ในวิชาเศรษฐมิติ
หาความสัมพันธ์ระหว่าง Y_t และ X_t ← ข้อมูลอนุกรมเวลา

เครื่องมือ
⇓
Regression

$Y_t = \beta_0 + \beta_1 X_t + u_t$

คุณสมบัติ อย่างไร
Stationary ?

EC435
บทที่ 6 โคอินทิเกรชัน (Cointegration)

เฉลิมพงษ์ คงเจริญ ©2556

คณะเศรษฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์

November 18, 2019

spurious regression

กรณีที่ตัวแปรทั้งหมดเป็น $I(1)$ และไม่ได้เป็น cointegrated ตัวทดสอบที่ได้จากแบบจำลองจะไม่เหมาะสม เราจะพบความสัมพันธ์ระหว่างตัวแปรทั้งๆที่จริงๆแล้วตัวแปรทั้งหลายไม่มีความสัมพันธ์กัน (spurious regression)

Y_t และ X_t เป็น unit root (nonstationary).

Y_t กับ X_t ไม่ได้สัมพันธ์กัน แต่เวลาที่เอา Run

Regression $Y_t = \beta_0 + \boxed{\beta_1} X_t + u_t$, $\beta_1 = 0$

$\hat{\beta}_1 \neq 0$ ตัวสถิติทดสอบ t-test และสรุปว่า $\boxed{\beta_1 \neq 0}$

เหมือนมีการที่พบความสัมพันธ์ที่ไม่มีจริง

spurious regression

กรณีที่ตัวแปรทั้งหมดเป็น $I(1)$ และไม่ได้เป็น cointegrated ตัวทดสอบที่ได้จากแบบจำลองจะไม่เหมาะสม เราจะพบความสัมพันธ์ระหว่างตัวแปรทั้งๆที่จริงๆแล้วตัวแปรทั้งหลายไม่มีความสัมพันธ์กัน (spurious regression)

หากเราพิจารณาอนุกรมเวลาสองกระบวนการที่เป็น $I(1)$ และเป็นอิสระต่อกัน เราสามารถแสดงความสัมพันธ์ของแต่ละตัวแปรได้ดังนี้

$$y_{1t} = y_{1,t-1} + \varepsilon_{1t} \qquad \varepsilon_{1t} \sim WN(0,1)$$

$$y_{2t} = y_{2,t-1} + \varepsilon_{2t} \qquad \varepsilon_{2t} \sim WN(0,1)$$

โดยที่ $Cov(\varepsilon_{1t}, \varepsilon_{2t}) = 0$ ซึ่งแสดงว่าทั้งสองอนุกรมเวลาเป็นอิสระต่อกัน

Spurious regression

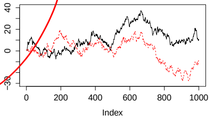
Figure: การจำลองกระบวนการ $I(1)$ สองอนุกรมเวลา

$y_{1t} = \beta_0 + \beta_1 y_{2t} + u_t$

$\beta_1 = 0$

โดยสร้าง → 0

จากการ สร้าง series



Spurious regression

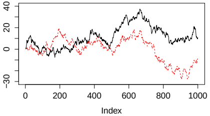


Figure: การจำลองกระบวนการ $I(1)$ สองอนุกรมเวลา

พิจารณาความสัมพันธ์เชิงถดถอย $y_{1t} = \alpha + \beta y_{2t} + u_t$

ใช้ Run OLS Regress

dep

explanatory

```
> summary(lm(y1~y2))
Call:
lm(formula = y1 ~ y2)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 11.43603    0.31329  36.503  < 2e-16 ***
y2           0.16237    0.02885   5.628 2.36e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '  1

Residual standard error: 9.905 on 998 degrees of freedom
Multiple R-squared: 0.03077,^^IAdjusted R-squared: 0.02979
F-statistic: 31.68 on 1 and 998 DF,  p-value: 2.363e-08
```

$\hat{\beta}_0 =$ $\hat{\beta}_1 =$

$H_0: \beta_1 = 0$ vs $H_1: \beta_1 \neq 0$

เมื่อ p-value < 0.05 เราปฏิเสธ H_0 และสรุปว่า $\beta_1 \neq 0$

ขัดแย้งกับโลกจริง $\beta_1 = 0$

[spurious regression]

ในทางปฏิบัติ เวลาที่เราพบว่า Y_t และ X_t เป็น unit root (non stationary)

เราไม่แน่ใจว่าความสัมพันธ์ที่พบ เป็น spurious หรือไม่

จาก spurious regression ⇒ คำแนะนำแรก

Spurious regression

⇒ ปรับตัวแปรให้เป็น stationary [Difference]

หากเราพิจารณาสมการ $\Delta y_{1t} = \beta \Delta y_{2t} + v_t$

Δ(y1) Δ(y2)

```
> summary(lm(diff(y1)~diff(y2)-1))

Call:
lm(formula = diff(y1) ~ diff(y2) - 1)

Coefficients:
           Estimate  Std. Error  t value  Pr(>|t|)
diff(y2)  0.003455    0.030827    0.112     0.911

Residual standard error: 0.9921 on 998 degrees of freedom
Multiple R-squared: 1.258e-05,^^IAdjusted R-squared: -0.0009894
F-statistic: 0.01256 on 1 and 998 DF,  p-value: 0.9108
```

ไม่มีจุดตัด

$H_0: \beta = 0$ vs $H_1: \beta \neq 0$

P-value > $\alpha(0.05)$ เราไม่สามารถปฏิเสธ H_0 ($\beta = 0$)

Spurious regression

หากเราพิจารณาสมการ $\Delta y_{1t} = \beta \Delta y_{2t} + v_t$

```
> summary(lm(diff(y1)~diff(y2)-1))

Call:
lm(formula = diff(y1) ~ diff(y2) - 1)

Coefficients:
          Estimate Std. Error t value Pr(>|t|)
diff(y2) 0.003455   0.030827   0.112    0.911

Residual standard error: 0.9921 on 998 degrees of freedom
Multiple R-squared: 1.258e-05,^^IAdjusted R-squared: -0.0009894
F-statistic: 0.01256 on 1 and 998 DF,  p-value: 0.9108
```

จากตัวอย่างข้างต้นเราจำเป็นต้องระมัดระวังในการพิจารณาความสัมพันธ์ระหว่างอนุกรมเวลาในกรณีที่ตัวแปรที่เราสนใจมีลักษณะเป็น $I(1)$ ซึ่งในกรณีทั่วไปเราจำเป็นต้องปรับให้ตัวแปรดังกล่าวเป็น $I(0)$

GDP - Y_t — $I(1)$

FDI - X_t — $I(1)$

$Y_t = \beta_0 + \beta_1 X_t + u_t$

$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X_t + u_t$ ไม่มีปัญหา Spurious Regression

Spurious regression

หากเราพิจารณาสมการ $\Delta y_{1t} = \beta \Delta y_{2t} + v_t$

```
> summary(lm(diff(y1)~diff(y2)-1))

Call:
lm(formula = diff(y1) ~ diff(y2) - 1)

Coefficients:
         Estimate Std. Error t value Pr(>|t|)
diff(y2) 0.003455   0.030827   0.112    0.911

Residual standard error: 0.9921 on 998 degrees of freedom
Multiple R-squared: 1.258e-05,^^IAdjusted R-squared: -0.0009894
F-statistic: 0.01256 on 1 and 998 DF,  p-value: 0.9108
```

จากตัวอย่างข้างต้นเราจำเป็นต้องระมัดระวังในการพิจารณาความสัมพันธ์ระหว่างอนุกรมเวลาในกรณีที่ตัวแปรที่เราสนใจมีลักษณะเป็น $I(1)$ ซึ่งในกรณีทั่วไปเราจำเป็นต้องปรับให้ตัวแปรดังกล่าวเป็น $I(0)$

ในบางกรณีที่เราอาจจะสามารถศึกษาความสัมพันธ์ระหว่างตัวแปรที่เป็น $I(1)$ ก็ได้ซึ่งในกรณีดังกล่าวคือกรณีที่ตัวแปรเป็น cointegration

บทนำ

$y_t = \beta_0 + \beta_1 x_t + u_t$

หากเราพิจารณาสมการถดถอยของตัวแปรทั้งสอง อาจจะมีเหตุการณ์เกิดขึ้นได้ 4 เหตุการณ์

ไม่มีปัญหา spurious

กรณีที่ 1: ทั้ง y และ x เป็น stationary เราสามารถใช้ OLS ปกติในการวิเคราะห์

บทนำ

หากเราพิจารณาสมการถดถอยของตัวแปรทั้งสอง อาจจะมีเหตุการณ์เกิดขึ้นได้ 4 เหตุการณ์

กรณีที่ 1: *ทั้ง y และ x เป็น stationary* เราสามารถใช้ OLS ปกติในการวิเคราะห์

กรณีที่ 2: *y และ x เป็น non-stationary แต่ integrated* ที่อันดับต่างกัน เราไม่สามารถใช้ OLS ศึกษาความสัมพันธ์ของทั้งสอง

เช่น $x_t \sim I(1)$

$y_t \sim I(2)$

- ทดสอบ unit root
- ทำ difference กี่ครั้ง ถึงจะเป็น stationary

[order of integration = ? , I(?)]

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta x_t + v_t$$ ← ไม่เกิดปัญหา spurious.

stationary

บทนำ

หากเราพิจารณาสมการถดถอยของตัวแปรทั้งสอง อาจจะมีเหตุการณ์เกิดขึ้นได้ 4 เหตุการณ์

กรณีที่ 1: *ทั้ง y และ x เป็น stationary* เราสามารถใช้ OLS ปกติในการวิเคราะห์

กรณีที่ 2: y และ x เป็น non-stationary แต่ integrated ที่อันดับต่างกัน เราไม่สามารถใช้ OLS ศึกษาความสัมพันธ์ของทั้งสอง

เป็น $X_t \sim I(1)$

$Y_t \sim I(1)$

กรณีที่ 3: y และ x เป็น non-stationary และ integrated ที่อันดับเดียวกัน แต่ ϵ เป็น *random walk* เราไม่สามารถใช้ OLS วิเคราะห์ แต่เราอาจจะพิจารณาสมการที่เรา difference ทั้งสองข้าง

$$\underset{I(1)}{Y_t} = \beta_0 + \beta_1 \underset{I(1)}{X_t} + \boxed{\varepsilon_t}$$

(stationary?)

เป็น Random Walk ?

Cointegration

กรณี 3

เป็น Random Walk

เราต้องแปลง Y_t และ X_t เป็น stationary

$$\Delta Y_t = \alpha_0 + \alpha_1 \Delta X_t + u_t$$

กรณี 4

เป็น stationary

เราสามารถ Run

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

ไม่ต้องแปลงโดยตัวแปรทั้งสอง

บทนำ

$$y_t = \beta_0 + \beta_1 X_t + \boxed{\varepsilon_t}$$

หากเราพิจารณาสมการถดถอยของตัวแปรทั้งสอง อาจจะมีเหตุการณ์เกิดขึ้นได้ 4 เหตุการณ์

กรณีที่ 1: *ทั้ง y และ x เป็น stationary* เราสามารถใช้ OLS ปกติในการวิเคราะห์

กรณีที่ 2: *y และ x เป็น non-stationary แต่ integrated* ที่อันดับต่างกัน เราไม่สามารถใช้ OLS ศึกษาความสัมพันธ์ของทั้งสอง

แปลงให้เป็น stationary ? ⇐ $y_t \sim I(2)$, $X_t \sim I(1)$ → spurious reg.

กรณีที่ 3: *y และ x เป็น non-stationary และ integrated* ที่อันดับเดียวกัน แต่ ϵ เป็น *random walk* เราไม่สามารถใช้ OLS วิเคราะห์ แต่เราอาจจะพิจารณาสมการที่เรา difference ทั้งสองข้าง

แปลง ⇒ y_t กับ X_t → Δy_t กับ ΔX_t

$y_t \sim I(1)$, $X_t \sim I(1)$

กรณีที่ 4: *y และ x เป็น non-stationary* โดยมีอันดับ *integrated* ที่เหมือนกัน และ *error term* เป็น *stationary*. เราเรียกว่า *y* และ *x* เป็น **cointegrated.**

$y_t \sim I(1)$, $X_t \sim I(1)$

เราสามารถประมาณ

$$\underbrace{y_t}_{I(1)} = \beta_0 + \beta_1 \underbrace{X_t}_{I(1)} + \varepsilon_t$$

ไม่มีปัญหา spurious regress

เราสามารถอธิบายความสัมพันธ์ระหว่าง y_t และ X_t ได้.

Cointegration

$$Y_t = \begin{pmatrix} y_{1t} \sim I(1) \\ \vdots \sim I(1) \\ y_{nt} \sim I(1) \end{pmatrix}$$

กำหนดให้ $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, ..., y_{nt})'$ แทนเวกเตอร์ $(n \times 1)$ ของอนุกรมเวลาที่เป็น $I(1)$ แล้ว $\boldsymbol{Y}_t$ จะเป็น cointegrated ถ้ามีเวกเตอร์ $\boldsymbol{\beta} = (\beta_1, ..., \beta_n)'$ ที่ทำให้

ค่าคงที่

$$\boldsymbol{\beta}'\boldsymbol{Y}_t = \beta_1 y_{1t} + ... + \beta_n y_{nt} \sim I(0) \qquad (6.1)$$

stationary แล้ว $Y_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{nt} \end{pmatrix}$ cointegrated กัน

ซึ่งหมายความว่าผลรวมเชิงเส้น (linear combination) ของตัวแปรที่เป็น $I(1)$ มีลักษณะเป็น $I(0)$

ผลรวมเชิงเส้นมักจะได้มาจากทฤษฎีทางเศรษฐศาสตร์ และเรียกสมการดังกล่าวว่าเป็นความสัมพันธ์ในดุลยภาพระยะยาว **(long-run equilibrium)**

เช่น $C = f(Y) \rightarrow C_t = \beta_0 + \beta_1 Y_t + u_t$ ⇐ สามารถประมาณค่าสัมพันธ์ได้ ถ้า C_t และ Y_t เป็น cointegration

$\sim I(1)$ $\quad GDP_t \sim I(1)$

$$\Delta C_t = \alpha_0 + \alpha_1 \Delta Y_t + u_t$$

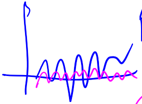
Cointegration

ไม่ได้มีรูปแบบเดียว

เวกเตอร์ β ไม่ได้มีลักษณะเป็นรูปแบบอย่างเดียว (unique) เนื่องจากเวกเตอร์ที่เป็นจำนวนเท่าของ $c\beta'$

$$\underline{\beta_1 Y_{1t} + \beta_2 Y_{2t}} \sim I(0)$$

$$\underbrace{\beta_1 Y_{1t} + \beta_2 Y_{2t}}_{\text{linear combination}} \sim I(0) \Rightarrow$$ Y_{1t} กับ Y_{2t} เป็น coingrated กัน.



$$\frac{\beta_1}{2} Y_{1t} + \frac{\beta_2}{2} Y_{2t} \sim I(0)$$

ทำให้ vector β มีได้หลายค่า

เพราะฉะนั้น กำหนดให้ $\beta = [1 \;\; -\beta_2 \;\; -\beta_3 \cdots -\beta_n]$

Cointegration

เวกเตอร์ β ไม่ได้มีลักษณะเป็นรูปแบบอย่างเดียว (unique) เนื่องจากเวกเตอร์ที่เป็นจำนวนเท่าของ $c\beta'$

β

เรามักจะสนใจในรูปแบบที่ normalized ซึ่งเขียนเป็นความสัมพันธ์ cointegration ได้เป็น

1

$$\beta' Y_t = y_{1t} - \beta_2 y_{2t} - \dots - \beta_n y_{nt} \sim I(0)$$

หรือเป็นสมการ

Reg

linear comb.

$$y_{1t} = \beta_2 y_{2t} + \dots + \beta_n y_{nt} + u_t \qquad (6.2)$$

$$y_{1t} - \beta_2 y_{2t} - \dots - \beta_n y_{nt} = \boxed{u_t}$$

โดยที่ $u_t \sim I(0)$ และเราเรียกพจน์ u_t ว่า disequilibrium error หรือ cointegrating residuals

$y_{1t}, \dots, y_{nt}$ เป็น cointegration.

Long-run

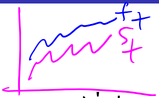
เมื่อ $u_t = 0$ [ตัวแปร Y_t อยู่ใน Eqm.]

(6.2) อธิบายความสัมพันธ์ระยะยาว

$$\boxed{y_{1t} = \beta_2 y_{2t} + \dots + \beta_n y_{nt}}$$

แต่ในระยะสั้น

$f_t = S_t + cost + u_t$



Cointegration

เวกเตอร์ $\boldsymbol{\beta}$ ไม่ได้มีลักษณะเป็นรูปแบบอย่างเดียว (unique) เนื่องจากเวกเตอร์ที่เป็นจำนวนเท่าของ $c\boldsymbol{\beta}'$

เรามักจะสนใจในรูปแบบที่ normalized ซึ่งเขียนเป็นความสัมพันธ์ cointegration ได้เป็น

$$\boldsymbol{\beta}'\boldsymbol{Y}_t = y_{1t} - \beta_2 y_{2t} - \ldots - \beta_n y_{nt} \sim I(0)$$

ตย. Spot price (S_t)
future price (f_t)
Cost of carrying
$f_t = 1S_t + cost$
$\sim I(1) \quad \sim I(1)$

หรือเป็นสมการ

$$y_{1t} = \beta_2 y_{2t} + \ldots + \beta_n y_{nt} + u_t \qquad (6.2)$$

โดยที่ $u_t \sim I(0)$ และเราเรียกพจน์ u_t ว่า disequilibrium error หรือ cointegrating residuals

ในดุลยภาพระยะยาวค่า disequilibrium error จะเท่ากับศูนย์ และความสัมพันธ์ระยะยาวจะเท่ากับ $y_{1t} = \beta_2 y_{2t} + \ldots + \beta_n y_{nt}$

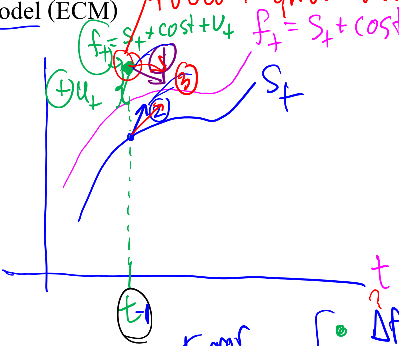
ตย. $C_t = \underbrace{\beta_0}_{\text{constant}} + \beta_1 Y_t + u_t$ → แล้วพบว่า $u_t \sim I(0)$
C_t กับ Y_t เป็น cointegration

ใน LR, $u_t = 0$, $\boxed{C_t = \beta_0 + \beta_1 Y_t}$

Error Correction Models

หากเกิดภาวะที่ในระยะสั้นตัวแปรขยับออกจากดุลยภาพในระยะยาว จะมีกลไกในการปรับตัวเข้าสู่ดุลยภาพระยะยาวอย่างไร ซึ่งสามารถอธิบายผ่านแบบจำลอง Error Correction Model (ECM)

ขยับออกจากดุลยภาพระยะยาว



① f_t ปรับตัวกลับเข้าสู่เส้นสมดุล

② S_t ปรับตัวสูงขึ้นจนเข้าสู่ดุลยภาพ

③ $f_t \downarrow$ $S_t \uparrow$ จนกระทั่งเข้าสู่ดุลยภาพ

สมการระยะยาว $f_t = \beta_0 + \overset{1}{\beta_1} S_t + u_t$

ถ้า S_t, f_t อยู่ในดุลยภาพ $u_t = 0$

ถ้า u_{t-1} มีค่า (+) $f_{t-1} > \beta_0 + \beta_1 S_{t-1}$

กลไกการปรับตัวระยะสั้น Error Correction Model (ECM)

- $\Delta f_t = f_t - f_{t-1} = \alpha_1 u_{t-1}$ (+) → f_t จะปรับตัวอย่างไรเมื่อเกิด shock
- $\Delta S_t = S_t - S_{t-1} = \alpha_2 u_{t-1}$ → S_t

Error Correction Models

หากเกิดภาวะที่ในระยะสั้นตัวแปรขยับออกจากดุลยภาพในระยะยาว จะมีกลไกในการปรับตัวเข้าสู่ดุลยภาพระยะยาวอย่างไร ซึ่งสามารถอธิบายผ่านแบบจำลอง Error Correction Model (ECM)

$y_{1t} = \beta_2 y_{2t} + u_t \Rightarrow u_t \sim I(0)$

หากเราพิจารณา cointegration ที่มีตัวแปรสองตัวแปรที่มีลักษณะ $I(1)$ คือ y_{1t} และ y_{2t} โดยมีความสัมพันธ์ cointegration เขียนได้เป็น $y_{1t} - \beta_2 y_{2t} \sim I(0)$ Engle and Granger (1987) แสดงให้เห็นว่ามีแบบจำลอง error correction model (ECM) ที่สามารถแสดงได้ในรูป

1) ทิศทาง ขนาดการปรับเร็ว

2) ขนาด ใหญ่ → ปรับเร็ว

u_{t-1} disequm ตอน t-1 (+)

$$\Delta y_{1t} = c_1 + \alpha_1 \underbrace{(y_{1,t-1} - \beta_2 y_{2,t-1})}_{u_{t-1}} + \sum_{j=1}^{p} \psi_{11}^{j} \Delta y_{1,t-j} + \sum_{j=1}^{p} \psi_{12}^{j} \Delta y_{2,t-j} + \varepsilon_{1t}$$

↓ ค่านี้ speed of adjustment

$$\Delta y_{2t} = c_2 + \alpha_2 (y_{1,t-1} - \beta_2 y_{2,t-1}) + \sum_{j=1}^{p} \psi_{21}^{j} \Delta y_{1,t-j} + \sum_{j=1}^{p} \psi_{22}^{j} \Delta y_{2,t-j} + \varepsilon_{2t}$$

$\Delta y_{1t} = y_{1t} - y_{1t-1} = c_1 + \alpha_1 \underbrace{(y_{1,t-1} - \beta_2 y_{2,t-1})}_{u_{t-1}}$

[สมการระยะยาว $y_{1t-1} = \beta y_{2t-1} + u_{t-1}$]

ทิศทาง $-0.5 / 0.9$

$y_{1t} \downarrow$ $+0.5$ $y_{1t} \uparrow$

ขนาด → speed

บอกถึง disequm ใน t-1 (+) 1

Error Correction Models

หากเกิดภาวะที่ในระยะสั้นตัวแปรขยับออกจากดุลยภาพในระยะยาว จะมีกลไกในการปรับตัวเข้าสู่ดุลยภาพระยะยาวอย่างไร ซึ่งสามารถอธิบายผ่านแบบจำลอง Error Correction Model (ECM)

หากเราพิจารณา cointegration ที่มีตัวแปรสองตัวแปรที่มีลักษณะ $I(1)$ คือ y_{1t} และ y_{2t} โดยมีความสัมพันธ์ cointegration เขียนได้เป็น $y_{1t} - \beta_2 y_{2t} \sim I(0)$ Engle and Granger (1987) แสดงให้เห็นว่ามีแบบจำลอง error correction model (ECM) ที่สามารถแสดงได้ในรูป

$$\Delta y_{1t} = c_1 + \alpha_1(y_{1,t-1} - \beta_2 y_{2,t-1}) + \sum_{j=1}^{p} \psi_{11}^{j} \Delta y_{1,t-j} + \sum_{j=1}^{p} \psi_{12}^{j} \Delta y_{2,t-j} + \varepsilon_{1t}$$

$$\Delta y_{2t} = c_2 + \alpha_2(y_{1,t-1} - \beta_2 y_{2,t-1}) + \sum_{j=1}^{p} \psi_{21}^{j} \Delta y_{1,t-j} + \sum_{j=1}^{p} \psi_{22}^{j} \Delta y_{2,t-j} + \varepsilon_{2t}$$

dynamic

สมการข้างต้นแสดงพฤติกรรมเชิงพลวัตรของ y_{1t} และ y_{2t} โดยแสดงการปรับตัวระยะสั้นเมื่อตัวแปรตัวใดตัวหนึ่งเคลื่อนตัวออกไปจากดุลยภาพ ซึ่งแบบจำลอง ECM ส่งผลให้เกิดประโยชน์ในการวิเคราะห์ข้อมูลการเงิน

ตัวอย่างที่ 6.2

ตัวอย่าง ECM สำหรับราคาหุ้นและเงินปันผล

กำหนดให้ล็อกของราคาหุ้น (s_t) และล็อกของเงินปันผล (d_t) เป็น $I(1)$ ถ้าล็อกของสัดส่วนเงินปันผลต่อราคา ($\log(D_t/S_t)$) เป็น $I(0)$ ดังนั้น $d_t - s_t \sim I(0)$ หรือเขียนเป็นสมการความสัมพันธ์ระยะยาวได้เป็น

$\log(D_t) - \log(S_t) \sim I(0)$

ดังนั้น -

$$d_t = s_t + \mu + u_t \qquad (6.4)$$

$u_t = d_t - s_t - \mu \sim I(0)$

ในระยะยาว ถ้า $u_t = 0 \Rightarrow d_t = s_t + \mu$

โดยที่ $u_t \sim I(0)$ และสามารถเขียน ECM ได้ดังนี้

เกิด $u_{t-1} \neq 0$

$$\Delta s_t = c_s + \alpha_s \overbrace{(d_{t-1} - s_{t-1} - \mu)}^{u_{t-1}} + \varepsilon_{st} \qquad (6.5)$$

$$\Delta d_t = c_d + \alpha_d \underbrace{(d_{t-1} - s_{t-1} - \mu)}_{u_{t-1}} + \varepsilon_{dt} \qquad (6.6)$$

ทิศทาง + ขนาด

ตัวอย่างที่ 6.2

ขนาด 0.1

u_{t-1} (+)

สมมุติให้ $\alpha_s = 0.5$ และ $\alpha_d = 0$ จะได้

$$\Delta s_t = c_s + 0.5(d_{t-1} - s_{t-1} - \mu) + \varepsilon_{st}$$
$$\Delta d_t = c_d + \varepsilon_{dt}$$

→ d_t ไม่ได้ช่วยปรับตัว.

ดังนั้นมีเพียงราคาหุ้นที่ตอบสนองต่อ disequilibrium error โดยที่ค่าเฉลี่ยของการเปลี่ยนแปลงของตัวแปรทั้งสองจะขึ้นอยู่กับข้อมูลในอดีต

$$E(\Delta s_t | F_{t-1}) = c_s + 0.5(d_{t-1} - s_{t-1} - \mu)$$
$$E(\Delta d_t | F_{t-1}) = c_d$$

multiple cointegrating relationship

$$Y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix}$$

เวกเตอร์ของอนุกรมเวลา $\boldsymbol{Y}_t = (y_{1t}, ..., y_{nt})'$ ซึ่งเป็นอนุกรมเวลาที่ cointegrated กัน อาจจะมีความสัมพันธ์ระยะยาวได้ r รูปแบบ ซึ่ง $o < r < n$ (จำนวนตัวแปร) เช่นกรณีที่ $\boldsymbol{Y}_t = (y_{1t}, y_{2t}, y_{3t})'$ อาจจะมีความสัมพันธ์ในรูป $\beta_{11}y_{1t} + \beta_{12}y_{2t} + \beta_{13}y_{3t} \sim I(0)$ และ $\beta_{21}y_{1t} + \beta_{22}y_{2t} + \beta_{23}y_{3t} \sim I(0)$ และเราเรียกเวกเตอร์ $(\beta_{11} \quad \beta_{12} \quad \beta_{13})'$ และ $(\beta_{21} \quad \beta_{22} \quad \beta_{23})'$ ว่า cointegrating vector

ต.ย. $Y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \end{pmatrix}$

กรณี 1 $r=1$ (มีความสัมพันธ์ 2 รูปแบบ) → 1 linear combination

$$y_{1t} + \beta_2 y_{2t} + \beta_3 y_{3t} \sim I(0)$$

กรณี 2 $r=2$ (มีความสัมพันธ์ 2 รูปแบบ)

เป็นอิสระ
$$\begin{cases} \beta_{11} y_{1t} + \beta_{12} y_{2t} + \beta_{13} y_{3t} \sim I(0) \\ \beta_{21} y_{1t} + \beta_{22} y_{2t} + \beta_{23} y_{3t} \sim I(0) \end{cases}$$

ต.ย. r_{1t} r_{2t} r_{3t}

$\boxed{r=2}$

- $r_{2t} = \beta_1 r_{1t} + 0 \cdot r_{3t}$
- $r_{3t} = \beta_2 r_{1t} + 0 r_{2t}$

การทดสอบ cointegration

หากเรากำหนดให้ $\boldsymbol{Y}_t$ เป็นเวกเตอร์ $n \times 1$ ของอนุกรมเวลา $I(1)$ ซึ่งมี $r\ (0 < r < n)$ cointegrating vectors ซึ่งแสดงเป็นสัญลักษณ์ได้ดังนี้

$$\begin{pmatrix} \boldsymbol{\beta}_1' \boldsymbol{Y}_t \\ \vdots \\ \boldsymbol{\beta}_r' \boldsymbol{Y}_t \end{pmatrix} = \begin{pmatrix} u_{1t} \\ \vdots \\ u_{rt} \end{pmatrix} \sim I(0) \qquad (6.7)$$

การทดสอบ cointegration อาจจะแบ่งได้เป็นสองกรณีคือ

1. การทดสอบว่ามีความสัมพันธ์ cointegrating อย่างน้อยหนึ่งรูปแบบหรือไม่ (Engle and Granger(1986))
   จำนวนความสัมพันธ์ ?
2. การทดสอบว่ามี cointegrating vector $0 \leq r < n$ หรือไม่ (Johansen (1988))

การทดสอบ cointegration

โดยใช้แนวคิดของ Engle กับ Granger

พื้นฐาน มาจาก นิยาม cointegration : $y_{1t} - \beta_2 y_{2t} - \cdots - \beta_n y_{nt} \sim I(0)$

$y_{1t} = \beta_2 y_{2t} + \cdots + \beta_n y_{nt} + u_t$, $u_t \sim I(0)$

การทดสอบในประเด็นแรกได้ถูกนำเสนอโดย Engle and Granger (1986) ซึ่งเสนอการทดสอบสองขั้นตอน (two-step cointegration test)ซึ่งประกอบด้วย

1. สร้าง cointegrating residuals โดยที่ $u_t = y_{1t} - \beta_2 y_{2t} - \ldots - \beta_n y_{nt}$
2. ทดสอบ unit root ของ u_t เพื่อแสดงว่า u_t เป็น $I(0)$ หรือไม่ โดยสมมุติฐานหลักของการทดสอบนี้คือไม่มี cointegration ซึ่งสมมุติฐานดังกล่าวสอดคล้องกับ u_t เป็น $I(1)$

Unit root test

การทดสอบ cointegration ตามวิธีการของ Engle and Granger สามารถแบ่งออกได้เป็นสองกรณีคือ

$H_0: u_t \sim I(1)$ (unit root) vs $H_1: u_t \sim I(0)$

[Y_t ไม่มี cointegration]

Y_t เป็น cointegration

ทราบ β – ทฤษฎี

กรณีที่หนึ่ง เราทราบความสัมพันธ์ของอนุกรมเวลาที่เราสนใจ

กรณีที่สอง เราจำเป็นต้องประมาณค่าสำหรับ cointegrating vector และสร้าง cointegrating vector จาก $\hat{u}_t = \hat{\boldsymbol{\beta}}' \boldsymbol{Y}_t$

$\boxed{\text{ไม่ทราบค่า } \beta}$

$C_t = \beta_0 + \beta_1 Y_t$

$\hat{u}_t = Y_{1t} - \hat{\beta}_2 Y_{2t} \cdots$

การทดสอบกรณีทราบ cointegration vector

รู้ β

① กรณี $u_t = Y_{1t} - \beta_2 Y_{2t} \cdots$

กำหนดให้ $\boldsymbol{Y}_t$ เป็นเวกเตอร์ $n \times 1$ ของ $I(1)$ และ cointegrated กันโดยที่มีความสัมพันธ์สามารถแสดงได้ด้วย cointegrating vector $\boldsymbol{\beta}$ ซึ่ง $u_t = \boldsymbol{\beta}' \boldsymbol{Y}_t$ เราต้องการ ② ทดสอบสมมุติฐานหลัก $H_0 : u_t = \boldsymbol{\beta}' \boldsymbol{Y}_t \sim I(1)$ หรือไม่มี cointegration กับสมมุติฐานรอง $H_1 : u_t = \boldsymbol{\beta}' \boldsymbol{Y}_t \sim I(0)$ หรือมี cointegration

เราสามารถทดสอบสมมุติฐานข้างต้นได้ด้วยการทดสอบยูนิทรูทเช่น ADF หรือ PP กับตัวแปร u_t ซึ่งเราจะสรุปว่า $\boldsymbol{Y}_t$ เป็น cointegrated ถ้าเราปฏิเสธสมมุติฐานหลัก

ตัวอย่างที่ 6.3

ในตัวอย่างนี้เราจะพิจารณาความสัมพันธ์ระหว่างราคาปัจจุบัน (spot price: S_t) และราคาฟิวเจอร์ (future price: F_t) ของอัตราแลกเปลี่ยนโดยที่ตัวแปรทั้งสองมีความสัมพันธ์ตามแบบจำลอง cost of carry model ซึ่งล็อกของราคาปัจจุบัน (s_t) จะเท่ากับล็อกของราคาฟิวเจอร์ (f_t) บวกกับค่าคงที่(ซึ่งสะท้อนต้นทุนในการถือสินทรัพย์)

Cost of carrying.

$$f_t = 1 s_t + c \qquad c = 0$$

สมการระยะยาว $f_t = s_t + u_t \Rightarrow \boxed{u_t = f_t - s_t}$

ดังนั้นในกรณีนี้เราสามารถเขียน $u_t = f_t - s_t$ หรือ $\beta = (1 \quad -1)'$

future กับ spot

```
> set50_m <- read.csv("C:/teaching/ec435/r/set50_m.csv")
> lfutures<-log(set50_m$futures)
> lspot<-log(set50_m$spot)
> u<-lfutures-lspot
> adfTest(u, lags=2, type=c("nc"))
Title:
 Augmented Dickey-Fuller Test
Test Results:
  PARAMETER:
    Lag Order: 2
  STATISTIC:
    Dickey-Fuller: -7.2144
  P VALUE:
    0.01
```

$\Rightarrow u_t = f_t - s_t$

① ขั้นตอน 1 สร้าง $u_t = f_t - s_t$

② ขั้นตอน 2 ทดสอบ u_t unit root

$H_0: u_t \sim I(1)$ vs $H_1: u_t \sim I(0)$

เมื่อพิจารณา p-value $< \alpha(0.05)$ ปฏิเสธ H_0 และยอมรับ H_1 ซึ่งทำ u_t เป็น stationary ($I(0)$)

$\Rightarrow$ แสดงว่า f_t กับ s_t cointegrated [มีความสัมพันธ์ระยะยาว]

กรณี 2 ไม่ทราบค่า β

การทดสอบกรณีที่เราไม่ทราบค่า cointegrating vector

ในกรณีนี้เราจำเป็นต้องประมาณค่า cointegrating vector ขึ้นมา

ขั้นตอนแรกของการทดสอบเราจำเป็นต้องประมาณค่า β โดยที่เราอาจจะเขียนในรูปบรรทัดฐาน (normalized) ซึ่งเขียนเป็นสมการได้

ขั้นตอน 1

$$\boxed{y_{1t} = c + \beta_2 y_{2t} + \dots + \beta_n y_{nt} + u_t} \qquad (6.8)$$

ขั้นตอน 2 ประมาณค่า ด้วย OLS → $\hat{\beta}_2, \dots, \hat{\beta}_n$ เอามาหาค่า ส่วน $\hat{u}_t$

การทดสอบ cointegration จะเป็นการทดสอบ unit root ของ $\hat{u}_t = y_{1t} - \hat{c} - \hat{\beta}_2 y_{2t} - \dots - \hat{\beta}_n y_{nt}$ โดยที่ $\hat{c}, \hat{\beta}_2, \dots, \hat{\beta}_n$ เป็นตัวประมาณค่า OLS ของสัมประสิทธิ์

$H_0: \hat{u}_t \sim I(1)$ vs $H_1: \hat{u}_t \sim I(0)$ $\hat{u}_t$ เป็นผลจากการประมาณค่า

Phillips and Ouliaris (1990) ได้แสดงให้เห็นว่าตัวทดสอบ ADF และ PP ของ cointegrating residual กับ $\hat{u}_t$ ในกรณีที่เราต้องประมาณค่าสัมประสิทธิ์ของ cointegrating vector ไม่ได้มีการแจกแจงแบบ ADF distribution แบบปกติ (เราไม่สามารถเปิดตาราง DF ปกติได้ → เปิดตารางใหม่ Phillips กับ Ouliaris.

P-O เสนอตารางใหม่ ขึ้นอยู่กับรูปแบบของสมการระยะยาว

1) ไม่มีค่าคงที่ [$y_{1t} = \beta_2 y_{2t} \cdots$]

2) มีค่าคงที่ [$y_{1t} = \boxed{\beta_1} + \beta_2 y_{2t} + \cdots$]

3) มีค่าคงที่ และเส้นแนวโน้ม [$y_{1t} = \alpha t + \beta_1 + \beta_2 y_{2t} + \cdots$]

ตัวอย่างที่ 6.4

ขั้นตอนที่ 1 $f_t = c + \beta_2 s_t + u_t$ ไม่ทราบ → ประมาณค่าด้วย OLS → $\hat{u}_t$.

ต่อเนื่องจากตัวอย่างที่ผ่านมา ในกรณีนี้สมมุติว่าเราไม่ทราบค่าของ cointegrating vector ดังนั้นสมมุติให้ $\boldsymbol{\beta} = (1 \quad -\beta_2)'$ ดังนั้นเราสามารถพิจารณาสมการถดถอย $s_t = c + \beta_2 f_t + u_t$ ซึ่งหลังจากที่เราประมาณค่าด้วย OLS เราสามารถที่จะหาค่า $\hat{u}_t = s_t - \hat{c} - \hat{\beta}_2 f_t$ ตามขั้นตอนในคำสั่งข้างล่าง

เก็บผลการประมาณค่า

```
> m1<-lm(lfutures~lspot)
> summary(m1)
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.044097   0.023113  -1.908   0.0587 .
lspot        1.006206   0.003535 284.610   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '  1

Residual standard error: 0.01349 on 124 degrees of freedom
Multiple R-squared:  0.9985,^^IAdjusted R-squared:  0.9985
F-statistic: 8.1e+04 on 1 and 124 DF,  p-value: < 2.2e-16

> uhat<-m1$residuals
```

$\hat{c}$, $\hat{\beta}_2$

→ > α(0.05) ไม่มีนัยสำคัญ constant

จากสมการ cointegration (LR) ที่ได้ออกมา

$\hat{u}_t$

ขั้นตอนที่ 2 ทดสอบ unit root กับ $\hat{u}_t$ $[H_0: \hat{u}_t \sim I(1)$ vs $H_1: \hat{u}_t \sim I(0)]$

↓

ADF

ตัวอย่างที่ 6.4

cointegration equation.

$\hat{u}_t$ เป็นผลจากสมการถดถอย

สมมติ unit root เราต้องเทียบ t-stat กับ c.v. จาก P-O table

```
> adfTest(uhat,lags=2, type=c("nc"))

Title:
 Augmented Dickey-Fuller Test

Test Results:
  PARAMETER:
    Lag Order: 2
  STATISTIC:
    Dickey-Fuller: -8.4232
  P VALUE:
    0.01
```

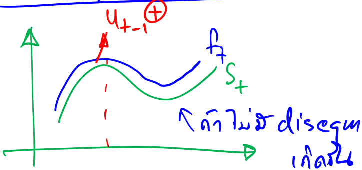
ดูแต่ละตัวกับ P-O table c.v. = -3.3654 (Constant, n-1 = 1)

ปฏิเสธ Ho

[เปรียบเทียบกับตาราง DF ปกติ]

t < c.v. เราปฏิเสธ Ho ที่ว่า $\hat{u}_t$ เป็น unit root -3.3654

แสดงว่า $\hat{u}_t$ เป็น stationary $\Rightarrow$ S_t กับ f_t เป็น cointegration



ในการอธิบาย การปรับตัว (ระยะสั้น) เพื่อเข้าสู่ดุลยภาพ

การประมาณค่า ECM ด้วย OLS

สมมุติว่าเราสนใจตัวแปร $(y_{1t} \quad y_{2t})'$ ซึ่งมีความสัมพันธ์ cointegration แสดงได้ด้วยสมการ $y_{1t} = \beta_2 y_{2t} + u_t$ (หรือ $y_{1t} - \beta_2 y_{2t} = u_t \sim I(0)$) และเรามีตัวประมาณค่า $\hat{\beta}_2$ ที่มีคุณลักษณะเป็น consistent และเราต้องประมาณค่า ECM ในรูปแบบ

LR

I(0)

$$\Delta y_{1t} = c_1 + \alpha_1 (\underbrace{y_{1,t-1} - \beta_2 y_{2,t-1}}_{u_{t-1}}) + \sum_{i=1}^{p} \psi_{11}^{j} \underbrace{\Delta y_{1,t-j}}_{I(0)} + \sum_{i=1}^{p} \psi_{12}^{j} \underbrace{\Delta y_{2,t-j}}_{I(0)} + \varepsilon_{1t} \qquad (6.9)$$

$$\Delta y_{2t} = c_2 + \alpha_2 (\underbrace{y_{1,t-1} - \beta_2 y_{2,t-1}}_{u_{t-1}}) + \sum_{i=1}^{p} \psi_{21}^{j} \Delta y_{1,t-j} + \sum_{i=1}^{p} \psi_{22}^{j} \Delta y_{2,t-j} + \varepsilon_{2t} \qquad (6.10)$$

เราสามารถแทนค่า $\hat{\beta}_2$ ลงในสมการเหมือนเป็นพารามิเตอร์ที่เราทราบค่า

เราสามารถพิจารณา disequilibrium error $(y_{1,t-1} - \beta_2 y_{2,t-1})$ เหมือนเป็นตัวแปรที่ทราบค่า เนื่องจากตัวแปรทุกตัวในสมการ (6.9) และ (6.10) เป็น $I(0)$ เราสามารถประมาณค่าสมการดังกล่าวด้วย OLS

แล้ว $\hat{c}_1, \hat{\alpha}_1, \hat{\psi}$, $\hat{c}_2, \hat{\alpha}_2, \hat{\psi}$ เป็น Consistent

ตัวอย่างที่ 6.5

ต่อเนื่องจาก ตัวอย่าง 6.4

ประมาณค่า $\hat{f}_t = 1.006 s_t - 0.044$

$\hat{u}_{t-1} = f_{t-1} - 1.006 s_{t-1} + 0.044 > 0$

$f_{t-1} > 1.006 s_t - 0.044$

เราต้องการประมาณค่าสมการ ECM สำหรับล็อกของราคาปัจจุบันและล็อกของราคาฟิวเจอร์ซึ่งเราจำเป็นต้องหาอันดับของค่าล่าของผลต่างที่เหมาะสม ซึ่งอาจจะใช้ AIC, BIC ในการเลือก แต่ในที่นี้เราสมมุติให้แบบจำลองของเป็นดังนี้

ECM

$$\underbrace{\Delta s_t}_{s_t - s_{t-1}} = c_1 + \alpha_1 \hat{u}_{t-1} + \underbrace{\psi_{11}\Delta s_{t-1} + \psi_{12}\Delta f_{t-1}}_{p=1} + \varepsilon_{ft} \qquad (6.11)$$

ECM

$$\underbrace{\Delta f_t}_{f_t - f_{t-1}} = c_2 + \alpha_2 \hat{u}_{t-1}) + \underbrace{\psi_{21}\Delta s_{t-1} + \psi_{22}\Delta f_{t-1}}_{p=1} + \varepsilon_{st} \qquad (6.12)$$

$f_t = 1.006 s_t - 0.044$

diseqm $t-1$ u_{t-1}

Δf_t $\quad \Delta f_{t-1}$ $\quad \Delta s_{t-1}$

```
> ecm1<-lm(dlfutures[2:125]~uhat[1:124]+dlfutures[1:124]+dlspot[1:124])
> summary(ecm1)
Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.004115   0.006281   0.655   0.5136
uhat[1:124]    -1.152182   0.598961  -1.924   0.0568 .
dlfutures[1:124] 0.071715  0.478109   0.150   0.8810
dlspot[1:124]   0.137282   0.498649   0.275   0.7836
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '  1

Residual standard error: 0.06973 on 120 degrees of freedom
Multiple R-squared:  0.06491,^^IAdjusted R-squared:  0.04154
F-statistic: 2.777 on 3 and 120 DF,  p-value: 0.04424
```

$\hat{c}_2$, $\hat{\alpha}_2$, $\hat{\psi}_{21}$, $\hat{\psi}_{22}$

$\hat{\alpha}_2 = -1.152$

ถ้า u_{t-1} (+)

$\underbrace{\Delta f_t}_{f_t - f_{t-1}} = -1.152 \cdot u_{t-1}$

$\boxed{\hat{\alpha}_1 = -0.2} \Rightarrow \Delta s_t \downarrow (s_t \downarrow)$

VAR models และ cointegration

ในกรณีที่เรามีตัวแปรมากกว่า 2 ตัวแปร
- เราอาจจะพบความสัมพันธ์ระยะยาว มากกว่า 1 ความสัมพันธ์.

Granger representation theorem เชื่อมโยง cointegration กับ error correction models

Johansen สร้างแบบจำลอง cointegration และ error correction models โดยใช้โครงร่าง vector autoregression

(ระยะยาว) (ระยะสั้น)

พิจารณาตัวแปรในระดับ level ในรูป $VAR(p)$ โดยมีตัวแปรอยู่ในรูปเวกเตอร์ $(n \times 1)$ เขียนเป็นสัญลักษณ์ $\boldsymbol{Y}_t = \begin{pmatrix} y_{1t} \\ \vdots \\ y_{nt} \end{pmatrix}$ เป็น unit root $(I(1))$

VAR(p)

$$\boldsymbol{Y}_t = \boldsymbol{\Pi}_1 \boldsymbol{Y}_{t-1} + \ldots + \boldsymbol{\Pi}_p \boldsymbol{Y}_{t-p} + \boldsymbol{\epsilon}_t, t = 1, \ldots, T$$

เดิมเราสมมุติให้ตัวแปรทุกตัวเป็น stationary.

ปกติ เราไม่สามารถประมาณค่าด้วย OLS ธรรมดา

ถ้า VAR(p) มีลักษณะเป็น unit root แสดงว่าตัวแปรบางตัวหรือทั้งหมดใน $\boldsymbol{Y}_t$ มีลักษณะเป็น $I(1)$ และตัวแปรเหล่านั้นอาจจะ cointegrate กันได้

$Y_t \to \Delta Y_t$

Vector Error Correction Model (VECM)

VAR(2)

$$Y_t = \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \varepsilon_t$$

$$\underbrace{Y_t - Y_{t-1}}_{\Delta Y_t} = \Pi_1 Y_{t-1} - Y_{t-1} + \Pi_2 Y_{t-2} + \varepsilon_t - \Pi_2 Y_{t-1} + \Pi_2 Y_{t-1}$$

$$= \underbrace{-(I - \Pi_1 - \Pi_2)}\, Y_{t-1} - \Pi_2 \underbrace{(Y_{t-1} - Y_{t-2})}_{\Delta Y_{t-1}} + \varepsilon_t$$

Cointegrated VAR

VECM(1)

$$\overset{I(0)}{\Delta Y_t} = \overset{I(0)?}{\boxed{\Pi Y_{t-1}}} + \overset{I(0)}{\Gamma_1 \Delta Y_{t-1}} + \epsilon_t$$

$\begin{pmatrix} \Delta Y_{1t} \\ \Delta Y_{nt} \end{pmatrix}$ α, β เดิม

เทียบกับสมการ ECM

ความสัมพันธ์ลักษณะ cointegrating สามารถพิจารณาได้อย่างชัดเจนขึ้นหากเราแปลง level VAR เป็น **vector error correction model (VECM)**

เปลี่ยนแปลงในอดีต.

long run + short run

VECM(P-1)

$$\Delta Y_t = \Pi Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \ldots + \Gamma_{p-1} \Delta Y_{t-p+1} + \epsilon_t$$

โดยที่เมตริกซ์ $\mathbf{\Pi}$ ถูกเรียกว่า *long run impact matrix* และ $\mathbf{\Gamma}_k$ ถูกเรียกว่า *short run impact matrices*

ในแบบจำลอง VECM เราพบว่า ΔY_t และ Lag ของ ΔY_t เป็น $I(0)$

จะเห็นได้ว่า เหลือเพียงพจน์ ΠY_{t-1} ที่มีโอกาสเป็น $I(1)$ และหากเราพิจารณา ΔY_t เราพบว่า ΔY_t จะเป็น $I(0)$ ถ้า ΠY_{t-1} เป็น $I(0)$

ดังนั้น ΠY_{t-1} จะเป็นเมตริกซ์ที่ระบุ cointegrating relations หากตัวแปร cointegrate กัน

ความสัมพันธ์ระยะยาว ของตัวแปร.

เพราะฉะนั้น Matrix $\Pi \rightarrow$ บอกถึงจำนวนความสัมพันธ์ระยะยาวของตัวแปร.

$$Y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{pmatrix} \qquad \Pi = \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix}$$

→ แถวที่เป็นอิสระจากแถวอื่น = ?

⇓ Rank ของ Matrix

Cointegrated VAR

ถ้า VAR(p) process มีลักษณะเป็น unit roots เราพบว่า $\mathbf{\Pi}$ จะเป็น singular matrix และถ้า $\mathbf{\Pi}$ เป็น singular แล้วมันจะมี reduced rank ($rank(\mathbf{\Pi}) = r < n$ ซึ่งค่าดังกล่าวเป็นไปได้สองกรณีคือ

ไม่มี cointegration $\Pi = 0$ สมการ VECM $\to$ $\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_p \Delta Y_{t-p+1} + \varepsilon_t$

$t-(p-1)$

1. $rank(\mathbf{\Pi}) = 0$ แสดงว่า $\mathbf{\Pi}$ และ $\boldsymbol{Y}_t$ เป็น $I(1)$ และตัวแปรมิได้ cointegrate กัน และแบบจำลอง VECM จะลดลงเหลือแค่ $\boxed{\text{VAR(p-1)}}$ ของตัวแปรในรูป first differences ΔY_t

มี cointegration

2. $0 < rank(\mathbf{\Pi}) = r < n$ กรณีนี้แสดงว่า $\boldsymbol{Y}_t$ เป็น $I(1)$ โดยที่มี r linear independent cointegrating vectors

จำนวนความสัมพันธ์ ระยะยาว.

จำนวนแถวที่เป็นอิสระจากแถวอื่น

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 5 & 7 & 8 \end{bmatrix} \overset{2r_1 - r_2}{\sim} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 5 & 7 & 8 \end{bmatrix} \Rightarrow \text{Rank} = 2$$

การทดสอบ Cointegration ตามแนวทางของ Johansen

Johansen's Methodology for Modeling Cointegration

ขั้นตอนในการสร้างแบบจำลอง Cointegration ของ Johansen คือ

AIC, BIC

1) ■ ระบุ Order และประมาณค่า VAR(p) สำหรับ Y_t

$Y_t = \Pi_1 Y_{t-1} + \dots + \Pi_p Y_{t-p} + \varepsilon_t$

ประมาณค่า VECM (p-1)

$\Delta Y_t = \boxed{\Pi} Y_{t-1} + \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_{p-1} \Delta Y_{t-p+1} + \varepsilon_t$

2) ■ สร้าง likelihood ratio tests สำหรับการทดสอบ rank of **Π** เพื่อระบุจำนวน cointegrating vectors

$r = ?$

■ หากจำเป็นเราอาจจะสมมุติ normalization และระบุเงื่อนไขของ cointegrating vectors

ประมาณ VECM โดยการระบุ rank ของ (2)

3) ■ จาก normalized cointegrating vectors ที่ได้ระบุไว้ เราสามารถประมาณค่า cointegrated VECM ด้วย MLE

อธิบาย 1) ความสัมพันธ์ระยะยาว Cointegration

2) การปรับตัวระยะสั้น

rank = r

rank = 1

$$\underset{n\times n}{\Pi} = \underset{n\times r}{\alpha}\ \underset{r\times n}{\beta}$$

ตย.

$$\underset{4\times 4}{\Pi} Y_{t-1} = \begin{bmatrix} \Pi_{11} & \Pi_{12} & \Pi_{13} & \Pi_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Y_{1t-1} \\ Y_{2t-1} \\ Y_{3t-1} \\ Y_{4t-1} \end{bmatrix}$$

$$\Delta Y_t = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ \vdots \\ Y_{4,t-1} \end{bmatrix}$$

Specification of Deterministic Terms

$$\begin{bmatrix} \Delta Y_{1t} \\ \Delta Y_{2t} \\ \Delta Y_{3t} \\ \Delta Y_{4t} \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} (\beta_1 Y_{1t-1} + \dots + \beta_4 Y_{4t-1})$$

Johansen (1995) ใช้ specification ของ VECM ในรูปแบบต่อไปนี้

ΠY_{t-1}

$$\boldsymbol{Y}_t = \boldsymbol{\Phi D}_t + \boldsymbol{\Pi}_1 \boldsymbol{Y}_{t-1} + \dots + \boldsymbol{\Pi}_p \boldsymbol{Y}_{t-p} + \boldsymbol{\epsilon}_t, t = 1, \dots, T$$

รูปแบบของสมการเส้นตรง

โดยที่ $\boldsymbol{\Phi D}_t = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ เป็น deterministic terms โดยที่พฤติกรรมของ deterministic term ของ $\boldsymbol{Y}_t$ สามารถแยกออกได้เป็น 5 กรณี:

1. Model $H_2(r)$: $\boldsymbol{\mu}_0$ (no constant) – ทั้งสมการระยะ ~~สั้น~~ ยาว + สั้น $(Y_1 = \alpha + \beta Y_2)$
2. Model $H_1^*(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0 = \boldsymbol{\alpha\rho}_0$ (restricted constant) – Constant ใน LR
3. Model $H_1(r)$: $\boldsymbol{\mu}_t = \boldsymbol{\mu}_0$ (unrestricted constant) – Constant ใน LR, VECM $\Delta Y_{1t} = c_1$
4. Model $H^*(r) = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\alpha\rho}_1 t$ (restricted trend) – สมการ LR มี trend
5. Model $H(r) = \boldsymbol{\mu}_t = \boldsymbol{\mu}_0 + \boldsymbol{\mu}_1 t$ (unrestricted constant and trend)

มี trend ทั้ง LR, ECM

$(Y_1 = \alpha + \gamma t + \beta Y_2)$

$\Delta Y_{1t} = c_1 + \gamma_1 t + \alpha_1 (\ LR\ ) + \dots$

จำนวนของความสัมพันธ์ ?

Likelihood Ratio Tests for the Number of Cointegrating Vectors

$r = 0, 1, 2, \ldots, n-1$

หากเราใช้ $H(r)$ เป็นสัญลักษณ์แทน unrestricted cointegrated VECM แล้ว แบบจำลอง $I(1)$ ของ $H(r)$ สามารถเขียนเป็นเงื่อนไขของ rank ของ $\mathbf{\Pi}$ มีค่าน้อยกว่าหรือเท่ากับ r เราสามารถสร้าง nested set ของแบบจำลอง

$$H(0) \subset \ldots \subset H(r) \subset \ldots \subset H(n)$$

โดยที่ $H(0)$ แสดงถึง non-cointegrated VAR model โดยที่ $\mathbf{\Pi} = \mathbf{0}$ และ $H(n)$ แสดงถึง unrestricted stationary VAR(p)

เนื่องจาก rank ของ $\mathbf{\Pi}$ เป็นตัวระบุจำนวน cointegrating relationships ในตัวแปร $\boldsymbol{Y}_t$ ดังนั้น Johansen ได้เสนอตัวทดสอบ LR statistics สำหรับจำนวน cointegrating relationships ซึ่งเป็น LR statistics สำหรับการระบุ rank ของ $\mathbf{\Pi}$

Johansen's Trace Statistics

การทดสอบของ Johensen มีพื้นฐานจากการตัวประมาณค่า eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_n$ ของเมตริกซ์ $\mathbf{\Pi}$ (rank ของเมตริกซ์ มีค่าเท่ากับจำนวนของ eigenvalue ที่มีค่าไม่เท่ากับศูนย์)

ได้ไปเรื่อย
$r = 0, 1, 2, \ldots$

Johansen's Trace Statistic

Johansen's LR statistic ทดสอบสมมุติฐานในลักษณะเป็น nested hypotheses

ค่าคงที่ (0, 1, 2, ...)

$$H_0(r_0) : r = r_0 \quad \text{vs} \quad H_1(r_0) : r > r_0$$

$r = 0$ $\quad$ $r > 0$

โดยที่ LR statistic ถูกเรียกว่า *trace statistic* และคำนวณได้ดังนี้

$$\boxed{LR_{trace}(r_0) = -T \sum_{i=r_0+1}^{n} \ln(1 - \hat{\lambda}_i)}$$

> c.v. เพราะปฏิเสธ H_0

ถ้า $rank(\mathbf{\Pi}) = r_0$ แล้ว $\hat{\lambda}_{r_0+1}, \ldots, \hat{\lambda}_n$ จะมีค่าเข้าใกล้ศูนย์ และ $LR_{trace}(r_0)$ ควรจะมีค่าน้อย

การแจกแจง(เมื่อตัวอย่างใหญ่ asymptotic) ของ $LR_{trace}(r_0)$ เมื่อ H_0 เป็นจริงจะเป็น multivariate distribution ของ Dickey-Fuller unit-root distribution

Sequential Procedure

กระบวนทดสอบ จะมีลักษณะ ต่อเนื่อง ไปเรื่อยๆ

Johansen เสนอการทดสอบในลักษณะที่เป็นอันดับ (sequential testing procedure) โดยที่กระบวนการดังกล่าวจะระบุจำนวนของ cointegrating vectors ได้อย่าง consistently

- ขั้นตอนแรกเราทดสอบ $H_0(r=0)$ กับ $H_1(r>0)$ ถ้าเราไม่สามารถปฏิเสธสมมุติฐานหลัก แสดงว่าไม่มี cointegrating vectors (ไม่มีสมการสัมพันธ์ $r=0$ ระยะยาว) ไม่มี cointegration → VAR(p-1) ของ ΔY_t
  ถ้าเราปฏิเสธ $H_0(r=0)$ เราสามารถสรุปได้ว่า มี cointegrating vector อย่างน้อยหนึ่งค่า และทดสอบในขั้นต่อไป $r>0$ [มีสมการสัมพันธ์ระยะยาว (cointegration) แต่ยังไม่ทราบว่า $r=?$]
- ทดสอบ $H_0(r=1)$ กับ $H_1(r>1)$ ถ้าเราไม่สามารถปฏิเสธสมมุติฐานหลัก เราสรุปว่ามี cointegrating vector เท่ากับหนึ่ง ($r=1$) → สรุปได้ว่า Y_t เป็น cointegration และมี 1 สมการสัมพันธ์
  แต่ถ้า เราปฏิเสธสมมุติฐานหลัก แสดงว่า มี cointegrating vectors อย่างน้อยสองค่า $(r>1)$
- กระบวนการที่เป็นอันดับดังกล่าวจะดำเนินไปเรื่อยๆจนกระทั่งสมมุติฐานหลักไม่สามารถปฏิเสธ สรุปว่ามีความสัมพันธ์เท่ากับ r_0 จำนวนของ r

$H_0(r=2)$: vs $H_1: r>2$

จนกระทั่งเราไม่สามารถปฏิเสธ H_0

Johansen's Maximum Eigenvalue Statistic

Johansen ยังได้เสนอการทดสอบ LR statistic สำหรับสมมุติฐานต่อไปนี้

$$H_0(r_0) : r = r_0 \quad \text{vs} \quad H_1(r_0) : r = r_0 + 1$$

โดยเราเรียก LR statistic ว่า *maximum eigenvalue statistic* และคำนวณได้ดังนี้

$$LR_{max}(r_0) = -T \ln(1 - \hat{\lambda}_{r_0+1})$$

ตัวอย่างที่ 6.6

การทดสอบของ Johansen

1) VAR(p)
2) ประมาณค่า VECM(p-1)
3) ทดสอบ Rank ของ Π
4) r=? ประมาณค่า VECM

เราจะพิจารณาข้อมูลในตัวอย่างที่ 6.5 ด้วยแบบจำลอง VECM

lfuture , lspot

(1)

```
> fsprice<-cbind(lfutures,lspot)
> library(vars)
> var.mod<-VAR(fsprice, lag.max=4, ic=c("SC"))
> var.mod

VAR Estimation Results:
=======================

Estimated coefficients for equation lfutures:
=============================================
Call:
lfutures = lfutures.l1 + lspot.l1 + const

lfutures.l1    lspot.l1       const
 0.07057597  0.91654897  0.08517149


Estimated coefficients for equation lspot:
==========================================
Call:
lspot = lfutures.l1 + lspot.l1 + const

lfutures.l1    lspot.l1       const
 -0.1508610   1.1346435   0.1098629
```

, BIC

VAR(1)
↓
VECM(0)

ตัวอย่างที่ 6.6

unit root

cointegration

แบบจำลองที่เหมาะสมคือ $VAR(1)$ package ที่ใช้ในการทดสอบ Johansen และประมาณค่า VECM คือ urca โดยที่คำสั่งที่ใช้ในการทดสอบ Johansen คือ *ca.jo* โดยที่เราต้องระบุข้อมูล (*fsprice*) วิธีในการทดสอบ (trace) รูปแบบของสมการ (*const*) และจำนวน lag ในแบบจำลอง *VAR* อย่างไรก็ตามใน package นี้กำหนด lag ขั้นต่ำเท่ากับ 2 (VECM(1))

ข้อมูล

ค่าสถิติที่ใช้ [Trace Statistic]

รูปแบบของสมการ (5 รูปแบบ)

Order ของ VAR (P)

VECM(P-1)

ค่า K ต่ำสุดคือ K=2 (ตามจะเลือก K=1)

แสดงผลการทดสอบ

```
> library(urca)
> fsprice.rc<-ca.jo(fsprice, type=c("trace"), ecdet=c("const"), K=2)
> summary(fsprice.rc)

######################
# Johansen-Procedure #
######################

Test type: trace statistic , without linear trend and constant in cointegration

Eigenvalues (lambda):
[1]  3.779375e-01  1.865708e-02 -2.281477e-17

Values of teststatistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  2.34  7.52  9.24 12.97
r = 0  | 61.20 17.85 19.96 24.60
```

0.1 C.V. 0.05 0.01

1 future และ 1 spot cointegrated กันและมี 1 สมการ ความสัมพันธ์ระยะยาว

① $H_0: r=0$ vs $H_1: r>0$

Trace statistic $= 61.20 >$ C.V. (19.96)

ปฏิเสธ H_0 ที่ว่า $r=0$ และสรุป $r>0$

② $H_0: r=1$ vs $H_1: r>1$

Trace stat $= 2.34 <$ C.V. (9.24)

เราไม่สามารถปฏิเสธ H_0 และสรุปว่า $r=1$

ตัวอย่างที่ 6.6

เราสามารถสร้างสมการเวกเตอร์เอเรอคอเรกชัน ได้ด้วยคำสั่ง *cajorls* โดยระบุว่าใช้รูปแบบจาก *fsprice.rc* และจำนวนความสัมพันธ์เท่ากับ 1 ($r = 1$)

(3) ประมาณค่า สมการ VECM ผลการทดสอบ [รูปแบบสมการที่ได้ ≤ const / lag ของ VECM

```
> fsprice.vecm<-cajorls(fsprice.rc, r=1)
> fsprice.vecm
$rlm

Call:
lm(formula = substitute(form1), data = data.mat)

Coefficients:
               lfutures.d  lspot.d
ect1           -1.1705     -0.2296
lfutures.dl1   -1.0964     -0.3301
lspot.dl1       1.3179      0.5226

$beta
                     ect1
lfutures.l2    1.00000000
lspot.l2      -1.00496480
constant       0.03639024
```

การทดสอบ Johansen

VECM

dis eqm (u_{t-1})

Δlf_{t-1}, Δls_{t-1}

Δ lfuture Δ lspot

$$\Delta lf_t = -1.17\, u_{t-1} - 1.09\, \Delta lf_{t-1} + 1.31\, \Delta ls_{t-1}$$

long run (cointegration)

$$\Delta ls_t = -0.22\, u_{t-1} - 0.33\, \Delta lf_{t-1} + 0.5\, \Delta ls_{t-1}$$

$$u_{t-1} = lf_{t-1} - 1.0049\, ls_{t-1} + 0.036$$

ถ้า $u_{t-1} = 0 \Rightarrow lf_{t-1} = 1.0049\, ls_{t-1} - 0.036$