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 $\mathbb{E}[Z^T AZ] = \mu^T A\mu + \mathrm{tr}(A\Sigma)$.

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

(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 $\mathrm{Cov}(y) = \sigma^2 I_{n\times n}$. Show that $\mathrm{Cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$.


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


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 $\mathrm{Var}(Z) = \underbrace{\Sigma}_{n\times n} = \mathbb{E}[\underbrace{(Z-\mu)}_{n\times 1}\underbrace{(Z-\mu)^T}_{1\times n}] = \mathbb{E}[Z \cdot Z^T] - \mu\mu^T$

with $$\Sigma = \begin{pmatrix} \mathrm{Var}(Z_1) & \cdots & \mathrm{Cov}(Z_1, Z_n) \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}(Z_n, Z_1) & \cdots & \mathrm{Var}(Z_n, Z_n) \end{pmatrix}$$

a) $$\mathbb{E}[Z^T A Z] = \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n a_{ij} \cdot X_i X_j\right] = \sum_{i=1}^n \sum_{j=1}^n \underset{\text{constant}}{a_{ij}} \cdot \mathbb{E}[\underset{\text{not I.I.D.}}{X_i \cdot X_j}] = \sum_{i=1}^n \sum_{j=1}^n a_{ij}(\mu_i \mu_j + \mathrm{Cov}(Z_i, Z_j))$$

$\mathbb{E}[X_i X_j] \Rightarrow \mathrm{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i] \cdot \mathbb{E}[X_j]$
$\Rightarrow \mathbb{E}(X_i, X_j) = \mathbb{E}(X_i)\mathbb{E}(X_j) + \mathrm{Cov}(X_i, X_j)$

$$= \sum_{i=1}^n \sum_{j=1}^n a_{ij} \mu_i \mu_j + \sum_{i=1}^n \sum_{j=1}^n a_{ij} \cdot \mathrm{Cov}(Z_i, Z_j) = \mu^T A \mu + \sum_{i=1}^n (A\Sigma)_{ii} = \mu^T A \mu + \mathrm{tr}(A\Sigma)$$

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

c) $\hat{\beta} = (X^T X)^{-1} X^T y$ From the definition, we have $y = X\beta + \varepsilon$
$= (X^T X)^{-1} X^T (X\beta + \varepsilon)$
$= \cancel{(X^T X)^{-1} X^T X}\beta + (X^T X)^{-1} X^T \varepsilon$ $(X^T X)^{-1} \cdot X^T X$ is the identity matrix
$= \beta + (X^T X)^{-1} X^T \varepsilon \quad / \cdot \mathrm{Cov}$
$= \mathrm{Cov}(\beta + (X^T X)^{-1} X^T \varepsilon)$
$= \mathrm{Cov}((X^T X)^{-1} \cdot X^T \cdot \varepsilon)$
$= (X^T X)^{-1} \cdot X^T \cdot \mathrm{Cov}(\varepsilon) \cdot X \cdot (X^T X)^{-1}$
$= (X^T X)^{-1} \cdot X^T \cdot \sigma^2 \cdot I \cdot X (X^T X)^{-1}$
$= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1}$
$= \sigma^2 (X^T X)^{-1}$

Q 3)

As per the given Q: Likelihood is

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

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

$$\frac{\partial \log(L(p))}{\partial p} = 0$$ 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

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

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

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

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

Page No

Date

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

as $x$, $y$ & $z$ are counts of people & deno's are square $\therefore$ the critical pt value of $p$ denotes the maximum.

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