

# 분산추정에 의한 LVQ 신경회로망의 최적 출력뉴런 분할에 관한 연구

°정 준 원      조 성 원

홍익대학교 공과대학 전자·전기공학부  
서울 마포구 상수동 72-1

A Study on Optimal Output Neuron Allocation  
of LVQ Neural Network using Variance Estimation

°Junwon Chung    Seongwon Cho

School of Electronic and Electrical Engineering, Hong Ik University

## 요약

본 논문에서는 BP(Back Propagation)에 비해서 빠른 학습시간과 다른 경쟁학습 신경회로망 알고리즘에 비해서 비교적 우수한 성능으로 패턴인식 등에 많이 이용되고 있는 LVQ(Learning Vector Quantization) 알고리즘의 성능을 향상시키기 위한 방법을 논의하고자 한다. 일반적으로 LVQ는 음(negative)의 학습을 하기 때문에 초기 가중치가 제대로 설정되지 않으면 발산할 수 있다는 단점이 있으며, 경쟁학습 계열의 신경망이기 때문에 출력 층의 뉴런 수에 따라 성능에 큰 영향을 받는다고 알려져 있다[1].

지도학습 형태를 지닌 LVQ의 경우에 학습패턴이  $n$ 개의 클래스를 가지고, 각 클래스 별로 학습패턴의 수가 같은 경우에 일반적으로 전체 출력뉴런에 대해서 (출력뉴런수/ $n$ )개의 뉴런을 각 클래스의 목표(desired) 클러스터로 할당하여 학습을 수행하는데, 본 논문에서는 각 클래스에 동일한 수의 출력뉴런을 할당하지 않고, 학습데이터에서 각 클래스의 분산을 추정하여 각 클래스의 추정분산에 비례하게 목표 출력뉴런을 할당하고, 초기 가중치도 추정분산에 비례하게 각 클래스의 초기 임의의 위치 입력벡터를 사용하여 학습을 수행하는 방법을 제안한다.

본 논문에서 제안하는 방법은 분류하고자 하는 데이터에 대해서 필요한 최적의 출력뉴런 수를 찾는 것이 아니라 이미 결정되어 있는 출력뉴런 수에 대해서 각 클래스에 할당할 출력 뉴런 수를 데이터의 추정분산에 의해서 결정하는 것으로, 추정분산이 크면 상대적으로 많은 출력 뉴런을 할당하고 작으면 상대적으로 적은 출력뉴런을 할당하고 초기 가중치도 마찬가지로 방법으로 결정하며, 이렇게 하면 정해진 출력뉴런 개수 안에서 각 클래스 별로 분류의 어려움에 따라서 출력뉴런을 할당하기 때문에 미학습 뉴런이 줄어들게 되어 성능의 향상을 기대할 수 있으며, 실험적으로 제안된 방법이 더 나은 성능을 보임을 확인했다.

## 1. 서론

LVQ(Learning Vector Quantization)는 Kohonen이 제안한 지도학습 방식의 경쟁학습 신경회로망으로 고전적인 Vector Quantization(VQ) 방식을 도입한 알고리즘이다. 이전부터 알려진 패턴분류 기법의 하나인 VQ는 패턴공간을 유한개의 참조벡터(reference vector)로 양자화하는 것이다. LVQ는 이러한 VQ기법을 발전시킨 것으로 참조벡터를 뉴런들간의 가중치 벡터로 대응시키고, 목적값을 가지고 지도학습을 통해 적절한 참조벡터를 구하는 방법이다. LVQ는 지도학습 방식을 사용하는 BP에 비해 빠른 학습시간과 다른 경쟁학습 신경회로망에 비해서 비교적 우수한 성능을 가지기 때문에 패턴인식 등의 분야에서 많이 이용되어 왔다.

LVQ는 다른 비지도학습 방식의 경쟁학습 알고리즘인 SCL(Simple Competitive Learning)이나 SOFM(Self-Organization Feature Map)에서와 다르게 학습 과정에서 학습패턴이 잘못 분류되는 경우에 대해서 음(negative)의 학습을 포함하고 있기 때문에 초기 가중치 벡터의 선정과 초기 학습률에 따라서 가중치 벡터가 패턴공간 밖으로 발산하는 경우가 생겨서 안정성에 문제가 발생할 수 있기 때문에 적절한 학습률과 초기 가중치 벡터 선정이 필요하며 이에 대한 많은 연구가 이루어져 왔다[4]. 또, LVQ는 다른 경쟁학습 신경회로망들과 마찬가지로 알고리즘의 성능이 신경회로망의 출력 뉴런수에 영향을 많이 받는다. 이에 대하여 가변 출력층 구조를 갖는 경쟁학습 신경회로망에 대한 연구가 많이 진행되어 왔으며[5], 이는 신경회로망의 성능을 향상시키기 위해서 필요한 출력뉴런의 수를 계속해서 늘려가는 방법을 이용하며, 대표적인 가변 출력층 구조의 경쟁학습 신경회로망으로는 ART(Adaptive Resonance Theory)가 있다[1][2].

일반적으로 LVQ는 각 클래스별로 학습패턴의 수가 같은 경우에는 학습시 이미 정해진 출력 뉴런을 각 클래스별로 동일한 수로 할당하여 학습을 수행한다. 본 논문에서는 위에서 서술한 가변 출력층 구조의 신경회로망을 이용해 학습 중에 필요한 출력 뉴런을 생성하는 것이 아니라 학습초기에 이미 정해진 출력 뉴런 수를 그대로 두고, 각 클래스별 학습패턴의 분류 어려움이 클래스별 분산에 비례한다고 보고, 각 클래스별 분산을 추정하여 그 분산에 비례하게 출력 뉴런 수를 분할하고, 가중치를 초기화하여 학습을 수행하는 방법에 대하여 논의하고자 한다.

## 2. LVQ 학습 알고리즘

지도학습 방식의 신경회로망인 LVQ는 고전적인 VQ 방식을 도입한 알고리즘이다. 이전부터 알려진 패턴분류 기법의 하나인 VQ는 패턴공간을 유한개의 참조벡터로 양자화하는 것이다. LVQ는 이러한 VQ기법을 발전시킨 것으로 참조벡터를 뉴런들간의 가중치 벡터로 대응시키고, 목적값(desired value)을 가지고 지도학습을 통해 적절한 참조벡터를 구하는 방법이다. 기본적인 구조는 다른 경쟁학습 신경회로망과 같이 단층 신경회로망 구조를 가지며, 승자 뉴런 선정 방법은 유클리디안 거리(Euclidean distance)거리가 가장 가까운 뉴런이 승자 뉴런으로 선정된다. 학습규칙은 지도학습 방식을 사용하여 (식 1)과 같이 입력벡터가 올바르게 분류되었는지 아닌지에 따라서 학습을 진행한다.

$$w_i(t+1) = \begin{cases} w_i(t+1) + \alpha(t) \cdot [x(t) - w_i(t)] & , \text{ if } T=c : x \text{ correctly classified} \\ w_i(t+1) - \alpha(t) \cdot [x(t) - w_i(t)] & , \text{ if } T \neq c : x \text{ incorrectly classified} \end{cases} \quad (\text{식 1})$$

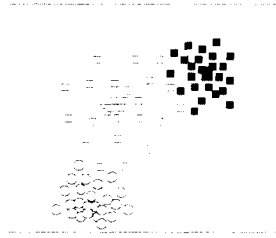
입력벡터에 대해 승자 뉴런(c)을 선정 후 해당 출력 뉴런의 목적(desired)클래스가 이미 알고 있는 입력벡터의 클래스 T와 같으면, 즉 입력벡터가 올바르게 분류되면 가중치 벡터가 입력벡터 쪽으로 가까워질 수 있도록 양(positive)의 학습을 수행하고, 분류가 올바르게 안되면 입력벡터와 멀어지도록 음(negative)의 학습을 수행하게 된다.

## 3. 제안된 출력뉴런 분할방법

### 3.1. 분산 추정

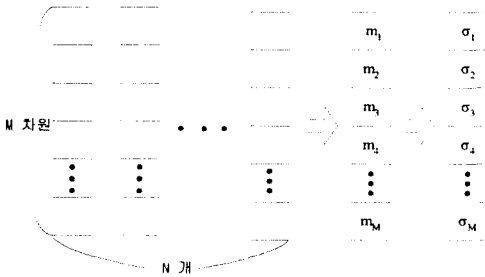
본 논문에서는 학습하고자 하는 패턴을 분류하는데 어려움의 기준에 따라서 이미 정해진 출력 뉴런수에서 좀더 분류가 어려운 클래스에 출력 뉴런을 비교적 많이 할당하고 적은 수의 가중치 벡터로 분류가 가능한 클래스는 비교적 적은 수의 출력 뉴런을 할당하여 가중치 벡터를 초기화하고 학습을 수행하여 알고리즘의 성능 향상을 도모한다. 여기서 학습패턴이 Gaussian분포를 가진다는 가정 하에 학습패턴이 분류

하기 어렵다는 것의 기준을 학습패턴의 분산과 밀접한 관계가 있다고 보고, 학습패턴의 각 클래스별 분산을 추정한다. 학습데이터가 (그림 1)과 같이 분포한다고 할 때, 데이터가 얼마나 넓게 분포되어 있는가를 나타내주는 분산에 따라서 각 클래스별로 비슷한 정도로 학습되도록 하려면 분산이 큰 경우에는 상대적으로 좀 더 많은 출력 뉴런이 필요함을 알 수 있다.



(그림 1. 입력데이터의 분포)

학습패턴의 각 클래스별 분산은 다음과 같은 방법으로 추정한다. 우선 (그림 2)와 같이 학습데이터의 입력벡터가 M개의 차원을 가지며, 총 N개로 구성될 때, 클래스별로 전체 N개의 입력벡터에 대해서 각 차원별로 평균을 구해서 각 클래스별 입력벡터의 평균벡터를 구한다.



(그림 2. 분산 추정 과정)

다음으로 입력벡터의 각 차원별로 분산을 구한다. 이렇게 구한 각 클래스의 입력차원별 분산에서 각 클래스별로 가장 큰 값을 가지는 분산 값의 제곱근을 사용하여 해당 클래스의 추정분산으로 결정한다. 여기서 제곱근 값을 사용한 것은 추정분산값에 의해서 상대적으로 출력 뉴런수를 결정하기 때문에 상대적으로 추정분산이 너무 작은 경우에 출력 뉴런이 아예 할당되지

않는 경우가 있기 때문에 추정분산의 영향을 적절하게 적용하기 위함이다.

### 3.2. 출력뉴런 분할

LVQ알고리즘은 지도방식의 학습 알고리즘이기 때문에 학습 시작 전에 각 출력 뉴런에 목적 클래스를 부여해주고 학습을 시작하게 된다. 일반적으로 출력 뉴런수는 학습 초기에 미리 결정하며, 일반적으로 각 출력 뉴런의 목적 클래스는 학습데이터의 각 클래스별로 동일한 수로 분할하여 학습을 수행한다.

본 논문에서는 학습시작 전에 미리 정해지는 출력 뉴런수가 K개라고 하고, 학습데이터의 클래스 수가 L개라고 할 때, 각 출력 뉴런의 목적클래스를 앞에서 구한 추정분산에 비례하도록 분할하는 방법을 제안한다.

클래스별로 구한 추정분산을  $\sigma_i$ 라고 하고, 클래스별로 가지는 출력 뉴런에 할당될 목적클래스 수를  $OUT_i$ 라고 하면,  $OUT_i$ 는 다음의 (식 2)와 같이 결정된다.

$$OUT_i = \frac{\sigma_i}{\sum_{n=0}^L \sigma_n} \cdot K \quad (\text{식 2})$$

위와 같은 방법으로 각 클래스별로 몇 개의 출력 뉴런을 할당할 것인지 결정하고 해당되는 개수에 따라서 출력 뉴런을 분할하고, 목적클래스를 할당한다. 출력 뉴런의 초기 가중치벡터는 해당 목적클래스의 학습데이터의 일부를 이용하여 초기화하고, 학습을 수행한다.

### 4. 실험결과

본 논문에서 실험에 사용한 데이터는 항공 촬영된 원격 탐사 데이터로서, 미국 인디애나주 남부지역의 농작물 재배지역을 촬영한 것이며, 학습패턴은 (표 1)과 같이 총 8개의 클래스로 구성되어 있으며, 하나의 입력벡터는 8차원이고, 각 클래스당 학습을 위한 데이터 200개와 테스트를 위한 데이터 375개로 구성되어

있다.

클래스	영역설명
1	풀밭
2	옥수수
3	귀리
4	붉은 클로버
5	콩
6	밀
7	공터
8	호밀

(표 1. 농작물 데이터의 클래스별 명칭)

본 실험에서는 일반적인 LVQ1를 사용한 실험결과와 제안된 방법에 의한 실험결과를 비교한다. LVQ1의 가중치 벡터 초기화는 학습데이터의 초기 임의 위치의 입력벡터를 사용하였고, 초기 학습율은 0.3, 출력 뉴런수는 40개, 80개의 경우에 대해서 나누어 실험하였으며, 실험결과는 다음과 같다.

(표 2)는 출력 뉴런수가 40, 80개 일 때, 분할된 출력 뉴런 개수를 나타내며, 일반적인 방법의 경우에는 각각 클래스당 5개, 10개 씩의 출력 뉴런으로 분할된다.

(표 3)는 출력 뉴런수가 80개인 경우에 대한 비교 실험이고, (표 4)은 출력 뉴런수가 40개인 경우에 대한 결과이다.

	1	2	3	4	5	6	7	8
80개	6	16	14	7	6	15	6	10
40개	3	8	7	4	3	7	3	5

(표 2. 출력 뉴런 분할)

기존의 방법	학습(%)	96.41
	테스트(%)	94.37
제안된 방법	학습(%)	97.53
	테스트(%)	94.44

(표 3. 출력 뉴런이 80개인 경우)

기존의 방법	학습(%)	94.77
	테스트(%)	92.45
제안된 방법	학습(%)	96.40
	테스트(%)	93.84

(표 4. 출력 뉴런이 40개인 경우)

실험결과에서 볼 수 있듯이 기존의 방법에 비해

제안된 방법이 좀 더 나은 성능을 보임을 알 수 있다. 특히, 출력 뉴런수가 적은 경우(40개)에 대해서 성능의 차이가 남을 볼 수 있다.

## 5. 결론

본 논문에서는 LVQ 알고리즘의 성능을 향상시키기 위해 각 클래스별 분산을 추정하여, 추정분산에 비례하게 출력 뉴런을 분할하고, 가중치 초기화 및 학습을 수행하는 방법을 제안한다. 이미 결정되어 있는 출력 뉴런수에서 최적의 성능을 낼 수 있도록 출력 뉴런을 분할하여 효율적이며, 미학습 뉴런의 발생이 줄어들게 된다. 실험을 통해서 기존의 방법보다 우수한 성능을 보임을 확인하였다.

## 참고문헌

- [1] Laurene Fausett, Fundamentals of Neural Networks, Prentice Hall, New Jersey, 1994.
- [2] James A. Freeman, David M. Skapura, Neural Networks Algorithms, Applications, and Programming Techniques, Addison-Wesley, 1991.
- [3] Philip D. Wasserman, "Advanced Methods in Neural Computing", Van Nostrand Reinhold, 1993.
- [4] 유광엽, "영상인식을 위한 주성분 분석 특징추출과 LVQ신경회로망의 균일분포 초기 가중치 선정방법", 홍익대학교 석사학위 논문, 1996.
- [5] 정성엽, "가변 출력층 구조의 경쟁학습 신경회로망을 이용한 패턴인식", 홍익대학교 석사학위 논문, 1996.