

웨이블렛 신경망의 성장 알고리즘

Growing Algorithm of Wavelet Neural Network

서재용*, 김성주*, 김성현**, 김용민***, 전홍태*

* 중앙대학교 전자전기공학부, ** 동원대학 전자공학과, *** 충청대학 전자공학과

Jae Yong Seo*, Seong Joo Kim*, Seong Hyun Kim**,
Yong Min Kim*** and Hong Tae Jeon*

*School of Electrical and Electronics Eng., Chung-Ang Univ.

**Dept. of Electronic Eng., Tongwon College

***Dept. of Electronic Communications, Chung Cheong College
(sjyong@ms.cau.ac.kr)

ABSTRACT

In this paper, we propose growing algorithm of wavelet neural network. It is growing algorithm that adds hidden nodes using wavelet frame which approximately supports orthogonality in wavelet neural network based on wavelet theory. The result of this processing can be reduced global error and progresses performance efficiency of wavelet neural network. We apply the proposed algorithm to approximation problem and evaluate effectiveness of proposed algorithm.

Keywords : Wavelet Neural Network, Growing Algorithm, Wavelet Frame, Function Approximation

I. 서론

웨이블렛 신경망은 RBFN의 기저함수들을 웨이블렛 함수(wavelet function)로 대체하여 직교성을 갖는 기저함수를 사용한다. 웨이블렛 신경망의 기저함수를 결정하는데 다중분해 해석이나 웨이블렛 이론들을 이용하여 알려지지 않은 함수에 대한 효율적이고 고유한 망구성이 가능하며, RBFN의 대부분의 장점을 유지할 수 있는 장점들이 있다. 이러한 장점 때문에 많은 응용사례가 보고되고 있다 [4][5]. 그러나 웨이블렛 함수들은 강한 직교 조건을 만족해야하며 이러한 직교 조건은 망의 유연성을 저해하는 요소로 작용하고, 또한 시간-스케일의 지역화에는 낮은 성능을 나타내는 단점으로 작용한다[1][2]. 이러한 문제점을 해결하기 위해 직교성을 갖는 웨이블렛 함수 대신에 망에 유연성을 제공하고 직교 웨이블렛 함수와 유사한 알고리즘을 제공하는 웨이블렛 프레임(Frame)이 많이 사용된다[3][6].

웨이블렛 신경망의 많은 장점에도 불구하고 대상 문제에 적합한 웨이블렛 신경망의 최적구조 설계 및 초기 파라미터 값의 설정 문제는 아직도 연구과제로 남아 있다.

따라서 본 논문에서는 이러한 단점을 극복할 수 있는 웨이블렛 신경망의 성장 알고리즘을 제안한다. 제안한 성장 알고리즘은 직교 웨이블렛을 사용하지 않고 망의 구성에 유연성을 제공하는 웨이블렛 프레임을 사용하며, 추가하는 웨이블렛 프레임의 근사적인 직교성을 보장하기 위해 F-투영(projection)법[7]을 사용하여 이동 단

의 크기(translation step size)를 결정한다. 초기 한 개의 노드로 시작하고 시스템의 전체에러를 감소시키기 위해 이동계수를 자동으로 결정하여 노드를 추가한다. 원하는 수행능력에 도달할 때까지 웨이블렛 프레임의 파라미터와 가중치의 초기 값을 결정하여 웨이블렛 신경망을 성장시킨다. 제안한 알고리즘을 함수근사화 문제에 적용하여 수행능력을 검증하였다.

II. 본론

2.1 기하학적인 성장 기준

데이터를 연속적으로 입력받는 연속추정 문제에서 망은 대상함수를 충분한 정확도로 근사화 할 수 있지만 최종추정에 기여도가 적은 은닉층의 노드를 계속 추가할 수 있다. 따라서 망의 복잡도가 불필요하게 증가하게 되고 계산속도가 느려지게 된다. 더욱이 데이터에 노이즈가 존재할 경우 좋은 추정기가 되기 위해서는 적은 수의 노드를 이용해서 우수한 매핑결과를 얻을 수 있어야 한다. 결국 망의 성장에 제약이 필요하다.

Hilbert 공간 H 를 고려하자. 망의 해는 무한차원 공간으로 확장될 수 있다. 만약 망이 K 개의 은닉층 노드 (K 개의 기저함수)로 구성되어 있고, 기저함수의 파라미터가 고정된다면 망의 해는 K 개의 기저함수로 형성된 K 차원의 H_K 의 부분공간에 존재하게 된다. 그림 1

은 3차원적으로 망의 해를 표현한 것이다[7].

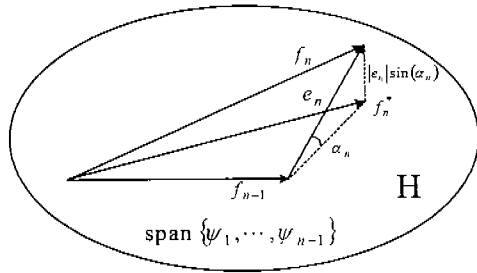


그림 1. Hilbert 공간에서 망의 해를 표현

사전추정 f_{n-1} 과 K 개의 기저함수를 이용해서 매핑한 사후추정 f_n^* 은 모두 H_k 공간에 존재하게 된다. 또한 사후추정 f_n 은 새로운 기저함수 ψ_n 을 추가해서 얻을 수 있다. f_n^* 은 H_k 에서 f_n 의 투영이기 때문에 f_n 에 가장 가까운 점이 된다. f_n 과 f_n^* 사이의 거리는 $\|f_n - f_n^*\|$ 로 정의할 수 있다. 이러한 거리는 새로운 기저함수를 추가하지 않았을 경우 잘못된 근사화 정도를 의미하기도 한다. 따라서 은닉층의 노드를 추가하기 위한 기준으로 이러한 거리가 특정한 문턱치(threshold)를 초과하는지 여부로 판단할 수 있다. 그림 1과 같이 망 해의 기하학적인 모양에서 새로운 은닉층 노드를 추가하기 위한 기준을 다음과 같이 표현할 수 있다.

$$\|f_n - f_n^*\| = |e_n| \sin(\alpha_n) > \epsilon \quad (1)$$

여기서 ϵ 은 문턱치, $|e_n| = \|f_n - f_{n-1}\|$, α_n 는 f_{n-1} 에서 K 개의 기저함수에 의해 정의된 부공간 H_K 에 새로운 기저함수 ψ_n 을 추가하여 얻어진 각도를 나타낸다. 기하학적인 기준은 다음과 같이 단순화해서 생각할 수 있다.

$$e_n > e_{\min} \quad (2)$$

$$\alpha_n > \alpha_{n_{\max}} \quad (3)$$

또한 이러한 기준은 예측 에러 기준(prediction error criterion)과 각도 기준(angle criterion)으로 생각할 수 있다. 일반적으로 각도 α_n 은 계산하기가 어렵기 때문에 새로운 기저함수와 이미 존재하는 모든 기저함수 사이의 가장 작은 각도가 임의의 문턱치를 초과하도록 각도 기준치를 근사적으로 선택하여 사용한다.

웨이블렛 신경망의 경우, $\{\psi_1, \dots, \psi_{n-1}\}$ 의 공간에 존재하는 다음 식과 같은 $\cos(\alpha_n)$ 을 계산하여 α_n 을 얻을 수 있다.

$$\cos(\alpha_n) = \sup_{i=1, \dots, n-1} \left\{ \left| \frac{\langle \psi_n, \psi_i \rangle}{\|\psi_n\| \|\psi_i\|} \right| \right\} \leq 1 - \delta \quad (4)$$

여기서 $\delta \leq 1$ 이다. 위의 식에서 각도 α_n 을 계산하기 위해 $\langle \psi_n, \psi_i \rangle$ 을 고려한다.

웨이블렛 프레임은 일반적으로 많이 사용되고 있는 Mexicanhat 웨이블렛 함수

$$(1 - x^2)e^{-\frac{x^2}{2}} \quad (5)$$

을 사용한다. 식 (4)에서 내적의 계산 결과는 $\forall j \in \{1, \dots, n-1\}$ 일 때 다음과 같다.

$$\langle \psi_{ij}, \psi_m \rangle = e^{-\frac{a^2 d_j^2}{2}} \left(1.0 - a^{2m} d_j^2 + \frac{1}{12} a^{4m} d_j^4 \right) \quad (6)$$

여기서 $d_j = x_{ji} - x_m$ 이다.

이 결과에서 만약 영점 교차점에서 거리 d 를 유지할 수 있다면 각도 $\alpha_n = 90^\circ$ 를 얻을 수 있고, 결국 두 노드사이의 직교성을 보장할 수 있다.

이러한 개념을 바탕으로 다음과 같은 조건을 만들 수 있다.

$$|d| > d_{\min}(m) \quad (7)$$

여기서 $d_{\min}(m)$ 은 팽창계수 m 이 주어지면 고정되는 값이다.

2.2 웨이블렛 신경망의 성장을 위한 지관적인 방법론(Heuristic Methodology)

가장 인접한 노드의 거리와 직교성을 유지하기 위해 정의한 식 (7)과 비교하여 on-line 상에서 웨이블렛 신경망의 노드를 추가할 수 있는 하나의 근거를 만들 수 있다. 성장 알고리즘의 단계별 구성은 다음과 같다. 각 차원은 독립적으로 구분하여 고려하고, 각 차원별로 동일한 단계를 반복한다.

[단계 1] 초기화

먼저 임의의 두 개의 노드 의해 표현되는 완전한 공간영역 $[x_u(l), x_v(l)]$ 을 포함하도록 주파수의 하한극한(lower limit) 값 $L(l)$ 을 정의하기 위해서 두 노드간의 폭을 다음과 같은 식으로 표현할 수 있다.

$$L(l) = -\frac{\log(x_u(l) - x_v(l)) - \log(b(l))}{\log(a)} \quad (8)$$

식 (8)을 초기 팽창계수로 다음과 같이 정의하고

$$m(l) = L(l), \quad (9)$$

팽창계수가 $m(l)$ 일 때 준직교성을 유지할 수 있는 두 노드사이의 거리 $d(l)$ 은

$$d(l) = \frac{a^{-m(l)} b(l)}{2} \quad (10)$$

로 정의한다. 또한 첫 번째 노드(이동계수)는 다음 식을 이용하여 정수에 가장 가까운 값을 선택한다.

$$n_{first(l)} = \frac{x(l)a^{L(l)}}{b(l)} \quad (11)$$

위의 초기화 단계에서 결정된 WNN을 이용하여 학습을 수행한다.

[단계 2] 새로운 노드 추가

새로운 이동계수를 가진 노드를 추가하기 위해 학습 데이터 t 에서 이미 존재하는 노드들 중에서 가장 인접한 노드의 거리 $d'_n(l)$ 을 계산한다. 만약 이전 거리와 비교하여 $d_n(l) > d(l)$ 이고, $|y'_d - y^t| > \epsilon_{MAE}$ 이면 다음과 같은 이동계수를 갖는 새로운 노드를 추가한다.

$$n(l) = \frac{x(l)a^{m(l)}}{b(l)} \quad (12)$$

여기서 ϵ_{MAE} 는 수행목표치를 의미한다.

[단계 3] 새로운 노드의 가중치 선택

새로 추가된 노드에 연결되는 가중치는 현재 입력되는 학습 데이터에 의해 발생하는 에러와 웨이블릿 함수의 출력을 고려하여 다음과 같은 식을 이용하여 결정한다.

$$w_{new} = \frac{|y'_d - y^t|}{\psi_p(x)} \quad (13)$$

여기서 팽창계수가 $m(l)$ 이고, 이동계수가 $n(l)$ 일 때 $\psi_p(x) = \psi_1(x) \cdots \psi_n(x)$ 이고, 분자항인 에러의 절대값은 웨이블릿 팽창계수를 고려하여 입력공간에 적용되는 폭을 결정하여 평균 에러를 사용한다.

[단계 4] 학습단계

현재 주어진 학습데이터를 이용하여 각 차원별로 [단계 2]와 [단계 3]을 반복 수행한다. 만약 더 이상 추가되는 노드가 없을 경우 역전파 학습 알고리즘이나 LMS 학습 알고리즘을 이용하여 가중치를 조절한다. 웨이블릿 신경망이 원하는 목표치의 출력을 내도록 학습을 수행한다. 만약 더 이상 에러가 감소하지 않을 경우 전체 에러가 원하는 목표치에 도달하는지 여부 ($|y_d - y| < \epsilon_{th}$)를 검사하여 만약 도달하지 않을 경우 다음 단계를 수행하고, 전역에러가 원하는 목표치에 도달하면 성장 알고리즘을 종료한다.

[단계 5] 새로운 팽창계수 정의

만약 웨이블릿 신경망이 충분히 학습되어 더 이상 에러가 감소하지 않을 경우 새로운 팽창계수를 다음과 같이 정의하고,

$$m(l) = m(l) + 1 \quad (13)$$

[단계 2]를 반복한다. 만약 웨이블릿 신경망의 출력이 원하는 문턱치 ϵ_{th} 에 도달하면 성장 알고리즘을 종료한다.

2.3 모의실험 및 검토

Mexicanhat 웨이블릿 함수에 대한 Daubechies의 계산결과[]를 참조하여 $a=2$ (최대 고주파 표현이 가능한 최소 정수 값)로 정의하고, $b \in (0.25, 1.875)$ [3]에 속하는 $b=1.0493$ 을 선택하였다. 근사화 대상식은 다음과 같은 연속함수인 Hermite 함수이고, 시간축 상에서 $[-4, 4]$ 를 고려하였다.

$$f(x) = 1.1(1 - x + 2x^2)e^{-\frac{x^2}{2}} \quad (14)$$

그림 3~5는 최대 절대치 에러(Maximum Absolute Error; MAE)인 $\epsilon_{MAE}=0.4$ 로 정의하고 성장 알고리즘을 적용하여 얻은 결과를 나타낸 것이다.

그림 3에서 웨이블릿 신경망에 추가된 웨이블릿 프레임의 공간-주파수 상에서 한정된 중심을 보인 것이다. 그림 4는 더 이상 노드가 추가되지 않고 성장 알고리즘을 종료한 후의 테스트 데이터를 이용하여 근사화를 수행한 결과이다. 그림 4는 성장 알고리즘이 적용되는 전체 반복동안 평균제곱근 에러(RMSE)와 추가된 노드 수를 동시에 나타내어 성장패턴을 보인 것이다.

그림 3과 그림 5에서 확인할 수 있듯이 2번째 성장 단계에서 추가된 3개의 노드는 근사화 대상함수의 복잡한 영역에서 실질적인 근사화에 실패하여 전체 에러를 감소시키는데 크게 기여하지 못함을 확인할 수 있다(그림 5 참조). 따라서 그림 3과 같이 3번째 성장단계에서 이전 단계에서 근사화를 달성하지 못한 영역에 새로운 노드가 자동으로 추가되어 전체 에러가 크게 감소하였다. 또한 지역에러(local error)가 ϵ_{MAE} 보다 작은 영역은 새로운 노드가 추가되지 않았다(그림 2 참조).

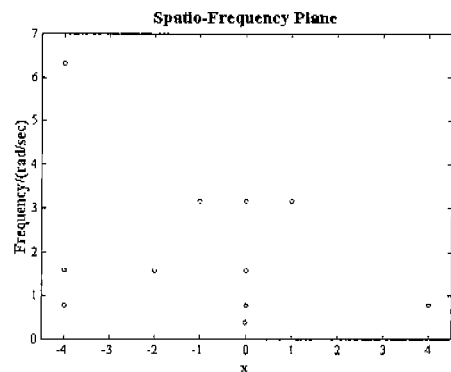


그림 2. 노드의 공간-주파수 분포(MAE : 0.4)

그림 2는 성장 알고리즘이 종료되고 테스트 데이터를 이용하여 근사화를 수행한 결과를 나

타낸 것이다. 결과에서 확인할 수 있듯이 함수의 시작영역에서 근사화가 잘 이루어지지 않았다.

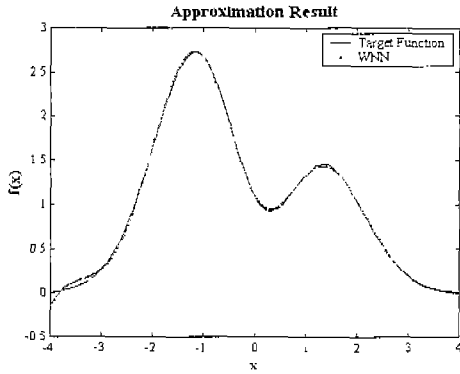


그림 3. 근사화 수행 결과(MAE : 0.4)

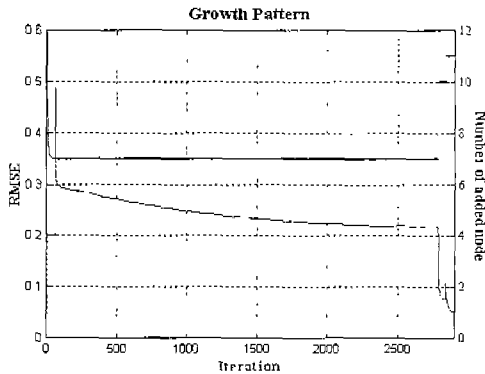


그림 4. 성장 패턴과 RMSE(MAE : 0.4)

또한 근사화 정확도를 높여 $\epsilon_{MAE} = 0.1$ 로 설정하고 모의실험을 수행하였다. 수행결과는 표 1에 나타내었다. 실험결과 $\epsilon_{MAE} = 0.4$ 인 경우보다 RMSE가 급격하게 감소함을 확인할 수 있다. 이러한 현상은 ϵ_{MAE} 가 적은 값으로 설정되어 입력공간 영역에 모든 노드가 추가되어 전체 에러를 감소하는데 기여했기 때문이다.

표 1은 최종 노드 수와 반복 횟수, RMSE를 비교하여 나타낸 것이다. 표에서 반복(iteration)은 그림 4에서 가로축과 같이 성장 알고리즘이 적용되는 전체의 수행 횟수를 나타낸다.

표 1. Hermite 함수에서 웨이블릿 신경망의 성장

MAE	최종 노드 개수	RMSE	Iteratio n	
Mexicanhat	0.4	11	0.054696	2905
	0.1	18	0.012759	1189

III. 결 론

기하학적인 방법을 사용하여 프레임의 직교기준을 구하고 근사적으로 직교하도록 이동단 크기를 결정하여

성장 알고리즘 적용시 프레임의 선택 기준으로 삼았다. 제안한 알고리즘을 연속함수, 부분적으로 연속인 1차원 함수와 2차원 함수의 근사화 문제어 적용하였을 경우 성공적인 수행을 검증하였다.

모의실험 결과에서 확인할 수 있듯이 웨이블릿 신경망의 성장과정에서 지역적인 에러를 보상하기 위해 필요한 웨이블릿 프레임이 자동으로 추가됨을 확인할 수 있다. 제안한 알고리즘의 장점은 원하는 수렴 한계치를 설계자가 임의로 선정할 수 있기 때문에 대상문제에 적합한 적응적인 웨이블릿 신경망의 설계가 가능하다는 것이다.

추후과제는 대상문제를 시스템 동정이나 제어 분야로 확대하여 적용분야를 확대하고 다양한 조건과 환경에서 제안한 알고리즘의 성능을 검증하는 것이다. 또한 다차원 문제에 적용할 경우 비교적 많은 노드가 사용되는 문제점을 해결하기 위한 하나의 방안으로 최적구조를 설계하기 위한 전지 알고리즘을 개발하는 것이다.

감사의 글 : 본 연구는 과학기술부의 뇌과학연구 프로그램에 의해 지원 받았습니다.

IV. 참고문헌

- [1] Simon Haykin, *Neural Networks - A Comprehensive Foundation*, Macmillian College Publishing Company Inc., 1994.
- [2] Y. C. Pati and P. S. Krishnaprasad, "Analysis and synthesis of feedforward neural network using discrete affine wavelet transformations," *IEEE Trans. on Neural Network*, vol. 4, pp. 73-85, Jan. 1993.
- [3] Ingrid Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trnas. on Information Theory*, Vol. 36, No. 5, pp. 961-1005, 1990.
- [4] Q. Zhang and A. Benveniste, "Wavelet networks," *IEEE Trans. Neural Networks*, vol. 3, pp. 889-898, Nov. 1992.
- [5] J. Zhang, G. G. Walter, Y. Miao and W. N. W. Lee, "Wavelet neural networks for function learning," *IEEE Trans. Signal Processing*, vol. 43, pp. 1485-1497, June. 1995.
- [6] Jaideva C.Goswami and Andrew K.Chan, *Fundamentals of Wavelets*, Wiley Interscience, 1999.
- [7] Visakan Kadiramanathan and Mehesan Niranja, "A Function Estimation Approach to Sequential Learning with Neural Networks," *Neural Computation*, Vol. 5, pp. 954-975, 1999.