STA 610 Homework 4

Yuren Zhou

October 4, 2024

Question 1. (2 points) For any $c \in \mathbb{R}^p$, since $\mathbb{V}[\check{\beta}] - \mathbb{V}[\widehat{\beta}]$ is positive definite, we have

$$\mathbb{V}[c^\top \widehat{\beta}] - \mathbb{V}[c^\top \check{\beta}] = -c^\top(\mathbb{V}[\check{\beta}] - \mathbb{V}[\widehat{\beta}])c < 0.$$

Question 2.

a. (3 points)

- Approach 1: The conditional distribution of $a|\cdot$ follows

$$\begin{aligned}
p(a|\cdot) &\propto p(y|\beta, a, \sigma^2)p(a|\Psi) \\
&\propto \exp\left(-\frac{\|y - X\beta - Za\|}{2\sigma^2} - \frac{1}{2}a^\top \Psi^{-1} a\right) \\
&\propto \exp\left(-\frac{1}{2}a^\top(\sigma^{-2}Z^\top Z + \Psi^{-1})a + \sigma^{-2}(y - X\beta)^\top Za\right) \\
&\sim N\left(\sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top(y - X\beta),\ (\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}\right).
\end{aligned}$$

- Approach 2: The joint distribution of a and $y - X\beta$ follows a multivariate normal distribution with parameters

$$\mathbb{E}[a] = 0, \quad \mathbb{E}[y - X\beta] = Z\mathbb{E}[a] + \mathbb{E}[\epsilon] = 0,$$

  and

$$\mathbb{V}[a] = \Psi, \quad \mathbb{V}[y - X\beta] = Z\Psi Z^\top + \sigma^2 I, \quad \mathrm{Cov}(a, y - X\beta) = \Psi Z^\top.$$

  Therefore, the conditional distribution of $a|y$ is als a normal distribution with parameters

$$\begin{aligned}
\mathbb{E}[a|y] &= \mathbb{E}[a] + \mathrm{Cov}(a, y - X\beta)\mathbb{V}[y - X\beta]^{-1}(y - X\beta - \mathbb{E}[y - X\beta]) \\
&= \Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-1}(y - X\beta),
\end{aligned}$$

  and

$$\begin{aligned}
\mathbb{V}[a|y] &= \mathbb{V}[a] - \mathrm{Cov}(a, y - X\beta)\mathbb{V}[y - X\beta]^{-1}\mathrm{Cov}(y - X\beta, a) \\
&= \Psi - \Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-1}Z\Psi.
\end{aligned}$$

We can easily verify that the normal distributions obtained from both approaches are the same, i.e.

$$\sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top = \Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-1}$$

and

$$(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1} = \Psi - \Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-1}Z\Psi.$$

b. (3 points) Taking expectations, we have $\mathbb{E}[y|a] = X\beta + Za$ and hence

$$\mathbb{E}[\widehat{a}|a] = \sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top Za,$$

$$\mathbb{E}[\check{a}|a] = (Z^\top Z)^{-1}Z^\top Za = a.$$

Taking variances and covariances, we have

$$\begin{aligned}\mathbb{V}[\widehat{a}|a] &= \mathbb{E}\left[\left(\sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top \epsilon\right)\left(\sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top \epsilon\right)^\top\right] \\ &= \sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top Z(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1},\end{aligned}$$

$$\mathbb{V}[\check{a}|a] = \mathbb{E}\left[\left((Z^\top Z)^{-1}Z^\top \epsilon\right)\left((Z^\top Z)^{-1}Z^\top \epsilon\right)^\top\right] = \sigma^2(Z^\top Z)^{-1},$$

$$\mathrm{Cov}(\widehat{a}, \check{a}|a) = \mathbb{E}\left[\left(\sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top \epsilon\right)\left((Z^\top Z)^{-1}Z^\top \epsilon\right)^\top\right] = (\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}.$$

If you are using approach 2 for part a, you might also obtain the equivalent forms

$$\mathbb{E}[\widehat{a}|a] = \Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-1}Za,$$

$$\mathbb{V}[\widehat{a}|a] = \sigma^2\Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-2}Z\Psi,$$

and

$$\mathrm{Cov}(\widehat{a}, \check{a}|a) = \sigma^2\Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-1}Z(Z^\top Z)^{-1}.$$

c. (3 points) Since $\mathbb{E}[a] = 0, \mathrm{Cov}(a) = \Phi$, we have $\mathbb{E}[y - X\beta] = 0$ and $\mathrm{Cov}(y - X\beta) = \mathbb{E}[(Za + \epsilon)(Za + \epsilon)^\top] = Z\Psi Z^\top + \sigma^2 I$. Therefore,

$$\mathbb{E}[\widehat{a}] = 0, \quad \mathbb{E}[\check{a}] = 0,$$

and

$$\begin{aligned}\mathbb{V}[\widehat{a}] &= \sigma^{-4}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top(Z\Psi Z^\top + \sigma^2 I)Z(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}, \\ \mathbb{V}[\check{a}] &= (Z^\top Z)^{-1}Z^\top(Z\Psi Z^\top + \sigma^2 I)Z(Z^\top Z)^{-1} = \Psi + \sigma^2(Z^\top Z)^{-1}, \\ \mathrm{Cov}(\widehat{a}, \check{a}) &= \sigma^{-2}(\sigma^{-2}Z^\top Z + \Psi^{-1})^{-1}Z^\top(Z\Psi Z^\top + \sigma^2 I)Z(Z^\top Z)^{-1}.\end{aligned}$$

If you are using approach 2 for part a, you might also obtain the equivalent forms

$$\mathbb{V}[\widehat{a}] = \Psi Z^\top(Z\Psi Z^\top + \sigma^2 I)^{-1}Z\Psi,$$

and

$$\mathrm{Cov}(\widehat{a}, \check{a}) = \Psi.$$