

패턴 분류를 위한 LVQ 네트워크의 성능 개선

정경권^{*} · 이정훈^{*} · 김주웅^{*} · 손동설^{**} · 엄기환^{*}

^{*}동국대학교 전자공학과 · ^{**}유한대학 전자과

Performance Improvement of LVQ Network for Pattern Classification

Kyung Kwon Jung^{*} · Jeong Hun Lee^{*} · Joo Woong Kim^{*} · Dong Seol Son^{**} · Ki Hwan Eom^{*}

^{*}Dongguk University · ^{**}Yuhan College

E-mail : kihwanum@dongguk.edu

요 약

본 논문에서는 LVQ 네트워크의 패턴 분류 성능을 향상시키기 위하여 n 차원 입력 벡터로 형성되는 초구면(hypersphere)에서의 반지름을 LVQ 네트워크의 학습 방식을 제안한다. 제안한 방식은 동일한 클래스에서 $n+1$ 개의 입력 벡터를 포함하는 초구면을 구성하여 반지름을 계산한다. 계산된 반지름을 이용하여 기준 벡터(reference vector)를 조정하는 방법이다. 제안한 방식의 유용성을 확인하기 위하여 Fisher의 IRIS 꽃의 분류 문제에 적용하여 기존의 LVQ 방식보다 우수한 분류 성능을 확인하였다.

ABSTRACT

In this paper, we propose a learning method of the performance improvement of the LVQ network using the radius of the hypersphere with the n -dimensional input vectors. The proposed method determines the reference vectors using the radius of the hypersphere include $n+1$ set of input vectors in the same class. In order to verify the effectiveness of the proposed method, we performed experiments on the Fisher's IRIS data. The experimental results showed that the proposed method improves considerably on the performance of the conventional LVQ network.

키워드

LVQ, pattern classification, hypersphere

1. 서 론

패턴인식의 정의는 데이터로부터 중요한 특징이나 속성을 추출하여 입력 데이터를 식별할 수 있는 부류로 식별하는 것을 말한다. 패턴 인식의 종류는 크게 지도 패턴인식(Supervised Pattern Recognition)과 비지도 패턴인식(Unsupervised Pattern Recognition)으로 구분된다. 비지도 패턴 인식은 임의의 그룹들을 서로 다르다고 구분만 하는 것을 말하며, 지도 패턴인식은 각 그룹에 명칭을 부여하는 것을 말한다. 패턴 인식시스템을 설계할 때의 기본적인 문제는 인식하고자 하는 대상으로부터 측정할 수 있는 입력 데이터를 표현하는 방법이다. 그 다음은 받아들인 입력 데이터의 특징과 속성을 추출하고 이들의 패턴 벡터의 차수를 줄이는 문제이며 이들의 구별 및 분류 과정에 최적의 결정 절차를 선택하는 문제이다. 이를 위하여 여러 가지 형태의 패턴 인식 알고리

즘이 제안되고 있다. 특히, 군집화와 신경 회로망은 이러한 문제를 해결하고 개선하는 데 많은 역할을 하고 있다[1-3].

패턴을 보다 정밀하게 분류하기 위하여 하나의 클러스터 혹은 클래스에 여러 개의 서브 클래스를 둘 수 있다. 따라서 학습을 통하여 서브 클래스들을 분류하고 분류된 서브 클래스들을 하나의 클래스로 지정한다면 패턴 분류의 성능은 보다 정밀하고 우수한 분류 성능을 갖게 된다. 이러한 클래스 속의 서브 클래스를 학습시켜서 클래스로 분류하는 신경회로망이 바로 LVQ(Learning Vector Quantization) 네트워크이다[4-5].

본 논문에서는 LVQ 네트워크의 성능을 향상시키기 위하여 n 차원 입력 벡터로 형성되는 초구면(hypersphere)에서의 반지름을 LVQ 네트워크의 학습에 이용하는 방식을 제안한다. 제안한 방식은

동일한 클래스에서 n+1개의 입력 벡터를 포함하는 초구면을 구성하여 반지름을 계산한다. 계산된 반지름을 이용하여 기준 벡터(reference vector)를 조정하는 방법이다. 제안한 방식의 유용성을 확인하기 위하여 Fisher의 IRIS 꽃의 분류 문제에 적용하여 기존의 LVQ 방식보다 우수한 분류 성능을 확인한다.

II. LVQ 네트워크

Kohonen이 제안한 LVQ 알고리즘은 클래스를 잘 구분 할 수 있도록 경계를 형성하는 것이 목적이다. Kohonen은 LVQ를 SOFM(Self-organizing Feature Map)에 분류 성능을 향상시키기 위해 제안하였으나, 단독으로 그림 1과 같은 LVQ 경쟁 신경회로망을 학습하는 데에 사용되기도 한다 [4-5].

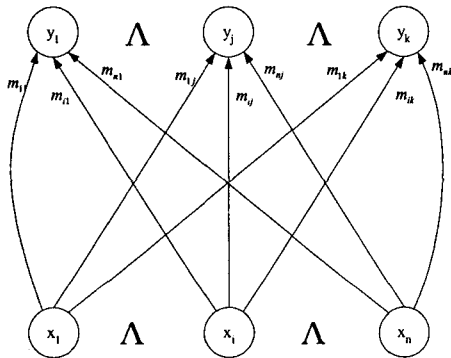


그림 1. LVQ 네트워크

경쟁 학습 신경회로망 중에서 음성 인식, 문자 인식, 지문 인식 등과 같은 패턴 인식뿐만 아니라, 제어, 통신 분야 등에서 널리 활용되고 있는 LVQ는 지도 학습과 비지도 학습이 동시에 존재하는 형태로 구성되어 있으며, 연결 벡터(weight vector)를 원형 벡터(prototype vector)로 하여 학습하는 알고리즘을 사용한다. LVQ의 학습은 입력과 출력 사이의 연결 벡터인 기준 벡터(reference vector)와 입력 벡터의 거리를 비교하여, 기준 벡터와의 거리가 가장 가까운 입력 벡터가 승리 벡터가 되고, 기준 벡터와 승리 벡터가 같은 클래스(class)로 분류되면 연결 벡터를 증가시키고, 다른 클래스로 분류되면 연결 벡터를 감소시킨다. 이러한 과정을 통해 입력 패턴은 분류가 되고 연결 벡터는 학습하게 된다[3][6-7].

LVQ는 잘못 분류된 출력 뉴런의 가중치 벡터에 대해서 결정 경계의 반대 방향으로 밀어냄으로써 분류 성능을 향상시키는 방법이다.

입력 벡터를 $x \in \mathbb{R}^n$ 이라 하고, LVQ 네트워크의 기준 벡터를 $m_i \in \mathbb{R}^n$ 이라 하면 LVQ 네트워

크는 입력 x 에 가장 근접한 m_c 를 찾는 것이다 [8-9].

$$\begin{aligned} \|x - m_c\| &= \min \|x - m_i\| \\ c &= \arg \min \|x - m_i\| \end{aligned} \quad (1)$$

왜곡 척도(distortion measure)를 식 (2)와 같이 정의하고, E 가 최소가 되도록 gradient를 계산하면 최적의 m_i 를 구할 수 있다.

$$E = \int \|x - m_c\|^2 p(x) dx \quad (2)$$

여기서 $p(x)$ 는 입력 벡터의 확률 밀도 함수이다.

입력 벡터 x 에 대해서 승리 뉴런이 i 이고 그 기준 벡터를 $m_i(t)$ 라고 할 때, 입력 벡터 x 와 출력 뉴런 i 의 클래스가 같은 경우에는 다음 식에 의해 기준 벡터를 조정한다.

$$m_i(t+1) = m_i(t) + \alpha(t)(x - m_i(t)) \quad (3)$$

입력 벡터 x 와 출력 뉴런 i 의 클래스가 다른 경우에는 아래 식에 의해 조정된다.

$$m_i(t+1) = m_i(t) - \alpha(t)(x - m_i(t)) \quad (4)$$

승리 뉴런 i 를 제외한 모든 출력 뉴런의 가중치 벡터는 갱신되지 않는다. $\alpha(t)$ 는 학습율이다.

$$m_j(t+1) = m_j(t) \quad \text{for } j \neq i \quad (5)$$

III. 제안한 방식

LVQ 네트워크의 성능을 향상시키기 위하여 n 차원 입력 벡터로 형성되는 초구면(hypersphere)에서의 반지름을 LVQ 네트워크의 학습에 이용하는 방식을 제안한다.

우선 n 차원 공간에서 $n \geq 4$ 일 때 n 개의 점 $x = (x_1, x_2, \dots, x_n)$ 에 대하여

$$x_1^2 + x_2^2 + \dots + x_n^2 = R^2 \quad (6)$$

을 초구면이라고 하며, R 은 초구면의 반지름이다. 만약 초구면의 중심점을 m_i 라 하면 식 (6)은

$$\|x - m_i\|^2 = R^2 \quad (7)$$

이 된다[10].

$n+1$ 개의 입력 벡터를 포함하는 초구면은 식 (8)과 같이 계산할 수 있다[11].

$$\begin{pmatrix} x_1^2 + x_2^2 + \dots + x_n^2 & x_1 & x_2 & \dots & x_n & 1 \\ x_1^2 + x_2^2 + \dots + x_n^2 & x_1 & x_2 & \dots & x_n & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ x_1^2 + x_2^2 + \dots + x_n^2 & x_1 & x_2 & \dots & x_n & 1 \\ x_1^2 + x_2^2 + \dots + x_n^2 & x_1 & x_2 & \dots & x_n & 1 \end{pmatrix} = 0 \quad (8)$$

2장의 식 (2)에서 가장 근접한 m_c 를 찾는 과정을 수행 할 때, 입력 벡터 x 와 가까이 있는 $n+1$ 개의 데이터 셋을 구성한다.

식 (8)의 첫 행을 기준으로 여인자(cofactor)를 계산한다.

$$C_{ij} = (-1)^{i+j} |M_{ij}| \quad (9)$$

여기서 $|M_{ij}|$ 는 소행렬식이다[12].

여인자를 이용하여 초구면의 중심을 다음과 같이 계산한다.

$$x_c = -\frac{1}{2C_{11}} \begin{pmatrix} C_{12} \\ C_{13} \\ \vdots \\ C_{1(n+1)} \end{pmatrix} \quad (10)$$

초구면의 중심 x_c 를 이용하여 학습식을 식 (11)과 같이 구성한다.

$$\begin{aligned} m_i(t+1) &= m_i(t) + \alpha(t)(x_c - m_i(t)) \\ m_i(t+1) &= m_i(t) - \alpha(t)(x_c - m_i(t)) \\ m_j(t+1) &= m_j(t) \quad \text{for } j \neq i \end{aligned} \quad (11)$$

IV. 실험

제안한 방식의 유용성을 확인하기 위하여 Fisher의 IRIS 데이터 분류를 시뮬레이션한다. 패턴 데이터를 훈련용과 시험용으로 나누고 분류 에러율을 가지고 기존 LVQ1, LVQ2와 제안한 방식의 분류 성능을 비교하였다.

1936년 Fisher는 붓꽃(iris)의 꽃받침(sepal)과 꽃잎(petal)의 크기와 길이를 조사하여 이를 데이터로 만들었다. 이렇게 해서 만들어진 데이터가 Fisher의 IRIS 데이터이다[52-55]. IRIS 데이터는 붓꽃의 3 종류인 Setosa, Versicolor, Virginica에

대해서 꽃받침과 꽃잎의 크기 및 길이에 대해 각 종류마다 50개씩의 데이터로 구성되었다. 이러한 IRIS 데이터는 분류나 클러스터링 해석을 위한 데이터로 널리 사용되고 있다. 이 3개의 종류 중에서 Setosa는 나머지 두 종류 Versicolor, Virginica와 선형적으로 분리가 가능하지만 Versicolor와 Virginica는 데이터가 서로 겹쳐서 있다.

각각의 종에서 25개씩 75개의 데이터를 가지고 학습을 하였고, 나머지 75개의 데이터로 시험을 하여 분류 성능을 비교하였다.

LVQ 네트워크는 입력 뉴런 4개, 출력 뉴런 3개로 구성하였고, 학습율은 0.9에서 학습을 하면서 감소하게 설정하였다. 초기 기준 벡터는 훈련 데이터의 최대값과 최소값의 평균으로 정하였다.

훈련 데이터와 시험 데이터의 분류 결과는 표 1, 2와 같다.

표 1. 훈련 데이터의 분류 결과

	학습횟수							
	100	200	300	400	500	1000	2000	5000
LVQ1	45	33	33	29	21	21	21	21
LVQ2	45	27	21	15	13	13	13	13
제안	45	23	14	11	10	10	10	10

표 2. 훈련 데이터의 분류 결과

	학습횟수							
	100	200	300	400	500	1000	2000	5000
LVQ1	53	47	40	36	29	29	29	29
LVQ2	53	34	31	19	17	16	16	16
제안	53	17	14	13	13	13	13	13

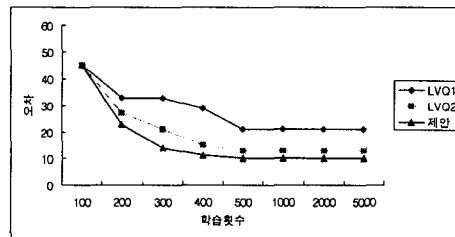


그림 2. 훈련 데이터의 분류 결과

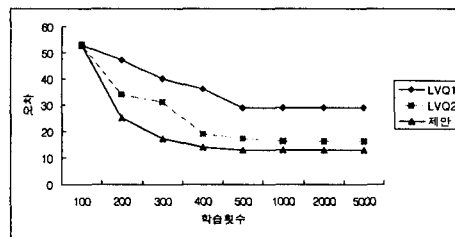


그림 3. 시험 데이터의 분류 결과
실험 결과 제안한 방식이 기존의 LVQ 방식보

다 분류 오차도 적고, 빠르게 수렴하는 것을 확인하였다.

V. 결 론

본 논문에서는 LVQ 네트워크의 성능을 향상시키기 위하여 n 차원 입력 벡터로 형성되는 초구면(hypersphere)에서의 반지름을 LVQ 네트워크의 학습에 이용하는 방식을 제안하였다. 제안한 방식은 동일한 클래스에서 거리가 가까운 $n+1$ 개의 입력 벡터를 결정하고, 이 입력 벡터들을 포함하는 초구면을 구성하여 반지름을 계산한다. 계산된 반지름을 이용하여 기준 벡터(reference vector)를 학습하는 방법이다. 제안한 방식의 유용성을 확인하기 위하여 Fisher의 IRIS 꽃의 분류 문제에 적용하여 기존의 LVQ 방식보다 우수한 분류 성능을 확인하였다.

참고문헌

- [1] Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, John Wiley & Sons, 2000.
- [2] 이성환, *패턴인식의 원리 I*, 홍릉과학출판사, 1994.
- [3] 김상운, *패턴인식 및 학습*, 홍릉과학출판사, 2003.
- [4] Fausett, Laurene, *Fundamentals of Neural Networks*, Prentice Hall, 1994.
- [5] Hagan, Martin T., *Neural Network Design*, Thomson Learning, 1996.
- [6] N. B. Karayiannis, "Soft learning vector quantization and clustering algorithms based on ordered weighted aggregation operators," *IEEE transactions on neural networks*, vol. 11, no. 5, pp.1093-1105, 2000.
- [7] S. C. Ahalt, A. K. Krishnamurthy, P. Chen, D. E. Melton, "Competitive Learning Algorithm for Vector Quantization," *Neural Networks*, vol. 3, pp. 277-290, 1990.
- [8] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464-1480, 1990.
- [9] T. Kohonen, "The self-organizing map," *Neurocomputing*, vol. 21, no. 1-3, pp. 1-6, 1998.
- [10] J. H. Conway, N. J. A. Sloane, *Sphere Packings, Lattices, and Groups*, Springer-Verlag, 1993.
- [11] D. Hilbert, S. Cohn-Vossen, *Geometry and the Imagination*, New York: Chelsea, 1999.
- [12] David C. Lay, *Linear Algebra and Its Applications*, Addison-Wesley Publishing, 2002.