Taller #5

Multicolinealidad

Econometría I 06169

Profesor: Julio César Alonso
Monitores: Ana María Lotero
Carlos I. Patiño

Un investigador es contratado para estimar un modelo sencillo que explique el consumo de una pequeña república caribeña. Para esto, cuenta con 20 observaciones correspondientes al Consumo interno $(CONS_t)$, Remuneración a asalariados del sector agrario (WA_t), Remuneración al capital en el resto de la economía (RNA_t) y Remuneración a los factores en el sector agrario (RA_t) durante el periodo 1928 – 1948. Todos los datos en miles de moneda local. A partir de la información contenida en el archivo "T5I-02-04.xls", el investigador plantea estimar el siguiente modelo:

$$CONS_t = \beta_1 + \beta_2 RA_t + \beta_3 WA_t + \beta_4 RNA_t + \varepsilon_t \qquad (1)$$

1. Estime el modelo (1) y reporte sus resultados en una tabla.
2. Analice la significancia individual y conjunta. ¿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. Corrija el problema encontrado 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 #5

Multicolinealidad

Respuestas Sugeridas
Econometría I 06169

Profesor: Julio César Alonso
Monitores: Ana María Lotero
Carlos I. Patiño

Un investigador es contratado para estimar un modelo sencillo que explique el consumo de una pequeña república caribeña. Para esto, cuenta con 20 observaciones correspondientes al Consumo interno $(CONS_t)$, Remuneración a asalariados del sector agrario (WA_t), Remuneración al capital en el resto de la economía (RNA_t) y Remuneración a los factores en el sector agrario (RA_t) durante el periodo 1928 – 1948. Todos los datos en miles de moneda local. A partir de la información contenida en el archivo "T5I-02-04.xls", el investigador plantea estimar el siguiente modelo:

$$CONS_t = \beta_1 + \beta_2 RA_t + \beta_3 WA_t + \beta_4 RNA_t + \varepsilon_t \qquad (1)$$

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

Los resultados se reportan en la Tabla 1

2. Analice la significancia individual y conjunta. ¿Presenta un buen ajuste el modelo? ¿Qué observa?

Únicamente el coeficiente asociado a la variable (WA_t) es significativo a un nivel del 1%. El resto de coeficientes (incluido el intercepto) no son significativos individualmente a ningún nivel. Para verificar la significancia conjunta, se emplea el estadístico F. Este es igual a 107.36 lo que permite rechazar la hipótesis nula de que todos los coeficientes son conjuntamente iguales a cero a un nivel de significancia del 1%.

El R^2 es igual a 0.9527, es decir, el 95.3% de la variabilidad en el consumo interno 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 la remuneración a los factores del sector agrario incluye la remuneración a los asalariados.

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

	VARIABLE DEPENDIENTE: *$CONS_t$* Estadísticos t entre paréntesis	
	Ecuación 1 1928 - 1948 MCO	Ecuación 2 1928 - 1948 MCO
Constante	7,80 (0,88)	7,779 (0,90)
RA_t	0,032 (0,03)	---
WA_t	1,085 (6,18) ***	1,089 (12,68) ***
RNA_t	0,409 (0,63)	0,406 (0,65)
R^2	0,9527	0,9527
R^2 Ajustado	0,9438	0,9471
F	107,36 ***	171,09 ***
# de Obs.	20	20

(*) nivel de significancia: 10%
(**) nivel de significancia: 5%
(***) nivel de significancia: 1%
MCO: Mínimos Cuadrados Ordinarios

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:

$$|R| = 2.4889 \times 0.4358 \times 0.0753 = 0.0816$$

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: $\lambda_1 = 2.4889$, $\lambda_2 = 0.4358$ y $\lambda_3 = 0.0753$. Se calcula el número de condición:

$$\kappa(X)=\frac{\sqrt{\lambda_{MAX}}}{\sqrt{\lambda_{MIN}}}=\sqrt{\frac{2.4889}{0.0753}}=5.75$$

Como el valor de $\kappa(X)$ es distinto de 1, existe multicolinealidad pero no alcanza un nivel elevado.

Matriz de correlación entre los coeficientes estimados.

Se determina a partir de la matriz de varianza y covarianza estimada. La matriz es igual a:

	$\hat\beta_2$	$\hat\beta_3$	$\hat\beta_4$	$\hat\beta_1$
$\hat\beta_2$	1	-0.863365	0.135355	0.068548
$\hat\beta_3$		1	-0.468629	0.133632
$\hat\beta_4$			1	-0.900887
$\hat\beta_1$				1

A partir de la matriz de correlación se determina que existe un fuerte correlación entre $\hat\beta_1$ y $\hat\beta_4$ cercana al 0.90. Adicionalmente, entre $\hat\beta_2$ y $\hat\beta_3$. Se observa entonces un problema de multicolinealidad.

4. Corrija el problema encontrado y reporte el modelo estimado. Explique claramente su decisión.

Se puede pensar en omitir la variable RA_t que está fuertemente correlacionada con WA_t ya que está contenida en esta. Adicionalmente, sus coeficientes asociados están fuertemente correlacionados. Por lo tanto el modelo a estimar sería:

$$CONS_t=\beta_1+\beta_2WA_t+\beta_3RNA_t+\varepsilon_t \qquad (2)$$

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

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

Matriz de Correlación de las X's.

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

$$|R|=1.7036\times0.2964=0.5049$$

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).

Los valores propios (eigenvalues) de la matriz son: $\lambda_1=1.7036$ y $\lambda_2=0.2964$. Se calcula el número de condición:

$$\kappa(X)=\frac{\sqrt{\lambda_{MAX}}}{\sqrt{\lambda_{MIN}}}=\sqrt{\frac{1.7036}{0.2964}}=2.3975$$

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

Matriz de correlación entre los coeficientes estimados.

Se determina a partir de la matriz de varianza y covarianza estimada. La matriz es igual a:

	$\hat\beta_2$	$\hat\beta_3$	$\hat\beta_1$
$\hat\beta_2$	1	-0,703625	0,383028
$\hat\beta_3$		1	-0,920785
$\hat\beta_1$			1

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

6. Interprete los coeficientes estimados.

$\beta_1=7.78$: Representa el consumo que no depende de la remuneración a los asalariados del sector agrario ni de la remuneración al capital en el resto de la economía.

$\beta_2=1.09$: Un incremento en la remuneración a los asalariados del sector agrario de 1000 unidades de moneda local genera un incremento de 1090 unidades de moneda local en el consumo interno.

$\beta_3=0.41$: Un aumento en la remuneración al capital del resto de la economía de 1000 unidades de moneda local genera un incremento en el consumo interno cercano a las 410 unidades de moneda local.