学习矢量量化的软竞争算法*

张志华, 郑南宁, 王天树

(西安交通大学 人工智能与机器人研究所,陕西 西安 710049)

E-mail: zhhaa@aiar.xjtu.edu.cn; nnzheng@xjtu.edu.cn; tswang@aiar.xjtu.edu.cn

http://www.xjtu.edu.cn

摘要: 尽管 FALVQ 算法的亏损因子为模糊隶属度函数,但由于它的尺度函数并不是模糊隶属度函数,使得算法的性能不稳定.为了克服这个问题,通过推广 FALVQ 中获胜亏损因子的定义,导出了广义 LVQ 的一类软竞争算法(SCALVQ),并且给出了它的 3 种具体形式.在 SCALVQ 中,亏损因子和对应的尺度函数是同一个模糊隶属度函数,它汲取了 FALVQ 和软竞争格式的优点,有效地克服了 FALVQ 存在的问题.

关 键 词: 模糊隶属度函数;亏损因子;尺度函数;干扰函数

中图法分类号: TP181 文献标识码: A

近 10 年来,由于 Kohonen 在学习矢量量化(learning vector quantization,简称 LVQ)算法和自组织特征映射(self-organization feature mapping,简称 SOFM)神经网络两方面[1,2]的许多开创性的工作,使得关于 Kohonen 聚类网络(Kohonen clustering network,简称 KCN)的研究变得十分流行.传统 KCN 的学习是基于胜者为王(winner-take-all,简称 WTA)或硬竞争(hard competitive,简称 HC)策略,它存在神经元未被充分利用以及输入样本和竞争神经元之间的信息被浪费等两个主要问题[3].而 SOFM 网络[1,2]需要选择学习率和被调整领域的大小等参数以及在学习过程中改变这些参数的策略.这些问题[4]同样制约了该网络的有效性.

目前,解决这些问题的主要方法是采用软竞争或模糊技术来学习码向量(表现型).比如,Yair 等人的软竞争格式[4](soft-competition scheme,简称 SCM)、Tsao 等人的模糊 Kohonen 聚类网络[5](fuzzy Kohonen clustering network,简称 FKCN)和 Chen 等人的无监督模糊竞争学习[3] (unsupervised fuzzy competitive learning,简称 UFCL)等等.广义学习矢量量化(generalized learning vector quantization,简称 GLVQ)网络是 Pal 等人[6]为改善 KCN 的性能提出来的.Karayiannis 等人通过对 GLVQ 网络的修正,提出了 GLVQ 网络的模糊算法(fuzzy algorithms for GLVQ,简称 FAGLVQ),即 GLVQ-F[7]和学习矢量量化的模糊算法(FALVQ)[8].

比较 SCM,FKCN,UFCL 和 FAGLVQ,我们发现:(1) 前者是一个启发性过程,后三者是最优化过程,其中 UFCL 和 FKCN 是一个约束最优化问题,FAGLVQ 是无约束最优化问题,但三者目标函数中的亏损因子都是模糊隶属度函数.进一步地说,UFCL 和 FKCN 的目标函数是关于亏损因子和表现型的函数,亏损因子满足某种约束条件,因而它们的学习过程是关于亏损因子和表现型交替迭代的过程;而在 FAGLVQ 中是事先按照某种规则选择好亏损因子,它的目标函数只是关于表现型的函数,其学习也就只能训练表现型.(2) 对于一个输入向量 x,上述 4 种算法的所有表现型都要进行训练,是软竞争学习.具体地说,前三者尺度函数值域为区间[0,1],且是模糊隶属度函数,而 SCM 的尺度函数还是一个概率密度函数;FAGLVQ 的尺度函数的取值范围复杂,通常不在区间[0,1]之间.(3) SCM,UFCL 和 FAGLVQ 采用在线学习方式,而 FKCN 是离线批处理方式.(4) SCM 的模糊(或软竞

* 收稿日期: 2000-03-08; 修改日期: 2000-11-07

基金项目: 国家自然科学基金资助项目(60175006);国家创新研究群体科学基金资助项目(60024301)

作者简介: 张志华(1969-),男,湖北阳新人,博士,讲师,主要研究领域为模式识别,智能信息处理;郑南宁(1952-),男,江苏南京人,博士,教授,博士生导师,中国工程院院士,主要研究领域为模式识别与人工智能,计算机视觉,数字电视技术;王天树(1977-),男,天津人,博士生,主要研究领域为数字视频,计算机视觉,模式识别.

争)含义体现在尺度函数是模糊隶属度函数;FAGLVQ 体现在亏损因子是模糊隶属度函数;而对于 FKCN 和 UFCL,模糊的含义表现在亏损因子和尺度函数都是模糊隶属度函数.

本文一方面利用 FAGLVQ 的无约束最优化模型等特征,同时吸取 SCM 的尺度函数是模糊隶属度函数,提出了一类新的 GLVQ 网络的模糊算法,称为 LVQ 的软竞争算法.

1 LVQ 的模糊算法(FALVQ)

考虑样本集合 X,它是一个 n 维实欧氏空间.设 $f(x)$是 $x\in X$ 的概率密度函数,LVQ 通常是基于最小化函数:

$$L(v_r,r=1,2,...,c)=\iint\cdots\int_{R^n}\sum_{r=1}^{c}u_r\cdot\|x-v_r\|^2\cdot f(x)\mathrm{d}x\,. \tag{1}$$

这里, $L(v_r,r=1,2,...,c)$ 表示由如下定义的亏损函数 $L_x=L_x(v_r,r=1,2,...,c)$ 的期望值.

$$L_x=L_x(v_r,r=1,2,...,c)=\sum_{r=1}^{c}u_r\cdot\|x-v_r\|^2\ . \tag{2}$$

以上定义中, $V=\{v_1,v_2,...,v_c\}$ 是我们需要设计的码本或表现型集合, $u_r=u_r(x)$,r=1,2,…,c 被称为对应于第 r 个表现型 v_r 的亏损因子.Pal 等人[6]基于求式(2)的最小值,利用梯度下降法导出了 GLVQ 网络.Karayiannis 等人[7~9]把 u_r 定义为一个模糊隶属度函数.Karayiannis 和 Pai[8]提出了选择隶属度函数的具体准则,根据这些准则,他们把 $u_r(x)$ 定义为如下形式：

$$u_r=u_r(x)=\begin{cases}1, & \text{if } r=i,\\ u\left(\dfrac{\|x-v_i\|^2}{\|x-v_r\|^2}\right)=u(z_r), & \text{if } r\neq i.\end{cases} \tag{3}$$

这里假设 v_i 是相对于输入向量 x 的获胜表现型, $z_r=\dfrac{\|x-v_i\|^2}{\|x-v_r\|^2}$,在本文其余部分都作如此假设.从式(3)可以看到, u_r 只与获胜表现型有直接关系,而与其他表现型无联系.基于上述,可导出网络的模糊算法的学习规则：

$$\nabla v_i L_x=-2\cdot(x-v_i)\cdot\omega_i\,, \tag{4}$$

$$\nabla v_j L_x=-2\cdot(x-v_j)\cdot\eta_j,(j=1,2,...,c;j\neq i)\,, \tag{5}$$

$$\omega_i=\left(1+\sum_{r\neq i}\omega_{ir}\right), \tag{6}$$

$$\omega_{ir}=u'(z_r),(r=1,2,...,c;r\neq i)\,, \tag{7}$$

$$\eta_j=u_j-z_j\omega_{ij}=u(z_j)-z_j\cdot u'(z_j)\,, \tag{8}$$

其中 ω_i 和 $\eta_r(r\neq i)$ 分别是获胜表现型和其他非获胜表现型的尺度函数, ω_{ir} 和 $\eta_r(r\neq i)$ 是干扰函数,它们分别表示非获胜表现型 v_r 对获胜表现型 v_i 的调整和获胜表现型 v_i 对非获胜表现型 v_r 的调整的干扰.只要事先根据某种规则选取 u_r ,就可以由式(6)~式(8)求出 ω_{ir} , ω_i 和 η_j $(j\neq i)$.文献[8]提出了 3 种算法 FALVQ1,FALVQ2 和 FALVQ3.设 u_r , ω_{ir} 和 η_r $(r\neq i)$ 的函数形式分别是 $u(z)$, $\omega(z)$ 和 $\eta(z)$ $(z\in(0,1))$,函数的表达式见表 1.

Table 1 Loss factors and interference functions for FALVQ

表 1 FALVQ 中的亏损因子和干扰函数

FALVQ family	$u(z)$	$\omega(z)$	$\eta(z)$
FALVQ1 ($0<\alpha<\infty$)	$z(1+\alpha z)^{-1}$	$(1+\alpha z)^{-2}$	$\alpha z^2(1+\alpha z)^{-2}$
FALVQ2 ($0<\beta<\infty$)	$z\exp(-\beta z)$	$(1-\beta z)\exp(-\beta z)$	$\beta z^2\exp(-\beta z)$
FALVQ3 ($0<\gamma<1$)	$z(1-\gamma z)$	$1-2\gamma z$	γz^2

1.1 基于干扰函数 $\eta_r(r\neq i)$ 构造FALVQ的方法

文献[9]给出了一种基于干扰函数 ω_{ir} 构造 FALVQ 的方法,这里我们提出一种新的基于干扰函数 $\eta_r(r\neq i)$ 构造 FALVQ 的方法.首先,由式(6)~式(8),可以直接得到下述关系：

$$\frac{\mathrm{d}\eta_j}{\mathrm{d}z_j} = -z_j \frac{\partial \omega_i}{\partial z_j} = -z_j \frac{\mathrm{d}\omega_{ij}}{\mathrm{d}z_j}, j = 1,2,\ldots,c; j \neq i\ . \tag{9}$$

也即

$$\eta'(z) = -z\omega'(z)\ . \tag{9$'$}$$

由式(9)可得 $\omega(z)$

$$\omega(z) = -\int z^{-1}\eta'(z)\mathrm{d}z + \mathrm{D} = -z^{-1}\eta(z) - \int \eta(z)\cdot z^{-2}\mathrm{d}z + \mathrm{D}\ , \forall z \in (0,1)\ , \tag{10}$$

D 是一个常数,它可以由 $\omega(z)$ 满足下述条件[9]来确定.

$$\lim_{z\to 0}\omega(z) = 1\ . \tag{11}$$

另外,由式(8)可得到 $u(z)$, $\omega(z)$ 和 $\eta(z)(z \in (0,1))$ 之间的关系：

$$u(z) = \eta(z) + z\cdot\omega(z)\ . \tag{12}$$

由此我们得到了 FALVQ 新的构造方法.简单地说,事先给定干扰函数 $\eta(z)$,然后利用式(10)和式(11),求出 $\omega(z)$,再根据式(12)确定 $u(z)$.

根据这个方法我们来分析 FALVQ2 算法.此时,已知 $\eta(z) = \beta z^2\cdot\exp(-\beta z), \forall z \in (0,1)$,由式(10)得到 $\omega(z) = (1-\beta z)\cdot\exp(-\beta z) + \mathrm{D}$.因为 $\lim_{z\to 0}\omega(z) = 1 + \mathrm{D} = 0$,所以 $\mathrm{D} = 0$.最后利用式(12),不难求得 $u(z) = z\cdot\exp(-\beta z), \forall z \in (0,1)$.其结果与表 1 是一致的.类似地,可以同理分析 FALVQ1 和 FALVQ3 的构造过程.

现在来讨论 η_j 的确定准则, η_j 事实上又是非获胜表现型的尺度函数.一种直观有意义的选择是 $\eta(z)$ 应满足下面条件(*):

$$0 < \eta(z) < \rho, \forall z \in (0,1)\ ; \lim_{z\to 0}\eta(z) = 0\ , \tag{*}$$

其中 ρ 是一个在区间(0,1]的常数,主要目的是控制非获胜表现型的尺度函数,使之与获胜表现型的尺度函数有一定的差距.

另一方面,倘若我们设

$$\eta(z) = z^2\cdot Q(z), \tag{13}$$

把它代入式(10),有

$$\omega(z) = -z\cdot Q(z) - \int Q(z)\mathrm{d}z + \mathrm{D}\ . \tag{14}$$

于是,若已知 $Q(z)$,由式(12)~式(14)便可以求出尺度函数和亏损因子,这简化了上述方法.选取 $Q(z)$ 是该方法的关键,由 $\eta(z)$ 满足的条件,直接得到 $Q(z)$ 应满足条件(**)：

$$0 \le Q(z) \le \frac{\rho}{z^2}, \forall z \in (0,1)\ ; \lim_{z\to 0} Q(z) = \text{不为 0 的常数 或 } \frac{1}{Q(z)} = o(z^2)\ . \tag{**}$$

1.2 FALVQ的性能分析

现在来分析 FALVQ 的学习性能.从式(4)和式(5)可知,若输入向量为 x,对于 FALVQ1 和 FALVQ2,由于 $0 \le \omega(z) \le 1$,所以 $1 \le \omega_i = 1 + \sum_{r\neq i}^{c}\omega_{ir} \le c$;而对于 FALVQ3,因为 $-1 \le \omega(z) \le 1$,则 $2-c \le \omega_i = 1 + \sum_{r\neq i}^{c}\omega_{ir} \le c$.因而它们不像 UFCL,FKCN 和 SCM 那样,尺度函数是模糊隶属度函数.从理论上来讲,如果算法是有效的,尺度函数在区间 $[0,1]$ 或 $[-1,1]$ 中取值是必要的,实际上我们还希望尺度函数是满足如下条件的模糊隶属度函数.

$$0 \le \eta_j \le \omega_i \le 1\ \ (j = 1,2,\ldots,c; j \neq i)\ . \tag{15}$$

所以,虽然把 $u_r (r = 1,2,\ldots,c)$ 定义为模糊隶属度函数,但是由此得到的学习算法却并不能令人满意.这是因为隶属度函数和尺度函数之间的关系很复杂.其实在应用中我们恰恰需要尺度函数为模糊隶属度函数,这一点在文献[10]中已得到证实,文献[10]直接把在 GLVQ-F 中用来定义亏损因子的模糊隶属度函数作为尺度函数,由此得到的算法的聚类性能较 GLVQ-F 有较大的改善,但遗憾的是此时算法变为一种启发式过程,也就是说,算法的调整方程不是通过最优化某个目标函数导出的,因而破坏了 GLVQ 网络的最优化过程这一显著优点.下面我们立足于 GLVQ 网络的基本思想,研究它的一类新的模糊算法.

2 GLVQ 网络一类新的模糊算法

给定一个输入向量,把与获胜表现型相对应的亏损因子定义为 1 是文献[8]提出的选取亏损因子的 4 个规则中的第 1 个规则,它的目的是为了保证获胜表现型较之非获胜表现型有一定的差距.而造成 FALVQ 上节所分析的问题恰恰主要是由于这个条件所导致的.因而,我们可以通过改变式(3)中 u_i 的定义来改善 FALVQ 的性能.

2.1 GLVQ的竞争算法的一般模型

对于亏损因子 $u_r(r=1,2,...,c)$ 的定义,式(3)的一个最直观的推广是

$$u_r=\begin{cases}u_i(z_1,z_2,...,z_c), & \text{if } r=i\\ u_r\left(\dfrac{\|x-v_i\|^2}{\|x-v_r\|^2}\right)=u(z_r), & \text{if } r\neq i\end{cases}, \tag{16}$$

也就是说 u_i 不一定等于 1,而是 $z_1,z_2,...,z_c$ 的一个函数.因而亏损函数式(2)变为

$$L_x=u_i\cdot\|x-v_i\|^2+\sum_{\substack{r=1\\r\neq i}}^{c}u_r\cdot\|x-v_r\|^2, \tag{17}$$

对 L_x 求导,经过简单计算,相应的学习规则为

$$\nabla v_jL_x=-2(x-v_j)\eta_j,\ (j=1,2,...,c;j\neq i). \tag{18}$$

$$\nabla v_iL_x=-2(x-v_i)\omega_i. \tag{19}$$

其中

$$\eta_j=u_j-\frac{\mathrm{d}u_j}{\mathrm{d}z_j}\cdot z_j-\frac{\partial u_i}{\partial z_j}\cdot {z_j}^2,\ j=1,2,...,c;\ j\neq i. \tag{20}$$

$$\omega_i=u_i+\sum_{\substack{r=1\\r\neq i}}^{c}\omega_{ir}. \tag{21}$$

$$\omega_{ir}=\frac{\mathrm{d}u_r}{\mathrm{d}z_r}+\frac{\partial u_i}{\partial z_r}\cdot z_r, r=1,2,...,c,\quad r\neq i. \tag{22}$$

上面导出的算法为 GLVQ 网络竞争算法的一般模型.选取不同的 u_i 就会产生不同形式的算法,当 $u_i=1$ 时,它就是 FALVQ,因而它推广了 FALVQ.

2.2 LVQ的软竞争算法(SCALVQ)

把 $u_r\ (r=1,2,...,c)$ 定义为式(16)推广了 FALVQ,虽然会使算法的形式更为丰富,但同时也使得亏损因子和尺度函数之间的关系变得更为复杂,而且如果亏损因子不是模糊隶属度函数,目标函数就失去了 FALVQ 中的物理意义.鉴于此,我们构造 GLVQ 算法的宗旨是通过对获胜亏损因子做某种限制或简化,使得到的尺度函数与对应的亏损因子一致,而且还是模糊隶属度函数.我们发现,若下式成立

$$\frac{\mathrm{d}u_j}{\mathrm{d}z_j}+\frac{\partial u_i}{\partial z_j}\cdot z_j=0,\ (j=1,2,...,c;\ j\neq i) \tag{23}$$

则可以改写式(18)、式(19)为

$$\begin{cases}\nabla v_jL_x=-2(x-v_j)\cdot u_j, & j\neq i\\ \nabla v_iL_x=-2(x-v_i)\cdot u_i, \end{cases}. \tag{24}$$

此时 $u_i=\omega_i$ 和 $u_j=\eta_j\ (j\neq i)$,这正是我们所期望的.如果亏损因子 $u_r\ (r=1,2,...,c)$ 同时还是模糊隶属度函数,就得到了 GLVQ 网络新的模糊算法.一方面该算法的亏损因子为隶属度函数,同时它的尺度函数也是隶属度函数,而且与亏损因子是同一个隶属度函数,所以它既保持了 FAGLVQ 特点,同时又具有 SCM 的特点,相对于 FALVQ 和 SCM,我们称它为学习矢量量化的软竞争算法(soft competitive algorithms for learning vector quantization,简称 SCALVQ).

2.3 构造SCALVQ的方法和3种SCALVQ

下面我们具体来讨论 SCALVQ 的构造方法.首先不难知道,条件(23)满足的必要条件是

$$u_i = \alpha_0 + \sum_{r \neq i} \alpha_r \cdot g_r(z_r), \tag{25}$$

其中 α_0 和 $\alpha_r (r \neq i)$ 是常系数, $g_r(z_r)$ 是一个只与 z_r 有关的干扰函数.为简化计算,令上式中的系数 $\alpha_0 = 1/c$ 和 $\alpha_r = 1/c$ $(r = 1,2,...,c; r \neq i)$,函数 $g_r(z_r) = g(z_r) = g_r$ $(r \neq i)$,则上式变为

$$u_i = \frac{1}{c} + \sum_{r \neq i} \frac{1}{c} \cdot g(z_r) . \tag{26}$$

于是式(23)就简化为

$$\frac{\mathrm{d}u_j}{\mathrm{d}z_j} + \frac{1}{c} \cdot \frac{\mathrm{d}g_j}{\mathrm{d}z_j} \cdot z_j = 0 , (j = 1,2,...,c; j \neq i) . \tag{27}$$

$$u'(z) + \frac{z}{c} \cdot g'(z) = 0 \quad \forall z \in (0,1) . \tag{27$'$}$$

根据上式,有

$$\int_0^z u(t)\mathrm{d}t = -\frac{1}{c}\int_0^z tg'(t)\mathrm{d}t = -\frac{tg(t)}{c}\bigg|_0^z + \frac{1}{c}\int_0^z g(t)\mathrm{d}t , \ \forall z \in (0,1) .$$

$$u(z) - u(0) = -\frac{zg(z)}{c} + \frac{1}{c}\int_0^z g(t)\mathrm{d}t .$$

由于 $u_j (j \neq i)$ 同时代表尺度函数,根据第 1.1 节的讨论,它应满足条件 $\lim_{z \to 0} u(z) = u(0) = 0$,于是

$$u(z) = -\frac{zg(z)}{c} + \frac{1}{c}\int_0^z g(t)\mathrm{d}t . \tag{28}$$

$$u(z_j) = -\frac{z_j g(z_j)}{c} + \frac{1}{c}\int_0^{z_j} g(t)\mathrm{d}t , \quad \forall \ j \neq i . \tag{28$'$}$$

由此我们得到了一个构造 SCALVQ 的方法事先选定 $g(z)$,然后由式(26)和式(28)′计算尺度函数 u_i 和 $u_j (j \neq i)$.由于 $u_j (j=1,2,\ldots,c)$ 还是模糊隶属度函数,所以要求得到的 $u_j (j=1,2,\ldots,c)$ 满足条件(15).下面的定理给出了基于这个要求选择 g_j 的条件.

定理 1. 假设函数 $g(z)$ 在区间 $[0,1]$ 上是递减函数,且满足 $0 < g(z) \leq 1$ ($\forall z \in [0,1]$).若 SCALVQ 中的获胜和非获胜尺度因子 u_i 和 $u_j (j \neq i)$ 分别由式(26)和式(28)′计算,则下述关系成立

(1) $\lim_{z \to 0} u(z) = u(0) = 0$;

(2) $\eta_j = u_j \leq u_r = \eta_r$,如果 $z_j \leq z_r, j \neq r \neq i$;

(3) $0 \leq \eta_j = u_j \leq u_i = \omega_i \leq 1$,对于任意的 $j \neq i$.

证明:因为 $g(z)$ 在区间 $[0,1]$ 上是递减函数,则 $g'(z) < 0$.又由式(28)和式(27)′,有 $u(0) = 0$ 且 $u'(z) > 0$,于是 $u(z)$ 在区间 $(0,1)$ 上是大于 0 的递增函数.再根据式(28)和定理 1 的条件,我们有

$$u(z) = -\frac{zg(z)}{c} + \frac{1}{c}\int_0^z g(t)\mathrm{d}t \leq \frac{z - zg(z)}{c} = \frac{z(1 - g(z))}{c} \leq \frac{1}{c} .$$

根据式(26)和定理的条件, $1/c \leq u_i \leq 1$ 显然成立,因而定理得证. □

这表明按照定理 1 的条件来选取干扰函数 $g(z)$,并根据式(26)和式(28)′来计算获胜和非获胜尺度因子 u_i 和 $u_j (j \neq i)$ 而确定的学习算法是有效的,它完全克服了在第 1.2 节所分析的 FALVQ 算法的弊端.根据定理 1,我们给出 3 种 SCALVQ,分别称之为 SCALVQ1,SCALVQ2 和 SCALVQ3.算法中的函数 $u(z)(\eta(z))$ 和 $g(z)$ 见表 2.

现仅以 SCALVQ 2 为例来讨论.首先给定 $g(z) = (1 - \beta z)\exp(-\beta z)$,因为 $\frac{\mathrm{d}g(z)}{\mathrm{d}z} = -\beta\exp(-\beta z)(2 - \beta z) \leq 0$,

$g(z)$是递减函数,且 $e^{-1} \leq g(z) \leq 1$,于是 $g(z)$满足定理条件.根据式(28), $u(z) = -\frac{z(1 - \beta z)\exp(-\beta z)}{c} + \frac{1}{c}\int_0^z (1 - \beta t)$ $\exp(-\beta t)\mathrm{d}t = \frac{1}{c}\beta z^2 \exp(-\beta z)$.又由于 $0 \leq \beta \leq 1$ 和 $0 < z < 1$, $\frac{\mathrm{d}u(z)}{\mathrm{d}z} = \frac{\beta z}{c} \cdot \exp(-\beta z)(2 - \beta z) \geq 0$.$u(z)$为区间(0,1)上

的递增函数,且 $0 \le u(z) \le \dfrac{\beta(2-\beta)}{c}\exp(-\beta) \le \dfrac{e^{-1}}{c}$, $\lim\limits_{z\to 0} u(z) = u(0) = 0$. $u(z)$ 满足条件(15).这从理论上验证了 SCALVQ 2 是有效的.

Table 2 Loss factors and interference functions for SCALVQ

表 2 SCALVQ 中的亏损因子和干扰函数

SCALVQ Family	$g(z)$	$u(z)\,(\eta(z))$
SCALVQ 1 ($0\le\alpha\le 1$)	$(1+\alpha z)^{-2}$	$\frac{1}{c}\alpha z^2(1+\alpha z)^{-2}$
SCALVQ 2 ($0\le\beta\le 1$)	$(1-\beta z)\cdot\exp(-\beta z)$	$\frac{1}{c}\beta z^2\exp(-\beta z)$
SCALVQ 3 ($0\le\gamma\le 1/2$)	$1-2\gamma z$	$\frac{1}{c}\gamma z^2$

3 数值实验及结果分析

这里我们用著名的 IRIS 数据来验证上述提出的算法的性能.IRIS 数据经常被用来检验聚类(无监督)或分类(有监督)模式识别方法的性能,它总共有 150 个数据,每个数据又由 4 个分量组成,IRIS 数据分为 3 类,每类有 50 个数据.对有监督设计标准的错误个数为 0~5,而无监督设计标准的个数在 15 左右.这里我们用 1-NP(最近邻)算法,SCALVQ1,SCALVQ2 以及 SCALVQ3 算法分别对 IRIS 数据进行聚类.在实验中,学习步长 $\alpha(0)=0.6$,最大迭代次数 T=200.表 3 给出了表现型的初始值[11],下面的实验都是根据这个初始值进行的.为了有一个比较准则,我们首先用最近邻法直接对样本进行聚类,结果见表 3.

Table 3 Iinitial centroids and the experimental results of 1-NP

表 3 数值实验中表现型的初试值及最近邻算法实验结果

Initial centroids①	Confusion matrix②			Error rate③(%)
5.006, 3.428, 1.462, 0.246	50	0	0	
5.936, 2.770, 4.260, 1.326	0	46	4	7.3
6.588, 2.974, 5.552, 2.026	0	7	43	

①初始中心,②混淆矩阵,③错误率.

表 4 给出了使用 SCALVQ1,SCALVQ2 和 SCALVQ3 三种算法的实验结果.为了更充分地验证这些算法的性能,每一种算法都取了不同参数值进行实验,结果表明,不但每一种算法对于不同参数值的聚类结果基本一致,而且 3 种算法的聚类结果也大致相同,这表明算法不但效果好,而且性能十分稳定,同时进一步验证出按照本文提出的理论框架构造的 GLVQ 网络的模糊算法 GFALVQ 是有效的.

Table 4 Experimental results of SCALVQ1, SCALVQ2 and SCALVQ3

表 4 SCALVQ1,SCALVQ2 和 SCALVQ3 三种算法的实验结果

SCALVQ1	α	1.0			0.5			0.0		
	Confusion matrix①	50	0	0	50	0	0	50	0	0
		0	43	7	0	43	7	0	43	7
		0	1	49	0	1	49	0	1	49
	Error rate②(%)	5.3			5.3			5.3		
SCALVQ2	β	1.0			0.5			0.0		
	Confusion matrix	50	0	0	50	0	0	50	0	0
		0	44	6	0	43	7	0	43	7
		0	1	49	0	1	49	0	1	49
	Error rate(%)	4.7			5.3			5.3		
SCALVQ3	γ	0.5			0.25			0.0		
	Confusion matrix	50	0	0	50	0	0	50	0	0
		0	43	7	0	43	7	0	43	7
		0	1	49	0	1	49	0	1	49
	Error rate(%)	5.3			5.3			5.3		

①混淆矩阵,②错误率.

4 结束语

GLVQ 网络是从最优化一个目标函数而导出的,为了克服 KCN 存在的问题提供了一个新的思路.Karayiannis 等人通过引入模糊隶属度函数来定义亏损因子提出 GLVQ 网络的模糊算法 GLVQ-F 和 FALVQ.为了有效地改善 FALVQ 的性能,本文基于把亏损因子与相应的尺度函数定义为相同的模糊隶属度函数的宗旨,提出了一套新的构造 GLVQ 网络的模糊算法的理论框架.在该框架的指导下,设计出一些高效的聚类方法并应用于图像处理问题中是我们今后需要进一步研究的工作.

References:

[1] Kohonen, T. Self-Organization Maps. Berlin: Springer-Verlag, 1995.

[2] Oja, E., Kaski, S. Kohonen Maps. Amsterdam: Elsevier, 1999.

[3] Chung, F.L, Lee, T. Fuzzy competitive learning. Neural Networks, 1994,7(3):539~551.

[4] Yair, E., Zeger, K., Gersho, A. Competitive learning and soft competition for vector quantizer design. IEEE Transactions on Signal Processing, 1992,40(2):294~308.

[5] Tsao, E.C.K, Bezdek, J.C, Pal, N.R. Fuzzy Kohonen clustering networks. Pattern Recognition, 1994,27(5):757~764.

[6] Pal, N.R., Bezdek, J.C., Tsao, E.C.K. Generalized clustering networks and Kohonen's self-organizing scheme. IEEE Transactions on Neural Networks, 1993,4(4):549~558.

[7] Karayiannis, N.B, Bezdek, J.C, Pal, N.R., *et al*. Repair to GLVQ: a new family of competitive learning schemes. IEEE Transactions on Neural Networks, 1996,7(5):1062~1071.

[8] Karayiannis, N.B, Pai, P.I. Fuzzy algorithms for learning vector quantization. IEEE Transactions on Neural Networks, 1996,7(5): 1196~1211

[9] Karayiannis, N.B. A methodology for constructing fuzzy algorithms for learning vector quantization. IEEE Transactions on Neural Networks, 1997,8(3):505~518

[10] Zhang, Zhi-hua, Zheng, Nan-ning, Wang, Tian-shu. Behavioral analysis and improvement of generalized LVQ neural network. Acta Automatica Sinica, 1999,25(5):583~589 (in Chinese).

[11] Bezdek, J.C, Pal, N.R. Two soft relatives of learning vector quantization. Neural Networks, 1995,8(5):729~743.

附中文参考文献:

[10] 张志华,郑南宁,王天树.广义 LVQ 神经网络的性能分析及及其改进.自动化学报,1999,25(5):583~589

Soft Competitive Algorithms for Learning Vector Quantization*

ZHANG Zhi-hua, ZHENG Nan-ning, WANG Tian-shu

(*Institute of Artificial Intelligence and Robotics*, *Xi'an Jiaotong University*, *Xi'an* 710049, *China*)

E-mail: zhhaa@aiar.xjtu.edu.cn; nnzheng@xjtu.edu.cn; tswang@aiar.xjtu.edu.cn

http://www.xjtu.edu.cn

Abstract: Though the loss factors in FALVQ algorithms are defined to fuzzy membership functions, the performance of the algorithms is not stable due to their scaling functions not being fuzzy membership functions. In this paper, a new family of fuzzy algorithms for the generalized LVQ network, called soft competitive algorithms for LVQ (SCALVQ), is derived from extending the definition of the loss factor corresponding to the winning prototype in FALVQ. Meanwhile, three concrete types of SCALVQ are given. In SCALVQ, the loss factors and the corresponding scaling function are both fuzzy membership functions, but an identical fuzzy membership function. Therefore, they absorb advantages of FALVQ and the soft-competition scheme, and overcome the disadvantages of the FALVQ.

Key words: fuzzy membership function; loss factor; scaling function; interference function

* Received March 8, 2000; accepted November 7, 2000

Supported by the National Natural Science Foundation of China under Grant No.60175006; the National Innovation Research Group Foundation of China under Grant No.60024301