บทที่ 4

ระเบียบวิธีการศึกษา

เนื่องจากงานวิจัยนี้มีวัตถุประสงค์ในการประมาณค่าแบบจำลองเศรษฐกิจมหภาคของประเทศไทยในภาคการลงทุน ซึ่งเป็นส่วนหนึ่งของแบบจำลองเศรษฐกิจมหภาค ที่ประกอบไปด้วย ภาคการผลิต แรงงาน และระดับราคา ภาคการบริโภค ภาครัฐบาล ภาคการส่งออกและนำเข้า และภาคการเงิน ดังนั้นแนวทางการศึกษาในแบบจำลองเศรษฐกิจมหภาคในส่วนต่างๆ จึงสอดคล้องและเป็นไปในทิศทางเดียวกัน อีกทั้งจำเป็นต้องใช้ข้อมูลทางเศรษฐกิจที่เป็นข้อมูลอนุกรมเวลา (time series data) ซึ่งตัวแปรเหล่านั้นมักจะมีลักษณะ non-stationary กล่าวคือ ค่าเฉลี่ย (mean) และค่าความแปรปรวน (variances) จะมีค่าไม่คงที่แต่เปลี่ยนแปลงไปตามกาลเวลา ทำให้ความสัมพันธ์ระหว่างตัวแปรของสมการมีความสัมพันธ์ไม่แท้จริง (spurious regression) จะสังเกตได้จากค่าสถิติบางอย่างอาทิ ค่า t-statistic จะไม่เป็นการแจกแจงที่เป็นมาตรฐาน และค่า R^2 ที่สูง ในขณะที่ค่า Durbin-Watson (DW) statistic อยู่ในระดับต่ำแสดงให้เห็นถึงการมีความสัมพันธ์กันเองในระดับสูง จึงเป็นการยากที่จะยอมรับได้ในทางเศรษฐศาสตร์ (Enders, 1995) และ (Johnston and DiNardo, 1997)

วิธีที่จะจัดการกับข้อมูลที่มีลักษณะเป็น non-stationary ที่ได้รับความนิยมแพร่หลาย คือ วิธี cointegration และ error correction mechanism (รังสรรค์ หทัยเสรี, 2538) เนื่องจากเป็นเครื่องมือในการวิเคราะห์ความสัมพันธ์เชิงดุลยภาพระยะยาว (cointegrating relationship) วิธีดังกล่าวมีขั้นตอนในการศึกษาดังต่อไปนี้

1. ทดสอบความเป็น stationarity ของตัวแปรที่นำมาทำการศึกษาโดยวิธี ADF (Augmented Dickey-Fuller Test)
2. นำตัวแปรที่ทำการทดสอบโดยวิธี ADF แล้ว มาพิจารณาดุลยภาพในระยะยาว ตามแนวทางของ Johansen คือ
   (1) พิจารณาความล่าของตัวแปร(lag length) โดยวิธี LR (likelihood ratio test) เลือกรูปแบบแบบจำลองที่เหมาะสม
   (2) คำนวณหาจำนวน cointegrating vectors โดยวิธี maximal eigenvalue statistic (λ_{Max}) หรือวิธี eigenvalue trace statistic (λ_{Trace})

(3) เมื่อพบว่าแบบจำลองมีความสัมพันธ์ในระยะยาวแล้ว ใช้วิธีการของ error correction mechanism (ECM) มาคำนวณหาลักษณะการปรับตัวในระยะสั้น

จากที่ได้กล่าวถึงระเบียบวิธีการศึกษาสังเขป ต่อไปจะเป็นการนำเสนอขั้นตอนการศึกษาในส่วนต่างๆ อย่างละเอียด ซึ่งมีลำดับดังต่อไปนี้

4.1 Unit Root Test

การทดสอบ Unit Root ถือเป็นขั้นตอนแรกในการศึกษาภายใต้วิธี cointegration and error correction mechanism ขั้นตอนนี้จะเป็นการทดสอบตัวแปรทางเศรษฐกิจต่างๆ ที่จะใช้ในสมการเพื่อดูความเป็น stationary [I (0); integrated of order 0] หรือ non-stationary [I (d); d > 0, integrated of order d] การศึกษาส่วนใหญ่ที่ผ่านมาจะนิยมการทดสอบ unit root ที่เสนอโดย David Dickey และ Wayne Fuller (Pindyck and Rubinfeld, 1998) ซึ่งรู้จักกันดีในชื่อของ Dickey-Fuller test สามารถแบ่งออกได้เป็น 2 วิธี คือ

1) Dickey-Fuller Test (DF) ทำการทดสอบตัวแปรที่เคลื่อนไหวไปตามช่วงเวลามีลักษณะเป็น autoregressive model โดยสามารถเขียนรูปแบบของสมการได้ออกเป็น 3 รูปแบบคือ

$$X_t = \rho X_{t-1} + \varepsilon_t \tag{4.1}$$

$$X_t = \alpha_0 + \rho X_{t-1} + \varepsilon_t \tag{4.2}$$

$$X_t = \alpha_0 + \alpha_2 t + \rho X_{t-1} + \varepsilon_t \tag{4.3}$$

โดยที่ X_t คือตัวแปรที่เราทำการศึกษา α, ρ คือค่าคงที่ t คือ แนวโน้มเวลา และ ε_t คือตัวแปรสุ่ม มีการแจกแจงแบบปกติที่เหมือนกันและเป็นอิสระต่อกัน (independent and identical distribution) โดยมีค่าเฉลี่ยเท่ากับ 0 และค่าความแปร ปรวนคงที่ เขียนแทนด้วยสัญลักษณ์ $\varepsilon_t \sim iid\ (0, \sigma_\varepsilon^2)$

สมการแรกจะเป็นสมการที่แสดงถึง กรณีรูปแบบของตัวแปรที่ไม่มีค่าคงที่ ขณะที่สมการที่สองจะเป็นรูปแบบของสมการที่ปรากฏค่าคงที่ และสมการสุดท้ายแสดงถึงรูปแบบของสมการที่มีทั้ง ค่าคงที่ และ แนวโน้มเวลา

ในการทดสอบว่า X_t มีลักษณะเป็น stationary process ($X_t \sim I(0)$) หรือไม่ ทำการทดสอบโดยการแปลงสมการทั้งสามรูปแบบให้อยู่ในรูปของ first differencing (ΔX_t) ได้ดังนี้

$$\Delta X_t = X_t - X_{t-1} = \gamma X_{t-1} + \varepsilon_t \qquad (4.4)$$

$$\Delta X_t = X_t - X_{t-1} = \alpha_0 + \gamma X_{t-1} + \varepsilon_t \qquad (4.5)$$

$$\Delta X_t = X_t - X_{t-1} = \alpha_0 + \alpha_2 t + \gamma X_{t-1} + \varepsilon_t \qquad (4.6)$$

โดยที่ $\gamma = (\rho - 1)$

2) Augmented Dickey-Fuller Test (ADF) เป็นการทดสอบ unit root อีกวิธีหนึ่งที่พัฒนามาจาก DF Test เนื่องจากวิธี DF ไม่สามารถทำการทดสอบตัวแปรในกรณีที่เป็น serial correlation ในค่า error term (ε_t) ที่มีลักษณะความสัมพันธ์กันเองในระดับสูง ซึ่งจะมีการเพิ่ม lagged change $\left[\sum_{j=1}^{p} \lambda_j \Delta X_{t-j}\right]$ เข้าไปในสมการทางด้านขวามือ จะได้ว่า

$$\Delta X_t = X_t - X_{t-1} = \gamma X_{t-1} + \sum_{j=1}^{P} \lambda_j \Delta X_{t-j} + \varepsilon_t \qquad (4.7)$$

$$\Delta X_t = X_t - X_{t-1} = \alpha_0 + \gamma X_{t-1} + \sum_{j=1}^{P} \lambda_j \Delta X_{t-j} + \varepsilon_t \qquad (4.8)$$

$$\Delta X_t = X_t - X_{t-1} = \alpha_0 + \alpha_2 t + \gamma X_{t-1} + \sum_{j=1}^{P} \lambda_j \Delta X_{t-j} + \varepsilon_t \qquad (4.9)$$

พจน์ที่เราใส่เข้าไปนั้น จำนวน lagged term (p) ก็ขึ้นอยู่กับความเหมาะสมของแต่ละงานวิจัย (Pindyck and Rubinfeld, 1998) หรือสามารถใส่ส่วนล่าช้าไปกระทั่งไม่เกิดปัญหา autocorrelation ในส่วนของ error term (พิเชษฐ์ พรหมผุย, 2540)

โดยในการทดสอบสมมติฐานทั้งวิธี Dickey-Fuller test และวิธี Augmented Dickey-Fuller test ทดสอบว่าตัวแปรที่เราสนใจ (X_t) นั้นมี unit root หรือไม่ สามารถพิจารณาได้จากค่า γ ถ้าค่า γ มีค่าเท่ากับ 0 แสดงว่า X_t นั้นมี unit root ซึ่งสามารถเขียนสมมติฐานในการทดสอบได้ดังนี้

$$H_0 \quad : \quad \gamma = 0$$

$$H_1 \quad : \quad |\gamma| < 1$$

ทดสอบสมมติฐาน โดยเปรียบเทียบค่า t-statistic ที่คำนวณได้กับค่าที่ในตาราง Dickey-Fuller tables (แสดงในภาคภาคผนวก ค.) ซึ่งค่า t-statistic ที่จะนำมาทำการทดสอบสมมติฐานในแต่ละรูปแบบนั้นจะต้องนำไปเปรียบเทียบกับตาราง Dickey-Fuller tables ที่ต่างกัน กล่าวคือใช้ค่า τ ในรูปแบบของสมการที่ (4.4) และ (4.7) τ_μ ในรูปแบบของสมการที่ (4.5) และ (4.8) และ τ_τ ในรูปแบบของสมการที่ (4.6) และ (4.9) ถ้าสามารถปฏิเสธสมมติฐานได้ แสดงว่าตัวแปรที่นำมาทดสอบเป็น Integrated of order 0 แทนได้ด้วย $X_t \sim I(0)$ ถ้าต้องการทดสอบกรณีที่ γ ร่วมกับ drift term หรือร่วมกับแนวโน้มเวลา หรือ ทดสอบ γ ร่วมกับ drift term และแนวโน้มเวลา ในขณะเดียวกัน สามารถทดสอบโดยใช้ค่า F-statistic ซึ่งเป็น joint hypothesis (Φ_1, Φ_2 และ Φ_3) เป็นสถิติทดสอบทำการเปรียบเทียบกับค่า Dickey-Fuller Tables (Enders, 1995) ซึ่งในการทดสอบสมการที่ (4.5) และ (4.8) ทดสอบภายใต้สมมติฐานที่ว่า $\gamma = \alpha_0 = 0$ จะใช้ Φ_1 statistic

ขณะที่สมการที่ (4.6) และ (4.9) ทดสอบภายใต้สมมติฐาน $\alpha_2 = \gamma = \alpha_0 = 0$ ใช้ Φ_2 statistic สำหรับการทดสอบภายใต้สมมติฐาน $\alpha_2 = \gamma = 0$ ใช้ Φ_3 statistic ในการทดสอบ ซึ่งค่าสถิติดังกล่าวสามารถคำนวณได้ดังนี้

$$\Phi_i = \frac{(N-k)(SSR_R - SSR_{UR})}{r(SSR_{UR})}$$

โดยที่ SSR_R = the sum of square of residuals from the restricted model

SSR_{UR} = the sum of square of residuals from the unrestricted model

N = number of observations

k = number of parameters estimated in the unrestricted model

r = number of restrictions

กรณีที่ผลการทดสอบสมมติฐานพบว่า X_t มี unit root นั้นเราจะต้องนำค่า ΔX_t มาทำ differencing ไปเรื่อยๆ จนสามารถปฏิเสธสมมติฐานที่ว่า X_t เป็น non-stationary process ได้ เพื่อทราบ order of integration (d) ว่าอยู่ในระดับใด ($X_t \sim I(d); d > 0$)

ถ้าหากพบว่าข้อมูลดังกล่าวเป็น non-stationary process และมีอันดับความสัมพันธ์ของข้อมูล (order of integration) ที่มากกว่า 0 (ทดสอบว่า $X_t \sim I(d)$) หรือไม่ จะทำการทดสอบตามรูปแบบสมการดังต่อไปนี้ (วิโชติ ตั้งสัภคาพร, 2540)

$$\Delta^{d+1} X_t = \alpha_0 + \alpha_2 t + \ \rho - 1 \ \Delta^d X_{t-1} + \sum_{j=1}^{p} \lambda_j \Delta^{d+1} X_{t-j} + \varepsilon_t \quad (4.10)$$

เดิมทีภายหลังจากทราบค่า d (order of integration) แล้วเราจะต้องทำการ differencing ตัวแปร (เท่ากับ d+1 ครั้ง) ตามกระบวนการของ Box-Jenkin's method (1970) ก่อนที่จะนำตัวแปรดังกล่าวมาทำการ regression เพื่อหลีกเลี่ยงปัญหา spurious regression แม้ว่าวิธีนี้จะได้รับความนิยมใช้กันอย่างแพร่หลาย แต่การกระทำดังกล่าวจะทำให้แบบจำลองที่ได้จากการประมาณขาดข้อมูลใน

ส่วนของการปรับตัวของตัวแปรต่างๆ เพื่อเข้าสู่ดุลยภาพระยะยาว (รังสรรค์ หทัยเสรี, 2535) และ (Hataiseree, 1996)

หลังจากนั้น ในปี 1987 Robert F. Engle และ Clive W. J. Granger ได้เสนอบทความทางวิชาการเรื่อง Cointegration and Error Correction: Representation, Estimation and Testing ซึ่ง Cointegration and Error Correction เป็นเศรษฐมิติแนวใหม่ที่ใช้กับข้อมูลอนุกรมเวลาในการหาดุลยภาพระยะยาวจากข้อมูล โดยไม่ต้องผ่านการทำ differencing รายละเอียดและวิธีการศึกษาจะกล่าวในส่วนต่อไป

4.2 Cointegration and Error Correction Mechanism

ขั้นตอนนี้เป็นของการทดสอบตัวแปรต่างๆ ที่นำมาใช้ ว่ามีความสัมพันธ์ในระยะยาวตามที่ระบุไว้ในทฤษฎีหรือไม่ และพบว่าจะมีอยู่ 2 วิธีที่นิยมใช้ในการทดสอบตัวแปร คือ วิธีของ Johansen and Juselius (1990) และวิธีTwo-step Approach ของ Engle-Granger (1987)

การทดสอบดุลยภาพระยะยาวนั้น วิธีของ Johansen-Juselius และวิธีของ Engle-Granger มีแนวการทดสอบที่แตกต่างกัน กล่าวคือตามกระบวนการของ Engle-Granger จะทำการทดสอบดุลยภาพระยะยาวจากค่า error term ว่า stationary หรือไม่ ขณะที่การทดสอบของ Johansen methodology จะพิจารณาจากค่า rank ของ π (ดูเพิ่มเติมในขั้นที่ 2 การประมาณแบบจำลองและหาจำนวน Cointegrating Vectors) แม้ว่าวิธีการของ Engle-Granger จะเป็นที่นิยม แต่ยังมีความไม่เหมาะสมในกรณีที่ตัวแปรมากกว่า 2 ตัวแปรขึ้นไป (G□len, 1996) คือ

วิธีของ Engle-Granger จะทำการระบุว่าตัวแปรใดเป็นตัวแปรตาม และตัวแปรใดเป็นตัวแปรอิสระ ซึ่งไม่สามารถแสดง multiple cointegrating vector ได้ กรณีมีรูปแบบของความสัมพันธ์มากกว่า 1 รูปแบบ

แม้ว่าวิธี Johansen จะไม่ระบุว่า ตัวแปรใดเป็นตัวแปรอิสระ หรือตัวแปรใดเป็นตัวแปรตาม เราก็ยังสามารถจะทดสอบว่าตัวแปรใดเป็นตัวแปรอิสระ ตัวแปรใดเป็นตัวแปรตามวิธีของ Granger รวมทั้งพิจารณาให้สอดคล้องกับทฤษฎีและหลักการทางเศรษฐศาสตร์

ดังนั้นงานวิจัยนี้จึงเลือกใช้วิธีของ Johansen and Juselius (1990) ซึ่งมีพื้นฐานการวิเคราะห์บน รูปแบบของ vector autoregressive (VAR) model และเป็นกระบวนการทดสอบ cointegration ที่มีตัวแปรหลายตัว (Wolter, 1998) ในการทดสอบหาดุลยภาพระยะยาวซึ่งมีขั้นตอนการศึกษาดังนี้

ขั้นที่ 1 ทดสอบหา order of integration และความยาวของ lag ของตัวแปร

เริ่มต้นจากการทดสอบหา Order of Integration ของตัวแปรทุกตัวและหากพบว่าตัวแปรแต่ละตัวมี order of integration ต่างกัน Johansen จะไม่รวมตัวแปรเหล่านั้นไว้ด้วยกัน จากนั้นทำการทดสอบหาความยาวของ lag ของตัวแปร ซึ่งมี 3 วิธีที่นิยมนำมาพิจารณา ได้แก่ AIC: Akaike Information Criterion (Johnston and Dinardo, 1997) LR: Likelihood Ratio Test และ SBC: Schwartz Bayesian Criterion (Enders, 1995) สามารถคำนวณได้ดังต่อไปนี้

$$AIC = T\log\left|\Sigma\right| + 2N \tag{4.11}$$

$$LR = \ T - c \quad \log\left|\Sigma_r\right| - \log\left|\Sigma_u\right| \tag{4.12}$$

$$SBC = T\log\left|\Sigma\right| + N\log(T) \tag{4.13}$$

โดยที่ T = number of observations

c = number of parameters in the unrestricted system

$\left|\Sigma\right|$ = determinant of variance/covariance matrices of the residuals

$\left|\Sigma_r\right|$ = determinant of variance/covariance matrices of the restricted system

$\left|\Sigma_u\right|$ = determinant of variance/covariance matrices of the unrestricted system

N = total number of parameters estimated in all equations

ทดสอบ สมมติฐานหลัก โดยกำหนดจำนวน lagged term เท่ากับ r ในกรณีที่มีข้อจำกัด ขณะที่ u เท่ากับจำนวน lagged term ทั้งหมดที่เป็นไปได้ (ซึ่งขึ้นอยู่กับลักษณะและระยะเวลาของข้อมูลจากงานวิจัยแต่ละชิ้น) แล้วใช้การแจกแจงแบบ Chi-square (χ^2) ทดสอบสมมติฐานว่ามีจำนวน lagged term เท่ากับ r โดยมีจำนวนระดับความเป็นอิสระ เท่ากับจำนวนสัมประสิทธิ์ที่เป็นข้อจำกัด (coefficient restrictions) ถ้าค่า χ^2 ที่คำนวณได้น้อยกว่าระดับนัยสำคัญ แสดงว่า ยอมรับสมมติฐานหลัก หรือทำการทดสอบโดยใช้ F-test ในแต่ละสมการก็จะได้ผลการทดสอบเช่นเดียวกับการทดสอบโดยใช้ Chi-square เช่นกัน และหากพบว่าตัวแปรสามารถใช้ lagged term ได้หลายจำนวนควรเลือกใช้เทอมที่ยาวที่สุด อย่างไรก็ตามเราควรคำนึงถึงระดับความเป็นอิสระด้วย เนื่อง

จากถ้าเราใช้จำนวน lagged term มากจนเกินความจำเป็นจะทำให้สูญเสียระดับความเป็นอิสระ (Enders, 1995) ส่งผลถึง critical value ทำให้การยอมรับหรือปฏิเสธสมมติฐานบิดเบือนไป ส่วนกรณีสมการที่เพิ่มตัวแปรหุ่นเข้ามา จะทำให้ค่า c = np + 1 + dummy variables กล่าวคือ ในแต่ละสมการจะมี Parameters ทั้งหมดเท่ากับ จำนวน lagged term (p) ของตัวแปร(n) รวมกับ ค่าคงที่ และตัวแปรหุ่น

อย่างไรก็ดีความยาวของ lag length เปลี่ยนแปลงได้ ขึ้นอยู่กับความเหมาะสม เนื่องจากการเพิ่มหรือลดความยาวของ lag length อาจจะมีผลกระทบต่อเครื่องหมายของตัวแปรต่างๆ (เปลี่ยนจากเครื่องหมายบวก เป็นเครื่องหมายลบ หรือในทางกลับกันเปลี่ยนจากเครื่องหมายลบเป็นเครื่องหมายบวก) ซึ่งส่งผลต่อการอธิบายตามหลักการทฤษฎีทางเศรษฐศาสตร์

ขั้นที่ 2 ประมาณแบบจำลองและหาจำนวน cointegrating vector

สร้างรูปแบบของแบบจำลองซึ่งสามารถพิจารณาได้เป็น 5 รูปแบบ ดังนี้

รูปแบบที่ 1 VAR Model ไม่ปรากฏทั้งค่าคงที่และแนวโน้มเวลา

$$X_t = \sum_{i=1}^{p} A_i X_{t-i} + \varepsilon_t \tag{4.14}$$

ดังนั้น

$$\Delta X_t = \pi X_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \tag{4.15}$$

ซึ่งมีค่า π, π_i ดังนี้

$$\pi = \sum_{i=1}^{p} A_i - I \tag{4.21}$$

$$\pi_i = \sum_{j=i+1}^{p} A_j \tag{4.22}$$

โดยที่ X_t = the (n x 1) vectors of variables $(x_{1t}, x_{2t}, \dots, x_{nt})'$

A_i = the (n x n) matrix of parameters

I = the (n x n) identity matrix

ε_t = the (n x 1) vectors of error term with multivariate white noise

รูปแบบที่ 2 VAR Model ไม่มี แนวโน้มเวลา แต่จำกัดค่าคงที่ ใน cointegrating vector

$$\Delta X_t = \pi^* X^*_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \tag{4.16}$$

โดยที่

$$\pi^* = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1n} & a_{01} \\ \pi_{21} & \pi_{22} & \dots & \pi_{2n} & a_{02} \\ . & & & & \\ . & & & & \\ \pi_{n1} & \pi_{n2} & \dots & \pi_{nn} & a_{0n} \end{bmatrix} \quad (4.17)$$

$$X^*_{t-1} = \quad x_{1t-1}, x_{2t-1}, \dots, x_{nt-1}, 1^{\prime} \quad (4.18)$$

รูปแบบที่ 3 VAR Model มีเฉพาะค่าคงที่

$$X_t = A_0 + \sum_{i=1}^{p} A_i X_{t-i} + \varepsilon_t \quad (4.19)$$

ดังนั้น

$$\Delta X_t = A_0 + \pi X_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \quad (4.20)$$

โดยที่ A_0 = the (n x 1) vectors of constants $(a_{01}, a_{02}, \dots, a_{0n})'$

รูปแบบที่ 4 VAR Model มีค่าคงที่และจำกัดแนวโน้มเวลาใน cointegrating vector

$$\Delta X_t = A_0 + \pi^{**} X^{**}_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \quad (4.23)$$

โดยที่

$$\pi^{**} = \begin{bmatrix} \pi_{11} & \pi_{12} & \dots & \pi_{1n} & t_{01} \\ \pi_{21} & \pi_{22} & \dots & \pi_{2n} & t_{02} \\ . & & & & \\ . & & & & \\ \pi_{n1} & \pi_{n2} & \dots & \pi_{nn} & t_{0n} \end{bmatrix} \quad (4.24)$$

$$X^{**}_{t-1} = (X_{1t-1}, X_{2t-1}, \dots, X_{nt-1}, T)'$$

T = 1,2,3,…,n

รูปแบบที่ 5 VAR Model ประกอบไปด้วย ค่าคงที่ และ แนวโน้มเวลา (Cointegration with unrestricted intercepts and unrestricted trends in the VAR)

$$\Delta X_t = A_0 + A_1 T + \pi X_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \qquad (4.25)$$

โดยที่ A_1 = the (n x 1) vectors of time trend coefficient $(t_{01}, t_{02}, \ldots, t_{0n})'$

จากนั้นทำการคำนวณหาค่า characteristic roots ของ π Matrix (λ_{ij}) ของแบบจำลองทั้ง 5 รูปแบบ (กรณีรูปแบบที่ 2 คือ π^* และกรณีรูปแบบที่ 4 คือ π^{**}) สามารถหาได้จาก $|\pi - \lambda I| = 0$ (Johnston and DiNardo, 1997) หรือ

$$\left|\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}\right| = 0 \qquad (4.26)$$

ขณะที่ S_{00}, S_{01}, S_{10}, S_{11} คือ product moment metrics of the residuals โดย

$$S_{ij} = \frac{\sum_{t=1}^{T} R_{it} R'_{jt}}{T} \quad ; \quad \forall\, i, j = 0, 1 \qquad (4.27)$$

R_{ot} คือ residuals จากการประมาณสมการ $\Delta X_t = \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + R_{ot}$

R_{1t} คือ residuals จากการประมาณสมการ $X_{t-1} = \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + R_{1t}$

แล้วทำการทดสอบว่าแบบจำลองควรจะมีรูปแบบใดโดยกรณีของการทดสอบว่าแบบจำลองจะมี drift term หรือมีค่าคงที่ใน cointegrating vector นั้นทำการทดสอบ โดยตั้ง สมมติฐานหลัก (H_0) ว่าแบบจำลองมีค่าคงที่ใน cointegrating vector แล้วพิจารณาผลจากค่าสถิติ

$$-T \sum_{i=r+1}^{n} \left[\ln\left(1 - \lambda_i^*\right) - \left(1 - \lambda_i\right)\right] \qquad (4.28)$$

โดยที่ T = number of observations

n = number of variables

r = rank of π

λ_i^* = characteristic roots of restricted model (model with intercept term in the cointegrating vector)

λ_i = characteristic roots of unrestricted model(model with drift term)

ใช้การแจกแจงแบบ χ^2 โดยมีระดับความเป็นอิสระเท่ากับ n-r หากค่าสถิติที่คำนวณได้มากกว่าค่าในตาราง χ^2 แสดงว่ารูปแบบของแบบจำลองจะไม่มีค่าคงที่ใน cointegrating vector แต่จะ ปรากฏอยู่ในรูปแบบของ drift term

เมื่อทราบรูปแบบของแบบจำลองที่จะใช้แล้วให้คำนวณหาจำนวน cointegrating vector ซึ่งมีค่าเท่ากับ rank (r) ของ π matrix โดยใช้ likelihood ratio test ประกอบด้วย eigenvalue trace statistic[1](λ_{trace}) และ maximal eigenvalue statistic[2] (λ_{max}) ซึ่งมีวิธีการคำนวณดังต่อไปนี้

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{n} \ln\left(1-\hat{\lambda}_i\right) \quad (4.29)$$

$$\lambda_{max}(r, r+1) = -T \ln\left(1-\hat{\lambda}_{r+1}\right) \quad (4.30)$$

โดยที่ T = the number of usable observations

r = rank of π

n = number of variables

$\hat{\lambda}_i$ = the estimated value of characteristic roots (eigenvalues) obtained from the estimated π matrix

วิธีการของ trace statistic จะเริ่มต้นจากการทำการทดสอบสมมติฐานหลัก (H_0) โดยเปรียบเทียบค่า λ_{trace} ที่คำนวณได้ ว่ามากกว่า Critical Value หรือไม่ เปรียบเทียบค่า Statistics ในตาราง distribution of λ_{max} and λ_{trace} statistics (Enders, 1995) ถ้าค่าที่คำนวณได้มากกว่าก็จะปฏิเสธ H_0 โดยเริ่มจาก $H_0 : r = 0$ และ $H_1 : r > 0$ ถ้าปฏิเสธ H_0 ก็ทำการเพิ่มค่า r ในสมมติฐานครั้งละ 1 ไปเรื่อยๆ จนกระทั่งยอมรับ H_0 ลักษณะการตั้งสมมติฐานแสดงได้ดังตาราง

---

[1] Eigenvalue Trace Statistic = Trace Statistic = Trace Test

[2] Maximal Eigenvalue Statistic = Max. Statistic = Max. Test

ตารางที่ 4.1 แสดงการทดสอบสมมติฐานการหาจำนวน Cointegrating Vectors

Eigenvalue Trace Statistic Hypothesis Testing		Maximal Eigenvalue Statistic Hypothesis Testing	
H_0	H_1	H_0	H_1
$r = 0$	$r > 0$	$r = 0$	$r = 1$
$r \leq 1$	$r > 1$	$r = 1$	$r = 2$
$r \leq 2$	$r > 2$	$r = 2$	$r = 3$
$r \leq 3$	$r > 3$	$r = 3$	$r = 4$
.	.	.	.
.	.	.	.

ที่มา: (Walter Enders, 1995)

ซึ่งค่า r ที่ได้ก็คือจำนวน cointegrating vector โดยพิจารณาได้ 2 กรณี คือ กรณีที่ $r = 0$ จะได้ว่า สมการที่นำมาทดสอบนั้นเป็น VAR in first difference คือตัวแปรที่นำมาทดสอบไม่ cointegrated กัน (there exists no linear combination of the elements of X_t that is stationary) และกรณี $0 < r < n$ แสดงว่ามีจำนวน cointegrating vectors เท่ากับ r (Enders, 1995) และ (Haug et al, 1999) เมื่อทราบว่าจำนวน cointegration relations ว่ามีค่าเท่ากับ r (จำนวน common trends เท่ากับ r) เราก็จะทราบจำนวน common stochastic trends ว่ามีค่าเท่ากับ n-r เช่นกัน (Wolters, 1998) และ (Clarida and Taylor, 1997)

ขั้นที่ 3 ทำการ normalized cointegrating vector(s) และ speed of adjustment coefficients

ทำการ normalized cointegrating vector(s) และ speed of adjustment coefficients เพื่อปรับ β และ α ให้สอดคล้องกับรูปแบบสมการที่ต้องการ โดยที่

$$\pi = \alpha \beta' \qquad (4.31)$$

โดยที่ β' = the (n x r) matrix of cointegrating parameters

α = the (n x r) matrix of speed of adjustment parameters in Δx_t

(กรณีรูปแบบที่ 2 คือ π^{*} และกรณีรูปแบบที่ 4 คือ π^{**})

จากนั้นจึงทดสอบความถูกต้องของสมการว่าควรจะมีค่าคงที่และเครื่องหมายของสัมประสิทธิ์ตรงตามทฤษฎีหรือไม่ ทดสอบโดย χ^2 ซึ่งมีค่า ระดับความเป็นอิสระ เท่ากับจำนวนข้อจำกัดในการทดสอบ ให้เริ่มทดสอบจากค่าคงที่ก่อนแล้วจึงทดสอบ สัมประสิทธิ์ของตัวแปรอื่นๆ จนครบทุกตัว โดย cointegrating vectors จะมีคุณสมบัติในการปรับค่าข้อมูลที่เป็น non-stationary process ให้เป็น stationary process ได้ เมื่ออยู่ในรูปแบบของ linear combination $\beta' X_t \sim I(0)$; $X_t \sim I(1)$ (Charemza and Deadman, 1992) แต่ในกรณีทั่วไป ถ้า $X_t \sim I(d)$ และ X_t cointegrated of order d และ b ($X_t \sim CI(d, b)$) จะมี linear combination ของตัวแปร ที่ทำให้ $\beta' X_t \sim I(d\text{-}b)$ โดยที่ $d \geq b > 0$ เมื่อ β คือ cointegrating vector

ทำการ normalized โดยสมมติว่ามี lag length เท่ากับ 1 และ rank =1 จะได้รูปแบบดังนี้

$$\Delta x_{1t} = \pi_{11} x_{1t-1} + \pi_{12} x_{2t-1} + \ldots + \pi_{1n} x_{nt-1} + \varepsilon_{1t} \tag{4.32}$$

ถ้าทำการ normalized โดยคำนึงถึงตัวแปร X_{1t-1} จะได้ว่า

$$\alpha_1 = \pi_{11} \text{ และ } \beta_{ij} = \frac{\pi_{ij}}{\pi_{11}} \tag{4.33}$$

$$\Delta x_{1t} = \alpha_1 (x_{1t-1} + \beta_{12} x_{2t-1} + \ldots + \beta_{1n} x_{nt-1}) + \varepsilon_{1t} \tag{4.34}$$

ฉะนั้น $x_{1t-1} + \beta_{12} x_{2t-1} + \ldots + \beta_{1n} x_{nt-1} = 0$ คือ long-run relationship

$\beta = 1 \ \ \beta_{12} \ \ \ldots . \ \ \beta_{1n}$ คือ cointegrating vector

α_1 คือ speed of adjustment coefficient

ค่าความเร็วในการปรับตัว หรือ speed of adjustment coefficient นั้น มีค่าอยู่ระหว่าง 0 และ −2 (Maddala and In-Moo, 1998) แต่มีการศึกษาแบบจำลองเศรษฐกิจมหภาคของ Federal Reserve Bank of ST. Louis เรื่อง A Vector Error-Correction Forecasting Model of the U.S. Economy ได้ทำการศึกษาโดยอาศัยวิธี Joahansen Methodology พบว่าผลของค่า speed of adjustment นั้นไม่ได้

อยู่ในช่วงดังที่กล่าวมา โดยบางส่วนนั้นมีค่าติดลบที่มากกว่า –2 และบางส่วนก็พบว่าสามารถเป็นค่าที่มากกว่าศูนย์ได้ (Hoffman and Rasche, 1997)

ขั้นที่ 4 ตรวจสอบสมการ

พิจารณา error correction model โดยใช้วิธี causality tests และให้เหตุผลทางเศรษฐศาสตร์ ตัวแปรใดเป็นตัวแปรตาม ตัวแปรใดเป็นตัวแปรอิสระ ซึ่งรูปแบบของสมการ error correction model จากสมการที่ (4.15), (4.16), (4.20), (4.23) และ (4.25) คือ

$$\Delta X_t = \pi X_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \qquad (4.35)$$

$$\Delta X_t = \pi^* X^*_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \qquad (4.36)$$

$$\Delta X_t = A_0 + \pi X_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \qquad (4.37)$$

$$\Delta X_t = A_0 + \pi^{**} X^{**}_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \qquad (4.38)$$

$$\Delta X_t = A_0 + A_1 T + \pi X_{t-1} + \sum_{i=1}^{p-1} \pi_i \Delta X_{t-i} + \varepsilon_t \qquad (4.39)$$

ขั้นที่ 5 ทำการทดสอบความสามารถในการอธิบายของแบบจำลอง โดยการทำ simulation ในโปรแกรม Eviews แล้วทำกราฟและหาค่าสถิติที่ใช้ในการทดสอบความสามารถในการอธิบาย ซึ่งได้แก่ root mean squared error, mean absolute error, mean absolute percentage error, Theil's inequality coefficient ซึ่งประกอบด้วย bias proportion, variance proportion และ covariance proportion ซึ่งมี สูตรในการคำนวณดังนี้

Root Mean Squared Error $= \sqrt{\dfrac{1}{h+1}\sum_{t=s}^{s+h}(\hat{y}_t - y_t)^2}$

Mean Absolute Error $= \dfrac{1}{h+1}\sum_{t=s}^{s+h}\left|\hat{y}_t - y_t\right|$

Mean Absolute Percentage Error $= \frac{1}{h+1}\sum_{t=s}^{s+h}\left|\frac{\hat{y}_t - y_t}{y_t}\right|$

Theil's Inequality Coefficient $= \dfrac{\sqrt{\frac{1}{h+1}\sum_{t=s}^{s+h}(\hat{y}_t - y_t)^2}}{\sqrt{\frac{1}{h+1}\sum_{t=s}^{s+h}(\hat{y}_t)^2} + \sqrt{\frac{1}{h+1}\sum_{t=s}^{s+h}(y_t)^2}}$

Bias Proportion $= \dfrac{(\bar{\hat{y}} - \bar{y})}{\sum(\hat{y}_t - y_t)^2 / h}$

Variance Proportion $= \dfrac{(s_{\hat{y}} - s_y)^2}{\sum(\hat{y}_t - y_t)^2 / h}$

Covariance Proportion $= \dfrac{2(1-r)s_{\hat{y}}s_y}{\sum(\hat{y}_t - y_t)^2 / h}$

โดยที่ $\hat{y}$ = forecasted value

y = actual value

$\bar{\hat{y}}$ = means of $\hat{y}$

$\bar{y}$ = means of y

$s_{\hat{y}}$ = standard deviations of $\hat{y}$

s_y = standard deviations of y

r = correlation between $\hat{y}$ and y

ซึ่งตัวที่พยากรณ์เริ่มนับค่า t ตั้งแต่ t = S, S+1 ,..., S+h