18-11-2021

REGRECION LINEAL MULTIPLE

CMASOO 1.2.1

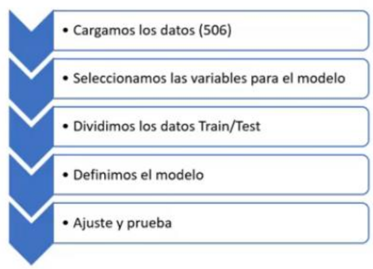
KEVIN ADRIAN CRUZ JACOBO
CDIA 301 COANLEP

Desripcion del dataset Boston

Predecir el valor de viviendas en Boston a partir de 13 variables independientes

Variables independientes — Variable dependiente

	LSTAT	RM	NOX	TAX	DIS	MEDV
0	4.98	6.575	0.538	296.0	4.0900	24.0
1	9.14	6.421	0.469	242.0	4.9671	21.6
2	4.03	7.185	0.469	242.0	4.9671	34.7
3	2.94	6.998	0.458	222.0	6.0622	33.4
4	5.33	7.147	0.458	222.0	6.0622	36.2



MEDV=13.634-0.700 X LSTAT+4.476 X RM-6.505 X NOX-0.0076 X TAX -0.968 X DIS

$\varepsilon^T \varepsilon = 4.76$ $R^2 = 0.71$

Fundamentación

Modelo de abstracción que representa la relación entre una variable dependiente Y_i (regresando) y otras variables explicativas independientes X_{ji} (regresores)

Se asume que las observaciones satisfacen una relación lineal

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_1 x_{pi} + \varepsilon_i$$

- Supongamos un conjunto de N observaciones en un modelo lineal P variables independientes

observación — Variables explicativas

$$\{x_{11}, x_{21}, x_{31} \dots, x_{p1}, y_1\}$$
$$\{x_{12}, x_{22}, x_{32} \dots, x_{p2}, y_2\}$$
$$\{x_{13}, x_{23}, x_{33} \dots, x_{p3}, y_3\}$$
$$\vdots$$
$$\{x_{1n}, x_{2n}, x_{3n} \dots, x_{pn}, y_n\}$$

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \cdots + \beta_1 x_{p1} + \varepsilon_1$$
$$y_2 = \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} + \cdots + \beta_1 x_{p2} + \varepsilon_2$$
$$y_3 = \beta_0 + \beta_1 x_{13} + \beta_2 x_{23} + \cdots + \beta_1 x_{p3} + \varepsilon_3$$
$$\vdots$$
$$y_n = \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \cdots + \beta_1 x_{pn} + \varepsilon_n$$

Función objetivo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_1 x_{pi} + \varepsilon_i$$

Representación en algebra matricial

$$Y = X\beta + \varepsilon$$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{11} & x_{21} & x_{31} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & x_{32} & & x_{p2} \\ 1 & x_{13} & x_{23} & x_{33} & & x_{p3} \\ & & \vdots & & & \\ 1 & x_{1n} & x_{2n} & x_{3n} & & x_{pn} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$n \times 1$ $\quad$ $n \times (p+1)$ $\quad$ $(p+1) \times 1$ $\quad$ $n \times 1$

$$min \sum_{i=1}^{n} \varepsilon_i^2 = \min(\varepsilon^T \varepsilon)$$

Minimizar la suma de los cuadrados de los errores

$$Y = X\beta + \varepsilon \begin{cases} \varepsilon = Y - X\beta \\ \varepsilon^T = (Y - X\beta)^T \end{cases}$$

$$min \sum_{i=1}^{n} \varepsilon_i^2 = \min(\varepsilon^T \varepsilon)$$

Minimizar la suma de los cuadrados de los errores

$$\varepsilon^T \varepsilon = (Y - X\beta)^T (Y - X\beta)$$
$$\varepsilon^T \varepsilon = (Y^T - (X\beta)^T)(Y - X\beta)$$
$$\varepsilon^T \varepsilon = (Y^T - \beta^T X^T)(Y - X\beta)$$
$$\varepsilon^T \varepsilon = Y^T Y - Y^T X\beta - \beta^T X^T Y + \beta^T X^T X\beta$$

$$\boxed{L(\beta) = Y^T Y - 2Y^T X\beta + \beta^T X^T X\beta}$$

Gradiente

$$\min \sum_{i=1}^{n} \varepsilon_i^2 = \min(\varepsilon^T \varepsilon)$$

Minimizar la suma de los cuadrados de los errores

$$\boxed{L(\beta) = Y^T Y - 2Y^T X\beta + \beta^T X^T X\beta}$$

$$Y = X\beta + \varepsilon \quad \begin{cases} \varepsilon = Y - X\beta \\ \varepsilon^T = (Y - X\beta)^T \end{cases}$$

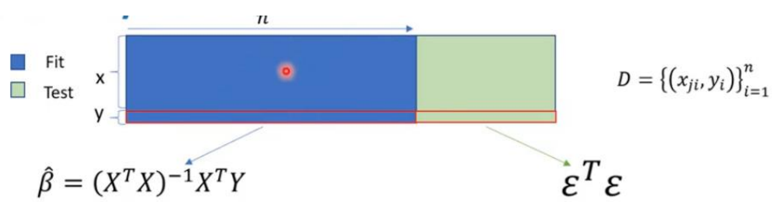
$$\frac{\partial L}{\partial \beta} = -2X^T Y + 2X^T X\beta = 0$$

$$X^T X\beta = X^T Y$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

La estimación de mínimos cuadros satisface esta expresión

Predicciones y error



Supongamos un conjunto de datos de prueba (Y_{TEST} , - X_{TEST})
Supongamos también un modelo determinado β = $(X^T X)^{-1} X^T X$
Entonces el error de predicción seria ε = $Y_{test} - X_{test}\,\beta$

$$\boxed{\varepsilon^t \varepsilon}$$