

MULTICOLINEALIDAD

Se da cuando algunas de las variables regresoras (explicativas) están correlacionadas , incumpliendo una de las hipótesis de partida

Se supone que es un problema de MUESTRA.

Si el modelo se utiliza para la predicción no será un problema importante si asumimos que la correlación entre las variables se mantiene en el futuro.

Tipos de Multicolinealidad

- **Perfecta:** rango de la matriz $X < K$.

La matriz no tiene rango completo; por lo tanto no podemos estimar el modelo por MCO. Sí podemos obtener estimadores MCO de los parámetros que son una combinación lineal de los parámetros originales.

Consecuencias: los parámetros estimados son INDETERMINADOS debido a que no es posible separar las influencias de las distintas variables explicativas debido a que están relacionadas linealmente.

- **Ortogonalidad (no existencia de multicolinealidad):** rango de la matriz $X = K$
En la realidad casi no ocurre. Implica que la estimación del vector de parámetros poblacionales es la misma tanto si estimamos el modelo de regresión múltiple como el modelo de regresión simple.
- **Imperfecta:** Es el caso en el que a la relación lineal entre las variables explicativas se le suma un término denominado error estocástico.

Consecuencias: podemos estimar los parámetros por MCO pero los valores estimados no son muy confiables. Cuanto mas grande es la correlación, más próximo a cero será el determinante de la matriz $X'X$ lo cuál incrementará las varianzas y covarianzas de los parámetros estimados.

Consecuencias de la Multicolinealidad

- Los estimadores MCO son Insesgados y Consistentes pero tienen varianzas grandes.
- Los intervalos de confianza son mas amplios.
- El modelo suele ser significativo (R^2 elevado) pero las variables individualmente no lo son (debido a que al ser grande la varianza el estadístico t de contraste es menor llevándonos a concluir que una variable es irrelevante cuando en realidad no lo es)
- Dificultad en la interpretación de los parámetros.

$$T = \frac{\text{S}}{\text{F}} = \frac{R^2/k-1}{(1-R^2)/(n-k)}$$

Técnicas de detección

- Analizar los coeficientes de correlación simple entre los regresores de 2 en 2. Si observamos una alta correlación nos estaría indicando la presencia de multicolinealidad.
- Analizar el determinante de la matriz de correlaciones". Si el valor del determinante es 0 hay MULTICOLINEALIDAD PERFECTA. Si es 1 NO EXISTE MULTICOLINEALIDAD.

- Estudiar los coeficientes de determinación de las regresiones en las cuales figura como variable endógena cada una de las variables explicativas del modelo. Un r^2 elevado implica MULTICOLINEALIDAD.
- También se puede detectar estudiando los parámetros estimados. Si éstos se ven muy alterados al efectuar pequeños cambios en las X, es un indicio de MULTICOLINEALIDAD.
- Si las estimaciones obtenidas contradicen la teoría económica puede deberse a la presencia de MULTICOLINEALIDAD.
- Analizando los R^2 de sucesivos MCO en que se elimina un regresor, si el R^2 no cambia puede ser debido a que la variable explicativa eliminada queda explicada por otras variables en cuyo caso es posible que exista MULTICOLINEALIDAD.
- Test de FARRAR–GLAUBER

Test de Farrar–Glauber

I ETAPA: Se busca evaluar la Ortogonalidad de las variables independientes.

Ho: Las x son Ortogonales

H1: Las x no son Ortogonales

Calc = - [$n - 1 - \frac{2k + 5}{6}$] De

6

n: tamaño de la muestra

k: coeficientes + intercepto

De: Valor absoluto del determinante de Rx

(Tabla) con $[k(k-1)]/2$ grados de libertad

Si $\text{calc} > \text{tabla}$, entonces rechazo Ho y estoy en presencia de Multicolinealidad.

II ETAPA: Si detectamos que las variables son NO ortogonales entonces se regresa cada variable explicativa contra el resto de variables independientes para ver cuál está más correlacionada conjuntamente.

Buscamos el coeficiente de determinación F de cada regresión y seleccionamos la que arroje el mayor F.

$X_2 = F(X_3; X_4 \dots)$, $R^2 (X_2 = X_3; X_4)$

$X_3 = F(X_2; X_4)$, $R^2 (X_3 = X_2; X_4)$

Ho: $R^2(X_2) = 0$

H1: $R^2(X_2) \neq 0$

$F_{\text{calc}} = \frac{R^2 X_i}{k-1}$

$1 - R^2 X_i / n - k$

Si $F_{calc} > F_{tabla}$, entonces rechazo H_0 y por lo tanto la variable X_2 está colineada con las demás variables explicativas y me demuestra la presencia de Multicolinealidad.

III ETAPA: Se hallan los coeficientes de correlación parcial para conocer con cual variable explicativa está mas correlacionada la variable seleccionada en la etapa anterior.

$H_0: r_{X_2X_j} \text{ (para todas las variables explicativas)} = 0$

$H_1: r_{X_2X_j} \neq 0$

$t_{calc} = \frac{r_{X_2X_j}}{\sqrt{1 - r_{X_2X_j}^2}} \sqrt{n - k}$

$(\text{raiz cuad}) \sqrt{1 - r_{X_2X_j}^2}$

t tabla con $n - k$ grados de libertad.

Si $t_{calc} > t_{tabla}$ entonces rechazo H_0 y no existe Multicolinealidad

Soluciones a la MULTICOLINEALIDAD

- Cambiar El modelo
- Ampliar la información
- Otros métodos de regresión (Regresión cresta*)
- Aceptar la Multicolinealidad y vigilar las consecuencias

* Regresión Cresta

Se supone que el valor reducido del determinante de $X'X$ causa problemas a la hora de aplicar MCO. Por este motivo se suma una cantidad determinada arbitrariamente (c) a los elementos de la diagonal principal.

$$\text{cresta } \&= [X'X + cI_k]^{-1} X'Y$$

" En cambio el estadístico de significación global F no es afectado.

" A diferencia del punto 1) en este caso se tienen en cuenta la correlación entre cualquier n° de variables.

&Para cada variable explicativa.

&Presenta un sesgo que aumenta con el valor de c .