

## บทที่ 3

### กรอบแนวความคิดและระเบียบวิธีวิจัย

#### 3.1 กรอบแนวคิด

กรอบแนวคิดแบ่งออกเป็น 4 ส่วน ได้แก่ แบบจำลองนิเวศน์เน็ตเวิร์ค, การทดสอบความนิ่งของข้อมูล (Unit Root Test), แบบจำลอง ARIMA และแบบจำลอง GARCH-M ซึ่งมีรายละเอียดดังต่อไปนี้

##### 3.1.1 แบบจำลองนิเวศน์เน็ตเวิร์ค

ขจรศักดิ์ กันทรพนิต และยิ่งศักดิ์ ศรีมารัตน์ (2542) ได้กล่าวไว้ว่า Artificial Neural Networks หรือข่ายสมองเทียม คือระบบที่มนุษย์สร้างขึ้นหรือพัฒนาขึ้นมาเพื่อเลียนแบบการทำงานของสมองมนุษย์ โดยอาจจะเป็นซอฟต์แวร์หรือฮาร์ดแวร์ หรือระบบผสมผสานระหว่างซอฟต์แวร์และฮาร์ดแวร์ก็ได้ โดยข่ายสมองต้องมีคุณสมบัติที่สำคัญๆของสมองได้แก่

- 1) สามารถเรียนรู้และเก็บความรู้ไว้ได้
- 2) สามารถนำความรู้ที่เก็บไว้นั้นมาใช้ประกอบในการตัดสินใจอย่างชาญฉลาด

โดยการทำงานของสมองมนุษย์นั้นประกอบขึ้นด้วยส่วนที่เล็กที่สุดคือเซลล์สมอง หรือนิวรอน (Neuron) จำนวนประมาณ  $10^{11}$  เซลล์ ต่อเชื่อมระหว่างกันด้วยไซแนปส์ (Synapses) โดยการเชื่อมต่อไปหลายๆ แบบและหลายๆ กลุ่มแล้วแต่การทำงานของแต่ละส่วนของสมอง โดยมีส่วนประกอบที่สำคัญคือตัวเซลล์ (Cell Body or Soma) มีนิวเคลียส (Nucleus) อยู่ตรงกลางทำหน้าที่ประมวลผลสัญญาณกระตุ้นทางไฟฟ้า (Electrical Impulses) ที่ส่งเข้ามาจากเซลล์อื่นผ่านทางเดนไดรต์ (Dendrites) ซึ่งต่อเชื่อมกับเดนไดรต์ของเซลล์อื่นๆ รอบๆ ข้าง หากมีสัญญาณกระตุ้นเข้ามาเพียงพอ เซลล์นี้ก็จะส่งสัญญาณกระตุ้นทางไฟฟ้าขนาดประมาณ 50 มิลลิโวลต์ในช่วงเวลาประมาณ 2 มิลลิวินาที ต่อไปให้เซลล์อื่นๆผ่านทางเอกซอน (Axon) ไปยังเดนไดรต์เพื่อส่งออกไปแก่เดนไดรต์ของเซลล์อื่น โดยการเชื่อมต่อระหว่างเดนไดรต์ของเซลล์ต่างๆ นั้นจะผ่านจุดปลายคือ Terminals โดยจะมีช่องว่างไซแนปส์ (Synaptic Gap) อยู่เล็กน้อย โดยการส่งสัญญาณไฟฟ้าของเซลล์สมองนั้นเกิดจากการไหลหรือถ่ายเทของสารเคมีชนิดต่างๆ ผ่านเดนไดรต์อีกทีหนึ่ง จากการทำงานของสมองมนุษย์นี้ จึงได้มีการพัฒนาเพื่อเลียนแบบการทำงานของ

สมองมนุษย์เพื่อนำไปประยุกต์ใช้ในงานวิจัยในสาขาต่างๆ โดยในทางเศรษฐศาสตร์นั้นก็เพิ่งเริ่มที่จะนำ Artificial Neural Networks มาใช้ในการวิจัยเมื่อไม่นานมานี้ ดังเช่นในงานวิจัยของคมสัน สุริยะ ซึ่งตีพิมพ์อย่างเป็นทางการใน Komsan (2006) ได้ทำการพยากรณ์ราคาน้ำมันดิบด้วยแบบจำลอง Neural Networks การศึกษานี้ใช้แบบจำลองนิวรอลเน็ตเวิร์กแบบแพร่ย้อนกลับ (Back Propagation) โดยมีการประเมินผลด้วย MAPE (Mean Absolute Percentage Error) มีสูตรการคำนวณดังนี้

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

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

โดยที่  $t_i$  = ค่าจริง

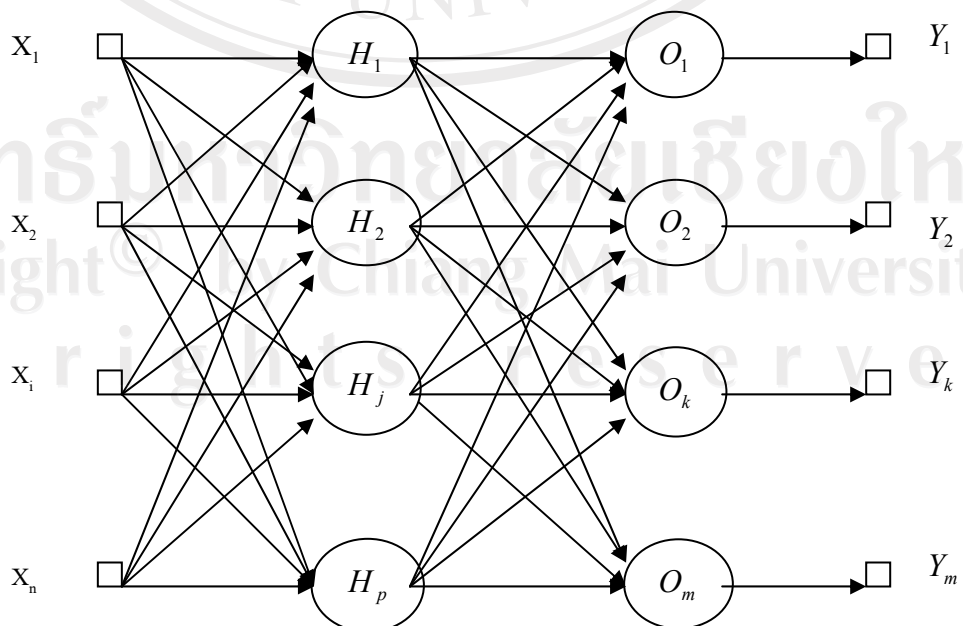
$y_i$  = ค่าพยากรณ์

$n$  = จำนวนข้อมูลของชุดข้อมูล Testing

นิวรอลเน็ตเวิร์กแบบแพร่ย้อนกลับ (Back Propagation) เป็นนิวรอลเน็ตเวิร์กแบบหนึ่งที่อยู่ในกลุ่มของ Multilayer Feed Forward Neural Networks (MLFF) ซึ่งจะรับข้อมูลนำเข้า (Input) แล้วทำการประมวลผลโดยนิวรอลในชั้น Hidden Layer และชั้น Output จากนั้นจะส่งผลลัพธ์ออกมาเป็นคำตอบ การที่สัญญาณนำเข้าถูกถ่ายทอดไปข้างหน้าเรื่อยๆ เรียกว่า Feed Forward และการที่มีจำนวนนิวรอลหลายชั้นเรียกว่า Multi Layer

นิวรอลเน็ตเวิร์กแบบแพร่ย้อนกลับสามารถแสดงได้ดังแผนภาพดังต่อไปนี้

รูปที่ 3.1 นิวรอลเน็ตเวิร์กแบบแพร่ย้อนกลับ



องค์ประกอบของนิเวศน์ตเวีรคแบบเพร่ย้อนกลับ ดังต่อไปนี้

1) Input

Input มีจำนวน  $n$  ตัว คือ  $X = \{ X_1, X_2, \dots, X_1, \dots, X_n \}$

2) Output

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

3) นิเวศน์ใน Hidden Layer

นิเวศน์ใน Hidden Layer มี  $p$  ตัว คือ  $H = \{ H_1, H_2, \dots, H_j, \dots, H_p \}$

4) นิเวศน์ใน output Layer

นิเวศน์ใน Output Layer มี  $m$  ตัว คือ  $O = \{ O_1, O_2, \dots, O_k, \dots, O_m \}$

5) ค่าน้ำหนักจาก Input layer สู่ hidden layer

ค่าน้ำหนักจาก input layer สู่ hidden layer สำหรับนิเวศน์แต่ละตัวใน hidden layer มีจำนวน  $n+1$  ตัว คือ  $W_{ij}^H = \{ W_{0j}^H, W_{1j}^H, W_{2j}^H, \dots, W_{ij}^H, W_{nj}^H \}$

ดังนั้นเมื่อมีจำนวน input จำนวน  $n$  ตัว และมีนิเวศน์ใน hidden layer จำนวน  $p$  ตัว แล้วจะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ  $(n+1)p$  ตัว

6) ค่าน้ำหนักจาก hidden layer สู่ output layer

ค่าน้ำหนักจาก Hidden layer สู่ output layer สำหรับนิเวศน์แต่ละตัวใน output layer มีจำนวน  $p+1$  ตัว คือ  $W_{jk}^O = \{ W_{0k}^O, W_{1k}^O, W_{2k}^O, \dots, W_{jk}^O, W_{pk}^O \}$

ดังนั้นเมื่อมีจำนวน Hidden layer จำนวน  $p$  ตัว และมีoutput จำนวน  $m$  ตัวแล้วจะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ  $(p+1)m$  ตัว

**การคำนวณแบบเพร่ย้อนกลับ (Back Propagation)**

ขั้นตอนของ Back-propagation Algorithm มีดังนี้

1) กำหนดค่าอัตราเร็วในการเรียนรู้ (rate parameter :  $r$ )

2) สำหรับแต่ละตัวอย่าง input ให้ทำตามขั้นตอนต่อไปนี้จนกว่าได้ระดับ performance ที่

ต้องการ

- คำนวณค่า output โดยใช้ค่าน้ำหนักเริ่มต้นซึ่งอาจได้จากการสุ่ม

- คำนวณค่า  $\beta$  : แทนประโยชน์ที่จะได้รับสำหรับการเปลี่ยนค่า output ของแต่ละ

node

จากกฎการปรับค่าน้ำหนักของ Gradient Descent

$$\Delta w = \eta(t - y)[\nabla_w f(a)] \quad (3.1)$$

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

ดังนั้น

$$\begin{aligned} \nabla_w f(a) &= \begin{pmatrix} \frac{\partial f(a)}{\partial w_0} \\ \frac{\partial f(a)}{\partial w_1} \\ \vdots \\ \frac{\partial f(a)}{\partial w_n} \end{pmatrix} = \frac{\partial f(a)}{\partial w} = \frac{\partial f(a)}{\partial a} \frac{\partial a}{\partial w} \\ &= f' \frac{\partial (w^T x)}{\partial w} = f' \frac{\sum_{i=0}^n w_i x_i}{\partial w_i} = f' x_i \\ &= f' \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = f' x \end{aligned}$$

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

**การปรับค่าน้ำหนักในชั้น Output**

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

$$\Delta W_{jk}^O = \eta(t_k - y_k) f' x_{jk}^O \quad (3.2)$$

แต่ Input ของชั้น Output คือ Output ของชั้น Hidden Layer ดังนั้น

$$\Delta W_{jk}^O = \eta(t_k - y_k) f' y_{jk}^H \quad (3.3)$$

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

### การปรับค่าน้ำหนักในชั้น Hidden Layer

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

$$\Delta W_{ij}^H = \eta(t_j - y_j)g'x_{ij}^H \quad (3.4)$$

เมื่อ  $g'$  คือ derivative ของ transformation function จากชั้น Input สู่อัน Hidden Layer (อรรถย เมืองใจ, 2548: 22)

แต่ปัญหาที่พบก็คือ เราไม่สามารถหาค่า  $t_j$  ได้เพราะว่าในชั้น Hidden layer เราไม่ทราบค่าของ  $y_j$  ที่นิเวศแต่ละตัวส่งออกมาว่าผิดหรือถูกเราจึงไม่สามารถหา error ได้ ดังนั้นจึงได้มีการสร้างค่า error เทียมออกมาโดยสมมติให้ นิเวศที่ส่งข้อมูลมาชั้นผลลัพธ์ (out put) มากก็ต้องรับผิดชอบค่า error ที่เกิดขึ้นจาก  $(t-y)$  มากตาม ดังนั้นความรับผิดชอบสำหรับค่า Output ที่นิเวศในชั้น Hidden layer ตัวที่  $j$  จะเป็นผลรวมของค่ารับผิดชอบของ Output ทั้งหมดจึงสามารถแก้ปัญหการปรับค่าน้ำหนักในชั้น Hidden Layer ได้ดังนี้

$$\Delta W_{ij}^H = \eta(t_j - y_j)g'x_{ij}^H \quad (3.5)$$

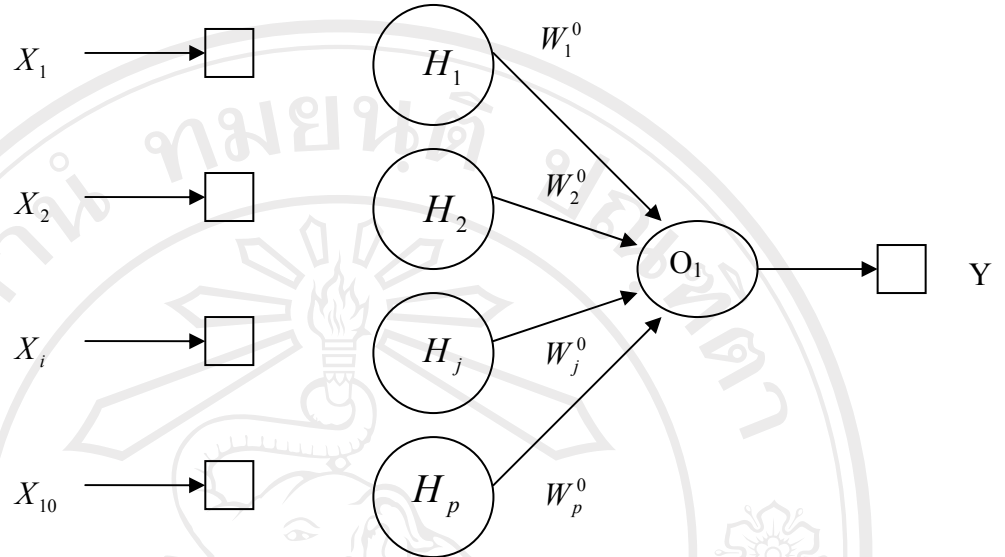
โดยที่  $t_j - y_j = \sum_{k=1}^m W_{jk}^0(t_k - y_k)f'$  ซึ่งเป็นผลรวมของความรับผิดชอบที่มีต่อ Output แต่ละตัว

ดังนั้นจึงได้กฎการปรับค่าน้ำหนักจากชั้น Input มายังชั้น Output ดังนี้

$$\Delta W_{ij}^H = \eta \left( \sum_{k=1}^m W_{jk}^0(t_k - y_k)f' \right) g'x_{ij}^H \quad (3.6)$$

ด้วยแนวคิดดังกล่าว จึงเขียนเป็นแผนผังได้ดังนี้

รูปที่ 3.2 นำหนักจากชั้น Hidden Layer สู่อินพุต Output



จากรูปจะเห็นว่า ความคลาดเคลื่อนของ Output แต่ละตัว สามารถแบ่งความรับผิดชอบไปยัง Neuron ใน Hidden Layer ทุก ๆ ตัวได้โดยผ่านค่าน้ำหนัก ซึ่งหาก Neuron ใน Hidden Layer ตัวใดมีค่าน้ำหนักมากก็ย่อมต้องรับผิดชอบกับความคลาดเคลื่อนที่เกิดขึ้นมาก

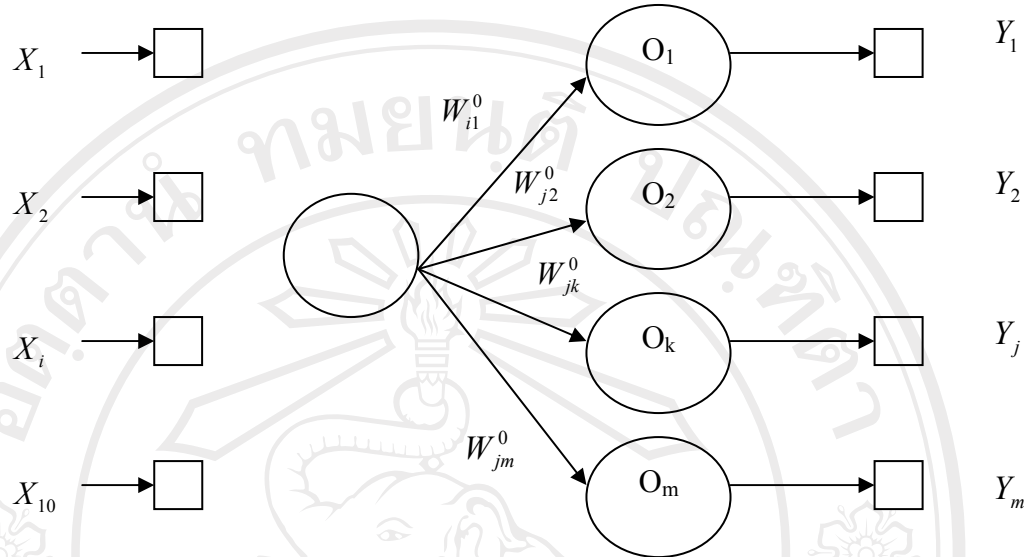
ดังนั้น ความคลาดเคลื่อนที่ Neuron ตัวที่  $j$  ใน Hidden Layer จะต้องรับผิดชอบสำหรับ Output ตัวที่ 1 ก็คือ  $W_{jl}^0(t_1 - y_1)f'$  หนึ่ง การที่ติดค่า derivative ของ transformation function มาด้วยนั้นก็เพราะการเปลี่ยนแปลงรูปแบบของ transformation function ย่อมมีผลกระทบต่อค่าน้ำหนักด้วย

ความรับผิดชอบรวมทั้งหมดที่ Neuron ตัวที่  $j$  ใน Hidden Layer จะต้องรับผิดชอบสำหรับ Output ทุกตัว ก็ย่อมเท่ากับ ผลบวกของความรับผิดชอบที่จะต้องรับผิดชอบต่อ Output แต่ละตัวนั่นเอง ดังนี้

$$\sum_{k=1}^m W_{jk}^0(t_k - y_k)f' \quad (3.7)$$

ซึ่งเขียนเป็นแผนภาพได้ดังนี้

รูปที่ 3.3 ความรับผิดชอบของ Neuron ตัวที่  $j$  ในชั้น Hidden Layer ที่มีต่อ Output



เมื่อนำเอาความรับผิดชอบที่ Neuron ตัวดังกล่าวต้องรับผิดชอบต่อ Output ทั้งหมดมาเป็น error เทียม จะทำให้สามารถแก้ปัญหาการปรับค่าน้ำหนักในชั้น Hidden Layer ได้ดังนี้

$$\Delta W_{ij}^H = \eta(t_j - y_j)g'x_{ij}^H \quad (3.8)$$

$$\text{แต่ } t_j - y_j = \sum_{k=1}^m W_{jk}^0(t_k - y_k)f'$$

ดังนั้นจึงจะได้กฎการปรับค่าน้ำหนักจากชั้น Input มายังชั้น Output ดังนี้

$$\Delta W_{ij}^H = \eta \left( \sum_{k=1}^m W_{jk}^0(t_k - y_k)f' \right) g'x_{ij}^H \quad (3.9)$$

#### การกำหนดจำนวนนิวรอนในชั้นซ่อนเร้น

การค้นหาค่าแบบจำลองที่ดีที่สุด ใช้การเปลี่ยนจำนวนนิวรอนในชั้นซ่อนเร้น ด้วยหลักการหาค่าต่ำสุดแบบ Quadratic interpolation ดังนี้

Quadratic interpolation เป็นการหาค่าแบบจำลองที่ดีที่สุด ในการกำหนดจำนวนนิวรอนในชั้นซ่อนเร้น โดยมีวิธีการหาค่าต่ำสุดด้วยวิธี Quadratic Interpolation ซึ่งจะได้เมื่อมีข้อมูลจำนวน 3 ชุดและมั่นใจว่าข้อมูลทั้งสามชุดอยู่บนฟังก์ชัน Quadratic เดียวกัน การมีข้อมูล 3 ชุดจะเพียงพอต่อการหา Unique Solution สำหรับฟังก์ชัน Quadratic ดังนี้



เมื่อมีข้อมูลดังต่อไปนี้

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

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

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

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

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

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

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

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

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

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่  
นอกจากนี้ยังพบว่า

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

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

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



แล้วค่า  $x$  ที่ทำให้พบกับจุดต่ำสุดของฟังก์ชันสามารถหาได้จากสูตรดังนี้

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

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

### 3.1.2 การทดสอบความนิ่งของข้อมูล (Unit Root Test)

การวิเคราะห์ข้อมูลทางเศรษฐศาสตร์ ซึ่งข้อมูลที่ใช้เป็นข้อมูลแบบอนุกรมเวลา (Time series data) มีความจำเป็นที่จะต้องมีการทดสอบว่า ข้อมูลที่เราใช้นั้นมีความนิ่ง (Stationary) หรือไม่ เพราะหากเราละเลยที่จะทดสอบข้อมูล แล้วข้อมูลเกิดมีความไม่นิ่ง (nonstationary) จะทำให้สมการถดถอยระหว่างตัวแปรอนุกรมเวลาสองตัวแปร จะได้ค่า  $R^2$  ที่มีค่าสูงมากและค่าสถิติ  $t$  จะมีนัยสำคัญ ทั้งที่ตัวแปรทั้งสองนั้น ไม่มีความสัมพันธ์กันเลยในทางเศรษฐศาสตร์ (Ender, 1995: 216; Gujarati, 1995: 709)

เหตุที่ทำให้ได้ค่า  $R^2$  สูงเช่นนี้เป็นเพราะอนุกรมเวลามีแนวโน้ม ไม่ใช่เนื่องจากความสัมพันธ์ที่แท้จริงระหว่างตัวแปรอนุกรมเวลาทั้งสองตัวแปร ส่วนกรณีที่ค่าสถิติ  $t$  มีนัยสำคัญ เป็นเพราะอนุกรมเวลาทั้งสองมีแนวโน้มที่แข็งแกร่งมาก (Strong Trend) โดยสรุปแล้วการละเลยการทดสอบความนิ่งของข้อมูลอาจนำไปสู่การลงความเห็นที่ผิดพลาดได้ในที่สุด

ข้อมูลที่มีลักษณะนิ่ง (Stationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) เท่ากันตลอดระยะเวลาที่ศึกษา

ส่วนข้อมูลที่มีลักษณะไม่นิ่ง (nonstationary) หมายถึง ข้อมูลอนุกรมเวลาที่มีค่าเฉลี่ย (Mean) และความแปรปรวน (Variance) ไม่เท่ากันตลอดระยะเวลาที่ศึกษา

ทั้งนี้การวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลาส่วนมากจะพบปัญหาความไม่นิ่งของข้อมูล ซึ่งสามารถแก้ไขได้ด้วยการทำให้ข้อมูลมีความนิ่งเสียก่อน โดยอาจใช้วิธีการหาผลต่าง (Difference) ของข้อมูล การแปลงให้อยู่ในรูป Logarithm หรือการทดสอบหาความสัมพันธ์ของตัวแปรในระยะยาว (cointegration) เป็นต้น

ในการทดสอบ unit root หรือ อันดับความสัมพันธ์ของข้อมูล (order of Integration) เป็นการทดสอบตัวแปรทางเศรษฐกิจต่างๆ ที่ใช้ในสมการเพื่อดูความเป็น station (I(0); integrated of order 0) หรือ non-stationary โดยส่วนมากแล้วจะนิยมการทดสอบโดยวิธี Dicky-Fuller test ซึ่งสามารถแบ่งออกได้เป็น 2 วิธี คือ

### วิธีที่ 1 Dickey-Fuller Test (DF)

วิธีนี้จะทำการทดสอบตัวแปรที่เคลื่อนไหวไปตามช่วงเวลาที่มิลักษณะเป็น autoregressive model โดยพิจารณาสมการ 3 รูปแบบที่แตกต่างกัน ดังนี้

$$\Delta X_t = \theta x_{t-1} + \varepsilon_t \quad (\text{random walk process})$$

$$\Delta X_t = \alpha + \theta x_{t-1} + \varepsilon_t \quad (\text{random walk with drift})$$

$$\Delta X_t = \alpha + \beta_t \theta x_{t-1} + \varepsilon_t \quad (\text{random walk with drift and linear time trend})$$

โดยที่  $\Delta X_t$  คือ ค่าความแตกต่างครั้งที่ 1 ของตัวแปรที่ทำการศึกษา  
 $\alpha, \beta, \theta$  คือ ค่าคงที่  
 $t$  คือ ค่าคงที่  
 $\varepsilon_t$  คือ ตัวแปรสุ่มที่มีค่าเฉลี่ยเท่ากับศูนย์ และค่าความแปรปรวนที่คงที่ หรือ  $\varepsilon_t \sim \text{iid}(0, \sigma_\varepsilon^2)$

การทดสอบ จะพิจารณาค่า  $\theta$  โดยเปรียบเทียบกับค่าสถิติ t (t-statistic) ที่คำนวณได้กับค่าที่เหมาะสมจากตาราง Dickey-Fuller ซึ่งมีสมมติฐานการทดสอบ ดังนี้

$$H_0: \theta = 0 \quad : \text{non-stationary}$$

$$H_1: \theta < 0 \quad : \text{stationary}$$

ถ้ายอมรับ  $H_0: \theta = 0$  จะได้ว่า ตัวแปรที่สนใจ ( $X_t$ ) มี unit root หรือ  $X_t$  มีลักษณะเป็น nonstationary

แต่ถ้ายอมรับ  $H_1: \theta \neq 0$  จะได้ว่าตัวแปรที่สนใจ ( $X_t$ ) ไม่มี unit root หรือ  $X_t$  มีลักษณะเป็น stationary

### วิธีที่ 2 Augmented Dickey-Fuller Test (ADF)

เป็นการทดสอบ unit root อีกวิธีหนึ่งที่พัฒนามาจาก DF Test เนื่องจากวิธี DF ไม่สามารถทำการทดสอบตัวแปรในกรณีที่เป็น serial correlation ในค่าความคลาดเคลื่อน error term ( $\varepsilon_t$ ) ที่มีลักษณะความสัมพันธ์กันเองในระดับสูง โดยมีสมการดังนี้

$$\Delta X_t = \theta x_t + \sum_{j=1}^p \phi_j \Delta X_{t-j} + \varepsilon_t$$

$$\Delta X_t = \alpha + \theta x_t + \sum_{j=1}^p \phi_j \Delta X_{t-j} + \varepsilon_t$$

$$\Delta X_t = \alpha + \beta t \theta x_t + \sum_{j=1}^p \phi_j \Delta X_{t-j} + \varepsilon_t$$

การใส่ค่า Lagged Term ( $\rho$ ) ว่ามีจำนวนเท่าใดจึงจะเหมาะสมสำหรับแต่ละข้อมูลอนุกรมนั้นมีหลักในการเลือก lag length ที่เสนอโดย Walter Ender (1995) ว่า ควรจะเริ่ม lag length ที่มีค่าที่มากพอ แล้วพิจารณาความมีนัยสำคัญทางสถิติ ที่ระดับนัยสำคัญต่างๆ ( $\sigma = 0.01, 0.05$  และ  $0.1$ ) เมื่อพบว่าที่ lag length ที่เลือกมีค่า  $t$ -statistic ที่ไม่มีนัยสำคัญทางสถิติ ณ ระดับนัยสำคัญ ร้อยละ 10 แล้วจึงทำการลด lag length ลงทีละ 1 ช่วง จนกระทั่งสามารถปฏิเสธสมมติฐานว่าง (Null Hypothesis)

### 3.1.3 แบบจำลอง ARIMA

แบบจำลอง ARIMA เป็นแบบจำลองที่ใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลา (Time series data) ที่ได้รับความนิยมและมีการพัฒนาปรับปรุงเป็น ARCH, GARCH และ EGARCH

การพยากรณ์ด้วยแบบจำลอง ARIMA ด้วยวิธีการ Box and Jenkins เป็นวิธีการพยากรณ์ค่าในอนาคตที่พัฒนาโดยนักสถิติผู้มีชื่อเสียงสองท่านคือ George E.P. Box และ Gwilym M. Jenkins ในปี ค.ศ. 1970 ซึ่งเป็นวิธีที่ทำให้ค่าพยากรณ์ที่ดี คือ ค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Square Error) ก่อนข้างต่ำ

การพยากรณ์อนุกรมเวลาโดยวิธี Box and Jenkins (ทรงศิริ แต่สมบัติ, 2539) ในรูปแบบ ARIMA (p,d,q) ต้องพิจารณาข้อมูลอนุกรม  $Y$  ว่ามีคุณสมบัติ stationary หรือไม่เสียก่อน โดยพิจารณาได้จาก

1) ค่าเฉลี่ย  $E(Y_t)$  คงที่ สำหรับทุกค่าของ  $t$  หรือไม่ ทำได้โดยการแบ่งอนุกรมเวลาออกเป็นส่วนๆ แล้วทำการหาค่าเฉลี่ยในแต่ละส่วน ถ้าค่าเฉลี่ยแต่ละส่วนย่อยไม่แตกต่างกันมาก จะสรุปได้ว่า  $E(Y_t)$  คงที่

2) ค่าความแปรปรวน  $[\text{Var}(Y_t)]$  คงที่ สำหรับทุกค่าของ  $t$  หรือทำได้โดยการแบ่งอนุกรมเวลาออกเป็นส่วนๆ แล้วทำการหาค่าความแปรปรวนแต่ละส่วน ถ้าค่าความแปรปรวนในแต่ละส่วนไม่มีความแตกต่างกันมากนัก จะสรุปได้ว่า  $\text{Var}(Y_t)$  คงที่

3) พิจารณาถึงแนวโน้มและปัจจัยของฤดูกาลว่ามีผลต่ออนุกรมเวลาหรือไม่

4) พิจารณา correlogram ของค่า autocorrelation ของตัวอย่าง ( $R_k$ ) กรณีที่อนุกรมเวลามีความนิ่ง (stationary) ค่า correlogram ของค่า autocorrelation จะมีค่าลดลงค่อนข้างเร็วเมื่อ  $k$  เพิ่มมากขึ้น ดังนั้นหากค่า autocorrelation มีค่าลดลงค่อนข้างช้าอธิบายได้ว่าอนุกรมเวลาชุดนี้มีแนวโน้ม (trend) แต่ถ้าค่า autocorrelation มีค่าลดลงค่อนข้างช้าและมีค่าค่อนข้างสูงที่  $k = L, 2L, 3L$  จะอธิบายได้ว่าอนุกรมเวลาชุดนี้มีแนวโน้มและอิทธิพลของฤดูกาล และหากการเคลื่อนไหวของ

correlogram ของค่า autocorrelation เป็นลักษณะลูกคลื่น โดยลูกคลื่นจะครบรอบในแต่ละช่วงเวลา แสดงว่ามีอิทธิพลของฤดูกาลเข้ามาเกี่ยวข้อง

หลังทำการตรวจสอบ หากพบว่าข้อมูลอนุกรมเวลาไม่มีความนิ่ง (nonstationary) จะต้องแปลงให้นิ่ง (stationary) เสียก่อนจากนั้นจึงเข้าสู่ 4 ขั้นตอนของ Box and Jenkins ดังนี้

1) การกำหนดรูปแบบ (identification) ให้กับอนุกรมเวลาที่มีความนิ่ง (stationary) เป็นการหารูปแบบ ARIMA (p,d,q) ที่คิดว่าเหมาะสมให้กับอนุกรมเวลาโดยที่ autocorrelation:  $P_k$  มีค่าอยู่ในช่วง  $[1,-1]$  โดยพิจารณาเปรียบเทียบกับค่า autocorrelation ( $R_k$ ) ของอนุกรมตัวอย่างกับค่า autocorrelation ( $P_k$ ) ของอนุกรมเวลาของประชากรที่มีช่วงเวลาย้อนหลังไป k หน่วย

Partial Autocorrelation:  $R_{kk}$  คือ การวัดความสัมพันธ์ของแต่ละช่วงเวลาโดยมีช่วงเวลาที่ย้อนหลังไป k หน่วยเวลาโดยพิจารณาเปรียบเทียบกับค่า Partial Autocorrelation ของอนุกรมตัวอย่าง ( $R_{kk}$ ) กับค่า Partial Autocorrelation ของอนุกรมเวลาประชากร ( $P_{kk}$ ) ที่มีช่วงเวลาย้อนหลังไป k หน่วยเวลา ซึ่งมีสูตรดังนี้

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

การพิจารณาแต่ละรูปแบบ ต้องพิจารณา  $R_k, R_{kk}$  กับ  $P_k, P_{kk}$  พร้อมกันหลายๆ ค่าจึงมักพิจารณาจากรูปแบบที่เรียกว่า correlogram ที่ได้จากการพล็อต  $R_k, R_{kk}$  กับ  $P_k, P_{kk}$  ในช่วงเวลาที่ k ด้วยเหตุนี้การเปรียบเทียบรูปแบบจึงเป็นการเปรียบเทียบ correlogram ของค่า autocorrelation ( $R_k$ ) ของอนุกรมตัวอย่างกับค่า autocorrelation ( $P_k$ ) ของอนุกรมเวลาของประชากร และ correlogram ของค่า Partial Autocorrelation ของอนุกรมตัวอย่าง ( $R_{kk}$ ) กับค่า Partial Autocorrelation ของอนุกรมเวลาประชากร ( $P_{kk}$ ) สำหรับแต่ละรูปแบบจะมี correlogram ของ  $P_k$  และ  $P_{kk}$  ต่างกัน อนุกรมเวลาที่จะนำมากำหนดรูปแบบจะต้องเป็นอนุกรมเวลาที่มีความนิ่ง (stationary) เท่านั้น ดังนั้นหากข้อมูลไม่ stationary จะต้องทำให้ stationary ก่อนทุกครั้ง

2) การประมาณค่าพารามิเตอร์ในรูปแบบ Estimator จะทำได้โดยการหาค่าประมาณแบบง่ายหรือค่าประมาณที่ได้จากการวิเคราะห์ตัวเลข สำหรับค่าประมาณแบบง่ายจะทำได้โดยการสร้างสมการที่ได้มาจากความสัมพันธ์ระหว่าง  $P_k$  และพารามิเตอร์ โคนสมการที่สร้างขึ้นจะมีจำนวนพารามิเตอร์ที่ต้องการประมาณ ส่วนค่าประมาณที่ได้จากการวิเคราะห์ตัวเลขจะทำได้จากการแก้สมการที่สร้างขึ้นจากวิธีกำลังสองน้อยที่สุด ในขั้นตอนของการวิเคราะห์ตัวเลขจะต้องมีการ

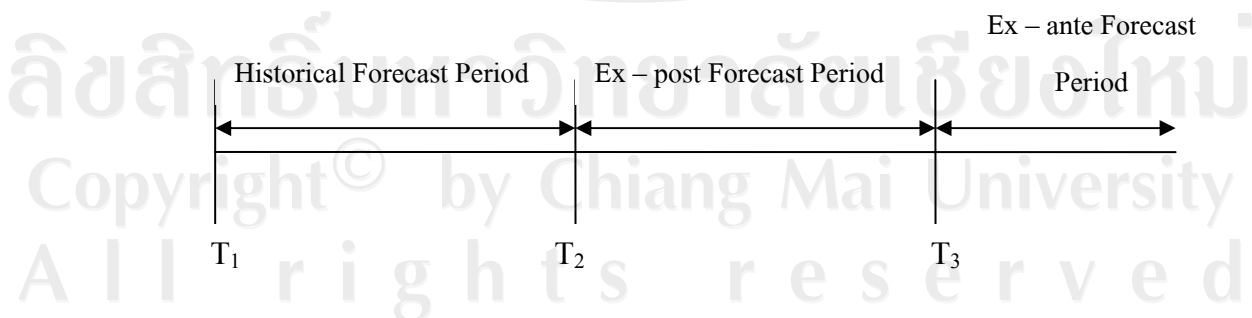
กำหนดค่าประมาณเริ่มต้นซึ่งส่วนใหญ่จะใช้การประมาณแบบง่ายเป็นจุดเริ่มต้น หลังวิเคราะห์เสร็จแล้วจะใช้ค่าประมาณสุดท้ายที่นำไปใช้ประโยชน์ในการสร้างสมการพยากรณ์

3) การตรวจสอบแบบจำลอง (Diagnostic checking) เมื่อกำหนดรูปแบบและประมาณค่าพารามิเตอร์ในแบบจำลอง จะต้องทำการตรวจสอบทุกครั้งว่ารูปแบบที่กำหนดนั้นมีความเหมาะสมจริงหรือไม่ โดยการตรวจสอบสามารถทำได้หลายวิธี ได้แก่ การพิจารณา correlogram ของ  $R_k$  หรือของความคลาดเคลื่อน การทดสอบค่าพารามิเตอร์ในแบบจำลองด้วยการทดสอบแบบ t-test และการทดสอบความเหมาะสมของแบบจำลองโดยการทดสอบของ Box และ Pierce หรือ การทดสอบของ Box และ Jung หากตรวจสอบพบว่าแบบจำลองนั้นมีความเหมาะสมแล้ว จะสามารถใช้รูปแบบนั้นพยากรณ์ต่อไป แต่หากแบบจำลองไม่มีความเหมาะสมต้องกลับไปเริ่มขั้นตอนที่ 1 ใหม่

4) การพยากรณ์ (Forecasting) สามารถทำได้ทั้งแบบจุด (Point forecast) และแบบช่วง (Interval forecast) โดยการพยากรณ์จะใช้สมการพยากรณ์ที่สร้างจากรูปแบบการพยากรณ์ที่กำหนดและผ่านการตรวจสอบในขั้นตอนที่แล้วมา

เมื่อได้แบบจำลองที่มีความเหมาะสมหลังจากการตรวจสอบความถูกต้องแล้วก็สามารถนำแบบจำลองไปใช้ในการพยากรณ์โดยแบ่งการพยากรณ์ออกเป็น 3 ช่วง ดังนี้

- 1) Historical Forecast – การพยากรณ์ตั้งแต่อดีตจนถึงช่วงเวลาที่เริ่มพิจารณา
- 2) Ex – post Forecast – การพยากรณ์โดยการตัดข้อมูลออกมาส่วนหนึ่ง แล้วทำการพยากรณ์เปรียบเทียบกับข้อมูลที่ได้จากการพยากรณ์
- 3) Ex – ante Forecast – การพยากรณ์ข้อมูลไปข้างหน้า โดยนำแบบจำลองที่หาได้ในช่วงที่ 2 มาทำการพยากรณ์



### 3.1.4 แบบจำลอง GARCH – in – mean (GARCH-M)

$$Y_t = \alpha + \sum_{i=1}^p Y_{t-i} + \sum_{j=1}^q \epsilon_{t-j} + \gamma \quad (3.10)$$

$$\sigma_t^2 = \alpha + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \quad (3.11)$$

จากสมการ (3.10) เรียกว่าสมการค่าเฉลี่ย (Mean Equation) และสมการที่ (3.11) เรียกว่า Generalized Autoregressive Conditional Heteroscedasticity: GARCH (p,q) โดยแบบจำลอง GARCH-M นั้น เป็นแบบจำลองที่ได้นำเอาค่าจากสมการ GARCH(p,q) ที่ได้ทำการยกกำลังเศษหนึ่งส่วนสองแล้วนำไปใส่ในสมการ Mean Equation เพื่อทำการประมาณค่าของตัวแปรที่ถูกศึกษา

### 3.2 วิธีการวิจัย

วิธีการวิจัยแบ่งได้ออกเป็น 3 ส่วนคือ วิธีการศึกษาด้วยแบบจำลองนิรอลเน็ตเวิร์ค, วิธีการศึกษาด้วยแบบจำลอง GARCH –M และวิธีการศึกษาด้วยแบบจำลอง ARIMA ซึ่งรายละเอียดดังต่อไปนี้

#### 3.2.1 วิธีการศึกษาด้วยแบบจำลองนิรอลเน็ตเวิร์ค

การศึกษาด้วยแบบจำลองนิรอลเน็ตเวิร์คประกอบไปด้วย 3 การทดลอง ดังนี้คือ

- 1) การทดลองเบื้องต้นด้วย Hidden Layer จำนวน 1 ชั้น
- 2) การทดลองปรับจำนวน Hidden Layer ให้เป็น 2 ชั้น
- 3) การทดลองปรับเปลี่ยนจำนวนนิรอลใน Hidden Layer ด้วยวิธี Quadratic

Interpolation

รายละเอียดของแต่ละการทดลอง มีดังนี้

##### 1) การทดลองเบื้องต้นด้วย Hidden Layer จำนวน 1 ชั้น

ขั้นที่ 1 แบ่งข้อมูลที่มีอยู่ออกเป็น 3 ช่วง คือ Training 450 ข้อมูล, Validation 50 ข้อมูล และ Testing 50 ข้อมูล

ขั้นที่ 2 จัดเรียงข้อมูลเพื่อกำหนด input ให้เท่ากับ 10 ตัว

ขั้นที่ 3 กำหนดจำนวนนิรอลในชั้น Hidden Layer เท่ากับ 10 โดยให้แบบจำลองเรียนรู้จากข้อมูลชุด Training จำนวน 100 รอบ และคำนวณค่า MSE ดังสมการที่ (3.12) ออกมาด้วยการทำนายค่าของชุดข้อมูล Validation



$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2 \quad (3.12)$$

โดยที่  $y_i$  = ค่าพยากรณ์

$t_i$  = ค่าจริง

$n$  = จำนวนข้อมูลของชุดข้อมูล Validation

- อีกครั้ง
- ขั้นที่ 4 ให้แบบจำลองเรียนรู้เพิ่มอีก 100 รอบ จากนั้นจึงคำนวณค่า MSE ออกมา
- ขั้นที่ 5 ทำซ้ำครั้งที่ 4 จนกระทั่งครบ 1000 รอบ
- ขั้นที่ 6 เปลี่ยนจำนวนนิวรอนในชั้น Hidden Layer ให้เท่ากับ 20 แล้วทำซ้ำขั้นที่ 4 และขั้นที่ 5
- ขั้นที่ 7 ทำซ้ำขั้นที่ 6 โดยเปลี่ยนจำนวนนิวรอนเป็น 30, 40, 50, 60, 70, 80, 90 และ 100
- ขั้นที่ 8 เปรียบเทียบค่า MSE ที่ได้ทั้งหมด แล้วเลือกแบบจำลองที่ดีที่สุดด้วยค่า MSE ที่น้อยที่สุด
- ขั้นที่ 9 ใช้แบบจำลองที่ดีที่สุดพยากรณ์ค่าของชุดข้อมูล Testing และคำนวณค่า MAPE ดังสมการที่ (3.13) ออกมา

$$MAPE = \frac{1}{n} \sum_{i=1}^n APE_i \quad (3.13)$$

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

โดยที่  $t_i$  = ค่าจริง

$y_i$  = ค่าพยากรณ์

$n$  = จำนวนข้อมูลของชุดข้อมูล Testing

หมายเหตุ : โครงสร้างของแบบจำลองนิวรอนเน็ตเวิร์คที่ใช้

จำนวนชั้น Hidden Layer : 1

จำนวนนิวรอนในชั้นแสดงผล : 1

ฟังก์ชันแปลงค่าในชั้นซ่อนเร้น : purelin

ฟังก์ชันแปลงค่าในชั้นแสดงผล : purelin

วิธีการเรียนรู้ : traingcf

โปรแกรมที่ใช้ : Matlab 6.5



## 2) การทดลองปรับจำนวน Hidden Layer ให้เป็น 2 ชั้น

ขั้นที่ 1 จากการทดลองเบื้องต้นด้วย Hidden Layer จำนวน 1 ชั้น นำค่าของจำนวนนิวรอนในชั้น Hidden Layer ที่ดีที่สุดมากำหนดไว้ให้เป็นค่าคงที่ แล้วทำการหาจำนวนนิวรอนที่เหมาะสมในชั้น Hidden Layer ที่สองโดยกำหนดให้มีค่าเท่ากับ 10 20 30 40 50 60 70 80 90 และ 100 และหาจำนวนรอบการเรียนรู้ที่เหมาะสม โดยกำหนดให้มีค่าเท่ากับ 100 200 300 400 500 600 700 800 900 และ 1000 รอบ

ขั้นที่ 2 เลือกจำนวนนิวรอนในชั้น Hidden Layer ที่สองและจำนวนรอบการเรียนรู้ที่เหมาะสม โดยพิจารณาจากค่า MSE ที่ต่ำที่สุด แล้วนำไปพยากรณ์ราคาทองคำ

ขั้นที่ 3 นำผลการพยากรณ์ราคาทองคำที่ได้คำนวณค่า MAPE

หมายเหตุ : โครงสร้างของแบบจำลองนิวรอนเน็ตเวิร์คที่ใช้

จำนวนชั้น Hidden Layer	: 2
จำนวนนิวรอนในชั้นแสดงผล	: 1
ฟังก์ชันแปลงค่าในชั้นซ่อนเร้นที่ 1	: purelin
ฟังก์ชันแปลงค่าในชั้นซ่อนเร้นที่ 2	: purelin
ฟังก์ชันแปลงค่าในชั้นแสดงผล	: purelin
วิธีการเรียนรู้	: traincgf
โปรแกรมที่ใช้	: Matlab 6.5

## 3) การทดลองปรับเปลี่ยนจำนวนนิวรอนใน Hidden Layer ด้วยวิธี Quadratic

### Interpolation

ขั้นที่ 1 กำหนดจำนวนข้อมูลนำเข้าให้เท่ากับ 10

ขั้นที่ 2 กำหนดจำนวนนิวรอนในชั้น Hidden Layer ที่หนึ่งเท่ากับ 10

ขั้นที่ 3 ทำการหาจำนวนนิวรอนในชั้น Hidden Layer ที่สอง ด้วยวิธี Quadratic Interpolation โดยกำหนดให้จำนวนนิวรอนในชั้น Hidden Layer ที่สองมีค่าเท่ากับ 10 30 และ 50

ขั้นที่ 4 เลือกจำนวนนิวรอนในชั้น Hidden Layer ที่สองที่ให้ค่า MSE ที่ต่ำที่สุด เพื่อใช้ในการคำนวณ Quadratic Interpolation ในขั้นต่อไป จนกว่าจะได้ค่า MSE ของจำนวนนิวรอนในชั้น Hidden Layer ที่สองที่มีค่า MSE ต่ำที่สุด

ขั้นที่ 5 เมื่อได้จำนวนนิวรอนในชั้น Hidden Layer ที่สองที่เหมาะสมแล้ว ทำการกำหนดให้มีค่าคงที่ แล้วทำการหาจำนวนนิวรอนในชั้น Hidden Layer ที่หนึ่งเช่นเดียวกับการหาจำนวนนิวรอนในชั้น Hidden Layer ที่สอง

ขั้นที่ 6 เมื่อได้จำนวนนิวรอลในชั้น Hidden Layer ที่หนึ่งและชั้นที่สองที่เหมาะสมแล้ว ให้ทำการหาจำนวนข้อมูลนำเข้าโดยวิธี Quadratic Interpolation โดยกำหนดให้จำนวนข้อมูลนำเข้ามีค่าระหว่าง 10 ถึง 50

ขั้นที่ 7 เมื่อได้จำนวนนิวรอลในชั้น Hidden Layer ที่ 1, ชั้นที่สองและจำนวนข้อมูลนำเข้าที่เหมาะสมแล้ว จึงนำไปพยากรณ์ราคาทองคำและคำนวณหาค่า MAPE

หมายเหตุ : โครงสร้างของแบบจำลองนิวรอลเน็ตเวิร์คที่ใช้

จำนวนชั้น Hidden Layer	: 2
จำนวนนิวรอลในชั้นแสดงผล	: 1
ฟังก์ชันแปลงค่าในชั้นซ่อนเร้นที่ 1	: purelin
ฟังก์ชันแปลงค่าในชั้นซ่อนเร้นที่ 2	: purelin
ฟังก์ชันแปลงค่าในชั้นแสดงผล	: purelin
วิธีการเรียนรู้	: traincgf
โปรแกรมที่ใช้	: Matlab 6.5

### 3.2.2 วิธีการศึกษาด้วยแบบจำลอง ARIMA

ขั้นที่ 1 ทดสอบความนิ่งของข้อมูลด้วยวิธี Unit Root Test

ขั้นที่ 2 กำหนดรูปแบบของแบบจำลอง ARIMA (p,d,q) โดยการพิจารณาจาก Correlogram เพื่อจะสามารถระบุได้ว่าแบบจำลองควรจะมี Autoregressive (q) เท่าใด และ Moving average (q) เท่าใด

ขั้นที่ 3 การประมาณค่าพารามิเตอร์ เพื่อนำค่าพารามิเตอร์ไปทำการพยากรณ์ราคา

ขั้นที่ 4 เปรียบเทียบแบบจำลองเพื่อนำแบบจำลองที่เหมาะสมที่สุดไปใช้ในการพยากรณ์

ขั้นที่ 5 พยากรณ์ราคาทองคำ 50 วัน

### 3.2.3 วิธีการศึกษาด้วยแบบจำลอง GARCH-M

ขั้นที่ 1 จากรูปแบบของแบบจำลอง ARIMA ในหัวข้อที่ 3.2.2 ที่เหมาะสมต่อการพยากรณ์ นำแบบจำลองที่ได้มาประมาณค่าพารามิเตอร์ โดยวิเคราะห์ตามวิธีแบบจำลอง GARCH-M เพื่อหารูปแบบที่เหมาะสมที่จะใช้ในการพยากรณ์

ขั้นที่ 2 ตรวจสอบความถูกต้องของแบบจำลองโดยพิจารณาจากค่า Q-Statistic

ขั้นที่ 3 พยากรณ์ราคาทองคำ 50 วัน