บทที่ 7

Autocorrelation

7.1 ความหมายและสาเหตุการเกิด Autocorrelation[1]

Autocorrelation คือเหตุการณ์ที่ตัวแปรสุ่ม u มีความสัมพันธ์ต่อกันหรือมี Covariation ระหว่าง u_i กับ u_j กล่าวคือ $E(u_i u_j) \neq 0, i \neq j$ ซึ่งเป็นเหตุการณ์ที่ขัดแย้งกับข้อตกลงสมการถดถอยที่กำหนดว่า $E(u_i u_j) = 0 \; ; \; i \neq j$ การคำนวณหาค่า Autocorrelation เราสามารถคำนวณหาได้จากภายในเวคเตอร์ U เองโดยคำนวณหาสหสัมพันธ์ระหว่างสมาชิกของ U ทีละคู่[2]ได้ทั้งสิ้น $\binom{n}{2} = \frac{n(n-1)}{2}$ คู่ ซึ่งเมื่อคำนวณหา $E(u_i u_j) \; ; \; i \neq j$ ครบ $\frac{n(n-1)}{2}$ คู่ ก็จะปรากฏค่าสหสัมพันธ์ซึ่งสามารถจัดวางไว้เหนือ (หรือใต้) แนวทะแยงของแมตริกซ์ V(U) ซึ่งมีผลให้ $V(U) \neq \sigma^2 I_n$ ตามข้อตกลงเดิม กล่าวคือ ถ้า $E(u_i u_j) \neq 0 \; ; \; i \neq j$ จะพบว่า

$$V(U) = \sigma^2 \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \cdots \rho_{1n} \\ \rho_{12} & 1 & \rho_{23} \cdots \rho_{2n} \\ \vdots & \vdots & \vdots \\ \rho_{1n} & \rho_{2n} & \rho_{3n} \cdots 1 \end{bmatrix} = \sigma^2 V \neq \sigma^2 I_n$$

---

[1] คำว่า Correlation หมายถึงสหสัมพันธ์ซึ่งโดยปกติจะเป็นสหสัมพันธ์ระหว่างตัวแปรสุ่ม 2 ตัว หรือตั้งแต่ 2 ตัวขึ้นไป แต่ถ้าเราคำนวณหาสหสัมพันธ์ภายในตัวแปรเดิมซึ่งกระทำโดยใช้วิธีแบ่งกลุ่มข้อมูลออกเป็น 2 ส่วน เช่น กลุ่มปัจจุบันกับกลุ่ม lag ค่าสหสัมพันธ์ที่ได้เรียกว่า Autocorrelation หรือ Intraclass Correletion หรือ Serial Correlation

[2] $U = (u_1, u_2, \ldots, u_n)'$ คือเวคเตอร์ของตัวแปรสุ่ม u_i ตัวแปรสุ่ม u_i แต่ละตัวจะมีค่าสังเกต (แต่วัดค่าไม่ได้) ของตนเอง (ขอให้นักศึกษาย้อนไปอ่านบทที่ 1-2 ซึ่งกล่าวถึงการทดลองซ้ำ ๆ ณ. ค่า $X_j = X_{jo}$ ซึ่งถ้าเราทดลองซ้ำ ๆ ทั้งสิ้น r ครั้ง ก็จะปรากฏค่าของ u ณ $X_j = X_{jo}$ รวม r ค่า

เมื่อ $\sigma^2\rho_{ij} = E(u_i u_j)$; $i \neq j$ [1]

ขอให้สังเกตว่า ถ้า $E(u_i u_j) \neq 0$; $i \neq j$ แล้ว จำนวนพารามิเตอร์ในแมตริกซ์ V(U) จะมีจำนวนมากถึง $\frac{n(n-1)}{2}$ ตัวรวมกับ σ^2 อีกตัวหนึ่งก็จะมีจำนวนพารามิเตอร์ที่ไม่ทราบค่าถึง $\frac{n(n-1)}{2}+1$ ตัว ซึ่งเป็นปัญหาสำคัญและก่อความยุ่งยากแก่นักวิจัยเป็นอย่างยิ่ง เพราะจะทำให้มีภาระที่จะต้องประมาณค่าพารามิเตอร์เหล่านี้ทั้งหมด แต่เราสามารถทำให้งานนี้ลดความยุ่งยากลงได้ด้วยการกำหนดแบบแผนของ Autocorrelation ขึ้นใช้ เช่น กำหนดให้เป็น Autoregressive Process of Order q (AR (q)) , Moving - Average - Process of Order p (MA (p)) หรือ การผสมระหว่าง AR กับ MA คือ Autoregressive Moving - Average of Order (q,p) ARMA (q,p)) ซึ่งจะทำให้จำนวนพารามิเตอร์ในแมตริกซ์ V ลดจำนวนจาก $\frac{n(n-1)}{2}$ ตัว ลงเหลือเพียง q ตัว p ตัว หรือ q+p ตัว แล้วแต่กรณีว่าจะใช้แบบแผน AR หรือ MA หรือ ARMA สำหรับการศึกษาในที่นี้ ซึ่งจะกล่าวถึงในรายละเอียดต่อไป ผู้เขียนจะใช้เพียง 1st Order Autoregressive Process (AR (1)) เท่านั้น ซึ่งในกรณีนี้นักวิจัยก็จะลดภาระของการประมาณค่าพารามิเตอร์ในแมตริกซ์ V จาก $\frac{n(n-1)}{2}$ ตัว เหลือเพียงตัวเดียวเท่านั้น

Autocorrelation เกิดขึ้นจากสาเหตุต่าง ๆ หลายสาเหตุดังนี้

1. การละเลยตัวแปรอิสระที่สำคัญ

โดยปกตินักวิจัยมีความจำเป็นต้องกำหนดตัวแปรอิสระเพื่อใช้เป็นปัจจัยในการบรรยายธรรมชาติของ Y ที่เราสนใจให้ครบถ้วนที่สุดเท่าที่จะมีความรู้ความสามารถกระทำได้ เพราะถ้าสามารถกำหนดตัวแปรอิสระได้ครบถ้วนมากเพียงใด เราย่อมทราบธรรมชาติตลอดจนพฤติกรรมในอดีต-ปัจจุบัน-อนาคตของ Y ได้มากเพียงนั้น แต่อย่างไรก็ตามแม้ว่านักวิจัยจะมีความตั้งใจดีเพียงใดความบกพร่องย่อมเกิดขึ้นได้ เพราะการจะกำหนดตัวแปรให้ได้ครบถ้วนนั้นเป็นเรื่องยาก

---

[1] จากนิยาม $\rho_{ij} = \frac{E(u_i u_j)}{\sqrt{E(u_i^2).E(u_j^2)}}$ และเมื่อเรากำหนดให้ $E(u_i^2)=\sigma^2; i=1,2,...,n$ ดังนั้น

$$E(u_i u_j)=\sigma^2\rho_{ij}$$

ถ้า $E(u_i^2) = \sigma_i^2; i=1,2,...,n$ และ $E(u_i u_j) \neq 0$; $i \neq j$ แสดงว่าเราพบทั้งปัญหา Autocorrelation และ Heteroscedasticity ทางออกของปัญหานี้คือให้ใช้ Variable Parameter Model โดยเฉพาะอย่างยิ่งคือ Time Varying parameter เช่น Cooley-Prescott Model, Rosenberg's Convergent parameter Model, Kalman Filter Model.

อาจเป็นไปได้ที่เราละเลยตัวแปรบางตัวหรือหลายตัวไปโดยไม่ตั้งใจ หรืออาจเนื่องมาจากความไม่รู้หรือคาดคิดไม่ถึง นอกจากนี้เราก็อาจตัดตัวแปรอิสระบางส่วนทิ้งด้วยเหตุจำเป็นเพราะไม่อาจวัดค่าหรือเสาะหาค่าสังเกตได้ ตัวแปรที่ถูกละเลยหรือตัดทิ้งนี้จึงแสดงบทบาทในการควบคุมความเคลื่อนไหวของ Y อยู่ภายนอกสมการ $Y = f(X's)$ โดยแอบแฝงส่งอิทธิพลผ่านตัวแปรสุ่ม u เมื่อเป็นเช่นนี้ตัวแปรสุ่ม u จึงขาดความเป็นอิสระ สหสัมพันธ์ระหว่างตัวแปร u ในต่างวาระจึงเกิดขึ้นอย่างหลีกเลี่ยงมิได้ Autocorrelation ลักษณะนี้จึงมิใช่ธรรมชาติแท้จริงของ u หรือมิใช่สหสัมพันธ์แท้ ๆ ของ u แต่เป็นอาการที่แสดงให้เห็นเสมือนว่า u มีสหสัมพันธ์เรียกว่า Quasi - Autocorrelation

2. การกำหนดความสัมพันธ์ทางคณิตศาสตร์ f ผิดพลาด

การระบุความสัมพันธ์ทางคณิตศาสตร์ระหว่างตัวแปรอิสระกับตัวแปรตามผิดจากสภาพความสัมพันธ์ที่แท้จริงจะมีผลให้ตัวแปรสุ่ม u เกิดความสัมพันธ์กันขึ้นได้ เช่น ความสัมพันธ์ที่แท้จริงของสมการ $Y = f(X's)$ คือโพลิโนเมียลดีกรี 2 แต่นักวิจัยกำหนดให้เป็นสมการเชิงเส้น ความคลาดเคลื่อนจึงเกิดขึ้น และความคลาดเคลื่อนเหล่านี้ล้วนผนวกไว้ในตัวแปรสุ่ม u ผลลัพท์ก็คือตัวแปรสุ่ม u ขาดความเป็นอิสระตามธรรมชาติเดิมและเกิดปัญหา Autocorrelation ขึ้น

3. การใช้ Interpolation Process และ Smoothing Process

ข้อมูลแหล่งที่ 2 โดยส่วนใหญ่แล้วเป็นข้อมูลที่ผู้จัดเก็บมิได้ดำเนินการจัดเก็บเองแต่มักอาศัยข้อมูลจากแหล่งต่าง ๆ มาประมวลเอาไว้เพื่อประโยชน์เฉพาะงานเฉพาะวัตถุประสงค์ เหตุที่มิได้เป็นผู้จัดเก็บเองขณะที่มีความจำเป็นต้องใช้ข้อมูล นักสถิติหรือผู้ที่เกี่ยวข้องจึงนิยมใช้วิธี Interpolation และ Extrapolation เพื่อให้ได้ข้อมูลในวาระที่สนใจ นอกจากนี้ในบางครั้งนักสถิติอาจมีความจำเป็นต้องปัดเศษหรือปรับค่าข้อมูลให้มี Significant Digit ที่เหมาะสมกับงานหรือในระดับที่พอยอมรับได้ การกระทำโดยวิธี Interpolation หรือ Extrapolation มีผลโดยตรงให้ข้อมูลรายการที่เพิ่มเข้ามาใหม่เกิดความสัมพันธ์ภายในกับข้อมูลเดิม ขณะที่การปรับข้อมูลให้กลมกลืนกันมีผลให้ข้อมูลเดิมมีความเกี่ยวเนื่องสัมพันธ์กัน ผลสะท้อนจากการที่ข้อมูลของตัวแปรมีความสัมพันธ์กันย่อมส่งผลให้ค่าของ u เกิดความสัมพันธ์กัน

4. การระบุข้อตกลงว่า $E(u_i u_j) = 0$ ทั้ง ๆ ที่ $E(u_i u_j) \neq 0$

โดยปกติแล้วตัวแปรเป็นจำนวนมิใช่น้อยมักมีธรรมชาติของความเกี่ยวเนื่องกันภายในตัวเองในระหว่างวาระอยู่แล้ว เช่นราคาสินค้าปัจจุบันสัมพันธ์ในทางบวกกับราคาในอดีต กล่าวคือ การกำหนดราคาสินค้า ผู้กำหนดราคานิยมยึดเอาราคาในอดีตเป็นฐาน ซึ่งการกระทำเช่นนี้มีผลให้ตัวแปรราคาเองมี Autocorrelation ขณะเดียวกันตัวแปรอิสระอาจมีความสัมพันธ์ภายในได้

ด้วย เหตุอื่น ๆ หลายประการซึ่งอาจเป็นไปตามธรรมชาติ หรือเพราะอุบัติการณ์ (Random Factor) บางประการเป็นตัวเร้า เช่นการนัดหยุดงานของกรรมกรมีผลให้ปริมาณการผลิต ยอดขาด สภาพคล่อง และอื่น ๆ กระทบกระเทือนเป็นลูกโซ่หลายวาระ (period) ด้วยเหตุที่ตัวแปรอิสระมีธรรมชาติของ Autocorrelation อยู่ในตัวเช่นนี้ผลสะท้อนย่อมปรากฏแก่ตัวแปรสุ่ม u คือมีผลให้ u มี Autocorrelation อยู่แล้วตามธรรมชาติ (True Autocorrelation) การควบคุมให้ $E(u_i u_j)=0$; $i \neq j$ จึงเป็นการบังคับให้ u ฝืนธรรมชาติของตนเอง

7.2 การกำหนดแบบแผนของ Autocorrelation

ดังที่ได้กล่าวมาแล้วในตอนต้นว่า ถ้า $E(u_i u_j) = \sigma^2 \rho_{ij}$; $i \neq j$ ปัญหาของนักวิจัยก็คือการมีภาระในการประมาณค่าพารามิเตอร์เพิ่มขึ้น และโดยปกติแล้วในงานทั่วไปที่มิใช่งานทดลองวิทยาศาสตร์ที่นักวิจัยสามารถควบคุมจำนวนซ้ำ (Replication) เพื่อให้ได้ค่าประมาณของ u_i ; $i=1,2,\ldots,n$ หลาย ๆ ค่าแล้ว การคำนวณหาค่า ρ_{ij} นับว่าเป็นสิ่งที่เป็นไปได้ยาก ทางออกของปัญหานี้ก็คือการกำหนดรูปแบบความสัมพันธ์ระหว่างตัวแปรสุ่ม u_i กับ u_j ขึ้นในลักษณะของ Regression Model ที่พารามิเตอร์ของ Regression Model สามารถใช้เป็นดัชนีวัด Covariation ระหว่าง u_i กับ u_j ได้ และ เพื่อความสะดวกและง่ายแก่การเชื่อมโยงถึงเรื่อง Distributed lag ผู้เขียนจะใช้ Subscript "t" แทน "i" หรือ "j" ดังนี้

1. AR (q) กำหนดให้ u_t มีความสัมพันธ์กันดังนี้

$$u_t = \Theta_1 u_{t-1} + \Theta_2 u_{t-2} + \ldots + \Theta_q u_{t-q} + v_t$$

สำหรับการศึกษาในที่นี้เราจะศึกษาเฉพาะ AR (1) เท่านั้น เพราะ AR (1) เป็นสมการที่อาศัยข้อมูลรายปีทั้งยังมีวิธีคำนวณที่ง่ายที่สุด นักศึกษาหรือนักวิจัยที่มีความสนใจ AR (2), AR (3), ... หรือ MA รวมถึง ARMA ให้ศึกษาได้จากเอกสารอ้างอิงท้ายเล่ม

สำหรับ AR (1) เรากำหนดให้ u_t มีความสัมพันธ์กันดังนี้

$$u_t = \rho u_{t-1} + v_t \quad \text{โดยที่} \quad |\rho|<1 \quad \text{หรือนัยหนึ่ง} \quad -1<\rho<1$$

2. MA (p) กำหนดให้ u_t มีความสัมพันธ์กันดังนี้

$$u_t = v_t + \alpha_1 v_{t-1} + \alpha_2 v_{t-2} + \ldots + \alpha_p v_{t-p}$$

3. ARMA (q, p) กำหนดให้ u_t มีความสัมพันธ์กันดังนี้

$$u_t = \theta_1 u_{t-1} + \theta_2 u_{t-2} + \ldots + \theta_q u_{t-q} + v_t + \alpha_1 v_{t-1} + \alpha_2 v_{t-2} + \ldots + a_p v_{t-p}$$

ทั้งนี้ถือเป็นข้อตกลงว่า $E(v_t)=0,\ E(v_t^2) = \sigma_v^2,\ E(v_t v_s)=0;\ t \neq s$[1] ขณะที่ θ และ α คือพารามิเตอร์ที่ไม่ทราบค่า

พิจารณา AR (1) คือ $u_t = \rho u_{t-1} + v_t$ จะเห็นว่าเราได้กำหนดรูปแบบความสัมพันธ์ของ u ให้แคบเข้าโดยศึกษาถึงเฉพาะกรณีของตัวแปรสุ่ม u คู่ใด ๆ ที่อยู่ติดต่อกันในเวคเตอร์ $U = (u_1, u_2, \ldots, u_n)'$ เท่านั้น นอกจากนี้ถ้าพิจารณา AR (1) ให้ดีจะพบว่า AR (1) คือสมการถดถอยที่ไม่มีเทอมคงที่ สัมประสิทธิ์ความถดถอยก็คือ ρ ซึ่งหมายถึงสหสัมพันธ์ระหว่าง u_t กับ u_{t-1} ไปในตัว[2] นอกจากนี้แมตริกซ์ V ซึ่งประกอบด้วย พารามิเตอร์ถึง $\frac{n(n-1)}{2}$ ตัวก็จะลดจำนวนพารามิเตอร์ที่ต้องประมาณค่าให้เหลือเพียงตัวเดียวคือ ρ

---

[1] เราถือเอาข้อตกลงของ Stationary เป็นเกณฑ์ หมายความว่า $E(v_t^2)$ คงที่เท่ากับ σ_v^2 เสมอ โดยที่ σ_v^2 จะไม่แปรเปลี่ยนไป กาลเวลา t ขณะเดียวกัน Covariation ของ u_t กับ u_s ก็เป็นเพียงฟังก์ชันของระยะห่างระหว่างวาระ มิได้ผันแปรไปตามกาลเวลา t (ดูการพิสูจน์และใช้ประโยชน์ของ Stationary Assumption เหล่านี้ในตอน 7.2.2)

[2] ความจริงแล้วสัมประสิทธิ์ความถดถอย $\hat{\rho}$ จากสมการ $u_t = \rho u_{t-1} + v_t$ มิใช่ดัชนีวัดค่าสหสัมพันธ์ได้ 100% แต่ถ้า $n \to \infty$ แล้ว $\hat{\rho}$ จะวัดสหสัมพันธ์ระหว่าง u_t กับ u_{t-1} ได้ เพื่อประโยชน์ในการพิสูจน์และป้องกันความสับสน ผู้เขียนจะใช้สมการ $u_t = \lambda u_{t-1} + v_t$ สำหรับ AR (1) และใช้ ρ ในความหมายของสหสัมพันธ์เช่นเดิมดังนี้

1. $\hat{\rho} = \sum_{t=2}^{n} u_t u_{t-1} \Big/ \sqrt{\sum_{t=1}^{n} u_t^2 \sum_{t=2}^{n} u_{t-1}^2} \cong \sum_{t=2}^{n} u_t u_{t-1} \Big/ \sum_{t=2}^{n} u_{t-1}^2$ เมื่อ $n \to \infty$

2. จากสมการ $u_t = \lambda u_{t-1} + v_t$ และสมมุติว่า u_t มีค่าที่วัดหรือสังเกตได้จะพบว่า

$\hat{\lambda} = \sum_{t=2}^{n} u_{t-1} u_t \Big/ \sum_{t=2}^{n} u_{t-1}^2$ จึงเห็นได้ว่า $\hat{\lambda} \to \hat{\rho}$ เมื่อ $n \to \infty$

7.2.1 AR (1) และ Markov Process

จาก AR (1) คือ $u_t = \rho u_{t-1}+v_t$ โดยที่ $|\rho|<1$ เราสามารถกำหนดโครงสร้างของ u_t ให้เป็นฟังก์ชั่นของ v_t โดยอาศัยการแทนค่าซ้ำ ๆ เนื่องจากอิทธิพลของ u มีต่อกันและกันในลักษณะผลกระทบแบบลูกโซ่ตามวิธีของ Markov ดังนี้

จากสมการ $\quad u_t = \rho u_{t-1}+v_t$

ถ้าเรา lag u_t ไปคราวละ 1 วาระจะพบว่า[1]

$$u_{t-1} = \rho u_{t-2}+v_{t-1} \qquad \ldots\ldots(1)$$

$$u_{t-2} = \rho u_{t-3}+v_{t-2} \qquad \ldots\ldots(2)$$

$$u_{t-3} = \rho u_{t-4}+v_{t-3} \qquad \ldots\ldots(3)$$

$$\vdots$$

$$u_{t-r} = \rho u_{t-(r+1)}+v_{t-r} \qquad \ldots\ldots(r)$$

ดังนั้นจากสมการ $u_t = \rho u_{t-1}+v_t$ แทนที่ u_{t-1} ด้วยสมการ (1) จะพบว่า

$$u_t = \rho(\rho u_{t-2}+v_{t-1})+v_t = \rho^2 u_{t-2}+(\rho v_{t-1}+v_t)$$

แทนที่ u_{t-2} ด้วยสมการ (2)

$$u_t = \rho^2(\rho u_{t-3}+v_{t-2})+(\rho v_{t-1}+v_t)$$

$$u_t = \rho^3 u_{t-3}+(\rho^2 v_{t-2}+\rho v_{t-1}+v_t)$$

แทนที่ u_{t-3} ด้วยสมการ (3)

$$u_t = \rho^3(\rho u_{t-4}+v_{t-3})+\rho^2 v_{t-1}+\rho v_{t-1}+v_t$$

$$u_t = \rho^4 u_{t-4}+(\rho^3 v_{t-3}+\rho^2 v_{t-1}+\rho v_{t-1}+v_t)$$

---

[1] คำว่า lag หมายถึงย้อนกาลเวลาไปในอดีต

แทนที่ u_{t-r} ด้วยสมการ (r)

$$u_t = \rho^{r+1}u_{t-(r+1)} + (v_t+\rho v_{t-1}+\rho^2 v_{t-2}+\rho^3 v_{t-3}+\ldots+\rho^r v_{t-r})$$

เนื่องจาก $-1<\rho<1$ และ $r\to\infty$ จะมีผลให้ $\rho^r \to 0$

ดังนั้น $u_t \cong v_t+\rho v_{t-1}+\rho^2 v_{t-2}+\rho^3 v_{t-3}+\ldots = \sum_{r=0}^{\infty} \rho^r v_{t-r}$

นั่นคือ โดยอาศัย Markov Process ตัวแปรสุ่ม u_t จากสมการ $u_t = \rho u_{t-1}+v_t$ สามารถเสนอในรูปฟังก์ชันของ laged value ของ v_t ได้เป็น $u_t \cong \sum_{r=0}^{\infty} \rho^r v_{t-r}$ ขอให้สังเกตว่า $u_t = \sum_{r=0}^{\infty} \rho^r v_{t-r} = v_t+\rho v_{t-1}+\rho^r v_{t-2}+\ldots$ นั้น u_t จะได้รับอิทธิพลจาก Residual u_t ในวาระปัจจุบันหรือใกล้ปัจจุบันมากที่สุด u_t ในวาระที่อยู่ในอดีตห่างไกลปัจจุบัน จะมีอิทธิพลลดลง และยิ่ง ρ มีค่าต่ำเพียงใด u_t ในอดีตย่อมมีอิทธิพลต่อ u_t น้อยลงมากเพียงนั้น[1]

7.2.2 คุณลักษณะทางประชากรของ u_t เมื่อเกิด Autocorrelation ใน AR (1)

ก. ค่าคาดหมายของ u_t

จาก $u_t \cong \sum_{r=0}^{\infty} \rho^r v_{t-r} = v_t+\rho v_{t-1}+\rho^2 v_{t-2}+\ldots\ldots$

จะพบว่า $E(u_t) = E(v_t)+\rho E(v_{t-1})+\rho^2 E(v_{t-2})+\ldots$

แต่ $E(v_t)=0;\ t=1,2,\ldots,n$

ดังนั้น $E(u_t)=0;\ t=1,2,\ldots,n$

---

[1] เนื่องจากสัมประสิทธิ์ของ v_{t-r} คือ ρ^r ; $r=0,1,2,\ldots$ ถ้า ρ มีค่าต่ำเพียงใด ρ^r ก็จะวิ่งเข้าสู่ 0 รวดเร็วมากเพียงนั้น เช่น $\rho = .10$ จะพบว่า $\rho^0=1.0, \rho^2=.01, \rho^3=.001, \rho^4=.0001$ ซึ่งแสดงให้เห็นว่า $v_{t-1}, v_{t-2}, v_{t-3}, v_{t-4}$ มีอิทธิพลต่อ u_t ลดน้อยลงตามลำดับเมื่อย้อนทวนอดีตให้ไกลปัจจุบันมาก ๆ

แสดงว่าแม้จะเกิดปัญหา Autocorrelation ใน AR (1) u_t ก็ยังคงมีคุณสมบัติสอดคล้องกับข้อตกลงเดิมว่า $E(u_t)=0$ ซึ่งแสดงว่า Regression Plane (หรือ Hyperplane) ยังคงพุ่งผ่านค่าเฉลี่ยของ Y เช่นเดิม และมีผลสะทอนให้ OLS - Estimator ของ β คือ $\hat{\beta} = (X'X)^{-1}X'Y$ ยังคงเป็นตัวประมาณค่าที่ปราศจากอคติของ β ซึ่งจะได้แสดงให้เห็นในตอนต่อไป

ข. ความแปรปรวนของ u_t

จาก $$u_t = \sum_{r=0}^{\infty} \rho^r v_{t-r}$$

จะพบว่า

$$E(u_t^2) = E\{\sum_{r=0}^{\infty} \rho^r v_{t-r}\}^2 = E\{v_t+\rho v_{t-1}+\rho^2 v_{t-2}+\ldots\}^2$$

$$= E[\{v_t^2+\rho^2 v_{t-1}^2+\rho^4 v_{t-2}^2+\ldots\}+\{2\rho v_t v_{t-1}+2\rho^2 v_t v_{t-2}+\ldots\}]$$

แต่ $E(v_t^2) = \sigma_v^2$ และ $E(v_t v_{t-s})=0 \ ; \ t\neq s \ ; \ t=1,2,\ldots,n$

ดังนั้น $$E(u_t^2) = (\sigma_v^2+\rho^2\sigma_v^2+\rho^4\sigma_v^2+\rho^6\sigma_v^2+\ldots)+(0+0+0+\ldots+0)$$

$$= \sigma_v^2(1+\rho^2+\rho^4+\rho^6+\ldots)$$

โดยอาศัยความรู้เรื่อง Geometric Progression $a, ar, ar^2, ar^3, \ldots$ เราทราบว่า $S_n=\frac{a(1-r^n)}{1-r}$ และถ้า $n\to\infty$ และ $|r|<1$ เราทราบว่า $S = \frac{a}{1-r}$

ดังนั้น $$1+\rho^2+\rho^4+\rho^6+\ldots = \frac{1}{1-\rho^2}$$

นั่นคือ $E(u_t^2) = \sigma_v^2(1+\rho^2+\rho^4+\rho^6+\ldots) = \sigma_v^2\frac{1}{1-\rho^2}$ หรือนัยหนึ่ง $V(u_t)=\frac{\sigma_v^2}{1-\rho^2}$

แสดงว่า แม้ u_t จะมีปัญหาของ Autocorrelation ใน AR (1) แต่ก็ไม่เกิดปัญหา Heteroscedasticity เพราะ $V(u_t)$ มีค่าคงที่เท่ากับ $\frac{\sigma_v^2}{1-\rho^2}$ เสมอในทุกค่าของ $t=1,2,\ldots,n$

ค. Covariation ของ u_t

ในที่นี้จะพิสูจน์ให้เห็นเฉพาะ $E(u_tu_{t-1})$ และ $E(u_tu_{t-2})$ แล้วค่อยขยายความสู่กรณีทั่วไปคือ $E(u_tu_{t-s})$; $t \neq s$

(1) $E(u_tu_{t-1})$

$$= E[\{v_t+(\rho v_{t-1}+\rho^2v_{t-2}+\rho^3v_{t-3}+\ldots)\}(v_{t-1}+\rho v_{t-2}+\rho^2v_{t-3}+\rho^3v_{t-4}+\ldots)]$$

$$= E[v_t(v_{t-1}+\rho v_{t-2}+\rho^2v_{t-3}+\ldots)+\rho(v_{t-1}+\rho v_{t-2}+\rho^2v_{t-3}+\ldots)^2]$$

$$= (0+0+0+\ldots+0) +\rho E(v_{t-1}+\rho v_{t-2}+\rho^2v_{t-3}+\ldots)^2$$

$$= \rho E[(v_{t-1}^2+\rho^2v_{t-2}^2+\rho^4v_{t-3}^2+\ldots)+(\rho v_{t-1}v_{t-2}+\rho^2v_{t-2}v_{t-3}+\ldots)]$$

$$= \rho[(\sigma_v^2+\rho^2\sigma_v^2+\rho^2\sigma_v^2+\ldots)+(0+0+\ldots+0)]$$

$$= \rho\sigma_v^2(1+\rho^2+\rho^4+\ldots)$$

$$= \rho\ \sigma_v^2\ \frac{1}{1-\rho^2}$$

$$= \rho\frac{\sigma_v^2}{1-\rho^2} \neq 0$$

$$= \rho\ \sigma_u^2$$ เมื่อ σ_u^2 คือ $V(u_t)$

(2) $E(u_tu_{t-2}) = E[(v_t+\rho v_{t-1}+\rho^2v_{t-2}+\ldots)(v_{t-2}+\rho v_{t-3}+\rho^2v_{t-4}+\ldots)]$

$$= E[v_t(v_{t-2}+\rho v_{t-3}+\rho^2v_{t-4}+\ldots)+\rho v_{t-1}(v_{t-2}+ \rho v_{t-3}+\rho^2v_{t-4}+\ldots)$$
$$+\rho^2(v_{t-2}+\rho v_{t-3}+\rho^2v_{t-4}+\ldots)^2]$$

$$= (0+0+\ldots+0)+(0+0+\ldots0)+\rho^2E(v_{t-2}+\rho v_{t-3}+\rho^2v_{t-4}+\ldots)^2$$

$$= \rho^2E[(v_{t-2}^2+\rho^2v_{t-3}^2+\rho^4v_{t-4}^2+\ldots)+(\rho v_{t-2}v_{t-3}+\rho^2v_{t-2}v_{t-4}+\ldots)]$$

$$= \rho^2[(\sigma_v^2+\rho^2\sigma_v^2+\rho^4\sigma_v^2+\ldots)+(0+0+\ldots+0)]$$

และ $W^{-1} = (1-\rho^2)V^{-1}$

ถ้าให้ $V(U) = \Phi = \sigma_u^2 V$ จะพบว่า

$$\Phi^{-1} = \frac{1}{\sigma_u^2} V^{-1}$$

$$= \frac{1}{\sigma_u^2} \cdot \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}$$

แต่ $\sigma_u^2 = \dfrac{\sigma_v^2}{1-\rho^2}$

ดังนั้น

$$\Phi^{-1} = \frac{1}{\sigma_v^2} \begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho^2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}$$

ความจริงเหล่านี้ยืนยันให้เห็นว่าเราไม่อาจใช้ OLS ประมาณค่า β ได้อย่างเหมาะสม วิธีประมาณค่าที่เหมาะสมกว่า OLS คือ GLS หรือ EGLS แล้วแต่กรณี

7.3 ผลกระทบของ Autocorrelation ต่อ OLS

เมื่อเกิดปัญหา Autocorrelation ขึ้นคือ $V(U) = \sigma_u^2 V = \dfrac{\sigma_v^2}{1-\rho^2} V \neq \sigma_u^2 I_n$ แต่ผู้วิจัยยังคงใช้ OLS ในการประมาณค่า β อยู่เช่นเดิม จะก่อให้เกิดผลเสียหายโดยเฉพาะในแง่ของคุณภาพของ $\hat{\beta}$ ดังนี้

1. $\hat{\beta}$ ยังเป็นตัวประมาณค่าที่ปราศจากอคติของ β แต่โดยทั่วไปจะมีประสิทธิภาพ (Efficiency) ต่ำกว่าวิธีอื่น ๆ ความจริงข้อนี้พิสูจน์ได้ดังนี้

$$E(\hat{\beta}) = E[(X'X)^{-1}X'Y] = E[(X'X)^{-1}X'(X\beta+U)] = \beta+(X'X)^{-1}X'E(U)$$

แต่เราได้พิสูจน์แล้วว่า $E(u_t) = 0$; $t=1,2,\ldots,n$ หรือ $E(U) = 0$

ดังนั้น $E(\hat{\beta}) = \beta$

2. ถ้าเกิดปัญหา Autocorrelation ขึ้นโดยเฉพาะอย่างยิ่งคือ Positive Autocorrelation[1] ค่าประมาณของ σ^2 ซึ่งประมาณขึ้นโดยอาศัย Residual คือ

$$\hat{\sigma}^2 = \frac{1}{n-k} \Sigma e_i^2$$

จะมีค่าต่ำกว่าความเป็นจริง ทั้งนี้เพราะในกรณีเช่นนี้ค่าจริงของ u จะจับกลุ่มกันห่างไกลจาก True Line หรือ Ture Hyperplane ขณะที่ค่าประมาณของ u คือ e จะจับกลุ่มกันใกล้ Estimated Line หรือ Estimated Hyperplane

3. เมื่อเกิดปัญหา Autocorrelation ขึ้น ค่าประมาณของ $V(\hat{\beta})$ คือ $\hat{V}(\hat{\beta})$ จะต่ำกว่าความเป็นจริงโดยเฉพาะอย่างยิ่งคือเมื่อเราประมาณค่า β โดยอาศัยวิธี OLS กล่าวคือ

$$\hat{V}(\hat{\beta}) = \hat{\sigma}^2(X'X)^{-1}$$

เหตุที่เป็นเช่นนี้เพราะ $\hat{\sigma}^2$ มีค่าต่ำกว่าความเป็นจริงประการหนึ่งและในกรณีของ Autocorrelation นั้น $E(u_i u_j) \neq 0$ แสดงว่า $Cov(u_i, u_j)$ จะต้องมีบทบาทในการกำหนดค่า $\hat{V}(\hat{\beta})$ ด้วย[2] ถ้าเราตัด $Cov(u_i, u_j)$ ทิ้งไปโดยถือเสมือน $Cov(u_i, u_j)=0$ ผลที่ตามมาก็คือ $\hat{V}(\hat{\beta})$ มีค่าต่ำ

---

[1] โดยทั่วไปมักเกิดปัญหา Positive Autocorrelation มากกว่า Negative Autocorrelation วิธีที่ง่ายที่สุดที่จะให้คำตอบว่าเป็น Positive Autocorrelation หรือ Negative Autocorrelation คือได้ตรวจดูเครื่องหมายของ e_i ; $i = 1, 2, \ldots, n$ ถ้า e_i มีเครื่องหมาย + และ - สลับกันแสดงว่าเป็น Negative Autocorrelation ถ้า e_i มีเครื่องหมายบวกต่อเนื่องกันไปแล้วติดตามด้วยเครื่องหมายลบต่อเนื่องกันไป หรือนัยหนึ่งคือ e_i เปลี่ยนเครื่องหมายช้ากว่า แสดงว่าเกิด Positive Autocorrelation

[2] อ่าน Koutsoyiannis, A., Op. cit., p.205

กว่าความเป็นจริง ค่า $\hat{V}(\hat{\beta})$ ที่ต่ำ ๆ ในกรณีนี้จึงเป็นเลขลวงตา หากนักวิจัยไม่ระมัดระวังและมีความเข้าใจในเรื่องเหล่านี้อาจเข้าใจผิดคิดว่างานของตนมีคุณภาพสูงทั้ง ๆ ที่ความจริงมิได้เป็นเช่นนั้น

4. เมื่อเกิดปัญหา Autocorrelation ขึ้น OLS - Estimator จะมีประสิทธิภาพต่ำ เพราะ $\hat{V}(\hat{\beta})$ ตามวิธี OLS จะมีค่าสูงกว่า $\hat{V}(\hat{\beta})$ ที่ประมาณได้โดยอาศัยวิธีอื่นที่เหมาะสมกว่า

7.4 การประมาณค่าเวคเตอร์ β เมื่อมี Autocorrelation ใน AR (1)

การประมาณค่าเวคเตอร์ β เมื่อมี Autocorrelation ใน AR (1) สามารถกระทำได้หลายวิธีดังนี้

7.4.1 GLS

จากแบบจำลอง $Y = X\beta+U$ เมื่อ $E(UU') = \sigma_u^2 V = \sigma_v^2 (\frac{1}{1-\rho^2} V) = \sigma_v^2 W \neq \sigma_u^2 \; I_n$

$$W = \frac{1}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}_{n \times n}$$

และ

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

ถ้าทราบค่าของ ρ ได้จาก Priori Information เราย่อมประมาณค่าเวคเตอร์ β ได้โดยง่ายดังนี้

$$\hat{\beta} = (X'W^{-1}X)^{-1}X'W^{-1}Y \ , \quad \hat{\sigma}^2 = \frac{1}{n-k}(Y'W^{-1}Y-\hat{\beta}'X'W^{-1}Y)$$

หรือจะอาศัยเทคนิคการแปลงรูปโดยใช้แมตริกซ์ P ที่ $P'P = W^{-1}$ อันจะมีผลให้
$V(PU) = E(PUU'P') = \sigma_v^2 In$ ก็ได้ แมตริกซ์ P สำหรับกรณีนี้คือ

$$P = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & -\rho & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix}_{nxn}$$

โดยที่ $P'P = (1-\rho^2)V^{-1} = \begin{bmatrix} 1 & -\rho & 0 & \cdots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \cdots & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} = W^{-1}$

ดังนั้น โดยอาศัยเทคนิคการแปลงรูปเราสามารถนำแมตริกซ์ P คูณด้านหน้าสมการ $Y = X\beta+U$ จะมีผลให้

$$PY = (PX)\beta + PU$$

ซึ่งในกรณีนี้จะพบว่า $V(PU) = \sigma_v^2 In$ โดยไม่ขัดกับข้อตกลงเดิมของสมการถดถอยแต่ประการใด

จากสมการ $PY = (PX)\beta+PU$ เราสามารถประมาณค่าเวคเตอร์ β ได้ดังนี้

$$\hat{\beta} = [(PX)'(PX)]^{-1}[(PX)'(PY)]=[X'(P'P)X]^{-1}[X'(P'P)Y]$$

$$= [(1-\rho^2)X'V^{-1}X]^{-1}[(1-\rho^2)X'V^{-1}Y]$$

$$= (X'W^{-1}X)^{-1}X'W^{-1}Y$$

$$\hat{\sigma}^2 = \frac{1}{n-k}[(PY)'(PY)-\hat{\beta}'(PX)'(PY)]$$

$$= \frac{1}{n-k}[(1-\rho^2)Y'V^{-1}Y-(1-\rho^2)\hat{\beta}'X'V^{-1}Y]$$

$$= \frac{1}{n-k}(Y'W^{-1}Y-\hat{\beta}'XW^{-1}Y)$$

อย่างไรก็ตาม นอกเหนือจากวิธีการทั้งสองนี้ เรายังสามารถวิเคราะห์ได้โดยอาศัยเทคนิคการ Lag ตัวแปรได้ดังนี้

จากสมการ $Y_t = \beta_1+\beta_2X_{2t}+\beta_3X_{3t}+\ldots+\beta_kX_{kt}+u_t \quad \ldots.(1)$

lag สมการ (1) ไป 1 วาระ

$$Y_{t-1} = \beta_1+\beta_2X_{2,t-1}+\beta_3X_{3,t-1}+\ldots+\beta_kX_{k,t-1}+u_{t-1} \quad \ldots(2)$$

คูณตลอดสมการที่ (2) ด้วย ρ

$$\Rightarrow \rho Y_{t-1} = \rho\beta_1+\rho\beta_2X_{2,t-1}+\rho\beta_3X_{3,t-1}+\ldots+\rho\beta_kX_{k,t-1}+\rho u_{t-1} \quad \ldots.(3)$$

(1) - (3)

$$\Rightarrow (Y_{t-1}-\rho Y_{t-1})=\beta_1(1-\rho)+\beta_2(X_{2,t-1}-\rho X_{2,t-1})+\beta_3(X_{3,t-1}-\rho X_{3,t-1})+\ldots+$$

$$\beta_k(X_{k,t-1}-\rho X_{k,t-1})+(u_t-\rho u_{t-1}) \quad \ldots.(4)$$

พิจารณาสมการ (4) จะพบว่า Residual คือ $u_t-\rho u_{t-1}$ ซึ่งก็คือ v_t ในสมการ AR (1) คือ $u_t=\rho u_{t-1}+v_t$ นั่นเอง แต่ $E(v_t^2)= \sigma_v^2$ ในทุกค่าของ t แสดงว่าสมการที่ (4) ไม่มี

Autocorrelation ดังนั้นเมื่อดำเนินการวิเคราะห์สมการถดถอยโดยใช้สมการ (4) เราก็สามารถแก้ปัญหา Autocorrelation ใน AR (1) ได้

สำหรับสมการ (4) คือ

$$(Y_t-\rho Y_{t-1})=\beta_1(1-\rho)+\beta_2(X_{2t}-\rho X_{2,t-1})+\ldots+\beta_k(X_{kt}-\rho X_{k,t-1})+v_t; t=2,3,\ldots,n$$

เมื่อจัดให้อยู่ในรูปแมตริกซ์จะพบว่า

$$\begin{bmatrix} Y_2-\rho Y_1 \\ Y_3-\rho Y_2 \\ \vdots \\ Y_n-\rho Y_{n-1} \end{bmatrix} = \begin{bmatrix} 1-\rho & X_{22}-\rho X_{21} & \cdots & X_{k2}-\rho X_{k1} \\ 1-\rho & X_{23}-\rho X_{22} & \cdots & X_{k3}-\rho X_{k2} \\ \vdots & \vdots & & \vdots \\ 1-\rho & X_{2n}-\rho X_{2,n-1} & \cdots & X_{kn}-\rho X_{k,n-1} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad ---(5)$$

หรือ $Y^* = X\beta^* + V$ โดยที่ Y^* และ V เป็นเวคเตอร์ขนาด $(n-1)\times 1$ ขณะที่ X^* เป็นแมตริกซ์ขนาด $(n-1)\times k$ และ เมื่อพิจารณาเวคเตอร์ Y^* และแมตริกซ์ X^* จะพบว่า

$$Y^* = \begin{bmatrix} Y_2-\rho Y_1 \\ Y_3-\rho Y_2 \\ Y_4-\rho Y_3 \\ \vdots \\ Y_n-\rho Y_{n-1} \end{bmatrix} = \begin{bmatrix} -\rho & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -\rho & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -\rho & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ \vdots \\ Y_n \end{bmatrix} = P_1 Y$$

ขณะเดียวกันแมตริกซ์ X^* ก็คือแมตริกซ์ P_1X โดยที่ P_1 คือแมตริกซ์ขนาด $(n-1)\times n$

ด้วยเหตุนี้ สมการที่ (5) คือ $Y^* = X^*\beta+V$ ก็คือสมการที่เกิดจากการแปลงรูปโดยอาศัยแมตริกซ์ P_1 กล่าวคือ แปลงรูปจากสมการ $Y = X\beta+U$ เป็น $P_1Y = (P_1X) + P_1U$ สำหรับแมตริกซ์ P_1 นั้นถ้าจะสังเกตให้ดีก็จะพบว่า P_1 ก็คือแมตริกซ์ที่เกิดจากการตัดแถวที่ 1 ของแมตริกซ์ P ทิ้งไปนั่นเอง การประมาณค่าเวคเตอร์ β โดยอาศัยสมการที่ (4) หรือ (5) แม้ไม่แตกต่างจาก 2 วิธีที่ผ่านมาอย่างเด่นชัด แต่ก็มีความแม่นยำต่ำกว่าเล็กน้อย เพราะต้องสูญเสีย df ไป

1 หน่วย เพราะวิธีนี้ให้ df ของ Residual เพียง (n-1)-k ขณะที่วิธี GLS ให้ df ของ Residual เท่ากับ n-k การปรับปรุงคุณภาพของวิธีนี้เพื่อให้มีคุณภาพทัดเทียมกับวิธี GLS จึงกระทำได้โดยง่ายเพียงแต่เพิ่มเวคเตอร์ $(\sqrt{1-\rho^2}, 0, 0, \ldots, 0, 0)$ ให้เป็นแถวที่ 1 ของแมตริกซ์ P_1 แมตริกซ์ P_1 ก็จะแปรสภาพเป็นแมตริกซ์ P ผลการวิเคราะห์ในทั้ง 3 วิธีก็จะตรงกันทุกประการ

7.4.2 EGLS

ในกรณีของการประมาณค่าเวคเตอร์ β ด้วย GLS นั้นเราถือว่าต้องทราบค่า ρ จาก Priori Information การทราบค่าของ ρ จะมีผลให้ทราบแมตริกซ์ P และ W^{-1} ทำให้ประมาณค่าเวคเตอร์ β ได้ แต่โดยทั่วไปเรามักไม่ทราบค่าของ ρ ด้วยเหตุนี้จึงได้มีความพยายามที่จะประมาณค่าของ ρ ขึ้นมาด้วยวิธีการที่เหมาะสม แล้วแทนที่ ρ ด้วย $\hat{\rho}$ แมตริกซ์ที่จำเป็นต้องใช้สำหรับ GLS คือ P และ W^{-1} จะกลายเป็น $\hat{P}$ และ $\hat{W}^{-1}$ ตามลำดับ จากนั้นจึงใช้แมตริกซ์ $\hat{P}$ และ $\hat{W}^{-1}$ ประมาณค่าเวคเตอร์ β ได้

$$\hat{\beta} = (X'\hat{W}^{-1}X)^{-1}X'\hat{W}^{-1}Y, \quad \hat{\sigma}^2 = \frac{1}{n-k}(Y'\hat{W}^{-1}Y-\hat{\beta}'X'\hat{W}^{-1}Y)$$

และ $SS(\beta_1) = n\bar{Y}^{*2}$ เมื่อ $\bar{Y}^*$ คือค่าเฉลี่ยจากสมาชิกของเวคเตอร์ Y^* เมื่อ $Y^* = \hat{P}Y$ การประมาณค่าวิธีนี้เรียกว่า EGLS(Extimated Generalized Least Square)

สำหรับการประมาณค่า ρ เราสามารถกระทำได้หลายวิธีดังนี้

วิธีที่ 1 ประมาณค่า ρ โดยใช้สหสัมพันธ์ r ระหว่าง OLS - Residual

จาก OLS - Residual $e = Y-X\hat{\beta}$ เมื่อ $e = (e_1, e_2, \ldots, e_n)'$ แยกเวคเตอร์ e ออกเป็น 2 กลุ่มคือ $e_{t-1} = (e_1, e_2, \ldots, e_{n-1})'$ และ $e_t = (e_2, e_3, \ldots, e_n)'$ แล้วคำนวณหาค่าสหสัมพันธ์เชิงเส้นระหว่าง e_t กับ e_{t-1}

$$r = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sqrt{\sum_{t=2}^{n-1} e^2 \sum_{t=2}^{n} e_{t-1}^2}} \cong \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2} = r_1 = \hat{\rho}$$

หรือนัยหนึ่งให้ใช้ r_1 แทน r เป็นค่าประมาณของ ρ

ก. วิธีของ Cochrane - Orcutt ใช้ r_1 เป็นตัวประมาณค่าของ ρ โดยแทนที่ ρ ในแมตริกซ์ P_1 ด้วย r_1 (เรียกว่า Cochrane - Orcutt Two - Step Procedure)

$$\hat{P}_1 = \begin{bmatrix} -r_1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -r_1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -r_1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -r_1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -r_1 & 1 \end{bmatrix}_{(n-1)xn}$$

แล้ววิเคราะห์สมการถดถอย $\hat{P}_1 Y = (\hat{P}_1 X)\beta + \hat{P}_1 U$ หรือนัยหนึ่งก็คือให้แทนที่ ρ ด้วย r_1 แล้ววิเคราะห์สมการถดถอย

$$Y_t - r_1 Y_{t-1} = \beta(1-r_1) + \beta_2(X_{2t} - r_1 X_{2,t-1}) + \ldots + \beta_k(X_{kt} - r_1 X_{k,t-1}) + (u_t - r_1 u_{t-1})$$

$$;\ t = 2,3,\ldots,n$$

ข. วิธีของ Prais - Winsten ใช้ r_1 เป็นตัวประมาณค่าของ ρ โดยแทนที่ ρ ในแมตริกซ์ P ด้วย r_1

$$\hat{P} = \begin{bmatrix} \sqrt{1-r_1^2} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -r_1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -r_1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -r_1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -r_1 & 1 \end{bmatrix}_{nxn}$$

แล้ววิเคราะห์สมการถอถอย $\hat{PY} = (\hat{PX})\beta+\hat{Pu}$ หรืออาศัยวิธี EGLS ประมาณค่าเวคเตอร์ β

คือ $\hat{\beta} = (X'\hat{W}^{-1}X)^{-1}X'\hat{W}^{-1}Y, \quad \hat{\sigma}^2 = \frac{1}{n-k}(Y'\hat{W}^{-1}Y-\hat{\beta}'X'\hat{W}^{-1}Y)$

วิธีที่ 1 ข. ควรใช้กับกรณีขนาดตัวอย่าง n มีค่าน้อยเพราะสงวน df ได้ดีกว่าวิธีที่ 1 ก

สำหรับวิธีที่ 1 นั้น นอกจากจะประมาณค่า ρ ด้วยสหสัมพันธ์ระหว่าง e_t กับ e_{t-1} แล้วเรายังสามารถใช้ สปส.ความถดถอยจากสมการ $\hat{u}_t = \rho\hat{u}_{t-1}+\hat{v}_t$ ซึ่งก็คือ $e_t = \rho e_{t-1}+\hat{v}_t$ เป็นตัวประมาณค่าของ ρ ได้กล่าวคือ

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_{t-1}e_t}{\sum_{t=2}^{n} e_{t-1}^2} \quad {}^{1}$$

และเมื่อ $n \to \infty$ สปส.สหสัมพันธ์ r_1 และ สปส.ความถดถอย ρ จะมีค่าเท่ากัน

นอกจากนี้ยังมีวิธีประมาณค่า ρ ตามนัยของวิธีที่ 1 อีกวิธีหนึ่งซึ่งเสนอโดยธีล (Theil, 1971) ธีลเพียงแต่ปรับค่า r_1 ให้เหมาะสมขึ้นโดยใช้ df เป็นตัวถ่วงน้ำหนักดังนี้

$$r_1^* = \frac{(n-k)\sum_{t=2}^{n} e_t e_{t-1}}{(n-1)\sum_{t=1}^{n} e_t^2}$$

วิธีที่ 2 ใช้ Durbin - Watson Statistics คือ $\hat{\rho} \cong \frac{1}{2}(1-d)$

จาก Durbin-Watson Statistics คือ $d = \frac{\sum_{t=2}^{n}(e_t-e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$ จะพบว่า

---

[1] จากสมการถอถอย $y = x\beta+u$ เราทราบว่า $\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$ ในที่นี้ e_{t-1} ทำหน้าที่ x, e_t ทำหน้าที่ของ y และ ρ ทำหน้าที่ของ β

$$d = \frac{\sum_{t=2}^{n}(e_t^2+e_{t-1}^2-2e_te_{t-1})^2}{\sum_{t=1}^{n}e_t^2} = \frac{\sum_{t=2}^{n}e_t^2+\sum_{t=1}^{n}e_{t-1}^2-2\sum_{t=2}^{n}e_te_{t-1}}{\sum_{t=1}^{n}e_t^2}$$

แต่เมื่อ $n \to \infty$ จะพบว่า $\sum_{t=2}^{n}e_t^2$, $\sum_{t=2}^{n}e_{t-1}^2$ และ $\sum_{t=1}^{n}e_t^2$

จะมีค่าใกล้เคียงกัน นั่นคือ เมื่อ $n \to \infty$ และจะมีผลให้

$$d \cong \frac{2\sum_{t=2}^{n}e_{t-1}^2-2\sum_{t=2}^{n}e_te_{t-1}}{\sum_{t=2}^{n}e_{t-1}^2} = 2-2\hat{\rho}$$ เมื่อ $\hat{\rho}$ คือค่าประมาณ

สปส.ความถดถอยจากสมการ $e_t = \rho e_{t-1}+\hat{v}_t$

นั่นคือ เมื่อ $n \to \infty$ เราทราบว่า $d \cong 2-2\hat{\rho}$ หรือ $\hat{\rho} \cong 1-\frac{d}{2}$ เมื่อ d คือ

Durbin - Watson Statistics

วิธีนี้เหมาะสำหรับกรณีของกลุ่มตัวอย่างขนาดใหญ่ เพราะถ้า n มีขนาดเล็ก ค่าของ $\hat{\rho}$ จะไม่ค่อยแม่นยำ

วิธีที่ 3 Theil - Nagar Modification

ธีลและนาการ์ (Theil, H. and Nagar, A.L., 1961)[1] ได้เสนอค่าประมาณของ ρ โดยปรับปรุงวิธีที่ 2 ให้เหมาะสมและรับกับสถานการณ์ที่ค่าของตัวแปรอิสระ X's ต่าง ๆ มีค่าใกล้เคียงกัน (คือ First Difference และ Second Difference มีค่าต่ำ) ดังนี้

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

การประมาณค่า ρ ตามวิธีนี้จะดีที่สุดถ้าหากว่า $\rho < 0.4$

---

[1] Judge, G.G., et al., **Op. Cit.** p. 183; 241

วิธีที่ 4 Durbin's Two - Step

เดอร์บิน (Durbin, J, 1960)[1] เสนอวิธีประมาณค่าของ ρ ใน AR (q) ซึ่งสามารถใช้กับกรณี AR (1) ได้ดังนี้

ขั้นที่ 1 ประมาณค่า ρ

จากสมการ

$$Y_t-\rho Y_{t-1} = \beta_1(1-\rho)+\beta_2(X_{2t}-\rho X_{2,t-1})+\dots+\beta_k(X_{kt}-\rho X_{k,t-1})+v_t \quad ; \; t = 2,3,\dots,n$$

ให้จัดรูปสมการเสียใหม่เป็น

$$Y_t = \beta_1(1-\rho)+\rho Y_{t-1}+\beta_2X_{2t}-\beta_2\rho X_{2,t-1}+\dots+\beta_kX_{kt}-\beta_k\rho X_{k,t-1}+v_t \quad ; \; t = 2,3,\dots,n$$

แล้วกำหนดพารามิเตอร์เสียใหม่เพื่อให้เป็น Linear Model คือ

กำหนดให้ $\beta_1(1-\rho) = \alpha_1$, $\beta_2 = \alpha_2$, $\beta_2\rho = \alpha_3 \cdots$ ดังนี้เรื่อยไป ดังนั้นสมการเดิมจึงเปลี่ยนรูปเป็น

$$Y_t = \alpha_1+\rho Y_{t-1}+\alpha_2X_{2t}-\alpha_3X_{3,t-1}+\dots+ v_t \; ; \; t = 2,3,\dots,n$$

จากนี้ให้วิเคราะห์สมการถดถอยโดยวิธี OLS ตัวประมาณค่าที่สนใจคือ $\hat{\rho}$ ซึ่งเป็นสัมประสิทธิ์ของ Y_{t-1}

ขั้นที่ 2 แทนค่า $\hat{\rho}$ ในที่ของ ρ ในสมการเดิม

$$Y_t-\hat{\rho}Y_{t-1} = \beta_1(1-\hat{\rho})+\beta_2(X_{2t}-\hat{\rho}X_{2,t-1})+\dots+\beta_k(X_{kt}-\hat{\rho}X_{k,t-1})+v_t$$

แล้วประมาณค่าเวคเตอร์ β ด้วยวิธี OLS ตามปกติ หรือจะใช้วิธีการแปลงรูปก็ได้คือประมาณค่า β. จากสมการ $\hat{P}_1Y = (\hat{P}_1X)\beta + \hat{P}_1U$

---

[1] **Ibid.**, p. 183 และ Koutsoyiannis, A., **Op. Cit.**, p. 217-8

วิธีที่ 5 วิธี Iteration

วิธีนี้เสนอโดย Cochrane and Orcutt[1] ซึ่งเพิ่มขึ้นจากวิธีที่ 1 (ก) เพียงเล็กน้อยกล่าวคือ วิธีที่ 1 (ก) นั้น เราประมาณค่า ρ จากสมการถดถอย $e_t = \rho e_{t-1} + \hat{v}_t$ เพียงครั้งเดียวเมื่อได้ค่า $\hat{\rho}$ ก็แทนค่า $\hat{\rho}$ ในแมตริกซ์ P_1 แล้วประมาณค่าเวคเตอร์ β จากสมการ

$$\hat{P}_1 Y = (\hat{P}_1 X)\beta + \hat{P}_1 U$$

แต่สำหรับวิธีที่ 5 เราจะดำเนินการเช่นเดียวกับวิธีที่ 1 (ก) แต่กระทำซ้ำ ๆ สลับกับการตรวจสอบนัยสำคัญว่า $\rho = 0$ หรือไม่ โดยจะหยุดดำเนินการเมื่อพบว่า ρ ไม่มีนัยสำคัญดังนี้

1. จาก OLS - Residual $e = Y - X\hat{\beta}$ ให้ประมาณค่าสัมประสิทธิ์ความถดถอย ρ จากสมการ $e_t = \rho e_{t-1} + \hat{v}_t$ ได้ $\hat{\rho}$

2. แทนค่า $\hat{\rho}$ ในที่ของ ρ ในสมการเดิมได้

$$Y_t - \hat{\rho}Y_{t-1} = \beta_1(1-\hat{\rho}) + \beta_2(X_{2t} - \hat{\rho}X_{2,t-1}) + \ldots + \beta_k(X_{kt} - \hat{\rho}X_{k,t-1}) + v_t^*$$

แล้วประมาณค่า $\beta_1, \beta_2, \ldots, \beta_k$ พร้อมทั้งหา Residual และอาศัย Residual นี้เป็นข้อมูลเพื่อหาค่า d แล้วตรวจสอบว่ามี Autocorrelation หรือไม่ตามวิธีของ Durbin - Watson

3. ถ้าพบว่า v_t^* ยังมี Autocorrelation ให้ประมาณค่า ρ จากสมการ

$$e_t^* = \rho\, e_{t-1}^* + w_t$$

ได้ค่าของ ρ คือ $\hat{\hat{\rho}}$ แทนค่า $\hat{\hat{\rho}}$ ในสมการเดิมได้

$$Y_t - \hat{\hat{\rho}}\, Y_{t-1} = \beta_1(1-\hat{\hat{\rho}}) + \beta_2(X_{2t} - \hat{\hat{\rho}}X_{2,t-1}) + \ldots + \beta_k(X_{kt} - \hat{\hat{\rho}}X_{k,t-1}) + v_t^{**}$$

แล้วประมาณค่า $\beta_1, \beta_2, \ldots, \beta_k$ และหา Residual e^{**} พร้อมทั้งตรวจสอบดูว่า v_t^{**} มี Autocorrelation หรือไม่โดยอาศัยข้อมูล e^{**} ตามวิธีของ Durbin - Watson

---

[1] Koutsoyiannis, A., **Op. Cit.**, p.216

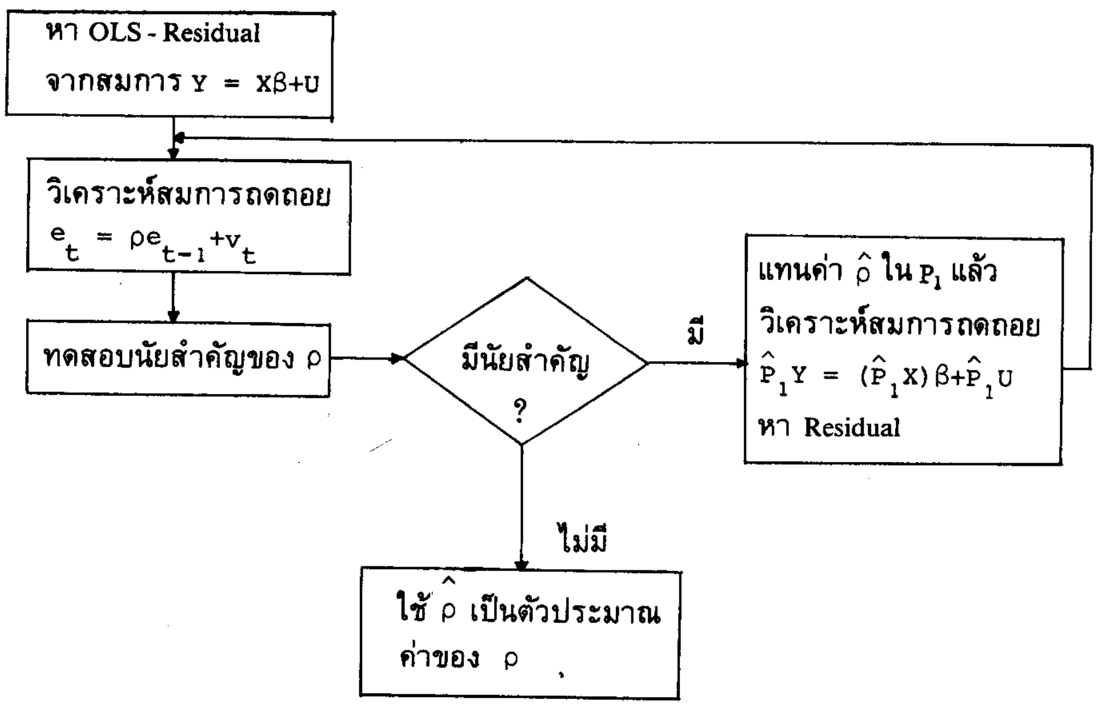
4. ให้ดำเนินการเช่นนี้เรื่อย ๆ ไป จนพบว่าไม่มี Autocorrelation ค่าประมาณของ ρ ที่มีผลให้ Residual จากสมการแปลงรูปไม่มี Autocorrelation ก็คือค่าประมาณของ ρ ตามต้องการ

ขอให้สังเกตว่าทั้ง Durbin และ Cochrane - Orcutt ต่างก็ประมาณค่า ρ จากสมการแปลงรูปคือ

$$Y_t-\rho Y_{t-1} = \beta_1(1-\rho)+\beta_2(X_{2t}-\rho X_{2,t-1})+..+\beta_k(X_{kt}-\rho X_{k,t-1})+v_t \quad ; t = 2,3,...,n$$

หรือ $P_1Y = (P_1X)\beta + P_1U$

สำหรับวิธีที่ 5 นี้เป็นวิธีที่พิสูจน์แล้วว่าเหมาะสมที่สุด นอกจากใช้เป็นวิธีประมาณค่าของ ρ ได้อย่างมีประสิทธิภาพแล้วยังสามารถใช้หลักเกณฑ์นี้กับ AR ดีกรีใด ๆ ก็ได้อีกทั้งยังเป็นวิธีที่สามารถช่วยให้นักวิจัยตัดสินใจว่าตนควรใช้ AR ดีกรีใด Programing Flow สำหรับวิธีนี้มีดังนี้



วิธีที่ 6 Search Procedure

วิธีเสนอโดยฮิลเดรธและลู (Chifford Hildreth and John Y. Lu, 1960)[1] กระบวนการ Search ใช้สมการเดียวกับวิธีที่ 4 และ 5 คือ

1. จากสมการ

$$Y_t-\rho Y_{t-1} = \beta_1(1-\rho)+\beta_2(X_{2t}-\rho X_{2,t-1})+\ldots+\beta_k(X_{kt}-\rho X_{k,t-1})+v_t$$

ให้ทดลองแปรค่า ρ ไปทีละน้อยจาก -1 ถึง 1 ในแต่ละครั้งให้คำนวณหา SS (Residual) คือ

$$SSR = \frac{1}{n-k}\,(Y'Y - \hat{\beta}'X'Y)$$

ค่าของ ρ ค่าใดที่มีผลให้ SSR มีค่าต่ำที่สุดว่าค่านั้นคือค่าประมาณชั้นต้น (First Approximator) ของ ρ สมมุติว่าเป็น ρ_0

2. จากค่าประมาณชั้นต้นของ ρ คือ ρ_0 ให้กำหนดช่วงแคบ ๆ ที่มีค่าใกล้เคียงกับ ρ_0 สมมุติว่าเป็น $(\rho_0-\delta, \rho_0+\delta)$ จากนั้นให้ดำเนินการเช่นเดียวกับข้อ 1 โดยทดลองแปรค่า ρ จาก $\rho_0-\delta$ ถึง $\rho_0+\delta$ ค่าของ ρ ที่มีผลให้ SSR ต่ำสุดถือว่าเป็นค่าประมาณของ ρ

จากการศึกษาโดยอาศัย Monte Carlo Experiment พบว่าวิธีที่ 6 นี้ ประมาณค่า ρ ได้ดีและมีประสิทธิภาพกว่าวิธีที่ 4 เพราะในบางครั้งพบว่าวิธีที่ 4 ให้ค่าประมาณของ ρ สูงกว่า 1 ซึ่งไม่ใช่สิ่งที่ถูกต้องในทางทฤษฎีสถิติ

นอกเหนือจากวิธีที่กล่าวแล้วยังมีวิธีอื่น ๆ อีกเช่นวิธี Nonlinear Least Square (NLS) วิธี Maximum Likelihood (ML) และวิธีอื่น นักศึกษาสามารถศึกษาเพิ่มเติมได้จากเอกสารอ้างอิง

---

[1] Maddala, G.S., **Op. Cit.**, p. 279

7.5 การทดสอบ Autocorrelation

การทดสอบว่าข้อมูลที่มีอยู่เอื้ออำนวยให้เกิดปัญหา Autocorrelation หรือไม่นั้น เราสามารถกระทำได้หลายวิธี ในที่นี้จะเสนอเป็น 2 แบบคือ test สำหรับกรณีทั่วไปซึ่งอาจเป็น AR, MA หรือ ARMA ดีกรีที่เท่าไรก็ได้ ซึ่งถือว่าเป็นตัวทดสอบแบบกลาง ๆ กับ test เฉพาะ AR (1)

7.5.1 วิธีทดสอบ Autocorrelation กรณีทั่วไป

7.5.1 ก. การตรวจสอบนัยสำคัญของสัมประสิทธิความถดถอย

การทดสอบวิธีนี้สามารถใช้กับ AR (q), MA (p) และ ARMA (q, p) ได้เพียงแต่เราระบุดีกรี q และ p ตามที่พิจารณาเห็นสมควรแก่สถานการณ์ลงไปให้ชัดเจนก็จะแก้ปัญหาได้เช่น

AR(q) คือ $e_t = \sum_{i=1}^{q} \theta_i e_{t-i} + v_t$

AR(1) คือ $e_t = \theta_1 e_{t-1} + v_t$ หรือที่เราใช้เป็น $e_t = \rho e_{t-1} + v_t$

MA(p) คือ $e_t = v_t + \sum_{i=1}^{p} \alpha_i v_{t-i}$

ARMA(q,p) คือ $e_t = \sum_{i=1}^{q} \theta_i e_{t-i} + v_t + \sum_{i=}^{p} \alpha_i v_{t-i}$

ในที่นี้จะกล่าวถึงวิธีปฏิบัติเฉพาะ AR (1) เท่านั้น ส่วน AR (q); q = 2, 3, ... หรือ MA (p) และ ARMA (q, p) ก็ปฏิบัติเช่นเดียวกัน ดังนี้

(1) จากสมการ $Y = X\beta + U$ ให้ประมาณค่า β ด้วย OLS และคำนวณหา Residual ได้ $e = Y - X\hat{\beta}$ แล้วนำเวคเตอร์ e มาจำแนกเป็น 2 กลุ่มคือ $e_t = (e_2, e_3, \ldots, e_n)'$ $e_{t-1} = (e_1, e_2, \ldots, e_{n-1})'$ โดยเวคเตอร์ e_t ทำหน้าที่ของเวคเตอร์ Y และเวคเตอร์ e_{t-1} ทำหน้าที่ของเวคเตอร์ X

(2) วิเคราะห์สมการถดถอย $e_t = \rho e_{t-1} + \hat{v}_t$ ตามวิธี OLS (Simple Linear Regression) ได้

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2} \quad , \quad \hat{\sigma}^2 = \frac{1}{n-1} \left(\sum_{t=1}^{n-1} e_t^2 - \hat{\rho} \sum_{t=2}^{n} e_t e_{t-1} \right)$$

$$\hat{V}(\hat{\rho}) = \frac{\hat{\sigma}^2}{\sum_{t=2}^{n} e_{t-2}^2} \quad \text{และ} \quad s_{\hat{\rho}} = \frac{\hat{\sigma}}{\sqrt{\sum_{t=2}^{n} e_{t-1}^2}}$$

(3) ทดสอบสมมุติฐาน $H_0 : \rho = 0$ VS $H_1 : \rho \neq 0$

โดยเราจะปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ $|t_c| > t_{n-1, 1-\frac{\alpha}{2}}$ เมื่อ $t_c = \frac{\hat{\rho}}{s_{\hat{\rho}}}$ ซึ่งในกรณีนี้ถ้าพบว่า ρ ไม่มีนัยสำคัญทางสถิติก็เป็นการยืนยันว่าข้อมูลที่มีอยู่ไม่เอื้ออำนวยให้มีปัญหา Autocorrelation ใน AR (1)

สำหรับการทดสอบปัญหา Autocorrelation ใน AR, MA และ ARMA ดีกรีอื่น ๆ ให้ใช้ F-test หรือจะยังคงใช้ t-test ก็ได้เช่นใน AR (2) คือ

$$e_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \hat{v}_t$$

เราต้องจำแนก OLS - Residual e ออกเป็น 3 กลุ่มคือ $e_t = (e_3, e_4, \ldots, e_n)'$, $e_{t-1} = (e_2, e_3, \ldots, e_{n-1})$ และ $e_{t-2} = (e_1, e_2, \ldots, e_{n-2})'$

จากนั้นให้วิเคราะห์สมการถดถอยตามวิธี Multiple Regression แล้วทดสอบสมมุติฐาน

$$H_0 : \theta_1 = 0, \ \theta_2 = 0 \quad \text{VS} \quad H_1 : \theta_1 \neq 0, \ \theta_2 \neq 0$$

7.5.2 วิธีทดสอบ Autocorrelation ใน AR (1)

7.5.2 ก. Durbin - Watson Test (DW)

DW เป็นเครื่องมือทดสอบ Autocorrelation ใน AR (1) ที่มีผู้นิยมใช้กันค่อนข้างกว้างขวาง แม้จะได้พัฒนามาตั้งแต่ปี 1950 เดอร์บินและวัตสันเสนอตัวทดสอบสำหรับสมมุติฐานต่อไปนี้ดังนี้

No Autocorrelation VS H_1 : **Positive** Autocorrelation

ปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α ถ้า $d < d_L$ ไม่อาจปฏิเสธสมมุติฐานหลัก (ยอมรับ) ถ้า $d > d_U$ และไม่อาจตัดสินใจได้ถ้า $d_L < d < d_U$

ในขณะเดียวกัน ถ้าสมมุติฐานรองเปลี่ยนเป็น H_1 Negative Autocorrelation ให้ปฏิเสธสมมุติฐานหลัก ถ้า $d < 4-d_L$ ไม่อาจปฏิเสธสมมุติฐานหลักถ้า $d < 4-d_U$ และไม่อาจตัดสินใจได้ ถ้า $4-d_U < d < 4-d_L$

โดยที่ $d = \dfrac{\sum_{t=2}^{n}(e_t-e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$ เมื่อ e คือ OLS - Residual Vector กล่าวคือ

$e = Y-X\hat{\beta}$ หรือเสนอในรูปแมตริกซ์ได้ดังนี้

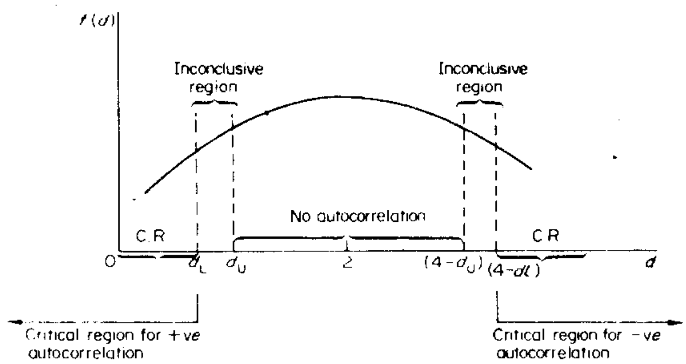
$$d = \frac{\sum_{t=2}^{n}(e_t-e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{e'Ae}{e'e} = \frac{U'MAMU}{U'MU}$$

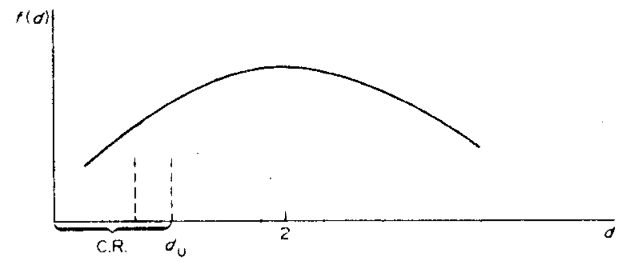
หมายเหตุ

$$d = \frac{\sum_{t=2}^{n}(e_t-e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} = \frac{\sum_{t=2}^{n} e_t^2 + \sum_{t=2}^{n} e_{t-1}^2 - 2\sum_{t=2}^{n} e_t e_{t-1}}{e'e} = \frac{a}{e'e}$$

พิจารณาเศษจะพบว่าสามารถจัดรูปเป็น Quadratic Form ได้ดังนี้

สำหรับเขตวิกฤติสามารถเสนอได้ดังภาพ





ข้อสังเกต

(1) จาก $d = \dfrac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$ เราสามารถพิสูจน์ได้ว่า $d \cong 2(1-\hat{\rho})$ ซึ่งแสดงว่าถ้าไม่มี Autocorrelation คือ $\rho = 0$ จะพบว่า $d \cong 2$ ความจริงเหล่านี้เป็นเครื่องยืนยันว่า ค่าของ d จะปรากฏอยู่ระหว่าง 0 ถึง 4 และเมื่อใดก็ตามที่ d มีค่าใกล้เคียง 2 เราจะพอเดาได้ว่าคงจะไม่มีปัญหา Autocorrelation ถ้า d มีค่าต่ำใกล้ 0 เราก็พอจะเดาว่าคงจะเกิดปัญหา Positive Autocorrelation (โดยทั่วไปเรามักประสบปัญหา Positive Autocorrelation มากกว่า Negative Autocorrelation) และ ถ้า d มีค่าสูงใกล้ 4 เราจะเดาได้ว่าขณะนี้คงเกิดปัญหา Negative

$$a = (e_2^2+e_3^2+\dots+e_n^2)+(e_1^2+e_2^2+\dots+e_{n-1}^2)-2(e_1e_2+e_2e_3+e_3e_4+\dots+e_{n-1}e_n)$$

$$= \{e_1^2+(2e_2^2+2e_3^2+\dots+2e_{n-1}^2)+e_n^2\} -2e_1e_2-2e_2e_3-2e_3e_4-\dots- 2e_{n-1}e_n$$

$$= (e_1,e_2,\dots,e_n)\begin{bmatrix} 1 & -1 & & & & & \\ -1 & 2 & -1 & & & 0 & \\ & -1 & 2 & -1 & & & \\ & & -1 & 2 & -1 & & \\ & & & \ddots & \ddots & \ddots & \\ & 0 & & & -1 & 2 & -1 \\ & & & & & -1 & 1 \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = e'Ae$$

ดังนั้น $d = \dfrac{e'Ae}{e'e}$ แต่เนื่องจาก $e = Y-X\hat{\beta} = [X\beta+U]-X[(X'X)^{-1}X'(X\beta+U)]$

$$= X\beta+U-X[\beta+(X'X)^{-1}X'U] = U-X(X'X)^{-1}X'U = [In-X(X'X)^{-1}X\ ]U = MU$$

ดังนั้น $$d = \frac{e'Ae}{e'e} = \frac{U'M'AMU}{U'M'MU} \quad \frac{U'MAMU}{U'MU}$$

โดยที่

$M = In -X(X'X)^{-1}X'$ และ

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

Autocorrelation การคาดเดาเหล่านี้เป็นสิ่งที่พอกระทำได้ แต่ก็มิใช่สิ่งที่ควรกระทำเป็นปกติเพราะในช่วง 0-4 ยังมีช่วงที่เป็น Indecision Interval อยู่

(2) Indecision Interval คือช่วงที่ $d_L<d<d_U$ และ $4-d_U<d<4-d_L$ เป็นช่วงอันตราย เพราะ ถ้า d ตกอยู่ใน Indecision Interval เราจะไม่อาจดำเนินการใด ๆ ได้นอกจากต้องเพิ่มขนาดตัวอย่างซึ่งมิใช่สิ่งที่สามารถกระทำได้ง่ายนัก โดยเฉพาะในการวิจัยที่อาศัยข้อมูลอนุกรมเวลา

(3) เดอร์บินและวัตสันมิได้กำหนด Sampling Distribution ของ d ไว้อย่างแน่ชัด ทั้งนี้เพราะค่าของ d ผูกพันอยู่กับลักษณะของแมตริกซ์ X กล่าวคือถ้าแมตริกซ์ X มีลักษณะเปลี่ยนไป f (d) ก็จะเปลี่ยนแปลงไป ทำให้ไม่อาจกำหนดรูปทั่วไปของ f(d) ได้

(4) DW ใช้ได้เฉพาะ AR (1) เท่านั้น ไม่อาจใช้กับ AR (q), q = 2, 3, ... ได้ นอกจากนี้ถ้ามี Lagged Variable ปะปนอยู่ในระหว่างตัวแปรอิสระ DW จะมี Power ต่ำกว่าวิธีอื่น ๆ

จากข้อสังเกตทั้ง 4 ประการจะพบว่าปัญหา Indecision เป็นปัญหาสำคัญที่จะต้องแก้ไข มิเช่นนั้น นักวิจัยจะประสบปัญหาในทางปฏิบัติ เกี่ยวกับปัญหานี้ได้มีผู้พยายามแก้ไขหลายท่านด้วยกัน รวมทั้งเดอร์บินและวัตสันด้วยดังนี้

วิธีที่ 1 ให้ใช้ d_U และ $4-d_U$ เป็นเขตวิกฤตกล่าวคือ ปฏิเสธ $H_0 : \rho = 0$ vs $H_1 : \rho > 0$ เมื่อ $d < d_U$ และปฏิเสธ $H_0 : \rho = 0$ vs $H_1 : \rho < 0$ เมื่อ $d > 4-d_U$

วิธีนี้เสนอโดยแฮนแนนและเทอร์เรล(Hannan, E.J. and Terrell, R.D., 1968)[1] หลักการของวิธีนี้ก็คือการผนวกเอา Indecision Interval เข้าเป็นเขตวิกฤต วิธีนี้แม้จะทำให้ระดับของ Probability of type I (α- error) สูงกว่าที่กำหนดไว้เดิม แต่ก็ใช้ได้ผลดีโดยเฉพาะอย่างยิ่งเมื่อค่าของตัวแปรอิสระไม่ผันผวนมากนัก

วิธีที่ 2 ประมาณการแจกแจงของ d ด้วยการแจกแจงของ $a+bd_U$

วิธีนี้เรียกว่า Durbin - Watson's $a+bd_U$ Approximation เสนอโดยเดอร์บินและวัตสันในปี 1971[2] ซึ่งมีวิธีการดังนี้

---

[1] Maddala, G.S., **Op. Cit.** p. 258

[2] Durbin, J. and Watson, G.S., "Testing for Serial Correlation in Least Square Regression III" (Biometrica, 1971, 58, 1), p. 17-18 (Appendix 3)

(1) คำนวณหาค่าของ E(d) และ V(d)

$E(d) = \frac{P}{n-k}$ เมื่อ $P = 2(n-1)-tr[X'AX(X'X)^{-1}]$ และ A คือแมตริกซ์ของ Quadratic Form ในตอน 7.5.2[1]

$$V(d) = \frac{2}{(n-k)(n-k+2)}[Q-PE(d)]$$

เมื่อ $Q = 2(3n-4)-2\ tr[X'A^2X(X'X)^{-1}]+tr\ [X'AX(X'X)^{-1}]^2$

(2) แก้สมการหาค่าของ a และ b จากสมการ $E(d) = a+bE(d_U)$ และ $V(d) = b^2V(d_U)$ เมื่อพิจารณาสมการคู่นี้จะพบว่าเราสามารถแก้สมการหาค่าของ a และ b ได้โดยง่ายดังนี้

$$b = \sqrt{\frac{V(d)}{V(d_U)}}$$

$$a = E(d)-bE(d_U) = E(d) - E(d_U)\sqrt{\frac{V(d)}{V(d_U)}}$$

---

1 $d = \frac{U'MAMU}{U'MU}$ จะเห็นได้ว่า

$$E(U'MAMU) = tr(MAM) = tr(AMM) = tr(AM) = tr\ A[In-X(X'X)^{-1}X']$$
$$= tr\ A- tr[AX(X'X)^{-1}X'] = tr\ A- tr[X'AX(X'X)^{-1}]$$

แต่ $tr\ A = 1+(2+2+...+2)+1 = \underbrace{2+2+2+...+2}_{n-1\text{ ครั้ง}} = 2(n-1)$

ดังนั้น $E(U'MAMU) = 2(n-1) - tr[X'AX(X'X)^{-1}]$,

ในทำนองเดียวกัน $E(U'MU) = tr\ M = tr[In-X'(X'X)^{-1}X'] = n-k$

แต่ U'MU และ d เป็นอิสระต่อกัน ดังนั้น d(U'MU) = U'MAMU ดังนั้น E(d) E(U'MU) = E(U'MAMU)

หรือ $E(d) = \frac{E(U'MAMU)}{E(U'MU)}$ (อ่าน Durbin J. and Watson, G.S., "Testing for Serial Correlation in Least Square Regression I", Biometrica, 1950, 37, P.409-428

ทั้งนี้เราทราบค่าของ E(d) และ V(d) ได้จากข้อ (1) ขณะที่เราสามารถหาค่า $E(d_U)$ และ $V(d_U)$ได้จากสมการต่อไปนี้[1]

$$E(d_U) = 2+ \frac{2}{n-k} \sum_{j=1}^{k-1} \text{Cos}(j\P/n)$$

$$V(d_U) \quad \frac{4[n-2-2\sum_{j=1}^{k-1} \text{Cos}^2(j\P/n)-2\{\sum_{j=1}^{k-1} \text{Cos}(j\P/n)\}^2/(n-k)]}{(n-k)(n-k+2)}$$

(3) ปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho > 0$ ณ ระดับนัยสำคัญ α เมื่อ $d < a+bd_U^0$ เมื่อ d_U^0 คือค่าวิกฤตของ d_U จากตารางท้ายเล่ม

วิธีที่ 3 ใช้ N(0,1) หรือ Z-test

แบลตเบอร์ก (Blattberg, R.C., 1972)[2] แนะนำให้ใช้ Z-test ตัดสินใจสำหรับงานทดสอบสมมุติฐาน $H_0 : \rho = 0$ VS $H_1 : \rho > 0$ หรือ $\rho < 0$ โดยให้ตัดสินใจดังนี้

ปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho > 0$ ณ. ระดับนัยสำคัญ α เมื่อ $Z_c > Z_{1-\alpha}$ และปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho < 0$ ณ. ระดับนัยสำคัญ α เมื่อ $Z_c < Z_\alpha$ โดยที่

$$Z_c = \frac{d-E(d)}{\sqrt{V(d)}}$$

ทั้งนี้ E(d) และ V(d) คำนวณได้จากวิธีที่ 2 ขณะที่ d คือ DW-Statistics

การตัดสินใจตามวิธีนี้นอกจากช่วยแก้ปัญหาเรื่อง Indecision Interval แล้วยังสามารถใช้ทดสอบได้ในทุกระดับนัยสำคัญ ขณะที่ DW-test ใช้ทดสอบได้เพียง 3 ระดับคือ 5%, 2.5% และ 1%

---

[1] Ibid p. 18

[2] Maddala, G.S., Op. Cit., p.286

7.5.2 ข Berenblut - Webb test

เบเรนบลูต และเวบบ์ (Berenblut, I, J, and Webb, G.I., 1973)[1] เสนอให้ใช้ตัวทดสอบ g ซึ่งสามารถใช้ได้ ทั้งในกรณีของ Stationary คือ $|\rho| < 1$ และ Non - Stationary คือ ρ อาจมีค่ามากกว่า ± 1 ได้

ก. $H_0 : \rho = 0$ VS $H_1 : \rho \neq 0$

$$g = \frac{(Y-X\tilde{\beta})'V^{-1}(\hat{\rho})(Y-X\tilde{\beta})}{(Y-X\hat{\beta})'(Y-X\hat{\beta})}$$

โดยที่

$$V^{-1}(\rho) = \begin{bmatrix} 1+\rho^2 & -\rho & 0 & 0 & \cdots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}$$

$\tilde{\beta} = [X'V^{-1}(\hat{\rho})X]^{-1}X'V^{-1}(\hat{\rho})Y$ และ $\hat{\beta} = (X'X)^{-1}X'Y$

ข. $H_0 : \rho = 0$ VS $H_1 : \rho > 0$

กรณีนี้ให้ถือเป็นกรณีเฉพาะของข้อ ก. โดยเพียงแต่แทนที่ $\hat{\rho}$ ด้วย 1 แมตริกซ์ $V^{-1}(\hat{\rho})$ จะกลายเป็นแมตริกซ์ B ดังนี้

$$B = \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

---

[1] Berenblut, I.J. and Webb, G.I, "A New Test for Autocorrelated Errors in the Linear Regression Model" (Journal of Royal Statistical Society, B, 35, 1973), p.33-50

ดังนั้น $\tilde{\beta} = (X'BX)^{-1}X'BY$ และ g จะเปลี่ยนรูปไปดังนี้

$$g = \frac{(Y-X\tilde{\beta})'B(Y-X\tilde{\beta})}{(Y-X\hat{\beta})'(Y-X\hat{\beta})} = \frac{Y'[B-BX(X'BX)^{-1}X'B]Y}{Y'MY}$$

ค. $H_0 : \rho = 0$ VS $H_1 : \rho < 0$

กรณีนี้ถือเป็นกรณีเฉพาะของข้อ ก โดยเพียงแต่แทนที่ ρ ด้วย -1 แมตริกซ์ $V^{-1}(\rho)$ จะกลายเป็นแมตริกซ์ B_1 และ g จะเปลี่ยนรูปไปดังนี้

$$g = \frac{Y'[B_1-B_1X(X'B_1X)^{-1}X'B_1]Y}{Y'MY}$$

จากการศึกษาของแบเรนบรูตและเวบบ์[1] พบว่าค่าของ g จะปรากฏในช่วงเดียวกันกับค่า d คือ $0 < g < 4$ การตัดสินใจจึงใช้ค่าของเดอร์บินและวัตสัน d_U และ d_L และตัดสินใจเช่นเดียวกับ DW กล่าวคือ

(1) ปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho > 0$ ณ ระดับนัยสำคัญ α ถ้า $g < d_L$

(2) ปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho < 0$ ณ ระดับนัยสำคัญ ถ้า $g > 4-d_L$

(3) ไม่ตัดสินใจถ้า $d_L < g < d_U$ และ $4-d_U < g < 4-d_L$ ซึ่งในกรณีเช่นนี้ให้แก้ปัญหาเช่นเดียวกับที่กล่าวมาแล้วในตอน 7.5.2 ก

(4) ปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho \neq 0$ ณ ระดับนัยสำคัญ $\frac{\alpha}{2}$ เมื่อ $g < d_L$ หรือ $d > 4-d_L$ ที่ระดับนัยสำคัญ $\frac{\alpha}{2}$

สำหรับในสถานการณ์ทางปฏิบัตินั้นเราสามารถคำนวณหาค่า g ได้โดยง่าย ขอให้พิจารณาตัวสถิติ g คือ

$$g = \frac{(Y-X\tilde{\beta})'V^{-1}(\hat{\rho})(Y-X\tilde{\beta})}{(Y-X\hat{\beta})'(Y-X\hat{\beta})} = \frac{a}{b}$$

แบเรนบรูต และเวบบ์แนะนำว่าค่าของ a ก็คือ Residual Sum Square ของสมการ[2]

$$CY = (CX)\beta + CU$$

---

[1] **Ibid.**, p.38

[2] Ibid., p.45

โดยที่

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix}_{n \times n} \quad \text{เมื่อ } C'C = V^{-1}(1)$$

ขอให้สังเกตว่า $CY = (CX)\beta + CU$ ก็คือการวิเคราะห์สมการถดถอย First Difference นั่นเอง คือ

$$\begin{bmatrix} Y_1 \\ Y_2 - Y_1 \\ Y_3 - Y_2 \\ \vdots \\ Y_n - Y_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 0 & X_{22}-X_{21} & X_{32}-X_{31} & \cdots & X_{k2}-X_{k1} \\ 0 & X_{23}-X_{22} & X_{33}-X_{32} & \cdots & X_{k3}-X_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & X_{2n}-X_{2,n-1} & X_{3n}-X_{3,n-1} & \cdots & X_{kn}-X_{k,n-1} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2-u_1 \\ u_3-u_2 \\ \vdots \\ u_n-u_{n-1} \end{bmatrix}$$

และถ้าตัดค่าสังเกต $Z_1 = (Y_1, X_{21}, X_{31}, \dots, X_{k1})$ ทิ้งซึ่งก็คือตัดแถวที่ 1 ของ C ทิ้งไป[1] Residual Sum Square จากสมการ $CY = (CX)\beta + CU$ ก็คือ Residual Sum Square จากสมการ

$$Y_t - Y_{t-1} = \beta_2(X_{2t} - X_{2,t-1}) + \beta_3(X_{3t} - X_{3,t-1}) + \dots + \beta_k(X_{k,t} - X_{k,t-1}) + (u_t - u_{t-1})$$

$$; \ t = 2,3,\dots,n$$

หรือ

$$\Delta Y_t = \beta_2 \Delta X_{2t} + \beta_3 \Delta X_{3t} + \dots + \beta_k \Delta X_{kt} + \Delta u_t \ ; \ t = 2,3,\dots,n$$

สมมุติว่า Residual Sum Square จากสมการ First Difference นี้คือ RSS* ดังนี้ แสดงว่า a = RSS*

---

[1] Ibid., p.45

ปัญหาของ Durbin's h Test มี 2 ประการคือ

1. h จะไม่ปรากฏค่าถ้า $n\hat{V}(\hat{\lambda}) > 1$

2. h ไม่เหมาะกับงานที่ใช้กลุ่มตัวอย่างขนาดเล็กซึ่งเป็นเหตุการณ์ปกติของงานวิจัยทั่วไป

สำหรับปัญหาประการแรกคือ เมื่อพบว่า $n\hat{V}(\hat{\lambda}) > 1$ ซึ่ง h จะไม่ปรากฏค่านั้น เดอร์บิน แนะนำให้แก้ไขปัญหาและใช้ทางเลือกดังนี้

(1) ให้วิเคราะห์หา Residual จากสมการเดิมสมมุติว่าสมการเดิมคือ Koyck 's Geometric Lag Model

$$Y_t = \beta_1+\beta_2X_{2t}+\lambda Y_{t-1}+u_t \quad ; \; t = 2,3,...,n$$

ก็ให้วิเคราะห์สมการดังกล่าวตามวิธีที่เหมาะสมแล้วคำนวณหา Residual[1]

(2) นำ Residual e มาวิเคราะห์สมการถดถอย

$$e_t = \beta_1+\beta_2X_{2t}+\lambda Y_{t-1}+\rho e_{t-1}+u_t \qquad ; \; t = 2,3,...,n$$

การวิเคราะห์ให้ใช้วิธีที่เหมาะสม จากนั้นให้ตรวจสอบนัยสำคัญของ ρ ถ้าพบว่าไม่มีนัยสำคัญ (ยอมรับ $H_0 : \rho = 0$) ก็แสดงว่าไม่มี Autocorrelation

สำหรับปัญหาประการที่ 2 คือกรณีที่ h - Test เหมาะสมสำหรับเฉพาะกรณีของงานที่ใช้กลุ่มตัวอย่างขนาดใหญ่ ซึ่งไม่เอื้ออำนวยให้สามารถใช้กับงานวิจัยทั่วไปที่มักเป็นงานที่ใช้กลุ่มตัวอย่างขนาดเล็กนั้น ปาร์ค (Park, Soo-Bin, 1975)[2] แนะนำให้ใช้ t-test โดย ปาร์ค ทำการทดลองเปรียบเทียบ Power ระหว่าง t - Test, h - Test และ DW - Test โดยวิธี Monte - Carlo Experiment กับสมการหลายแบบ สำหรับกรณีของขนาดตัวอย่างขนาดเล็กเช่น

$$Y_t = \alpha+\beta_1Y_{t-1}+\beta_2X_{t-1}+u_t$$

$$Y_t = \alpha+\beta_1Y_{t-1}+\beta_2X_{t-1}+\beta_3X_{t-1}+u_t$$

$$Y_t = \alpha+\beta_1Y_{t-1}+u_t$$

---

[1] จะกล่าวถึงเทคนิคการวิเคราะห์สมการถดถอยกรณีของ Lagged Variable และ Lagged Dependent Variable ในบทต่อไป

[2] Park, Soo-Bin, "On the Small-Sample Power of Durbin's h-Test", (Jour. of Amer. Stat. Assoe., 70, 1975), p. 60-73

ขณะเดียวกัน เมื่อพิจารณาส่วนของ g คือ $b = (Y-X\hat{\beta})'(Y-X\hat{\beta})$ จะพบว่า b ก็คือ $e'e$ ซึ่งก็คือ Residual Sum Square จากสมการ $Y = X\beta+U$ นั่นเอง

ดังนั้น รูปที่ง่ายที่สุดของ g ก็คือ

$$g = \frac{RSS^*}{RSS}$$

จากการศึกษาพบว่า g-test มี Power สูงกว่า DW เมื่อ $\rho > 0.5$ และพบว่าโดยทั่วไปแล้ว g-test เป็นเครื่องมือทดสอบปัญหา Autocorrelation ได้ดี และใช้แทน DW ได้

7.5.2 ค Durbin's h Test

ในกรณีที่มี lagged dependent variable คือ Y_{t-1} ปรากฏร่วมเป็นตัวแปรอิสระเช่นสมการ

$$Y_t = \beta_1+\beta_2X_{2t}+\lambda Y_{t-1}+u_t \; ; \; t = 2,3,\ldots,n \qquad \text{หรืออื่น ๆ}$$

ในกรณีเช่นนี้เดอร์บินแนะนำให้ใช้สถิติ h สำหรับทดสอบ $H_0 : \rho = 0$ ดังนี้[1]

$$h = r_1 \frac{n}{\sqrt{1-n\hat{V}(\hat{\lambda})}}$$

โดยที่ r_1 คือค่าประมาณของ ρ ซึ่งประมาณได้โดยอาศัยวิธีที่กล่าวมาแล้วในตอน 7.4.2 n คือขนาดตัวอย่างและ $\hat{V}(\hat{\lambda})$ คือค่าประมาณ Variance ของ λ ซึ่งเป็นสัมประสิทธิ์ของ Y_{t-1}

เดอร์บินพบว่า $h \underset{ass}{\sim} N(0,1)$ กล่าวคือ เมื่อ $n \to \infty$ แล้ว $h \sim N(0,1)$ ซึ่งแสดงว่า h-test ใช้ทดสอบได้วิธีเดียวกับ Z-test และเหมาะสำหรับงานที่ใช้กลุ่มตัวอย่าง (n) ขนาดใหญ่กล่าวคือ

(1) ปฏิเสธ $H_0 : \rho = 0$ vs $H_1 : \rho > 0$ ณ.ระดับนัยสำคัญ α เมื่อ $h > Z_{1-\alpha}$
(2) ปฏิเสธ $H_0 : \rho = 0$ vs $H_1 : \rho < 0$ ณ.ระดับนัยสำคัญ α เมื่อ $h < Z_{\alpha}$
(3) ปฏิเสธ $H_0 : \rho = 0$ vs $H_1 : \rho \neq 0$ ณ.ระดับนัยสำคัญ α เมื่อ $|h| > Z_{1-\frac{\alpha}{2}}$

---

[1] Durbin, J. "Testing for Serial Correlation in Least Square Regression when Some of the Regressors are Lagged Dependent Variables", (Econometrica, 38, May, 1970) p. 410-21 ความจริงแล้วเดอร์บินประมาณค่า ρ ด้วย $1-\frac{1}{2}d$ เมื่อ d คือ DW - Statistics

โดยถือว่า u_t มี Autocorrelation ใน AR (1) ผลการศึกษาครั้งนี้พบว่า

(1) สถานการณ์ที่ $n\hat{V}(\hat{\lambda}) > 1$ มักปรากฏขึ้นเมื่อ $|\rho| < 0.2$ หรือ $n \cong 20$ แต่เหตุการณ์เช่นนี้เกิดขึ้นได้น้อยมากคือเพียง 0.84% เท่านั้น

(2) t-test จะมี Power สูงกว่า h-test เมื่อ $\rho < 0$ และ h-test จะมี power สูงกว่า t-test เมื่อ $\rho > 0$

(3) ทั้ง t-test และ h-test มีPower สูงกว่า DW-test ยกเว้นเฉพาะเมื่อ $|\rho| \leq 0.40$ และเมื่อ $n = 30$

(4) ในทางปฏิบัติเราควรเลือกใช้ h-test เพราะวิเคราะห์ได้ง่ายกว่าเว้นแต่เมื่อ $n\hat{V}(\hat{\lambda}) \geq 1$ เท่านั้นที่เราควรเปลี่ยนไปใช้ t-test

7.5.2 ง Wallis test

วอลลิส (Wallis, Kenneth F., 1972)[1] เสนอตัวสถิติ d_4 เพื่อใช้ตรวจสอบปัญหา Autocorrelation ในกรณีผู้วิจัยดำเนินการวิจัยโดยใช้ข้อมูลรายไตรมาส (Quarterly Time Series Data) โดยขยายแนวคิดมาจากวิธีการของเดอร์บิน-วัตสัน และ AR (4) ที่ใช้วอลลิสใช้รูปง่ายคือ

$$u_t = \rho u_{t-4} + v_t$$

สถิติที่ใช้ทดสอบสมมุติฐาน $H_0 : \rho = 0$ VS $H_1 : \rho > 0$ และ $H_1 : \rho < 0$ คือ

$$d_4 = \sum_{t=5}^{n} (e_t - e_{t-4})^2 / \sum_{t=1}^{n} e_t^2$$

แล้วเปรียบเทียบค่าสถิติ d_4 กับค่าวิกฤต $d_{4,L}$, $d_{4,U}$ กล่าวคือ

ก. ปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho > 0$ ณ ระดับนัยสำคัญ 5% เมื่อ $d_4 < d^*_{4,L}$ ไม่อาจปฏิเสธ H_0 ได้ถ้า $d_4 > d^*_{4,U}$ และไม่อาจตัดสินใจได้ถ้า $d^*_{4,L} < d_4 < d^*_{4,U}$

ข. ปฏิเสธ $H_0 : \rho = 0$ VS $H_1 : \rho < 0$ ณ ระดับนัยสำคัญ 5% เมื่อ $d_4 > 4 - d^*_{4,L}$ ไม่อาจปฏิเสธ H_0 ได้ถ้า $d_4 < 4 - d^*_{4,U}$ และไม่อาจตัดสินใจได้ถ้า $4 - d^*_{4,U} < d_4 < 4 - d^*_{4,L}$

---

[1] Wallis, K. F., "Testing For Fourth - Order Autocorrelation in Quarterly Regression Equation", (Econometrica, 40, 1972), P.617-636

วอลลิสเสนอตารางค่าวิกฤตของ d_4 ไว้ 2 ชุด คือตารางที่ 1 สำหรับกรณีที่ไม่มี Quarterly Dummy Variable และตารางที่ 2 สำหรับกรณีที่มี Quarterly Dummy Variable และสมการมีเทอมคงที่ β_1 ตารางทั้งสองนี้ใช้ระดับนัยสำคัญ 5% และจัดไว้เฉพาะงานที่ใช้ตัวแปรอิสระไม่เกิน 5 ตัว และวอลลิส พิสูจน์ให้เห็นว่า ถึงแม้ในสมการจะมี Lagged Dependent Variable คือ Y_{t-1} ปรากฏร่วมเป็นตัวอิสระ ก็ยังคงใช้ตัวสถิติ d_4 ตัดสินใจได้เช่นเดิมโดยมิได้มีปัญหาอะไรที่รุนแรง[1]

7.6 การพยากรณ์

7.6.1 การพยากรณ์ในกรณี GLS

จากสมการของ True Relationship

$$Y_t = \beta_1+\beta_2X_{2t}+\beta_3X_{3t}+\ldots+\beta_kX_{kt}+u_t; \quad t = 1,2,\ldots,n$$

หรือ $$Y = X\beta+U$$

โดยที่ $E(U) = 0$ และ $E(UU') = \Phi \neq \sigma^2 In$

สมมุติว่าเราทราบ True Relationship และทราบค่าของ X's ในอนาคตคือทราบว่า $X_0 = (1, X_{20}, X_{30}, \ldots, X_{k0})$ เราย่อมพยากรณ์ค่าในอนาคตของ Y ได้ดังนี้คือ

$$Y_0 = X_0\beta+u_0$$

หรือ $$Y_0 = \beta_1+\beta_2X_{20}+\beta_3X_{30}+\ldots+\beta_kX_{k0}+u_0$$

โดยที่ u_0 คือ Prediction Disturbance และ

$$E(u_0) = 0, \quad E(u_0^2) = \sigma_0^2$$

และความผันแปรร่วมระหว่าง Prediction Disturbance กับ Disturbance ใน Sample Period คือ

---

[1] **Ibid.**, p. 628

$$E(u_0U) = Eu_0\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = E\begin{bmatrix} u_0u_1 \\ u_0u_2 \\ \vdots \\ u_0u_n \end{bmatrix} = W$$

ปัญหาของนักวิจัยก็คือ เราจะสามารถหา Best Linear Unbiased Predictor (BLUP) ของ Y_0 ได้หรือไม่

การเสาะหา BLUP เราจะอาศัยเทคนิคเดียวกันกับที่ใช้หา BLUE โดยดำเนินการเป็น 2 ระยะดังนี้คือ

(1) สมมุติ Linear Combination ใด ๆ ของ $Y_1, Y_2, \ldots, Y_n$ ขึ้นมาชุดหนึ่งแล้วตรวจสอบดูว่า Linear combination ดังกล่าวเป็น Unbiased Predictor ของ Y_0 หรือไม่ ถ้าเป็นจะต้องเป็นไปภายใต้เงื่อนไขใด

(2) Linear Combination ดังกล่าวเป็น Best Predictor ของ Y หรือไม่ (คือให้ Minimum Variance หรือไม่) ถ้าเป็น Best Predictor จะต้องเป็นไปภายใต้เงื่อนไขใด

ขั้นที่ 1 สมมุติว่า $P = c_1Y_1+c_2Y_2+\ldots+c_nY_n = C'Y$ เป็น Unbiased predictor ของ Y_0 โดยที่ $C = (c_1, c_2, \ldots, c_n)$ เป็นเวคเตอร์ของค่าคงที่ใด ๆ ที่เราต้องการเสาะหาโครงสร้างที่พึงปรารถนา (คือ C มีโครงสร้างที่มีผลให้ P เป็น BLUP ของ Y_0)

เมื่อกำหนดให้ $P = C'Y$ เป็น Unbiased Predictor ของ Y_0 ก็แสดงว่า $E(P) = Y_0$ หรือ $E(P-Y_0) = 0$ แต่เมื่อทดลองตรวจสอบ $E(P-Y_0)$ กลับพบว่า

$$\begin{aligned} E(P-Y_0) &= E(C'Y-Y_0) \\ &= E[C'(X\beta+U)-X_0\beta+u_0] \\ &= E[(C'X-X_0)\beta+C'U-u_0] \\ &= (C'X-X_0)\beta \quad \neq 0 \end{aligned}$$

หรือ $$E(P) = Y_0+(C'X-X_0)\beta \neq Y_0$$

แสดงว่า $E(P-Y_0)$ จะเท่ากับ 0 หรือ P จะเป็น Unbiased Predictor ของ Y_0 (คือ $E(P) = Y_0$) ได้ก็ต่อเมื่อ $(C'X-X_0)\beta = 0$ หรือ $(C'X-X_0) = 0$ เท่านั้น[1]

นั่นคือ $(C'X-X_0) = 0$ คือเงื่อนไขที่ทำให้ $P = C'Y$ เป็น Unbiased Predictor ของ Y_0

ขั้นที่ 2[2] พิจารณา V(p) พบว่า

$$
\begin{aligned}
V(P) &= E[P-E(P)]^2 \\
&= E[P-\{Y_0+(C'X-X_0)\beta\}]^2
\end{aligned}
$$

แต่ $(C'X-X_0) = 0$

ดังนั้น

$$
\begin{aligned}
V(P) &= E(P-Y_0)^2 \\
&= E[(P-Y_0)(P-Y_0)']
\end{aligned}
$$

พิจารณา $P-Y_0$ พบว่า

$$P-Y_0 = C'Y-Y_0 = C'(X\beta+U)-X_0\beta-u_0 = (C'X-X_0)\beta+C'U-u_0$$

แต่ $(C'X-X_0) = 0$

ดังนั้น $P-Y_0 = C'U-u_0$

นั่นคือ

$$
\begin{aligned}
V(P) &= E[(C'U-u_0)(C'U-u_0)'] \\
&= E[(C'U-u_0)(U'C-u_0)] \\
&= E[C'UU'C-C'Uu_0-u_0U'C+u_0^2] \\
&= E[C'UU'C-2C'Uu_0+u_0^2] \\
&= C'E(UU')C-2C'E(Uu_0)+E(u_0^2) \\
&= C'\Phi C-2C'W+\sigma_0^2
\end{aligned}
$$

---

[1] $(C'X-X_0)\beta = 0$ แสดงว่า $(C'X-X_0) = 0$ หรือ $\beta = 0$ อย่างใดอย่างหนึ่งแต่ $\beta \neq 0$ เพราะเป็นพารามิเตอร์สิ่งที่เป็นไปได้ก็คือ $(C'X-X_0) = 0$

[2] ตรวจสอบดูว่า ภายใต้เงื่อนไขว่า $(C'X-X_0) = 0$ นั้น $V(P)$ มีค่าต่ำสุดเมื่อไร ภายใต้ข้อจำกัดใด

นั่นคือ $V(P) = C'\Phi C-2C'W+\sigma_0^2$

ผลการพิสูจน์ทั้งสองตอนนี้เป็นเพียงเครื่องชี้เงื่อนไขว่า $P = C'Y$ จะเป็น Unbiased Predictor ได้ภายใต้เงื่อนไขใด และมีผลให้ V(p) มีรูปแบบอย่างใด แต่ยังไม่อาจให้ค่า C ที่เหมาะสมตามเงื่อนไขคือทำให้ V(p) มีค่าต่ำสุด

ขั้นที่ 3 หาค่าที่เหมาะสมของ C โดยวิธี Lagrange Multiplier Method

จากเงื่อนไข $(C'X-X_0)=0$ และ Objective Function คือ $V(p) = C'\Phi C-2C'W+\sigma_0^2$ จะพบว่า

$$F = V(P)-2(C'X-X_0)\lambda \text{ เมื่อ } \lambda = (\lambda_1,\lambda_2,\ldots,\lambda_k)' \text{ คือ Lagrange multiplier vector}$$

$$= C'\Phi C-2C'W+\sigma_0^2-2(C'X-X_0)\lambda^{1}$$

ค่าที่เหมาะสมของ $C = (c_1,c_2,\ldots,c_n)'$ คือค่าที่ได้จากระบบสมการ

$$\frac{\partial F}{\partial C_i} = 0 \quad ; i = 1,2,\ldots,n \text{ และ } \frac{\partial F}{\partial \lambda_i} = 0 \; ; i = 1,2,\ldots,k$$

หรือ $\frac{\partial F}{\partial C} = 0$, $\frac{\partial F}{\partial \lambda} = 0$ เมื่อ $C = (c_1,c_2,\ldots,c_n)'$, $\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_k)'$

และพบว่า

$$\frac{\partial F}{\partial C} = 0 \Rightarrow 2\Phi C-2W-2X\lambda = 0 \qquad \ldots.(1)$$

$$\frac{\partial F}{\partial \lambda} = 0 \Rightarrow C'X-X_0 = 0 \qquad \ldots.(2)$$

จาก (1) พบว่า $\Phi C - X\lambda = W$ ....(3)

จาก (2) พบว่า $X'C - X_0' = 0$

หรือ $X'C - 0\lambda = X_0'$ ....(4)

---

[1] คูณด้วย 2 เพื่อประโยชน์ในการจัดรูปสมการ

จาก (3) และ (4) จัดรูปแมตริกซ์ได้ดังนี้คือ

$$\begin{bmatrix} \Phi & X \\ X' & 0 \end{bmatrix} \begin{bmatrix} C \\ -\lambda \end{bmatrix} = \begin{bmatrix} W \\ X_0' \end{bmatrix} \quad \ldots. (5)$$

โดยอาศัยความรู้เรื่อง Inverse of Partitioned Matrix[1] พบว่า

$$\hat{C} = \Phi^{-1}[In - X(X'\emptyset^{-1}X)^{-1}X'\Phi^{-1}]W + \Phi^{-1}X(X'\Phi^{-1}X)^{-1}X_0'$$

คือค่าของเวคเตอร์ C ที่เหมาะสมที่มีผลให้ $P = C'Y$ เป็น BLUP

---

[1] ถ้าแบ่งแมตริกซ์ A ขนาด $n \times n$ ออกเป็นแมตริกซ์ย่อย $A_{11}, A_{12}, A_{21}, A_{22}$ โดยที่ A_{11} และ A_{22} ต้องเป็น Square Matrix ดังนี้

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

สมมุติว่า -

$$A^{-1} = \begin{bmatrix} E & F \\ G & H \end{bmatrix}$$ ซึ่งมีผลให้ $AA^{-1} = In$

เราสามารถพิสูจน์ได้ว่า

$$A^{-1} = \begin{bmatrix} E & -E\,A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}E & A_{22}^{-1} + A_{22}^{-1}A_{21}EA_{12}A_{22}^{-1} \end{bmatrix}$$

โดยที่ $E = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1}$

นั่นคือ $\hat{P} = \hat{C}'Y$

$$= [\Phi^{-1}\{In-X(X'\Phi^{-1}X)^{-1}X'\Phi^{-1}\}W]'Y+\{\Phi^{-1}X(X'\Phi^{-1}X)^{-1}X_0'\}'Y$$

$$= [W'\{In-\Phi^{-1}X(X'\Phi^{-1}X)^{-1}X'\}\Phi^{-1}]Y+\{X_0(X'\Phi^{-1}X)^{-1}X'\Phi^{-1}\}Y$$

$$= W'\Phi^{-1}Y-W'\Phi^{-1}X(X'\Phi^{-1}X)^{-1}X'\Phi^{-1}Y+X_0(X'\Phi^{-1}X)^{-1}X'\Phi^{-1}Y$$

แต่ $\hat{\beta}_{GLS} = (X'\Phi^{-1}X)^{-1}X'\Phi^{-1}Y$

ดังนั้น $\hat{P} = W'\Phi^{-1}Y-W'\Phi^{-1}X\hat{\beta}_{GLS}+X_0\hat{\beta}_{GLS}$

$$= X_0\hat{\beta}_{GLS}+W'\Phi^{-1}(Y-X\hat{\beta}_{GLS})$$

$= X_0\hat{\beta}_{GLS}+W^{-1}\Phi e$ โดยที่ $e = Y-X\hat{\beta}_{GLS}$ คือ GLS-Residual

นั่นคือ BLUP ของ Y_0 คือ $\hat{P}$ โดยที่ $\hat{P} = X_0\hat{\beta}_{GLS}+W'\Phi^{-1}e$ ซึ่งเป็นรูปที่ใช้ได้กับทุกกรณี

7.6.2 การพยากรณ์ในกรณีมี Autocorrelation ใน AR (1)

จากกรณีของ Autocorrelation ใน AR (1) เราทราบว่า

$$E(u_t u_{t-s}) = \begin{cases} \rho^s\sigma_u^2 \ ; & s \neq t \\ \sigma_u^2 = \dfrac{\sigma_v^2}{1-\rho^2} \ ; & s = t \end{cases}$$

หรือสามารถเสนอ V(U) ได้ดังนี้

$$V(U) = \sigma_u^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \ldots & \rho^{n-1} \\ \rho & 1 & \rho & \rho^2 \ldots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \rho \ \ldots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \rho^{n-4} \ldots & 1 \end{bmatrix}$$

พิจารณาสูตร $\hat{Y}_0 = X_0\hat{\beta}_{GLS}+W'\phi^{-1}e$ สมมุติว่าเราต้องการพยากรณ์ค่า Y ใน period ที่ 1 นอก Sample Period จะพบว่า

$$\hat{Y}_{n+1} = X_{n+1}\hat{\beta}_{GLS}+W'\phi^{-1}e$$

โดยในที่นี้จะพบว่า

$$X_{n+1} = (1,X_{2,n+1},X_{3,n+1},\ldots,X_{k,n+1}),\hat{\beta} = (\hat{\beta}_1,\hat{\beta}_2,\ldots,\hat{\beta}_k)'$$

$$X_{n+1}\hat{\beta}_{GLS} = \hat{\beta}_1+\hat{\beta}_2X_{2,n+1}+\hat{\beta}_3X_{3,n+1}+\cdots+\hat{\beta}_kX_{k,n+1}$$

$$W = E\begin{bmatrix} u_1u_{n+1} \\ u_2u_{n+1} \\ u_3u_{n+1} \\ \vdots \\ u_nu_{n+1} \end{bmatrix} = E\begin{bmatrix} u_{n+1}u_{n+1-n} \\ u_{n+1}u_{n+1-(n-1)} \\ u_{n+1}u_{n+1-(n-2)} \\ \vdots \\ u_{n+1}u_{n+1-1} \end{bmatrix} = \begin{bmatrix} \rho^n\sigma_u^2 \\ \rho^{n-1}\sigma_u^2 \\ \rho^{n-2}\sigma_u^2 \\ \vdots \\ \rho\ \sigma_u^2 \end{bmatrix} = \rho\sigma_u^2\begin{bmatrix} \rho^{n-1} \\ \rho^{n-2} \\ \rho^{n-3} \\ \vdots \\ \rho \\ 1 \end{bmatrix}$$

แสดงว่า W สำหรับการพยากรณ์ไปในอนาคต 1 วาระ มีค่าเท่ากับ ρ เท่าของสดมภ์สุดท้ายของแมตริกซ์ V(U) และสามารถพิสูจน์ได้ว่า W สำหรับการพยากรณ์ไปในอนาคต r วาระมีค่าเท่ากับ ρ^r เท่าของสดมภ์สุดท้ายของ V(U) และพบว่าในการพยากรณ์ Y_{n+1} นั้น $W'\phi^{-1}e$ คือ

$$\rho\sigma_u^2(\rho^{n-1},\rho^{n-2},\rho^{n-3},\ldots,\rho,1)\frac{1}{\sigma_u^2}\frac{1}{1-\rho^2}\begin{bmatrix} 1 & -\rho & 0 & \ldots & 0 & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & \ldots & 0 & 0 & 0 \\ 0 & -\rho & 1+\rho^2 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -\rho & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & \ldots & 0 & -\rho & 1 \end{bmatrix}\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}$$

$$= \frac{\rho\ \sigma_u^2}{(1-\rho^2)\sigma_u^2}\ (0,0,0,\ldots,1-\rho^2)\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}$$

$$= \frac{\rho}{1-\rho^2}(1-\rho^2)e_n$$

$$= \rho e_n$$

ดังนั้น $\hat{Y}_{n+1} = X_{n+1}\hat{\beta} + \rho\ e_n$ และในกรณีที่ไม่ทราบค่า ρ ให้ใช้ค่าของ ρ คือ $\hat{\rho}$ แทน

ดังนั้น $\hat{Y}_{n+1} = X_{n+1}\hat{\beta} + \rho\ e_n$ คือ BLUP ของ Y_{n+1}

หรือ $\hat{Y}_{n+1} = (\hat{\beta}_1+\hat{\beta}_2X_{2,n+1}+\hat{\beta}_3X_{3,n+1}+\ldots+\hat{\beta}_kX_{k,n+1})+\hat{\rho}e_n$

สำหรับการพยากรณ์ค่าของ Y ในวาระใด ๆ (สมมุติว่าเป็นวาระที่ r) คือ $\hat{Y}_{n+r}$ นอก Sample Period ก็พัฒนาได้ในทำนองเดียวกัน และเราสามารถพิสูจน์ได้โดยง่ายว่า

$$\hat{Y}_{n+r} = X_{n+r}\hat{\beta}+\rho^r e_n \ ;\ r=1,2,3,\ldots \quad \text{คือ BLUP ของ } Y_{n+r}$$

หมายเหตุ ถ้า $V(U) = \sigma_u^2 In$ หรือ $E(u_iu_j)=0\ ;\ i \neq j$ จะพบว่า $W = 0$ ซึ่งมีผลให้ $W'\phi^{-1}e = 0(\sigma_u^2 In)e = 0$ และ BLUP ของ Y_{n+r} คือ $X_{n+r}\hat{\beta}$