

Matrices de covarianzas

Contents

Sean $m + 1$ variables, con n observaciones cada una:

y	x_1	x_2	\cdots	x_m
y_1	x_{11}	x_{21}	\cdots	x_{m1}
y_2	x_{12}	x_{22}	\cdots	x_{m2}
\cdots	\cdots	\cdots	\cdots	\cdots
y_n	x_{1n}	x_{2n}	\cdots	x_{mn}

se define la covarianza entre las variables x_j e y como:

$$\text{cov}(x_j, y) = s_{jy} = \frac{\sum_{i=1}^n (x_{ji} - \bar{x}_j)(y_i - \bar{y})}{n - 1}, \quad j \in [1, m], \quad i \in [1, n]$$

y la covarianza entre las variables x_j e x_k como:

$$\text{cov}(x_j, x_k) = s_{jk} = \frac{\sum_{i=1}^n (x_{ji} - \bar{x}_j)(x_{ki} - \bar{x}_k)}{n - 1}, \quad j, k \in [1, m], \quad i \in [1, n]$$

Se define la matriz de covarianzas para x_j y x_k como:

$$S_{xx} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{21} & s_{22} & \cdots & s_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ s_{m1} & s_{m2} & \cdots & s_{mm} \end{bmatrix}$$

Se define la matriz de covarianzas entre x_j e y como:

$$S_{xy} = \begin{bmatrix} s_{1y} \\ s_{2y} \\ \cdots \\ s_{my} \end{bmatrix}$$

Las matrices S_{xx} y S_{xy} se pueden calcular a partir de la matriz

$$X = \begin{bmatrix} y_1 - \bar{y} & x_{11} - \bar{x}_1 & x_{21} - \bar{x}_2 & \cdots & x_{m1} - \bar{x}_m \\ y_2 - \bar{y} & x_{12} - \bar{x}_1 & x_{22} - \bar{x}_2 & \cdots & x_{m2} - \bar{x}_m \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ y_n - \bar{y} & x_{1n} - \bar{x}_1 & x_{2n} - \bar{x}_2 & \cdots & x_{mn} - \bar{x}_m \end{bmatrix}$$

Efectivamente:

$$\frac{1}{n-1}X^TX = \left[\begin{array}{c|ccccc} s_y^2 & | & s_{1y} & s_{1y} & \cdots & s_{1y} \\ \hline - & - & - & - & - & - \\ s_{1y} & | & s_{11} & s_{12} & \cdots & s_{1m} \\ s_{2y} & | & s_{12} & s_{22} & \cdots & s_{2m} \\ \cdots & | & \cdots & \cdots & \cdots & \cdots \\ s_{my} & | & s_{m1} & s_{m2} & \cdots & s_{mm} \end{array} \right] = \left[\begin{array}{c|c} s_y^2 & S_{xy}^T \\ \hline S_{yx} & S_{xx} \end{array} \right]$$