บทที่ 3

กรอบแนวคิดทางทฤษฎี

สำหรับการเปรียบเทียบความแม่นยำในการพยากรณ์ราคาน้ำมันดิบดูไบในครั้งนี้ จะสร้างระบบพยากรณ์โดยใช้ตัวแบบ 3 ประเภทคือตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) ตัวแบบการ์ช (GARCH Model) และตัวแบบโครงข่ายใยประสาทประดิษฐ์ (ANNs Model) ซึ่งแนวคิดและทฤษฎีที่เกี่ยวข้องกับการสร้างระบบพยากรณ์ด้วยตัวแบบทั้งสาม มีรายละเอียดดังต่อไปนี้

3.1 การพยากรณ์ด้วยตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model)

ตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) ถูกพัฒนาโดย George Box และ Gwilym Jenkins (1976) ซึ่งตัวแบบนี้ได้รับความนิยม และเหมาะสมสำหรับการพยากรณ์ไปข้างหน้าในช่วงเวลาสั้น ๆ เนื่องจากตัวแบบดังกล่าวให้ค่าเฉลี่ยของค่าคลาดเคลื่อนกำลังสอง (Mean Square Error : MSE) ของการพยากรณ์ค่อนข้างต่ำ

โดยที่ Spyros Makridakis, et al (1998) เสนอว่าการพยากรณ์โดยใช้ตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) มี 5 ขั้นตอน คือ การเตรียมข้อมูล (Data Preparation) การเลือกตัวแบบ (Model Selection) การประมาณค่าพารามิเตอร์ (Parameter Estimation) การตรวจสอบความถูกต้อง (Model Checking) และการพยากรณ์ (Forecasting) ซึ่งขั้นตอนต่างๆ มีรายละเอียดดังนี้

3.1.1 การเตรียมข้อมูล (Data Preparation)

การพยากรณ์ด้วยตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) มีข้อกำหนดที่สำคัญคือ อนุกรมเวลาที่พิจารณาต้องมีคุณสมบัติหยุดนิ่ง (Stationary) ซึ่งคุณสมบัติดังกล่าวมีเงื่อนไขว่าอนุกรมเวลา (Z_t) ที่พิจารณาต้องมีการกระจายตัวรอบ ๆ ค่าคงที่ ค่าใดค่าหนึ่ง [$E(Z_t) = \mu$] โดยมีขอบเขตที่จำกัด [$E(Z_t - \mu)^2 = \sigma^2$] และความแปรปรวนร่วม (Covariance) ขึ้นอยู่กับระยะห่างของข้อมูลเท่านั้นไม่ได้ขึ้นกับเวลา [$E(Z_t - \mu)(Z_{t-k} - \mu) = \gamma_k$] กล่าวคือคุณสมบัติทางสถิติของอนุกรมเวลาในช่วงใดช่วงหนึ่ง จะคงอยู่และไม่เปลี่ยนแปลงไปตามเวลา ซึ่งคุณสมบัติหยุดนิ่ง (Stationary) ของอนุกรมเวลา สามารถตรวจสอบได้จากการทดสอบดิกกี้-ฟุลเลอร์ฉบับขยาย (Augmented Dickey-Fuller : ADF Test)

3.1.1.1 การทดสอบดิกกี้-ฟุลเลอร์ฉบับขยาย (Augmented Dickey-Fuller : ADF Test)

การทดสอบดิกกี้-ฟุลเลอร์ฉบับขยายถูกนำเสนอโดย David Dickey และ Waynes Fuller (1979) เพื่อตรวจสอบว่าอนุกรมเวลา (Z_t) ที่พิจารณามีคุณสมบัติหยุดนิ่ง (Stationary) หรือไม่ โดยกำหนดสมมติฐานว่าง (Null Hypothesis) ว่าอนุกรมเวลาที่พิจารณามียูนิตรูท [๑] (Unit Root)

ซึ่งการตรวจสอบคุณสมบัติหยุดนิ่ง (Stationary) ของอนุกรมเวลา จะต้องพิจารณารายละเอียด 3 แนวทางตามลำดับคือ

1) อนุกรมเวลา (Z_t) มีค่าแนวโน้มและจุดตัดแกน (Trend and Intercept) กล่าวคือเป็นแนวเดินสุ่มที่มีความโน้มเอียงทั่วไปและมีค่าแนวโน้มตามเวลาเชิงเส้น (Random Walk with Drift and Linear Time Trend) ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

$$\Delta Z_t = \alpha + \beta + \gamma Z_t + \phi_1 \Delta Z_{t-1} + \phi_2 \Delta Z_{t-2} + \ldots + \phi_p \Delta Z_{t-p} + \varepsilon_t$$

2) อนุกรมเวลา (Z_t) มีจุดตัดแกน (Intercept) กล่าวคือเป็นแนวเดินสุ่มที่มีความโน้มเอียงทั่วไป (Random Walk with Drift) สามารถเขียนเป็นสมการได้ดังนี้

$$\Delta Z_t = \alpha + \gamma Z_t + \phi_1 \Delta Z_{t-1} + \phi_2 \Delta Z_{t-2} + \ldots + \phi_p \Delta Z_{t-p} + \varepsilon_t$$

3) อนุกรมเวลา (Z_t) ไม่มีค่าแนวโน้มและจุดตัดแกน (None Trend and Intercept) กล่าวคือเป็นแนวเดินสุ่มที่ปราศจากความโน้มเอียงทั่วไปและค่าแนวโน้มตามเวลาเชิงเส้น (Random Walk without Drift and Linear Time Trend) สามารถเขียนเป็นสมการได้ดังนี้

$$\Delta Z_t = \gamma Z_t + \phi_1 \Delta Z_{t-1} + \phi_2 \Delta Z_{t-2} + \ldots + \phi_p \Delta Z_{t-p} + \varepsilon_t$$

โดยที่ α คือ พจน์คงตัว (Constant Term)

β คือ สัมประสิทธิ์ (Coefficient) ของค่าแนวโน้มตามเวลาเชิงเส้น (Linear Time Trend)

---

[๑] ในกระบวนการเฟ้นสุ่มเชิงเส้น (Linear Stochastic Process) ค่าราก (Root) หรือคำตอบของสมการลักษณะเฉพาะ (Characteristic Equation) จะมีค่าเท่ากับ 1 ($\gamma = 1$) ถ้ากระบวนการดังกล่าวเป็นกระบวนการที่ขาดคุณสมบัติหยุดนิ่ง (Non-Stationary)

ϕ_p คือ พารามิเตอร์อัตถดถอย (Autoregressive) อันดับที่ p

γ คือ สัมประสิทธิ์ (Coefficient) ของค่าสังเกต ณ เวลา t

ε_t คือ ตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance)

สำหรับการตรวจสอบสมมติฐาน (Hypothesis) จะใช้ค่าสถิติที (t-Statistic) ที่มีสูตรการคำนวณดังนี้

$$t_{\hat{\gamma}} = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$$

โดยที่ $\hat{\gamma}$ คือ ค่าประมาณของ γ

$SE(\hat{\gamma})$ คือ สัมประสิทธิ์ (Coefficient) ความเคลื่อนคลาดมาตรฐาน (Standard Error)

ซึ่งอนุกรมเวลาที่มีคุณสมบัติหยุดนิ่ง (Stationary) จะต้องปฏิเสธสมมติฐานว่าง (Null Hypothesis) ที่ว่าอนุกรมเวลามียูนิตรูท (Unit Root) กล่าวคือค่าสถิติที (t-Statistic) ที่คำนวณได้ จะต้องมีค่าน้อยกว่าค่าวิกฤตแมคคินนอน (MacKinnon Critical Value)

3.1.1.2 กระบวนการมีอันดับ (Integrated Process : I_d)

เมื่อตรวจสอบพบว่าอนุกรมเวลาที่พิจารณาขาดคุณสมบัติหยุดนิ่ง (Non-Stationary) จะต้องแปลงข้อมูล (Transformation) ดังกล่าวให้มีคุณสมบัติหยุดนิ่งเสียก่อน โดยใช้กระบวนการมีอันดับ (Integrated Process : I_d) ซึ่งเป็นการกรองเอาคุณลักษณะที่หยุดนิ่งของอนุกรมเวลาออกมาด้วยการหาผลต่าง (Differencing) ระหว่างข้อมูล ณ เวลาปัจจุบันกับข้อมูลเมื่อถอยหลังไป d คาบเวลา ซึ่งสามารถเขียนในรูปสมการได้ดังนี้

$$I_d \text{ คือ } \Delta_d Z_t = \Delta_{d-1}(Z_t - Z_{t-1})$$

$$I_d = (1-B)_d Z_t$$

สำหรับการแปลงข้อมูลด้วยการหาผลต่างครั้งที่หนึ่ง (First Differencing : I_1) สามารถเขียนเป็นรูปแบบสมการได้ดังนี้

$$I_1 \text{ คือ } \Delta_1 Z_t = (Z_t - Z_{t-1})$$

$$I_1 = (1-B)Z_t$$

หากข้อมูลผ่านการหาผลต่างครั้งที่หนึ่ง (First Differencing : I_1) แล้ว แต่เมื่อนำไปตรวจสอบด้วยค่าสถิติเอดีเอฟ (ADF Test) ยังพบว่าอนุกรมดังกล่าวมียูนิตรูท (Unit Root) อยู่ ดังนั้นจะต้องใช้การหาผลต่างครั้งที่สอง (Second Differencing : I_2) ซึ่งเขียนเป็นรูปแบบสมการได้ดังนี้

$$I_2 \text{ คือ } \Delta_2 Z_t = \Delta_1 (Z_t - Z_{t-1})$$

$$I_2 = (1-B)_2 Z_t$$

โดยที่ B คือ ตัวดำเนินการเลื่อนเวลาย้อนหลัง (Backward Shift Operator)

$(1-B)_d Z_t$ คือ ผลต่างอันดับที่ d

3.1.2 การเลือกตัวแบบ (Model Selection)

เมื่ออนุกรมเวลาที่พิจารณามีคุณสมบัติหยุดนิ่ง (Stationary) แล้ว ขั้นตอนต่อไปคือการหากระบวนการอัตถดถอย (Autoregressive : AR) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average : MA) ที่เหมาะสมให้กับอนุกรมเวลา

3.1.2.1 กระบวนการอัตถดถอย (Autoregressive : AR)

กระบวนการอัตถดถอย (Autoregressive) ถูกพัฒนาโดย George Yule (1927) ซึ่งมีแนวคิดว่าค่าสังเกตที่ต้องการพยากรณ์ ขึ้นกับค่าสังเกตในอดีต p หน่วยเวลาย้อนหลัง (โดยที่ p = 1, 2, 3, ...) และตัวรบกวนเฟ้นสุ่ม [๒] (Stochastic Disturbance) ณ เวลาปัจจุบัน กล่าวคือ กระบวนการอัตถดถอยอันดับที่ p (AR(p)) เป็นการอธิบาย Z_t ในเทอมของ $Z_{t-1}, Z_{t-2}, \ldots, Z_{t-p}$ และ ε_t ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

---

[๒] ตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance) เป็นตัวแปรของตัวแบบทางคณิตศาสตร์ (Mathematical Model) ที่ถูกสร้างขึ้นเพื่อชดเชยความสัมพันธ์ที่ไม่สมบูรณ์ (Incomplete Relationship) ระหว่างตัวแปรอิสระ (Independent Variable) และตัวแปรตาม (Dependent Variable) ซึ่งการวิเคราะห์การถดถอย (Regression Analysis) บางครั้งเรียกตัวแปรดังกล่าวแตกต่างกันไป เช่น พจน์คลาดเคลื่อน (Error Term) ส่วนตกค้าง (Residual) เศษเหลือ (Remainder) ตัวเสริม (Innovation) ฯลฯ

$$\text{AR(p) คือ} \quad Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t$$

โดยที่ μ คือ พจน์คงตัว (Constant Term)

ϕ_p คือ พารามิเตอร์อัตถดถอยตัวที่ p

ε_t คือ ตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance)

สำหรับกรณีของกระบวนการอัตถดถอยอันดับที่หนึ่ง (AR(1)) คือการพิจารณาให้ค่าปัจจุบันขึ้นอยู่กับค่าอดีตย้อนหลังไป 1 คาบเวลาและตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance) ณ เวลาปัจจุบัน ซึ่งเขียนเป็นรูปแบบสมการได้ดังนี้

$$Z_t = \mu + \phi_1 Z_{t-1} + \varepsilon_t$$

ในกรณีของกระบวนการอัตถดถอยอันดับที่สอง (AR(2)) จะเป็นการพิจารณาให้ค่าปัจจุบันขึ้นอยู่กับค่าอดีตย้อนหลังไป 2 คาบเวลาและตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance) ณ เวลาปัจจุบัน เขียนรูปแบบสมการได้ดังนี้

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \varepsilon_t$$

3.1.2.2 กระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average : MA)

สำหรับกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average) ถูกนำมาใช้เป็นครั้งแรกโดย Evgeny Slutzky (1936) ซึ่งมีแนวคิดว่าค่าปัจจุบัน (Current Value) ของอนุกรมเวลาชุดหนึ่งๆ สามารถอธิบายได้ด้วยความผิดปกติเชิงสุ่ม (Random Shock) หรือตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance) ในอดีต q หน่วยเวลาย้อนหลัง (โดยที่ q = 1, 2, 3, ...) และตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance) ณ เวลาปัจจุบัน กล่าวคือกระบวนการค่าเฉลี่ยเคลื่อนที่อันดับที่ q (MA(q)) เป็นการอธิบาย Z_t ในเทอมของ ε_{t-1} , ε_{t-2} , ... , ε_{t-q} และ ε_t ซึ่งสามารถเขียนในรูปสมการได้ดังนี้

$$\text{MA(q) คือ} \quad Z_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

โดยที่ μ คือ พจน์คงตัว (Constant Term)

θ_q คือ พารามิเตอร์ค่าเฉลี่ยเคลื่อนที่ตัวที่ q

ε_t คือ ตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance)

สำหรับกรณีของกระบวนการค่าเฉลี่ยเคลื่อนที่อันดับที่ 1 (MA(1)) หรือการพิจารณาให้ค่าปัจจุบันขึ้นอยู่กับตัวรบกวนเฟ้นสุ่มในอดีตย้อนหลังไป 1 คาบเวลาและตัวรบกวนเฟ้นสุ่ม ณ เวลาปัจจุบัน ซึ่งเขียนรูปแบบสมการได้ดังนี้

$$Z_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

สำหรับกรณีของกระบวนการค่าเฉลี่ยเคลื่อนที่อันดับที่ 2 (MA(2)) หรือการพิจารณาให้ค่าปัจจุบันขึ้นอยู่กับพจน์คลาดเคลื่อนในอดีตย้อนหลังไป 2 คาบเวลาและพจน์คลาดเคลื่อน ณ เวลาปัจจุบัน ซึ่งเขียนรูปแบบสมการได้ดังนี้

$$Z_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

โดยการกำหนดระยะความล้าช้า (Lag Length) ที่เหมาะสมให้กับกระบวนการอัตถดถอย (Autoregressive : AR) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average : MA) จะใช้ฟังก์ชันอัตสหสัมพันธ์ (Autocorrelation Function : ACF) ร่วมกับฟังก์ชันอัตสหสัมพันธ์บางส่วน (Partial Autocorrelation Function : PACF) ในการพิจารณา

3.1.2.3 ฟังก์ชันอัตสหสัมพันธ์ (Autocorrelation Function : ACF)

ฟังก์ชันอัตสหสัมพันธ์เป็นมาตรวัดสหสัมพันธ์ (Correlation) ระหว่างค่าสังเกต (Z_t) ในอนุกรมเวลาที่อยู่ห่างกัน k ช่วงเวลา กล่าวคือเป็นการวัดสหสัมพันธ์ระหว่าง Z_t และ Z_{t-k} โดยที่สหสัมพันธ์ดังกล่าวไม่ได้ขึ้นอยู่กับตัวแปรเวลา t แต่ขึ้นอยู่กับระยะที่ห่างกัน k หน่วยเท่านั้น ซึ่งฟังก์ชันอัตสหสัมพันธ์มีสูตรการคำนวณดังนี้

$$\rho_k = \frac{\sum_{t-k-1}^{n} (Z_t - \bar{Z})(Z_{t-k} - \bar{Z})}{\sum_{t-k-1}^{n} (Z_t - \bar{Z})^2}$$

3.1.2.4 ฟังก์ชันอัตสหสัมพันธ์บางส่วน (Partial Autocorrelation Function : PACF)

ในการพิจารณาสหสัมพันธ์ (Correlation) ระหว่างตัวแปร Z_t และ Z_{t-k} มีความเป็นไปได้ว่าสหสัมพันธ์ดังกล่าวอาจเป็นผลสืบเนื่องมาจากสหสัมพันธ์ของ $Z_{t-1}, Z_{t-2}, \ldots, Z_{t-k}$ ด้วย ดังนั้นเพื่อขจัดขจัดอิทธิพลของความล่าช้า (Lag) ทางเวลาอื่น ๆ ที่ไม่เกี่ยวข้องในการหาสหสัมพันธ์ระหว่าง Z_t และ Z_{t-k} จึงใช้ฟังก์ชันอัตสหสัมพันธ์บางส่วน (Partial Autocorrelation Function : ρ_{kk}) เป็นมาตรวัดสหสัมพันธ์แทน ซึ่งสามารถเขียนเป็นสมการได้ดังนี้

$$\rho_{kk} = \begin{cases} \rho_{kk} & \text{เมื่อ } k = 1 \\ \dfrac{\rho_k - \rho_{k-1}}{1 - \rho_{k-1}} & \text{เมื่อ } k = 2, 3, \ldots, n \end{cases}$$

สำหรับรูปแบบการพยากรณ์ที่เหมาะสมควรเน้นหลักของการมัธยัสถ์ (Parsimonious) กล่าวคือ พารามิเตอร์ของตัวแบบต้องสามารถอธิบายการเคลื่อนไหวของอนุกรมเวลาได้เป็นอย่างดี โดยใช้จำนวนพารามิเตอร์น้อยที่สุด ซึ่งโดยทั่วไปจะใช้กระบวนการอัตถดถอย (Autoregressive) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average) ไม่เกินอันดับ (Order) ที่สอง

ตารางที่ 3-1 หลักการเลือก AR(p) และ MA(q) โดยพิจารณาจากลักษณะของ ACF และ PACF

กระบวนการ	ACF	PACF
AR(1)	ค่า ACF ลดลงเข้าสู่ศูนย์อย่างรวดเร็วแบบเลขชี้กำลัง (Exponential)	ค่า PACF ณ ความล่าช้าที่ 1 ไม่เป็นศูนย์ แต่ค่า PACF ณ ความล่าช้าอื่นๆ เป็นศูนย์
AR(2)	ค่า ACF ลดลงเข้าสู่ศูนย์อย่างรวดเร็วแบบเลขชี้กำลัง (Exponential)	ค่า PACF ณ ความล่าช้าที่ 1 และ 2 ไม่เป็นศูนย์ แต่ค่า PACF ณ ความล่าช้าอื่นๆ เป็นศูนย์
MA(1)	ค่า ACF ณ ความล่าช้าที่ 1 ไม่เป็นศูนย์ แต่ค่า ACF ณ ความล่าช้าอื่นๆ เป็นศูนย์	ค่า PACF ลดลงเข้าสู่ศูนย์อย่างรวดเร็วแบบเลขชี้กำลัง (Exponential)
MA(2)	ค่า ACF ณ ความล่าช้าที่ 1 และ 2 ไม่เป็นศูนย์ แต่ค่า ACF ณ ความล่าช้าอื่นๆ เป็นศูนย์	ค่า PACF ลดลงเข้าสู่ศูนย์อย่างรวดเร็วแบบเลขชี้กำลัง (Exponential)

3.1.2.5 กระบวนการอาร์มา (ARMA Process)

หากอนุกรมเวลาที่พิจารณามีคุณสมบัติหยุดนิ่ง (Stationary) ณ ระดับ (At Level) ตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) จะถูกจัดรูปแบบให้อยู่ในสมการของกระบวนการอาร์มา (ARMA Process) กล่าวคือสามารถอธิบายค่าปัจจุบัน (Current Value) โดยใช้กระบวนการอัตถดถอย (Autoregressive : AR) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average : MA) ร่วมกัน ซึ่งเขียนเป็นสมการได้ดังนี้

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

โดยที่ Z_t คือ ค่าสังเกต ณ เวลาปัจจุบัน

μ คือ พจน์คงตัว (Constant Term)

ϕ_p คือ พารามิเตอร์อัตถดถอยตัวที่ p

θ_q คือ พารามิเตอร์ค่าเฉลี่ยเคลื่อนที่ตัวที่ q

Z_{t-p} คือ ค่าสังเกตใน p คาบเวลาที่ผ่านมา

ε_{t-q} คือ ค่าคลาดเคลื่อนใน q คาบเวลาที่ผ่านมา

ε_t คือ ค่าคลาดเคลื่อน ณ เวลาปัจจุบัน

3.1.2.6 กระบวนการอารีมา (ARIMA Process)

ถ้าอนุกรมเวลาที่พิจารณาขาดคุณสมบัติหยุดนิ่ง (Non-Stationary) ณ ระดับ (At Level) การสร้างระบบพยากรณ์ด้วยตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) จะต้องจัดรูปแบบสมการเป็นกระบวนการอารีมา (ARIMA Process) กล่าวคือต้องนำกระบวนการมีอันดับ (Integrated : I_d) ซึ่งเป็นการหาผลต่าง (Differencing) ระหว่างข้อมูล ณ เวลาปัจจุบันกับข้อมูลเมื่อถอยหลังไป d คาบเวลา มาใช้ร่วมกับกระบวนการอัตถดถอย (Autoregressive) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average) ในการอธิบายอนุกรมเวลาที่ขาดคุณสมบัติหยุดนิ่ง ณ ระดับ (At Level) ซึ่งกระบวนการอารีมา (ARIMA Process) สามารถจัดรูปแบบสมการ ได้ดังนี้

$$\Delta_d Z_t = \mu + \phi_1 \Delta_d Z_{t-1} + \phi_2 \Delta_d Z_{t-2} + \ldots$$
$$+ \phi_p \Delta_d Z_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$

โดยที่ $\Delta_d Z_t$ คือ ค่าสังเกต ณ เวลาปัจจุบันที่ผ่านการหาผลต่างมาแล้ว d ครั้ง

μ คือ พจน์คงตัว (Constant Term)

ϕ_p คือ พารามิเตอร์อัตถดถอยตัวที่ p

θ_q คือ พารามิเตอร์ค่าเฉลี่ยเคลื่อนที่ตัวที่ q

$\Delta_d Z_{t-p}$ คือ ค่าสังเกตใน p คาบเวลาที่ผ่านมาที่ผ่านการหาผลต่างมาแล้ว d ครั้ง

ε_{t-q} คือ ค่าคลาดเคลื่อนใน q คาบเวลาที่ผ่านมา

ε_t คือ ค่าคลาดเคลื่อน ณ เวลาปัจจุบัน

3.1.2.7 พจน์คงตัว (Constant Term)

พจน์คงตัว (Constant Term) ในตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) ทำหน้าที่เป็นค่าเฉลี่ย (Mean) ของสมการ ซึ่งการพิจารณาว่าต้องใส่พจน์คงตัว (Constant Term) เข้าไปในตัวแบบ (Model) หรือไม่นั้น มีหลักสำหรับการพิจารณาดังนี้

1) หากอนุกรมเวลาที่พิจาณาใช้กระบวนการอาร์มา (ARMA Process) กล่าวคือมีคุณสมบัติหยุดนิ่ง (Stationary) ณ ระดับ (At Level) สมการพยากรณ์จำเป็นต้องใส่พจน์คงตัว(Constant Term) เข้าไปด้วย

2) อนุกรมเวลาที่พิจารณาใช้กระบวนการอารีมา (ARIMA Process) โดยใช้กระบวนการมีอันดับลำดับที่หนึ่ง (Integrated : I_1) กล่าวคืออนุกรมเวลามีคุณสมบัติหยุดนิ่ง (Stationary) เมื่อหาผลต่างครั้งที่หนึ่ง (First Difference) ซึ่งการกำหนดว่าจะใส่พจน์คงตัว (Constant Term) หรือไม่นั้น สามารถแยกพิจารณาได้ดังนี้

2.1) หากทดสอบสมมติฐานแล้วพบว่าค่าเฉลี่ยของอนุกรมเวลาเป็นศูนย์ (Zero Mean) สมการพยากรณ์ไม่จำเป็นต้องใส่พจน์คงตัว (Constant Term) เข้าไปด้วย

2.2) หากทดสอบสมมติฐานแล้วพบว่าค่าเฉลี่ยของอนุกรมเวลาไม่เป็นศูนย์ (Non-Zero Mean) สมการพยากรณ์จะต้องใส่พจน์คงตัว (Constant Term) เข้าไปด้วย

3) อนุกรมเวลาที่พิจารณาใช้กระบวนการอารีมา (ARIMA Process) โดยใช้กระบวนการมีอันดับลำดับที่สอง (Integrated : I_2) กล่าวคืออนุกรมเวลามีคุณสมบัติหยุดนิ่ง (Stationary) เมื่อหาผลต่างครั้งที่สอง (Second Difference) สมการพยากรณ์ไม่จำเป็นต้องมีพจน์คงตัว (Constant Term)

3.1.3 การประมาณค่าพารามิเตอร์ (Parameter Estimation)

เมื่อกำหนดรูปแบบของกระบวนการอัตถดถอย (Autoregressive : AR) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average : MA) ที่เหมาะสมให้กับอนุกรมเวลาเรียบร้อยแล้ว ขั้นตอน

ต่อไปคือการประมาณค่าสัมประสิทธิ์ (Coefficient) ของพารามิเตอร์ ซึ่งวิธีที่ได้รับความนิยมและใช้กันโดยทั่วไปคือวิธีกำลังสองน้อยที่สุดสามัญ (Ordinary Least Square : OLS Method) ซึ่งค่าประมาณของพารามิเตอร์ที่ได้จากวิธีดังกล่าว เป็นค่าประมาณที่ให้ผลรวมกำลังสองของค่าคลาดเคลื่อน (Sum of Square Error : SSE) มีค่าน้อยที่สุด

เมื่อประมาณค่าพารามิเตอร์ของตัวแบบแล้ว ต้องทดสอบว่าสัมประสิทธิ์ (Coefficient) ที่ประมาณค่าได้แตกต่างจากศูนย์อย่างมีนัยสำคัญหรือไม่ โดยกำหนดสมมติฐานว่าง (Null Hypothesis) ว่าสัมประสิทธิ์ (Coefficient) ที่ประมาณค่าได้ไม่แตกต่างจากศูนย์อย่างมีนัยสำคัญ ซึ่งพารามิเตอร์ที่เหมาะสมสำหรับนำไปสร้างสมการพยากรณ์ ต้องปฏิเสธสมมติฐานว่าง (Null Hypothesis) กล่าวคือค่าความน่าจะเป็น (Probability-Value : Prob) ของค่าสถิติที (t-Statistic) มีค่าน้อยกว่าระดับนัยสำคัญ (Significant : α) ที่กำหนด

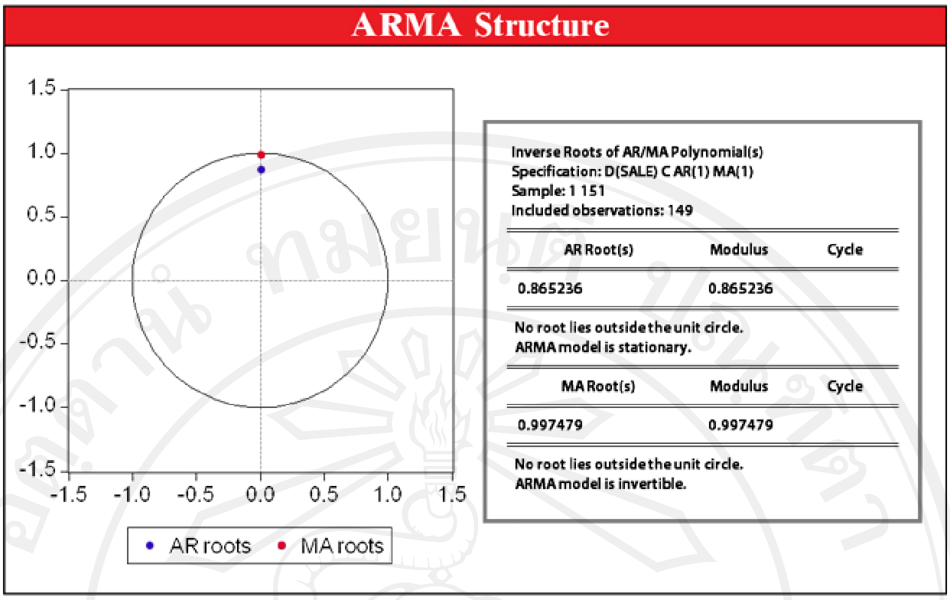
3.1.4 การตรวจสอบความถูกต้อง (Model Checking)

เมื่อเลือกรูปแบบของกระบวนการอัตถดถอย (Autoregressive : AR) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average : MA) ที่เหมาะสม รวมทั้งประมาณค่าพารามิเตอร์และตรวจสอบว่าแตกต่างจากศูนย์อย่างมีนัยสำคัญเรียบร้อยแล้ว ขั้นตอนต่อไปคือการวิเคราะห์เพื่อตรวจสอบความถูกต้องของตัวแบบ เพื่อการสนับสนุนว่าตัวแบบที่ได้ มีความเหมาะสมสำหรับนำไปสร้างสมการพยากรณ์ ซึ่งการตรวจสอบความถูกต้องจะแบ่งการทดสอบออกเป็น 2 ลักษณะ คือ การทดสอบโครงสร้างอาร์มา (ARMA Structure) และการทดสอบส่วนตกค้าง (Residual Test)

3.1.4.1 การทดสอบโครงสร้างอาร์มา (ARMA Structure)

เมื่อได้ตัวแบบที่เหมาะสมแล้ว ต้องทำการตรวจสอบโครงสร้างอาร์มา (ARMA Structure) ว่ากระบวนการอัตถดถอย (Autoregressive) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average) ที่เลือกมานั้น มีคุณสมบัติหยุดนิ่ง (Stationary) และคุณสมบัติผกผัน (Invertible) ซึ่งกระบวนการ (Process) ที่มีคุณสมบัติดังกล่าว จะต้องไม่มีรากผกผัน [๓] (Inverse Root) อยู่นอกวงกลมหนึ่งหน่วย (Unit Circle) ดังภาพที่ 3-1

[๓] การทดสอบโครงสร้างอาร์มา (ARMA Structure) โดยใช้ซอฟต์แวร์ EViews จะพิจารณารากผกผัน (Inverse Root) ซึ่งเป็นคำตอบของสมการพหุนามลักษณะเฉพาะ (Characteristic Polynomial) ต้องอยู่ภายในวงกลมหนึ่งหน่วย (Unit Circle) ขณะที่การสรุปว่ารูปแบบ AR และ MA ใด ๆ จะมีคุณสมบัติหยุดนิ่ง (Stationary) และผกผัน (Invertible) โดยทั่วไป จะพิจารณาจากรากลักษณะเฉพาะ (Characteristic Root) ของสมการพหุนามลักษณะเฉพาะ (Characteristic Polynomial) ที่อยู่นอกวงกลมหนึ่งหน่วย (Unit Circle)



ภาพที่ 3-1 การทดสอบโครงสร้างอาร์มา

3.1.4.2 การทดสอบส่วนตกค้าง (Residual Test)

ภายใต้ข้อสมมติของเศรษฐมิติแบบดั้งเดิม (Traditional Econometric) ส่วนตกค้าง (Residual) หรือพจน์คลาดเคลื่อน (Error Term) ที่ประมาณค่ามาได้ จะต้องไม่มีปัญหาอัตสหสัมพันธ์ (Autocorrelation) รวมทั้งมีการแจกแจงแบบปกติ (Normal Distribution) และไม่มีปัญหาความแปรปรวนของพจน์คลาดเคลื่อนไม่คงที่ (Heteroscedasticity) หรือกล่าวได้ว่าส่วนตกค้าง (Residual) มีคุณสมบัติเช่นเดียวกับสัญญาณรบกวนขาว[๔] (White Noise) ซึ่งคุณสมบัติต่างๆ

---

[๔] ในงานด้านวิศวกรรมการสื่อสาร (Communication Engineering) สัญญาณรบกวนขาว (White Noise) คือลักษณะของคลื่นเสียงที่เป็นสัญญาณสุ่ม (Random Signal) มีความหนาแน่นสเปรกตรัม (Spectrum Density) เท่าๆ กันหมดตลอดย่านความถี่ (Bandwidth) ตั้งแต่ 20-20,000 Hz ส่วนสาเหตุที่เรียกว่าสัญญาณรบกวนขาว เนื่องจากมีสเปรกตรัมแบบเดียวกันกับแสงสีขาว (White Light)

โดย Ron Mittelhammer (2000) ได้เปรียบเทียบระหว่างการส่งสัญญาณวิทยุกับการสร้างตัวแบบทางเศรษฐมิติว่า การส่งสัญญาณวิทยุจะประกอบด้วยองค์ประกอบหลักคือสัญญาณเสียงที่ต้องการให้ผู้รับได้ฟัง และส่วนที่เป็นสัญญาณรบกวน (Noise) ดังนั้นหากสัญญาณเสียงมีการกำหนดลักษณะการส่งมาเป็นอย่างดี สัญญาณรบกวนจะไม่สามารถตัดทอนหรือลดความสำคัญของสัญญาณเสียงที่เป็นองค์ประกอบหลักได้เลย

เช่นเดียวกันกับการสร้างตัวแบบทางเศรษฐมิติที่ประกอบด้วยส่วนที่เป็นเชิงกำหนด (Deterministic) คือค่าสังเกต และส่วนที่เป็นเชิงเฟ้นสุ่ม (Stochastic) ที่เรียกว่าตัวรบกวน (Disturbance) หรือส่วนตกค้าง (Residual) ซึ่งตัวแบบที่ดีนั้นไม่เพียงแต่จะนำเสนอส่วนที่เป็นโครงสร้างหลักที่มีความหมายทางทฤษฎีเท่านั้น แต่ส่วนที่เป็นตัวรบกวนต้องไม่ทำให้ประสิทธิภาพ (Efficiency) ของตัวแบบลดลง ดังนั้นจึงต้องมีการตรวจสอบคุณสมบัติต่างๆ ของตัวรบกวน

สำหรับตัวรบกวนที่มีความเหมาะสมจะต้องมีลักษณะเป็นตัวแปรสุ่ม (Random Variable) และเป็นอิสระ (Un-Correlated) ต่อกัน (ตัวรบกวนเฟ้นสุ่มในคาบเวลาที่แตกต่างกันไม่มีความสัมพันธ์กัน) โดยมีการแจกแจงแบบปกติ (Normal Distribution) ที่มีค่าเฉลี่ยเป็นศูนย์ (Zero Mean) และความแปรปรวนคงที่ (Uniform Variance) [$E(\varepsilon_t) = \mu = 0$ และ $VAR(\varepsilon_t) = \sigma^2_{\varepsilon}$]

ในข้างต้น สามารถแบ่งการตรวจสอบได้ 3 รูปแบบ คือ หนึ่ง การทดสอบค่าสถิติคิว (Q-Statistic Test) สอง การทดสอบรูปแบบปกติ (Normality Test) และสาม การทดสอบอาร์ช-แอลเอ็ม (ARCH-LM Test) อย่างไรก็ตามการสร้างระบบพยากรณ์ด้วยตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) ถ้าส่วนตกค้าง (Residual) ไม่มีปัญหาอัตสหสัมพันธ์ (Autocorrelation) ก็ถือว่ามีคุณสมบัติใกล้เคียง (Nearly) กับสัญญาณรบกวนขาว (White Noise) และสามารถนำไปใช้ในการพยากรณ์ได้ ซึ่งรายละเอียดของการทดสอบส่วนตกค้าง (Residual Test) มีดังนี้

1) การทดสอบค่าสถิติคิว (Q-Statistic Test)

การทดสอบค่าสถิติคิว (Q-Statistic Test) ถูกนำเสนอโดย Greta Ljung และ George Box (1979) เพื่อทดสอบว่าส่วนตกค้าง (Residual) หรือพจน์คลาดเคลื่อน (Error Term) ที่

ซึ่งสัญญาณรบกวนขาว (White Noise) ที่มีคุณสมบัติดังกล่าวครบถ้วนเรียกว่าสัญญาณรบกวนขาวแบบเกาส์เซียน (Gaussian White Noise) หรือสัญญาณรบกวนขาวที่แข็งแกร่ง (Strong White Noise) อย่างไรก็ตามคุณสมบัติของส่วนตกค้าง (Residual) ที่ประมาณค่ามาได้จากการวิเคราะห์อนุกรมเวลา (Time Series Analysis) มักจะเป็นสัญญาณรบกวนขาวที่อ่อนกำลัง (Weak White Noise) กล่าวคือเป็นตัวแปรสุ่ม (Random Variable) ที่ไม่มีปัญหาสหสัมพันธ์เชิงอันดับ (Serial Correlation) หรือไม่มีปัญหาอัตสหสัมพันธ์ (Autocorrelation) เพียงเท่านั้น

ภาพที่ 3-2 การเปรียบเทียบลักษณะของสัญญาณรบกวนขาวกับกราฟของส่วนตกค้าง

ภาพที่ 3-2 เป็นการเปรียบเทียบให้เห็นระหว่างสเปกตรัมกำลัง (Power Spectrum) ของสัญญาณรบกวนขาว (White Noise) กับกราฟของส่วนตกค้าง (Residual) ที่ได้จากการถดถอยอนุกรมเวลาว่ามีลักษณะที่ใกล้เคียงกัน โดยกราฟของส่วนตกค้าง (Residual) มีการเคลื่อนไหวแบบไม่มีแบบแผน (Pattern) และมีการเปลี่ยนแปลงมากบ้างน้อยบ้างรอบระดับค่าเฉลี่ย

ห่างกัน j ช่วงเวลา (โดยที่ j = 1 , 2 , ... , K) ไม่มีความสัมพันธ์ (Un-Correlated) กัน โดยกำหนดสมมติฐานว่าง (Null Hypothesis) ว่าส่วนตกค้าง (Residual) ไม่มีปัญหาอัตสหสัมพันธ์ (Autocorrelation) ซึ่งค่าสถิติที่ใช้ทดสอบคือ

$$Q_{LB} = n(n+2) \sum_{j=1}^{K} \frac{\zeta_j^2}{n-j}$$

โดยที่ Q_{LB} คือ ค่าสถิติคิว (Q-Statistic Test) มีการแจกแจงแบบไคกำลังสอง(Chi-Square Distribution) โดยที่มีองศาเสรี (Degree of Freedom) เท่ากับ K-p-q (เมื่อ p คืออันดับของกระบวนการอัตถดถอย และ q คืออันดับของกระบวนการเฉลี่ยเคลื่อนที่)

n คือ จำนวนส่วนตกค้าง (Residual) ทั้งหมดที่ได้จากตัวแบบที่พิจารณา

K คือ จำนวนของสัมประสิทธ์ที่ทดสอบปัญหาอัตสหสัมพันธ์

ζ_ϵ คือ สัมประสิทธิ์อัตสหสัมพันธ์ (Autocorrelation Coefficient)

หากค่าความน่าจะเป็น (Probability-Value : Prob) ของค่าสถิติคิว (Q-Statistic) มีค่ามากกว่าระดับนัยสำคัญ (Significant : α) จะยอมรับสมมติฐานว่าง (Null Hypothesis) ว่าส่วนตกค้าง (Residual) หรือพจน์คลาดเคลื่อน ณ ห้วงเวลาต่างกันไม่มีความสัมพันธ์ (Un-Correlated) กัน ซึ่งการกำหนดความล่าช้า (Lag) สำหรับการทดสอบนั้น หากจำนวนของความล่าช้า (Lag) ต่ำจนเกินไป จะไม่สามารถค้นหา (Detect) ปัญหาอัตสหสัมพันธ์อันดับสูง (High Order Autocorrelation) ได้ ส่วนความล่าช้า (Lag) ที่สูงเกินไป ค่าสถิติดังกล่าวจะสูญเสียอำนาจ (Power) ในการอธิบายผล เนื่องจากสหสัมพันธ์ที่มีนัยสำคัญ (Significant Correlation) ณ ความล่าช้า (Lag) หนึ่ง อาจถูกลดทอน (Washed Out) จากสหสัมพันธ์ที่ไม่มีนัยสำคัญ (Insignificant Correlation) ของความล่าช้า (Lag) อื่น

2) การทดสอบรูปแบบปกติ (Normality Test)

การทดสอบรูปแบบปกติ (Normality Test) เป็นการตรวจสอบว่าส่วนตกค้าง (Residual) มีการแจกแจงแบบปกติ[๕] (Normal Distribution) หรือไม่ โดยสามารถพิจารณาได้จากภาพแท่งความถี่ (Histogram) และค่าสถิติฌักค์-เบอรา (Jarque-Bera Statistic : JB)

สำหรับค่าสถิติฌักค์-เบอรา (JB) ถูกนำเสนอโดย Carlos Jarque และ Anil Bera (1980) ซึ่งค่าสถิติดังกล่าวถูกคำนวณจากความเบ้ (Skewness) และภาวะยอดมน (Kurtosis) ของส่วนตกค้าง (Residual) โดยเขียนในรูปสมการได้ดังนี้

$$JB = \frac{n}{6}\left[\varsigma^2 + \frac{(\kappa - 3)^2}{4}\right]$$

โดยที่ JB คือ ค่าสถิติฌักค์-เบอรา (Jarque-Bera Statistic) ที่มีการแจกแจงแบบไคกำลังสอง (Chi-Square Distribution) โดยที่มีองศาอิสระ (Degree of Freedom) เท่ากับ 2

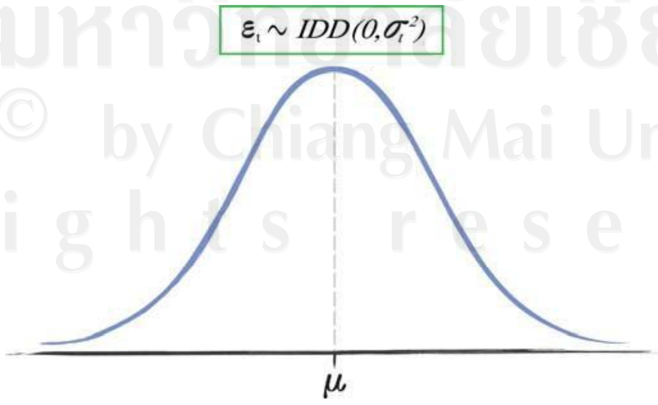
ς คือ ความเบ้ (Skewness) ของส่วนตกค้าง

κ คือ ภาวะยอดมน (Kurtosis) ของส่วนตกค้าง

N คือ จำนวนส่วนตกค้าง (Residual) ทั้งหมดที่ได้จากตัวแบบที่พิจารณา

---

[๕] ส่วนตกค้าง (Residual) ที่แจกแจงแบบปกติ (Normal Distribution) อาจพิจารณาได้จากโมเมนต์ (Moment) ทั้ง 4 คือ โมเมนท์ที่หนึ่ง (1st Moment) ใช้เพื่อตรวจสอบความถูกต้อง หรือกล่าวได้ว่าค่าเฉลี่ย (Mean) เป็น 0 โมเมนต์ที่สอง (2nd Moment) ใช้เพื่อวัดการกระจาย หรือกล่าวได้ว่าความแปรปรวน (Variance) คงที่ โมเมนต์ที่สาม (3rd Moment) ใช้วัดความสมมาตร หรือกล่าวได้ว่าความเบ้ (Skewness) เป็น 0 และโมเมนต์ที่สี่ (4th Moment) ใช้วัดความโด่ง หรือกล่าวได้ว่าภาวะยอดมน (Kurtosis) เท่ากับ 3



ภาพที่ 3-3 ส่วนตกค้างที่แจกแจงแบบปกติ

สำหรับการทดสอบรูปแบบปกติ (Normality Test) จะกำหนดสมมติฐานว่าง (Null Hypothesis) ว่าส่วนตกค้าง (Residual) มีการแจกแจงแบบปกติ หากค่าความน่าจะเป็น (Probability-Value : Prob) ของค่าสถิติฌักค์-เบอรา (Jarque-Bera Statistic) มีค่ามากกว่าระดับนัยสำคัญ (Significant : α) จะยอมรับสมมติฐานว่าง (Null Hypothesis) ว่าส่วนตกค้าง (Residual) มีการแจกแจงแบบปกติ (Normal Distribution)

3) การทดสอบอาร์ช-แอลเอ็ม (ARCH-LM Test)

การทดสอบดังกล่าวถูกนำเสนอโดย Robert Engle (1982) เพื่อตรวจสอบปัญหาความแปรปรวนของพจน์คลาดเคลื่อนไม่คงที่อย่างมีเงื่อนไข (Conditional Heteroscedasticity) โดยมีแนวคิดว่าหากขนาด (Magnitude) ของส่วนตกค้าง (Residual) ในปัจจุบันไม่เป็นอิสระ (Independent) กับขนาดของส่วนตกค้าง (Residual) ในคาบเวลา (Period) ก่อนหน้าจะทำให้ความแปรปรวนของส่วนตกค้างไม่เท่ากับค่าคงที่ (Constant) ค่าหนึ่ง หรือกล่าวได้ว่าเกิดปัญหาความแปรปรวนของพจน์คลาดเคลื่อนไม่คงที่ (Heteroscedasticity) ซึ่งในการสร้างสมการทดสอบจะใช้ตัวคูณลากรานจ์ (Lagrange Multiplier : LM) ที่คำนวณด้วยวิธีถดถอยประกอบ (Auxiliary Regression) โดยเขียนเป็นสมการได้ดังนี้

$$\varepsilon_t^2 = \omega + \sum_{s=1}^{q} \alpha_s \varepsilon_{t-s}^2 + \ell_t$$

โดยที่ ε_t^2 คือ ค่าคลาดเคลื่อนกำลังสอง (Square Error) ณ เวลาปัจจุบัน

ω คือ พจน์คงตัว (Constant Term)

α_s คือ พารามิเตอร์อัตถดถอยตัวที่ s

ε_{t-s}^2 คือ ค่าคลาดเคลื่อนกำลังสอง (Square Error) ใน s คาบเวลาที่ผ่านมา

ℓ_t คือ ตัวรบกวนเฟ้นสุ่ม (Stochastic Disturbance) ที่ได้จากการถดถอย

ซึ่งการทดสอบอาร์ช-แอลเอ็ม [๖] (ARCH-LM Test) มีการกำหนดสมมติฐานว่าง (Null Hypothesis) ว่าไม่มีกระบวนการอาร์ชลำดับที่ q (ARCH(q) Process) หลงเหลืออยู่ใน

[๖] การทดสอบอาร์ช-แอลเอ็ม (ARCH-LM) เป็นการพิจารณาปัญหาความแปรปรวนของพจน์คลาดเคลื่อนอย่างมีเงื่อนไข (Conditional Heteroscedasticity) ซึ่งบางครั้งการทดสอบดังกล่าวใช้เพื่อตรวจว่าส่วนตกค้าง (Residual) สามารถสร้างเป็นตัวแบบความผันผวน (Volatility Model) ที่ตัวแปรความล่าช้า (Lag Variable) อยู่ในรูปพจน์อาร์ช (ARCH Term) หรือพจน์การ์ช (GARCH Term) ได้

ส่วนตกค้าง (Residual) ดังนั้นหากอนุกรมเวลาที่พิจารณาไม่มีปัญหาความแปรปรวนของพจน์คลาดเคลื่อนไม่คงที่ (Heteroscedasticity) จะต้องยอมรับสมมติฐานว่าง กล่าวคือค่าความน่าจะเป็น (Probability-Value : Prob) ของค่าสถิติทดสอบอาร์ช-แอลเอ็ม (ARCH-LM Test Statistic) มีค่ามากกว่าระดับนัยสำคัญ (Significant : α) ที่กำหนด

3.1.5 การพยากรณ์ (Forecasting)

หากเป็นการพยากรณ์ค่าในช่วงที่ใช้จำลองแบบ (Historical Forecasting) จะใช้เงื่อนไขของอากาอิเกะ (Akaike's Information Criterion : AIC) หรือเงื่อนไขของชวาร์ซ (Schwarz's Bayesian Criterion : SBC) ในการเลือกตัวแบบที่เหมาะสมเพียงตัวแบบเดียว อย่างไรก็ตาม การศึกษาครั้งนี้มีวัตถุประสงค์เพื่อหาตัวแบบที่มีประสิทธิภาพมากที่สุดในการพยากรณ์ไปข้างหน้าเพื่อทดสอบผล (Ex-post Forecasting) ดังนั้นในการเลือกตัวแทนของตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) จะทดลองพยากรณ์ไปข้างหน้าในช่วงสั้น ๆ แล้วเลือกตัวแบบที่ให้ค่าร้อยละสัมบูรณ์ความคลาดเคลื่อนเฉลี่ย (Mean Absolute Percentage Error : MAPE) น้อยที่สุด สำหรับนำไปเปรียบเทียบความแม่นยำในการพยากรณ์กับตัวแบบประเภทอื่นๆ ซึ่งค่าเกณฑ์วัด (Criterion) ดังกล่าวมีสูตรการคำนวณดังนี้

$$\text{MAPE} = \frac{100}{h} \sum_{t=T+1}^{T+h} \left| \frac{\hat{Y}_t - Y_t}{Y_t} \right|$$

โดยที่ Y_t คือ ค่าจริง (Actual Value)

$\hat{Y}_t$ คือ ค่าพยากรณ์

h คือ จำนวนค่าพยากรณ์ทั้งหมด

3.2 การพยากรณ์ด้วยตัวแบบการ์ช (GARCH Model)

ข้อสมมติของตัวแบบเศรษฐมิติแบบดั้งเดิม (Traditional Econometric Model) เกี่ยวกับคุณสมบัติของส่วนตกค้าง (Residual) ว่าต้องมีความเป็นอิสระและกระจายอย่างเท่ากัน (Independent and Identically Distribution : i.i.d) มักจะถูกละเมิด (Violate) อยู่บ่อยครั้ง เนื่องจากความแปรปรวน (Variance) ของตัวแปรทางการเงิน (Financial Variables) มีลักษณะแปรเปลี่ยนตามกาลเวลา (Time Varying) ดังนั้นจึงมีการพัฒนาตัวแบบที่พิจารณาความแปรปรวนของพจน์

คลาดเคลื่อนไม่คงที่อย่างมีเงื่อนไข (Conditional Heteroscedasticity) นั่นคือตัวแบบอาร์ช (ARCH Model) ต่อมาได้มีการปรับปรุงตัวแบบดังกล่าวให้มีความยืดหยุ่น (Flexible) มากขึ้น และทำให้ภาระในการคำนวณ (Computational Burden) ลดลง โดยสร้างความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) ให้อยู่ในรูปแบบนัยทั่วไป (Generalized Form) ที่เรียกว่าตัวแบบการ์ช (GARCH Model) ซึ่งขั้นตอนการสร้างระบบพยากรณ์ด้วยตัวแบบการ์ช (GARCH Model) มีดังนี้

3.2.1 การสร้างสมการค่าเฉลี่ย (Mean Equation)

การพยากรณ์ด้วยตัวแบบการ์ช (GARCH Model) จะประกอบด้วย 2 ส่วนที่สำคัญคือ สมการค่าเฉลี่ย (Mean Equation) และสมการความแปรปรวน (Variance Equation) ซึ่งในการศึกษาครั้งนี้จะสร้างสมการค่าเฉลี่ยจากกระบวนการอัตถดถอย (Autoregressive : AR) และกระบวนการค่าเฉลี่ยเคลื่อนที่ (Moving Average) ของตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) ที่ผ่านการตรวจสอบความถูกต้อง (Model Checking) แล้ว โดยอนุกรมเวลาที่มีคุณสมบัติหยุดนิ่ง (Stationary) ณ ระดับ (At Level) สามารถเขียนเป็นสมการค่าเฉลี่ยได้ดังนี้

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \nu_t \sigma_t$$

โดยที่ Z_t คือ ค่าสังเกต ณ เวลา t

μ คือ พจน์คงตัว (Constant Term)

ϕ_p คือ พารามิเตอร์อัตถดถอยตัวที่ p

θ_q คือ พารามิเตอร์ค่าเฉลี่ยเคลื่อนที่ตัวที่ q

Z_{t-p} คือ ค่าสังเกตใน p คาบเวลาที่ผ่านมา

ε_{t-q} คือ ค่าคลาดเคลื่อนใน q คาบเวลาที่ผ่านมา

ν_t คือ กระบวนการเฟ้นสุ่ม (Stochastic Process) ณ เวลา t

σ_t คือ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) อย่างมีเงื่อนไขของพจน์คลาดเคลื่อน

สำหรับอนุกรมเวลาที่ขาดคุณสมบัติหยุดนิ่ง (Stationary) จะต้องทำการหาผลต่าง (Differencing) จนอนุกรมเวลาดังกล่าวมีคุณสมบัติหยุดนิ่ง ซึ่งสามารถเขียนเป็นสมการค่าเฉลี่ยได้ดังนี้

$$\Delta_d Z_t = \mu + \phi_1 \Delta_d Z_{t-1} + \phi_2 \Delta_d Z_{t-2} + \dots + \phi_p \Delta_d Z_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \nu_t \sigma_t$$

โดยที่ $\Delta_d Z_t$ คือ ค่าสังเกต ณ เวลา t ที่ผ่านการหาผลต่างมาแล้ว d ครั้ง

μ คือ พจน์คงตัว (Constant Term)

ϕ_p คือ พารามิเตอร์อัตถดถอยตัวที่ p

θ_q คือ พารามิเตอร์ค่าเฉลี่ยเคลื่อนที่ตัวที่ q

$\Delta_d Z_{t-p}$ คือ ค่าสังเกตใน p คาบเวลาที่ผ่านมา ที่ผ่านการหาผลต่างมาแล้ว d ครั้ง

ν_t คือ กระบวนการเฟ้นสุ่ม (Stochastic Process) ณ เวลา t

σ_t คือ ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) อย่างมีเงื่อนไขของพจน์คลาดเคลื่อน

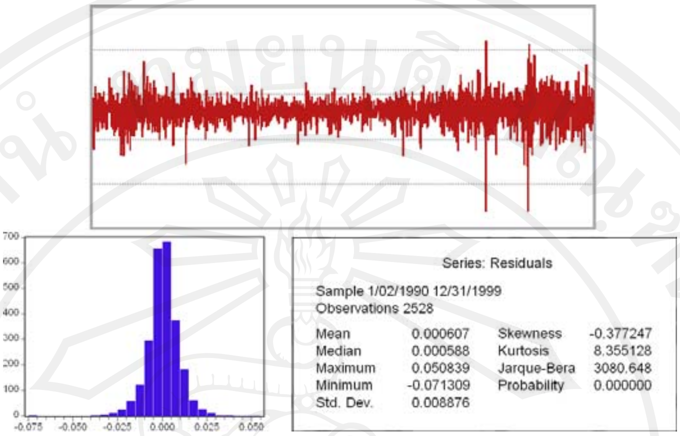
3.2.2 การสร้างสมการความแปรปรวน (Variance Equation)

เมื่อสร้างสมการค่าเฉลี่ย (Mean Equation) เสร็จแล้ว ขั้นตอนต่อไปคือการสร้างสมการความแปรปรวน (Variance Equation) โดยพิจารณาตัวแปรความล่าช้า (Lag Variable) จากแผนภาพสหสัมพันธ์ (Correlogram) ของส่วนตกค้างกำลังสอง (Square Residual) สำหรับการศึกษาครั้งนี้กำหนดรูปแบบของสมการความแปรปรวน (Variance Equation) โดยใช้กระบวนการอาร์ช (ARCH Process) กระบวนการการ์ช (GARCH Process) และกระบวนการอีการ์ช (EGARCH Process) ซึ่งรายละเอียดของกระบวนการต่างๆ มีดังนี้

3.2.2.1 กระบวนการอาร์ช (ARCH Process)

ตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) เป็นหนึ่งในเครื่องมือทางเศรษฐศาสตร์ที่มีประโยชน์มากสำหรับการพยากรณ์อนุกรมเวลาแบบทั่วไป (General Time Series) โดยข้อสมมติ (Assumption) ที่สำคัญเกี่ยวกับคุณสมบัติของพจน์คลาดเคลื่อน (Error Term) คือต้องมีความเป็นอิสระและกระจายอย่างเท่ากัน (Independent and Identically Distribution : i.i.d) หรือกล่าวได้ว่ามีคุณสมบัติเช่นเดียวกันกับสัญญาณรบกวนขาว (White Noise) นั่นคือพจน์คลาดเคลื่อน

เป็นฟังก์ชันของกระบวนการสุ่ม (Random Process) มีค่าเฉลี่ยเป็นศูนย์ (Zero Mean) และมีความแปรปรวนคงที่ (Homoscedasticity)



ภาพที่ 3-4 พจน์คลาดเคลื่อนที่ละเมิดข้อสมมติ i.i.d

อย่างไรก็ตามการวิเคราะห์ตัวแปรทางการเงิน (Financial Variable) ส่วนใหญ่นั้น มักจะละเมิด (Violated) คุณสมบัติดังกล่าวของตัวแบบเศรษฐมิติแบบดั้งเดิม (Traditional Econometric Model) ดังภาพที่ 3-4 เนื่องจากการเคลื่อนไหวของอนุกรมเวลาทางการเงินจะประกอบด้วยคาบเวลาที่รุนแรง (Swing) สลับกับคาบเวลาที่สงบเงียบ (Calm) ทำให้ความผันผวนมีลักษณะกระจุกตัว (Volatility Clustering) กล่าวคือแม้จะมีค่าเฉลี่ยคงที่ (Constant) แต่ความแปรปรวน (Variance) เปลี่ยนแปลงตามกาลเวลา (Varies Over Time) นั่นหมายความว่าความแปรปรวนของพจน์คลาดเคลื่อน (Error Term) ซึ่งเป็นปัจจัยที่ไม่สามารถคาดคะเนล่วงหน้าได้ขึ้นอยู่กับค่าในอดีตของพจน์คลาดเคลื่อนนั้น ๆ โดยความผันผวน (Volatility) ที่สูงหรือต่ำเกินปกติ (Extreme Value) มักจะนำไปสู่การแจกแจงแบบหางอ้วน [๗] (Fat Tailed Distribution)

[๗] การแจกแจงแบบหางอ้วน (Fat Tailed Distribution) บางครั้งเรียกว่าการแจกแจงแบบหางหนัก (Heavy Tailed Distribution) หรือการแจกแจงแบบหางหนา (Thick Tailed Distribution) ซึ่งการแจกแจงดังกล่าวมักเกิดขึ้นกับข้อมูลที่มีค่านอกกลุ่ม (Outlier) เป็นจำนวนมาก ทำให้มวลที่หางสูงกว่าข้อมูลที่แจกแจงแบบปกติ (Normal Distribution) โดยทั่วไปการแจกแจงแบบหางอ้วนจะมีภาวะยอดมน (Kurtosis) มากกว่า 3 ทำให้ภาพแท่งความถี่ (Histogram) มีลักษณะเป็นโค้งสูงแหลม (Leptokurtic) ดังภาพที่ 3-5

ซึ่ง Robert Engle (1982) ได้พัฒนาแนวทางใหม่ของกระบวนการเฟ้นสุ่ม (Stochastic Process) ที่เรียกว่ากระบวนการอาร์ช (Autoregressive Conditional Heteroscedasticity : ARCH Process) โดยผ่อนปรน (Relax) เงื่อนไขความแปรปรวนของพจน์คลาดเคลื่อนคงที่ (Homoscedasticity) และเสนอแนวคิดเกี่ยวกับความแปรปรวนของตัวรบกวนอย่างมีเงื่อนไข (Conditional Disturbance Variance) ว่าปัญหาความแปรปรวนของพจน์คลาดเคลื่อนไม่คงที่ (Heteroscedasticity) สามารถเกิดขึ้นได้ หากกำหนดให้ความแปรปรวนของพจน์คลาดเคลื่อนขึ้นอยู่กับขนาด (Magnitude) ของค่าคลาดเคลื่อนกำลังสอง (Squared Error) ในอดีต ซึ่งความสัมพันธ์ดังกล่าวสามารถเขียนในรูปสมการได้ดังนี้

$$\text{ARCH (q)} \quad ; \quad \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \upsilon_{t-i}^2$$

โดยที่ σ_t^2 คือ ความแปรปรวนของตัวรบกวนอย่างมีเงื่อนไข

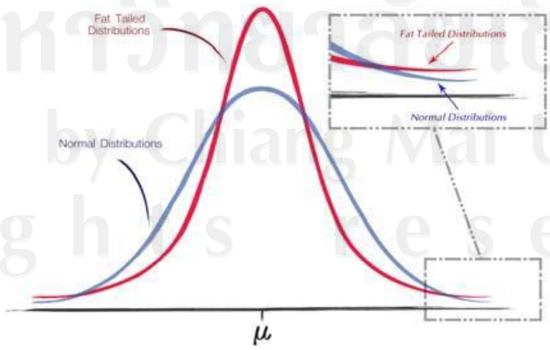
ω คือ พจน์คงตัว (Constant Term)

α_i คือ ความล่าช้าอัตถดถอย ~ (Autoregressive Lag) ลำดับที่ i

υ_{t-i}^2 คือ ค่าคลาดเคลื่อนกำลังสอง (Squared Error) ใน i คาบเวลาที่ผ่านมา

สำหรับพจน์อาร์ชอันดับที่หนึ่ง (ARCH(1)) นั่นคือความแปรปรวนของตัวรบกวนอย่างมีเงื่อนไข (Conditional Disturbance Variance) ขึ้นอยู่กับค่าคลาดเคลื่อนกำลังสอง (Squared Error) ในหนึ่งคาบเวลาที่ผ่านมา โดยสามารถเขียนเป็นสมการได้ดังนี้

$$\sigma_t^2 = \omega + \alpha_1 \upsilon_{t-1}^2$$



ภาพที่ 3-5 การแจกแจงแบบหางอ้วน (Fat Tailed Distribution) เปรียบเทียบกับการแจกแจงแบบปกติ (Normal Distribution)

~ ความล่าช้าอัตถดถอย (Autoregressive Lag) หรือผลของอาร์ช (ARCH Effect) ใช้สำหรับวัดผลกระทบของค่าคลาดเคลื่อนกำลังสอง (Square Error) ในอดีตต่อความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance)

ซึ่งรูปแบบของกระบวนการอาร์ช (ARCH Process) เป็นการประเมินค่าความผันผวน (Volatility) โดยวัดค่าออกมาในรูปความแปรปรวน (Variance) ดังนั้นจึงต้องมีการใส่ข้อจำกัดต่างๆ ให้กับพารามิเตอร์ของสมการความแปรปรวนเพื่อให้เกิดเสถียรภาพ (Stability) และไม่ทำให้ความแปรปรวนอย่างมีเงื่อนไขมีค่าติดลบ (Non-Negative) โดยข้อจำกัดดังกล่าวกำหนดให้พจน์คงตัว (Constant Term) มีค่ามากกว่าศูนย์ ($\omega > 0$) และความล่าช้าอัตถดถอย (Autoregressive Lag) ลำดับที่ i ต้องไม่ติดลบ ($\alpha_i \geq 0$)

3.2.2.2 กระบวนการ์ช (GARCH Process)

จากคุณสมบัติของกระบวนการอาร์ช (ARCH Process) ที่มีการผ่อนปรนเงื่อนไขความแปรปรวน (Variance) ของพจน์คลาดเคลื่อน (Error Term) ให้มีลักษณะเป็นความแปรปรวนไม่คงที่ (Heteroscedasticity Variance) ทำให้กระบวนการอาร์ช (ARCH Process) ถูกนำไปใช้อย่างแพร่หลายในการอธิบายปรากฏการณ์ (Phenomena) ต่าง ๆ ของตัวแปรทางการเงิน เช่น การวิเคราะห์ความเสี่ยง (Risk Analysis) การวิเคราะห์ความผันผวน (Volatility) รวมทั้งการเลือกพอร์ทการลงทุน (Portfolio Selection)

อย่างไรก็ตามจากการศึกษาเชิงประจักษ์ (Empirical Work) หลาย ๆ ครั้ง พบว่ากระบวนการอาร์ช (ARCH Process) มีข้อจำกัดหลายประการ กล่าวคือการกำหนดให้ความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) หรือความผันผวน (Volatility) ขึ้นอยู่กับค่าคลาดเคลื่อนกำลังสอง (Squared Error) ในอดีตเป็นจำนวน q ความล่าช้า (Lag) นั้น บางครั้งมีความจำเป็นต้องใช้อันดับ q สูงมาก [๙] (Very High Order ARCH(q))

ซึ่งการกำหนดโครงสร้างความล่าช้า (Lag Structure) โดยใช้พารามิเตอร์หลายตัว นอกจากจะยากต่อการประมาณค่า (Approximation) แล้ว ยังมักนำไปสู่การละเมิด (Violate) เงื่อนไขบังคับของพารามิเตอร์ที่ต้องไม่เป็นจำนวนติดลบ [๑๐] (Non-Negative Constraint) ดังนั้นเพื่อการหลีกเลี่ยง (Circumvent) ปัญหาดังกล่าว ทำให้เกิดแนวคิดในการพัฒนากระบวนการอาร์ช (ARCH Process) บนหลักของการมัธยัสถ์ (Parsimonious) กล่าวคือใช้จำนวนความล่าช้า (Lag) ให้น้อยที่สุดสำหรับการอธิบายสมการความผันผวน (Volatility)

---

[๙] รูปแบบแผนภาพสหสัมพันธ์ (Correlogram) ของส่วนตกค้างกำลังสอง (Square Residual) อาจแสดงความสัมพันธ์ในลักษณะที่ต้องใช้อันดับ (Order) ของพจน์อาร์ช (ARCH Term) เป็นจำนวนมากในการอธิบายสมการความแปรปรวน

[๑๐] ตัวแบบความผันผวน (Volatility Model) เป็นการวัดค่า (Mesure) ในรูปความแปรปรวน (Variance) ดังนั้นต้องมีการกำกับพารามิเตอร์ไม่ให้เป็นจำนวนติดลบ (Non-Negative) เพื่อให้สมการความแปรปรวน (Variance Equation) มีค่าเป็นบวกเสมอ

โดย Tim Bollerslev (1986) เสนอว่าควรแปลงอันดับของกระบวนการอาร์ช (ARCH Process) จากที่เป็นความล่าช้าอัตถดถอย (Autoregressive Lag) ให้มีความยืดหยุ่น (Flexible) มากขึ้น โดยเพิ่มองค์ประกอบของความล่าช้าค่าเฉลี่ยเคลื่อนที่ (Moving Average Lag) เข้าไป กล่าวคือการสร้างสมการความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) กำหนดให้ตัวแปรความล่าช้า (Lag Variable) มีลักษณะเป็นกระบวนการอาร์มา (ARMA Process) แทนการใช้ความล่าช้าอัตถดถอย (Autoregressive Lag) อันดับสูง (High-Order) เพียงอย่างเดียว ซึ่งความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) นอกจากจะขึ้นอยู่กับค่าคลาดเคลื่อนกำลังสอง (Squared Error) ในอดีตแล้ว ยังอนุญาต (Allow) ให้ขึ้นอยู่กับความผันผวน (Volatility) ในอดีตอีกด้วย โดยกระบวนการดังกล่าวเรียกว่ากระบวนการการ์ช[๑๑] (Generalized Autoregressive Conditional Heteroscedasticity : GARCH Process) ซึ่งสามารถเขียนในรูปสมการได้ดังนี้

$$\text{GARCH (p,q)} \quad ; \quad \sigma_t^2 = \omega + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 + \sum_{i=1}^{q} \alpha_i \upsilon_{t-i}^2$$

โดยที่ σ_t^2 คือ ความแปรปรวนของตัวรบกวนอย่างมีเงื่อนไข

ω คือ พจน์คงตัว (Constant Term)

α_i คือ ความล่าช้าอัตถดถอย (Autoregressive Lag) ลำดับที่ i

β_j คือ ความล่าช้าค่าเฉลี่ยเคลื่อนที่[๑๒] (Moving Average Lag) ลำดับที่ j

υ_{t-i}^2 คือ ค่าคลาดเคลื่อนกำลังสอง (Squared Error) ใน i คาบเวลาที่ผ่านมา

σ_{t-j}^2 คือ ความผันผวน (Volatility) ใน j คาบเวลาที่ผ่านมา

สำหรับพจน์อาร์ชและพจน์การ์ชอันดับที่หนึ่ง (GARCH(1,1)) คือการกำหนดให้ความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) ขึ้นอยู่กับค่าคลาดเคลื่อนกำลังสอง (Squared Error) และความผันผวน (Volatility) ในหนึ่งคาบเวลาที่ผ่านมาโดยสามารถเขียนเป็นสมการได้ดังนี้

$$\sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + \alpha_1 \upsilon_{t-1}^2$$

[๑๑] กระบวนการการ์ช (GARCH Process) บางครั้งเรียกว่ากระบวนการอาร์ชนัยทั่วไป (Generalized ARCH Process)

[๑๒] ความล่าช้าค่าเฉลี่ยเคลื่อนที่ (Moving Average Lag) หรือผลของการ์ช (GARCH Effect) ใช้สำหรับวัดผลกระทบของผันผวน (Volatility) ในอดีต ต่อความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance)

โดยกระบวนการการ์ช (GARCH Process) จะมีการกำหนดข้อจำกัดของพารามิเตอร์เช่นเดียวกับกระบวนการอาร์ช (ARCH Process) กล่าวคือพจน์คงตัว (Constant Term) มีค่ามากกว่าศูนย์ ($\omega > 0$) รวมทั้งความล่าช้าอัตถดถอย (Autoregressive Lag) ลำดับที่ i และความล่าช้าค่าเฉลี่ยเคลื่อนที่ (Moving Average Lag) ลำดับที่ j ต้องไม่ติดลบ ($\alpha_i \geq 0$ and $\beta_j \geq 0$) นอกจากนี้กระบวนการการ์ช (GARCH Process) จะมีเสถียรภาพถ้าผลรวมระหว่างพารามิเตอร์ทั้งสองในข้างต้นมีค่าน้อยกว่าหนึ่ง ($\alpha_i + \beta_j < 1$)

3.2.2.3 กระบวนการอีการ์ช (EGARCH Process)

กระบวนการการ์ช (GARCH Process) เป็นรูปแบบในการวัด (Measure) ความผันผวน (Volatility) ที่ถูกขยาย (Extended) มาจากรูปแบบดั้งเดิม (Original) ของกระบวนอาร์ช (ARCH Process) เพื่อให้สามารถประยุกต์ใช้งานได้เป็นการทั่วไป (Generalize) โดยกระบวนการการ์ช (GARCH Process) ได้รับความนิยมสำหรับศึกษาพฤติกรรมการแกว่งตัว (Fluctuation) ของตัวแปรในตลาดการเงิน (Financial Market) อาทิ อัตราผลตอบแทนของสินทรัพย์ (Asset Returns) ราคาสัญญาออปชัน (Option Pricing) อัตราแลกเปลี่ยนเงินตราต่างประเทศ (foreign exchange) ดัชนีราคาหลักทรัพย์ (Stock Index) เป็นต้น

อย่างไรก็ตามการทำงานของกระบวนการการ์ช (GARCH Process) มักก่อให้เกิดความผิดพลาด (Failure) เมื่อนำไปใช้ศึกษาพฤติกรรมของตัวแปรที่อ่อนไหว (Sensitive) ต่อข่าวสาร เนื่องจากกระบวนการการ์ช (GARCH Process) กำหนดว่าอิทธิพลของความผิดปกติ (Shock) ต่อความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) มีลักษณะสมมาตร (Symmetric) โดยขึ้นอยู่กับขนาด (Magnitude) เท่านั้น กล่าวคือถ้าความผิดปกติที่เกิดขึ้นทั้งเชิงบวก (Positive) และเชิงลบ (Negative) มีขนาดเท่ากัน ความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) ที่เกิดขึ้นก็จะเพิ่มขึ้นหรือลดลงในขนาด (Magnitude) ที่เท่ากันด้วย

ซึ่งในความเป็นจริงแล้วตัวแปรทางการเงิน (Financial Variable) มักจะมีผลของอัตราทด [๑๓] (Leverage Effect) ดังนั้นการใช้กระบวนการการ์ช (GARCH Process) สำหรับการ

---

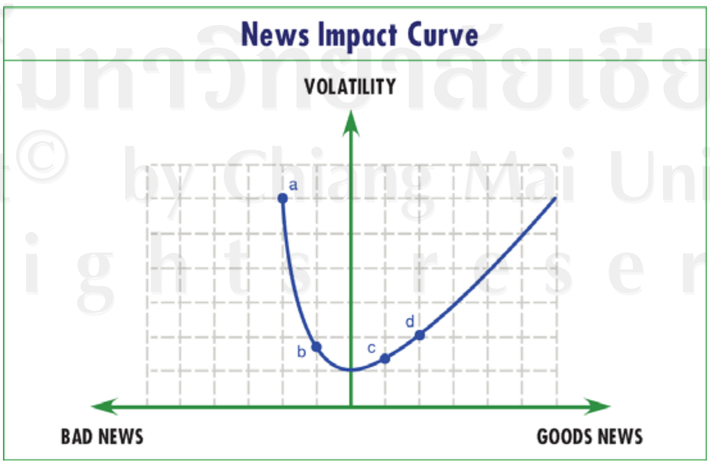
[๑๓] ผลของอัตราทด (Leverage Effect) หรือผลของภาวะตลาดตกต่ำ (Down-Market Effect) คือความสัมพันธ์ระหว่างความผันผวน (Volatility) และความผิดปกติ (Shock) ซึ่งมีลักษณะไม่สมมาตร (Asymmetric) กล่าวคือข่าวร้ายหรือความผิดปกติเชิงลบ (Negative Shock) จะส่งผลต่อความผันผวน (Volatility) มากกว่าข่าวดีหรือความผิดปกติเชิงบวก (Positive Shock)

วิเคราะห์อนุกรมเวลาจึงมักจะขาดประสิทธิภาพในการจับ (Capture) ปรากฏการณ์ที่ไม่สม่ำเสมอ (Irregular Phenomena) ดังกล่าว

โดย Daniel Nelson (1992) ได้พัฒนากระบวนการการ์ช (GARCH Process) ที่สามารถจับ (Capture) ผลของความไม่สมมาตร (Asymmetric Effect) ในความผันผวน (Volatility) ได้ โดยเรียกกระบวนการดังกล่าวว่ากระบวนการอีการ์ช (Exponential Generalized Autoregressive Conditional Heteroscedasticity : EGARCH Process) ซึ่งกระบวนการดังกล่าวผ่อนปรนเงื่อนไขความสัมพันธ์ของความผันผวน (Volatility) กับความผิดปกติ (Shock) ให้มีลักษณะที่ยืดหยุ่นมากขึ้น โดยความผันผวนขึ้นอยู่กับขนาด (Magnitude) และสัญญาณ (Sign) ทั้งเชิงบวกและลบ ซึ่งเขียนรูปแบบสมการความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) ได้ดังนี้

$$\text{EGARCH (p,q)} \quad ; \quad \log(\sigma_t^2) \;=\; \omega + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2) + \sum_{i=1}^{q} \alpha_i \left| \frac{\upsilon_{t-i}}{\sigma_{t-i}} \right| + \sum_{i=1}^{q} \gamma_i \frac{\upsilon_{t-i}}{\sigma_{t-i}}$$

โดยสัมประสิทธิ์ γ_i สามารถจับ (Capture) ผลของความไม่สมมาตร (Asymmetric Effect) จากความผิดปกติ (Shock) ในอดีตได้ ซึ่งค่าของสัมประสิทธิ์ γ_i มักจะมีค่าเป็นลบ อันมีนัย (Imply) ว่าความผิดปกติเชิงบวก ส่งผลกระทบต่อความผันผวน (Volatility) น้อยกว่าความผิดปกติเชิงลบ เมื่อความผิดปกติทั้งสองมีขนาด (Magnitude) เท่ากัน หรือกล่าวได้ว่าอนุกรมเวลาที่พิจารณามีผลของอัตราทด (Leverage Effect) โดยกระบวนการอีการ์ช (EGARCH



ภาพที่ 3-6 ความสัมพันธ์ระหว่างความผันผวน (Volatility) กับความผิดปกติ (Shock)

Process) จะมีเสถียรภาพถ้าสัมประสิทธิ์ (Coefficient) ของความผันผวน (Volatility) ใน j คาบเวลาที่ผ่านมาน้อยกว่าหนึ่ง ($\beta_j < 1$)

นอกจากนี้กระบวนการอีการ์ช (EGARCH Process) ยังผ่อนปรนเงื่อนไขบังคับของพารามิเตอร์ที่ต้องไม่เป็นจำนวนติดลบ (Non-Negative Constraint) เนื่องจากกระบวนการอีการ์ช (EGARCH Process) จะพิจารณาความแปรปรวนอย่างมีเงื่อนไข (Conditional Variance) ในรูปของลอการิทึม (Logarithm) ดังนั้นไม่ว่าสัมประสิทธิ์ที่ประมาณค่าได้ในสมการความแปรปรวน (Variance Equation) จะเป็นบวกหรือลบ ก็จะไม่ละเมิดข้อจำกัดดังกล่าว

3.2.3 การทดสอบคุณสมบัติของส่วนตกค้าง (Residual)

ขั้นตอนนี้จะคล้ายกับการตรวจสอบความถูกต้อง (Model Checking) ของตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) โดยมีวัตถุประสงค์เพื่อทดสอบว่าส่วนตกค้าง (Residual) มีคุณสมบัติใกล้เคียงกับสัญญาณรบกวนขาว (White Noise) โดยแบ่งการตรวจสอบได้ดังนี้

3.2.3.1 การทดสอบค่าสถิติคิว (Q-Statistic Test)

สำหรับการสร้างระบบพยากรณ์ด้วยตัวแบบการ์ช (GARCH Model) การทดสอบค่าสถิติคิว (Q-Statistic Test) เป็นการตรวจสอบว่าส่วนตกค้าง (Reidual) ของสมการค่าเฉลี่ย (Mean Equation) ไม่มีปัญหาอัตสหสัมพันธ์ (Autocorrelation) หรือสหสัมพันธ์เชิงอันดับ (Serial Correlation) ซึ่งหลักเกณฑ์ในการทดสอบและตั้งสมมติฐานว่าง (Null Hypothesis) เหมือนกับการทดสอบค่าสถิติคิว (Q-Statistic Test) ของตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) (ดูในการทดสอบค่าสถิติคิว, หน้า 31)

3.2.3.2 การทดสอบอาร์ช-แอลเอ็ม (ARCH-LM Test)

การทดสอบอาร์ช-แอลเอ็ม (ARCH-LM Test) ของตัวแบบการ์ช (GARCH Model) เป็นการตรวจสอบว่าส่วนตกค้างในรูปคะแนนมาตรฐาน [๑๔] (Standardized Residual) ของ

[๑๔] ส่วนตกค้างในรูปคะแนนมาตรฐาน (Standardized Residual) ใช้สำหรับการวิเคราะห์ความแปรปรวนของพจน์คลาดเคลื่อน (Variance of Error) ที่มีลักษณะไม่คงที่ (Not Uniform) ซึ่งส่วนตกค้างในรูปคะแนนมาตรฐานสามารถเขียนได้ในรูปของผลหาร (Quotient) ระหว่างส่วนที่เหลือ (Residual) ต่อส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation)

สมการความแปรปรวน (Variance Equation) ไม่มีปัญหาความแปรปรวนของพจน์คลาดเคลื่อนไม่คงที่อย่างมีเงื่อนไข (Conditional Heteroscedasticity) กล่าวคือไม่มีกระบวนการอาร์ช (ARCH Process) หลงเหลืออยู่ในส่วนตกค้างในรูปคะแนนมาตรฐาน (Standardized Residual) ซึ่งหลักเกณฑ์ในการทดสอบและตั้งสมมติฐานว่าง (Null Hypothesis) เหมือนกับการทดสอบอาร์ช-แอลเอ็ม (ARCH-LM Test) ของตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) (ดูในการทดสอบอาร์ช-แอลเอ็ม, หน้า 34)

3.2.3.3 การทดสอบรูปแบบปกติ (Normality Test)

การทดสอบรูปแบบปกติ (Normality Test) ของตัวแบบการ์ช (GARCH Model) เป็นการตรวจสอบว่าส่วนตกค้างในรูปคะแนนมาตรฐาน (Standardized Residual) ของสมการความแปรปรวน (Variance Equation) มีการแจกแจงแบบปกติ (Normal Distribution) ซึ่งหลักเกณฑ์ในการทดสอบและตั้งสมมติฐานว่าง (Null Hypothesis) เหมือนกับการทดสอบรูปแบบปกติ (Normality Test) ของตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) (ดูในการทดสอบรูปแบบปกติ, หน้า 33)

3.2.4 การพยากรณ์ (Forecasting)

เมื่อสร้างสมการค่าเฉลี่ย (Mean Equation) และสมการความแปรปรวน (Variance Equation) รวมทั้งทดสอบคุณสมบัติของส่วนตกค้าง (Residual) แล้ว ขั้นตอนนี้เป็นการทดสอบความสามารถในการพยากรณ์ของตัวแบบการ์ช (GARCH Model) เพื่อหาตัวแบบที่ดีที่สุดเพียงตัวแบบเดียว ในการนำไปเปรียบเทียบความแม่นยำกับตัวแบบประเภทอื่นๆ ซึ่งขั้นตอนนี้มีการดำเนินการคล้ายคลึงกับขั้นตอนการพยากรณ์ของตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) (ดูในการพยากรณ์, หน้า 35)

3.3 การพยากรณ์ด้วยตัวแบบโครงข่ายใยประสาทประดิษฐ์ (ANNs Model)

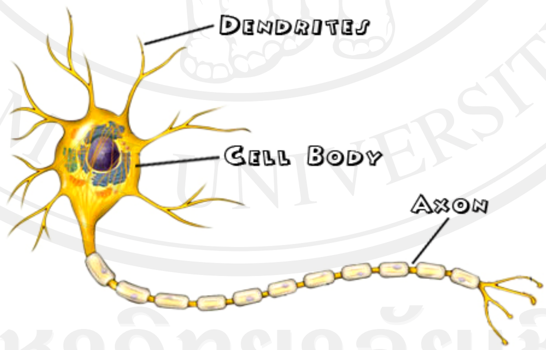
แม้ว่าการวิจัยทางด้านเศรษฐศาสตร์ จะมีเครื่องมือที่ทันสมัยสำหรับการวิเคราะห์อนุกรมเวลา (Time Series) ที่ได้รับการประเมินและปรับปรุง รวมไปถึงการนำไปใช้จริงจนเป็นหลักการปฏิบัติระดับมาตรฐาน อาทิ ตัวแบบบ็อกซ์และเจนคินส์ (Box&Jenkins Model) และตัวแบบที่กำหนดความแปรปรวนของพจน์คลาดเคลื่อนไม่คงที่อย่างมีเงื่อนไข (Modeling Conditional Heteroscedasticity) แต่ในปัจจุบันพบว่าเริ่มมีการนำเสนอแนวคิดเกี่ยวกับตัวแบบโครงข่ายใย

ประสาทประดิษฐ์ (Artificial Neural Networks : ANNs Model) เพื่อประยุกต์ใช้กับงานวิจัยด้านเศรษฐศาสตร์

3.3.1 แนวคิดเบื้องต้นของตัวแบบโครงข่ายใยประสาทประดิษฐ์ (ANNs Model)

โครงข่ายใยประสาทประดิษฐ์ (Artificial Neural Networks : ANNs Model) คือตัวแบบทางคณิตศาสตร์ (Mathematical Model) ที่ระบบการประมวลผลมีการเชื่อมโยงกัน (Connectionist) เป็นโครงข่ายเสมือนการทำงานของเซลล์สมอง ซึ่งวัตถุประสงค์สำหรับการจำลองแบบการทำงานของโครงข่ายในสมองมนุษย์ ก็เพื่อการสร้างตัวแบบที่มีความสามารถเทียบเท่ากับสมองของมนุษย์ กล่าวคือเป็นแบบจำลองที่มีความสามารถในการจดจำรูปแบบ (Pattern Recognization) และการนิรนัยความรู้ (Knowledge Deduction) ได้

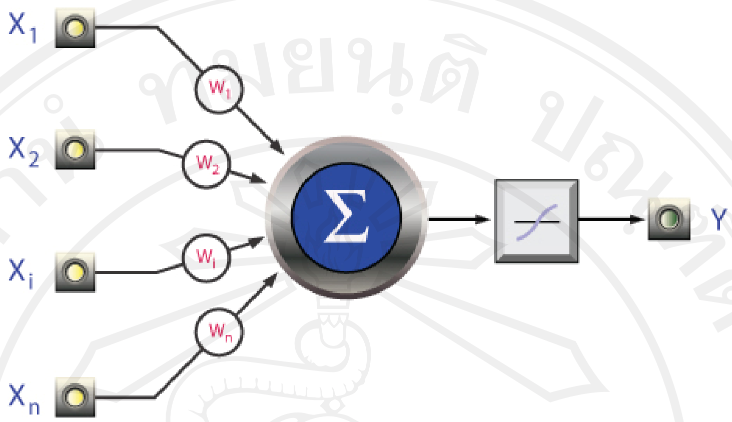
สำหรับการสร้างตัวแบบทางคณิตศาสตร์ใด ๆ ขึ้นมา เพื่อให้มีคุณสมบัติในการเลียนแบบการทำงานของสมองได้นั้น มีความจำเป็นที่จะต้องศึกษาถึงรูปแบบการทำงานจริงๆ ของสมองมนุษย์ในขั้นต้นก่อน โดยสมองของมนุษย์นั้นประกอบด้วยเซลล์ประสาทหรือนิวรอน (Neuron) ประมาณ 100,000 ล้านเซลล์ ดังนั้นเพื่อให้ง่ายต่อการทำความเข้าใจ จึงควรเริ่มต้นจากการศึกษากระบวนการทำงานของเซลล์สมองทีละเซลล์ก่อน



ภาพที่ 3-7 เซลล์สมอง 1 เซลล์

จากภาพที่ 3-7 แสดงให้เห็นถึงการทำงานของเซลล์สมองแต่ละเซลล์ว่า แท้ที่จริงแล้วก็คือการรับและส่งสัญญาณไฟฟ้าชีวภาพ (Bioelectric) ระหว่างกันนั่นเอง โดยที่ใยประสาทนำเข้า (Dendrite) ซึ่งมีลักษณะเป็นแขนงย่อย จะทำหน้าที่เป็นตัวรับสัญญาณที่ถูกส่งเข้ามาจากเซลล์สมองอื่น จากนั้นสัญญาณดังกล่าวจะถูกส่งเข้ามาประมวลผลที่ตัวเซลล์ (Cell Body or Soma) จากนั้น

แกนประสาทนำออก (Axon) ที่มีลักษณะเป็นแขนงใหญ่จะเป็นตัวส่งสัญญาณที่ได้รับการประมวลผลแล้วไปยังใยประสาทนำเข้า (Dendrites) ของเซลล์สมองอื่นๆ ต่อไป



ภาพที่ 3-8 การสร้างระบบประมวลผลทางคณิตศาสตร์โดยจำลองจากเซลล์สมอง 1 เซลล์

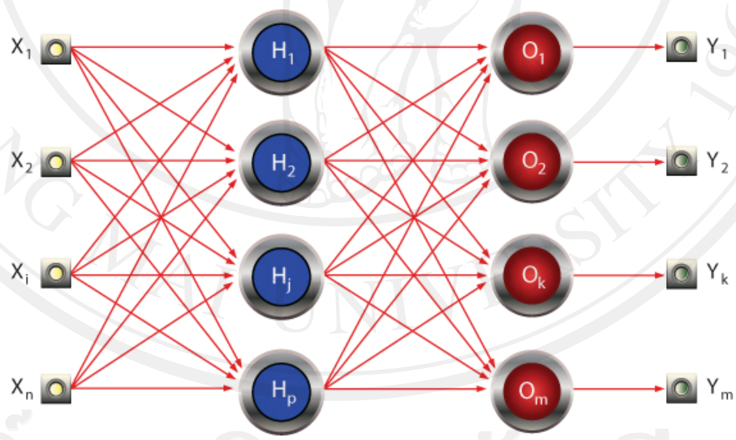
ซึ่ง Warren McCulloch และ William Pitt (1943) ได้สร้างตัวแบบทางคณิตศาสตร์จากกระบวนการทำงานของเซลล์สมอง 1 เซลล์ ดังภาพที่ 3-8 โดยเริ่มต้นมีข้อมูลนำเข้า (Input) คือ x ทั้งหมด n ข้อมูล ซึ่งจะถูกส่งไปรวมกันอยู่ที่ฟังก์ชันผลรวม (Summation Function : Σ) อย่างไรก็ตามข้อมูลนำเข้า (Input) แต่ละตัวอาจมีความสำคัญมากน้อยแตกต่างกันไป ดังนั้นจึงต้องมีการกำหนดค่าถ่วงน้ำหนัก (Weight) ให้แก่ข้อมูลนำเข้า (Input) แต่ละตัวด้วยค่า w จากนั้นจะถูกส่งไปประมวลผลที่ฟังก์ชันการแปลงค่า (Transfer Function) โดยข้อมูลที่ได้รับการแปลงแล้วจะถูกส่งออกมาเป็นข้อมูลส่งออก (Output) ซึ่งก็คือ y นั่นเอง

3.3.2 สถาปัตยกรรมโครงข่าย (Architecture of Networks)

จากหัวข้อที่ผ่านมานั้น แสดงให้เห็นถึงกระบวนการทำงานของเซลล์สมองเพียงเซลล์เดียว ซึ่งในความเป็นจริงแล้ว เซลล์สมองของมนุษย์มีการเชื่อมต่อกันเป็นโครงข่าย โดยที่เซลล์ประสาท (Neuron) แต่ละตัวก็จะนำใยประสาทนำเข้า (Dendrites) ของตัวเองไปเชื่อมกับแกนประสาทนำออก (Axon) ของเซลล์ประสาทอื่นเพื่อรับสัญญาณเข้ามา และนำใยประสาทนำเข้า (Dendrites) ของตัวเองไปเชื่อมกับแกนประสาทนำออก (Axon) ของเซลล์ประสาทอื่นเพื่อส่งส่งสัญญาณออกไป ซึ่งกระบวนการทำงานที่ซับซ้อนขึ้นของเซลล์สมองแสดงดังภาพที่ 3-9

ภาพที่ 3-9 กระบวนการเชื่อมโยงของเซลล์สมองแบบเป็นโครงข่าย

โดยกระบวนการทำงานของสมองที่มีลักษณะการเชื่อมโยงเป็นโครงข่ายนั้น สามารถจำลองเป็นระบบประมวลผลทางคณิตศาสตร์ที่ประกอบด้วยชั้นข้อมูลนำเข้า (Input Layer) ชั้นซ่อนเร้น (Hidden Layer) ชั้นข้อมูลส่งออก (Output Layer) และเส้นการต่อเชื่อม (Interconnection) ได้ดังภาพที่ 3-10



ภาพที่ 3-10 ระบบประมวลผลทางคณิตศาสตร์ที่จำลองการทำงานของเซลล์สมองที่เชื่อมโยงเป็นโครงข่าย

สำหรับกระบวนการทำงานของตัวแบบโครงข่ายใยประสาทประดิษฐ์ (ANNs Model) ประกอบด้วยส่วนหลัก ๆ 4 ส่วนคือ ชั้นข้อมูลนำเข้า (Input Layer) ชั้นซ่อนเร้น [๑๕] (Hidden Layer)

[๑๕] ชั้นซ่อนเร้น (Hidden Layer) มีหน้าที่รับสัญญาณที่ถูกส่งมาจากชั้นข้อมูลนำเข้า (Input Layer) เพื่อนำมาประมวลผลและส่งต่อไปยังชั้นข้อมูลส่งออก (Output Layer) โดยชั้นซ่อนเร้น (Hidden Layer) มีลักษณะเสมือนกล่องดำ (Black Box) กล่าวคือพารามิเตอร์ต่าง ๆ ที่ถูกสร้างขึ้นไม่สามารถอธิบายความหมายได้เหมือนกับการวิเคราะห์อนุกรมเวลา (Time Series Analysis) เนื่องจากพารามิเตอร์ดังกล่าวมีหน้าที่ในการคำนวณเท่านั้น

ชั้นข้อมูลส่งออก (Output Layer) และเส้นการต่อเชื่อม (Interconnections) ซึ่งรายละเอียดของส่วนประกอบต่าง ๆ มีดังนี้

3.3.2.1 ชั้นข้อมูลนำเข้า (Input Layer)

สำหรับการกำหนดรูปแบบข้อมูลนำเข้า (Input) นั้น ไม่ได้มีหลักการตายตัวแน่นอน ซึ่ง Milton Boyds (1997) เสนอว่าการสร้างชุดการเรียนรู้ (Learning Set) เป็นขั้นตอนที่มีความสำคัญมาก โดยข้อมูลนำเข้า (Input) ที่เลือกมาต้องมีความสามารถในการอธิบายค่าเป้าหมาย (Target Response) ได้เป็นอย่างดี

3.3.2.2 ชั้นซ่อนเร้น (Hidden Layer)

ชั้นซ่อนเร้น (Hidden Layer) เป็นชั้นที่อยู่ระหว่างชั้นข้อมูลนำเข้า (Input Layer) และชั้นข้อมูลส่งออก (Output Layer) โดยโครงข่าย (Networks) จะสามารถประมวลผลหาฟังก์ชันที่เหมาะสมได้หากมีชั้นซ่อนเร้น (Hidden Layer) ที่มากพอ อย่างไรก็ตามหากมีชั้นซ่อนเร้นมากเกินไป นอกจากจะไม่ก่อให้เกิดประสิทธิภาพที่สูงขึ้นแล้ว ยังทำให้โครงข่ายเสียเวลาในการเรียนรู้นานอีกด้วย โดย Kurt Hornik, et al (1989) ได้นำเสนอทฤษฎีบทลักษณะทั่วไปเกือบจะเหมือน [๑๖] (Universal Approximation Theorem) ว่าตัวแบบโครงข่ายใยประสาทประดิษฐ์ (ANNs Model) ที่มีชั้นซ่อนเร้น (Hidden Layer) เพียง 1 ชั้น ก็เพียงพอแล้วสำหรับสร้างฟังก์ชันที่มีข้อผิดพลาดน้อยมาก จนแทบจะเป็นแบบเดียวกับฟังก์ชันจริงเลย

ภาพที่ 3-11 ระบบการประมวลผลที่เหมือนกล่องดำ

สำหรับการสร้างระบบพยากรณ์ด้วยตัวแบบโครงข่ายใยประสาทประดิษฐ์ (ANNs Model) ส่วนที่เป็นจุดเริ่มต้น (Origin) และส่วนที่เป็นจุดปลาย (Destination) ซึ่งก็คือข้อมูลนำเข้า (Input) และข้อมูลส่งออก (Output) ตามลำดับ เป็นส่วนที่โครงข่าย (Networks) สามารถแสดงให้เห็นเป็นตัวเลขได้ชัดเจน ขณะที่ชั้นซ่อนเร้น (Hidden Layer) ซึ่งทำหน้าที่เสมือนหน่วยประมวลผลกลาง (Central Processing Unit : CPU) ในเครื่องคอมพิวเตอร์ ไม่สามารถแสดงค่าใด ๆ ออกมาเป็นตัวเลขได้ดังภาพที่ 3-11

[๑๖] ลักษณะทั่วไปเกือบจะเหมือน (Universal Approximation) เป็นการพิสูจน์ทางคณิตศาสตร์ที่ขยายมาจากทฤษฎีบทไซเบนโก (Cybenko Theorem) โดยสรุปได้ว่าโครงข่ายใยประสาทประดิษฐ์ที่ใช้การส่งข้อมูลไปข้างหน้าเพียงอย่างเดียวแบบหลายชั้น (Multilayer Feed Forward Neural Network : MLFNN) สามารถสร้างรูปแบบที่ดีที่สุดได้ โดยใช้ชั้นซ่อนเร้นเชิงเดี่ยว (Single Hidden Layer) หรือกำหนดชั้นซ่อนเร้น (Hidden Layer) เพียงหนึ่งชั้น

นอกจากนี้ Biemo Soemardi (1997) ได้เสนอว่าจำนวนนิวรอน [๑๗] (Neuron) ในชั้นนี้ไม่ควรกำหนดให้มากจนเกินไป เนื่องจากจะทำให้โครงข่ายพบกับปัญหาเฉพาะเจาะจงกับข้อมูล [๑๘] (Over Fitting)

3.3.2.3 ชั้นข้อมูลส่งออก (Output Layer)

สำหรับชั้นข้อมูลส่งออก (Output Layer) ทำหน้าที่ในการรับสัญญาณที่ถูกส่งมาจากชั้นซ่อนเร้น (Hidden Layer) แล้วนำมาประมวลผลยังฟังก์ชันผลรวม (Summation Function) จากนั้นจึงถูกนำไปแปลงค่าสัญญาณออกมาเป็นค่าพยากรณ์ที่ฟังก์ชันการแปลงค่า (Transfer Function)

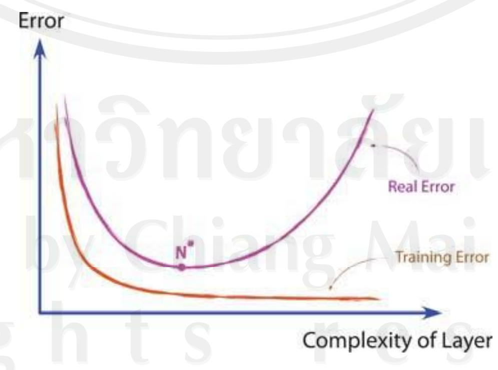
3.3.2.4 เส้นการต่อเชื่อม (Interconnections)

เส้นการต่อเชื่อมทำหน้าที่สำหรับประสานองค์ประกอบต่างๆ ของโครงข่ายเข้าไว้ด้วยกัน โดยเส้นเชื่อมแต่ละเส้นจะมีค่าถ่วงน้ำหนัก (Weight) ของตัวมันเอง ซึ่งค่าน้ำหนักเริ่มต้น (Initial Weight) ถูกกำหนดแบบสุ่ม จากนั้นเมื่อโครงข่าย (Networks) เริ่มทำการฝึกสอน (Training) ค่าน้ำหนักเหล่านี้จึงถูกปรับ (Adjusted) ให้มีความเหมาะสม

---

[๑๗] นิวรอน (Neuron) ในสาขาวิชาวิศวกรรมศาสตร์เรียกว่าบัพ (Node) มีหน้าที่เป็นจุดเชื่อมต่อ โดยอาจเป็นจุดกระจายซ้ำหรือจุดปลายทางของการส่งข้อมูลก็ได้

[๑๘] ปัญหาเฉพาะเจาะจงกับข้อมูล (Over Fitting) คือพฤติกรรมการเรียนรู้ของโครงข่าย (Network) ที่ยึดติดกับนิวรอน (Neuron) ที่มีจำนวนมากเกินไป จนไม่สามารถหาคำตอบที่ดีได้



ภาพที่ 3-12 ปัญหาเฉพาะเจาะจงกับข้อมูล

จากภาพที่ 3-12 สำหรับการสร้างฟังก์ชันการฝึกสอน (Training Function) ใดๆ ยิ่งโครงข่ายมีความซับซ้อนมาก จะส่งผลให้ค่าคลาดเคลื่อน (Error) น้อยลงตามไปด้วย อย่างไรก็ดีเมื่อนำรูปแบบดังกล่าวไปใช้เพื่อขยายผลจริงเช่นการพยากรณ์ พบว่าโครงข่ายที่มีความซับซ้อนเท่ากับ N^* เท่านั้น ที่ให้ค่าคลาดเคลื่อน (Error) น้อยที่สุด หากโครงข่ายมีความซับซ้อนมากหรือน้อยกว่านี้จะทำให้ประสิทธิภาพในการพยากรณ์ลดลง

3.3.2.5 ฟังก์ชันผลรวม (Summation Function)

ฟังก์ชันผลรวม (Summation Function) เป็นฟังก์ชันที่ทำหน้าที่สำหรับในการหาค่าเฉลี่ย ค่าถ่วงน้ำหนัก (Weight) ของทุก ๆ นิวรอน (Neuron) ที่เชื่อมต่อกัน กล่าวคือเมื่อโครงข่าย (Networks) เริ่มทำงานข้อมูลนำเข้า (Input) แต่ละนิวรอนจะถูกคูณเข้ากับค่าถ่วงน้ำหนัก จากนั้นฟังก์ชันผลรวม (Summation) จะรวมผลลัพธ์ของทุกนิวรอนเข้าไว้ด้วยกัน ซึ่งฟังก์ชันผลรวมจะมีอยู่ในชั้นซ่อนเร้น (Hidden Layer) และชั้นข้อมูลส่งออก (Output Layer)

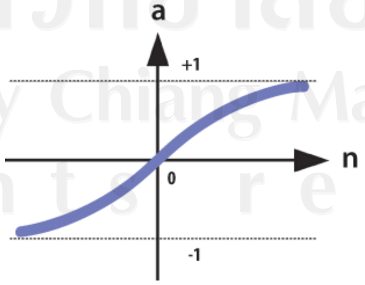
3.3.2.6 ฟังก์ชันการแปลงค่า (Transfer Function)

ฟังก์ชันการแปลงค่า [๑๕] (Transfer Function) คือความสัมพันธ์ระหว่างระดับการกระตุ้นภายในนิวรอน (Neuron) กับค่าเป้าหมาย (Target Response) โดยการศึกษาครั้งนี้จะใช้ฟังก์ชันการแปลงค่า 2 ประเภทดังนี้

1) ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ซิกมอยด์

โดยทั่วไปการกำหนดฟังก์ชันการแปลงค่า (Transfer Function) ในชั้นซ่อนเร้น (Hidden Layer) มักจะใช้ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ซิกมอยด์ (Hyperbolic Tangent Sigmoid : Tan-Sigmoid Function) เนื่องจากมีลักษณะไม่เป็นเชิงเส้น มีความต่อเนื่อง สามารถหาอนุพันธ์ได้ และง่ายต่อการคำนวณ ซึ่งฟังก์ชันแทนซิกมอยด์ (Tan-Sigmoid Function) จะแปลงค่าและส่งไปยังชั้น (Layer) ต่อๆ ไป โดยให้ช่วงข้อมูลตั้งแต่ -1 ถึง 1 โดยฟังก์ชันแทนซิกมอยด์ (Tan-Sigmoid Function) มีรูปแบบอัลกอริทึม (Algorithm) ดังนี้

$$\text{tansig}(n) = \frac{2}{[1+\exp(-2n)]-1}$$



ภาพที่ 3-13 ฟังก์ชันไฮเพอร์โบลิกแทนเจนต์ซิกมอยด์

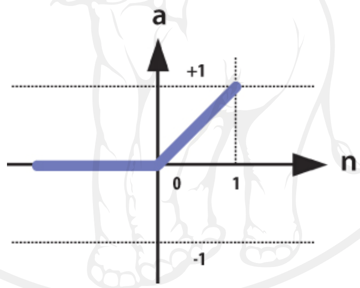
---

[๑๕] ฟังก์ชันการแปลงค่า (Transfer Function) บางครั้งเรียกว่าฟังก์ชันการกระตุ้น (Activation Function)

2) ฟังก์ชันโพสลิน

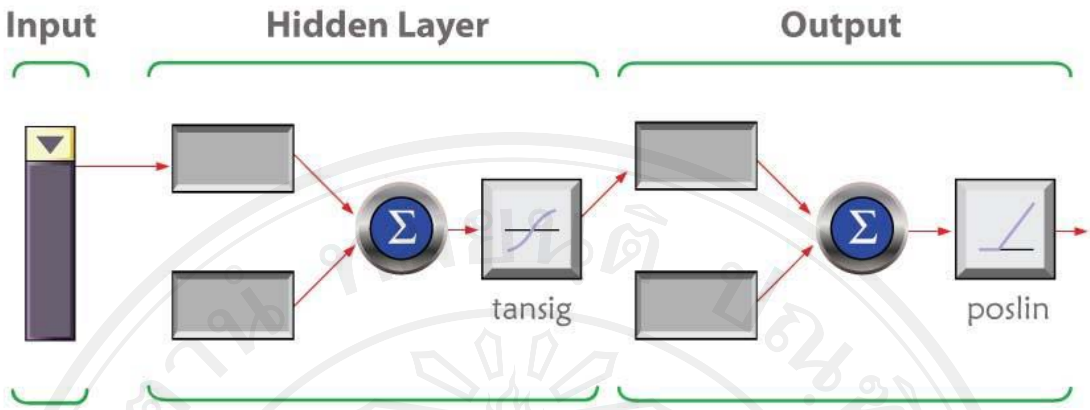
การกำหนดฟังก์ชันการแปลงค่า (Transfer Function) ในชั้นข้อมูลส่งออก (Output Layer) เป็นขั้นตอนที่มีความสำคัญมาก เนื่องจากข้อมูลที่ถูกส่งออกจากชั้นดังกล่าวคือค่าพยากรณ์ ดังนั้นเพื่อให้การแปลงข้อมูลสอดคล้องกับความเป็นจริง จึงเลือกใช้ฟังก์ชันโพสลิน (Positive Linear Transfer Function : Poslin Function) เนื่องจากสามารถแปลงข้อมูลและส่งออกมาเป็นค่าพยากรณ์ที่เป็นจำนวนเต็มบวกเท่านั้น โดยมีรูปแบบอัลกอริทึม (Algorithm) ดังนี้

$$\mathrm{poslin}(n) = \begin{cases} n & \text{เมื่อ } n \geq 0 \\ 0 & \text{เมื่อ } n \leq 0 \end{cases}$$



ภาพที่ 3-14 ฟังก์ชันโพสลิน

จากภาพที่ 3-15 แสดงถึงสถาปัตยกรรมโครงข่าย (Architecture of Networks) ซึ่งประกอบด้วยชั้นข้อมูลนำเข้า (Input Layer) ชั้นซ่อนเร้น (Hidden Layer) ชั้นข้อมูลส่งออก (Output layer) เส้นการต่อเชื่อม (Interconnections) ฟังก์ชันผลรวม (Summation Function) และฟังก์ชันการแปลงค่า (Transfer Function)



ภาพที่ 3-15 สถาปัตยกรรมโครงข่าย

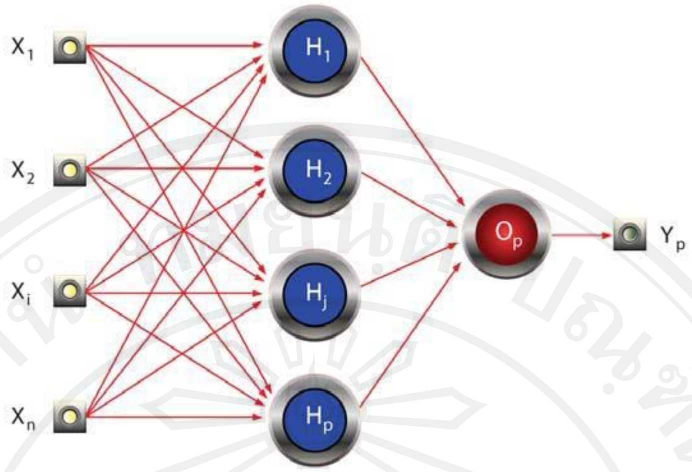
3.3.3 ขั้นตอนการเรียนรู้ (Learning Algorithms)

กระบวนการเรียนรู้ของโครงข่าย (Networks) คือการนำเอาข้อผิดพลาดจากการคำนวณในครั้งก่อน มาปรับแก้ค่าถ่วงน้ำหนัก (Weight) สำหรับการคำนวณในรอบต่อไป โดยมีวัตถุประสงค์เพื่อให้คำตอบที่ได้มีความถูกต้องมากขึ้น ซึ่งแบบการเรียนรู้ (Learning Paradigm) ประกอบด้วยวิธีการต่าง ๆ ดังนี้

3.3.3.1 การส่งข้อมูลไปข้างหน้าเพียงอย่างเดียวแบบหลายชั้น (Multilayer Feed Forward)

วิธีการนี้เป็นการส่งข้อมูลนำเข้า (Input) เข้าไปยังโครงข่าย (Networks) ผ่านช่องทางต่างๆ โดยเป็นการส่งไปข้างหน้าเพียงอย่างเดียว โดยไม่มีการส่งข้อมูลย้อนหลังหรือส่งข้อมูลไปยังนิวรอน (Neuron) อื่นๆ ที่อยู่ในชั้น (Layer) เดียวกัน กล่าวคือเมื่อโครงข่าย (Networks) เริ่มทำงาน ข้อมูลนำเข้า (Input) ที่มีสมาชิกทั้งหมด n ตัว คือ $X = \{X_1 , X_2 , ... , X_i , X_n\}$ จะถูกส่งต่อไปข้างหน้าผ่านเส้นเชื่อมโยง (Interconnection) แต่ละเส้นที่มีการกำหนดค่าน้ำหนัก (Weight) เพื่อนำไปประมวลผลยังชั้นซ่อนเร้น (Hidden Layer) ที่มีนิวรอน (Neuron) ทั้งหมด p ตัวคือ $H = \{H_1 , H_2 , ... , H_j , H_p\}$ ซึ่งค่าน้ำหนักทั้งหมดในชั้นนี้คือ $W_{qr}^{H} = \{W_{11}^{H} , W_{12}^{H} , W_{1j}^{H} , ... , W_{ni}^{H} , W_{np}^{H}\}$ โดยที่ q = 1 , 2 , i , ... , n และ r = 1 , 2 , j , ... , p

เมื่อข้อมูลถูกประมวลผลแล้วในชั้นซ่อนเร้น (Hidden Layer) จะถูกส่งออกไปประมวลผลต่อในชั้นข้อมูลส่งออก (Output Layer) โดยที่เส้นเชื่อมทั้งหมดในชั้นนี้มีค่าน้ำหนักคือ $W_{rs}^{O} = \{W_{1s}^{O} , W_{2s}^{O} , ... , W_{ps}^{O}\}$ เมื่อกระบวนการต่างๆ ดำเนินการจนครบแล้ว จะได้ค่าข้อมูลส่งออก (Output) คือ Y นั่นเอง ดังภาพที่ 3-16



ภาพที่ 3-16 การส่งข้อมูลไปข้างหน้าเพียงอย่างเดียวแบบหลายชั้น

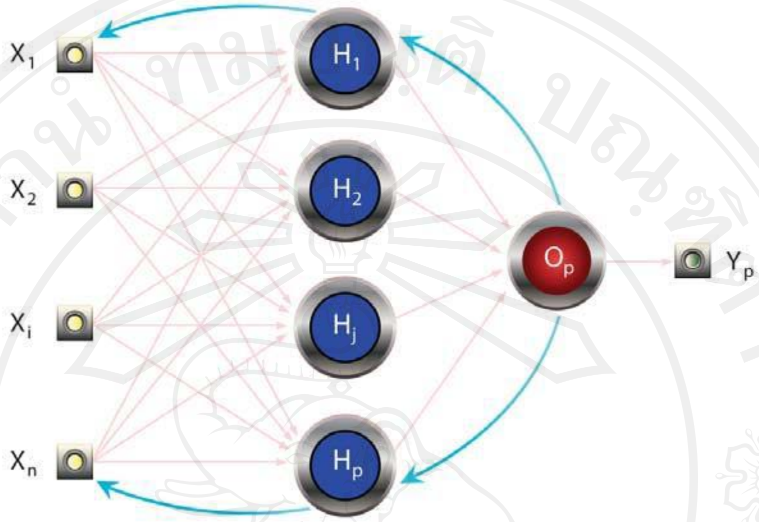
3.3.3.2 การเรียนรู้แบบมีการสอน (Supervised Learning)

เพื่อให้วิธีการส่งข้อมูลไปข้างหน้าเพียงอย่างเดียวแบบหลายชั้น (Multilayer Feed Forward) มีประสิทธิภาพมากขึ้น Laurene Fausett (1994) จึงเสนอให้ใช้งานควบคู่กับการเรียนรู้แบบมีการสอน (Supervised Learning) กล่าวคือมีการสร้างแบบการเรียนรู้ (Learning Set) ที่มีการตรวจคำตอบ เพื่อให้โครงข่ายสามารถปรับตัวได้ โดยเมื่อข้อมูลนำเข้า (Input) ถูกส่งเข้าไปในโครงข่าย (Network) ซึ่งภายในจะมีแบบการฝึกสอน (Training) ที่มีลักษณะเป็นฟังก์ชันสำหรับการคำนวณ เมื่อได้ค่าผลลัพธ์ออกมา จึงนำไปเปรียบเทียบกับค่าที่แท้จริง (True Value) เพื่อหาค่าความคลาดเคลื่อน (Error) ซึ่งค่าคลาดเคลื่อนดังกล่าวจะถูกส่งกลับเข้าไปยังโครงข่าย พร้อมกับการปรับแก้ค่าถ่วงน้ำหนัก (Weight) เพื่อคำนวณผลลัพธ์ใหม่อีกครั้งจนกว่าจะได้ค่าความคลาดเคลื่อนที่ต่ำที่สุด ซึ่งกระบวนการที่ส่งผ่านข้อมูลย้อนกลับดังกล่าวเรียกว่าอัลกอริทึมแพร่ย้อนกลับ (Backpropagation Algorithm)

3.3.3.3 วิธีแพร่ย้อนกลับแบบแอลเอ็ม (LM Backpropagation)

ในปัจจุบันมีการพัฒนาอัลกอริทึมแพร่ย้อนกลับ (Backpropagation Algorithm) มากมายหลายวิธี อาทิ วิธีความชันสังยุคของเฟล็ตท์เชอร์-รีฟส์ (Fletcher-Reeves Conjugate Gradient : CGF) วิธีเกาส์เซียน-นิวตัน (Gaussian-Newton Method : BGF) วิธีเคลื่อนลงตามความชัน (Gradient Descent : GD) วิธีแพร่ย้อนกลับแบบแอลเอ็ม (LM Backpropagation) ฯลฯ ซึ่งการวิจัยในครั้งนี้จะเลือกใช้วิธีแพร่ย้อนกลับแบบแอลเอ็ม (LM Backpropagation) สำหรับการปรับแก้ค่าถ่วงน้ำหนัก (Weight) ของโครงข่าย เนื่องจากเป็นวิธีที่ได้รับการทดสอบแล้วว่าสามารถปรับค่า

น้ำหนักจนเกิดความผิดพลาด (Error) น้อยที่สุด รวมทั้งใช้ระยะเวลาในการลู่เข้าสู่เป้าหมายได้รวดเร็วที่สุด



ภาพที่ 3-17 การทำงานของอัลกอริทึมแพร่ย้อนกลับ

วิธีแพร่ย้อนกลับแบบแอลเอ็ม (LM Backpropagation) ถูกพัฒนาโดย Kenneth Levenberg และ Donald Marquandt (1978) โดยนำเอาข้อดีของวิธีเคลื่อนลงตามความชัน (Gradient Descent : GD) และวิธีเกาส์เซียน-นิวตัน (Gaussian-Newton Method : BGF) มารวมกัน ซึ่งอัลกอริทึมแพร่ย้อนกลับ (Backpropagation Algorithm) ดังกล่าว จะทำการปรับแก้ค่าถ่วงน้ำหนัก (Weight) ตั้งแต่ชั้นข้อมูลส่งออก (Output Layer) แล้วปรับย้อนถอยหลังกลับมาเรื่อยๆ จนถึงชั้นข้อมูลนำเข้า (Input Layer) ดังภาพที่ 3-17 โดยสมการที่ใช้คำนวณเขียนได้ดังนี้

$$JJ = J(W) \times J(W)$$

$$JW = J(W) \times E$$

$$\Delta W = (-JJ + \lambda I) \times JE\text{-}1$$

โดยที่ J คือ เมทริกซ์จาโคเบียน (Jacobian Matrix)

E คือ เวกเตอร์ค่าคลาดเคลื่อน (Vector Error)

W คือ ค่าถ่วงน้ำหนัก (Weight)

I คือ เมทริกซ์เอกลักษณ์ (Identity Matrix)

λ คือ อัตราการเรียนรู้ (Learning Rate)

ถ้าอัตราการเรียนรู้ (Learning Rate) มีค่าต่ำ ๆ รูปแบบของอัลกอริทึมแพร่ย้อนกลับ (Backpropagation Algorithm) จะกลายเป็นวิธีเกาส์เซียน-นิวตัน (Gaussian-Newton Method : BGF) ขณะที่อัตราการเรียนรู้ที่มีค่ามาก ๆ รูปแบบจะเปลี่ยนเป็นวิธีเคลื่อนลงตามความชัน (Gradient Descent : GD) ซึ่งอัตราการเรียนรู้ (Learning Rate) จะถูกปรับเปลี่ยนเองโดยอัตโนมัติ เพื่อให้ลู่เข้าสู่เป้าหมายได้อย่างถูกต้องและรวดเร็ว