

文章编号: 1001-0920(2011)11-1721-05

基于可拓理论的 RBF 神经网络研究及其应用

徐 圆, 冯 晶, 朱群雄

(北京化工大学 信息科学与技术学院, 北京 100029)

摘 要: 针对径向基函数(RBF)神经网络构造时其结构和参数难以确定的问题, 结合可拓理论对输入样本和基函数的中心向量建立物元模型, 并借鉴第 2 类型可拓神经网络(ENN2)的聚类思想, 根据样本分布, 采用可拓分析及可拓变换动态调整隐节点数目和基函数中心, 从而提出基于可拓理论的 RBF(ERBF)神经网络. 同时, 通过 UCI 标准数据集进行了测试, 并通过应用实例进行了验证, 结果表明, ERBF 结构和参数的确定方法简单、收敛速度快, 且泛化精度、鲁棒性和稳定性均显著提高.

关键词: 径向基函数神经网络; 可拓理论; 回归; 建模

中图分类号: TP29

文献标识码: A

Research and application of extension theory-based radial basis function neural network

XU Yuan, FENG Jing, ZHU Qun-xiong

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China.

Correspondent: ZHU Qun-xiong, E-mail: zhuqx@mail.buct.edu.cn)

Abstract: During the construction process of radial basis function(RBF) neural network, the structure and parameters are hard to be determined. Therefore, combining with the extension theory, an extension theory-based RBF(ERBF) neural network is proposed, in which the matter-element models including input samples and center vectors of the basis function are established, the clustering method of extension neural network type 2(ENN2) is introduced, and the hidden layer nodes number and center vectors of the basis function are dynamically adjusted by using extension analysis and extension transformation according to the sample distribution. Meanwhile, UCI standard data sets are tested, and application object is validated. Through the verification and comparison, the proposed ERBF algorithm has the advantages of simple calculation and fast convergence, which significantly enhances the generalization accuracy, robustness and stability.

Key words: radial basis function neural network; extension theory; regression; modeling

1 引 言

过程工业具有高维数、强关联和非线性等特点, 其过程模型越来越复杂, 求解难度越来越大. 以网络结构简单、非线性全局逼近能力强以及收敛速度快等优点著称的径向基函数(RBF)神经网络^[1]在建模中得到了广泛应用, 而决定 RBF 结构的主要因素有 2 个: 网络隐含层神经元的数目及其中心向量和宽度; 隐含层与输出层的连接权值^[2-3]. 然而, 如何有效地确定其网络结构和参数, 虽然已有大批学者致力于此, 但一直没有得到公认的、有效的解决方法.

可拓理论^[4]作为新型的人工智能技术, 采用形式化模型讨论事物拓展的可能性以及开拓创新的规

律和方法, 近几年在经济、管理、控制和决策等领域得到研究和应用. 其中基于可拓理论的神经网络技术^[5]是继模糊神经网络、进化神经网络之后的又一新网络. 以台湾学者王孟辉为代表的多位专家近几年通过融合可拓理论和神经网络技术的优点, 提出了双权值连接结构, 并形成了多种 ENN^[6-8]类型. 与传统神经网络采用梯度下降法不同, ENN 权值训练过程基于可拓距离根据相应公式进行调整, 该类型网络结构简单, 权值意义清晰明了, 对基于区间的分类和聚类问题特别有效, 可细分为本质为分类的监督学习算法 ENN 1 和本质为聚类的非监督学习算法 ENN 2. 其中 ENN 2 无需事先确定聚类数目, 对新数据有很好

收稿日期: 2010-12-23; 修回日期: 2011-03-23.

基金项目: 国家自然科学基金项目(60774079, 61074153); 国家 863 计划项目(2006AA04Z184).

作者简介: 徐圆(1983—), 女, 讲师, 博士, 从事可拓工程等研究; 朱群雄(1960—), 男, 教授, 博士生导师, 从事人工智能、数据挖掘等研究.

的自适应性,在聚类过程中保持了稳定性和灵活性.而RBF研究的重点是对输入样本的聚类.基于此,本文通过建立输入变量和隐含层基函数中心向量的物元模型,采用可拓分析和可拓变换,根据样本分布动态调整隐含层节点数目和基函数中心向量,进而计算每个基函数的宽度,最终探索出基于可拓理论的RBF神经网络.同时采用UCI标准数据集进行验证,并推广应用于精对苯二甲酸(PTA)溶剂系统过程建模.结果表明,此方法合理、有效、实用.

2 基于可拓理论的RBF神经网络

RBF为前馈式神经网络,只有一个隐含层,其隐含层节点数目通常根据经验或试凑确定,因此带有很大的主观性.针对此问题,借鉴ENN2聚类思想,运用可拓变换动态调整隐含层节点数目和基函数的中心,形成基于可拓理论的RBF神经网络(ERBF).其中可拓理论采用形式化的模型(物元、事元、关系元),利用其可拓性和基本变换描述事物变化,为问题求解提供策略.可拓性(发散性、可扩性、相关性、蕴含性)用于拓展解决问题的方向,基本变换(置换、增删、扩缩和分解),变换运算(积变换、与变换、或变换和逆变换)及复合变换用于消解问题解决过程中发生的矛盾,从而生成多种解决方案^[9-10].

RBF神经网络研究的重点是输入空间的数据聚类.本文首先采用多维物元模型描述各输入样本,并

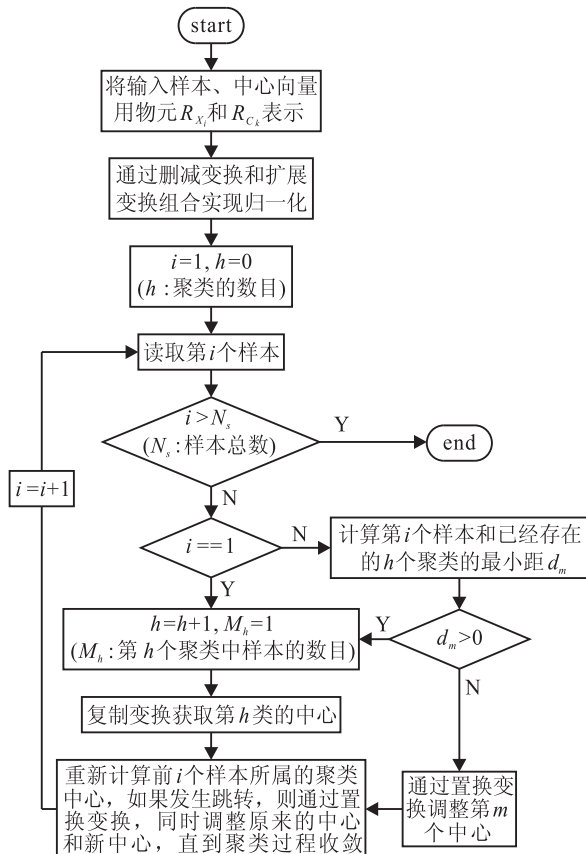


图1 可拓聚类过程流程

运用可拓理论中变换的可组合性,通过删减变换和扩缩变换组合实现归一化;其次,采用多维物元模型描述中心向量,设置距离参数阈值,通过对第1个样本进行复制变换获得第1个聚类中心,同时对中心物元的对象实施主动变换,引起对应特征值的传导变换;再次,将可拓学中距的定义扩展到 n 维物元中,得到扩展的可拓距公式,并根据此公式计算出样本和所有已存在中心的最小可拓距,据此判断样本是否属于当前已存在的类,然后分情况运用置换变换进行中心和权值的调整,直到所有的聚类过程收敛;最后,计算宽度,并运用LMS算法进行隐含层和输出层的权值调整.其中可拓聚类过程的流程如图1所示.

基于以上可拓聚类流程,ERBF算法步骤如下:

1) 建立输入样本的多维物元模型,并求取各个特征的最大、最小值.第 i 个输入样本的多维物元模型表示如下:

$$R_{X_i} = \begin{bmatrix} X_i & a_1 & x_{i1} \\ & a_2 & x_{i2} \\ & \vdots & \vdots \\ & a_n & x_{in} \end{bmatrix}, i = 1, 2, \dots, N_s; \quad (1)$$

$$x_{j,\max} = \max\{x_{ij}\}, x_{j,\min} = \min\{x_{ij}\},$$

$$j = 1, 2, \dots, n. \quad (2)$$

其中: n 为样本特征个数, N_s 为样本总数, X_i 为第 i 个样本, a_j 为 X_i 的第 j 个特征, x_{ij} 为第 i 个样本关于第 j 个特征的值, $x_{j,\max}$ 和 $x_{j,\min}$ 为第 j 个特征的最大、最小值.

因各变量描述对象不同,量纲不统一,将导致各变量对网络训练的贡献不同,因此通过对 R_{X_i} 实施删减变换和扩缩变换的组合实现归一化.

$$x_{ij}^{(1)} = T_1 x_{ij} = x_{ij} - x_{j,\min};$$

$$x_{ij}^{(2)} = T_2 x_{ij}^{(1)} = \alpha x_{ij}^{(1)}, \alpha = 1/(x_{j,\max} - x_{j,\min});$$

$$R_{NX_i} = \begin{bmatrix} NX_i & a_1 & x_{i1}^{(2)} \\ & a_2 & x_{i2}^{(2)} \\ & \vdots & \vdots \\ & a_n & x_{in}^{(2)} \end{bmatrix} = T_1 T_2 R_{X_i} = \begin{bmatrix} NX_i & a_1 & (x_{i1} - x_{1,\min})/(x_{1,\max} - x_{1,\min}) \\ & a_2 & (x_{i2} - x_{2,\min})/(x_{2,\max} - x_{2,\min}) \\ & \vdots & \vdots \\ & a_n & (x_{in} - x_{n,\min})/(x_{n,\max} - x_{n,\min}) \end{bmatrix}. \quad (3)$$

其中: T_1 为删减变换, T_2 为扩缩变换, $T_1 T_2$ 为变换 T_1 与 T_2 的组合, NX_i 为归一化后的第 i 个样本, a_j 为 NX_i 的第 j 个输入特征, $x_{ij}^{(2)}$ 为第 i 个样本关于第 j 个特征的值.

2) 采用多维物元模型规范化描述基函数中心向

量, 运用经典域描述各特征取值的上、下限, 如下式所示:

$$R_{C_k} = \begin{bmatrix} C_k & a_1 & c_{k_1} \\ & a_2 & c_{k_2} \\ & \vdots & \vdots \\ & a_n & c_{k_n} \end{bmatrix} = \begin{bmatrix} C_k & a_1 & \langle c_{k_1}^L, c_{k_1}^U \rangle \\ & a_2 & \langle c_{k_2}^L, c_{k_2}^U \rangle \\ & \vdots & \vdots \\ & a_n & \langle c_{k_n}^L, c_{k_n}^U \rangle \end{bmatrix}, \quad k = 1, 2, \dots, h. \quad (4)$$

其中: C_k 为第 k 个聚类的中心, 即第 k 个基函数的中心向量; a_j 为 C_k 的第 k 个特征; c_{k_j} 为第 k 个聚类关于第 j 个特征的值; $V_{k_j} = \langle c_{k_j}^L, c_{k_j}^U \rangle$ 为特征 a_j 的经典域, L 为下限值, U 为上限值; h 为聚类中心的数目.

3) 为度量聚类中心和期望边界距离, 设置距离参数阈值 θ , 此值一般根据样本分布等先验知识或实验法确定. 定义 $c_{k_j}^L = c_{k_j} - \theta, c_{k_j}^U = c_{k_j} + \theta, M_h$ 为第 h 个聚类中样本的数目.

读取第 i ($i = 1$) 个样本, 创建第 1 个类, 初始化 $h = 1, M_h = 1$, 采用 R_{NX_1} 复制变换获得第 1 个类中心, 即对对象 NX_i 实施主动变换 $T_{NX_i} NX_i = C_h$, 引起量值传导变换 $NX_i T_{x_{i_j}^{(2)}} x_{i_j}^{(2)} = C_{h_j}$.

$$R_{C_h} = T_{NX_i \rightarrow C_h} R_{NX_i} = \begin{bmatrix} C_h & a_1 & x_{i_1}^{(2)} \\ & a_2 & x_{i_2}^{(2)} \\ & \vdots & \vdots \\ & a_n & x_{i_n}^{(2)} \end{bmatrix}. \quad (5)$$

4) 定义多维物元空间中点和区间的距. 可拓学通过定义点 x 和区间 $X_0 = \langle a, b \rangle$ 的距 ρ , 将点与区间的位置关系用定量形式精确刻画, 使人们从“类内即为同”发展到类内也有程度区别的定量描述. 但是, 此公式只适用于一维物元, 无法描述多维物元空间中的点与区间的关系. 为克服这一局限性, 借鉴可拓学中距的概念, 本文对距的定义进行扩展, 提出了 n 维空间中点与区间距的表达式如下:

$$d(NX_i, C_k) = \sum_{j=1}^n \alpha_j \rho(x_{i_j}^{(2)}, \langle c_{k_j}^L, c_{k_j}^U \rangle) = \sum_{j=1}^n \alpha_j [|x_{i_j}^{(2)} - c_{i_j}| - (c_{k_j}^U - c_{k_j}^L)/2]. \quad (6)$$

其中: α_j 为权系数, 用以衡量各个变量重要性的程度, 当各变量的重要性程度差别较大时, 可考虑采用层次分析法等来确定. 为了避免盲目设置权系数带来的种种问题, 这里将权重平均分配, 选取 $\alpha_j = 1/n$.

与 ENN2^[7] 中提出的可拓距离公式相比, 尽管本文的公式更加简化, 但能够实现相同的目的, 即衡量样本与中心的距离时, 同时引入了权系数, 考虑问题更加全面. 由式 (6) 可知, 当样本 X_i 越接近聚类中心

C_k 时, 可拓距越小; 越远离 C_k 时, 可拓距越大; 当样本的值等于 C_k 时, 可拓距最小, 为 $-\theta$.

5) 读入下一个样本向量 X_i , 对已经存在的 h 个聚类, 采用式 (6) 计算第 i 个样本与其中第 k 个聚类中心的距离. 从获取的 h 个距中, 采用下式查找与第 i 个样本距离最近的聚类编号 m :

$$d_m = \min\{d(NX_i, C_k)\}, k = 1, 2, \dots, h. \quad (7)$$

其中: $d(NX_i, C_k)$ 为第 i 个样本与第 k 个聚类中心的距离, d_m 为第 i 个样本与所有已经存在聚类的最小距离.

6) 如果 $d_m > 0$, 表明第 i 个样本不属于已经存在的所有聚类, 需增加 1 个新聚类, 则 $h = h + 1, M_h = 1$, 采用式 (5) 的复制变换确定该聚类中心; 反之, 表明第 i 个样本属于第 m 个聚类中心, 则 $M_m = M_m + 1$. 根据当前的样本分布特征, 通过下式所示的置换变换更新第 m 个聚类中心:

$$c_{m_j} = Tc_{m_j} = \frac{c_{m_j} * (M_m - 1) + x_{i_j}^{(2)}}{M_m}, \quad R_{C_m} = TR_{C_m} = \begin{bmatrix} C_m & a_1 & \frac{c_{i_1} * (M_m - 1) + x_{i_1}^{(2)}}{M_m} \\ & a_2 & \frac{c_{i_2} * (M_m - 1) + x_{i_2}^{(2)}}{M_m} \\ & \vdots & \vdots \\ & a_n & \frac{c_{i_n} * (M_m - 1) + x_{i_n}^{(2)}}{M_m} \end{bmatrix}. \quad (8)$$

7) 重新判断第 i 个样本之前所有样本的聚类中心是否发生跳转. 若其中第 j 个样本从聚类中心 k 变化到 m , 则 $M_m = M_m + 1, M_k = M_k - 1$, 采用式 (8) 更新第 m 个聚类中心, 同时采用加减号互换修正式 (8), 据此更新第 k 个聚类中心. 如此循环, 直到所有样本聚类中心不发生变动. 学习过程只需调整聚类 k 和聚类 m 所对应的权值, 能够快速适应新的和重要的信息.

8) 设置 $i = i + 1$, 重复第 5)~第 7) 步, 直到对比完所有样本和已经存在的聚类中心. 如果聚类过程收敛, 则进入下一步; 否则, 回到第 3) 步, 调整阈值.

9) 设置隐含层节点数目等于聚类的数目 h , 选择高斯函数作为基函数, 将各个聚类中心设为基函数中心向量, 按最近邻中心距离的 1~2 倍计算每个基函数宽度 δ_i ; 然后, 采用最小均方算法求取隐含层到输出层之间的权值.

从上述步骤可以看出, 相比 RBF, ERBF 对输入样本的聚类过程进行了优化, 允许对重要的和新的数据进行自适应学习, 兼顾了稳定性和灵活性, 物理意义清晰, 能更好地模拟人脑神经系统的思维过程, 具有较强的学习和推理能力, 并且能在学习过程中不断积

累和完善,具有较强的自适应能力.

3 实验测试

为验证 ERBF 算法的有效性和稳定性,选取 UCI 数据库中的 Friedman#1 函数及 Housing 数据集进行测试,其各自属性如表 1 所示.其中 Friedman #1 数据

集中每个输入向量含有 5 个分量,由下式产生:

$$y = 10 \sin(\pi x_1 x_2) + 20(x_3 - 0.5)^2 + 10x_4 + 5x_5 + \varepsilon. \quad (9)$$

式中:分量 $X_i (i=1, 2, \dots, 5)$ 服从 $[0, 1]$ 上的均匀分布, ε 服从正态分布 $N(0, 1)$.

表 1 数据测试集的属性

Test Data Sets	Total Sample Number (Training / Generalization)	Input Variable Number	Output Variable Number
Friedman#1	1 400 (934 / 466)	5	1
Housing	506 (400 / 106)	13	1

3.1 Friedman #1 实验测试

选取 Friedman #1 数据集进行函数拟合测试,表 2 列出了 ERBF 网络的各项性能指标随参数 θ 的变化情况.

表 2 参数 θ 与隐含层节点数目的关系

θ	0.1	0.15	0.2	0.25
Hidden Layer Nodes Number	302	85	34	20
Recall Error / %	4.31	6.55	10.14	18.31
Recall Standard Deviation	0.038 6	0.060 5	0.094 0	0.151 2
Generalization Error / %	6.72	7.76	11.29	17.66
Generalization Standard Deviaton	0.068 3	0.071 8	0.114 9	0.141 4

θ 是 ERBF 中的重要参数,其选取直接影响网络的结构和精度.从表 2 可以看出,随着 θ 增大,隐含层节点数目减少,回想和泛化误差、标准差增大,精度降低,稳定性变差.可见 θ 的选取要综合权衡网络复杂度和精度,即在达到精度要求的情况下尽可能使网络结构简单.

为验证 ERBF 算法性能,将 ERBF 与基于 k 均值聚类的 RBF 神经网络设计方法进行比较.表 3 记录了两种算法第 2 阶段训练的迭代次数、回想的相对误差和标准偏差、泛化的相对误差和标准偏差.其中: θ 设置为 0.15, ERBF 隐含层节点数目为 85.为更好地比较算法精度,将 RBF 隐含层节点数目也设置为 85,以保证网络结构相同.

表 3 Friedman #1 数据集上 RBF 和 ERBF 测试结果对比

Friedman #1 Dataset	RBF	ERBF
Iteration Times	1 000	243
Recall Error / %	8.31	6.55
Recall Standard Deviation	0.072 4	0.060 5
Generalization Error / %	9.51	7.76
Generalization Standard Deviation	0.140 0	0.071 8

从表 3 可以看出,在 Friedman#1 数据集上,ERBF 算法相比 RBF 具有更高的训练和泛化精度,且训练时间更短,标准偏差更小,稳定性更佳.ERBF 只需 243 次训练即可达到精度要求,反映出 ERBF 在聚类阶段确定的中心向量更接近真实值,为第 2 阶段训练奠定了基础,故在网络结构相同的条件下,可以更快收敛.

3.2 Housing 实验测试

从 UCI 机器学习数据库中选择 Housing 数据集,采用 ERBF 算法与文献 [11] 提出的两种输入变量和基函数同步选择的 RBF 改进算法进行对比,其中 θ 设置为 0.07.表 4 列出了各算法中选择的输入变量和基函数的数目以及回想相对误差和泛化相对误差.

表 4 Housing 数据集上实验结果对比

Housing Datasets	Algorithm 1 ^[11]	Algorithm 2 ^[11]	ERBF
Number of Inputs	12	13	13
Number of Basic Functions	152	227	66
Recall Error / %	0.13	0.12	0.12
Generalization Error / %	0.15	0.18	0.146

由表 4 可看出,与文献 [11] 提出的 RBF 改进算法相比,ERBF 网络结构更加简单,回想和泛化误差更小.ERBF 回想标准差为 0.126 8,泛化标准差为 0.154 9.标准差很小说明该算法稳定性很好.

4 应用实例

为进一步验证 ERBF 的实用价值,将其推广应用用于中石化某 PTA 溶剂系统过程建模,用于预测 PTA 溶剂系统中共沸塔塔顶醋酸含量,为 PTA 生产工艺优化奠定基础.PTA 溶剂系统是 PTA 生产过程中的一个重要环节,主要用于提纯 PTA 氧化工段的醋酸.选取影响醋酸消耗量的 17 个主要因素^[12]作为输入变量.由于醋酸含量不易测量,常采用共沸精馏塔塔顶的电导率来间接反应醋酸含量的变化,这里以此作为输出变量.

采集现场的实时数据,通过对其进行数据融合、缺失数据填充、滤波、稳态检测等数据预处理,将处理后的数据按 2:1 的比例分为训练和泛化样本集进行测试,RBF 和 ERBF 测试结果如表 5 所示,其相应模型泛化曲线如图 2 所示.

从应用结果可以看出,相比于 RBF,ERBF 的泛化精度从 0.98% 提高到了 0.70%,而且标准差从 0.008 1 提高到了 0.004 9,显然算法泛化精度和稳定性都得到了明显改善.在 RBF 训练预测值中有个别预测值偏离真实值比较大,而 ERBF 不存在这个问题,从而增

强了鲁棒性和可信度. 在实际生产过程中, 可以更准确地预测醋酸含量的变化趋势, 以便后续的工艺改造和优化.

表 5 PTA 溶剂系统 RBF 和 ERBF 测试结果对比

PTA Solvent System	RBF	ERBF
Iteration Times	1 000	142
Train Error / %	0.98	0.70
Train Standard Deviation	0.008 1	0.004 9
Generalization Error / %	1.05	0.68
Generalization Standard Deviation	0.008 5	0.004 9

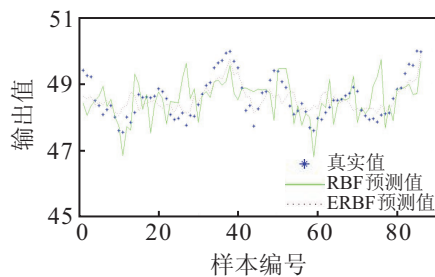


图 2 PTA 溶剂系统泛化结果图

5 结 论

本文提出了一种基于可拓理论的 RBF 神经网络学习算法, 运用可拓分析和可拓变换方法动态调整 RBF 神经网络的结构和参数, 优化了输入样本的聚类过程. 通过 UCI 标准数据集测试以及对 PTA 溶剂系统过程建模的应用表明, ERBF 不仅易于实现, 具有较强的语义解释性, 而且结构简单、收敛速度快、泛化性能好. 将所提出的方法应用于工业建模, 可大大提高建模的效率.

参考文献(References)

- [1] Wan C H. Self-configuring radial basis function neural networks for chemical pattern recognition[J]. J of Chemical Information and Computer Science, 1999, 39(6): 1049-1056.
- [2] Du D J, Li K, Fei M R. A fast multi-output RBF neural network construction method[J]. Neurocomputing, 2010, 73(10/11/12): 2196-2202.
- [3] Feng H M. Self-generation RBFNs using evolutionary PSO learning[J]. Neurocomputing, 2006, 70(1/2/3): 241-251.
- [4] 蔡文, 石勇. 可拓学的科学意义与未来发展[J]. 哈尔滨工业大学学报, 2006, 38(6): 1079-1086.
(Cai W, Shi Y. Extenics: Its signification in science and prospects in application [J]. J of Harbin Institute of Technology, 2006, 38(6): 1079-1086.)
- [5] 孙佰清, 邢爱国, 张积宾, 等. 可拓神经网络模型的设计与实现[J]. 哈尔滨工业大学学报, 2006, 38(7): 1156-1159.
(Sun B Q, Xing A G, Zhang J B, et al. Design and implementation of extension neural network model[J]. J of Harbin Institute of Technology, 2006, 38(7): 1156-1159.)
- [6] Wang M H, Hung C P. Extension neural network and its applications[J]. Neural Networks, 2003, 16(5/6): 779-784.
- [7] Wang M H. Extension neural network-type 2 and its applications[J]. IEEE Trans on Neural Networks, 2005, 16(6): 1352-1452.
- [8] Wang M H. Extension neural network type 3[J]. Lecture Notes in Computer Science, 2005, 3496(1): 503-508.
- [9] 顾祥柏, 朱群雄. 基于物元的流程报警系统重构算法及应用[J]. 控制与决策, 2005, 20(9): 1006-1011.
(Gu X B, Zhu Q X. Alarm system reconstruction algorithm and application for process industry based on matter-element[J]. Control and Decision, 2005, 20(9): 1006-1011.)
- [10] Xu Y, Zhu Q X. A new design method for energy saving and consumption reducing of process industry based on extension theory[J]. Int J of Innovative Computing, Information and Control, 2010, 6(4): 1571-1582.
- [11] Jarkko T. Simultaneous input variable and basis function selection for RBF networks[J]. Neurocomputing, 2009, 72(10/11/12): 2649-2658.
- [12] 徐圆, 朱群雄. 基于物元分析的过程工业报警优化[J]. 化工学报, 2008, 59(7): 1609-1614.
(Xu Y, Zhu Q X. Alarm optimization for process industry based on matter-element analysis[J]. J of Chemical Industry and Engineering, 2008, 59(7): 1609-1614.)

(上接第 1720 页)

- [9] Huang Y, Zhang W, Feng G. Infinite horizon H_2/H_∞ control for stochastic systems with Markovian jumps[J]. Automatica, 2008, 44(3): 857-863.
- [10] Ho D W C, Lu G. Robust stabilization for a class of discrete-time nonlinear system via output feedback: The unified LMI approach[J]. Int J of Control, 2003, 76(2): 105-115.
- [11] Kolmanovskii V B, Myshkis A. Applied theory of functional differential equations[M]. Dordrecht: Kluwer, 1992.