Stat873 HW04

1. Consider Model $Y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2\Sigma)$.

   (1) Among all maximum likelihood estimators for β, point out the one with minimum norm.

   In $\text{MLE}(\beta) = \text{GLSE}_{\Sigma^{-1}}(\beta) = \left(\Sigma^{-1/2}X\right)^+ \Sigma^{-1/2}Y + \mathcal{N}(X)$
   $\widehat{\beta} = \left(\Sigma^{-1/2}X\right)^+ \Sigma^{-1/2}Y$ has minimum norm.

   (2) Which norm was used in (1)? Why not $\|\cdot\|_{\Sigma^{-1}}$?

   The norm used in (1) is Frobenius norm $\|u\| = \sqrt{u'u}$.
   Norm $\|v\|_{\Sigma^{-1}} = \sqrt{v'\Sigma^{-1}v}$ is for vectors in R^n. But β and its estimators are in R^p.
   So $\|\cdot\|_{\Sigma^{-1}}$ can not be used.

   (3) Suppose X has full column rank. Find the distribution for the estimator in (1).

   $\widehat{\beta} = AY$ where $A = \left(\Sigma^{-1/2}X\right)^+ \Sigma^{-1/2}$. So $E(\widehat{\beta}) = AX\beta = \left(\Sigma^{-1/2}X\right)^+ \left(\Sigma^{-1/2}X\right)\beta = \beta$ since $\left(\Sigma^{-1/2}X\right)^+$ is a right-inverse of $\Sigma^{-1/2}X$.

   $$\begin{aligned}\text{Cov}(\widehat{\beta}) &= A\sigma^2\Sigma A' = \sigma^2 \left(\Sigma^{-1/2}X\right)^+ \left[\left(\Sigma^{-1/2}X\right)'\right]^+ = \sigma^2 \left(X'\Sigma^{-1/2}\Sigma^{-1/2}X\right)^+ \\ &= \sigma^2 \left(X'\Sigma^{-1}X\right)^+ = \sigma^2 \left(X'\Sigma^{-1}X\right)^{-1}.\end{aligned}$$

   Hence $\widehat{\beta} \sim N\left(\beta, \sigma^2(X'\Sigma^{-1}X)^{-1}\right)$.

2. In Model $Y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2 I_n)$, X has full column rank, and $X'X = P\Lambda P'$ is the EVD.

   (1) Let $\widehat{\beta}$ be the MVUE for β. Write out the expression for $\widehat{\beta}$ and its distribution.

   $\widehat{\beta} = X^+Y = (X'X)^{-1}X'Y \sim N(\beta, \sigma^2(X'X)^{-1})$.

   (2) Let $\widehat{\beta}(K) = [P(\Lambda + K)P']^{-1}X'Y$ be the ridge estimator for β. Express matrix A via P, Λ and K such that $\widehat{\beta}(K) = A\widehat{\beta}$.

   $$\begin{aligned}\widehat{\beta}(K) &= [P(\Lambda + K)P']^{-1}X'Y = P(\Lambda + K)^{-1}P'(X'X)(X'X)^{-1}X'Y \\ &= P(\Lambda + K)^{-1}P'P\Lambda P'\widehat{\beta} = P(\Lambda + K)^{-1}\Lambda P'\widehat{\beta}.\end{aligned}$$

   So $\widehat{\beta}(K) = A\widehat{\beta}$ where $A = P(\Lambda + K)^{-1}\Lambda P'$.

   (3) Find the expression for $\text{Cov}(\widehat{\beta}(K))$ via σ^2, P, Λ and K only.
   Hint: $\text{Cov}(\widehat{\beta}(K)) = A[\text{Cov}(\widehat{\beta})]A'$.

   $$\begin{aligned}\text{Cov}[\widehat{\beta}(K)] &= A[\text{Cov}(\widehat{\beta})]A' = P(\Lambda + K)^{-1}\Lambda P'\sigma^2(X'X)^{-1}P\Lambda(\Lambda + K)^{-1}P' \\ &= \sigma^2 P(\Lambda + K)^{-1}\Lambda P'(P\Lambda^{-1}P')\Lambda(\Lambda + K)^{-1}P' \\ &= P(\Lambda + K)^{-1}\Lambda(\Lambda + K)^{-1}P'.\end{aligned}$$

(4) Based on (3) find $\mathrm{tr}[\mathrm{Cov}(\widehat{\beta}(K))]$ via σ^2, Λ and K only.

$$\begin{aligned}\mathrm{tr}[\mathrm{Cov}(\widehat{\beta}(K))] &= \mathrm{tr}[\sigma^2 P(\Lambda + K)^{-1}\Lambda(\Lambda + K)^{-1}P'] = \sigma^2\mathrm{tr}[(\Lambda + K)^{-1}\Lambda(\Lambda + K)^{-1}] \\ &= \sigma^2 \sum_i \frac{\lambda_i}{(\lambda_i + k_i)^2}.\end{aligned}$$