Proceedings of KFIS 2001 Spring Conference, 2001. 5. 19

구간회귀 신경망의 속도개선

A Note for Speed-Up of Interval Regression Neural Network

이중우 · 권순학
영남대학교 전자정보공학부

Jung W. Lee and Soon H. Kwon
EECS, Yeungnam University
E-Mail : infoic@chunma.yeungnam.ac.kr

ABSTRACT

This paper deals with the speed-up of interval regression neural network. We propose an improved method of adjusting the parameter alpha used in the interval regression neural network to improve the learning speed and regression performance. Finally, we provide numerical examples to evaluate the performance of the proposed method.

Keywords : Interval regression, Neural Network

I. 서 론

프로세스로부터 측정된 자료를 바탕으로 프로세스의 거동을 관측하고 모델링하기 위하여 다양한 기법들이 제안되어 사용되어지고 있다. 대표적 기법으로는 i) 전통적인 수학적 모델을 사용한 기법, ii) 인간의 논리적 경험적 지식을 이용할 수 있는 퍼지 이론에 바탕을 둔 언어적 표현에 의한 모델링 방법 [1] 및 iii) 관측된 자료에 대한 학습을 통하여 프로세스의 구조를 일반화시켜 그 거동을 모델화 하려는 신경망 [2]과 같은 지능적 기법들이 많은 관심을 끌고 있으며 또한 좋은 결과들이 많이 발표되고 있다.

그러나, 이들 대부분의 연구가 측정된 자료로부터 프로세스의 거동을 보다 정확하게 모델링 하고자 하는데 그 초점을 두고 있다. 이와는 달리, 프로세서의 거동 범위에 대한 정보도 제어, 인식, 의사결정 등의 다양한 분야에서 유용하게 사용될 수 있기 때문에 프로세서의 거동 범위 선정에 관한 연구가 진행되어 왔으며, 그 대표적 방법으로는 Ishibuchi등에 의해 제안된 신경망에 의한 구간회귀(interval regression) 방법 [3]이 있다. Ishibuchi등의 신경망에 의한 구간회귀 방법은 프로세서의 거동범위의 상한과 하한을 일반화시키는 좋은 방법이지만 식(4) 및 식(5)에서 처럼 신경망의 출력과 데이터의 오차의 부호에 따라 평가함수(cost function)의 어느 한쪽의 기울기가 급격히 작아 지도록 평가함수를 조절하기 때문에 학습속도가 빠르지 못하고 학습에 소요되는 시간이 균일하지 않아서 적용에 많은 제한이 따르게 된다.

이런 문제점을 부분적으로 해결하기 위하여 Ishibuchi등은 학습 횟수에 따라서 평가함수의 한쪽 기울기를 점진적으로 감소시키는 방법을 제안하고 있으나, 이 또한 오차의 변화에 관계없이 임의로 기울기를 조절하므로 좋은 결과를 얻기에는 제한이 따른다.

본 논문에서는, 이러한 학습속도와 관련된

문제점을 개선하기 위하여 학습오차에 의한 구간회귀 매개변수 값의 변경에 관한 방법을 제안하고, 이러한 방법이 학습속도의 개선뿐만 아니라 학습후의 구간회귀 성능의 개선에도 기여함을 보이고자 한다.

II. 구간회귀 모델의 개념

Ishibuchi등에 의해 제안된 두 개의 층을 가지는 오차 역전파 신경망에 의한 구간회귀의 기본 개념은 상한회귀용(BPN^+)과 하한회귀용(BPN^-)의 두 개의 신경망을 독립적으로 학습시키는데, 주어진 입출력 패턴을

$$\begin{aligned}&(\mathbf{x}_k, d_k),\ \ k=1,2,\cdots,p\\&\mathbf{x}_k=(x_{k1},x_{k2},\cdots,x_{kn})^T \text{ 는 } k\text{번째 입력벡터}\end{aligned} \tag{1}$$

라 하면, 입력 $\mathbf{x}_k$에 대한 BPN^+와 BPN^-의 출력 $g^+(\mathbf{x}_k)$와 $g^-(\mathbf{x}_k)$가 근사적으로 만족해야 할 조건은

$$g^-(\mathbf{x}_k)\le d_k \le g^+(\mathbf{x}_k),\ \ k=1,2,\cdots,p \tag{2}$$

이 된다. 그리고 신경망을 학습시키기 위한 평가함수를

$$E=\sum_{k=1}^{p}E_k=\frac{1}{2}\sum_{k=1}^{p}\alpha_k[d_k-g^-(\mathbf{x}_k)]^2 \tag{3}$$

로 두면, 여기서 평가함수의 기울기에 대한 가중치 α_k는 BPN^+에 대해서는

$$\alpha_k=\begin{cases}1 & \text{if } d_k > g^+(\mathbf{x}_k)\\ \alpha & \text{if } d_k\le g^+(\mathbf{x}_k)\end{cases} \tag{4}$$

가 되고, BPN^-에 대해서는

$$\alpha_k=\begin{cases}\alpha & \text{if } d_k\ge g^-(\mathbf{x}_k)\\ 1 & \text{if } d_k < g^-(\mathbf{x}_k)\end{cases} \tag{5}$$

가 되는데, α_k는 (0,1)의 작은 양의 실수이다. 오차 역전파 학습을 위한 학습알고리즘은 식(3)의 평가함수로부터 구할 수 있다.

$$\Delta w_j=\eta\left(-\frac{\sigma E_k}{\sigma w_j}\right)=\eta\delta_k y_{kj} \tag{6}$$

$$\Delta w_{ji}=\eta\left(-\frac{\sigma E_k}{\sigma w_{ji}}\right)=\eta\delta_{kj}x_{ki} \tag{7}$$

$$\delta_k=\alpha_k(d_k-y_k)y_k(1-y_k) \tag{8}$$

$$\delta_{kj}=y_{kj}(1-y_{kj})\delta_k w_j \tag{9}$$

여기서, w_j는 은닉층과 출력층의 연결가중치, w_{ji}는 입력층과 은닉층의 연결가중치, η는 학습률, $y_k=g^+(\mathbf{x}_k)$는 신경망 출력, y_{kj}는 은닉층의 j번째 뉴런의 출력이다. y_k및 y_{kj}를 위한 활성함수로는 식(10)과 같은 단극 S-자형 함수(unipolar sigmoid function)를 사용한다.

$$y(net)=\frac{1}{1+e^{-\lambda\cdot net}}$$

$$\frac{dy(net)}{dnet}=y(net)(1-y(net)) \tag{10}$$

Ishubichi등은 매개변수 a를 경험적으로 0.01정도가 적절하다고 제시하고 있으나, 학습속도가 느린 단점이 있어, 아래의 식(11)로 표현되는, 즉, 학습횟수에 따른 새로운 매개변수 선정 방법을 제시하였다.

$$\alpha(t)=\frac{1}{[1+(t/2000)]^3} \tag{11}$$

여기서, t는 학습횟수.

III. 제안된 학습방법의 개념

Ishubichi등에 의해 제안된 매개변수 설정방법은 신경망의 학습오차와 관계없이 a 값을 결정하므로 여전히 많은 문제점을 가지고 있다. 본 논문에서는, 이러한 불합리한 점을 개선하기 위하여 신경망의 학습오차 E에 비례한 매개변수 a를 결정하는 방법과 활성함수의 재선정 및 학습순서의 최적화를 통하여 학습속도와 학습후의 결과가 더욱 향상될 수 있음을 보이고자 한다.

우선, Ishubichi등이 사용한 S-자형 활성함수를 은닉층에는 식(13)과 같은 양극 S-자형 활성함수 (bipolar sigmoid activation function)로

출력층에는 식(15)와 같은 선형 활성함수 (linear activation function)로 바꾼다.

$$\delta_{kj}=\frac{1}{2}(1-y_{kj}^2)\delta_k w_j, \tag{12}$$

$$y(net)=\frac{2}{1+e^{-\lambda\cdot net}}-1$$
$$\frac{dy(net)}{dnet}=\frac{1}{2}(1-y^2(net)) \tag{13}$$

$$\delta_k=\alpha_k(d_k-y_k), \tag{14}$$

$$y(net)=net$$

$$\frac{dy(net)}{dnet}=1 \tag{15}$$

그러므로, 은닉층과 출력층의 d값은 각각 식 (12) 및 (13)과 같이 된다.

신경망을 임의의 초기치에서 양의 작은 a값을 가지는 구간회귀 신경망을 학습할 경우 학습속도가 느리게 되므로 이를 보완하기 위해 먼저 a=1로 두고 충분히 일반화 되도록 학습시킨 후 BPN'와 BPN⁻을 일반화된 신경망에 기반하여 계속 학습시킨다.

$$\Delta E^{(t)}=\log\left(\frac{E^{(t-1)}}{E^{(t)}}\right) \tag{16}$$

$$\alpha^{(t)}(E)=\begin{cases}(1-\alpha_{\min})\dfrac{E^{(t)}}{E_g}+\alpha_{\min} & 0<\Delta E^{(t)}<\Delta E_{\min}\\ \alpha^{(t-1)} & otherwise\end{cases} \tag{17}$$

여기서, t는 학습횟수, $E^{(t)}$는 t에서의 학습오차, $\Delta E^{(t)}$는 t에서의 학습오차의 변화율, $\Delta E_{\min}$은 학습오차 변화율의 하한값, $\alpha_{\min}$은 $\alpha^{(t)}$의 하한값, E_g는 $\alpha=1$로 일반화 되도록 학습한 최종 학습오차.

Ⅳ. 모의실험 결과 및 고찰

Ishubichi등에 의해 제안된 방법과 본 논문에서 제안된 개선된 방법의 결과를 비교하기 위한 학습 자료로는 Ishubichi등의 논문에서 사용된 식(18)을 통해 발생되는 자료를 사용하였다.

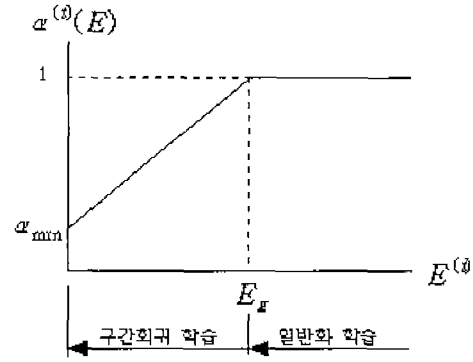

그림 1. 학습오차에 따른 α의 변경

$$x_k=0.02(k-1),\quad k=1,2,\cdots,51,$$

$$d_k=0.2\sin(2\pi x_k)+0.2x_k^2+0.3+(0.1x_k^2+0.05)rnd[-1,1]. \tag{18}$$

모의실험은 은닉층에 세 개의 뉴런을 가진 2층 전방향 신경망을 OSS(One step secant backpropagation) 방식으로 학습하였다.

모의실험은 3가지 경우, 즉, (a) 식(4),(5)와 같이 고정된 α_k값을 사용한 경우, (b) 식(11)과 같이 학습횟수에 따라 $\alpha(t)$를 변화 시킨 경우, 그리고 (c) 본 논문에서 제안된 방법에 의한 경우에 대하여 모의실험 도구(MatLab)으로 10회 반복하고 그 결과를 그래프와 소요된 평균연산횟수 (MFlops)로 비교하였다. 여기서, 학습횟수는 BPN⁺,BPN⁻와 일반화 학습에 사용된 총 학습횟수를 말한다. 그리고 200회의 학습후 $\alpha(t)=0.01$에 도달하게 하려면 식(11)의 2000은 부적절하므로 식(19)에 의해서 $Const$=55를 사용하였다.

$$\alpha(t)=0.01=\frac{1}{[1+(t/Const)]^3}\Big|_{t=200} \tag{19}$$

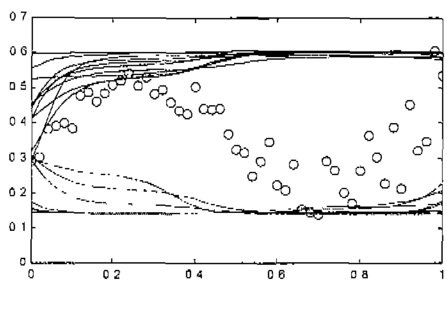

그림 2. 경우(a)의 학습결과

그림 2.에서 알 수 있듯이 고정된 α_k를 사

용할 경우 2000회의 학습에도 불구하고 만족할 만한 결과를 얻을 수 없었다. Ishubichi등의 논문에서 제시된 바와 같이 만족스런 결과를 얻기 위해서는 10000회 이상의 반복학습을 요한다.

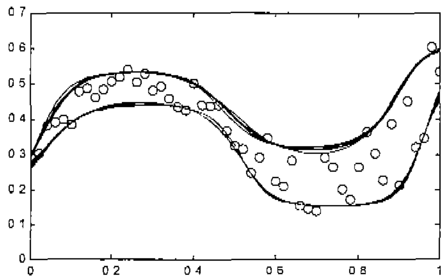

그림 3. 경우(b)의 학습결과

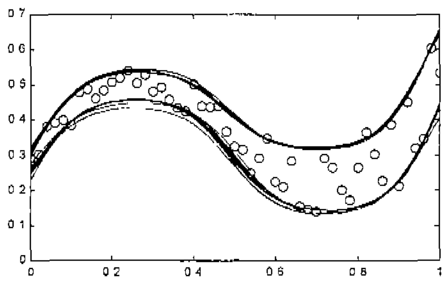

그림 4. 경우(c)의 학습결과

반면에 학습횟수에 따라 매개변수 $\alpha(t)$를 변경시킬 때는 그림 3에서와 같이 상당히 개선된 결과를 보이고 있지만, 학습 자료가 삼각함수에 잡음을 첨가한 형태임을 고려할때 그림 4.의 $\alpha^{(t)}(E)$를 사용한 경우가 더욱 더 학습 자료의 경향을 충실하게 나타냄을 볼 수 있다.

표 1. 모의실험 결과비교

실험방법	경우(a)	경우(b)	경우(c)
MFlops	26.97	10.08	2.39
학습횟수	2000	400	400
사용된 파라미터	$\alpha_k = 0.01$	$\alpha(t) = \frac{1}{[1-(t/55)]^3}$	$\alpha_{\min} = 10^{-2}$ $\Delta E_{\min} = 10^{-5}$

각각의 매개변수 변경 방법을 적용하여 10회 반복한 결과를 나타내는 표1에서 알 수 있듯이 본 논문에서 제시된 방법에 의한 구간회귀 신경망의 학습은 고정된 매개변수를 사용한 경우 (a)와 비교할 때 최소 11배 이상, 학습횟수에 의한 매개변수를 변경하는 경우(b)와 비교할 때 약 4배 이상의 학습속도 향상을 나타낼 뿐만 아니라 그림 2, 3 및 그림 4에서 알 수 있듯이 학습 자료의 경향을 더욱 충실히 나타내는 결과를 얻어 그 효용성이 뛰어남을 알 수 있다.

V. 결 론

본 논문에서는 프로세스로부터 측정된 자료의 거동범위에 대한 정보를 얻기 위한 구간회귀 신경망의 학습속도 개선방법에 대하여 논의하였다. 구간회귀에 사용되는 신경망의 은닉층과 출력층의 뉴런의 적절한 선정, 신경망을 먼저 일반화 학습시킨 후 구간회귀 학습하는 방법, 학습오차에 의한 매개변수의 변경이 학습 후의 구간회귀 성능과 학습시간의 측면에서 개선된 결과를 나타냄을 보였다.

향후 연구과제는 구간회귀 신경망의 학습에 있어서 매개변수 a에 영향을 받지 않는 성능 비교 기준의 제시 및 응용에 관한 연구라 할 수 있다.

VI. 참고문헌

[1] E. Mamdani, "Application of fuzzy logic to approximate reasoning using linguistic systems," Fuzzy Sets and Systems, Vol. 26, pp. 1182-1191, 1977.

[2] K. S. Narendra, K.Parthasarathy, "Identification and control of dynamical systems using neural networks," IEEE Trans, Neural Networks, Vol. 1, pp. 4-12, 1990.

[3] H. Ishibuchi, H. Tanaka, "Determination of fuzzy regression models by neural networks," Tokyo: IOS Press, Fuzzy Engineering toward Human Friendly Systems, pp. 523-534, 1992.