

TAREA 4
Multicolinealidad
Profesor Javier Olaya

Presentación

El problema de multicolinealidad en un modelo de regresión múltiple se presenta cuando existe una alta dependencia lineal entre las variables regresoras (X_j). Este problema no es deseable en el ajuste del modelo por mínimos cuadrados ordinarios (OLS), ya que conduce a estimaciones poco fiables de los coeficientes de regresión, pues tendrán varianzas y covarianzas muy grandes.

Las estimaciones de OLS en presencia de multicolinealidad son insesgadas, pero pierden eficiencia. Para enfrentar esta dificultad, se han propuesto algunas técnicas, dentro de las que se encuentran la regresión por componentes principales (PCR), la regresión ridge (RR) y la regresión por mínimos cuadrados parciales (PLS).

Esta tarea tiene como propósito principal hacer conciencia de la importancia de identificar este problema de multicolinealidad en un conjunto de datos, aprender a detectar si existe o no y evaluar las estimaciones obtenidas utilizando PLS, PCR, RR y OLS.

Temas de presentación

En primer lugar se debe verificar la presencia de multicolinealidad en el conjunto de datos. En caso que se detecte multicolinealidad en los datos, se debe utilizar una de las soluciones propuestas y comparar los resultados obtenidos vía OLS con los obtenidos con cada una de las otras técnicas.

Metodología

Se utilizará la misma base de datos *cadata*. Cada grupo usará todas las variables con su propio conjunto de datos de la tarea 3 y a cada uno se le asignará uno de los métodos para comparar con OLS.

El informe escrito debe:

- Describir los métodos de detección de multicolinealidad y usarlos para identificar si este fenómeno afecta o no estos datos en particular.
- Presentar cómo se estiman los coeficientes de regresión y sus varianzas, usando el método asignado.
- Comparar los resultados obtenidos por OLS y por el método asignado.
- Contener una sección dedicada a la implementación en **R**. No se incluyen los scripts en el informe, pero el análisis de los datos debe realizarse en **R**.
- Incluir conclusiones propias
- Detallar la bibliografía consultada