STA553 HW04 KEY

Ex-3a-7 $\sum \hat{Y}_i(Y_i - \hat{Y}_i) = \hat{Y}'(Y - \hat{Y}) = Y'P(I_n - P)Y = Y'(P - P^2)Y = Y'0_{n\times n}Y = 0$.

Ex-3d-1 Using terms of Theorem 3.5, $X = 1_n, \beta = \theta, p = 1$. Because $\hat{\theta} = \hat{\beta} = (X'X)^{-1}X'Y = \frac{1}{n}X'Y = \bar{Y}$, RSS $= (Y - X\hat{\beta})'(Y - X\hat{\beta}) = (Y - 1_{n\times 1}\bar{Y})'(Y - 1_{n\times 1}\bar{Y}) = \sum_i (Y_i - \bar{Y})^2 = Q$. Due to (iii) and (iv) of Theorem 3.5, the conclusions are correct.

Ex-3e-3 (a) This is because

$$\begin{pmatrix} I_p & 0 \\ -x^{(p)\prime}W(W'W)^{-1} & 1 \end{pmatrix}\begin{pmatrix} W'W & W'x^{(p)} \\ x^{(p)\prime}W & x^{(p)\prime}x^{(p)} \end{pmatrix} = \begin{pmatrix} W'W & W'x^{(p)} \\ 0 & x^{(p)\prime}x^{(p)} - x^{(p)\prime}W(W'W)^{-1}W'x^{(p)} \end{pmatrix}$$

(b) $x^{(p)\prime}W(W'W)^{-1}W'x^{(p)} \geq 0$ implies $\frac{\det(W'W)}{\det(X'X)} \geq \frac{1}{x^{(p)\prime}x^{(p)}}$. Since $\hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1})$, then $\mathrm{var}(\hat{\beta}_p) = \sigma^2[(X'X)^{-1}]_{p+1,p+1} = \sigma^2\frac{\det(W'W)}{\det(X'X)} \geq \frac{\sigma^2}{x^{(p)\prime}x^{(p)}}$. with equality if and only if $x^{(p)\prime}W(W'W)^{-1}W'x^{(p)} = 0$ which is equivalent to $P_W x^{(p)} = 0$, i.e., $x^{(p)}$ is orthogonal to the space spanned by $\{x^{(i)} : i = 0, 1, \ldots, p-1\}$. Or because $X'X$ is positive definte, then $(X'X)^{-1}$ is positive definite, then $x^{(p)\prime}W(W'W)^{-1}W'x^{(p)} = 0$ iff $W'x^{(p)} = 0$.

4. (a) $\hat{\theta}_i^{\text{ridge}}(k) = \frac{1}{1+k}Y_i$ where $k > 0$.

   (b) $MSE(\hat{\theta}^{\text{LS}}) = 1$. $MSE(\hat{\theta}^{\text{ridge}}) = \frac{1}{(1+k)^2} + \frac{k^2}{(1+k)^2}\frac{1}{n}\sum_{i=1}^n \theta_i^2$.

   (c) $MSE(\hat{\theta}^{\text{LS}}) = 1$. $MSE(\hat{\theta}^{\text{ridge}}) = \frac{1}{(1+k)^2} + \frac{k^2}{(1+k)^2}\epsilon\theta_0^2$.

   (d) As $n \to \infty$, $\epsilon\theta_0^2 = n^{-\beta}2r\log(n) \to 0$. Then $MSE(\hat{\theta}^{\text{ridge}}) \to \frac{1}{(1+k)^2} < 1 = MSE(\hat{\theta}^{\text{LS}})$. Ridge regression is better in this situation.

5. (a)

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

   (b) $\mathrm{rank}(X) = 7$.

(c) $c'\theta$, $\forall c \in C(X')$. Because $C(X') \perp \mathcal{N}(X), C(X') + \mathcal{N}(X) = R^8$ and $\mathcal{N}(X) = \text{span}\{(1,1,1,1,-1,-1,-1,-1)'\}$, so the simpler form is

$$\{c'\theta : \ (1,1,1,1,-1,-1,-1,-1)'c = 0\}.$$

6. (a) $a'\beta$ is estimable iff $a \in C(X')$, that is, there exists $c_1, c_2 \in R$ such that $a = c_1(1,1,2)' + c_2(1,0,3)'$.

   (b) $c_1 = 1, c_2 = 0$. $l' = (1/2, 0, 1/2, 0)$.

7. (a) Without loss of generality, let $i = 1$ and $e_1 = (1, 0, \ldots, 0)'$. Then β_1 is estimable $\iff$ $e_1 \in C(X')$ $\iff$ $C(e_1, X') = C(X')$ $\iff$ row rank of $\binom{e_1'}{X}$ equals to row rank of X $\iff$ rank of $\binom{e_1'}{X}$ equals to rank of X. Now

   $$\text{rank}\begin{pmatrix} e_1' \\ X \end{pmatrix} = \text{rank}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ x_{(1)}^{\perp} & x_{(2)} & \cdots & x_{(p)} \end{pmatrix} = \text{rank}(x_{(2)}, \ldots, x_{(p)}) + 1$$

   Therefore, β_1 is estimable if and only if $x_{(1)}^{\perp} \neq 0$.

   Proof 2: β_1 is estimable $\iff$ $e_1 \in C(X')$ $\iff$ $\exists a \in R^p,\ e_1' = a'X$ $\iff$ $\exists a \in R^p,\ a'x_{(1)} = 1, a'x_{(2)} = 0, \ldots, a'x_{(p)} = 0$ $\iff$ $\exists a \in R^p,\ a'x_{(1)}^{\perp} \neq 0, a \perp \text{span}\{x_{(2)}, \ldots, x_{(p)}\}$ $\iff$ $x_{(1)}^{\perp} \neq 0$.

   (b) Use result of question 3.

   Proof 2: Let $a = x_{(1)}^{\perp}/\|x_{(1)}^{\perp}\|^2$, then $e_1' = a'X, \beta_1 = a'X\beta, \hat{\beta}_1 = a'PY$. Hence, $\text{Var}(\hat{\beta}_1) = \sigma^2 a'PP'a = \sigma^2\|Pa\|^2 = \sigma^2\|a\|^2 = \sigma^2/\|x_{(1)}^{\perp}\|^2$. The equality $Pa = a$ is due to the fact $a \in \text{span}\{x_{(1)}, x_{(2)}, \ldots, x_{(p)}\}$.

8. The BLUE of $w_2 + w_3$ is $\hat{Y}_2 = l'Y$, where $l' = (1/5, 2/5, -2/5, 1/5)$.

9. Prove by induction in m. Let v be a vector that is not in $A_1, \cdots, A_{m-1}$. If $v \notin A_m$, we are done. Now suppose $v \in A_m$. Since A_m does not fill the whole space, we can find $w \notin A_m$. Now consider the line $v + tw$, where t is a real number. For $i = 1, \cdots, m-1$, since $v \notin A_i$, the line can only intersect A_i once (this can be proved by setting up an equation system that the solution is the intersection). The line intersects A_m only at $t = 0$. So we can choose some t such that $v + tw$ is not in any of the subspaces.