

Chương 8

LỰA CHỌN MÔ HÌNH HỒI QUY

1. Các tiêu chuẩn của mô hình

Tính tiết kiệm : mô hình càng đơn giản càng tốt

Tính đồng nhất : các tham số ước lượng là duy nhất cho cùng một tập hợp số liệu

Tính thích hợp : R^2 và R^2 hiệu chỉnh càng gần 1 càng tốt

Tính bền vững : mô hình phải dựa trên một cơ sở lý thuyết nào đó

Có khả năng dự báo tốt : mô hình cho kết quả dự báo sát với thực tế

2. Cách tiếp cận để lựa chọn mô hình

a. Xác định số biến độc lập

Có hai hướng tiếp cận

- Từ đơn giản đến tổng quát : *Bổ sung biến độc lập từ từ vào mô hình*
- Từ tổng quát đến đơn giản : *Đầu tiên, xét mô hình đầy đủ các biến độc lập đã được xác định. Sau đó tiến hành loại trừ những biến không quan trọng ra khỏi mô hình*

2. Cách tiếp cận để lựa chọn mô hình

b. Kiểm tra mô hình có vi phạm giả thiết hay không

Kiểm tra các “bệnh của mô hình”

- Đa cộng tuyến
- Tự tương quan
- Phương sai thay đổi

2. Cách tiếp cận để lựa chọn mô hình

c. Chọn dạng hàm

Cần dựa vào

- Các lý thuyết kinh tế
- Các kết quả thực nghiệm
- Đồ thị biểu diễn

2. Cách tiếp cận để lựa chọn mô hình

d. Một số tiêu chuẩn khác

Giá trị của hàm hợp lý log-likelihood(L)

$$L = -\frac{n}{2} \ln \sigma^2 - \frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum U_i^2$$

- Giá trị của L càng lớn chứng tỏ mô hình càng phù hợp

2. Cách tiếp cận để lựa chọn mô hình

d. Một số tiêu chuẩn khác

Tiêu chuẩn AIC (Akaike info criterion)

$$AIC = \frac{RSS}{n} e^{2k/n}$$

➤ Giá trị của AIC càng nhỏ chứng tỏ mô hình càng phù hợp

2. Cách tiếp cận để lựa chọn mô hình

d. Một số tiêu chuẩn khác

Tiêu chuẩn Schwarz (Schwarz criterion)

$$SC = \frac{RSS}{n} n^{2k/n}$$

➤ Giá trị của SC càng nhỏ chứng tỏ mô hình càng phù hợp

2. Cách tiếp cận để lựa chọn mô hình

d. Một số tiêu chuẩn khác

Nếu chú ý đến độ phức tạp của mô hình thì thường chú ý đến tiêu chuẩn SC

Nếu xét số liệu theo thời gian thì thường dùng tiêu chuẩn AIC

Lưu ý là biến phụ thuộc xuất hiện trong mô hình phải cùng dạng

Kết quả hồi quy bằng Eviews như sau :

Equation: UNTITLED - Workfile: HOIQUYBO1:Unlitled				
View Proc Object Print Name Freeze Estimate Forecast Stats Resids				
Dependent Variable: Y				
Method: Least Squares				
Date: 04/01/09 Time: 10:15				
Sample: 1 10				
Included observations: 10				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14.99215	2.922713	5.129531	0.0014
X2	0.761780	0.283438	2.687639	0.0312
X3	-0.589005	0.244593	-2.408105	0.0469
R-squared	0.960934	Mean dependent var	16.50000	
Adjusted R-squared	0.949773	S.D. dependent var	2.549610	
S.E. of regression	0.571362	Akaike info criterion	1.961807	
Sum squared resid	2.285340	Schwarz criterion	2.052582	
Log likelihood	-6.809035	Hannan-Quinn criter.	1.862226	
F-statistic	86.09278	Durbin-Watson stat	1.800073	
Prob(F-statistic)	0.000012			

3. Các sai lầm thường gặp khi chọn mô hình

a. Bỏ sót biến thích hợp

Giả sử mô hình đúng là :

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + U_i \quad (a)$$

Nhưng ta lại chọn mô hình :

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + V_i \quad (b)$$

→ hậu quả :

3. Các sai lầm thường gặp khi chọn mô hình

b. Thừa biến

Giả sử mô hình đúng là :

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

Nhưng ta lại chọn mô hình (có thêm X_3):

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + V_i \quad (b)$$

→ hậu quả :

4. Phát hiện những sai lầm

a. Phát hiện thừa biến

Xét hàm hồi qui : $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + U_i$

- Trường hợp nghi ngờ X_5 là biến thừa \rightarrow kiểm định

$$H_0 : \beta_5 = 0 \text{ (Kiểm định bằng cách nào?)}$$

Nếu chấp nhận $H_0 \rightarrow X_5$ không cần thiết. (Có thể sử dụng *redundant test của Eviews*)

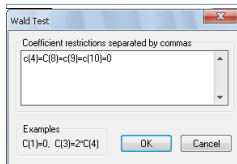
Trường hợp nghi ngờ X_3 và X_5 là các biến không cần thiết \rightarrow kiểm định giả thiết đồng thời

$$H_0 : \beta_3 = \beta_5 = 0$$

(Sử dụng kiểm định Wald)

Kiểm định Wald cho mô hình sau .

Dependent Variable: Y Method: Least Squares Date: 11/24/08 Time: 08:16 Sample: 1 106 Included observations: 106				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.206821	0.386368	18.65274	0.0000
X2	-0.000749	0.000252	-2.969765	0.0038
X3	0.000550	0.000225	2.447388	0.0162
X4	-0.007568	0.053629	-0.141111	0.8891
D1	-0.760411	0.279496	-2.720645	0.0077
D2	0.200568	0.173248	1.157691	0.2499
D3	-0.153518	0.060297	-2.546040	0.0125
D4	-0.047001	0.264910	-0.177423	0.8595
D5	-0.106459	0.261804	-0.406637	0.6852
D6	0.095354	0.380443	0.250640	0.8026
R-squared	0.233728		Mean dependent var	6.638113
Adjusted R-squared	0.161891		S.D. dependent var	0.761072
S.E. of regression	0.696749		Akaike info criterion	2.204804
Sum squared resid	46.60404		Schwarz criterion	2.456072
Log likelihood	-105.8546		F-statistic	3.263652
Durbin-Watson stat	1.362529		Prob(F-statistic)	0.001717

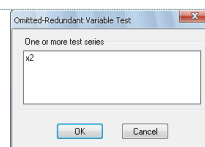


$P_value = 0,9688 > 0,05$
 \rightarrow chấp nhận H_0

Wald Test:
Equation: EQ01

Test Statistic	Value	df	Probability
F-statistic	0.136763	(4, 96)	0.9683
Chi-square	0.547052	4	0.9688

Redundant variables Test



$P_value = 0.0022 < 0,05$
 \rightarrow bác bỏ $H_0 \Rightarrow$ Không thừa biến X_2

Redundant Variables: X2

F-statistic	8.819505	Prob. F(1,96)	0.003764
Log likelihood ratio	9.316518	Prob. Chi-Square(1)	0.002271

4. Phát hiện những sai lầm

b. Kiểm định các biến bị bỏ sót

Xét mô hình : $Y_i = \beta_1 + \beta_2 X_i + U_i$ (*)

Giả sử nghi ngờ mô hình đã bỏ sót biến $Z \rightarrow$ kiểm tra bằng cách :

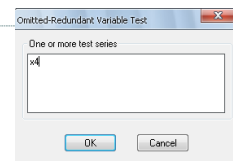
- Nếu có số liệu của Z :

+ Hồi qui mô hình $Y_i = \beta_1 + \beta_2 X_i + \beta_3 Z_i + U_i$

+ Kiểm định $H_0 : \beta_3 = 0$. Nếu bác bỏ H_0 thì mô hình ban đầu đã bỏ sót biến Z .

- Hoặc dùng Omitted variable test

Omitted variables Test



$P_value = 0.8200 > 0,05$
 \rightarrow chấp nhận $H_0 \Rightarrow$ Không bỏ sót biến X_4

Omitted Variables: X4

F-statistic	0.049798	Prob. F(1,102)	0.823861
Log likelihood ratio	0.051739	Prob. Chi-Square(1)	0.820065

Kiểm định RESET của Ramsey :

Nếu không có số liệu của Z : dùng kiểm định RESET của Ramsey.

Ramsey đề xuất sử dụng \hat{Y}_i^2, \hat{Y}_i^3 làm xấp xỉ cho Z_i .

Bước 1 : Hồi qui mô hình (*), thu lấy \hat{Y}_i^2, \hat{Y}_i^3

Bước 2 : Hồi qui Y_i theo các biến độc lập trong (*) và \hat{Y}_i^2, \hat{Y}_i^3 (mô hình này gọi là mô hình (new)) .

Bước 3 : Kiểm định H_0 : các hệ số của \hat{Y}_i^2, \hat{Y}_i^3 đồng thời bằng 0.

Nếu bác bỏ $H_0 \rightarrow$ mô hình (*) đã bỏ sót biến.

Kiểm định RESET của Ramsey :

Ramsey RESET Test:

F-statistic	8.513130	Prob. F(4,195)	0.000002
Log likelihood ratio	33.79987	Prob. Chi-Square(4)	0.000001

$P_value = 0.0000 < 0,05 \rightarrow$ mô hình ban đầu bỏ sót biến.

Hết