

## บทที่ 3

### กรอบแนวความคิดและวิธีวิจัย

#### 3.1 กรอบแนวคิด

##### 3.1.1 แนวคิดการพยากรณ์ข้อมูลอนุกรมเวลา

ข้อมูลอนุกรมเวลา (Time series data) คือค่าสังเกต (Observation) ชุดหนึ่งซึ่งถูกกำหนดขึ้น ณ เวลาต่างๆ ถ้าค่าสังเกตกระทำในเวลาต่อเนื่องกันจะเรียกว่า ข้อมูลอนุกรมเวลาต่อเนื่อง แต่ถ้าค่าสังเกตกระทำ ณ จุดเวลาที่ไม่ต่อเนื่องกัน เรียกว่า ข้อมูลอนุกรมเวลาไม่ต่อเนื่อง ดังนั้นการวิเคราะห์ ข้อมูลอนุกรมเวลาจึงเป็นการวิเคราะห์ค่าสังเกตที่มีการเปลี่ยนแปลงไปตามเวลาที่กระทำ และลักษณะของการเปลี่ยนแปลงอาจจะมีรูปแบบหรือไม่มีก็ได้ แต่ถ้าข้อมูลอนุกรมเวลามีลักษณะการเปลี่ยนแปลงที่มีรูปแบบก็จะสามารถพยากรณ์รูปแบบในอนาคตได้ดีขึ้น

##### 3.1.2 การพยากรณ์ด้วยแบบจำลอง ARIMA

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

การพยากรณ์ด้วยวิธี Box and Jenkins เราต้องทราบว่าอนุกรมเวลาที่นำมาศึกษามีลักษณะเป็น Stationary หรือไม่ โดยพิจารณาจาก  $E(Y_t)$  และ  $Var(Y_t)$  คงที่สำหรับทุกค่าของ  $t$  หรือไม่ ถ้าอนุกรมเวลาเป็น Stationary และไม่มีอิทธิพลของฤดูกาลและแนวโน้ม สามารถกำหนดรูปแบบ ARMA(p,q) (ทรงศิริ แต่สมบัติ, 2539)

ถ้า  $Y_t$  เป็นอนุกรมเวลาที่คงที่ (Stationary) ซึ่งตัวแปร  $Y_t$  มีคุณสมบัติ คือ

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

$$\text{Variance} : Var(Y_t) = E(Y_t - \mu)^2 = \sigma^2$$

$$\text{Covariance} : E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k$$

ถ้า  $Y_t$  เป็นอนุกรมเวลาที่ไม่คงที่ (Nonstationary) ซึ่งตัวแปร  $Y_t$  มีคุณสมบัติ คือ

$$\text{Mean} : E(Y_t) = t\mu$$

$$\text{Variance} : \text{Var}(Y_t) = E(Y_t - \mu)^2 = t\sigma^2$$

$$\text{Covariance} : E[(Y_t - \mu)(Y_{t+k} - \mu)] = t\gamma_k$$

ถ้าอนุกรมเวลามีลักษณะเป็น Nonstationary ต้องแปลงอนุกรมเวลาให้เป็นอนุกรมเวลาใหม่ให้เป็น Stationary ด้วยวิธีการต่างๆดังนี้

วิธีที่ 1 การหาผลต่าง (regular difference) ถ้าอนุกรมเวลาได้รับอิทธิพลของแนวโน้ม จะต้องแปลงให้เป็นอนุกรมเวลาใหม่ที่ไม่มีแนวโน้ม

โดย  $Z_t = \Delta^d Y_t$  d : ลำดับของการหาผลต่าง

$$\text{เมื่อ } d=1 \quad Z_t = \Delta Y_t = Y_t - Y_{t-1}$$

$$\text{เมื่อ } d=2 \quad Z_t = \Delta^2 Y_t = \Delta(Y_t - Y_{t-1}) = \Delta Y_t - \Delta Y_{t-1} = Y_t - 2Y_{t-1} + Y_{t-2}$$

เมื่อรูปแบบของ Box and Jenkins อยู่ในรูปแบบ Backward shift operator ซึ่งกำหนดสัญลักษณ์เป็น “B” กำหนดให้  $BY_t = Y_{t-1}$  และ  $B^d Y_t = Y_{t-d}$  จาก

$$\Delta Y_t = Y_t - Y_{t-1}$$

$$\Delta Y_t = Y_t - BY_t$$

$$\Delta Y_t = (1 - B)Y_t$$

### 3.1.3 การทดสอบความนิ่งของข้อมูล ( UNIT ROOT TEST)

เป็นการทดสอบว่าข้อมูลนั้นมีค่าเฉลี่ยและค่าความแปรปรวนของตัวแปรคงที่หรือไม่ ซึ่งก็คือการทดสอบว่าข้อมูลมีความนิ่ง (Stationary) หรือไม่ เรียกว่า การทดสอบ Unit Root โดยสมมติและตั้งสมมติฐานว่าง (Null Hypothesis) ของการทดสอบ คือ  $H_0 = \alpha = 1$  ซึ่งสมการที่ใช้ในการทดสอบ คือ  $\mu_t = \alpha\mu_{t-1} + \varepsilon$  การทดสอบนี้เรียกว่า Dickey and Fuller test (Dickey and Fuller, 1981) นอกจากนั้นยังเรียกว่า การทดสอบ Augmented Dickey and Fuller test (Said and Dickey, 1984) ดังมีรายละเอียดดังนี้

#### ก. การทดสอบ DF (Dickey and Fuller test)

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

ผลการทดสอบ Unit Root ถ้า  $|\rho| < 1$   $X_t$  จะมีลักษณะนิ่ง แต่ถ้า  $|\rho| = 1$   $X_t$  จะมีลักษณะไม่นิ่ง และจากสมการข้างต้น สามารถเปลี่ยนรูปแบบสมการได้เป็น

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

จากสมการดังกล่าว จะสามารถพิจารณาความนิ่งของข้อมูลอนุกรมเวลาได้ เมื่อ  $\rho=(1+\theta)$  ซึ่ง ถ้า  $\theta$  มีค่าเป็นลบ จะได้ว่า  $\rho$  มีค่าน้อยกว่า 1

ดังนั้นสรุปได้ว่าการปฏิเสธสมมติฐานว่าง(Null Hypothesis)  $H_0 : \theta = 0$  ถือว่าเป็นการยอมรับสมมติฐาน  $H_a : \theta < 0$  แสดงว่า  $\rho < 1$  และ  $X_t$  มีลักษณะ Integration of order zero นั่นคือ  $X_t$  มีลักษณะนิ่ง ในทางตรงกันข้ามถ้ายอมรับ  $H_0 : \theta = 0$  หมายความว่า  $X_t$  มีลักษณะไม่นิ่ง อย่างไรก็ตาม ถ้าอนุกรมเวลามรลักษณะเป็น Random Walk ซึ่งมีแนวโน้มรวมอยู่ด้วย สามารถเขียนเป็นแบบจำลองที่อยู่ในรูป First Difference ดังนี้

$$\text{กรณีทั่วไป} \quad \Delta X_t = \theta X_{t-1} + \varepsilon_t$$

$$\text{กรณีมีค่าคงที่} \quad \Delta X_t = \alpha + \theta X_{t-1} + \varepsilon_t$$

$$\text{กรณีมีค่าคงที่และแนวโน้ม} \quad \Delta X_t = \alpha + \beta t + \theta X_{t-1} + \varepsilon_t$$

#### ข. การทดสอบ ADF (Augmented Dickey and Fuller test)

การทดสอบอนุกรมเวลาที่มีปัญหา Serial Correlation ในค่า Error term ( $\varepsilon_t$ ) เป็นการทดสอบที่ DF-Test ไม่สามารถทำได้ ดังนั้น ADF จึงเพิ่มค่า Lagged Change  $\left[ \sum_{j=1}^p \lambda_j \Delta X_{t-j} \right]$  เข้า

$$\text{ไปในสมการดังนี้} \quad \Delta X_t = \theta X_{t-1} + \sum_{i=1}^p \phi_i \Delta X_{t-i} + \varepsilon_t$$

$$\Delta X_t = \alpha + \theta X_{t-1} + \sum_{i=1}^p \phi_i \Delta X_{t-i} + \varepsilon_t$$

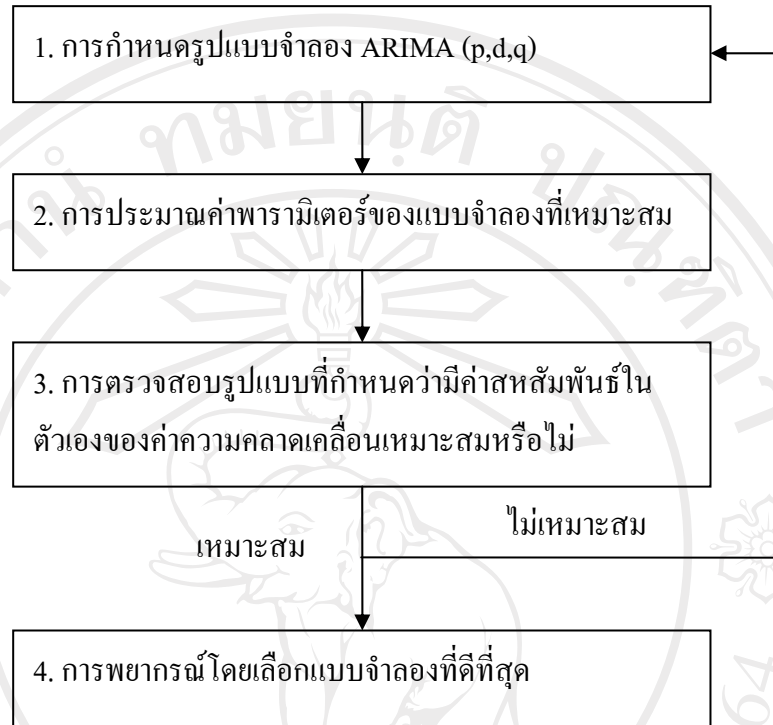
$$\Delta X_t = \alpha + \beta t + \theta X_{t-1} + \sum_{i=1}^p \phi_i \Delta X_{t-i} + \varepsilon_t$$

การใส่ค่า Lagged Term ( $\rho$ ) ว่ามีจำนวนเท่าใดจึงจะเหมาะสมสำหรับแต่ละข้อมูลอนุกรมนั้นมีหลักการในการเลือก lag length ที่เสนอโดย Enders (1995) ว่า ควรจะเริ่ม lag length ที่มีค่าที่มากพอ แล้วพิจารณาความมีนัยสำคัญทางสถิติ ที่ระดับนัยสำคัญต่าง ๆ ( $\sigma = 0.01, 0.05$  และ  $0.1$ ) เมื่อพบว่าที่ lag length ที่เลือกมีค่า t-statistic ที่ไม่มีนัยสำคัญทางสถิติ ณ ระดับนัยสำคัญร้อยละ 10 แล้วจึงทำการลด lag length ลงทีละ 1 ช่วง จนกระทั่งสามารถปฏิเสธสมมติฐานว่าง

#### 3.1.4 การพยากรณ์โดยวิธี Box – Jenkins

การพยากรณ์ด้วยวิธี Box and Jenkins หรือเรียกว่า ARIMA ใช้กับอนุกรมเวลาที่มีลักษณะเป็น Stationary และไม่มีอิทธิพลของฤดูกาลและแนวโน้ม ซึ่งสามารถทำตามขั้นตอนของ Box and Jenkins ได้ดังนี้

## ขั้นตอนการพยากรณ์โดยวิธี ARIMA



รูปที่ 3.1 ขั้นตอนการพยากรณ์โดยวิธี ARIMA

### ขั้นที่ 1 การกำหนดรูปแบบ (Identification)

การกำหนดรูปแบบจำลอง ARIMA(p,d,q) โดยการพิจารณาคอเรลโลแกรม Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF) เพื่อจะสามารถระบุได้ว่าแบบจำลองควรมี Autoregressive (p) เท่าใด และ Moving average (q) เท่าใด โดยเลือกรูปแบบจำลองที่หลากหลายเพื่อหาแบบจำลองที่เหมาะสมที่สุด ที่คาดว่าเหมาะสมให้กับอนุกรมเวลา โดยพิจารณาเปรียบเทียบจาก คอเรลโลแกรมของค่า  $r_k$  และ  $r_{kk}$  ของอนุกรมเวลา และอาศัยการพิจารณาค่าสถิติเพื่อประกอบการตัดสินใจ เช่น Root Mean Squared Error (RMSE) ค่า Theil's inequality coefficient ค่า Adjusted  $R^2$  ค่า Akaike Information Criterion (AIC) และค่า Schwarz's Bayesian Information Criterion (SBC)

#### ก. ค่า Root Mean Squared Error ( RMSE )

RMSE คือ การวัดค่าความแตกต่างระหว่างค่าจริง และค่าที่ประมาณได้จากแบบจำลอง หาก RMSE มีค่าน้อย แสดงว่าแบบจำลองสามารถประมาณค่าประมาณได้ใกล้เคียง

กับค่าจริง (Pindyck, 1998) ดังนั้นหาก RMSE มีค่าเท่ากับศูนย์แล้ว จะหมายความว่า ไม่เกิดความคลาดเคลื่อนในแบบจำลองนี้เลย โดย RMSE คำนวณได้ดังนี้

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t^s - Y_t^a)^2}$$

โดยกำหนด  $Y_t^s$  = ค่าประมาณจากแบบจำลอง

$Y_t^a$  = ค่าที่แท้จริง

T = จำนวนคาบเวลาที่ใช้ในการประมาณแบบจำลอง

#### ข. ค่า Adjusted $R^2$

Adjusted  $R^2$  คือ การพิจารณาว่าตัวแปรอิสระสามารถที่จะอธิบายถึงการเปลี่ยนแปลงของตัวแปรตามได้มากน้อยเพียงใด ถ้าหากค่า Adjusted  $R^2$  มีค่าเท่ากับ 1 แสดงว่า ตัวแปรอิสระสามารถอธิบายตัวแปรตามได้ทั้งหมด แต่ถ้าหากค่า Adjusted  $R^2$  มีค่าเท่ากับ 0 หมายความว่า ตัวแปรอิสระไม่สามารถอธิบายตัวแปรตามได้เลย ซึ่งค่า Adjusted  $R^2$  นี้เป็นค่าสถิติที่ประยุกต์มาจากค่า  $R^2$  ซึ่งถ้ามีการเพิ่มตัวแปรอิสระเข้าไปในสมการมากขึ้นก็จะทำให้ค่า R Square สูงขึ้น ดังนั้น จึงมีการเพิ่มระดับความเป็นอิสระในสมการ ซึ่งเรียกว่า Adjusted  $R^2$  (Gujarati, 2003) โดยสามารถพิจารณาความสัมพันธ์ของ  $R^2$  และ Adjusted  $R^2$  ได้ดังสมการ

$$R^2 = 1 - \frac{\sum \hat{u}_i}{\sum y_i^2}$$

$$\overline{R^2} = 1 - \frac{\sum u_i^2 / (n - k)}{\sum y_i^2 / (n - 1)}$$

#### ค. ค่า Akaike Information Criterion (AIC)

ค่า AIC เป็นค่าสถิติที่มีการประยุกต์คล้ายกับค่า Adjusted  $R^2$  แต่มีการถ่วงน้ำหนักมากกว่า Adjusted  $R^2$  และยังมีการใช้ลอการิทึมฐานธรรมชาติ (Natural logarithm : ln) ดังนั้นถ้าค่า AIC น้อยจึงหมายถึงแบบจำลองสามารถเป็นตัวแทนข้อมูลจริงได้ดี (Gujarati, 2003) และยังสามารถนำค่า AIC ไปใช้ในการหาค่าย้อนหลัง (lag length) ที่เหมาะสมได้อีก สามารถคำนวณค่า AIC ได้ดังนี้

$$\ln AIC = \left( \frac{2k}{n} \right) + \ln \left( \frac{\sum \hat{\mu}_i^2}{n} \right)$$

โดยกำหนดให้  $\sum \hat{\mu}_i^2 =$  ผลรวมของค่าความคลาดเคลื่อนกำลังสอง  
 $N =$  ค่าสังเกตทั้งหมด

### ง. ค่า Schwarz's Bayesian Information Criterion (SBC)

ค่า SBC นำเสนอโดย Schwarz (1990) เป็นเกณฑ์ที่ใช้การเลือกแบบจำลอง เหมือนค่า AIC คือ ถ้าค่า SBC น้อย ก็หมายถึงแบบจำลองสามารถเป็นตัวแทนข้อมูลจริงได้ดีสามารถคำนวณค่า SBC ได้ดังนี้

$$\ln SBC = \log \hat{\sigma}^2 + 2 \frac{p+q}{T} \log T$$

โดยกำหนดให้  $T =$  จำนวนคาบเวลาที่ใช้ในการประมาณแบบจำลอง

### ขั้นที่ 2 การประมาณค่าพารามิเตอร์ของแบบจำลองที่เหมาะสม (Estimation)

การประมาณค่า คือการประมาณค่าสัมประสิทธิ์ที่มาจากรูปแบบการถดถอยในตัวเอง และรูปแบบการเคลื่อนที่ของค่าความคลาดเคลื่อน ซึ่งสามารถเลือกใช้ วิธีการถดถอยเชิงเส้นอย่างง่าย (Ordinary least squares) หรือใช้วิธีการถดถอยแบบไม่เป็นเชิงเส้น (Nonlinear) เพื่อสร้างความสัมพันธ์ของสมการที่จะสามารถนำไปใช้ในการพยากรณ์ได้เมื่อรูปแบบความสัมพันธ์นั้นเป็นรูปแบบที่มีความเหมาะสมที่สุด

### ขั้นที่ 3 การวิเคราะห์ความถูกต้อง (Diagnostic Checking)

การวิเคราะห์ความถูกต้อง หมายถึง การตรวจสอบรูปแบบจำลองว่ามีความเหมาะสมหรือไม่ โดยการพิจารณาจากคอเรโลแกรมของอัตราสัมพันธ์ของกลุ่มตัวอย่าง ซึ่งเป็นการทดสอบโดยใช้ การทดสอบของ Box and Pierce ซึ่งพิจารณาจาก ค่า Q-statistic (Gujarati, 1995) ดังสมการด้านล่าง ซึ่งมีการแจกแจงแบบ Chi-square และมี Degree of freedom เท่ากับ  $m$  โดยมีสมมติฐานว่าง (Null Hypothesis) คือ พจน์ความคลาดเคลื่อนที่ได้จากการประมาณที่มีลักษณะเป็น white noise ซึ่งหมายถึง แบบจำลองไม่มีอัตราสัมพันธ์ ถ้าหากแบบจำลองมีอัตราสัมพันธ์จะตัวเองกลับไปทำการกำหนดรูปแบบจำลองใหม่ แต่หากแบบจำลองไม่มีอัตราสัมพันธ์จะใช้แบบจำลองนั้นทำการพยากรณ์ต่อไป

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

โดยกำหนดให้  $n$  = จำนวนของข้อมูล

$m$  = ค่า lag length

#### ขั้นที่ 4 การพยากรณ์ (Forecasting)

เมื่อได้แบบจำลองที่เหมาะสมภายหลังจากการวิเคราะห์ความถูกต้องแล้วก็สามารถนำแบบจำลองดังกล่าวใช้ในการพยากรณ์

#### 3.1.5 แบบจำลอง Autoregressive Conditional Heteroskedastic (ARCH)

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

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

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

อย่างไรก็ตามถ้าการพยากรณ์อย่างไม่มีเงื่อนไขถูกนำมาใช้ ก็หมายถึงค่าเฉลี่ยของ  $\{X_t\}$

ในระยะยาวเท่ากับ  $\frac{a_0}{(1-a_1)}$  จะได้ว่าพยากรณ์อย่างไม่มีเงื่อนไขของค่าความแปรปรวนของความคลาดเคลื่อนคือ

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

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

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

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

$$\begin{aligned} \text{Var}(X_{t+1} | X_t) &= E[(X_{t+1} - a_0 - a_1 X_t)^2] \\ &= E_t \varepsilon_{t+1}^2 \end{aligned}$$

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

การใช้ลำดับของ Autoregressive หรือ AR(q) โดยการยกกำลังสองของส่วนเหลือที่ถูกประมาณค่า ดังนี้

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + V_t$$

เมื่อ  $V_t$  คือ White noise process

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

### 3.1.6 แบบจำลอง Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

แบบจำลอง Generalized Autoregressive Conditional Heteroscedasticity (GARCH) เป็นแบบจำลองที่ Bollerslev ได้ขยายมาจาก ARCH model โดยเสนอแนวคิดที่ว่า การพยากรณ์ค่าความแปรปรวนอย่างมีเงื่อนไขสามารถอธิบายจาก ค่าความแปรปรวนและค่าความคลาดเคลื่อน ที่มีลักษณะ ARMA process พร้อมกันได้ ต่อไปนี้

$$X_t = a_0 + a_1 X_{t-1} + \varepsilon_t \quad (3.16)$$

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

จากสมการ (3.16) เป็นสมการค่าเฉลี่ย (Mean Equation) ที่ถูกกำหนดจากตัวแปรภายนอกและค่าความคลาดเคลื่อน  $\varepsilon_t$  ส่วนสมการ (3.17) เป็นสมการค่าความแปรปรวนในคาบเวลาปัจจุบัน ที่ถูกกำหนดโดย 3 ฟังก์ชัน คือ ค่าคงที่  $\alpha_0$  เหตุการณ์ที่สำคัญที่ส่งผลต่อความคลาดเคลื่อนในคาบเวลาที่ผ่านมา โดยถูก กำหนดให้อยู่ในรูป กำลังสองของส่วนเหลือจากสมการ



ค่าเฉลี่ย  $E\varepsilon_{t-i}^2$  ซึ่งก็คือ ARCH term นั่นเองค่าพยากรณ์ความแปรปรวนในคาบเวลาที่ผ่านมา  $\sigma_{t-i}^2$  ซึ่งก็คือ GARCH term นั่นเอง

ดังนั้นจากสมการ (3.17) จึงถูกเรียกว่า Generalized Autoregressive Conditional Heteroscedasticity :GARCH (p,q) ซึ่งได้เปิดโอกาสให้มีส่วนประกอบที่เป็น Autoregressive Moving Average ในความแปรปรวนที่มีลักษณะ Heteroscedastic Variance จะเห็นได้ว่า ถ้า  $p=0$  และ  $q=1$  เราจะได้แบบจำลอง GARCH(0,1) ซึ่งก็คือ ARCH(1) หรือ ARCH( $q=1$ ) นั่นเอง โดยสรุปว่าถ้า  $\beta_i$  ทุกตัวมีค่าเท่ากับศูนย์ แบบจำลอง GARCH ก็คือ ARCH( $q$ ) นั่นเอง

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

Engle (1987. อ้างถึงใน จิตติ ดันเสนีย์, 2549) ได้ทำการขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชันของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักในชื่อของ GARCH-in-Mean หรือ GARCH-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

### 3.1.8 Artificial neural network

โครงข่ายประสาทเทียม (Artificial neural network) หรือที่มักจะเรียกสั้น ๆ ว่า ข่ายงานประสาท (neural network หรือ neural net) คือโมเดลทางคณิตศาสตร์ สำหรับประมวลผลสารสนเทศด้วยการคำนวณแบบคอนเนกชันนิสต์ (connectionist) เพื่อจำลองการทำงานของเครือข่ายประสาทในสมองมนุษย์ ด้วยวัตถุประสงค์ที่จะสร้างเครื่องมือซึ่งมีความสามารถในการเรียนรู้การจดจำแบบรูป (Pattern Recognition) และการอุปมานความรู้ ( Knowledge deduction) เช่นเดียวกับความสามารถที่มีในสมองมนุษย์ แนวคิดเริ่มต้นของเทคนิคนี้ได้มาจากการศึกษาข่ายงานไฟฟ้าชีวภาพ (bioelectric network) ในสมอง ซึ่งประกอบด้วย เซลล์ประสาท หรือ “นิวรอน” (neurons) และ จุดประสานประสาท (synapses) แต่ละเซลล์ประสาทประกอบด้วยปลายในการรับกระแสประสาท เรียกว่า “เดนไดรท์” (Dendrite) ซึ่งเป็น input และปลายในการส่งกระแสประสาท

เรียกว่า "แอกซอน" (Axon) ซึ่งเป็นเหมือน output ของเซลล์ เซลล์เหล่านี้ทำงานด้วยปฏิกิริยาไฟฟ้าเคมี เมื่อมีการกระตุ้นด้วยสิ่งเร้าภายนอกหรือกระตุ้นด้วยเซลล์ด้วยกัน กระแสประสาทจะวิ่งผ่านเดนไดรต์เข้าสู่นิวเคลียสซึ่งจะเป็นตัวตัดสินใจว่าต้องกระตุ้นเซลล์อื่น ๆ ต่อหรือไม่ ถ้ากระแสประสาทแรงพอ นิวเคลียสก็จะกระตุ้นเซลล์อื่น ๆ ต่อไปผ่านทางแอกซอนของมัน

ตามโมเดลนี้ทำงานประสาทเกิดจากการเชื่อมต่อระหว่างเซลล์ประสาท จนเป็นเครือข่ายที่ทำงานร่วมกัน

### 3.1.9 Artificial neural network แบบ Feed Forward

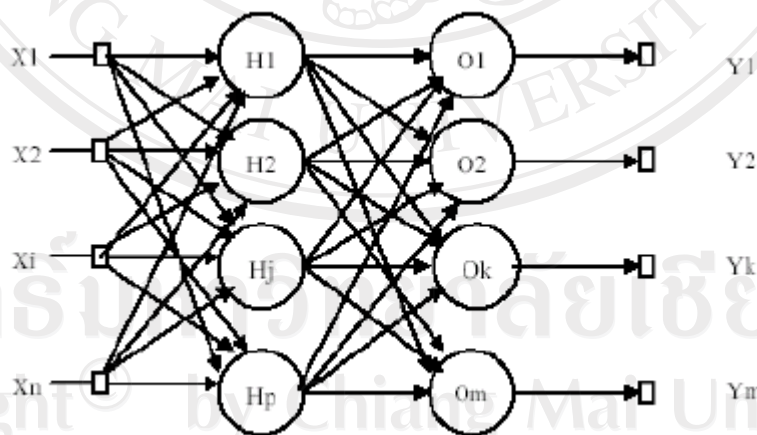
Neural Networks แบบ feed forward จะรับข้อมูลนำเข้า (input) แล้วทำการประมวลผลโดย นิวรอล ในชั้น hidden layer และชั้น output จากนั้นจึงจะส่งผลลัพธ์ออกมาเป็นคำตอบ การที่สัญญาณนำเข้าถูกถ่ายทอดไปข้างหน้าเรื่อยๆเรียกว่า feed forward และการที่ใช้จำนวนนิวรอลหลายชั้นเรียกว่า multi layer (คมตัน สุริยะ, 2548: 3)

องค์ประกอบของ Artificial neural network แบบ Feed Forward มีดังนี้

1) input

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

รูปที่ 3.2 แสดงแบบจำลอง Multilayer Feed Forward แบบ Feed Forward



2) Output

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

3) Neuron ใน Hidden Layer

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

4) Neuron ใน Output Layer

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

5) คำนวณน้ำหนักจากชั้น Input สู่อัน Hidden Layer

คำนวณน้ำหนักจากชั้น Input สู่อัน Hidden Layer สำหรับ Neuron แต่ละตัวในชั้น Hidden Layer มีจำนวน  $n+1$  ตัว คือ  $W_{ij}^H = \{ W_{0j}^H, W_{1j}^H, W_{2j}^H, \dots, W_{ij}^H, \dots, W_{nj}^H \}$

ดังนั้น เมื่อมีจำนวน Input จำนวน  $n$  ตัว และมีนิวรอนใน Hidden Layer จำนวน  $p$  ตัว แล้วจะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ  $(n+1)p$  ตัว

6) คำนวณน้ำหนักจากชั้น Hidden Layer สู่อัน Output

คำนวณน้ำหนักจากชั้น Hidden Layer สู่อัน Output สำหรับ Neuron แต่ละตัวในชั้น Output มีจำนวน  $m+1$  ตัว คือ  $W_{ik}^O = \{ W_{0k}^O, W_{1k}^O, W_{2k}^O, \dots, W_{ik}^O, \dots, W_{nk}^O \}$

ดังนั้น เมื่อมีจำนวนนิวรอนใน Hidden Layer จำนวน  $p$  ตัว และมี Output จำนวน  $m$  ตัว แล้วจะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ  $(p+1)m$  ตัว (จิตติ ดันเสนีย์, 2549)

### 3.1.10 การคำนวณแบบแพร่ย้อนกลับ (Back Propagation)

ขั้นตอนของ Back-propagation Algorithm มีดังนี้

1. กำหนดค่าอัตราเร็วในการเรียนรู้ (rate parameter :  $r$ )
2. สำหรับแต่ละตัวอย่าง input ให้ทำตามขั้นตอนต่อไปนี้จนกว่าได้ระดับ performance ที่ต้องการ

- คำนวณค่า output โดยใช้ค่าน้ำหนักเริ่มต้นซึ่งอาจได้จากการสุ่ม
- คำนวณค่า  $\beta$  : แทนประโยชน์ที่จะได้รับสำหรับการเปลี่ยนค่า output ของแต่ละ node

จากกฎการปรับค่าน้ำหนักของ Gradient Descent

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

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

ดังนั้น

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

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

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

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

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

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 สู่อัน Output

### การปรับค่าน้ำหนักในชั้น Hidden Layer

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

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

เมื่อ  $g'$  คือ derivative ของ transformation function จากชั้น Input สู่อัน Hidden Layer แต่ปัญหาที่พบก็คือ เราไม่สามารถหาค่า  $t_j$  ได้ เพราะว่าในชั้น 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

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

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

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

**การกำหนดจำนวนนิวรอนในชั้นซ่อนเร้น**

การค้นหาลักษณะการเชื่อมต่อที่ดีที่สุด ใช้การเปลี่ยนจำนวนนิวรอนในชั้นซ่อนเร้น ด้วยหลักการหาค่าต่ำสุดแบบ Quadratic interpolation ดังนี้

*Quadratic interpolation* เป็นการหาแบบจำลองที่ดีที่สุด ในการกำหนดจำนวนนิวรอนในชั้นซ่อนเร้น โดยมีวิธีการหาค่าต่ำสุดด้วยวิธี Quadratic Interpolation ซึ่งจะทำให้ได้เมื่อมีข้อมูลจำนวน 3 ชุดและมั่นใจว่าข้อมูลทั้งสามชุดอยู่บนฟังก์ชัน Quadratic เดียวกัน การมีข้อมูล 3 ชุดจะเพียงพอต่อการหา Unique Solution สำหรับฟังก์ชัน Quadratic ดังนี้

เมื่อมีข้อมูลดังต่อไปนี้

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

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

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

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

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

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

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

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

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

$$\text{แล้วจะได้ว่า } \psi = (\alpha - \beta)(\beta - \gamma)(\gamma - \alpha)$$

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  $x$  ที่ทำให้พบกับจุดต่ำสุดของฟังก์ชันสามารถหาได้จากสูตรดังนี้

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

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

### 3.2 วิธีการวิจัย

#### 3.2.1 วิธีการศึกษาด้วยแบบจำลอง ARIMA

การพยากรณ์ราคาอัตราแลกเปลี่ยนระหว่างเงินบาทกับดอลลาร์สหรัฐฯ โดยวิธี ARIMA เป็นการนำเอาข้อมูลอนุกรมเวลามารูปแบบแบบจำลองที่เหมาะสมที่สุด ซึ่งมีขั้นตอนดังต่อไปนี้

- 1) การทดสอบความนิ่งของข้อมูลเป็นการพิจารณาว่าข้อมูลอนุกรมเวลามีลักษณะนิ่งหรือไม่ โดยการทดสอบ unit root

2) การกำหนดรูปแบบจำลอง ARIMA (p,d,q) โดยการพิจารณาคอเรลโลแกรม Autocorrelation Function (ACF) และ Partial Autocorrelation Function (PACF) เพื่อจะสามารถระบุได้ว่าแบบจำลองควรมี Autoregressive (p) เท่าใด และ Moving average (q) เท่าใด โดยเลือกสร้างไว้หลายแบบจำลองเพื่อหาแบบจำลองที่เหมาะสมที่สุด

3) การประมาณค่าพารามิเตอร์ เพื่อนำค่าพารามิเตอร์ที่ได้นั้นนำไปทำการพยากรณ์ราคาต่อไปได้

4) การตรวจสอบความถูกต้อง เมื่อทำการหาแบบจำลองที่เหมาะสมและประมาณค่าพารามิเตอร์แล้ว จึงทำการทดสอบแบบจำลองโดยพิจารณาจากค่า Q-statistic จากคอเรลโลแกรม

5) การพยากรณ์

### 3.2.2 วิธีการศึกษาด้วยแบบจำลอง GARCH-M

ในขั้นนี้จะทำการพยากรณ์ราคาอัตราแลกเปลี่ยนระหว่างค่าเงินบาทกับเงินดอลลาร์สหรัฐฯ โดยใช้แบบจำลอง GARCH-M เป็นการศึกษาความสัมพันธ์ของราคาปิดของราคาซื้อขายอัตราแลกเปลี่ยนในปัจจุบันกับราคาปิดของราคาอัตราแลกเปลี่ยนในอดีต ซึ่งสามารถสรุปขั้นตอนการสร้างแบบจำลองได้ดังนี้ คือ

1) นำข้อมูลที่ผ่านมาการศึกษาด้วยแบบจำลอง ARIMA ข้อ 1 และ ข้อ 2 มาประมาณค่าพารามิเตอร์ โดยวิเคราะห์ตามวิธีแบบจำลอง GARCH-M เพื่อหารูปแบบที่เหมาะสมที่จะใช้ในการพยากรณ์

2) ตรวจสอบความถูกต้อง โดยพิจารณาจากค่า Q-Statistic จากคอเรลโลแกรม

3) การพยากรณ์

### 3.2.3 วิธีการศึกษาด้วยแบบจำลอง Neural Networks

การศึกษาด้วยแบบจำลอง Neural Networks ประกอบด้วย 3 การทดลอง ดังนี้

1. การทดลองเบื้องต้นด้วย Hidden layer จำนวน 1 ชั้น

2. การทดลองปรับเปลี่ยนปรับเปลี่ยนจำนวนนิวรอนในชั้น Hidden layer ด้วยวิธี

Quadratic Interpolation

3. การทดลองปรับเปลี่ยนจำนวนข้อมูลนำเข้า ด้วยวิธี Quadratic Interpolation

1) การทดลองเบื้องต้นด้วย Hidden layer จำนวน 1 ชั้น ประกอบด้วยขั้นตอน ดังต่อไปนี้

1.1) กำหนดข้อมูลนำเข้า (input) ในแบบจำลองที่ใช้ในการศึกษากำหนดให้มีจำนวน 10 ค่า ซึ่งเป็นราคาอัตราแลกเปลี่ยนเงินตราต่างประเทศรายวันย้อนหลัง 10 วัน

1.2) กำหนดนิวรอลในชั้น Hidden layer โดยเริ่มต้นที่ 10 และเพิ่มขึ้นทีละ 10 จนถึงจำนวนนิวรอลที่ 100 พร้อมกันนั้นก็กำหนดจำนวนรอบของการเรียนรู้ (Epochs) เริ่มต้นที่ 100 Epoch และเพิ่มขึ้นทีละ 100 Epoch จนถึง 1000 Epoch

1.3) จากข้อ 2 จะได้ค่า MSE ของ Validation ตรวจสอบว่าค่า MSE ต่ำสุด จะได้จำนวน hidden layer และจำนวนรอบของการเรียนรู้ที่เท่าไร

สูตร Mean Standard Error (MSE)

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

โดยที่  $y_i$  = ค่าพยากรณ์

$t_i$  = ค่าจริง

$n$  = จำนวนข้อมูลที่ใช้ในการพยากรณ์

1.4) นำจำนวนนิวรอลในชั้น hidden layer และจำนวนรอบของการเรียนรู้ (Epoch) ไปพยากรณ์

2) การทดลองปรับเปลี่ยนปรับเปลี่ยนจำนวนนิวรอลในชั้น Hidden layer ด้วยวิธี Arbitrary ประกอบด้วยขั้นตอน ดังต่อไปนี้

2.1) กำหนดข้อมูลนำเข้า (input) ให้มีจำนวน 10 ค่า

2.2) กำหนดนิวรอลในชั้น Hidden layer ด้วยวิธี Arbitrary โดยเริ่มต้นที่ 10 และเพิ่มขึ้นทีละ 5 จนกระทั่งถึงจำนวนนิวรอลที่ 100

2.3) หาจำนวนนิวรอลที่ให้ค่า MSE ต่ำที่สุด 3 ค่าที่ใกล้เคียงกัน

2.4) หาจำนวนนิวรอลที่เหมาะสมโดยวิธี Quadratic Interpolation

2.5) พยากรณ์

3) การทดลองปรับเปลี่ยนจำนวนข้อมูลนำเข้า ด้วยวิธี Quadratic Interpolation ประกอบด้วยขั้นตอน ดังต่อไปนี้

3.1) จากการทดลองที่ 2 ได้จำนวนนิวรอลในชั้น hidden layer ที่เหมาะสมที่สุดคือ 60 กำหนดนิวรอลในชั้น hidden layer คงไว้ที่ 60

3.2) เปลี่ยนจำนวนข้อมูลนำเข้า(input) จาก 10 เป็น 20, 30, 40, 50, ตามลำดับ

3.3) หาจำนวนข้อมูลนำเข้า(input) ที่ให้ค่า MSE ต่ำที่สุด 3 ค่าที่ใกล้เคียงกัน

3.4) หาจำนวนข้อมูลนำเข้า(input) ที่เหมาะสมโดยวิธี Quadratic Interpolation

การศึกษานี้จะทำการประเมินผลด้วย MAPE ( Mean Absolute Percentage Error) มีสูตรการคำนวณ ดังนี้



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

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

เมื่อ  $t$  คือ ค่าที่เกิดขึ้นจริง

$y$  คือ ค่าที่ได้จากการพยากรณ์

$n$  คือ จำนวนวันที่พยากรณ์



ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่  
Copyright© by Chiang Mai University  
All rights reserved