Nathália Dantas Handam Nunes
STA13824: Prof. Dr. Fabio Alexander Fajardo Molinares
May 12, 2025

Desafio 1

Question 1. Considere $Y = \mathbf{X}\beta + \epsilon$, com Y um vetor aleatório com distribuição normal multivariada com vetor de médias $\mathbf{X}\beta$ e matriz de variâncias e covariâncias $\sigma^2\mathbf{I}$. Assumindo as Suposições A1-A5:

(a) Calcule Estimador de Máxima Verossimilhança de σ^2;

1. Função de Verossimilhança
   O modelo é:
   $$Y = \mathbf{X}\beta + \epsilon,$$
   onde:
   - Y é um vetor $n \times 1$ de variáveis dependentes (observações).
   - $\mathbf{X}$ é uma matriz $n \times p$ de variáveis explicativas (fixas ou condicionalmente fixas).
   - β é um vetor $p \times 1$ de parâmetros desconhecidos.
   - ϵ é um vetor $n \times 1$ de erros aleatórios.

   A densidade de uma normal multivariada $N(\mu, \Sigma)$ é:
   $$f(Y) = (2\pi)^{-n/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(Y-\mu)'\Sigma^{-1}(Y-\mu)\right).$$
   No caso:
   - $\mu = \mathbf{X}\beta$,
   - $\Sigma = \sigma^2\mathbf{I}_n$, então $|\Sigma| = (\sigma^2)^n$ e $\Sigma^{-1} = \frac{1}{\sigma^2}\mathbf{I}_n$.

   Substituindo:
   $$f(Y|\mathbf{X}, \beta, \sigma^2) = (2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}(Y-\mathbf{X}\beta)'(Y-\mathbf{X}\beta)\right).$$
2. Log-Verossimilhança
   A log-verossimilhança é definida como:
   $$\ell(\beta, \sigma^2) = \log f(Y|\mathbf{X}, \beta, \sigma^2).$$
   Substituindo a expressão da verossimilhança:
   $$\ell(\beta, \sigma^2) = \log\left[(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}(Y-\mathbf{X}\beta)'(Y-\mathbf{X}\beta)\right)\right].$$
   $$\ell(\beta, \sigma^2) = \log\left[(2\pi\sigma^2)^{-n/2}\right] + \log\left[\exp\left(-\frac{1}{2\sigma^2}(Y-\mathbf{X}\beta)'(Y-\mathbf{X}\beta)\right)\right].$$
   $$\ell(\beta, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(Y-\mathbf{X}\beta)'(Y-\mathbf{X}\beta).$$
   $$\ell(\beta, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(Y-\mathbf{X}\beta)'(Y-\mathbf{X}\beta).$$
3. Maximização em Relação a β
   Para maximizar em relação a β, ignoramos os termos que não dependem de β (constantes), pois não afetam a solução. Assim, o problema se reduz a:
   $$\arg\max_{\beta}\left[-\frac{1}{2\sigma^2}(Y-\mathbf{X}\beta)'(Y-\mathbf{X}\beta)\right].$$
   Como $\sigma^2 > 0$, maximizar essa expressão equivale a:
   $$\arg\min_{\beta}(Y-\mathbf{X}\beta)'(Y-\mathbf{X}\beta),$$
   que é exatamente o MQO.

Queremos encontrar β que minimiza:

$$\begin{aligned} S(\beta) = \epsilon'\epsilon &= (Y - \mathbf{X}\beta)'(Y - \mathbf{X}\beta) \\ &= Y'Y - Y'\mathbf{X}\beta - \beta'\mathbf{X}'Y + \beta'\mathbf{X}'\mathbf{X}\beta \\ &= Y'Y - 2\beta'\mathbf{X}'Y + \beta'\mathbf{X}'\mathbf{X}\beta. \end{aligned}$$

Para encontrar o mínimo, derivamos RSS em relação a β e igualamos a zero:

$$\frac{\partial S(\beta)}{\partial \beta} = -2\mathbf{X}'Y + 2\mathbf{X}'\mathbf{X}\beta = 0.$$

Resolvendo para β:

$$\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'Y.$$

Assumindo que $\mathbf{X}'\mathbf{X}$ tem inversa (posto completo, suposição A2), multiplicamos ambos os lados por $(\mathbf{X}'\mathbf{X})^{-1}$:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y.$$

Para garantir que essa solução é um mínimo (e não um máximo ou ponto de sela), calcula-se a segunda derivada da função objetivo $S(\beta)$ e verifica-se se a matriz resultante é definida positiva. Assim,

$$\frac{\partial^2 S(\beta)}{\partial \beta \partial \beta'} = 2\mathbf{X}'\mathbf{X}.$$

Como $\mathbf{X}'\mathbf{X}$ é positiva definida (sob a suposição de posto completo), a solução é de fato um mínimo global.

4. Maximização em Relação a σ^2
   Maximizar a log-verossimilhança condicional (já com β substituído por seu EMV $\hat{\beta}$) em relação a σ^2. A log-verossimilhança concentrada (após substituir $\hat{\beta}$) é:

$$\ell(\sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}S(\beta),$$

onde $S(\beta) = \epsilon'\epsilon = (Y - \mathbf{X}\hat{\beta})'(Y - \mathbf{X}\hat{\beta})$ é a soma dos quadrados dos resíduos.
Para maximizar $\ell(\sigma^2)$, derivamos em relação a σ^2 e igualamos a zero:

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{S(\beta)}{2\sigma^4} = 0.$$

$$-\frac{n}{2\sigma^2} + \frac{S(\beta)}{2\sigma^4} = 0 \implies \frac{S(\beta)}{2\sigma^4} = \frac{n}{2\sigma^2}.$$

$$\frac{S(\beta)}{\sigma^2} = n \implies \sigma^2 = \frac{S(\beta)}{n}.$$

Para garantir que é um máximo, calculamos a segunda derivada:

$$\frac{\partial^2 \ell}{(\partial \sigma^2)^2} = \frac{n}{2\sigma^4} - \frac{S(\beta)}{\sigma^6}.$$

Substituindo $\sigma^2 = \frac{S(\beta)}{n}$:

$$\left.\frac{\partial^2 \ell}{(\partial \sigma^2)^2}\right|_{\sigma^2 = \frac{S(\beta)}{n}} = \frac{n^3}{2S(\beta)^2} - \frac{n^3}{S(\beta)^2} = -\frac{n^3}{2S(\beta)^2} < 0.$$

Como $n > 0$ e $S(\beta) > 0$ (a menos que o ajuste seja perfeito, o que é raro), o termo $-\frac{n^3}{2S(\beta)^2}$ é negativo.
Como a segunda derivada é negativa, a solução é de fato um máximo.

(b) Calcule o viés e o Erro Quadrático Médio do estimador obtido no item anterior

Para calcular o **viés** do estimador de máxima verossimilhança (EMV) de σ^2, precisamos comparar o valor esperado do EMV com o verdadeiro valor de σ^2.

Conforme derivado, o EMV de σ^2 é:

$$\hat{\sigma}^2 = \frac{S(\hat{\beta})}{n} = \frac{(Y - \mathbf{X}\hat{\beta})'(Y - \mathbf{X}\hat{\beta})}{n} = \hat{\epsilon}'\hat{\epsilon}.$$

onde:
- $\hat{\epsilon} = Y - \mathbf{X}\hat{\beta}$
- $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y$

Substituindo $\hat{\beta}$, temos:

$$\hat{\epsilon} = Y - \mathbf{X}\hat{\beta} = Y - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y.$$

Inicialmente, definindo as matrizes:

$$H = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'.$$

$$M = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{I} - H$$

As matrizes H e M são simétricas e idempotentes, isto é,

$$H' = H$$
$$M' = M$$
$$HH = H$$
$$MM = M$$

Verifica-se também que:

$$H\mathbf{X} = \mathbf{X} \quad \text{ou} \quad \mathbf{X}'H = \mathbf{X}'$$

e

$$M\mathbf{X} = \mathbf{0} \quad \textbf{ou} \quad \mathbf{X}'M = \mathbf{0}$$

onde $\mathbf{0}$ é uma matriz de zeros.

Temos que os resíduas podem ser escritos como:

$$\hat{\epsilon} = Y - X\hat{\beta} = Y - HY = (\mathbf{I} - H)Y = MY = M(\mathbf{X}\beta + \epsilon)$$

Considerando $MX = 0$:

$$\hat{\epsilon} = M\epsilon$$

Então,

$$S(\hat{\beta}) = \hat{\epsilon}'\hat{\epsilon} = \epsilon' M\epsilon$$

Como $\epsilon'\epsilon$ é uma matriz com apenas um elemento temos que

$$\epsilon'\epsilon = \text{tr}(\epsilon' M\epsilon)$$

O teorema da álgebra de matrizes estabelece que o traço de um produto de matrizes não é afetado por uma mudança na ordem dos fatores, desde que o novo produto também seja definido, assim:

$$\epsilon'\epsilon = \text{tr}(\epsilon\epsilon' M)$$

De acordo com **A4** $E[\epsilon\epsilon|X] = \sigma^2\mathbf{I}$, assim:

$$\mathbb{E}(S(\beta)|\mathbf{X}) = \mathbb{E}(\epsilon'\epsilon|\mathbf{X}) = \mathbb{E}[\text{tr}(\epsilon\epsilon' M)|\mathbf{X}] = \text{tr}[M \cdot \mathbb{E}(\epsilon\epsilon')] = \text{tr}(M\sigma^2\mathbf{I}) = \sigma^2\text{tr}(M)$$

Como $M = \mathbf{I} - H$:

$$\mathrm{tr}(M) = \mathrm{tr}(\mathbf{I}) - \mathrm{tr}(H) = n - p$$

pois H tem posto p (assumindo $\mathbf{X}$ tem posto completo) e $\mathrm{tr}(H) = p$.

Portanto:

$$E[S(\hat{\beta})] = \sigma^2(n - p).$$

O EMV de σ^2 é:

$$\hat{\sigma}^2 = \frac{S(\hat{\beta})}{n},$$

portanto:

$$E[\hat{\sigma}^2] = \frac{E[S(\hat{\beta})]}{n} = \frac{\sigma^2(n-p)}{n} = \sigma^2\left(1 - \frac{p}{n}\right).$$

Assim, o viés é definido como:

$$\text{Viés}(\hat{\sigma}^2) = E[\hat{\sigma}^2] - \sigma^2 = \sigma^2\left(1 - \frac{p}{n}\right) - \sigma^2 = -\frac{p}{n}\sigma^2.$$

Para calcular o **Erro Quadrático Médio (EQM)** do estimador de máxima verossimilhança (EMV) de σ^2, $\hat{\sigma}^2 = \frac{S(\hat{\beta})}{n}$, usamos a definição geral do EQM para um estimador $\hat{\theta}$ de um parâmetro θ:

$$\text{EQM}(\hat{\theta}) = \text{Var}(\hat{\theta}) + \left[\text{Viés}(\hat{\theta})\right]^2.$$

Já sabemos que o viés de $\hat{\sigma}^2$ é:

$$\text{Viés}(\hat{\sigma}^2) = -\frac{p}{n}\sigma^2.$$

Agora, precisamos calcular a variância de $\hat{\sigma}^2$.

Sabemos que:

$$\hat{\sigma}^2 = \frac{S(\hat{\beta})}{n} = \frac{\epsilon' M \epsilon}{n},$$

onde M é idempotente e simétrica, e $\epsilon \sim N(0, \sigma^2 I)$.

Sob normalidade dos erros ($\epsilon \sim N(0, \sigma^2 I)$), a forma quadrática $\epsilon' M \epsilon$ segue uma distribuição qui-quadrado escalada:

$$\frac{\epsilon' M \epsilon}{\sigma^2} \sim \chi^2_{n-p},$$

pois $\mathrm{rank}(M) = n - p$.

Assim:

$$S(\hat{\beta}) = \epsilon' M \epsilon \sim \sigma^2 \chi^2_{n-p}.$$

A variância de uma variável aleatória $Q \sim \sigma^2 \chi^2_k$ é:

$$\text{Var}(Q) = 2k\sigma^4.$$

Portanto:

$$\text{Var}(S(\hat{\beta})) = 2(n-p)\sigma^4.$$

Como $\hat{\sigma}^2 = \frac{S(\hat{\beta})}{n}$, temos:

$$\text{Var}(\hat{\sigma}^2) = \frac{2(n-p)\sigma^4}{n^2}.$$

Substituindo o viés e a variância na fórmula do EQM:

$$\text{EQM}(\hat{\sigma}^2) = \text{Var}(\hat{\sigma}^2) + \left[\text{Viés}(\hat{\sigma}^2)\right]^2,$$

$$\text{EQM}(\hat{\sigma}^2) = \frac{2(n-p)\sigma^4}{n^2} + \left(-\frac{p}{n}\sigma^2\right)^2,$$

$$\text{EQM}(\hat{\sigma}^2) = \frac{2(n-p)\sigma^4}{n^2} + \frac{p^2\sigma^4}{n^2},$$

$$\mathrm{EQM}(\hat{\sigma}^2) = \frac{2(n-p)+p^2}{n^2}\sigma^4.$$

Simplificando:

$$\mathrm{EQM}(\hat{\sigma}^2) = \frac{2n-2p+p^2}{n^2}\sigma^4.$$

Universidade Federal do Espirito Santo