REVISTA DE MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA (24). Páginas 169–189.
Diciembre de 2017. ISSN: 1886-516X. D.L: SE-2927-06.
www.upo.es/revistas/index.php/RevMetCuant/article/view/2886

Métodos cuantitativos para un modelo de regresión lineal con multicolinealidad. Aplicación a rendimientos de letras del tesoro

Salmerón Gómez, Román
Departamento de Métodos Cuantitativos para la Economía y la Empresa
Universidad de Granada (España)
Correo electrónico: romansg@ugr.es

Rodríguez Martínez, Eduardo
Máster en Técnicas Cuantitativas en Gestión Empresarial
Universidad de Granada (España)
Correo electrónico: edurodr@correo.ugr.es

RESUMEN

Es conocido que, cuando en el modelo de regresión lineal existe un alto grado de multicolinealidad, los resultados obtenidos a partir del método de mínimos cuadrados ordinarios (MCO) son inestables. Como solución a esta situación, en este trabajo se presentan los métodos de alzado, cresta y variables ortogonales como alternativa a la estimación por MCO. También se muestra que la regresión con variables ortogonales tiene sentido independientemente de la existencia de multicolinealidad grave, ya que permite dar respuesta a cuestiones no accesibles con el modelo original. Dichas metodologías se aplican a un conjunto de datos de rendimientos de letras del tesoro.

Palabras claves: modelos de regresión; multicolinealidad; regresión alzada; regresión cresta; regresión con variables ortogonales.
Clasificación JEL: C13; C21.
MSC2010: 62J05; 62J07; 62P25.

Artículo recibido el 14 de septiembre de 2016 y aceptado el 1 de marzo de 2017.

Quantitative Methods for a Linear Regression Model with Multicollinearity. Application to Yields of Treasury Bills

ABSTRACT

It is known that, when in the linear regression model there is a high degree of multicollinearity, the results obtained by using the Ordinary Least Squares (OLS) method are unstable. As a solution to this situation, in this paper we present the raised method, the ridge method and the orthogonal variables method as an alternative to the estimate by OLS. It is also shown that regression with orthogonal variables makes sense regardless of the existence of serious multicollinearity because it allows to answer questions which are not accessible when using the original model. These methodologies are applied to a data set of yields of treasury bills.

Keywords: regression models; multicollinearity; raised regression; ridge regression; regression with orthogonal variables.
JEL classification: C13; C21.
MSC2010: 62J05; 62J07; 62P25.

1. Introducción

La regresión lineal es una herramienta estadística ampliamente usada para analizar cómo influyen (si es que lo hacen) un conjunto de variables (independientes o explicativas) en otra (dependiente o explicada), permitiendo la estimación numérica de los signos y magnitudes de los coeficientes en una relación lineal previamente establecida. Una de las características de las variables económicas es la posible correlación entre ellas debido a la existencia de determinantes comunes, por lo que la multicolinealidad entre variables explicativas en una regresión múltiple debe considerase una situación habitual.

Multicolinealidad. Según Novales [7], "la presencia de multicolinealidad en un modelo lineal puede revestir dos formas: la que se conoce como multicolinealidad exacta, que ocurre cuando una de las variables explicativas es combinación lineal determinista de todas las demás (o de algunas de ellas), y la multicolinealidad aproximada, que ocurre cuando una de las variables es aproximadamente igual a una combinación lineal de las restantes"; es decir, dado el siguiente modelo de regresión lineal con n observaciones y p variables:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \tag{1}$$

donde $\mathbf{u}$ representa la perturbación (que se presupone esférica) y se asume la independencia lineal entre las variables independientes presentes en $\mathbf{X}$; cuando esta condición no se verifica se dice que en el modelo hay multicolinealidad. Por tanto, la multicolinealidad describe la situación de ausencia de ortogonalidad entre las variables independientes del modelo de regresión.

En el primer caso (multicolinealidad perfecta), el modelo no satisface la condición de rango completo y conduce a infinitas estimaciones de los coeficientes del modelo de regresión, mientras que en el segundo caso (multicolinealidad aproximada[1]), aunque dicha condición es satisfecha, la estimación será inestable ya que se pueden presentar los siguientes problemas:

- Coeficientes estimados sensibles a pequeños cambios en los datos;
- Varianzas elevadas de los estimadores;
- Tendencia a no rechazar la hipótesis nula al efectuar los contrastes de significación individual; y
- Un coeficiente de determinación elevado y, por tanto, tendencia a rechazar la hipótesis nula en el constraste de significación conjunta.

El segundo caso es el problemático ya que pone en entredicho las conclusiones obtenidas en el análisis realizado. Por tanto, la existencia de un alto grado de multicolinealidad no es una cuestión baladí cuando se estima el modelo de regresión por el método de mínimos cuadrados ordinarios (MCO); de ahí la importancia de detectarla y tratarla de forma adecuada.

[1]De aquí en adelante, cuando se haga referencia al término de multicolinealidad se ha de sobreentender que se hace referencia a la multicolinealidad aproximada.

Causas. Las causas que producen multicolinealidad en un modelo son diversas; sin embargo, podrían dividirse en los dos siguientes bloques (ver Spanos y McGuirk [11]):

Multicolinealidad sistemática: debida a un problema estructural; es decir, a la alta correlación lineal de las variables explicativas consideradas.

Multicolinelidad errática: debida a un problema puramente numérico; es decir, a un mal condicionamiento de los datos considerados por escasa variabilidad de las observaciones y/o reducido tamaño de la muestra.

Detección. Por lo comentado anteriormente, estar ante un modelo donde los coeficientes no son significativamente distintos de cero, si lo es el modelo conjuntamente, y con un coeficiente de determinación alto, son síntomas que hacen pensar en la existencia de multicolinealidad grave. Sin embargo, basarse únicamente en dichos síntomas no es adecuado, por lo que se hace necesario disponer de otras herramientas. Las medidas más usadas para detectar el grado de multicolinealidad presente en un modelo de regresión son el factor de inflación de la varianza (FIV) y el número de condición (NC). Por desgracia, estas herramientas no son un contraste estadístico aplicado a la detección de la multicolinealidad, sino una regla práctica que trata de analizar la existencia de multicolinealidad aproximada en el modelo y determinar las variables implicadas en la misma.

El FIV es una de las medidas más usadas para detectar si el grado de multicolinealidad presente en el modelo es preocupante. Para cada una de las variables explicativas del modelo (1), se obtiene a partir de la expresión:

$$FIV(i) = \frac{var\left(\widehat{\beta}_i\right)}{var\left(\widehat{\beta}_i^o\right)} = \frac{1}{1-R_i^2}, \quad i = 2, \dots, p, \tag{2}$$

siendo $\widehat{\boldsymbol{\beta}}$ el estimador por MCO del modelo (1), $\widehat{\boldsymbol{\beta}}^o$ el estimador por MCO del modelo (1) suponiendo que las variables explicativas son ortogonales, y R_i^2 el coeficiente de determinación de la regresión auxiliar:

$$\mathbf{X}_i = \mathbf{X}_{-i}\boldsymbol{\delta} + \mathbf{w},$$

donde $\mathbf{X}_{-i}$ es igual a la matriz $\mathbf{X}$ tras eliminar la variable $\mathbf{X}_i$ para $i = 2, \dots, p$. Como $0 \leq R_i^2 \leq 1$, se verifica que $FIV(i) \geq 1$ para $i = 2, \dots, p$.

Puesto que el FIV se obtiene como el cociente entre la varianza observada y la varianza que se hubiera obtenido en el caso de que $\mathbf{X}_i$ estuviese incorrelada con el resto de variables explicativas del modelo, entonces muestra en qué medida se agranda la varianza del estimador como consecuencia de la relación lineal existente entre los regresores.

Es comúnmente aceptado que valores del FIV superiores a 10 indicarían que el grado de multicolinealidad presente en el modelo es preocupante ya que implicaría un coeficiente de determinación auxiliar superior a 0,9.

Por otro lado, dado el modelo de regresión lineal múltiple (1), el NC se define como:

$$K(\mathbf{X}) = \frac{\mu_{max}}{\mu_{min}}, \tag{3}$$

donde μ_{max} y μ_{min} son, respectivamente, los valores singulares máximo y mínimo de la matriz $\mathbf{X}$.

Puesto que la descomposición en valores singulares de $\mathbf{X}$ corresponde a $\mathbf{X} = \mathbf{UDV}^t$, donde $\mathbf{U}^t\mathbf{U} = \mathbf{I}$, $\mathbf{V}^t\mathbf{V} = \mathbf{I}$ siendo $\mathbf{I}$ la matriz identidad (de dimensiones adecuadas en cada caso) y $\mathbf{D} = diag(\mu_1 \dots \mu_p)$, con μ_i, $i = 1, \dots, p$, los valores singulares de la matriz $\mathbf{X}$, se tiene que:

$$\mathbf{X}^t\mathbf{X} = \mathbf{VDU}^t\mathbf{UDV}^t = \mathbf{VD}^2\mathbf{V}^t.$$

En tal caso se verifica que los autovalores de la matriz $\mathbf{X}^t\mathbf{X}$ coinciden con el cuadrado de los valores singulares de $\mathbf{X}$; es decir, $\xi_i = \mu_i^2$ para $i = 1, \dots, p$. Entonces, la expresión (3) es equivalente a:

$$K(\mathbf{X}) = \sqrt{\frac{\xi_{max}}{\xi_{min}}}, \tag{4}$$

donde ξ_{max} y ξ_{min} son, respectivamente, los autovalores máximo y mínimo de la matriz $\mathbf{X}^t\mathbf{X}$.

Según Belsley [1], valores de $K(\mathbf{X})$ entre 20 y 30 suponen multicolinealidad moderada y valores superiores a 30, multicolinealidad grave. Además, los datos deben tener longitud unidad; es decir, han de ser divididos por la raíz cuadrada de la sumatoria de cada uno de sus elementos elevados al cuadrado.

Soluciones. Entre las diversas soluciones que se suelen barajar, se tienen aquellas relacionadas con los datos de los que se dispone (como, por ejemplo, mejora del diseño muestral, incorporar más observaciones a la muestra o el uso de información *a priori*), la aplicación de técnicas de estimación alternativas a MCO (como, por ejemplo, regresión cresta, regresión alzada, regresión de componentes principales, regresión con variables ortogonales, regresión LASSO o máxima entropía) o incluso la eliminación del modelo de las variables que se consideran que provocan el problema.

Por desgracia, mejorar la calidad de los datos disponibles no siempre es posible; por lo que antes de eliminar variables del modelo (es conocido que esta elección podría conducir a la aparición de otros problemas como la heterocedasticidad y autocorrelación), sería aconsejable intentar usar alguna técnica alternativa a MCO para la estimación del modelo. En este sentido, el presente trabajo se centra en la aplicación de los métodos de alzado, de cresta y de variables ortogonales a modelos donde la presencia de multicolinealidad es grave.

Estructura del trabajo. En la sección 2, se obtienen las estimaciones por MCO del modelo (1) así como las expresiones referentes a la significación individual y conjunta (véase Novales *et al.* [8]). En la sección 3, se presentan los métodos de alzado, de cresta y de variables ortogonales. Esta sección no es más que una recopilación bibliográfica de los contenidos presentes en los trabajos de Salmerón *et al.* [10] y Novales *et al.* [8]. Adviértase que para obtener expresiones explícitas se ha considerado el caso en el que $p = 2$ y los datos están estandarizados, es decir, se les resta su media y, posteriormente, se divide entre $\sqrt{n}$ veces la desviación típica. En la sección 4, se aplican los resultados mostrados en la sección anterior a un conjunto de datos reales correspondiente a tipos de interés medios de letras del tesoro. Finalmente, en la sección 5, se destacan los principales resultados obtenidos.

2. Preliminares

En el presente apartado se realiza un resumen de las principales características que se obtienen al estimar por MCO el siguiente modelo de regresión lineal con 2 variables estandarizadas y n observaciones:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} = \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \mathbf{u}, \tag{5}$$

donde $\mathbf{y}$ representa a un vector $n \times 1$ que contiene las observaciones de la variable explicada, $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2]$ es una matriz $n \times 2$ que contiene las observaciones de las variables explicativas, $\boldsymbol{\beta} = (\beta_1 \ \beta_2)'$ es un vector 2×1 que contiene los coeficientes de las variables explicativas y $\mathbf{u}$ es un vector $n \times 1$ de perturbaciones aleatorias. Se supone que la perturbación aleatoria está centrada y es esférica. Además, por estar las variables estandarizadas, se tiene:

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}, \tag{6}$$

donde ρ corresponde a la correlación entre las variables $\mathbf{x}_1$ y $\mathbf{x}_2$ y γ_i a la correlación entre $\mathbf{x}_i$ e $\mathbf{y}$, para $i = 1, 2$.

Estimación y bondad de ajuste. En tal caso, se verifica que la estimación por mínimos cuadrados ordinarios (MCO) de los coeficientes de los regresores del modelo (5) será:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{pmatrix} \frac{\gamma_1 - \rho\gamma_2}{1-\rho^2} \\ \frac{\gamma_2 - \rho\gamma_1}{1-\rho^2} \end{pmatrix}, \tag{7}$$

cuya matriz de varianzas covarianzas estimada es:

$$\widehat{Var(\widehat{\boldsymbol{\beta}})} = \widehat{\sigma}^2 (\mathbf{X}'\mathbf{X})^{-1} = \widehat{\sigma}^2 \begin{pmatrix} \frac{1}{1-\rho^2} & -\frac{\rho}{1-\rho^2} \\ -\frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} \end{pmatrix}, \tag{8}$$

donde:

$$\widehat{\sigma}^2 = \frac{SCR}{n-2} = \frac{1}{n-2} \cdot \frac{1 - \rho^2 - \gamma_1^2 - \gamma_2^2 + 2\rho\gamma_1\gamma_2}{1-\rho^2}.$$

Ya que $SCT = 1$ (al estar el modelo estandarizado) y $SCE = \widehat{\boldsymbol{\beta}}^t \mathbf{X}^t \mathbf{y} = \gamma_1 \widehat{\beta}_1 + \gamma_2 \widehat{\beta}_2 = \frac{\gamma_1^2 + \gamma_2^2 - 2\rho\gamma_1\gamma_2}{1-\rho^2}$,[2] se tiene que:

$$R^2 = \frac{SCE}{SCT} = \frac{\gamma_1^2 + \gamma_2^2 - 2\rho\gamma_1\gamma_2}{1-\rho^2}.$$

Inferencia individual y conjunta. A partir de la expresión (8), los intervalos de confianza al nivel $1 - \alpha$ correspondientes a los coeficientes de los regresores son:

$$\widehat{\beta}_i \pm t_{n-2}\left(1 - \frac{\alpha}{2}\right) \widehat{\sigma} \sqrt{\frac{1}{1-\rho^2}}, \quad i = 1, 2, \tag{9}$$

[2] SCR denota la suma de cuadrados de los residuos, SCT la suma de cuadrados totales y SCE la suma de cuadrados explicada.

donde $t_{n-2}\left(1-\frac{\alpha}{2}\right)$ corresponde al punto de una t-Student con $n-2$ grados de libertad que deja a la izquierda una probabilidad $1-\frac{\alpha}{2}$.

Mientras que las regiones de rechazo de los contrastes de significación individual (aquellos que tienen como hipótesis nula $H_0 : \beta_i = 0$ y alternativa $H_1 : \beta_i \neq 0$ para $i = 1, 2$) serán:

$$t_{exp} = \left| \frac{\widehat{\beta}_i}{\widehat{\sigma}\sqrt{\frac{1}{1-\rho^2}}} \right| > t_{n-2}\left(1-\frac{\alpha}{2}\right), \quad i = 1, 2,$$

la región de rechazo del contraste de significación conjunta (aquel que tiene como hipótesis nula $H_0 : \beta_1 = \beta_2 = 0$ frente a la alternativa $H_1 : \exists i$ tal que $\beta_i \neq 0$ con $i = 1, 2$) será:

$$F_{exp} = \frac{SCE}{\widehat{\sigma}^2} = (n-2) \cdot \frac{\gamma_1^2 + \gamma_2^2 - 2\rho\gamma_1\gamma_2}{1 - \rho^2 - \gamma_1^2 - \gamma_2^2 + 2\rho\gamma_1\gamma_2} > F_{2,n-2}(1-\alpha),$$

donde $F_{2,n-2}(1-\alpha)$ es el punto de una F de Snedecor con 2 y $n-2$ grados de libertad que deja a su izquierda una probabilidad $1-\alpha$.

Detección de la multicolinealidad. Finalmente, el FIV asociado al modelo (5) se obtiene, por ejemplo, a partir del coeficiente de determinación de la regresión auxiliar $\mathbf{x}_1 = \alpha\mathbf{x}_2 + \mathbf{v}$. Puesto que en la regresión simple el coeficiente de deteminación coincide con el cuadrado del coeficiente de correlación de las variables, se obtiene que:

$$\text{FIV} = \frac{1}{1-\rho^2}. \tag{10}$$

Por otro lado, el número de condición se obtendrá a partir de los autovalores de la matriz $\mathbf{X}'\mathbf{X}$.

3. Técnicas de estimación bajo multicolinealidad

Como se ha comentado, en presencia de multicolinealidad grave en el modelo de regresión, el análisis por MCO de la sección anterior queda en entredicho por lo que podrían aplicarse técnicas alternativas a los MCO como la estimación por componentes principales, la regresión LASSO, la estimación alzada, la cresta, con variables ortogonales, etc. Este apartado se centra en éstas tres últimas. Para más detalles, véanse los trabajos de Salmerón *et al.* [10] y Novales *et al.* [8].

3.1. Método de regresión alzada

Dado el modelo (5), la regresión alzada trata el problema de la multicolinealidad desde un punto de vista geométrico.

La colinealidad surge porque el vector $\mathbf{x}_1$ y el vector $\mathbf{x}_2$ están muy cerca geométricamente; es decir, el ángulo θ_1 que determinan ambos vectores es muy pequeño (véase Figura 1). Para corregir este problema antes de proceder a la estimación, vamos a separarlos a través de la siguiente regresión:

$$\mathbf{x}_1 = \alpha\mathbf{x}_2 + \boldsymbol{\epsilon}, \tag{11}$$

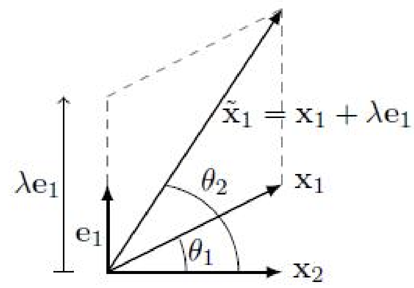

Figura 1: Representación del método alzado para dos variables independientes

cuya estimación por MCO es $\hat{\alpha} = (\mathbf{x}_2'\mathbf{x}_2)^{-1}\mathbf{x}_2'\mathbf{x}_1$, de forma que se verifica que $\hat{\alpha} = \rho$ (ya que los datos están estandarizados). En tal caso, se tiene que $\mathbf{x}_1 = \rho\mathbf{x}_2 + \mathbf{e}_1$ con $\mathbf{e}_1 \perp \mathbf{x}_2$ donde $\mathbf{e}_1$ es el residuo obtenido de la regresión (11). Luego, el vector alzado se define como:

$$\tilde{\mathbf{x}}_1 = \mathbf{x}_1 + \lambda\mathbf{e}_1. \tag{12}$$

De la Figura 1, se evidencia que el ángulo θ_2 entre $\tilde{\mathbf{x}}_1$ y $\mathbf{x}_2$ será mayor que el ángulo θ_1. Por lo tanto, sustituyendo el vector $\mathbf{x}_1$ por el vector alzado $\tilde{\mathbf{x}}_1$ en el modelo (5), la correlación entre ambos vectores será menor y el problema de la multicolinealidad disminuirá. Cuanto mayor sea el parámetro λ, mayor será el ángulo entre los vectores y menor será la correlación.

Estimación. Por tanto, el nuevo modelo viene dado por:

$$\mathbf{y} = \beta_1(\lambda)\tilde{\mathbf{x}}_1 + \beta_2(\lambda)\mathbf{x}_2 + \mathbf{w}. \tag{13}$$

Mediante el uso de la estimación MCO en el modelo (13), podemos obtener los siguientes estimadores:

$$\hat{\boldsymbol{\beta}}(\lambda) = \left(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}\right)^{-1}\tilde{\mathbf{X}}'\mathbf{y} = \left(\begin{array}{c} \frac{\gamma_1 - \rho\gamma_2}{(1+\lambda)(1-\rho^2)} \\ \frac{(1+\lambda)\gamma_2 - \rho\gamma_1 - \rho^2\gamma_2\lambda}{(1+\lambda)(1-\rho^2)} \end{array}\right), \tag{14}$$

donde $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1\mathbf{x}_2]$.

Bondad de ajuste e inferencia. Las sumas de los cuadrados en la regresión alzada coinciden con los de la estimación por MCO (véase Salmerón *et al.* [10]). Por esta razón, el coficiente de determinación, la estimación de la varianza de la perturbación aleatoria y la F experimental serán las mismas para los modelos MCO y alzado:

$$R^2(\lambda) = R^2, \quad \widehat{\sigma}^2(\lambda) = \widehat{\sigma}^2, \quad F_{exp}(\lambda) = F_{exp}.$$

Las expresiones de las t experimentales son:

$$t_{exp}(\beta_1(\lambda)) = \frac{\gamma_1 - \rho\gamma_2}{\sqrt{\hat{\sigma}^2(1-\rho^2)}}, \quad t_{exp}(\beta_2(\lambda)) = \frac{(1+\lambda)\gamma_2 - \rho\gamma_1 - \rho^2\gamma_2\lambda}{\sqrt{\hat{\sigma}^2(1-\rho^2)[(1+\lambda)^2 - \rho^2(2\lambda+\lambda^2)]}}. \tag{15}$$

Podemos ver que la correspondiente t experimental de la variable alzada, $t_{exp}(\beta_1(\lambda))$, no depende de λ. Por lo tanto se mantiene constante, mientras que $t_{exp}(\beta_2(\lambda))$ depende de λ y variará ante incrementos de esta.

Debemos recordar que uno de los síntomas de la multicolinealidad grave son los modelos significativos globalmente pero no de forma individual. Aplicando la regresión alzada, la condición inicial de significación conjunta se respeta y la de no significación individual puede variar de forma beneficiosa.

Factor de inflación de la varianza. Una vez aplicado el procedimiento alternativo a los MCO, es necesario comprobar si es efectivo; lo que consiste, por ejemplo, en calcular el FIV.

El FIV asociado a la regresión alzada viene dado por:

$$\mathrm{FIV}(\lambda) = \frac{1}{1 - \frac{\rho^2}{(1+\lambda)^2 - (\lambda^2+2\lambda)\rho^2}} = \frac{(1+\lambda)^2(1-\rho^2)+\rho^2}{(1+\lambda)^2(1-\rho^2)} = 1 + \frac{\rho^2}{(1+\lambda)^2(1-\rho^2)}, \tag{16}$$

de forma que, para un λ concreto, se puede comprobar si la multicolinealidad se ha mitigado o no.

A partir de (16) es claro que el $\mathrm{FIV}(\lambda)$ es siempre mayor o igual a 1, es decreciente en λ y continuo en $\lambda = 0$ (coincidiendo con el obtenido para el modelo (5)). Para más detalles, véase Salmerón *et al.* [10].

3.2. Método de regresión cresta

La regresión cresta trata la multicolinealidad desde un punto de vista sistemático (es decir, puramente numérico), introduciendo una cantidad positiva en la diagonal principal de la matriz $\mathbf{X}'\mathbf{X}$.

Estimación. El estimador cresta $\hat{\boldsymbol{\beta}}(k)$, con $k \geq 0$, se expresa de la siguiente forma:

$$\hat{\boldsymbol{\beta}}(k) = \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}\mathbf{X}'\mathbf{y} = \begin{pmatrix} \frac{(1+k)\gamma_1 - \rho\gamma_2}{(1+k)^2-\rho^2} \\ \frac{(1+k)\gamma_2 - \rho\gamma_1}{(1+k)^2-\rho^2} \end{pmatrix}, \tag{17}$$

con matriz de varianzas covarianzas:

$$\mathrm{var}(\hat{\boldsymbol{\beta}}(k)) = \sigma^2 \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1} \mathbf{X}'\mathbf{X} \left(\mathbf{X}'\mathbf{X} + k\mathbf{I}\right)^{-1}. \tag{18}$$

Marquardt [5] muestra que la estimación dada en (17) se puede obtener estimando por MCO el modelo aumentado:

$$\mathbf{y}_A = \mathbf{X}_A\beta + \mathbf{u}_A, \tag{19}$$

donde:

$$\mathbf{y}_A = \begin{pmatrix} \mathbf{y}_{n\times 1} \\ \mathbf{0}_{p\times 1} \end{pmatrix}, \quad \mathbf{X}_A = \begin{pmatrix} \mathbf{X}_{n\times p} \\ \sqrt{k}\mathbf{I}_{\mathbf{p}\times\mathbf{p}} \end{pmatrix}.$$

Test de significación individual. Los valores t experimentales para el test de significación individual de los estimadores cresta se pueden obtener, para $i, j = 1, 2$ e $i \neq j$, mediante:

$$t_{exp}(\beta_i(k)) = \frac{\hat{\beta}_i(k)}{\sqrt{\hat{\sigma}^2(k) b_{i,i}}} = \frac{(1+k)\gamma_i - \rho\gamma_j}{\hat{\sigma}(k)\sqrt{(1+k)^2 - 2(1+k)\rho^2 + \rho^2}}, \tag{20}$$

donde $b_{i,i}$ es el elemento (i, i) de la matriz $(\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}'\mathbf{X} (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}$ y la estimación de la varianza de la perturbación aleatoria se puede obtener a partir de:

$$\hat{\sigma}^2(k) = \frac{SCR(k)}{n-2} = \frac{(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}(k))'(\mathbf{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}(k))}{n-2}.$$

Se puede apreciar que ambas t experimentales dependen del parámetro k, por lo que variarán conforme cambie éste parámetro.

Bondad de ajuste y test de significación global. McDonald [6] obtuvo la siguiente expresión para calcular el coeficiente de determinación del estimador cresta:

$$R^2(k) = \frac{\left(\hat{\boldsymbol{\beta}}(k)'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}(k) + k\hat{\boldsymbol{\beta}}(k)'\hat{\boldsymbol{\beta}}(k)\right)^2}{\hat{\boldsymbol{\beta}}(k)'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}(k)}, \tag{21}$$

demostrando que es decreciente en k. A partir de esta expresión, es posible obtener el valor experimental para el test de significación global:

$$F_{exp}(k) = \frac{R^2(k)}{\frac{1-R^2(k)}{n-2}}, \tag{22}$$

que de igual forma, también es decreciente en k.

Por tanto, aumentando el valor de k, se tiene que decrecen el coeficiente de determinación y el valor experimental de la significación global. Es decir, puede ocurrir que, para un valor de k dado, se haya mitigado el problema de la multicolinealidad pero que el modelo no sea válido globalmente.

Factor de inflación de la varianza. Siguiendo a Marquardt [5], el elemento de la diagonal principal de la matriz $(\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}'\mathbf{X} (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}$ son los factores de inflación de la varianza[3], es decir:

$$\text{FIV}(k) = \frac{(1+k)^2 - 2(1+k)\rho^2 + \rho^2}{[(1+k)^2 - \rho^2]^2}. \tag{23}$$

Para $k = 0$, FIV(k) coincide con el FIV del modelo (5) y con el FIV de la regresión alzada para $\lambda = 0$. García y otros [2] mostraron que esta expresión presenta algunas anomalías, como por ejemplo, que toma valores inferiores a 1, que hacen desaconsejable su aplicación. En García *et al.* [4] se muestra cómo ha de calcularse el FIV de forma adecuada en la regresión cresta.

[3]Siempre y cuando los datos estén estandarizados.

3.3. Método de regresión con variables ortogonales

En la regresión con variables ortogonales se sustituye una de las variables explicativas del modelo (5) por el componente de la misma que no está relacionado con la otra. Este componente se obtiene de la estimación de la regresión de la variable elegida sobre las demás variables explicativas; más concretamente, el residuo de dicha regresión es el componente que buscamos.

De esta forma, supongamos que en el modelo (5) queremos ortogonalizar la variable $\mathbf{x}_2$, en tal caso se plantea la siguiente regresión auxiliar:

$$\mathbf{x}_2 = \alpha \mathbf{x}_1 + \mathbf{w}. \tag{24}$$

A continuación sustituimos en el modelo (5) la variable $\mathbf{x}_2$ por los residuos obtenidos en la regresión auxiliar:

$$\mathbf{y} = \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{e} + \mathbf{u}. \tag{25}$$

Ahora las variables explicativas del modelo son ortogonales entre sí, por lo que el problema de multicolinealidad aproximada se habría resuelto (el FIV es igual a 1). En el caso de tener un modelo con más variables, la multicolinealidad habría descendido; si no fuera suficiente, se procedería a ortogonalizar otra variable.

Si bien, en la estimación alzada y cresta, la interpretación de los coeficientes estimados no cambia, en este caso sí que lo hace ya que el residuo $\mathbf{e}$ es la parte de $\mathbf{x}_2$ que no está relacionada con $\mathbf{x}_1$. Entonces, β_2 se interpreta como el efecto que sufre la variable explicada debido a la parte de $\mathbf{x}_2$ que no está relacionada con $\mathbf{x}_1$. Esta interpretación abre la puerta a nuevas investigaciones; por lo que la regresión ortogonal, además de mitigar el problema de multicolinealidad, nos brinda otra visión e interpretación de los modelos que puede resultar muy útil a los investigadores.

Estimación. Si llamamos $\mathbf{X}_O = [\mathbf{x}_1 \ \mathbf{e}]$, se tiene que la estimación de los coeficientes por MCO del modelo (25) será:

$$\widehat{\boldsymbol{\beta}}_O = (\mathbf{X}_O^t \mathbf{X}_O)^{-1} \mathbf{X}_O^t \mathbf{y} = \left(\begin{array}{cc} 1 & 0 \\ 0 & 1-\rho^2 \end{array} \right)^{-1} \left(\begin{array}{c} \gamma_1 \\ \gamma_2 - \rho\gamma_1 \end{array} \right) = \left(\begin{array}{c} \gamma_1 \\ \frac{\gamma_2 - \rho\gamma_1}{1-\rho^2} \end{array} \right). \tag{26}$$

Bondad de ajuste e inferencia. La matriz de varianzas covarianzas estimada es:

$$\widehat{Var(\widehat{\boldsymbol{\beta}}_O)} = \widehat{\sigma}_O^2 (\mathbf{X}_O^t \mathbf{X}_O)^{-1} = \widehat{\sigma}_O^2 \left(\begin{array}{cc} 1 & 0 \\ 0 & \frac{1}{1-\rho^2} \end{array} \right), \tag{27}$$

donde $\widehat{\sigma}_O^2 = \widehat{\sigma}^2$ ya que $SCR_O = SCR$ (véase Novales *et al.* [8]). Además, se verifica que $SCT_O = 1$ y $SCE_O = SCE$ (ver Novales *et al.* [8]), por lo que se tiene que $R_O^2 = R^2$ y que las regiones de rechazo del contraste de significación conjunta coincidirán ($F_{exp,O} = F_{exp}$).

Test de significación individual. A partir de la expresión (27), los intervalos de confianza al nivel $1-\alpha$ correspondientes a los coeficientes de los regresores son:

$$\widehat{\beta}_{1O} \pm t_{n-2}\left(1 - \frac{\alpha}{2}\right) \widehat{\sigma}, \qquad \widehat{\beta}_2 \pm t_{n-2}\left(1 - \frac{\alpha}{2}\right) \widehat{\sigma} \sqrt{\frac{1}{1-\rho^2}}, \tag{28}$$

y las regiones de rechazo de los contrastes de significación individual serán:

$$t_{exp} = \left| \frac{\widehat{\beta}_{1O}}{\widehat{\sigma}} \right| > t_{n-2}\left(1 - \frac{\alpha}{2}\right), \qquad t_{exp} = \left| \frac{\widehat{\beta}_2}{\widehat{\sigma}\sqrt{\frac{1}{1-\rho^2}}} \right| > t_{n-2}\left(1 - \frac{\alpha}{2}\right).$$

Se observa que las únicas diferencias entre el análisis de los modelos (5) y (25) las encontramos en la estimación del primer coeficiente y de su correspondiente intervalo de confianza y contraste de significación individual; todo lo demás permanece intacto.

4. Aplicación práctica

El objetivo del siguiente ejemplo es poner en práctica los resultados obtenidos en el apartado anterior sobre los métodos de estimación alternativos a los MCO cuando en el modelo de regresión existe un grado de multicolinealidad grave.

Destacar que en estos modelos se deben comprobar también las hipótesis básicas de homocedasticidad e incorrelación, aunque no se ha hecho debido a que no es el objetivo del presente trabajo y no deseamos alejarnos de lo que constituye la principal aportación del mismo, que no es más que la aplicación de las distintas metodologías estudiadas.

Finalmente, indicar que, aunque los resultados anteriores se han obtenido tras estandarizar las variables, estos siguen siendo válidos sin haber realizado transformación alguna. Así, en el siguiente ejemplo se usan los datos originales.

4.1. Rendimientos de letras del tesoro

En este modelo se analizan los tipos de interés medio de las letras del tesoro. Se utiliza para ello tres series temporales mensuales que abarcan desde febrero de 2013 hasta febrero de 2016; es decir, se cuentan con 37 observaciones de los tipos de interés medio de las letras del tesoro a 6 meses, 9 meses y 12 meses. Como variable dependiente se usa el tipo de interés medio a 12 meses (**TI12**) y como variables independientes el tipo de interés medio a 6 (**TI6**) y 9 meses (**TI9**).

En el gráfico de la Figura 2, se puede observar que las tres series tienen un tendencia muy similar, lo que nos hace sospechar que las variables pueden presentar un nivel alto de colinealidad. Por tanto, procedemos a calcular los FIV, obteniéndose que $FIV = 20{,}782$. Puesto que es superior a 10, podemos afirmar que el modelo presenta multicolinealidad grave y entonces la estimación MCO podría conducir a errores y a conclusiones inestables.

El modelo a estimar queda de la siguiente forma:

$$\mathbf{TI12} = \beta_0 + \beta_1 \mathbf{TI6} + \beta_2 \mathbf{TI9} + \mathbf{u}, \tag{29}$$

donde se considera que la perturbación aleatoria $\mathbf{u}$ está centrada y es homocedástica e incorrelada.

Si aplicamos la regresión por MCO, los resultados que se obtienen se muestran en el Cuadro 1. Podemos observar que los coeficientes de las variables independientes son significativamente distintos de cero, excepto el término independiente que no lo es, y el modelo es significativo globalmente.

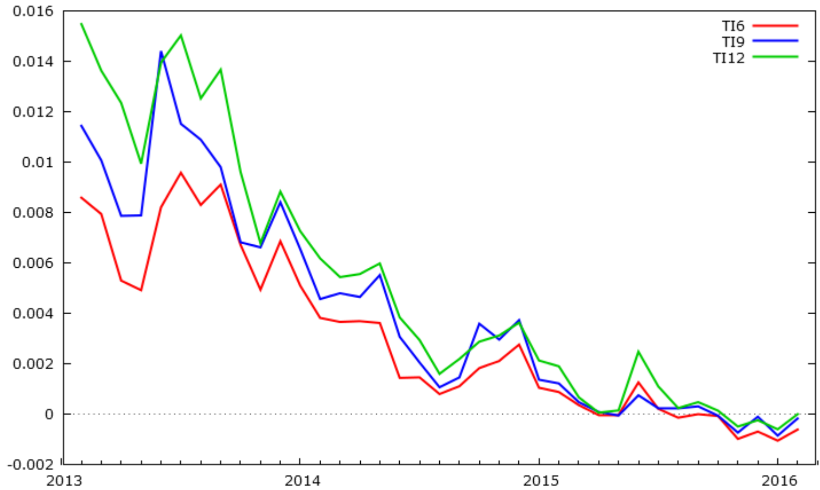

Figura 2: Tipos de interés medio de las letras del tesoro

Término independiente	**TI6**	**TI9**	R^2	F de ANOVA (2,34)
0,000343968 (0,000215074)	0,862294 (0,214776)	0,531898 (0,16668)	0,969096	533,0873

Cuadro 1: Estimación por MCO del modelo de rendimientos del tesoro (entre paréntesis la desviación típica estimada)

Adviértase que, en O'Brien [9], se muestra que la varianza de los coeficientes estimados depende de diversos factores:

$$\widehat{\text{var}\left(\widehat{\beta}_i\right)} = \frac{\widehat{\sigma}^2}{n \cdot \text{var}(\mathbf{X}_i)} \cdot FIV(i),$$

por lo que un FIV alto no implica necesariamente que no se vaya a rechazar la hipótesis nula en los contrastes de significación individual.

4.1.1. Aplicación de la regresión alzada

Procedemos a realizar la siguiente regresión auxiliar sobre el modelo (29) con el fin de alzar la variable **TI9**:

$$\mathbf{TI9} = \alpha_0 + \alpha_1 \mathbf{TI6} + \boldsymbol{\epsilon}. \tag{30}$$

A partir de los residuos de la regresión anterior, **e**, el vector alzado se define como:

$$\widetilde{\mathbf{TI9}} = \mathbf{TI9} + \lambda \mathbf{e}. \tag{31}$$

A continuación, el nuevo modelo queda de la siguiente forma:

$$\mathbf{TI12} = \beta_0(\lambda) + \beta_1(\lambda)\mathbf{TI6} + \beta_2(\lambda)\widetilde{\mathbf{TI9}} + \mathbf{w}. \tag{32}$$

Los resultados de estimar el modelo anterior por MCO se muestran en el Cuadro 2 para distintos valores de λ. Se puede observar que:

- El coeficiente de determinación y la significación conjunta se mantienen intactas con respecto al modelo de MCO.
- La t experimental de la variable alzada, **TI9**, no cambia; pero por el contrario, sí cambia la t experimental del resto de variables, en este caso de **TI6**, y va en aumento cuanto mayor es el parámetro λ.
- El FIV decrece a medida que aumentamos el parámetro λ y, por lo tanto, el grado de multicolinealidad va disminuyendo.
- Si nos quedamos con $\lambda = 0{,}5$, ya que es primer valor en el que los FIV presentan valores por debajo de 10, la estimación de los coeficientes refleja que:
  - cuando **TI6** aumenta un 1 %, manteniendose **TI9** constante, **TI12** aumenta un 1,08519 %.
  - cuando **TI9** aumenta en un 1 % y se mantiene constante **TI6**, **TI12** aumenta un 0,354598 %.

4.1.2. Aplicación de la regresión cresta

Siguiendo Marquardt [5], el modelo aumentado, véase expresión (19), que se ha utilizado para la regresión cresta es el siguiente:

$$\mathbf{TI12}_A = \beta_0 + \beta_1 \mathbf{TI6}_A + \beta_2 \mathbf{TI9}_A + \mathbf{u}_A. \tag{33}$$

Los resultados obtenidos al estimar dicho modelo se muestran en el Cuadro 3. Se puede observar que:

λ	Término independiente	**TI6**	$\widetilde{\textbf{TI9}}$	R^2	F de ANOVA (2,34)	t exp **TI6**	t exp $\widetilde{\textbf{TI9}}$	FIV **TI6**	FIV **TI9**
0,1	0,00035912 (0,00021397)	0,923084 (0,196235)	0,483543 (0,151527)	0,969096	533,0873	4,704	3,191	17,349	17,349
0,2	0,000371747 (0,000213127)	0,973742 (0,180865)	0,443248 (0,1389)	0,969096	533,0873	5,384	3,191	14,738	14,738
0,3	0,000382431 (0,000212468)	1,01661 (0,167932)	0,409152 (0,128215)	0,969096	533,0873	6,054	3,191	12,705	12,705
0,4	0,000391589 (0,000211944)	1,05335 (0,156915)	0,379927 (0,119057)	0,969096	533,0873	6,713	3,191	11,093	11,093
0,5	0,000399526 (0,000211521)	1,08519 (0,147427)	0,354598 (0,11112)	0,969096	533,0873	7,361	3,191	9,792	9,792
0,6	0,000406471 (0,000211173)	1,11305 (0,139182)	0,332436 (0,104175)	0,969096	533,0873	7,997	3,191	8,727	8,727
0,7	0,000412599 (0,000210885)	1,13764 (0,131958)	0,312881 (0,098047)	0,969096	533,0873	8,621	3,191	7,845	7,845
0,8	0,000418045 (0,000210643)	1,15949 (0,125586)	0,295499 (0,0926)	0,969096	533,0873	9,233	3,191	7,106	7,106
0,9	0,000422919 (0,000210438)	1,17904 (0,119928)	0,279946 (0,0877263)	0,969096	533,0873	9,831	3,191	6,480	6,480
1	0,000427305 (0,000210263)	1,19664 (0,114878)	0,265949 (0,08334)	0,969096	533,0873	10,42	3,191	5,964	5,964
50	0,000507375 (0,000208638)	1,51787 (0,0472916)	0,0104294 (0,00326823)	0,969096	533,0873	32,10	3,191	1,008	1,008
100	0,000508993 (0,000208636)	1,52436 (0,0471585)	0,00526631 (0,0016503)	0,969096	533,0873	32,32	3,191	1,002	1,002
500	0,00051031 (0,000208635)	1,52964 (0,0471147)	0,00106167 (0,000332695)	0,969096	533,0873	32,47	3,191	1	1
1000	0,000510476 (0,000208635)	1,53031 (0,0471133)	0,000531366 (0,000166513)	0,969096	533,0873	32,48	3,191	1	1

Cuadro 2: Estimación por MCO del modelo (32) sobre rendimientos del tesoro (entre paréntesis la desviación típica estimada)

- El coeficiente de determinación (R^2) y la significación global disminuyen a medida que aumenta el parámetro k, lo cual puede provocar problemas a la hora de validar el modelo.
- La significación individual (t_{epx}) también decrece para las dos variables independientes de manera que los coeficientes de las dos variables explicativas dejan de ser significativamente distintos de cero.
- En este modelo nos quedamos con un valor de $k = 0{,}03$, ya que es el primer valor en el que los FIV presentan valores por debajo de 10 (véase Cuadro 4). En este caso, no tiene sentido interpretar los coeficientes obtenidos, debido a que no son significativos individualmente.

En el Cuadro 4 se muestran los FIV calculados para la estimación cresta a partir de las expresiones (23) y (2) dadas por Marquardt [5] y Zhang e Ibrahim [12], respectivamente. Los FIV decrecen a medida que aumentamos el parámetro k, llegando a mitigar el problema de multicolinealidad; pero como hemos comentado, provocando otros efectos adversos. Además se observa que, siguiendo los cálculos de Marquardt, se pueden obtener FIV menores que 1, lo cual va en contradicción con la definición dada en (2).

4.1.3. Aplicación de la regresión con variables ortogonales

Para sacarle el mayor partido posible a la regresión ortogonal, vamos a proceder a ortogonalizar la variable **TI9**. De esta forma, se va a poder analizar el efecto del tercer trimestre sobre la variable dependiente, **TI12**; es decir, del séptimo, octavo y noveno mes, ya que ésta es la parte de **TI9** que no está relacionada con **TI6**. Para ello, debemos plantear la siguiente regresión auxiliar:

$$\mathbf{TI9} = \alpha_0 + \alpha_1 \mathbf{TI6} + \mathbf{w}, \tag{34}$$

ya que sus residuos, **e**, se interpretan como el tercer trimestre al ser la parte de **TI9** que no tiene relación con **TI6**.

Si sustituimos en el modelo (29) la variable **LT9** por los residuos de la regresión auxiliar, se tiene el modelo con variables ortogonales:

$$\mathbf{TI12} = \beta_0 + \beta_1 \mathbf{TI6} + \beta_2 \mathbf{e} + \mathbf{u}, \tag{35}$$

cuya estimación por MCO se muestra en el Cuadro 5. Se puede observar que:

- El coeficiente de determinación (R^2) y la significación global del modelo permanecen inalteradas con respecto a la estimación por MCO del modelo (29) (véase Cuadro 1).
- El coeficiente de la variable ortogonalizada, **TI9**, también permanece igual al del modelo estimado por MCO; sin embargo, la estimación de la otra variable, **TI6**, sí sufre cambios (véase Cuadro 1).
- Los coeficientes de las variables explicativas son significacivamente distintos de cero (t experimentales igual a 32,5 y 3,191, respectivamente). El modelo es significativo globalmente.
- La estimación de los coeficientes refleja que cuando **TI6** (los dos primeros trimestres) aumenta un 1 %, **TI12** aumenta un 1,53098 %. En el modelo original es un 0,8622 % (véase Cuadro 1). Al aislar los dos primeros trimestres, su efecto sobre **TI12** prácticamente se duplica. Por otra parte, cuando **e** aumenta en un 1 % (el tercer trimestre), **TI12** aumenta un 0,531898 %.

k	Término independiente	**TI6**	**TI9**	R^2	F de ANOVA (2,34)	t exp **TI6**	t exp **TI9**
0,03	0,0049842 (0,0008294)	0,0190850 (0,0285177)	0,0245165 (0,0283986)	0,5248	13,62	0,6692335	0,8632996
0,1	0,005084 (0,000830)	0,005870 (0,015880)	0,007543 (0,015860)	0,513	12,99	0,3696474	0,475599
0,2	0,0050949 (0,0008301)	0,0029574 (0,0112829)	0,0038010 (0,0112757)	0,5093	12,8	0,2621135	0,3370966
0,3	0,0050893 (0,0008301)	0,0019800 (0,0092356)	0,0025452 (0,0092317)	0,5071	12,69	0,2143878	0,2757022
0,4	0,0050797 (0,0008302)	0,0014901 (0,0080137)	0,0019157 (0,0080112)	0,5054	12,6	0,1859441	0,2391277
0,5	0,0050686 (0,0008302)	0,0011957 (0,0071799)	0,0015375 (0,0071780)	0,5038	12,52	0,1665344	0,2141962
0,6	0,0050567 (0,0008302)	0,0009993 (0,0065646)	0.0012851 (0,0065632)	0.5023	12.45	0.1522256	0.1958039
0,7	0,0050444 (0,0008302)	0,0008589 (0,0060867)	0,0011047 (0,0060856)	0,5009	12,38	0,1411109	0,1815269
0,8	0,0050320 (0,0008302)	0,0007535 (0,0057018)	0,0009693 (0,0057009)	0,4994	12,31	0,1321513	0,1700258
0,9	0,0050193 (0,0008302)	0,0006715 (0,0053833)	0,0008640 (0,0053825)	0,4981	12,24	0,1247376	0,1605202
1	0,0050067 (0,0008302)	0,0006059 (0,0051141)	0,0007797 (0,0051134)	0,4967	12,17	0,1184764	0,1524817

Cuadro 3: Estimación por MCO del modelo aumentado (33) para los rendimientos del tesoro (entre paréntesis la desviación típica estimada)

- Como las variables independientes son ortogonales entre sí (los FIV asociados toman el valor 1), el problema de multicolinealidad ha sido eliminado. Además, dicha ortogonalidad implica que en este modelo realmente se verifique el *ceteris paribus*; es decir, cuando aumenta **TI6** se verifica que **e** permanece constante y viceversa.

- La estimación de las variables inalteradas coincide prácticamente con la obtenida para las mismas variables cuando se aplica la regresión alzada para valores altos de λ (véase Cuadro 2).

5. Conclusiones

En el presente trabajo, en primer lugar se ha realizado una recopilación bibliográfica sobre técnicas de estimación alternativas a los mínimos cuadrados ordinarios cuando en el modelo de regresión hay multicolinealidad grave. Más concretamente, se han estudiado los métodos de alzado, cresta y variables ortogonales obteniéndose que:

- En la regresión alzada se mantienen inalterados tanto el coeficiente de determinación (R^2) como la significación global del modelo (F_{exp}) con respecto a la estimación por MCO. La significación individual (t_{exp}) de la variable alzada tampoco sufre cambios; sin embargo, la significación individual de las variables que permanecen intactas sí varía, lo que puede ser positivo para el modelo si se consigue que los coeficientes que no eran significativamente distintos de cero pasen a serlo.

- En la regresión cresta, el coeficiente de determinación (R^2) y la significación conjunta (F_{exp}) decrecen a medida que aumenta el parámetro k, lo cual es un aspecto negativo para el modelo, pues con el fin de mitigar el problema de multicolinealidad de las variables podemos incurrir en un modelo no válido. En el caso de la significación individual (t_{exp}), esta cambia con respecto a las estimación por MCO y, al igual que en el caso anterior, esto puede ser positivo si se consigue que los coeficientes sean significativamente distintos de cero.

- La regresión con variables ortogonales, al igual que la alzada, mantiene el coeficiente de determinación (R^2) y la significación global (F_{exp}) sin cambios con respecto a la estimación MCO. La variable ortogonalizada mantiene inalterada la significación individual (t_{exp}), así como el coeficiente estimado y su desviación típica estimada correspondiente. La significación individual de las demás variables sí que sufre cambios. En el caso de un modelo con dos variables explicativas, la multicolinealiad es eliminada. Además, la nueva interpretación de la variable ortogonalizada puede servir para dar respuestas no accesibles con el modelo original.

En segundo lugar, se han aplicado los resultados anteriores a un conjunto de datos reales. Se observa que:

- Hemos visto de forma empírica que la regresión con variables ortogonales es el límite de la alzada cuando λ tiende a infinito. Luego, para valores de λ razonablemente altos, ambas técnicas de estimación pueden proporcionar resultados similares.

- La regresión con variables ortogonales y alzada son preferibles a la regresión cresta, ya que las primeras no alteran tanto el modelo de partida y la última puede llegar a “destrozarlo”. Esta es una cuestión importante ya que la regresión cresta ha sido universalmente usada cuando en el modelo ha existido multicolinealidad grave.

	Marquardt [5]		Zhang e Ibrahim [12]	
k	**TI6**	**TI9**	**TI6**	**TI9**
0,01	10,568020	10,568020	14,949795	14,949795
0,02	6,437767	6,437767	11,747760	11,747760
0,03	4,367304	4,367304	9,724347	9,724347
0,04	3,183468	3,183468	8,330125	8,330125
0,05	2,443372	2,443372	7,311231	7,311231
0,06	1,949742	1,949742	6,534182	6,534182
0,07	1,603892	1,603892	5,922079	5,922079
0,08	1,352013	1,352013	5,427502	5,427502
0,09	1,162733	1,162733	5,019616	5,019616
0,1	1,016761	1,016761	4,677505	4,677505
0,2	0,4506237	0,4506237	2,938800	2,938800
0,3	0,3065038	0,3065038	2,277991	2,277991
0,4	0,2426567	0,2426567	1,932521	1,932521
0,5	0,2054679	0,2054679	1,721608	1,721608
0,6	0,1801438	0,1801438	1,580335	1,580335
0,7	0,1611913	0,1611913	1,479668	1,479668
0,8	0,1461387	0,1461387	1,404683	1,404683
0,9	0,1337086	0,1337086	1,346932	1,346932
1	0,1231677	0,1231677	1,301279	1,301279

Cuadro 4: FIV de la estimación cresta del modelo de rendimientos del tesoro

Término independiente	TI6	e	R^2	F de ANOVA (2,34)
0,000510643 (0,000208635)	1,53098 (0,0471128)	0,531898 (0,16668)	0,969096	533,0873

Cuadro 5: Estimación con variables ortogonales del modelo de rendimientos de las letras del tesoro (entre paréntesis las desviaciones típicas estimadas)

- Sin embargo, no queda clara la preferencia entre la regresión alzada y con variables ortogonales. Seguramente el investigador deba decidir cuál usar dependiendo de las características del modelo que desea analizar.
- La regresión con variables ortogonales es de especial interés independientemente de que exista el problema de multicolinealidad en el modelo de regresión ya que puede ser usado en casos concretos donde el investigador quiera dar respuesta a preguntas que no se pueden contestar mediante el modelo original.

Finalmente, hay dos cuestiones no abordadas en este trabajo:

- Por un lado, no se han establecido criterios sobre las variables que se han de alzar u ortogonalizar. Esta decisión depende de diversos factores y el investigador debe tomarla dependiendo del modelo con el que trabaje en cada caso. Así por ejemplo, son variables candidatas a ser transformadas, aquellas cuyos coeficientes sean significativamente distintos de cero inicialmente (ya que esa característica no cambiará) o que tenga un mayor FIV. En la regresión ortogonal hay que buscar también que los residuos de la regresión auxiliar usada tengan sentido. Así, por ejemplo, en este trabajo se ha decidido ortogonalizar la variable **TI9** ya que es la que permite establecer una interpretación coherente de los residuos. También se ha alzado esta variable para comparar los resultados proporcionados por ambas técnicas para valores altos de λ.
- Por otro lado, tampoco se han comentado posibles criterios para determinar cuáles son los valores óptimos de k y λ. Siguiendo a García *et al.* [3], se ha considerado el primer valor de k y λ que tienen un FIV asociado menor que 10.

Agradecimientos

Los autores desean agradecer al Departamento de Métodos Cuantitativos para la Economía y la Empresa de la Universidad de Granada la provisión de fondos para la investigación. También agradecemos a los revisores anónimos sus sugerencias y comentarios en el proceso de evaluación.

Referencias

[1] Belsley, D.A. (1982). "Assessing the presence of harmful collinearity and other forms of weak data through a test for signal-to-noise". *Journal of Econometrics*, 20, 211–253.

[2] García, C.B.; García, J.; López Martín, M.M. y Salmerón, R. (2015). "Collinearity: revisiting the variance inflation factor in ridge regression". *Journal of Applied Statistics*, 42(3), 648–661.

[3] García, C.; Salmerón, R.; García, J. y López, M.M. (2016). "Estimación cresta y alzada: selección de k y λ a partir del coeficiente de correlación". *Anales de Economía Aplicada*, XXX, 949–955.

[4] García, J.; Salmerón, R.; García, C. y López Martín, M.M. (2016). "Standardization of Variables and Collinearity Diagnostic in Ridge Regression". *International Statistical Review*, 84(2), 245–266.

[5] Marquardt, D.W. (1970). "Generalized inverses, ridge regression, biased linear estimation and nonlinear estimation". *Technometrics*, 12(3), 591–612.

[6] McDonald, G.C. (2010). "Tracing ridge regression coefficients". *Wiley Interdisciplinary Reviews: Computational Statistics*, 2, 695–703.

[7] Novales, A. (1993). *Econometría*. McGraw-Hill, 2ª edición.

[8] Novales, A.; Salmerón, R.; García, C.; García, J. y López, M.M. (2015). "Tratamiento de la multicolinealidad aproximada mediante variables ortogonales". *Anales de Economía Aplicada*, XXIX, 1212–1227.

[9] O'Brien, R.M. (2007). "A caution regarding rules of thumb for variance inflation factors". *Quality and Quantity*, 41, 673–690.

[10] Salmerón, R.; García, C.; García, J. y López, M. M. (2017). "The raise estimators. Estimation, inference and properties". *Communications in Statistics-Theory and Methods*, 46(13), 6446–6462.

[11] Spanos, A. y McGuirk, A. (2002). "The problem of near-multicollinearity revisited: erratic vs systematic volatily". *Journal of Econometrics*, 108 (2), 365–393.

[12] Zhang, J. e Ibrahim, M. (2005). "A simulation study on SPSS ridge regression and ordinary least squares regression procedures for multicollinearity data". *Journal of Applied Statistics*, 32(6), 571–588.