

2. 6 points

Let Z be a multivariate random variable with expectation $\mu \in \mathbb{R}^n$ and covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$. Let $A \in \mathbb{R}^{n \times n}$ be a matrix.

(a) Show that $\mathbb{E}[Z^T A Z] = \mu^T A \mu + \text{tr}(A \Sigma)$.

(b) Show that $\text{Cov}(AZ) = A^T \Sigma A$.

(c) Let $\hat{\beta} = (X^T X)^{-1} X^T y$ be the LS-estimator from the linear regression lecture, i.e. for $y = X\beta + \varepsilon$ and $\text{Cov}(y) = \sigma^2 I_{n \times n}$. Show that $\text{Cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$.

Group Members:
Dhvaniben Jasoliya
Leutrim Uka
Nicola Horst
Tauqeer Rumaney
Yuvraj Dhepe

I. 3/6
II. 3/6
III. 6/6

12/18

Since Z is multivariate, $Z = (z_1, z_2, \dots, z_n)$ & $\mathbb{E}[Z] = \mu = (\mathbb{E}[z_1], \dots, \mathbb{E}[z_n])$
and $\text{Var}(Z) = \Sigma = \mathbb{E}[(Z - \mu)(Z - \mu)^T] = \mathbb{E}[Z Z^T] - \mu \mu^T$

- Wrong path for X.txt and Y.txt in your python code
- your $\hat{\sigma}^2_{adj}$ value is wrong by 0.05

with $\Sigma = \begin{pmatrix} \text{Var}(z_1) & \dots & \text{Cov}(z_1, z_n) \\ \vdots & \ddots & \vdots \\ \text{Cov}(z_n, z_1) & \dots & \text{Var}(z_n, z_n) \end{pmatrix}$

a) $\mathbb{E}[Z^T A Z] = \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j\right] = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \mathbb{E}[x_i x_j] = \sum_{i=1}^n \sum_{j=1}^n a_{ij} (\mu_i \mu_j + \text{Cov}(z_i, z_j))$

$\mathbb{E}[x_i x_j] \Rightarrow \text{Cov}(x_i, x_j) = \mathbb{E}[x_i x_j] - \mathbb{E}[x_i] \mathbb{E}[x_j]$
 $\Rightarrow \mathbb{E}[x_i x_j] = \mathbb{E}[x_i] \mathbb{E}[x_j] + \text{Cov}(x_i, x_j)$



2/2

$= \sum_{i=1}^n \sum_{j=1}^n a_{ij} \mu_i \mu_j + \sum_{i=1}^n \sum_{j=1}^n a_{ij} \text{Cov}(z_i, z_j) = \mu^T A \mu + \sum_{i=1}^n (A \Sigma)_{ii} = \mu^T A \mu + \text{tr}(A \Sigma)$



0/2

b) $\text{Cov}(AZ) = \mathbb{E}[(AZ)(AZ)^T] = \mathbb{E}[AZ Z^T A^T] = A \mathbb{E}[Z Z^T] A^T = A \Sigma A^T$

Falsch!



c) $\hat{\beta} = (X^T X)^{-1} X^T y$ From the definition, we have $y = X\beta + \varepsilon$

$= (X^T X)^{-1} X^T (X\beta + \varepsilon)$

$= (X^T X)^{-1} X^T X \beta + (X^T X)^{-1} X^T \varepsilon$

$(X^T X)^{-1} X^T X$ is the identity matrix

$= \beta + (X^T X)^{-1} X^T \varepsilon$

What !!!!

* $= \text{Cov}(\beta + (X^T X)^{-1} X^T \varepsilon)$

$= \text{Cov}((X^T X)^{-1} X^T \varepsilon)$

$= (X^T X)^{-1} X^T \text{Cov}(\varepsilon) X (X^T X)^{-1}$

$= (X^T X)^{-1} X^T \sigma^2 I X (X^T X)^{-1}$

$= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1}$

$= \sigma^2 (X^T X)^{-1}$

1/2

* $\text{Cov}(\hat{\beta}) = \text{Cov}((X^T X)^{-1} X^T y)$
 $\hat{\beta}$

Q3)

As per the given Q_3 : Likelihood is

$$L(p) = [(1-p)]^{2x} [2p(1-p)]^y [p^2]^z$$

$$\log(L(p)) = 2x \log(1-p) + y \log 2p + y \log(1-p) + 2z \log p$$

$$\frac{\partial \log(L(p))}{\partial p} = 0 \quad \text{to find MLE, we get}$$

$$-\frac{2x}{1-p} + \frac{y}{p} - \frac{y}{1-p} + \frac{2z}{p} = 0$$

Multiplying with $p(1-p)$ on both sides

$$y(1-p) + 2z(1-p) = 2px + yp$$

$$y - py + 2z - 2zp = 2px + yp$$

$$y + 2z = 2px + 2py + 2zp$$

$$p = \frac{y + 2z}{2x + 2y + 2z}$$

$$\because \frac{\partial^2 \log(L(p))}{\partial p^2} = -\frac{2x}{(1-p)^2} - \frac{y}{p^2} - \frac{y}{(1-p)^2} - \frac{2z}{p^2} < 0$$

as x , y & z are counts of people
& denoms are square \therefore the critical p
value of p denotes the maximum.

$$\therefore \text{MLE is } \frac{y+z}{2x+2y+2z}.$$

