Homework Sheet 4

Group members:
Dhvaniben Jasoliya
Leutrim Uka
Nicola Horst
Taqueer Rumaney
Yuvraj Dhepe

Task 1

1. **8 points**
A square matrix M is called idempotent if $M^2 = M$.

(a) Show that $C = I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\mathbf{T}$ is idempotent, where I_n is the identity matrix and $\mathbf{1}$ is a vector of ones. Also show that $C\mathbf{1} = \mathbf{0}$.

$$C = I_n - \frac{1}{n}\vec{1}\vec{1}^T \quad \text{where} \quad \vec{1} = [1 \;\; 1 \;\; \cdots \;\; 1]_n^T, \qquad I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n \times n}$$

$$C^2 = \left(I_n - \frac{1}{n}\vec{1}\vec{1}^T\right)\cdot\left(I_n - \frac{1}{n}\cdot\vec{1}\vec{1}^T\right) =$$

$$= \underbrace{I_n^2}_{I_n} - I_n\cdot\frac{1}{n}\cdot\vec{1}\vec{1}^T - \frac{1}{n}\vec{1}\vec{1}^T\cdot I_n + \frac{1}{n}\vec{1}\vec{1}^T\cdot\frac{1}{n}\vec{1}\vec{1}^T$$

$$= I_n - \frac{1}{n}\cdot\vec{1}\vec{1}^T - \frac{1}{n}\vec{1}\vec{1}^T + 1$$

$$= I_n - \underbrace{\frac{2}{n}\vec{1}\vec{1}^T}_{1} + 1$$

$$= I_n - 2 + 1$$

$$= I_n - \underbrace{1}_{\frac{1}{n}\vec{1}\vec{1}^T}$$

$$\boxed{= I_n - \frac{1}{n}\vec{1}\vec{1}^T}$$

We proved that $C^2 = C$. Therefore, C is idempotent.

- Next part is to show that $C\cdot\vec{1} = \vec{0}$

$$C = I_n - \frac{1}{n}\vec{1}\vec{1}^T \quad /\cdot\vec{1} \qquad \text{where } \vec{1} = [1 \;\; 1 \;\cdots\; 1]_n^T$$

$$C\cdot\vec{1} = \left(I_n - \frac{1}{n}\vec{1}\vec{1}^T\right)\cdot\vec{1}$$

$$C\cdot\vec{1} = I_n\cdot\vec{1} - \frac{1}{n}\cdot n$$

$$C\cdot\vec{1} = \vec{1} - \underbrace{1}_{\text{scalar}}$$

$$C\cdot\vec{1} = \vec{0}\;; \quad \text{where } \vec{0} = [0 \;\; 0 \;\cdots\; 0]_n^T$$

(b) Let $a \in \mathbb{R}^{n\times 1}$, show that

$$Ca = \begin{bmatrix} a_1 - \bar{a} \\ \vdots \\ a_n - \bar{a} \end{bmatrix}.$$

$$C = I_n - \frac{1}{n}\vec{1}\vec{1}^T \quad \Big|\cdot\vec{a}$$

$$C\cdot\vec{a} = \left(I_n - \frac{1}{n}\cdot\vec{1}\vec{1}^T\right)\cdot\vec{a}$$

$$= I_n\cdot\vec{a} - \frac{1}{n}\cdot\vec{1}\vec{1}^T\vec{a}$$

$$= \vec{a} - \underbrace{\frac{1}{n}\sum_{i=1}^{n} a_i}_{\bar{a}}$$

$$= \underbrace{\vec{a}}_{\text{vector}} - \underbrace{\bar{a}}_{\text{scalar}} \implies C\vec{a} = \begin{pmatrix} a_1 - \bar{a} \\ a_2 - \bar{a} \\ \vdots \\ a_n - \bar{a} \end{pmatrix}$$

Let $A \in \mathbb{R}^{n\times k}$, show that

$$CA = \begin{bmatrix} a_{11}-\bar{a}_1 & \cdots & a_{1k}-\bar{a}_k \\ \vdots & \vdots & \vdots \\ a_{n1}-\bar{a}_1 & \cdots & a_{nk}-\bar{a}_k \end{bmatrix}$$

row vector, where j-th member is the mean of the j-th column of the A matrix

$$C\cdot A = \left(I_n - \frac{1}{n}11^T\right)A = I_n \cdot A_{n\times k} - \frac{1}{n}11^T A = A_{n\times k} - \left[\frac{1}{n}\cdot\sum_{j=1}^{k} a_j\right]_{1\times k}$$

$$\Rightarrow C\cdot A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix} - \begin{bmatrix} \bar{a}_1 & \bar{a}_2 & \cdots & \bar{a}_k \end{bmatrix}$$

$$\Rightarrow C\cdot A = \begin{bmatrix} a_{11}-\bar{a}_1 & a_{12}-\bar{a}_2 & \cdots & a_{1k}-\bar{a}_k \\ \vdots & & & \\ a_{n1}-\bar{a}_1 & a_{n2}-\bar{a}_2 & \cdots & a_{nk}-\bar{a}_k \end{bmatrix}$$

(c) Show that

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 = x^T C x, \quad \text{where} \quad x = [x_1, \dots, x_n]^T.$$

$$x^T C x = \vec{x}^T\left(I_n - \frac{1}{n}\vec{1}\vec{1}^T\right)\vec{x} = \vec{x}^T I_n \vec{x} - \vec{x}^T\cdot\frac{1}{n}\vec{1}\vec{1}^T\vec{x} =$$

$$= \underbrace{\vec{x}^T\vec{x}}_{\sum_{i=1}^{n} x_i\cdot x_i = \sum_{i=1}^{n} x_i^2} - \frac{1}{n}\vec{x}^T\vec{1}\vec{1}^T x = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 = \sum_{i=1}^{n} x_i^2 - \frac{1}{n} n^2\cdot\bar{x}^2$$

$$= \sum_{i=1}^{n} x_i^2 - n\cdot\bar{x}^2 = \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n}\bar{x}^2 = \sum_{i=1}^{n}(x_i^2 - \bar{x}^2)$$

$$= \sum_{i=1}^{n}\left(x_i^2 - 2x_i\bar{x} + \bar{x}^2 + 2x_i\bar{x} - 2\bar{x}^2\right)$$

$$= \sum_{i=1}^{n}(x_i-\bar{x})^2 + 2\cdot\bar{x}\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})}_{0}$$

$$\boxed{= \sum_{i=1}^{n}(x_i-\bar{x})^2}$$

(d) Let $z_1, \ldots, z_n \in \mathbb{R}^p$ be realisations of a multivariate random variable Z. Express the sample mean $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$ and the sample covariance $S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}(z_i - \bar{z})\sum_{j=1}^{n}(z_j - \bar{z})^T\right)$ as matrix operation, using the above results and the design matrix $T = (x_1 \ldots x_n) \in \mathbb{R}^{p\times n}$.

- Sample mean

$$\bar{z} = \frac{1}{n}\cdot T \cdot \vec{1}$$

$$\text{features}\left\{ \overset{T}{\overbrace{\begin{matrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & x_{p3} & \cdots & x_{pn} \end{matrix}}^{\text{observations}}} \right. \times \overset{\vec{1}}{\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{p\times 1}}$$

$$x_{11}\cdot 1 + x_{12}\cdot 1 + \ldots + x_{1n}\cdot 1 = \sum_{i=1}^{n} x_{1i}$$

- Sample covariance

$S^2 = \frac{1}{n-1}\cdot \bar{T}\cdot \bar{T}^{-1}$, where $\bar{T}$ is the centered matrix $T - \bar{z}\cdot \vec{1}$