บทที่ 3
กรอบแนวคิดและระเบียบวิธีวิจัย

3.1 กรอบแนวคิด

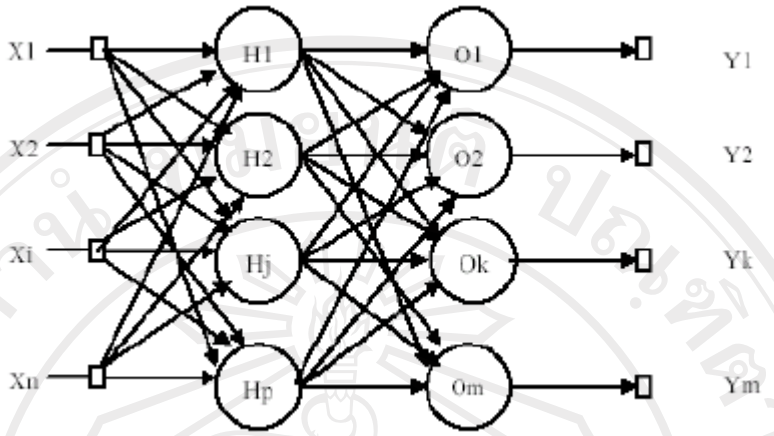
ในการศึกษาครั้งนี้ต้องการที่จะทำการเปรียบเทียบความแม่นยำในการพยากรณ์ราคาน้ำมันดิบ ซึ่งข้อมูลมีความเป็นอนุกรมเวลา (Time Series) ดังนั้นจึงได้นำแนวคิดและทฤษฎีต่างๆ ได้แก่ แบบจำลอง Neural Networks การทดสอบความนิ่ง (Stationary) ของข้อมูล และการทดสอบ Unit Root , แบบจำลอง Autoregressive intregrated moving average (ARIMA) , แบบจำลอง Autoregressive Conditional Heteroskedasticity (ARCH), แบบจำลอง Generalized Autoregressive Conditional Heteroskedasticity (GARCH), แบบจำลอง GARCH-in-Mean (GARCH-M), เพื่อนำมาพยากรณ์ราคาน้ำมันดิบและการตรวจสอบความแม่นยำของแบบจำลองต่างๆ ดังต่อไปนี้

3.1.1 Artificial Neural Networks แบบ Multilayer Feed Forward

การพยากรณ์โดยแบบจำลอง Artificial Neural Networks หรือโครงข่ายประสาทเทียม เป็นการจำลองวิธีการทำงานของสมองมนุษย์ ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำแบบรูป(Pattern Recognition) และการอุปมานความรู้ (Knowledge deduction) โดยในการศึกษาครั้งนี้จะใช้แบบจำลอง Multi Layer Feed Forward (MLFF) ซึ่งประกอบด้วย 3 Layers คือ Input layer Hidden layer และ Output layer และในแต่ละชั้นจะทำการปรับค่าน้ำหนักโดยการคำนวณแบบแพร่ย้อนกลับ (Back Propagation) สำหรับวิธีการทำงานของ MLFF นั้นจะมีรับข้อมูลนำเข้า (input) ประมวลผลในนิวรอลเข้าสู่ชั้น Hidden layer และชั้น Output layer ตามลำดับ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

รูปที่ 3.1 แสดงแบบจำลอง Multilayer Feed Forward แบบการแพร่ย้อนกลับ



องค์ประกอบของนิวรอลเน็ตเวิร์คแบบ Multilayer Feed Forward

- แบบจำลองนิวรอลเน็ตเวิร์คมีองค์ประกอบ 6 ส่วนดังต่อไปนี้

1. Input

Input มีจำนวน n ตัว คือ $X = \{ X_1, X_2, X_3, \ldots, X_i, \ldots, X_n \}$

2. Output

Output มีจำนวน m ตัว คือ $Y = \{ Y_1, Y_2, Y_3, \ldots, Y_k, \ldots, Y_m \}$

3. Neuron ใน Hidden Layer

Neuron ใน Hidden Layer มี p ตัว คือ $H = \{H_1, H_2, H_3, \ldots, H_j, \ldots, H_p\}$

4. Neuron ใน Output Layer

Neuron ใน Output Layer มี m ตัว คือ $O = \{ O_1, O_2, O_3, \ldots, O_k, \ldots, O_m \}$

5. ค่าน้ำหนักจากชั้น Input สู่ชั้น Hidden Layer

ค่าน้ำหนักจากชั้น Input สู่ชั้น Hidden Layer สำหรับ Neuron แต่ละตัวในชั้น Hidden layer มีจำนวน n+1 ตัวคือ $W_{ij}^{H} = \left\{W_{0j}^{H}, W_{1j}^{H}, W_{2j}^{H}, \ldots, W_{ij}^{H}, \ldots, W_{nj}^{H}\right\}$

ดังนั้น เมื่อมีจำนวนInputจำนวน n ตัว และมีนิวรอลใน Hidden Layerจำนวน p ตัว แล้วจะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ (n+1)p ตัว

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

6. ค่าน้ำหนักจากชั้น Hidden Layer สู่ชั้น Output

ค่าน้ำหนักจากชั้น Hidden Layer สู่ชั้น Output สำหรับ Neuron แต่ละตัว ในชั้น Output มีจำนวน m+1 ตัว คือ $W_{ik}^{O} = \{W_{0k}^{O}, W_{1k}^{O}, W_{2k}^{O}, ..., W_{ik}^{O}, ..., W_{nk}^{O}\}$

ดังนั้น เมื่อมีจำนวนนิวรอลใน Hidden Layer จำนวน p ตัว และมี Output จำนวน m ตัวแล้วจะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ (p+1)m ตัว

- การคำนวณแบบแพร่ย้อนกลับ (Back Propagation)

Back-propagation เป็นอัลกอริทึมที่นิยมนำมาใช้กับ MLFF มีกฎการปรับค่าน้ำหนักที่ต่างกันใน Layer แต่ละชั้น โดยการปรับค่าขึ้นอยู่กับความแตกต่างของผลลัพธ์ (output) กับค่าผลลัพธ์(output) ที่เกิดขึ้นจริง ซึ่งทั่วไปแล้วการปรับน้ำหนักใช้กฎการปรับค่าน้ำหนักของ Gradient Descent ดังต่อไปนี้

จากกฎการปรับค่าน้ำหนักของ Gradient Descent

$$\Delta w = \eta(t-y)[\nabla_w f(a)]$$

โดยกำหนด $y = f(a) = f(w^t x) = \frac{1}{1+e^{-a}}$ เมื่อ $a = w^T x$

ดังนั้น

$$\nabla_w f(a) = \begin{pmatrix} \partial f(a)/\partial w_0 \\ \partial f(a)/\partial w_1 \\ . \\ . \\ . \\ \partial f(a)/\partial w_n \end{pmatrix} = \frac{\partial f(a)}{\partial w} = \frac{\partial f(a)}{\partial a}\frac{\partial a}{\partial w}$$

$$= f'\frac{\partial(w^T x)}{\partial w} = f'\frac{\sum_{i=0}^{n} w_i x_i}{\partial w_i} = f' x_i$$

$$= f'\begin{bmatrix} x_0 \\ x_1 \\ . \\ . \\ . \\ x_n \end{bmatrix} = f' x$$

จึงสรุปได้ว่า $\Delta w = \eta(t-y) f' x$ (จิตติ ตันเสนีย์, 2549)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

การปรับค่าน้ำหนักในชั้น Output

ค่าน้ำหนักแต่ละตัวที่เชื่อมโยงจาก Neuron ใน Hidden Layer มายังชั้น Output คือ W_{jk}^{H} จึงจะมีการปรับค่าน้ำหนักดังนี้

$$\Delta W_{jk}^{O} = \eta(t_k - y_k) f' x_{jk}^{O}$$

แต่ Input ของชั้น Output คือ Output ของชั้น Hidden Layer ดังนั้น

$$\Delta W_{jk}^{O} = \eta(t_k - y_k) f' y_{jk}^{H}$$

เมื่อ f' คือ derivative ของ transformation function จากชั้น Hidden Layer สู่ชั้น Output

การปรับค่าน้ำหนักในชั้น Hidden Layer

ค่าน้ำหนักแต่ละตัวที่เชื่อมโยงจาก input มายัง Neuron ใน Hidden Layer คือ W_{ij}^{H} จึงจะมีการปรับน้ำหนักดังนี้

$$\Delta W_{ij}^{H} = \eta(t_j - y_j) g' x_{ij}^{H}$$

เมื่อ g' คือ derivative ของ transformation function จากชั้น Input สู่ชั้น Hidden Layer (อรทัย เมืองใจ, 2548: 22)

แต่ปัญหาที่พบก็คือ เราไม่สามารถหาค่า t_j ได้เพราะว่าในชั้น Hidden layer เราไม่ทราบค่าของ y_j ที่นิวรอลแต่ละตัวส่งออกมาว่าผิดหรือถูกเราจึงไม่สามารถหา error ได้ดังนั้นจึงได้มีการสร้างค่า error เทียมออกมาโดยสมมติให้ นิวรอลที่ส่งข้อมูลมาชั้นผลลัพธ์ (output) มากก็ต้องรับผิดชอบค่า error ที่เกิดขึ้นจาก (t-y) มากตาม ดังนั้นความรับผิดชอบสำหรับค่า Output ที่นิวรอลในชั้น Hidden layer ตัวที่ j จะเป็นผลรวมของค่ารับผิดชอบของ Output ทั้งหมดจึงสามารถแก้ปัญหาการปรับค่าน้ำหนักในชั้น Hidden Layer ได้ดังนี้

$$\Delta W_{ij}^{H} = \eta(t_j - y_j) g' x_{ij}^{H}$$

โดยที่ $t_j - y_j = \sum_{k=1}^{m} W_{jk}^{0}(t_k - y_k) f'$ ซึ่งเป็นผลรวมของความรับผิดชอบที่มีต่อ Output แต่ละตัว

ดังนั้นจึงจะได้กฎการปรับค่าน้ำหนักจากชั้น Input มายังชั้น Output ดังนี้

$$\Delta W_{ij}^{H} = \eta\left(\sum_{k=1}^{m} W_{jk}^{0}(t_k - y_k) f'\right) g' x_{ij}^{H}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

การกำหนดจำนวนนิวรอลในชั้นซ่อนเร้น

การค้นหาแบบจำลองที่ดีที่สุด ใช้การเปลี่ยนจำนวนนิวรอลในชั้นซ่อนเร้น ด้วยหลักการหาค่าต่ำสุดแบบ Quadratic interpolation ดังนี้

Quadratic interpolation เป็นการหาแบบจำลองที่ดีที่สุด ในการกำหนดจำนวนนิวรอลในชั้นซ่อนเร้นโดยมีวิธีการหาค่าต่ำสุดด้วยวิธี Quadratic Interpolation ซึ่งจะทำได้เมื่อมีข้อมูลจำนวน 3 ชุดและมั่นใจว่าข้อมูลทั้งสามชุดอยู่บนฟังก์ชั่น Quadratic เดียวกัน การมีข้อมูล 3 ชุดจะเพียงพอต่อการหา Unique Solution สำหรับฟังก์ชั่น Quadratic ดังนี้

เมื่อมีข้อมูลดังต่อไปนี้

ชุดที่ $1 : \{x, y\} = \{\alpha, f(\alpha)\}$

ชุดที่ $2 : \{x, y\} = \{\beta, f(\beta)\}$

ชุดที่ $3 : \{x, y\} = \{\gamma, f(\gamma)\}$

ที่จุด $\alpha : f(\alpha) = a\alpha^2 + b\alpha + c$

ที่จุด $\beta : f(\beta) = a\beta^2 + b\beta + c$

ที่จุด $\gamma : f(\gamma) = a\gamma^2 + b\gamma + c$

การแก้สมการเพื่อหา Unique Solution

$$\begin{bmatrix} \alpha^2 & \alpha & 1 \\ \beta^2 & \beta & 1 \\ \gamma^2 & \gamma & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} f(\alpha) \\ f(\beta) \\ f(\gamma) \end{bmatrix}$$

แก้สมการ

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \alpha^2 & \alpha & 1 \\ \beta^2 & \beta & 1 \\ \gamma^2 & \gamma & 1 \end{bmatrix}^{-1} \begin{bmatrix} f(\alpha) \\ f(\beta) \\ f(\gamma) \end{bmatrix}$$

กำหนดให้ $\psi = \det \begin{bmatrix} \alpha^2 & \alpha & 1 \\ \beta^2 & \beta & 1 \\ \gamma^2 & \gamma & 1 \end{bmatrix}$

แล้วจะได้ว่า $\psi = (\alpha - \beta)(\beta - \gamma)(\gamma - \alpha)$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

นอกจากนั้นยังจะพบว่า

$$a = \frac{1}{\psi}[(\gamma-\beta)f(\alpha)+(\alpha-\gamma)f(\beta)+(\beta-\gamma)f(\gamma)]$$

$$b = \frac{1}{\psi}[(\beta^2-\gamma^2)f(\alpha)+(\gamma^2-\alpha^2)f(\beta)+(\alpha^2-\beta^2)f(\gamma)]$$

$$c = \frac{1}{\psi}[\beta\gamma(\gamma-\beta)f(\alpha)+\gamma\alpha(\alpha-\gamma)f(\beta)+\alpha\beta(\beta-\gamma)f(\gamma)]$$

แล้วค่า x ที่ทำให้พบกับจุดต่ำสุดของฟังก์ชั่นสามารถหาได้จากสูตรดังนี้

$$x^* = -\frac{b}{2a}$$

ดังนั้น $x^* = \frac{1}{2}\left[\frac{(\beta^2-\gamma^2)f(\alpha)+(\gamma^2-\alpha^2)f(\beta)+(\alpha^2-\beta^2)f(\gamma)}{(\beta-\gamma)f(\alpha)+(\gamma-\alpha)f(\beta)+(\alpha-\beta)f(\gamma)}\right]$

การวัดประสิทธิภาพของแบบจำลอง Multi Layer Feed Forward

การศึกษานี้จะทำการการประเมินผลด้วย MAPE (Mean Absolute Percentage Error) มีสูตรการคำนวณ ดังนี้

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|t_i - y_i|}{t_i} * 100$$

เมื่อ t คือ ค่าที่เกิดขึ้นจริง

y คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวันที่พยากรณ์

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3.1.2 การทดสอบความนิ่งของข้อมูล (Stationary) และการทดสอบ Unit Root

ข้อมูลอนุกรมเวลา (Time Series) นั้นเป็นข้อมูลหรือค่าสังเกตที่มีการเปลี่ยนแปลงของตัวแปรในช่วงเวลาที่ผ่านมา ลักษณะของการเปลี่ยนแปลงอาจมีหรือไม่มีรูปแบบก็ได้ แต่ถ้าอนุกรมเวลาแสดงให้เห็นรูปแบบการเปลี่ยนแปลงในช่วงเวลาที่ผ่านมาในอดีต ก็จะทำให้สามารถคาดการณ์ได้ว่าในอนาคตลักษณะการเปลี่ยนแปลงควรอยู่ในรูปแบบใด และสามารถพยากรณ์การ

เปลี่ยนแปลงข้อมูลของในอนาคตได้ การวิเคราะห์ข้อมูลเวลาจะขึ้นอยู่กับการเปลี่ยนแปลงของเวลาในอดีตเป็นพื้นฐาน (ภัทร ตั้งตระกูล, 2546)

สำหรับการศึกษาข้อมูลที่เป็นอนุกรมเวลา (Time Series Data) นั้นเรามีข้อสมมุติว่าข้อมูลมีลักษณะ “นิ่ง” (Stationary) เนื่องด้วยโดยทั่วไปข้อมูลอนุกรมเวลามีลักษณะ “ไม่นิ่ง” (Nonstaionary) เมื่อทำการหาสมการถดถอยระหว่างตัวแปรอนุกรมเวลา 2 ตัวแปรออกแล้วเรามักจะพบว่าได้ R^2 มีค่าสูงมากและค่าสถิติ t จะมีนัยสัมพันธ์ของตัวแปรทั้งสองตัวโดยที่ในทางทฤษฎีแล้วไม่มีความหมายในทางเศรษฐศาสตร์ (ทรงศักดิ์ ศรีบุญจิตต์, 2547) เนื่องจากการทดสอบความสัมพันธ์แบบถดถอยของตัวแปรเราใช้ตาราง t ที่มีการแจกแจงแบบมาตรฐานแต่ค่าสถิติที่คำนวณได้จากความสัมพันธ์แบบถดถอยของตัวแปรที่ไม่นิ่งนั้นใช้ค่าสถิติ t (t satatistic) ที่มีการแจงแจงแบบไม่มาตรฐานจึงจะนำไปสู่การลงความเห็นที่ผิดพลาดได้

ข้อมูลที่มีลักษณะนิ่ง (Stationary) หมายถึงข้อมูลที่มีค่าเฉลี่ย (means) และความแปรปรวนที่มีค่าคงที่ (constant) เมื่อเวลาเปลี่ยนไป ในขณะที่ค่าความแปรปรวนร่วมเกี่ยว (covariance) ระหว่างสองคาบเวลาจะขึ้นอยู่กับช่องว่าง (gap) ระหว่างคาบเวลาเท่านั้น (ทรงศักดิ์ ศรีบุญจิตต์, 2547) ถ้าข้อมูลไม่ได้เป็นไปตามเงื่อนไขนี้ให้ลงความเห็นว่าข้อมูลมีลักษณะไม่นิ่ง (Nonstationary) เราสามารถเขียนให้อยู่ในรูปของคณิตศาสตร์ได้ดังต่อไปนี้

กระบวนการเชิงสุ่ม (X_t) จะถูกเรียกว่า “นิ่ง” (Stationay) ถ้า

Mean :$E(X_t) = \text{constant} = \mu$

Vatiance :$V(X_t) = \text{constant} = \sigma^2$

Covariance :$\text{cov}(X_t , X_{t+k}) = E(X_t - \mu)(X_{t+k} - \mu) = \sigma_k - \mu$

เนื่องด้วยข้อมูลที่ได้มาจากกระบวนการเชิงสุ่ม(Random Process) เราจึงต้องนำข้อมูลนั้นมาทดสอบว่ามีความนิ่งหรือไม่ โดยใช้การทดสอบ Unit Root ดังสมการต่อไปนี้

$$X_t = \rho X_{t-1} + \varepsilon_t \tag{3.1}$$

สมมติฐานว่าง $H_0 : \rho = 1$

$H_1 : |\rho| < 1$

ถ้ามีการยอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่ง (Nonstationary) แต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลมีลักษณะนิ่ง

สำหรับการทดสอบแบบ Dickey-Fuller Test (DF) ซึ่งถูกนำเสนอโดย Dickey และ Fuller ในปี 1981เพื่อให้ง่ายต่อการทดสอบ Unit Root ของข้อมูลอนุกรมเวลาว่ามีความนิ่งหรือไม่ ได้เสนอให้ θ มีค่าเท่ากับ ศูนย์ โดยนำ X_{t-1} ลบทั้งสองข้างของสมการที่ (3.1) ดังที่จะแสดงต่อไปนี้

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

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

$$\Delta X_t = (\rho - 1) X_{t-1} + \varepsilon_t$$

$$\Delta X_t = \theta X_{t-1} + \varepsilon_t \qquad (3.2)$$

ให้ สมมติฐานว่าง H_0: $\theta = 0$

H_1: $\theta < 0$

ถ้ามีการยอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่งแต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลมีลักษณะนิ่ง และการทดสอบนี้ยังสามารถแปลงสมการได้ดังต่อไปนี้ คือ

กรณีไม่มีค่าคงที่และแนวโน้มเวลา $\Delta X_t = \theta_{t-1} + \varepsilon_t$ (3.3)

กรณีมีค่าคงที่ $\Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t$ (3.4)

กรณีมีทั้งค่าคงที่และแนวโน้มเวลา $\Delta X_t = \alpha + \beta_t + \theta X_{t-1} + \varepsilon_t$ (3.5)

ต่อมาในปี 1984 Said และ Dickey ได้นำเสนอวิธี Augmented Dickey-Fuller Test (ADF) เพิ่มจำนวนของ lagged difference terms เข้าไปในสมการเพื่อแก้ปัญหา Serial Correlation ดังสมการต่อไปนี้

กรณีไม่มีทั้งค่าคงที่และแนวโน้มเวลา $\Delta X_t = \theta X_{t-1} + \sum_{i=1}^{\rho} \phi_t \Delta X_{t-i} + \varepsilon_t$ (3.6)

กรณีมีเฉพาะค่าคงที่ $\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^{\rho} \phi_i \Delta X_{t-i} + \varepsilon_t$ (3.7)

กรณีมีทั้งค่าคงที่และแนวโน้มเวลา $\Delta X_t = \alpha + \beta_t + \theta X_{t-1} + \sum_{i=1}^{\rho} \phi_t \Delta X_{t-i} + \varepsilon_t$ (3.8)

ในการตรวจสอบว่าข้อมูลมีลักษณะนิ่งหรือไม่นิ่งนั้นสามารถเปรียบเทียบค่าสถิติ t ที่คำนวณได้กับค่าวิกฤติ (Critical Value) ในตาราง ADF

3.1.3 แบบจำลอง Autoregressive Integrated Moving Average (ARIMA)

แบบจำลอง Autoregressive/Integrated/Moving Average(ARIMA) ได้มีการศึกษาโดย Grorge Box และ Gwilym Jenkins (1976) แต่ Wold (1938) ได้เป็นผู้ให้พื้นฐานทางทฤษฎีของกระบวนการหรือระบบ ARIMA บนพื้นฐานของ Wold แบบจำลอง ARIMA ได้ถูกพัฒนาขึ้นในสามทิศทาง ซึ่งได้แก่ ขั้นตอบสนองการประมาณค่าและการบ่งชี้ที่มีประสิทธิภาพ (efficient identification and estimation procedures) (สำหรับกระบวนการหรือระบบ AR,MA และ ARMA) การครอบคลุมไปถึงผลลัพธ์ที่ได้รวบรวมเอาอนุกรมเวลาเชิงฤดูกาล (seasonal time series) และ การขยายขอบเขตไปเพื่อรวมเอากระบวนการหรือระบบไม่นิ่ง (nonstationary process (ARIMA)) เข้าไว้ด้วย (ทรงศักดิ์ ศรีบุญจิตต์, 2547)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

โดยทั่วไปแล้วข้อมูลอนุกรมเวลาส่วนใหญ่มีลักษณะไม่นิ่ง(nonstationary) เนื่องจากข้อมูลอนุกรมเวลานั้นมาจากกระบวนการเชิงสุ่ม (random process) แต่ด้วยทฤษฎีของ AR และ MA หมายถึงข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง(stationary) ดังนั้นเมื่อข้อมูลที่รวบรวมได้มีลักษณะไม่นิ่ง เราจึงต้องทำการหาผลต่าง(differencing)

ความนิ่งและความไม่นิ่ง (Stationarity and Nonstationarity)

เครื่องมือทางด้านสัญลักษณ์ที่มีประโยชน์มากก็คือ backward shift operator, B. หรือ lag operator, L.(ซึ่งบางครั้งเราก็อาจใช้สัญลักษณ์ B หรือสัญลักษณ์ L สลับกันไปมาได้มีความหมายเหมือนกัน) ซึ่งถูกนำมาใช้ดังนี้

$$BX_t = X_{t-1} \tag{3.9}$$

ซึ่งถ้า B อยู่หน้า X_t จะมีผลต่อการ shift ข้อมูลถอยหลังไปหนึ่งคาบเวลา และถ้าเรามี

$$B(BX_t) = B_2X_t = X_{t-2} \tag{3.10}$$

ซึ่งหมายความว่า X_t ได้ถูก shift ถอยหลังไปสองคาบเวลา

ผลต่างที่หนึ่ง (first difference)

$$X_t' = X_t - X_{t-1} \tag{3.11}$$

ถ้าเราใช้ backward shift operator จะได้

$$X_t' = X_t - BX_t = (1-B)X_t \tag{3.12}$$

ผลต่างอันดับที่สอง (second-order difference)

$$\begin{aligned} X_t'' &= X_t' - X_{t-1}' \\ &= (X_t - X_{t-1}^t) - (X_{t-1} - X_{t-2}) \\ &= X_t - 2X_{t-1} + X_{t-2} \\ &= (1 - 2B + B^2)X_t \\ &= (1-B)^2 X_t \end{aligned} \tag{3.13}$$

$(1-B)^2$ คือ ผลต่างอันดับที่สอง (second-order difference)

$1-B^2$ คือ ผลต่างที่สอง (second difference) ซึ่งไม่เหมือนกัน

$(1-B)^d X_t$ คือ ผลต่างอันดับที่ d

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

กระบวนการหรือระบบอัตถดถอย (Autoregressive Processes)

กระบวนการหรือระบบ AR (p) ซึ่งคือกระบวนการหรือระบบ AR ที่มีอันดับที่ p เขียนในรูปของ ARIMA(p,d,q) ได้ดังนี้คือ

ARIMA(p,o,o) ซึ่งคือ

$$X_t = \mu' + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + ... + \varphi_p X_{t-p} + e_t \qquad (3.14)$$

โดยที่ μ' คือ พจน์คงที่หรือคงตัว (constant term)

φ_j คือ พารามิเตอร์อัตถดถอยตัวที่ j

e_t คือ พจน์ความคลาดเคลื่อน ณ เวลา t

กระบวนการหรือระบบเฉลี่ยเคลื่อน (Moving Average Processes)

กระบวนการหรือระบบ MA(q) ซึ่งก็คือกระบวนการหรือระบบ MA ที่มีอันดับ q เขียนในรูปของ ARIMA(p,d,q) ได้ดังนี้ คือ

ARIMA(o,o,q)

$$X_t = \mu - e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - ... - \theta_q e_{t-q} \qquad (3.15)$$

โดยที่ μ' คือ พจน์คงที่หรือคงตัว (constant term)

φ_j คือ พารามิเตอร์อัตถดถอยตัวที่ j

e_t คือ พจน์ความคลาดเคลื่อน ณ เวลา t

ดังนั้นการผสมกันระหว่าง AR และ MA ในรูปของกระบวนการ หรือระบบ ARIMA สำหรับข้อมูลที่มีลักษณะนิ่ง(stationary) จะมีรูปแบบเป็น ARIMA (p,0,q) สมมุติให้ AR(1) และ MA(1) เราสามารถเขียนในรูป ARIMA ได้คือ ARIMA (1,0,1) ดังจะแสดงในสมาการต่อไปนี้

$$X_t = \mu' + \theta_t X_{t-1} + e_t - \theta_t e_{t-1}$$

หรือ $$(1 - \theta_1 B) X_t = \mu' + (1 - \theta_1 B) e_t$$

AR(1) MA(1)

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

แต่ถ้าข้อมูลมีลักษณะไม่นิ่ง (nonstationary) จะต้องหาผลต่าง (differencing) d ครั้ง เพื่อให้ข้อมูลมีลักษณะนิ่ง ดังนี้

ARIMA (1,1,1)

$$(1-B)(1-\varphi_1 B)X_t = \mu' + (1-\theta_1 B)e_t$$

First difference — AR(1) — MA(1)

หรือ

$$\left[1 - B(1+\varphi_1) + \varphi_1 B^2\right]X_t = \mu' + e_t - \theta_1 e_{t-1}$$

$$X_t = (1+\varphi_1)X_{t-1} - \varphi_1 X_{t-2} + \mu' + e_t - \theta_1 e_{t-1}$$

- *แบบจำลองการพยากรณ์โดยวิธีของ Box-Jenkins*

การพยากรณ์อนุกรมเวลาโดยวิธีของ Box-Jenkins ในรูปแบบของ (p,d,q) ต้องพิจารณาก่อนว่าข้อมูลอนุกรมเวลามีลักษณะนิ่ง แล้วนำเข้าสู่ขั้นตอนของ Box-Jenkins ดังนี้

1) การกำหนดรูปแบบ (Indentification) ให้กับอนุกรมเวลาที่มีความนิ่ง (stationary) เป็นการหารูปแบบARIMA (p,d,q) ที่เหมาะสมให้กับอนุกรมเวลาโดยที่ autocorrelation: P_k มีค่าอยู่ในช่วง [1,-1] โดยพิจารณาเปรียบเทียบค่า autocorrelation (R_k) ของอนุกรมตัวอย่างกับค่า autocorrelation (P_k) ของอนุกรมเวลาของประชากรที่มีช่วงเวลาย้อนหลังไป k หน่วยเวลาย้อนหลังไป k หน่วยเวลา ซึ่งมีสูตร ดังนี้ $Y_t = \sum nt = a(Y_t)$

โดยที่ q = จำนวนเวลาสุดท้ายที่ย้อนหลัง

n = จำนวนข้อมูล

t = ระยะเวลา

Partial Autocorrelation : R_{kk} คือ การวัดความสัมพันธ์ของแต่ละช่วงเวลา โดยมีช่วงเวลาที่ย้อนหลังไป k หน่วยเวลาโดยพิจารณาเปรียบเทียบค่า Partial Autocorrelation ของอนุกรมเวลาตัวอย่าง (R_{kk}) กับค่า Partial Autocorrelation ของอนุกรมเวลาประชากร (P_{kk}) ที่มีช่วงเวลาย้อนหลังไป k หน่วยเวลา ซึ่งมีสูตร ดังนี้

$$R_k = \frac{\sum_{t-a}^{n-k}(Y_t - q)(Y_t + k - q_t)}{\sum_{t-a}^{n}(Y_t - q)^2}$$

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

การพิจารณาแต่ละรูปแบบ ต้องพิจารณา R_k, R_{kk} กับ P_k, P_{kk} พร้อมกันหลายๆค่า จึงมักพิจารณาจากรูปที่เรียกว่าคอเรลโลแกรม ที่ได้จากการพล็อต R_k, R_{kk} กับ P_k, P_{kk} ในช่วงเวลาที่ k ด้วยเหตุนี้การเปรียบเทียบรูปแบบจึงเป็นการเปรียบเทียบคอเรลโลแกรมของค่า autocorrelation(R_k) ของอนุกรมตัวอย่างกับค่า autocorrelation(P_k) ของอนุกรมเวลาของประชากร และคอเรลโลแกรมของค่า Partial Autocorrelation ของอนุกรมเวลาตัวอย่าง (R_{kk}) กับค่า Partial Autocorrelation ของอนุกรมเวลาประชากร (P_{kk}) สำหรับแต่ละรูปแบบจะมีคอเรลโลแกรมของ P_k และ P_{kk} ต่างกัน

2) การประมาณค่าพารามิเตอร์ (Estimation) ของแบบจำลอง โดยการหาค่าประมาณแบบง่ายหรือค่าประมาณที่ได้จากการวิเคราะห์ตัวเลข สำหรับค่าประมาณแบบง่ายจะทำโดยการสร้างสมการที่ได้มาจากความสัมพันธ์ระหว่าง P_k และพารามิเตอร์ โดยสมการที่สร้างขึ้นจะมีจำนวนพารามิเตอร์ที่ต้องการประมาณ ส่วนค่าประมาณที่ได้จากการวิเคราะห์ตัวเลขจะได้จากการแก้สมการที่สร้างขึ้นจากวิธีกำลังสองน้อยที่สุด ในขั้นตอนของการวิเคราะห์ตัวเลขจะต้องมีการกำหนดค่าประมาณเริ่มต้น ซึ่งส่วนใหญ่จะใช้การประมาณแบบง่ายเป็นจุดเริ่ม หลังวิเคราะห์เสร็จแล้วจะใช้ค่าประมาณสุดท้ายที่นำไปใช้ประโยชน์ในการสร้างสมการพยากรณ์

3) การตรวจสอบแบบจำลอง (Diagnostic checking) เมื่อกำหนดรูปแบบและประมาณค่าพารามิเตอร์ในแบบจำลอง จะต้องทำการตรวจสอบทุกครั้งว่ารูปแบบที่กำหนดนั้นมีความเหมาะสมจริงหรือไม่ โดยการตรวจสอบสามารถทำได้หลายวิธี ได้แก่ การพิจารณาคอเรลโลแกรมของ R_k หรือของความคลาดเคลื่อน การทดสอบค่าพารามิเตอร์ในแบบจำลองด้วยการทดสอบแบบ t-test และการทดสอบความเหมาะสมของแบบจำลองโดยการทดสอบของ Box และ Pierce ซึ่งพิจารณาจากค่า Q-statistic ดังสมการนี้

$$Q = n\sum_{k-1}^{m} \hat{\rho}_k^{\,2}$$

โดยที่ n= จำนวนของข้อมูล

m= ค่า lag length

ซึ่งค่า Q ที่ได้มีการแจกแจงแบบ Chi-square และมี Degree of freedom เท่ากับ m โดยให้สมมติฐานว่าง เป็นพจน์ของความคลาดเคลื่อนที่ได้จากการประมาณที่มีลักษณะเป็น White noise หมายถึงแบบจำลองที่ไม่มีอัตตสหสัมพันธ์ ถ้าหากแบบจำลองที่ได้ไม่มีอัตตสหสัมพันธ์ให้ใช้แบบจำลองนี้ไปพยากรณ์ต่อไป แต่ถ้าหากแบบจำลองมีอัตตสหสัมพันธ์ให้กลับไปกำหนดรูปแบบตามข้อที่ 1 ใหม่

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

4) การพยากรณ์ (Forecasting) เมื่อได้แบบจำลองที่เหมาะสม สามารถทำการพยากรณ์ได้ทั้งแบบจุด (Point forecast) และแบบช่วง (Interval forecast) ซึ่งในการพยากรณ์แบบช่วงนั้นปกติจะมีอยู่ 3 ช่วงคือ ช่วง Historical forecast, Ex-post forecast และช่วง Ex-ante forecast

3.1.4 แบบจำลอง Autoregressive Conditional Heteroskedastic (ARCH)

ในปี 1982 Engle ได้แสดงแนวคิดถึงความเป็นไปได้ที่จะสร้างแบบจำลองที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) ขึ้นพร้อมกัน โดยในขั้นตอนแรกต้องทำความเข้าใจก่อนว่าทำไมถึงต้องใช้การพยากรณ์อย่างมีเงื่อนไข

จากแบบจำลองอารีมา $X_t = a_0 + a_1 X_{t-1} + \varepsilon_t$ และต้องการพยากรณ์ X_{t+1} ดังนั้นเงื่อนไขเฉลี่ยการพยากรณ์ของ X_{t+1} คือ

$$E_t X_{t+1} = a_0 + a_1 X_t$$

ถ้าใช้ค่าเฉลี่ยอย่างมีเงื่อนไขมาพยากรณ์ X_{t+1} ฉะนั้นการพยากรณ์ค่าความแปรปรวนของความคลาดเคลื่อน คือ

$$E_t\left[\left(X_{t+1} - a_0 - a_1 X_t\right)^2\right] = E_t \varepsilon_{t+1}^2 = \sigma^2$$

อย่างไรก็ตามถ้าการพยากรณ์อย่างไม่มีเงื่อนไขถูกนำมาใช้ ก็หมายถึงค่าเฉลี่ยของ $\{X_t\}$ ในระยะยาวเท่ากับ $\frac{a_0}{(1-a_1)}$ จะได้ว่าการพยากรณ์อย่างไม่มีเงื่อนไขของค่าความแปรปรวนของความคลาดเคลื่อนคือ

$$E\left\{\left(X_{t+1} - \frac{a_0}{(1-a_1)}\right)^2\right\} = E\left[\left(\varepsilon_{t+1} + a_1\varepsilon_t + a_1^2\varepsilon_{t-1} + a_1^3\varepsilon_{t-2} + \ldots\right)^2\right]$$

$$= \frac{\sigma^2}{(1+a_1)^2}$$

จะเห็นว่า $\frac{1}{(1+a_1)^2}$ ที่ได้มีค่ามากกว่า 1 ซึ่งหมายถึงการพยากรณ์อย่างมีเงื่อนไข มีความแปรปรวนมากกว่าการพยากรณ์อย่างไม่มีเงื่อนไข ดังนั้นการพยากรณ์อย่างมีเงื่อนไขจึงถูกนำมาใช้ในการพยากรณ์มากว่า

โดยแบบจำลอง Autoregressive Conditional Heteroskedastic (ARCH) เกิดจากการให้ข้อสังเกตุว่าค่าความแปรปรวนของความคลาดเคลื่อน $\{\varepsilon_t\}$ ไม่คงที่ แต่สามารถประมาณค่าแนวโน้มของการเคลื่อนที่ของความแปรปรวนจากแบบจำลองอารีมาได้ ตัวอย่างเช่น ให้ $\{\varepsilon_t\}$ ที่ได้จากการประมาณค่าของส่วนเหลือ (Estimated Residauls) จากสมการ $X_t = a_0 + a_1 X_{t-1} + \varepsilon_t$ ดังนั้นค่าความแปรปรวนอย่างมีเงื่อนไขของ X_{t+1} คือ

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved

$$Var\left(X_{t+1}\mathrm{lX_t}\right) = E\left[\left(X_{t+1} - a_0 - a_1 X_t\right)^2\right]$$
$$= E_t \varepsilon_{t+1}^2$$

จะได้ว่า $E_t \varepsilon_{t+1}^2$ เท่ากับค่าความแปรปรวนที่ (σ^2) คงที่ แต่ถ้าสมมติให้ค่าความแปรปรวนอย่างมีเงื่อนไขไม่คงที่ วิธีการหนึ่งสำหรับที่จะนำมาพยากรณ์ค่าความแปรปรวนอย่างมีเงื่อนไข คือ

การใช้ลำดับของ Autoregressive หรือ AR(q) โดยการยกกำลังสองของส่วนเหลือที่ถูกประมาณค่า ดังนี้ $\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \ldots + \alpha_q \hat{\varepsilon}_{t-q}^2 + V_t$

เมื่อ V_t คือ White noise process

ซึ่งก็คือสมการของ Autoregressive Conditional Heteroskedastic (ARCH: AR(q)) นั้นเอง

3.1.5 แบบจำลอง Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

แบบจำลอง Generalized Autoregressive Conditional Herroscedasticity (GARCH) เป็นแบบจำลองที่ Bollerslev (1986. อ้างถึงใน ทรงศักดิ์ ศรีบุญจิต, 2547) ได้ขยายมาจาก ARCH model โดยเสนอแนวคิดที่ว่าการพยากรณ์ค่าความแปรปรวนอย่างมีเงื่อนไขสามารถอธิบายจาก ค่าความแปรปรวนและค่าความคลาดเคลื่อน ที่มีลักษณะ ARMA process พร้อมกันได้ ต่อไปนี้

$$X_t = a_0 + a_1 X_{t-1} + \varepsilon_t \tag{3.16}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{3.17}$$

จากสมการ (3.16) เป็นสมการค่าเฉลี่ย (Mean Equation) ที่ถูกกำหนดจากตัวแปรภายนอกและค่าความคลาดเคลื่อน ε_t ส่วนสมการ (3.17) เป็นสมการค่าความแปรปรวนในคาบเวลาปัจจุบันที่ถูกกำหนดโดย 3 ฟังก์ชัน คือ

- ค่าคงที่ α_0
- เหตุการณ์ที่สำคัญที่ส่งผลต่อความคลาดเคลื่อนในคาบเวลาที่ผ่านมา โดยถูกกำหนดให้อยู่ในรูป กำลังสองของส่วนเหลือจากสมการค่าเฉลี่ย $E\varepsilon_{t-i}^2$ ซึ่งก็คือ ARCH term นั่นเอง
- ค่าพยากรณ์ความแปรปรวนในคาบเวลาที่ผ่านมา σ_{t-i}^2 ซึ่งก็คือ GARCH term นั่นเอง

ดังนั้นจากสมการ (3.17) จึงถูกเรียกว่า Generalized Autoregressive Conditional Heteroscedasticity :GARCH (p,q) ซึ่งได้เปิดโอกาสให้มีส่วนประกอบที่เป็น Autoregressive Moving Average ในความแปรปรวนที่มีลักษณะ Hetoroscedastic Variance จะเห็นได้ว่า ถ้า p=0

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

และ q=1 เราก็จะได้แบบจำลอง GARCH(0,1) ซึ่งก็คือ ARCH(1) หรือ ARCH(q=1) นั่นเอง โดยสรุปว่าถ้า β_i ทุกตัวมีค่าเท่ากับศูนย์ แบบจำลอง GARCH ก็คือ ARCH(q) นั่นเอง

3.1.6 แบบจำลอง Generalized Autoregressive Conditional Heteroskedastic in Mean (GARCH-M)

Engle (1987. อ้างถึงใน จิตติ ตันเสนีย์, 2549) ได้ทำการขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชันของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักในชื่อของ GARCH-in-Mean หรือGARCH-M ดังสมการต่อไปนี้

$$X_t = \alpha_0 + \alpha X_{t-1} + \delta_1 \sigma_t^{1/2} + \varepsilon_t \quad (3.18)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \quad (3.19)$$

ซึ่งสมการ (3.18) ถูกเรียกว่า Generalized Autoregressive Conditional Heteroscedasticity in Mean เนื่องจากค่าความแปรปรวนที่ได้ถูกนำไปใส่ในสมการค่าเฉลี่ย หรือ สมการ (3.18) เพื่อการประมาณค่าของตัวแปรที่ถูกศึกษา

3.2 วิธีการวิจัย

3.2.1 วิธีการศึกษาด้วยแบบจำลองนิวรอลเน็ตเวิร์ค

การศึกษาด้วยแบบจำลองนิวรอลเน็ตเวิร์คประกอบไปด้วยการทดลอง 3 การทดลอง ดังนี้

1) การทดลองเบื้องต้นด้วย Hidden Layer จำนวน 1 ชั้น ด้วยวิธี Arbitrary

2) การทดลองปรับเปลี่ยนจำนวนนิวรอลชั้นซ่อนเร้นด้วยวิธี Quadratic Interpolation

3) การทดลองปรับเปลี่ยนจำนวนข้อมูลนำเข้าด้วยวิธี Quadratic Interpolation

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

รายละเอียดของแต่ละการทดลอง มีดังนี้

1) การทดลองเบื้องต้นด้วย Hidden Layer จำนวน 1 ชั้น ด้วยวิธี Arbitrary

ขั้นที่1 แบ่งข้อมูลที่มีอยู่ออกเป็น 3 ช่วง คือ Training 467 ข้อมูล, Validation 50 ข้อมูล และ Testing 50 ข้อมูล

ขั้นที่ 2 จัดเรียงข้อมูลนำเข้า เพื่อกำหนด input 10 ตัว

ขั้นที่ 3 หาจำนวนรอบการเรียนรู้ (Epochs) โดยกำหนด ให้มีค่าเท่ากับ 100 200 300 400 500 600 700 800 900 และ 1000 รอบ และหาจำนวนนิวรอลในชั้น Hidden Layer โดยกำหนดให้มีค่าเท่ากับ 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 และ 95 โดยเลือกค่าการเรียนรู้ (Epochs) และจำนวนนิวรอลที่เหมาะสมจากการพิจารณาค่า Mean Standard Error (MSE) ซึ่งมีสูตรดังนี้

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - t_i)^2$$

โดยที่ y_i = ค่าพยากรณ์

t_i = ค่าจริง

n = จำนวนข้อมูลที่ใช้ในการพยากรณ์

ขั้นที่ 4 เมื่อได้ค่า MSE ที่ต่ำที่สุดของจำนวนรอบการเรียนรู้ (Epochs) และจำนวนนิวรอลในชั้น Hidden Layer แล้ว นำค่าจำนวนรอบการเรียนรู้และจำนวนนิวรอลที่เหมาะสมไปพยากรณ์ราคาน้ำมันดิบเบรนท์ข้างหน้าทีละวัน จำนวน 50 วัน แล้วเปรียบเทียบกับค่าจริง

ขั้นที่ 5 นำผลการพยากรณ์ที่ได้ไปคำนวณค่า Mean Absolute Percentage Error (MAPE) ดังสูตรการคำนวณดังนี้

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} APE_i$$

$$APE_i = \frac{|t_i - y_i|}{t_i} \times 100$$

โดยที่ t_i = ค่าจริง

y_i = ค่าพยากรณ์

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

ขั้นที่ 6 หาจำนวนรอบการเรียนรู้ (Epochs) โดยกำหนด ให้มีค่าเท่ากับ 100 200 300 400 500 600 700 800 900 และ 1000 รอบ และหาจำนวนนิวรอลในชั้น Hidden Layer โดยกำหนดให้มีค่าเท่ากับ 100 200 300 400 500 600 700 800 900 และ 1000โดยเลือกค่าการเรียนรู้ (Epochs) และจำนวนนิวรอลที่เหมาะสมจากการพิจารณาค่า Mean Standard Error (MSE) ซึ่งมีสูตรดังนี้

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - t_i)^2$$

โดยที่ y_i = ค่าพยากรณ์

t_i = ค่าจริง

n = จำนวนข้อมูลที่ใช้ในการพยากรณ์

แล้วทำซ้ำขั้นที่ 4 และขั้นที่ 5

2) การทดลองปรับเปลี่ยนจำนวนนิวรอลชั้นซ่อนเร้นด้วยวิธี Quadratic Interpolation

ขั้นที่ 1 กำหนดจำนวนข้อมูลนำเข้าให้เท่ากับ 10

ขั้นที่ 2 ทำการหาจำนวนนิวรอลในชั้นซ่อนเร้นด้วยวิธี Quadratic Interpolation โดยกำหนดให้จำนวนนิวรอลในชั้นซ่อนเร้น ที่สองมีค่าเท่ากับ 70 75 และ 80

ขั้นที่ 3 เลือกจำนวนนิวรอลในชั้นซ่อนเร้น ที่ให้ค่า MSE ที่ต่ำที่สุด เพื่อใช้ในการคำนวณ Quadratic Interpolation ในขั้นต่อไป จนกว่าจะได้ค่า MSE ของจำนวนนิวรอลในชั้น Hidden Layer ที่มีค่า MSE ต่ำที่สุด

ขั้นที่4 ทำการหาจำนวนนิวรอลในชั้นซ่อนเร้น ด้วยวิธี Quadratic Interpolation โดยกำหนดให้จำนวนนิวรอลในชั้นซ่อนเร้น ที่สองมีค่าเท่ากับ 373 379 และ383 และทำซ้ำในขั้นที่ 3

3) การทดลองปรับเปลี่ยนจำนวนข้อมูลนำเข้าด้วยวิธี Quadratic Interpolation

ขั้นที่ 1 เมื่อได้จำนวนนิวรอลในชั้นซ่อนเร้นที่เหมาะสมแล้ว ให้ทำการหาจำนวนข้อมูลนำเข้าด้วยวิธี Quadratic Interpolation โดยการกำหนดให้จำนวนข้อมูลนำเข้า 10 20 30 40 และ 50 ที่ให้ค่า MSE ที่ต่ำที่สุด เพื่อใช้ในการคำนวณ Quadratic Interpolation ในขั้นต่อไป จนกว่าจะได้ค่า MSE ของจำนวนนิวรอลในชั้น Hidden Layer ที่มีค่า MSE ต่ำที่สุด

ขั้นที่ 2 เมื่อได้จำนวนนิวรอลในชั้นซ่อนเร้น และจำนวนข้อมูลนำเข้าที่เหมาะสมที่สุดแล้ว นำไปพยากรณ์ราคาน้ำมันดิบเบรนท์ไปข้างหน้าทีละวัน จำนวน 50 วัน ทำการเปรียบเทียบกับค่าจริง เพื่อคำนวณหาค่า MAPE

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3.2.2 การศึกษาด้วยแบบจำลองอารีมา

ได้ทำการแบ่งข้อมูลอนุกรมเวลาของราคาน้ำมันดิบเบรนท์ออกเป็น 2 ชุด คือ ชุดสำหรับนำไปสร้างแบบจำลอง และชุดค่าจริงสำหรับการทดสอบความแม่นยำในการพยากรณ์ จากนั้นจึงนำข้อมูลที่ได้ทำเข้าสู่กระบวนการศึกษาแบบจำลองอารีมา ต่อไปดังนี้

1) การทดสอบความนิ่ง (Stationary) ของข้อมูลอนุกรมเวลา ด้วยวิธี Unit Root Test ว่าข้อมูลอนุกรมเวลาของราคาน้ำมันดิบเบรนท์มีลักษณะนิ่งหรือไม่ ถ้าข้อมูลมีลักษณะไม่นิ่ง (Nonstationary) ให้ทำการหาผลต่างลำดับที่ 1 (First Difference) ของข้อมูล แล้วทำการทดสอบ Unit Root อีกครั้งว่าผลต่างลำดับที่ 1 มีลักษณะนิ่งหรือไม่ถ้า ไม่นิ่งให้หาผลต่างลำดับที่ 2 (Second Difference) ไปเรื่อยๆจนกว่าข้อมูลจะมีลักษณะนิ่ง จึงนำไปคำนวณต่อไป

2) การกำหนดแบบจำลอง โดยการพิจารณาจาก Correlogram จากค่า ACF และ PACF เพื่อที่จะระบุว่าแบบจำลองควรมี Autoregressive (p) และ Moving Average (q) เท่าใด

3) การประมาณค่าแบบจำลอง เป็นการนำเอารูปแบบ ARIMA(p,d,q) จากแบบจำลองที่พิจารณาจากค่า ACF และ PACF มาประมาณค่าพารามิเตอร์ ประกอบการพิจารณาความมีนัยสำคัญทางสถิติด้วยค่า t-statistic โดย การสร้างหลายๆแบบจำลอง แล้วทำโดยการพิจารณาค่า Akaike Information Criterion (AIC) และ Schwarz Criterion (SIC) ที่มีค่าต่ำที่สุด

4) การตรวจสอบความถูกต้องของแบบจำลองว่ามีความถูกต้องสามารถนำแบบจำลองไปพยากรณ์ได้หรือไม่ จากการพิจารณาค่า Q-statistic หลังจากนั้นนำแบบจำลองที่ผ่านการพิจารณาค่า Q-statistic ทำการพยากรณ์ช่วง Ex-ante Forecast ทีละวันจำนวน 50 วัน เปรียบเทียบกับค่าจริงเพื่อคำนวณหาค่า MAPE

3.2.3 การศึกษาด้วยแบบจำลองการ์ชเอ็ม

หลังจากที่ทดสอบความเป็นนิ่งของข้อมูลอนุกรมเวลาของราคาน้ำมันดิบเบรนท์ จึงได้เข้าสู่ขั้นตอนการหาแบบจำลองการ์ชเอ็ม

1) กำหนดแบบจำลองการ์ชเอ็มโดยพิจารณาจาก Correlogram เพื่อกำหนดระบุค่าความแปรปรวนอย่างมีเงื่อนไข (q) และค่าความคลาดเคลื่อนกำลังสอง (p)

2) การประมาณค่าแบบจำลอง เป็นการนำเอารูปแบบ GARCH-M (p,q) จากแบบจำลองที่พิจารณาจากค่า ACF และ PACF มาประมาณค่าพารามิเตอร์ ประกอบการพิจารณาความมีนัยสำคัญทางสถิติด้วยค่า t-statistic โดย การสร้างหลายๆแบบจำลอง แล้วทำการพิจารณาค่า Akaike Information Criterion (AIC) และ Schwarz Criterion (SIC) ที่มีค่าต่ำที่สุด

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved

3) การตรวจสอบความถูกต้องของแบบจำลองว่ามีความถูกต้องสามารถนำแบบจำลองไปพยากรณ์ได้หรือไม่ จากการพิจารณาค่า Q-statistic หลังจากนั้นนำแบบจำลองที่ผ่านการพิจารณาค่า Q-statistic ทำการพยากรณ์ช่วง Ex-ante Forecast ทีละวันจำนวน 50 วัน เปรียบเทียบกับค่าจริงเพื่อคำนวณหาค่า MAPE

หลังจากการหารูปแบบที่เหมาะสมของแต่ละแบบจำลองแล้วทำการพยากรณ์เปรียบเทียบกับค่าจริงแล้ว ต่อมาจึงได้ทำการเปรียบเทียบความแม่นยำของแบบจำลองแต่ละแบบจำลอง โดยพิจารณาค่า MAPE ที่มีค่าต่ำที่สุด ดังสมการต่อไปนี้

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|t_i - y_i|}{t_i} * 100$$

เมื่อ t คือ ค่าที่เกิดขึ้นจริง

y คือ ค่าที่ได้จากการพยากรณ์

n คือ จำนวนวันที่พยากรณ์

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright© by Chiang Mai University
All rights reserved