Revista Colombiana de Estadística
Volumen 24 (2001) N° 1, páginas 45 *a* 57

UNA PROPUESTA PARA LA MAXIMIZACIÓN DE LA ESTADÍSTICA Q_K

JOSÉ A. JIMÉNEZ M.*

Resumen

En este artículo mediante el método de los multiplicadores de Lagrange se presenta una forma de maximizar la Estadística Q_k, y de esta manera cuantificar el impacto que ejerce en la suma de cuadrados de los residuales un grupo de observaciones previamente seleccionadas si son corregidas o modificadas.
Palabras claves Modelos Lineales, Mínimos cuadrados, Observaciones Influyentes, Estadística Q_k, Multiplicadores de Lagrange

1. Introducción

Para el modelo de regresión lineal múltiple

$$\vec{Y} = X\vec{\beta} + \vec{\epsilon} \tag{1}$$

Draper and John (1981) desarrollaron una metodología para detectar un grupo de k observaciones influyentes o atípicas, equivalente a la propuesta por Bartlett (1937), citada en Little and Rubin (1987), para estimar los parámetros del modelo de regresión lineal cuando existen observaciones faltantes en la variable respuesta. En la propuesta de Draper and John (1981) se analiza el modelo (1)

*Profesor asistente, Universidad Nacional de Colombia, Departamento de matemáticas; e-mail: jjimenez@matematicas.unal.edu.co

particionado

$$\begin{bmatrix}\vec{Y_1}\\ \vec{Y_2}\end{bmatrix} = \begin{bmatrix}X_1 & I\\ X_2 & 0\end{bmatrix}\begin{bmatrix}\vec{\beta}\\ \vec{\gamma}\end{bmatrix} + \begin{bmatrix}\vec{\epsilon_1}\\ \vec{\epsilon_2}\end{bmatrix} \tag{2}$$

donde $\vec{Y_1}$ es el bloque conformado por las observaciones consideradas atípicas. Las estimaciones $\vec{\beta}$ y $\vec{\gamma}$ del modelo (2) están definidas por:

$$\hat{\beta} = (X_2'X_2)^{-1}X_2'\vec{Y_2} \qquad \hat{\gamma} = (I - H_{11})^{-1}\hat{\epsilon}_1$$

donde $H_{ij} = X_i(X'X)^{-1}X_j'$, es submatriz de la matriz $H = X(X'X)^{-1}X'$ la cual se conoce como "Matriz Hat", Tukey (1977); $X = \begin{bmatrix}X_1\\ X_2\end{bmatrix}$ y el cambio en la suma de cuadrados de residuales está dado por la estadística

$$Q_k = \hat{\epsilon}_1'(I - H_{11})^{-1}\hat{\epsilon}_1$$

En resumen, el método descrito permite detectar el grupo de observaciones atípicas en base al cambio en la suma de cuadrados de residuales, lo cual se cuantifica con la estadística Q_k. Sin embargo, el método no mide la influencia de estas observaciones en la estimación de los parámetros. En este artículo se muestra la maximización de dicha estadística sujeta a algunas condiciones.

2. Nueva expresión de la Estadística $\mathbf{Q_k}$

Para el modelo de regresión lineal múltiple definido en (1) el método de estimación mínimos cuadrados proporciona el estimador $\hat{\beta}$ de los parámetros, los valores estimados $\hat{Y}$, los errores estimados $\hat{\epsilon}$ y la suma de cuadrados de los residuales SCE según las siguientes expresiones

$$\begin{aligned}\hat{\beta} &= (X'X)^{-1}X'\vec{Y}\\ \hat{Y} &= X\hat{\beta} = X(X'X)^{-1}X'\vec{Y} = H\vec{Y}\\ \hat{\epsilon} &= \vec{Y} - \hat{Y} = \vec{Y} - H\vec{Y} = (I - H)\vec{Y}\\ SCE &= \hat{\epsilon}'\hat{\epsilon} = [(I - H)\vec{Y}]'(I - H)\vec{Y} = \vec{Y}'(I - H)\vec{Y}\end{aligned}$$

En Jiménez (1999) se plantea el modelo

$$\vec{Y}^* = X\vec{\beta}^* + \vec{\epsilon}^* \tag{3}$$

siendo $\vec{Y}^* = \vec{Y} + \vec{\gamma}$ con $\vec{\gamma} \in \mathbb{R}^n$ un vector arbitrario no nulo; bajo estos criterios se establece por el método de mínimos cuadrados las expresiones de los estimadores del modelo (3), en función de $\vec{\gamma}$, $\vec{Y}$ y de los estimadores descritos anteriormente. Así, los nuevos estimadores están dados por

$$\begin{aligned}\hat{\beta}^* =& \hat{\beta} + (X'X)^{-1}X'\vec{\gamma}\\ \hat{Y}^* =& \hat{Y} + H\vec{\gamma}\\ \hat{\epsilon}^* =& \hat{\epsilon} + (I-H)\vec{\gamma}\\ SCE^* =& SCE + 2\vec{\gamma}'\hat{\epsilon} + \vec{\gamma}'(I-H)\vec{\gamma}\end{aligned}$$

de la última ecuación se obtiene que

$$Q_k = SCE - SCE^* = -2\vec{\gamma}'\hat{\epsilon} - \vec{\gamma}'(I-H)\vec{\gamma} \tag{4}$$

esta nueva expresión de la estadística Q_k tiene la ventaja de estar en términos de los residuales estimados del modelo (1) y del $\vec{\gamma} \neq \vec{0}$ arbitrario. Para los objetivos de este trabajo es más atractiva, ya que permite maximizarla sujeta a las restricciones que se presentaran a continuación

3. Maximización de la estadística $\mathbf{Q_k}$

Para la maximización de esta estadística se particiona el modelo (3) como:

$$\begin{bmatrix}\vec{Y}_1\\ \vec{Y}_2\end{bmatrix} + \begin{bmatrix}\vec{\gamma}_1\\ \vec{\gamma}_2\end{bmatrix} = \begin{bmatrix}X_1\\ X_2\end{bmatrix}\vec{\beta}^* + \begin{bmatrix}\vec{\epsilon}_1^*\\ \vec{\epsilon}_2^*\end{bmatrix}$$

y de manera análoga a la propuesta de Draper and Jhon (1981) se asume que el bloque $\vec{Y}_1$ de dimensión $k \times 1$, $k < n$, esta conformado por las observaciones atípicas; el interés en este artículo es presentar valores óptimos para $\vec{\gamma}$, que permitan corregir dichas observaciones de tal manera que la nueva suma de cuadrados del error sea mínima. Para ello se consideran las siguientes restricciones:

3.1. Restricción sobre el vector $\vec{\gamma}$

Si se toma como condición que $\vec{\gamma} = \begin{bmatrix}\vec{\gamma}_1\\ \vec{0}\end{bmatrix}$, entonces para que Q_k dada en (4) tenga un extremo en $\hat{\gamma}_1$, es necesario que

$$\frac{\partial Q_k}{\partial \vec{\gamma}} = \vec{0}$$

al efectuar dicha derivada se tiene que

$$\frac{\partial Q_k}{\partial \vec{\gamma}} = -2\hat{\epsilon} - 2(I-H)\hat{\gamma} = \vec{0}$$

si se despeja $\hat{\gamma}$ queda

$$\begin{bmatrix} I - H_{11} & -H_{12} \\ -H_{21} & I - H_{22} \end{bmatrix} \begin{bmatrix} \hat{\gamma}_1 \\ \vec{0} \end{bmatrix} = - \begin{bmatrix} \hat{\epsilon}_1 \\ \hat{\epsilon}_2 \end{bmatrix}$$

$$\begin{bmatrix} I - H_{11} & -H_{12} \\ -H_{21} & I - H_{22} \end{bmatrix} \begin{bmatrix} \hat{\gamma}_1 \\ \vec{0} \end{bmatrix} = - \begin{bmatrix} I - H_{11} & -H_{12} \\ -H_{21} & I - H_{22} \end{bmatrix} \begin{bmatrix} \vec{Y}_1 \\ \vec{Y}_2 \end{bmatrix}$$

como el rango de la matriz $(I-H)$ es $n-r$ donde r es el número de parámetros del modelo, entonces esta matriz no tiene inversa; por lo tanto para encontrar las componentes de $\hat{\gamma}_1$ que maximiza a Q_k se resuelve el sistema que aparece a continuación, el cual se obtiene de efectuar los respectivos productos

$$(I - H_{11})\hat{\gamma}_1 = - (I - H_{11})\vec{Y}_1 + H_{12}\vec{Y}_2 \tag{5}$$

$$-H_{21}\hat{\gamma}_1 = H_{21}\vec{Y}_1 - (I - H_{22})\vec{Y}_2 \tag{6}$$

si se premultiplica (6) por H_{12} y se suma con (5) se obtiene

$$\begin{aligned} \left[(I - H_{11}) - H_{12}H_{21}\right] \hat{\gamma}_1 &= - \left[(I - H_{11}) - H_{12}H_{21}\right] \vec{Y}_1 + H_{12}H_{22}\vec{Y}_2 \\ \hat{\gamma}_1 &= - \vec{Y}_1 + \left[(I - H_{11}) - H_{12}H_{21}\right]^{-1} H_{12}H_{22}\vec{Y}_2 \end{aligned} \tag{7}$$

pero como la matriz $(I - H)$ es idempotente, se puede probar fácilmente que

$$H_{12}H_{21} = (I - H_{11})H_{11} \tag{8}$$

$$H_{12}H_{22} = (I - H_{11})H_{12} \tag{9}$$

además, como $HX = X$, se puede verificar de manera trivial que

$$H_{21}X_1 = (I - H_{22})X_2 \tag{10}$$

$$H_{12}X_2 = (I - H_{11})X_1 \tag{11}$$

reemplazando las expresiones (8) y (9) en la ecuación (7), se tiene que

$$\hat{\gamma}_1 = -\vec{Y}_1 + [I - H_{11}]^{-1}H_{12}\vec{Y}_2$$

utilizando el hecho de que

$$(I - H_{11})^{-1} = I + X_1(X_2^{'}X_2)^{-1}X_1^{'}$$

se obtiene que

$$\begin{aligned} \hat{\gamma}_1 &= - \vec{Y}_1 + [I + X_1(X_2^{'}X_2)^{-1}X_1^{'}]H_{12}\vec{Y}_2 \\ &= - \vec{Y}_1 + H_{12}\vec{Y}_2 + X_1(X_2^{'}X_2)^{-1}[H_{21}X_1]^{'}\vec{Y}_2 \end{aligned} \tag{12}$$

en la última ecuación se utilizó el hecho de que $H_{ij}^{'} = H_{ji}$, luego, reemplazando la expresión (10) en (12) se tiene:

$$\begin{aligned}\hat{\gamma}_1 =& -\vec{Y}_1 + H_{12}\vec{Y}_2 + X_1(X_2^{'}X_2)^{-1}X_2^{'}(I - H_{22})\vec{Y}_2 \\ =& -\vec{Y}_1 + H_{12}\vec{Y}_2 + X_1(X_2^{'}X_2)^{-1}X_2^{'}\vec{Y}_2 - X_1(X_2^{'}X_2)^{-1}X_2^{'}H_{22}\vec{Y}_2\end{aligned}$$

por otra parte, como la matriz $H_{22} = X_2(X^{'}X)^{-1}X_2$, se obtiene que

$$\begin{aligned}\hat{\gamma}_1 =& -\vec{Y}_1 + H_{12}\vec{Y}_2 + X_1(X_2^{'}X_2)^{-1}X_2^{'}\vec{Y}_2 - H_{12}\vec{Y}_2 \\ =& -\vec{Y}_1 + X_1(X_2^{'}X_2)^{-1}X_2^{'}\vec{Y}_2 \qquad (13)\end{aligned}$$

Nótese que esta última expresión no depende de la estimación del vector de parámetros $\hat{\beta}$, ni de la estimación del vector de errores $\hat{\epsilon}$.

3.2. Restricción sobre el cambio en el vector de parámetros $\vec{\beta}$

Si se asume como requisito que los parámetros del modelo (1) se modifican en un vector fijo de constantes al sumársele al bloque $\vec{Y}_1$ el vector $\vec{\gamma}_1$, entonces se debe maximizar (4) bajo la siguiente restricción

$$X[\hat{\beta} - \hat{\beta}^*] + X(X'X)^{-1}X'\vec{\gamma} = \vec{0}$$

Sea $\vec{b}$ un vector de constantes que indica la diferencia entre $\hat{\beta}$ y $\hat{\beta}^*$; entonces la condición dada anteriormente se puede expresar como:

$$X\vec{b} + H\vec{\gamma} = \vec{0}$$

si se denota por $\vec{\lambda}$ el vector de Lagrange, la función a maximizar es:

$$\begin{aligned}z =& Q_k - \vec{\lambda}'(X\vec{b} + H\vec{\gamma}) \\ =& -2\vec{\gamma}'\hat{\epsilon} - \vec{\gamma}'(I - H)\vec{\gamma} - \vec{\lambda}'(X\vec{b} + H\vec{\gamma})\end{aligned}$$

luego, para que Q_k tenga un extremo en $\hat{\gamma}_1$, es necesario que

$$\frac{\partial z}{\partial \vec{\gamma}} = \vec{0} \qquad \text{y} \qquad \frac{\partial z}{\partial \vec{\lambda}} = \vec{0}$$

por consiguiente si se hacen las respectivas derivadas y se igualan a cero se tiene:

$$\frac{\partial z}{\partial \vec{\gamma}} = -2\hat{\epsilon} - 2(I - H)\hat{\gamma} - H\vec{\lambda} = \vec{0} \qquad (14)$$

$$\frac{\partial z}{\partial \vec{\lambda}} = -X\vec{b} - H\hat{\gamma} = \vec{0} \qquad (15)$$

pero como $\hat{\epsilon} = (I - H)\vec{Y}$, si se reemplaza en (14) se obtiene que

$$-2\begin{bmatrix} I - H_{11} & -H_{12} \\ -H_{21} & I - H_{22} \end{bmatrix}\begin{bmatrix} \vec{Y}_1 \\ \vec{Y}_2 \end{bmatrix} - 2\begin{bmatrix} I - H_{11} & -H_{12} \\ -H_{21} & I - H_{22} \end{bmatrix}\begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}\begin{bmatrix} \vec{\lambda}_1 \\ \vec{\lambda}_2 \end{bmatrix}$$

dividiendo esta expresión por 2 y efectuando los productos correspondientes se tiene:

$$(I - H_{11})\hat{\gamma}_1 - H_{12}\hat{\gamma}_2 = -(I - H_{11})\vec{Y}_1 + H_{12}\vec{Y}_2 - \frac{1}{2}H_{11}\vec{\lambda}_1 - \frac{1}{2}H_{12}\vec{\lambda}_2 \quad (16)$$

$$(I - H_{22})\hat{\gamma}_2 - H_{21}\hat{\gamma}_1 = -(I - H_{22})\vec{Y}_2 + H_{21}\vec{Y}_1 - \frac{1}{2}H_{21}\vec{\lambda}_1 - \frac{1}{2}H_{22}\vec{\lambda}_2 \quad (17)$$

si se premultiplica (17) por H_{12} y se suma con (16) se obtiene

$$\begin{aligned}[(I - H_{11}) - H_{12}H_{21}]\,\hat{\gamma}_1 - H_{12}H_{22}\hat{\gamma}_2 = &-(I - H_{11})\vec{Y}_1 + H_{12}H_{21}\vec{Y}_1 + H_{12}H_{22}\vec{Y}_2 \\ &-\frac{1}{2}\left[H_{11} + H_{12}H_{21}\right]\vec{\lambda}_1 - \frac{1}{2}\left[H_{12} + H_{12}H_{22}\right]\vec{\lambda}_2\end{aligned}$$

reemplazando las expresiones dadas en (8) y (9) en esta última ecuación, queda

$$\begin{aligned}(I - H_{11})\left[(I - H_{11})\hat{\gamma}_1 - H_{12}\hat{\gamma}_2\right] = &-(I - H_{11})(I - H_{11})\vec{Y}_1 + (I - H_{11})H_{12}\vec{Y}_2 \\ &-H_{11}\vec{\lambda}_1 - H_{12}\vec{\lambda}_2 + \frac{1}{2}H_{11}H_{11}\vec{\lambda}_1 + \frac{1}{2}H_{11}H_{12}\vec{\lambda}_2\end{aligned}$$

si se sustituye (16) en esta ecuación y se simplifican términos, se llega a

$$H_{11}\vec{\lambda}_1 = -H_{12}\vec{\lambda}_2 \quad (18)$$

por otra parte, de (15) se tiene que:

$$-\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\vec{b} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}\begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{bmatrix}$$

si se efectúan los productos indicados, da como resultado

$$\begin{aligned}-X_1\vec{b} &= H_{11}\hat{\gamma}_1 + H_{12}\hat{\gamma}_2 \\ -X_2\vec{b} &= H_{21}\hat{\gamma}_1 + H_{22}\hat{\gamma}_2\end{aligned} \quad (19)$$

reescribiendo (16) se obtiene que

$$\begin{aligned}(I - H_{11})\hat{\gamma}_1 - H_{12}\hat{\gamma}_2 &= -(I - H_{11})\vec{Y}_1 + H_{12}\vec{Y}_2 - \frac{1}{2}\left[H_{11}\vec{\lambda}_1 + H_{12}\vec{\lambda}_2\right] \\ \hat{\gamma}_1 - [H_{11}\hat{\gamma}_1 + H_{12}\hat{\gamma}_2] &= -(I - H_{11})\vec{Y}_1 + H_{12}\vec{Y}_2 - \frac{1}{2}\left[H_{11}\vec{\lambda}_1 + H_{12}\vec{\lambda}_2\right]\end{aligned}$$

reemplazando (18) y (19) en la última ecuación, se llega a

$$\hat{\gamma}_1 = -X_1\vec{b} - \left[(I - H_{11})\vec{Y}_1 - H_{12}\vec{Y}_2\right] \tag{20}$$

obsérvese que la última expresión no queda ligada a la estimación del vector de parámetros $\hat{\beta}$; pero si depende del cambio en los parámetros del modelo completo y de la estimación del vector de errores $\vec{\epsilon}_1$.

3.3. Restricción sobre el cambio en el vector de errores $\vec{\epsilon}$

Si se toma como condición que los residuales del modelo (1) cambian en un vector fijo de constantes, cuando se le suma al bloque $\vec{Y}_1$ el vector $\vec{\gamma}_1$, entonces (4) se debe maximizar bajo la siguiente restricción

$$[\hat{\epsilon} - \hat{\epsilon}^*] + (I - H)\vec{\gamma} = \vec{0}$$

sea $\vec{u}$ el vector de constantes que indica la diferencia entre $\hat{\epsilon}$ y $\hat{\epsilon}^*$; entonces la condición dada anteriormente se puede expresar como:

$$\vec{u} + (I - H)\vec{\gamma} = \vec{0}$$

si se considera $\vec{\lambda}$ como el vector de Lagrange, la función a maximizar es:

$$\begin{aligned} z =& Q_k - \vec{\lambda}'(\vec{u} + (I - H)\vec{\gamma}) \\ =& -2\vec{\gamma}'\hat{\epsilon} - \vec{\gamma}'(I - H)\vec{\gamma} - \vec{\lambda}'(\vec{u} + (I - H)\vec{\gamma}) \end{aligned}$$

luego, para que Q_k tenga un extremo en $\hat{\gamma}_1$, es necesario que

$$\frac{\partial z}{\partial \vec{\gamma}} = \vec{0} \qquad \text{y} \qquad \frac{\partial z}{\partial \vec{\lambda}} = \vec{0}$$

por consiguiente, realizando las respectivas derivadas e igualando a cero, se tiene:

$$\frac{\partial z}{\partial \vec{\gamma}} = -2\hat{\epsilon} - 2(I - H)\hat{\gamma} - (I - H)\hat{\lambda} = \vec{0} \tag{21}$$

$$\frac{\partial z}{\partial \vec{\lambda}} = -\vec{u} - (I - H)\vec{\gamma} = \vec{0} \tag{22}$$

si se multiplica (22) por (-2) y se suma a (21) se obtiene que

$$\begin{aligned} \frac{\partial z}{\partial \vec{\gamma}} - 2\frac{\partial z}{\partial \vec{\lambda}} =& 2\vec{u} - 2\hat{\epsilon} - (I - H)\hat{\lambda} = \vec{0} \\ (I - H)\hat{\lambda} =& 2\vec{u} - 2\hat{\epsilon} \end{aligned}$$

si se reemplaza este resultado en (21) se llega a:

$$\begin{aligned} -2\hat{\epsilon} - 2(I-H)\hat{\gamma} - [2\vec{u} - 2\hat{\epsilon}] &= \vec{0} \\ (I-H)\hat{\gamma} &= -\,\vec{u} \end{aligned} \tag{23}$$

como se necesita que únicamente cambien los primeros k residuales $\vec{u} = \begin{bmatrix} \vec{u}_1 \\ \vec{0} \end{bmatrix}$, es decir en los $n-k$ residuales siguientes no se desea ningún cambio, luego se debe hallar el $\hat{\gamma}_1$ que permita esto; si se reescribe (23) queda

$$\begin{bmatrix} I-H_{11} & -H_{12} \\ -H_{21} & I-H_{22} \end{bmatrix} \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{bmatrix} = -\begin{bmatrix} \vec{u}_1 \\ \vec{0} \end{bmatrix}$$

como la matriz $(I-H)$ no es de rango completo, entonces no tiene inversa, por lo tanto para obtener los valores de $\hat{\gamma}_1$ que maximizan a Q_k se resuelve el siguiente sistema que se obtiene de efectuar los correspondientes productos

$$(I-H_{11})\hat{\gamma}_1 - H_{12}\hat{\gamma}_2 = -\,\vec{u}_1 \tag{24}$$

$$-H_{21}\hat{\gamma}_1 + (I-H_{22})\hat{\gamma}_2 = \vec{0} \tag{25}$$

si se asume que el rango de la matriz $(I-H_{22})$ es menor o igual que $n-r$; se puede despejar $\hat{\gamma}_2$ de (25) y al reemplazarlo en (24) se tiene que:

$$\begin{aligned} (I-H_{11})\hat{\gamma}_1 - H_{12}\left[(I-H_{22})^{-1}H_{21}\hat{\gamma}_1\right] &= -\vec{u}_1 \\ \hat{\gamma}_1 &= -\left[(I-H_{11}) - H_{12}(I-H_{22})^{-1}H_{21}\right]^{-1}\vec{u}_1 \end{aligned} \tag{26}$$

en este caso, la inversa de $(I-H_{22})$ está dada por

$$(I-H_{22})^{-1} = I + X_2(X_1^{'}X_1)^{-1}X_2^{'}$$

la cual al reemplazarse en (26) se obtiene que

$$\begin{aligned} \hat{\gamma}_1 &= -\left[(I-H_{11}) - H_{12}[I + X_2(X_1^{'}X_1)^{-1}X_2^{'}]H_{21}\right]^{-1}\vec{u}_1 \\ &= -\left[(I-H_{11}) - H_{12}H_{21} - H_{12}X_2(X_1^{'}X_1)^{-1}X_2^{'}H_{21}\right]^{-1}\vec{u}_1 \end{aligned}$$

si se sustituye en esta última fórmula las expresiones dadas en (8) y (11), se establece que

$$\begin{aligned} \hat{\gamma}_1 &= -\left[(I-H_{11})^2 - (I-H_{11})X_1(X_1^{'}X_1)^{-1}X_1^{'}(I-H_{11})\right]^{-1}\vec{u}_1 \\ &= -\left[(I-H_{11})\left(I - X_1(X_1^{'}X_1)^{-1}X_1^{'}\right)(I-H_{11})\right]^{-1}\vec{u}_1 \end{aligned}$$

utilizando el hecho de que

$$(I - H_{11})\left(I - X_1(X_1^{'}X_1)^{-1}X_1^{'}\right) = I - X_1(X_1^{'}X_1)^{-1}X_1^{'}$$

se obtiene finalmente que

$$\hat{\gamma}_1 = -\left[I - X_1(X_1^{'}X_1)^{-1}X_1^{'}\right]^{-1}\vec{u}_1 \tag{27}$$

como era de esperarse en la última expresión no interviene $\vec{\gamma}_2$, es decir que en el cambio de la suma de cuadrados únicamente interesa la modificación que se le haga a los primeros k residuales.

Por otra parte, cuando la matriz $(I - H_{22})$ no sea de rango completo, la ecuación (27) queda expresada de la siguiente manera:

$$\begin{aligned}\hat{\gamma}_1 &= -\left[I - H_{11}\right]^{-1}\vec{u}_1\\ &= -\left[I + X_1(X_2^{'}X_2)^{-1}X_1^{'}\right]\vec{u}_1\end{aligned}$$

3.4. Restricciones sobre el cambio en el vector de parámetros $\vec{\beta}$ y el vector de errores $\vec{\epsilon}$

Si se consideran las dos últimas condiciones simultáneamente, se debe maximizar la estadística Q_k dada en (4) sujeta a:

$$\begin{aligned}X[\hat{\beta} - \hat{\beta}^*] + X(X'X)^{-1}X'\vec{\gamma} &= 0\\ [\hat{\epsilon} - \hat{\epsilon}^*] + (I - H)\vec{\gamma} &= 0\end{aligned}$$

denotando como $\vec{b}$ el vector de constantes que indica la diferencia entre $\hat{\beta}$ y $\hat{\beta}^*$; y $\vec{u}$ el vector de constantes que representa la diferencia entre $\hat{\epsilon}$ y $\hat{\epsilon}^*$, entonces las condiciones anteriores se pueden expresar como:

$$\begin{aligned}X\vec{b} + H\vec{\gamma} &= 0\\ \vec{u} + (I - H)\vec{\gamma} &= 0\end{aligned}$$

sean $\vec{\mu}$ y $\vec{\lambda}$ los vectores de Lagrange, luego la función a maximizar es:

$$\begin{aligned}z =& Q_k - \vec{\mu}'\left(X\vec{b} + H\vec{\gamma}\right) - \vec{\lambda}'\left[\vec{u} + (I - H)\vec{\gamma}\right]\\ =& -2\vec{\gamma}'\hat{\epsilon} - \vec{\gamma}'(I - H)\vec{\gamma} - \vec{\mu}'\left(X\vec{b} + H\vec{\gamma}\right) - \vec{\lambda}'\left[\vec{u} + (I - H)\vec{\gamma}\right]\end{aligned}$$

para que Q_k tenga un extremo en $\hat{\gamma}_1$, es necesario que

$$\frac{\partial z}{\partial \vec{\gamma}} = \vec{0}, \qquad \frac{\partial z}{\partial \vec{\mu}} = \vec{0} \qquad \text{y} \qquad \frac{\partial z}{\partial \vec{\lambda}} = \vec{0}$$

efectuando las respectivas derivadas e igualando a cero se obtiene:

$$\frac{\partial z}{\partial \vec{\gamma}} = -2\hat{\epsilon} - 2(I-H)\hat{\gamma} - H\vec{\mu} - (I-H)\vec{\lambda} = \vec{0} \tag{28}$$

$$\frac{\partial z}{\partial \vec{\mu}} = -X\vec{b} - H\hat{\gamma} = \vec{0} \tag{29}$$

$$\frac{\partial z}{\partial \vec{\lambda}} = -\vec{u} - (I-H)\hat{\gamma} = \vec{0} \tag{30}$$

si se suman las ecuaciones (29) y (30) se tiene que

$$\frac{\partial z}{\partial \vec{\mu}} + \frac{\partial z}{\partial \vec{\lambda}} = -X\vec{b} - \vec{u} - \hat{\gamma} = \vec{0}$$

despejando de aquí $\vec{\gamma}$ se tiene que el $\hat{\gamma}_1$ esta dado por:

$$\hat{\gamma}_1 = -X_1\vec{b} - \vec{u}_1$$

como era de esperarse, en esta última expresión no intervienen las estimaciones de los parámetros ni la de los residuales.

4. Ejemplo

Para el conjunto de 21 observaciones (x, y) dado por Mickey, Dunn, and Clark (1967), tabla 1, se presentan los siguientes resultados.

1. La estimación del modelo de regresión lineal, con las 21 observaciones
2. Los términos h_{ii}, las estimaciones de los γ_i y el cálculo de la estadística $\mathbf{Q_1}$ al eliminar el i-ésimo dato.
3. La estimación del modelo de regresión lineal, después de modificar con el $\widehat{\gamma}_i$ la observación que tiene el $\mathbf{Q_1}$ más grande.
4. La estimación del modelo de regresión lineal, después de eliminar la observación con el $\mathbf{Q_1}$ más grande.

Tabla 1. Datos de Mickey, Dunn, and Clark (1967)

Obs	x	y	**Obs**	x	y	**Obs**	x	y
1	15	95	8	11	100	15	11	102
2	26	71	9	8	104	16	10	100
3	10	83	10	20	94	17	12	105
4	9	91	11	7	113	18	42	57
5	15	102	12	9	96	19	17	121
6	20	87	13	10	83	20	11	86
7	18	93	14	11	84	21	10	100

1.

Fuente de variación	Grados libertad	Suma de cuadrados	Cuadrados Medios	F	Valor crítico de F
Regresión	1	1604.0809	1604.0809	13.2018	0.00177
Residuos	19	2308.5858	121.5045		
Total	20	3912.6667			

	Coeficientes	Error típico	Estadístico t	Probabilidad
Intercepto	109.8738	5.0678	21.6808	7,30934E-15
Variable X	-1.1270	0.3102	-3.6334	0,00177

Coeficiente de determinación $R^2 = 0,409971261$
Error típico $\widehat{\sigma} = 11{,}0229086$
2. Para el valor de $\widehat{\gamma}_i$ dado en (13), se tiene que

Obs. Mod.	h_{ii}	$\widehat{\gamma}_i$	Q_k $(k=1)$	**Obs. Mod.**	h_{ii}	$\widehat{\gamma}_i$	Q_k $(k=1)$
1	0.0479	-2.1332	4,333	12	0.0705	4.0141	14,976
2	0.1545	11.3214	108,370	13	0.0628	16.6498	259,803
3	0.0628	16.6498	259,803	14	0.0567	14.2866	192,540
4	0.0705	9.3936	82,015	15	0.0567	-4.7948	21,687
5	0.0479	-9.4856	85,664	16	0.0628	-1.4896	2,080
6	0.0726	0.3602	0,120	17	0.0521	-9.1255	78,936
7	0.0580	-3.6220	12,358	18	0.6516	15.9026	88,105
8	0.0567	-2.6746	6,748	19	0.0531	-31.9816	968,562
9	0.0799	-3.4148	10,729	20	0.0567	12.1664	139,634
10	0.0726	-7.1879	47,914	21	0.0628	-1.4896	2,080
11	0.0908	-12.1145	133,443				

En los resultados anteriores se verifica que los valores de la estadística Q_1, corresponden a la expresión

$$Q_1 = (1 - h_{ii})\widehat{\gamma}_i^2$$

3. Cuando se modifica la observación 19 con el correspondiente $\widehat{\gamma}_i$, se obtiene

Fuentes de variación	Grados libertad	Suma de cuadrados	Cuadrados Medios	F	Valor crítico de F
Regresión	1	1798,4323	1798,4323	25,4997	7,11533E-05
Residuos	19	1340.0238	70,5276		
Total	20	3138,4561			

	Coeficientes	Error típico	Estadístico t	Probabilidad
Intercepto	109.30468	3,8610	28,3097	5,32302E-17
Variable X	-1.19331	0,2363	-5,0497	7,11533E-05

Coeficiente de determinación $R^2 = 0,573030892$
Error típico $\widehat{\sigma} = 8,39806936$

El modelo que se obtiene al modificar la pareja $(17, 121)$ por (17; 89,0184) es mejor que el modelo original; pues el nuevo coeficiente de determinación es superior al del modelo inicial, el valor crítico de la F es también inferior al valor crítico que se determinó en el análisis de varianza del modelo inicial y además se logra minimizar la SCE lo cual implica que el nuevo cuadrado medio del error (CME) sea menor que el CME del modelo original.

4. Eliminando la observación 19 que tiene el $\mathbf{Q_1}$ más alto, se tiene

Fuentes de variación	Grados libertad	Suma de cuadrados	Cuadrados Medios	F	Valor crítico de F
Regresión	1	1788.17619	1788.17619	24.01985	0.0001151
Residuos	18	1340.02381	74.44577		
Total	19	3128.2			

	Coeficientes	Error típico	Estadístico t
Intercepto	109.30468	3.96996	27.5329
Variable X	-1.19331	0.24348	-4.9010

Coeficiente de determinación $R^2 = 0,57163103$
Error típico $\widehat{\sigma} = 8{,}628196$
Cambio en la suma de los residuales $Q_k = 968,5619674$

El modelo que se obtiene al eliminar la pareja $(17, 121)$ es mejor que el modelo completo; pues el nuevo coeficiente de determinación es superior al del modelo inicial, el valor crítico de la F es también inferior al valor crítico que se determinó en el análisis de varianza del modelo inicial y además el cuadrado medio del error (CME) fue menor que el CME del modelo completo.

5. Conclusiones

Si el propósito es minimizar la suma de cuadrados de los residuales de un modelo de regresión cuando se tiene un bloque $\vec{Y}_1$ de observaciones influyentes, la recomendación que se hace en este artículo es la de corregirlo con alguno de los vectores $\vec{\gamma}_1$ propuestos, de tal manera que no sea necesario eliminar dicho bloque y por consiguiente no se pierdan grados de libertad.

Referencias

[1] Bartlett, M.S. *Some examples of statistical methods of research in agriculture and applied botany.* J. Roy. Statist. Soc. Vol. B4, 1937, pag. 137-170

[2] Belsley, D. et al. *Regression diagnostics.* John Wiley, New York, 1980

[3] Draper, N.R. and John J.A. *Influential observations and outliers in Regression.* Technometrics, vol. 23, 1981, pag. 21

[4] Jiménez, J.A. *Propuesta Metodológica para Imputar Valores no Influyentes en Modelos de Regresión Lineal Múltiple con Información Incompleta.* Universidad Nacional de Colombia, Tesis de Maestría, 1999

[5] Little R.J., and Rubin D.B. *Statistical Analysis with Missing Data.* John Wiley, vol. 25, 1987, pag. 119

[6] Mickey, M. R., Dunn, O. J., and Clark, V. (1967) *Note on the use of stepwise regression in detecting outliers.* Computers and Biomedical Research, 1, pag 105-111

[7] Searle, S. R. *Linear Models.* John Wiley & Sons, New York, 1971

[8] Tukey, J. *Exploratory Data Analysis.* Addison-Wesley Publishing Company, United States of America, 1977