NOTAS SOBRE COLINEALIDAD EN EL MODELO DE REGRESIÓN LINEAL DE TRES VARIABLES

FERNANDO CORTÉS C.*
Y ROSA MA. RUBALCAVA
El Colegio de México

INTRODUCCIÓN

ESTE TRABAJO INVESTIGA LAS PECULIARIDADES que caracterizan la estimación de un modelo de regresión parcial del tipo:

(1) $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + U_i \qquad (i = 1, 2, \ldots, n)$

en el que se ha constatado una estrecha relación lineal entre las variables explicativas.

A modo de ejemplo, podría considerarse el ajuste de una ecuación del nivel de precios en función de los salarios y la cantidad de dinero.

Una situación típica que también podría responder al caso esbozado sería la de la estimación de una función en donde una de las variables fuese el tamaño o número de las unidades de observación. Tal sería el caso de una función consumo agregada en que X_{2i} podría representar el nivel del ingreso disponible, X_{3i} la población y Y_i el valor del consumo personal.

La existencia de un cierto grado de independencia entre las variables explicativas permite establecer que:

$$\frac{\partial Y}{\partial X_2} = \beta_2 \qquad\qquad \frac{\partial Y}{\partial X_3} = \beta_3$$

Vale decir, el coeficiente de regresión β_2 indicaría el efecto que tendría un cambio unitario en X_2 sobre la variable Y, bajo el supuesto de que X_3 se mantiene constante. Para β_3 se puede realizar una interpretación análoga.

Pero si existiera una relación lineal entre X_2 y X_3, la interpretación de los coeficientes ya no sería tan directa.

En efecto, si

$$X_{2i} = a + b X_{3i} + V_i$$

*Facultad Latinoamericana de Ciencias Sociales, sede México. (FLACSO-MEXICO); profesor en El Colegio de México.

y si la reemplazamos en (1), tendremos:

$$(2) \quad Y_i = (\beta_1 + \beta_2\, a) + (\beta_3 + \beta_2\, b) X_{3i} + (\beta_2\, V_i + U_i)\ (i = 1, 2, \ldots, n)$$

y el impacto lineal de X_3 sobre Y será:

$$\frac{\partial\ Y}{\partial\ X_3} = \beta_3 + \beta_2\, b^1$$

Este resultado indica que, en la medida que haya una relación lineal estrecha entre X_2 y X_3, difícilmente podremos evaluar el efecto de una variable explicativa sobre Y, manteniendo la otra constante. Si la correlación entre X_2 y X_3 se aproxima a la unidad, entonces toda variación en X_3 será seguida por una en X_2 y viceversa, por ello nos veremos incapacitados para alterar los valores de una de las variables manteniendo la otra constante.

La derivada muestra que el coeficiente de regresión de X_3 en la ecuación (2) mezcla el impacto que tiene X_3 sobre $Y\ (\beta_3)$, con el que tiene X_2 sobre Y a través de X_3 $(\beta_2\, b)$.

En métodos no experimentales (como lo es el subyacente al modelo de regresión) el control de variables sólo se puede realizar a través de la consideración explícita de ellas.[2] En modelos de regresión es usual incluir una tercera variable al estudiar la relación entre otras dos, ya sea a través de su incorporación en un modelo parcial (tal como el de la ecuación (1)) o bien, establecer uno donde las variables se definan como razones. Este último procedimiento es de uso bastante frecuente cuando la tercera variable es la población o el nivel de precios.

Así la ecuación:

$$(3) \quad \frac{Y_i}{X_{3i}} = \delta_1 + \delta_2 \frac{X_{2i}}{X_{3i}} + U_i \quad (i = 1, 2, \ldots, n)$$

en que Y_i/X_{2i} es el consumo per cápita y X_{2i}/X_{3i} es el ingreso per cápita, estaría conformada por una variable explicada y una explicativa, definidas ambas como razones entre variables, lo que supuestamente permitiría controlar el efecto de la población sobre el consumo y el ingreso.

Pero, si nuestro pensamiento teórico nos conduce a suponer que sobre el nivel de consumo actúan tanto el crecimiento económico como el demográfico y que sus efectos son aditivos, entonces el modelo correctamente especificado será el representado por la ecuación (1). Sin embargo, es probable que al intentar ajustar este modelo teórico nos encontremos con una colinealidad estrecha entre el

[1] Dado que $b = r_{x_2 x_3} \frac{Sx_2}{Sx_3}$; si no hay relación lineal entre x_2 y x_3 $(r_{x_2 x_3} = 0)$ entonces $\frac{\partial y}{\partial x_3} = \beta_3$, es decir el coeficiente de regresión que afecta a x_3 en la ecuación (2) mide el impacto que tiene esta variable sobre y. En la medida que la relación lineal sea más estrecha b se alejará más de cero y $\frac{\partial y}{\partial x_3}$ recogerá el efecto combinado de x_2 y x_3.

[2] Stinchcombe, Arthur. *La construcción de Teorías Sociales*. Ediciones Nueva Visión, 1970, pág. 49.

ingreso personal disponible y la población, lo cual introduce una serie de complicaciones en la estimación.

Ante la dificultad de medir el efecto de una variable explicativa manteniendo la otra constante –debido a la fuerte correlación– se puede intentar controlar el efecto de la población a través de una transformación en las variables que permita extraer la "contaminación" inducida por ella. Esto se puede lograr dividiendo la iguadad (1) entre X_3.

$$(4) \quad \frac{Y_i}{X_{3i}} = \beta_1 \cdot \frac{1}{X_{3i}} + \beta_2 \frac{X_{2i}}{X_{3i}} + \beta_3 + U_i \cdot \frac{1}{X_{3i}} \quad (i = 1, 2, \ldots, n)$$

Al comparar (3) con (4) notamos tres diferencias importantes:

a) La variable $1/X_3$ aparece en la última ecuación y no en la primera. Esto quiere decir que si el modelo correctamente especificado es el que representa la igualdad (1), al ajustar (3) cometemos un error de especificación lo que introduce un sesgo en los estimadores.[3]

b) Si el término estocástico de (1) es homocedástico,[4] la división entre X_3 llevaría a que $W_i = U_i/X_{3i}$ fuese heterocedástico.

Var $W_i = (1/X_{3i}^2)$.Var U_i bajo el supuesto que X_{3i} no es una variable aleatoria.

Pero si U_i tuviese la forma heterocedástica: $U_i = X_{3i}\ V_i$, donde $E\,(V^2) = \sigma^2$, entonces W_i sería homocedástica:

$$\text{Var } W_i = (1/X_{3i}^2) \quad \text{Var } U = \sigma^2$$

c) La constante (β_3) del modelo (4) permite estimar el efecto que tiene X_3 sobre Y (la población sobre el consumo).

En resumen, el modelo parcial (1) permite estimar el efecto de la variable X_2, controlando el de X_3 y viceversa, sólo en caso de que haya una relativa independencia lineal entre ambas. Si por el contrario, hubiese una fuerte colinealidad entre las variables explicativas –lo cual implica que sería muy difícil, en una situación no experimental, separar el efecto de una variable manteniendo constante la otra– se puede proceder a depurarlas, para lo cual las expresamos como razón. Pero, la manera adecuada de hacerlo sería a través de la ecuación de ajuste (4) que correspondería al modelo teórico expresado matemáticamente por la igualdad (1).

Ahora bien, si debemos emprender el proceso de estimación en una situación como la planteada, ¿cómo sabemos que el procedimiento utilizado para eliminar o disminuir la contaminación es realmente efectivo?

En los términos en que hemos presentado el problema, esta pregunta constituye otra manera de interrogarse acerca del efecto que tendría sobre la colineali-

3 Theil Henry. *Principles of Econometrics.* John Wiley. 1971, págs. 548-551 y Przeworski, Adam y Cortés, Fernando. Comparing Partial and Ratio Regression Models. *Political Methodology,* 1977.

4 Nos parece poco útil conservar a medias la grafía inglesa y escribir homoscedástico. Más aún tratándose de un neologismo, creemos que las variantes "homosedástico" y "homocedástico" dependerán del criterio del autor, hasta que el uso acuñe definitivamente la palabra en nuestra lengua.

dad la transformación no lineal aplicada en (1), ya que si la correlación entre las variables $1/X_3$ y X_2/X_3 de la igualdad (4) permanece igual o es mayor que la existente entre X_2 y X_3, no habremos conseguido depurar las variables y los problemas de estimación persistirán o se agravarán.

Pero antes de abordar el estudio comparativo de las correlaciones entre esos pares de variables procederemos a resumir de manera muy sintética las principales consecuencias que tiene en el proceso de estimación, la colinealidad estrecha.

COLINEALIDAD Y ESTIMACIÓN

El caso límite de colinealidad es aquel en el que la relación entre las dos variables explicativas es perfectamente lineal ($r_{23} = \pm 1$). Esta situación conduce a que sea imposible realizar las estimaciones punto e intervalo del modelo.

En efecto, si la correlación entre las variables explicativas es unitaria, la matriz de observaciones X no tendrá rango completo (su rango será dos en lugar de tres) por lo cual $(X'X)$ será una matriz singular y por consecuencia será imposible obtener.

$$\hat{B} = (X'X)^{-1} X'Y$$

y

$$\text{Var } \hat{B} = \sigma^2 \quad (X'X)^{-1}$$

La experiencia señala que en el proceso de ajuste de modelos lineales es relativamente poco frecuente enfrentarse con una situación de colinealidad perfecta. Por esto nuestro interés es enfocar el caso en que la relación lineal entre las variables explicativas es estrecha pero no perfecta.

Las consecuencias fundamentales de una correlación fuerte entre X_2 y X_3 se manifiestan en una precisión decreciente de las estimaciones: *a*) A mayor grado de colinealidad mayores serán las varianzas muestrales de los coeficientes de regresión. *b*) Lo cual lleva, a través de la covariación entre los estimadores mínimo cuadráticos ordinarios, a que si un estimador sobreestima su correspondiente parámetro el otro tenderá a subestimarlo.[5]

Las consecuencias de la colinealidad estrecha se manifiestan en la significación estadística de los coeficientes. En efecto una varianza grande (grande en relación al valor del estimador) tenderá a disminuir el valor del coeficiente *t* de Student calculado, lo que puede llevar a eliminar alguna variable del modelo aun cuando tenga un papel explicativo. Lo que acontece es que el modelo no puede captar su efecto debido al conjunto de datos disponible. Al decidir, erróneamente, eliminar una variable explicativa, cometemos un error de especificación lo que implica introducir un sesgo en los estimadores.

[5] Johnston J. Econometric Methods. MacGraw Hill, 1972, pág. 160.

Ahora bien, si encontrásemos una manera de estimar los parámetros del modelo de regresión parcial que implicara una caída en la correlación entre los regresores, probablemente[6] haríamos bajar la varianza de los estimadores con lo cual ganaríamos en la precisión de los mismos, y disminuiríamos el riesgo de cometer error de especificación. Justamente ésta será la idea que orientará la próxima sección.

Para finalizar este apartado, plantearemos de manera sintética la naturaleza del problema:

a) En primer lugar suponemos que la especificación correcta de nuestro pensamiento teórico está dada por el modelo de regresión parcial (1):

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + U_i \; (i = 1, 2, \ldots, n)$$

y además que entre X_2 y X_3 hay una correlación lineal estrecha.

b) Para evitar esta "contaminación" –cuando una de las variables es el tamaño de las unidades de observación– debería dividirse toda la ecuación entre una de ellas (por ej. X_3) generándose de este modo el modelo de ajuste (4) correspondiente al modelo teórico (1):

$$\frac{Y_i}{X_{3i}} = \beta_1 . \frac{1}{X_{3i}} + \beta_2 \frac{X_{2i}}{X_{3i}} + \beta_3 + W_i; \; (i = 1, 2, \ldots, n)$$

en que $W_i = U_i / X_{3i}$. Las variables serán expresadas en términos per cápita si las unidades son personas, por establecimientos si son empresas, etcétera.

c) Esta manera particular de "controlar" el efecto de una tercera variable será exitosa si realmente logra disminuir el tamaño de las varianzas de los estimadores.

Ahora bien, entre los factores de los cuales dependen estas varianzas queremos destacar la naturaleza homo o heterocedástica del término de error, y la correlación entre los regresores.

Lo que resta de este trabajo se dedicará principalmente al estudio del último de estos aspectos. También realizaremos algunas consideraciones sobre heterocedasticidad aun cuando sólo tendrán calidad de acotaciones marginales.

La correlación entre los regresores

Nuestro problema se reduce a encontrar una función que ligue la correlación entre $1/X_3$ y X_2/X_3 (que simbolizaremos por r) con la correlación entre X_2 y X_3 (simbolizada por r_{23}).

Para obtener tal función recurriremos a los desarrollos de Karl Pearson[7] que establecen que si definimos una variable Z como una relación entre otras dos

[6] La palabra probablemente se utiliza para señalar que el efecto resultante no sólo depende de la correlación entre los regresores.

[7] Pearson Karl. On a form of spurious correlation wich may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, No. 60.

$$Z = \frac{X_k}{X_j}$$

su promedio y varianza serán aproximadamente:

$$M\ [Z] \doteq \frac{\overline{X}_k}{\overline{X}_j} \quad [1 - V_k\ V_j\ r_{kj} + V_j^2]$$

$$\text{Var}\ [Z] \doteq \frac{\overline{X}_k^2}{\overline{X}_j^2} [V_k^2 - 2\ V_k\ V_j\ r_{kj} + V_j^2$$

donde $\overline{X}_k$ y $\overline{X}_j$ representan los promedios de las variables k y j respectivamente. V_k y V_j simbolizan los coeficientes de variación y r_{kj} el coeficiente de correlación entre X_k y X_j.

Estas fórmulas corresponden a una aproximación proveniente de un desarrollo en serie de funciones[8] en que no se han tomado en cuenta todos los términos de orden superior a dos, bajo la condición de convergencia:

$$(5) \quad \left(\frac{X_{ji} - \overline{X}_j}{\overline{X}_j} \right)^2 < 1 \text{ para todo } i \text{ y } j,$$

la cual establece que cada valor de la variable que se usó como denominador de Z debe ser estrictamente menor que el doble de la media de X_j.

Al aplicar suma sobre ambos lados de la desigualdad y despejar, la condición de convergencia asume la forma:

$$V_j^2 < 1$$

vale decir, que el cuadrado del coeficiente de variación de X_j debe ser menor que uno. Esta restricción, que asegura que la serie no diverja, es una condición necesaria pero no suficiente, ya que puede ser menor que uno aun cuando algún valor X_{ji} de la variable sea mayor que dos veces la media aritmética.

El cálculo del coeficiente de correlación entre las variables de la ecuación de ajuste (4) ($Z_1 = 1/X_{3i}$ y $Z_2 = X_{2i}/X_{3i}$) se puede realizar a través de:

$$n\,r\,S_1\ S_2 = \Sigma\ (Z_1 - \overline{Z}_1)\ (Z_2 - \overline{Z}_2)$$

donde r representa el coeficiente de correlación y S_1 y S_2 las desviaciones típicas de Z_1 y Z_2, respectivamente.

Después de una serie de desarrollos[9] se llega en definitiva a

8 Ver apéndice 1.
9 Ver apéndice 1.

(6) $$r^2 \doteq \frac{(V_3 - r_{23} V_2)^2}{V_2^2 - 2r_{23} V_2 V_3 + V_3^2}$$

donde V_3 y V_2 son los coeficientes de variación de X_3 y X_2 respectivamente y r_{23} su correspondiente coeficiente de correlación.

Esta fórmula sólo constituye una aproximación obtenida por medio de un desarrollo en serie de funciones en la que el tamaño del error que se cometa dependerá tanto del número de términos considerados como de la condición de convergencia dada por la desigualdad.

$$\left(\frac{X_{3i} - \bar{X}_3}{\bar{X}_3}\right)^2 < 1 \quad (i = 1, 2, \ldots, n)$$

que sólo es una especificación de (5).

En el caso de colinealidad perfecta entre las variables explicativas del modelo (1) r_{23}^2 será igual a 1 y, según (6)

$$r^2 = \frac{(V_3 - V_2)^2}{V_2^2 - 2 V_2 V_3 + V_3^2} = 1$$

Es decir, la transformación de las variables es inoperante, por cuanto no reduce el grado de relación entre los regresores del modelo de ajuste representado por la ecuación (4).

Podemos preguntar a continuación acerca del locus en que se cumple $r_{23}^2 = r^2$. Reemplazando esta última igualdad en la ecuación (6) y despejando convenientemente se llega a

$$r^3 - \frac{V_3}{2 V_2} r^2 - r + \frac{V_3}{2 V_2} = 0$$

si reemplazamos $V_3/2V_2$ por a, tendremos la ecuación cúbica

$$r^3 - ar^2 - r + a \doteq 0$$

donde conocemos dos de las tres soluciones: $r_1 = 1$ y $r_2 = -1$. De aquí que resulte sencillo obtener la tercera solución.

$$r_3 = a \doteq \frac{V_3}{2 V_2}$$

Sobre la base del conocimiento de estas características hemos procedido al estudio numérico de la igualdad (6). Los resultados alcanzados se encuentran en el apéndice 2 de los cuales hemos seleccionado algunos para exponerlos gráficamente en el apéndice 4.

En la gráfica debemos distinguir la zona formada por el triángulo OBC en que cualquier punto en ella se caracteriza por $r^2_{23} < r^2$, del área incluida en el interior del triángulo OBA donde $r^2_{23} > r^2$. Resulta evidente que sólo podremos mitigar los efectos de la colinealidad si al aplicar la transformación no lineal $(1/X_3)$ al modelo (1) nos ubicamos en un punto que esté dentro del triángulo inferior.

Las condiciones para que la correlación entre las variables X_2/X_3 y $1/X_3$ sea menor que la existente entre X_2 y X_3 han resultado ser:

(i) $r^2_{23} \geqslant a^2$ donde $a^2 \doteq \left(\dfrac{V_3}{V_2}\right)^2$

(ii) El signo de r_{23} debe ser igual al signo de

$V_3/2\,V_2$.

Además debemos agregar las restricciones provenientes del desarrollo en serie de funciones:

(iii) Condición necesaria de convergencia

$$V^2_3 < 1$$

(iv) Condición necesaria y suficiente de convergencia:

$$\left(\frac{X_{3i} - \bar{X}_3}{\bar{X}_3}\right)^2 < 1 \quad (i = 1, \ldots, n)$$

Es decir si existe una relación lineal estrecha entre las variables del modelo (1) no siempre se puede disminuir el grado de colinealidad a través del expediente de generar una ecuación de ajuste dividiendo dicha igualdad entre una de las variables (por ejemplo X_3). Este procedimiento será efectivo para rebajar las consecuencias de la colinealidad si se cumplen todas las condiciones que hemos establecido. En cualquier otro caso la estrategia conduciría a intensificar la relación.

Un ejemplo*

Sobre la base de datos trimestrales para el período 1970-1977 para la economía mexicana, se ha decidido estimar una ecuación para el nivel de precios en función

* Los resultados que se exponen en esta sección han sido tomados de un trabajo presentado al Curso de Econometría del programa de maestría en Economía de El Colegio de Mé-

de los salarios y de la cantidad de dinero.[10] Al realizar el ajuste mínimo cuadrático ordinario se detectó autocorrelación entre los residuos, lo que condujo a realizar una regresión en diferencias finitas de primer orden generalizadas, con los siguientes resultados:

$$(7) \qquad Y_i^* = 13.585 + 0.156\, X_{21}^* + 1.225\, X_{3i}^*$$

$$(S_{\hat{\beta}_2} = 0.082) \quad (S_{\hat{\beta}_3} = 0.186)$$

$$(t = 1.917) \quad (t = 6.581)$$

$R^2 = 0.9622$ $\qquad F = 343.967 \qquad$ y $r_{23} = 0.9471$

donde

$Y_1^* = Y_i - \rho Y_{i-1}$, y Y_i es el índice de precios de los bienes no duraderos en el tiempo i.

$X_{2i}^* = X_{2i} - \rho X_{2,\, i-1}$. X_{2i} es un índice de salario en el trimestre i.

$X_{3i}^* = X_{3i} - \rho X_{3,\, i-1}$. X_{3i} es la cantidad de circulante en i.

Además, ρ se define en la ecuación $e_i = \rho e_{i-1} + V_i$, en que e_i son los residuos observados y V_i representa los errores carentes de autocorrelación.

Los resultados de este ajuste parecen ser consistentes en la medida que los coeficientes tendrían el sentido que era de esperar desde el punto de vista teórico. El valor del coeficiente de determinación parecería estar fuertemente influido por la intensa colinealidad existente entre los regresores.[11] Además, esta misma relación lineal estrecha podría ser la causa de que el efecto de los salarios sobre el nivel de precios sea estadísticamente igual a cero.

Para intentar disminuir las consecuencias de la colinealidad podríamos dividir toda la ecuación entre X_3^* ya que los datos muestran que ningún valor de X_3^* es mayor que dos veces la media aritmética con lo cual se cumple la condición necesaria y suficiente de convergencia. Ambos coeficientes de variación son menores que uno.

$V_2 = 0.4643$ y $V_3 = 0.4229$

---

xico, por el estudiante Rodrigo Quintanilla. En este artículo, no nos interesa discutir la naturaleza del modelo ni sus resultados sustantivos, sólo queremos mostrar la bondad del método que estamos proponiendo.

10 Los detalles referidos a la información se encuentran en el apéndice 3.

11 Ya que $R^2 = \dfrac{r_{12}^2 + r_{13}^2 - 2 r_{12} r_{13} r_{23}}{1 - r_{23}^2}$.

De donde resulta que:

$$a^2 = \frac{(V_3\ )^2}{(2\ V_2)^2} = 0.2074$$

Por lo tanto,

$$r^2_{23} > a^2$$

Cumpliéndose de esta manera la tercera condición. Por último, también resulta que se satisface la condición de signos iguales para a y r_{23}.

De acuerdo con los desarrollos de la sección anterior la división de la ecuación (7) entre $X_{3\ i}$ debería conducir a una disminución de la colinealidad. Según la fórmula (6) el r^2 –la correlación entre $1/X^*_3$ y X^*_2/X^*_3 de la ecuación transformada– debería ser aproximadamente igual a:

$$r^2 \doteq \frac{[0.4229 - (0.9471)\ (0.4643)]^2}{(0.4229)^2 - 2\,(0.9471)\,(0.4229)\,(0.4643) + (0.4643)^2}$$

$$r^2 \doteq 0.0126$$

$$r \doteq 0.1122$$

El ajuste de la ecuación transformada ha dado los siguientes resultados.

$$(8)\quad \frac{Y^*_i}{X^*_{3i}} = \underset{\substack{(S_{\hat{\beta}_1} = 1.962)\\ (t = 5.685)}}{11.152} \left(\frac{1}{X^*_{3i}}\right) + \underset{\substack{(S_{\hat{\beta}_2} = 0.087)\\ (t = 1.366)}}{0.119} \left(\frac{X^*_{2i}}{X^*_{3i}}\right) + \underset{\substack{(S_{\hat{\beta}_3} = 0.199)\\ (t = 6.555)}}{1.291}$$

$R^2 = 0.5672$ $F = 17.692$ y $r_{23} = 0.060$[12]

La caída sustancial del coeficiente de determinación se explica básicamente por la disminución en la correlación entre los regresores.[13]

Los coeficientes de las ecuaciones (7) y (8) no difieren esencialmente. Debe recordarse que el término libre de (8) es el coeficiente de X_3 y que el coeficiente de $1/X_3$ es la constante de la ecuación (7).

12 Si se quisiera lograr una mejor aproximación entre el valor de r dado por (6) y la correlación entre los regresores del modelo de ajuste, debería incorporarse un número mayor de términos en el desarrollo en serie de funciones.

13 Un buen estudio de los factores de que depende el coeficiente de determinación y sus correspondientes interpretaciones se encuentra en: Duncan O. D., Partials, Partitions Paths, en *Sociologycal Methodology 1970.* Borgatta E. F. y Bohrnstedt G. Eds. Jossey Bass, 1970.

A pesar de haber reducido sustancialmente el grado de colinealidad entre los regresores, los errores estándares de los estimadores experimentaron tanto un alza absoluta como en relación al valor de la estimación, lo que se refleja en que los coeficientes t de (8) son inferiores a los de la ecuación (7). *Esto nos muestra que la disminución de la correlación entre los regresores no necesariamente se refleja en una mayor precisión en las estimaciones.*

En consecuencia, la transformación no lineal aplicada sobre el modelo original disminuirá el nivel de la colinealidad si se cumplen las condiciones que hemos estipulado, pero ello no siempre se traduce en una disminución de los errores estándares de los estimadores, ya que al dividir entre X_3 se puede introducir heterocedasticidad en modelos homocedásticos. En la próxima sección expondremos algunos muy breves alcances sobre esta temática.

A MODO DE CONCLUSIÓN: MULTICOLINEALIDAD, HETEROCEDASTICIDAD Y VARIANZA DE LOS ESTIMADORES

En los casos que aplicamos la división entre una de las variables explicativas y que efectivamente logramos disminuir el nivel de la colinealidad, la caída de la correlación entre los regresores tendrá como efecto la disminución en la varianza de los estimadores. Pero la transformación aplicada sobre el modelo tendrá también un impacto sobre los errores de los estimadores dependiendo de la naturaleza del término de error.

En efecto, si el modelo además de presentar una relación lineal estrecha entre las variables explicativas, tiene:

(*i*) Errores homocedásticos. La división entre X_3 introducirá heterocedasticidad, lo cual hará aumentar la varianza de los estimadores.

Que la varianza de los estimadores de la ecuación transformada sea mayor o menor que la de los estimadores del modelo original, será la resultante de la disminución operada por la caída en la colinealidad y el aumento causado por la heterocedasticidad.

(*ii*) Errores heterocedásticos del tipo:

$$U_i = X_{3i} V_i$$

donde $EV_i^2 = \sigma^2$ para todo i.

En este caso la división entre X_3 transforma un modelo heterocedástico en homocedástico, con lo cual se debe producir un aumento en la precisión de las estimaciones.

Los efectos de disminución del nivel de colinealidad y el hecho de pasar de un modelo hetero a uno homocedástico se refuerzan mutuamente para hacer bajar la varianza de los estimadores.

(*iii*) Errores heterocedásticos de cualquier tipo.

En esta situación la precisión de las estimaciones será la resultante del aumento inducido por la menor correlación entre los regresores y el efecto que se produzca sobre ellas vía el alejamiento o cercanía respecto a la homocedasticidad.

Estas consideraciones que han vinculado el tratamiento del problema de la colinealidad con el de la heterocedasticidad, proveen un criterio para seleccionar la variable que se usará como divisor en el modelo; *se debe elegir aquella variable que además de cumplir con las condiciones de convergencia de las series, sea la que introduce heterocedasticidad en el modelo.*

La generalización de este estudio al caso de modelos con más de dos variables explicativas, se hace bastante compleja desde el punto de vista matemático. Sin embargo, hemos aplicado el procedimiento descrito a modelos lineales de dicho tipo y en algunas ocasiones se ha logrado aumentar los niveles de precisión en las estimaciones de los parámetros.

Apéndice 1

Media aritmética y varianza de razón de variables. Correlación de variables de razón.

1. *Media aritmética*

Sea

$$Z_i = \frac{X_{ki}}{X_{ji}} \qquad i = 1, 2, \ldots n$$

$$M[Z] = \frac{1}{n} \sum_{i=1}^{n} \frac{X_{ki}}{X_{ji}}$$

Definamos

$$x_{ki} = X_{ki} - \bar{X}_k \qquad \text{y} \qquad x_{ji} = X_{ji} - \bar{X}_j$$

Por lo tanto

$$M[Z] = \frac{1}{n} \sum \frac{X_{ki}}{X_{ji}} = \frac{\bar{X}_k}{n\bar{X}_j} \sum \left(1 + \frac{x_{ki}}{\bar{X}_k}\right) \left(1 + \frac{x_{ji}}{\bar{X}_j}\right)^{-1}$$

Pero

$$\left(1 + \frac{x_{ji}}{\bar{X}_j}\right) = 1 - \frac{x_{ji}}{\bar{X}_j} + \frac{x_{ji}^2}{\bar{X}_j^2}$$

La condición de convergencia de la serie es $\left(\frac{x_{ji}}{\bar{X}_j}\right)^2 < 1$.

Ahora bien, supongamos que $\left(\frac{x_{ji}}{\bar{X}_j}\right)^2$ es lo suficientemente pequeño como para no considerar las potencias mayores que dos. Reemplazando y desarrollando se llega a

$$M[Z] \doteq \frac{\bar{X}_k}{\bar{X}_j} \left[1 - r_{kj} V_k V_j + V_j^2\right]$$

2. *La varianza*

$$\mathrm{Var}\,[Z] = \frac{1}{n} \sum \left(\frac{X_{ki}}{X_{ji}} \right)^2 - \left\{ M\,[Z] \right\}^2$$

$$\mathrm{Var}\,[Z] + \left\{ M\,[Z] \right\}^2 = \frac{1}{n} \sum \left(\frac{X_{ki}}{X_{ji}} \right)^2$$

$$= \frac{1}{n} \cdot \frac{\bar{X}_k^2}{\bar{X}_j^2} \sum \left(1 + \frac{x_{ki}}{\bar{X}_k} \right) \left(1 + \frac{x_{ji}}{\bar{X}_j} \right)^{-2}$$

Pero

$$\left(1 + \frac{x_{ji}}{\bar{X}_j} \right)^{-2} = 1 - \frac{2x_{ji}}{\bar{X}_j} + \frac{3x_{ji}^2}{\bar{X}_j^2} - \ldots\ldots\ldots$$

Con la condición de convergencia $\left(\frac{x_{ji}}{\bar{X}_j} \right)^2 < 1$. Bajo el supuesto que esta última expresión es lo suficientemente pequeña como para lograr una buena aproximación aun cuando se eliminen las potencias mayores que dos.

Después de reemplazar y desarrollar se concluye que

$$\mathrm{Var}\,[Z] = \frac{\bar{X}_k^2}{\bar{X}_j^2} \left[V_k^2 - 2\, r_{kj}\, V_k\, V_j + V_j^2 \right]$$

3. *El coeficiente de correlación*

Sea r el coeficiente de correlación entre las variablez Z_k y Z_m y S_k y S_m sus desviaciones típicas.

$$n\, r\, S_k\, S_m = \sum (Z_{ki} - \bar{Z}_k)\,(Z_{mi} - \bar{Z}_m)$$

donde

$$Z_{ki} = \frac{X_{ki}}{X_{ji}} \qquad \text{y} \qquad Z_m = \frac{X_{mi}}{x_{ji}}$$

$$\sum (Z_{ki} - \bar{Z}_k)\,(Z_{mi} - \bar{Z}_m) = \sum \frac{X_{ki}\, X_{mi}}{X_{ji}^2} - n\, \bar{Z}_k\, \bar{Z}_m$$

$$\sum \frac{X_{ki}\, X_{mi}}{X_{ji}^2} = \frac{\bar{X}_k\, \bar{X}_m}{\bar{X}_j^2} \sum \left(1 + \frac{x_k}{\bar{X}_k} \right) \left(1 + \frac{x_m}{\bar{X}_m} \right) \left(1 + \frac{x_j}{\bar{X}_j} \right)^{-2}$$

Pero

$$\left(1 + \frac{x_{ji}}{\bar{X}_j}\right)^{-2} = 1 - \frac{2x_{ji}}{\bar{X}_j} + \frac{3x_{ji}^2}{\bar{X}_j^2} - \ldots$$

Con la condición de convergencia $\left(\frac{x_{ji}}{x_j}\right)^2 < 1$ y suponiendo que la tasa de convergencia es rápida se pueden eliminar todas las potencias superiores a dos.

Después de reemplazar la serie y de operar algebraicamente se concluye en

$$r^2 \doteq \frac{[r_{km} V_k V_m - r_{kj} V_k V_j - r_{mj} V_m V_j + V_j^2]^2}{[V_k^2 - 2r_{kj} V_k V_j + V_j^2]\,[V_m^2 - 2r_{mj} V_m V_j + V_j^2]}$$

Donde V_k, V_m y V_j simbolizan los coeficientes de variación de las variables X_k, X_m y X_j, respectivamente.

Ahora bien, a nosotros nos interesa el caso particular

$$Z_{ki} = Z_1 = \frac{1}{X_{3i}} \qquad \text{y} \qquad Z_{mi} = Z_2 = \frac{X_{2i}}{X_{3i}}$$

Como $V_k = V_1 = 0$ ya que el numerador de Z_1 es una constante, desaparecen de r^2 todos los términos multiplicados por V_k

$$r^2 \doteq \frac{(V_3^2 - r_{23} V_2 V_3)^2}{V_3^2\,[V_2^2 - 2\,r_{23} V_2 V_3 + V_3^2]}$$

$$r^2 \doteq \frac{(V_3 - r_{23} V_2)^2}{V_2^2 - 2\,r_{23} V_2 V_3 + V_3^2}$$

Si reemplazamos $a = \dfrac{V_3}{2V_2}$

Llegamos a:

$$r^2 \doteq \frac{(2a - r_{23})^2}{1 - 4a\,r_{23} + 4a^2}$$

Con esta fórmula se realizaron los cálculos del apéndice 2 y la gráfica del apéndice 4.

Apéndice 2. Tabla de r^2 en función de r_{23} y a[1]

r_{23} –	−1.0	−0.9	−0.8	−0.7	−0.6	−0.5	−0.4	−0.3	−0.2	−0.1	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
r^2	1.00	0.81	0.64	0.49	0.36	0.25	0.16	0.09	0.04	0.01	0.00	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00
a																					
−2.0	1.000	0.981	0.966	0.955	0.948	0.942	0.939	0.938	0.938	0.939	0.941	0.944	0.948	0.953	0.958	0.964	0.971	0.977	0.985	0.992	1.000
−1.5	1.000	0.959	0.931	0.912	0.900	0.893	0.889	0.889	0.891	0.895	0.900	0.907	0.914	0.923	0.932	0.942	0.953	0.964	0.976	0.988	1.000
−1.0	1.000	0.864	0.800	0.768	0.754	0.750	0.753	0.761	0.771	0.785	0.800	0.817	0.834	0.853	0.873	0.893	0.914	0.935	0.956	0.978	1.000
−0.9	1.000	0.810	0.735	0.703	0.692	0.693	0.700	0.712	0.727	0.745	0.764	0.785	0.806	0.829	0.852	0.876	0.900	0.925	0.949	0.975	1.000
−0.8	1.000	0.721	0.640	0.614	0.610	0.617	0.632	0.650	0.671	0.694	0.719	0.745	0.771	0.799	0.826	0.855	0.883	0.912	0.914	0.970	1.000
−0.7	1.000	0.568	0.500	0.490	0.500	0.519	0.543	0.571	0.600	0.631	0.662	0.694	0.727	0.761	0.794	0.828	0.862	0.896	0.931	0.965	1.000
−0.6	1.000	0.321	0.308	0.329	0.360	0.395	0.432	0.471	0.510	0.550	0.590	0.631	0.671	0.712	0.753	0.794	0.835	0.876	0.917	0.959	1.000
−0.5	9.000[2]	0.050	0.100	0.150	0.200	0.250	0.300	0.350	0.400	0.450	0.500	0.550	0.600	0.650	0.700	0.750	0.800	0.850	0.900	0.950	1.000
−0.4	1.000	0.050	0.000	0.019	0.059	0.107	0.160	0.216	0.273	0.331	0.390	0.450	0.510	0.571	0.632	0.693	0.754	0.815	0.877	0.938	1.000
−0.3	1.000	0.321	1.100	0.019	0.000	0.013	0.045	0.090	0.143	0.202	0.265	0.331	0.400	0.471	0.543	0.617	0.692	0.768	0.845	0.922	1.000
−0.2	1.000	0.568	0.308	0.150	0.059	0.013	0.000	0.011	0.040	0.083	0.138	0.202	0.273	0.350	0.432	0.519	0.610	0.703	0.800	0.899	1.000
−0.1	1.000	0.721	0.500	0.329	0.200	0.107	0.045	0.011	0.000	0.010	0.038	0.083	0.143	0.216	0.300	0.395	0.500	0.614	0.735	0.864	1.000
0.0	1.000	0.810	0.640	0.490	0.360	0.250	0.160	0.090	0.040	0.010	0.000	0.010	0.040	0.090	0.160	0.250	0.360	0.490	0.640	0.810	1.000
0.1	1.000	0.864	0.735	0.614	0.500	0.395	0.300	0.216	0.143	0.083	0.038	0.010	0.000	0.011	0.045	0.107	0.200	0.329	0.500	0.721	1.000
0.2	1.000	0.899	0.800	0.703	0.610	0.519	0.432	0.350	0.273	0.202	0.138	0.083	0.040	0.011	0.000	0.013	0.059	0.150	0.308	0.568	1.000
0.3	1.000	0.922	0.845	0.768	0.692	0.617	0.543	0.471	0.400	0.331	0.265	0.202	0.143	0.090	0.045	0.013	0.000	0.019	0.100	0.321	1.000
0.4	1.000	0.938	0.877	0.815	0.754	0.693	0.632	0.571	0.510	0.450	0.390	0.331	0.273	0.216	0.160	0.107	0.059	0.019	0.000	0.050	1.000
0.5	1.000	0.950	0.900	0.850	0.800	0.750	0.700	0.650	0.600	0.550	0.500	0.450	0.400	0.350	0.300	0.250	0.200	0.150	0.100	0.050	9.000[2]
0.6	1.000	0.959	0.917	0.876	0.835	0.794	0.753	0.712	0.671	0.631	0.590	0.550	0.510	0.471	0.432	0.395	0.360	0.329	0.308	0.321	1.000
0.7	1.000	0.965	0.931	0.896	0.862	0.828	0.794	0.761	0.727	0.694	0.662	0.631	0.600	0.571	0.543	0.519	0.500	0.490	0.500	0.568	1.000
0.8	1.000	0.970	0.941	0.912	0.883	0.855	0.826	0.799	0.771	0.745	0.719	0.694	0.671	0.650	0.632	0.617	0.610	0.614	0.640	0.721	1.000
0.9	1.000	0.975	0.949	0.925	0.900	0.876	0.852	0.829	0.806	0.785	0.764	0.745	0.727	0.712	0.700	0.693	0.692	0.703	0.735	0.810	1.000
1.0	1.000	0.978	0.956	0.935	0.914	0.893	0.873	0.853	0.834	0.817	0.800	0.785	0.771	0.761	0.753	0.750	0.754	0.768	0.800	0.864	1.000
1.1	1.000	0.981	0.962	0.943	0.925	0.907	0.889	0.873	0.857	0.842	0.829	0.817	0.806	0.799	0.794	0.794	0.800	0.815	0.845	0.889	1.000
1.2	1.000	0.983	0.966	0.950	0.934	0.918	0.903	0.889	0.876	0.863	0.852	0.842	0.834	0.829	0.826	0.828	0.835	0.850	0.877	0.922	1.000

1 La fórmula para calcular r^2 aparece al final del apéndice 1.

2 El valor de r^2 para $r_{23} = -1$ y $r_{23} = 1$; a $= 0.5$ está indeterminado

Apéndice 3

LOS DATOS.

AÑO	Trimestre	Indice de precios al consumidor de los bienes no duraderos[1]	Indice de salarios[2]	Dinero (Miles de mill.)[2]
1970	1	100.0	100.0	44.47
	2	100.4	101.0	45.55
	3	102.0	103.1	46.40
	4	102.6	110.3	49.44
1971	1	104.7	102.1	47.81
	2	105.8	111.3	48.90
	3	106.3	111.3	49.93
	4	107.0	119.6	53.21
1972	1	108.0	112.4	53.47
	2	109.4	114.4	55.45
	3	111.2	118.6	57.77
	4	112.6	127.8	61.15
1973	1	116.2	119.6	64.69
	2	120.9	125.8	68.05
	3	130.0	129.9	72.50
	4	138.9	157.7	74.70
1974	1	154.8	151.5	77.99
	2	160.1	156.7	82.65
	3	165.5	169.1	85.86
	4	173.9	203.1	89.65
1975	1	178.2	190.7	95.26
	2	183.4	202.1	100.94
	3	189.4	212.4	103.32
	4	192.2	222.7	108.67
1976	1	198.0	220.6	112.00
	2	202.2	238.1	117.79
	3	208.0	254.6	130.14
	4	236.1	322.7	140.04
1977	1	256.6	311.3	146.05
	2	268.8	325.9	147.78
	3	280.1	344.5	160.70

1 Indices Económicos. Banco de México.
2 Financial Statistics. Fondo Monetario Internacional.

Apéndice 4. Gráfica de la relación entre r^2 y $r^2{}_{23}$ para tres valores particulares de a