Método de los mínimos cuadrados

UCO ONLINE

Método de los mínimos cuadrados

UNIVERSIDAD DE CÓRDOBA

1. Descripción del método de los mínimos cuadrados

Tal y como se explicó en la lección anterior, el objetivo de un método de regresión es estimar el intercepto y la pendiente de la recta de regresión que mejor aproxime los valores de la variable independiente. En otras palabras, queremos obtener los coeficientes de aquella recta que pasa lo más cerca posible de todas las observaciones (x, y). Existen varias formas con las que podemos calcular esa "cercanía", por ejemplo, en base a los residuos. El método de los mínimos cuadrados, propuesto por Carl Friedrich Gauss en 1809, se basa en esta idea. Más concretamente, este método obtiene los valores β_0 y β_1 que minimizan la suma de los cuadrados de los residuos, abreviado como RSS (*residual sum of squares*).

En esta lección, se describe cómo este método estima los valores de los coeficientes y cómo debe interpretarse el modelo obtenido. Para finalizar, se presenta un ejemplo de aplicación del método.

2. Estimación de los coeficientes

En la lección anterior, se definió el concepto de residuo como la diferencia entre el valor real y el valor estimado de la variable independiente para cada observación: $\varepsilon_i = y_i - \hat{y}_i$. La suma de los errores cuadráticos se define como:

$$RSS = \varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2 \tag{1}$$

De forma equivalente, podemos expresar RSS como:

$$RSS = (y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 \ldots + (y_n - \hat{y}_n)^2 \tag{2}$$

Si sustituimos $\hat{y}_i$ por la función de regresión lineal definida en la lección anterior, obtenemos:

$$RSS = (y_1 - \hat{\beta}_0 - \hat{\beta}_1 x_1)^2 + (y_2 - \hat{\beta}_0 - \hat{\beta}_1 x_2)^2 \ldots + (y_n - \hat{\beta}_0 - \hat{\beta}_1 x_n)^2 \tag{3}$$

En la ecuación (3) tenemos definido RSS en base a el conjunto de observaciones (x, y), que son valores conocidos, y a los dos coeficientes a estimar (desconocidos). Para determinar los valores $\hat{\beta}_0$ y $\hat{\beta}_1$ que minimizan RSS, igualamos las derivadas parciales a cero:

$$\frac{\partial RSS}{\partial \hat{\beta}_0} = 0 \qquad \frac{\partial RSS}{\partial \hat{\beta}_1} = 0 \tag{4}$$

No abordaremos aquí el desarrollo completo de derivadas parciales. Despejando $\hat{\beta}_0$ y $\hat{\beta}_1$, se obtiene que los valores que minimizan RSS se obtienen como:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{5}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} \tag{6}$$

Donde $\bar{x}$ e $\bar{y}$ representan las medias de los valores x e y en el conjunto de observaciones, también llamado muestra:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{7}$$

3. Interpretación del método de los mínimos cuadrados

A partir de lo estudiado en el apartado anterior, vemos que la estimación de los coeficientes por medio del método de los mínimos cuadrados es dependiente de la muestra de observaciones (x, y). Tanto $\hat{\beta}_0$ como $\hat{\beta}_1$se calculan en base a la media de los valores x e y. Si disponemos de pocos puntos que, además, no sean representativos de la población completa, esto es, todos los posibles pares (x, y), la curva de regresión obtenida no será una buena aproximación a la curva de regresión real, también llamada línea de regresión de la población.

El método de los mínimos cuadrados se basa en la misma idea que utilizamos en estadística para inferir propiedades de una población a partir de una muestra. En estadística, como no disponemos de todas las posibles observaciones de una población, utilizamos mediciones sobre una muestra para estimar los parámetros de la población. Un ejemplo habitual es la media de una muestra. Según la muestra tomada, la media de la población puede estar sobreestimando o infraestimando el valor real de la media de la población. Sin embargo, es sabido que si calculamos la media sobre un número elevado de muestras, y a esas estimaciones le calculamos la media, el valor obtenido es equivalente a la muestra real de la población.

Bajo este mismo supuesto, los coeficientes $\hat{\beta}_0$ y $\hat{\beta}_1$obtenidos por medio del método de los mínimos cuadrados son una estimación de los valores reales (que escribiríamos β_0 y β_1). Por tanto, la recta de regresión obtenida por medio del método de los mínimos cuadrados para un conjunto de observaciones puede diferir de la recta de regresión obtenida para otro conjunto de observaciones y, en general, para la población completa. No obstante, si calculamos $\hat{\beta}_0$ y $\hat{\beta}_1$ sobre un conjunto independiente de muestras y las representamos gráficamente, veremos que "en media" se aproximan bastante bien a la línea de regresión de la población. Como es lógico, disponer de una muestra amplia de observaciones nos dará más garantías de que la estimación que estamos realizando es una buena aproximación a la recta de regresión de la población.

En la siguiente lección, se profundizará en cómo se puede evaluar la calidad de la estimación de los coeficientes de regresión. También se explicarán medidas de evaluación para analizar el error cometido por los modelos de regresión a la hora de predecir la variable independiente ante nuevos valores de x. Antes de ello, finalizaremos esta lección con el desarrollo de un ejemplo completo.

4. Ejemplo

En este apartado se presenta un ejemplo donde se aplican las fórmulas de estimación de los coeficientes $\hat{\beta}_0$ y $\hat{\beta}_1$ según el método de los mínimos cuadrados. También se ilustra gráficamente la diferencia entre la recta de regresión obtenida y la recta de regresión de la población.

Vamos a suponer que tenemos una población donde la relación entre las variables es la siguiente: $y = f(x) = 1 + 2x$. Esta relación no suele ser conocida, pero nos va a permitir comparar la estimación de coeficientes $\hat{\beta}_0$ y $\hat{\beta}_1$ con los valores reales ($\beta_0 = 1$ y $\beta_1 = 2$).

Para aplicar el método de los mínimos cuadrados, partimos de la siguiente muestra:

$$(1,3), (3,7), (4,9), (8,17), (10,21)$$

Todas estas observaciones responden de forma exacta a la relación $y = 1 + 2x$. Veamos si el método de los mínimos cuadrados es capaz de estimar los coeficientes a partir de tan solo estas 5 observaciones.

En primer lugar, calculamos la media de x e y, pues las necesitamos para obtener ambos coeficientes:

$$\bar{x} = \frac{1 + 3 + 4 + 8 + 10}{5} = 5.2$$

$$\bar{y} = \frac{3 + 7 + 9 + 17 + 21}{5} = 11.4$$

A continuación, vamos a obtener $\hat{\beta}_1$ utilizando la ecuación (5):

$$\hat{\beta}_1 = \frac{(1-5.2)(3-11.4) + (3-5.2)(7-11.4) + (4-5.2)(9-11.4) + (8-5.2)(17-11.4) + (10-5.2)(21-11.4)}{(1-5.2)^2 + (3-5.2)^2 + (4-5.2)^2 + (8-5.2)^2 + (10-5.2)^2}$$
$$= \frac{109.6}{54.8} = 2$$

Una vez calculado $\hat{\beta}_1$, podemos obtener $\hat{\beta}_0$ mediante la ecuación (6):

$$\hat{\beta}_0 = 11.4 - 2 * 5.2 = 1$$

Podemos comprobar que, con tan solo cinco pares de valores, hemos obtenido los valores exactos de la recta de regresión de la población.
Vamos a repetir el proceso pero considerando que la muestra de observaciones contiene errores fruto, por ejemplo, de alguna imprecisión en la medición. En este caso, la muestra de partida es:

$$(1.1{,}3.1), (2.8{,}7.3), (4.2{,}8.7), (8.3{,}16.9), (10.2{,}20.8)$$

Repetimos el proceso, calculando primero las medias:

$$\bar{x} = \frac{1.1 + 2.8 + 4.2 + 8.3 + 10.2}{5} = 5.32$$

$$\bar{y} = \frac{3.1 + 7.3 + 8.7 + 16.9 + 2}{5} = 11.36$$

Vemos que los valores son ligeramente diferentes a los obtenidos anteriormente. A partir de las medias y de la muestra, obtenemos las siguientes estimaciones de los coeficientes:

$$\hat{\beta}_1 = \frac{110.644}{58.108} = 1.904$$

$$\hat{\beta}_0 = 11.36 - 1.904 * 5.32 = 1.23$$

Como era de esperar, la estimación obtenida para la nueva muestra no es exacta, pero se acerca bastante a los valores reales de la recta de regresión.
Finalmente, podemos representar gráficamente ambas rectas de regresión para comprobar cómo se ajustan a cada una de las muestras. En el primer caso, el ajuste es perfecto, ya que la recta pasa por todos los puntos (ver Figura 1). Por el contrario, en la segunda muestra existen puntos que no son atravesados por la recta, aunque los residuos que se aprecian son muy pequeños (ver Figura 2).

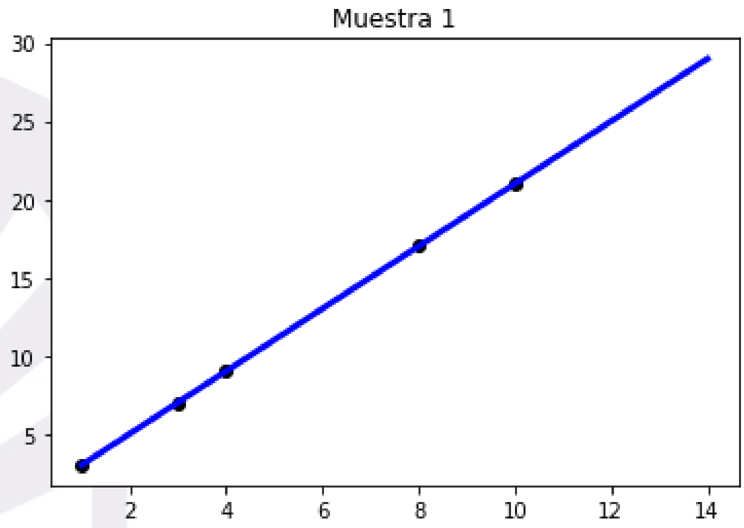

Figura 1. Recta de regresión para la primera muestra de observaciones

Figura 2. Recta de regresión para la segunda muestra de observaciones

Referencias

B. Caffo. “Regression Models for Data Science in R”. Leanpub (CCA-NC 3.0), 129 páginas. 2015. Disponible en: https://github.com/bcaffo/regmodsbook
G. Hackeling. “Mastering Machine Learning with Scikit-Learn”. Pactkt Publishing, 221 páginas. 2014.
T. Hastie, R. Tibshirani, J. Friedman. “The Elements of Statistical Learning: Data Mining, Inference, and Prediction”. Springer Series in Statistics, 2ª edición, 745 páginas. 2017.
G. James, D. Witten, R. Tibshirani, T. Hastie. “An Introduction to Statistical Learning with Applications in R”. Springer Texts in Statistics, 1ª edición (7ª impresión), 426 páginas. 2017. Disponible en: https://www.statlearning.com/
D. Peña. “Regresión y diseño de experimentos”. Alianza Editorial, 744 páginas. 2010.