

Multicolinealidad

Román Salmerón Gómez

Universidad de Granada

Contenidos

Concepto y causas

Multilinealidad exacta: efectos

Multilinealidad aproximada: efectos

Detección de la Multilinealidad

Soluciones al problema de multilinealidad

Concepto y causas

Multilinealidad exacta: efectos

Multilinealidad aproximada: efectos

Detección de la Multilinealidad

Soluciones al problema de multilinealidad

Contenidos

Concepto y causas

Multicolinealidad

Multicolinealidad exacta:
efectos

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

Concepto y causas

El problema de multicolinealidad consiste en la existencia de relaciones lineales entre dos o más variables independientes del modelo lineal uniecuacional múltiple.

Dependiendo de cómo sea dicha relación lineal hablaremos de multicolinealidad perfecta o aproximada.

Las principales causas que producen multicolinealidad en un modelo son:

- relación causal entre variables explicativas del modelo.
- escasa variabilidad en las observaciones de las variables independientes.
- reducido tamaño de la muestra.

En definitiva, la multicolinealidad suele ser un problema muestral que se presenta normalmente en datos con el perfil de series temporales.

Así, por ejemplo, la edad y la experiencia suelen presentar una alta relación ya que ambas evolucionan conjuntamente: a mayor edad se presupone mayor experiencia. Por tal motivo será difícil separar el efecto de cada una sobre la variable dependiente y que se produzca multicolinealidad debido a la relación causal existente entre dichas variables (series temporales).

Supongamos ahora que nos pasan una encuesta donde hay que valorar las siguientes afirmaciones en una escala de 1 a 5 donde 1 significa que estamos totalmente en desacuerdo y 5 totalmente de acuerdo:

- Seguro que saco un 10 en Econometría.
- No me gusta la Econometría.

Para la primera afirmación, variable que llamaremos X , tendremos valores concentrados alrededor del 1, mientras que para la segunda, que llamaremos Y , obtendremos valores alrededor del 5. Por tanto, tendríamos dependencia lineal ya que $Y \cong 5 \cdot X$.

Estas variables se podrían usar en un modelo donde la variable dependiente es la calificación obtenida en la asignatura de Econometría: X podría ser un indicio de la calificación esperada e Y del grado de afinidad a la materia.

Como se puede observar, la multicolinealidad de este ejemplo se debe a problemas con las observaciones disponibles (escasa variabilidad o reducido tamaño de la muestra). Por tanto, si se es capaz de mejorar estos problemas muestrales se evitaría la presencia de multicolinealidad entre dichas variables.

Contenidos

Concepto y causas

Multilinealidad exacta:
efectos

Multilinealidad
exacta o perfecta

Ejemplo

Multilinealidad
aproximada: efectos

Detección de la
Multilinealidad

Soluciones al problema
de multilinealidad

Multilinealidad exacta: efectos

Multicolinealidad exacta o perfecta

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

Multicolinealidad
exacta o perfecta

Ejemplo

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

La multicolinealidad exacta o perfecta hace referencia a la existencia de una relación lineal exacta entre dos o más variables independientes.

Dicho tipo de multicolinealidad se traduce en el incumplimiento de una de las hipótesis básicas del modelo uniecuacional múltiple: la matriz X no es de rango completo por columnas, esto es, $rg(X) < k$.

El incumplimiento de dicha hipótesis no permite invertir la matriz $X^t X$, por lo que el sistema normal

$$X^t X \cdot \beta = X^t y,$$

es compatible indeterminado, es decir, es imposible obtener una solución única para $\hat{\beta}$ (hay infinitas).

¿Qué hacer ante esta situación? Evidentemente no se podrán estimar los coeficientes de las variables independientes, sin embargo, si se podrá estimar una combinación lineal de los mismos. Y en tal caso no tenemos garantizado que se puedan recuperar a partir de éstas las estimaciones de los parámetros originales.

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

Multicolinealidad
exacta o perfecta

Ejemplo

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

Consideremos el modelo

$$Y_t = \beta_1 + \beta_2 X_{t2} + \beta_3 X_{t3} + u_t,$$

donde $X_{t2} - X_{t3} = 1$.

Entonces, sin más que sustituir $X_{t2} = 1 + X_{t3}$ en el modelo original:

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 \cdot (1 + X_{t3}) + \beta_3 X_{t3} + u_t \\ &= (\beta_1 + \beta_2) + (\beta_2 + \beta_3) \cdot X_{t3} + u_t, \end{aligned}$$

obtenemos que las combinaciones lineales estimables de los parámetros originales son:

$$\beta_1 + \beta_2, \quad \beta_2 + \beta_3.$$

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

**Multicolinealidad
aproximada: efectos**

Multicolinealidad
aproximada

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

Multicolinealidad aproximada: efectos

La multicolinealidad aproximada hace referencia a la existencia de una relación lineal aproximada entre dos o más variables independientes.

En este caso, no se incumplirá la hipótesis básica de que la matriz X sea completa por columnas ($rg(X) = k$), por lo que se podrá invertir $X^t X$ y obtener los estimadores por mínimos cuadrados ordinarios. Sin embargo, el determinante de $X^t X$ será muy próximo a cero, por lo que $(X^t X)^{-1}$ tenderá a tener valores altos.

En consecuencia, cuando existe un problema de multicolinealidad no perfecta se presentan los siguientes problemas:

- las varianzas de los estimadores son muy grandes.
- al efectuar contrastes de significación individual no se rechazará la hipótesis nula, mientras que al realizar contrastes conjuntos si.
- los coeficientes estimados serán muy sensibles ante pequeños cambios en los datos.
- un coeficiente de determinación elevado.

Contenidos

Concepto y causas

Multilinealidad exacta:
efectos

Multilinealidad
aproximada: efectos

Detección de la
Multilinealidad

Soluciones al problema
de multilinealidad

Detección de la Multilinealidad

Detección de la multicolinealidad

Contenidos

Concepto y causas

Multicolinealidad exacta: efectos

Multicolinealidad aproximada: efectos

Detección de la Multicolinealidad

Soluciones al problema de multicolinealidad

Basarse en los síntomas enumerados anteriormente para la detección de la multicolinealidad no es un procedimiento fiable ya que es subjetivo.

Por tal motivo, para la detección de la multicolinealidad usaremos los métodos:

- Número de condición.
- Factor de agrandamiento de la varianza.

El número de condición, $k(X)$, se define como la raíz cuadrada del cociente entre el autovalor más grande de $X^t X$, λ_{max} , y el más pequeño, λ_{min} . Esto es:

$$k(X) = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$$

Si dicho número de condición toma un valor entre 20 y 30 estamos ante un problema de multicolinealidad probable y se considera seguro si supera 30.

Detección de la multicolinealidad

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

El factor de agrandamiento de la varianza, FAV , se define para cada uno de los coeficientes como:

$$FAV(\hat{\beta}_i) = \frac{1}{1 - R_i^2}, \quad i = 2, \dots, k,$$

donde R_i^2 es el coeficiente de determinación obtenido al efectuar la regresión de X_i sobre el resto de las variables independientes del modelo.

El FAV se interpreta como la razón entre la varianza observada y la que habría sido en caso de que X_i estuviera incorrelacionada con el resto de variables independientes del modelo, es decir, muestra en qué medida se agranda la varianza del estimador como consecuencia de la relación de los regresores.

Valores del FAV superiores a 10 hacen pensar en la posible existencia de multicolinealidad en el modelo.

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

Soluciones

Soluciones al problema de multicolinealidad

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

Soluciones

Algunas de las posibles soluciones al problema de multicolinealidad son las siguientes:

- mejora del diseño muestral estrayendo la información máxima de las variables observadas.
- eliminación de las variables que se sospechan son causantes de la multicolinealidad.
- en caso de disponer de pocas observaciones, aumentar el tamaño de la muestra.
- utilizar la relación extramuestral que permita realizar relaciones entre los parámetros (información a priori) que permita estimar el modelo por mínimos cuadrados restringidos.

Por otro lado, algunos autores sugieren tratar el problema de la multicolinealidad de forma mecánica y puramente numérica proponiendo una técnica conocida como regresión alomada. Sin embargo, esta técnica tiene dos problemas importantes: es arbitraria y los estimadores obtenidos no son interpretables.

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

Soluciones

- [1] Salmerón, R. y Gómez, S. (2012). Relación entre los factores institucionales y el emprendimiento: análisis mediante técnicas cuantitativas. *Revista de Métodos Cuantitativos para la Economía y la Empresa*, Número 13, Páginas 54-72.
- [2] Novales, A. (1993). *Econometría*. McGraw Hill. Capítulo 1 (obtención de autovalores de una matriz).

Puedes encontrarlas en la dirección web:

<http://www.ugr.es/local/romansg/material/WebEco/index.html>

Contenidos

Concepto y causas

Multicolinealidad exacta:
efectos

Multicolinealidad
aproximada: efectos

Detección de la
Multicolinealidad

Soluciones al problema
de multicolinealidad

Soluciones

- [1] Gujarati, D. (1997). *Econometría*. Ed. McGraw Hill. Capítulo 8.
- [2] Novales, A. (1993). *Econometría*. McGraw Hill. Capítulo 10.
- [3] Uriel, E., Contreras, D., Moltó, M.L. y Peiró, A. (1990). *Econometría. El Modelo Lineal*. Editorial AC. Capítulo 7.