Journal of the Korean
Data & Information Science Society
1998, Vol. 9, No. 2, pp. 345 ~ 355

로버스트 다층전방향 신경망을 이용한 패턴인식

황 창 하 [1] · 김 상 민 [2]

요약

다층전방향 신경망을 학습시키기 위해 역전파 알고리즘이 널리 사용되고 있으나 이 알고리즘은 긴 훈련시간, 극소점 문제, 이상치에 민감하다는 단점을 가지고 있다. 한편 실제문제에서는 많은 경우에 자료에 과대오차와 이상치가 포함되게 된다. 따라서 과대오차에 민감하지 않고, 이상치의 영향을 최소화시키는 로버스트 역전파 알고리즘의 필요성이 대두되었다. 본 논문에서는 기존의 두종류의 로버스트 역전파 알고리즘을 이론적으로 비교하고 비선형 회귀함수추정과 문자인식과 같은 패턴인식 문제에 적용하여 실험결과를 분석한다. 그리고 향후 연구과제로 신경망 학습을 위해 베이지안 기법의 사용을 제안한다.

주제어: 다층전방향 신경망, 로버스트, 문자인식, 역전파 알고리즘, 회귀함수

1. 서론

생물의 신경계가 정보를 어떻게 저장하고 다루는가에 관한 지식을 이용하려는 시도로부터 신경계산방법이 부상하게 되었고, 신경망이라 불리우는 인공신경계의 연구분야를 이끌어 내었다. 이러한 연구는 심리학, 신경과학, 인지학 및 시스템이론과 같은 매우 다양한 분야의 연구를 혼합한 것으로 최근에 새롭게 상당한 주목을 받고 있다. 신경망은 비교적 새로운 계산방식이다. 전통적인 계산방식에 대한 신경망의 장점, 단점, 응용성 및 관련성이 이해되지 않은 상태에 있으므로 신경망 영역에 대한 기대는 매우 높다. 신경망은 비교적 간단하면서도 많은 노드들이 상호 연결된 망으로 구성되어 있으며 사전지식과 내부시스템 동작에 대한 상세한 지식이 거의 필요 없기 때문에 패턴인식을 구현하기 위한 좋은 도구로 사용 되고있다.

패턴인식의 제 문제들은 패턴들의 군집(class)들이 이미 정의되어 있고 새로운 패턴이 주어지면 이를 어떤 군집으로 분류(classification) 하는 문제이다. 이런 문제들을 분류 또는 감

[1] 712-702 경북 경산시 하양읍 대구효성가톨릭대학교 정보통계학과 부교수
[2] 740-200 경북 김천시 삼락동 김천대학 전산정보처리과 조교수

독패턴인식(supervised pattern recognition)이라 부른다. 문제해결을 위해서는 먼저 누군가가 군집을 결정해야 하는데 이와 같이 패턴의 그룹을 찾는 것을 군집분석(cluster analysis) 또는 무감독패턴인식(unsupervised pattern recognition) 이라 부른다. 다층전방향 신경망을 이용한 감독패턴인식의 기본이론은 회귀분석이다.

컴퓨터의 급속한 발전으로 인하여 커널, 스플라인, PPR(Projection Pursuit Regression), 다층전방향 신경망 방법 등의 다양한 비모수 회귀함수 추정법들이 제안되었다. 커널 또는 스플라인 방법은 독립변수가 여러 개인 경우에는 차원문제(curse of dimensionality)로 자료의 개수가 적으면 고차원 공간에서 회귀함수를 잘 추정하지 못하는 경향이 있다. 한편 PPR, 신경망과 같은 투영(projection)을 이용한 방법은 차원문제를 극복하기 때문에 고차원 공간에서 회귀함수를 추정하기 위해서는 투영방법이 효과적이다. 패턴인식의 제 문제들은 고차원 공간에서 회귀함수를 추정하는 것과 연관이 있다.

비모수적 또는 준모수족(semiparametric) 방법으로 분류되는 다층전방향 신경망은 좀 더 유용한 모형을 만들기 위하여 적응모수(adaptive parameter)의 개수를 증가시킬 수 있는 일반적 함수형태를 취하며 모수의 개수는 자료의 크기와는 독립적으로 변화를 줄 수 있으므로 모수적 방법과 비모수적 방법의 장점을 동시에 가지고 있다. 그러나 다층전방향 신경망을 학습시키기 위해 주로 사용되는 역전파 알고리즘은 긴 훈련시간, 극소점 문제, 이상치에 민감하다는 단점을 가지고 있다. 한편 실제문제에서는 많은 경우에 자료에 과대오차와 이상치가 포함되게 된다. 따라서 과대오차에 민감하지 않으며 이상치의 영향을 최소화시키는 로버스트 역전파 알고리즘의 필요성이 대두되었다. Chen과 Jain(1994)은 Hampel의 tanh 함수를 이용하여 이상치에 덜 민감한 로버스트 역전파 알고리즘을 제안하였으며 황창하 등(1997) 은 통계물리에서 많이 응용되는 베이지안 이론을 적용하여 로버스트 역전파 알고리즘을 제안하였다.

본 논문에서는 황창하 등(1997)에 의해 제안된 로버스트 역전파 알고리즘을 조금 보완하여 두종류의 로버스트 역전파 알고리즘을 이론적으로 비교 분석하며, Wahba와 Wold(1975)가 사용한 일변량 비선형함수인 French curve와 Hwang 등(1994)이 사용한 5개의 이변량 비선형함수를 추정하는 문제와 박희주(1995)에 의해 제안된 필기체 한글인식 문제에 적용하여 성능을 평가하고자 한다. 그리고 향후 연구과제로 패턴인식의 제 문제에 베이지안 학습기법을 적용한 결과와 비교하는 것을 제안하고자 한다.

2. 통계물리에 근거한 로버스트 역전파 알고리즘

다음의 내용은 황창하 등(1997)에서 인용하였다. 자세한 내용은 논문을 참고하라. 로버스트 역전파 알고리즘을 유도하기 위해서 다음과 같은 일반화된 에너지함수를 사용한다.

$$E(V, \mathbf{W}) \;=\; \sum_{p=1}^{P}\sum_{j=1}^{m} V_p\, z\,(y_{pj}, \hat{y}_{pj}) \;+\; E_{prior}(\mathbf{V})$$

여기서 $z(y_{pj}, \hat{y}_{pj}) = \frac{1}{2}(y_{pj} - \hat{y}_{pj})^2$이고 V_p는 0 또는 1의 값을 가지는 확률변수이며 $\mathbf{V} = \{V_p, p = 1, \cdots, P\}$ 이다. 즉, V_p는 입력자료가 이상치이면 0, 아니면 1의 값을 갖는다. 한편 $E_{prior}(\mathbf{V})$는 $\{V_p\}$의 사전분포에 의해 공헌되어진 에너지의 양을 나타내며, 이것의 일반적인 선택은 다음과 같다.

$$E_{prior}(\mathbf{V}) = \eta \sum_{p=1}^{P} (1 - V_p).$$

이것은 $z(y_{pj}, \hat{y}_{pj}) < \eta$ 이면, $V_p = 1$이 되어 주어진 관측치가 표본으로 간주되고, 그렇치 않으면 $V_p = 0$이 되어 이상치로 간주된다.

한편 η는 지정된 반복회수에 도달할 때마다 계산되는 값으로 황창하 등(1997)에서는 η를 우측경계값으로 정의하였다. 그러나 여기서는 $3 \times s$로 수정하여 다소 향상된 결과를 얻었다. 그런데 s 대신에 모표준편차의 M-추정량을 사용해도 결과에는 별다른 차이를 발견하지 못했다. 목표는 V_p가 이진 값을 가진다는 제약조건 하에서 $\{V_p\}$와 $\mathbf{W}$에 관해 $E(\mathbf{V}, \mathbf{W})$를 최소화 시키는 것이다. 그런데 이 문제는 연속형변수와 이산형변수가 혼합된 경우의 최적화 문제이기 때문에 해석적인 해를 구할 수 없을 뿐 아니라 최급강하법을 사용하여 해를 구하는 것도 쉽지는 않다. 이런 문제점을 해결하기 위해 통계물리에서 많이 사용되는 Gibbs 분포, 즉 $P[\mathbf{V}, \mathbf{W}] = 1/Ze^{-\beta E[\mathbf{V}, \mathbf{W}]}$을 사용한다. 이때, Z는 관계식 $\sum_{\mathbf{V}} \int_{\mathbf{W}} P[\mathbf{V}, \mathbf{W}] = 1$을 만족한다. 따라서 $E(\mathbf{V}, \mathbf{W})$를 최소화하는 문제는 $P[\mathbf{V}, \mathbf{W}]$를 최대화하는 문제로 귀착된다. 그러나 이것 또한 연속형변수와 이산형변수가 혼합된 경우의 최적화 문제이기 때문에 어려움이 따른다. 따라서 이런 문제에 대한 하나의 해결책으로는 $\mathbf{W}$의 주변분포 $P_{margin}[\mathbf{W}]$를 구하여 이것을 최대화 시키는 것이다. 따라서 일반적인 역전파 알고리즘 중에서 delta를 조정하는 부분만을 수정하여 로버스트 역전파 알고리즘을 구현할 수 있다. 사실 이 방법을 가중치 벡터의 M-추정량을 구하는 방법으로 간주할 수 있기 때문에 로버스트성(robustness)에 대한 이론적인 연구가 더욱 필요하며 나아가 다음에 설명될 Chen과 Jain의 알고리즘과 이론적으로 비교 분석하면 좋을 것이다.

(1) 출력층 노드의 경우:

$$\begin{aligned}
\Delta v_{kj} &= \alpha \sum_{p=1}^{P} \frac{1}{1 + exp(\beta\{\sum_{j=1}^{m} \frac{1}{2}(y_{pj} - \hat{y}_{pj})^2 - \eta\})} \times \{y_{pk} - \hat{y}_{pk}\} \times h_{pj} \\
&\equiv \alpha \delta_{pk} h_{pj}, \\
\Delta \theta_k &= \beta \delta_{pk}
\end{aligned}$$

(2) 은닉층 노드의 경우:

$$\Delta w_{ji} = \alpha \sum_{k} \delta_{pk} v_{kj} \hat{y}_{pi}$$

$$\begin{aligned} &\equiv \alpha\delta_{pj}\hat{y}_{pi}, \\ \Delta\theta_j &= \beta\delta_{pj} \end{aligned}$$

여기서, α, β는 학습률이다. 한편 β는 신경망의 복잡도를 결정하는 모수로서 베이지안 기법을 사용하여 이론적으로 추정할 수 있으나 실제문제에서 그 값은 다소 경험적으로 결정된다. 자세한 내용을 위해 MacKay(1995)를 참고하라.

3. Chen과 Jain의 로버스트 역전파 알고리즘

Chen과 Jain(1994)은 일반적인 역전파 알고리즘이 이상치에 민감한것을 보완하여 로버스트 알고리즘을 유도하기 위해 Hampel의 tanh 함수를 이용한 에너지 함수 $E_R(\mathcal{W},\mathcal{T},t) = \sum_{p=1}^{P}\sum_{j=1}^{m}\phi_t(r_{pj})$를 사용한다. 여기서, $r_{pj} = y_{pj} - \hat{y}_{pj}$이고 $\phi_t(\cdot)$는 시간에 의존하는 함수로서 관계식 $\psi_t(r) = d\phi_t(r)/dr$을 만족한다. 그리고 $\psi_t(\cdot)$는 Hampel의 tanh 함수를 사용하여 다음과 같이 정의된다.

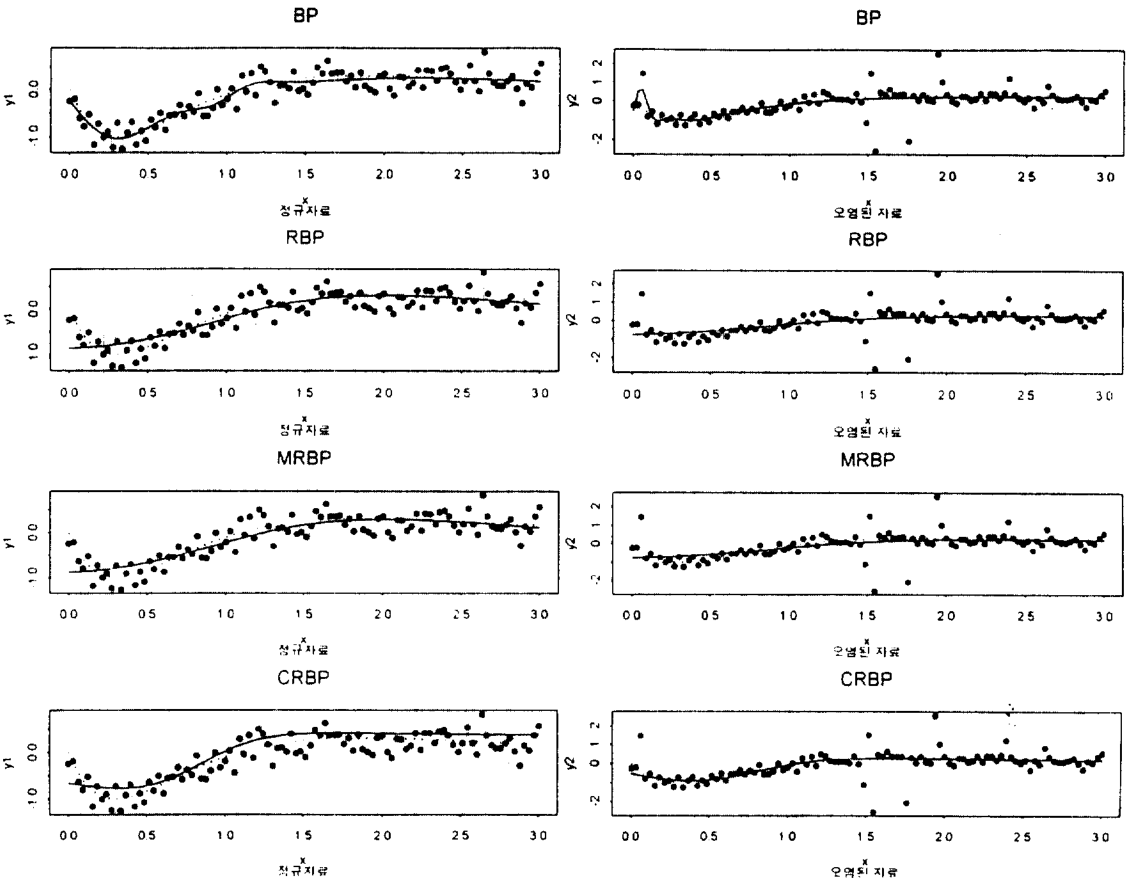
$$\psi_t(r) = \begin{cases} r, & |r| \le a(t) \\ c_1 tanh(c_2(b(t) - |r|))sign(r), & a(t) < |r| \le b(t) \\ 0, & |r| > b(t) \end{cases}$$

여기서 $a(t), b(t)$는 적절한 반복회수가 되면 잔차들 중에서 이상치 유무를 조사하여 이상치의 영향을 최소화 시키도록 조정되는 상수이며 c_1, c_2는 잔차들의 분포에 따라 결정되는 상수이다. 이 방법은 편의와 분산을 균형되게 통제하는 최적의 로버스트 M-추정량을 구하는 방법으로 알려져 있기 때문에 통계물리 이론을 이용해 구한 로버스트 역전파 알고리즘 보다 더 좋은 결과를 준다고 기대할 수 있다. 따라서 실제 패턴인식 문제에서 두 알고리즘의 성능을 비교하는 것은 의미가 있을 것이다. Chen과 Jain(1994)은 처음 일정한 학습회수(1000번) 동안에는 기존의 역전파 알고리즘을 사용하고 그후 가중치가 어느 정도 수렴한 상태에서 로버스트 역전파 알고리즘을 실행시켜 원하는 결과를 얻게된다. 자세한 내용을 위해 Chen과 Jain(1994)를 참고하라.

4. 패턴인식 예제

본 절에서는 일반적인 역전파 알고리즘(BP), 황창하 등(1997)에 의해 제안된 로버스트 역전파 알고리즘(RBP), 수정된 로버스트 역전파 알고리즘(MRBP) 및 Chen과 Jain의 로버스트 역전파 알고리즘(CRBP)을 Wahba와 Wold(1975)의 일변량 비선형함수인 French curve와 Hwang 등(1994)이 사용한 5개의 이변량 비선형함수를 추정하는 회귀분석 문제와 박희주(1995)에 의해 제안된 자소분리 기법을 이용한 필기체 한글인식 시스템에 적용하여

그림1: 일변량 비선형 회귀함수 추정 결과



알고리즘의 성능을 평가하고자 한다. 컴퓨터 실험을 위해 S-PLUS와 MATLAB을 사용하였다.

4.1 일변량 비선형 회귀함수 추정

본 절에서는 Wahba와 Wold(1975)가 제안한 French curve

$$y_i = f_{true}(x_i) + \epsilon_i, \quad i = 1, \ldots, n$$

에 대하여 다층전방향 신경망이 어떻게 일변량 비선형 회귀함수를 추정하는지 모의실험을 통하여 비교 분석한다. 여기서 $f_{true}(x) = 4.26(e^{-x} - 4e^{-2x} + 3e^{-3x})$ 이며, $\epsilon_i \sim N(0, 0.2^2)$ 이다. 입력값 x는 실수구간 상의 임의의 실수들로서 본 논문에서는 1과 3사이에서 랜덤으로 발생된 100개 값이다. 한편, 오염된 자료의 생성을 위해 $N(0, 9 \times 0.2^2)$에서 10개의 오차항을 생성하고 $N(0, 0.2^2)$에서 90개의 오차항을 생성하였다.

커널, 스플라인, PPR 및 베이지안 이론에 기초를 둔 신경망의 가중치감소법(weight decay) 등을 사용해서 실험을 하였는데 대체로 가중치감소법이 커널, 스플라인 및 PPR 보

다 더 좋은 결과를 보여주었다. 그러나 지면 관계상 결과를 생략하고 로버스트 다층전방향 신경망과 관련된 결과만을 언급한다. 그림에서 점선은 참 함수를 나타내고 실선은 추정된 함수를 나타낸다. RBP와 MRBP는 정규자료와 오염된 자료에 대해서 대체로 비슷한 결과를 보여준다. CRBP는 정규자료에 대해서 RBP, MRBP 보다 좋은 추정 결과를 보여준다고 단언하여 말할 수 없지만 전체적인 추세는 더 잘 파악한다고 말할 수 있다. 그리고 오염된 자료에 대해서는 CRBP가 다른 방법들에 비해서 더 좋은 추정 결과를 보여준다.

4.2 이변량 비선형 회귀함수 추정

알고리즘의 성능을 비교 분석하기 위해 Hwang 등(1994)이 사용한 5개의 이변량 비선형 함수를 사용하며, 측도로 FVU(fraction of variance unexplained)를 사용한다. 함수형태와 실험조건에 대한 자세한 내용을 위해 Hwang 등(1994) 또는 황창하 등(1997)을 참고하라.

황창하 등(1997)은 대부분의 함수들의 경우에 오차를 포함하지 않는 훈련자료 및 검정자료 그리고 오차를 포함하는 훈련자료에 대해서는 일반적 역전파 알고리즘과 RBP는 거의 비슷한 결과를 보여주고, 오차를 포함하는 검정자료에 대해서는 RBP가 다소 좋은 결과를 보여준다고 설명하였다. 표1은 모의실험의 결과를 설명한다. 함수의 종류에 따라 다른 결과를 보여주지만 전체적으로 볼때 MRBP와 CRBP는 거의 비슷한 결과를 보여주는 경향이 있으며 RBP 보다는 조금 향상된 결과를 보여준다.

4.3 필기체 한글인식

본 모의실험에서는 박희주(1995)에 의해 제안된 필기체 한글인식 시스템에 기존의 시스템과 동일한 조건하에서 실험하여 인식률을 분석하였다. 즉, Fahlman의 수정형 역전파 알고리즘을 사용한 기존의 필기체 한글 인식시스템과 RBP, MRBP 및 CRBP로 수정한 시스템을 자소인식 성능에 대해 모의실험을 통해 결과를 비교 분석하였다. 이에 사용되는 자료로 기존의 시스템에서 사용한 4명의 필기자가 작성한 1600자의 한글을 이용하여 문서에 대한 인식을 시도하여 오인식 문자수의 변화를 확인하였다.

표2는 필기자 1과 2의 필기체 한글을 학습표본으로 사용하여 신경망을 학습시킨 후에 필기자 1, 2, 3, 4의 필기체 한글을 검정한 결과를 설명한다. 표2에서 O3/T3은 MRBP로 학습하지 않은 필기자3과 필기자4를 포함하여 검정할 때 총자소 중 인식한 자소수, E3은 오인식 자소수, RT3은 인식률, O4/T4는 CRBP로 학습하지 않은 필기자3과 필기자4를 포함하여 검정할 때 총자소 중 인식한 자소수, E4는 오인식 자소수, RT4는 인식률을 나타낸다. 오광식 등(1998)의 실험결과에 의하면 RBP의 인식률은 95.4 % 이었다. 그러므로 RBP, MRBP 보다 CRBP가 조금 향상된 자소인식 능력을 보여주며 RBP 보다는 MRBP가 약간 더 좋은 결과를 보여주는 것을 알 수 있다.

표 1: FVU에 의해 결정되는 정확도

		오차가 없는 자료			오차가 있는 자료		
함수	방법	노드수	훈련표본	검정표본	노드수	훈련표본	검정표본
g(1)	RBP	5	0.000686	0.001755	5	0.057235	0.071344
		10	0.000581	0.002182	10	0.067744	0.079234
	MRBP	5	0.000678	0.001776	5	0.057167	0.071652
		10	0.000481	0.001824	10	0.067876	0.079440
	CRBP	5	0.000850	0.002155	5	0.058322	0.075243
		10	0.000481	0.002121	10	0.063257	0.075512
g(2)	RBP	5	0.013394	0.021717	5	0.122900	0.131182
		10	0.013820	0.023580	10	0.123967	0.131090
	MRBP	5	0.013394	0.022586	5	0.080296	0.088417
		10	0.013820	0.028300	10	0.080099	0.087122
	CRBP	5	0.014475	0.049152	5	0.072480	0.080160
		10	0.010884	0.026171	10	0.072651	0.079825
g(3)	RBP	5	0.326175	0.483644	5	0.427944	0.626164
		10	0.177148	0.294066	10	0.373888	0.550518
	MRBP	5	0.298503	0.465517	5	0.269904	0.496508
		10	0.118797	0.238858	10	0.239402	0.444205
	CRBP	5	0.298471	0.465887	5	0.269904	0.499790
		10	0.118793	0.253875	10	0.239402	0.456173
g(4)	RBP	5	0.033983	0.033271	5	0.093983	0.094813
		10	0.016381	0.019167	10	0.077504	0.081574
	MRBP	5	0.027429	0.026236	5	0.086860	0.089106
		10	0.012840	0.014748	10	0.076280	0.080741
	CRBP	5	0.027429	0.026211	5	0.082452	0.087213
		10	0.012805	0.016615	10	0.071150	0.080224
g(5)	RBP	5	0.284076	0.309024	5	0.329137	0.336245
		10	0.240227	0.262676	10	0.334221	0.362864
	MRBP	5	0.223241	0.252381	5	0.267965	0.294188
		10	0.197427	0.216802	10	0.263774	0.286464
	CRBP	5	0.223052	0.255942	5	0.260019	0.268621
		10	0.197363	0.220246	10	0.240722	0.251081

표 2: 필기자 1, 2 학습 후, 필기자 1, 2, 3, 4로 검정한 결과

유형			MRBP			CRBP		
TYPE	갯수	Jaso	O3/T3	E3	RT3	O4/T4	E4	RT4
유형1	80	초	309/320	11	96.6	310/320	10	96.9
		중	311/320	9	97.2	312/320	8	97.5
유형2	44	초	173/176	3	98.3	173/176	3	98.3
		중	167/176	9	94.9	170/176	6	96.6
유형3	15	초	58/60	2	96.7	58/60	2	96.7
		중	53/60	7	88.3	56/60	4	93.3
유형4	162	초	609/648	39	94.0	612/648	36	94.4
		중	638/648	10	98.5	637/648	11	98.3
		종	639/648	9	98.6	640/648	8	98.8
유형5	84	초	326/336	10	97.0	326/336	10	97.0
		중	308/336	28	91.7	310/336	26	92.3
		종	322/336	14	95.8	321/336	15	95.5
유형6	15	초	57/60	3	95.0	58/60	2	96.7
		중	57/60	3	95.0	57/60	3	95.0
		종	58/60	2	96.7	58/60	2	96.7
계	400	평균		10.6	95.6		9.7	96.3

5. 결 론

다층전방향 신경망은 패턴인식을 위한 도구로서 제안되었으며 이 신경망의 학습 알고리즘으로 역전파 알고리즘이 널리 이용되어져 왔다. 역전파 알고리즘은 최급강하법을 이용하여 오차합이 최소화 되도록 가중치를 반복적으로 변화시켜 나가는 알고리즘이다. 패턴인식의 실제문제에서 많은 경우에 학습자료에 이상치 또는 과대오차가 존재한다. 본 논문에서는 이상치 또는 과대오차에 로버스트한 에너지 함수를 사용하여 유도된 두종류의 로버스트 역전파 알고리즘에 대해 이론적으로 분석하고 회귀함수추정과 문자인식에 적용하여 성능을 확인하였다. 실험 결과 대체로 Chen과 Jain(1994)의 로버스트 역전파 알고리즘이 통계물리를 이용한 황창하 등(1997)의 로버스트 역전파 알고리즘 보다 이상치에 덜 민감하고 더 좋은 결과를 제공하는 것을 확인하였다.

최근에 다층전방향 신경망의 학습을 위해 베이지안 이론을 많이 사용한다. 베이지안 이론은 신경망의복잡도(complexity)를 조정하여 과대적합을 피하도록 한다. 베이지안 이론을 사용한 신경망 학습은 주어진 자료에 대해 가중치벡터의 사후분포를 구하는 것을 말하는데 정규근사(normal approximation) 또는 MCMC(Markov Chain Monte Carlo) 기법을 사용한다. 자료에 이상치 또는 과대오차가 포함되어 있을 때 MCMC 기법을 사용하여 이들의 영향을 최소화하는 가중치벡터의 사후분포를 유도하여 회귀함수추정 또는 문자인식 등과 같은 패턴인식의 제 문제에 적용하여 로버스트 에너지 함수를 사용하여 유도된 로버스트 역전파 알고리즘의 결과와 비교하는 것은 향후의 연구과제로서 의미가 있을 것으로 생각된다.

참고문헌

1. Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford.

2. Chen, D. S. and Jain, R. C. (1994). A Robust Back Propagation Learning Algorithm for Function Approximation, *IEEE Transactions on Neural Networks*, 5, 467-479.

3. Müller, P. and Insua, D. R. (1998). Issues in Bayesian Analysis of Neural Network Models, *Neural Computation*, 10, 749-770.

4. Hampel, F. R., Rousseeuw, P. J., Ronchetti, E.M. and Stahel, W.A. (1986). *Robust Statistics-The Approach Based on Influence Function*, John Wiley, New York.

5. Hwang, J. N., Lay, S. R., Maechler, M., Martin, D. and Schimert, J. (1994). Regression Modeling in Back-Propagation and Projection Pursuit Learning, *IEEE Transactions on Neural Networks*, 5, 342-353.

6. MacKay, D. J. C.(1995). *Bayesian Methods for Neural Networks: Theory and Applications*, Course notes for Neural Networks Summer School, Cambridge University.

7. Neal, R. M. (1996). *Bayesian learning for neural networks*, Springer-Verlag, New York.

8. Ripley, B. D. (1995). *Statistical ideas for selecting network architectures*, in: Kappen, B. and Gielen, S. (eds), Neural Networks: Artificial Intelligence and Industrial Applications, Springer, London.

9. Wahba, G. and Wold, S. (1975). A completely automatic French curve, *Communications in Statistics*. 4, 1-17.

10. Xu, L. and Yuille, A. L. (1995) Robust Principal Component Analysis by Self-Organizing Rules Based on Statistical Physics Approach, *IEEE Transactions on Neural Networks*, 6, 131-143.

11. 박희주. (1995) 자소분리 알고리즘과 계층구조 신경회로망을 이용한 필기체 한글인식 시스템, 대구효성가톨릭대학교 대학원 박사학위논문.

12. 오광식, 김상민, 이동로 (1998). 문자인식을 위한 로버스트 역전파 알고리즘, 통계이론방법연구, 8권 2호 163-171.

13. 황창하, 김상민, 박희주. (1997). 회귀분석을 위한 로버스트 역전파 알고리즘, 한국통계학회논문집, 4권 2호 327-332.

Pattern Recognition using Robust Feedforward Neural Networks

Changha Hwang [3] · **Sangmin Kim** [4]

Abstract

The back propagation(BP) algorithm allows multilayer feedforward neural networks to learn input-output mappings from training samples. It iteratively adjusts the network parameters(weights) to minimize the sum of squared approximation errors using a gradient descent technique. However, the mapping acquired through the BP algorithm may be corrupt when errorneous training data are employed.

In this paper two types of robust backpropagation algorithms are discussed both from a theoretical point of view and in the case studies of nonlinear regression function estimation and handwritten Korean character recognition. For future research we suggest Bayesian learning approach to neural networks and compare it with two robust backpropagation algorithms.

[3] Associate Professor, Dept. of Statistical infomation, Catholic University of Taegu-Hyosung, Kyungbuk 712-702
[4] Assistant Professor, Dept. of Computer Information Processing, Kimchon University, Kyungbuk 740-200