

## บทที่ 2

### ทฤษฎีเวฟเล็ตและโครงข่ายเวฟเล็ต

บทนี้จะกล่าวถึงทฤษฎีเวฟเล็ต การแปลงเวฟเล็ตชนิดต่อเนื่อง รวมทั้งชนิดไม่ต่อเนื่อง โดยจะเริ่มจากกรณีฟังก์ชันที่จะถูกแปลงมีมิติเดียวไปจนถึงกรณีฟังก์ชันที่มีมิติมากกว่าหนึ่ง จากนั้นจะกล่าวถึงลักษณะโครงสร้างของโครงข่ายเวฟเล็ต รวมทั้งวิธีการปรับค่าพารามิเตอร์ของโครงข่ายในการเรียนรู้ความสัมพันธ์ใด ๆ ที่ต้องการ ดังนี้

#### ทฤษฎีเวฟเล็ต

ทฤษฎีเวฟเล็ต (Wavelet Theory) เป็นทฤษฎีทางคณิตศาสตร์ที่เกี่ยวข้องกับการสร้างแบบจำลองสำหรับสัญญาณ (signals), ระบบ (systems) หรือกระบวนการ (processes) ด้วยเซต (set) ของสัญญาณที่เกิดจากการกระทำทางคณิตศาสตร์กับสัญญาณพิเศษ สัญญาณพิเศษดังกล่าวจะมีลักษณะคล้ายคลื่นเล็ก ๆ (little wave : ซึ่งเป็นที่มาของคำว่า 'wavelet' [7]) นั่นคือมีการสั่น (oscillate) หรือมีความเป็นคลื่น (wave) และต้องมีขนาดลดลง (decay) อย่างรวดเร็วในทิศทางทั้งสองด้านของแกนตัวแปรอิสระ

การกระทำทางคณิตศาสตร์กับสัญญาณพิเศษที่มีลักษณะคล้ายคลื่นเล็ก ๆ ข้างต้น คือ การเลื่อน (translation) และการถ่าง (dilation) (ดังจะกล่าวถึงในหัวข้อถัดไป) ดังนั้นการแทนสัญญาณ ระบบ หรือกระบวนการด้วยเซตของสัญญาณที่เกิดจากการเลื่อนและการถ่างต่าง ๆ กัน จึงเป็นการแปลงสัญญาณ ระบบ หรือกระบวนการจากโดเมน (domain) หนึ่ง ไปสู่อีกโดเมนหนึ่ง

ต่อไปจะกล่าวถึงฟังก์ชันที่มีลักษณะพิเศษสำหรับการใช้ในการแปลงเวฟเล็ต การแปลงเวฟเล็ตชนิดต่อเนื่อง การแปลงชนิดไม่ต่อเนื่อง ซึ่งจะรวมทั้งการปรับวิธีการแปลงเพื่อใช้กับกรณีเมื่อฟังก์ชันที่ต้องการแปลงมีมิติมากกว่าหนึ่งขึ้นไป ดังนี้

#### ฟังก์ชันมูลฐานในการแปลงเวฟเล็ต

การแปลงเวฟเล็ต (Wavelet Transform) เป็นการแทนความสัมพันธ์ ไม่ว่าจะเป็ นฟังก์ชันหรือสัญญาณใด ๆ ด้วยเซตของฟังก์ชันมูลฐานที่เกิดจากการเลื่อน และการถ่าง ณ ค่าการ

เลื่อนและการต่างต่าง ๆ กัน โดยฟังก์ชันมูลฐานที่ใช้อาจเป็นฟังก์ชันใด ๆ ก็ได้ที่มีคุณสมบัติยอมรับได้ (admissible) ดังจะกล่าวถึงต่อไป ถ้าจะเปรียบเทียบกับ การแปลงฟูเรีย (Fourier Transform) แล้ว การแปลงฟูเรีย นั้นเป็นการแทนความสัมพันธ์ด้วยฟังก์ชันมูลฐานที่เป็นรูปไซน์ (sinusoid) ซึ่งมีค่าความถี่ต่าง ๆ กัน นั่นคือการแปลงฟูเรีย เป็นการแปลงความสัมพันธ์จากโดเมนเวลา (time domain) ไปสู่โดเมนความถี่ (frequency domain) ส่วนการแปลงเวฟเลตนั้นเป็นการแปลงความสัมพันธ์จากโดเมนเวลาไปสู่โดเมนของการเลื่อนและโดเมนของการต่างในขณะเดียวกัน ซึ่งจะรวมเรียกว่าโดเมนเวฟเลต (wavelet domain)

ฟังก์ชันที่ใช้เป็นฟังก์ชันมูลฐานในการแปลงเวฟเลตนั้น จำเป็นอย่างยิ่งที่จะต้องมีความสมบัติยอมรับได้ คุณสมบัตินี้คือ ฟังก์ชัน  $\psi(t) \in L^2(\mathbb{R})$  ใด ๆ จะยอมรับได้เมื่อ

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (2-1)$$

เมื่อ  $\Psi(\omega)$  เป็นการแปลงฟูเรียของ  $\psi(t)$  และเรียกค่า  $C_\psi$  ว่าค่าคงที่ยอมรับได้ (admissibility constant) จาก (2-1) สามารถกล่าวอีกนัยหนึ่งได้ว่า

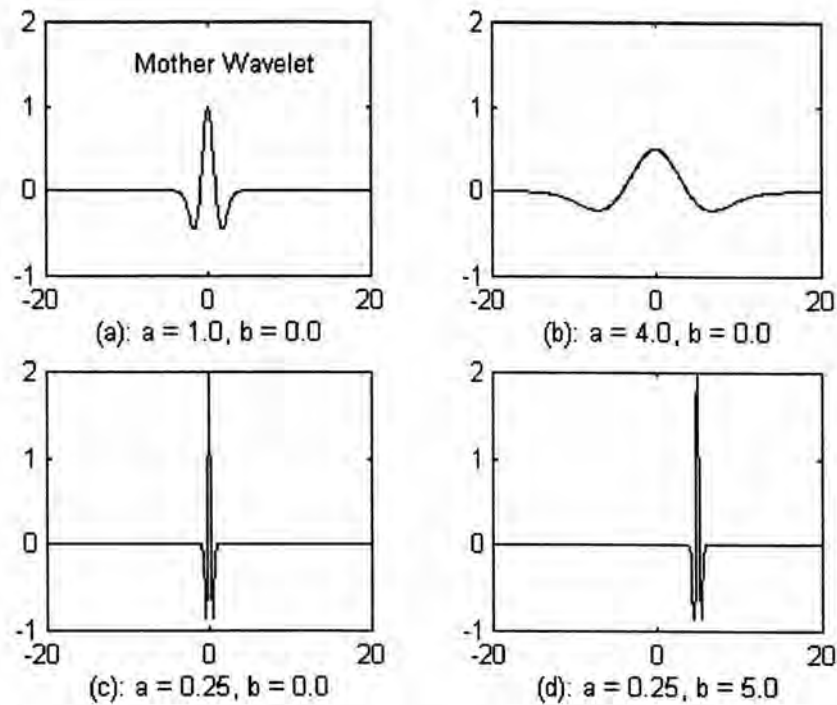
$$\int_{-\infty}^{+\infty} \psi(t) \cdot dt = 0 \quad (2-2)$$

จากข้อกำหนดข้างต้น จะเห็นว่าฟังก์ชันใดก็ตามที่มีคุณสมบัติสอดคล้องกับ (2-1) สามารถใช้เป็นฟังก์ชันมูลฐานในการแปลงเวฟเลตได้ทั้งสิ้น ฟังก์ชัน  $\psi(t)$  ดังกล่าวมีชื่อเรียกว่าเวฟเลตแม่ (mother wavelet)

โดยค่าการเลื่อน (translate)  $b \in \mathbb{R}$  และค่าการต่าง (dilate)  $a \in \mathbb{R} - \{0\}$  จะได้ตระกูลเวฟเลต (wavelet families) คือ

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \cdot \psi\left(\frac{t-b}{a}\right) \quad (2-3)$$

สำหรับตัวคูณ  $\frac{1}{\sqrt{|a|}}$  นั้นมีเพียงเพื่อให้ค่าพลังงานของ  $\psi_{a,b}(t)$  มีค่าเท่ากับค่าพลังงานของเวฟเลตแม่  $\psi(t)$



รูปที่ 2-1 เวฟเล็ตแม่  $\psi(t)$  และผลการต่างและการเลื่อนของเวฟเล็ตแม่  $\psi_{a,b}(t)$

ในที่นี้จะขอใช้ฟังก์ชันรูปทรงหมวกเม็กซิกัน (Mexican hat function) เพื่อเป็นตัวอย่าง สำหรับฟังก์ชันที่มีคุณสมบัติยอมรับได้ดังรูปที่ 2-1(a) ซึ่งมีสมการเป็น

$$\psi(t) = (1 - t^2) \cdot \exp\left(-\frac{t^2}{2}\right) \quad (2-4)$$

ส่วนในรูปที่ 2-1(b), รูปที่ 2-1(c), และ รูปที่ 2-1(d) เป็นผลที่เกิดจากการเลื่อนและการต่างเวฟเล็ตแม่ (2-4) ที่ค่าการเลื่อน  $b$  และการต่าง  $a$  ต่าง ๆ กันดังแสดงในภาพ

### การแปลงเวฟเล็ตชนิดต่อเนื่อง

ดังได้กล่าวมาแล้วว่าการแปลงเวฟเล็ตนั้นเป็นการแปลงความสัมพันธ์จากโดเมนเวลาไปสู่โดเมนเวฟเล็ตอันประกอบด้วยการต่างและการเลื่อน ค่าสัมประสิทธิ์ (coefficient) ของการแปลงเวฟเล็ตชนิดต่อเนื่อง (Continuous Wavelet Transform) ของฟังก์ชัน  $f(t) \in L^2(\mathcal{R})$  ณ ค่าการต่าง  $a \in \mathcal{R} - \{0\}$  และค่าการเลื่อน  $b \in \mathcal{R}$  หาได้จาก

$$\begin{aligned}
 CWT_{\psi} f(a,b) &= \langle f, \psi_{a,b} \rangle \\
 &= \frac{1}{\sqrt{|a|}} \cdot \int_{-\infty}^{+\infty} f(t) \cdot \psi^* \left( \frac{t-b}{a} \right) \cdot dt
 \end{aligned} \tag{2-5}$$

โดยที่เครื่องหมายคอนจูเกตแทนคอนจูเกตเชิงซ้อน (complex conjugate) ในกรณีที่เวฟเล็ตแม่เป็นฟังก์ชันเชิงซ้อน จากสมการข้างต้นอาจกล่าวได้ว่าค่าสัมประสิทธิ์  $CWT_{\psi} f(a,b)$  เป็นค่าวัดความคล้ายกันระหว่าง  $\psi_{a,b}(t)$  กับฟังก์ชัน  $f$  กล่าวคือ ถ้าฟังก์ชันทั้งสองคล้ายกันมาก ค่าสัมประสิทธิ์ก็จะมีค่าสูงขึ้น

ส่วนการแปลงกลับ (inverse transform) ของการแปลงเวฟเล็ตชนิดต่อเนื่องนั้นเป็นการแปลงพื้นที่  $CWT_{\psi} f(a,b), \forall a \in \mathcal{R} - \{0\}, \forall b \in \mathcal{R}$  ในโดเมนการเลื่อน-การถ่วงกลับไปสู่โดเมนเวลาดังนี้

$$f(t) = C_{\psi}^{-1} \cdot \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} CWT_{\psi} f(a,b) \cdot \psi_{a,b}(t) \cdot \frac{da \cdot db}{a^2} \tag{2-6}$$

นั่นคือการแปลงกลับสามารถทำได้อย่างสมบูรณ์

เมื่อฟังก์ชันที่ต้องการทำการแปลงมีมิติ (dimension) มากกว่าหนึ่ง นั่นคือ  $f(\mathbf{t}) \in L^2(\mathcal{R}^n); n > 1$  สิ่งที่ต้องเปลี่ยนไปคือเวฟเล็ตแม่จะต้องมีมิติเพิ่มขึ้นตามไปด้วย สำหรับรูปแบบหนึ่งของเวฟเล็ตแม่มิติ  $n$  เป็นดังสมการ (2-7) ซึ่งถ้ากำหนดให้  $\psi(\mathbf{t}) \in L^2(\mathcal{R}^n)$  และ

$$\psi(\mathbf{t}) = \psi_s(t_1) \cdot \psi_s(t_2) \cdot \dots \cdot \psi_s(t_n) \tag{2-7}$$

โดยที่  $\psi_s(\bullet)$  เป็นเวฟเล็ตแม่ที่มีมิติ 1 และ  $\mathbf{t}^T = [t_1 \ t_2 \ \dots \ t_n]$  แล้ว จะได้ว่า  $\psi(\mathbf{t})$  มีคุณสมบัติยอมรับได้ ซึ่งสามารถใช้เป็นเวฟเล็ตแม่ในกรณีที่ฟังก์ชันมีมิติมากกว่าหนึ่งได้ และจะได้ตระกูลเวฟเล็ตสำหรับการแปลงเวฟเล็ตชนิดต่อเนื่องเป็น

$$\psi_{a,b}(\mathbf{t}) = \frac{1}{\sqrt{|a|^n}} \cdot \psi \left( \frac{\mathbf{t}-\mathbf{b}}{a} \right) \tag{2-8}$$

โดยที่  $a \in \mathcal{R} - \{0\}$  และ  $\mathbf{b} \in \mathcal{R}^n$  สำหรับสูตรการแปลงและการแปลงกลับนั้น จะเป็นดัง (2-9) และ (2-10) ตามลำดับดังนี้ [8]

$$CWT_{\psi} f(a, \mathbf{b}) = \langle f, \psi_{a, \mathbf{b}} \rangle \quad (2-9)$$

$$f(\mathbf{t}) = C_{\psi}^{-1} \cdot \int_{-\infty}^{+\infty} \left\{ \int_{\mathcal{R}^n} CWT_{\psi} f(a, \mathbf{b}) \cdot \psi_{a, \mathbf{b}}(\mathbf{t}) \cdot d\mathbf{b} \right\} \cdot \frac{da}{a^{n+1}} \quad (2-10)$$

### การแปลงเวฟเล็ตชนิดไม่ต่อเนื่อง

การแปลงเวฟเล็ตชนิดไม่ต่อเนื่อง (Discrete Wavelet Transform) เป็นการแปลงความสัมพันธ์จากโดเมนเวลา (ซึ่งต่อเนื่อง) ไปยังโดเมนการถ่วงและโดเมนการเลื่อนซึ่งไม่ต่อเนื่อง คล้ายกับอนุกรมฟูเรียร์ (Fourier Series) มีตำราบางเล่มเช่น Young [7] เรียกการแปลงนี้ว่า อนุกรมเวฟเล็ตต่อเนื่องทางเวลา (Continuous Time Wavelet Series)

การแปลงเวฟเล็ตชนิดนี้ได้ผลการแปลงที่มีจำนวนสัมประสิทธิ์นับได้ในโดเมนเวฟเล็ต การแปลงชนิดนี้เป็นการแปลงความสัมพันธ์จาก  $L^2(\mathcal{R})$  ไปยัง  $l^2(\mathcal{Z}^2)$  นั่นคือในโดเมนเวฟเล็ตจะมีตัวแปรสองตัวที่เป็นจำนวนเต็มได้แก่  $m \in \mathcal{Z}$  เป็นอันดับของการถ่วง และ  $n \in \mathcal{Z}$  เป็นอันดับของการเลื่อน จาก (2-3) โดยแทนค่าการถ่วงในการแปลงเวฟเล็ต  $a$  ด้วย  $a_0^m, a_0 > 1$  และแทนค่าการเลื่อน  $b$  ด้วย  $nb_0 a_0^m, b_0 > 0$  ค่า  $a_0$  และ  $b_0$  เป็นระยะก้าวของการถ่วง (dilation step size) และระยะก้าวของการเลื่อน (translation step size) ตามลำดับ อันจะทำให้ตระกูลเวฟเล็ตชนิดไม่ต่อเนื่องกลายเป็น

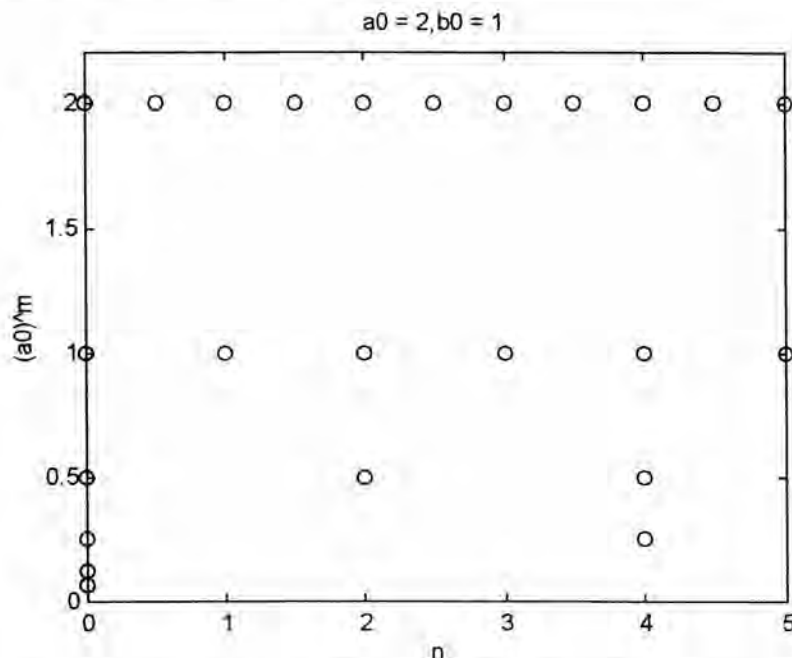
$$\psi_{m, n}(t) = \frac{1}{\sqrt{|a_0^m|}} \cdot \psi\left(\frac{t - n \cdot b_0 \cdot a_0^m}{a_0^m}\right) \quad (2-11)$$

และสามารถคำนวณอนุกรมเวฟเล็ตต่อเนื่องทางเวลาได้ดังนี้

$$\begin{aligned} CTWS_{\psi} f(m, n) &= \langle f, \psi_{m, n} \rangle \\ &= \frac{1}{\sqrt{|a_0^m|}} \cdot \int_{-\infty}^{+\infty} f(t) \cdot \psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) \cdot dt \end{aligned} \quad (2-12)$$

จากสมการดังกล่าวจะเห็นว่ายิ่ง  $a_0$  เข้าใกล้ 1.0 และ  $b_0$  เข้าใกล้ 0.0 การแปลงชนิดไม่ต่อเนื่องจะยิ่งคล้ายกับการแปลงชนิดต่อเนื่องมากยิ่งขึ้น

การแปลงเวฟเล็ดชนิดไม่ต่อเนื่องเป็นการแทนฟังก์ชัน  $f$  ด้วยเซตของค่าสัมประสิทธิ์ของเวฟเล็ดในจำนวนที่นับได้ จะเห็นได้ว่า ณ ค่าการถ่าง  $m$  ต่ำ ๆ (ความถี่สูง) เวฟเล็ดแม่จะถูกเลื่อนด้วยก้าวสั้น ๆ เพื่อให้ครอบคลุมแกนเวลาทั้งหมด ในขณะที่ค่าการถ่างสูง (ความถี่ต่ำ) เวฟเล็ดแม่จะถูกเลื่อนด้วยก้าวที่ยาวขึ้นดังแสดงในรูปที่ 2-2



รูปที่ 2-2 ตำแหน่งของเวฟเล็ดที่ถูกเลื่อนและถ่างด้วยค่าการเลื่อนและการถ่างต่าง ๆ กัน

ในกรณีของการแปลงเวฟเล็ดชนิดไม่ต่อเนื่องนั้น ตระกูลเวฟเล็ด  $\psi_{m,n}(t)$  สามารถเป็น bi-orthogonal, semi-orthogonal หรือแม้กระทั่ง orthonormal ได้ [1] นั่นคือสามารถทำให้ฟังก์ชันมูลฐานที่ใช้ในอนุกรมเวฟเล็ดต่อเนื่องทางเวลาให้มีความเป็นอิสระต่อกันมากขึ้นได้ตามความต้องการในการใช้งานโดยการกำหนดค่า  $a_0$  และ  $b_0$  ให้เหมาะสม และความเป็นอิสระดังกล่าวยังขึ้นกับเวฟเล็ดแม่อีกด้วย

ค่าความไม่เป็นอิสระเชิงเส้นต่อกันจะถูกเรียกว่า ‘ซ้ำซ้อน’ (redundant) นั่นคือจะมีการซ้อนทับกันของตระกูลเวฟเล็ดในโดเมนเวฟเล็ด ซึ่งความซ้ำซ้อนนี้จะช่วยเพิ่มความคงทน (robustness) ในการเก็บหรือคำนวณค่าสัมประสิทธิ์เวฟเล็ด  $\langle f, \psi_{m,n} \rangle$  ด้วยค่าความเที่ยงตรงต่ำ (low precision) แล้วสามารถนำค่าสัมประสิทธิ์เหล่านี้กลับมาคำนวณหา  $f$  ได้ด้วยค่าความเที่ยง

ตรงสูงเมื่อเทียบกับการเก็บหรือคำนวณค่าสัมประสิทธิ์เวฟเล็ด  $\langle f, \psi_{m,n} \rangle$  ด้วยความซ้ำซ้อนต่ำ [8]

ในการแปลงเวฟเล็ดชนิดไม่ต่อเนื่องนั้น ถ้าค่าความซ้ำซ้อนมีค่าเหมาะสม จะทำให้การแปลงกลับมีเสถียรภาพ กล่าวคือสามารถคำนวณหา  $f$  จากค่าสัมประสิทธิ์เวฟเล็ด  $\langle f, \psi_{m,n} \rangle$  ได้ถูกต้องมากขึ้น แต่ถ้ามีความซ้ำซ้อนต่ำ (นั่นคือตระกูลเวฟเล็ดแทบจะไม่ซ้อนทับกันเลย) ค่าสัมประสิทธิ์เวฟเล็ด  $\langle f, \psi_{m,n} \rangle$  จะไม่สามารถเก็บข้อมูล (information) หรือลักษณะ (characteristic) ของ  $f$  ได้ทั้งหมด ยังผลให้การคำนวณ  $f$  ในการแปลงกลับจากค่าสัมประสิทธิ์เวฟเล็ดมีความผิดพลาดสูงจนกระทั่งไม่สามารถยอมรับได้

สำหรับการแปลงกลับแบบเวฟเล็ดชนิดไม่ต่อเนื่องทำได้ดังนี้

$$f(t) = k \cdot \sum_m \sum_n \{CTWS_\psi(m,n) \cdot \psi_{m,n}(t)\} + R \quad (2-13)$$

โดยค่า  $R$  คือค่าความผิดพลาดจากการแปลงกลับ ค่า  $k$  เป็นค่าคงที่ ๆ ขึ้นกับค่าความซ้ำซ้อนในการแปลง ซึ่งมีค่าเป็น

$$k = \frac{2}{A+B} \quad (2-14)$$

โดยค่า  $A > 0$  และ  $B < \infty$  มีชื่อว่า 'ขอบเฟรม' (frame bounds) คำนวณได้จาก

$$A \cdot \|f\|^2 \leq \sum_{m,n} |\langle f, \psi_{m,n} \rangle|^2 \leq B \cdot \|f\|^2 \quad (2-15)$$

และ  $\{\psi_{m,n}\}$  ถูกเรียกว่า 'เฟรม' [8] อย่างไรก็ตามค่า  $k$  ดังกล่าวมักถูกละเลยในการใช้งานเนื่องจากเป็นเพียงตัวคูณเพื่อเพิ่มขนาดสัญญาณเท่านั้น

ในกรณีของการแปลงหลายมิตินั้น ค่า  $a$  และ  $b$  ในตระกูลเวฟเล็ด (2-8) ถูกแทนที่ด้วยตัวแปรจำนวนเต็ม  $m \in Z$  และตัวแปรเมตริกซ์ทแยง (diagonal matrix) ที่มีสมาชิกเป็นจำนวนเต็ม  $N = \text{diag}(n_1, n_2, \dots, n_n); n_i \in Z$  คล้ายกับกรณีฟังก์ชันที่มีมิติเดียวดังนี้  $a = a_0^m, a_0 > 1$  และ  $b = N \cdot b_0 \cdot a_0^m; b_0 \in \mathcal{R}^n$  ซึ่งจะได้สูตรการแปลงและการแปลงกลับชนิด

ไม่ต่อเนื่องจากกรณีมิติเดียว (2-12) และ (2-13) เป็นกรณีหลายมิติ ดังสมการที่ (2-16) และ (2-17) ตามลำดับ

$$CTWS_{\psi} f(m, \mathbf{N}) = \langle f, \psi_{m, \mathbf{N}} \rangle \quad (2-16)$$

$$f(\mathbf{t}) \cong k \cdot \sum_m \sum_{\mathbf{N}} CTWS_{\psi} f(m, \mathbf{N}) \cdot \psi_{m, \mathbf{N}}(\mathbf{t}) \quad (2-17)$$

### โครงข่ายเวฟเล็ต

โครงข่ายเวฟเล็ต (Wavelet Networks) เป็นการดัดแปลงวิธีการแปลงเวฟเล็ตให้สามารถใช้แทนโครงข่ายประสาทเทียม (Artificial Neural Networks) ในการประมาณค่าฟังก์ชันไม่เชิงเส้นใด ๆ ดังที่ Zhang กับ Benveniste [5] ได้เสนอไว้ และได้กล่าวถึงไว้ดังนี้

กำหนดให้เซต

$$\Phi = \left\{ \begin{array}{l} \det(\mathbf{D}_k)^{\frac{1}{2}} \cdot \psi(\mathbf{D}_k \cdot (\mathbf{x} - \mathbf{t}_k)) : \mathbf{t}_k \in \mathfrak{R}^n, \\ \mathbf{D}_k = \text{diag}(\mathbf{d}_k), \mathbf{d}_k \in \mathfrak{R}^n, k \in Z \end{array} \right\} \quad (2-18)$$

มีคุณสมบัติสอดคล้องกับ

$$A \cdot \|f\|^2 \leq \sum_{\phi \in \Phi} |\langle f, \phi \rangle|^2 \leq B \cdot \|f\|^2 \quad (2-19)$$

เมื่อ  $A > 0$  และ  $B < \infty$  สำหรับทุก ๆ ฟังก์ชัน  $f(\mathbf{x}) \in L^2(\mathfrak{R}^n)$  จะได้ว่าเซตดังกล่าวจะหนาแน่น (dense) ใน  $L^2(\mathfrak{R}^n)$  กล่าวคือ ฟังก์ชันใน  $L^2(\mathfrak{R}^n)$  ใด ๆ สามารถแทนได้ด้วยผลรวมถ่วงน้ำหนักเชิงเส้น (linear weighted combination) ของ  $\phi \in \Phi$  ซึ่งจาก

$$f(\mathbf{x}) = C_{\psi}^{-1} \cdot \int \int_{\mathfrak{R}^n \mathfrak{R}^n} (\det(\mathbf{D}))^{\frac{1}{2}} \cdot \psi(\mathbf{D} \cdot (\mathbf{x} - \mathbf{t})) \cdot (H(\mathbf{d}, \mathbf{t}) d d d t) \quad (2-20)$$

$$H(\mathbf{d}, \mathbf{t}) = \int_{\mathfrak{R}^n} f(\mathbf{x}) \cdot \det(\mathbf{D})^{\frac{1}{2}} \cdot \psi(\mathbf{D} \cdot (\mathbf{x} - \mathbf{t})) \cdot dx \quad (2-21)$$

โดยที่  $\mathbf{D} = \text{diag}(\mathbf{d})$  จึงทำให้สามารถประมาณ  $f(\mathbf{x})$  ได้เป็น



$$f(\mathbf{x}) \cong g(\mathbf{x}) = \sum_{i=1}^N \omega_i \cdot \det(\mathbf{D}_i)^{\frac{1}{2}} \cdot \psi(\mathbf{D}_i \cdot (\mathbf{x} - \mathbf{t}_i)) \quad (2-22)$$

โดย  $N$  เป็นจำนวนเต็มบวกซึ่งจะถูกเรียกว่า ‘จำนวนเวฟลอน’ (number of wavelons) หมายถึงจำนวนพจน์ที่ใช้ในการประมาณ และแต่ละพจน์  $\omega_i \cdot \det(\mathbf{D}_i)^{\frac{1}{2}} \cdot \psi(\mathbf{D}_i \cdot (\mathbf{x} - \mathbf{t}_i))$  ถูกเรียกว่า ‘เวฟลอน’ ยิ่ง  $N$  มีค่ามากพอ ค่าผิดพลาดในการประมาณจะยิ่งต่ำ และในกรณีที่ฟังก์ชัน  $f(\mathbf{x})$  มีค่าเฉลี่ยที่ไม่เป็นศูนย์ จึงมีการเพิ่มค่าคงที่เข้าไปดังนี้

$$g(\mathbf{x}) = \sum_{i=1}^N \omega_i \cdot \det(\mathbf{D}_i)^{\frac{1}{2}} \cdot \psi(\mathbf{D}_i \cdot (\mathbf{x} - \mathbf{t}_i)) + \bar{g} \quad (2-23)$$

และเพื่อช่วยให้มีการจัดเรียงตัวที่ดีขึ้นของเวกเตอร์  $\mathbf{x} - \mathbf{t}_i$  จึงเพิ่มเมตริกซ์การหมุน (rotation matrix)  $\mathbf{R}_i$  ทำให้สมการที่ (2-23) กลายเป็น

$$g(\mathbf{x}) = \sum_{i=1}^N \omega_i \cdot \det(\mathbf{D}_i)^{\frac{1}{2}} \cdot \psi(\mathbf{D}_i \cdot \mathbf{R}_i \cdot (\mathbf{x} - \mathbf{t}_i)) + \bar{g} \quad (2-24)$$

ซึ่งเป็นการประมาณฟังก์ชัน  $f(\mathbf{x})$  ด้วย  $g(\mathbf{x})$  รูปที่ 2-3 แสดงโครงสร้างของโครงข่ายเวฟเล็ด

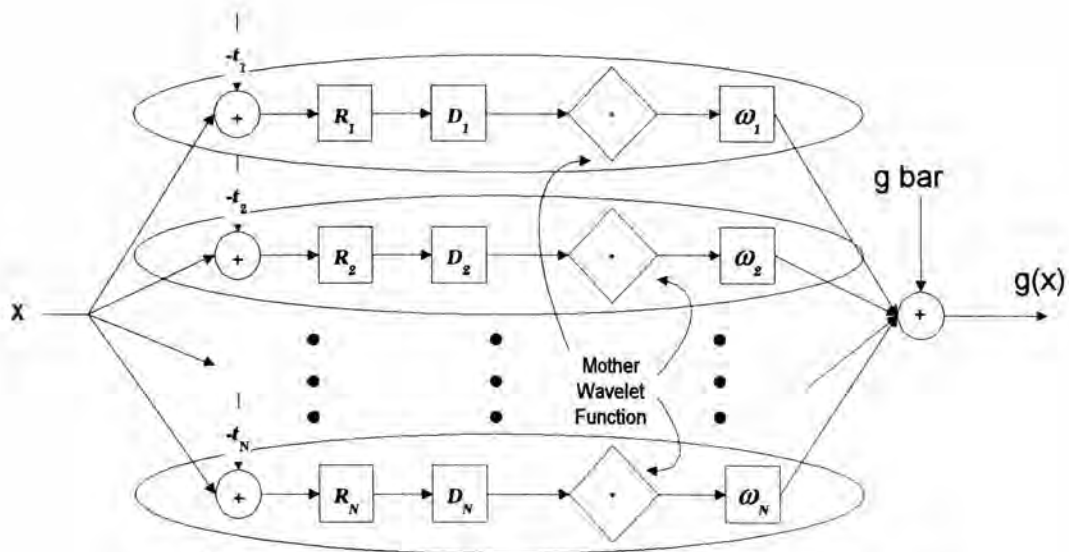
ในการใช้งานโครงข่ายเวฟเล็ดนั้น จำเป็นต้องปรับค่าพารามิเตอร์ต่าง ๆ อันได้แก่  $\omega_i$ ,  $\mathbf{D}_i$ ,  $\mathbf{R}_i$ ,  $\mathbf{t}_i$  และ  $\bar{g}$  เพื่อให้ผลการประมาณ  $g(\mathbf{x})$  มีค่าใกล้เคียงกับฟังก์ชัน  $f(\mathbf{x})$  มากที่สุด ซึ่งการปรับค่าพารามิเตอร์ดังกล่าวคล้ายกับการเรียนรู้ของโครงข่ายประสาทเทียมดังจะกล่าวถึงต่อไป

### การเรียนรู้ของโครงข่ายเวฟเล็ด

ในการเรียนรู้ของโครงข่ายเวฟเล็ดนั้น สามารถแบ่งได้เป็นสองขั้นตอนดังนี้

#### 1. การเรียนรู้ก่อนการใช้งานจริง

การเรียนรู้ก่อนการใช้งานจริงนั้น เป็นการปรับพารามิเตอร์ของโครงข่ายไปที่ค่าเริ่มต้นค่าหนึ่งในการที่จะป้อนสัญญาณควบคุมแก่กระบวนการเพื่อให้ฟังก์ชันต้นทุนมีค่าต่ำสุด โดยใช้สัญญาณอ้างอิงหนึ่งในช่วงเวลาชักตัวอย่างที่กำหนด ทั้งนี้สัญญาณอ้างอิงที่ใช้ควรมีความสามารถในการกระตุ้นกระบวนการให้มีการกระจายของข้อมูลให้มากเพียงพอกับสภาพใช้งานจริง



รูปที่ 2-3 โครงสร้างของโครงข่ายเวฟเล็ตในลักษณะ block diagram

การปรับค่าพารามิเตอร์ของโครงข่ายก่อนการใช้งานจริงนั้น สามารถใช้วิธี 'ออปติไมเซชัน' (optimization) โดยทั่วไปเช่น Steepest Descent, Conjugate Gradient, Newton Method หรือ ฯลฯ ได้ตามปกติ

Zhang และ Benveniste ได้เสนอการใช้ back-propagation ในการปรับค่าพารามิเตอร์ อย่างไรก็ตาม วิทยานิพนธ์นี้ได้ทดลองใช้วิธี Steepest Descent และ Conjugate Gradient โดยมีฟังก์ชันต้นทุน (cost function) เป็นครึ่งหนึ่งของผลรวมของค่าผิดพลาดยกกำลังสองของการประมาณ ณ จุดข้อมูลที่ต้องการให้โครงข่ายเรียนรู้ดังนี้

$$J(\theta) = \frac{1}{2} \cdot \sum_{k=1}^l (y_k - g_\theta(\mathbf{x}_k))^2 \quad (2-25)$$

เมื่อ  $l$  คือจำนวนคู่ของข้อมูล  $\{(\mathbf{x}_k, y_k), y_k = f(\mathbf{x}_k)\}$  และ  $\theta$  เป็นพารามิเตอร์ทั้งหมดที่ต้องปรับในการเรียนรู้ของโครงข่าย อันได้แก่  $\omega_i, D_i, R_i, t_i$  และ  $\bar{g}$  นั้นเอง

จาก (2-25) สามารถประยุกต์ใช้โครงข่ายในระบบควบคุมโดยการปรับค่าพารามิเตอร์ของโครงข่ายซึ่งมีฟังก์ชันต้นทุนเป็น

$$J(\theta) = \frac{1}{2} \cdot \sum_{k=1}^m (\|e(k)\|^2) \quad (2-26)$$

เมื่อ  $m$  คือจำนวนขั้นเวลาที่ใช้ในการเรียนรู้ และ  $\mathbf{e}(\bullet)$  เป็นค่าความผิดพลาดที่ต้องการให้มีค่าน้อยที่สุด ณ เวลาชกตัวอย่างต่าง ๆ ทั้ง  $m$  ค่า

## 2. การเรียนรู้ในขณะใช้งาน

ในการปรับค่าพารามิเตอร์ในขณะใช้งานนั้น ไม่สามารถใช้วิธีการออปติไมเซชันตามปกติได้ วิทยาพนธ์นี้จึงเสนอวิธีโมเมนตัม (Momentum) และ วิธีเดลต้าบาร์เดลต้า (Delta Bar Delta: DBD) ซึ่งเสนอโดย Jacobs [9] ดังจะกล่าวถึงต่อไปในภาคผนวก โดยมีฟังก์ชันต้นทุนเป็น

$$J_{On-line}(k) = \frac{1}{2} \cdot (y(k) - g_{\theta(k)}(\mathbf{x}(k)))^2 \quad (2-27)$$

ในที่นี้  $\{(\mathbf{x}(k), y(k)); k \in Z, k > 0\}$  เป็นคู่ของข้อมูลที่ใช้ในการปรับโครงข่าย และ  $\theta(k)$  เป็นค่าพารามิเตอร์ ณ เวลาชกตัวอย่างที่  $k$

เช่นเดียวกับการเรียนรู้ก่อนการใช้งานจริง การเรียนรู้ในขณะใช้งานนั้นก็สามารปรับใช้ในระบบควบคุมได้ในลักษณะคล้ายกัน กล่าวคือ ในขณะใช้งาน จะต้องปรับค่าพารามิเตอร์ของโครงข่ายเพื่อให้ฟังก์ชันต้นทุนต่อไปนี้มีค่าต่ำสุด

$$J_{On-line}(k) = \frac{1}{2} \cdot (\|\mathbf{e}(k)\|^2) \quad (2-28)$$

ซึ่งจะกล่าวถึงโดยละเอียดในบทต่อไป