Revista Venezolana de Análisis de Coyuntura, 2006, Vol. XII, No. 1 (ene-jun), pp. 283-296
recibido: 09-11-05 / arbitrado: 17-05-06

COLINEALIDAD Y MÍNIMOS CUADRADOS PONDERADOS

Gustavo Ramírez Valverde[1]
Benito Ramírez Valverde[2]
COLEGIO DE POSTGRADUADOS, MÉXICO

Resumen:

Dentro de las técnicas econométricas, uno de los métodos más usuales es el método de mínimos cuadrados ordinarios (MCO). La inferencia basada en este método tiene como uno de los supuestos básicos, la homogeneidad de varianzas. En caso de no cumplirse este supuesto, el estimador de MCO es insesgado pero ineficiente. Uno de los métodos alternativos de análisis cuando no se cumple este supuesto, es el método de mínimos cuadrados ponderados (MCP), que proporciona el mejor estimador linealmente insesgado (suponiendo que la forma funcional de la varianza es conocida), sin embargo, el condicionamiento de la matriz de información se ve afectado por las magnitudes de los ponderadores. De tal forma que pueden presentarse los problemas asociados a la colinealidad, aún cuando la matriz con las variables explicativas esté bien condicionada. En el presente trabajo se estudia la influencia de los ponderadores en el condicionamiento de la matriz de información, se muestra que los métodos de diagnóstico basados en la matriz con las variables explicativas podrían ser inadecuados, así como las soluciones que se derivan de esta.

Palabras claves: Heterogeneidad de varianzas, mínimos cuadrados ponderados, modelos econométricos, colinealidad.

INTRODUCCIÓN

El modelo uniecuacional de regresión es una de las herramientas estadísticas más comúnmente usadas en los modelos econométricos, la cual utiliza generalmente, estimación de mínimos cuadrados ordinarios (MCO). Para realizar inferencias con los estimadores MCO se requiere se cumplan una serie de supuestos básicos.

Cuando estos supuestos no se cumplen, se presentan algunos problemas en las propiedades de los_estimadores MCO; en particular cuando no se tienen varianzas homogéneas, el estimador de MCO sigue siendo insesgado y consistente, pero ya no es eficiente ni siquiera asintóticamente (Gujarati, 2004,

---

[1] gramirez@colpos.mx ; [2] bramirez@colpos. mx

Greene 1999). Bajo estas características, ya no es el mejor estimador lineal e insesgado (MELI).

Cuando la matriz de varianzas-covarianzas es singular, esto es, cuando las columnas de la matriz de varianzas-covarianzas son linealmente dependientes, no existe una solución única para el estimador de MCO, a este problema se le denomina colinealidad exacta. Cuando nos acercamos a la situación anterior, es decir, cuando las columnas de la matriz de varianzas-covarianzas están cerca de ser linealmente dependientes, el estimador de MCO existe y es MELI, pero presentan los siguientes problemas inferenciales (Gujarati, 2004; Greene, 1999; Tu *el al*, 2005; Mela *et al*, 2002):

1. Varianzas de gran magnitud
2. Poca potencia de las pruebas de hipótesis
3. Inestabilidad en el signo de los parámetros estimados.
4. La esperanza de la norma del vector de parámetros estimados es grande.
5. Sesgo al omitir una variable colineal

Esta situación es definida como el problema de colinealidad. Mela *et al*, (2002) califica este problema como uno de los mas endémicos que se han presentado en estudios empíricos sobre mercados, para dimensionar su magnitud, en una búsqueda de la revista "Applied Economics" (que usa Infotrac) obtuvo 221 resultados desde 1991.

Una de las técnicas más usadas para remediar los problemas de heterogeneidad de varianzas es la estimación de Mínimos Cuadrados Ponderados (MCP), que bajo ciertas condiciones genera estimadores MELI; sin embargo, el condicionamiento de la matriz de información se ve afectado por las magnitudes de los ponderadores. De tal forma que pueden presentarse los problemas asociados a la colinealidad, aún cuando las variables explicativas no son linealmente dependientes.

En el presente trabajo se estudia la influencia de los ponderadores en el condicionamiento de la matriz de información, se muestra que los métodos de diagnóstico basados en la matriz de las variables explicativas podrían ser inadecuados, así como las soluciones que se derivan de estos. Este error no lo hace evidente el software estadístico comúnmente empleado, que solo pide las variables explicativas para el diagnóstico de colinealidad, pero a la hora de realizar MCP, solo pide el ponderador, sin tomar en cuenta que el posible diagnóstico de colinealidad debería ser ahora en la matriz de información y no en la matriz con las variables explicativas.

MATERIALES Y MÉTODOS

El modelo de regresión

El modelo de regresión puede escribirse

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \text{......} (1)$$

donde **y** es un vector de la variable respuesta, **X** es una matriz n x (p+1) con el intercepto y las p variables, β es el vector de parámetros y ε es el vector de errores que tiene distribución Normal n-variada con vector de medias 0 y matriz de varianzas y covarianzas $\mathbf{I}\sigma^2$, con **I** la matriz identidad nxn.

El estimador de MCO es: $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$, con $\mathrm{var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^T\mathbf{X})^{-1}\sigma^2$

El problema de colinealidad

Aplicando la descomposición espectral a la matriz $\mathbf{X}^T\mathbf{X}$ se tiene:

$$\mathbf{X}^T\mathbf{X} = \Gamma\Lambda\Gamma^T,$$

donde $\Gamma = (e_1, e_2 \ldots e_{p+1})$ tiene los autovectores de la matriz **X** y $\Lambda = \mathrm{diag}(\lambda_i)$ con $\lambda_1 > \lambda_2 > \ldots > \lambda_{p+1}$) los autovalores de la matriz $\mathbf{X}^T\mathbf{X}$, entonces el estimador de mínimos cuadrados ordinarios (MCO) puede escribirse como:

$$\hat{\boldsymbol{\beta}} = \Gamma\Lambda^{-1}\Gamma^T\mathbf{X}^T\underline{\mathbf{y}}, \text{ con}$$

$$\mathrm{var}(\hat{\boldsymbol{\beta}}) = \Gamma\Lambda^{-1}\Gamma^T\sigma^2, \text{ y}$$

$$V(\hat{\beta}_j) = \sigma^2 \sum_{i=1}^{P+1} \frac{e_{ij}}{\lambda_i}, \qquad (2)$$

donde e_{ij} es el elemento i del vector e_j. El problema de colinealidad exacta se presenta cuando la matriz es singular y al menos un λ_i es cero. Esto ocurre cuando existen dependencias lineales entre las columnas de **X** y se dice que el estimador de MCO no existe. A medida que la matriz $\mathbf{X}^T\mathbf{X}$ se acerca a la singularidad, se está cercano a la colinealidad exacta, esto es, existen casi dependencias lineales entre las columnas de **X** y al menos un λ_i está cercano a cero, en este caso se presentan problemas inferenciales. La varianza del estimador de mínimos cuadrados de β está dado por:

$$\mathrm{var}(\hat{\boldsymbol\beta}) = E\{ [\beta - E(\hat{\boldsymbol\beta})] [\beta - E(\hat{\boldsymbol\beta})]^T \} = (\mathbf{X}^T\mathbf{X})^{-1} \sigma^2$$

Tomando la traza en ambos lados:

$$\mathrm{tr}[\ E\{ [\beta - E(\hat{\boldsymbol\beta})] [\beta - E(\hat{\boldsymbol\beta})]^T \}] = E\{\mathrm{tr}[\ [\beta - E(\hat{\boldsymbol\beta})] [\beta - E(\hat{\boldsymbol\beta})]^T] \}$$

$$= E\ \{ [\beta - E(\hat{\boldsymbol\beta})]^T [\beta - E(\hat{\boldsymbol\beta})] \ \}$$

$$= E[\hat{\boldsymbol\beta}^T \hat{\boldsymbol\beta}\ ] - E(\hat{\boldsymbol\beta})^T E(\hat{\boldsymbol\beta}) = \mathrm{tr}(\mathbf{X}^T\mathbf{X})^{-1} \sigma^2$$

pero, $E(\hat{\boldsymbol\beta})^T E(\hat{\boldsymbol\beta}) \geq 0$ entonces:

$$E[\hat{\boldsymbol\beta}^T \hat{\boldsymbol\beta}\ ] \geq \mathrm{tr}(\mathbf{X}^T\mathbf{V}\mathbf{X})^{-1} \sigma^2 = \mathrm{tr}(\Lambda)\sigma^2 = \sigma^2 \sum_{i=1}^{p+1} \frac{1}{\lambda_i}$$

De la ecuación anterior se deduce que el valor esperado de la traza cuadrada tiende a ser grande cuando un λ_i tiende a cero.

En la ecuación 2, se puede observar que cuando existe colinealidad, la varianza tiende a infinito. Varianzas grandes afectan las predicciones en algunas regiones y a la potencia de las pruebas estadísticas (Mandel, 1982).

Detección de colinealidad

En regresión lineal, han sido propuestos diversos métodos para detectar colinealidad. En este trabajo se utilizaran los más recomendados por la literatura especializada (Tu et. al. 2005, Belsley, 1991), los cuales se describen a continuación:

1) **El autovalor más pequeño de la matriz $\mathbf{X}^T\mathbf{X}$**. Cuando el mínimo autovalor es cercano a cero, se dice que el problema de colinealidad existe, sin embargo, es un problema de grado: a medida que sea más cercano a cero, el problema es más severo.

2) **El número condición de la matriz $\mathbf{X}^T\mathbf{X}$ (k_X)**. Sin pérdida de generalidad, consideramos las variables explicativas estandarizadas a longitud unitaria, entonces el número condición de la matriz $\boldsymbol{X}^T\boldsymbol{X}$ es definido por:

$$k_X = \left(\frac{\lambda_1}{\lambda_p}\right)^{.5}$$

Donde λ_1 y λ_p son los autovalores máximo y mínimo de la matriz $\mathbf{X}^T\mathbf{X}$ respectivamente.

Entonces: Si $k_X > 30$, existen problemas de colinealidad. A mayor número condición más severo es el problema.

3) **Los Indices de Condición** (C_j) (Stewart, 1987)

$$C_i = \left(\frac{\lambda_1}{\lambda_i}\right)^{.5}$$

Es definida como el i-ésimo índice condición, donde λ_i es el i-ésimo autovalor de la matriz $\boldsymbol{X}^T\boldsymbol{X}$ y λ_1 es el autovalor máximo. El criterio empleado para la detección de problemas de colinealidad sigue la misma regla que el número condición.

4) **Proporción de varianzas** (p_{ij}) **(**Silvey, 1969**)**

Donde p_{ij} es la proporción de la varianza de $\hat{\beta}_j$ atribuido al autovalor λ_i y se calcula con:

$$p_{ij} = \frac{e_{ij}^2 \, / \, \lambda_i}{\sum_{k=1}^{p+1} \frac{e_{kj}^2}{\lambda_k}}$$

La matriz de proporción de varianza puede ser construida y se presenta en el cuadro 1. Entonces un criterio de diagnóstico de colinealidad es, si existe un autovalor pequeño (en relación con el máximo), que es responsable de al menos dos proporciones grandes, entonces existe evidencia de problemas de colinealidad, y las variables explicativas asociadas son involucradas en una relación colineal.

Mínimos cuadrados ponderados

El modelo de regresión puede ser extendido a situaciones donde las observaciones no son independientes o en presencia de Heterocedasticidad de varianzas. Este modelo puede escribirse como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ , con}$$

$$\varepsilon \sim N(0, \sigma^2 V) \text{ ....(3)}$$

El Cuadro 1 presenta en su primera columna los autovalores ordenados y la matriz de proporción de varianzas, donde cada columna desglosa la proporción de varianza atribuido a cada autovalor.

Cuadro 1. Autovalores ordenados y la matriz de proporción de varianzas

Autovalores ordenados	*Proporción*			
	var($\hat{\beta}_1$)	*var*($\hat{\beta}_2$)	...	*var*($\hat{\beta}_{p+1}$)
λ_1	p_{11}	p_{12}	...	$p_{1\,p+1}$
λ_2	p_{21}	p_{22}	...	$p_{2\,p+1}$
.				
.				
.				
λ_{p+1}	$p_{p+1\,1}$	$p_{p+1\,2}$	...	$p_{p+1\,p+1}$

Donde **V** es una matriz simétrica y definida positiva (si la matriz **V** es la identidad, entonces este modelo se reduce al modelo lineal de la ecuación 1). Un método comúnmente usado de estimación en este modelo, es el de mínimos cuadrados generalizados, y consiste en encontrar una transformación tal que, el modelo resultante cumpla con las especificaciones del modelo lineal dado por la ecuación 1. La transformación del estimador de mínimos cuadrados generalizados (EMCG) se basa en el siguiente lema:

lema: Si la **A** es simétrica y definida positiva, entonces existe una matriz P tal que $\mathbf{A} = \mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T$

La matriz **P** es conocida como la matriz raiz cuadrada de **A**. Nótese que la matriz **V** asociada al modelo descrito en la ecuación 3, es simétrica y definida positiva, entonces existe **P** tal que $\mathbf{V} = \mathbf{P}^T\mathbf{P}$. Si se premultiplica $\mathbf{P}^{-1}$ a ambos lados del modelo en la ecuación 3 se tiene:

$$\mathbf{P}^{-1}\,\mathbf{y} = \mathbf{P}^{-1}\,\mathbf{X}\beta + \mathbf{P}^{-1}\,\varepsilon \qquad \text{..(4)}$$

si se construyen las siguientes variables:

$\mathbf{y}^* = \mathbf{P}^{-1}\,\mathbf{y}$; $\mathbf{X}^* = \mathbf{P}^{-1}\,\mathbf{X}$; y $\varepsilon^* = \mathbf{P}^{-1}\,\varepsilon$, entonces la ecuación 4 se reduce a:

$$\mathbf{y}^* = \mathbf{X}^*\beta + \varepsilon^*, \text{ donde:}$$

$$E(\varepsilon^*) = E(\mathbf{P}^{-1}\varepsilon\,) = 0\text{ , y}$$

$$\mathbf{var}(\varepsilon^*) = \mathbf{var}(\mathbf{P}^{-1}\varepsilon^*\,) = \mathbf{P}^{-1}\,\mathbf{var}(\varepsilon\,)\,(\mathbf{P}^{-1})^T = \sigma^2\mathbf{P}^{-1}\,\mathbf{V}\,(\mathbf{P}^{-1})^T = \sigma^2$$

El estimador de MCO del modelo transformado es:

$$\hat{\boldsymbol{\beta}}_{mcg} = (\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\,\mathbf{X}^{*T}\mathbf{y}^*$$

con:

$$\mathbf{var}(\hat{\boldsymbol{\beta}}_{mcg}) = (\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\sigma^2 = (\mathbf{X}^T\mathbf{V}\mathbf{X})^{-1}\sigma^2$$

Este estimador es conocido como el Estimador de Mínimos Cuadrados Generalizados (MCG) y es el MELI del vector de parámetros β. Esto se sigue fácilmente del hecho de que el modelo transformado cumple con las condiciones del modelo de la ecuación 1.

El estimador de Mínimos Cuadrados Ponderados es un caso particular del MCG donde $\mathbf{V}$ = diag(f($\mathbf{x}_i$)), además, $\mathbf{x}_i$ es un vector conteniendo las variables explicativas asociadas a la variable respuesta en la observación i, y f(.) es una función conocida, esto es, supone un modelo lineal donde las observaciones son independientes, pero se tienen heterogeneas y las varianzas dependen de las variables explicativas a través de la función f(.), entonces para la i-ésima observación se tendrá que :

$$\text{var}(\varepsilon_i) = \sigma^2 f(\mathbf{x}_i)$$

Las funciones f(.) más usadas son:

1) $f(\mathbf{x}_i) = (x_{ij})^2$

2) $f(\mathbf{x}_i) = (x_{ij})^{-2}$

Donde x_{ij} es el valor de la i-ésima observación de la variable explicativa j.

En el caso del estimador MCP, la matriz raíz cuadrada de **V** es **P** = diag(f($\mathbf{x}_i$)$^{-0.5}$). Entonces, para obtener el estimador de mínimos cuadrados ponderados se multiplica a la variable respuesta y a cada una de las variables explicativas (incluyendo el intercepto) por el recíproco de la raíz cuadrada de f($\mathbf{x}_i$).

Datos Analizados

En este trabajo se consideran dos casos construidos artificialmente, de tal manera que presentan situaciones en las cuales el diagnóstico de colinealidad basado en la matriz **X**, presenta resultados contradictorios.

Caso 1. Las ponderaciones inducen colinealidad

Este caso se construyó con un tamaño de muestra de 50 y tres variables explicativas. Primeramente se determinaron las variables explicativas, buscando que no presentaran problemas en el condicionamiento de la matriz **X** (el número condición de la matriz $\mathbf{X}^T\mathbf{X}$ fue de 10.53). Posteriormente se definió la variable respuesta mediante la ecuación:

$$y = 4 + 18 * X_1 + 8 * X_2 + 8 * X_3 + \varepsilon$$

donde

$$\varepsilon \sim N\,(0,\, 9 * X_1^2)$$

Este caso presenta varianzas heterogéneas que dependen de X_1, los datos se presentan en el anexo 1.

Caso 2. Las ponderaciones enmascaran la colinealidad

En el caso anterior, la matriz **V** con los ponderadores inducían colinealidad en la matriz de información de los estimadores de MCP, sin embargo no es la única posibilidad de resultados divergentes al hacer el diagnóstico basándose en la matriz ***X***.

Para ilustrar lo anterior se diseñó un ejemplo con un tamaño de muestra de 50 observaciones y dos variables explicativas. Donde se construyó la matriz ***X*** con problemas severos de colinealidad, los datos se presentan en el anexo 2.

RESULTADO Y DISCUSIÓN

Caso 1

El análisis estadístico se realizó utilizando el paquete estadístico SAS. La tabla de análisis de varianza, para el estimador de mínimos cuadrados ordinarios se presenta en el Cuadro 2. En esta tabla se puede observar que el modelo resultó significativo (0.0001), con un R^2 aceptable.

Cuadro 2. Tabla de Análisis de Varianza (MCO)

Fuente de Variación	*GL*	*Suma de Cuadrados*	*Cuadrado Medio*	*F*	*Prob>F*
Modelo	*3*	*40126173.781*	*13375391.260*	*81.143*	*0.0001*
Error	*46*	*7582514.0244*	*164837.26140*		
Total	*49*	*47708687.806*			

$(CME)^{.5}$= 406.00155; CV= 18.61427; R^2 = 0.8411; R^2 Aj. = 0.8307

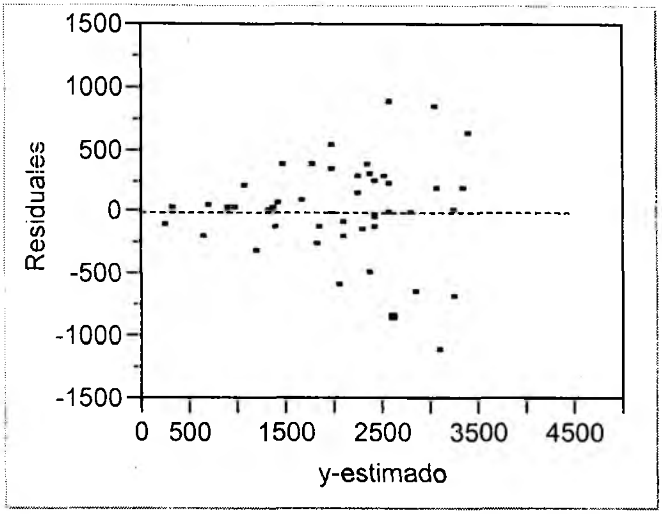
Si el investigador deseara verificar la presencia de colinealidad en su modelo, realizaría un diagnóstico como el presentado en el cuadro 3.

Cuadro 3. Diagnóstico de colinealidad

Número	*Autovalor*	*Número Condición*	*Proporción Varianza Intercepto*	*Proporción Varianza X1*	*Proporción Varianza X3*	*Proporción Varianza X2*
1	3.76215	1.00000	0.0032	0.0093	0.0071	0.0048
2	0.13354	5.30773	0.0033	0.6671	0.2998	0.0028
3	0.07041	7.30983	0.0252	0.2552	0.5090	0.5098
4	0.03390	10.53481	0.9683	0.0685	0.1840	0.4826

De este cuadro se desprende que no existen problemas serios de colinealidad, como puede notarse en los valores obtenidos en el número condición, por lo que el investigador concluiría que este problema no se encuentra presente y continuaría su análisis verificando la homogeneidad de varianza, empleando gráfica de residuales. Los residuales evidenciaron (tal como se esperaba, por la forma en que los datos fueron construidos), que existe problemas con la homogeneidad de las varianzas, como se muestra en la figura 1.

Figura 1. Gráfica de Residuales



Ante el problema de falta de homogeneidad de varianza, se presenta la alternativa del uso de estimadores de mínimos cuadrados ponderados, asumiendo que la varianza de los errores es proporcional a X_1^2 y realizando nuevamente el análisis de varianza, se obtienen los siguientes resultados.

Cuadro 4. Análisis de varianza usando mínimos cuadrados ponderados

Fuente de Variación	*GL*	*Suma de Cuadrados*	*Cuadrado Medio*	*F*	*Prob>F*
Modelo	3	4454.68452	1484.89484	174.839	0.0001
Error	46	390.67514	8.49294		
Total	49	4845.35966			

$(CME)^{.5}$= 2.91426; CV= 14.47590; R^2 = 0.9194; R^2 Aj. = 0.9141

En el cuadro 5, se puede observar que los valores obtenidos no son del todo satisfactorios, a pesar de que los estimadores son MELI, ya que se realizó el análisis adecuado para remediar las varianzas heterogéneas de acuerdo a la construcción del ejemplo. Esto se manifiesta en la magnitud de los parámetros estimados, el signo equivocado en uno de sus estimadores, el gran tamaño de sus varianzas y la falta de significancia en algunos de ellos. Todos los problemas presentes son síntomas clásicos de colinealidad, lo cual contradice el resultado obtenido en el diagnóstico de colinealidad basado en la matriz **X**.

Cuadro 5. Pruebas de hipótesis sobre los parámetros individuales

Variable	*GL*	*Estimador*	*Error estándar*	*Valor de T*	*Prob > |T|*
Intercepto	1	17.615098	0.47209870	37.312	0.0001
X_1^{-1}	1	-222.446981	167.31049654	-1.330	0.1902
X_2/X_1	1	77.291677	169.59092761	0.456	0.6507
X_3/X_1	1	20.818627	6.73648305	3.090	0.0034

El diagnóstico de colinealidad basado en la matriz **X**, no es el adecuado cuando se usan los estimadores de MCP, ya que la matriz de información de estos estimadores no es la matriz $\mathbf{X}^T\mathbf{X}$, sino la matriz $\mathbf{X}^T\mathbf{V}\mathbf{X}$, donde **V** es una matriz diagonal, asociada a la varianza de los errores, por lo que el diagnóstico apropiado no debe basarse en la matriz **X**, sino en la matriz de información $\mathbf{X}^T\mathbf{V}\mathbf{X}$. El cuadro 6, muestra el resultado del diagnóstico sobre esta matriz, en el cual se aprecia claramente que existen problemas de colinealidad, ya que el número condición es 69.81125 y puede apreciarse que el autovalor más pequeño tiene asociadas dos proporciones de varianza grandes.

Cuadro 6. Diagnóstico de Colinealidad de la matriz de información X^TVX

Número	*Autovalor*	*Índice de Condición*	*Proporción Varianza Intercepto*	*Proporción Varianza X_1^{-1}*	*Proporción Varianza X_2/X_1*	*Proporción Varianza X_3/X_1*
1	*3.10045*	*1.0000*	*0.0116*	*0.0001*	*0.0002*	*0.0003*
2	*0.8958*	*1.8604*	*0.8103*	*0.0000*	*0.0000*	*0.0001*
3	*0.00311*	*31.54991*	*0.0197*	*0.0045*	*0.2437*	*0.5708*
4	*0.0006362*	*69.81125*	*0.1584*	*0.9954*	*0.7560*	*0.4288*

Todos los criterios para diagnóstico de colinealidad definidos anteriormente para regresión lineal con MCO pueden ser empleados con MCP. Para estos estimadores, la matriz diseño ponderada se obtiene con la ecuación $\mathbf{X}_W = \mathbf{V}^{1/2}\,\mathbf{X}$, donde $\mathbf{V} = \text{diag}(v_i)$, con $v_i = f(\mathbf{x}_i)$, es más importante que **X** en determinar el condicionamiento de la matriz de información. Todos los criterios descritos anteriormente pueden ser usados con la substitución de la matriz diseño **X** por la matriz diseño ponderada $\mathbf{X}_W$. En este caso, nuestro interés es en la dependencia linear en la matriz diseño ponderada.

CASO 2

Los resultados del diagnóstico de colinealidad basado en la matriz **X** se presentan en el cuadro 7. Donde se observa un número condición de 58.3363, indicando colinealidad en la matriz **X**.

Cuadro 7. Diagnóstico de colinealidad basado en la matriz X

Número	*Autovalor*	*Número Condición*	*Proporción Varianza Intercepto*	*Proporción Varianza X_1*	*Proporción Varianza X_2*
1	1.99941	1.00000	0.0000	0.0003	0.0003
2	1.00000	1.41401	0.9992	0.0000	0.0000
3	0.0005875	58.33632	0.0008	0.9997	0.9997

Si asumimos que las varianzas son inversas a los pesos asignados (anexo 2) y usamos nuevamente mínimos cuadrados ponderados, al construir la matriz de información $\mathbf{X}^T\mathbf{V}\mathbf{X}$ de estos estimadores y realizar un diagnóstico de colinealidad se puede observar lo siguiente :

Cuadro 8. Diagnóstico de colinealidad basado en la matriz de información X^TVX

Número	*Autovalor*	*Índice de Condición*	*Proporción Varianza Intercepto*	*Proporción Varianza X_1*	*Proporción Varianza X_2*
1	1.53362	1.00000	0.2311	0.1286	0.1286
2	1.03030	1.22005	0.0000	0.3936	0.3936
3	0.43608	1.87533	0.7689	0.4778	0.4778

Como puede apreciarse en el cuadro 8, contrario a lo acontecido en el caso 1, la matriz información del estimador de mínimos cuadrados ponderados no presenta problemas de colinealidad, por lo que en este caso la matriz **V** con los ponderadores no induce colinealidad, sino que corrige el problema manifestado en la matriz **X**.

CONCLUSIONES

El diagnóstico de colinealidad en la matriz **X**, no es el más apropiado cuando se usan mínimos cuadrados ponderados. Se recomienda usar los mismos métodos de diagnóstico de colinealidad que con MCO, pero con los estimadores MCP, se debe usar la matriz de información $\mathbf{X}^T\mathbf{V}\mathbf{X}$.

La matriz diagonal **V** influye fuertemente en la presencia de colinealidad, ya sea disminuyendo o agudizando este problema por lo que es necesario verificar siempre el papel que está jugando esta matriz en nuestro modelo.

REFERENCIAS BIBLIOGRÁFICAS

Belsley, D. A. (1991), *Conditioning Diagnostics: Collinearity and Weak Data in Regression*, New York: Wiley,

Greene W. H (1999), *Análisis Econométrico*, Tercera Edición, Prentice Hall.

Gujarati D. N. (2004), Basic Econometrics, Cuarta Edición, McGraw-Hill, México.

Mandel, J. (1982), "Use of the Singular Value Decomposition in Regression Analysis", *The American Statistician*, 1:15-24.

Mela C. F.; Kopalle P. K. (2002), "The impact of collinearity on regression analysis: the asymmetric effect of negative and positive correlations", *Applied Economics*, 34: 667-677.

Silvey, S. D. (1969), "Multicollinearity and Imprecise Estimation", *Journal of the Royal Statistical Society*, Series B. 31:539-552.

Sewart G. W. (1987), "Collinearity and Least Squares Regression", *Statistical Science*, 1:68-100.

Tu.Y-K; Kellett M. Clerehugh V.; Gilthorpe M. S. (2005), "Problems of correlations between explanatory variables in multiple regression analyses in the dental literature", *British Dental Journal*, 7: 457-466.

ANEXO 1

OBS	Y	X1	X2	X3
1	155.41	1.819	139.486	0.71376
2	449.55	17.193	174.398	0.51074
3	353.11	20.954	78.643	0.68537
4	763.37	37.545	105.835	0.64101
5	930.28	47.602	121.327	0.90806
6	991.72	50.380	119.707	0.69612
7	1398.49	65.194	170.824	102.266
8	1508.73	66.391	167.650	0.50753
9	1273.06	67.014	67.474	0.44243
10	1333.02	68.192	133.333	0.66332
11	897.70	75.197	79.412	0.72579
12	1273.64	78.356	102.018	0.37566
13	1774.60	84.268	155.736	0.45149
14	1572.26	91.221	170.672	0.63151
15	1731.83	91.285	175.642	0.71941
16	1863.28	94.753	63.624	0.63633
17	1722.84	95.443	152.142	0.55785
18	2160.84	99.332	124.988	0.78038
19	1461.74	102.523	175.377	0.52438
20	2024.43	104.690	176.333	0.38932
21	1893.11	108.087	161.030	0.46492
22	2525.74	111.112	187.363	0.39325
23	2325.36	114.725	115.605	0.88358
24	1971.45	118.754	86.137	0.44684
25	2165.70	119.345	175.121	0.68617
26	2512.12	124.381	68.000	0.60097
27	750.94	126.793	155.389	0.59181
28	2295.24	128.066	177.041	106.958
29	2416.39	128.355	113.559	0.26785
30	2563.86	131.114	191.117	0.52010
31	2668.02	131.911	139.517	0.77780
32	1884.96	133.658	126.180	0.52894
33	1755.20	137.814	165.974	0.34680
34	2790.65	138.748	143.466	0.56359
35	3466.40	141.288	134.548	0.02305
36	2670.02	141.350	122.122	0.93743
37	2797.37	146.467	131.063	0.69541
38	2363.96	148.324	65.581	0.13815
39	2793.41	154.734	149.368	0.61156
40	3883.76	170.466	147.066	0.60878
41	4017.78	174.901	218.703	0.60537
42	3247.45	175.196	175.429	0.44787
43	3260.45	177.063	127.159	0.60806
44	1989.39	180.435	115.786	0.51273
45	2569.50	182.304	160.574	0.81168
46	2198.17	184.348	39.739	0.56949
47	3537.13	185.454	174.367	0.97077
48	3645.18	187.919	184.490	0.77695
49	4204.07	189.301	172.141	0.40331
50	4253.81	215.129	191.128	111.861

ANEXO 2

OBS	X1	X2	V
1	-98	-100	0.00707
2	-98	-100	0.00707
3	-98	-100	0.00707
4	-98	-100	0.00707
5	-98	-100	0.00707
6	-98	-100	0.00707
7	-98	-100	0.00707
8	-98	-100	0.00707
9	-98	-100	0.00707
10	-98	-100	0.00707
11	-98	-100	0.00707
12	-98	-100	0.00707
13	-98	-100	0.00707
14	-98	-100	0.00707
15	-98	-100	0.00707
16	-98	-100	0.00707
17	-98	-100	0.00707
18	-2	2	0.49497
19	-2	2	0.49497
20	-2	2	0.49497
21	-2	2	0.49497
22	-2	2	0.49497
23	-2	2	0.49497
24	-2	2	0.49497
25	-2	2	0.49497
26	-2	2	0.49497
27	-2	2	0.49497
28	-2	2	0.49497
29	-2	2	0.49497
30	-2	2	0.49497
31	-2	2	0.49497
32	-2	2	0.49497
33	-2	2	0.49497
34	100	98	0.00707
35	100	98	0.00707
36	100	98	0.00707
37	100	98	0.00707
38	100	98	0.00707
39	100	98	0.00707
40	100	98	0.00707
41	100	98	0.00707
42	100	98	0.00707
43	100	98	0.00707
44	100	98	0.00707
45	100	98	0.00707
46	100	98	0.00707
47	100	98	0.00707
48	100	98	0.00707
49	100	98	0.00707
50	100	98	0.00707