

Análisis multivariado

Emmanuel Alcalá

Fuentes bibliográficas:

Introducción

Notación

El análisis multivariado tiene como objeto el análisis y entendimiento de datos en varias dimensiones. Partimos de un conjunto $\{x_i\}_{i=1}^n$ de un arreglo (un vector variable) $\mathbf{X} \in \mathbb{R}^p$

Cada *observación* x_j tiene p dimensiones. Denotamos como x_{jk} a la k variable observada en la j observación.

$$x = \begin{bmatrix} x_{i1} & x_{i2} & \dots & x_{ip} \end{bmatrix}$$

	X_1	X_2	\dots	X_k	\dots	X_p
Obs 1	x_{11}	x_{12}	\dots	x_{1k}	\dots	x_{1p}
Obs 2	x_{11}	x_{12}	\dots	x_{1k}	\dots	x_{1p}
\vdots	\vdots	\vdots	\dots	\vdots	\dots	\vdots
Obs j	x_{j1}	x_{j2}	\dots	x_{jk}	\dots	x_{jp}
\vdots	\vdots	\vdots	\dots	\vdots	\dots	\vdots
Obs n	x_{n1}	x_{n2}	\dots	x_{nk}	\dots	x_{np}

Que podemos resumir por el arreglo rectangular \mathbf{X} con columnas \mathbf{X}_k

$$\mathbf{X} = \begin{bmatrix} X_1 & X_2 & \dots & X_k & \dots & X_p \end{bmatrix}$$

Ejemplo:

Una selección de 4 recibos de una librería se obtuvieron para evaluar la venta de libros. Cada recibo tenía el número de libros vendidos y la cantidad de cada venta. Sea la primera variable la venta en dólares y la segunda el número de libros. Podemos considerar el número correspondiente de recibos como cuatro mediciones de dos variables. Los datos en forma tabular serían

Variable 1 (dólares):	42	52	48	58
Variable 2 (n libros):	4	5	4	3

Con la notación anterior tendríamos

$x_{11} = 42$	$x_{21} = 52$	$x_{31} = 48$	$x_{41} = 58$
$x_{12} = 4$	$x_{22} = 5$	$x_{32} = 4$	$x_{42} = 3$

o

$$\mathbf{X} = \begin{bmatrix} 42 & 4 \\ 52 & 5 \\ 48 & 4 \\ 58 & 3 \end{bmatrix}$$

Descriptivos
