บทที่ 11
ตัวแปรย้อนเวลา
(Distributed Lag; Lag Variable)

11.1 ความหมายและความจำเป็น

ตัวแปรย้อนเวลาได้รับความนิยมอย่างกว้างขวาง และโดยปกตินิยมนำมาใช้กรณีที่แบบจำลองนั้นมีส่วนเกี่ยวข้องผูกพันกับพฤติกรรมของมนุษย์ที่มักชลอเวลาในการตัดสินใจดำเนินการเรื่องหนึ่งเรื่องใดจนกว่าสถานการณ์จะเหมาะสม (ความหมายนี้คือความหมายของ Lead ถ้าถือเอาจุดใดจุดหนึ่งในอดีตเป็นจุดเริ่มพิจารณา) หรือหากจะมองในมุมกลับ มนุษย์จะตัดสินใจดำเนินการเรื่องหนึ่งเรื่องใดได้ก็ต่อเมื่อได้พิจารณาทบทวนสถานการณ์ในอดีตอย่างถี่ถ้วนเสียก่อน (ความหมายนี้คือความหมายของ Lag ถ้าถือเอาจุดในปัจจุบันเป็นจุดเริ่มต้น) แบบจำลองที่นำพฤติกรรมของมนุษย์มาแทรกปนอยู่จึงมีผลให้แบบจำลองมีลักษณะที่เคลื่อนไหวไปได้ตามกาลเวลา (Dynamic Model) สอดคล้องกับพฤติกรรมของมนุษย์ที่ต้องอาศัยเวลา (Timing) ในการตัดสินใจ เช่น ถ้าเรามีรายได้สูงขึ้นเราอาจตัดสินใจเปลี่ยนรถยนต์จากยี่ห้อโตโยต้าเป็นรถเบ้นซ์ แต่ชลอเวลาตัดสินใจสักระยะหนึ่งเพื่อคอยดูว่ารถเบ้นซ์รุ่นที่จะออกใหม่มีสิ่งอำนวยความสะดวกอะไรบ้าง หรือเมื่อเรามีรายได้สูงขึ้นเราอาจตัดสินใจขยายบ้านให้กว้างขึ้น หากมองในมุมกลับ ณ จุดที่ตัดสินใจแล้ว (หรือกำลังตัดสินใจ) จะพบว่าเราจะคอยเวลาจนกว่าจะมีรายได้สูงพอเสียก่อนจึงค่อยตัดสินใจเปลี่ยนมาใช้รถเบ้นซ์หรือเราจะคอยเวลาจนกว่าจะมีรายได้สูง และบ้านเก่าเกินไปจึงค่อยตัดสินใจขยายบ้านและตกแต่งภายในใหม่ ขอให้สังเกตว่าการรอเวลาตัดสินใจนั้นมิได้เกิดขึ้น ณ จุดใดจุดหนึ่ง (Point of Time) แต่จะกระจายตัว (Distribute) ไปทั่วทุกจุดของเวลาในอดีต (ถ้ามองในรูปของ Lag) หรือทั่วทุกจุดของเวลาในอนาคต (ถ้ามองในรูปของ Lead) ด้วยเหตุนี้เราจึงเรียกตัวแปรชนิดนี้ว่า Distributed Lag และขอให้สังเกตว่าโดยทั่วไปแล้ว เรามักไม่ทราบระยะเวลารอคอยเพื่อชลอการตัดสินใจ (Lag Length) อย่างแจ้งชัด ซึ่งอาจเป็นระยะเวลาสั้น ๆ หรือยาวนานเพียงใดก็ได้ ด้วยเหตุนี้การศึกษา Distributed Lag จึงต้องศึกษาทั้งในแง่ของ Finite Lag และ Infinite Lag ประกอบกับมนุษย์พิจารณาเห็นความสำคัญและอิทธิพลของเวลาในแต่ละจุดต่างกัน เช่นบางคนอาจเห็นว่าราคาน้ำมันดิบเมื่อปีกลายมีส่วนต่อแนวโน้มของราคาในปัจจุบันสูงที่สุด บางคนอาจเห็นว่าราคาเมื่อ 5 ปี น่าจะมีผลต่อราคาปัจจุบันที่สุดส่วนราคาในปีอื่น ๆ ก็มีอิทธิพลแต่ค่อย ๆ ลดหลั่นและมีความสำคัญน้อยลง ด้วยเหตุนี้รูปทรงหรือโครงสร้างของ Lag จึงผิดแปลกแตกต่างกันไปในหลายลักษณะ เพื่อให้สอดคล้องกับสถานการณ์เฉพาะแบบเฉพาะตัว ซึ่งเราได้ศึกษากันโดยละเอียดต่อไป

กล่าวโดยสรุปแล้ว Distribute Lag ก็คือฟังก์ชั่นที่ช่วยบรรยายภาพความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระว่า ตัวแปรอิสระเดิมในอดีตส่งผลกระทบต่อความผันแปรค่าของตัวแปรตาม (Y) หรือไม่ และขอให้สังเกตว่าในเรื่องนี้กาลเวลาได้รับการพิจารณาในฐานะปัจจัยที่ควบคุมความเคลื่อนไหวของตัวแปรตาม แต่ไม่ปรากฏโฉมหน้าออกมาในรูปของตัวแปรอิสระ t หากแฝงอยู่ในตัวแปรอิสระในต่างกาลต่างวาระ

11.2 Finite Distributed Lag

11.2.1 ลักษณะทั่วไป

Finite Distributed Lag Model คือแบบจำลองที่นำเอาอิทธิพลทั้งปวงในอดีตของตัวแปรอิสระมาร่วมพิจารณาในฐานะเป็นตัวแปรอิสระ หรือกล่าวอีกนัยหนึ่งก็คือตัวแปรอิสระ X_t มีอิทธิพลย้อนทวนไปในอดีตมากเพียงใดก็ให้นำตัวแปรอิสระ X_t ในอดีตมาร่วมเป็นตัวแปรอิสระมากเพียงนั้น ทั้งนี้ถือว่า Lag Length เป็นจำนวนที่วัดค่าได้ เช่น เราทราบว่าราคาสินค้า ก. ที่นับย้อนไปถึง 5 ปี ยังมีอิทธิพลต่อความเคลื่อนไหวในยอดขายสินค้า ก. ในกรณีตัวแปรนี้ $X_t, X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}, X_{t-5}$ จะทำหน้าที่เป็นตัวแปรอิสระ สมการถดถอยสำหรับกรณีนี้คือ

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + \beta_5 X_{t-5} + u_t \; ; \; t = 6,7,\ldots,T$$

เมื่อพิจารณาค่าของตัวแปรแต่ละตัวจะพบลักษณะข้อมูลดังตาราง

	period ที่	Y_t	X_t	X_{t-1}	X_{t-2}	X_{t-3}	X_{t-4}	X_{t-5}
ค่าสังเกต Z_t ไม่สมบูรณ์	1	Y_1	X_1					
	2	Y_2	X_2	X_1				
	3	Y_3	X_3	X_2	X_1			
	4	Y_4	X_4	X_3	X_2	X_1		
	5	Y_5	X_5	X_4	X_3	X_2	X_1	
ค่าสังเกต Z_t สมบูรณ์ทั้งสิ้น T − n ชุดในที่นี้ n = 5	6	Y_6	X_6	X_5	X_4	X_3	X_2	X_1
	7	Y_7	X_7	X_6	X_5	X_4	X_3	X_2
	8	Y_8	X_8	X_7	X_6	X_5	X_4	X_3
	9	Y_9	X_9	X_8	X_7	X_6	X_5	X_4
	10	Y_{10}	X_{10}	X_9	X_8	X_7	X_6	X_5
	11	Y_{11}	X_{11}	X_{10}	X_9	X_8	X_7	X_6
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	T − 5	Y_{T-5}	X_{T-5}	X_{T-6}	X_{T-7}	X_{T-8}	X_{T-9}	X_{T-10}
	T − 4	Y_{T-4}	X_{T-4}	X_{T-5}	X_{T-6}	X_{T-7}	X_{T-8}	X_{T-9}
	T − 3	Y_{T-3}	X_{T-3}	X_{T-4}	X_{T-5}	X_{T-6}	X_{T-7}	X_{T-8}
	T − 2	Y_{T-2}	X_{T-2}	X_{T-3}	X_{T-4}	X_{T-5}	X_{T-6}	X_{T-7}
	T − 1	Y_{T-1}	X_{T-1}	X_{T-2}	X_{T-3}	X_{T-4}	X_{T-5}	X_{T-6}
	T	Y_T	X_T	X_{T-1}	X_{T-2}	X_{T-3}	X_{T-4}	X_{T-5}

ขอให้สังเกตว่าตามตัวอย่างนี้มี Lag Length เท่ากับ 5 ข้อมูลของค่าสังเกต Z_t ใน 5 period แรกจะไม่มีสมบูรณ์ซึ่งจะตัดทิ้งคือ

$$Z_1 = (Y_1, X_1, --, --, --, --, --)$$

$$Z_2 = (Y_2, X_2, X_1, --, --, --, --)$$

$$Z_3 = (Y_3, X_3, X_2, X_1, --, --, --)$$

$$Z_4 = (Y_4, X_4, X_3, X_2, X_1, --, --)$$

$$Z_5 = (Y_5, X_5, X_4, X_3, X_2, X_1, --)$$

ขนาดตัวอย่าง T จึงลดลงไป 5 หน่วยเหลือเพียงเฉพาะค่าสังเกตที่สมบูรณ์ (Complete Sample) เพียง T-5 ชุดคือ $Z_6, Z_7, \ldots, Z_T$ ขอให้สังเกตว่าเมื่อกำหนด Lag Length เท่ากับ 5 ค่าสังเกตที่สมบูรณ์จะเริ่มต้นที่ Z_{5+1} เป็นต้นไป ถึง Z_T ซึ่งถ้าหากเป็นกรณีทั่วไปที่กำหนดให้มี Lag Length เท่ากับ n ค่าสังเกตที่สมบูรณ์จะเริ่มต้นที่ Z_{n+1} เป็นต้นไป ถึง Z_T

สำหรับกรณีทั่วไปเมื่อมี Lag Length เท่ากับ n และกำหนดให้มีค่าสังเกต T ชุด คือ t = 1, 2, ..., T เราสามารถเสนอรูปทั่วไปของ Finite Distributed Lag Function ได้ดังนี้

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \ldots + \beta_n X_{t-n} + u_t \quad ; \ t = n+1, n+2, \ldots, T$$

หรือ

$$Y_t = \sum_{i=0}^{n} \beta_i X_{t-i} + u_i \quad ; \ t = n+1, n+2, \ldots, T$$

ซึ่งสามารถเสนอในรูปของแมตริกซ์ได้ดังนี้คือ

$$Y = X\beta + U$$

โดยที่ Y เป็นเวคเตอร์ขนาด $(T-n)\times 1$ แมตริกซ์ X เป็น Design Matrix ขนาด $(T-n)\times(n+1)$ และ β เป็นเวคเตอร์ของพารามิเตอร์หรือตัวถ่วงน้ำหนักขนาด $(n+1)\times 1$ ดังรายละเอียดต่อไปนี้

$$\begin{bmatrix} Y_{n+1} \\ Y_{n+2} \\ Y_{n+3} \\ \vdots \\ Y_Y \end{bmatrix} = \begin{bmatrix} X_{n+1} & X_n & X_{n-1} & \cdots & X_1 \\ X_{n+2} & X_{n+1} & X_n & \cdots & X_2 \\ X_{n+3} & X_{n+2} & X_{n+1} & \cdots & X_3 \\ \vdots & \vdots & \vdots & & \vdots \\ X_T & X_{T-1} & X_{T-2} & \cdots & X_{T-n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} u_{n+1} \\ u_{n+2} \\ u_{n+3} \\ \vdots \\ u_T \end{bmatrix}$$

จากสมการ $Y = X\beta + U$ ข้างต้นนี้เราสามารถประมาณค่าเวคเตอร์ β ได้โดยอาศัย OLS และสามารถพิสูจน์ได้โดยง่ายว่า $\hat{\beta} = (X'X)^{-1}X'Y$ เป็น BLUE ของ β และถ้าถือว่า $U \sim N(0,\sigma^2 I)$ เราย่อมพิสูจน์ได้ว่า $\hat{\beta} \sim N[\beta,\sigma^2(X'X)^{-1}]$ ซึ่งเอื้ออำนวยให้เราสามารถทดสอบสมมุติฐาน และคำนวณหาเขตเชื่อมั่นของ β ได้

อย่างไรก็ตาม เนื่องจากตัวแปรอิสระก็คือ X_t และ Lagged X_t คือ $X_{t-1}, X_{t-2}, \ldots, X_{t-n}$ ซึ่งเป็นตัวแปรเดียวกัน เพียงแต่เริ่มนับค่าในต่างจุดเวลา (Point of Time) ด้วยเหตุนี้ตัวแปรอิสระ $X_t, X_{t-1}, \ldots, X_{t-n}$ จึงมีลักษณะใกล้ที่จะมี Linearly Dependent หรือแมตริกซ์ X มีลักษณะของ Near Singularity ซึ่งเป็นดัชนีชี้ให้เห็นว่าเรากำลังเผชิญกับปัญหา Multicolineary ซึ่งในกรณีนี้ OLS - Estimator จะให้ค่าประมาณที่ไม่ถูกต้องแม่นยำ หนทางแก้ปัญหานี้อาจมีอยู่หลายทาง เช่น การใช้ Ridge Regression การใช้ Principle Component Method การใช้ Restricted Least Square โดยกำหนดรูปแบบของ Priori แบบต่าง ๆ กันเช่น Exact Restriction, Stochastic restriction, Prior Distribution เป็นต้น และสำหรับในการณีของ Distributed Lag Model เรานิยมแก้ปัญหา Multicolinearity ด้วยการกำหนดข้อจำกัด (Restriction) ให้ในรูปของการกำหนดรูปทรงของ Lag Distribution หรือรูปแบบที่น่าจะเป็นไปได้ของพารามิเตอร์ โดยเรากำหนดการแจกแจง[1]ดังกล่าวได้หลายลักษณะแตกต่างกันไปดังต่อไปนี้ ซึ่งผู้เขียนจะแนะนำไปเพียง 3 แบบ นอกเหนือไปจากที่กล่าวไว้แล้วนี้ผู้สนใจสามารถติดตามศึกษาได้จากเอกสารอ้างอิง

•

11.2.2 Arithematic Lag

Arthematic Lag ถือว่าอิทธิพลของ Lagged X_t จะค่อย ๆ ลดลงตามลำดับในลักษณะของเส้นตรงตามลำดับกาลเวลาที่ย้อนไกลไปในอดีตดังนี้

ให้ $$\beta_j = \begin{cases} (n+1-i)\alpha & \text{เมื่อ } i = 0,1,2,\ldots,n \\ 0 & \text{เมื่อ } i \text{ มีค่าอื่น ๆ} \end{cases}$$

ด้วยเหตุนี้ค่าของ β จึงเป็นอนุกรมเลขคณิตเมื่อแทนค่า $i = 0, 1, 2, \ldots, n$ ดังนี้

เมื่อ $i = 0$ จะได้ $\beta_0 = (n+1)\alpha$

เมื่อ $i = 1$ จะได้ $\beta_1 = (n-0)\alpha$

เมื่อ $i = 2$ จะได้ $\beta_2 = (n-1)\alpha$

เมื่อ $i = 3$ จะได้ $\beta_3 = (n-2)\alpha$

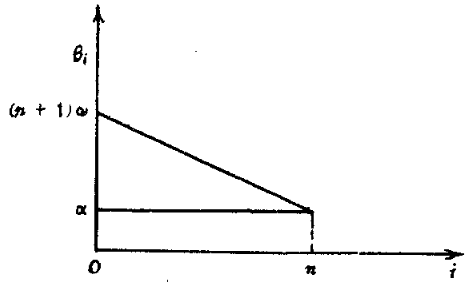
$\vdots$

เมื่อ $i = n$ จะได้ $\beta_n = (n+1-n)\alpha$

---

[1] เรียก β_j ใน Distributed Lag Model ว่า Lag Weight

ซึ่งเมื่อทดลองนำค่าของ β ไปพล๊อตโดยที่แกนนอนคือแกน i แกนตั้งคือแกน β จะพบว่าทางเดินของจุด (i,β_i) เป็นเส้นตรงที่ลาดลงทางขวาดังภาพ ซึ่งแสดงว่า X_{t-1} มีอิทธิพลต่อ Y_t สูงที่สุด X_{t-2} มีอิทธิพลต่อ Y_t ในลำดับที่ 2 .... X_{t-n} มีอิทธิพลต่อ Y_t น้อยที่สุด



An arithmetic distributed lag.

สำหรับการประมาณค่าของ β_i ให้ดำเนินการประมาณค่า α เสียก่อนดังนี้

แทนค่า $\beta_i = (n+1-i)\alpha;\ i = 0,1,\ldots,n$ ลงในสมการ

$Y_t = \sum_{i=0}^{n} \beta_i X_{t-i} + u_t$ จะได้สมการใหม่ดังนี้คือ

$$Y_t = (n+1)\alpha X_t + n\alpha X_{t-1} + (n-1)\alpha X_{t-2} + \ldots + \alpha X_{t-n} + u_t \quad ;\ t = n+1, n+2, \ldots, T$$

$$= \alpha[(n+1)X_t + nX_{t-1} + (n-1)X_{t-2} + \ldots + X_{t-n}] + u_t$$

หรือ $$Y_t = \alpha \sum_{i=0}^{n} (n+1-i)X_{t-i} + u_t \quad ;\ t = n+1, n+2, \ldots, T$$

ซึ่งเป็นสมการที่มี Unknown Parameter ที่จะต้องประมาณค่าเพียงตัวเดียวกัน α ซึ่งในกรณีเช่นนี้เราย่อมประมาณค่า α ได้โดยอาศัย OLS ดังนี้

ให้ $$W_t = \sum_{i=0}^{n} (n+1-i)X_{t-i} \quad ;\ t = n+1, n+2, \ldots, T$$

ดังนั้นสมการถดถอยที่ต้องการคือ

$$Y_t = \alpha W_t + u_t \quad ;\ t = n+1, n+2, \ldots, T$$

ดังนั้น $\hat{\alpha} = \Sigma\, w_t y_t \,/\Sigma\, w_t^2$ และ $\hat{\alpha} \sim N(\alpha, \sigma^2/\Sigma w_y^2)$ โดยที่

$$\hat{\sigma}^2 = \frac{1}{(T-n)-1}\, [\Sigma y_t^2 - \hat{\alpha}\Sigma\, w_t y_t$$

และเนื่องจาก $\beta_i = (n + 1 - i)\alpha \quad , i = 0,1,2,\ldots,n$

ดังนั้น $\hat{\beta}_i = (n + 1 - i)\hat{\alpha}$; $i = 1,2,\ldots,n$ และ $\hat{V}(\hat{\beta}_i) = (n + 1 - i)^2\, \hat{\sigma}^2/\Sigma w_t^2$

และ $\hat{\beta}_i \sim N[\beta_i, (n + 1 - i)^2\, \hat{\sigma}^2/\Sigma w_t^2]$

และสมการประมาณค่าที่ต้องการคือ

$$Y_t = \hat{\alpha}[(n+1)X_t + nX_{t-1} + (n-1)X_{t-2} + \ldots + X_{t-n}]$$

ปัญหาที่สำคัญของวิธีนี้คือเรามักไม่ทราบ Lag Length (n) ซึ่งจะได้พิจารณาปัญหานี้อีกครั้งหนึ่งเมื่อศึกษาถึง Almon Distributed Lag ทั้งนี้เพราะ Arithematic Lag เป็นกรณีเฉพาะของ Almon Distributed Lag

11.2.3 Inverted - V Lag

ในกรณีที่งานวิจัยต้องอาศัยข้อมูลรายเดือน (Monthly Data) หรือข้อมูลรายไตรมาส (Quarterly Data) ซึ่งโดยธรรมชาติมักจะมีลักษณะที่มีแนวโน้มที่จะ Boom ในช่วงกลางปี กล่าวคือข้อมูลจะค่อย ๆ มีค่าสูงขึ้น ๆ จนถึงจุดสูงสุดในระยะกลางปีจากนั้นจึงค่อย ๆ ลดค่าลงจนถึงจุดต่ำสุดในวาระปลายปี เช่นในอุตสาหกรรมการผลิตสินค้าประเภทต่าง ๆ เราจะพบว่าความต้องการซื้อสินค้าประเภททุนสำหรับการผลิตทางการเกษตร เช่น ปุ๋ย ยาปราบศัตรูพืช เครื่องมือการเกษตร จะมีอยู่สูงที่สุดในวาระกลางปีซึ่งเป็นฤดูการผลิต โดยความต้องการซื้อจะค่อย ๆ ทวีขึ้นจากปริมาณต่ำในช่วงต้นปีเรื่อยไปจนถึงวาระประมาณกลางปี จากนั้นจึงค่อย ๆ ลดปริมาณลงเรื่อย ๆ หากธรรมชาติของข้อมูลมีลักษณะเช่นนี้การกำหนดให้ Lag Distribution มีรูปทรงแบบ Arithematic Lag ย่อมเป็นเรื่องที่ไม่ถูกต้อง ทางที่ถูกที่ควรกำหนดให้มีรูปทรงแบบอักษร V กลับคือ Λ ซึ่งจะสอดคล้องกับธรรมชาติของข้อมูลมากกว่า และเนื่องจาก β_i คือพารามิเตอร์ที่แสดงระดับอิทธิพลของตัวแปรอิสระ X_{t-i} การกำหนดความเคลื่อนไหวหรือรูปแบบของ β_i ย่อมหมายถึงการกำหนดรูปแบบอิทธิพลของตัวแปร X_{t-i} ที่มีต่อตัวแปรสุ่ม Y_t ไปในขณะเดียวกัน

สำหรับ Inverted - V Lag นั้นเราจะกำหนดให้มี Lag Length เป็นค่าคงที่ n และเพื่อความสะดวกเราจะกำหนดให้ n เป็นเลขคู่

กำหนดให้ $s = \frac{n}{2}$ คือ Period ที่เป็นจุดกึ่งกลางของ Lag Length อันเป็นจุดที่ β_i มีค่าสูงที่สุด และเพื่อที่จะกำหนดค่าของ β_i ให้มีลักษณะสอดคล้องกับการแจกแจงรูปอักษร V คว่ำ เราจึงกำหนดค่า β_i คราวละครึ่งช่วงของ Lag Length ดังนี้

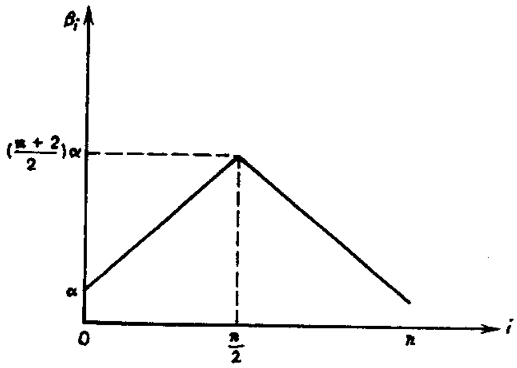
$$\text{ให้} \quad \beta_i = \begin{cases} (i+1)\alpha & \text{ถ้า } 0 \leqslant i \leqslant s \\ (n+1-i)\alpha & \text{ถ้า } i \geqslant s+1 \\ 0 & \text{ถ้า } i \text{ มีค่าปรากฏ ณ ตำแหน่งอื่น} \end{cases}$$

ขอให้สังเกตว่า ถ้า $i = 0, 1, 2, \ldots, s$ แล้ว β_i จะมีลักษณะเป็นอนุกรมเลขคณิตที่ทางเดินของจุด (i, β_i) เป็นเส้นตรงที่เอียงขึ้นดังนี้คือ

$$\alpha,\ 2\alpha,\ 3\alpha,\ 4\alpha, \ldots, (s+1)\alpha$$

ขณะเดียวกัน ถ้า $i = s+1, s+2, s+3, \ldots, n$ แล้ว β_i จะมีลักษณะเป็นอนุกรมเลขคณิตที่ทางเดินของจุด (i, β_i) เป็นเส้นตรงที่ลาดลงดังนี้คือ

$$(n-s)\alpha, (n-s-1)\alpha, (n-s-2)\alpha, \ldots, \alpha$$



The inverted *V* lag.

สำหรับการประมาณค่าสมการถดถอย $Y_t = \sum_{i=0}^{n} \beta_i X_{t-i} + u_i$ ให้ดำเนินการโดยการแทนค่า $\beta_i = (i+1)\alpha;\ 0 \leqslant i \leqslant s$ และ $\beta_i = (n+1-i)\alpha\ ;\ i \geqslant s+1$ ลงในที่ของ β_i ในสมการ $Y_t = \sum_{i=0}^{n} \beta_i X_{t-i} + u_i$ ซึ่งจะมีผลให้ได้สมการใหม่ดังนี้

$$\begin{aligned} Y_t &= [\alpha \sum_{i=0}^{s} (i+1) X_{t-i} + \alpha \sum_{i=s+1}^{n} (n+1-i) X_{t-i}] + u_t \\ &= \alpha W_t + u_t \qquad ;\ t = n+1, n+2, \ldots, T \end{aligned}$$

โดยที่ $w_t = \sum_{i=0}^{s} (i+1)X_{t-i} + \sum_{i=s+1}^{n} (n+1-i)X_{t-i}$

จากสมการ $Y_t = \alpha w_t + u_t$ เราสามารถประมาณค่า α ได้โดยง่ายดังนี้

$$\hat{\alpha} = \Sigma w_t y_t / \Sigma w_t^2$$

$\hat{\alpha} \sim N(\alpha, \hat{\sigma}/\Sigma w_t^2)$, $\hat{\sigma}^2 = \frac{1}{(T-n)-1}[\Sigma y_t^2 - \hat{\alpha}\Sigma w_t y_t]$ และค่าประมาณของ β_i ปรากฏดังนี้

$$\hat{\beta}_i = (i+1)\hat{\alpha} \text{ , } \hat{V}(\hat{\beta}_i) = (i+1)^2 \hat{\sigma}^2/\Sigma w_t^2 \quad ; i = 0,1,2,...,s$$

และ $\hat{\beta}_i = (n+1-i)\hat{\alpha}$, $\hat{V}(\hat{\beta}_i) = (n+1-i)^2 \hat{\sigma}^2/\Sigma w_t^2 \quad ; i = s+1,s+2,...,n$

หรือสามารถเสนอค่าประมาณของ $\hat{\beta}_i$ ในทั้งสองกลุ่มร่วมกันโดยอาศัย Indicator Function $I_{[\cdot]}(i)$ ได้ดังนี้[1]

$$\hat{\beta}_i = I_{[0,s]}(i)(i+1)\hat{\alpha} + I_{[s+1,n]}(i)(n+1-i)\hat{\alpha} \text{ , } i = 0,1,2,...,n$$

และ $\hat{\beta}_i \sim N[\beta_i, (\hat{\sigma}^2/\Sigma w_t^2)(I_{[0,s]}(i)(i+1)^2 + I_{[s+1,n]}(i)(n+1-i)^2)]$

สมการประมาณค่าที่ต้องการคือ

$$Y_t = \sum_{i=0}^{n} \hat{\beta}_i X_{t-i}$$

หรือ $Y_t = \hat{\alpha}[\sum_{i=0}^{s} (i+1)X_{t-i} + \sum_{i=s+1}^{n} (n+1-i)X_{t-i}]$

---

[1] Indicator Function $I_{[\cdot]}(i)$; $i = 0,1,2,..,n$ คือฟังก์ชั่นที่มีค่าเป็น 0 หรือ 1 อย่างใดอย่างหนึ่ง กล่าวคือถ้าค่าของ i ตกอยู่ในช่วง $[\cdot]$ ที่กำหนด $I_{[\cdot]}(i)$ จะมีค่าเท่ากับ 0 ตามรูปข้างต้น ถ้า s = 8 และ i = 3 จะพบว่า i มีค่าตกอยู่ในช่วง $[0,8]$ ดังนั้น $I_{[0,8]}(3) = 1$ และ $I_{[9,16]}(3) = 0$ ดังนั้น $\hat{\beta}_3 = 4\hat{\alpha}$

11.2.4 Almon Distributed Lag

Almon distributed Lag เป็นกรณีทั่วไปของทั้ง Arithematic Lag และ Inverted - V Lag โดย Almon Distributed Lag ถือว่า β_i เป็นโพลีโนเมียลลำดับที่ q ของ i กล่าวคือ

$$\beta_i = \alpha_0+\alpha_1 i+\alpha_2 i^2 +\ldots+\alpha_q i^q \quad ; \; i = 0,1,\ldots,n$$

สำหรับการประมาณค่า β_i ให้ดำเนินการประมาณค่า α_i เสียก่อน แล้วจึงค่อยย้อนมาคำนวณหาค่าประมาณของ β_i ในภายหลังดังนี้

1. จากสมการ $\beta_i = \alpha_0+\alpha_1 i+\alpha_2 i^2 +\ldots+\alpha_q i^q \; ; \; i = 0,1,\ldots,n$ ซึ่งสามารถจัดเป็นรูปแมตริกซ์ได้ดังนี้

$$\underset{(n+1)\times 1}{\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_n \end{bmatrix}} = \underset{(n+1)\times(q+1)}{\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 1 & 1 & 1 & \ldots & 1 \\ 1 & 2 & 2^2 & \ldots & 2^q \\ 1 & 3 & 3^2 & \ldots & 3^q \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & n & n^2 & \ldots & n^q \end{bmatrix}} \underset{(q+1)\times 1}{\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_q \end{bmatrix}}$$

หรือ $\beta = H\alpha$

2. จากสมการ $Y_t = \sum_{i=0}^{n} \beta_i X_{t-i} + u_t \; ; \; t = n+1, n+2, \ldots, T$ หรือในรูปแมตริกซ์คือ $Y = X\beta + U$

แทนที่ β ในสมการ $Y = X\beta + U$ ด้วย $\beta = H\alpha$ จะได้สมการใหม่ดังนี้คือ

$$Y = XH\alpha + U$$

และเมื่อกำหนดให้ $Z = XH$ จะพบว่า Z เป็น Design Matrix ขนาด $(T-n)\times(q+1)$ และ

$$Y = Z\alpha + U$$

โดยที่

$$X = \begin{bmatrix} X_{n+1} & X_n & X_{n-1} & \cdots & X_1 \\ X_{n+2} & X_{n+1} & X_n & \cdots & X_2 \\ X_{n+3} & X_{n+2} & X_{n+1} & \cdots & X_3 \\ \vdots & \vdots & \vdots & & \vdots \\ X_T & X_{T-1} & X_{T-2} & \cdots & X_{T-n} \end{bmatrix}_{(T-n)X(n+1)}$$

3. จากสมการ $Y = Z\alpha + U$ เราสามารถประมาณค่า α โดยอาศัย OLS กล่าวคือ

$$\hat{\alpha} = (Z'Z)^{-1}Z'Y, \quad \hat{\sigma}^2 = \frac{1}{(T-n)-(q+1)}[Y'Y-\hat{\alpha}'Z'Y], \quad \hat{\alpha} \sim N[\alpha, \hat{\sigma}^2(Z'Z)^{-1}]$$

4. จากสมการ $\beta = H\alpha$ เมื่อแทนที่ α ด้วย $\hat{\alpha}$ เราย่อมคำนวณหาค่าประมาณของ β ได้โดยง่ายคือ $\hat{\beta} = H\hat{\alpha}$

และ

$$\begin{aligned} \hat{V}(\hat{\beta}) &= E(\hat{\beta}\hat{\beta}') = E[(H\hat{\alpha})(H\hat{\alpha})'] \\ &= H\,E(\hat{\alpha}\hat{\alpha}')H' = H\,\hat{V}(\hat{\alpha})\,H' \\ &= \hat{\sigma}^2\,H(Z'Z)^{-1}\,H' \end{aligned}$$

ดังนั้น

$$\hat{\beta} \sim N[\beta, \hat{\sigma}^2\,H(Z'Z)^{-1}H]$$

Almon Distribution Lag เป็นวิธีที่ยืดหยุ่นได้มากเพราะเราสามารถกำหนดดีกรีของโพลีโนเมียล $\beta_i = \alpha_0+\alpha_1 i + \alpha_2 i^2 + \ldots + \alpha_q i^q$ ได้ตามความเหมาะสม แต่อย่างไรก็ตาม ปัญหาในทางปฏิบัติที่เกิดขึ้นอยู่เสมอก็คือ

1. เราไม่อาจทราบว่าในสถานการณ์หนึ่งที่เกี่ยวข้องอยู่นั้นควรใช้โพลีโนเมียลดีกรีใด

2. เรามักไม่ทราบ Lag Length ซึ่งปัญหานี้ก็เป็นปัญหาของ Arithematic Lag และ Inverted - V Lag ด้วย

11.2.5 การประมาณค่าของ Lag Length (n)

การประมาณค่าของ Lag Length และดีกรีของโพลีโนเมียลสามารถกระทำได้หลายวิธี ในที่นี้จะเสนอเฉพาะวิธีของ พากาโน และฮาร์ตลี (Pagano, M. and Hartley, M.J., 1975)[1] ซึ่งมีวิธีการดังนี้

1. ให้ Lag Length ที่ยาวที่สุดมีค่าเท่ากับ N ดังนั้นสมการสำหรับ Finite Distributed Lag คือ

$$Y_t = \sum_{i=o}^{N} \beta_i X_{t-i} + u_t \qquad ; \ t = N+1, N+2, .., T$$

หรือ $$Y_N = X_N \beta_N + U_N$$

หมายเหตุ Subscript "N" แสดงว่าแบบจำลองนี้ใช้ Lag Length เท่ากับ N และมีค่าสังเกต Z ทั้งสิ้น T-N ชุด

2. จากสมการ $Y_N = X_N \beta_N + U_N$ อาศัย Gram - Schmidt Process[2] เพื่อจัดเวคเตอร์ (Column Vector) ในแมตริกซ์ X_N ออกเป็น Orthonormal Vector และแยกแฟคเตอร์แมตริกซ์

---

[1] Judge, G.G. et. al., p. 650

[2] อ่านมนตรี พิริยะกุล, คณิตศาสตร์สำหรับสถิติ (พีชคณิตเชิงเส้น)(สำนักพิมพ์กราฟิคอาร์ต กรุงเทพฯ, 2525), หน้า 245-8

X_N ออกเป็นผลคูณระหว่างแมตริกซ์ของ Orthonormal Vector และ Upper Triangular Matrix[1]

ให้ P_N และ R_N คือองค์ประกอบ (Factor) ของแมตริกซ์ X_N โดยที่ P_N คือแมตริกซ์ของ Orthonormal Vector และ R_N คือ Upper Triangular Matrix ขนาด $(N+1)\times(N+1)$

ดังนั้น $\quad X_N = P_N R_N$

นั่นคือ $\quad Y_N = P_N R_N \beta_N + U_N$

ให้ $\quad R_N \beta_N = \alpha_N$

$\Rightarrow \quad Y_N = P_N \alpha_N + U_N$ เมื่อ P_N คือแมตริกซ์ของ Orthonormal Vector[2] โดยอาศัย OLS จะพบว่า

---

[1] ตัวอย่างเช่น

$$A = \begin{matrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{matrix}$$

เมื่ออาศัย Gram-Scmidt Process จะได้แมตริกซ์ P ซึ่งเป็น Orthogonal Matrix (แต่ละเวคเตอร์เป็น Orthonormal Vector) ดังนี้

$$P = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix}$$

จากแมตริกซ์ P จะพบว่า $P'P = I$ ดังนี้ถ้าให้ B เป็นองค์ประกอบ (Factor) หนึ่งของ A จะพบว่า $A = PB$ ซึ่งเราสามารถแสดงให้เห็นได้โดยง่ายว่า B เป็น Upper Triangular Matrix ดังนี้

จาก $A = PB$

ดังนั้น $P'A = B$ (ทั้งนี้เพราะ $P'A = P'PB$ และ $P'P = I$)

$$B = P'A = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}$$

แสดงว่า

$$A = PB \quad \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}$$

[2] จาก $Y_N = P_N \alpha_N + U_N$ เนื่องจาก P_N คือแมตริกซ์ของ Orthonormal Vector ผลที่ตามมาก็คือ $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \ldots, \hat{\alpha}_N$ เป็นอิสระต่อกัน

$$\hat{\alpha}_N = (P_N'P_N)^{-1}P_N'Y_N = P_N'Y_N$$

3. ให้ดำเนินการทดสอบสมมุติฐาน หรือตรวจสอบนัยสำคัญของพารามิเตอร์ α ทีละตัวเรื่อยไปจนกระทั่งพบว่า α ไม่มีนัยสำคัญ กล่าวคือให้ดำเนินการตรวจสอบดูว่า $\alpha_N = 0$ หรือไม่ ถ้า $\alpha_N = 0$ ให้ตรวจสอบดูว่า $\alpha_{N-1} = 0$ หรือไม่ ถ้า $\alpha_{N-1} = 0$ ให้ตรวจสอบดูว่า $\alpha_{N-2} = 0$ หรือไม่ ดังนี้เรื่อยไปจนกระทั่งสมมุติพบว่า $\alpha_{20} \neq 0$ ก็ให้สรุปว่า Lag Length เท่ากับ 20

นั่นคือให้ดำเนินการทดสอบสมมุติฐานต่อไปนี้อย่างต่อเนื่องกันไป[1]

$$H_{0j} : \alpha_{N-j} = 0 \quad \text{vs} \quad H_{1j} : \alpha_{N-j} \neq 0 \quad ; \; j = 0,1,2,\ldots,N-1$$

โดยเราจะปฏิเสธสมมุติฐานหลัก H_{0j} ณ ระดับนัยสำคัญเมื่อ

$$F_{N-j} > F_{j+1,T-2N,1-\alpha}$$

$$F_{N-j} = \sum_{k=N-j}^{n} \frac{\hat{\alpha}_k^2}{(j+1)\hat{\sigma}^2}$$

ตัวอย่างเช่นกำหนดให้ Lag Length ไว้ยาวที่สุดเท่ากับ 15 และดำเนินการบันทึกข้อมูลได้ค่าสังเกต Z ทั้งสิ้น 40 ชุด ดังนั้นสมการถดถอยที่ต้องการคือ

$$Y_t = \sum_{i=0}^{15} \beta_i X_{t-i} + u_t \quad ; \quad t = 16,17,..,40$$

หรือ $\quad Y_{15} = X_{15}\beta_{15} + U_{15}$

โดยที่ X_{15} คือ Design Matrix ขนาด $(T-N)\times(N+1)$ ในที่นี้ X_{15} คือแมตริกซ์ขนาด 25×16 $\beta_{15} = [\beta_0,\beta_1,\ldots,\beta_{15}]$, $U_{15} = [u_{16},u_{17},\ldots,u_{40}]$ และ $Y_{15} = [Y_{16},Y_{17},\ldots,Y_{40}]$

จาก

$$Y_{15} = X_{15}\beta_{15} + U_{15}$$

---

[1] ถ้า $j = 0,1,2,\ldots,N-1$ สมมุติฐานหลักจะปรากฏดังนี้คือ $H_0 : \alpha_N = 0$, $H_0 : \alpha_{N-1} = 0$, $H_0 : \alpha_{N-2} = 0$, $H_0 : \alpha_{N-3} = 0$, ..., $H_0 : \alpha_1 = 0$

ให้แยกองค์ประกอบแมตริกซ์ X_{15} ออกเป็นผลคูณระหว่างแมตริกซ์ P_{15} ซึ่งเป็นแมตริกซ์ขนาด 25×16 ของ Orthonormal Vector กับแมตริกซ์ R_{15} ซึ่งเป็น Upper Triangular Matrix ขนาด 16×16

ดังนั้น $Y_{15} = P_{15}R_{15}\beta_{15} + U_{15} = P_{15}\alpha_{15} + U_{15}$ และโดยอาศัย OLS จะพบว่า

$$\hat{\alpha}_{15} = P_{15}' Y_{15}$$

โดยที่ $\hat{\alpha}_{15} = [\hat{\alpha}_0, \hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_{15}]'$ และ $Y_{15} = \hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1} + \hat{\alpha}_2 X_{t-2} + \ldots + \hat{\alpha}_{15} X_{t-15}$
และสมมุติฐานที่ต้องการทดสอบก็คือ

H_{00} : $\alpha_{15} = 0$ vs H_{10}: $\alpha_{15} \neq 0$ ถ้าปฏิเสธ H_0 แสดงว่า Lag Length = 15

H_{01} : $\alpha_{14} = 0$ VS H_{11}: $\alpha_{14} \neq 0$ ถ้าปฏิเสธ H_0 แสดงว่า Lag Length = 14

H_{02} : $\alpha_{13} = 0$ VS H_{12} : $\alpha_{13} \neq 0$ ถ้าปฏิเสธ H_0 แสดงว่า Lag Length = 13

$\vdots$

H_{014}: $\alpha_1 = 0$ vs H_{114}: $\alpha_1 \neq 0$ ถ้าปฏิเสธ H_0 แสดงว่า Lag Length = 1

จากรูปทั่วไปคือเราจะปฏิเสธ H_{0j} : $\alpha_{N-j} = 0$ Vs H_{1j} : $\alpha_{N-j} \neq 0$; $j = 0,1,2,\ldots,N-1$ เมื่อ

$$F_{N-j} > F_{j+1,T-2N,1-\alpha}$$

โดยที่ $$F_{N-j} = \sum_{k=N-j}^{N} \frac{\hat{\alpha}_k^2}{(j+1)\hat{\sigma}^2}$$

(1) สำหรับ H_{00} : $\alpha_{15} = 0$ **vs** H_{10} : $\alpha_{15} \neq 0$ เราจะปฏิเสธ H_0 ณ ระดับนัยสำคัญ α เมื่อ

$$F_{15} > F_{1,10,1-a} \quad \text{เมื่อ} \quad F_{15} = \frac{\ddot{\sigma}^2_{15}}{\hat{\sigma}^2}$$

เมื่อ $\hat{\sigma}^2$ คือ MSE จากสมการ $Y_t = \sum_{i=0}^{15} \beta_i X_{t-i} + u_t$; $t = N+1, N+2, \ldots, T$

(2) ถ้ายอมรับ H_{00} : $\alpha_{15} = 0$ แสดงว่า X_{t-15} ไม่ควรปรากฏอยู่ในแบบจำลอง หรือ Lag Length ควรน้อยกว่า 15 ให้ทดสอบว่า Lag Length เท่ากับ 14 หรือไม่โดยทดสอบ

$$H_{01} : \alpha_{14} = 0 \quad \text{VS} \quad H_{11} : \alpha_{14} \neq 0$$

โดยเราจะปฏิเสธ H_{01} ณ ระดับนัยสำคัญ α เมื่อ

$$F_{14} > F_{2,10,1-\alpha} \quad \text{เมื่อ} \quad F_{14} = \frac{\hat{\sigma}^2_{14}}{2\hat{\sigma}^2} + \frac{\hat{\sigma}^2_{15}}{\hat{\sigma}^2}$$

(3) ถ้ายอมรับ $H_{01} : \alpha_{14} = 0$ แสดงว่า X_{t-14} ไม่ควรปรากฏในแบบจำลอง หรือ Lag Length ควรน้อยกว่า 14 ให้ทดสอบว่า Lag Length เท่ากับ 13 หรือไม่ โดยทดสอบ

$$H_{02} : \alpha_{13} = 0 \quad \text{VS} \quad H_{12} : \alpha_{13} \neq 0$$

โดยเราจะปฏิเสธ H_{02} ณ ระดับนัยสำคัญ α เมื่อ

$$F_{13} > F_{3,10,1-\alpha} \quad \text{เมื่อ} \quad F_{13} = \frac{\hat{\sigma}^2_{13}}{3\hat{\sigma}^2} + \frac{\hat{\sigma}^2_{14}}{2\hat{\sigma}^2} + \frac{\hat{\sigma}^2_{15}}{\hat{\sigma}^2}$$

$$\vdots$$

ให้ดำเนินการทดสอบเช่นนี้เรื่อยไปตราบใดก็ตามที่ยังคงยอมรับสมมุติฐานหลักอยู่ และหยุดดำเนินการเมื่อสามารถปฏิเสธสมมุติฐานหลักได้ เช่นทดสอบ $H_{09} : \alpha_6 = 0$ VS $H_{19} : \alpha_6 \neq 0$ พบว่า $F_6 > F_{6,10,1-\alpha}$ โดยที่ $F_6 = \frac{\hat{\sigma}^2_6}{7\hat{\sigma}^2} + \frac{\hat{\sigma}^2_7}{8\hat{\sigma}^2} + \frac{\hat{\sigma}^2_8}{9\hat{\sigma}^2} + \ldots + \frac{\hat{\sigma}^2_{15}}{\hat{\sigma}^2}$ ให้หยุดดำเนินการและสรุปว่า Lag Length เท่ากับ 6 หรือ Finite Distributed Lag ที่ควรกำหนดขึ้นใช้สำหรับสถานการณ์นี้คือ $Y_t = \sum_{i=0}^{6} \beta_i X_{t-i} + u_i \; ; \; i = 7,8,9,\ldots,40$ จากนั้นจึงดำเนินการประมาณค่าสมการดังกล่าวต่อไปตามวิธีที่เห็นว่าเหมาะสม

11.2.6 สรุป

การใช้ Finite Distributed Lag Model เพื่อช่วยให้แบบจำลองมีธรรมชาติที่เคลื่อนไหวสอดคล้องกับพฤติกรรมของมนุษย์นั้นแม้ว่าจะสอดคล้องกับพฤติกรรมมนุษย์ที่มักชลอเวลาไปในช่วงหนึ่งที่จำกัด (มิใช่ชลอเวลาไปอย่างไม่มีขีดจำกัด) แต่ก็มีปัญหาคือ โดยปกติเราไม่อาจทราบ Lag Length ได้ แม้ว่าจะสามารถประมาณค่าของ Lag Length ได้ แต่ก็ยังพบว่าวิธีการดังกล่าวขาดคุณสมบัติของตัวประมาณค่าที่ดี อนึ่งการประมาณค่าพารามิเตอร์ β_i จากสมการ $Y_t = \sum_{i=0}^{n} \beta_i X_{t-i} + u_t$ เป็นเรื่องที่กระทำโดยตรงมิได้เพราะจะเกิดปัญหา Multicolinearity

ทำให้เราต้องเพิ่มเงื่อนไขเกี่ยวกับแบบแผนการแจกแจงของ β_i ไว้ ทำให้นักวิจัยประสบปัญหาเรื่องการเลือกแบบแผนให้สอดคล้องกับงานของตน เรื่องราวของ Finite Distributed Lag จึงนับว่ามีปัญหาอยู่บ้าง

อย่างไรก็ตาม ทางออกทางหนึ่งสำหรับปัญหาเรื่อง Lag Length ก็คือการใช้ Infinite Distribute Lag ซึ่งถือว่าตัวแปร X ในอดีตทุกวาระล้วนส่งผลต่อความผันแปรในค่าของ Y อย่างไม่มีขีดจำกัดว่าจะต้องทวนกาลเวลาไปนานเท่าไร วิธีนี้แม้จะฝืนความรู้สึกอยู่บ้าง แต่ก็สามารถแก้ปัญหาเรื่อง Lag Length อย่างได้ผล สำหรับการกำหนดแบบแผนการแจกแจงของ β_i ก็ยังคงต้องมีอยู่เพื่อแก้ปัญหา Multicolinearity ผู้วิจัยจำเป็นต้องใช้วิจารณญาณประกอบความรู้ในการเลือกแบบแผนดังกล่าวมาใช้เองตามความเหมาะสมเป็นราย ๆ ไป

11.3 Infinite Distributed Lag

ในกรณีของ Infinite Distributed Lag Model เราถือว่าความเปลี่ยนแปลงในค่าของตัวแปรอิสระจะส่งผลกระทบต่อความเคลื่อนไหวของ Y ตลอดไปอย่างไม่มีที่สิ้นสุด กล่าวอีกนัยหนึ่งก็คือไม่ว่าเวลาจะล่วงเลยมานานสักเพียงใดก็ตามอิทธิพลของตัวแปร X ในอดีตก็ยังคงมีอิทธิพลในการควบคุมความเคลื่อนไหวในค่าของตัวแปรตาม Y แต่ทั้งนี้อิทธิพลดังกล่าวจะค่อย ๆ ลดลงตามลำดับ กล่าวคือยิ่งย้อนเวลาไปนานเพียงใด ตัวแปร X ก็ยิ่งลดอิทธิพลลงมากเพียงนั้น แต่จะไม่สูญหายหรือขาดตอน

แบบจำลองสำหรับ Infinite Distributed Lag มีลักษณะดังนี้

$$Y_t = \sum_{i=0}^{\infty} \beta_i X_{t-i} + u_i$$

แสดงว่าในกรณีนี้นักวิจัยต้องเผชิญกับปัญหาการประมาณค่าพารามิเตอร์ β จำนวนมหาศาลและต้องเผชิญปัญหา Multicolinearity ถ้าหากจะใช้ OLS โดยตรง หรือหากจะมองในรูปของความเป็นไปได้จะพบว่าเราไม่อาจใช้ OLS ประมาณค่าเวคเตอร์ $\beta = [\beta_0, \beta_1, \beta_2, \ldots]'$ ได้เลย ทางออกสำหรับปัญหานี้ก็คือการกำหนดแบบแผนการแจกแจงของ β_i ตามความเหมาะสมซึ่งกระทำได้หลายวิธีเช่น Geometric Distribution, Gamma Distribution, Pascal Distribution, Exponential Distribution และอื่น ๆ ซึ่งในที่นี้จะกล่าวถึงเฉพาะ Geometric Lag เท่านั้น หากนักศึกษาเข้าใจเรื่องราวเกี่ยวกับ Geometric Lag ได้ Infinite Lag Pattern อื่นก็มิใช่เรื่องยาก

11.3.1 Geometric Lag หรือ Koyck's Geometric Lag Scheme

Geometric Lag ถือว่า Lag Weight β_i มีค่าที่ค่อยๆลดลงในรูปอนุกรมเรขาคณิตคือ

$$\beta_i = \alpha\lambda^i \quad ; 0 < \lambda < 1 \; ; i = 0,1,2,\ldots$$

กล่าวคือเมื่อ i = 0, 1, 2, 3, ... β_i จะมีค่าเป็นอนุกรมเรขาคณิตคือ

$$\alpha, \alpha\lambda, \alpha\lambda^2, \alpha\lambda^3, \ldots$$

เมื่อ $0 < \lambda < 1$ ก็แสดงว่า Lag Weight จะมีค่าค่อย ๆ ลดลงตามลำดับดังภาพ

A geometric lag.

ดังนั้นเมื่อแทนค่า β_i ด้วย $\alpha\lambda^i$ ลงในสมการ Infinite Distributed Lag จะได้สมการใหม่ดังนี้

$$Y_t = \sum_{i=0}^{\infty} (\alpha\lambda^i) X_{t-i} + u_t$$

$$= \alpha \sum_{i=0}^{\infty} \lambda^i X_{t-i} + u_t \qquad \ldots\ldots(1)$$

หรือ $Y_t = \alpha X_t + \alpha\lambda X_{t-1} + \alpha\lambda^2 X_{t-2} + \alpha\lambda^3 X_{t-3} + \ldots + u_t \qquad \ldots\ldots(2)$

เพื่อความสะดวกและประหยัดเวลาในการศึกษา เราจะนิยาม Lag Operator L และฟังก์ชั่นของ L คือ W(L) ดังนี้

$(L)X_t = X_{t-1}$, $(L^2)X_t = X_{t-2}$, $(L^3)X_t = X_{t-3}, \ldots$ หรือเสนอในรูปทั่วไปคือ $(L^i)X_t = X_{t-i} \quad ; i = 0,1,2,3,\ldots$

และให้ W(L) คือฟังก์ชั่นของ Lag Operator ดังนิยาม

$$W(L) = 1 + \lambda L + \lambda^2 L^2 + \lambda^3 L^3 + \ldots = (1 - \lambda L)^{-1}$$

ดังนั้นโดยอาศัยนิยามของ Lag Operator และฟังก์ชั่น W(L) เราสามารถจัดรูปสมการที่ (1) หรือ (2) ให้เป็นรูปง่ายได้ดังนี้

$$
\begin{aligned}
Y_t &= \alpha X_t + \alpha\lambda X_{t-1} + \alpha\lambda^2 X_{t-2} + \alpha\lambda^3 X_{t-3} + \ldots + u_t \\
&= \alpha(X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + \lambda^3 X_{t-3} + \ldots) + u_t \\
&= \alpha(1 + \lambda L + \lambda^2 L^2 + \lambda^3 L^3 + \ldots) X_t + u_t \\
&= \alpha W(L) X_t + u_t \\
&= \alpha\ (1-\lambda L)^{-1} X_t + u_t
\end{aligned}
$$

$$(1-\lambda L) Y_t = \alpha\ X_t + (1-\lambda L) u_t \qquad \ldots\ldots(3)$$

หรือ

$$Y_t - \lambda Y_{t-1} = \alpha X_t + (u_t - \lambda u_{t-1})$$

$$Y_t = \alpha X_t + \lambda Y_{t-1} + (u_t - \lambda u_{t-1}) \qquad \ldots\ldots(4)$$

จะเห็นได้ว่าสมการที่ (4) คือสมการลดรูปของสมการที่ (1) หรือ (2) ซึ่งกระทำได้โดยอาศัยข้อจำกัดในรูปแบบการแจกแจงของ β_i ในลักษณะของ Geometric Distribution ขอให้สังเกตว่าจำนวนพารามิเตอร์ β ในสมการที่ (1) ซึ่งมีจำนวนมากลดลงเหลือเพียง 2 ตัว คือ α และ λ แต่จากสมการที่ (4) จะพบว่าปัญหา Multicolinearity ไม่ปรากฏอีกต่อไปเพราะ X_t และ X_{t-1} ไม่เกี่ยวเนื่องกัน แต่เรากลับพบปัญหาใหม่คือปัญหา Autocorrelation[1] ขอให้สังเกตว่า Stochastic Term ในสมการที่ (1) คือ u_t ขณะที่ Stochastic Term ในสมการที่ (4) คือ $(u_t - \lambda u_{t-1})$ การประมาณค่า α และ λ จึงกระทำโดยอาศัย OLS โดยตรงมิได้ เว้นแต่จะมีข้อตกลงเป็นการเฉพาะ ด้วยเหตุนี้จึงมีเทคนิคการประมาณค่าพารามิเตอร์ α และ λ ได้มากมายหลายแบบขึ้นอยู่กับข้อตกลงเกี่ยวกับ Stochastic Term

---

[1] ให้ $v_t = (u_t - \lambda u_{t-1})$ จะพบว่า Covariance ของ v_t คือ

$$
\begin{aligned}
E(v_t v_{t-1}) &= E[(u_t - \lambda u_{t-1})(u_{t-1} - \lambda u_{t-2})] \\
&= E[u_t u_{t-1} - \lambda u_t u_{t-2} - \lambda u_{t-1}^2 + \lambda^2 u_{t-1} u_{t-2}] \\
&= 0 - 0 - \lambda\sigma^2 + 0 \\
&= -\lambda\sigma^2 \neq 0
\end{aligned}
$$

แสดงว่าสมการลดรูปมีผลให้เกิดปัญหา Autocorrelation ทั้ง ๆ ที่สมการเดิมมิได้มีปัญหานี้

11.3.1 ก เมื่อถือว่า v_t ไม่มี Autocorreletion

วิธี OLS ใช้ได้ในกรณีที่ถือว่า $(u_t - \lambda u_{t-1})$ ไม่มีปัญหา Autocorrelation ซึ่งเป็นข้อตกลงที่นำไปใช้ในกรณีของ Partial Adjustment Model[1]

จากสมการ $Y_t = \alpha X_t + \lambda Y_{t-1} + v_t$; $t = 2,3,\ldots,T$

เมื่อ $v_t = (u_t - u_{t-1})$ และสามารถจัดในรูปแมตริกซ์ได้ดังนี้

$$Y = Z\delta + V$$

โดยที่ $Y = [Y_2, Y_3, \ldots, Y_t]'$, $\delta = [\alpha, \lambda]'$, $V = [v_2, v_3, \ldots, v_T]$ และ

$$Z = \begin{bmatrix} X_2 & Y_1 \\ X_3 & Y_2 \\ \vdots & \vdots \\ X_T & Y_{T-1} \end{bmatrix}$$

ถ้า $V \sim N(0, \sigma^2 I)$ แล้ว $\hat{\delta} = (Z'Z)^{-1}Z'Y$ จะเป็น BLUE ของ δ และ $\hat{\delta} \sim N[\delta, \sigma^2(Z'Z)^{-1}]$ แต่ถ้า $(u_t - \lambda u_{t-1})$ มี Autocorrelation $\hat{\delta}$ จะเป็น Inconsistent และ Inefficient Estimator ของ δ

สำหรับกรณีที่ $(u_t - \lambda u_{t-1})$ มี Autocorrelation เราสามารถประมาณค่าพารามิเตอร์ α และ λ ด้วยวิธีการต่าง ๆ ตามข้อตกลงเกี่ยวกับ v_t และ u_t คือ

ก. v_t มี Autocorrelation เมื่อ $v_t = u_t - \lambda u_{t-1}$

ข. u_t ไม่มี Autocorrelation

ค. u_t มี Autocorrelation

---

[1] ให้ Y_t^* คือค่าที่พึงหวัง (Desired Value) ของตัวแปรสุ่ม Y_t โดยถือว่า $Y_t^* = \alpha_0 + \alpha_1 X_t + u_t$ และเชื่อว่าค่าของ Y_t จะได้รับการปรับปรุงเข้าสู่ระดับที่พึงหวังนั้นขึ้นอยู่กับอัตราการเปลี่ยนแปลงค่าของ Y_t ในวาระที่ t และ $t-1$ นั่นคือถือว่า $Y_t - Y_{t-1} = \gamma Y_t^* - Y_{t-1})$ เมื่อ $0 < \gamma < 1$

ดังนั้น Partial Adjustment Model คือ

$$Y_t = \alpha_0\gamma + \alpha_1\gamma X_t + (1-\gamma)Y_{t-1} + \gamma u_t$$

11.3.1 ข เมื่อถือว่า v_t มี Autocorrelation

ในกรณีนี้ถือว่า $v_t = u_t - \lambda u_{t-1}$ มี Autocorrelation เราสามารถประมาณค่า α และ λ ในสมการ $Y_t = \alpha X_t + \lambda Y_{t-1} + (u_t - \lambda u_{t-1})$ ได้หลายวิธี ในที่นี้จะเสนอไว้เพียง 2 วิธีคือ วิธี Search and Iteration และวิธี Wallis Three - Step

ก. Search and Iteration

จากสมการ $Y_t = \alpha X_t + \lambda Y_{t-1} + v_t$ หรือ $v_t = Y_t - \alpha X_t - \lambda Y_{t-1}$ ..... (1)

เมื่อ $v_t = (u_t - \lambda u_{t-1})$

ให้ $v_t = \rho v_{t-1} + \varepsilon_t$ โดยถือว่า $\varepsilon_t \sim N(0, \sigma^2)$ และ ρ คือสัมประสิทธิ์ Autocorrelation ..... (2)

จาก (1) เมื่อย้อนเวลาไป 1 ช่วงเวลาจะพบว่า

$$v_{t-1} = Y_{t-1} - \alpha X_{t-1} - \lambda Y_{t-2}$$

และ

$$\rho v_{t-1} = \rho Y_{t-1} - \rho\alpha X_{t-1} - \rho\lambda Y_{t-2} \quad \text{..... (3)}$$

แทนค่า v_t และ ρv_{t-1} จากสมการที่ (1) และ (3) ลงในสมการที่ (2) จะได้สมการใหม่ดังนี้

$$Y_t - \alpha X_t - \lambda Y_{t-1} = \rho Y_{t-1} - \rho\alpha X_{t-1} - \rho\lambda Y_{t-2} + \varepsilon_t \quad \text{..... (4)}$$

ซึ่งเราสามารถจัดรูปสมการที่ (4) เพื่อนำไปใช้ในกระบวนการ Search and Iteration ได้ 3 แบบตามวิธีการ Search and Iteration ที่จะเสนอไว้ 3 วิธีดังนี้ (ขอให้สังเกตว่าสมการ (ก) - (ค) คือสมการเดียวกัน

$$Y_t = \alpha X_t - \rho\alpha X_{t-1} + (\rho+\lambda) Y_{t-1} - \rho\lambda Y_{t-2} + \varepsilon_t \quad \text{..... (ก)}$$

$$(Y_t - \rho Y_{t-1}) = \alpha (X_t - \rho X_{t-1}) + \lambda (Y_{t-1} - \rho Y_{t-2}) + \varepsilon_t \quad \text{..... (ข)}$$

$$(Y_t - \alpha X_t - \lambda Y_{t-1}) = \rho (Y_{t-1} - \alpha X_{t-1} - \lambda Y_{t-2}) + \varepsilon_t \quad \text{..... (ค)}$$

วิธีที่ 1 Durbin's Two Stage Procedure

จากสมการ (ก) คือ

$$Y_t = \alpha X_t - \rho\alpha X_{t-1} + (\rho+\lambda) Y_{t-1} - \rho\lambda Y_{t-2} + \varepsilon_t \; ; \; t = 3,4,\ldots,T \quad \text{..... (ก)}$$

ให้ดำเนินการเป็น 2 ขั้นตอนดังนี้

ขั้นที่ 1 แปลงรูปพารามิเตอร์ให้เป็น Linear Parameter คือกำหนดให้

$$a = \alpha \quad , b = -\rho\alpha \quad , c = (\rho+\lambda) \quad , d = -\rho\lambda$$

สมการ (ก) จะเปลี่ยนรูปจาก Nonlinear Model เป็น Linear Model ซึ่งสามารถประมาณค่า a, b, c และ d ได้โดยอาศัย OLS ดังนี้

$$Y_t = aX_t+bX_{t-1}+cY_{t-1}+dY_{t-2}+\varepsilon_t \; ; \; t = 3,4,\ldots,T$$

จากนั้นจึงประมาณค่า a, b, c และ d ตามวิธี OLS ได้ $\hat{a}$, $\hat{b}$, $\hat{c}$ และ $\hat{d}$ และประมาณค่าของ Autocorrelation Coefficient จากความสัมพันธ์

$$\hat{\rho} = -\frac{\hat{b}}{\hat{a}}$$

ข้อสังเกต จากสมการ (ก) จะพบว่า $a = \alpha$ และ $b = -\rho\alpha$ ดังนั้น $\frac{b}{a} = \frac{-\rho\alpha}{\alpha} = -\rho$ ค่าประมาณของ ρ ในที่นี้จึงใช้ $\hat{\rho} = -\frac{\hat{b}}{\hat{a}}$ เรียกว่า Nagative Ratio

ขั้นที่ 2 แทนค่า $\hat{\rho}$ ที่ได้จากขั้นที่ 1 ในสมการ (ก) ตามเดิมได้

$$Y_t = \alpha X_t-\hat{\rho}\alpha X_{t-1}+(\hat{\rho}+\lambda)Y_{t-1}-\hat{\rho}\lambda Y_{t-2}+\varepsilon_t$$

เมื่อจัดรูปสมการเสียใหม่จะได้รูปสมการเช่นเดียวกับสมการ (ข) ดังนี้

$$(Y_t-\hat{\rho}Y_{t-1}) = \alpha(X_t-\hat{\rho}X_{t-1})+\lambda(Y_{t-1}-\hat{\rho}Y_{t-2})+\varepsilon_t$$

แล้วประมาณค่า α และ λ ตามวิธี OLS

วิธีที่ 2 จากสมการ (ข) คือ

$$(Y_t-\rho Y_{t-1}) = \alpha(X_t-\rho X_{t-1})+\lambda(Y_{t-1}-\rho Y_{t-2})+\varepsilon_t \; ; \; t = 3,4,\ldots,T$$

ซึ่งสามารถจัดเป็นรูปแมตริกซ์ $Y = X\beta + \varepsilon$ ได้โดยที่

$$Y = [Y_3-\rho Y_2, Y_4-\rho Y_3, \ldots, Y_T-\rho Y_{T-1}]' \quad , \beta = [\alpha,\lambda]'$$

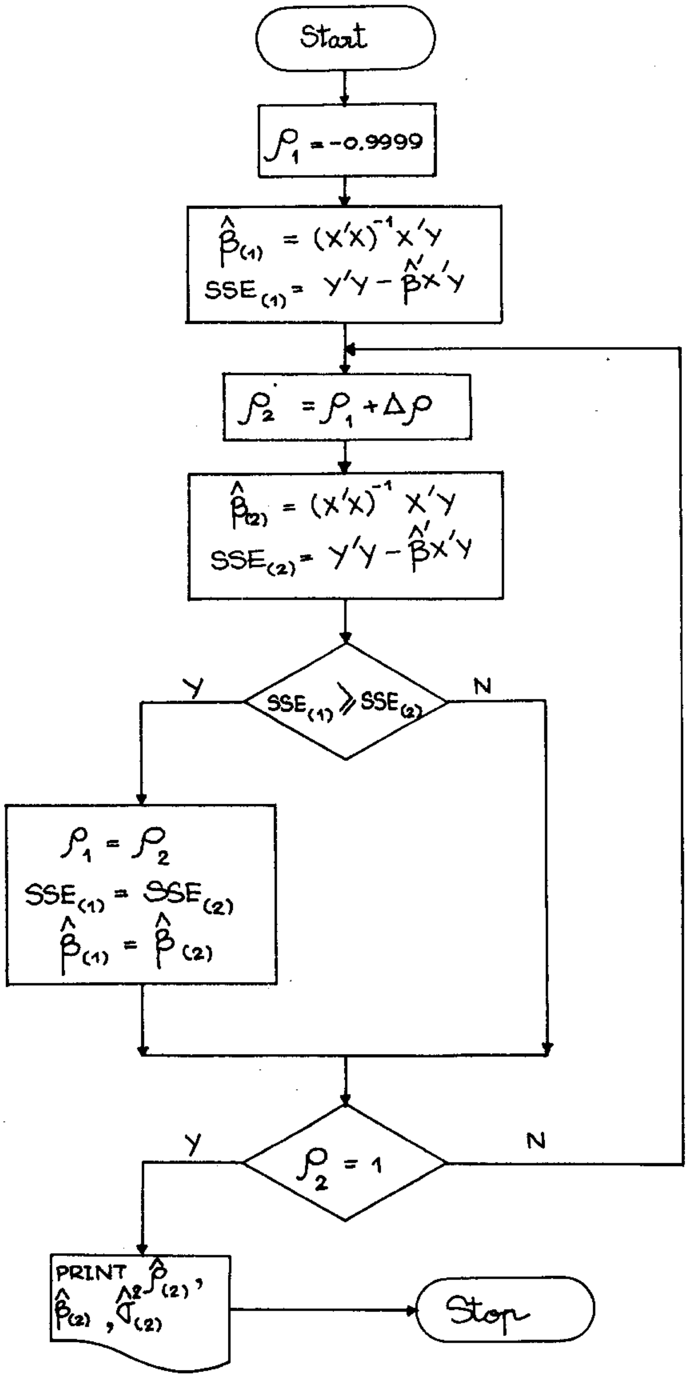
$$\varepsilon = [\varepsilon_3,\varepsilon_4,\ldots,\varepsilon_T]'$$

$$X = \begin{matrix} X_3-\rho X_2 & Y_2-\rho Y_1 \\ X_4-\rho X_3 & Y_3-\rho Y_2 \\ X_5-\rho X_4 & Y_4-\rho Y_3 \\ \vdots & \vdots \\ X_T-\rho X_{T-1} & Y_{T-1}-\rho Y_{T-2} \end{matrix}$$

โดยที่ $\hat{\beta} = (X'X)^{-1}X'Y$, $\hat{\sigma}^2 = \frac{1}{(T-2)-2}[Y'Y-\hat{\beta}'X'Y]$ หรือ $SSE = Y'Y-\hat{\beta}'X'Y$
การประมาณค่า α และ λ ให้ดำเนินการเป็นขั้นตอนตามอัลกอลิทึมต่อไปนี้

1. กำหนดให้ ρ มีค่าระหว่าง -.99 ถึง +.99 โดยเริ่มกำหนดค่า ρ ครั้งแรกที่ ρ = -.99
2. แทนค่า ρ ในสมการ $Y = X\beta$ แล้วประมาณค่า β และ σ^2 หรือ SSE
3. กำหนดค่า ρ เพิ่มขึ้นทีละน้อย แล้วประมาณค่า β และ σ^2 หรือ SSE
4. ดำเนินการขั้นที่ 3 ซ้ำ ๆ จนกระทั่งพบว่า $\hat{\sigma}^2$ หรือ SSE มีค่าต่ำที่สุด จึงหยุดดำเนินการ และ $\hat{\alpha}$ $\hat{\lambda}$ ที่สอดคล้องกับค่า ρ ที่มีผลให้ $\hat{\sigma}^2$ มีค่าต่ำที่สุดถือว่า เป็นค่าประมาณที่เหมาะสมของ α และ λ
5. การตรวจสอบนัยสำคัญของ α และ λ ให้ใช้สมการ $Y = X\beta + \varepsilon$ ที่เป็นจริงตามค่า ρ ในขั้นที่ 4

FLOW CHART สำหรับวิธีที่ 2 ปรากฏดังนี้



วิธีที่ 3 Cochrane - Orcutt Iterative Procedure

มีวิธีการประมาณค่า α และ λ โดยดำเนินการประมาณค่า (α,λ) และ ρ สลับกันไปมาจนกระทั่งค่าประมาณของ (α,λ) คงที่ในระหว่างสมการ (ข) และ (ค) คือ

$$Y_t-\rho Y_{t-1} = \alpha(X_t-\rho X_{t-1})+\lambda(Y_{t-1}-\rho Y_{t-2})+\varepsilon_t \qquad \text{.....(ข)}$$

และ $$(Y_t-\alpha X_t-\lambda Y_{t-1}) = \rho(Y_{t-1}-\alpha X_{t-1}-\lambda Y_{t-2})+\varepsilon_t \qquad \text{.....(ค)}$$

ดังนี้

1. กำหนดค่าเริ่มต้นของ ρ ขึ้นใช้โดยที่ $-1 < \rho < 1$
2. แทนค่า ρ ลงในสมการ (ข) แล้วประมาณค่า α และ λ โดยวิธี OLS ได้ $\hat{\alpha}$ และ $\hat{\lambda}$
3. แทนค่า $\hat{\alpha}$ และ $\hat{\lambda}$ จากขั้นที่ 2 ลงในสมการ (ค) แล้วประมาณค่า ρ ตามวิธี OLS ได้ $\hat{\rho}$

$$\hat{\rho} = \frac{\Sigma w_t z_t}{\Sigma w_t^2}$$

โดยที่ $w_t = W_t - \bar{W}$ เมื่อ $W_t = (Y_{t-1} - \hat{\alpha}X_{t-1} - \hat{\lambda}Y_{t-2})$ และ $z_t = Z_t - \bar{Z}$

เมื่อ $Z_t = (Y_t - \alpha X_t - \lambda Y_{t-1})$

4. แทนค่า $\hat{\rho}$ ลงในสมการ (ข) แล้วประมาณค่า α และ λ โดยวิธี OLS ได้ $\hat{\hat{\alpha}}$ และ $\hat{\hat{\lambda}}$
5. แทนค่า $\hat{\hat{\alpha}}$ และ $\hat{\hat{\lambda}}$ ลงในสมการ (ค) แล้วประมาณค่า ρ ตามวิธี OLS ได้ $\hat{\hat{\rho}}$
6. ดำเนินการแทนค่าซ้ำ ๆ สลับไปมาเช่นนี้เรื่อยไปจนกว่าค่าประมาณของ α และ λ จะมีค่าคงตัว (Convergence)

ข. Wallis Three - Step Procedure หรือ Three - Pass Least Square

วิธีการของวอลลิสที่ใช้ประมาณค่า α และ λ จากสมการ

$$Y_t = \alpha X_t+\lambda Y_{t-1}+v_t \quad ; \quad t = 2,3,\dots,T$$

ดำเนินการเป็น 3 ขั้นตอนดังนี้

1. ใช้ X_{t-1} เป็น Instrumental Variable ของ Y_{t-1} แล้วประมาณค่า α, λ และ $\hat{v}_t$ เพื่อนำ $\hat{v}_t$ ไปใช้ประมาณค่า ρ ดังนี้

ให้ $\hat{\alpha}$ และ $\hat{\lambda}$ เป็นตัวประมาณค่าของ α และ λ ทำให้ได้สมการ Basic Form ดังนี้

$$Y_t = \hat{\alpha}X_t + \hat{\lambda}Y_{t-1} \qquad \ldots\ldots(1)$$

คูณตลอดสมการ (1) ด้วย X_t แล้วรวมตลอดทุกค่าของ t ได้

$$\Sigma X_t Y_t = \hat{\alpha}\Sigma X_t^2 + \hat{\lambda}\Sigma X_t Y_{t-1} \qquad \ldots\ldots(2)$$

คูณตลอดสมการ (1) ด้วย X_{t-1} แล้วรวมตลอดทุกค่าของ t ได้

$$\Sigma X_{t-1} Y_t = \hat{\alpha}\Sigma X_{t-1} X_t + \hat{\lambda}\Sigma X_{t-1} Y_{t-1} \qquad \ldots\ldots(3)$$

จาก Normal Equation (2) และ (3) ให้ประมาณค่า α และ λ และ $\hat{v}_t$ ดังนี้

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \Sigma X_t^2 & X_t Y_{t-1} \\ \Sigma X_{t-1} X_t & X_{t-1} Y_{t-1} \end{bmatrix}^{-1} \begin{bmatrix} \Sigma X_t Y_t \\ \Sigma X_{t-1} Y_t \end{bmatrix}$$

และ $\quad \hat{v}_t = Y_t - \hat{Y}_t = Y_t - (\hat{\alpha}X_t + \hat{\lambda}Y_{t-1}) \ ; \ t = 2,3,\ldots,T$

2. ประมาณค่า ρ โดยอาศัย Residual $\hat{v}_t$ ดังนี้

$$\hat{\rho} = \frac{\sum_{t=2}^{T} \hat{v}_t \hat{v}_{t-1} \,/(T-1)}{\sum_{t=1}^{T} v_t^2 \,/\, T} + \frac{k}{T}$$

โดยที่ k = จำนวนพารามิเตอร์ ในที่นี้ $k = 2$

และแทนค่า $\hat{\rho}$ ในที่ของ ρ ในแมตริกซ์ W ได้ $\hat{W}$

$$\hat{W} = \frac{1}{1-\hat{\rho}^2}\begin{bmatrix} 1 & \hat{\rho} & \hat{\rho}^2 & \cdots & \hat{\rho}^{T-1} \\ \hat{\rho} & 1 & \hat{\rho} & \cdots & \hat{\rho}^{T-2} \\ \hat{\rho}^2 & \hat{\rho} & 1 & \cdots & \hat{\rho}^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}^{T-1} & \hat{\rho}^{T-2} & \hat{\rho}^{T-3} & \cdots & 1 \end{bmatrix}_{(T-1)X(T-1)}$$

$$\hat{W}^{-1} = \begin{bmatrix} 1 & -\hat{\rho} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\hat{\rho} & 1+\hat{\rho}^2 & -\hat{\rho} & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\hat{\rho} & 1+\hat{\rho}^2 & -\hat{\rho} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\hat{\rho} & 1+\hat{\rho}^2 & -\hat{\rho} \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\hat{\rho} & 1 \end{bmatrix}$$

3. เนื่องจาก $v_t = (u_t - \lambda u_{t-1})$ มีปัญหา Autocorrelation การประมาณ α และ λ จากสมการ

$$Y_t = \alpha X_t + \lambda Y_{t-1} + v_t \quad ; \; t = 2,3,\ldots,T$$

จึงใช้วิธี EGLS ดังนี้

$$\hat{\beta} = (X'\hat{W}^{-1}X)^{-1}X'\hat{W}^{-1}Y \quad ,\hat{\sigma}^2 = \frac{1}{(T-1)-2}\,[Y'\hat{W}^{-1}Y - \hat{\beta}'X\hat{W}^{-1}Y]$$

โดยที่ $\beta = [\alpha,\lambda]'$ $\quad ,Y = [Y_2,Y_3,\ldots,Y_T]'$

$$X = \begin{bmatrix} X_2 & Y_1 \\ X_3 & Y_2 \\ X_4 & Y_3 \\ \vdots & \vdots \\ X_T & Y_{T-1} \end{bmatrix}$$

วิธีนี้ $\hat{\alpha}$ และ $\hat{\lambda}$ เป็น Inefficient Estimator เพราะ ρ เป็นพารามิเตอร์ไม่ทราบค่า นอกจากวิธีทั้งสองนี้ วิธีหนึ่งที่มีประสิทธิภาพคือ Baysian Estimation

11.3.1 ค เมื่อถือว่า u_t ไม่มี Autocorrelation

การประมาณค่า α และ λ เมื่อถือว่า u_t ไม่มี Autocorrelation สามารถกระทำได้หลายวิธีดังนี้

ก. Cagan's Search Procedure

วิธีของคาแกนเป็นวิธี Search ที่อาศัยการแปลงรูปตัวแปรอิสระและการทดลองแปรค่าของ λ เข้าช่วย ทั้งนี้คาแกนดำเนินการจากสมการของ Geometric Lag เดิม มิได้ใช้สมการที่แปลงรูป กล่าวคือ

จากสมการ $$Y_t = \alpha \sum_{i=0}^{\infty} \lambda^i X_{t-i} + u_t \qquad \text{.....(1)}$$

ให้แปลงรูปตัวแปรอิสระเดิมสู่ตัวแปรอิสระตัวใหม่คือ

$$X_t^* = \sum_{i=0}^{\infty} \lambda^i X_{t-i} \qquad \text{.....(2)}$$

ดังนั้นสมการแปลงรูปคือ

$$Y_t = \alpha X_t^* + u_t \qquad \text{.....(3)}$$

การประมาณค่า α ในสมการที่ (3) ให้ดำเนินการดังนี้

1. กำหนดค่าให้แก่ λ จาก 0 ถึง 1 โดยเพิ่มค่าคราวละน้อยเริ่มจาก 0
2. แทนค่า λ จากขั้นที่ 1 ในสมการที่ (2) เพื่อหาค่า X_t^*
3. ประมาณค่า α ของสมการที่ (3) ด้วยวิธี OLS แล้วคำนวณหา R^2
4. ดำเนินการขั้นที่ 1-3 ซ้ำ ๆ จนกระทั่งพบค่าของ λ ที่มีผลให้ R^2 มีค่าสูงที่สุด ให้ใช้ $\hat{\lambda}$ และ $\hat{\alpha}$ ดังกล่าวเป็นค่าประมาณของ λ และ α สำหรับสมการที่ (1)

วิธีนี้มีข้อเสียที่เห็นได้ชัดคือ เราต้องอาศัยกลุ่มตัวอย่างขนาดใหญ่มาก และกระบวนการประมาณค่าค่อนข้างจะกินแรง แต่ก็มีข้อดีคือสามารถใช้ได้กับสมการเดิมโดยตรง

ข. Klein's MLE

จากสมการ $$Y_t = \alpha \sum_{i=0}^{\infty} \lambda^i X_{t-i} + u_t \qquad \ldots\ldots(1)$$

ให้หยุด (Truncate) การย้อนเวลาไว้ ณ วาระที่ t ทำให้สมการที่ (1) เปลี่ยนรูปเป็น

$$Y_t = \alpha \sum_{i=0}^{t-1} \lambda^i X_{t-i} + \alpha \sum_{i=0}^{\infty} \lambda^{t+i} X_{t-i} + u_t$$

$$= \alpha[X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + \ldots + \lambda^{t-1} X_{t-(t-1)}]$$

$$+ \alpha[\lambda^t X_t + \lambda^{t+1} X_{t-1} + \ldots] + u_t$$

หรือ $$Y_t = \alpha \sum_{i=0}^{t-1} \lambda^i X_{t-i} + \eta_0 \lambda^t + u_t \qquad \ldots\ldots(2)$$ [1]

โดยที่ $$\eta_0 = \alpha \sum_{i=0}^{\infty} \lambda^i X_{t-i}$$

พิจารณาสมการที่ (2) จะพบว่าพารามิเตอร์ที่เกี่ยวข้องคือ α, λ และ η_0 ซึ่งถ้าถือว่า $u_t \sim N(0,\sigma^2)$ เราย่อมประมาณค่า α, λ และ η_0 ได้โดยวิธี MLE เพียงแต่ $u_t = Y_t - [\alpha \sum_{i=0}^{t-1} \lambda^i X_{t-i} + \eta_0 \lambda^t]$ เป็น Nonlinear Function ซึ่งหากเรากำหนด λ โดยให้ λ แปรค่าเรื่อย ๆ จาก 0 ถึง 1 ก็จะสามารถประมาณค่า α และ λ ได้ตามต้องการ

การประมาณค่าพารามิเตอร์ α และ λ ในที่นี้สามารถกระทำได้โดยง่ายดังนี้

1. กำหนดให้ $0 < \lambda < 1$ โดยค่อย ๆ แปรค่า λ ไปคราวละน้อยเริ่มจาก 0
2. แทนค่า λ ในสมการที่ (2) แล้วประมาณค่า α และ η_0 และหา SSE โดยวิธี OLS
3. ดำเนินการขั้นที่ 1 และ 2 เรื่อยไปจนกระทั่งพบค่าประมาณของ α และ η_0 ที่มีผลให้ SSE มีค่าต่ำที่สุด $\hat{\alpha}$, $\hat{\lambda}$ และ $\hat{\eta}_0$ ที่สอดคล้องกับ SSE ที่มีค่าต่ำที่สุดดังกล่าว คือค่าประมาณของ α , λ และ η_0 ตามต้องการ

---

[1] เรียกเทอม $\eta_0 \lambda^t$ ว่า Truncation Remainder

วิธีนี้ให้ค่าประมาณที่ให้ผลตรงกันกับวิธี OLS ทุกประการอีกทั้งค่าประมาณของ α , λ และ η_0 ยังเป็น Consistent Estimator ปัญหาที่อาจพบในทางปฏิบัติก็คือนักวิจัยต้องกำหนดจุดของการหยุดย้อนเวลาเองโดยอาศัยความรอบรู้ และวิจารณญาณประกอบเช่น ถ้านักวิจัยต้องการหยุดการย้อนเวลาได้ 4 วาระสมการที่ (2) จะมีลักษณะดังนี้

$$Y_t = \alpha \sum_{i=0}^{5-1} \lambda^i X_{t-i} + \eta_0 \lambda^t + u_t \qquad ; \ t = 5,6,\ldots,T$$

$$= \alpha \sum_{i=0}^{4} \lambda^i X_{t-i} + \eta_0 \lambda^t + u_t \qquad ; \ t = 5,6,\ldots,T$$

$$\begin{bmatrix} Y_5 \\ Y_6 \\ Y_7 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} X_5+\lambda X_4+\lambda^2 X_3+\lambda^3 X_2+\lambda^4 X_1 & \lambda^5 \\ X_6+\lambda X_5+\lambda^2 X_4+\lambda^3 X_3+\lambda^4 X_2 & \lambda^6 \\ X_7+\lambda X_6+\lambda^2 X_5+\lambda^3 X_4+\lambda^4 X_3 & \lambda^7 \\ \vdots & \vdots \\ X_T+\lambda X_{T-1}+\lambda^2 X_{T-2}+\lambda^3 X_{T-3}+\lambda^4 X_{T-4} & \lambda^T \end{bmatrix} \begin{bmatrix} \alpha \\ \eta_0 \end{bmatrix} + \begin{bmatrix} u_5 \\ u_6 \\ u_7 \\ \vdots \\ u_T \end{bmatrix}$$

หรือ $\quad Y = X\beta + U$

สมการประมาณค่าที่ได้รับจากกระบวนการ Iteration คือ

$$Y_t = \hat{\alpha} \sum_{i=0}^{4} \hat{\lambda}^i X_{t-i} + \hat{\eta}_0 \hat{\lambda}^t$$

นอกเหนือจากวิธีประมาณค่าทั้งสองข้างต้นแล้วยังมีวิธีประมาณค่าแบบอื่น ๆ อีกเช่น Koyck's Two - Step Procedure และ Baysian Method ผู้สนใจสามารถติดตามศึกษาได้จากเอกสารอ้างอิง

11.3.1 ง. เมื่อถือว่า u_t มี Autocorrelation

ในกรณีที่ u_t มีปัญหา Autocorrelation เรามีวิธีประมาณค่าพารามิเตอร์ได้หลายวิธี เช่น Spectral Analysis, Two-Step Gauss - Newton Method, Dhrymes' MLE, Zellner - Geisel Iteration

และ Baysian Estimation ในที่นี้จะกล่าวถึงเฉพาะ Zellner - Geisel Interation เท่านั้น

ก. Zellner - Geisel Interative Procedure

จากสมการ $Y_t = \alpha \sum_{i=0}^{t-1} \lambda^i X_{t-i} + \eta_0 \lambda^t + u_t$ .....(1)

และกำหนดให้ u_t มี Autocorrelation ใน AR (1) คือ

$$u_t = \rho u_{t-1} + \varepsilon_t \quad ; \ 0 < \rho < 1 \qquad \text{.....(2)}$$

โดยถือเป็นข้อตกลงว่า $\varepsilon_t \sim NID(0,\sigma^2)$ ซึ่งมีผลให้ $\varepsilon_t = (u_t - \rho u_{t-1})$ ไม่มี Autocorrelation

จาก (1) ให้ย้อนเวลาไป 1 วาระแล้วคูณตลอดด้วย ρ

$$\rho Y_{t-1} = \rho\alpha \sum_{i=0}^{t-1} \lambda^i X_{t-i-1} + \rho\lambda^t \eta_{-1} + \rho u_{t-1} \qquad \text{.....(3)}$$

โดยที่ $\eta_0 = \alpha \sum_{i=0}^{\infty} \lambda^i X_{t-i}$ และ η_{-1} คือ lagged η_0 กล่าวคือ $\eta_{-1} = \alpha \sum_{i=0}^{\infty} \lambda^i X_{t-i-1}$

$$(1)-(3) = (Y_t - \rho Y_{t-1}) = \alpha \sum_{i=0}^{t-1} \lambda^i (X_{t-i} - \rho X_{t-i-1}) + \lambda^t(\eta_0 - \rho\eta_{-1}) + (u_t - u_{t-1}) \qquad \text{.....(4)}$$

ให้ $\eta_t(\rho) = (\eta_t - \rho\eta_{t-1})$ และ $\eta_0(\rho) = (\eta_0 - \rho\eta_{-1})$ แล้วกำหนดให้ $\eta_0(\rho)$ ซึ่งเป็นฟังก์ชั่นของ ρ เป็นพารามิเตอร์ที่ต้องประมาณค่า ดังนั้นสมการที่ (4) จึงเปลี่ยนรูปเป็น

$$(Y_t - \rho Y_{t-1}) = \alpha \sum_{i=0}^{t-1} \lambda^i (X_{t-i} - \rho X_{t-i-1}) + \lambda^t \eta_0(\rho) + \varepsilon_t \qquad \text{.....(5)}$$

กระบวนการประมาณค่าพารามิเตอร์ α และ $\eta_0(\rho)$ ให้อาศัยวิธี Iteration ดังนี้

1. กำหนดให้ $0 \leqslant \lambda \leqslant 1$ และ $-1 \leqslant \rho \leqslant 1$ โดยทดลองแปรค่า λ และ ρ ไปคราวละน้อย λ มีค่าเริ่มต้นที่ 0 และ ρ มีค่าเริ่มต้นที่ -1

2. แทนค่า λ และ ρ ในสมการที่ (5) แล้วประมาณค่า α และ $\eta_0(\rho)$ พร้อมทั้งคำนวณหา SSE หรือ MSE

3. ดำเนินขั้นที่ 1 และ 2 ติดต่อกันไปเรื่อย ๆ จนกระทั่งพบค่าประมาณของ α และ $\eta_0(\lambda)$ ที่มีผลให้ SSE มีค่าต่ำที่สุด $\hat{\alpha}$, $\hat{\lambda}$, $\hat{\rho}$,$\hat{\eta}_0(\hat{\rho})$ ที่สอดคล้องกับ SSE ที่มีค่าต่ำสุด คือค่าประมาณของ α , λ,ρ และ $\eta_0(\rho)$ ตามต้องการ

ตัวอย่างเช่น ถ้านักวิจัยต้องการหยุดการย้อนเวลาไปเพียง 4 วาระ สมการที่ (5) จะมีลักษณะดังนี้คือ

$$(Y_t-\rho Y_{t-1}) = \alpha\sum_{i=0}^{4}\lambda^i(X_{t-i}-\rho X_{t-i-1})+\lambda^t\eta_0(\rho)+\varepsilon_t \;;\; t = 5,6,\ldots,T$$

ซึ่งสามารถเสนอได้ในรูปแมตริกซ์ $Y^* = X^*\beta + \varepsilon$ ดังนี้

$$\begin{bmatrix} Y_5-\rho Y_4 \\ Y_6-\rho Y_5 \\ Y_7-\rho Y_6 \\ \vdots \\ Y_T- Y_{T-1} \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{4}\lambda^i(X_{5-i}-\rho X_{4-i}) & \lambda^5 \\ \sum_{i=0}^{4}\lambda^i(X_{6-i}-\rho X_{5-i}) & \lambda^6 \\ \sum_{i=0}^{4}\lambda^i(X_{7-i}-\rho X_{6-i}) & \lambda^7 \\ \vdots & \vdots \\ \sum_{i=0}^{4}\lambda^i(X_{T-i}-\rho X_{T-i-1}) & \lambda^T \end{bmatrix} \begin{bmatrix} \alpha \\ \eta_0(\rho) \end{bmatrix} + \begin{bmatrix} \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \vdots \\ \varepsilon_T \end{bmatrix}$$

เมื่อกำหนดค่าให้แก่ (λ,ρ) โดยที่ $0 \leqslant \lambda \leqslant 1$ และ $-1 \leqslant \rho \leqslant 1$ ครั้งหนึ่งให้ประมาณค่าเวคเตอร์ $\beta = [\alpha,\eta_0(\rho)]'$ และ SSE ครั้งหนึ่ง โดยที่

$$\hat{\beta} = (X^{*\prime}X^*)^{-1}X^{*\prime}Y^* \quad , \text{SSE} = Y^{*\prime}Y^* - \hat{\beta}^{*\prime}X^{*\prime}Y^*$$

เมื่อดำเนินการเช่นนี้เรื่อยไปจนพบค่าประมาณ $\hat{\beta}$ ที่มีผลให้ SSE มีค่าต่ำที่สุด เราก็จะได้ค่าประมาณของ λ, ρ, α และ $\eta_0(\rho)$ ตามต้องการ และสมการประมาณค่าที่ต้องการคือ

$$Y_t - \hat{\rho}Y_{t-1} = \hat{\alpha}\sum_{i=0}^{4}\hat{\lambda}^{i}(X_{t-i} - \hat{\rho}X_{t-i-1}) + \hat{\lambda}^{t}\hat{\eta}_0(\hat{\rho})$$

11.3.1 จ เมื่อไม่สนใจข้อตกลงเกี่ยวกับ u_t หรือ v_t

จากสมการ $Y_t = \alpha X_t + \alpha Y_{t-1} + v_t$; $t = 2,3,\ldots,T$ หากเราไม่สนใจในข้อตกลงเกี่ยวกับ v_t บางข้อโดยเฉพาะอย่างยิ่งคือข้อตกลงเกี่ยวกับ $E(v_t v_s)$ เราจะสามารถประมาณค่าพารามิเตอร์ α และ λ ได้ 2 วิธีคือ วิธี Instrumental Variable Technique และวิธี Two - Stage LS ดังนี้

ก. Instrumental Variable Technique (IV)

หลักการโดยสรุปของ IV ก็คือให้หาตัวแปรภายนอกที่เป็นอิสระกับ v_t แต่สัมพันธ์กับ Y_{t-1} มาเป็นตัวแปรเพื่อทำหน้าที่แทน Y_{t-1} ใน Normal Equation[1] สำหรับสมการข้างต้นคือ

$$Y_t = \alpha X_t + \lambda Y_{t-1} + v_t \qquad ; \; t = 2,3,\ldots,T$$

เราทราบว่า $E(v_t Y_{t-1}) \neq 0$ ดังนั้น IV ที่เหมาะสมคือ X_{t-1} เพราะ X_{t-1} โดยปกติจะเป็นอิสระกับ v_t และสัมพันธ์ในรูปตัวแปรอิสระของ Y_{t-1} วิธีปฏิบัติโดยทั่วไปก็คือให้ใช้ Lagged X ที่ใกล้ปัจจุบันมากที่สุดเป็น IV เช่นสมการข้างต้น Lagged X ที่ใกล้ปัจจุบันที่สุดคือ X_{t-1} หรือในสมการ

$$Y_t = \beta_0 X_t + \beta_1 X_{t-1} + \lambda Y_{t-1} + v_t \; ; \; t = 2,3,\ldots,T$$

IV ที่เข้าใกล้ปัจจุบันที่สุดคือ X_{t-2} ดังนี้เป็นต้น

พิจารณาสมการ $$Y_t = \alpha X_t + \lambda Y_{t-1} + v_t \qquad \ldots\ldots(1)$$

---

[1] $E(v_t Y_{t-1}) \neq 0$

สมมุติว่า $\hat{\alpha}$ และ $\hat{\lambda}$ เป็นตัวประมาณค่าที่ปราศจากอคติของ α และ λ เมื่อแทนที่ α และ λ ของสมการ (1) ด้วย $\hat{\alpha}$ และ $\hat{\lambda}$ จะได้สมการ Basis Form ดังนี้

$$Y_t = \hat{\alpha}X_t + \hat{\lambda}Y_{t-1} \qquad \ldots\ldots (2)$$

จะเห็นว่าตัวแปรอิสระในสมการ (2) คือ X_t และ Y_{t-1} ให้นำ X_t คูณสมการที่ (2) แล้วรวมตลอดในทุกค่าของ t จะได้ Normal Equation ที่ 1 และให้ X_{t-1} เป็น IV ของ Y_{t-1} แล้วนำ X_{t-1} คูณตลอดสมการที่ (2) แล้วรวมตลอดในทุกค่าของ t จะได้ Normal Equation ที่ 2 ดังนี้[1]

$$\Sigma X_tY_t = \hat{\alpha}\Sigma X_t^2 + \hat{\lambda}\Sigma X_tY_{t-1} \qquad \ldots\ldots (3)$$

$$\Sigma X_{t-1}Y_t = \hat{\alpha}\Sigma X_{t-1}X_t + \hat{\lambda}\Sigma X_{t-1}Y_{t-1} \qquad \ldots\ldots (4)$$

เมื่อจัดเป็นรูปแมตริกซ์จะพบว่า

$$\begin{bmatrix} \Sigma X_t^2 & \Sigma X_tY_{t-1} \\ \Sigma X_{t-1}\Sigma X_t & \Sigma X_{t-1}Y_{t-1} \end{bmatrix}\begin{bmatrix} \hat{\alpha} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \Sigma X_tY_t \\ \Sigma X_{t-1}Y_t \end{bmatrix}$$

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \Sigma X_t^2 & \Sigma X_tY_{t-1} \\ \Sigma X_{t-1}\Sigma X_t & \Sigma X_{t-1}Y_{t-1} \end{bmatrix}^{-1}\begin{bmatrix} \Sigma X_tY_t \\ \Sigma X_{t-1}Y_t \end{bmatrix}$$

สำหรับการหา $V(\hat{\alpha})$, $V(\hat{\lambda})$ และการทดสอบสมมุติฐานให้อาศัยหลักและวิธีของ OLS

---

[1] วิธีนี้เป็นวิธีหา Normal Equation อีกวิธีหนึ่งที่ให้ผลตรงกันกับวิธี Minimize Σe_i^2 หลักการโดยสรุปก็คือให้สร้าง Basic Form ขึ้นก่อน จากนั้นจึงนำตัวแปรอิสระทุกตัวใน Basic Form คูณสมการ Basic Form ตลอดแล้วรวมตลอดในทุกค่าของ t (หรือ i) การคูณด้วยตัวแปรอิสระตัวหนึ่งตัวใด แล้วรวมตลอดในทุกค่าของ t ครั้งหนึ่งจะได้ Normal Equation หนึ่งสมการเสมอ

ปัญหาของ IV ก็คือนักวิจัยสามารถสร้าง IV ขึ้นใช้แทน Y_{t-1} ได้มากมายหลายตัว ขอแต่เพียงให้ IV ดังกล่าวมีความสัมพันธ์กับ Y_{t-1} ก็นับว่าสามารถเป็น IV ได้ ปัญหานี้เรียกว่าปัญหา Overidentification และเพื่อลดทอนปัญหานี้ให้น้อยลงนักวิจัยพึงใช้ Lagged X ที่ใกล้ปัจจุบันที่สุดเป็น IV โดยให้พิจารณาดูตัวแปรอิสระ X ในสมการเสียก่อนว่าย้อนเวลาไปกี่วาระ ถ้าย้อนไป 1 วาระ ให้ใช้ X_{t-2} เป็น IV ของ Y_{t-1} ถ้าย้อนไป 2 วาระ ให้ใช้ X_{t-3} เป็น IV ที่เป็นเช่นนี้เพราะได้มีการศึกษาวิจัยพบว่า Lagged X ที่ใกล้ปัจจุบันกว่าจะสัมพันธ์กับ Y_{t-1} มากกว่า Lagged X ที่ไกลปัจจุบัน และโดยปกติวิธี IV จะให้ Consistent Estimator แต่ Inefficient

ข. Two - Stage Least Square Method (2LS)

จากวิธี IV เราทราบว่ามี IV เป็นจำนวนมากมายมหาศาลที่สามารถใช้แทน Y_{t-1} โดยที่ IV ดังกล่าวเป็นอิสระกับ v_t เช่นในสมการ

$$Y_t = \alpha X_t + \lambda Y_{t-1} + v_t$$

IV ที่สามารถใช้แทน Y_{t-1} ได้ก็คือ $X_{t-1}, X_{t-2}, X_{t-3}, \ldots$ การเลือกใช้เพียง X_{t-1} เป็น IV เพราะถือว่าสัมพันธ์กับ Y_{t-1} สูงที่สุด แต่ก็มิใช่ว่า X_{t-1} จะใช้เป็นตัวแทนของ Y_{t-1} ได้ครบถ้วน วิธีที่ดีกว่าใช้เพียง X_{t-1} เป็น IV ก็คือการใช้ $X_{t-1}, X_{t-2}, X_{t-3}, \ldots$ ทุกตัวร่วมกันเป็น IV แต่ทั้งนี้จะต้องมีจำนวนค่าสังเกตมากพอหรือขนาดตัวอย่าง T ต้องใหญ่พอ วิธีการดังกล่าวนำไปสู่กระบวนการประมาณค่าที่เรียกว่า 2LS ซึ่งมีขั้นตอนดังนี้

ขั้นที่ 1 วิเคราะห์สมการถดถอย $Y_{t-1} = f(X_{t-1}, X_{t-2}, X_{t-3}, \ldots)$ แล้วคำนวณหาค่าของ $\hat{Y}_{t-1}$

สมมุติว่าเราใช้ $X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}$ เป็น IV ดังนั้น สมการถดถอยที่ต้องการในขั้นที่ 1 คือ

$$Y_{t-1} = \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + u_t \ ; \ t = 5,6,7,\ldots,T$$

แล้วใช้วิธี OLS ประมาณค่า $\beta_1, \beta_2, \beta_3$ และ β_4

ดังนั้น $\hat{Y}_{t-1}$; $t = 5,6,\ldots,T$ สามารถคำนวณหาได้จากสมการประมาณค่าดังนี้

$$\hat{Y}_{t-1} = \hat{\beta}_1 X_{t-1} + \hat{\beta}_2 X_{t-2} + \hat{\beta}_3 X_{t-3} + \hat{\beta}_4 X_{t-4} \ ; \ t = 5,6,7,\ldots,T$$

ขั้นที่ 2 ใช้ $\hat{Y}_{t-1}$ เป็นตัวแปรอิสระแทน Y_{t-1} แล้ววิเคราะห์สมการถดถอย

$$Y_t = \alpha X_t + \lambda \hat{Y}_{t-1} + v_t \qquad ; \ t = 2,3,\ldots,T$$

โดยวิธี OLS

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \Sigma X_t^2 & \Sigma X_t \hat{Y}_{t-1} \\ \Sigma \hat{Y}_{t-1} X_t & \Sigma \hat{Y}_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma X_t Y_t \\ \Sigma Y_{t-1} Y_t \end{bmatrix}$$

การดำเนินการอื่น ๆ เช่น การคำนวณหา R^2, $\hat{\sigma}^2$ และการทดสอบสมมุติฐานให้ใช้วิธี OLS

วิธีปฏิบัติสำหรับขั้นที่ 1 เพื่อให้สามารถทราบจำนวนตัวแปรอิสระ $Y = f(X_{t-1}, X_{t-2}, X_{t-3}, \ldots)$ ก็คือให้ทดลองเพิ่มตัวแปรอิสระคราวละ 1 ตัวพร้อมทดสอบนัยสำคัญของ β และวัดค่า R^2 จำนวนตัวแปรอิสระที่ควรคงไว้คือตัวแปรอิสระที่มีนัยสำคัญ และให้ R^2 สูง แต่ในทางปฏิบัติเรานิยมใช้เพียง $X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}$ ไม่เกินไปกว่านี้

11.3.1 ฉ การพัฒนา Geometric Lag วิธีอื่น และ Modified Geometric Lag

ก. การพัฒนา Reduced Geometric Lag Model

จากสมการ $Y_t = \alpha X_t + \alpha\lambda X_{t-1} + \alpha\lambda^2 X_{t-2} + \alpha\lambda^3 X_{t-3} + \ldots + u_t$ .....(1)

เราสามารถลดรูปสมการ (1) ให้เข้าสู่ Reduced Form โดยไม่ต้องอาศัย Lag Operator ดังนี้

จากสมการ (1) ให้ย้อนเวลาไป 1 วาระ แล้วคูณตลอดด้วย λ

$$\lambda Y_{t-1} = \alpha\lambda X_{t-1} + \alpha\lambda^2 X_{t-2} + \alpha\lambda^3 X_{t-3} + \ldots + \lambda u_{t-1} \qquad \ldots\ldots(2)$$

นำสมการ (2) ไปหักลบออกจากสมการ (1)

$$Y_t - \lambda Y_{t-1} = \alpha X_t + u_t - \lambda u_{t-1}$$

หรือ $$Y_t = \alpha X_t + \lambda Y_{t-1} + (u_t - \lambda u_{t-1}) \qquad \ldots\ldots(3)$$

จะเห็นว่าสมการ (3) ก็คือสมการเดียวกันกับสมการ (4) ในตอน 11.3.1 ขอให้สังเกตว่าการลดรูปสมการ Geometric Lag ตามแนวนี้ง่ายกว่าการพัฒนาโดยอาศัย Lag Operator

ข. Modified Geometric Lag

ในกรณีที่เรามั่นใจว่า Initial Lag บางส่วนควรปรากฏอยู่ขณะที่ Lagged X ส่วนที่เหลือค่อย ๆ ลดอิทธิพลลงตามลักษณะของ Geometric Scheme หรือเราปรารถนาที่จะเพิ่มตัวแปรอิสระอื่นเข้าสู่ Lag Model นอกเหนือไปจากการมีเฉพาะตัวแปรอิสระ X การพัฒนา Reduced Lag Model ตามนัยนี้ผู้วิจัยสามารถพัฒนาได้โดยง่ายดังนี้

ก) ในกรณีที่มั่นใจว่า X_t, X_{t-1} และ X_{t-2} ปรากฏอยู่ส่วน Lagged X อื่น ๆ ลดอิทธิพลลงตามนัยของ Geometric Scheme

จาก $$Y_t = \beta_0 X_t+\beta_1 X_{t-1}+\beta_2 X_{t-2}+\beta_3 X_{t-3}+\ldots+u_t \qquad \ldots\ldots(1)$$

ให้ $$\beta_{i+2} = \beta_2\lambda^i \quad ; \quad 0 < \lambda < 1 \ ; \ i = 0,1,2,\ldots$$

สมการ (1) จะเปลี่ยนรูปไปดังนี้

$$Y_t = \beta_0 X_t+\beta_1 X_{t-1}+[\beta_2 X_{t-2}+\beta_2\lambda X_{t-3}+\beta_3\lambda^2 X_{t-4}+\ldots]+u_t \qquad \ldots\ldots(2)$$

ย้อนเวลาสมการ (2) ไป 1 ช่วงเวลาแล้วคูณตลอดด้วย λ

$$\lambda Y_{t-1} = \beta_0\lambda X_{t-1}+\beta_1\lambda X_{t-2}+\beta_2\lambda X_{t-3}+\beta_2\lambda^2 X_{t-4}+\ldots+\lambda u_{t-1} \qquad \ldots\ldots(3)$$

(2)-(3)

$$Y_t-\lambda Y_{t-1} = \beta_0 X_t+(\beta_1-\beta_0\lambda)X_{t-1}+(\beta_2-\beta_1\lambda)X_{t-2}+(u_t-\lambda u_{t-1}) \qquad \ldots\ldots(4)$$

$$\Rightarrow \qquad Y_t = \beta_0 X_t+\theta_1 X_{t-1}+\theta_2 X_{t-2}+\lambda Y_{t-1}+v_t \qquad \ldots\ldots(5)$$

โดยที่ $$\theta_1 = (\beta_1-\beta_0\lambda) \quad , \quad \theta_2 = (\beta_2-\beta_1\lambda)$$

สมการ (5) เรียกว่า Modified Koyck Transformation และในกรณีที่ต้องการคง Lagged X อื่น ๆไว้มากกว่า 2 ตัวก็ให้พัฒนาโดยนัยเดียวกันนี้ อนึ่งนักศึกษาสามารถพัฒนา Modified Koyck Transformation ได้โดยอาศัย Lag Operater ได้เช่นเดียวกัน

ข) ในกรณีที่มีตัวแปรอิสระมากกว่า 1 ตัวใน Koyck's Lag Model

ให้ Infinite Lag Model ประกอบด้วยตัวแปรอิสระ W_t, Z_t และ X_t ดังนี้

$$Y_t = \sum_{i-0}^{\infty} a_i W_{t-i} + \sum_{i=0}^{\infty} b_i Z_{t-i} + \sum_{i=0}^{\infty} c_i X_{t-i} + u_t \qquad \ldots\ldots (1)$$

ให้ $a_i = \alpha\lambda^1$, $b_i = \beta\lambda^1$, $c_i = \theta\lambda^1$; $0 < \lambda < 1$; $i = 0,1,2,\ldots$ $\ldots\ldots(2)$

แทนค่า a_i, b_i, c_i จากสมการ (2) ในสมการ (1)

$$\Rightarrow \quad Y_t = \alpha\sum_{i=0}^{\infty} \lambda^1 W_{t-i} + \beta \sum_{i=0}^{\infty} \lambda^i Z_{t-i} + \theta \sum_{i=0}^{\infty} \lambda^i X_{t\ i} + u_t$$

$$= \alpha(1-\lambda L)^{-1} W_t + \beta(1-\lambda L)^{-1} Z_t + \theta(1-\lambda L)^{-1} X_t + u_t$$

เมื่อ $(1-\lambda L)^{-1} = W(L) = 1+\lambda L+\lambda^2 L^2+\lambda^3 L^3+\ldots$

$$\Rightarrow \quad (1-\lambda L) Y_t = \alpha W_t + \beta Z_t + \theta X_t + (1-\lambda L) u_t$$

$$\Rightarrow \quad Y_t = \alpha W_t + \beta Z_t + \theta X_t + \lambda Y_{t-1} + (u_t - \lambda u_{t-1}) \qquad \ldots\ldots (3)$$

สมการที่ (3) คือ Modified Koyck Transformation ที่มีตัวแปรอิสระ W_t, Z_t และ X_t ขอให้สังเกตว่าสมการ (3) พัฒนาโดยอาศัย Lag Operator และถ้านักวิจัยปรารถนาจะเพิ่มตัวแปรอิสระอื่น ๆ มากกว่านี้ก็สามารถกระทำได้โดยนัยเดียวกัน