

模糊超球神经网络的条件重叠学习算法*

华 强, 李师贤
(中山大学计算机科学系, 广东 广州 510275)

摘 要: 模糊超球神经网络的聚类学习算法采用形状因子 θ 实行有指导的学习。针对 θ 的引入所带来的问题, 提出一种新的无指导学习算法——条件重叠学习算法。算法不受模糊超球形状因子 θ 的影响, 学习速度快, 学习后的模糊超球个数更少, 识别正确率更高。

关键词: 模式识别; 模糊超球神经网络; 学习算法

中图分类号: TP183 **文献标识码:** A **文章编号:** 0529-6579 (2004) 01-0001-04

1 模糊超球神经网络及聚类学习算法

三层模糊超球神经网络的输入层输入 n 维模式 X_h , 其隐层用 n 维超球来定义某个类的样本群在 n 维模式空间中的一个全隶属区, 然后用聚类学习算法在输出层把独立的同类邻近超球并联起来, 形成该模式类的非规则边界。模糊极大极小神经网络等可视作模糊超球神经网络的相近形式。模糊超球以及相近类型的神经网络在医学图像分割和识别^[1]、变推力火箭发动机故障识别^[2,3]、飞机机型雷达信号识别^[4]、伽马望远镜图像识别^[5]等模式聚类和识别的实际问题中有着重要的应用。

模糊超球隶属函数 $b_j(X_h, W_j, v_j)$ 用来度量输入样本 $X_h = (x_{j1}, x_{j2}, \dots, x_{ji}, \dots, x_{jn})^T$ 与由球心 $W_j = (w_{j1}, w_{j2}, \dots, w_{ji}, \dots, w_{jn})^T$ 和半径 v_j 确定的模糊超球 HS_j 之间的相似关系。

$$b_j(X_h, W_j, v_j) = \exp\{-(\|X_h - W_j\| - v_j)^\gamma\} \\ \text{当} (\|X_h - W_j\| - v_j)^\gamma \geq 0 \\ b_j(X_h, W_j, v_j) = 1 \\ \text{当} (\|X_h - W_j\| - v_j)^\gamma < 0 \quad (1)$$

式中, $\|X_h - W_j\| = [(x_1 - w_1)^2 + (x_2 - w_2)^2 + \dots + (x_n - w_n)^2]^{1/2}$; 灵敏参数 γ 反映 X_h 远离超球 HS_j 时隶属度的降低速率^[6]。

模糊超球聚类学习算法对 $\|X_h - W_j\| > v_j$ 的同类训练样本 X_h 调整球心 W_j 和半径 v_j , 扩张超球 HS_j ; 并在扩张后进行异类超球间的重叠测试; 若存在异类超球重叠则进行超球收缩, 收缩的方法为重叠的异类超球分别收缩到重叠区的大约中点处,

或其中某个超球收缩后完全退出重叠区。

模糊超球聚类学习算法将模糊超球定义为模糊类隶属函数的核^[2], 不同类属的超球相互间显然不应重叠。然而, 以图1所示的二维超球为例, 一旦超球在扩张中与异类超球发生重叠, 采用收缩策略并不能保证超球中不含有异类训练样本, 同时还可能有部分同类训练样本被排除在收缩后的超球之外, 这些都会导致识别率的下降。算法为此引入形状因子 θ 来将模糊超球的扩张范围限制在 $\|X_h - W_j\| < \theta$ 的范围内, 实现有指导学习, 当识别率不满足要求时降低 θ 值; 识别率高于要求时提高 θ 值以减少超球个数^[6]。但是, 由于“模糊超球神经网络对输入样本非常敏感”^[3], θ 的优化值与训练样本集的大小和分布密切相关。这使得必须多次学习才能得到较好的网络结构, 有时甚至“很难为不同分类找到一个统一的 θ 值”^[4], 使得学习更为复杂。本文提出一种新的无指导条件重叠学习算法, 算法不受模糊超球形状因子的影响, 学习速度快, 学习后得到的模糊超球个数更少, 识别的正确率更高。

2 模糊超球的条件重叠

模糊超球聚类学习算法将模糊超球定义为模糊类隶属函数的核, 模糊类隶属函数和模糊超球隶属函数是等同的, 即模糊超球 HS_j 内的任何模式必须完全属于某个模式类 C_k , 前述不足即由此而生。为此, 首先将模糊类隶属函数和模糊超球隶属函数区分开来, 并重新定义模糊类隶属函数。仍用式(1)度量样本 $X_h = (x_{j1}, x_{j2}, \dots, x_{ji}, \dots, x_{jn})^T$

* 收稿日期: 2003-03-10

基金项目: 国家自然科学基金与 RGC 联合科研基金资助项目 (79910161989)

作者简介: 华强 (1944年生), 男, 教授; E-mail: hua_qiang@21cn.com

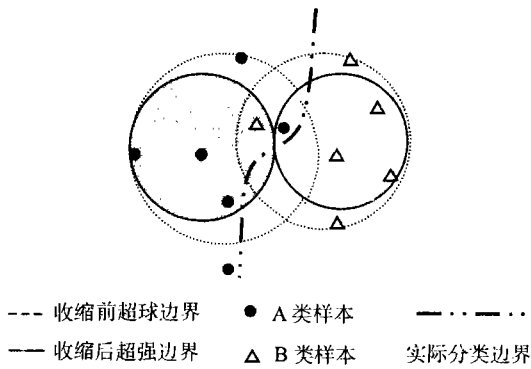


图1 二维模糊超球扩张产生重叠后的收缩

Fig.1 The contraction of 2-d fuzzy hyperspheres with overlap while expanding

与由球心 $\mathbf{W}_j = (w_{j1}, w_{j2}, \dots, w_{jn}, \dots, w_{jm})^T$ 和半径 v_j 确定的模糊超球 HS_j 间的相似关系即模糊超球隶属函数。重新定义模糊超球神经网络的模糊类隶属函数为:

$$\begin{aligned} bhs_j(\mathbf{X}_h, \mathbf{W}_j, v_j) = & \exp\{-(\|\mathbf{X}_h - \mathbf{W}_j\| - v_j)^\gamma\} = \\ & \exp\{-(\|\mathbf{X}_h - \mathbf{W}_j\| \alpha_j)\} \times \\ & \exp\{-(\|\mathbf{X}_h - \mathbf{W}_j\| - v_j)^\gamma\} \end{aligned} \quad (2)$$

式中, 子项 $\exp\{-(\|\mathbf{X}_h - \mathbf{W}_j\| - v_j)^\gamma\}$ 反映 $\mathbf{X}_h$ 离开超球 HS_j 后在超球外部模糊类隶属度的降低, 参数 γ 仍反映 $\mathbf{X}_h$ 远离超球 HS_j 时在超球外部隶属度的降低速率; 而子项 $\exp\{-(\|\mathbf{X}_h - \mathbf{W}_j\| \alpha_j)\}$ 则反映 $\mathbf{X}_h$ 离开超球 HS_j 球心 $\mathbf{W}_j$, 但仍在超球内部时其模糊类隶属度的降低, 参数 α_j 反映 $\mathbf{X}_h$ 在超球 HS_j 内隶属度的降低速率。

对所有超球而言, 当 $\|\mathbf{X}_h - \mathbf{W}_j\| - v_j = 0$ 时, 要求 $\exp\{-(\|\mathbf{X}_h - \mathbf{W}_j\| \alpha_j)\} = \lambda, 0 < \lambda < 1$, 所以

$$\alpha_j = \frac{\ln(-\ln \lambda)}{\ln(\|\mathbf{X}_h - \mathbf{W}_j\|)} = \frac{\ln(-\ln \lambda)}{\ln(v_j)} \quad (3)$$

如此一来, 模糊超球只是模糊类隶属函数的 λ 截集, 异类超球就可有条件的重叠。只要所有超球均取相同的 $\lambda (0 < \lambda < 1)$ 值, 任一模式 $\mathbf{X}_h$ 若落在超球 HS_j 的非重叠区内时, 必有 $\lambda < bhs_j(\mathbf{X}_h, \mathbf{W}_j, v_j) \leq 1$ 和 $bhs_k(\mathbf{X}_h, \mathbf{W}_k, v_k) < \lambda, \forall k \neq j$ 。因而必有

$$bhs_k(\mathbf{X}_h, \mathbf{W}_k, v_k) < bhs_j(\mathbf{X}_h, \mathbf{W}_j, v_j), \forall k \neq j \quad (4)$$

$\mathbf{X}_h$ 若落在超球 HS_j 与其他异类超球(设为异类超球 $HS_k, k \neq j$)的重叠区内时, 则可以通过比较对不同超球主要归属类的类隶属函数的大小链决定分类。因此算法中异类超球可重叠的充分必要条件是重叠区内不存在与此相矛盾的训练样本。为简化

计算, 学习算法中异类超球可以扩张(不论有无重叠发生)的充分条件是扩张后超球内不存在异类的训练样本(见图2)。

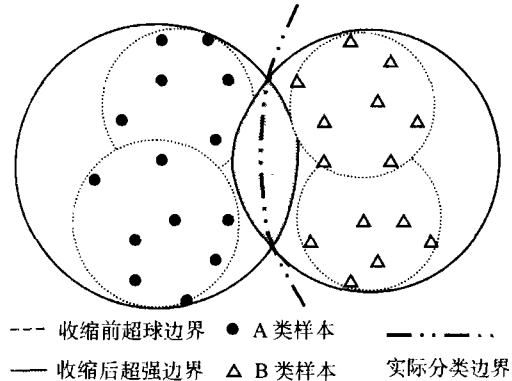


图2 二维模糊超球合并时的条件重叠

Fig.2 The conditional overlap of 2-d fuzzy hyper-spheres while merging

为此, 可定义样本模式到超球球心的 n 维距离与超球半径之比为比半径

$$D_j = \frac{\|\mathbf{X}_h - \mathbf{W}_j\|}{v_j} \quad (5)$$

模糊类隶属度 $bhs_j(\mathbf{X}_h, \mathbf{W}_j, v_j)$ 随比半径 D_j 上升而下降, 在超球内部从球心处的 1 逐渐下降到球面处的 λ 。如让模糊类隶属度在超球外部继续按各自球内的下降速率从球面处的 λ 再逐渐下降到 0, 式(2)可以进一步简化。令 $bhs_j(\mathbf{X}_h, \mathbf{W}_j, v_j)$ 正比于 $\exp\{-(\|\mathbf{X}_h - \mathbf{W}_j\| \alpha_j)\}$, 将式(2)再简化为

$$\begin{aligned} bhs_j(\mathbf{X}_h, \mathbf{W}_j, v_j) = & \text{Max}\left\{0, 1 - (1 - \lambda) \frac{\|\mathbf{X}_h - \mathbf{W}_j\|}{v_j}\right\} = \\ & \text{Max}\{0, 1 - (1 - \lambda) * D_j\} \end{aligned} \quad (6)$$

式(6)中 λ 是一个共同的、但无需具体设定的值。如此一来, $\mathbf{X}_h$ 的类属只需比较对不同模糊超球球心距离的比半径就可判别。例如, 当 $D_j > D_k$ 时有

$$bhs_j(\mathbf{X}_h, \mathbf{W}_j, v_j) < bhs_k(\mathbf{X}_h, \mathbf{W}_k, v_k)$$

反之亦然。

3 无指导条件重叠学习算法

条件重叠学习算法先将训练样本集中的每个样本 (n 维论域空间中的一个点) 作为一个模糊超球, 为其设置主动状态标志, 显然其中不存在异类样本; 在保持所有模糊超球中不含异类训练样本的前提下, 将具有主动状态标志的模糊超球尽量与其它模糊超球合并, 被合并的模糊超球改为被动状态标志, 这样将同类超球尽可能合并成更大的模糊超

球。算法不限制超球合并的范围, 即不对学习进行指导。而由于不产生不满足前述充分条件的重叠, 因而也就不需要进行收缩。若模式类数 $P = 1, \dots, C$; 每类训练样本数 $Q = 1, \dots, M$; 超球 bhs_j 的主动状态标志为 $F(bhs_j) = 1$, 条件重叠学习算法为:

```
for(P = 1; P <= C; P++)
{
  for(Q = 1; Q <= M; Q++) if(F(bhs_Q) = 1)
  {
    先生成一个临时模糊超球 bhs_T = bhs_Q;
    for(S = Q + 1; S <= M; S++) if(F(bhs_S) = 1)
    {
      扩张 bhs_T, 将 bhs_S 试并入 bhs_T; 检查是否
      有任一异类训练样本被包含在超球
      bhs_T 中,
      if(任一异类训练样本被包含在超球 bhs_T
      中) {取消该次合并; continue;}
      else {正式扩张 bhs_Q = bhs_T, 合并 bhs_S;
    }
  }
}
```

$$F(bhs_S) = 0; \{ \} \{ \} \{ \}$$

4 学习和分类的实验效果

在归一化二维空间 $X - Y$ 中以式(7) - (9)所示的圆、椭圆和双曲线作为类 A 、 B 的边界。

$$(X - 0.5)^2 + (Y - 0.5)^2 = 0.01 \quad (7)$$

$$9.0 \times (X - 0.5)^2 + 4.0 \times (Y - 0.5)^2 = 0.36 \quad (8)$$

$$(X - 0.5)^2 - (Y - 0.5)^2 = 0.01 \quad (9)$$

以晶种 $r = 5$ 的伪随机函数 $r = (r * 123 + 59) \% 65536 / 65536$ 产生每类样本数同为 $S = 20, 50, 100, 300$ 的 $2 * S$ 大小的训练样本集和检验样本集。分别用不同 θ 值的模糊超球聚类学习算法和条件重叠学习算法进行学习和分类的实验。得出最后的 A 、 B 类隐层节点数 $Ah + Bh$ 和对训练集/检验集的识别正确率再进行对比, 以不失一般性。学习和分类的结果见表 1。

表 1 两种学习算法的训练集/检验集识别正确率和隐层节点数 ($Ah + Bh$) 的比较

Tab.1 The comparison of the classification accuracy to training/testing sets and the node number ($Ah + Bh$) in implicit layer between two learning algorithms

学习算法		模糊超球的原聚类学习算法, θ					条件重叠学习算法
		0.05	0.15	0.25	0.35	0.45	无指导
圆 边 界	20 + 20	100/85.0	100/85.0	100/85.0	100/85.0	100/82.5	100/85.0
		8 + 18	5 + 13	4 + 9	4 + 7	4 + 6	1 + 4
	50 + 50	100/93.0	100/95.0	100/90.0	89.0/85.0	89.0/82.0	100/93.0
		13 + 39	8 + 20	9 + 13	9 + 9	5 + 9	1 + 6
	100 + 100	100/96.0	100/96.5	100/94.5	97.0/94.5	92.5/91.5	100/97.0
		19 + 75	14 + 30	14 + 21	10 + 16	5 + 18	1 + 6
	300 + 300	99.8/98.5	98.0/97.3	87.7/87.2	96.8/95.0	79.5/81.2	100/98.7
		39 + 144	25 + 64	23 + 42	10 + 33	6 + 26	3 + 6
	20 + 20	100/100	100/100	100/95.0	92.5/97.5	82.5/67.5	100/80.0
		15 + 20	7 + 13	6 + 9	5 + 8	3 + 5	2 + 6
椭 圆 边 界	50 + 50	100/98.0	100/100	97.0/92.0	85.0/84.0	84.0/87.0	100/92.0
		27 + 45	12 + 21	8 + 16	6 + 11	4 + 8	2 + 8
	100 + 100	100/97.0	98.0/98.5	99.5/97.5	88.0/87.0	89.5/89.0	100/95.5
		38 + 74	16 + 28	9 + 20	6 + 14	4 + 13	4 + 10
	300 + 300	99.7/98.0	96.2/96.7	94.5/95.2	89.7/89.5	84.0/83.0	100/95.8
		75 + 142	29 + 54	11 + 33	6 + 20	5 + 15	10 + 13
	20 + 20	100/85.0	97.5/82.5	95.0/87.5	95.0/85.0	92.5/85.0	100/92.5
		18 + 18	12 + 10	9 + 6	7 + 6	5 + 5	4 + 3
	50 + 50	100/95.0	98.0/93.0	95.0/95.0	95.0/90.0	93.0/89.0	100/89.0
		42 + 40	15 + 17	13 + 11	8 + 9	7 + 7	6 + 7
双 曲 边 界	100 + 100	100/93.0	99.0/93.0	96.0/92.0	92.0/88.5	93.5/88.5	100/93.5
		62 + 57	26 + 25	17 + 17	11 + 13	9 + 10	9 + 10
	300 + 300	99.8/97.2	96.7/94.7	93.5/92.3	93.8/92.0	92.7/91.2	100/95.5
		120 + 108	45 + 41	28 + 26	15 + 14	12 + 14	17 + 17
	平均识别正确率/%	99.95/94.6	98.6/94.3	96.5/92.0	92.8/89.5	89.5/84.8	100/92.3
	100 + 100 样本组	40 + 69	19 + 28	13 + 19	9 + 14	6 + 14	5 + 9
	平均超球个数						

5 讨论

由表1, 样本数较大时, 聚类学习算法产生超球重叠, 训练集识别率低于100%, 检验集平均识别率也较低; 而条件重叠学习算法训练集识别率均为100%。检验集平均识别率也高于聚类学习算法。

超球聚类学习算法的扩张范围受形状因子 θ 限制, 留下的隐层节点数较多。如要达到与条件重叠学习算法相当的识别率, 所需隐层节点数要比条件重叠学习算法多, 有时甚至要多到几倍; 而当取隐层节点数相近时, 条件重叠学习算法的识别正确率高于超球聚类学习算法。

超球聚类学习算法受形状因子 θ 影响大, 且不同样本集大小和分布下形状因子 θ 的优化值不同。 θ 小时识别率较高但隐层节点数大; θ 较大时节点数虽较少但识别率则显著下降。为此聚类学习算法必须根据实际样本集大小和分布进行有指导的多次学习, 以逐步选出 θ 的较优值, 甚至需对不

同分类使用不同的 θ , 学习过程慢; 而条件重叠学习算法学习快, 一次学习即可得出所要求的结果。

参考文献:

- [1] 梁夷龙, 王子罡. 基于超椭圆模糊聚类的人脑磁共振图像分割[J]. 软件学报, 1998, 9(9): 683-689.
- [2] 黄敏超, 张育林. 模糊超球神经网络在模式聚类中的应用[J]. 自动化学报, 1997, 23(2): 279-282.
- [3] 黄敏超, 吴建军. 模糊超体神经网络及其在火箭发动机故障分离中的应用[J]. 航空动力学报, 1997, 12(1): 79-82.
- [4] 孙真真, 陈曾平, 庄钊文. 模糊极大极小神经网络雷达目标识别方法研究[J]. 现代雷达, 1998, (6): 29-35.
- [5] JIRINA M, JIRINA M Jr. Neural network classifier based on growing hyperspheres[J]. Neural Network World, 2000, 10(3): 417-428.
- [6] ANAGNOSTOPOULOS G C, GEORGIPOPOULOS M. Hypersphere ART and ARTMAP for unsupervised and supervised, incremental learning[J]. Proceedings of International Joint Conference on Neural Networks, 2000: 59-64.

A Conditional Overlap Learning Algorithm of the Fuzzy Hyperspheres Neural Networks

HUA Qiang, LI Shi-xian

(Department of Computer Sciences, Sun Yat-sen University, Guangzhou 510275, China)

Abstract: The class-aggregating algorithm of the fuzzy hypersphere neural networks use a form factor θ to realize a supervised learning. Appointing to the problems produced by θ , a new unsupervised learning algorithm is developed. The new algorithm has nothing to do with the form factor, achieve better classification accuracy with less training time and a fewer number of hyperspheres.

Key words: pattern recognition; fuzzy hypersphere neural networks; learning algorithm