유전 알고리듬을 이용한 웨이블릿 신경회로망의 최적 구조 설계

Optimal structure of wavelet neural network systems using genetic algorithm

이창민, 서재용, 전홍태
Chang-Min Lee, Jae-Yong Seo, Hong-Tae Jeon
중앙대학교 전기전자공학부 지능제어 연구실
E-Mail: lechami@ms.cau.ac.kr, sjyong@ms.cau.ac.kr

ABSTRACT

In order to approximate a nonlinear function, wavelet neural networks combining wavelet theory and neural networks have been proposed as an alterantive to coventional multi-layered neural networks. Wavelet neural networks provide better approximating performance than conventional neural networks. In this paper, an effective method to construct an optimal wavelet neural network is proposed using genetic algorithm. This is verified through experimental results.

1. 서론

최근에, 신경회로망(Neural Network)은 다소 복잡한 형태의 함수를 학습할 수 있는 능력 때문에 비선형 시스템의 제어에 인기있는 도구가 되었다. 그 중에서도 역전파 학습 알고리듬(Back Propagation Learning)을 이용한 다층퍼셉트론(Multi-Layer Perceptron) 구조의 신경회로망은 실제 응용에 있어서 자주 사용되는 신경회로망의 형태이다. 그러나, 다층화된 구조와 역전파 알고리듬의 성질 때문에 학습과정은 종종 바람직스럽지못한 지역 최소치에 빠지거나 수렴이 너무 느리게 되는 수가 있다. 이런 MLP의 대안으로 등장한 RBF(Radial Basis Function) 회로망은 더욱 간단한 (1개의 은닉층) 구조를 갖고 그 학습 역시 MLP보다는 더 쉬워진다.

함수 근사의 측면에서 볼 때, RBF 회로망은 지역성을 갖는 기저함수(basis function)를 사용하여 함수를 근사화하는 형태이다. 이런 기저함수의 지역성으로인하여 RBF 회로망은 지역적 변화와 불연속을 갖는 함수를 학습하는데 더 적합하다. 이러한 장점에도 불구하고 RBF 신경회로망은 기준함수들이 일반적으로 직교하지 않는다는 이유때문에 필요 이상의 기준함수들을 사용하게 되고 이것은 곧 효율적이고 고유한 망구성이 이루어지지 않을 수 있음을 의미한다. 이러한 문제들을 극복하기 위해 기존의 신경회로망과 웨이블릿 이론을 결합한 웨이블릿 신경회로망이 도입되었다[1][2]. 직교성을 갖는 웨이블릿 함수를 기준함수로 사용하는 웨이블릿 신경회로망은 RBF 신경회로망의 대부분의 장점을 유지함과 동시에, 알려지지 않은 함수에 대하여 효율적이고 고유한 망구성을 제공한다. 그러나 직교성을 갖는 웨이블릿 함수들은 강한 직교 조건을 만족해야하며 이러한 직교 조건은 망의 유연성을 저해하는 요소로 작용하기도 하며 또한 시간-스케일 지역화에는 낮은 성능을 나타낸다[1][3].

이에 본 논문에서는 직교성을 갖는 웨이블릿 함수 대신에 망에 유연성을 제공하면서도 직교 웨이블릿 함수와 비슷한 알고리듬을 제공하는 웨이블릿 프레임(Frame) 함수로 웨이블릿 신경회로망을 구성한 뒤, 망의 각 파라미터들을 적절히 구성하고 학습시키는 효율적인 방법을 제

안한다. 제안한 웨이블릿 신경회로망을 학습시키는 과정은 다음의 두과정으로 나눌수 있다. 유전 알고리듬을 이용하여 웨이블릿 신경회로망의 기준함수들을 최적으로 구성하는 과정과 이를 이용하여 기존의 역전파 학습 방법으로 가중치를 조절하는 과정이 그것이다. 기존의 MLP 신경회로망과 직교 웨이블릿 신경회로망과의 모의실험 비교를 통하여 제안한 방법의 유효성을 검증한다.

본 논문은 다음과 같이 구성되어있다. 2절,3절에서 웨이블릿 이론과 웨이블릿 신경회로망, 유전 알고리듬에 관하여 알아본 후, 4절에서는 유전 알고리듬을 이용한 웨이블릿 신경회로망의 구성 방법에 대하여 알아본다. 5절에서 임의의 비선형 함수에 대한 근사 실험 결과를 기술한다.

2. 웨이블릿 이론과 웨이블릿 신경회로망

웨이블릿 이론에 의하면, 다음과 같이 $L^2(R)$ 공간을 직교 분해할 수 있는 웨이블릿 함수 $\psi(t)$가 존재한다[3].

$$L^2(R) = \cdots \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus \cdots \quad (1)$$

$$\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n) \quad (2)$$

여기서 W_m은 $[2^{m/2}\psi(2^m t-n)]_{n=-\infty}^{n=+\infty}$으로 표시할 수 있는 부분공간이며, m,n은 정수이다.

웨이블릿 $\psi(t)$는 스케일링 함수라 알려진 $\varphi(t)$로부터 구할 수 있으며 스케일링 함수의 신축항과 이동항은 $L^2(R)$의 다중 분해 분석(multiresoulution analysis, MRA)을 가능케하고 다음과 같은 관계를 성립케한다[3].

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots \subset L^2 \quad (3)$$

여기서 V_m은 $[2^{m/2}\varphi(2^m t-n)]_{n=-\infty}^{n=+\infty}$으로 표시할 수 있는 부분공간이다.

또한, V_m은 다음과 같이 W_m과 연관되어진다.

$$V_{m+1} = V_m \oplus W_m \quad (4)$$

이로부터 $L^2(R)$공간의 임의의 함수 f(t)가 다음과 같이 분해됨을 알 수 있다[3].

$$f(t) = \sum_n \langle f, \varphi_{m_0,n}\rangle \varphi_{m_0,n}(t) + \sum_{m \ge m_0, n} \langle f, \psi_{m,n}\rangle \psi_{m,n}(t) \quad (5)$$

또는 $f(x) = \sum_{m,n} \langle f, \psi_{m,n}\rangle \psi_{m,n}(t)$ (6)

여기서 $\langle \cdot, \cdot \rangle$는 내적을 표시하고, m_0는 가장 낮은 분해능(resolution)을 표시하는 임의의 정수이다. 단일 차원에서의 이와 같은 결과는 다음과 같은 텐서(tensor) 곱 형태로 다차원으로 확장할 수 있다[1][2][3].

$$\psi_d(\boldsymbol{x}) = \prod_{j=1}^{d} \psi(x_j) \quad (7)$$

상기의 논의로부터 다음과 같이 웨이블릿 신경회로망을 구성할 수 있다.

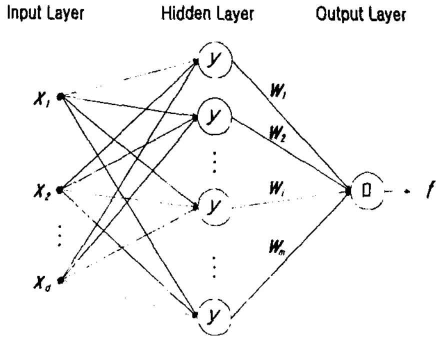


그림 1. 웨이블릿 신경회로망.

$$f(\boldsymbol{x}) = \sum_i w_i \psi_d(\boldsymbol{a}_i \boldsymbol{x} - \boldsymbol{t}_i) \quad (8)$$

여기서, $\boldsymbol{a} \in \boldsymbol{R}^d$, $\boldsymbol{t} \in \boldsymbol{R}^d$는 각각 신축과 이동 변수이고, $\boldsymbol{x} \in \boldsymbol{R}^d$는 웨이블릿 함수의 입력값이며, $\boldsymbol{w} = (w_1, w_2, \cdots, w_m)^T$는 가중치 벡터를 나타낸다. 한편, 직교 웨이블릿 함수는 앞서 서술한 바와 같이 실제 응용에 있어서 부적절한 면이 있으므로, 웨이블릿 프레임 함수를 사용한다. 프레임 함수의 정의는 다음과 같다[1].

정의 1: Hilbert 공간 $H(L^2(R^d))$과 벡터 $\{h_k\}_{k=-\infty}^{\infty} \subset H$의 수열이 주어졌을 때, 다음 식을 만족하는 상수 $A > 0$고 $B < 0$이 존재하면 $\{h_k\}_{k=-\infty}^{\infty}$를 프레임(frame)이라 부른다.

$$A\|f\|^2 \le \sum_k |\langle f, h_k\rangle|^2 \le B\|f\|^2 \quad (9)$$

여기서, $f \in H$이고, A와 B는 프레임 한계(frame bound)라고 부른다. 프레임 한계 $A = B$인 프레임 $\{h_k\}$는 엄격한(tight) 프레임이라 부른다. 만약 $A = B = 1$이면 엄격한 프레임은 직교 기준이 된다.

3. 유전 알고리듬

유전알고리듬은 자연생태계의 진화 매카니즘인 '적자 생존'의 원리를 모방하여, 최적화를

목적으로 만들어진 알고리듬이다. 유전 알고리듬을 문제해결에 적용시키기 위해서는 '염색체상에 문제의 해를 부호화 하는 방법' 과 '문제속에서 염색체의 가치를 측정하는 평가함수 설정 방법'의 2가지 매카니즘을 사용해야 한다. 첫번째의 경우 부호화하는 기술은 문제마다 다르며 부호화는 비트스트링(bit string) 으로 수행한다. 두번째의 경우 평가함수는 유전 알고리듬과 해결해야 할 문제를 결합시켜주며 또한 염색체를 입력으로 받아 해결할 문제에 대한 염색체의 적합도의 측정을 제공한다. 이러한 유전 알고리듬은 미지의 함수 Y = G(x)의 최적해를 발견하는 모의진화형의 탐색 알고리듬의 성격을 갖고 있으며, 또한 유전 알고리듬은 세가지의 주요 기능자 즉, 재생산, 교배, 돌연변이 기능자를 갖는다[4]. 본 논문에서 사용된 유전알고리듬의 구체적인 처리과정은 다음 절에서 설명한다.

4. 유전 알고리듬을 이용한 웨이블릿 신경회로망의 설계

본 논문에서 사용된 유전 알고리듬의 목표는 근사화 대상함수에 대하여 식(8)의 웨이블릿 신경회로망의 신축항과 이동항의 최적값을 찾아내는 것이다. 이를 위하여, 다음과 같이 유전알고리듬을 수행한다.

[단계 1] 필요한 신축항과 이동항을 모두 표현할수 있도록 스트링의 길이를 결정한 후 이를 2진수로 표현하여 초기 스트링을 구성한다.

[단계 2] *단계 1*의 초기 스트링을 사용하여 초기 집단을 구성한다.

[단계 3] 집단에 있는 각각의 스트링을 10진수로 변환하여 웨이블릿 신경회로망에 대입하여 각 집단의 적합도 값을 계산한다. 변환 시, 신축항의 최소값은 적은 수의 웨이블릿 기저 함수를 사용하여도 근사화 대상 함수의 영역을 충분히 포함할 수 있도록 설정하며 최대값은 대상 함수의 입력 데이타 분포에 맞게 적절히 설정한다. 또한, 이동항은 해당되는 신축항 값에 따라서 웨이블릿 기저 함수가 이동하여도 대상함수의 영역을 벗어나지 않도록 최대값, 최소값을 설정하여 변환되도록 한다. 각 집단의 적합도 값은 변환된 신축항과 이동항으로 웨이블릿 신경회로망을 고정시킨 후에 은닉층과 출력층 사이의 가중치 값을 역전파 학습 방법을 이용하여 학습하여 구한다. 역전파 학습시 너무 많은 반복 횟수는 많은 시간이 소요되므로 각 집단의 적합도 차이를 알 수 있는 선에서 학습을 중단한다. 적합도는 다음 식에 의하여 구한다.

$$\text{fitness} = \frac{1}{\sqrt{\sum_{i=1}^{n}(f(x_i)-g(x_i))^2/n}} \tag{10}$$

여기서 $f(x_i)$와 $g(x_i)$는 각각 i번째 입력의 근사화 대상함수와 웨이블렛 신경회로망의 출력이다.

[단계 4] 유전자 집단에서 높은 적합도 값을 갖는 스트링에 높은 선택 확률을 주고 유전자들 사이에서 교배나 돌연변이의 진화과정 을 통해 새로운 집단을 얻는다.

[단계 5] 새로운 최고의 스트링이나 원하는 적합도 값이 얻어질 때까지 상기 단계를 반복한다.

[단계 6] 최종적으로 얻어진 스트링으로 망을 고정시킨 후 은닉층과 출력층 사이의 가중치 값을 역전파 학습 방법을 이용하여 충분히 학습시켜 웨이블릿 신경회로망의 최종 형태를 완성한다.

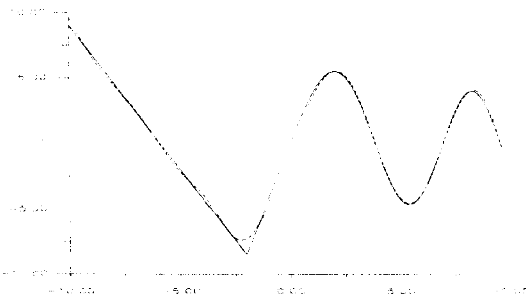
5. 모의 실험 결과

1) 1차원 함수: 웨이블릿 신경회로망에 사용된 웨이블릿 프레임 함수는 $\psi(x) = x\,e^{-\frac{1}{2}x^2}$이고 그림 2에서 다음 함수의 근사 결과를 보였다.

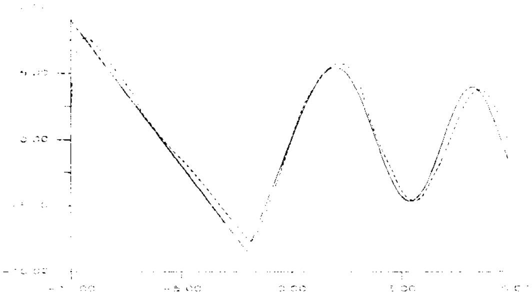
$$f(x) = \begin{cases} -2.186x - 12.864 & [-10 \le x < -2] \\ 4.246x & [-2 \le x < 0] \\ 10\,e^{-0.05x-0.5}\cdot\sin[(0.03x+0.7)x] & [0 \le x \le 10] \end{cases} \tag{11}$$

비교를 위하여 시그모이드 함수에 기반한 MLP 신경회로망과 직교 기준 스케일링 함수를 사용한(신축항과 이동항이 고정된) 웨이블릿 신경회로망의 결과도 함께 보였다. 결과에서 볼 수 있듯이 유전 알고리듬에 의하여 구성된 웨이블릿 신경회로망이 적은수의 뉴런으로도 더 좋은 근

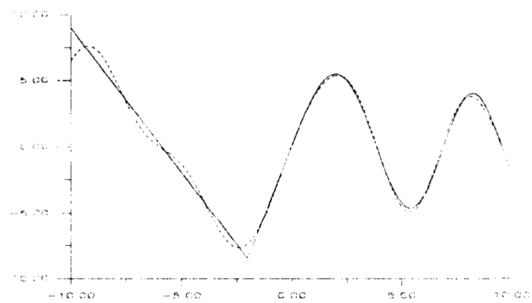
사 결과를 나타냄을 알 수 있다.



유전 알고리듬으로 구성된 8뉴런 웨이블릿 신경회로망의 근사 결과



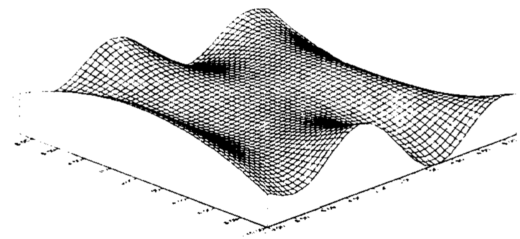
10뉴런 MLP 신경회로망의 근사 결과



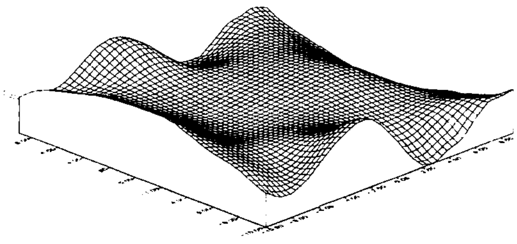
11뉴런 직교기준 웨이블릿 신경회로망의 근사 결과

그림2. 1차원함수의 근사 결과

2) 2차원 함수의 근사: 웨이블릿 프레임 함수로 1차 웨이블릿 프레임 함수의 텐서 곱 형태인 $\psi(x) = x_1 x_2 e^{-\frac{1}{2}(x_1^2 + x_2^2)}$이 사용되었으며, 근사대상 함수로 영역[-10,10]×[-10,10]사이에 있는 $f(x) = (x_1^2 - x_2^2)\sin(0.5x_1)$가 사용되었다. 직교기준 웨이블릿 신경회로망의 경우 121개의 뉴런이 사용되어야 유전 알고리듬으로 구성된 64 뉴런의 웨이블릿 신경회로망의 결과와 유사한 결과를 얻을 수 있음을 알 수 있었다.



2차원 근사화 대상 함수



64 뉴런 웨이블릿 신경회로망의 근사 결과

그림 3. 2차원 함수의 근사 결과

6. 결론

본 논문에서는 웨이블릿 신경회로망의 각 파라미터들을 최적으로 구성할 수 있는 방법을 제안하였다. 이 방법은 유전 알고리듬과 역전파 학습 방법의 병행으로 가능하며, 모의실험 결과에서 기존의 신경회로망들보다 우수한 성능을 나타냄을 알 수 있다. 그러나, 차원이 높아질수록 많은 수의 뉴런이 필요하며 이는 곧 유전 알고리듬이 탐색해야할 해공간이 늘어나게 되어 최적해를 구하는데 많은 시간과 함께 그 가능성도 떨어짐을 의미한다. 따라서, 향후에는 이런 문제점을 극복할 수 있는 방향으로의 연구가 필요하다.

7. 참고문헌

[1] Q. Zhang and A. Benveniste, "Wavelet networks," *IEEE Trans. Neural Networks*, vol. 3, pp. 889-898, Nov. 1992.
[2] 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.
[3] Jaideva C.Goswami and Andrew K.Chan, *Fundamentals of Wavelets*, Wiley Interscience, 1999.
[4] David E.Goldberg, *Genetic Algorithms in Search, Optimization,and Machine Learning*, Addison-Wesley Publishing, 1989.