Matemática para ciencias de los datos: Trabajo práctico 1

M. Sc. Saúl Calderón Ramírez
Instituto Tecnológico de Costa Rica,
Escuela de Computación, Programa de Ciencias de Datos,
PAttern Recongition and MAchine Learning Group (PARMA-Group)

15 de mayo de 2021

Fecha de entrega: Domingo 15 de Mayo 2021
Entrega: Un archivo .zip con el el pdf generado con lyx o latex, y un jupyter notebook en Pytorch, debidamente documentado, con una función definida por ejercicio que lo necesite. A través del TEC-digital.
Modo de trabajo: Grupos de 3 personas.
Integrantes del grupo: Daniel Madriz Granados, Noelia Rojas Ramírez, Steven Jiménez

Resumen

En el presente trabajo práctico se repasarán aspectos básicos del algebra lineal, relacionados con los conceptos a desarrollar a lo largo del curso, mezclando aspectos teóricos y prácticos, usando el lenguaje Python con la librería Pytorch.

1. Operaciones y propiedades básicas de matrices

1. **(10 puntos)** Demuestre que $(A\,B)^T = B^T A^T$. Se recomienda usar la notación de filas y columnas compacta.

   a) Se tiene que

$$A \in \mathbb{R}^{m\times n}$$

$$B \in \mathbb{R}^{n\times p}$$

$$A = \begin{bmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{bmatrix}$$

$$B = \begin{bmatrix} b_{1,1} & \ldots & b_{1,p} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \ldots & b_{n,p} \end{bmatrix}$$

Parte izquierda de la ecuación

$$AB = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} b_{1,1} & \dots & b_{1,p} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \dots & b_{n,p} \end{bmatrix} = \begin{bmatrix} c_{1,1} & \dots & c_{1,p} \\ \vdots & \ddots & \vdots \\ c_{m,1} & \dots & c_{m,p} \end{bmatrix}$$

$$Nota : c_{m,p} = \sum_{k=1}^{n} A_{m,k} B_{k,,p}$$

$$(AB)^T = \begin{bmatrix} c_{1,1} & \dots & c_{1,m} \\ \vdots & \ddots & \vdots \\ c_{p,1} & \dots & c_{p,m} \end{bmatrix}$$

Parte derecha de la ecuación:

$$B^T A^T = \begin{bmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{p,1} & \dots & b_{p,n} \end{bmatrix} \cdot \begin{bmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,m} \end{bmatrix} = \begin{bmatrix} c_{1,1} & \dots & c_{1,p} \\ \vdots & \ddots & \vdots \\ c_{m,1} & \dots & c_{m,p} \end{bmatrix}$$

La parte izquierda y derecha de la ecuación son iguales

2. **(15 puntos)** Para la siguiente matriz:

$$A = s \begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

defina un valor del scalar s que haga la matriz ortonormal, y verifiquelo haciendo $U^T = U^{-1}$.

3. **(15 puntos)** Sean las matrices $A, B \in \mathbb{R}^{n \times n}$, suponiendo que $A + B = I$ y $AB = 0_n$, demuestre que $A^2 = A$ y$B^2 = B$.

   Sea $A = I - B$.

   Retomando la ecuación $AB = 0_n$

$$(I - B)B = 0_n$$

$$IB - BB = 0_n$$

$$IB = 0_n + BB$$

$$B = 0_n + BB$$

$$B = BB$$

$$B = B^2$$

4. **(5 puntos extra)** Sea la matriz $A \in \mathbb{R}^{n\times n}$ simétrica y una matriz cualquiera (no puede asumir que es simétrica) $P \in \mathbb{R}^{n\times n}$. Demuestre que la matriz $P^T AP$ es también simétrica.

$$(P^T AP)^T =$$
$$(AP)^T(P^T)^T =$$
$$P^T A^T P =$$
$$P^T AP$$

Pues A es simetrica, se tiene que $P^T A^T P = P^T AP$ como se muestra a continuación

Se tiene que

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix}$$

$$P = \begin{bmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{bmatrix}$$

Ahora,

$$P^T A^T P = \begin{bmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{bmatrix} \cdot \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{bmatrix}$$

$$P^T AP = \begin{bmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{bmatrix} \cdot \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix} \cdot \begin{bmatrix} p_{1,1} & \dots & p_{1,n} \\ \vdots & \ddots & \vdots \\ p_{n,1} & \dots & p_{n,n} \end{bmatrix}$$

2. Ecuaciones vectoriales-matriciales

1. **(15 puntos)** Muestre con un ejemplo numérico que para un vector $\overrightarrow{x} \in \mathbb{R}^n$ y una matriz simétrica $A \in \mathbb{R}^{n\times n}$ que:

$$\left(\vec{x}^T \, A \, \vec{x}\right)^T = \vec{x}^T \, A \, \vec{x}$$

   a) Incluya el código en Pytorch que permita corroborar tal igualdad para cualquier matriz simétrica A y vector $\overrightarrow{x} \in \mathbb{R}^n$, generado aleatoriamente. Recuerde que a partir de cualquier matriz cuadrada A generada aleatoriamente puede calcularse una matriz simétrica, según lo discutido en clase.

2. **(15 puntos)** Demuestre que la siguiente ecuación matricial:

$$\left\| A\,\vec{x} - \vec{b} \right\|^2 - \left\| \vec{b} \right\|^2$$

con $\overrightarrow{x} \in \mathbb{R}^n$, $\overrightarrow{b} \in \mathbb{R}^m$ y $A \in \mathbb{R}^{m \times n}$,se puede reescribir como sigue:

$$\vec{x}^T\, A^T A\, \vec{x} - 2\,\vec{b}^T A\, \vec{x}$$

Sean:

$$\overrightarrow{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}; \overrightarrow{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix};$$

$$A = \begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} -- & \overrightarrow{a_{1,:}} & -- \\ -- & \vdots & -- \\ -- & \overrightarrow{a_{m,:}} & -- \end{bmatrix}$$

Primero calculamos $A\overrightarrow{x} - \overrightarrow{b}$

$$A\overrightarrow{x} - \overrightarrow{b} = \begin{bmatrix} -- & \overrightarrow{a_{1,:}} & -- \\ -- & \vdots & -- \\ -- & \overrightarrow{a_{m,:}} & -- \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$A\overrightarrow{x} - \overrightarrow{b} = \begin{bmatrix} -- & \overrightarrow{a_{1,:}^T} \cdot \overrightarrow{x} & -- \\ -- & \vdots & -- \\ -- & \overrightarrow{a_{m,:}^T} \cdot \overrightarrow{x} & -- \end{bmatrix} - \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$A\overrightarrow{x} - \overrightarrow{b} = \begin{bmatrix} -- & \overrightarrow{a_{1,:}^T} \cdot \overrightarrow{x} - b_1 & -- \\ -- & \vdots & -- \\ -- & \overrightarrow{a_{m,:}^T} \cdot \overrightarrow{x} - b_m & -- \end{bmatrix}$$

$$||A\overrightarrow{x} - \overrightarrow{b}|| = \sqrt{\left(\overrightarrow{a_{1,:}^T} \cdot \overrightarrow{x} - b_1\right)^2 + \cdots + \left(\overrightarrow{a_{m,:}^T} \cdot \overrightarrow{x} - b_m\right)^2}$$

$$||A\overrightarrow{x} - \overrightarrow{b}||^2 = \left(\overrightarrow{a_{1,:}^T} \cdot \overrightarrow{x} - b_1\right)^2 + \cdots + \left(\overrightarrow{a_{m,:}^T} \cdot \overrightarrow{x} - b_m\right)^2$$

$$||A\overrightarrow{x} - \overrightarrow{b}||^2 = \sum_{i=1}^{m} \left(\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x} - b_i\right)^2$$

$$||A\overrightarrow{x} - \overrightarrow{b}||^2 = \sum_{i=1}^{m} ((\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})^2 - 2(\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})b_i + b_i^2)$$

Luego tenemos $\left\|\vec{b}\right\|^2$que

$\left\|\vec{b}\right\|^2 = \sum_{i=1}^{m} b_i^2$

Por lo que

$||A\overrightarrow{x} - \overrightarrow{b}||^2 - \left\|\vec{b}\right\|^2 = \sum_{i=1}^{m} ((\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})^2 - 2(\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})b_i + b_i^2) - \sum_{i=1}^{m} b_i^2$

$||A\overrightarrow{x} - \overrightarrow{b}||^2 - \left\|\vec{b}\right\|^2 = \sum_{i=1}^{m} ((\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})^2 - 2(\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})b_i) + \cancel{\sum_{i=1}^{m} b_i^2} - \cancel{\sum_{i=1}^{m} b_i^2}$

$||A\overrightarrow{x} - \overrightarrow{b}||^2 - \left\|\vec{b}\right\|^2 = \sum_{i=1}^{m} ((\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})^2 - 2(\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})b_i)$

Ahora debemos comprobar que

$\vec{x}^T A^T A\vec{x} - 2\,\vec{b}^T Ax = ||A\overrightarrow{x} - \overrightarrow{b}||^2 - ||\overrightarrow{b}||^2 = \sum_{i=1}^{\vec{m}} ((\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})^2 - 2(\overrightarrow{a_{i,:}^T} \cdot \overrightarrow{x})b_i)$

Sean:

$$\overrightarrow{x}^T = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix};\ \overrightarrow{b}^T = \begin{bmatrix} b_1 & \cdots & b_m \end{bmatrix};$$

$$A^T = \begin{bmatrix} | & \cdots & | \\ \overrightarrow{a_{1,:}} & \cdots & \overrightarrow{a_{m,:}} \\ | & \cdots & | \end{bmatrix}$$

Tenemos que $\vec{x}^T A^T$

$$\vec{x}^T A^T = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} | & \cdots & | \\ \overrightarrow{a_{1,:}} & \cdots & \overrightarrow{a_{m,:}} \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} \overrightarrow{x}^T\overrightarrow{a_{1,:}} & \cdots & \overrightarrow{x}^T\overrightarrow{a_{m,:}} \end{bmatrix}$$

Tenemos que $A\overrightarrow{x}$

$$A\overrightarrow{x} = \begin{bmatrix} -- & \overrightarrow{a_{1,:}^T} \cdot \overrightarrow{x} & -- \\ -- & \vdots & -- \\ -- & \overrightarrow{a_{m,:}^T} \cdot \overrightarrow{x} & -- \end{bmatrix}$$

Por lo que $\vec{x}^T A^T A\,\vec{x}$

$$\vec{x}^T A^T A\,\vec{x} = \begin{bmatrix} \overrightarrow{x}^T\overrightarrow{a_{1,:}} & \cdots & \overrightarrow{x}^T\overrightarrow{a_{m,:}} \end{bmatrix} \begin{bmatrix} -- & \overrightarrow{a_{1,:}^T} \cdot \overrightarrow{x} & -- \\ -- & \vdots & -- \\ -- & \overrightarrow{a_{m,:}^T} \cdot \overrightarrow{x} & -- \end{bmatrix}$$

$\vec{x}^T A^T A\,\vec{x} = \overrightarrow{x}^T\overrightarrow{a_{1,:}}\overrightarrow{a_{1,:}^T} \cdot \overrightarrow{x} + ... + \overrightarrow{x}^T\overrightarrow{a_{m,:}}\overrightarrow{a_{m,:}^T} \cdot \overrightarrow{x}$

Se observa que $\overrightarrow{x}^T \cdot \overrightarrow{a_{1,:}} = \overrightarrow{a_{1,:}} \cdot \overrightarrow{x}$,por lo tanto

$\vec{x}^T A^T A\,\vec{x} = \sum_{i=1}^{m} \left(\overrightarrow{x}^T\overrightarrow{a_{1,:}}\right)^2$

Luego para $2\,\vec{b}^T Ax$

$$2\,\vec{b}^T Ax = 2\begin{bmatrix} b_1 & \cdots & b_m \end{bmatrix}\begin{bmatrix} -- & \overrightarrow{a^T_{1,:}}\cdot\overrightarrow{x} & -- \\ -- & \vdots & -- \\ -- & \overrightarrow{a^T_{m,:}}\cdot\overrightarrow{x} & -- \end{bmatrix}$$

$$2\,\vec{b}^T Ax = 2\left[b_1(\overrightarrow{a^T_{1,:}}\cdot\overrightarrow{x}) + ... + b_m(\overrightarrow{a^T_{m,:}}\cdot\overrightarrow{x})\right] = \sum_{i=1}^{m} 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})$$

$$2\,\vec{b}^T Ax = \sum_{i=1}^{m} 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})$$

Uniendo la ecuación completa $\vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax$ se tiene que

$$\vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax = \sum_{i=1}^{m}\left(\overrightarrow{x}^T\overrightarrow{a_{1,:}}\right)^2 - \sum_{i=1}^{m} 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})$$

$$\vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax = \sum_{i=1}^{m}\left(\left(\overrightarrow{x}^T\overrightarrow{a_{1,:}}\right)^2 - 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})\right)$$

Por lo tanto se comprueba que $\left\|A\,\vec{x} - \vec{b}\right\|^2 - \left\|\vec{b}\right\|^2$ se puede reescribir como $\vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax$

$$||A\overrightarrow{x} - \overrightarrow{b}||^2 - \left\|\vec{b}\right\|^2 = \sum_{i=1}^{m}((\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})^2 - 2(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})b_i)$$

$$\vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax = \sum_{i=1}^{m}\left(\overrightarrow{x}^T\overrightarrow{a_{1,:}}\right)^2 - \sum_{i=1}^{m} 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})$$

$$\vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax = \left(\left(\overrightarrow{x}^T\overrightarrow{a_{1,:}}\right)^2 - 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})\right)$$

$$\vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax = \sum_{i=1}^{m}\left(\left(\overrightarrow{x}^T\overrightarrow{a_{1,:}}\right)^2 - 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})\right)$$

$$||A\overrightarrow{x} - \overrightarrow{b}||^2 - \left\|\vec{b}\right\|^2 = \vec{x}^T\, A^T A\,\vec{x} - 2\,\vec{b}^T Ax = \sum_{i=1}^{m}\left(\left(\overrightarrow{x}^T\overrightarrow{a_{1,:}}\right)^2 - 2b_i(\overrightarrow{a^T_{i,:}}\cdot\overrightarrow{x})\right)$$

3. **(10 puntos)** Sea $f(\vec{x}) = \frac{1}{2}\vec{x}^T A\vec{x} + \vec{b}^T\vec{x} + \vec{b}\cdot\vec{x}$ con A una matriz simétrica, calcule $\nabla_{\vec{x}} f(\vec{x})$.

AA

4. **(10 puntos extra)** Sea $f(\vec{x}) = g(h(\vec{x}))$ con $g : \mathbb{R} \to \mathbb{R}$ y $h : \mathbb{R}^n \to \mathbb{R}$ calcule $\nabla f(\vec{x})$.

$$\nabla f(\vec{x}) = g'(h(\vec{x}))\, h'(\vec{x}) = \begin{bmatrix} \frac{\partial g(h(\vec{x}))}{\partial x_1} \frac{\partial h(\vec{x})}{\partial x_1} \\ \frac{\partial g(h(\vec{x}))}{\partial x_2} \frac{\partial h(\vec{x})}{\partial x_2} \\ \vdots \\ \frac{\partial g(h(\vec{x}))}{\partial x_n} \frac{\partial h(\vec{x})}{\partial x_n} \end{bmatrix}$$

3. (20 puntos) La matriz de covarianza

1. Para dos variables aleatorias X e Y se define la covarianza como el valores esperado de la diferencia de una variable aleatoria y su esperanza (media):

$$\Sigma_{X,Y} = \text{cov}\left(X,Y\right) = \mathbb{E}\left[\left(X - \mathbb{E}\left[X\right]\right)\left(Y - \mathbb{E}\left[Y\right]\right)\right]$$

y mide la variación conjunta de tales variables aleatorias. Para el caso de contar con arreglos de muestras $h\left[u\right]$ y $g\left[u\right]$ para las variables aleatorias X e Y respectivamente, se tiene que la covarianza de tales variables aleatorias está dada por:

$$\Sigma_{X,Y} \cong \frac{1}{N-1}\sum_{u=1}^{N}\left(h\left[u\right] - \mu_X\right)\left(g\left[u\right] - \mu_Y\right).$$

con las medias o esperanzas $\mu_X = \mathbb{E}\left[X\right]$ y $\mu_Y = \mathbb{E}\left[Y\right]$, componentes del vector medio

$$\overrightarrow{\mu} = \begin{bmatrix}\mu_X\\ \mu_Y\end{bmatrix}$$

La matriz de covarianza para n variables aleatorias $X_1, X_2, \ldots, X_n$ se define como:

$$\Sigma = \begin{bmatrix}\mathbb{E}\left[\left(X_1 - \mathbb{E}\left[X_1\right]\right)\left(X_1 - \mathbb{E}\left[X_1\right]\right)\right] & \ldots & \mathbb{E}\left[\left(X_1 - \mathbb{E}\left[X_1\right]\right)\left(X_n - \mathbb{E}\left[X_n\right]\right)\right]\\ \vdots & \ddots & \vdots\\ \mathbb{E}\left[\left(X_n - \mathbb{E}\left[X_n\right]\right)\left(X_1 - \mathbb{E}\left[X_1\right]\right)\right] & \ldots & \mathbb{E}\left[\left(X_n - \mathbb{E}\left[X_n\right]\right)\left(X_n - \mathbb{E}\left[X_n\right]\right)\right]\end{bmatrix},$$

observe que en la diagonal de la matriz Σ (entrada $\Sigma_{i,i}$) se tiene que

$$\mathbb{E}\left[\left(X_i - \mathbb{E}\left[X_i\right]\right)\left(X_i - \mathbb{E}\left[X_i\right]\right)\right] = \sigma_{X_i}^2,$$

por lo que entonces la matriz de covarianza se puede reescribir como:

$$\Sigma = \begin{bmatrix}\sigma_{X_1}^2 & \ldots & \mathbb{E}\left[\left(X_1 - \mathbb{E}\left[X_1\right]\right)\left(X_n - \mathbb{E}\left[X_n\right]\right)\right]\\ \vdots & \ddots & \vdots\\ \mathbb{E}\left[\left(X_n - \mathbb{E}\left[X_n\right]\right)\left(X_1 - \mathbb{E}\left[X_1\right]\right)\right] & \ldots & \sigma_{X_n}^2\end{bmatrix}.$$

Además, la matriz de covarianza Σ presenta la propiedad de ser simétrica, puesto que $\mathbb{E}\left[\left(X_i - \mathbb{E}\left[X_i\right]\right)\left(X_j - \mathbb{E}\left[X_j\right]\right)\right] = \mathbb{E}\left[\left(X_j - \mathbb{E}\left[X_j\right]\right)\left(X_i - \mathbb{E}\left[X_i\right]\right)\right] \Rightarrow \Sigma_{X_i,X_j} = \Sigma_{X_j,X_i}$.

Ejemplo

Suponga que se desea encontrar la matriz de covarianza para tres variables aleatorias X_1, X_2 y X_3, para las cuales se han recabado los siguientes arreglos de muestras para $N = 4$ experimentos, respectivamente:

$$h_1 = \begin{bmatrix} 2 & 4 & 6 & 8 \end{bmatrix}$$
$$h_2 = \begin{bmatrix} 4 & 8 & 12 & 16 \end{bmatrix}$$
$$h_3 = \begin{bmatrix} 12 & 10 & 5 & 9 \end{bmatrix}$$

En términos de muestras se tienen 4 muestras

$$U = \{\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4\} = \begin{bmatrix} | & | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 & \vec{u}_4 \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} 2 & 4 & 6 & 8 \\ 4 & 8 & 12 & 16 \\ 12 & 10 & 5 & 9 \end{bmatrix}$$

con $u_i \in \mathbb{R}^3$, donde cada dimensión es una variable aleatoria, y $U \in \mathbb{R}^{3\times 4}$.

Observe en estos datos, que la dimensión 1 y 2 son combinación lineal, por lo que la covarianza de ambas dimensiones debe ser alta, no así la dimensión 1 con la 3 o la 2 con la 3. Además

Se procede entonces a calcular las entradas Σ_{X_1,X_2}, Σ_{X_1,X_3} y Σ_{X_2,X_3}, además de los valores de la diagonal $\sigma^2_{X_1}$, $\sigma^2_{X_2}$ y $\sigma^2_{X_3}$, teniendo en cuenta que $\mu_{X_1} = 5 = \frac{2+4+6+8}{4}$, $\mu_{X_2} = 10$ y $\mu_{X_3} = 9$, con lo que entonces:

$$\overrightarrow{\mu} = \begin{bmatrix} \mu_{X_1} = 5 \\ \mu_{X_2} = 10 \\ \mu_{X_3} = 9 \end{bmatrix}$$

y haciendo los cálculos respectivos:

$$\begin{aligned} \Sigma_{X_1,X_2} &= \tfrac{1}{4-1}\left((5-2)(10-4) + (5-4)(10-8) + (5-6)(10-12) + (5-8)(10-16)\right) \\ \Sigma_{X_1,X_3} &= \tfrac{1}{4-1}\left((5-2)(9-12) + (5-4)(9-10) + (5-6)(9-5) + (5-8)(9-9)\right) \\ \Sigma_{X_2,X_3} &= \tfrac{1}{4-1}\left((10-4)(9-12) + (10-8)(9-10) + (10-12)(9-5) + (10-16)(9-9)\right) \\ \sigma^2_{X_1} &= \tfrac{1}{4-1}\left((5-2)^2 + (5-4)^2 + (5-6)^2 + (5-8)^2\right) \\ \sigma^2_{X_2} &= \tfrac{1}{4-1}\left((10-4)^2 + (10-8)^2 + (10-12)^2 + (10-16)^2\right) \\ \sigma^2_{X_3} &= \tfrac{1}{4-1}\left((9-12)^2 + (9-10)^2 + (9-5)^2 + (9-9)^2\right) \end{aligned}$$

lo cual desarrollado corresponde a:

$$\begin{aligned} \Sigma_{X_1,X_2} &= \tfrac{1}{3}\left(3\cdot 6 + 1\cdot 2 + -1\cdot -2 + -3\cdot -6\right) = \tfrac{40}{3} = 13{,}333 \\ \Sigma_{X_1,X_3} &= \tfrac{1}{3}\left(3\cdot -3 + 1\cdot -1 + -1\cdot 4 + -3\cdot 0\right) = -\tfrac{14}{3} = -4{,}667 \\ \Sigma_{X_2,X_3} &= \tfrac{1}{3}\left(6\cdot -3 + 2\cdot -1 + -2\cdot 4 + -6\cdot 0\right) = -\tfrac{28}{3} = -9{,}333 \\ \sigma^2_{X_1} &= \tfrac{1}{3}\left(9+1+1+9\right) = \tfrac{20}{3} = 6{,}667 \\ \sigma^2_{X_2} &= \tfrac{1}{3}\left(36+4+4+36\right) = \tfrac{80}{3} = 26{,}667 \\ \sigma^2_{X_3} &= \tfrac{1}{3}\left(9+1+16+0\right) = \tfrac{14}{3} = 8{,}667. \end{aligned}$$

Por lo que se obtiene la matriz de covarianza:

$$\Sigma = \begin{bmatrix} \frac{20}{3} & \frac{40}{3} & -\frac{14}{3} \\ \frac{40}{3} & \frac{80}{3} & -\frac{28}{3} \\ -\frac{14}{3} & -\frac{28}{3} & \frac{14}{3} \end{bmatrix} = \begin{bmatrix} 6{,}667 & 13{,}333 & -4{,}667 \\ 13{,}333 & 26{,}666 & -9{,}333 \\ -4{,}667 & -9{,}333 & 8{,}667 \end{bmatrix}.$$

También la matriz de covarianza se puede escribir, para un conjunto de muestras $X = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\}$, con $\vec{x}_i \in \mathbb{R}^{n\times 1}$, como:

$$\Sigma = \frac{1}{m-1}\sum_{i=1}^{m} (\vec{x}_i - \vec{\mu})(\vec{x}_i - \vec{\mu})^T$$

donde $\vec{\mu} \in \mathbb{R}^{n\times 1}$ es la muestra promedio del conjunto de datos X (donde cada componente es el valor medio de cada dimensión).

a) Escriba la función en Pytorch sin usar estructuras de repetición *for* o *while*, usando únicamente las funciones para cálculo de vector medio, y multiplicaciones matriciales. Documente su uso con el ejemplo descrito en este documento. Compruebe su correcto funcionamiento, implementando el ejemplo desarrollado en este documento.

Solucion

```
import torch
import numpy as np

#Variable del problema/enunciado
h1 = torch.tensor([2.0,4,6,8])
#Variable del problema/enunciado
h2 = torch.tensor([4.0,8,12,16])
#Variable del problema/enunciado
h3 = torch.tensor([12.0,10,5,9])
#Creacion de la matriz K compuesta por h1, h2, h3 print("K",K)
K = torch.stack([h1,h2,h3],dim=0)

def covMatrix(U):
  #calcular la cantidad de observaciones
  #por cada variable/dimension
  ObservationsQty = U.size(1)

  #calcular la media para cada variable
  #(fila) de la matriz U
  Means = torch.mean(U,1,True)

  #Calcular la diferencia entre las
  #observaciones de cada variable y su media
  difference = U-Means
```

```
    # calculo de la matriz de covarianza
    COVmatrixresult = ((1)/ (observationsQty -1)) *
           difference.mm(difference.transpose(0,1))


    #retornar el resultado de la matriz de covarianza
    return COVmatrixresult


#comprobar el funcionamiento
#de la matriz de covarianza
print("Covariance matrix for input matrix",covMatrix(K))
```

Resultado

```
K tensor
([[ 2.,  4.,  6.,  8.],
  [ 4.,  8., 12., 16.],
  [12., 10.,  5.,  9.]])


Covariance matrix for input matrix tensor
([[ 6.6667, 13.3333, -4.6667],
  [13.3333, 26.6667, -9.3333],
  [-4.6667, -9.3333,  8.6667]])
```