Tự tương quan
(Autocorrelation)

Đinh Công Khải

Tháng 04/2016

Nội dung

1. Tự tương quan là gì?
2. Hậu quả của việc ước lượng bỏ qua tự tương quan?
3. Làm sao để phát hiện tự tương quan?
4. Các biện pháp khắc phục?

Tự tương quan là gì?

- Giả thuyết không có tự tương quan của mô hình CLRM

  $E(u_i\ u_j) = 0$ với $i \neq j$

- Có tự tương quan

  $E(u_i\ u_j) \neq 0$ với $i \neq j$

Tự tương quan là gì?

- Tự tương quan (autocorrelation)

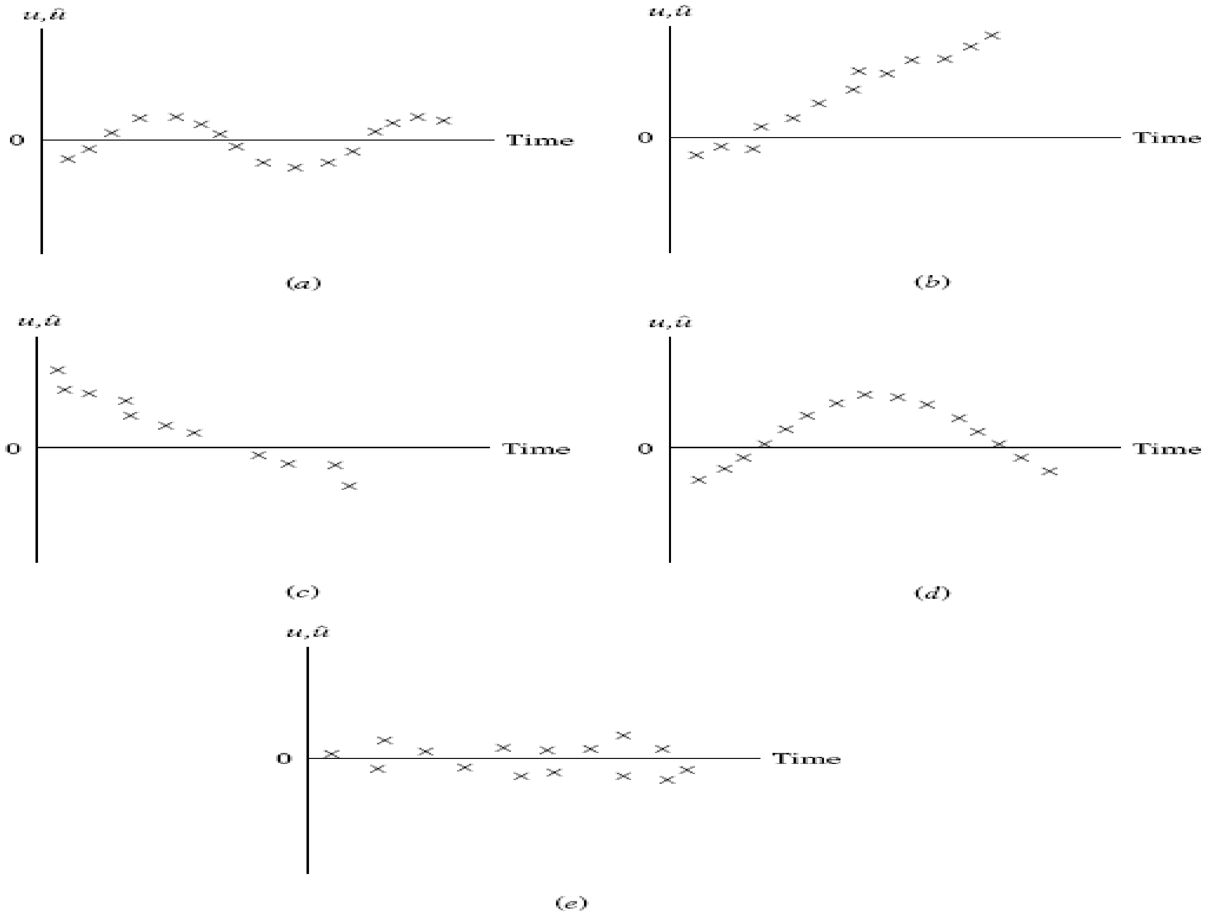
  $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + .. + \rho_p u_{t-p} + \varepsilon_t$

  AR(p): cơ chế tự hồi qui bậc p (p-order autoregressive scheme)

- Tương quan chuỗi (serial correlation)

  $u_t = v_t + \lambda v_{t-1} + \varepsilon_t$

Tự tương quan



12.1 Patterns of autocorrelation and nonautocorrelation.

Nguyên nhân của tự tương quan là gì?

- Quán tính (GDP_t, CPI_t, …)
- Bỏ sót các biến quan trọng

Hàm đúng: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + u_t$

Hàm thiếu biến: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + v_t$

$v_t = \beta_4 X_{4t} + u_t$

Nguyên nhân của tự tương quan là gì?

- Dạng hàm không đúng

Hàm đúng: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X^2_{3i} + u_i$

Hàm sai: $Y_i = \beta_1 + \beta_2 X_{2i} + v_i$

$v_i = \beta_3 X^2_{3i} + u_i$

- Hiện tượng Coweb

$Q^s_t = \beta_1 + \beta_2 P_{t-1} + u_t$

- Các độ trễ

Tiêu dùng $_t = \beta_0 + \beta_1$ Thu nhập$_t + \beta_2$Tiêu dùng $_{t-1} + u_t$

Ước lượng OLS khi có tự tương quan

- Giả định: $Y_t = \beta_1 + \beta_2 X_{2t} + u_t$

$$u_t = \rho_1 u_{t-1} + \varepsilon_t \qquad\qquad AR(1)$$

trong đó sai số ngẫu nhiên ε_t có tính nhiễu trắng khi:

$$E(\varepsilon_t) = 0$$

$$E(\varepsilon^2_t) = \sigma^2 = const$$

$$E(\varepsilon_t \varepsilon_{t-s}) = 0 \text{ với } s \neq 0$$

Ước lượng OLS khi có tự tương quan

- Trong trường hợp có AR các ước lượng OLS vẫn không thiên lệch.
- Tuy nhiên nếu sử dụng OLS không tính đến tự tương quan

$$\operatorname{var}(\hat{\beta}_2)_{OLS} = \frac{\sigma^2}{\Sigma x_t^2}$$

- Sử dụng OLS có tính đến AR

$$\operatorname{var}(\hat{\beta}_2)_{AR(1)} = \frac{\sigma^2}{\Sigma x_t^2} + \frac{2\sigma^2}{\Sigma x_t^2}\left(\rho \frac{\Sigma x_t x_{t+1}}{\Sigma x_t^2} + \ldots + \rho^{n-1}\frac{x_1 x_n}{\Sigma x_t^2}\right)$$

Ước lượng OLS khi có tự tương quan

- Trong trường hợp ρ > 0 và các quan sát X tương quan nghịch biến hoặc ρ < 0 và các quan sát X tương quan đồng biến

$$\operatorname{var}(\hat{\beta}_2)_{OLS} > \operatorname{var}(\hat{\beta}_2)_{AR(1)}$$

- Các kiểm định giả thuyết t và F không còn hiệu lực
- Phương pháp GLS sẽ cho ước lượng BLUE

$$\operatorname{var}(\hat{\beta}_2)_{GLS} < \operatorname{var}(\hat{\beta}_2)_{OLS}, \operatorname{var}(\hat{\beta}_2)_{AR(1)}$$

Kiểm định tự tương quan

1) Kiểm định bằng phương pháp đồ thị
2) Kiểm định Durbin-Watson (d)

Điều kiện áp dụng:

- Các nhiễu được tạo từ AR(1): $u_t = \rho u_{t-1} + \varepsilon_t$
- Không áp dụng cho mô hình có biến độc lập Y_{t-1}

$$Y_t = \beta_1 + \beta_2 X_{2t} + \ldots + \beta_k X_{kt} + \gamma Y_{t-1} + \varepsilon_t$$

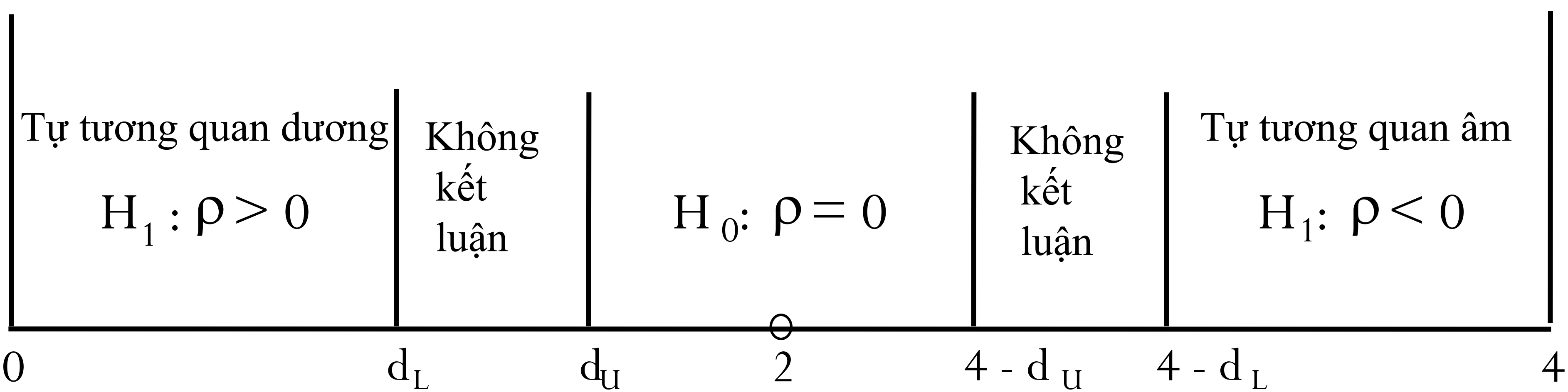
- Trị kiểm định

Kiểm định Durbin_Watson (Nguồn: Cao Hào Thi)

$$d = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2} \approx 2(1 - \hat{\rho})$$

$$\hat{\rho} \approx \frac{\sum_{t=2}^{n} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^{n} \hat{u}_t^2}$$

Giả thuyết H0: ρ=0



Kiểm định tự tương quan

3) Kiểm định tiệm cận (mẫu lớn)

Giả thuyết H_0: $\rho = 0$

$$\sqrt{n}\,\hat{\rho} \sim N(0,1)$$

4) Kiểm định Breusch-Godfrey (BG) (Kiểm định nhân tử Lagrance)

- Áp dụng cho
- ✓ $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + .. + \rho_p u_{t-p} + \varepsilon_t$ AR(p)
- ✓ Hàm hồi quy chứa các giá trị trễ của biến phụ thuộc (Y_{t-1}, Y_{t-2},..)
- ✓ $u_t = \varepsilon_t + \lambda_1 \varepsilon_{t-1} + .. + \lambda_p \varepsilon_{t-p}$

Các bước kiểm định BG

PRF: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_2 X_{3t} + \ldots + \beta_k X_{kt} + u_t$ (1)

với $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + .. + \rho_p u_{t-p} + \varepsilon_t$

Kiểm định giả thuyết:

$H_0 : \rho_1 = \rho_2 = \ldots = \rho_p = 0 \Rightarrow$ Không có AR(p)

H_1 : Có ít nhất 1 $\rho_j \neq 0$ (j = 1, p) $\Rightarrow$ Có AR(p)

<u>Bước 1</u>: Thực hiện hồi qui OLS (1) tính phần dư u_t^

<u>Bước 2</u>: Tính các giá trị trễ của u_t^

Các bước kiểm định BG

<u>Bước 3</u>: Thực hiện hồi qui phụ

$$\hat{u}_t = \alpha_1 + \alpha_2 X_{2t} + \ldots + \alpha_K X_{Kt} + \rho_1 u_{t-1} + \rho_2 u_{t-2} + \ldots. + \rho_p u_{t-p} + \upsilon_t$$

Xác định R^2_{hqp}

Trị kiểm định: $(n-p)*R^2_{hqp} \sim \chi 2\ (p)$

$(n-p)*R^2_{hqp} > \chi^2_{p,\alpha}$ hoặc p-value $< \alpha \Rightarrow$ Bác bỏ H_0

Các biện pháp khắc phục

1. Thay đổi dạng hàm số
2. Lấy sai phân

- Trong trường hợp biết trước ρ: $u_t = \rho\, u_{t-1} + \varepsilon_t$ $[\varepsilon \sim N(0, \sigma^2)]$

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

➔ $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$

➔ $\rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho u_{t-1}$

➔ $Y_t - \rho Y_{t-1} = \beta_1(1 - \rho) + \beta_2 (X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})$

➔ $Y^*_t = \beta^*_1 + \beta^*_2 X^*_t + \varepsilon_t$

Các ước lượng β^*_1 và β^*_2 là BLUE (phương pháp GLS)

Các biện pháp khắc phục

- Trong trường hợp không biết trước

  Giả định $\rho=1$ tức $u_t = u_{t-1} + \varepsilon_t$

  $Y_t = \beta_1 + \beta_2 X_t + u_t$

➔ $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$

➔ $Y_t - Y_{t-1} = \beta_2 (X_t - X_{t-1}) + (u_t - u_{t-1})$

➔ $\Delta Y_t = \beta_2 \Delta X_t + \varepsilon_t$

Chú ý: Mô hình hồi qui qua gốc tọa độ

Các biện pháp khắc phục

- Kiểm định Berenblutt-Webb (H0: $\rho=1$)

  Trị kiểm định

$$g = \frac{\sum_{t=2}^{n} \hat{e}_t^2}{\sum_{t=1}^{n} \hat{u}_t^2}$$

- $u^\wedge_t$ là phần dư của hồi qui OLS của mô hình ban đầu
- $e^\wedge_t$ là phần dư của hồi qui OLS của mô hình sai phân

➔ Sử dụng phương pháp Durbin-Watson để kiểm định

- Xác định ρ từ trị kiểm định DW

  d = 2(1- ρ)

Thủ tục COCHRANE – ORCUTT để ước lượng ρ

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \ldots + \beta_k X_{kt} + u_t \qquad (1)$$

Giả sử, $u_t = \rho u_{t-1} + \varepsilon_t$

➔ $Y_{t-1} = \beta_1 + \beta_2 X_{2(t-1)} + \beta_3 X_{3(t-1)} + \ldots + \beta_k X_{k(t-1)} + u_{t-1}$

➔ $Y_t - \rho Y_{t-1} = \beta_1(1-\rho) + \beta_2[X_{2t} - \rho X_{2(t-1)}] + \beta_3[X_{3t} - \rho X_{3(t-1)}]$

$$+ \ldots + \beta_k[X_{kt} - \rho X_{k(t-1)}] + \varepsilon_t$$

➔ $Y^*_t = \beta^*_1 + \beta^*_2 X^*_{t2} + \ldots + \beta^*_k X^*_{tk} + \varepsilon^*_t \qquad (2)$

Thủ tục COCHRANE – ORCUTT để ước lượng ρ

1. Ước lượng (1) bằng OLS tính $u\hat{}_t$ và $u\hat{}_{t-1}$

2. Hồi quy

$$\hat{u}_t = \hat{\rho}\hat{u}_{t-1} + \varepsilon_t$$

3. Dùng ρ^ trong mô hình (2) dưới đây và ước lượng β^*_k

$Y^*_t = \hat{\beta}_1^* + \hat{\beta}_2^* X^*_{t2} + \ldots + \hat{\beta}_k^* X^*_{tk} + \varepsilon\hat{}^*_t$

Thủ tục COCHRANE – ORCUTT để ước lượng ρ

4. Thay $\hat{\beta}_k^*$ vào trong (1) để tính u^**$_t$

$$\hat{u}_t^{**} = Y_t - \hat{\beta}_1^* - \hat{\beta}_2^* X_{2t} - \ldots - \hat{\beta}_k^* X_{kt}$$

5. Tiếp tục bước 2, ước lượng ρ^**, so sánh với giá trị ρ^ đã tính

$$\hat{u}_t^{**} = \hat{\rho}^{**} \hat{u}_{t-1}^{**} + w_t$$

Chú ý: Dừng quá trình khi sự thay đổi giá trị của ρ^ là không quá 0.01 (thường quá trình lặp từ 3-4 lần là đủ)