

多元统计 A 2025-2026

1. $X = (x_1 : x_2) \in \mathbb{R}^p$. $x_1 \in \mathbb{R}^q$. $x_2 \in \mathbb{R}^{p-q}$. $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. $\mu = (\mu_1 : \mu_2)^T$

(1) 证 $\mu' \Sigma^{-1} \mu \leq \mu \Sigma^{-1} \mu$

(2) 证 $x_2 | x_1$ 分布

2. 正态假设下, 导出 Σ 相同时 $\mu_1 = \mu_2$ 的似然比检验

3. 证明 PCA 第一主成分等价于 $\max_{\|v\|=1} v' S v$
以及确定主成分个数 k 的方法

4. 岭回归 $\beta = (X'X + kI_p)^{-1} X'Y$ 公式推导, 无偏或有偏

5. $\Sigma_1 = \Sigma_2$ 时, Bayes 判别法公式推导, 与 LDA 关系
高维时 LDA 不稳定原因? 怎么解决

6. 计算层次聚类, 画图

分布式 2025-2026 (记不清了, 只写了有印象的)

1. Map-Reduce 设计函数: 给出 n 个人的好友列表, 找共同好友

2. 正态条件分布下的 Gibbs 采样

3. 查看 hdfs 的端口

4. 作业题: 计算 $E \begin{pmatrix} \sin(x^2) & \sin(xy) \\ \cos(xy) & \cos(y^2) \end{pmatrix}$