

Taller #7
Multicolinealidad
Econometría 06169

Profesor: Julio César Alonso
Monitores: Hernán Betancur
David Valencia

Notas:

- Recuerde que tres preguntas, seleccionadas al azar, serán calificadas.
- Este taller es para ser entregado entre las 8:30 am y 9:30 am del 12 de febrero en mi oficina.

INSTRUCCIONES:

- Este taller debe ser escrito en computador.
 - Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.
- Este taller es un trabajo en pareja. Por tanto el taller debe reflejar **únicamente**

Usted como jefe del área financiera de la empresa el Maracuya.com solicita un estudio para determinar cuál ha sido el comportamiento del gasto total ($Gtot_t$) en la empresa desde que ésta ha sido fundada. Solicita ud entonces cierta información al área de estadística, la cual rápidamente le es enviada. Esta información consta de 18 observaciones correspondientes a la Inversión en maquinaria ($InvM_t$) en miles de millones de moneda local, gasto en planta (Gpl_t) también en miles de millones de moneda local y por último los gastos en administración $Gadm_t$ medidas en miles de millones de moneda local durante el periodo 1988 – 2006.

A partir de la información contenida en el archivo “T6-01-06.xls”, ud debe estimar el siguiente modelo que desafortunadamente no cuenta con ninguna fundamentación teórica:

$$Gtot_t = \beta_1 + \beta_2 InvM_t + \beta_3 Gpl_t + \beta_4 Ln(Gadm)_t + \varepsilon_t \quad (1)$$

1. Estime el modelo (1) y reporte sus resultados en una tabla.
2. Continuando con la pregunta anterior:
 - a) Analice la significancia individual y conjunta.
 - b) ¿Presenta un buen ajuste el modelo? ¿Qué observa?
3. Determine si existe o no multicolinealidad (muestre todo su trabajo y como llega a sus conclusiones).
4. Basándose en las pruebas efectuadas anteriormente corrija el problema encontrado (si es que este existe y si es que se puede corregir) y reporte el modelo estimado. Explique claramente su decisión.
5. Compruebe que el nuevo modelo no tiene problemas de multicolinealidad.
6. Interprete los coeficientes estimados.

Taller #7
Respuestas Sugeridas
Multicolinealidad
Econometría 06169

Profesor: Julio César Alonso
Monitores: Hernán Betancur
David Valencia

Notas:

- Recuerde que tres preguntas, seleccionadas al azar, serán calificadas.
- Este taller es para ser entregado entre las 8:30 am y 9:30 am del 12 de febrero en mi oficina.

INSTRUCCIONES:

- Este taller debe ser escrito en computador.
- Cuando sea posible, debe mostrar el procedimiento efectuado para llegar a sus resultados.

Este taller es un trabajo en pareja. Por tanto el taller debe reflejar **únicamente**

Usted como jefe del area financiera de la empresa el Maracuya.com solicita un estudio para determinar cual ha sido el comportamiento del gasto total $Gtot_t$ en la empresa desde que esta ha sido fundada, solicita ud entonces cierta informacion al area de estadistica, la cual rapidamente le es enviada, la cual consta de 18 observaciones correspondientes a la Inversion en maquinaria ($InvM_t$) en miles de millones de moneda local, gasto en planta (Gpl_t) también en miles de millones de moneda local y por último los gastos en administracion $Gadm_t$ medidas en miles de millones de moneda local durante el periodo 1988 – 2006.

A partir de la información contenida en el archivo “T6-01-06.xls”, ud debe estimar el siguiente modelo:

$$Gtot_t = \beta_1 + \beta_2 InvM_t + \beta_3 Gpl_t + \beta_4 Ln(Gadm)_t + \varepsilon_t \quad (1)$$

1. Estime el modelo (1) y reporte sus resultados en una tabla.

Los resultados se reportan en la Tabla 1

2. Continuando con la pregunta anterior:
 - a) Analice la significancia individual y conjunta.

El coeficiente asociado a la variable ($InvM_t$) es significativo a un nivel del 1%, mientras que el intercepto y los coeficientes asociados a las variables (Gpl_t) y ($Gadm_t$) no son significativos ni al 10%. Para verificar la significancia conjunta, se emplea el

estadístico F. Este es igual a 7917.52 lo que permite rechazar la hipótesis nula de que todos los coeficientes son conjuntamente iguales a cero a un nivel de significancia del 1%.

b) ¿Presenta un buen ajuste el modelo? ¿Qué observa?

El R^2 es igual a 0.9994, es decir, el 99.9% de la variabilidad en el PIB es explicada por las variables incluidas en el modelo. Aparentemente el ajuste del modelo es muy alto. Todo lo anterior hace pensar que puede existir un grave problema de multicolinealidad. Esto puede deberse a la definición de las variables ya que gran parte de la información proporcionada por la variable (inversión en maquinaria) es suministrada a su vez por la variable (gasto en planta).

	VARIABLE DEPENDIENTE: PIB _t		
	Estadísticos t entre paréntesis		
	Ecuación 1 1987 - 2005 MCO	Ecuación 2 1987 - 2005 MCO	Ecuación 3 1987 - 2005 MCO
Constante	-11.70 (-0.75)	-5.648 (-0.38)	-88.142 (-1.62)
InvM_t	3.533 (13.94) ***	3.824 (148.12) ***	- -
Gpl_t	0.288 (1.16)	- -	3.748 (40.01) ***
LnGadm_t	4.041 (0.88)	2.258 (0.51)	26.868 (1.67) *
R ²	0.9994	0.9994	0.9912
R ² Ajustado	0.9993	0.9993	0.9901
F	7,917.52 ***	11,617.02 ***	848.16 ***
# de Obs.	18	18	18

(*) nivel de significancia: 10%

(**) nivel de significancia: 5%

(***) nivel de significancia: 1%

MCO: Mínimos Cuadrados Ordinarios

Tabla 1: Estimación ecuaciones (1) y (2).

3. Determine si existe o no multicolinealidad (muestre todo su trabajo y como llega a sus conclusiones).

Se emplean las pruebas vistas en clase para determinar la presencia de multicolinealidad.

Matriz de Correlación de las X's.

Se calcula el determinante de la matriz de correlación $|R|$ empleando los valores propios obtenidos:

$$|R| = 2.08014 \times 0.91501 \times 0.00485 = 0.009$$

El valor que toma el determinante es cercano a cero lo que implica un problema de multicolinealidad en el modelo.

Medida de Besley, Kuck y Welsch (1980).

Los valores propios (eigenvalues) de la matriz son: $\lambda_1 = 1.9944$, $\lambda_2 = 0.00551$ y Se calcula el número de condición:

$$\kappa(X) = \frac{\sqrt{\lambda_{MAX}}}{\sqrt{\lambda_{MIN}}} = \sqrt{\frac{2.08014}{0.00485}} = 20.714$$

Como el valor de $\kappa(X)$ es distinto de 1, existe multicolinealidad alcanzando un nivel preocupante, puesto que entre 20 y 30.

Matriz de correlación entre los coeficientes estimados.

	β_{hat_2}	β_{hat_3}	β_{hat_4}	β_{hat_0}
β_{hat_2}	1	-0.994911438	-0.355100033	0.351597463
β_{hat_3}	-	1	0.33470594	-0.33592357
β_{hat_4}	-	-	1	-0.998763769
β_{hat_0}	-	-	-	1

A partir de la matriz de correlación se determina que existe un fuerte correlación entre β_2 y β_3 cercana al -0.99. Se observa entonces un problema de multicolinealidad.

4. Basándose en las pruebas efectuadas anteriormente corrija el problema encontrado (si es que este existe) y reporte el modelo estimado. Explique claramente su decisión.

En este caso se puede pensar en omitir la variable $InvM_t$ o Gpl_t que se encuentran altamente correlacionadas ya que la teoría económica no brinda en este caso una guía clara en cuanto a esta decisión, y es fácil deducir que sin importar la decisión que se tome, la estimación del modelo no deberá variar, ya que ambas variables varían casi al unísono. Por lo tanto si el modelo a estimar es:

$$Gtot_t = \beta_1 + \beta_2 InvM_t + \beta_3 Ln(Gadm)_t + \varepsilon_t \tag{2}$$

Los resultados de la estimación de la ecuación (2) se reportan en la Tabla 1.

Y si el modelo a estimar es:

$$Gtot_t = \beta_1 + \beta_2 Gpl_t + \beta_3 Ln(Gadm)_t + \varepsilon_t \tag{3}$$

Los resultados de la estimación de la ecuación (3) se reportan en la Tabla 1.

5. Compruebe que el nuevo modelo no tiene problemas de multicolinealidad.

Matriz de Correlación de las X's.

Para la ecuación (2) se calcula el determinante de la matriz de correlación $|R|$ empleando los valores propios obtenidos:

$$|R| = 1.2327 \times 0.7672 = 0.9458$$

El valor que toma el determinante es relativamente cercano a uno lo que implica que aparentemente el problema de multicolinealidad queda solucionado. Es necesario llevar a cabo las demás pruebas para corroborar los resultados.

Para la ecuación (3) se calcula el determinante de la matriz de correlación $|R|$ empleando los valores propios obtenidos:

$$|R| = 1.197 \times 0.8026 = 0.961$$

El valor que toma el determinante es relativamente cercano a uno lo que implica que aparentemente el problema de multicolinealidad queda solucionado. Es necesario llevar a cabo las demás pruebas para corroborar los resultados.

Medida de Besley, Kuck y Welsch (1980).

Para la ecuación (2) se toman los valores propios (eigenvalues): $\lambda_1 = 1.2327$ y $\lambda_2 = 0.7672$. Se calcula el número de condición:

$$\kappa(X) = \frac{\sqrt{\lambda_{MAX}}}{\sqrt{\lambda_{MIN}}} = \sqrt{\frac{1.2327}{0.7672}} = 1.267$$

Como el valor de $\kappa(X)$ es cercano a 1, no existe multicolinealidad.

Para la ecuación (3) se toman los valores propios (eigenvalues): $\lambda_1 = 1.197$ y $\lambda_2 = 0.8026$. Se calcula el número de condición:

$$\kappa(X) = \frac{\sqrt{\lambda_{MAX}}}{\sqrt{\lambda_{MIN}}} = \sqrt{\frac{1.197}{0.8056}} = 1.22$$

Como el valor de $\kappa(X)$ es cercano a 1, no existe multicolinealidad.

Matriz de correlación entre los coeficientes estimados.

Para la ecuación (2). Se obtiene:

	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_0$
$\hat{\beta}_2$	1	-0.232744669	0.183177557
$\hat{\beta}_3$	-	1	-0.998608105
$\hat{\beta}_0$	-	-	1

Se observa que la correlación entre los dos coeficientes asociados a las variables explicativas no es mayor a 0.8 por lo tanto no existe un problema grave.

Para la ecuación (3). Se obtiene:

	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_0$
$\hat{\beta}_2$	1	-0.197343004	0.14720872
$\hat{\beta}_3$	-	1	-0.998595399
$\hat{\beta}_0$	-	-	1

Se observa que la correlación entre los dos coeficientes asociados a las variables explicativas no es mayor a 0.8 por lo tanto no existe un problema grave

6. Interprete los coeficientes estimados.

Para la ecuación (2)

$\beta_1 = -5.648$: El total de gastos que no depende de la inversión en maquinaria ni de los gastos en administración.

$\beta_2 = 3.824$: Un incremento en la inversión en maquinaria de mil millones de unidades de moneda local genera un incremento de 3.824 miles de millones de unidades de moneda local en los gastos totales de la empresa.

$\beta_3 = 2.258$: Un aumento de un punto porcentual en los gastos en administración generan un incremento de 2.258 miles de millones de unidades de moneda local en los gastos totales de la empresa.

Para la ecuación (3)

$\beta_1 = -88.14$: El total de gastos que no depende de los gastos en planta ni de los gastos en administración.

$\beta_2 = 3.748$: Un incremento en los gastos en planta de mil millones de unidades de moneda local genera un incremento de 3.784 miles de millones de unidades de moneda local en los gastos totales de la empresa.

$\beta_3 = 26.86$: Un aumento de un punto porcentual en los gastos en administración genera un incremento de 26.86 miles de millones de unidades de moneda locales en los gastos totales de la empresa.