การตรวจเอกสาร

วิธีประมาณค่าพารามิเตอร์นั้นมีหลายวิธีด้วยกัน ในการศึกษาวิจัยครั้งนี้ใช้วิธีการหาค่าที่เหมาะสมซึ่งอาศัยหลักการทางคณิตศาสตร์เพื่อเปรียบเทียบกับวิธีที่อาศัยหลักการทางสถิติ

การวิเคราะห์อนุกรมเวลาที่ไม่เป็นอิสระจากกัน

การศึกษาอนุกรมเวลาที่ค่าสังเกตไม่เป็นอิสระจากกันเริ่มในช่วงต้นคริสต์ศตวรรษที่ 19 โดยใน ค.ศ. 1923 ได้มีการพิจารณากระบวนการ Markov ซึ่งก็คือกระบวนการถดถอยบนตัวเอง (Autoregressive Process) ที่มีลำดับเป็น 1 และในปี ค.ศ. 1927 George Udny Yule ได้วิเคราะห์ข้อมูลที่เป็นกระบวนการถดถอยบนตัวเองที่มีลำดับเป็น 2 ในปี ค.ศ. 1931 Gilbert Walker ได้พิจารณาข้อมูลที่เป็นกระบวนการถดถอยบนตัวเองที่มีลำดับเป็น 4

หลังจากนั้นการศึกษาอนุกรมเวลาชนิดนี้มีการพัฒนาไปมากในปี ค.ศ. 1970 โดย Box และ Jenkins ได้เสนอระเบียบวิธีใช้วิเคราะห์อนุกรมเวลาที่ค่าสังเกตไม่เป็นอิสระจากกัน (วิชิต, 2548) โดยใช้ความสัมพันธ์ระหว่างค่าสังเกตในอนุกรมเวลาที่เป็นฟังก์ชั่นสหสัมพันธ์ในตัวเอง (autocorrelation function) เพื่อที่จะมาช่วยในการพยากรณ์ อนุกรมเวลาที่ใช้เทคนิคของ Box-Jenkins ในการพยากรณ์ จะเป็นอนุกรมเวลาในกระบวนการความน่าจะเป็นเชิงเส้นที่ไม่ต่อเนื่อง (discrete linear stochastic process) โดยเริ่มจากการพิจารณาอนุกรมเวลาที่มาจากกระบวนการที่มีสเตชั่นนารี นั่นคือ ถ้า X_t เป็นชุดของอนุกรมเวลาที่มีการแจกแจงแบบตัวแปรเดียว (univariate distribution) แต่มีสหสัมพันธ์กัน อนุกรม X_t จะมีการแจกแจงค่าความถี่ไม่เปลี่ยนแปลงในทุกช่วงเวลา t และทุกๆ ค่า j ที่ต่อเนื่องเป็นชุดกันของ $X_{t+1}, X_{t+2}, \ldots, X_{t+j}$ และมีชนิดของการแจกแจงแบบเดียว นั่นคือการแจกแจงความน่าจะเป็น $f(x)$ ไม่เปลี่ยนแปลงสำหรับทุกค่าของ t และชนิดของการแจกแจงความน่าจะเป็นก็ไม่เปลี่ยนแปลง ค่าคาดหมายของกระบวนการสเตชั่นนารี่เป็นค่าคงที่ และกระบวนการมีการแกว่งไปมารอบๆค่าคาดหมายด้วยความแปรปรวนที่คงที่ ซึ่งจะได้ค่าคาดหมายของกระบวนการเป็น

$$E(X_t) = \mu = \int_{-\infty}^{\infty} X_t \quad f(X_t) \; dx$$

และมีค่าความแปรปรวนของกระบวนการสเตชันนารีดังนี้

$$E(X_t-\mu)^2 \;=\; \delta_x^2 = \int_{-\infty}^{\infty}(X-\mu)^2 \;\; f(x)dx$$

ซึ่งค่าความแปรปรวนร่วมในตัวเองระหว่าง X_t และ X_{t+j} เมื่อ $j > 0$ มีค่าเป็น

$$Cov(X_t, X_{t+j}) \;=\; \gamma_j = E(X_t-\mu)(X_{t+j}-\mu)$$

สหสัมพันธ์ในตัวเอง (ρ) ของค่าสังเกตที่ตามกัน j หน่วยเวลามีค่าเท่ากับ

$$\rho_j \;=\; \frac{\gamma_j}{\gamma_0} = \frac{\gamma_j}{\delta_x^2}$$

และมีค่า $-1 \le \rho_j \le 1$ ถ้า m เป็นค่าเฉลี่ยของชุดอนุกรมนี้ซึ่งเป็นค่าจำกัด ที่มีค่าสังเกต n เทอม และ n อาจมีค่าเป็นอนันต์ได้ จะได้ว่า (Kendall, 1976)

$$\lim_{n\to\infty}\frac{1}{n}\;\sum_{j=1}^{n}\rho_j = 0$$

แล้ว $m = \mu$ ซึ่งในกรณีทั่วไปของสมการข้างบน ρ_k จะลดลงจนเป็นศูนย์อย่างรวดเร็วหลังจากจุด k และค่าเฉลี่ยของ ρ_k มีแนวโน้มเข้าหาศูนย์

พิจารณาอนุกรมเวลาซึ่งมีค่าสังเกตเป็นตัวแปรสุ่มที่เป็นอิสระซึ่งอยู่ในรูปผลบวกเชิงเส้นของ $\in_t, \in_{t-1}, \in_{t-2}, \ldots$ ซึ่งมีการแจกแจงความน่าจะเป็นที่คงที่ และมีค่าเฉลี่ยเป็นศูนย์ ค่าความแปรปรวนคงที่ ($\delta_\in^2$) และมักจะสมมติให้การแจกแจงความน่าจะเป็นของ $\in_t$ เป็นแบบปกติ (normal distribution) ตัวแปรสุ่ม $\in_t, \in_{t-1}, \in_{t-2}, \ldots$, นี้เรียกว่ากระบวนการไวท์น้อยส์ (White noise process) ซึ่งผลบวกเชิงเส้น (Montgomery, 1990) ของ $\{\in_i\}$ คือ

$$x_t = \mu + \psi_0 \in_t + \psi_1 \in_{t-1} + \psi_2 \in_{t-2} + \ldots.. \qquad (1)$$

หรือ

$$x_t = \mu + \sum_{j=0}^{\infty}\Psi_j \in_{t-j}$$

เมื่อ

ψ_j เรียกว่าค่าถ่วงน้ำหนัก (weight)ของพจน์ที่ j เมื่อ $j = 0, 1, 2, \ldots$

μ คือ ค่าคงที่แสดงระดับที่อยู่ (level) ของกระบวนการ

โดยทั่วไป ψ_0 = 1 และจะเขียนสมการ (1) ได้ใหม่ในเทอมของการแทนค่าแบบย้อนกลับ B (backward shift operator B) ได้ดังนี้

$$\mathrm{B} \in_t = \in_{t-1}$$

หรือ

$$\mathrm{B}^{\mathrm{j}} \in_t = \in_{t-j}$$

แทนค่าลงในสมการ (1) จะได้

$$x_t = \mu + (\psi_0 B^0 + \psi_1 B^1 + \psi_2 B^2 + ...) \in_t$$

หรือ

$$x_t = \mu + \Psi(B) \in_t$$

เมื่อ

$$\Psi(\mathrm{B}) = \Psi_0 B^0 + \Psi_1 B^1 + \Psi_2 B^2 + ... \qquad \text{และ } \Psi_0 = 1$$

สมการ(1) นี้เรียกว่า ตัวกรองเชิงเส้น (linear filter) จากสมการข้างต้นจะเห็นได้ว่า ค่าสังเกตในอนุกรมเวลา x_t นั้นไม่เป็นอิสระจากกันเพราะค่าสังเกต x_t จะหาได้จากค่า$\in_{\mathrm{i}}$ในช่วงเวลาก่อนหน้านั้น ถ้า$\in_{\mathrm{i}}$มีการแจกแจงแบบปกติแล้ว x_t ก็จะมีการแจกแจงแบบปกติด้วย

ถ้าพิจารณาจากตัวแบบตัวกรองเชิงเส้นแล้วก็อาจให้นิยามกับตัวแบบอนุกรมเวลาว่าเป็นฟังก์ชันที่ได้จากการแปลงเทอมของกระบวนการไวท์น้อยส์ ของข้อมูลอนุกรมเวลานั้น Box-Jenkins ได้รวบรวมตัวแบบเหล่านี้ไว้และพัฒนาการใช้งานตัวแบบในลักษณะนี้ ดังนั้นตัวแบบที่ได้จากสมการตัวกรองเชิงเส้นมักจะเรียกว่าตัวแบบของ Box-Jenkins

ตัวแบบที่ได้จากสมการตัวกรองเชิงเส้นที่ใช้กับอนุกรมเวลาที่เป็นสเตชันนารี นั้นก็คาดว่าค่าเฉลี่ยคงที่ (constant mean) จะมีการแกว่งอย่างสุ่ม

<u>กระบวนการถดถอยบนตัวเอง (Autoregressive Process)</u>

สมการตัวกรองเชิงเส้นถูกพัฒนาเป็นตัวแบบที่กระชับ ซึ่งเหมาะที่จะอธิบายอนุกรมเวลาที่มีสหสัมพันธ์กัน

กระบวนการถดถอยบนตัวเองเป็นกรณีหนึ่งที่พัฒนาจากตัวกรองเชิงเส้น ซึ่งอาจเขียนได้เป็น (Montgomery, 1990)

$$x_t = \xi + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} + \in_t \qquad (2)$$

โดย $\phi_1, \phi_2, \ldots, \phi_p$ เป็นพารามิเตอร์ จำนวน p ค่า และ ξ คือ ค่าคงที่ ส่วน $\in_t$ คือ กระบวนการไวท์น้อยส์ (white noise process) ณ เวลา t

ค่าสังเกต ณ เวลาปัจจุบัน (x_t) นั้นจะ"ถดถอย"บนค่าสังเกตที่เวลาก่อนหน้ามันคือ x_{t-1}, x_{t-2}, …, x_{t-p} กระบวนการถดถอยบนตัวเองนี้มีลำดับ เป็น p หรือเขียนเป็นสัญลักษณ์ย่อๆ ว่า AR(p)

ซึ่งสามารถเขียนกระบวนการถดถอยบนตัวเองนี้โดยใช้ตัวกรองเชิงเส้นด้วยการแทนค่า x_{t-1}, x_{t-2} ,.. ลงในสมการ (2) โดยค่า x_{t-1}, x_{t-2} มีค่าดังนี้คือ

$$x_{t-1} = \xi + \phi_1 x_{t-2} + \phi_2 x_{t-3} + \ldots + \phi_p x_{t-p-1} + \in_{t-1}$$

และ

$$x_{t-2} = \xi + \phi_1 x_{t-3} + \phi_2 x_{t-4} + \ldots + \phi_p x_{t-p-2} + \in_{t-2}$$

ตอนนี้ค่า x_{t-1} และ x_{t-2} ก็จะถูกกำจัดไป พจน์อื่นๆถัดจาก x_{t-2} ก็ทำในทำนองเดียวกัน และสามารถเขียน AR(p) ในรูปการแทนค่าแบบย้อนกลับได้ดังนี้

$$x_t = \xi + (\phi_1 B^1 + \phi_2 B^2 + \ldots\ldots \phi_p B^p) x_t + \in_t$$

หรือ

$$x_t (1 - \phi_1 B^1 - \phi_2 B^2 - \ldots. - \phi_p B^p) x_t = \xi + \in_t \qquad (3)$$

ถ้าให้

$$\phi_p(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - \ldots. - \phi_p B^p$$

จะได้สมการ (3) เป็น

$$\phi_p(B)x_t = \xi + \in_t$$

ถ้าให้ $\tilde{x}_t = x_t - \mu$ สำหรับทุกช่วงเวลา t จะได้ AR(p) เป็น

$$\phi_p(B)\tilde{x}_t = \in_t$$

นั่นคือค่าเฉลี่ย (μ) ของ AR(p) มีค่าเท่ากับ

$$\mu = \frac{\xi}{1 - \sum_{j=1}^{p} \phi_j}$$

ตัวแบบถดถอยบนตัวเองนี้ประยุกต์ใช้กับอนุกรมเวลาที่สเตชั่นนารี ถ้าค่ารากของโพลี่โน-เมียล $\phi_p B = 0$ อยู่นอกวงกลมหนึ่งหน่วย (unit circle) ก็จะได้อนุกรมเวลาที่เป็นแบบสเตชั่นนารี ซึ่งเงื่อนไขนี้ได้จากค่า $\sum_{j=0}^{\infty} \psi_j$ จะต้องมีการลู่เข้า

<u>ขั้นตอนในการพยากรณ์ด้วยตัวแบบ AR(p)</u>

ขั้นตอนในการพยากรณ์ตัวแบบ AR(p) นี้สรุปตามวิธีของ Box-Jenkins ได้เป็นสี่ขั้นตอนคือ

1. การวิเคราะห์ค่าสหสัมพันธ์ในตัวเองและสหสัมพันธ์ในตัวเองบางส่วนของข้อมูลอนุกรมเวลา
2. เลือกตัวแบบจากผลการวิเคราะห์ข้อมูลที่ได้
3. การประมาณค่าพารามิเตอร์ในตัวแบบ
4. ตรวจสอบความถูกต้องของตัวแบบด้วยการนำตัวแบบที่ได้ไปทำการพยากรณ์และวิเคราะห์ตัวแบบจากค่าเศษส่วนเหลือ (residual) จากการพยากรณ์