

第六次作业

1907402030 熊雄

2022 年 1 月 18 日

Question 1.

假设样本 $(x_1, \dots, x_p) \sim N(0, I_p)$, 其中 $p \in \{10, 20\}$. 建立线性回归模型 $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$, 其中回归系数 $\beta_0, \beta_1, \beta_2, \beta_3 \sim Unif(1, 2)$ 为随机数, 误差项满足正态性假设, 即 $\varepsilon \sim N(0, 1)$.

请重复下述过程 n 次, 记录正确, 多选, 少选, 错选的次数, 其中 $n \in \{200, 500, 1000\}$.

1. 生成 $\beta_0, \beta_1, \beta_2, \beta_3$;
2. 生成 ε 和 X 数据;
3. 生成 Y 数据;
4. 分别利用 AIC 准则, BIC 准则, R_{adj} , C_p 统计量寻找最优子集.

Answer.

```
1  n <- 200
2  p <- 10
3  correct = 0
4  add = 0
5  less = 0
6  error = 0
7  for (i in 1:500)
8  {
```