บทที่ 3

ทฤษฎีและแนวคิดที่ใช้ในการศึกษา

3.1 Artificial Neural Networks แบบ Feed Forward

การพยากรณ์โดยใช้แบบจำลอง Artificial Neural Networks หรือ โครงข่ายประสาทจำลองเป็นแบบจำลองที่ได้รับการพัฒนาโดยการเลียนแบบการทำงานของเซลล์สมองของมนุษย์ ซึ่งเชื่อว่าผลการตัดสินใจของแบบจำลองจะคล้ายคลึงการคิดและตัดสินใจของมนุษย์

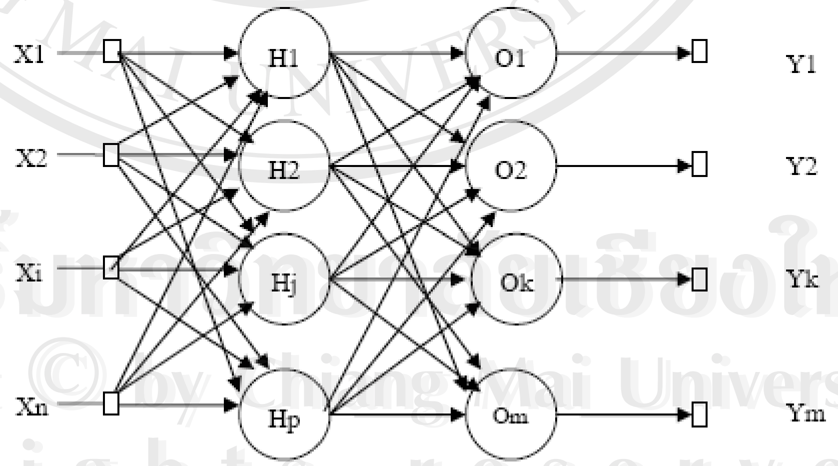
Neural Networks แบบ feed forward จะรับข้อมูลนำเข้า (input) แล้วทำการประมวลผลโดยนิวรอลใน hidden layer และ output layer จากนั้นจึงจะส่งผลลัพธ์ออกมาเป็นคำตอบ การที่สัญญาณนำเข้าถูกถ่ายทอดไปข้างหน้าเรื่อยๆ เรียกว่า feed forward และการที่มีจำนวนนิวรอลหลายชั้นเรียกว่า multi layer (คมสัน สุริยะ, 2548: 3)

องค์ประกอบของ Artificial Neural Networks แบบ feed forward มีดังต่อไปนี้

1) Input

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

รูปที่ 3.1 แบบจำลองของ Artificial Neural Networks แบบ feed forward



2) Output

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

3) นิวรอลใน Hidden Layer

นิวรอลใน hidden layer มี p ตัว คือ $H = \{H_1,H_2,\ldots,H_j,\ldots,H_p\}$

4) นิวรอลใน Output Layer

นิวรอลใน output layer มี m ตัว คือ O = $\{O_1, O_2, \ldots, O_k, \ldots, O_m\}$

5) ค่าน้ำหนักจาก Input Layer สู่ Hidden Layer

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

ดังนั้นเมื่อมีจำนวน input จำนวน n ตัว และมีนิวรอลใน hidden layer จำนวน p ตัวแล้ว จะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ (n+1)p ตัว

6) ค่าน้ำหนักจาก hidden layer สู่ Output Layer

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

ดังนั้นเมื่อมีจำนวนนิวรอลใน hidden layer จำนวน p ตัว และมี output จำนวน m ตัว แล้ว จะได้ว่าจำนวนค่าน้ำหนักในส่วนนี้มีจำนวนทั้งหมดเท่ากับ (p+1)m ตัว

7) การคำนวณแบบแพร่ย้อนกลับ (Back Propagation)

นิวรอลใน output layer แต่ละตัวจะรับสัญญาณจากนิวรอลใน hidden layer แล้วจึงผลิตสัญญาณออกไป การพิจารณาการปรับค่าน้ำหนักต้องเริ่มพิจารณาจากนิวรอลใน output layer ก่อน ดังนี้

จากกฎการปรับค่าน้ำหนักของ Gradient Descent

$$\Delta \mathrm{w} = \eta\left(\mathrm{t}-\mathrm{y}\right)\left[\nabla_{w} f\left(a\right)\right]$$

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

ดังนั้น

$$\nabla_{w} f\left(a\right) = \begin{pmatrix} \partial f\left(a\right)/\partial w_0 \\ \partial f\left(a\right)/\partial w_1 \\ \vdots \\ \partial f\left(a\right)/\partial w_n \end{pmatrix} = \frac{\partial f\left(a\right)}{\partial w} = \frac{\partial f\left(a\right)}{\partial a}\frac{\partial a}{\partial w}$$

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

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

จึงสรุปได้ว่า $\Delta w = \eta\left(t - y\right) f'x$

การปรับค่าน้ำหนักใน output layer

ค่าน้ำหนักที่เชื่อมจากนิวรอลใน hidden layer มายัง output layer คือ W_{jk}^{O} ซึ่งมีกฎการปรับค่าน้ำหนัก ดังนี้

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

แต่ input ของ output layer (x_{jk}^{O}) คือ output ของ hidden layer (y_{jk}^{H})

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

เมื่อ f' คือ derivartive ของ transformation function จาก hidden layer สู่ output layer

การปรับค่าน้ำหนักใน hidden layer

ค่าน้ำหนักแต่ละตัวที่เชื่อมโยงจาก input มายังนิวรอลใน hidden layer คือ W_{ij}^{H} ซึ่งมีกฎการปรับค่าน้ำหนัก ดังนี้

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

เมื่อ g' คือ derivartive ของ transformation function จากชั้น input สู่ hidden layer

แต่เราไม่สามารถที่จะบอกได้ว่าค่า y_j ที่นิวรอลแต่ละตัวสร้างขึ้นมานั้นถูกหรือผิด เนื่องจากเราไม่มีค่า t_j ทำให้ไม่สามารถหา error ได้ จึงได้มีการสร้าง error เทียมขึ้นมา โดยอาศัยหลักการว่านิวรอลตัวใดส่งข้อมูลไปยัง output layer มากก็ต้องรับผิดชอบในความคลาดเคลื่อนที่เกิดขึ้นในตอนท้ายที่สุดมากกว่า

ความคลาดเคลื่อน (error) ของ output แต่ละตัว สามารถแบ่งความรับผิดชอบไปยังนิวรอลใน hidden layer ทุกๆตัวได้โดยผ่านค่าน้ำหนัก ถ้านิวรอลใน hidden layer ใดมีค่าน้ำหนักมากย่อมต้องรับผิดชอบกับความคลาดเคลื่อนที่เกิดขึ้นมาก

ความคลาดเคลื่อนที่นิวรอลตัวที่ j ใน hidden layer ต้องรับผิดชอบสำหรับ output ตัวที่ 1 คือ $W_{j1}^{O}\left(t_1 - y_1\right) f'$ ความรับผิดชอบรวมทั้งหมดที่ นิวรอล ตัวที่ j ใน hidden layer จะต้องรับผิดชอบสำหรับ output ทุกตัว คือ $\sum_{k=1}^{m} W_{jk}^{O}\left(t_k - y_k\right) f'$

เมื่อนำความรับผิดชอบที่นิวรอลตัวดังกล่าวต้องรับผิดชอบต่อ output ทั้งหมดมาเป็น error เทียม จะทำให้สามารถแก้ปัญหาการปรับค่าน้ำหนักในชั้น hidden layer ได้ดังนี้

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

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

ดังนั้นจึงได้กฎการปรับน้ำหนักจาก input layer มายัง hidden layer ดังนี้

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

3.2 แบบจำลอง ARIMA, ARCH, GARCH และ EGARCH

แบบจำลอง ARIMA เป็นแบบจำลองที่ใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลา (time series data) ที่ได้รับการความนิยมและมีการพัฒนาปรับปรุงเป็น ARACH, GARCH และ EGARCH

3.2.1 การทดสอบความนิ่งของข้อมูล (stationary) และการทดสอบ Unit Root

ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง (stationary) คือข้อมูลที่ค่าเฉลี่ยและความแปรปรวนของกระบวนการเชิงสุ่มนั้นมีค่าคงที่เมื่อเวลาได้เปลี่ยนไปและค่าความแปรปรวนระหว่างสองคาบเวลาขึ้นอยู่กับความล้าหลัง (lag) ระหว่างคาบเวลาทั้งสอง (ทรงศักดิ์ ศรีบุญจิตต์ และ อารี วิบูลย์พงศ์, 2542. อ้างถึงใน สธนพล วิเชียรรัตนพันธ์, 2547: 27) โดยสามารถเขียนเป็นสมการทางคณิตศาสตร์ตามสมการ

ค่าเฉลี่ย (mean) : $\mathrm{E}\left(x_t\right)\mathrm{constant}=\mu$

ค่าความแปรปรวน (variance) : $\mathrm{V}\left(\mathrm{x_t}\right)\mathrm{constant}=\sigma^2$

ความแปรปรวนร่วม(covariance): $\mathrm{cov}\left(x_t, x_{t+k}\right) = \mathrm{E}\left(x_t - \mu\right)\left(x_{t+k} - \mu\right) = \sigma_k - \mu$

โดยที่ $\mathrm{x_t}$ แทนข้อมูลอนุกรมเวลาซึ่งเป็นกระบวนการเชิงสุ่ม

ในการวิเคราะห์ข้อมูลที่เป็นอนุกรมเวลานั้น ข้อมูลจะต้องมีลักษณะนิ่ง เนื่องด้วยข้อมูลอนุกรมเวลานั้นมาจากกระบวนการเชิงสุ่ม (random process) การนำข้อมูลอนุกรมเวลาไปใช้โดยไม่ได้ทำการตรวจสอบว่าข้อมูลนั้นมีลักษณะนิ่งนั้น ถ้าข้อมูลอนุกรมเวลานั้นไม่นิ่งจะทำให้ค่าสถิติที่เกิดขึ้นมีการแจกแจงไม่มาตรฐาน (nonstandard distributions) ซึ่งทำให้การนำไปใช้เปรียบเทียบกับค่าในตารางมาตรฐาน (standard tables) ไม่ถูกต้องเนื่องจากค่าต่างๆนั้นมีสมมติฐานว่าข้อมูลนั้นมีการแจกแจงมาตรฐาน (standard distributions) ทำให้นำไปสู่การลงความเห็นที่ผิดพลาดและความสัมพันธ์ที่ไม่แท้จริง (spurious regression) กล่าวคือค่า R^2 มีค่าสูงมากและได้ค่าสถิติ t มีนัยสำคัญหรือสูงเกินกว่าความเป็นจริง (ทรงศักดิ์ ศรีบุญจิตต์ และ อารี วิบูลย์พงศ์, 2542. อ้างถึงใน สธนพล วิเชียรรัตนพันธ์, 2547: 28)

ในการใช้ข้อมูลอนุกรมเวลาจึงต้องทำการทดสอบว่าข้อมูลที่นำมาใช้มีลักษณะนิ่งหรือไม่ ซึ่งจะใช้การทดสอบ unit root โดยสามารถทดสอบได้โดยใช้การทดสอบ DF (Dicky-Fuller (DF) test) (Dickey and Fuller, 1981. อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2547: 476) และการทดสอบ ADF (Augmented Dicky-Fuller (ADF) test) (Said and Dickey, 1984. อ้างถึงใน ทรงศักดิ์ ศรีบุญจิตต์, 2547: 476)

การทดสอบ DF (DF test) มีสมมติฐานหลักคือ $\mathrm{H}_0 : \rho = 1$ และสมมติฐานรอง $\mathrm{H}_1 : |\rho| < 1$ จากสมการ $x_t = \rho x_{t-1} + \varepsilon_t$

ถ้ายอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่ง แต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง และการทดสอบนี้ยังสามารถแปลงสมการได้ดังนี้

กรณีไม่มีค่าคงที่และแนวโน้มเวลา $\Delta x_t = \theta x_{t-1} + \varepsilon_t$ (1)

กรณีมีเฉพาะค่าคงที่ $\Delta x_t = \alpha + \theta x_{t-1} + \varepsilon_t$ (2)

กรณีมีทั้งค่าคงที่และแนวโน้มเวลา $\Delta x_t = \alpha + \beta_t + \theta x_{t-1} + \varepsilon_t$ (3)

โดยกำหนดสมมติฐานหลัก $\mathrm{H}_0 : \theta = 0$

และสมมิติฐานรอง $\mathrm{H}_1 : \theta < 0$

การยอมรับ H_0 แสดงว่าข้อมูลมีลักษณะไม่นิ่ง แต่ถ้าปฏิเสธ H_0 แสดงว่าข้อมูลนั้นมีลักษณะนิ่ง นอกจากนี้ถ้าสมการ (1) , (2) และ (3) นำไปเข้ากระบวนการอัตถดถอย (autoregressive process) จะได้ดังนี้

กรณีไม่มีค่าคงที่และแนวโน้มเวลา $\Delta x_t = \theta x_{t-1} + \sum_{i=1}^{\rho} \Phi_i \Delta x_{i-1} + \varepsilon_t$

กรณีมีเฉพาะค่าคงที่ $\Delta x_t = \alpha + \theta x_{t-1} + \sum_{i=1}^{\rho} \Phi_i \Delta x_{i-1} + \varepsilon_t$

กรณีมีทั้งค่าคงที่และแนวโน้มเวลา $\Delta x_t = \alpha + \beta_t + \theta x_{t-1} + \sum_{i=1}^{\rho} \Phi_i \Delta x_{i-1} + \varepsilon_t$

แต่ละสมการเป็นการทดสอบ Augmented Dicky-Fuller Test (ADF) นั่นเอง ในการตรวจสอบว่าข้อมูลมีลักษณะนิ่งหรือไม่นิ่ง โดยการเปรียบเทียบค่าสถิติ t ที่คำนวณได้กับค่าวิกฤต (critical value) ในตาราง ADF

3.2.2 แบบจำลอง ARIMA

ในแบบจำลอง Autoregressive / Integrated / Moving Average (ARIMA) มีการใช้สัญลักษณ์ ARIMA(p,d,q) ซึ่งหมายถึง ข้อมูลดังกล่าวมีลักษณะ AR(p) และ MA(q) โดยต้องหาผลต่าง d ครั้งข้อมูลดังกล่าวจึงจะมีลักษณะนิ่ง

ความนิ่งและความไม่นิ่ง (stationary and nonstationary)

เมื่อพบว่าข้อมูลมีลักษณะไม่นิ่งต้องหาผลต่าง (difference) ของข้อมูลก่อน เพื่อให้ได้ข้อมูลอนุกรมเวลาที่มีลักษณะนิ่ง ซึ่งจะใช้ backward shift operator, B เป็นการแสดงการถอยหลังข้อมูลไปหนึ่งคาบเวลา ดังนี้

$$\mathrm{BX}_t = \mathrm{X}_{t-1}$$

ถ้าถอยหลังไปสองคาบเวลา ได้ว่า

$$B\left(BX_t\right) = B^2X_t = X_{t-2}$$

ผลต่างที่หนึ่ง (first difference)

$$X_t' = X_t - X_{t-1} = X_t - BX_t = \left(1-B\right)X_t$$

ผลต่างอันดับที่สอง (second-order difference)

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

และ $\left(1-B\right)^d X_t$ คือผลต่างอันดับที่ d

กระบวนการอัตถดถอย (Autoregressive Processes)

กระบวนการ AR(p) คือ กระบวนการ AR ที่มีอันดับที่ p สามารถเขียน ARIMA(p,0,0) ได้ดังนี้คือ

$$X_t = \mu' + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_p X_{t-p} + e_t$$

โดยที่ μ' = พจน์คงที่หรือคงตัว (constant term)

Φ_j = พารามิเตอร์อัตถดถอยตัวที่ j

e_t = พจน์ความคลาดเคลื่อน ณ เวลา t

หรือเขียนได้เป็น

$$\left(1-\Phi_1 B - \Phi_2 B^2 - \ldots - \Phi_p B^p\right) = \mu' + e_t$$

กระบวนการเฉลี่ยเคลื่อนที่ (Moving Average Processes)

กระบวนการ MA(q) คือ กระบวนการ MA ที่มีอันดับที่ q สามารถเขียน ARIMA(0,0,q) ได้ดังนี้คือ

$$X_t = \mu + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_p e_{t-p}$$

โดยที่ μ' = พจน์คงที่หรือคงตัว (constant term)

θ_j = พารามิเตอร์เฉลี่ยเคลื่อนที่ตัวที่ j

e_t = พจน์ความคลาดเคลื่อน ณ เวลา t

หรือเขียนได้เป็น

$$X_t = \mu + \left(1-\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q\right)e_t$$

3.2.3 แบบจำลอง Autoregressive Conditional Heteroscedasticity (ARCH)

ในการวิเคราะห์อนุกรมเวลาส่วนใหญ่แล้วจะมีการกำหนด stochastic variable ให้มีความแปรปรวนคงที่ (homoscedastic) ซึ่งในการประยุกต์ใช้กับบางข้อมูลนั้นค่าความแปรปรวนของค่าความคลาดเคลื่อน (error term) จะไม่ใช่ฟังก์ชั่นของตัวแปรอิสระแต่มีค่าเปลี่ยนแปลงไปตามช่วงเวลาขึ้นอยู่กับขนาดของความคลาดเคลื่อนที่เกิดขึ้นในอดีต และในบางงานศึกษา เช่น แบบจำลองของเงินเฟ้อ อัตราดอกเบี้ยหรือผลตอบแทนจากตลาดหลักทรัพย์ ในบางคาบเวลาจะมีค่าความผันผวน (volatility) สูง (และความคลาดเคลื่อนขนาดใหญ่) ตามด้วยคาบเวลาที่มีค่าความผันผวน (volatility) ต่ำ (และมีค่าความคลาดเคลื่อนที่มีขนาดเล็ก) สรุปได้ว่าค่าความแปรปรวนของค่าความคลาดเคลื่อนจากการถดถอยจะขึ้นอยู่กับค่าความผันผวน (volatility) ของความคลาดเคลื่อนในอดีตที่ผ่านมา (Enders,1995. อ้างถึงใน สธนพล วิเชียรรัตนพันธ์, 2547: 29)

ความเป็นไปได้ในการหาค่าเฉลี่ยและความแปรปรวนของอนุกรมเวลาไปพร้อมกันนั้นในขั้นต้นการพยากรณ์อย่างมีเงื่อนไขจะมีความแม่นยำเหนือกว่าการพยากรณ์อย่างไม่มีเงื่อนไขมาก ซึ่งจากแบบจำลอง Autoregressive Moving Average (ARMA) ซึ่งแสดงดังสมการ

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

และการพยากรณ์อย่างมีเงื่อนไขของ X_{t+1} ดังนี้คือ

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

ถ้าเราใช้ค่าเฉลี่ยอย่างมีเงื่อนไขในการพยากรณ์ X_{t+1} ค่าความคลาดเคลื่อนของความแปรปรวนอย่างมีเงื่อนไขที่พยากรณ์ได้ดังนี้คือ

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

ถ้าเปลี่ยนไปใช้การพยากรณ์อย่างไม่มีเงื่อนไขแล้ว ผลที่จะใช้เป็นค่าเฉลี่ยในช่วง Long-Run ของลำดับ $\{X_t\}$ ซึ่งเท่ากับ $\dfrac{a_0}{(1-a_1)}$ จะได้ค่าความคลาดเคลื่อนของการพยากรณ์อย่างไม่มีเงื่อนไขตามสมการ ดังนี้

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

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

เนื่องจาก $\dfrac{1}{(1-a_1)^2} > 1$ เพราะฉะนั้นความแปรปรวน (variance) จากการพยากรณ์แบบไม่มีเงื่อนไข (unconditional variance) จึงมีค่าสูงกว่าความแปรปรวน ของการพยากรณ์แบบมี

เงื่อนไข ในลักษณะเดียวกันถ้าความแปรปรวนของ $\{\varepsilon_t\}$ ไม่เป็นค่าคงที่ จะสามารถประมาณค่าแนวโน้มของการเปลี่ยนแปลงความแปรปรวนโดยใช้ ARMA Model ความแปรปรวนแบบมีเงื่อนไขของ X_{t+1} คือ

$$\mathrm{Var}\left(X_{t+1}|X_t\right)=E\left[\left(X_{t+1}-a_0-a_1X_t\right)^2\right]$$
$$=E_t\varepsilon_{t+1}^2$$

และจากที่ให้ $E_t\varepsilon_{t+1}^2=\sigma_{t+1}^2$ จึงแสดงว่าความแปรปรวนอย่างมีเงื่อนไขไม่ใช่ค่าคงที่และจะได้แบบจำลองในการประมาณค่าส่วนที่เหลือ (residuals) ออกมาดังสมการ

$$\hat{\varepsilon}_t^2=\alpha_0+\alpha_1\hat{\varepsilon}_{t-1}^2+\ldots+\alpha_q\hat{\varepsilon}_{t-q}^2+v_t \qquad (4)$$

เมื่อ v_t = white noise process

ถ้าค่าของ $\alpha_1,\alpha_2,\ldots,\alpha_q$ มีค่าเท่ากับศูนย์ ความแปรปรวนที่ประมาณค่ามาได้ (estimated variance) จะมีค่าคงที่หรือคงตัว (constant variance) α_0 อีกนัยหนึ่ง คือ ค่าความแปรปรวนอย่างมีเงื่อนไขของ Xt จะมีการเปลี่ยนแปลงสอดคล้องกับในสมการ (4) และค่าพยากรณ์สามารถเขียนได้ดังนี้

$$E\hat{\varepsilon}_{t+1}^2=\alpha_0+\alpha_1\hat{\varepsilon}_t^2+\alpha_2\hat{\varepsilon}_{t-1}^2+\ldots+\alpha_q\hat{\varepsilon}_{t+1-q}^2 \qquad (5)$$

สมการที่มีลักษณะเช่นสมการ (4) เราเรียกว่า an autoregressive conditional heteroscedastic (ARCH) model และสมการ (5) เป็น ARCH(q) โดยค่า $E_t\hat{\varepsilon}_{t+1}^2$ หรือ σ_{t+1}^2 จะประกอบด้วย 2 องค์ประกอบ คือค่าคงที่และความผันผวน (volatility) ในคาบเวลาที่ผ่านมา ซึ่งเขียนได้เป็นส่วนเหลือกำลังสองของคาบในอดีต (ARCH term) ส่วนค่าสัมประสิทธิ์ $\left(\alpha_0,\alpha_1,\ldots,\alpha_q\right)$ สามารถหาค่าได้โดยใช้วิธี Maximum Likelihood

3.2.4 แบบจำลอง Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

แบบจำลอง ARCH ของ Engle ได้มีการพัฒนาต่อโดย Bollerslev(1986. อ้างถึงใน สธนพล วิเชียรรัตนพันธ์, 2547: 34) ด้วยการให้ความแปรปรวนแบบมีเงื่อนไข (conditional varience) มีลักษณะเป็น ARMA process โดยที่ให้ error process มีลักษณะดังนี้

$$\varepsilon_t=v_t\sqrt{c_t}$$

โดยที่ความแปรปรวนของ $v_t=\sigma_v^2=1$

และ
$$c_t=\alpha_0+\sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2+\sum_{i=1}^{p}\beta_i c_{t-i} \qquad (6)$$

เมื่อ $\{v_t\}$ คือ white noise process ที่เป็นค่าอิสระจากเหตุการณ์ในอดีต $\left(\varepsilon_{t-1}\right)$ ค่าเฉลี่ยเคลื่อนที่อย่างมีเงื่อนไขและไม่มีเงื่อนไขของ ε_t จะมาจาก c_t ในสมการ (6)

GARCH(p,q) นั้นใช้กระบวนการ Autoregressive และ Moving Average ในการหา Heteroscedastic Variance ได้ดังนี้

$$E_{t-1}\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i c_{t-i}$$

ถ้ากำหนดให้ค่า p=0 และ q=1 จะได้เป็น ARCH(1) หรือถ้าค่า β_i ทั้งหมดมีค่าเป็นศูนย์ แบบจำลอง GARCH(p,q) จะเทียบเท่ากับแบบจำลอง ARCH(q) คุณสมบัติที่สำคัญของแบบจำลอง GARCH คือค่าความแปรปรวนอย่างมีเงื่อนไขของ disturbances ของค่า Xt สร้างขึ้นมาจากกระบวนการ ARMA จึงสามารถคาดได้ว่าส่วนเหลือจากการทำ ARMA จะแสดงถึงรูปแบบคุณลักษณะเดียวกัน เช่น ถ้าการประมาณค่า {Xt} ด้วยกระบวนการ ARMA ค่า Autocorrelation Function (ACF) ซึ่งเป็นสหสัมพันธ์ระหว่างตัวแปรสุ่มที่หน่วยเวลาห่างกันของกระบวนการเดียวกันและ Partial Autocorrelation Function (PACF) ของส่วนที่เหลือ (residuals) ควรจะบ่งถึงกระบวนการ white noise และ ACF ของกำลังสองของส่วนเหลือ (squared residuals) นำมาช่วยในการระบุถึงลำดับ (order) ของกระบวนการ GARCH

3.2.5 แบบจำลอง ARCH in mean (ARCH-M)

พื้นฐานจากแนวคิด ARCH ได้มีการขยายเพื่อให้ค่าเฉลี่ยของลำดับอนุกรมขึ้นอยู่กับความแปรปรวนอย่างมีเงื่อนไขของตัวเอง ซึ่งเรียกว่า ARCH-M เพื่อที่จะมีความเหมาะสมในการศึกษา โดยมีหลักการพื้นฐาน คือผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยง (risk averter) จะต้องการชดเชยสำหรับการถือสินทรัพย์ที่มีความเสี่ยง หรือ risk premium เมื่อให้ความเสี่ยงในทรัพย์สินสามารถวัดค่าได้จากความแปรปรวนของผลตอบแทน ค่าชดเชยความเสี่ยงจะเป็นฟังก์ชั่นเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของผลตอบแทน Engle(1987. อ้างถึงใน สธนพล วิเชียรรัตนพันธ์, 2547: 32) ได้แสดงแนวคิดนี้โดยนำส่วนที่เหลือ (residual) ของผลตอบแทนจากการถือครองทรัพย์สินที่มีความเสี่ยงดังสมการ

$$x_t = \mu_t + \varepsilon_t \qquad (7)$$

เมื่อ x_t = ผลตอบแทนส่วนเกินจากการถือครองทรัพย์สินในระยะยาวเปรียบเทียบกับหนึ่งคาบเวลาของผลตอบแทนอ้างอิง

μ_t = ค่าชดเชยความเสี่ยงที่จำเป็นในการโน้มน้าวผู้ที่มีลักษณะหลีกเลี่ยงความเสี่ยงให้ถือครองทรัพย์สินในระยะยาวแทนที่จะเป็นแค่คาบเวลาเดียว

ε_t = unforecastable shock ของส่วนที่เหลือ (residual) ของผลตอบแทนในการถือสินทรัพย์ในระยะยาว

สมการ (7) จะให้ค่าคาดหวังของส่วนที่เหลือ (residual) ของผลตอบแทนจากการถือครองทรัพย์สินในระยะยาวมีค่าเท่ากับค่าชดเชยความเสี่ยงดังสมการ (8)

$$E_{t-1}x_t=\mu_t \tag{8}$$

สมมติว่าค่าชดเชยความเสี่ยงเป็นฟังก์ชันเพิ่มของความแปรปรวนอย่างมีเงื่อนไขของ ε_t อีกนัยหนึ่ง คือยิ่งค่าความแปรปรวนอย่างมีเงื่อนไขของผลตอบแทนมีขนาดใหญ่ ผลตอบแทนของการโน้มน้าวให้คนหันมาถือครองทรัพย์สินในระยะยาวก็จะยิ่งมีค่ามากขึ้นตาม ถ้า c_t เป็นความแปรปรวนอย่างมีเงื่อนไขของ ε_t ค่าชดเชยความเสี่ยงสามารถแสดงได้เป็น

$$\mu_t=\beta+\delta c_t,\delta>0 \tag{9}$$

เมื่อ c_t คือ ARCH(q) process

$$c_t=\alpha_0+\sum_{i=0}^{q}\alpha_i\varepsilon_{t-1}^2 \tag{10}$$

สมการ (8) , (9) และ (10) ประกอบขึ้นเป็นพื้นฐานของแบบจำลอง ARCH-M จากสมการ (8) และ (9) ค่าเฉลี่ยอย่างมีเงื่อนไขของ x_t จะขึ้นอยู่กับค่าความแปรปรวนอย่างมีเงื่อนไข c_t จากสมการ (10) ความแปรปรวนอย่างมีเงื่อนไขได้จากกระบวนการ ARCH(q) ซึ่งเป็นการอธิบายว่าถ้าความแปรปรวนอย่างมีเงื่อนไขเป็นค่าคงที่ (เช่น ถ้า $\alpha_1=\alpha_2=\ldots=\alpha_q=0$) แบบจำลอง ARCH-M จะกลับไปสู่กรณีเดิมของค่าชดเชยความเสี่ยงคงที่

3.2.6 แบบจำลอง GRACH-in-mean (GARCH-M)

จากแบบจำลอง GARCH (p,q) ในสมการ (6) ความแปรปรวนอย่างมีเงื่อนไขของข้อมูลอนุกรมเวลาขึ้นอยู่กับส่วนที่เหลือกำลังสอง (squared residual) ของกระบวนการนี้ (Bollerslev,1986. อ้างถึงใน สธนพล วิเชียรรัตนพันธ์, 2547: 34)

Engle (1987. อ้างถึงใน สธนพล วิเชียรรัตนพันธ์, 2547: 34) ขยายแนวคิดนี้โดยให้ค่าเฉลี่ยอย่างมีเงื่อนไขเป็นฟังก์ชันของความแปรปรวนอย่างมีเงื่อนไขโดยรู้จักกันในชื่อ GARCH-in-mean หรือ GARCH-M ผลตอบแทนจากหลักทรัพย์สามารถแสดงในรูปของแบบจำลอง GARCH (p,q)-M ดังสมการ (11) , (12) และ (13)

$$x_t=\mu_t+\delta_1 c_t^{\frac{1}{2}}+\varepsilon_t \tag{11}$$

$$\varepsilon_t/\psi_{t-1}\sim N(0,c_t) \tag{12}$$

$$c_t=v+\sum_{i=1}^{p}\beta_i c_{t-1}+\sum_{i=1}^{q}\alpha_i(\varepsilon_{t-1})^2 \tag{13}$$

เมื่อ x_t คือผลตอบแทนจากหลักทรัพย์

μ_t คือค่าเฉลี่ย x_t อย่างมีเงื่อนไขในอดีต (ψ_{t-1}) และตามสมการข้อจำกัด $\omega>0, \alpha_i>0$ และ $\beta_i \geq 0$ เพื่อให้แน่ใจว่าค่าความแปรปรวนอย่างมีเงื่อนไข (c_t) นั้นเป็นบวก

$c_t^{\frac{1}{2}}$ ในสมการ(11) เพื่อใช้แสดงความสัมพันธ์โดยตรงถึง trade off ระหว่างความเสี่ยงและผลตอบแทนที่คาดหวัง

อิทธิพลอย่างมีนัยสำคัญของความผันผวนในผลตอบแทนจากหลักทรัพย์ถูกตรวจจับด้วยสัมประสิทธิ์ $c_t^{\frac{1}{2}}$ $(\delta 1)$ ในสมการ (11) ซึ่งอธิบายแทนดัชนีของความสัมพันธ์ของการหลีกเลี่ยงความเสี่ยง ค่าสัมประสิทธิ์ $\delta 1$ ที่เป็นบวกอย่างมีนัยสำคัญบอกถึงผู้ลงทุนในตลาดหลักทรัพย์เมื่อเผชิญกับระดับความเสี่ยงที่สูงขึ้นก็ต้องการค่าชดเชยความเสี่ยงที่มากขึ้นตามไปด้วย

3.2.7 แบบจำลอง EGARCH

EGARCH หรือ Exponential GARCH model ถูกเสนอโดย Nelson (1991) EGARCH model ให้ความแปรปรวนอย่างมีเงื่อนไขมีค่าดังสมการ (14)

$$\log\left(\sigma_t^2\right) = \omega + \beta \log\left(\sigma_{t-1}^2\right) + \alpha \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \qquad (14)$$

ข้อแตกต่างของ EGARCH model ระหว่างโปรแกรม EViews กับ โมเดลเดิมของ Nelson มีอยู่ 2 ข้อคือ ข้อแรก Nelson ได้ตั้งสมมติฐานว่า error มีการแจกแจงแบบ general distribution ส่วน EView ได้ตั้งสมมติฐานว่า error มีการแจกแจงแบบ normal distribution ข้อที่สอง ค่า log ของความแปรปรวนอย่างมีเงื่อนไขของ Nelson แตกต่างกับของ EView เล็กน้อยดังสมการ (15)

$$\log\left(\sigma_t^2\right) = \omega + \beta \log\left(\sigma_{t-1}^2\right) + \alpha \left(\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| - \sqrt{\frac{2}{\pi}}\right) + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \qquad (15)$$

การประมาณค่าสมการภายใต้สมมติฐานที่ error มีการแจกแจงแบบปกติจะให้ค่าที่เหมือนกันยกเว้นค่า intercept term ω ที่แตกต่างกันเท่ากับ $\alpha\sqrt{\frac{2}{\pi}}$

และการประมาณค่า EGACH model โดยโปรแกรม EViews ได้ดังสมการ (16)

$$\log\left(\sigma_t^2\right) = \omega + \sum_{j=1}^{p} \beta_j \log\left(\sigma_{t-j}^2\right) + \sum_{i=1}^{q} \left(\alpha_i \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma_i \frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right) \qquad (16)$$