ANALYSIS OF LEARNING VECTOR QUANTIZATION ALGORITHMS FOR PATTERN CLASSIFICATION

Ce Zhu [1], Jun Wang [2], and Taijun Wang [2]
[1]AI & PR Open Lab., Department of Electronic Engineering
Shantou University, Shantou 515063, Guangdong Prov., P.R. China
[2]Dept. of Radio Engineering, Southeast University
Nanjing 210096, P.R.China

ABSTRACT

Although the family of LVQ algorithms have been widely used for pattern classification and have achieved a great success, the rigirous theoretical studies on the classification performance of LVQ algorithms have seldom been made. In this paper, the asymptotical performance of LVQ1, LVQ2 and LVQ2.1 algorithms have been studied thoroughly, and three significant conclusions have been achieved respectively. Furthermore, a simple modification scheme to LVQ2 algorithm has been developed and analyzed on the asymptotical performance, which can produce the optimal or nearly-optimal classifier in the stable equilibrium state for the classification problems with classes overlapping.

1. INTRODUCTION

Learning Vector Quantization(LVQ) algorithms developed by Kohonen are a family of training algorithms for the nearest-neighbor classifiers, which include LVQ1, LVQ2 and its improved versions LVQ2.1, LVQ3 algorithms [1-3]. The family of LVQ algorithms are widely used for pattern classification such as speech recognition, and the satisfactory results are obtained [4]. Furthermore, in many cases, LVQ algorithms (e.g., LVQ2 algorithm) can achieve better results than other neural network classifiers in spite of their simple and time-efficient training process.

Although LVQ algorithms have achieved a great success in the application of pattern classification , and it is also generally thought that LVQ algorithms such as LVQ2 algorithm can produce the optimal classifier, the rigirous theoretical analysis of the classification performance for LVQ algorithms is seldom made [5,6], especially of the asymptotical performance for classification. In this paper, the asymptotical analysis of LVQ1, LVQ2 and LVQ2.1 algorithms for classification is made thoroughly, three significant conclusions have been derived respectively in Section 2, and an effective modification scheme to LVQ2 algorithm is developed and analyzed in Section 3

2. ASYMPTOTICAL ANALYSIS OF LVQ ALGORITHMS FOR CLASSIFICATION

Without loss of generality, in one-dimensional case, suppose there are two classes C_1 and C_2 with their respective mean vectors m'_1 and m'_2 . Class C_i is represented by its reference vector m_i , $i=1,2$. The family of LVQ algorithms are aimed at adjusting the reference vectors with different learning schemes to make them form the optimal classifier.

2.1. LVQ1 Algorithm

Firstly, we discuss the classification performance of LVQ1 algorithm for the linearly-separable problem (as shown in Fig.1) in the asymptotical sense. For LVQ1 algorithm, the equilibrium state equation can be derived

0-7803-2431-5/95 $4.00

as (suppose the learning step $\alpha(t) > 0$) [1]:

$$\int_{V_i} p(x|C_i)P(C_i)(x-m_i)dx - \sum_{j\neq i}\int_{V_i} p(x|C_j)P(C_j)(x-m_i)dx = 0 \quad (1)$$

where $i,j=1,2$, V_1 and V_2 are the nearest partition regions of m_1 and m_2 respectively. In the equilibrium state, suppose the reference vectors m_1 and m_2 converge to the optimal locations and form the optimal classifier with no classification error. Then we have $p(x|C_1)|_{x\in V_2}=0$, and $p(x|C_2)|_{x\in V_1}=0$. So Eq.(1) can be simplified as:

$$\int_{V_i} p(x|C_i)P(C_i)(x-m_i)dx = 0, \qquad (i=1,2) \qquad (2)$$

$$m_i = \frac{\int_{V_i} xp(x|C_i)P(C_i)dx}{\int_{V_i} p(x|C_i)P(C_i)dx}, \qquad (i=1,2) \qquad (3)$$

From Eq.(3), it can be seen that the values of the reference vectors m_1 and m_2 are actually equal to those of the mean vectors m'_1 and m'_2 respectively. That is to say, if the reference vectors m_1 and m_2 form an optimal classifier with no classification error in the equilibrium state, then the mean vectors m'_1 and m'_2 will also form the optimal classifier. Therefore, we can see that, if the mean vectors m'_1 and m'_2 can not make up an optimal classifier, LVQ1 algorithm will not produce such reference vectors as to form the optimal classifier in the equilibrium state. We can easily generalize and develop a significant conclusion as follows:

Theorem 1. For linearly-separable pattern classification problems, LVQ1 algorithm will not produce the optimal classifier unless the mean vectors of each pattern class can make up the optimal classifier.

2.2. LVQ2 Algorithm

For LVQ2 algorithm, we consider the classification problem with classes overlapping, as shown in Fig.2. In the same way, the equilibrium state equation of LVQ2 algorithm can be derived as:

$$\begin{cases} \int_{V_2} p(x|C_1)P(C_1)(x-m_1)dx - \int_{V_1} p(x|C_2)P(C_2)(x-m_1)dx = 0 & (4a) \\ \int_{V_1} p(x|C_2)P(C_2)(x-m_2)dx - \int_{V_2} p(x|C_1)P(C_1)(x-m_2)dx = 0 & (4b) \end{cases}$$

where V_1 and V_2 depicted in Fig.2 are two symmetric regions in the "window" defined in LVQ2 algorithm with respect to the centroid of the "window". Combining Eq.(4a) with Eq.(4b) as Eq.(4a)+Eq.(4b), we can obtain that $(m_1 \neq m_2)$:

$$\begin{cases} \int_{V_2} p(x|C_1)P(C_1)dx = \int_{V_1} p(x|C_2)P(C_2)dx & (5a) \\ \int_{V_2} xp(x|C_1)P(C_1)dx = \int_{V_1} xp(x|C_2)P(C_2)dx & (5b) \end{cases}$$

Make Eq.(5b) ÷ Eq.(5a), then

$$E[x|x \text{ belongs to class} C_1 \text{ and } x \text{ is located in the region} V_2] = E[x|x \text{ belongs to class} C_2 \text{ and } x \text{ is located in the region} V_1] \quad (6)$$

where $E[\cdot]$ denotes the *expectation operator*. Obviously, the above Eq.(6) is definitely not true. Therefore, we can see that Eq.(4a) and Eq.(4b) can not be simultaneously satisfied. The conclusion can be stated as:

Theorem 2. For the classification problem with classes overlapping, LVQ2 algorithm will not lead to any stable equilibrium state.

From the theorem, it can be inferred that LVQ2 algorithm can only be applied very limited times for the classes-overlapping classification problems, because no stable equilibrium state exists in that case. Otherwise, it will result in a very detrimental effect that the distance between the reference vectors m_1 and m_2 will become closer and closer [3], even turn to zero.

2.3. LVQ2.1 Algorithm

LVQ2.1 algorithm is an improved version of LVQ2 algorithm which aims at eliminating the detrimental effect described above[2]. Similarly, the equilibrium state equation of LVQ2.1 algorithm can be described as (for the same classification problem as shown in Fig.2):

$$\begin{cases} \int_{V_1+V_2} p(x|C_1)P(C_1)(x-m_1)dx \\ -\int_{V_1+V_2} p(x|C_2)P(C_2)(x-m_1)dx = 0 & (7a) \\ \int_{V_1+V_2} p(x|C_2)P(C_2)(x-m_2)dx \\ -\int_{V_1+V_2} p(x|C_1)P(C_1)(x-m_2)dx = 0 & (7b) \end{cases}$$

By means of the same derivation method as that for LVQ2 algorithm, we can obtain that the reference vectors m_1 and m_2 will reach the stable equilibrium state if and only if both of the following equations are simultaneously satisfied,

$$\begin{cases} \int_{V_1+V_2} p(x|C_1)P(C_1)dx = \int_{V_1+V_2} p(x|C_2)P(C_2)dx & (8a) \\ E[x|x \text{ belongs to } C_1 \text{ and } x \text{ is located in the "window" of } V_1+V_2] \\ = E[x|x \text{ belongs to } C_2 \text{ and } x \text{ is located in the "window" of } V_1+V_2] \\ & (8b) \end{cases}$$

Theorem 3. For the classification problem with classes overlapping, LVQ2.1 algorithm can lead to a stable equilibrium state if and only if Eq.(8a) and Eq.(8b) are simultaneously satisfied.

In the case of classes-overlapping with only one crossing point existing for the probability density functions of two classes (e.g., the case as shown in Fig.2), it can be also proved that Eqs.(8a) and (8b) will not be simultaneously satisfied. It is very rare that Eqs.(8a) and (8b) are simultaneously satisfied for the classes-overlapping classification problems (except for some special cases). Furthermore, LVQ2.1 algorithm will result in the effect of the distance between m_1 and m_2 becoming farther and farther for the problem as shown in Fig.2, which is opposite to that of LVQ2 algorithm.

3. A MODIFICATION SCHEME TO LVQ2 ALGORITHM

In order that the classification algorithms can lead to a stable equilibrium state corresponding to the optimal classifier, a slight modification to LVQ2 algorithm should be made that the values of adjustment for the two reference vectors m_1 and m_2 are taken the same during each step of iteration, denoted as $\Delta(t)$ (>0, where t denotes the t-th iteration). A simple selection for the adjustment value $\Delta(t)$ can be made as $\Delta(t) = (|x(t) - m_1(t)| + |x(t) - m_2(t)|)/2$. The direction of adjustment is the same as that for LVQ2 algorithm. In contrast to LVQ2 algorithm, the equilibrium state equation can be derived as:

$$\Delta_1 \int_{V_2} p(x|C_1)P(C_1)dx - \Delta_2 \int_{V_1} p(x|C_2)P(C_2)dx = 0 \quad (9)$$

where Δ_1 and Δ_2 denote the mean values of adjustment during the whole iterative procedure for m_1 and m_2 respectively. To clarify the meaning of Eq.(9) more intensively, the equation are simplified by applying the Intermediate Value Theorem of Integration as:

$$\Delta_1 p(x_2|C_1)P(C_1)\Delta V_2 = \Delta_2 p(x_1|C_2)P(C_2)\Delta V_1 \quad (10)$$

where x_1 and x_2 are two points in the region V_1 and V_2 respectively, ΔV_1 and ΔV_2 are the sizes of V_1 and V_2, and $\Delta V_1 = \Delta V_2$. Suppose $\Delta_1 = \Delta_2$, and because of the very narrow "window", x_1 and x_2 are very close to x_0 which is the centroid of the "window" (depicted in Fig.2), so the above equation can be further approximated as:

$$p(x_0|C_1)P(C_1) = p(x_0|C_2)P(C_2) \quad (11)$$

which is just the equation required by Bayes decision rule. Therefore, the modification scheme to LVQ2 algorithm will ultimately lead to a stable equilibrium state corresponding to the optimal or nearly-optimal classifier, which can obtain the optimal classifier by finely adjusting the value Δ_1/Δ_2.

4. CONCLUSION

In this paper, the asymptotical performance of LVQ1, LVQ2 and LVQ2.1 algorithms has been studied thoroughly in one-dimensional case, and three significant conclusions have been achieved respectively. Furthermore, a simple modification scheme to LVQ2 algorithm has been presented and studied on the asymptotical performance. The analysis shows that the modification scheme to LVQ2 algorithm can lead to a stable equilibrium state corresponding to the optimal or nearly-optimal classifier for the classes-overlapping classification problems. All the conclusions can be verified by experiments. As for the multi-dimensional cases, we can treat them as multiple one-dimensional cases.

References

[1] T.Kohonen, Self-Organization and Associative Memory. New York: Springer-Verlag,Third Edition, 1989.

[2] T.Kohonen, "Improved version of learning vector quantization," Proc. of IJCNN, San Diego, 1990, vol.1, pp.545-550.

[3] T.Kohonen, "The self-organizing map," Proc. IEEE, vol.78, no.9, pp.1464-1480, Sept.1990.

[4] E.McDermott and S.Katagiri, "LVQ-based shift-tolerant phoneme recognition," IEEE Trans. on Signal Processing, vol.39, no.6, pp.1398-1411, 1991 [5] Z.P.Lo, Y.Q.Yu, and B.Bavarian, "Derivation of learning vector quantization algorithms," Proc. of IJCNN, Maryland, 1992, vol.3, pp.561-566.

[6] Z.P.Lo, Y.Q.Yu, and B.Bavarian, "Analysis of a learning algorithm for neural network classifiers," Proc. of IJCNN, Maryland, 1992, vol.1, pp.589-594.

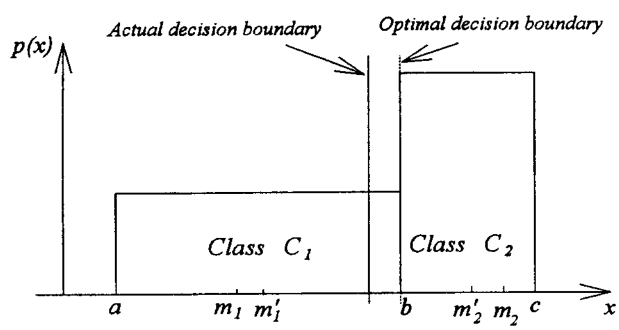

Fig.1 A linearly-separable classification example

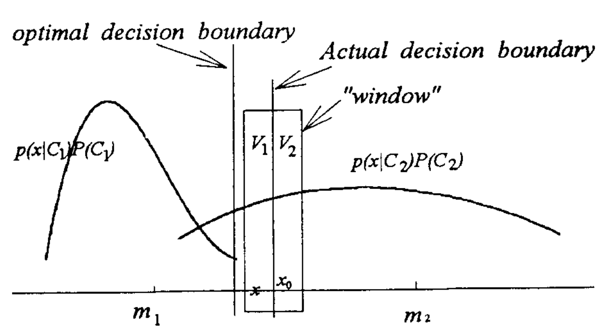

Fig.2 A classification example with classes overlapping