

自组织区间二型模糊神经网络及其自适应学习算法

姚兰^{1,2†}, 肖建¹, 王嵩¹, 蒋玉莲³

(1. 西南交通大学 电气工程学院, 四川 成都 610031; 2. 成都信息工程学院 控制工程学院, 四川 成都 610225;

3. 西南交通大学 交通运输与物流学院, 四川 成都 610031)

摘要: 针对复杂不确定非线性系统的辨识问题, 提出一种基于聚类的自组织区间二型模糊神经网络学习算法. 首先采用具有两个不同加权参数的FCM算法对输入数据进行划分来获取规则前件的不确定均值, 同时结合聚类有效性标准确定模糊规则数目, 从而自动完成神经网络的结构辨识和规则前件参数辨识; 随后给出了基于梯度下降法和Lyapunov函数稳定收敛定理的规则后件权向量学习速率的自适应学习算法. 通过非线性系统辨识实例, 验证了该算法与其他方法相比具有更快的收敛速度和更高的逼近精度; 并且利用该算法建立了某市电力短期负荷预测模型, 结果表明该模型具有较高的预测精度, 泛化性能更佳.

关键词: 自组织; 区间二型模糊神经网络; 梯度下降法; 自适应学习算法

中图分类号: TP273 **文献标识码:** A

Interval type-2 fuzzy neural networks with self-organizing structure and adaptive learning algorithm

YAO Lan^{1,2†}, XIAO Jian¹, WANG Song¹, JIANG Yu-lian³

(1. College of Electrical Engineering, Southwest Jiaotong University, Chengdu Sichuan 610031, China;

2. College of Control Engineering, Chengdu University of Information Technology, Chengdu Sichuan 610225, China;

3. College of Transportation and Logistics, Southwest Jiaotong University, Chengdu Sichuan 610031, China)

Abstract: For the identification problems of nonlinear system with complex uncertainties, an interval type-2 fuzzy neural network with self organizing structure and learning algorithm is proposed. Firstly, the fuzzy c -means algorithm with two different weighting parameters are used to partition the input data to obtain the uncertainty means of rule antecedent; meanwhile, according to the cluster validity criterion, the number of fuzzy rules is determined. Thus, the structure and parameters of rule antecedent identification are automatically completed. Then, based on the gradient descent method and Lyapunov function stability convergence theorem, the adaptive learning algorithm for weight vectors of rules consequent is presented. Finally, the experimental results of nonlinear system identification indicate that the proposed algorithm has faster convergence rate and higher approximation precision than other algorithms. In addition, based on the power load data of a city, a short-term load-forecasting model is developed by the algorithm, which has higher prediction precision and better generalization.

Key words: self-organizing; interval type-2 fuzzy neural networks; gradient-descent algorithm; adaptive learning algorithm

1 引言(Introduction)

近年来, 模糊神经网络在非线性系统辨识和自适应控制问题上已取得了一些显著成果^[1-2]. 但在实际工程与科学实验中, 构建模糊神经网络的规则及隶属函数所需要的专家经验及大量可知信息含有各种形式的复杂不确定性、模糊性和不精确性^[3-5], 以往基于一型模糊集的模糊神经网络很难能得到满意的解决方案. Zadeh等提出的二型模糊集具有更强描述与处理不确定性的能力^[6-8], 于是, 将二型模糊集与神经网络结合研究基于复杂不确定非线性系统问题是值得关注的研究领域. 文献[9]提出了自组织递

归区间二型模糊神经网络, 利用网络结构算法建立网络结构, 前后件参数学习分别采用梯度下降算法和有序规则卡尔曼滤波算法. 不足之处在于没有分析网络结构的收敛性能. 文献[10]提出3种不同结构的T-S型区间二型模糊神经网络, 主要区别在于输入层模糊化方式和推理结果的降型方法, 然后都采用梯度下降算法对参数进一步优化. 但模糊规则是预先根据经验确定的, 没有解决规则自动产生的问题. 文献[11]采用不对称隶属函数提出动态递归区间二型模糊神经网络, 基于Lyapunov稳定性理论得出了参数学习算法稳定收敛的条件, 提高了网络的稳定

性和对动态系统的辨识能力. 文献[12]根据输入输出数据相似性比较自动调整网络结构和参数, 然后采用粒子群和最小二乘法相结合的混合方法进行参数优化. 文献[13]基于聚类提出自组织二型模糊神经网络, 能够根据输入输出数据实现规则的自动生成. 但没有涉及适当规则数目的判断准则以及网络结构的收敛性.

结合已有文献的优势与不足, 本文设计了一种基于Mamdani自组织区间二型模糊神经网络(self-organizing interval type-2 fuzzy neural network, SOIT2FNN). 借鉴文献[14]中模糊C均值(fuzzy c-means, FCM)聚类算法的加权参数对聚类模糊程度有重要调节作用的结论, 采用具有两个不同加权参数的FCM算法进行划分从而获得了隶属函数显著变化的区域, 并与区间二型模糊集的不确定迹的本质相结合, 提出了一种新的区间二型模糊规则前件不确定均值的确定方法, 同时利用聚类有效性标准来判断最佳聚类数目, 即模糊规则数目, 自动实现了网络的结构辨识和规则前件参数辨识; 随后给出了基于梯度下降法和Lyapunov函数稳定收敛定理的规则后件权向量学习速率的自适应学习算法, 动态地对非线性不确定性模型进行建模.

2 区间二型模糊神经网络的结构(Structure of interval type-2 fuzzy neural networks)

由于一个多输入多输出模糊系统可由多个多输入单输出系统合成, 不失一般性, 本文考虑一个多输入单输出模糊系统, 第j条模糊规则为

$$\tilde{R}^j : \text{IF } x_1 \text{ is } \tilde{F}_1^j \text{ and } \dots \text{ and } x_n \text{ is } \tilde{F}_n^j, \\ \text{THEN } y \text{ is } [w_1^j, w_r^j], j = 1, 2, \dots, M,$$

式中: $x = \{x_i\}_{i=1, \dots, n}$ 是输入变量, y 是输出变量, \tilde{F}_i^j 为前件的区间二型模糊集, $[w_1^j, w_r^j]$ 是后件区间二型模糊集的质心集, 也称为加权区间集.

该网络结构如图1所示, 由5层组成.

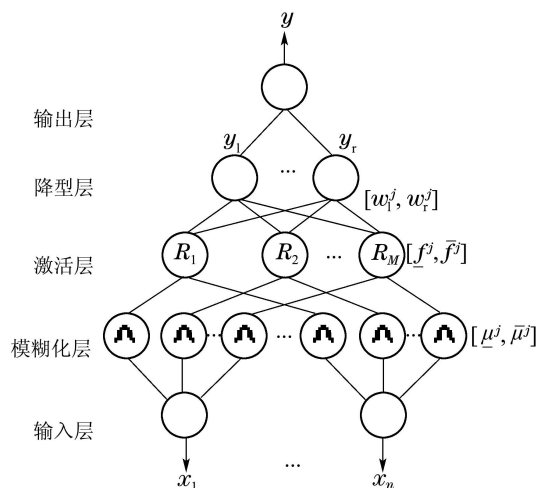


图1 自组织区间二型模糊神经网络结构
Fig. 1 Structure of SOIT2FNN

第1层为输入层, 该层的结点直接与输入变量连接, 将输入值传送到下一层.

第2层为模糊化层, 该层节点被分为M组, 每一组表示一条模糊规则的前件部分, 每个节点代表一个对称的区间二型隶属函数. 根据输入变量与均值的关系, 该层节点的输出可利用区间二型模糊集的不确定迹(foot of uncertain, FOU)的上界和下界隶属函数直接计算得到, 表示如下:

$$y_{ij}^{(2)} = \begin{cases} \bar{\mu}_{ij}(x_i) = \begin{cases} N(\underline{m}_{ij}, \sigma; x_i), & x_i \leq \underline{m}_{ij}, \\ 1, & \underline{m}_{ij} < x_i \leq \bar{m}_{ij}, \\ N(\bar{m}_{ij}, \sigma; x), & x_i > \bar{m}_{ij}, \end{cases} \\ \underline{\mu}_{ij}(x_i) = \begin{cases} N(\bar{m}_{ij}, \sigma; x), & x_i \leq \frac{\underline{m}_{ij} + \bar{m}_{ij}}{2}, \\ N(\underline{m}_{ij}, \sigma; x), & x_i > \frac{\underline{m}_{ij} + \bar{m}_{ij}}{2}, \end{cases} \end{cases} \quad (1)$$

式中: $\bar{\mu}_{ij}$ 和 $\underline{\mu}_{ij}$ 是上界和下界隶属函数, $[\underline{m}_{ij}, \bar{m}_{ij}]$ 和 σ 分别是隶属函数的不确定均值和标准差.

第3层为激活层, 每个节点对应一条模糊规则, 通过对第2层节点的输出执行product t-norm规则下的meet运算得到相应的激活强度, 表示如下:

$$y_j^{(3)} = \prod_i y_{ij}^{(2)} = \begin{cases} \bar{f}^j = \prod_i \bar{\mu}_{ij}(x_i), \\ \underline{f}^j = \prod_i \underline{\mu}_{ij}(x_i). \end{cases} \quad (2)$$

第4层为降型层, 该层节点为后件节点, 通过对前件激活强度和后件权值执行Center-of-sets降型处理得到输出变量, 表示如下:

$$y_j^{(4)} = \begin{cases} y_r^{(4)} = \frac{\sum_{j=1}^M w_r^j y_j^{(3)}}{\sum_{j=1}^M y_j^{(3)}}, \\ y_1^{(4)} = \frac{\sum_{j=1}^M w_1^j y_j^{(3)}}{\sum_{j=1}^M y_j^{(3)}}. \end{cases} \quad (3)$$

通常采用Karnik-Mendel(KM)迭代算法^[15]来计算 $y_r^{(4)}$ 和 $y_1^{(4)}$. 以 $y_r^{(4)}$ 的计算为例, KM方法的关键是在初始规则后件参数按照升序排列的情况下确定适当R, 当 $j \leq R$, $y_j^{(3)} = \underline{f}^j$; 当 $j > R$, $y_j^{(3)} = \bar{f}^j$, 则 $y_r^{(4)}$ 表示如下:

$$y_r^{(4)} = \frac{\sum_{j=1}^R \underline{f}^j w_r^j + \sum_{j=R+1}^M \bar{f}^j w_r^j}{\sum_{j=1}^R \underline{f}^j + \sum_{j=R+1}^M \bar{f}^j} =$$

$$\sum_{j=1}^R \underline{q}_r^j w_r^j + \sum_{j=R+1}^M \bar{q}_r^j w_r^j, \quad (4)$$

式中:

$$\underline{q}_r^j = \underline{f}^j / (\sum_{j=1}^R \underline{f}^j + \sum_{j=R+1}^M \bar{f}^j),$$

$$\bar{q}_r^j = \bar{f}^j / (\sum_{j=1}^R \underline{f}^j + \sum_{j=R+1}^M \bar{f}^j).$$

等式(4)可写成向量与矩阵形式:

$$y_r^{(4)} = \mathbf{q}_r^T \mathbf{w}_r, \quad (5)$$

式中:

$$\mathbf{q}_r = [\underline{q}_r^1 \cdots \underline{q}_r^R \bar{q}_r^{R+1} \cdots \bar{q}_r^M]^T, \quad (6)$$

$$\mathbf{w}_r = [w_r^1 \cdots w_r^M]^T. \quad (7)$$

同理可得 $y_1^{(4)}$ 的计算表达式的向量描述形式:

$$y_1^{(4)} = \mathbf{q}_1^T \mathbf{w}_1, \quad (8)$$

式中:

$$\mathbf{q}_1 = [\underline{q}_1^1 \cdots \bar{q}_1^L \underline{q}_1^{L+1} \cdots \underline{q}_1^M]^T, \quad (9)$$

$$\mathbf{w}_1 = [w_1^1 \cdots w_1^M]^T. \quad (10)$$

第5层为输出层, 通过计算 $y_r^{(4)}$ 和 $y_1^{(4)}$ 的平均值获得网络的精确输出, 表示如下:

$$y^{(5)} = \frac{1}{2}(y_r^{(4)} + y_1^{(4)}) = \frac{1}{2}(\mathbf{q}_r^T \mathbf{w}_r + \mathbf{q}_1^T \mathbf{w}_1). \quad (11)$$

从以上分析可知, 建立SOIT2FNN需要辨识的参数包括: 模糊规则数目, 前件部分隶属函数的不确定均值和标准差以及后件的加权区间集. 于是, 本文提出一种基于聚类的自组织区间二型模糊神经网络学习算法.

3 学习算法(Learning algorithm)

自组织区间二型模糊神经网络学习算法包括自组织结构学习和参数学习两部分, 其中自组织结构学习通过对输入空间合理划分获得尽可能少的模糊规则, 保证网络逼近性能的同时使得结构更加紧凑; 参数学习利用Lyapunov稳定性理论给出了具有自适应学习速率的梯度下降法调整后件参数, 从而保证了参数学习过程中的有效性和收敛性.

3.1 自组织结构学习(Learning of self-organizing structure)

为了得到结构紧凑的区间二型模糊神经网络, 以往常利用FCM及其改进算法确定模糊规则. 一般在目标函数中加权参数 $m = 2$ 的条件下, 利用预先设定的聚类数目或根据目标函数的收敛速度所得的最佳聚类数目确定规则数目, 然后在聚类中心上叠加偏差如 ± 0.1 来设定前件不确定均值. 这样的方法存在两方面问题: 1) 参数 $m = 2$ 的FCM算法只适用于聚类分组包含样本数相等的情况; 2) 错误地估

计聚类数目 c 将使算法不能揭示输入样本的聚类结构. 结合区间二型模糊集不确定迹的特点, 本文对基于FCM算法的区间二型神经网络结构学习方法做了一些改进.

1) 对于参数 m 对FCM算法性能的影响, 文献[14]的研究成果表明, 在聚类分组包含的样本数不一致的情况下, 参数 m 的取值控制着样本在不同分组间的模糊程度, 直接影响隶属函数的几何形状. 当 $m \rightarrow 1$ 时, FCM算法变成硬C均值划分, 所得隶属函数转变为普通集合的特征函数; 当 $m \rightarrow \infty$ 时, FCM划分最模糊, 样本属于各分组的隶属度值均相等, 已经失去了划分特性. 同时, 文献[16]通过聚类有效性方面的实验研究指出了当 m 的取值范围为[1.4, 2.6]时, 利用FCM算法所得到样本隶属度值变化最为显著, 形成了样本隶属函数变化的区域, 这正好与区间二型模糊集的主隶属函数所形成的不确定迹的本质相符合, 从而可获得不确定迹的上界和下界隶属函数. 于是, 本文提出一种新的确定区间二型模糊集隶属函数的方法, 分别采用 $m_1 = 1.4$ 和 $m_2 = 2.6$ 的FCM算法对输入样本进行划分, 利用所得的聚类中心作为前件不确定均值.

2) 对于最佳聚类数目, Bezdek曾提出了划分系数和划分熵两个有效性标准进行确定. 但这两个有效性标准具有随聚类数目增加而单调递减的趋势. 本文采用文献[17]给出的有效性标准确定最佳聚类数目, 表示如下:

$$V(u, c) = \frac{1}{n} \sum_{k=1}^n \max_i (u_{ik}) - \frac{1}{K} \sum_{i=1}^{c-1} \sum_{j=i+1}^c \left[\frac{1}{n} \sum_{k=1}^n \min(u_{ik}, u_{jk}) \right], \quad (12)$$

式中 $K = \sum_{i=1}^{c-1} i$. 该有效性标准由分组的组内紧致性程度和组间分离性程度两者决定. 如果样本划分效果越好, 则样本 x_k 越靠近聚类中心 v_i , 前者 $\max_i (u_{ik})$ 的值越接近1, 即反映了组内的紧致性程度; 同时, 后者 $\min(u_{ik}, u_{jk})$ 的值越接近0, 表明分组 u_i 和 u_j 是明显分离. 但如果样本划分效果越模糊时, 样本 x_k 属于各分组的隶属度值均接近 $1/c$, 前后两者的值基本相等. 由此可见, 最终该有效性标准的取值范围为[0, 1], 取值最大时所对应的分类数目即为最佳分类数目.

综上所述, 可获得区间二型模糊神经网络的自组织结构学习方法的具体步骤如下:

Step 1 设置初始聚类数目 $c = 2$ 和最大聚类数目 c_{\max} , 并确定终止条件 ε ;

Step 2 为 c 个聚类中心赋初值;

Step 3 设置 $m_1 = 1.4$, 采用FCM算法对样本进

行聚类, 获取聚类中心值 $\{v_i^{(1)}\}_{i=1,2,\dots,c}$;

Step 4 设置 $m_2 = 2.6$, 采用FCM算法对样本进行聚类, 获取聚类中心值 $\{v_i^{(2)}\}_{i=1,2,\dots,c}$;

Step 5 按式(12)计算聚类有效性指标 V_{m1} 和 V_{m2} , 如果 $c < c_{\max}$, 令 $c = c + 1$, 并转到Step 2; 否则, 聚类停止, 最佳聚类数目按下式选取:

$$c_{\text{opt}} = \arg\{\max_{\forall c}\{\min\{V_{m1}(u, c), V_{m2}(u, c)\}\}\}. \quad (13)$$

根据最佳聚类数目 c_{opt} 与所得聚类中心 $\{v_i^{(1)}\}_{i=1,2,\dots,c_{\text{opt}}}$ 和 $\{v_i^{(2)}\}_{i=1,2,\dots,c_{\text{opt}}}$, 可以确定前件不确定均值为

$$[\underline{m}_{ij}, \bar{m}_{ij}] = [\min(v_i^{(1)}, v_i^{(2)}), \max(v_i^{(1)}, v_i^{(2)})]. \quad (14)$$

对于规则前件高斯函数的标准差, 一般取值范围在 $[-1, 1]$ 内, 研究表明, 标准差在有限范围内的变化对系统性能的影响不大, 本文选取标准差 $\sigma = 0.5$. 由此, 通过自组织结构学习便完成了区间二型模糊神经网络的结构和前件参数辨识.

3.2 后件参数自适应学习算法及稳定性 (Adaptive learning algorithm and stability of consequent parameters)

对于任意一个观测数据 $\{x_i^k\}_{i=1,\dots,n}$, 定义性能指标函数为

$$E(k) = \frac{1}{2}[(y_d(k) - y(k))^2] = \frac{1}{2}e(k)^2, \quad (15)$$

式中 $y_d(k)$, $y(k)$ 和 $e(k)$ 分别为模型预期输出、模型实际输出和辨识误差. 本文采用负梯度下降法及链式法可得到后件权向量更新公式如下:

$$\begin{aligned} \mathbf{w}_r(k+1) &= \\ \mathbf{w}_r(k) + \eta_r(k) \left[-\frac{\partial E(k)}{\partial \mathbf{w}_r(k)} \right] &= \\ \mathbf{w}_r(k) + \frac{1}{2} \eta_r(k) \mathbf{q}_r^T e(k), \end{aligned} \quad (16)$$

$$\begin{aligned} \mathbf{w}_1(k+1) &= \\ \mathbf{w}_1(k) + \eta_l(k) \left[-\frac{\partial E(k)}{\partial \mathbf{w}_1(k)} \right] &= \\ \mathbf{w}_1(k) + \frac{1}{2} \eta_l(k) \mathbf{q}_l^T e(k), \end{aligned} \quad (17)$$

式中 $\eta_r(k)$ 和 $\eta_l(k)$ 分别为后件权向量在 k 时刻的学习速率. 为了保证迭代算法的收敛, 下面采用Lyapunov稳定收敛定理导出最佳学习速率.

定理 1 设 $\eta_{r,\text{opt}}(k)$ 和 $\eta_{l,\text{opt}}(k)$ 分别表示后件权向量 \mathbf{w}_r 和 \mathbf{w}_1 在 k 时刻的最佳速率. 如学习速率按下方式选择:

$$\eta_{r,\text{opt}}(k) = (CD - 2AE)/(4AB - C^2), \quad (18)$$

$$\eta_{l,\text{opt}}(k) = (CE - 2BD)/(4AB - C^2), \quad (19)$$

则迭代算法收敛, 其中:

$$\begin{aligned} A &= \frac{1}{16} \mathbf{q}_l^T \mathbf{q}_l \mathbf{q}_l^T \mathbf{q}_l, \quad B = \frac{1}{16} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_r^T \mathbf{q}_r, \\ C &= \frac{1}{8} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_l^T \mathbf{q}_l, \quad D = -\frac{1}{2} \mathbf{q}_l^T \mathbf{q}_l, \quad E = -\frac{1}{2} \mathbf{q}_r^T \mathbf{q}_r. \end{aligned}$$

证 为求迭代算法稳定时的最佳学习速率, 选择Lyapunov函数如下:

$$V(k) = \frac{1}{2} [(y_d(k) - y(k))^2] = \frac{1}{2} e(k)^2, \quad (20)$$

则Lyapunov函数的偏差为

$$\begin{aligned} \Delta V(k) &= \frac{1}{2} [e(k+1)^2 - e(k)^2] = \\ &= \frac{1}{2} \Delta e(k) [2e(k) + \Delta e(k)]. \end{aligned} \quad (21)$$

学习过程产生的误差偏差可表示如下:

$$\Delta e(k) = e(k+1) - e(k) = \left[\frac{\partial e(k)}{\partial \mathbf{w}} \right]^T \Delta \mathbf{w}, \quad (22)$$

式中:

$$\left[\frac{\partial e(k)}{\partial \mathbf{w}} \right]^T = \left[\frac{\partial e(k)}{\partial \mathbf{w}_r} \frac{\partial e(k)}{\partial \mathbf{w}_l} \right]^T, \quad (23)$$

$$\begin{aligned} \Delta \mathbf{w} &= [\Delta \mathbf{w}_r \quad \Delta \mathbf{w}_l] = \\ &= e(k) \left[\eta_r \frac{\partial y(k)}{\partial \mathbf{w}_r} \quad \eta_l \frac{\partial y(k)}{\partial \mathbf{w}_l} \right]. \end{aligned} \quad (24)$$

将式(22)-(24)代入式(21)得到

$$\begin{aligned} \Delta V(k) &= \frac{1}{2} \Delta e(k) [2e(k) + \Delta e(k)] = \\ &= \frac{1}{2} \{ e(k) \eta_r \left[\frac{\partial e(k)}{\partial \mathbf{w}_r} \right]^T \left[\frac{\partial y(k)}{\partial \mathbf{w}_r} \right] + \\ &= e(k) \eta_l \left[\frac{\partial e(k)}{\partial \mathbf{w}_l} \right]^T \left[\frac{\partial y(k)}{\partial \mathbf{w}_l} \right] \} \times \{ 2e(k) + \\ &= e(k) \eta_r \left[\frac{\partial e(k)}{\partial \mathbf{w}_r} \right]^T \left[\frac{\partial y(k)}{\partial \mathbf{w}_r} \right] + \\ &= e(k) \eta_l \left[\frac{\partial e(k)}{\partial \mathbf{w}_l} \right]^T \left[\frac{\partial y(k)}{\partial \mathbf{w}_l} \right] \}. \end{aligned} \quad (25)$$

由式(11)求偏导可得

$$\frac{\partial y(k)}{\partial \mathbf{w}_r} = \frac{1}{2} \mathbf{q}_r^T, \quad \frac{\partial y(k)}{\partial \mathbf{w}_l} = \frac{1}{2} \mathbf{q}_l^T. \quad (26)$$

由式 $e(k) = y_d(k) - y(k)$ 求偏导可得

$$\left[\frac{\partial e(k)}{\partial \mathbf{w}_r} \right]^T = \left[-\frac{\partial y(k)}{\partial \mathbf{w}_r} \right]^T = -\frac{1}{2} \mathbf{q}_r, \quad (27)$$

$$\left[\frac{\partial e(k)}{\partial \mathbf{w}_l} \right]^T = \left[-\frac{\partial y(k)}{\partial \mathbf{w}_l} \right]^T = -\frac{1}{2} \mathbf{q}_l. \quad (28)$$

代入式(25)后, 整理得到

$$\begin{aligned} \Delta V(k) &= \\ &= -\frac{1}{2} e(k)^2 \left[\frac{1}{4} \eta_r \mathbf{q}_r^T \mathbf{q}_r + \frac{1}{4} \eta_l \mathbf{q}_l^T \mathbf{q}_l \right] \times \\ &= \left[2 - \frac{1}{4} \eta_r \mathbf{q}_r^T \mathbf{q}_r - \frac{1}{4} \eta_l \mathbf{q}_l^T \mathbf{q}_l \right] = \end{aligned}$$

$$\frac{1}{2}e(k)^2 \left(\frac{1}{16} \mathbf{q}_1^T \mathbf{q}_1 \mathbf{q}_1^T \mathbf{q}_1 \eta_1^2 + \frac{1}{16} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_r^T \mathbf{q}_r \eta_r^2 + \frac{1}{8} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_1^T \mathbf{q}_1 \eta_r \eta_1 - \frac{1}{2} \mathbf{q}_1^T \mathbf{q}_1 \eta_1 - \frac{1}{2} \mathbf{q}_r^T \mathbf{q}_r \eta_r \right). \quad (29)$$

设

$$A = \frac{1}{16} \mathbf{q}_1^T \mathbf{q}_1 \mathbf{q}_1^T \mathbf{q}_1, \quad B = \frac{1}{16} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_r^T \mathbf{q}_r, \\ C = \frac{1}{8} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_1^T \mathbf{q}_1, \quad D = -\frac{1}{2} \mathbf{q}_1^T \mathbf{q}_1, \quad E = -\frac{1}{2} \mathbf{q}_r^T \mathbf{q}_r,$$

则有

$$\Delta V(k) = \frac{1}{2}e(k)^2 (A\eta_1^2 + B\eta_r^2 + C\eta_r\eta_1 + D\eta_1 + E\eta_r). \quad (30)$$

由Lyapunov稳定性定理可知, 如果 $\Delta V(k) < 0$ 则系统是稳定的. 要使 $\Delta V(k) < 0$, 即满足

$$G = A\eta_1^2 + B\eta_r^2 + C\eta_r\eta_1 + D\eta_1 + E\eta_r < 0. \quad (31)$$

对于式(31)所示的二次函数, $A > 0, B > 0$, 且由柯西不等式可得

$$C^2 = \left(\frac{1}{8} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_1^T \mathbf{q}_1 \right)^2 < 4 \times \left(\frac{1}{16} \mathbf{q}_1^T \mathbf{q}_1 \mathbf{q}_1^T \mathbf{q}_1 \right) \times \left(\frac{1}{16} \mathbf{q}_r^T \mathbf{q}_r \mathbf{q}_r^T \mathbf{q}_r \right) = 4AB, \quad (32)$$

则该式必然存在极小值. 分别对 η_r 和 η_1 取偏导求极值, 得

$$\begin{cases} \frac{\partial G}{\partial \eta_r} = 2B\eta_r + C\eta_1 + E = 0, \\ \frac{\partial G}{\partial \eta_1} = 2A\eta_1 + C\eta_r + F = 0. \end{cases} \quad (33)$$

由此便可解得最佳学习速率, 定理得证.

4 仿真实验(Simulation)

为了进一步验证算法的有效性, 下面首先通过一个非线性系统辨识实例, 证明该算法与其他算法相比具有收敛速度快和精度高的优点. 然后, 将该算法应用于一个实际的电力短期负荷预测, 仿真结果证明了所建立的预测模型具有一定的实用价值.

4.1 非线性动态系统辨识(Nonlinear system identification)

选取以下非线性动态系统进行辨识:

$$y(k) = \frac{y(k-1)}{1+y(k-1)^2} + \mu(k)^3, \quad k \in [1, 400], \quad (34)$$

式中: 当 $k < 100$, $\mu(k)$ 服从 $[-1, 1]$ 的均匀分布, 其余 $\mu(k) = \sin(\pi k/45)$, 获取400个样本作为训练样本对系统进行辨识. 采用本文提出的聚类算法对系统输入量分别进行划分, 其中输入量 $y(k-1)$ 的聚类有效性标准随聚类数目变化的曲线如图2所示, 从中可以判断最佳聚类数为3; 同时对另一个输入量 $\mu(k)$ 进行分析确定其最佳聚类数目也为3, 由此共产生了9

条模糊规则, 所形成的规则前件不确定均值如表1所示. 规则后件的初始参数在 $[-1, 1]$ 范围内随机选取, 并利用自适应学习算法对 \mathbf{w}_r 和 \mathbf{w}_1 进行训练. 图3给出了SOIT2FNN与文献[18]中的T1FNN和IT2FNN经过10步训练的均方根误差曲线, 显然SOIT2FNN与其他算法相比具有更快的收敛速度, 且学习误差最小, 表明其在确定网络结构和模型训练方面的优越性. 训练中每一步都以最优的学习速率进行学习迭代, 学习速率 η_r 和 η_1 的变化曲线如图4所示.

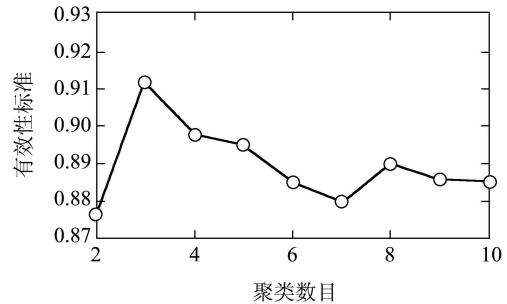


图 2 有效性标准随聚类数目变化的曲线
Fig. 2 The curve of cluster validity criterion

表 1 规则前件不确定均值

Table 1 Uncertain means of rule antecedent

	$y(k-1)$	$\mu(k)$
第1组	[-1.1087, -1.0774]	[-0.8331, -0.8179]
第2组	[-0.0798, -0.0550]	[0.0076, 0.0167]
第3组	[1.979, 1.1324]	[0.8270, 0.8455]

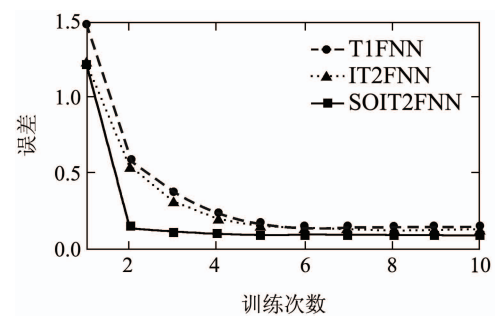


图 3 训练误差曲线

Fig. 3 The curve of training error

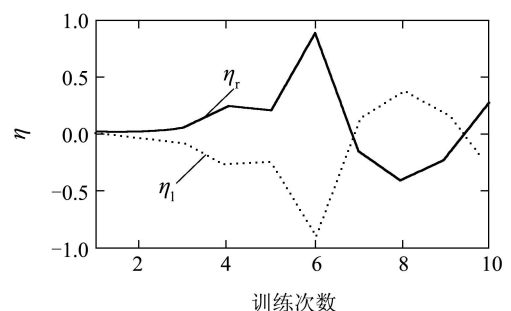


图 4 学习速率变化

Fig. 4 The change of learning rates

网络训练完成后,选取不同的系统输入量产生测试样本以便验证网络的泛化性能,如下所示:

$$u(k) = \begin{cases} -0.7 + \frac{\text{mod}(k, 50)}{40}, & k \leq 80, \\ \text{rands}(1, 1), & 80 < k \leq 130, \\ -0.7 - \frac{\text{mod}(k, 180)}{180}, & 130 < k \leq 250, \\ 0.6 \cos \frac{\pi k}{50}, & k > 250. \end{cases} \quad (35)$$

测试样本的实际值和模型输出值如图5所示,可见模型具有较好的逼近精度和泛化能力.训练样本和测试样本的各项性能指标如表2所示.

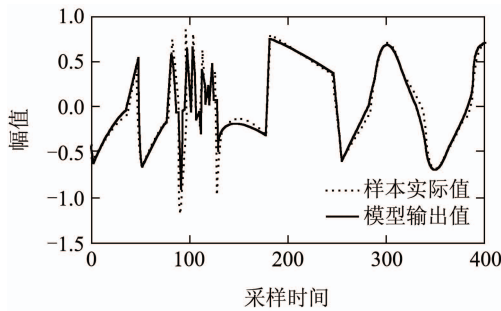


图5 测试样本辨识效果

Fig. 5 Identification result of testing samples

表2 训练样本和测试样本的性能指标
Table 2 Performance index of training and testing samples

性能指标	训练样本 模型值(实际值)	测试样本 模型值(实际值)
REMS	0.06002	0.07587
APE	2.6112%	3.247%
mean	14.83(14.83)	0.02657(0.02658)
min	13.97(13.8)	-0.9124(-1.185)
max	15.85(16.09)	0.7539(0.8498)

表2中: RMSE为均方根误差, APE为平均百分误差, mean为均值, min为最小值, max为最大值.

4.2 在电力负荷预测中的应用(Application in short-term power load forecasting)

电力负荷预测是电力系统规划的重要组成部分,准确的电力负荷预测是合理进行电力系统规划、建设、生产、调度以及检修的重要依据.然而,电力负荷的波动受诸多自然、社会因素不同程度的影响,且各种影响因素具有一定的复杂不确定性,增加了其准确预测的难度.

利用本文的算法来建立电力负荷预测模型, $x = [x(k-3), x(k-2), x(k-1), x(k)]^T$ 为输入变量,

$x(k+1)$ 为输出变量.选取我国某市2011年5月份各小时负荷时间序列作为实验数据,共744个数据,其中前504个(前21天)用来建立预测模型,后240个(后10天)用来检验预测结果.经过本文聚类算法处理确定每个输入变量的最佳聚类数目为3,可构成 $3^4 = 81$ 条规则,共有 $2 \times 3 \times 4 + 2 \times 81 = 186$ 个需要辨识的参数.经过200步训练后,训练结果和测试结果如图6和图7所示,可见该模型获得较高的逼近精度,能较好地跟踪负荷曲线的实际走向.表3给出了SOIT2FNN与其他几种方法的电力负荷预测性能的比较结果,从表中可以看出,SOIT2FNN具有更高的预测精度和更低的泛化误差,虽然比SVD-BP和SVD-SM-BP产生的规则数多,但需要辨识的参数个数明显更少.

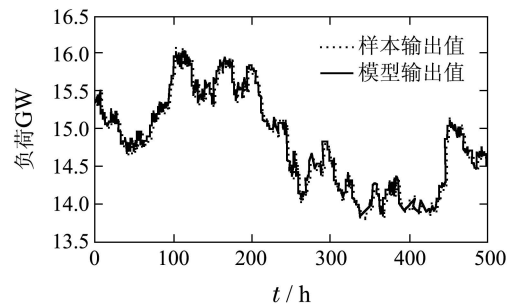


图6 训练结果

Fig. 6 Training result

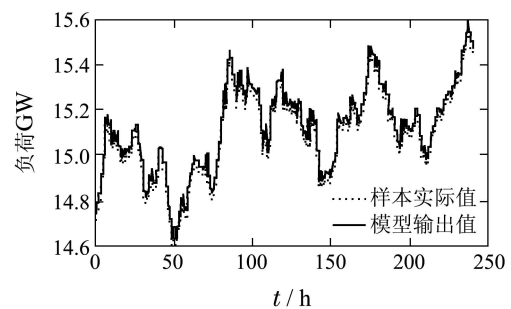


图7 测试结果

Fig. 7 Testing result

表3 SOIT2FNN与其他模型性能比较

Table 3 Performance comparison of SOIT2FNN with other models

模型	规则数量	参数个数	训练REMS	测试REMS	测试APE/%
MPSO-FNN ^[19]	—	—	—	1.4274	1.182
PSO-BP ^[19]	—	—	—	2.0552	1.705
SVD-BP ^[20]	34	484	—	—	2.58
SVD-SM-BP ^[20]	31	442	—	—	2.58
SOIT2FNN	81	186	0.01549	0.02165	1.4

5 结论(Conclusions)

本文利用模糊聚类中加权参数控制着聚类模糊

性的结论, 提出了一种自组织区间二型模糊神经网络, 能够基于输入输出数据自动完成结构辨识和参数辨识. 主要特点在于: 1) 提出了一种新的确定区间二型模糊规则前件参数的方法; 2) 利用聚类有效性标准自动确定规则数目, 避免了由于错误估计聚类数目而使算法不能正确揭示数据聚类结构的问题; 3) 基于梯度下降法和Lyapunov函数稳定收敛定理给出规则后件权向量学习速率的自适应学习算法, 由于每一步训练都是以最优学习速率进行学习迭代, 因而加快了模型训练的收敛速度. 通过系统辨识实例验证了该方法与其他方法相比具有收敛速度快、逼近精度高的优点; 利用该算法建立的电力短期负荷预测模型具有较高的预测精度和较强的泛化能力, 表明了该预测模型的可行性和有效性. 下一步本文将研究适用于前件参数进一步优化的算法, 同时将研究成果应用到含有复杂不确定性因素的电力系统故障诊断和预测等.

参考文献(References):

- [1] JUANG C F. A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithm [J]. *IEEE Transactions on Fuzzy Systems*, 2002, 10(2): 155 – 170.
- [2] ABIYEV R H, KAYNAK O. Fuzzy wavelet neural networks for identification and control of dynamic plants—a novel structure and a comparative study [J]. *IEEE Transactions on Industrial Electronics*, 2008, 55(8): 3133 – 3140.
- [3] MENDEL J M, WU D R. Perceptual reasoning for perceptual computing [J]. *IEEE Transactions on Fuzzy Systems*, 2008, 16(6): 1550 – 1564.
- [4] KARNIK N N, MENDEL J M. Type-2 fuzzy logic systems [J]. *IEEE Transactions on Fuzzy Systems*, 1999, 7(6): 643 – 657.
- [5] MENDEL J M, JOHN R I B. Type-2 fuzzy sets made simple [J]. *IEEE Transactions on Fuzzy Systems*, 2002, 10(2): 117 – 127.
- [6] ZADEH L A. The concept of linguistic variable and its application to approximate reasoning [J]. *Information Sciences*, 1975, 8(3): 199 – 249.
- [7] 潘永平, 孙宗海, 黄道平. II型模糊集合与系统研究进展 [J]. *控制理论与应用*, 2011, 28(12): 1693 – 1703.
(PAN Yongping, SUN Zonghai, HUANG Daoping. A survey of type-2 fuzzy sets and systems [J]. *Control Theory & Applications*, 2011, 28(12): 1693 – 1703.)
- [8] COUPLAND S, JOHN R. Type-2 fuzzy logic and the modeling of uncertainty [M] // *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models*. Berlin: Springer-Verlag, 2008, 220: 3 – 22.
- [9] JUANG C F, TSAO Y W. A self-evolving interval type-2 fuzzy neural network with online structure and parameter learning [J]. *IEEE Transactions on Fuzzy Systems*, 2008, 16(6): 1411 – 1424.
- [10] CASTRO J R, CASTILLO O, MELIN P, et al. A hybrid learning algorithm for interval type-2 fuzzy neural networks: the case of time series prediction [M] // *Soft Computing for Hybrid Intelligent Systems*. Berlin: Springer-Verlag, 2008, 154: 363 – 386.
- [11] LEE C H, HU T W, LEE C T, et al. A recurrent interval type-2 fuzzy neural network with asymmetric membership functions for nonlinear system identification [C] // *IEEE World Congress on Computational Intelligence*. Hong Kong: IEEE, 2008, 6: 1496 – 1502.
- [12] YE H C Y, JENG W R, LEE S J. Data-based system modeling using a type-2 fuzzy neural network with a hybrid learning algorithm [J]. *IEEE Transactions on Neural Networks*, 2011, 22(12): 2296 – 2309.
- [13] ABIYEV R H, KAYNAK O. A type-2 neural fuzzy system based on clustering and gradient techniques applied to system identification and channel equalization [J]. *Applied Soft Computing*, 2011, 11(1): 1396 – 1406.
- [14] HWANG C, RHEE F C H. Uncertain fuzzy clustering: interval type-2 fuzzy approach to C-means [J]. *IEEE Transactions on Fuzzy Systems*, 2007, 15(1): 107 – 120.
- [15] WU D R, MENDEL J M. Enhanced Karnik-Mendel algorithms [J]. *IEEE Transactions on Fuzzy Systems*, 2009, 17(4): 923 – 934.
- [16] OZKAN I, TURKSEN I B. Upper and lower values for the level of fuzziness in FCM [J]. *Information Sciences*, 2007, 177(23): 5143 – 5152.
- [17] CHEN M Y, LINKENS D A. Rule-base self-generation and simplification for data-driven fuzzy models [J]. *Fuzzy Sets and Systems*, 2004, 142(2): 243 – 265.
- [18] LEE C H, HONG J L, LIN Y C, et al. Type-2 fuzzy neural network systems and learning [J]. *International Journal of Computational Cognition*, 2003, 1(4): 79 – 90.
- [19] 师彪, 李郁侠, 于新花, 等. 基于改进粒子群—模糊神经网络的短期电力负荷预测 [J]. *系统工程理论与实践*, 2010, 30(1): 157 – 166.
(SHI Biao, LI Yuxia, YU Xinhua, et al. Short-term load forecasting based on modified particle swarm optimizer and fuzzy neural network model [J]. *Systems Engineering—Theory & Practice*, 2010, 30(1): 157 – 166.)
- [20] ZHENG G. *Studies on fuzzy measure of type-2 fuzzy logic system and its application to power load forecasting* [D]. Chengdu: Southwest Jiaotong University, 2011.

作者简介:

姚 兰 (1980–), 女, 博士研究生, 主要研究方向为模糊系统辨识, E-mail: dancyao@163.com;

肖 建 (1950–), 男, 教授, 主要研究方向为模糊控制及计算机控制, E-mail: jxiao@home.swjtu.edu.cn;

王 嵩 (1983–), 男, 博士研究生, 主要研究方向为电机控制与粘着控制等, E-mail: songwang@swjtu.edu.cn;

蒋玉莲 (1978–), 女, 博士研究生, 主要研究方向为系统工程与分布式滤波等, E-mail: alian96432@yahoo.com.cn.