

基于禁忌搜索的模糊神经网络结构优化^{*})

肖丽¹ 刘光远² 贺一^{1,3} 邱玉辉¹

(西南师范大学计算机与信息科学学院 重庆 400715)¹ (西南师范大学电子信息工程学院 重庆 400715)²
(重庆师范大学管理学院 重庆 400047)³

摘要 本文提出了一种基于禁忌搜索的模糊神经网络自动优化学习方法(fuzzy neural network based on tabu search, FNNTS)。该方法利用禁忌搜索算法搜索最优的模糊神经网络结构,并结合最小二乘法和梯度下降法对网络参数进行学习,大大减少了对专家知识的依赖。非线性函数逼近的实验结果表明,所提出的方法能获得更精练的网络结构和更小的误差,从而验证了本文方法的有效性和可行性。

关键词 禁忌搜索,模糊神经网络,函数逼近

Structure Optimization of Fuzzy Neural Network Based on Tabu Search

XIAO LI¹ LIU Guang-Yuan² HE Yi^{1,3} QIU Yu-Hui¹

(Faculty of Computer and Information Science, Southwest Normal University, Chongqing 400715)¹
(School of Electronic and Information Engineering, Southwest Normal University, Chongqing 400715)²
(Dept. of Modern Information Management, Chongqing Normal University, Chongqing 400047)³

Abstract In this paper, a method based on tabu search (TS) is proposed for constructing fuzzy neural network automatically. TS is used to search the minimal and optimal structure of fuzzy neural network, after which least squares and gradient descend method are combined to adjust the parameters of the network. A simulation for a nonlinear function approximation is presented and the experimental results show that the proposed algorithm can generate a more compact structure with a lower average percentage error.

Keywords Tabu search, Fuzzy neural network, Function approximation

模糊神经网络(Fuzzy Neural Network, FNN)作为模糊技术和神经网络的结合,因具有模糊推理的可理解性和神经网络的学习能力而得到广泛的应用。模糊神经网络的性能在很大程度上受网络结构的制约,一个拥有大量模糊规则的模糊神经网络,不仅在学习速度和精确度受到冗余信息的影响,还往往与一个“黑箱”神经网络没有多大的区别,严重地妨碍了人们对整个系统更好的了解。所以在模糊神经网络设计时,如何构成良好的网络结构,以便利于神经网络实现模糊输入、模糊推理、网络中的传播和最终结果的理解等等已经成为很多学者关心并加以研究的问题。目前已经提出了很多方法如相似性分析^[1]、专家经验^[5]和遗传算法^[2,3]、训练剪枝法^[12]等。

禁忌搜索算法^[4](Tabu Search, TS)由 Fred Glover 于 1986 年提出,它通过引入一个灵活的存储结构和相应的禁忌准则来避免迂回搜索,并通过藐视准则来放宽一些被禁忌的优良状态,进而保证多样化的有效探索以最终实现全局优化,已经在很多领域得到成功应用。本文提出一种基于 TS 的模糊神经网络的自动优化学习方法(fuzzy neural network based on tabu search, FNNTS)。该方法利用 TS 搜索最优的模糊神经网络结构,并结合最小二乘法和梯度下降方法调整网络参数,大大减少了对专家知识的依赖。实验结果表明,与文^[8,9,12]的方法相比, FNNTS 能使模糊神经网络获得更精练的结构和更高的精确度。

1 模糊神经网络

文中采用基于 Takagi-Sugeno-Kang (TSK)模糊推理模型的模糊神经网络(图 1)。

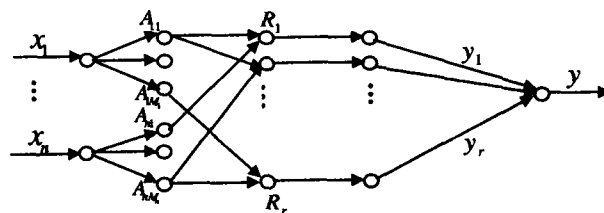


图 1 模糊神经网络结构

TSK 模糊推理系统的模糊规则的形式为:

R_k : if x_1 is A_{1m_1} and x_2 is A_{2m_2} , ..., and x_n is A_{nm_n} then y_k
 $= w_{0k} + w_{1k}x_1 + w_{2k}x_2 + \dots + w_{nk}x_n$

式中, R_k 表示第 k 条规则, x_1, x_2, \dots, x_n 是输入变量, y 是输出变量。 A_{m_i} 是输入变量 x_i 的模糊子集, 规则结论 y_k 则是输入变量的线性函数。

FNN 中各层的功能如下:

第一层是输入层, 该层的节点直接将输入变量值传给下一层。

第二层的每个节点代表一个模糊子集。该层的节点用来

^{*}) 受到教育部科学技术重点项目(104262)资助。肖丽 硕士研究生, 主要研究方向: 神经网络与计算智能。刘光远 博士, 教授, 博导, 主要研究方向: 神经网络、模糊控制、现代优化方法及智能信息处理等。贺一 副教授, 博士研究生, 主要研究方向: 神经网络与计算智能。

计算各输入变量属于各模糊子集的隶属度函数,如高斯型函数:

$$u_{ij}(x_i) = \exp[-(x_i - c_{ij})^2 / \sigma_{ij}^2], i = 1, 2, \dots, n, j = 1, 2, \dots, M_i$$

其中 c_{ij} 和 σ_{ij} 分别是隶属度函数 $u_{ij}(x_i)$ 的中心点和宽度值。

第三层的节点称为规则节点,每个节点代表一条模糊规则,用于计算每条规则的激励强度,

$$f_k = \prod_{i=1}^n u_{ij}(x_i), i = 1, 2, \dots, n, k = 1, 2, \dots, r, j = 1, 2, \dots, M_i$$

其中 r 是规则节点数。

第四层实现的是归一化计算。

$$\Psi_k = f_k / \sum_{k=1}^r f_k, k = 1, 2, \dots, r$$

第五层是输出层,进行清晰化计算。

$$y = \sum_{k=1}^r \psi_k y_k$$

2 TS 用于 FNN 的结构优化

TS 用于优化设计模糊神经网络的结构,包括 FNN 第二层中各变量的模糊子集数和第三层的规则节点数以及两层之间的连接。算法的具体实现如下:

2.1 解的结构

解由两部分组成(图 2),前半部分是 FNN 第二层各输入变量的模糊子集数的十进制序列, M_i 是变量 x_i 的模糊子集数;后半部分是一个二进制序列, $R_{j_1 j_2 \dots j_n}$ 代表一个规则节点, $j_1 j_2 \dots j_n$ 表示该节点与变量 x_i 的第 j_i 个模糊子集节点连接。 $R_{j_1 j_2 \dots j_n}$ 为 1 表示该节点存在,0 表示不存在。序列长度即规则节点的最大个数为 $\prod_{i=1}^n M_i$,因此不同的解可能具有不同的长度。

模糊子集序列				规则节点序列			
M_1	M_2	...	M_n	$R_{11\dots 1}$	$R_{11\dots 2}$...	$R_{M_1 M_2 \dots M_n}$

图 2 解的结构

为了保证每一个解唯一地对应一个网络结构,我们定义:一个合法的解中的每个模糊子集节点至少和一个规则节点相连接,即:对于任意 $j_i (j_i = 1, 2, \dots, M_i)$,存在 $\sum R_{j_1 j_2 \dots j_n} \geq 1$ 。

2.2 邻域

根据解的结构,邻域解的产生分为两步:

(1)为当前解的模糊子集序列的每个分量 M_i 分别产生一个随机数 $\delta(i) (0 < \delta(i) < 1)$,如果 $\delta(i)$ 大于一个给定阈值 P1,则对该分量等概率进行加 1 或减 1 操作。

(2)根据模糊子集序列得到邻域解可能的规则节点以及序列长度。如果邻域解的规则节点 R' 与当前解中规则节点 R 具有相同的下标 $j_1 j_2 \dots j_n$,则令 $R' = R$,最后以概率 P2 对规则节点序列的每个分量进行取反操作。

最后判断解的合法性,如果存在无规则节点连接的模糊子集,则随机地增加一个与其连接的规则节点直到满足合法条件为止。

2.3 目标函数

因为 FNNTS 是为了使 FNN “透明”且易理解,我们选用的目标函数 G 考虑了网络输出误差以及网络复杂性两个方面:

$$G = RESE + \alpha * N$$

其中的 α 是一个常数,用于平衡 RMSE 和 N 的影响度。 N 是 FNN 中的模糊规则数,即网络第三层的规则节点数。RMSE 是网络输出 $O(i)$ 和期望输出 $T(i)$ 的均方根误差,即:

$$RESE = \sqrt{\frac{d}{d} \sum_{i=1}^d (T(i) - O(i))^2 / d}$$

其中 d 为训练样本的总数。

2.4 禁忌目标和禁忌准则

禁忌目标选用目标函数值 G 。但是由于 G 是一个实数,直接对其禁忌很难奏效,因此我们采用了文[10]中的禁忌准则:以禁忌表中的 G 为中心,左右偏移一个很小的范围以形成一个禁忌区间,落入该区间的目标函数值则被禁忌。

2.5 TS 算法基本流程

STEP1:随机产生初始解,令禁忌表为空;

STEP2:按 2.2 中的方法生成当前解的邻域;

STEP3:如果邻域中的最优解满足特赦规则,令其为当前解并转到 STEP5;

STEP4:从邻域中选取没有被禁忌的解组成候选集,选出候选集中的局部最优解作为当前解;

STEP5:更新禁忌表;

STEP6:判断是否满足终止条件,若满足则终止算法,输出最优解,否则转到 Step2。

3 FNN 的参数学习

FNN 的参数学习分成规则前件的隶属函数参数学习和规则结论参数学习两个步骤进行。首先固定隶属函数的中心点和宽度值,采用最小二乘法来学习结论部分的参数;然后固定结论部分,采用基于梯度下降的误差反传算法学习前提参数。

第 i 个训练样本的网络输出 $O(i)$ 可以用下式计算:

$$O(i) = \sum_{k=1}^r \psi_k * y_k = \sum_{k=1}^r \psi_k * (w_{0k} + w_{1k} x_1(i) + w_{2k} x_2(i) + \dots + w_{nk} x_n(i)) = \Psi(i)w$$

其中 $\psi(i) = [\Psi_{11} x(i) \psi_{12} x(i) \dots \Psi_{r1} x(i)]$, ψ_k 是第 i 个样本在归一化层第 k 个节点的输出, $x(i) = [1 \ x_1(i) \dots x_n(i)]$, $w = [w(1) \ w(2) \dots w(r)]^T$, $w(k) = [w_{0k} \ w_{1k} \dots w_{nk}]$, $k = 1, 2, \dots, r$ 。

令 $T = [T(1) \ T(2) \dots T(d)]^T$, $E = [e(1) \ e(2) \dots e(d)]^T$, $\phi = [\psi(1)^T \ \psi(2)^T \dots \psi(d)^T]^T$,

其中 $e(i) = T(i) - O(i)$,则有 $T = \phi w + E$,根据最小二乘法的原理,当 $\phi^T \phi$ 非奇异时,使 E 取极小值的参数 $w = (\phi^T \phi)^{-1} \phi^T T$ 。

选用梯度下降的误差反传算法时,在时刻 t ,隶属函数的中心和宽度值由以下公式更新:

$$c_{ij}(t) = c_{ij}(t-1) - \eta \frac{\partial G}{\partial c_{ij}}, \sigma_{ij}(t) = \sigma_{ij}(t-1) - \eta \frac{\partial G}{\partial \sigma_{ij}}$$

其中 η 为学习速率。

4 仿真实验

为验证本文算法的有效性,选用文[8,9]中的三输入非线性函数进行函数逼近:

$$t = (1 + x^{0.5} + y^{-1} + z^{-1.5})^2$$

在输入空间 $[1, 6] \times [1, 6] \times [1, 6]$ 中随机地产生 216 个训练样本。使用本文所提出的 FNNTS 获得的 FNN 结构为: 3-6-5-5-1,其中输入变量 x, y 和 z 各有 2 个模糊子集,所得到的 5 条模糊规则见表 1。

表 1 由 FNNTS 得到模糊规则

模糊规则数	前提参数			结论参数
	<i>x</i>	<i>y</i>	<i>z</i>	
1	A (3.816, 2.742)*	A (2.781, -4.245)	A (2.510, -4.532)	t=15.64x+29.1y-10.08z+24.16
2	A (3.816, 2.742)	A (1.346, 2.686)	A (1.205, 2.385)	t=4.282x-0.072y-0.076z+8.989
3	A (3.337, 5.522)	A (2.781, -4.245)	A (1.205, 2.385)	t=0.402x+1.377y-0.605z+42.72
4	A (3.337, 5.522)	A (1.346, 2.686)	A (2.510, -4.532)	t=-0.644x-2.041y-2.45z+99.39
5	A (3.337, 5.522)	A (1.346, 2.686)	A (1.205, 2.385)	t=2.29x-0.5825y-0.462z-9.647

* A(σ, c)表示的是宽度值为 σ 和中心点为 c 的高斯函数。

为了方便比较,我们采用与文[8,9],文[12]相同的性能指标 APE(average percentage error)。

$$APE = \frac{1}{d} \sum_{i=1}^d (|T(i) - O(i)| / |T(i)|) \times 100\%$$

另外从相同的输入空间里随机选取 125 个检验样本用来测试 FNN 的泛化性能。与自适应神经-模糊推理系统(Adaptive-network-based fuzzy inference systems, ANFIS)^[8], 正交最小平方(orthogonal least squares, OLS)^[9] 和泛化动态模糊神经网络(generalized dynamic FNN, GD-FNN)^[9], 自组织模糊神经网络(self-organising fuzzy neural network, SOFNN)^[12] 的比较结

果如表 2 所示。表中, APE_{trn} 和 APE_{chk} 分别表示训练样本的 APE 和检验样本的 APE。从表中可以看出, 与其他方法相比, 虽然 FNNTS 的训练样本平均误差大于 ANFIS, 但是值得注意的是由 FNNTS 优化学习得到的 FNN 具有最少的网络参数和最小的检验误差, 由此证明其具有较好的泛化性。

表 2 FNNTS 和其他方法的比较

学习算法	训练样本数	检验样本数	APE _{trn} (%)	APE _{chk} (%)	参数个数
ANFIS	216	125	0.043	1.066	50
OLS	216	125	2.430	2.560($e_m=1.3949$) *	66
GD-FNN	216	125	2.110	1.540($e_m=0.8781$)	64
SOFNN	216	125	1.138	1.1214	60
FNNTS	216	125	0.157	0.108($e_m=0.1966$)	32

* e_m 表示的是网络输出和期望输出的最大误差。

结论 本文详细讨论了基于 TS 的模糊神经网络的自动优化学习方法 FNNTS, 利用 TS 的全局逐步寻优、随机操作和局部搜索能力强等特点能有效地构建 FNN 的结构, 不需要任何专家经验, 尤其适合在系统较为复杂, 专家经验不易总结的情况下对系统进行建模。仿真结果表明, 由 FNNTS 得到的模糊神经网络结构精练且有较高的精确度, 从而验证了本算法的有效性和可行性, 并有望在其他实际问题中得到更广泛的应用。

但是由于禁忌搜索在模糊神经网络领域的研究还属于起步阶段, 虽然训练精度较高, 但其在搜索过程中串行移动的特点使得训练时间较长, 同时梯度算法也有易陷入极小值的不足, 这些问题都还有待进一步的探讨和研究。

参 考 文 献

- Setnes M, Koene A, Babuska R, Bruijn P M. Data-Driven Initialization and Structure Learning in Fuzzy Neural Networks. IEEE World Congress on Computational Intelligence, Anchorage, USA, 1998
- 马铭, 徐岩, 张利彪. 基于模糊加权神经网络的模糊规则自动获取. 计算机应用, 2003, 23(11): 15~17
- Zhou Z J, Mao Z Y. On Designing An Optimal Fuzzy Neural Network Controller Using Genetic Algorithms. In: the 3rd World Congress on Intelligent Control and Automation, Hefei, China, 2000
- Glover F, Laguna M. Tabu Search. Boston, Kluwer Academic Publishers, 1997

- Lin Y H, Cunningham III G A. A New Approach to Fuzzy-neural System Modeling. IEEE Trans. on Fuzzy Syst., 1999, 3(2): 190~198
- Liu Guangyuan, Fang Yonghui, Zheng Xufei, et al. Tuning Neuro-fuzzy Function Approximator by Tabu Search. In: Intl. Symposium on Neural Networks, Dalian, China, 2004
- 贺一, 刘光远, 邱玉辉. Tabu Search 中集中性与多样性的自适应搜索策略. 计算机研究与发展, 2004, 30(5): 162~166
- Jang J-S R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Trans. on Syst., Man and Cybernetics, 1993, 23(3): 665~684
- Wu Shiqian, Member M J Er, et al. A Fast Approach for Automatic Generation of Fuzzy Rules by Generalized Dynamic Fuzzy Neural Networks. IEEE Trans. on Fuzzy Syst., 2001, 9(4): 578~594
- Sexton R S, Alidaee B, Dorsey R E, et al. Global Optimization for Artificial Neural Networks: A Tabu Search Application. European Journal of Operational Research, 1998, 106: 570~584
- Takagi T, Sugeno M. Fuzzy Identification of Systems and Its Applications to Modeling and Control. IEEE Trans. on Syst., Man and Cybernetics, 1985, 15(1): 116~132
- Leng Gang, Thomas Martin McGinnity and Girijesh Prasad; An Approach for On-line Extraction of Fuzzy Rules Using A Self-organising Fuzzy Neural Network. Fuzzy Sets and Syst., 2005, 150(2): 211~243