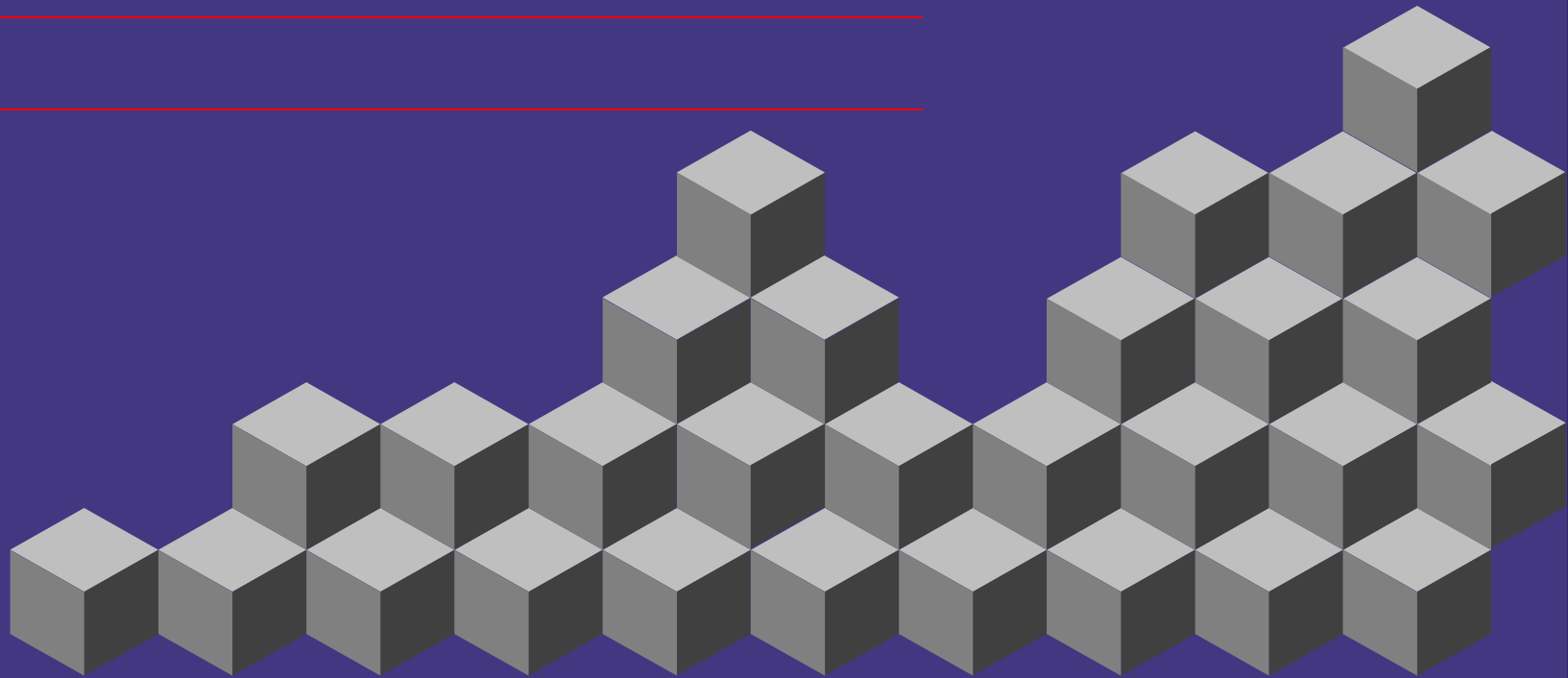


# CHƯƠNG 8

## HIỆN TƯỢNG TỰ TƯƠNG QUAN (Autocorrelation)



# TỰ TƯƠNG QUAN

## MỤC TIÊU

1. Hiểu bản chất và hậu quả của tự tương quan
2. Biết cách phát hiện tự tương quan và biện pháp khắc phục

# NỘI DUNG

- 1 Bản chất hiện tượng hiện tượng tự tương quan
- 2 Hậu quả
- 3 Cách phát hiện tự tương quan
- 4 Cách khắc phục tự tương quan

## 8.1 Bản chất

### 1. Tự tương quan là gì ?

Trong mô hình hồi qui tuyến tính cổ điển, giả định rằng không có tương quan giữa các sai số ngẫu nhiên  $u_i$ , nghĩa là:

$$\text{cov}(u_i, u_j) = 0 \quad (i \neq j)$$

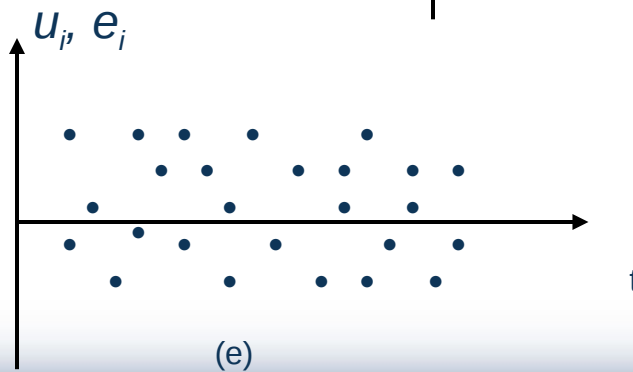
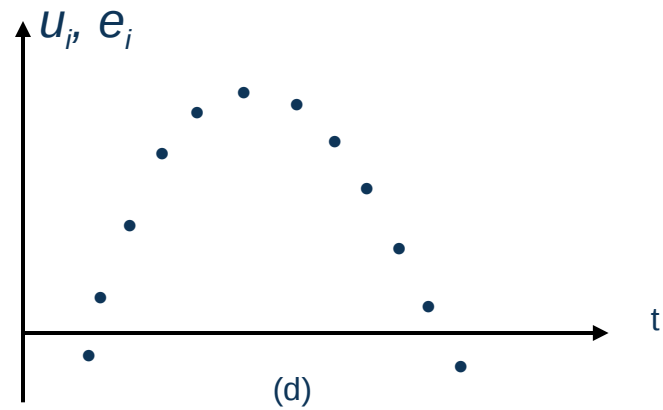
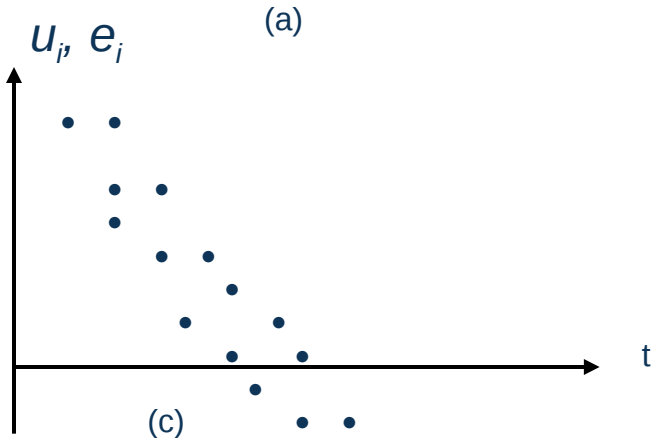
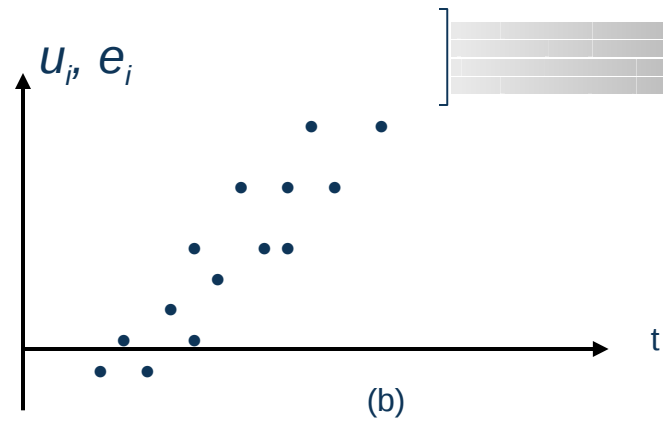
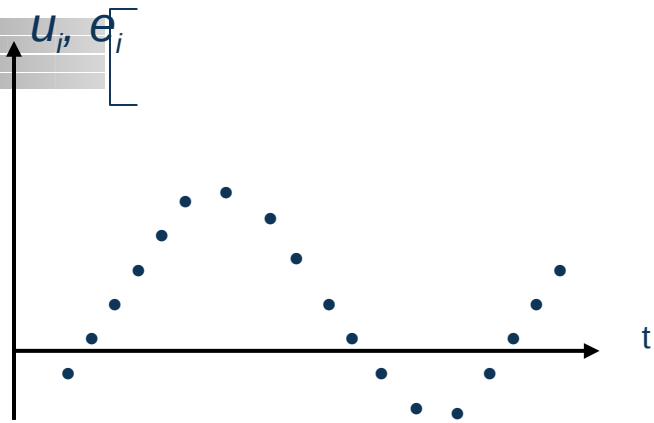
Tuy nhiên trong thực tế có thể xảy ra hiện tượng mà sai số của các quan sát lại phụ thuộc nhau, nghĩa là:

$$\text{cov}(u_i, u_j) \neq 0 \quad (i \neq j)$$

Khi đó xảy ra hiện tượng tự tương quan.

## 8.1 Bản chất

- ❖ Sự tương quan xảy ra đối với những quan sát theo không gian gọi là “**tự tương quan không gian**”.
- ❖ Sự tương quan xảy ra đối với những quan sát theo chuỗi thời gian gọi là “**tự tương quan thời gian**”.



Hình 8.1 Một số dạng biến thiên của nhiễu theo thời gian

# Nguyên nhân

## Nguyên nhân khách quan:

- ❖ Quán tính: các chuỗi thời gian mang tính chu kỳ, VD: các chuỗi số liệu thời gian về GDP, chỉ số giá, sản lượng, tỷ lệ thất nghiệp...
- ❖ Hiện tượng mạng nhện: phản ứng của cung của nông sản đối với giá thường có một khoảng trễ về thời gian:

$$Q_{St} = \beta_1 + \beta_2 P_{t-1} + u_t$$

- ❖ Độ trễ: tiêu dùng ở thời kỳ hiện tại phụ thuộc vào thu nhập và chi tiêu tiêu dùng ở thời kỳ trước đó:  $C_t = \beta_1 + \beta_2 I_t + \beta_3 C_{t-1} + u_t$

# Nguyên nhân

## Nguyên nhân chủ quan

- ❖ Hiệu chỉnh số liệu: do việc “làm trơn” số liệu → loại bỏ những quan sát “gai góc”.
- ❖ Sai lệch do lập mô hình: bỏ sót biến, dạng hàm sai.
- ❖ Phép nội suy và ngoại suy số liệu



## 8.2 Hậu quả của tự tương quan

Áp dụng OLS thì sẽ có các hậu quả:

- ❖ Các ước lượng không chệch nhưng không hiệu quả (vì phương sai không nhỏ nhất)
- ❖ Phương sai của các ước lượng là các ước lượng chệch, vì vậy các kiểm định t và F không còn hiệu quả.

## 8.2 Hậu quả của tự tương quan

- ❖  $\hat{\sigma}^2$  là ước lượng chệch của  $\sigma^2$
- ❖  $R^2$  của mẫu là ước lượng chệch (dưới) của  $R^2$  tổng thể
- ❖ Các dự báo về  $Y$  không chính xác

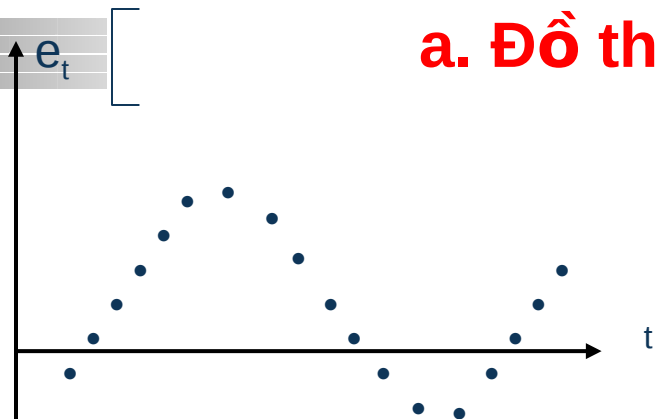
## 8.3 Cách phát hiện tự tương quan

### a. Đồ thị

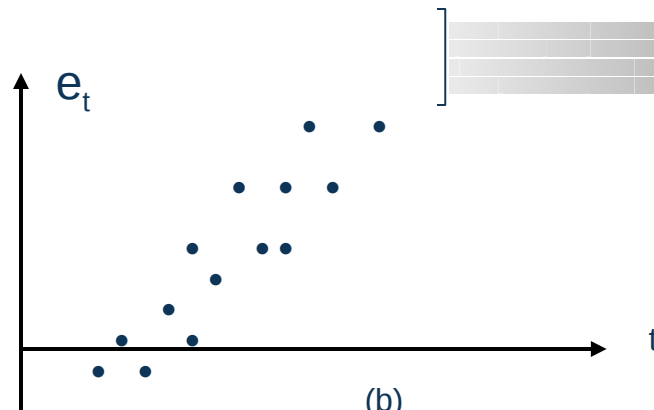
Chạy OLS cho mô hình gốc và thu thập  $e_t$ .

Vẽ đường  $e_t$  theo thời gian. Hình ảnh của  $e_t$  có thể cung cấp những gợi ý về sự tự tương quan.

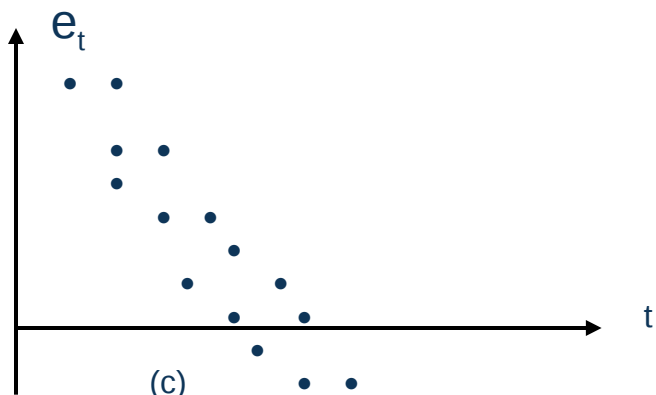
## a. Đồ thị



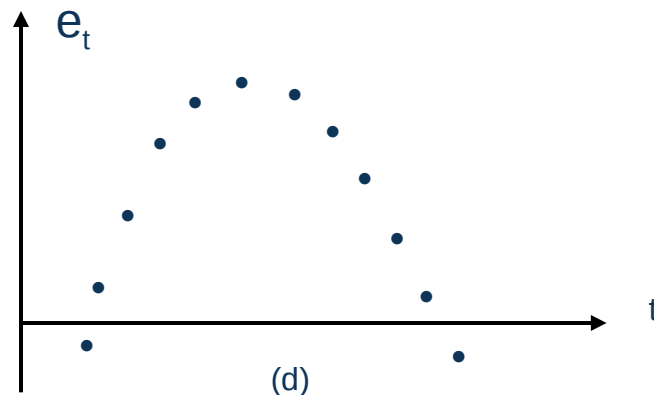
(a)



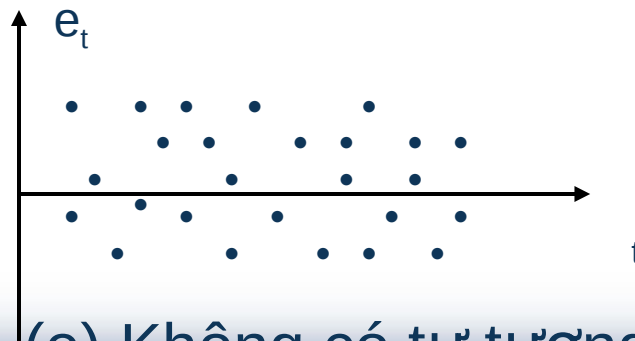
(b)



(c)



(d)



(e) Không có tự tương quan

## b. Dùng kiểm định d của Durbin – Watson

Thống kê d của Durbin – Watson

$$d = \frac{\sum (e_i - e_{i-1})^2}{\sum e_i^2}$$

Khi n đủ lớn thì  $d \approx 2(1-\rho)$  với  $\rho = \frac{\sum e_i e_{i-1}}{\sum e_i^2}$

do  $-1 \leq \rho \leq 1$ , nên  $0 \leq d \leq 4$ :

$\rho = -1 \Rightarrow d = 4$ : tự tương quan hoàn hảo âm

$\rho = 0 \Rightarrow d = 2$ : không có tự tương quan

$\rho = 1 \Rightarrow d = 0$ : tự tương quan hoàn hảo dương

## b. Dùng kiểm định d của Durbin – Watson

Bảng thống kê Durbin cho giá trị tới hạn  $d_U$  và  $d_L$  dựa vào 3 tham số:

$\alpha$ : mức ý nghĩa

$k'$ : số biến độc lập của mô hình

$n$ : số quan sát

Có tự tương quan dương	Không quyết định được	Không có tự tương quan bậc nhất	Không quyết định được	Có tự tương quan âm		
0	$d_L$	$d_U$	2	$4-d_U$	$4-d_L$	4

## b. Dùng kiểm định $d$ của Durbin – Watson

### Các bước thực hiện kiểm định $d$ của Durbin – Watson:

1. Chạy mô hình OLS và thu thập phần sai số  $e_t$ .
2. Tính  $d$  theo công thức trên.
3. Với cỡ mẫu  $n$  và số biến giải thích  $k$ , tìm giá trị tra bảng  $d_L$  và  $d_U$ .
4. Dựa vào các quy tắc kiểm định trên để ra kết luận.

## b. Dùng kiểm định d của Durbin – Watson

Nếu  $d$  thuộc vùng chưa quyết định, sử dụng quy tắc kiểm định cải biên:

$$1. H_0: \rho = 0; H_1: \rho > 0$$

Nếu  $d < d_U$  : bác bỏ  $H_0$  và chấp nhận  $H_1$  (với mức ý nghĩa  $\alpha$ ), nghĩa là có tự tương quan dương.

Có tự tương quan dương

Không có tự tương quan dương

$d_U$



## b. Dùng kiểm định d của Durbin – Watson

2.  $H_0: \rho = 0; H_1: \rho < 0$

Nếu  $d > 4 - d_U$  : bác bỏ  $H_0$  và chấp nhận  $H_1$   
(với mức ý nghĩa  $\alpha$ ), nghĩa là có tự tương quan âm.

Không có tự tương quan âm

Có tự tương quan âm

$4 - d_U$

## b. Dùng kiểm định d của Durbin – Watson

3.  $H_0: \rho = 0$ ;  $H_1: \rho \neq 0$

Nếu  $d < d_U$  hoặc  $d > 4 - d_U$  : bác bỏ  $H_0$  và chấp nhận  $H_1$  (với mức ý nghĩa  $2\alpha$ ), nghĩa là có tự tương quan (âm hoặc dương).

Có tự tương quan dương	Không có tự tương quan	Có tự tương quan âm
$d_U$	$4 - d_U$	

## **b. Dùng kiểm định d của Durbin – Watson**

➤ Lưu ý khi áp dụng kiểm định d:

1. Mô hình hồi quy phải có hệ số chặn.
2. Các sai số ngẫu nhiên có tương quan bậc nhất:

$$u_t = \rho u_{t-1} + e_t$$

1. Mô hình hồi quy không có chứa biến trễ  $Y_{t-1}$ .
2. Không có quan sát bị thiếu (missing).

## c. Dùng kiểm định Breusch – Godfrey (BG)

Xét mô hình:

$$Y_t = \beta_1 + \beta_2 X_t + u_t \quad (8.1)$$

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

Kiểm định giả thiết

$H_0: \rho_1 = \rho_2 = \dots = \rho_p = 0$ , có nghĩa là không tồn tại tự tương quan ở bất kỳ bậc nào trong số từ bậc 1 đến bậc  $p$ .

### c. Dùng kiểm định Breusch – Godfrey (BG)

**Bước 1:** Ước lượng (8.1) bằng OLS, tìm phần dư  $e_t$

**Bước 2:** Dùng OLS để ước lượng mô hình

$$e_t = \beta_1 + \beta_2 X_t + \rho_1 e_{t-1} + \rho_2 e_{t-2} + \dots + \rho_p e_{t-p} + \varepsilon_t$$

từ đây thu được  $R^2$ .

**Bước 3:** với  $n$  đủ lớn,  $(n-p)R^2$  có phân phối xấp xỉ  $\chi^2(p)$  với  $p$  là bậc tương quan.

- Nếu  $(n-p)R^2 > \chi^2_\alpha(p)$ : Bác bỏ  $H_0$ , nghĩa là có tự tương quan ít nhất ở một bậc nào đó.

- Nếu  $(n-p)R^2 \leq \chi^2_\alpha(p)$ : Chấp nhận  $H_0$ , nghĩa là không có tự tương quan.

## c. Dùng kiểm định Breusch – Godfrey (BG)

Kiểm định BG có đặc điểm:

- Áp dụng cho mẫu có kích thước lớn
- Áp dụng cho mô hình có biến độc lập có dạng  $Y_{t-1}$ ,  $Y_{t-2}$  ..
- Kiểm định được bậc tương quan bất kỳ

## 8.4 Khắc phục

Các bước tiến hành

- 1) Ước lượng giá trị  $\rho$
- 2) Dùng giá trị  $\rho$  vừa được ước lượng để chuyển đổi mô hình hồi quy

## 8.4 Khắc phục

### 1. Trường hợp đã biết cấu trúc của tự tương quan: Phương pháp GLS:

- ❖  $u_t$  tự hồi quy bậc  $p$ , AR( $p$ )

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

với  $\rho$ : hệ số tự tương quan;  $|\rho| < 1$

- ❖ Giả sử  $u_t$  tự hồi quy bậc nhất AR(1)

$$u_t = \rho u_{t-1} + e_t \quad (*)$$

$e_t$ : sai số ngẫu nhiên (nhiều trắng), thỏa mãn những giả định của OLS:

$$E(e_t) = 0; \quad \text{Var}(e_t) = \sigma_\varepsilon^2; \quad \text{Cov}(e_t, e_{t+s}) = 0$$



## 8.4 Khắc phục

Xét mô hình hai biến:

$$y_t = \alpha_1 + \beta_1 x_t + u_t \quad (8.2)$$

Nếu (8.2) đúng với  $t$  thì cũng đúng với  $t - 1$

$$y_{t-1} = \alpha_1 + \beta_1 x_{t-1} + u_{t-1} \quad (8.3)$$

Nhân hai vế của (8.3) với  $\rho$

$$\rho y_{t-1} = \rho \alpha_1 + \rho \beta_1 x_{t-1} + \rho u_{t-1} \quad (8.4)$$

Trừ (8.2) cho (8.4)

$$\begin{aligned} y_t - \rho y_{t-1} &= \alpha_1(1 - \rho) + \beta_1 (x_t - \rho x_{t-1}) + (u_t - \rho u_{t-1}) \\ &= \alpha_1(1 - \rho) + \beta_1 (x_t - \rho x_{t-1}) + e_t \end{aligned} \quad (8.5)$$

## 8.4 Khắc phục

(8.5) gọi là phương trình sai phân tổng quát

Đặt:  $\alpha_1^* = \alpha_1 (1 - \rho)$

$$\beta_1^* = \beta_1$$

$$y_t^* = y_t - \rho y_{t-1}$$

$$x_t^* = x_t - \rho x_{t-1}$$

Khi đó (8.5) thành

$$y_t^* = \alpha_1^* + \beta_1^* x_t^* + e_t \quad (8.5^*)$$

## 8.4 Khắc phục

Vì  $e_t$  thoả mãn các giả định của phương pháp OLS nên các ước lượng tìm được là BLUE

- ❖ Phương trình hồi qui 8.5\* được gọi là phương trình sai phân tổng quát (Generalized Least Square – GLS).
- ❖ Để tránh mất mát một quan sát, quan sát đầu của  $y$  và  $x$  được biến đổi như sau:

$$y_1^* = y_1 \sqrt{1 - \rho}$$

$$x_1^* = x_1 \sqrt{1 - \rho}$$

## 2. Trường hợp $\rho$ chưa biết

### 2.1 Phương pháp sai phân cấp 1

❖ Nếu  $\rho = 1$ , thay vào phương trình sai phân tổng quát (8.5)

$$\begin{aligned}y_t - y_{t-1} &= \beta_1(x_t - x_{t-1}) + (u_t - u_{t-1}) \\ &= \beta_1(x_t - x_{t-1}) + e_t\end{aligned}$$

Hay: 
$$\Delta y_t = \beta_1 \Delta x_t + e_t \quad (8.6)$$

(8.6) phương trình sai phân cấp 1

$\Delta$  toán tử sai phân cấp 1

Sử dụng mô hình hồi qui qua gốc tọa độ để ước lượng hồi qui (8.6)

## 2.1 Phương pháp sai phân cấp 1

Giả sử mô hình ban đầu

$$y_t = \alpha_1 + \beta_1 x_t + \beta_2 t + u_t \quad (8.7)$$

Trong đó

$t$  biến xu thế

$u_t$  theo mô hình tự hồi qui bậc nhất

Thực hiện phép biến đổi sai phân cấp 1 đối với (8.7)

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 + e_t$$

trong đó:  $\Delta y_t = y_t - y_{t-1}$

$$\Delta x_t = x_t - x_{t-1}$$

## 2.1 Phương pháp sai phân cấp 1

❖ Nếu  $\rho = -1$ , thay vào phương trình sai phân tổng quát (8.5)

$$y_t + y_{t-1} = 2\alpha_1 + \beta_1(x_t + x_{t-1}) + e_t$$

Hay:

$$\frac{y_t + y_{t-1}}{2} = \alpha_1 + \beta_1 \frac{x_t + x_{t-1}}{2} + \frac{e_t}{2} \quad (*)$$

Mô hình \* gọi là mô hình hồi qui trung bình trượt.

## 2.2 Ước lượng $\rho$ dựa trên thống kê d-Durbin-Watson

$$d \approx 2(1 - \hat{\rho}) \text{ hay } \hat{\rho} \approx 1 - \frac{d}{2}$$

Đối với các mẫu nhỏ có thể sử dụng thống kê d cải biên của Theil – Nagar.

$$\hat{\rho} = \frac{n^2(1 - d/2) + k^2}{n^2 - k^2}$$

Dùng giá trị  $\rho$  vừa được ước lượng để chuyển đổi số liệu như mô hình 8.5

## 2.3 Thủ tục lặp Cochran – Orcutt để ước lượng $\rho$

Giả sử có mô hình hai biến

$$y_t = \alpha_1 + \beta_1 x_t + u_t \quad (8.8)$$

Mô hình  $u_t$  tự tương quan bậc nhất AR(1)

$$u_t = \rho u_{t-1} + e_t \quad (8.9)$$

Các bước ước lượng  $\rho$

**Bước 1:** Ước lượng mô hình (8.8) bằng phương pháp OLS và thu được các phần dư  $e_t$ .



## 2.3 Thủ tục lặp Cochran – Orcutt để ước lượng $\rho$

**Bước 2:** Sử dụng các phần dư để ước lượng hồi qui:

$$e_t = \hat{\rho}e_{t-1} + v_t \quad (8.10)$$

Do  $e_t$  là ước lượng vững của  $u_t$  thực nên ước lượng  $\rho$  có thể thay cho  $\rho$  thực.

**Bước 3:** Sử dụng  $\hat{\rho}$  thu được từ (8.10) để ước lượng phương trình sai phân tổng quát (8.5)

$$Y_t - \hat{\rho}Y_{t-1} = \alpha_1(1 - \hat{\rho}) + \beta_1(X_t - \hat{\rho}X_{t-1}) + (u_t - \hat{\rho}u_{t-1})$$

$$\text{Hay } y_t^* = \alpha_1^* + \beta_1^* x_t^* + v_t \quad (8.11)$$

## 2.3 Thủ tục lặp Cochran – Orcutt để ước lượng $\rho$

**Bước 4:** Vì chưa biết  $\hat{\rho}$  thu được từ (8.10) có phải là ước lượng tốt nhất của  $\rho$  hay không nên thế giá trị ước lượng của  $\alpha_1^*$  và  $\beta_1^*$  từ (8.11) vào hồi qui gốc (8.8) và được các phần dư mới  $e_t^*$ :

$$e_t^* = y_t - (\alpha_1^* + \beta_1^* x_t) \quad (8.12)$$

Ước lượng phương trình hồi qui tương tự với (8.10)

$$e_t^* = \hat{\rho} e_{t-1}^* + w_t \quad (8.13)$$

(8.13) là ước lượng vòng 2 của  $\rho$ .

Thủ tục này tiếp tục cho đến khi các ước lượng kế tiếp nhau của  $\rho$  khác nhau một lượng rất nhỏ, chẳng hạn nhỏ hơn 0,05 hoặc 0,005.

## 2.4 Phương pháp Durbin – Watson 2 bước để ước lượng $\rho$

Viết lại phương trình sai phân tổng quát

$$y_t = \alpha_1(1 - \rho) + \beta_1 x_t - \rho\beta_1 x_{t-1} + \rho y_{t-1} + e_t$$

(8.14)

Thủ tục Durbin – Watson 2 bước để ước lượng  $\rho$ :

### Bước 1:

1. Hồi qui (8.14)  $y_t$  theo  $x_t$ ,  $x_{t-1}$  và  $y_{t-1}$

2. Xem giá trị ước lượng hệ số hồi qui của  $y_{t-1}$  ( $= \hat{\rho}$ ) là ước lượng của  $\rho$

## 2.4 Phương pháp Durbin – Watson 2 bước để ước lượng $\rho$

**Bước 2:** Sau khi thu được  $\hat{\rho}$ , thay

$$y_t^* = y_t - \hat{\rho} \cdot y_{t-1}; x_t^* = x_t - \hat{\rho} \cdot x_{t-1}$$

và ước lượng hồi qui (8.5\*) với các biến đã được biến đổi như trên.