



模糊超球神经网络在模式聚类中的应用¹⁾

黄敏超 张育林 陈启智

(国防科技大学航天技术系 长沙 410073)

摘要 提出和实现了用于模式聚类的无监督模糊超球神经网络. 模式集是一个具有超球核的用隶属函数表示的模糊集, 模式集又可以合并成模式类. 模糊超球神经网络学习算法能在几次循环学习中形成模式集, 无需对已知模式集重新训练就可融合新样例和精炼已存在的模式集. 模式聚类的数值仿真解释了模糊超球聚类神经网络的优越性能.

关键词 模式聚类, 模糊集, 神经网络, 分层聚类, 液体推进剂火箭发动机

1 引言

模糊集理论和神经网络的引入, 使得模式聚类的研究更具活力, 发展至今比较出名的聚类分析方法有模糊 ART^[1]、模糊 C 均值聚类^[2]和自适应模糊首项聚类^[3]等. 但这些聚类分析方法都有不足之处, 其中, 模糊 ART 没有适当地把神经网络和模糊集结合起来, 它只用模糊集来进行极大极小操作, 这在实际应用中没有多大意义, 另外, 模糊 ART 容许一个样本以全隶属度属于多个模式集, 减弱了聚类结果的可靠性; 模糊 C 均值算法用于聚类分析之前, 必须已知模式集的数目, 限制了模糊 C 均值聚类算法的应用范围; 自适应模糊首项聚类只用一个点代表模式集的中心, 并在学习过程中, 不断调整中心的位置, 这样会引起聚类学习算法不稳定问题. 本文提出的模糊超球神经网络聚类算法明确把神经网络与模糊集有机结合起来, 每个模式集节点定义一个 n 维超球空间范围(称为全隶属区), 模式集节点数由实际问题自适应形成, 模糊超球神经网络在学习过程中不断调整超球的中心和半径, 以优化模式集的边界. 为了提高聚类精度, 进一步考虑了对模式集的聚类问题, 即模糊超球神经网络增加一层高标准聚类网络, 以实现分层聚类算法.

2 模糊超球聚类神经网络

2.1 超球隶属函数

超球隶属函数 $b_j(A_h, V_j, W_j)$ 用来度量输入样本 $A_h = (a_{h1}, a_{h2}, \dots, a_{hn})^T$ 与由半径 V_j 和球心 $w_j = (w_{j1}, w_{j2}, \dots, w_{jn})^T$ 构成的超球之间的相似关系, 当 A_h 远离超球时, 超球隶属

1) 本文系国家自然科学基金项目.

收稿日期: 1994-12-26

函数接近0;当 A_h 靠近超球时,超球隶属函数接近1;超球包含这点时, $b_j(A_h, V_j, W_j) \equiv 1$, 由此定义超球隶属函数如下:

$$b_j(A_h, V_j, W_j) = \begin{cases} e^{-(\|A_h - W_j\| - V_j)^\gamma}, & 0 \leq \|A_h - W_j\| - V_j, \\ 1, & (\|A_h - W_j\| - V_j) < 0, \end{cases} \quad (1)$$

$$\|A_h - W_j\| = \sqrt{(a_{h1} - w_{j1})^2 + (a_{h2} - w_{j2})^2 + \dots + (a_{hm} - w_{jm})^2}$$

式中灵敏参数 γ 反映输入样本在远离超球核时,它们之间的隶属度降低速率.

2.2 神经网络聚类器

神经网络拓扑结构由实际问题自适应形成,包括 F_B 层和 F_C 层(见图1). F_B 层节点代

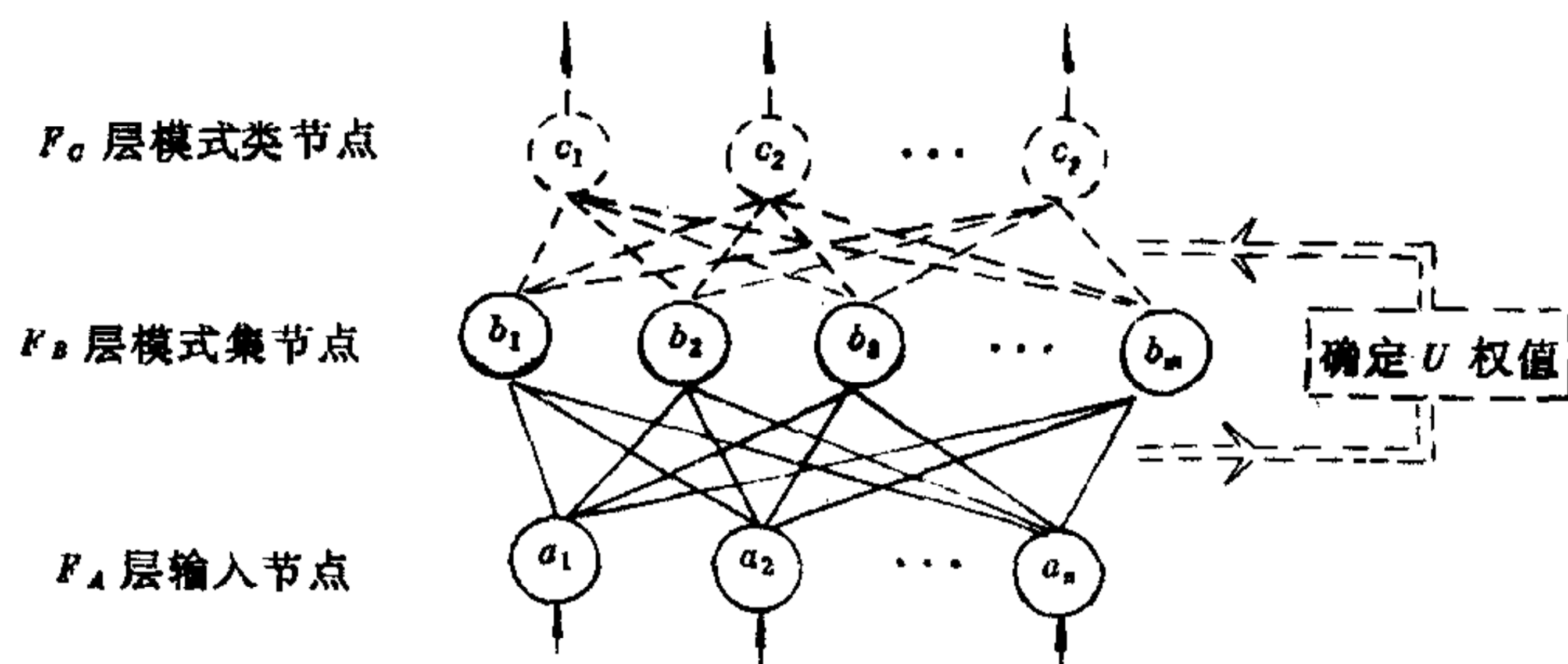


图1 三层模糊超球聚类神经网络系统

表超球模糊集(即模式集), F_A 层与 F_B 层联结权值是超球的球心和半径, F_B 层传递函数是由(1)式定义的超球隶属函数,半径贮存在 $m \times 1$ 维矩阵 V 中,球心贮存在 $m \times n$ 维矩阵 W 中. F_B 层与 F_C 层联结权值是二值矢量 $\{0, 1\}^p$, 并存在 $p \times m$ 维矩阵 U 中. 根据第 3.3 节提出的高标准聚类学习算法,通过对由半径 V_j 和球心 W_j 决定的超球进一步聚类来确定矩阵 U . F_C 层节点的传递函数是同一类超球模糊集的并集, 定义为

$$C_k = \max_{j=1}^m b_j u_{kj}, \quad (2)$$

式中 b_j 是 F_B 层的第 j 个节点, c_k 是 F_C 层的第 k 个节点. 每个 F_C 层节点代表一种模式类, F_C 层节点输出值表示输入样本 A_h 与模式类 k 之间的隶属程度. u_{kj} 是 b_j 与 c_k 之间的联结权值, 若超球模糊集 b_j 属于模式类 c_k , 则 $u_{kj} = 1.0$, 否则 $u_{kj} = 0.0$.

3 聚类神经网络学习算法

3.1 超球扩张

已知某一样本 A_h , 找到最近似的超球 B_j . 若要超球 B_j 扩张到能包含 A_h , 必须首先满足以下条件限制:

$$\|A_h - W_j\| < \theta, \quad (3)$$

式中 θ 是选定常数, 用来限制超球大小. 其次判断 A_h 是否是超球 B_j 外的一点, 即表示为

$$\|A_h - W_j\| > V_j, \quad (4)$$

若式(1), (2)的条件满足, 则超球 B_j 的球心和半径调整公式为

$$V_j^{new} = \frac{1}{2} (\|A_h - W_j^{old}\| + V_j^{old}), \quad (5)$$

$$W_j^{new} = W_j^{old} + \frac{V_j^{new} - V_j^{old}}{\|A_h - W_j^{old}\|} (A_h - W_j^{old}). \quad (6)$$

3.2 超球重叠测试与收缩

为了测试是否产生重叠,须在两模糊超球之间进行比较.如果超球 B_j 与超球 B_k 发生重叠,必须消除超球之间的重叠.3种测试条件及相应的超球收缩算法为

$$1) \quad |V_j - V_k| < \|W_j - W_k\| < V_j + V_k,$$

$$V_j^{new} = \frac{1}{2} (V_j^{old} - V_k^{old} + \|W_j^{old} - W_k^{old}\|),$$

$$V_k^{new} = \frac{1}{2} (V_k^{old} - V_j^{old} + \|W_j^{old} - W_k^{old}\|);$$

$$2) \quad \|W_j - W_k\| < V_j - V_k,$$

$$V_j^{new} = \frac{1}{2} (V_j^{old} - V_k^{old} + \|W_j^{old} - W_k^{old}\|),$$

$$W_j^{new} = W_k^{old} + \frac{V_j^{old} + V_k^{old} + \|W_j^{old} - W_k^{old}\|}{2\|W_j^{old} - W_k^{old}\|} (W_j^{old} - W_k^{old});$$

$$3) \quad \|W_j - W_k\| < V_k - V_j,$$

$$V_k^{new} = \frac{1}{2} (V_k^{old} - V_j^{old} + \|W_j^{old} - W_k^{old}\|),$$

$$W_k^{new} = W_j^{old} + \frac{V_k^{old} + V_j^{old} + \|W_j^{old} - W_k^{old}\|}{2\|W_j^{old} - W_k^{old}\|} (W_k^{old} - W_j^{old}).$$

3.3 U 权值的确定

模糊超球聚类学习算法用超球来划分模式空间的非规则样本群的分布,若阈值 θ 取得比较大,会引起超球内(隶属度为 1.0)一些空间不是训练样本的实际分布,降低了聚类精度.可行的方法是把阈值 θ 选得比较小,使聚集形成的超球充分反映训练样本的实际空间分布,然后采用高标准聚类学习算法把独立的邻近超球并联起来,形成模式类的非规则边界. F_B 层与 F_C 层联结权值 U 用高标准聚类学习算法确定,具体算法实现为

1)设置所有权值 $U=0$,令 F_C 层节点数 $p=0$,标识变量 $d_i=1, i=1, 2, \dots, m, m$ 是模式集总数.

2)若能找到第一个 $d_i=1 (i=1, 2, \dots, m)$ 的模式集, $p \leftarrow p+1$,并令 $u_{pi}=1, j=i$. 否则结束.

3)计算 $ERR_{ij} = \sqrt{\sum_{k=1}^n (w_{ik} - w_{jk})^2 + V_i + V_j}$,如果 ERR_{ij} 小于阈值 σ ,令标识变量 $d_i=0$ 和联结权值 $u_{pj}=1$,转2).

4)若 $j < m, j \leftarrow j+1$,转3). 否则转2).

4 模式聚类的数值仿真

4.1 二维模式空间聚类

由 30 个点构成的 2 维数据群用来说明不同阈值 θ 的模糊超球聚类性能. 这些数据选取考虑了以下两个主要问题:1)有一群数据可以聚集成一个模式集或两个模式集;2)有

一些数据介于两个主要数据群之间. 阈值在 0.2到0.05之间取值, 数据群被聚类4次. 当阈值等于 0.08时, 聚类结果比较恰当. 这个例子表明阈值 θ 是非常重要的参数, 它限制模式集大小和决定模式决策精度.

4.2 变推力火箭发动机故障识别

变推力火箭发动机通过调节活门开度控制流量大小来实现推力变化, 在此考虑了两种变推力发动机故障模式. 针对不同的扩张标准 θ 用热试车 RS73 数据训练模糊聚类神经网络, 得到不同 F_B 层节点输出. 阈值 $\theta=0.3$ 时, 模糊神经网络形成 3 个超球, b_1 和 b_3 用于识别发动机两种故障模式, b_2 反映发动机工况的过渡特性. 当阈值 $\theta=0.05$ 时, 对于同一种故障模式, F_B 层将同时有几个节点输出值接近1, 必须在模糊神经网络 F_B 层之上加一层 F_C , 把表示同一种故障的超球合并成模式类. 当阈值 σ 取为 0.3, 0.2, 0.1 时, 三层模糊神经网络可以正确识别发动机故障模式, 当阈值 σ 小于 0.05时, 又出现几个 F_C 层节点输出接近1. 通过以上分析, 针对本文考虑的变推力发动机, 推荐阈值 θ 取值介于 0.02—0.3, 阈值 σ 介于 0.1—0.3.

参 考 文 献

- [1] Carpenter G, Grossberg S, Rosen D. Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks*, 1991, 4: 759—771.
- [2] Davis J, Warms T, Winters W, A neural network implementation of the fuzzy c-means clustering algorithm. In: Proc. Int. Joint Conf. Neural Networks (Seattle), 1991, 11: A-953.
- [3] Scott C Nweton, Surya Pemmaraju, Sunanda Mitra. Adaptive fuzzy leader clustering of complex data sets in pattern recognition. *IEEE Trans. Neural Networks*, 1992, 3(5).

FUZZY HYPERSPHERE NEURAL NETWORK AND ITS APPLICATION ON PATTERN CLUSTERING

HUANG MINCHAO ZHANG YULIN CHEN QIZHI

(Department of Aerospace Technology, National University of Defense Technology, Changsha, 410073)

Abstract An unsupervised hypersphere neural network that is used for pattern clustering is presented. Each clustered pattern set is a fuzzy set hypersphere with a corresponding membership function. A pattern class is the combination of the pattern sets. The center and radius of the hypersphere are determined using the fuzzy hypersphere learning algorithm, an expansion-contraction process that can learn nonlinear pattern set boundaries in few passed through the data and provides the ability to incorporate new and refine existing pattern sets without re-training. The simulation of clustering demonstrates the superiority of the fuzzy hypersphere clustering neural network.

Key words Pattern clustering, fuzzy set, neural network, hierarchy clustering, liquid propellant rocket engine.