

확장 칼만 필터를 이용한 웨이블릿 신경 회로망의 학습 방법에 관한 연구

*김경주, **최윤희, *박진배
*연세대학교 전기전자공학과, **경기대학교 전자공학부

A Study on Training Method of Wavelet Neural Network Using Extended Kalman Filter

*Kyung Ju Kim, **Yoon Ho Choi, *Jin Bae Park
*Department of Electrical and Electronic Engineering, Yonsei University.
**School of Electronic Engineering, Kyonggi University.

Abstract - 시간 및 주파수 특성 분석이 용이한 웨이블릿 변환을 네트워크화 시킨 웨이블릿 신경 회로망의 파라미터 학습 방법에는 오차 역전파 알고리즘 및 유전 알고리즘 등이 있으나 이러한 학습 방법들은 수렴 시간이 오래 걸리는 단점이 있다. 따라서 본 논문에서는 웨이블릿 신경 회로망의 최적 파라미터를 결정하기 위한 학습 방법으로 확장 칼만 필터 알고리즘을 제안한다. 일반적으로 확장 칼만 필터 알고리즘은 복잡한 연산 과정에도 불구하고 적은 학습 횟수로 빠른 수렴 특성을 가진다. 제안한 방법의 효율성을 검증하기 위해 확장 칼만 필터 학습 방법을 적용한 웨이블릿 신경 회로망을 혼돈 시스템 동정에 적용하여 경사 하강법을 사용한 기존의 신경 회로망에 비해 더 양호한 성능을 가짐을 검증하고자 한다.

분야에서 푸리에(Fourier) 해석의 단점을 보완하기 위해 제안되었다[4]. 웨이블릿(wavelet)은 이동(translation) 인자와 확장(dilation) 인자로 구성되며, 식 (1)과 같은 모체 웨이블릿(mother wavelet)함수로 표현된다.

$$\phi(z) = -z \exp(-\frac{1}{2}z^2) \quad \text{with } z = \frac{x-m}{d} \quad (1)$$

여기서, m 과 d 는 각각 이동, 확장 인자를 나타내며, 모체 웨이블릿은 가우시안(Gaussian) 함수의 1차 미분형이다. 그림 1은 N_x 개의 입력과 1개의 출력을 갖는 웨이블릿 신경 회로망의 구조를 나타낸다.

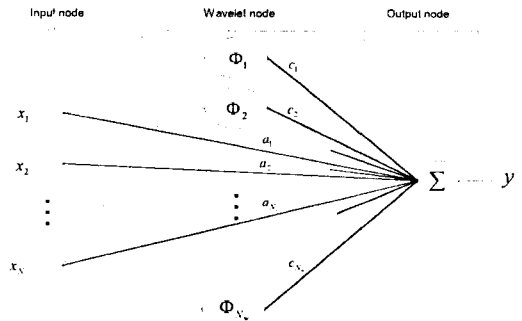


그림 1 웨이블릿 신경회로망 구조

그림에서 웨이블릿 노드의 각 웨이블릿은 식 (2)와 같이 각각의 입력 x_k 에 대한 모체 웨이블릿의 곱으로 표현된다.

$$\Phi_j(x) = \prod_{k=1}^{N_x} \phi(z_{jk}), \quad \text{with } z_{jk} = \frac{x_k - m_{jk}}{d_{jk}} \quad (2)$$

여기서, $k=1, \dots, N_x, j=1, \dots, N_x$ 이며, N_x 는 입력 수, N_x 는 웨이블릿 노드의 웨이블릿 수이다.

한편 그림에서 웨이블릿 신경 회로망의 출력 값은 식 (3)과 같이 각각의 입력에 가중치를 적용한 값과 웨이블릿 노드의 출력에 가중치를 곱한 값의 합으로 나타낸다.

$$y = \Psi(x, \theta) = \sum_{j=1}^{N_x} c_j \Phi_j(x) + \sum_{k=1}^{N_x} a_k x_k \quad (3)$$

여기서, a_k 는 입력 노드와 출력 노드 사이의 가중치 (weighting), c_j 는 웨이블릿 노드와 출력 노드 사이의 가중치를 나타내고 θ 는 학습을 통하여 갱신되는 웨이블릿 신경 회로망의 파라미터 집합이다.

$$\theta = \{a_k, c_j, m_{jk}, d_{jk}\} \quad (4)$$

1. 서 론

비선형 시스템에서 발생하는 혼돈 현상은 일반적으로 시스템의 초기 조건에 민감하고 예측 불가능하다. 이러한 혼돈 시스템을 제어하기 위한 기존의 제어 방법들[1]~[3]은 시스템의 정확한 수학적 모델이나 비선형 방정식을 필요로 하므로, 수학적으로 정확한 모델을 알 수 없는 시스템에 적용하기 힘든 단점이 있다. 이러한 기존 혼돈 제어 방법들의 단점을 극복하기 위해 파라미터 추정을 통해 동정 모델을 생성하여 제어하는 적응 제어 기법과 신경 회로망, 퍼지 이론 등이 도입 되었다. 신경 회로망의 경우 다층 신경 회로망(Multilayer Perceptron Neural Network)은 신경 회로망의 기본이라고 할 수 있으며, 최근에는 다양한 신경 회로망들이 연구되고 있다. 이와 같이 다양한 신경 회로망들 중에 본 논문에서 사용할 웨이블릿 신경 회로망(Wavelet Neural Network: WNN)은 다층 신경 회로망과는 달리 신경 회로망 내부에 웨이블릿 함수를 사용함으로써 적은 계층(layer) 수를 가지기 때문에 연산 시간을 줄이면서도 수렴 시간이 비교적 빠르며 급격한 신호의 변화도 잘 추정하는 특성을 가진다. 한편 이와 같은 웨이블릿 신경 회로망의 기존 학습 방법에는 경사 하강법 (gradient descent method), 유전 알고리즘 등의 방법들이 있으나, 이와 같은 학습 방법들은 일반적으로 수렴 속도가 느린 단점을 가진다. 따라서 본 논문에서는 빠른 수렴 특성을 갖는 칼만 필터 알고리즘을 웨이블릿 신경 회로망의 학습 방법으로 적용함으로써 수렴 속도가 빠른 신경 회로망 학습 방법을 제안한다. 한편 제안한 학습 방법의 효율성을 검증하기 위해 혼돈 시스템의 동정에 적용하여 경사 하강법을 사용한 기존의 신경 회로망에 비해 더 양호한 성능을 가짐을 검증하고자 한다

2. 본 론

2.1 웨이블릿 신경회로망

웨이블릿 이론은 다해상도 해석(multi-resolution analysis)

2.2 확장 칼만 필터

확장 칼만 필터(Extended Kalman Filter: EKF)는 비선형 확률 미분 방정식으로 주어지는 이산 모델을 추정할 때 유용하게 사용된다. 비선형 확률 미분 방정식으로 주어지는 상태 방정식은 식 (5)와 (6)같이 주어진다.

$$x_k = f(x_{k-1}, u_{k-1}, u_{k-1}) \quad (5)$$

$$z_k = h(x_k, v_k) \quad (6)$$

여기서, $x \in R^n$, $z \in R^m$ 이고, 랜덤 노이즈 변수 u_k 와 v_k 는 각각 프로세스 노이즈(process noise)와 측정 노이즈(measurement noise)를 나타낸다. $f(\cdot)$ 는 k 번째 단계인 현재 상태와 이전 단계인 $k-1$ 번째 상태의 관계를 나타낸다. $h(\cdot)$ 는 상태 x_k 와 출력 z_k 의 관계를 나타낸다.

실제 상황에서는 u_k 와 v_k 의 값을 정확히 알 수는 없지만, 식 (5)와 (6)은 식 (7) 및 (8)과 같이 근사화 될 수 있다[6].

$$\tilde{x}_k = f(\tilde{x}_{k-1}, u_{k-1}, 0) \quad (7)$$

$$\tilde{z}_k = h(\tilde{x}_k, 0) \quad (8)$$

여기서 \tilde{x}_k , \tilde{z}_k 는 각각 근사화된 상태벡터와 측정 벡터이다. 예측 오차 \hat{e}_x 와 측정 오차 \hat{e}_z 는 식 (9)와 같이 정의된다.

$$\hat{e}_x \equiv x_k - \tilde{x}_k \quad (9)$$

$$\hat{e}_z \equiv z_k - \tilde{z}_k$$

그러나, 식 (7)과 (8)은 근사화된 식이므로 원래의 상태 벡터는 식 (10)으로 나타난다.

$$x_k = \tilde{x}_k + \hat{e}_k \quad (10)$$

이때 \hat{e}_k 는 식 (11)과 같다.

$$\hat{e}_k = K_k \hat{e}_z \quad (11)$$

따라서, 식 (11)을 식 (10)에 대입하면 식 (12)가 된다.

$$\begin{aligned} \hat{x}_k &= \tilde{x}_k + K_k \hat{e}_z \\ &= \tilde{x}_k + K_k(z_k - \tilde{z}_k) \end{aligned} \quad (12)$$

오차를 최소화하는 칼만 이득 K_k 는 식 (13)과 같고,

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \quad (13)$$

이때 P_k 와 \hat{x}_k 는 각각 식(14)와 (15)로 나타난다.

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - h(\hat{x}_k^-, 0)) \quad (14)$$

$$P_k = P_k^- - K_k H_k P_k^- \quad (15)$$

여기서 P_k 는 예측 에러 분산을 나타내고, K_k 는 칼만 이득을 나타낸다.

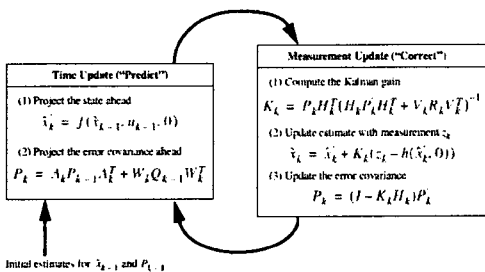


그림 2 확장 칼만 필터 연산 과정

확장 칼만 필터 예측 과정을 나타내는 완전한 식은 그림 2와 같다. 그림에서 A , W 는 $f(\cdot)$ 를 각각 상태 벡터 x 와 프로세스 노이즈 w 에 대해 편미분한 행렬이다. H , V 는 $h(\cdot)$ 를 각각 상태 벡터 x 와 측정 노이즈 v 에 대한 편미분 행렬이다.

2.3 신경 회로망 학습

칼만 필터 알고리즘을 적용하기 위해 웨이블릿 신경 회로망의 파라미터들의 집합 θ 를 상태 변수로 놓는다.

$$\theta = [a_1, \dots, a_k, c_1, \dots, c_j, \quad (16)$$

$$m_{11}, \dots, m_{jk}, d_{11}, \dots, d_{jk}]^T$$

벡터 θ 는 신경 회로망의 모든 파라미터들을 포함하고 있고, 시스템 모델은 식 (17)과 같은 칼만 필터 형태로 표현할 수 있다.

$$\theta(n) = \theta(n-1), n > 2 \quad (17)$$

$$\hat{y}(n) = h(\theta, v)$$

$$\begin{aligned} &= \sum_{j=1}^{N_x} c_{j,n} \Phi_{j,n}(x(n)) \\ &+ \sum_{k=1}^N a_{k,n} x_{k,n} + v(n) \end{aligned} \quad (18)$$

여기서, $v(n)$ 는 측정 노이즈를 나타낸다.

식 (17)과 (18)에서 정확한 상태 벡터 θ 를 추정하기 위해 확장 칼만 필터를 적용하면 식 (20)~(25)와 같다.

예측 과정 :

$$\hat{\theta}_n^- = \hat{\theta}_{n-1} \quad (20)$$

$$P_n^- = A_n P_{n-1} A_n^T + W_n Q_{n-1} W_n^T \quad (21)$$

측정 후 보정 과정 :

$$K_n = P_n^- H_n^T (H_n P_n^- H_n^T + V_n R_n V_n^T)^{-1} \quad (22)$$

$$\hat{y}_n = \Psi(x, \hat{\theta}_n^-) \quad (23)$$

$$\hat{\theta}_n = \hat{\theta}_n^- + K_n (y_n - \hat{y}_n) \quad (24)$$

$$P_n = P_n^- + K_n H_n P_n^- \quad (25)$$

한편 신경 회로망의 파라미터들은 서로 독립적이므로, 본 논문에서는 A , W , V 는 단위행렬로 놓는다. 또한 H , V 는 각각 y 의 θ , v 에 대한 자코비안(Jacobian) 행렬이다.

$$H = \frac{\partial y}{\partial \theta}(x, \theta_n, 0) \quad (26)$$

$$V = \frac{\partial y}{\partial v}(x, \theta_n, 0) \quad (27)$$

H 를 구하기 위한 파라미터 집합 θ 에 대한 시스템 출력 함수 y 의 편미분은 다음과 같이 한다.

$$\frac{\partial y(n)}{\partial a_k(n)} = x_k, k = 1, 2, \dots, N_x \quad (28)$$

$$\frac{\partial y(n)}{\partial c_j(n)} = \Phi_j(x), j = 1, 2, \dots, N_w \quad (29)$$

$$\begin{aligned} \frac{\partial y(n)}{\partial m_{jk}(n)} &= -\frac{c_j}{d_{jk}} \frac{\partial \Phi_j}{\partial z_{jk}}, \\ &k = 1, 2, \dots, N_x, j = 1, 2, \dots, N_w \end{aligned} \quad (30)$$

여기서 $\frac{\partial \Phi_j}{\partial z_{jk}} = \phi(z_{j1})\phi(z_{j2}) \cdots \phi'(z_{jk}) \cdots \phi(z_{jN})$ 이며,

$\phi'(z_{jk})$ 는 z_{jk} 에서의 스칼라 모체 웨이블릿의 z 에 대한 미분 값을 나타내며 $\phi'(z_{jk})$ 는 식 (31)과 같다.

$$\phi'(z_{jk}) = (z_{jk}^2 - 1) \exp(-0.5z_{jk}^2) \quad (31)$$

$$\frac{\partial y(n)}{\partial d_{jk}(n)} = -\frac{c_j}{d_{jk}} \frac{\partial \Phi_j}{\partial z_{jk}}, \quad (32)$$

$$k = 1, 2, \dots, N_j, j = 1, 2, \dots, N_w$$

3. 모의 실험

본 논문에서는 제안된 학습 방법의 성능을 평가하기 위하여 대표적인 연속 시간 혼돈 비선형 시스템인 Duffing 시스템에 대한 모의 실험을 수행하였으며, Duffing 시스템의 상태 방정식은 다음과 같다.

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \end{bmatrix} = \begin{bmatrix} y(t) \\ -p_1 x(t) - x^3(t) - p_2 y(t) + q \cos(ut) \end{bmatrix} \quad (33)$$

여기서, 파라미터 집합은 다음과 같다.

$$\{p_1, p_2, q, u\} = \{0.4, -1.1, 1.8, 1.8\} \quad (34)$$

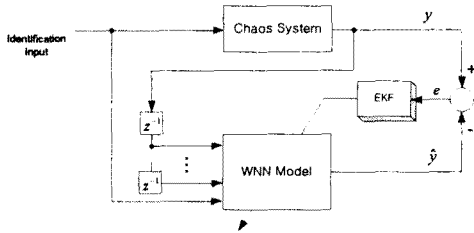


그림 3 웨이블릿 신경회로망 동정 구조

그림 3은 본 논문에서 제안한 확장 칼만 필터를 적용한 웨이블릿 신경 회로망의 동정 구조를 나타내고 있다. 웨이블릿 신경 회로망 모델의 입력은 혼돈 시스템의 출력의 과거 값을 사용한다.

모의 실험을 위해 경사 하강법을 이용한 웨이블릿 신경 회로망의 학습률 η 는 0.03으로 고정하였고, 확장 칼만 필터 알고리즘을 이용한 신경 회로망에서 측정 오차 분산 R 은 0.01로 고정하였다. 두 신경 회로망의 학습은 0.05초 간격으로 30초까지 수행하였다.

그림 4는 경사 하강법을 적용한 웨이블릿 신경 회로망의 Duffing 시스템 동정 결과를 나타내며, 그림 5는 확장 칼만 필터를 적용한 웨이블릿 신경 회로망의 동정 결과를 나타낸다. 확장 칼만 필터 알고리즘을 적용한 웨이블릿 신경 회로망이 혼돈 시스템의 동특성에 빨리 수렴하고 수렴 후에도 오차가 거의 없음을 알 수 있다. 그림 4와 5를 비교해 보면 확장 칼만 필터를 적용한 웨이블릿 신경 회로망이 더 빠른 수렴 특성을 보이고 있으며, 표 2에서 보이는 바와 같이 평균 제곱 오차도 경사 하강법을 적용한 경우보다 작음을 확인할 수 있다.

4. 결 론

본 논문에서는 확장 칼만 필터 알고리즘을 적용한 웨이블릿 신경 회로망의 새로운 학습 방법을 제안하였다. 제안된 학습 방법의 성능을 확인하기 위해 연속 시간 혼돈 시스템인 Duffing 시스템의 동정을 위한 웨이블릿 신경 회로망의 학습 방법에 적용하였다. 웨이블릿 신경 회로망 학습 방법으로 본 논문에서 제안한 확장 칼만 필터를 적용한 경우 기존의 경사 하강법을 적용한 경우에 비해 적은 회수의 학습으로도 빠른 수렴 특성을 가짐을 알 수

있었고, 또한 작은 평균 제곱 오차를 가짐을 확인할 수 있었다.

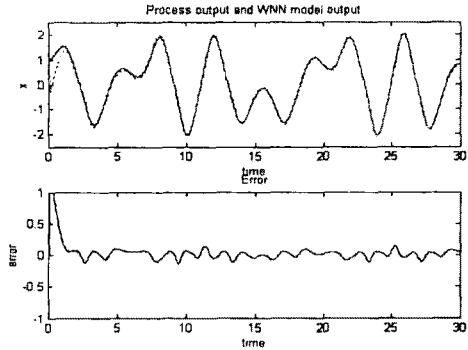


그림 4 경사 하강법을 이용한 웨이블릿 신경회로망의 동정 결과

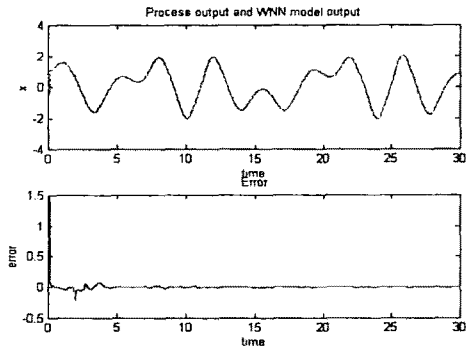


그림 5 확장 칼만 필터를 이용한 웨이블릿 신경 회로망의 동정 결과

표 2 웨이블릿 신경 회로망 모델의 모의 실험 환경

	GD	EKF
웨이블릿 노드 수	5	5
입력 노드 수	2	2
샘플링 주기	0.05	0.05
MSE	0.0237	0.0017

본 논문은 한국과학기술재단의 목적기초연구사업(R01-2001-000-00316)에 의해 지원되었습니다.

[참 고 문 헌]

- [1] E. Ott, C. Grebogi and J. A. Yorke, "Controlling Chaos", Phys. Rev. Lett., Vol. 64, No. 11, pp. 1196-1199, 1990.
- [2] E. A. Jackson and A. W. Hübler, "Periodic Entrainment of Chaotic Logistic Map Dynamics", Physica D, Vol. 44, pp. 407-420, 1990.
- [3] G. Chen and X. Dong, "On Feedback Control of Chaotic Continuous-Time Systems", IEEE Trans. on Circuits and Systems, Vol. 40, No. 9, pp. 591-601, 1993.
- [4] S. Mallat, "A theory for multiresolution signal decomposition: The Wavelet Transform", IEEE Trans. on Pattern Anal. Machine Intell., Vol. 7, No. 7, pp. 674-693, 1989.
- [5] Sharad Singhal and Lance Wu, "Training feed-forward network with the extended kalman algorithm", ICASSP-89, Vol. 2, pp. 1187-1190, 1989.
- [6] Mohinder S. Grewal and Angus P. Andrews, "Kalman Filtering: Theory and Practice", A Wiley-Interscience Publication, 2001.
- [7] 박상우, 최종태, 윤태성, 박진배, 최윤호, "웨이블릿 신경회로망을 이용한 혼돈 비선형 시스템의 모델링", 대한전기학회 하계학술대회 논문집, pp. 2034-2036, 2002.