

신경망을 이용한 고신뢰성의 회귀분석 모델

조 용 현†

요 약

본 논문에서는 기울기하강과 동적터널링이 조합된 학습알고리즘의 다층신경망을 이용한 고신뢰성의 회귀분석 모델을 제안하였다. 기울기하강은 빠른 수렴속도의 최적화가 가능하도록 하기 위함이고, 동적터널링은 국소최적해를 만났을 때 이를 벗어난 새로운 연결가중치를 설정하여 전역최적해로 수렴되도록 하기 위함이다. 또한 대용량의 입력 데이터를 통계적으로 독립인 특징들의 집합으로 변환시키는 주요성분분석 기법의 속성을 살려 학습데이터의 차원을 감소시킴으로서 고차원의 학습데이터에 따른 회귀분석 모델의 제약도 동시에 해결하였다. 제안된 기법의 신경망을 3개의 독립변수 패턴을 가진 암모니아 제조공정문제와 10개의 독립변수 패턴을 가진 자동차 연비문제에 각각 적용하여 시뮬레이션한 결과, 기존의 역전파 알고리즘의 신경망이나 주요성분분석에 의한 차원을 감소시키지 않은 학습패턴을 이용한 신경망보다 각각 더욱 우수한 학습능력과 회귀성능이 있음을 확인할 수 있었다. 또한 학습패턴의 영평균 정규화로 회귀용 신경망의 성능을 더욱 더 개선하였다.

Regression Model With High Reliability by Using Neural Networks

Yong-Hyun Cho†

ABSTRACT

This paper proposes a regression analysis with high reliability using neural networks of a hybrid learning algorithm. The proposed algorithm is a learning algorithm based on the steepest descent and dynamic tunneling. The steepest descent is applied for high-speed optimization, and the dynamic tunneling is also applied for converging to the global optimum by estimating the new initial weights for escaping the local minimum. And the adaptive principal component analysis (PCA) is also applied for solving the limits of regression model by reducing efficiently the dimension by extracting the features of the given independent variables. The proposed neural networks has been applied to regress the Ammonia producing process (3-independent variable) and the mobile oil cost (10-independent variable). The simulation results show that the proposed models have better performances of the learning and the regression, in comparison with those using the conventional backpropagation-based multilayer perceptron (BPMLP) and the BPMLP using the learning patterns without reducing the dimension by PCA, respectively. And we confirm that the zero mean normalization of learning patterns makes better performances of the regression neural networks.

키워드 : 회귀분석(regression analysis), 다층신경망(multilayer perceptron), 동적터널링(dynamic tunneling), 주요성분분석(principal component analysis), 영평균 정규화(zero mean normalization)

1. 서 론

회귀분석은 하나의 종속변수가 다른 독립변수들에 의해 어떻게 설명 또는 예측되는지를 알아보기 위해 적절한 함수로 표현하여 자료분석을 하는 통계적인 기법이다[1-6]. 이러한 회귀분석은 모수추정을 포함한 관계의 기술이나 새로운 것에 대한 예측, 그리고 통계적 제어 등에 널리 이용되고 있다.

기존의 수치적인 기법에 기초한 회귀분석에는 독립변수의 개수나 종속변수와 독립변수들의 관계에 따라 여러 가지 회귀기법들로 나누어진다[1, 2]. 일반적으로 다항회귀(poly-nomial regression)와 다중선형회귀분석(multiple linear regression)이 많이 이용되고 있다. 이들 기법들에서는 주어진 자

료들로부터 독립변수와 종속변수의 상관관계에 대한 사전 지식을 통하여 회귀분석 방정식의 모델을 설정하며 회귀계수들의 값은 통계적으로 결정하였다. 하지만 문제에 따라서는 이러한 모델의 설정이나 계수들의 결정이 매우 힘든 제약들은 여전히 존재한다.

이러한 기존 수치적 기법들이 가지는 제약들을 해결하기 위해서 입력과 출력간의 사상을 가능하게 하며, 대규모 분산병렬처리와 학습, 그리고 일반화 등의 속성을 가지는 신경망이 널리 이용되고 있다[3-8]. 신경망을 이용하는 기법 중에서 입력층과 출력층 사이에 은닉층을 가지는 다층전향 신경망은 입력층 데이터 내에 포함된 어떤 비선형 연속함수도 근사화하거나 재구성할 수 있어 대단히 일반적이면서도 융통성 있는 속성을 가지고 있다. 이러한 신경망에는 일반회귀신경망(general regression neural network : GRNN)[5, 6]과 역전파(backpropagation : BP) 알고리즘의 다층신경망

* 본 연구는 2000학년도 대구가톨릭대학교 연구비 지원에 의한 것임.

† 정 회 원 : 대구가톨릭대학교 컴퓨터정보통신공학부 교수
논문접수 : 2001년 7월 31일, 심사완료 : 2001년 8월 23일

(multilayer perceptron : MLP)[3,4]이 이용된다.

이들 중 GRNN은 입력층간 확률밀도함수의 추정에 기반을 둔 memory-based 전향신경망으로 4개의 층 구조를 가지는 우수한 성능의 회귀분석 신경망이다[3-6]. 특히 패턴층의 출력은 각 뉴런이 가지는 중앙값과 입력값의 차의 제곱이나 절대치를 구하여 활성화 함수를 통해 생성된다. 여기서의 활성화 함수로는 다양한 커널(kernel) 함수들이 이용된다. 이는 GRNN에서 학습데이터의 입력과 출력사이 확률밀도함수가 커널함수의 가중된 합으로 추정되는 것을 의미하며, 그 함수로는 정규화된 가우스(Gaussian) 함수가 가장 널리 이용되고 있다. GRNN은 학습알고리즘이 대단히 간단하고 구현이 용이하며, 학습시간이 요구되지 않는 매우 빠르면서도 고신뢰성을 가지는 회귀분석 신경망이다. 하지만 독립변수 패턴의 고차원 문제, 최적의 평활요소(smoothing factor) 및 중앙값 설정, 회귀분석을 위한 데이터 내 변화를 정확하게 측정하기 위한 많은 학습패턴 요구 등의 제약들이 있다.

한편, MLP는 충분한 뉴런을 가지고 있을 때 어떤 임의의 함수도 근사화할 수 있다고 알려져 있지만 유용한 모델을 얻기 위해서는 많은 학습데이터와 시험데이터가 요구된다. 특히, 학습알고리즘으로 널리 이용되고 있는 기울기하강(steepest descent)의 속성을 이용하는 역전파 알고리즘은 학습파라미터의 설정에 따라 수렴속도와 견실성 중 하나 이상의 제약을 가지며, 전역최적해로의 수렴이 보장되어 있지 않다. 이러한 문제를 해결하기 위한 여러 방법들이 연구되어 왔다[7,8]. 특히 독립변수 패턴의 고차원은 학습시간의 증가와 함께 과학습(overlearning)이 발생되어 회귀분석 성능의 저하를 가져오게 된다[1,2,5,6]. 이를 해결하기 위해 기존의 방법들에서는 분산팽창인자(variance inflation factor), 상태수, 상태지표, 그리고 분산비율 등의 척도에 의한 다중공선성 분석이나 변수변환 등의 방법들을 이용하고 있다[1,2]. 그러나 척도들을 계산하는 데는 표본의 평균과 표준편차, 독립변수 행렬의 고유벡터(eigen vector)와 고유값(eigen value) 등의 계산이 요구되며, 그 계산 또한 매우 복잡하여 계산시간이 오래 걸린다. 따라서 독립변수들 간의 상호관계를 잘 나타내는 주요 특징들을 효과적으로 추출하여 회귀분석 데이터로 이용하고, 좀 더 빠른 수렴속도와 학습성능을 가지는 학습알고리즘의 신경망이 개발된다면 효과적인 고신뢰성의 회귀분석이 이루어질 수 있을 것이다.

본 연구에서는 수치적인 데이터 집합의 차원을 해석하여 이를 감소시키는 것으로 널리 알려진 적응적 주요성분분석(principal component analysis : PCA) 기법[8-11]으로 독립변수 벡터의 특징을 추출한 다음, 추출된 특징을 입력으로 하는 기울기하강과 동적터널링(dynamic tunneling)[12,13]을 조합한 역전파 알고리즘의 신경망을 이용한 고신뢰성의 회귀분석 모델을 제안한다. 또한 이용되는 회귀자료들을 영평균(zero mean) 정규화함으로써 알고리즘을 더욱 간략화하였다[14]. 제안된 조합형 학습알고리즘의 신경망을 3개의

독립변수 패턴을 가진 암모니아 제조공정문제[1]와 10개의 독립변수 패턴을 가진 자동차 연비문제[2]에 각각 적용하여 시뮬레이션하고 그 타당성을 확인하였으며, 기존의 BP를 이용한 MLP 및 독립변수의 수를 감소시키지 않은 패턴을 이용할 경우와 그 성능을 비교 고찰하였다.

2. 적응적 특징추출을 위한 주요성분분석

PCA는 n 차원 입력공간의 데이터를 m 차원 출력공간의 데이터로 투영시키는 것이다. 여기서 $m < n$ 이며, 이는 입력데이터 벡터의 대부분 내부정보를 유지하도록 차원의 감소를 얻는 것이다. 자기상관행렬(auto-correlation matrix) $\mathbf{R}_{xx} = \langle \mathbf{x}\mathbf{x}^T \rangle$ 를 가진 평균이 영인 입력벡터 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ 에 대해서 생각해 보자. 여기서 $\langle \mathbf{x} \rangle$ 는 기대치를 나타낸다. 또한 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$ 이 \mathbf{R}_{xx} 의 고유벡터와 직교되는 연결가중치 벡터라 할 때, $\mathbf{w}_1 = [w_{11}, w_{12}, \dots, w_{1n}]^T$ 는 가장 큰 고유치 λ_1 과 일치하며, $\mathbf{w}_2 = [w_{21}, w_{22}, \dots, w_{2n}]^T$ 는 두 번째로 큰 고유치 λ_2 와 $\mathbf{w}_n = [w_{n1}, w_{n2}, \dots, w_{nn}]^T$ 는 가장 작은 고유치 λ_n 과 각각 일치한다. 이상의 관계를 행렬방정식으로 나타내면

$$\mathbf{R}_{xx} \mathbf{w}_j = \lambda_j \mathbf{w}_j \quad (j = 1, 2, \dots, n) \quad (1)$$

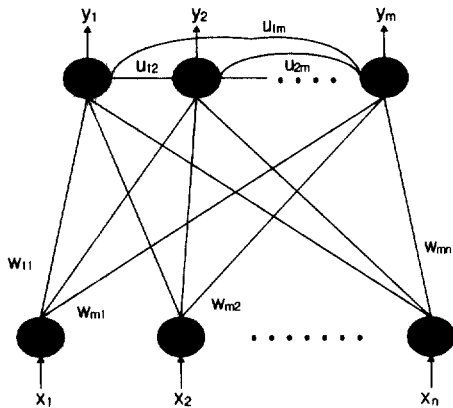
이다. 여기서 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$ 이다.

한편, 입력되는 데이터의 주요특징들을 추출하기 위한 기존의 PCA 방법들을 살펴보면, 수치적 방법에서는 먼저 자기상관행렬 \mathbf{R}_{xx} 를 계산한 다음 그것과 일치하는 고유벡터를 계산한다. 그러나 입력데이터가 대규모의 차원을 가진다면, 상관행렬은 대단히 크게 되어 요구되는 고유벡터를 찾는 것도 매우 복잡한 일이다. 결국 기존의 수치적 기법들은 대규모의 실시간 처리가 요구되는 응용문제에서는 매우 비효율적이다. 이를 해결하기 위한 대안으로 상관행렬의 고유벡터를 실시간으로 추정하는 적응적 학습알고리즘의 신경망을 이용하는 방법들이 제안되었다[8-11]. 여기서는 입력데이터 벡터로부터 직접 고유벡터를 추정할 수 있기 때문에 상관행렬의 추정과정이 요구되지 않는다. 이때 이용되는 신경망은 주로 입력층과 출력층으로 구성된 단층구조이다.

Oja[8]는 정규화된 헤비안규칙(normalized Hebbian rule)의 적응학습 방법을 이용한 간단한 뉴런모델을 제안하고 정상입력벡터 계열의 주요 특징으로 연결가중치 벡터가 수렴됨을 증명하였다. 그러나 이 방법에서는 한 개의 중요한 주요특징 추출만 가능하다. Sanger 등[8]은 일반화된(generalized) 헤비안규칙을 이용함으로써 정상과정의 m 개의 가장 중요한 주요특징들을 계산하기 위한 다중 뉴런모델을 제안하였다. 한편, Foldiak [8,9,11]은 망의 입력과 출력사이의 연결가중치 경신에는 정규화된 헤비안규칙을 이용하고, 망의 출력사이의 측면연결 가중치 경신에는 반 헤비안규칙(anti-Hebbian rule)을 함께 이용한 학습 알고리즘을

제안하였다. 일반적으로 Sanger 등에 의해 제안된 학습규칙보다는 Foldiak에 의해 제안된 학습규칙이 수렴속도 면에서 더 우수한 것으로 알려져 있다[8, 11].

(그림 1)은 n개의 입력뉴런과 m개의 출력뉴런으로 구성된 입력과 출력뉴런간 및 출력뉴런 상호간의 측면연결을 가진 단층신경망의 구조이다.



(그림 1) m개의 첫 번째 주요특징 추출을 위한 측면연결의 단층신경망

그림에서 입력과 출력의 관계를 나타내면 다음과 같다. 즉

$$y_i = \sum_{j=1}^n w_{ij} x_j + \sum_{k=1}^{i-1} u_{ik} y_k \quad (i = 1, 2, \dots, m) \quad (2)$$

이다. 여기서 w_{ij} 는 입력뉴런과 출력뉴런을 연결하는 연결가중치이고, u_{ik} 는 출력뉴런 상호간의 측면연결 가중치이다. 이때 출력뉴런 간의 상호연결을 보면 뉴런 i 는 $h < i$ 인 뉴런에만 연결된다. 따라서 신경망을 학습시켜 입출력 뉴런간 및 출력뉴런 상호간의 연결가중치를 각각 구함으로써 주어진 n개의 입력벡터 \mathbf{x} 로부터 m개의 주요특징벡터인 출력벡터 \mathbf{y} 를 구할 수 있다. 입력과 출력뉴런 간의 연결가중치 경신에는 정규화된 헤비안규칙을 이용하고 출력뉴런 간의 측면 연결가중치 경신에는 반헤비안규칙을 이용하는 Foldiak의 학습알고리즘은 다음과 같다. 여기서 입출력 뉴런간의 연결가중치 w_{ij} 의 경신규칙은

$$w_{ij}(t+1) = w_{ij}(t) + \eta [y_i(t) x_j(t) - w_{ij}(t) y_i(t)^2], \quad (i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n) \quad (3)$$

이고, 출력뉴런 간의 측면 연결가중치 u_{ik} 의 경신규칙은

$$u_{ik}(t+1) = u_{ik}(t) + \rho y_i(t) y_k(t), (i > k) \quad (4)$$

이다. 여기서 식 (3)의 η 와 식 (4)의 ρ 는 각각 학습율이다. 따라서 식 (3)과 식 (4)에 따라 입력과 출력뉴런 간의 연결가중치 및 출력뉴런 상호간의 측면 연결가중치를 각각 경신시켜 식 (2)에 대입하면 m개의 주요특징들을 추출할 수 있다.

결국 PCA는 높은 차원의 입력공간을 더 낮은 차원의 표

현공간으로 사상시켜 입력데이터가 가지는 두드러진 특징들을 추출하는 기법이다. 이는 입력 데이터 내에 존재하는 주요특징들을 추출함으로써 입력의 개수를 감소시키는데 이용될 수 있다. 따라서 학습데이터의 독립변수 패턴 각각에 적용적 PCA를 이용하여 특징을 구한 다음 이를 회귀분석 신경망의 입력으로 이용한다. 이렇게 하면 MLP 입력층 뉴런의 개수가 줄어들어 학습을 위한 연결가중치의 감소로 학습시간이 감소되고, 학습데이터의 고차원에 따른 독립변수들 상호 간의 다중공선성 등으로 인한 회귀성능의 저하와 과학습과 같은 제약들을 해결할 수 있어 MLP가 가지는 우수한 속성을 더욱 더 잘 살릴 수 있다.

3. 기울기하강과 동적터널링을 조합한 학습알고리즘

다층신경망은 입력층과 출력층 간에 하나 이상의 은닉층을 가지는 구조의 신경망이다. 은닉층은 단층신경망이 가지는 여러 가지 한계점을 극복하기 위해서 제안되었다. 다층신경망의 학습알고리즘으로 역전파 알고리즘이 제안되어 널리 이용되고 있다. 제안된 학습알고리즘은 기울기하강과 동적터널링에 기초를 둔 역전파 알고리즘이다. 먼저 기존의 역전파 알고리즘에 대하여 간략하게 설명한다.

i_{jp} 와 y_{kp} 를 각각 p번째 입력 신호 x_{ip} 에 대한 은닉층과 출력층의 j번째와 k번째 뉴런 출력으로 하고, d_{kp} 를 p번째 입력에 대한 k번째 출력층 뉴런의 원하는 출력이라 할 때, 기존의 기울기하강에 의한 출력층의 연결가중치는

$$w_{kj}(t+1) = w_{kj}(t) + \eta \delta_{kp} i_{jp} + \alpha \Delta_p w_{kj}(t-1) \quad \delta_{kp} = (y_{kp} - d_{kp}) f'_k(\sum_i w_{ki} i_{jp}) \quad (5)$$

와 같이 경신되며, 또한 은닉층의 연결가중치는

$$w_{ji}(t+1) = w_{ji}(t) + \eta \delta_{jp} x_{ip} + \alpha \Delta_p w_{ji}(t-1) \quad \delta_{jp} = f'_j(\sum_i w_{ji} x_{ip}) \sum_k \delta_{kp} w_{ki} \quad (6)$$

와 같이 경신된다. 여기에서도 η 는 학습율이고 α 는 모멘트이다. 즉, 연결가중치의 경신은 위의 식 (5)과 식 (6)에 따라 후향으로 이루어진다.

한편, 동적터널링 알고리즘은 Yao에 의해서 제안되었다 [12]. 이는 동적최적화와 동적터널링으로 구성된다. 여기서 동적최적화는 초기 상태로부터 국소최소값 중의 한 값을 찾는 과정이며, 동적터널링은 찾은 국소최소값보다 하위유역에 위치한 새로운 초기 상태를 찾는 과정이다. 즉, 동적터널링은 동적최적화가 국소최소점에 빠지는 것을 막아준다.

따라서, 신경망에서 연결가중치 \mathbf{w} 에 대한 오차함수 $E(\mathbf{w})$ 를 최소화하는 문제를 살펴보자. 이 문제는 $(\partial E_p(\mathbf{w})/\partial \mathbf{w})$ 를 이용하여 $E(\mathbf{w})$ 의 국소최소값 \mathbf{w}^* 를 찾는 것으로 변형될 수 있다. 즉, $\Delta p \mathbf{w} = -\eta (\partial E_p(\mathbf{w})/\partial \mathbf{w})$ 는 동적최적화와 동일하다. 동적터널링은 동적최적화로부터 구해진 \mathbf{w}^* 에다 결정론적으로나 임의로 설정되는 교란벡터(perturbation vector) ϵ

이 추가된 새로운 시작점에서 출발하는 것으로 다음과 같은 미분 방정식의 형태로 주어진다.

$$\Delta_p \mathbf{w} = -(\partial E_p(\mathbf{w})/\partial \mathbf{w})[1/[(\mathbf{w}-\mathbf{w}^*)^T (\mathbf{w}-\mathbf{w}^*)]^\gamma - \xi E_p^*(E_p^*(\mathbf{w}))] \quad (7)$$

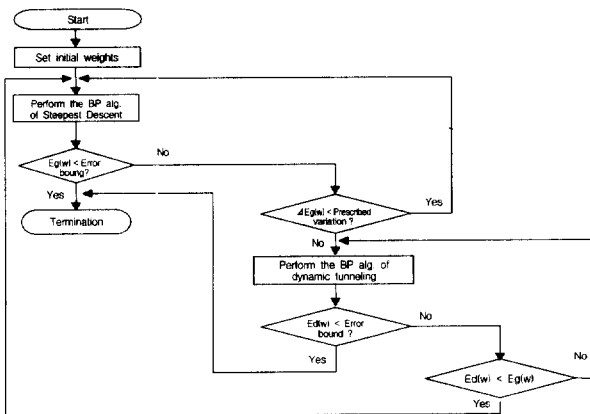
여기서, ξ 는 터널링페널티(tunneling penalty)로 $\|\frac{\partial E_p}{\partial \mathbf{w}}\|/\lambda$ 보다 같거나 큰 값을 가지나 실험에서는 적당한 상수로 설정하였으며, λ 는 작은 양수값이다. 그리고 γ 의 값은 $(\partial E_p(\mathbf{w})/\partial \mathbf{w})$ 의 원점 \mathbf{w}^* 의 차수보다 크거나 같은 값을 가진다. 또한 $E_p^*(\mathbf{w}) = E_p(\mathbf{w}) - E_p(\mathbf{w}^*)$ 로 정의되며, $E_p^*(\cdot)$ 는 다음과 같이 두 개의 분리된 구분선형(piecewise linear) 함수로 정의된다. 즉, $z > 0$ 이면 $E^*(z) = z$, 그렇지 않으면 0의 값을 가진다. 식 (7)에서, 우편 첫 번째 항의 분모 $[(\mathbf{w}-\mathbf{w}^*)^T (\mathbf{w}-\mathbf{w}^*)]^\gamma$ 항은 동적최적화로부터 구해진 국소최소점인 평형상태 \mathbf{w}^* 를 제거하기 위한 것이며, 두 번째 항은 하위유역에 존재하는 즉, $E_p(\mathbf{w}(0)) \leq E_p(\mathbf{w}^*)$ 인 새로운 초기 상태 $\mathbf{w}^{(0)}$ 를 찾기 위한 것이다. 그러므로 동적터널링에 의한 학습알고리즘의 연결가중치 경신식은 식 (7)과 같으며, 여기서 출력층과 은닉층 각각의 $(\partial E_p(\mathbf{w})/\partial \mathbf{w})$ 는 다음과 같이 정의된다. 즉,

$$\begin{aligned} \partial E_p(\mathbf{w})/\partial \mathbf{w}_{kj} &= -\eta \delta_{kp} i_{jp} \\ \partial E_p(\mathbf{w})/\partial \mathbf{w}_{ji} &= -\eta \delta_{jp} x_{ip} \end{aligned} \quad (8)$$

이다. 여기서 δ_{kp} 와 δ_{jp} 는 각각 식 (5)와 식 (6)으로 주어진다. 따라서 식 (7)과 식 (8)이 조합된 동적터널링에 기초한 제안된 다층신경망의 연결가중치 경신식은 다음과 같다. 즉,

$$\begin{aligned} \mathbf{w}(t+1) &= \mathbf{w}(t) + \eta \delta_i [1/[(\mathbf{w}-\mathbf{w}^*)^T (\mathbf{w}-\mathbf{w}^*)]^\gamma \\ &\quad - \xi E_p^*(E_p^*(\mathbf{w}))] + \alpha \Delta_p \mathbf{w}(t-1) \end{aligned} \quad (9)$$

이다. 결국 제안된 다층신경망의 학습알고리즘은 먼저 기울기하강의 식 (5)와 식 (6)의 경신식으로 학습을 수행하다 오차의 변화가 일어나지 않거나 그 변화량이 설정치 이하이면 국소최소값에 수렴된 것으로 가정하고 다음의 동적터널링의 식 (9) 경신식으로 학습을 수행한다. 이 과정의 반복수행으로 신경망을 학습하면 빠른 시간 내에 전역최적의 학습이 가능할 것이다. (그림 2)는 제안된 학습 알고리즘의 흐름도를 요약



(그림 2) 제안된 학습알고리즘의 흐름도

도시한 것이다. 그림에서 $E_g(\mathbf{w})$ 와 $E_d(\mathbf{w})$ 는 각각 기울기하강과 모멘트 및 동적터널링에 의해서 계산되는 전체 오차 함수 값이다.

4. 시뮬레이션 결과 및 분석

제안된 기법의 회귀용 신경망의 성능을 평가하기 위해 Foldiak 학습알고리즘의 단층신경망과 3층 전향신경망을 각각 구성하였다. 이용된 단층신경망 및 다층신경망에서 각 뉴런층 사이의 초기 연결가중치는 각각 랜덤시드(random seed)를 이용하여 -1에서 +1 사이의 임의 값으로 설정하였다. 학습은 전체 학습반복수가 20,000이상이거나 전체 오차 값이 허용치 10^{-4} 이하일 때 종료되도록 하였다. 또한 학습 알고리즘에서 학습율과 모멘트는 각각 0.1과 0.5로 하였다.

제안된 기법을 이용한 신경망을 21개의 패턴데이터(3개 독립변수)를 가진 암모니아 제조공정문제[1]와 32개의 패턴 데이터(10개 독립변수)를 가지는 자동차 연비문제[2]를 대상으로 펜티엄III-700 컴퓨터를 이용하여 시뮬레이션한 후 그 타당성을 확인하였으며, 기존의 독립변수를 그대로 이용하는 신경망 및 PCA에 의해 차원이 감소된 학습패턴을 이용한 신경망과 그 성능을 각각 비교 고찰하였다.

4.1 암모니아 제조공정문제

실험에 이용된 암모니아 제조공정문제는 공정 입력조건에 대한 공정의 출력특성 관계를 얻기 위한 문제이다. 이 문제는 독립변수로 3개의 공정 입력조건 x_1, x_2, x_3 와 종속변수로 1개의 출력특성 y 를 가지는 21개의 패턴들로 구성되며, 실험에서 16개는 학습패턴으로 나머지 5개는 시험패턴으로 이용하였다. 이용된 MLP에서 입력층 뉴런수는 입력조건인 독립변수의 개수와 동일하게 하고, 은닉층과 출력층 뉴런수는 각각 3개와 1개로 구성하였다. 다음의 <표 1>은 실험에 이용된 자료의 일부를 나타낸 것이다. 실험에서는 입·출력 조건을 각각 최대값으로 나누어 1이하의 값으로 단순히 정규화한 데이터와 각각의 조건에서 평균값을 뺀 차를 정규화한 데이터를 각각 이용하였다. 여기서 후자는 영평균을 위한 것으로 데이터를 중앙(centering)으로 모아 1차적 통계성을 고려함으로써 알고리즘을 더욱 간략화하기 위함이다.

<표 1> 3개의 독립변수와 1개의 종속변수를 가진 암모니아 제조공정자료

패턴 수	입력조건			출력특성 y
	x_1	x_2	x_3	
1	80	27	89	42
2	80	27	88	37
⋮	⋮	⋮	⋮	⋮
15	50	18	89	8
16	50	18	86	7
⋮	⋮	⋮	⋮	⋮
21	50	20	91	15

<표 2>는 역전파 알고리즘의 다층신경망에서 입력조건

의 단순한 정규화와 영평균 정규화 패턴 각각에 대하여 PCA를 적용한 경우와 그렇지 않은 경우에 대한 학습반복수 N_i 와 학습에 소요된 CPU 시간 t_{cpu} , 그리고 출력 y 와 학습 후 출력 y' 와의 절대오차(absolute error : AE, $AE = \sum_{i=1}^n |y - y'|$)로 표현되는 학습오차 E_l , 시험오차 E_t 및 합오차 E_s 를 나타낸 것이다. 한편, 실험에서 PCA의 적용은 입력조건들의 차원감소가 아닌 선형화를 위함이다. 이는 패턴 내의 상관성을 줄이고 분산을 1로 함으로써 입력조건들의 독립성을 증가시키기 위함이다. 표에서 보면 정규화 및 영평균 정규화 패턴에 대하여 PCA의 적용 유무에 관계없이 학습의 종료조건은 만족되지 않는다. 이는 이용된 다층신경망의 학습알고리즘으로 기존의 역전파 알고리즘을 이용하였기 때문이며, 역전파 알고리즘은 학습파라미터나 신경망의 구조에 따라서는 국소최적해에 수렴됨을 알 수 있다. 표에서 절대오차 중 학습오차를 보면, 영평균 정규화 패턴이 단순한 정규화 패턴에 비해 최대 약 2.2배 정도, PCA를 적용한 경우가 적용하지 않은 경우보다 최대 약 1.4배 정도 더 작은 값을 가진다. 이는 동일한 학습반복수에서 영평균 정규화 패턴에 PCA를 적용하면 학습오차면에서 우수한 학습성능이 있음을 확인할 수 있다. 한편 5개의 시험패턴에 대한 시험오차 측면에서도 영평균 정규화 패턴에 PCA를 적용한 경우가 가장 작은 오차를 나타낸다. 이는 영평균 정규화와 PCA를 적용하면 회귀분석의 성능도 향상됨을 알 수 있다. 한편 학습오차와 시험오차의 합으로 표현되는 합오차에서도 PCA와 정규화를 이용한 패턴이 PCA를 적용하지 않은 단순 정규화 패턴보다 약 1.8배 정도 작은 값을 가짐을 알 수 있다. 따라서 회귀분석을 위한 신경망의 입력패턴은 정규화뿐만 아니라 PCA를 적용하면 입력패턴 상호간의 종속성을 더욱 감소시켜 학습성능과 회귀성능이 개선됨을 알 수 있다.

<표 2> 입력조건들의 정규화 및 영평균 정규화 패턴에 대한 PCA 적용유무에 따른 학습결과

	Normalized pattern data		Centering & normalized pattern data		
	non-PCA	PCA	non-PCA	PCA	
학습반복수, N_i	20000	20000	20000	20000	
학습시간, t_{cpu}	35	35	35	35	
절대 오차, AE	학습오차, E_l	0.071116	0.065015	0.040833	0.029748
	시험오차, E_t	0.567587	0.509047	0.371189	0.329392
	합 오차, E_s	0.638703	0.574062	0.412022	0.35914

<표 3>은 회귀용 신경망을 대상으로 자료데이터를 영평균 정규화한 후 PCA를 적용한 데이터를 입력패턴으로 하여 기존의 역전파 알고리즘과 제안된 알고리즘으로 학습한 결과를 나타낸 것이다. 여기에서도 PCA의 적용은 입력조건들의 차원감소가 아닌 선형화를 위함이다. 이때 제안된 알고리즘에서 기울기하강으로부터 동적터널링으로의 전환을 위한 조건은 오차의 변화가 10^{-6} 이하일 때로 하였다. 이는 기울기하강에 의한 오차의 변화가 10^{-6} 이하이면 역전파 알고

리즘은 국소최적해로 수렴된 경우로 가정한다. 표에서 보면 기존 역전파 알고리즘의 경우 학습반복수가 20000번이 되어도 학습오차는 10^{-4} 이하로 감소되지 않아 국소최적해에 수렴된 경우이다. 특히 역전파 알고리즘의 경우 학습반복수를 60000번까지 지속하였으나 발전만 계속되고 오히려 학습오차는 증가되어 과학습이 이루어짐을 알 수 있다. 이는 기존역전파 알고리즘의 경우 기울기하강에 기초하여 연결가중치를 경신함으로써 국소최적해에 수렴되면 이를 벗어날 수 없기 때문이다. 하지만 제안된 알고리즘의 경우는 18987번의 기울기하강과 834번의 동적터널링이 조합된 19821번의 총 학습반복수에서 학습이 종료되었다. 이는 제안된 알고리즘에서 기울기하강에 의해 국소최적으로 수렴되면 동적터널링이 적용되며, 이들이 반복적으로 적용되어 결국 전역최적해로 수렴됨을 확인할 수 있다. 한편 종료조건으로 최대 학습반복수를 20000번으로 제한할 경우에도 학습시간과 학습반복수에서 개선된 결과를 보이며, 특히 학습오차와 시험오차에서는 매우 향상된 성능이 있음을 알 수 있다. 합오차를 기준으로 보면 제안된 알고리즘은 기존의 역전파 알고리즘보다 약 17.1배 정도 개선된 우수한 성능이 있음을 확인할 수 있다.

<표 3> 역전파 알고리즘과 제안된 학습알고리즘의 학습결과

	BP Algorithm		Proposed Algorithm	
	20000	60000	20000	
학습종료조건	20000	60000	20000	
학습반복수, N_i	20000	60000	19821(18987,834)	
학습시간, t_{cpu}	35	112	34	
절대 오차, AE	학습오차, E_l	0.029748	0.04154	0.0001
	시험오차, E_t	0.329392	0.451907	0.020863
	합 오차, E_s	0.35914	0.483447	0.020963

<표 4>는 역전파 알고리즘과 제안된 알고리즘을 각각을 대상으로 PCA를 적용하여 독립변수 개수 N 을 조정하면서 학습한 결과를 나타낸 것이다. 여기에서도 <표 3>에서처럼 제안된 알고리즘은 학습반복수와 시간, 그리고 절대오차에서 기존의 역전파 알고리즘에 비해 개선된 결과들을 보인다. 이는 제안된 알고리즘을 이용하면 상대적으로 역전파 알고리즘을 이용할 때보다 우수한 학습성능과 회귀성능이 있음을 알 수 있다. 한편 표에서는 제안된 알고리즘이나 역전파 알고리즘 모두 차원감소에 따라 구해진 2개와 1개의 특징들을 독립변수로 이용하는 경우가 3개의 독립변수를 입력조건으로 이용하는 경우보다 합오차가 작아 우수한 회귀성능이 있음을 알 수 있다. 이는 문제에 이용되는 3개의 독립변수 상호간에는 선형적인 연관성이 많음을 추측할 수 있다. 특히, 독립변수가 2개인 경우보다는 1개만을 이용하는 경우가 약간 우수한 회귀성능을 가진다. 이는 암모니아 제조공정자료에서 입력조건인 독립변수 상호간에는 많은 상관성이 있음을 다시 한번 더 확인할 수 있으며, 결국 독립변수 상호간의 상관성은 회귀성능을 저하시키는 결과를 초래함을 확인할 수 있다. 따라서 회귀분석 신경망에서 독립변수의 고차원에 따른 회귀성능의 저하와 같은 제약들은

제안된 기법의 PCA를 이용한 특징추출로 차원을 감소시키므로써 해결될 수 있을 것이다. 또한 표에서는 독립변수의 차원을 감소시키면 신경망의 학습성능도 좀 더 향상됨을 알 수 있다. 이는 입력조건의 차원을 감소시키면 신경망의 학습을 위한 데이터의 양이 줄어들어 학습시간이 감소되며, 또한 독립변수 상호간의 의존성에 따른 공선성을 줄일 수 있어 학습오차도 감소되기 때문이다.

〈표 4〉 입력조건의 변화에 따른 역전파 알고리즘과 제안된 알고리즘의 학습결과

		BP Algorithm			Proposed Algorithm		
독립변수 수, N		3	2	1	3	2	1
학습반복수, N _i		20000	20000	20000	19821 (18967,834)	17539 (16935,604)	17332 (16263,1069)
학습시간, t _{cpu}		35	34	33	34	32	32
절대 오차, AE	학습오차, E _t	0.029748	0.022531	0.074366	0.0001	0.0001	0.0001
	시험오차, E _t	0.329392	0.301452	0.239795	0.020863	0.019024	0.018457
합 오차, E _s		0.35914	0.323983	0.314161	0.020963	0.019124	0.018557

4.2 자동차 연비문제

자동차 연비문제는 10개의 독립변수에 대하여 1개의 종속 변수를 가진 회귀분석 문제이다. 이 문제는 실린더 수(x₁), 입방인치 단위의 배기량(x₂), 마력(x₃), 최중기어비인 최종 구동장치비(x₄), 자동차 무게(x₅), 순간가속도(x₆), V자형(0) 또는 직선형(1)의 엔진형태(x₇), 자동(0)이나 수동(1)이냐의 기어의 종류(x₈), 기어속도의 수(x₉), 그리고 기화기의 밸브수(x₁₀)에 대한 자동차 연비(y)와의 관계를 나타낸 것이다. 본 실험에서는 32개의 패턴데이터 중에서 25개는 학습 패턴으로 나머지 7개는 시험패턴으로 이용하였다. 이용된 MLP에서 입력층 뉴런수는 입력조건인 독립변수의 개수와 동일하게 하고, 은닉층과 출력층 뉴런수는 각각 10개와 1개로 구성하였다. 다음의 <표 5>는 실험에 이용된 자료의 일부를 나타낸 것이다. 여기에서도 입·출력 조건을 각각 최대값으로 나누어 1이하의 값으로 단순히 정규화한 데이터와 영평균 정규화 데이터를 각각 실험에 이용하였다.

〈표 5〉 자동차 연비자료

패턴수	입력 조건										출력 특성 y
	x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	x ₇	x ₈	x ₉	x ₁₀	
1	6	160	110	3.9	2.62	16.46	0	1	4	4	21
2	6	160	110	3.9	2.875	17.02	0	1	4	4	21
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
24	8	350	245	3.73	3.84	17.05	0	0	3	4	13.3
25	8	400	175	3.08	3.845	18.9	0	0	3	2	19.2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
32	4	121	109	4.11	2.78	18.6	1	1	4	2	21.4

<표 6>은 10개의 입력 및 은닉뉴런을 가진 다층신경망에 역전파 알고리즘을 적용하여 학습시킬 때, 입력조건의 단순한 정규화 및 영평균 정규화 패턴 각각에 대해 PCA 적용 유무에 따른 학습결과를 나타낸 것이다. 여기에서도 <표 2>에서처럼 PCA의 적용 유무에 관계없이 학습의 종

료조건은 만족되지 않았다. 또한 PCA를 적용하는 경우가 하지 않는 경우보다 더 작은 학습오차와 시험오차를 가지며, 입력조건의 단순한 정규화보다는 영평균 정규화의 패턴 데이터를 이용할 경우에 더 작은 합오차를 가짐을 알 수 있다. 이는 입력조건인 독립변수 값의 평균을 각각 구하여 영평균을 수행한 다음, 이들을 중앙으로 모은 후 정규화한 패턴 데이터를 이용하면 상대적으로 우수한 회귀성능의 신경망이 구현될 수 있음을 보여주는 것이다. 표에서도 학습 오차를 살펴보면, 영평균 정규화 패턴은 단순한 정규화 패턴에 비해 최대 약 168.7배 정도, PCA를 적용한 경우가 적용하지 않은 경우에 비해 최대 약 3.1배 정도 더 작은 값을 가진다. 한편 학습오차와 시험오차의 합으로 표현되는 합오차에서도, 영평균 정규화와 PCA를 적용한 패턴을 이용한 경우가 PCA를 적용하지 않은 단순 정규화 패턴을 이용한 경우에 비해 최대 약 4.7배 정도 작은 값을 가진다. 한편 3개의 독립변수를 가지는 암모니아 제조공정 문제와 비교할 때, 10개의 독립변수를 가지는 자동차 연비문제에서 학습성능과 회귀성능의 개선 정도가 더욱 더 증가됨을 알 수 있다.

〈표 6〉 입력조건의 정규화 및 영평균 정규화 패턴에 대한 PCA 적용유무에 따른 학습결과

		Normalized pattern data		Centering & normalized pattern data	
		non-PCA	PCA	non-PCA	PCA
학습반복수, N _i		20000	20000	20000	20000
학습시간, t _{cpu}		37	37	37	37
절대 오차, AE	학습오차, E _t	0.042676	0.024126	0.000447	0.000143
	시험오차, E _t	1.952232	0.857857	0.67318	0.425316
합 오차, E _s		1.994908	0.881983	0.673627	0.425459

이는 입력조건이 많은 복잡한 문제일수록 패턴데이터의 영평균 정규화의 효과는 크며, PCA의 적용에 따른 효과도 더욱 더 현저하게 개선됨을 확인할 수 있다.

<표 7>은 회귀용 신경망을 대상으로 자동차연비 자료 데이터를 영평균 정규화한 후 PCA를 적용한 데이터를 입력패턴으로 기존의 역전파 알고리즘과 제안된 알고리즘으로 각각 학습한 결과를 나타낸 것이다. 여기에서도 기울기하강으로부터 동적터널링으로의 전환조건은 오차의 변화가 10⁻⁶ 이하일 때로 하였다. 표에서 보면 기존의 역전파 알고리즘의 경우 학습반복수가 28467번에서 학습오차가 10⁻⁴ 이하로 감소되어 학습이 종료되었다. 이는 이용된 학습패턴에 대한 신경망의 학습에는 국소최적해가 존재하지 않기 때문으로 추측된다. 하지만 제안된 알고리즘의 경우는 18528번의 기울기하강과 211번의 동적터널링이 수행되어 학습이 종료되었다. 여기에서도 제안된 알고리즘에서는 기울기하강과 동적터널링의 반복적인 적용으로 결국 전역최적해로 수렴될 수 있음을 확인할 수 있다. 종료조건으로 학습오차가 10⁻⁴ 이하 일때를 기준으로 설정할 경우 학습반복수와 학습시간에서는 각각 약 1.5배와 약 1.1배 정도 개선되었으며, 합오차를 기준으로 보면 약 1.1배 정도 제안된 알고리즘은 기존의 역전파 알고

<표 7> 역전파 알고리즘과 제안된 학습알고리즘의 학습결과

		BP Algorithm		Proposed Algorithm
학습종료조건		20000	60000	20000
학습반복수, N_i		20000	28467	18739(18528,211)
학습시간, t _{cpu}		37	52	35
절대오차, AE	학습오차, E_i	0.000143	0.0001	0.0001
	시험오차, E_t	0.425316	0.092643	0.090545
	합 오차, E_s	0.425459	0.092743	0.090645

리즘에 비해 개선된 성능이 있음을 확인할 수 있다. 한편 종료조건을 학습반복수로 설정할 경우에도 합오차를 보면 약 4.7배 정도 개선된 성능이 있다. 결과적으로 압모니아 제조공정이나 자동차 연비 문제에서 제안된 알고리즘은 역전파 알고리즘에 비해 상대적으로 더 개선된 학습 성능과 회귀성능이 있음을 알 수 있다.

<표 8>은 역전파 알고리즘과 제안된 알고리즘 각각을 대상으로 PCA를 적용하여 독립변수 수 N을 조정하면서 학습한 결과를 나타낸 것이다. 표에서 독립변수 10은 원래의 입력조건 개수와 동일하며, 나머지 9에서 2까지는 독립변수에 대한 차원을 감소시켜 이를 입력조건으로 학습한 경우이다. 여기에서도 <표 7>에서처럼 제안된 알고리즘은 학습반복수와 학습시간, 그리고 절대오차에서 기존의 역전파 알고리즘에 비해 감소된 결과들을 보여 우수한 학습성능과 회귀성능이 있음을 알 수 있다. 먼저 역전파 알고리즘의 결과를 표에서 보면 독립변수가 10개인 경우는 학습조건을 만족하지 못하였으며, 절대오차로 표현되는 학습오차와 시험오차도 상대적으로 큰 값을 가짐을 알 수 있다. 그러나 PCA를 이용하여 차원을 감소시켜 학습한 경우 학습반복수나 학습시간에서 상대적으로 개선된 학습성능과 시험오차도 더욱 작음을 알 수 있다. 이는 독립변수 상호간의 의존성에 따라서는 학습성능의 저하뿐만 아니라 회귀성능도 감소되기 때문이다. 또한 제안된 알고리즘에서는 모든 독립변수의 개수에 대하여 전역최적해로 수렴된 결과를 보인다. 이는 제안된 알고리즘에서는 기울기하강에 따른 국소최적해로의 수렴제약은 동적터널링으로 해결될 수 있기 때문이다. 여기에서도 역전파 알고리즘에

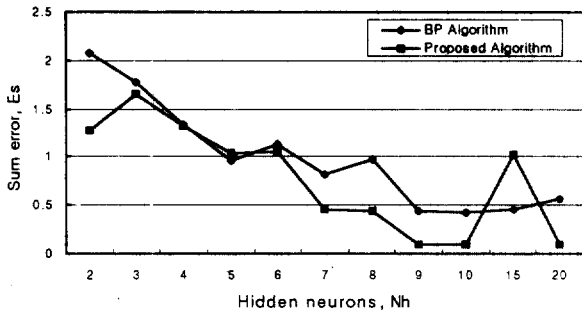
서와 같이 차원을 감소시켜 학습시키면 좀 더 개선된 학습 성능과 회귀성능을 얻을 수 있다. 특히 독립변수의 개수가 8개와 7개인 경우에는 각각 역전파 알고리즘과 동일한 결과를 보이는데 이는 이용된 학습패턴에 대해서 신경망을 학습할 경우에 국소최적해가 존재하지 않아 단순히 기울기하강만이 적용된 경우로 별도의 동적터널링이 적용되지 않았기 때문이다. 하지만 독립변수의 수가 5개 이하일 때에는 학습조건은 만족되나 시험오차가 증가되어 과학습에 따른 일반화 성능이 감소된 것으로 추측된다. 한편 두 알고리즘 모두의 결과들을 살펴보면, 독립변수의 개수가 9개에서 5개 사이는 학습성능과 회귀성능이 상대적으로 우수하지만 독립변수가 4개 이하인 경우는 상대적으로 저하된 성능을 보인다. 이는 전체 10개의 입력조건들 상호간의 속성들이 추출된 4개 이하의 독립변수에는 충분히 반영되지 않았기 때문으로 추측된다. 따라서 PCA에 의한 독립변수의 차원 감소는 각 변수 상호간에 많은 다중공선성을 가지는 회귀분석 문제에서는 우수한 성능이 있음을 추측할 수 있으며, 특히 입력조건 상호간에 선형적 연관성이 많을수록 더욱 더 우수한 성능이 있을 것이다. 이는 PCA 기법이 데이터 상호간의 선형성에 바탕을 두고 주요성분인 특징을 추출하여 차원을 감소시키기 때문이다.

(그림 3)은 영평균 정규화와 PCA를 적용한 10개의 독립변수를 가진 학습패턴으로 신경망을 학습시킬 경우 은닉층 뉴런수 N_h 에 따른 학습오차 E_i 와 시험오차 E_t 의 합오차 E_s 의 변화를 나타낸 것이다. 그림에서 보면 역전파 알고리즘이나 제안된 알고리즘 모두 은닉층의 뉴런수가 증가할 수록 합오차는 감소함을 알 수 있다. 그러나 실험에서는 은닉층 뉴런수가 증가되면 상대적으로 학습을 위한 계산시간이 증가되어 그에 따른 학습시간도 증가됨을 확인하였다. 그림에서 보면 은닉층 뉴런의 수가 8개에서 15개 사이가 상대적으로 좋은 결과를 보이며, 특히 9개와 10개일 경우가 가장 우수한 성능이 있음을 알 수 있다. 또한 전체적으로 보면 신경망의 구조에 따른 학습성능이나 회귀성능에서도 제안된 알고리즘은 역전파 알고리즘에 비해 우수한 학습알고리즘임을 알 수 있다.

<표 8> 입력조건의 변화에 따른 역전파 알고리즘과 제안된 알고리즘의 학습결과

		BP Algorithm								
독립변수 수, N		10	9	8	7	6	5	4	3	2
학습반복수, N_i		20000	20000	19957	19881	20000	20000	20000	20000	20000
학습시간, t _{cpu}		37	37	35	34	34	34	33	33	33
절대오차, AE	학습오차, E_i	0.000143	0.000128	0.0001	0.0001	0.000119	0.0009311	0.024663	0.517977	0.838667
	시험오차, E_t	0.425316	0.420921	0.106432	0.091259	0.37485	1.295529	1.485278	2.498355	2.148351
	합 오차, E_s	0.425459	0.421049	0.106532	0.091359	0.374969	1.30484	1.509941	3.016332	2.987018

		Proposed Algorithm								
독립변수 수, N		10	9	8	7	6	5	4	3	2
학습반복수, N_i		18739 (18528,211)	17542 (17404,138)	19957 (19957,0)	19881 (19881,0)	18534 (18087,447)	19623 (18571,1052)	16321 (16049,772)	18957 (17831,1126)	19362 (16263,1518)
학습 시간, t		35	33	35	34	33	35	30	33	33
절대오차, AE	학습오차, E_i	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
	시험오차, E_t	0.090545	0.090297	0.106432	0.091259	0.195732	0.943825	1.104726	2.362714	2.39416
	합 오차, E_s	0.090645	0.090397	0.106532	0.091359	0.195832	0.943925	1.104826	2.362814	2.39426



(그림 3) 은닉층 뉴런수 Nh에 따른 합오차 Es의 변화

따라서 이상의 실험결과들로부터 학습패턴 데이터의 단순한 정규화보다는 영평균 정규화 및 독립변수의 차원감소는 상대적으로 우수한 학습성능과 회귀성능을 보임을 확인할 수 있다. 또한 회귀용 신경망에 기울기하강과 동적터널링이 조합된 제안된 알고리즘의 이용은 기울기하강에 기초한 기존의 역전파 알고리즘을 이용할 때보다 상대적으로 우수한 학습성능과 회귀성능이 있음을 알 수 있다. 결국 영평균 정규화 및 차원을 감소시킨 학습패턴으로 본 연구에서 제안된 알고리즘의 회귀용 신경망을 이용한다면 좀 더 우수한 성능의 회귀분석이 이루어질 수 있다.

5. 결 론

본 논문에서는 다층신경망의 학습성능을 개선시키기 위한 조합형 학습알고리즘과 학습데이터의 차원을 감소시킨 고신뢰성의 신경망 기반 회귀분석 모델을 제안하였다. 조합된 알고리즘은 기울기하강과 동적터널링을 이용한 학습알고리즘으로 기울기하강의 빠른 수렴속도와 동적터널링의 전역최적해로의 수렴 특성을 동시에 살리기 위함이다. 또한 주요성분분석 기법이 가지는 대용량의 입력 데이터를 통계적으로 독립인 특징들의 집합으로 변환시키는 속성을 살려 학습패턴의 차원을 감소시킴으로써 데이터의 고차원에 따른 회귀분석 신경망의 제약을 해결하였다. 특히 학습데이터를 영평균 정규화함으로써 신경망의 학습을 좀 더 간략화하였다.

제안된 기법의 신경망을 3개의 독립변수를 가진 21개 학습패턴의 암모니아 제조공정문제와 10개의 독립변수를 가진 32개 학습패턴의 자동차 연비문제에 각각 적용하여 시뮬레이션한 결과, 학습패턴의 영평균 정규화가 단순한 정규화에 비해 개선된 성능이 있음을 확인하였다. 또한 역전파 알고리즘을 이용한 신경망이나 학습패턴의 차원을 감소시키지 않은 경우와 각각 비교할 때 학습성능과 절대오차로 나타나는 회귀성능에서 더욱 우수함을 확인할 수 있었다. 특히 2가지 회귀분석 문제에서 문제의 규모가 증가할 수록 제안된 알고리즘의 회귀용 신경망은 더욱 더 성능개선의 정도도 증가됨을 확인하였다.

향후 제안된 기법의 신경망을 좀 다양한 회귀분석 문제에 적용하고 회귀검증도 함께 고려되는 연구가 계속되어야 할 것이다. 또한 효율적인 응용을 위해서는 제안된 알고리즘의 신경망에서 다른 성능결정 요소들에 대한 연구도 뒤

따라야 할 것이다.

참 고 문 헌

- [1] 허명희 외 1, "SAS 회귀분석", 자유아카데미, 1996.
- [2] 강명욱 외 3, "회귀분석 : 모형개발과 진단", 울곡출판사, 1997.
- [3] S. Haykin, "Neural Networks : A Comprehensive Foundation," Prentice-Hall, London, 1999.
- [4] T. Masters, "Advanced Algorithms for Neural Networks : A C++ Sourcebook," John Wiley & Sons, Inc., New York, 1995.
- [5] D. F. Specht, "A General Regression Neural Network," IEEE Trans. on Neural Networks, Vol.2, No.6, pp.568-576, Nov. 1991.
- [6] M. Catusus, W. Branagh, and E. D. Salin, "Improved Calibration for Inductively Coupled Plasma-Atomic Emission Spectrometry Using Generalized Regression Neural Networks," Applied Spectroscopy, Vol.49, No.6, pp.798-807, 1995.
- [7] J. A. Freeman and D. M. Skapura, "Neural Networks : Algorithms, Applications, and Programming Techniques," Addison Wesley, London, 1991.
- [8] A. Cichock and R. Unbehauen, "Neural Networks for Optimization and Signal Processing," John Wiley & Sons., New York, 1993.
- [9] P. Foldiak, "Adaptive Network for Optimal Linear Feature Extraction," International Joint Conference on Neural Networks, Washington D.C., Vol.1, pp.401-406, June, 1989.
- [10] S. Bannour, A. Mahmood, and A. Sadjadi, "Principal Component Extraction Using Recursive Least Squares Learning," IEEE Trans. on Neural Networks, Vol.6, No.2, pp. 457-469, Mar. 1995.
- [11] K. I. Diamantaras and S. Y. Kung, "Principal Component Neural Networks : Theory and Applications, Adaptive and Learning Systems for Signal Processing, Communications, and Control," John Wiley & Sons, Inc., 1996.
- [12] Y. Yao, "Dynamic Tunneling Algorithm for Global Optimization," IEEE Trans. On Systems, Man, and Cybernetics, Vol.19, No.5, pp.1222-1230, Sept/Oct. 1989.
- [13] 조용현, 윤중환, 박용수, "조합형 학습알고리즘의 신경망을 이용한 데이터의 효율적인 특징추출", 정보처리학회논문지, 제8-B권 제2호, pp.130-136, April, 2001.
- [14] A. Hyvarinen and E. Oja, "Independent Component Analysis : Algorithms and Applications," Neural Networks, Vol. 13, No.4-5, pp.411-430, 2000.



조 용 현

e-mail : yhcho@cuth.cataegu.ac.kr

1979년 경북대학교 공과대학 전자공학과 (공학사)

1981년 경북대학교 대학원 전자공학과 (공학석사)

1993년 경북대학교 대학원 전자공학과 (공학박사)

1983년~1984년 삼성전자(주) 사원

1984년~1987년 한국전자통신연구소 연구원

1987년~1997년 영남이공대학 전자과 부교수

1997년~현재 대구가톨릭대학교 컴퓨터정보통신공학부 부교수

관심분야 : 신경망, 병렬분산처리, 신호처리, 영상인식, 교환기 등