

Cálculo de la matriz de covarianza Σ

Dr. Mario A. Gómez Rodríguez

5 de junio de 2020

Índice

1. Fórmula y nomenclatura	3
1.1. Ejemplo	3

1. Fórmula y nomenclatura

$$\sum_{j,k} = \frac{1}{n-1} \sum_{i=1}^n (X_{i,j} - \overline{X_j})(X_{i,k} - \overline{X_k})$$

\sum : matriz de covarianza.

X : conjunto de datos.

$\sum_{j,k}$: elemento en la fila j y columna k de la matriz de covarianza. Se corresponde con la covarianza entre los atributos (o columnas) j y k del conjunto de datos (X).

n : número de objetos de datos en el conjunto de datos X .

j, k : denotan dos cosas: 1) elemento en la fila j y columna k de la matriz de covarianza; 2) atributos (o columnas) j y k del conjunto de datos X .

$X_{i,j}$: valor del atributo (o columna) j en el objeto de datos i , del conjunto de datos X . Se corresponde con la fila i y columna j del conjunto de datos X .

$X_{i,k}$: valor del atributo (o columna) k en el objeto de datos i , del conjunto de datos X . Se corresponde con la fila i y columna k del conjunto de datos X .

$\overline{X_j}$: promedio del atributo (o columna) j del conjunto de datos X .

$\overline{X_k}$: promedio del atributo (o columna) k del conjunto de datos X .

$$\begin{bmatrix} & \textcolor{red}{j} & \textcolor{red}{k} \\ 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \leftarrow \text{Conjunto de datos } X.$$

$$\textcolor{red}{j} \begin{bmatrix} & & \textcolor{red}{k} \\ \mathbb{R} & \mathbb{R} & \mathbb{R} \\ \mathbb{R} & \mathbb{R} & \sum_{j,k} \\ \mathbb{R} & \mathbb{R} & \mathbb{R} \end{bmatrix} \leftarrow \text{Matriz de covarianza } \sum.$$

1.1. Ejemplo

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \leftarrow \text{Conjunto de datos } X.$$

$$j = 2$$

$$k = 3$$

$$n = 3$$

$$\textcolor{red}{j} = 2 \begin{bmatrix} & & \textcolor{red}{k} = 3 \\ \mathbb{R} & \mathbb{R} & \mathbb{R} \\ \mathbb{R} & \mathbb{R} & \sum_{2,3} \\ \mathbb{R} & \mathbb{R} & \mathbb{R} \end{bmatrix} \leftarrow \text{Matriz de covarianza } \sum.$$

$$\sum_{2,3} = \frac{1}{3-1} \sum_{i=1}^3 (X_{i,2} - \overline{X_2})(X_{i,3} - \overline{X_3})$$

$$\overline{X_2} = \frac{1}{3} \sum_{i=1}^3 X_{i,2} = \frac{2+5+8}{3} = \frac{15}{3} = 5$$

$$\overline{X_3} = \frac{1}{3} \sum_{i=1}^3 X_{i,3} = \frac{3+6+9}{3} = \frac{18}{3} = 6$$

$$\begin{array}{lclclclclcl} i=1) & 2-5 & \times & 3-6 & = & -3 & \times & -3 & = & 9 \\ i=2) & 5-5 & \times & 6-6 & = & 0 & \times & 0 & = & 0 \\ i=3) & 8-5 & \times & 9-6 & = & 3 & \times & 3 & = & 9 \end{array}$$

$$\sum_{2,3} = \frac{1}{3-1} (9+0+9) = \frac{18}{2} = 9$$

$$\sum_{2,3} = 9$$

$$j=2 \left[\begin{array}{ccc} \mathbb{R} & \mathbb{R} & \overset{k=3}{\mathbb{R}} \\ \mathbb{R} & \mathbb{R} & 9 \\ \mathbb{R} & \mathbb{R} & \mathbb{R} \end{array} \right] \leftarrow \text{Matriz de covarianza } \Sigma.$$