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 $\mathrm{Cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$.

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

constant   not I.I.D.

a) $\mathbb{E}[Z^T A Z] = \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}\cdot X_i X_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\cdot \mathbb{E}[X_i\cdot X_j] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}\left(\mu_i \mu_j + \mathrm{Cov}(Z_i, Z_j)\right)$

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

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

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

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

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

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

$= \cancel{(X^T X)^{-1}}\,\cancel{X^T X}\beta + (X^T X)^{-1} X^T \varepsilon$ $(X^T X)^{-1}\cdot X^T X$ is the identity matrix

$= \beta + (X^T X)^{-1} X^T \varepsilon \quad /\cdot \mathrm{Cov}$

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

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

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

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

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

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

3. **6 points**
Every human is a carrier of one of the three genotypes AA, Aa, or aa. The genotypes are occurring with the probabilities $(1-p)^2$, $2p(1-p)$, and p^2 whereas $0 \leq p \leq 1$ and testing of n persons yielded:

- x persons had the genotype AA
- y persons had the genotype Aa
- z persons had the genotype aa

Describe the corresponding statistical model and determine the Maximum Likelihood Estimator for p.

$$L(p) = (1-p)^{2 \cdot x} \cdot (2p(1-p))^y \cdot p^{2z} \quad / \log$$

$$\log L(p) = \log(1-p)^{2x} + \log(2p(1-p))^y + \log(p^{2z}) \quad / \frac{d}{dp}$$

$$\Rightarrow \frac{d}{dp}\left[\log(1-p)^{2x}\right] + \frac{d}{dp}\left[\log(2p(1-p))^y\right] + \frac{d}{dp}\left[\log(p^{2z})\right]$$

$$\Rightarrow 2x \cdot \frac{d}{dp}\left(\log(1-p)\right) + y\left(\frac{d}{dp}\left(\log(2p(1-p))\right)\right) + 2z \cdot \frac{d}{dp}\left(\log(p)\right)$$

$$\Rightarrow 2x \cdot \frac{1}{1-p}(-1) + y \cdot \frac{1}{\cancel{2}p(1-p)} \cancel{2} \cdot \frac{d}{dp}\left(p(1-p)\right) + 2z \cdot \frac{1}{p}$$

$$\Rightarrow -\frac{2}{1-p} x + \frac{1-2p}{p(1-p)} y + \frac{2}{p} z \overset{!}{=} 0 \qquad (*)$$

$$\boxed{p = \frac{y+2z}{2(x+y+z)}} \quad \text{MLE for } p$$

Second derivative of $(*)$ should be negative

$$\frac{d^2}{dp^2} \log L(p) = -\frac{2}{1-p} x + \frac{1-2p}{p(1-p)} y + \frac{2}{p} z \quad / \frac{d}{dp}$$

$$= -2x \frac{d}{dp}\left(\frac{1}{1-p}\right) + y \cdot \frac{d}{dp}\left(\frac{1-2p}{p(1-p)}\right) + 2z \frac{d}{dp}\left(\frac{1}{p}\right)$$

$$= \underbrace{-\underbrace{2x}_{+} \underbrace{\frac{1}{(1-p)^2}}_{+}}_{-} - \underbrace{\underbrace{\frac{\overset{+}{2p^2} - \overset{+}{2p} + 1}{(p-1)^2 p^2}}_{+} \underbrace{y}_{+}}_{+} - \underbrace{\frac{\overset{+}{2z}}{\underbrace{p^2}_{+}}}_{+}$$

The second derivative is negative.