随机矩阵 W 就是一个标量 $W = \sum_{a=1}^{k} Z_a^2$。
  那么 $\frac{W}{\sigma^2} = \sum_{a=1}^{k}\left(\frac{Z_a}{\sigma}\right)^2$。

由于 $Z_a/\sigma \sim N(0,1)$，所以 $\frac{W}{\sigma^2} \sim \chi_k^2$ (自由度为 k 的卡方分布)。
因此，$W_1(k,\sigma^2)$ 分布等价于 $\sigma^2\chi_k^2$ 分布。

5. Wishart 分布的性质

- **期望：**
  如果 $W \sim W_p(k,\Sigma)$，则 $E(W) = k\Sigma$。
  这意味着样本离差阵 $S \sim W_p(n-1,\Sigma)$ 的期望是 $E(S) = (n-1)\Sigma$。
  因此，样本协方差矩阵 $V = \frac{1}{n-1}S$ 的期望是 $E(V) = \frac{1}{n-1}E(S) = \Sigma$，这表明 V 是 Σ 的无偏估计。
- **可加性：**
  如果 $W_1 \sim W_p(k_1,\Sigma)$ 和 $W_2 \sim W_p(k_2,\Sigma)$ 是独立的 Wishart 随机矩阵（具有相同的参数矩阵 Σ），则它们的和也服从 Wishart 分布：
  $W_1 + W_2 \sim W_p(k_1 + k_2,\Sigma)$
- **线性变换：**
  如果 $W \sim W_p(k,\Sigma)$ 并且 C 是一个 $q \times p$ 的常数矩阵，则：
  $CWC' \sim W_q(k, C\Sigma C')$
  特别地，如果 C 是一个 $p \times p$ 的非奇异矩阵 ($q = p$)，则 $CWC' \sim W_p(k, C\Sigma C')$。
- **正定性：**
  如果 Σ 是正定的，则 $W \sim W_p(k,\Sigma)$ 的随机矩阵 W 是几乎必然正定的当且仅当自由度 $k \geq p$。
  如果 $k < p$，则 W 是奇异的（行列式为0，秩小于 p)。
  对于样本离差阵 $S \sim W_p(n-1,\Sigma)$，它几乎必然正定当且仅当 $n-1 \geq p$，即 $n > p$。
- **对角元素的分布：**
  如果 $W \sim W_p(k,\Sigma)$，则 W 的第 j 个对角元素 W_{jj} 服从 $\sigma_{jj}\chi_k^2$ 分布，其中 σ_{jj} 是 Σ 的第 j 个对角元素 (即第 j 个基础正态变量的方差)。
  这是因为 $W_{jj} = \sum_{a=1}^k Z_{aj}^2$，而 $Z_{aj} \sim N(0,\sigma_{jj})$。所以 $\sum_{a=1}^k (Z_{aj}/\sqrt{\sigma_{jj}})^2 \sim \chi_k^2$。
- **行列式的分布：**
  Wishart 矩阵的行列式 $|W|$ 的分布也是已知的，它与一系列独立的卡方随机变量的乘积有关。

6. 非中心Wishart分布

如果定义 Wishart 矩阵的基础正态向量 Z_a 的均值不为零，即 $Z_a \sim N_p(\mu_a,\Sigma)$，则
$ W = \sum_{a=1}^k Z_a Z_a' $
服从**非中心Wishart分布**，记为 $W_p(k,\Sigma,\mathbf{M})$，其中 $\mathbf{M} = \sum_{a=1}^k \mu_a\mu_a'$ 是非中心参数矩阵。
非中心Wishart分布在假设检验的备择假设下或某些模型的推导中会出现。它的期望是 $E(W) = k\Sigma + \mathbf{M}$。

7. 应用

- **假设检验：** 用于构造关于协方差矩阵的检验统计量，例如检验 $\Sigma = \Sigma_0$ 或检验不同总体的协方差矩阵是否相等 (如 Box's M-test)。
- **置信区间/区域：** 用于构造协方差矩阵元素的置信区间或整个协方差矩阵的置信区域。
- **多元方差分析 (MANOVA)：** Wishart 分布出现在MANOVA的检验统计量中，用于比较多个总体的均值向量。
- **贝叶斯统计：** Wishart 分布常被用作协方差矩阵的共轭先验分布。
- **随机矩阵理论：** Wishart 矩阵是随机矩阵理论中一类重要的研究对象。