

MÔ HÌNH HỒI QUY TUYẾN TÍNH TỪNG KHÚC

Nhóm 4

Võ Thị Thái

Nguyễn Bảo Việt

Tạ Thị Quỳnh

Đỗ Thị Phương

Đoàn Minh Phương

Nguyễn Thị Thu Vân

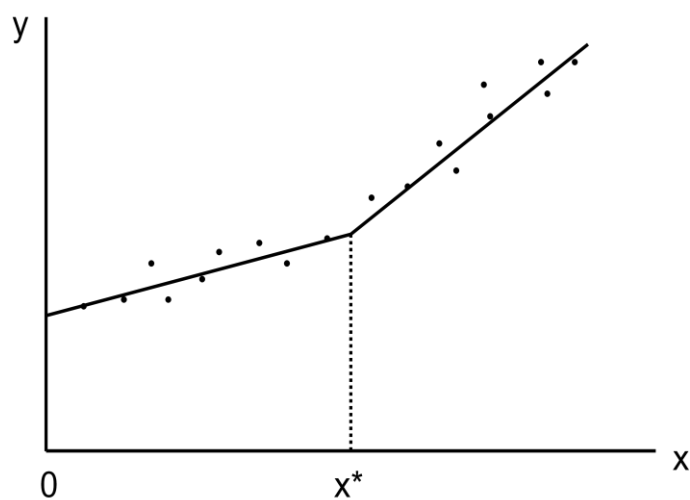
I. ĐẶT VẤN ĐỀ

Trong phân tích hồi quy, kỹ thuật biến đổi không những cho phép chúng ta lượng hóa được những thuộc tính về chất mà nó còn được sử dụng như là một trong những công cụ giải thích cho sự khác nhau về hệ số chặn hay độ dốc hoặc cả hệ số chặn và độ dốc.

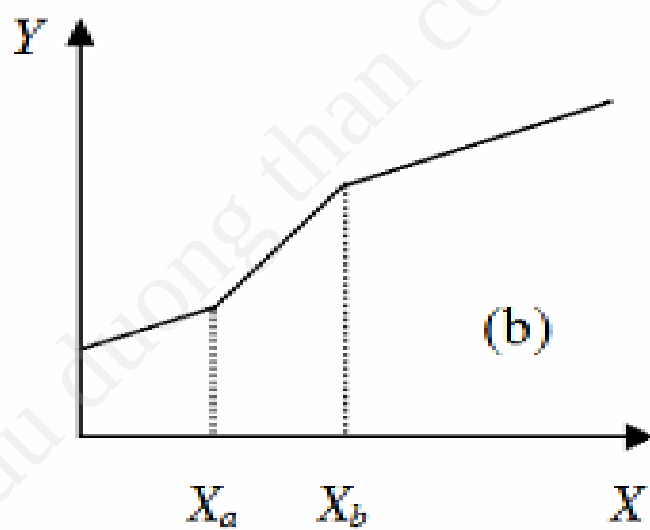
Trong trường hợp hàm hồi quy có sự thay đổi về mặt cấu trúc với từng khoảng giá trị khác nhau của biến độc lập, nghĩa là mỗi một khoảng tương ứng với một cấu trúc khác nhau của hàm hồi quy, lúc đó ta có thể dùng hàm hồi quy tuyến tính từng khúc để biểu diễn chung cho các cấu trúc khác nhau. Mô hình hồi quy tuyến tính từng khúc được phân tích sau đây là một ví dụ điển hình cho sự thay đổi độ dốc (có sự thay đổi về kết cấu) được giải thích thông qua kỹ thuật biến đổi hạn chế trong trường hợp mô hình được ước lượng vẫn đúng là một mô hình liên tục.

Việc phân tích mô hình hồi quy tuyến tính từng khúc là hoàn toàn phù hợp và cần thiết. Bởi lẽ, trong thực tế, một sự kiện, biến cố hay một quyết định, chính sách,... được đưa ra đều có thể ảnh hưởng đến cấu trúc của mô hình. Ví dụ như công ty trả tiền hoa hồng cho các đại lý dựa vào doanh thu (trong đó cách tính tiền hoa hồng là khác nhau khi doanh thu đạt dưới hoặc trên một mức nhất định); tác động của thuế nhập khẩu đến mức tiêu dùng một nước trong thời kỳ trước và sau khi chuyển đổi hoặc trước và sau khủng hoảng, chiến tranh...

- Một số mô hình hồi quy tuyến tính từng khúc thực quan



Hình 1 : Mô hình có một thay đổi về cấu trúc



Hình 2 : Mô hình có hai thay đổi về cấu trúc

II. PHƯƠNG PHÁP ƯỚC LƯỢNG VÀ PHÂN TÍCH MÔ HÌNH.

Xét mô hình tác động của thunhập đến tiết kiệm của Mỹ trong thời kỳ 1973-1992.

Trong thời kỳ này vào năm 1982 Mỹ rơi vào khủng hoảng kinh tế.

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 (X_t - X_{t_0}) D_t + U_t$$

Y_t : tiết kiệm

X_t : thunhập

X_{t_0} : thunhập trong năm khủng hoảng 1982

t_0 : năm 1982

$$D_t = \begin{cases} 1 & \text{nếu } t > t_0 \\ 0 & \text{nếu } t \leq t_0 \end{cases}$$

Với giả thiết $E(U_t) = 0$, chúng ta

thấy ngay rằng trung bình của tiết kiệm trong những năm trước khủng hoảng kinh tế ở Mỹ là:

$$E(Y_t | D_t = 0, X_t) = \beta_1 + \beta_2 X_t$$

Với $D_t = 1$ hay với những năm sau khủng hoảng, ta có:

$$E(Y_t | D_t = 1, X_t) = \beta_1 - \beta_3 X_{t_0} + (\beta_2 + \beta_3) X_t$$

Vậy β_2 cho độ dốc của đường hồi quy trước khủng hoảng.

$(\beta_2 + \beta_3)$ cho độ dốc của đường hồi quy sau khủng hoảng.

Một điều cần lưu ý rằng, ở đây không có sự gián đoạn vì:

$$\begin{aligned} E(Y_{t_0}) &= \beta_1 + \beta_2 X_{t_0} \\ &= \beta_1 - \beta_3 X_{t_0} + (\beta_2 + \beta_3) X_{t_0} \end{aligned}$$

Khi $\beta_3 = 0$ thì mô hình ta đang xét trở thành mô hình của đường thẳng, vì thế $\beta_3 = 0$ sẽ cung cấp cho ta kiểm định đơn giản về sự thay đổi cấu trúc.

$$\begin{cases} H_0: \beta_3 = 0 \\ H_1: \beta_3 \neq 0 \end{cases}$$

Nếu mô hình có nhiều thay đổi về cấu trúc với t_0 và t_1 thì mô hình thích hợp sẽ là:

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 (X_t - X_{t_0}) D_1 + \beta_4 (X_t - X_{t_1}) D_2 + U_t$$

Trong đó:

$$D_1 = \begin{cases} 1 & \text{nếu } t > t_0 \\ 0 & \text{nếu } t \leq t_0 \end{cases}$$

$$D_2 = \begin{cases} 1 & \text{nếu } t > t_1 \\ 0 & \text{nếu } t \leq t_1 \end{cases}$$

Vậy phương trình cho 3 giai đoạn trên như sau:

$$E(Y_t) = \begin{cases} \beta_1 + \beta_2 X_t & \text{với } 0 < t < t_0 \\ \beta_1 - \beta_3 X_{t_0} + (\beta_2 + \beta_3) X_t & \text{với } t_0 < t < t_1 \\ \beta_1 - \beta_3 X_{t_0} - \beta_4 X_{t_1} + (\beta_2 + \beta_3 + \beta_4) X_t & \text{với } t > t_1 \end{cases}$$

III. ÁP DỤNG

Thu nhập và tiết kiệm của Mỹ trong thời kỳ 1973-1992 được cho bởi bảng số liệu sau:

Năm	Tiết kiệm (Y_t)	Thu nhập (X_t)	D_t	Chênh lệch so với X_{t_0}	$(X_t - X_{t_0})D_t$
1973	89,6	965	0	-1382,3	0
1974	97,6	1054,2	0	-1293,1	0
1975	104,4	1159,2	0	-1188,1	0
1976	96,4	1273	0	-1074,3	0
1977	92,5	1401,4	0	-945,9	0
1978	112,6	1580,1	0	-767,2	0
1979	130,1	1769,5	0	-577,8	0
1980	161,8	1973,3	0	-374	0
1981	199,1	2200,2	0	-147,1	0
1982	205,5	2347,3	0	0	0
1983	167	2522,4	1	175,1	175,1
1984	235,7	2810	1	462,7	462,7
1985	206,2	3002	1	654,7	654,7
1986	196,5	3187,6	1	840,3	840,3
1987	168,4	3363,1	1	1015,8	1015,8
1988	189,1	3640,8	1	1293,5	1293,5
1989	187,8	3894,5	1	1547,2	1547,2
1990	208,7	4166,8	1	1819,5	1819,5
1991	246,4	4343,7	1	1996,4	1996,4
1992	272,6	4613,7	1	2266,4	2266,4
$X_{1982} = X_0 = 2347,3$					

Trong đó:

- t_0 : năm 1982, năm xảy ra khủng hoảng kinh tế ở Mỹ
- X_{t_0} : thu nhập tại năm khủng hoảng
- Y_t : tiết kiệm
- X_t : thu nhập
- $D_t = \begin{cases} 1 & \text{nếu } t > 0 \\ 0 & \text{nếu nhận các giá trị khác} \end{cases}$

Vấn đề cần giải quyết

a. Ước lượng hàm hồi quy tuyến tính dựa theo mẫu số liệu trên

b. Cuộc khủng hoảng năm 1982 có ảnh hưởng đến tiêu dùng của Mỹ hay không (kiểm định giả thuyết $H_0: \beta_3 = 0$)

1. Tìm mô hình hồi quy và kiểm định $\beta_3 = 0$ theo phương pháp ma trận

• Tìm mô hình hồi quy

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 (X_t - X_{t0}) D_t + U_t (*)$$

$$\text{Đặt } X_t = X_{2i}; (X_t - X_{t0}) D_t = X_{3i}$$

Ta viết lại phương trình (*):

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_t + \hat{\beta}_3 (X_t - X_{t0}) D_t$$

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} \quad Y = \begin{bmatrix} Y_1 \\ \dots \\ Y_{26} \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_1 & X_{31} \\ \dots & \dots & \dots \\ 1 & X_{220} & X_{320} \end{bmatrix} \quad X' = \begin{bmatrix} 1 & \dots & 1 \\ X_{21} & \dots & X_{220} \\ X_{31} & \dots & X_{320} \end{bmatrix}$$

$$X'X = \begin{bmatrix} 20 & \sum X_{2i} & \sum X_{3i} \\ \sum X_{2i} & \sum X_{2i}X_{2i} & \sum X_{2i}X_{3i} \\ \sum X_{3i} & \sum X_{3i}X_{2i} & \sum X_{3i}X_{3i} \end{bmatrix} = \begin{bmatrix} 20 & 51267,8 & 12071,6 \\ 51267,8 & 157523790,4 & 47246736,06 \\ 12071,6 & 47246736,06 & 18911069,38 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} \sum Y_i \\ \sum X_{2i}Y_i \\ \sum X_{3i}Y_i \end{bmatrix} = \begin{bmatrix} 3368 \\ 9723297,7 \\ 2634107,16 \end{bmatrix}$$

$$\text{Ta có: } X'X \hat{\beta} = X'Y$$

$$\Rightarrow \hat{\beta} = (X'X)^{-1} X'Y$$

$$A_{jj} = \begin{bmatrix} 7,466892609 * 10^{14} & -3,991852237 * 10^{11} & 5,206720268 * 10^{11} \\ -3,991852237 * 10^{11} & 232497861 & -326050346,7 \\ 5,206720268 * 10^{11} & -326050346,7 & 522088491,2 \end{bmatrix}$$

$$\Rightarrow \text{Det}(X'X) = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$$

$$= 20 * 7,466892609 * 10^{14} - 51267,8 * 3,991852237 * 10^{11} +$$

$$12071,6 * 5,206720268 * 10^{11} = 7,537814451 * 10^{14}$$

$$\Rightarrow \hat{\beta} = \frac{1}{7,537814451 * 10^{14}} \begin{bmatrix} 7,466892609 * 10^{14} & -3,991852237 * 10^{11} & 5,206720268 * 10^{11} \\ -3,991852237 * 10^{11} & 232497861 & -326050346,7 \\ 5,206720268 * 10^{11} & -326050346,7 & 522088491,2 \end{bmatrix} X'Y$$

$$\hat{\beta} = \frac{1}{7,537814451 \cdot 10^{14}} \begin{bmatrix} 4,95857704 \cdot 10^{15} \\ 5,733853094 \cdot 10^{13} \\ -4,142416707 \cdot 10^{13} \end{bmatrix} = \begin{bmatrix} 6,578268903 \\ 0,076067846 \\ -0,054955142 \end{bmatrix}$$

=> Ta có hàm hồi quy mẫu:

$$\hat{Y} = 6,578268903 + 0,076067846X_t - 0,054955142(X_t - X_{t0})Dt$$

- Trung bình của tiết kiệm trong những năm trước khủng hoảng kinh tế ở Mỹ là:

$$E(Y_t | D_t = 0, X_t) = 6,578268903 + 0,076067846X_t$$

- Ước lượng trung bình tiết kiệm của những năm sau khủng hoảng:

$$E(Y_t | D_t = 1, X_t) = \beta_1 - \beta_3 X_{10} + (\beta_2 + \beta_3)X_t = 135,5744737 + 0,021112704X_t$$

• Kiểm định giả thuyết

$$\begin{cases} H_0: \beta_3 = 0 \\ H_1: \beta_3 \neq 0 \end{cases}$$

Xây dựng tiêu chuẩn kiểm định :

$$T = \frac{\hat{\beta}_3 - 0}{se(\hat{\beta}_3)}, \text{ nếu } H_0 \text{ đúng thì } T \sim T^{(n-k)}$$

→ Tìm được $t_{\alpha/2}^{(n-k)}$ sao cho $P(|T| > t_{\alpha/2}^{(n-k)}) = \alpha$

→ Miền bác bỏ : $W_\alpha = \{t_{qs} : |t_{qs}| > t_{\alpha/2}^{(n-k)}\}$

$$\text{Ta có : } \sum e_i^2 = Y'Y - \hat{\beta}'X'Y = \sum Y_i^2 - \hat{\beta}'X'Y$$

$$= 626497,36 - [6,578268903 \quad 0,076067846 \quad -0,054955142] * \begin{bmatrix} 3368 \\ 9723297,7 \\ 2634107,16 \end{bmatrix}$$

$$= 626497,36 - 617028,1887 = 9469,1713$$

$$\rightarrow Se(\hat{\beta}_3) = \sqrt{C_{jj} * \frac{\sum e_i^2}{n-k}} = \sqrt{\frac{522088491,2}{7,537814451 \cdot 10^{14}} * \frac{9469,1713}{17}} = 0,019641781$$

$$\rightarrow t_{qs} = \frac{-0,054955142}{0,019641781} = -2,797869603$$

$$t_{\alpha/2}^{(n-k)} = t_{0,025}^{17} = 2,110$$

→ t_{qs} thuộc miền bác bỏ $W_\alpha \rightarrow$ bác bỏ H_0

→ Kết luận : Với mức ý nghĩa là 0,05 ta có thể kết luận rằng cuộc khủng hoảng năm 1982 có ảnh hưởng đến thặng dư của Mỹ và lúc đó cấu trúc của mô hình tiết kiệm do

tác động của thính nghiệm có sự thay đổi trước và sau khủng hoảng
(độ dốc của đường lượng mô hình thay đổi độ dốc sau khi khủng hoảng xảy ra)

2. Ứng dụng phần mềm Eviews

Sau khi chạy phần mềm Eviews, kết quả thu được được thể hiện như sau:

- Ước lượng các hệ số β_j

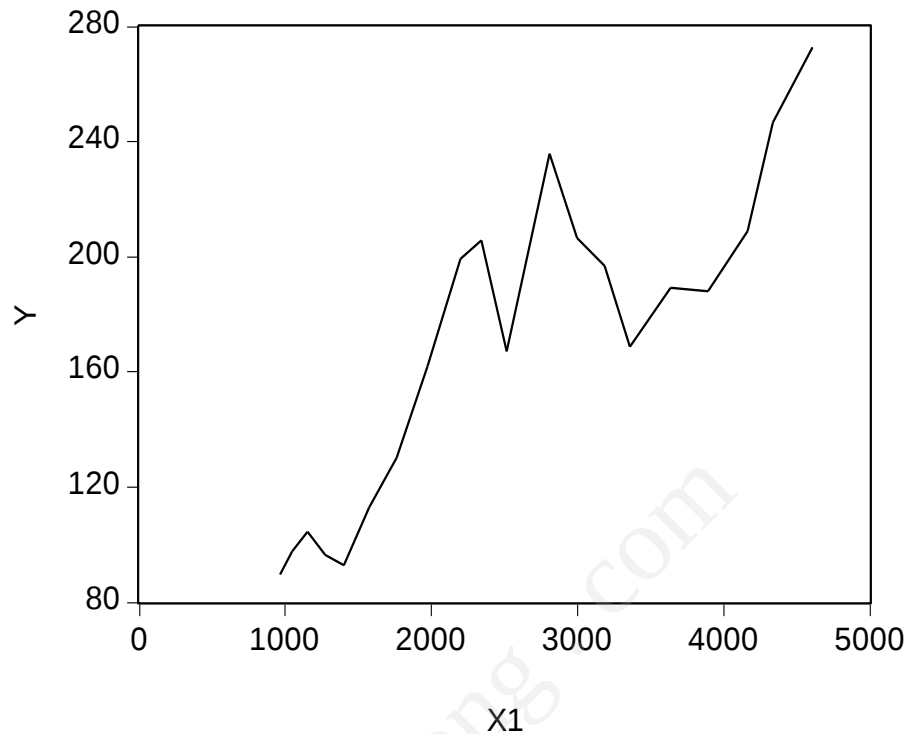
Dependent Variable: Y
Method: Least Squares
Date: 11/09/11 Time: 14:20
Sample: 1 20
Included observations: 20

x	Coefficient	Std. Error	t-Statistic	Prob.
C	6.578267	23.48977	0.280048	0.7828
X1	0.076068	0.013107	5.803404	0.0000
X2	-0.054955	0.019642	-2.797870	0.0124
R-squared	0.840388	Mean dependent var	168.4000	
Adjusted R-squared	0.821610	S.D. dependent var	55.87870	
S.E. of regression	23.60106	Akaike info criterion	9.297942	
Sum squared resid	9469.172	Schwarz criterion	9.447301	
Log likelihood	-89.97942	F-statistic	44.75412	
Durbin-Watson stat	1.193294	Prob(F-statistic)	0.000000	

→ Hàm hồi quy được ước lượng:

$$\hat{Y}_t = 6.578267 + 0.076068X_t - 0.054955(X_t - X_{t0})D_t$$

- Kết quả được thể hiện trên đồ thị :



- Kiểm định giả thuyết

$$\begin{cases} H_0: \beta_3 = 0 \\ H_1: \beta_3 \neq 0 \end{cases}$$

Ta có :

Với $\beta_3 = -0.054955$ thì $P_{\text{value}} = 0.0124 < 0.05$

→ Giả thuyết H_0 bị bác bỏ

→ Kết luận : Với mức ý nghĩa là 0,05 ta có thể kết luận rằng cuộc khủng hoảng năm 1982 có ảnh hưởng đến thủng lỗ của Mỹ và lúc đó cấu trúc của mô hình tiết kiệm do tác động của thủng lỗ có sự thay đổi trước và sau khủng hoảng (độ dốc của đường ước lượng mô hình thay đổi độ dốc sau khi khủng hoảng xảy ra)

IV. Nhận xét

- Mô hình hồi quy tuyến tính từng khúc là một trong những trường hợp khai thác được ưu thế của kỹ thuật biến đổi
- Công cụ hiệu quả và có ý nghĩa thực tế trong việc kiểm định sự thay đổi cũng như hướng thay đổi trong mô hình kinh tế do tác động ở những mức độ khác nhau của các yếu tố ảnh hưởng đến nó.