Introducción al modelo lineal
Soluciones a los ejercicios opcionales

Francesc Carmona

14 de febrero de 2023

Ejercicios del libro de Faraway

Los ejercicios 3, 4 y 5 del capítulo 1 son similares a los ejercicios 1 y 2.

Ejercicios del libro de Carmona

Ejercicio 1.1

Hallar las estimaciones de los parámetros en un modelo de regresión lineal simple, minimizando la suma de los cuadrados de los errores:

$$S = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

La solución se puede ver en el Apéndice A del módulo **El modelo lineal**.

Ejercicio 1.2

Hallar las estimaciones de los parámetros en un modelo de regresión parabólico, minimizando la suma de los cuadrados de los errores:

$$S = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)^2$$

Del mismo modo que en el ejercicio 1.1 vamos a derivar respecto a cada uno de los parámetros

$$\begin{aligned}
\frac{\partial S}{\partial \beta_0} &= 2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)(-1) \\
\frac{\partial S}{\partial \beta_1} &= 2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)(-x_i) \\
\frac{\partial S}{\partial \beta_2} &= 2\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i - \beta_2 x_i^2)(-x_i^2)
\end{aligned}$$

y al igualar a cero obtenemos

$$\begin{aligned}
-\sum_{i=1}^{n} y_i + n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i + \beta_2 \sum_{i=1}^{n} x_i^2 &= 0 \\
-\sum_{i=1}^{n} x_i y_i + \beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2 + \beta_2 \sum_{i=1}^{n} x_i^3 &= 0 \\
-\sum_{i=1}^{n} x_i^2 y_i + \beta_0 \sum_{i=1}^{n} x_i^2 + \beta_1 \sum_{i=1}^{n} x_i^3 + \beta_2 \sum_{i=1}^{n} x_i^4 &= 0
\end{aligned}$$

que son las llamadas *ecuaciones normales*

$$\begin{pmatrix} n & \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 \\ \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i^3 & \sum_{i=1}^n x_i^4 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n x_i^2 y_i \end{pmatrix}$$

La solución de este sistema proporciona las estimaciones MC de los parámetros del modelo parabólico. Dado que $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} - \hat{\beta}_2(\frac{1}{n}\sum_{i=1}^n x_i^2)$, las predicciones son

$$\begin{aligned} \hat{y}_i &= \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 \\ &= \bar{y} - \hat{\beta}_1\bar{x} - \hat{\beta}_2(\frac{1}{n}\sum_{i=1}^n x_i^2) + \hat{\beta}_1 x_i + \hat{\beta}_2 x_i^2 \\ &= \bar{y} + \hat{\beta}_1(x_i - \bar{x}) + \hat{\beta}_2(x_i^2 - \frac{1}{n}\sum_{i=1}^n x_i^2) \end{aligned}$$

y los residuos

$$e_i = y_i - \hat{y}_i = y_i - \bar{y} - \hat{\beta}_1(x_i - \bar{x}) - \hat{\beta}_2(x_i^2 - \frac{1}{n}\sum_{i=1}^n x_i^2)$$