บทที่ 4

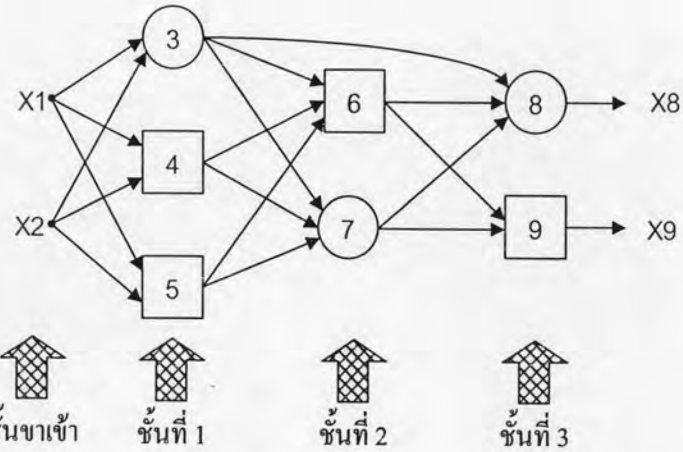
ระบบอนุมานนิวโรฟัซซีแบบปรับตัวได้

ระบบอนุมานฟัซซี (Fuzzy Inference System,FIS) เป็นระบบอนุมานค่าขาออกโดยอาศัยแบบจำลองฟัซซีของขาเข้าและขาออกซึ่งประกอบด้วยฟัซซีเซตที่แทนปริมาณในรูปแบบค่าภาษา ด้วยลักษณะของระบบที่สามารถกำหนดขาเข้าด้วยค่าภาษานี้เองทำให้ผู้ที่ใช้งานระบบอนุมานฟัซซีสามารถสร้างระบบอนุมานได้ด้วยความเข้าใจของมนุษย์

ในเวลาต่อมาจึงได้มีการประยุกต์ใช้ระบบอนุมานฟัซซีไปเป็นระบบที่สามารถปรับพารามิเตอร์ภายในได้หลังจากมีการฝึกสอนระบบด้วยข้อมูลตัวอย่างที่ได้จากระบบจริง เรียกระบบนี้ว่าระบบอนุมานนิวโรฟัซซีแบบปรับตัวได้ หรือ ANFIS (Adaptive Neuro-Fuzzy Inference System) [17] ทั้งนี้ ANFIS จัดเป็นโครงข่ายปรับตัวได้ [20] ชนิดหนึ่งที่มีพื้นฐานมาจากระบบอนุมานฟัซซีแบบ Sugeno ซึ่งพารามิเตอร์ของ ANFIS ที่เปลี่ยนไปเมื่อผ่านการฝึกสอนจะมีอยู่ 2 ส่วนคือ พรีมิสพารามิเตอร์ (premise parameter) และคอนซีเควนท์พารามิเตอร์ (consequence parameter) ที่อยู่ในฟัซซีเซตขาเข้าและอยู่ในสมการโพลิโนเมียลขาออกของกฎตามลำดับ ทั้งนี้หลักการพื้นฐานและวิธีการฝึกสอน ANFIS จะรวบรวมไว้ในเนื้อหาบทนี้

4.1 โครงข่ายปรับตัวได้ (adaptive networks)

โครงข่ายปรับตัวได้ คือโครงข่ายที่มีโครงสร้างประกอบด้วยโหนดจำนวนหนึ่งเชื่อมโยงกัน โดยแต่ละโหนดจะเป็นหน่วยประมวลผลซึ่งจะรับขาเข้าเพื่อคำนวณขาออกส่งไปยังโหนดที่อยู่ในชั้นถัดไป การประมวลผลในโหนดแต่ละโหนดจะเป็นไปตามฟังก์ชันทางคณิตศาสตร์และพารามิเตอร์ที่เก็บไว้ในโหนดนั้น ๆ



รูปที่ 4.1 ตัวอย่างของโครงข่ายปรับตัวได้

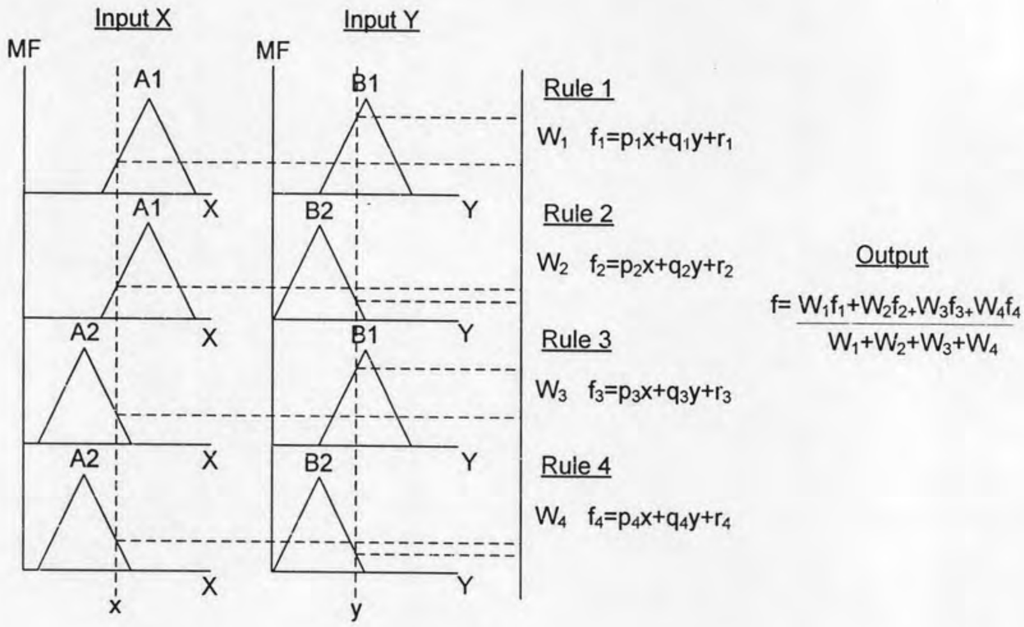
ค่าพารามิเตอร์ภายในโหนดสามารถปรับตัวได้โดยการฝึกสอนโครงข่ายปรับตัวได้ด้วยข้อมูลตัวอย่าง ซึ่งการฝึกสอนก็คือการปรับพารามิเตอร์ของโครงข่ายปรับตัวได้เพื่อให้ขาออกที่ได้ใกล้เคียงกับขาออกตัวอย่างมากที่สุด

ประโยชน์ของโครงข่ายปรับตัวได้คือสามารถใช้ในการคำนวณขาออกของระบบที่มีความสลับซับซ้อน โครงข่ายปรับตัวได้จะปรับพารามิเตอร์ภายในเพื่อให้ได้ขาออกเหมือนกับขาออกของระบบจริงโดยไม่ต้องสร้างแบบจำลองทางคณิตศาสตร์ที่ยุ่งยาก และการคำนวณก็ใช้เวลาน้อยเนื่องจากกระบวนการภายในโหนดเป็นการแทนค่าขาเข้าลงในสมการ โครงข่ายปรับตัวได้ที่เป็นที่รู้จักและมีการนำไปใช้ประยุกต์ใช้อย่างหลากหลายคือ โครงข่ายประสาทเทียม (Neural Networks, NNs) [13] อย่างไรก็ตามการใช้งานโครงข่ายประสาทเทียมมีความยุ่งยากในการเลือกจำนวนชั้นภายในและการเลือกจำนวนโหนดในแต่ละชั้นเนื่องจากไม่มีหลักการที่แน่นอนในการเลือกใช้ [15] ต่อมาจึงเกิดโครงข่ายปรับตัวได้ชนิดใหม่ขึ้นโดยอาศัยหลักการของระบบอนุมานฟัซซี นั่นก็คือ ANFIS โดย ANFIS มีฟังก์ชันภายในโหนดที่ทำหน้าที่แตกต่างกันชัดเจน และสามารถจำลองระบบที่มีความซับซ้อนมากขึ้นได้โดยง่ายด้วยการเพิ่มจำนวนฟัซซีเซตของขาเข้าให้มากขึ้น ทำให้มีการประยุกต์ใช้ ANFIS ในงานวิศวกรรมอย่างแพร่หลายในเวลาต่อมา [14]

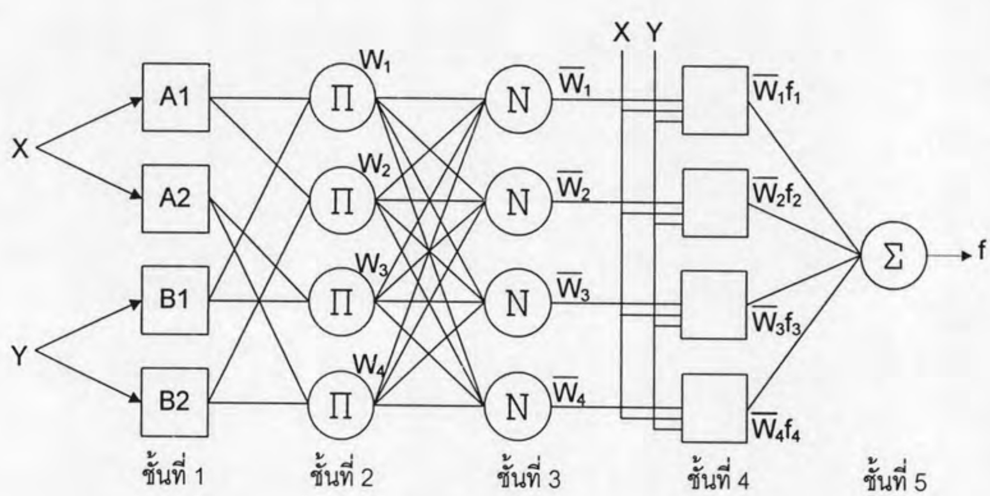
4.2 โครงสร้างของ ANFIS

ANFIS เป็นระบบที่มีพื้นฐานมาจากระบบอนุมานฟัซซีแบบ Sugeno และจัดเป็นโครงข่ายปรับตัวได้ชนิดหนึ่ง โดย ANFIS สามารถปรับค่าพารามิเตอร์ได้ด้วยการนำข้อมูลขาเข้าขาและขาออกตัวอย่างมาฝึกสอนระบบ ANFIS ทำให้สามารถแปลงขาเข้าให้เป็นขาออกที่สอดคล้องกับ

ตัวอย่างที่ ดังนั้นจึงสามารถใช้ ANFIS ในการจำลองการคำนวณของระบบที่มีความสลับซับซ้อนได้



รูปที่ 4.2 ระบบอนุมานฟัซซีแบบ Sugeno ที่มีจำนวนกฎ 4 กฎ



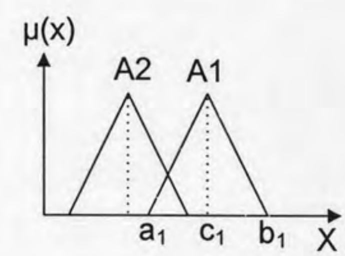
รูปที่ 4.3 ระบบอนุมานนิวโรฟัซซีแบบปรับตัวได้ (ANFIS)

โครงสร้าง ANFIS เป็นการปรับปรุงระบบอนุมานฟัซซีแบบ Sugeno ให้มาอยู่ในรูปแบบโครงข่ายปรับตัวได้ ซึ่งแบ่งเป็นชั้นต่าง ๆ ในแต่ละชั้นจะประกอบด้วยหน่วยการคำนวณที่เรียกว่าโหนด ซึ่งแต่ละโหนดในชั้นนั้นจะบรรจุฟังก์ชันการคำนวณเพื่อส่งขาออกจากโหนดนั้นไปเป็นขา

เข้าของโหนดในชั้นถัดไป จากระบบอนุมานฟัซซีแบบ Sugeno ในรูปที่ 4.2 สามารถแสดงในรูปโครงสร้าง ANFIS ได้ดังรูปที่ 4.3 ซึ่งเป็นตัวอย่าง ANFIS ที่มีขาเข้า 2 ตัว โดยในแต่ละชั้นของ ANFIS มีการคำนวณดังนี้

<u>ชั้นที่1</u>

โหนดในชั้นที่ 1 จะเก็บฟัซซีเซตในรูปแบบฟังก์ชันความเป็นสมาชิกของฟัซซีเซต ซึ่งรูปร่างของฟัซซีเซต 1 เซตจะกำหนดโดยพารามิเตอร์ ซึ่งเรียกว่าพรีมิสพารามิเตอร์ ตัวอย่างเช่นรูปด้านล่างเป็นฟัซซีเซตที่มีฟังก์ชันความเป็นสมาชิกเป็นรูปสามเหลี่ยม ซึ่งฟัซซีเซต 1 เซตจะมีพารามิเตอร์จำนวน 3 ค่า โดย ฟัซซีเซต A1 จะมีพารามิเตอร์คือ a_1 b_1 และ c_1



รูปที่ 4.4 ฟัซซีเซตที่มีฟังก์ชันความเป็นสมาชิกรูปสามเหลี่ยม

จำนวนพรีมิสพารามิเตอร์ที่อยู่ในชั้นที่ 1 จะขึ้นอยู่กับจำนวนขาเข้า (N_{INPUT}),จำนวนพารามิเตอร์ของฟังก์ชันความเป็นสมาชิก (N_f) และการกำหนดจำนวนฟัซซีเซตในขาเข้าแต่ละตัว (m_k) จำนวนมิสพารามิเตอร์ทั้งหมด (N_{PRE}) สามารถคำนวณได้ดังนี้

$$N_{PRE} = N_f \sum_{k=1}^{N_{INPUT}} m_k \tag{4.1}$$

และสามารถคำนวณจำนวณกฎ (N_{RULE}) ได้ดังนี้

$$N_{RULE} = \prod_{k=1}^{N_{INPUT}} m_k \tag{4.2}$$

ค่าขาเข้าจากภายนอกที่เข้ามายังชั้นนี้จะแปลงไปเป็นค่าฟัซซี ซึ่งขาออกที่ออกจากชั้นนี้จะเป็นค่าความเป็นสมาชิกของฟัซซีเซตที่อยู่ในโหนดนั้น จากรูปที่ 4.3 ขาออกของโหนดแต่ละโหนดในชั้นนี้แสดงได้ดังนี้

$$
\begin{aligned}
O_{1,i} &= \mu_{Ai}(x) \qquad ;i=1,2,or \\
O_{1,i} &= \mu_{Bi-2}(y) \qquad ;i=3,4
\end{aligned}
\tag{4.3}
$$

โดย

$O_{1,i}$ คือ ขาออกของโหนดที่ i ในชั้นที่ 1

$\mu_{Ai}(x)$ คือ ฟังก์ชันความเป็นสมาชิกของขาเข้า x บนโหนดที่ i

$\mu_{Bi-2}(y)$ คือ ฟังก์ชันความเป็นสมาชิกของขาเข้า y บนโหนดที่ i

<u>ชั้นที่2</u>

จำนวนโหนดในชั้นนี้มีจำนวนเท่ากับผลคูณของจำนวนฟัซซีเซตสำหรับแต่ละขาเข้า จากตัวอย่างในรูปที่ 4.3 จะมีจำนวนโหนดในชั้นที่ 2 ทั้งหมด 4 โหนด โดยชั้นนี้จะนำขาออกที่ได้จากชั้นที่ 1 มาคูณกันเพื่อให้ได้ค่าถ่วงน้ำหนักดังนี้

$$O_{2,i} = W_i = \mu_{An}(x)\mu_{Bn}(y) \quad ; n=1,2 \; ; i=1,...,4 \tag{4.4}$$

โดย

$O_{2,i}$ คือ ขาออกของโหนดที่ i ในชั้นที่ 2

W_i คือ ค่าถ่วงน้ำหนักของโหนดที่ i

$\mu_{An}(x)$ คือ ฟังก์ชันความเป็นสมาชิกของขาเข้า x บนฟัซซีเซตที่ n

$\mu_{Bn}(y)$ คือ ฟังก์ชันความเป็นสมาชิกของขาเข้า y บนฟัซซีเซตที่ n

<u>ชั้นที่3</u>

ชั้นนี้จะรับค่าจากชั้นที่ 2 มาทำการนอมัลไลซ์และส่งขาออกออกมาดังนี้

$$O_{3,i} = \overline{W}_i = \frac{W_i}{\sum_{i=1}^{4} W_i} \quad ;i=1,...,4 \tag{4.5}$$

โดย

$O_{3,i}$ คือ ขาออกของโหนดที่ i ในชั้นที่ 3

W_i คือ ค่าถ่วงน้ำหนักของโหนดที่ i

$\overline{W}_i$ คือ ค่าถ่วงน้ำหนักที่ผ่านการนอมัลไลซ์แล้ว

<u>ชั้นที่4</u>

ชั้นนี้จะนำค่าขาเข้าที่ป้อนให้กับ ANFIS มาคำนวณผ่านสมการโพลิโนเมียลลำดับที่ 1 โดยสมการโพลิโนเมียลจะประกอบด้วยพารามิเตอร์ที่มีชื่อว่าคอนซีเควนท์พารามิเตอร์ เป็นสัมประสิทธิ์ของสมการ สำหรับ ANFIS ที่มีคอนซีเควนท์เป็นโพลิโนเมียลลำดับที่ 1 มีจำนวนขาเข้าคือ N_{INPUT} มีจำนวนกฎคือ N_{RULE} สามารถหาจำนวนคอนซีเควนท์พารามิเตอร์ทั้งหมด (N_{CONS}) ในชั้นที่ 4 ได้ดังนี้

$$N_{CONS} = (N_{INPUT} + 1) \times N_{RULE} \tag{4.6}$$

สมการโพลิโนเมียลภายในชั้นที่ 4 จะมีจำนวนสัมประสิทธิ์แปรตามจำนวนขาเข้า โดยมีจำนวนเท่ากับจำนวนขาเข้าบวก 1 ค่าที่ได้จากสมการจะคูณด้วยค่าถ่วงน้ำหนักที่ส่งมาจากชั้นที่ 3 จาก ANFIS ในรูปที่ 4.3 ขาออกที่ได้จากโหนดในชั้นนี้คือ

$$O_{4,i} = \overline{W}_i f_i = \overline{W}_i (p_i x + q_i y + r_i) \tag{4.7}$$

โดย

$O_{4,i}$ คือ ขาออกของโหนดที่ i ในชั้นที่ 4

$\overline{W}_i$ คือ ค่าถ่วงน้ำหนักที่ผ่านการนอมัลไลซ์แล้ว

f_i คือ ค่าที่ได้จากการคำนวณสมการโพลิโนเมียลในโหนดที่ i

p_i, q_i, r_i คือ คอนซีเควนท์พารามิเตอร์ในโหนดที่ i

<u>ชั้นที่5</u>

ชั้นนี้นำขาออกที่ได้จากชั้นที่ 4 มารวมกัน ขาออกของชั้นนี้จะเป็นคำตอบของโครงสร้าง ANFIS ที่ต้องการ

$$O_{5,i} = \sum_{i=1}^{4} \overline{W}_i f_i \tag{4.8}$$

โดย

$O_{5,i}$ คือ ขาออกของโหนดที่ i ในชั้นที่ 5

$\overline{W}_i$ คือ ค่าถ่วงน้ำหนักที่ผ่านการนอมัลไลซ์แล้ว

f_i คือ ค่าที่ได้จากการคำนวณสมการโพลิโนเมียลในโหนดที่ i

4.3 การฝึกสอน ANFIS

ANFIS จะสามารถทำงานได้อย่างถูกต้อง จะต้องฝึกสอน ANFIS ด้วยขาเข้าขาออกตัวอย่าง เมื่อระหว่างการฝึกสอน ANFIS จะเปลี่ยนแปลงพารามิเตอร์ภายในเพื่อให้สามารถคำนวณขาออกได้ใกล้เคียงกับตัวอย่างที่นำมาฝึกสอน โดยพารามิเตอร์ที่มีการเปลี่ยนแปลงระหว่างการฝึกสอน ANFIS มีอยู่ 2 ชุดคือ พรีมิสพารามิเตอร์ และคอนซีเควนท์พารามิเตอร์ ที่อยู่ในชั้นที่ 2 และชั้นที่ 4 ของโครงสร้าง ANFIS ตามลำดับ

4.3.1 การหาคอนซีเควนท์พารามิเตอร์

หลักการในการฝึกสอนคือการทำให้ขาออกที่ได้จากการคำนวณโดยโครงสร้างANFIS มีค่าใกล้เคียงกับขาออกตัวอย่างที่นำมาฝึกสอน การหาคอนซีเควนท์พารามิเตอร์ทำได้โดยการแก้ปัญหาออปติไมเซชัน เพื่อให้ได้พารามิเตอร์ที่ให้ค่าความคลาดเคลื่อนกำลังสูงมีค่าน้อยที่สุดดังนี้

$$\min \quad E(\theta) = \sum_{p=1}^{P} (f_{pt} - f_{pA}(\theta))^2 \tag{4.9}$$

โดย

$\theta = \left[\theta_1, ..., \theta_n\right]^T$ คือ คอนซีเควนท์พารามิเตอร์ที่ต้องการหา

P คือ จำนวนขาเข้าและขาออกตัวอย่าง

$E(\theta)$ คือ ความคลาดเคลื่อนกำลังสอง

f_{pt} คือ ขาออกตัวอย่างลำดับที่ p

$f_{pA}(\theta)$ คือ ขาออกที่คำนวณได้จาก ANFIS ของข้อมูลตัวอย่างที่ p

จากโครงสร้าง ANFIS ในรูปที่ 4.3 จะได้คอนซีเควนท์พารามิเตอร์คือ

$$\theta = \left[p_1, q_1, r_1, ..., p_4, q_4, r_4\right]^T \tag{4.10}$$

และค่าขาออกของ ANFIS ในข้อมูลตัวอย่างที่ p คือ

$$f_{pA}(\theta) = a_p^T \theta$$

$$= \begin{bmatrix} W_{1p} x_p & W_{1p} y_p & W_{1p} & \dots & W_{4p} x_p & W_{4p} y_p & W_{4p} \end{bmatrix} \cdot$$
$$\begin{bmatrix} p_1 & q_1 & r_1 & \dots & p_4 & q_4 & r_4 \end{bmatrix}^T \tag{4.11}$$

จาก 4.9 จะได้

$$E(\theta) = \sum_{p=1}^{P} (f_{pt} - a_p^T \theta)^2$$
$$= (f_t - A\theta)^T (f_t - A\theta) \tag{4.12}$$

สำหรับตัวอย่างรูปที่ 4.3 จะได้ A เป็นเมตริกซ์ขนาด Px12 ดังนี้

$$A = \begin{bmatrix} W_{11} x_1 & W_{11} y_1 & W_{11} & \dots & W_{41} x_1 & W_{41} y_1 & W_{41} \\ & & & \vdots & & & \\ W_{1p} x_p & W_{1p} y_p & W_{1p} & \dots & W_{4p} x_p & W_{4p} y_p & W_{4p} \\ & & & \vdots & & & \\ W_{1P} x_P & W_{1P} y_P & W_{1P} & \dots & W_{4P} x_P & W_{4P} y_P & W_{4P} \end{bmatrix} \tag{4.13}$$

และ f_t คือเมตริกซ์ขนาด Px1 ประกอบด้วยค่าขาออกตัวอย่างทั้งหมดดังนี้

$$f_t = \begin{bmatrix} f_{1t} & \dots & f_{pt} & \dots & f_{Pt} \end{bmatrix}^T \tag{4.14}$$

สมการที่ 4.12 อยู่ในรูปสมการควอดราติกซึ่งสามารถแก้ปัญหาออปติไมเซชันได้ด้วยวิธี LSE หรือ Least-squares estimator [18] ซึ่งทำให้สามารถหาค่าคอนซีเควนพารามิเตอร์ได้ดังนี้

$$\theta = (A^T A)^{-1} A^T f_t \tag{4.15}$$

ค่าคอนซีเควนท์พารามิเตอร์ที่ได้จะเป็นค่าพารามิเตอร์ที่ทำให้ ANFIS คำนวณขาออกออกมามีค่าความคลาดเคลื่อนน้อยที่สุดเทียบกับขาออกตัวอย่างที่นำมาฝึกสอน

4.3.2 การหาพรีมิสพารามิเตอร์

เนื่องจากค่าพรีมิสพารามิเตอร์อยู่ในชั้นที่ 2 ของ ANFIS สมการความคลาดเคลื่อนกำลังสองขาออกที่เป็นฟังก์ชันของ พรีมิสพารามิเตอร์จึงไม่ใช่สมการควอดราติกเหมือนกับ 4.12 การแก้ปัญหาออปติไมเซชันเพื่อหาพรีมิสพารามิเตอร์จึงต้องใช้การทำออปติไมเซชันด้วยวิธีเกรเดียน ซึ่งต้องหาอนุพันธ์ของฟังก์ชันความคลาดเคลื่อนกำลังสองเทียบกับพรีมิสพารามิเตอร์ วิธีที่ช่วยหาอนุพันธ์ดังกล่าวคือ Backpropagation ซึ่งต้องใช้หลักการของอนุพันธ์ลำดับ (ordered derivative)

4.3.2.1 อนุพันธ์ลำดับ (ordered derivative) [19]

อนุพันธ์ลำดับเป็นอนุพันธ์ของฟังก์ชันเทียบกับตัวแปรทั้งที่เป็นตัวแปรโดยตรงและไม่ใช่ตัวแปรโดยตรงของฟังก์ชัน ซึ่งแตกต่างจากอนุพันธ์ย่อย (partial derivative) ที่เป็นอนุพันธ์ที่เทียบกับตัวแปรที่อยู่ในฟังก์ชันโดยตรงเท่านั้น

กำหนดให้ z เป็นฟังก์ชันของ x และ y กำหนดให้ y เป็นฟังก์ชันของ x ดังนี้

$$\begin{cases} z = g(x, y) \\ y = f(x) \end{cases} \tag{4.16}$$

จะได้อนุพันธ์ย่อยของ z เทียบกับ x คือ

$$\frac{\partial z}{\partial x} = \frac{\partial g(x, y)}{\partial x} \tag{4.17}$$

อนุพันธ์ลำดับของ z เทียบกับ x สามารถแสดงได้ดังต่อไปนี้

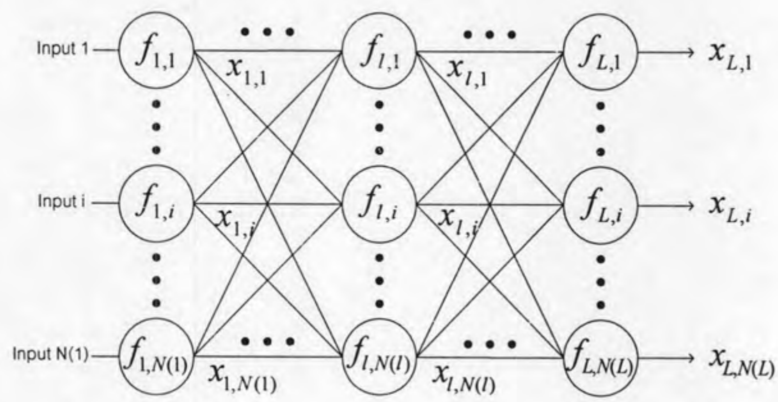
$$\begin{aligned} \frac{\partial^+ z}{\partial x} &= \frac{\partial g(x, f(x))}{\partial x} \\ &= \left.\frac{\partial g(x, y)}{\partial x}\right|_{y=f(x)} + \left.\frac{\partial g(x, y)}{\partial y}\right|_{y=f(x)} \frac{\partial f(x)}{\partial x} \end{aligned} \tag{4.18}$$

ดังนั้นจากหลักการนี้จึงสามารถนำไปหาอนุพันธ์ของฟังก์ชันความคลาดเคลื่อนกำลังสองเทียบกับพรีมิสพารามิเตอร์ที่อยู่ภายใน ANFIS ได้

4.3.2.2 Backpropagation

วิธีการ Backpropagation [20] เป็นวิธีการที่ใช้กับโครงข่ายปรับตัวได้ทั่วไป ใช้สำหรับหาค่าอนุพันธ์ของฟังก์ชันความคลาดเคลื่อนกำลังสองเทียบกับพารามิเตอร์ที่ต้องการ เพื่อนำค่าอนุพันธ์ที่ได้ไปปรับปรุงค่าพารามิเตอร์ในการแก้ปัญหาออปติไมเซชัน การหาค่าอนุพันธ์ลำดับของฟังก์ชันจะใช้วิธีการแพร่ย้อนกลับของสัญญาณคลาดเคลื่อน (error signal, $\varepsilon_{l,i}$)[20]

สำหรับโครงข่ายปรับตัวได้ใด ๆ ที่มีจำนวนชั้น L ชั้น กำหนดให้ l เป็นหมายเลขชั้นซึ่งมีจำนวนโหนดคือ $N(l)$ และโหนดต่าง ๆ มีหมายเลขโหนดคือ i ฟังก์ชันในแต่ละโหนดคือ $f_{l,i}$ และขาออกแต่ละโหนดคือ $x_{l,i}$ ดังแสดงในรูปด้านล่าง



รูปที่ 4.5 โครงข่ายปรับตัวได้

ค่าความคลาดเคลื่อนกำลังสองของขาออกตัวอย่างที่ p จากขาเข้าขาออกตัวอย่างจำนวน P ($1 \le p \le P$) คือ

$$E_p = \sum_{i=1}^{N(L)} (d_i - x_{L,i})^2 \tag{4.19}$$

ซึ่ง d_i คือขาออกลำดับที่ i และ E_p เป็นความคลาดเคลื่อนกำลังสองเฉพาะตัวอย่างที่ p ดังนั้นความคลาดเคลื่อนกำลังสองโดยรวมทุก ๆ ข้อมูลตัวอย่างคือ

$$E = \sum_{p=1}^{P} E_p \tag{4.20}$$

สัญญาณคลาดเคลื่อนมีนิยามคืออนุพันธ์ลำดับของฟังก์ชันความคลาดเคลื่อนเทียบกับขาออกของโหนดต่าง ๆ ดังนี้

$$\varepsilon_{l,i} = \frac{\partial^{+} E_p}{\partial x_{l,i}} \tag{4.21}$$

ที่ชั้นสุดท้าย (ชั้นที่ L) ของโครงข่ายปรับตัวได้ ขาออกของโหนดจะเป็นฟังก์ชันโดยตรงกับฟังก์ชันความคลาดเคลื่อนกำลังสอง (E_p) ดังนั้นสัญญาณคลาดเคลื่อน ($\varepsilon_{L,i}$) จะเป็นอนุพันธ์ย่อยของฟังก์ชันความคลาดเคลื่อนกำลังสองเทียบกับขาออกของโหนดที่ชั้นนี้ ดังนี้

$$\varepsilon_{L,i} = \frac{\partial^{+} E_p}{\partial x_{L,i}} = \frac{\partial E_p}{\partial x_{L,i}} = -2(d_i - x_{L,i}) \tag{4.22}$$

แต่สำหรับชั้นที่ l ใด ๆ สัญญาณคลาดเคลื่อนของขาออกของโหนดต่าง ๆ ($\varepsilon_{l,i}$) ในชั้นนี้สามารถเขียนให้อยู่ในรูปแบบสัญญาณคลาดเคลื่อนของโหนดในชั้นถัดไป ($\varepsilon_{l+1,m}$) ได้ดังนี้

$$\varepsilon_{l,i} = \underbrace{\frac{\partial^{+} E_p}{\partial x_{l,i}}}_{\substack{\text{สัญญาณ} \\ \text{คลาดเคลื่อน} \\ \text{ที่ชั้น } l \text{ โหนด} \\ \text{ที่ } i}} = \sum_{m=1}^{N(l+1)} \underbrace{\frac{\partial^{+} E_p}{\partial x_{l+1,m}}}_{\substack{\text{สัญญาณ} \\ \text{คลาดเคลื่อน} \\ \text{ที่ชั้น } l+1 \\ \text{โหนดที่ } m}} \frac{\partial f_{l+1,m}}{\partial x_{l,i}} = \sum_{m=1}^{N(l+1)} \varepsilon_{l+1,m} \frac{\partial f_{l+1,m}}{\partial x_{l,i}} \tag{4.23}$$

ในการหาสัญญาณคลาดเคลื่อนที่ชั้น l สามารถทำได้โดยการหาสัญญาณคลาดเคลื่อนที่ชั้นสุดท้าย (ชั้นที่ L) ด้วย 4.22 และใช้ 4.23 หาสัญญาณคลาดเคลื่อนในชั้นก่อนหน้านี้ ย้อนไปจนถึงชั้นที่ l

สัญญาณคลาดเคลื่อน $\varepsilon_{l,i}$ ที่ได้เป็นอนุพันธ์ลำดับเทียบกับขาออกของโหนดนั้น ($\partial^{+} E_p / \partial x_{l,i}$) ซึ่งสามารถนำค่า $\varepsilon_{l,i}$ เพื่อหาค่าอนุพันธ์ลำดับเทียบกับพารามิเตอร์ในโหนดนั้น ($\partial^{+} E_p / \partial \alpha$) ดังนี้

$$\frac{\partial^{+}E_{p}}{\partial\alpha}=\frac{\partial^{+}E_{p}}{\partial x_{l,i}}\frac{\partial f_{l,i}}{\partial\alpha}=\varepsilon_{l,i}\frac{\partial f_{l,i}}{\partial\alpha} \tag{4.24}$$

และจะได้อนุพันธ์ผลรวมของความคลาดเคลื่อนกำลังสองเทียบกับพารามิเตอร์ α คือ

$$\frac{\partial^{+}E}{\partial\alpha}=\sum_{p=1}^{P}\frac{\partial^{+}E_{p}}{\partial\alpha} \tag{4.25}$$

ซึ่งสามารถนำสมการข้างต้นไปใช้เพื่อหาค่าพรีมิสพารามิเตอร์ด้วยการแก้ปัญหาออปติไมเซชันได้ในหัวข้อต่อไป

4.3.2.3 การคำนวณพรีมิสพารามิเตอร์

การหาค่าพรีมิสพารามิเตอร์ทำได้โดยการแก้ปัญหาออปติไมเซชันฟังก์ชันความคลาดเคลื่อนกำลังสองเช่นเดียวกันกับการหาค่าคอนซีเควนท์พารามิเตอร์ดังนี้

$$\min \quad E=\sum_{p=1}^{P}(f_{pt}-f_{pA})^{2} \tag{4.26}$$

โดย

E คือ ความคลาดเคลื่อนกำลังสอง

P คือ จำนวนขาเข้าและขาออกตัวอย่าง

f_{pt} คือ ขาออกตัวอย่างลำดับที่ p

f_{pA} คือ ขาออกที่คำนวณได้จาก ANFIS ของข้อมูลตัวอย่างที่ p ซึ่งมีความสัมพันธ์กับ พรีมิสพารามิเตอร์ (α)

สมการที่ 4.26 มีข้อแตกต่างจาก 4.12 คือค่าพรีมิสพารามิเตอร์ (α) อยู่ในชั้นที่ 2 ของ ANFIS ดังนั้นรูปสมการของความคลาดเคลื่อนกำลังสอง (E) จึงไม่ใช่สมการควอดราติก แต่เป็นสมการไม่เชิงเส้นในรูปแบบที่ไม่เจาะจง โดยเราสามารถแก้ปัญหาออปติไมเซชันสมการไม่เชิงเส้นนี้ได้ด้วยวิธีเกรเดียน หรืออีกชื่อหนึ่งเรียกว่า Steepest-descent [21] การแก้ปัญหาจะตั้งพรีมิสพารามิเตอร์เริ่มต้นแล้วมีการปรับปรุงค่า จนกว่าจะได้ค่าฟังก์ชันเป้าหมายต่ำที่สุด วิธีการปรับปรุงค่าเป็นไปตามสมการดังนี้

$$\alpha_{k+1} = \alpha_k + \Delta\alpha \tag{4.27}$$

โดย α_k และ α_{k+1} คือค่าพรีมิสพารามิเตอร์จากการทำซ้ำครั้งที่ k และ k+1 ตามลำดับ $\Delta\alpha$ คือค่าที่ใช้ปรับปรุงคำตอบในการทำซ้ำ ซึ่งในการหาพรีมิสพารามิเตอร์ $\Delta\alpha$ เป็นอนุพันธ์ลำดับของความคลาดเคลื่อนกำลังสอง (E) เทียบกับพรีมิสพารามิเตอร์ที่พิจารณา (α) ดังนี้

$$\Delta\alpha = \frac{\partial^{+} E}{\partial\alpha} \tag{4.28}$$

ดังนั้น $\Delta\alpha$ จึงสามารถหาได้โดยใช้วิธี Backpropagation ตามหัวข้อที่กล่าวมาข้างต้น

4.3.3 การฝึกสอนแบบไฮบริด (hybrid training)

การฝึกสอน ANFIS จะทำให้ได้ค่าพารามิเตอร์ที่ให้ขาออกของ ANFIS สอดคล้องกับตัวอย่างที่นำมาใช้ฝึกสอน โดยการฝึกสอนที่ได้กล่าวมาข้างต้นแบ่งออกเป็น 2 ชุด คือส่วนคอนซีเควนท์พารามิเตอร์และพรีมิสพารามิเตอร์ ในการฝึกสอนนั้นอาจจะฝึกสอนเพียงแค่คอนซีเคนท์พารามิเตอร์ก็ได้ แต่เพื่อความแม่นยำในการคำนวณมากขึ้นจึงนิยมฝึกสอนเพื่อให้ได้พารามิเตอร์ทั้ง 2 ชุด ซึ่งเรียกการฝึกสอนลักษณะนี้ว่าการฝึกสอนแบบไฮบริดซึ่งมีขั้นตอนดังนี้

1.) ตั้งค่าพรีมิสพารามิเตอร์เริ่มต้น
2.) คงค่าพรีมิสพารามิเตอร์แล้วฝึกสอนเพื่อให้ได้ค่าคอนซีเควนท์พารามิเตอร์ทั้งหมดด้วยวิธี LSE ตามหัวข้อ 4.3.1
3.) ฝึกสอนให้ได้พรีมิสพารามิเตอร์ตามหัวข้อ 4.3.2 โดยใช้วิธี Backpropagation โดยหาพรีมิสพารามิเตอร์ทีละค่า โดยคงค่าพรีมิสพารามิเตอร์อื่นและคงค่าคอนซีเควนท์พารามิเตอร์ไว้
4.) จากขั้นตอนที่ 1-3 ถือเป็นการคำนวณ 1 รอบ (epoch) จากนั้นทำซ้ำขั้นตอนที่ 1-3 อีกครั้งจนกว่าค่าความคลาดเคลื่อนกำลังสองจะมีค่าอยู่ในระดับที่ยอมรับได้

อย่างไรก็ตามในการฝึกสอนแบบไฮบริดในรอบถัดไปอาจจะให้ค่าความคลาดเคลื่อนของขาออกมีค่ามากกว่ารอบก่อนหน้า ดังนั้นในการฝึกสอนอาจจะมีการกำหนดจำนวนรอบในการฝึกสอนและเลือกใช้ ANFIS ที่ผ่านการฝึกสอนในรอบที่ให้ค่าความคลาดเคลื่อนน้อยที่สุดมาใช้งาน

4.4 สรุป

จากที่กล่าวมาจะเห็นได้ว่า ANFIS มีโครงสร้างพื้นฐานที่มาจากระบบอนุมานฟัซซีแบบ Sugeno แต่พารามิเตอร์ภายในจะได้มาจากการนำข้อมูลตัวอย่างมาฝึกสอน ANFIS ทำให้ ANFIS สามารถแปลงค่าขาเข้าของระบบใด ๆ ให้กลายเป็นขาออกที่คล้ายคลึงกับระบบนั้น ๆ ได้ โดยไม่ต้องสร้างแบบจำลองของระบบขึ้นมาโดยตรง ถึงแม้ว่ากระบวนการฝึกสอนของ ANFIS จะมีวิธีการที่ยุ่งยากเพื่อที่จะหาพารามิเตอร์ที่เหมาะสม อย่างไรก็ตามกระบวนการคำนวณภายในของ ANFIS ก็เป็นการส่งค่าไปยังชั้นต่อไปเรื่อย ๆ ANFIS ที่ผ่านการฝึกสอนแล้วจึงสามารถคำนวณขาออกได้อย่างรวดเร็ว ซึ่งเป็นข้อดีที่จะนำ ANFIS ไปใช้งานต่อไป