Chương 6
Đa cộng tuyến

I. Bản chất của đa cộng tuyến

Đa cộng tuyến là tồn tại mối quan hệ tuyến tính giữa một số hoặc tất cả các biến độc lập trong mô hình.

Xét hàm hồi qui k biến :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + U_i$$

- Nếu tồn tại các số $\lambda_2, \lambda_3, \ldots, \lambda_k$ không đồng thời bằng 0 sao cho :

CuuDuongThanCong.com https://fb.com/tailieudientucntt

$$\lambda_2 X_{2i} + \lambda_3 X_{3i} + \ldots + \lambda_k X_{ki} + a = 0$$

(a : hằng số)

Thì giữa các biến độc lập xảy ra hiện tượng ***đa cộng tuyến hoàn hảo***.

- Nếu tồn tại các số $\lambda_2, \lambda_3, \ldots, \lambda_k$ không đồng thời bằng 0 sao cho :

$$\lambda_2 X_{2i} + \lambda_3 X_{3i} + \ldots + \lambda_k X_{ki} + V_i = 0$$

(V_i : sai số ngẫu nhiên)

Thì giữa các biến độc lập xảy ra hiện tượng ***đa cộng tuyến không hoàn hảo***.

CuuDuongThanCong.com https://fb.com/tailieudientucntt

<u>Ví dụ</u> : $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + U_i$

Với số liệu của các biến độc lập :

X_2	10	15	18	24	30
X_3	50	75	90	120	150
X_4	52	75	97	129	152

Ta có : $X_{3i} = 5X_{2i}$ → có hiện tượng cộng tuyến hoàn hảo giữa X_2 và X_3 và $r_{23} = 1$

$X_{4i} = 5X_{2i} + V_i$ → có hiện tượng cộng tuyến không hoàn hảo giữa X_2 và X_3, có thể tính được $r_{24} = 0.9959$.

CuuDuongThanCong.com https://fb.com/tailieudientucntt

II. Ước lượng trong trường hợp có đa cộng tuyến

1.Trường hợp có đa cộng tuyến hoàn hảo

Xét mô hình :$Y_i = \beta_1+\beta_2X_{2i}+\beta_3X_{3i}+ U_i$ (1)

Giả sử : $X_{3i} = \lambda X_{2i} \rightarrow x_{3i} = \lambda x_{2i}$. Theo OLS:

$$\hat{\beta}_2 = \frac{\sum x_{2i} y_i \sum x_{3i}^2 - \sum x_{2i} x_{3i} \sum x_{3i} y_i}{\sum x_{2i}^2 \sum x_{3i}^2 - (\sum x_{2i} x_{3i})^2}$$

$$\hat{\beta}_3 = \frac{\sum x_{3i} y_i \sum x_{2i}^2 - \sum x_{2i} x_{3i} \sum x_{2i} y_i}{\sum x_{2i}^2 \sum x_{3i}^2 - (\sum x_{2i} x_{3i})^2}$$

CuuDuongThanCong.com https://fb.com/tailieudientucntt

Thay $x_{3i} = \lambda_2 x_{2i}$ vào công thức :

$$\hat{\beta}_2 = \frac{\sum x_{2i} y_i (\lambda^2 \sum x_{2i}^2) - (\lambda \sum x_{2i}^2)(\lambda \sum x_{2i} y_i)}{\sum x_{2i}^2 (\lambda^2 \sum x_{2i}^2) - \lambda^2 (\sum x_{2i}^2)^2} = \frac{0}{0}$$

Tương tự : $\hat{\beta}_3 = \frac{0}{0}$

Tuy nhiên nếu thay $X_{3i} = \lambda X_{2i}$ vào hàm hồi qui (1), ta được :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 \lambda X_{2i} + U_i$$

Hay $$Y_i = \beta_1 + (\beta_2 + \lambda\beta_3) X_{2i} + U_i \quad (2)$$

Ước lượng (2), ta có : $\hat{\beta}_1, \hat{\beta}_0 = \hat{\beta}_2 + \lambda \hat{\beta}_3$

cuu duong than cong . com

- Tóm lại, khi có đa cộng tuyến hoàn hảo thì không thể ước lượng được các hệ số trong mô hình mà chỉ có thể ước lượng được một tổ hợp tuyến tính của các hệ số đó.

2. Trường hợp có đa cộng tuyến không hoàn hảo

Thực hiện tương tự như trong trường hợp có đa cộng tuyến hoàn hảo nhưng với $X_{3i} = \lambda X_{2i} + V_i$ → Vẫn có thể ước lượng được các hệ số trong mô hình.

cuu duong than cong . com

III. Hậu quả của đa cộng tuyến

1. Phương sai và hiệp phương sai của các ước lượng OLS lớn.
2. Khoảng tin cậy rộng hơn
3. Thống kê t nhỏ nên tăng khả năng các hệ số ước lượng không có ý nghĩa
4. R^2 cao nhưng thống kê t nhỏ.
5. Dấu của các ước lượng có thể sai.

cuu duong than cong . com

6. Các ước lượng OLS và sai số chuẩn của chúng trở nên rất nhạy với những thay đổi nhỏ trong dữ liệu.

7. Thêm vào hay bớt đi các biến cộng tuyến với các biến khác, mô hình sẽ thay đổi về dấu hoặc độ lớn của các ước lượng.

cuu duong than cong . com

IV. Cách phát hiện đa cộng tuyến

1. Hệ số R^2 lớn nhưng thống kê t nhỏ.

2. Tương quan cặp giữa các biến giải thích (độc lập) cao.

Ví dụ : $Y_i = \beta_1+\beta_2X_{2i}+\beta_3X_{3i}+ \beta_4X_{4i} + U_i$

Nếu r_{23} hoặc r_{24} hoặc r_{34} cao → có ĐCT. Tuy nhiên điều ngược lại không đúng, nếu các r nhỏ thì chưa biết có đa cộng tuyến hay không.

3. Sử dụng mô hình hồi qui phụ.

cuu duong than cong . com

Xét : $Y_i = \beta_1+\beta_2X_{2i}+\beta_3X_{3i}+ \beta_4X_{4i} + U_i$

<u>Cách sử dụng mô hình hồi qui phụ như sau</u> :

- Hồi qui mỗi biến độc lập theo các biến độc lập còn lại. Tính R^2 cho mỗi hồi qui phụ :

Hồi qui $X_{2i} = \alpha_1+\alpha_2X_{3i}+\alpha_3X_{4i}+u_{2i} \rightarrow R_2^2$

Hồi qui $X_{3i} = \lambda_1+ \lambda_2X_{2i}+ \lambda_3X_{4i}+u_{3i} \rightarrow R_3^2$

Hồi qui $X_{4i} = \gamma_1+\gamma_2X_{2i}+ \gamma_3X_{3i}+u_{4i} \rightarrow R_4^2$

- Kiểm định <u>các</u> giả thiết

$$H_0 : R_j^2 = 0 \quad \forall j = 2 \ldots 4$$

- Nếu chấp nhận <u>các</u> giả thiết trên thì không có đa cộng tuyến giữa các biến độc lập.

CuuDuongThanCong.com https://fb.com/tailieudientucntt

4. Sử dụng nhân tử phóng đại phương sai

$$VIF_j = \frac{1}{1 - R_j^2}$$

R_j^2 là hệ số xác định của mô hình hồi qui phụ X_j theo các biến độc lập khác.

Nếu có đa cộng tuyến thì VIF lớn.

$VIF_j > 10$ thì X_j có đa cộng tuyến cao với các biến khác.

* Với mô hình 3 biến thì $VIF = \frac{1}{1 - r_{23}^2}$

cuu duong than cong . com