

结合 SVD_TLS 及 EKF 算法的动态自组织模糊神经网络在动态系统中的应用

李 云

(中国矿业大学计算机科学与技术学院 徐州 221116) (宿迁高等师范学院计算机学院 宿迁 223800)

摘 要 针对如何优化模糊神经网络的规则及如何合理地调整非线性参数及线性参数等问题,提出了将奇异值分解_总体最小二乘法(SVD_TLS)及扩展卡尔曼滤波(EKF)相结合的动态自组织模糊神经网络(STD_DSFNN)。首先给出了 STD_DSFNN 的结构及各层的含义;其次,用 EKF 算法学习非线性参数,SVD_TLS 算法学习线性参数的同时提取重要模糊规则;最后,通过典型的 Machey-Glass 时间序列预测实例验证 SVD_TLS 及 EKF 相结合的动态自组织模糊神经网络(STE_DSFNN),同时与 DFNN、ANFIS 及 UKF_DFNN 相对比,结果表明 STE_DSFNN 网络结构更紧凑,具有更好的泛化能力。

关键词 奇异值分解_总体最小二乘法,扩展卡尔曼滤波,动态自组织模糊神经网络

中图分类号 TP394.1, TH691.9 **文献标识码** A

Dynamic Self-organizing Fuzzy Neural Network Combined with SVD_TLS and EKF Algorithm Used for Dynamic System Processing

LI Yun

(School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)

(School of Computer Science, Suqian Higher Normal School, Suqian 223800, China)

Abstract This paper proposed SVD_TLS and EKF based dynamic self-organizing fuzzy neural network (STD_DSFNN) for optimizing fuzzy rules and Adjusting nonlinear and linear parameters reasonably. Firstly, the structure and meanings of each layer are given. Then nonlinear parameters are learned by using EKF algorithm, the linear parameters are learned by using SVD_TLS algorithm which also extract important rules at the same time. At last, the STE_DSFNN is verified through the typical Machey-Glass time series prediction examples. The results show that the STE_DSFNN network structure is more compact and has better generalization ability compared with the DFNN, ANFIS and UKF_DFNN.

Keywords SVD_TLS algorithm, EKF algorithm, Dynamic self-organizing fuzzy neural network

1 引言

近年来,模糊神经网络的研究得到了极大的发展。模糊神经网络的提出是要获取模糊逻辑系统及神经网络的优点而克服各自的特点。把神经网络应用于模糊系统,可以解决模糊系统中的知识抽取问题;把模糊系统应用于神经网络,其神经网络就不再是黑箱了,人类知识就很容易融合到神经网络中,避免了初值选择的任意性。

目前,模糊神经网络研究的热点是如何生成最优的模糊规则数及模糊规则的自动生成和修剪,以及模糊神经网络的参数调整问题。文献[1]提出了动态模糊神经网络,其结构辨识与参数估计同时进行,采用误差下降率作为修剪策略,得到了较为紧凑的结构,但其假设规则中输入变量的所有高斯隶属函数具有相同宽度,与现实不符;文献[2]提出了广义动态模糊神经网络,其采用椭圆基函数,学习效率和性能都优于动态模糊神经网络,文献[1,2]对线性部分的参数进行了学习,而对非线性部分的参数只是有规律地扩大,因此,结构的紧凑和网络的泛化能力都不是最佳的;文献[3,4]采用扩展卡尔曼

滤波器对结构已确定的模糊神经网络非线性部分的参数进行了学习,网络性能有所提高,但 EKF 存在明显的缺陷:非线性模型的线性化过程容易引入误差,无法保证估计精度,计算高维非线性函数的雅可比矩阵非常烦琐而且容易出错;文献[5]提出基于竞争学习算法用于模糊神经网络的修剪,其没有考虑输入间的关系,而且每个 RBF 单元的贡献不仅由它的权重系数决定,还取决于自身的输出,因而其结果并非最优。

为了改善动态模糊神经网络对参数调整的准确性,文献[6]提出了基于 RLS_UKF 参数的调整策略,RLS 学习线性参数,UKF 学习非线性参数,其经过算法得到的模糊神经网络结构较为紧凑,泛化性能较好,但其计算量较大,比较占用内存容量。本文提出基于 SVD_TLS 与 EKF 相结合的动态自组织模糊神经网络,其中 EKF 用于非线性参数调整,SVD_TLS 在进行线性参数调整的同时还用来提取重要模糊规则,在学习开始之前,没有一条模糊规则,其模糊规则在学习过程中逐渐增长而形成,规则生成后采用 SVD_TLS 算法来实现动态自组织模糊神经网络的规则修剪。所得到的模糊神经网络具有结构小的特点,避免了出现过拟合的现象,且得到的网

络结构更紧凑,泛化能力更强。

2 动态模糊神经网络结构

动态模糊神经网络结构如图 1 所示,其本质意义是代表一个基于 TSK 模型的模糊系统。在图中 x_1, x_2, \dots, x_r 是输入的语言变量, y 是系统的输出, MF_{ij} 是第 i 个输入变量的第 j 个隶属函数, R_j 表示第 j 条模糊规则, N_j 是第 j 个归一化节点, ω_j 是第 j 个规则的结果参数或者连接权, 指系统总的规则数。下面对该网络各层的含义做详细的描述。

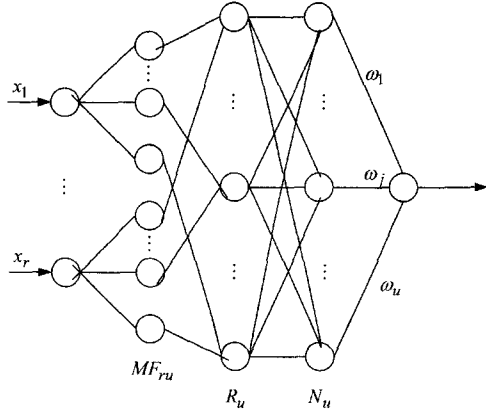


图 1 动态模糊神经网络结构图

第一层为输入层,每个节点分别表示一个输入的语言变量。

第二层为隶属函数层,每个节点分别代表一个隶属函数,该隶属函数用如下的高斯函数表示为:

$$\mu_{ij} = \exp\left[-\frac{(x_i - c_{ij})^2}{\sigma_j^2}\right], i=1, \dots, r, j=1, \dots, u \quad (1)$$

式中, μ_{ij} 是 x_i 的第 j 个隶属函数, c_{ij} 是 x_i 第 j 个高斯隶属函数的中心, σ_j 是 x_i 的第 j 个高斯隶属函数的宽度, r 是输入变量数, u 是隶属函数的数量, 也代表系统总的规则数。

第三层为 T-范数层,每个节点分别代表一个可能的模糊规则中的 IF-部分,因此该层节点数反映了模糊规则数。第 j 个规则 R_j 的输出为:

$$\phi_j = \exp\left[-\frac{\sum_{i=1}^r (x_i - c_{ij})^2}{\sigma_j^2}\right] = \exp\left[-\frac{\|X - C_j\|^2}{\sigma_j^2}\right], \quad j=1, \dots, u \quad (2)$$

式中, $X = (x_1, x_2, \dots, x_r)$, $C_j = (c_{1j}, c_{2j}, \dots, c_{rj})$ 是第 j 个 RBF 单元的中心, 从上式可以看到, 该层的每个节点均代表了一个 RBF 单元。

第四层为归一化层,称这些节点为 N 节点, N 节点数与模糊规则节点数相等,第 j 个节点 N_j 的输出为:

$$\psi_j = \frac{\phi_j}{\sum_{k=1}^u \phi_k} \quad (3)$$

第五层为输出层,该层中的每个节点分别表示一个输出量,该输出是所有输入信号的叠加:

$$y(X) = \sum_{k=1}^u \omega_k \cdot \psi_k \quad (4)$$

式中, y 是变量的输出, ω_k 是规则中的 THEN-部分, 或者是第 k 个规则的连接权。对于 TSK 模型:

$$\omega_k = a_{k0} + a_{k1}x_1 + \dots + a_{kr}x_r, k=1, 2, \dots, u \quad (5)$$

3 参数学习及规则修剪算法

参数学习采用交替学习的策略。首先,通过 SVD_TLS 学习线性参数,然后通过 EKF 学习非线性参数,使动态模糊神经网络更快地收敛,获得好的泛化性能。

3.1 非线性参数学习算法 EKF

EKF 作为一种非线性更新算法,可以用来调节动态模糊神经网络的所有参数。但是,若采用 EKF 算法调节所有非线性及线性参数将涉及大矩阵运算,遇到巨大计算负担以及占用大量的内存。由式(4)可知,输出 y_j 与权值参数 ω_j 及 RBF 神经元的中心参数 C_j , 宽度参数 σ_j 相关,研究表明 RBF 神经元的宽度及中心参数对系统性能影响较大,因此采用 EKF 对其进行更新处理,而采用 SVD_TLS 来更新网络权值,学习流程如图 2 所示。

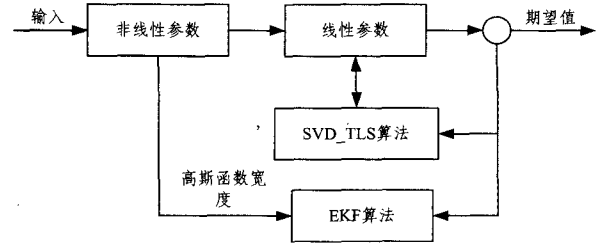


图 2 SVD_TLS 及 EKF 学习流程方框图

高斯宽度的更新是非线性的,可以用 EKF 作如下优化:

$$K_i^* = \frac{S_{i-1} F_i}{1 + F_i^T S_{i-1} F_i} \quad (6)$$

$$S_i = S_{i-1} - K_i^* F_i^T S_{i-1}$$

$$\Sigma_i = \Sigma_{i-1} + K_i^* (T_i^T - \phi_i^T \omega_{i-1})$$

初始调节 $\Sigma_0 > 0$, $S_0 = \rho I$, K_i^* 是第 i 次观测的增益矩阵, S_i 是第 i 次观测误差协方差矩阵, $\Sigma_i = (\sigma_1, \dots, \sigma_u)$ 表示经过 i 次迭代后的高斯宽度向量, $F_i = (\delta_{\sigma_1}, \dots, \delta_{\sigma_u})$ 是第 i 次观测宽度的梯度向量:

$$F_i = \frac{\partial Y_i^T}{\partial \Sigma_{i-1}} \Big|_{\Sigma = \Sigma_{i-1}} = 12 \Sigma_{(i-1)} \|X - C_j\|^2 \psi_j \sum_{k=1}^u (\omega_j - \omega_k) \phi_k \quad (7)$$

3.2 线性参数学习及规则修剪算法 SVD_TLS

假设理想输出为 $D = T^T$, $H = \phi^T = (h_1 \dots h_v)$ 为回归向量, $\theta = W^T$ 包含线性参数,假定 E 是与回归量不相关的一个误差向量,则有线性回归模型为:

$$D = H\theta + E$$

下面解析通过 SVD_TLS 方法从一个给定的规则库中提取最重要的模糊规则的同时,对线性参数进行学习:

步骤 1 计算 $[H \ D]$ 的奇异值分解 SVD, 保存 $\tilde{\Sigma}$ 和 \tilde{S} 。

把 SVD 应用于一个定义为 $[HD]$ 扩展的矩阵, 即 $[HD] = \tilde{\Sigma} \tilde{S}$, 其中 $\tilde{\Sigma} = \text{diag}(\tilde{\sigma}_1, \dots, \tilde{\sigma}_{v+1})$, 把 \tilde{S} 划分为:

$$\tilde{S} = \begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ \tilde{S}_{21} & \tilde{S}_{22} \\ \tilde{S}_{31} & \tilde{S}_{32} \end{bmatrix} \begin{matrix} b \\ v-b \\ 1 \end{matrix} \quad (8)$$

步骤 2 检查 $\tilde{\Sigma} = \text{diag}(\tilde{\sigma}_1, \dots, \tilde{\sigma}_{v+1})$ 的奇异值并确定用来构造模型的模糊规则数(定义为 b), 其中 $r \ll \text{rand}[H \ D]$ 。

步骤 3 根据式(8)划分 \tilde{S} 并保存 $[\tilde{S}_{11} \ \tilde{S}_{12}]$, 对 $[\tilde{S}_{11}$

\tilde{S}_{12}^r 运用QR列主元同时得到交换矩阵 $\tilde{\Pi}$ 。前 b 列中“1”的位置表示了规则库中 b 个最重要的模糊规则。同时也可以选择式(11)中定义的测量指数向量 I_{TLS} 来确定规则库中 r 个最重要的模糊规则位置。

$$\tilde{Q}^T [\tilde{S}_{11} \quad \tilde{S}_{12}^r] \tilde{\Pi} = \tilde{Q}^T [\tilde{S}_{11}^r \quad \tilde{S}_{12}^r] [\tilde{R}_{11} \quad \tilde{R}_{12}] \begin{matrix} b & v-b \end{matrix} \quad (9)$$

上式运用QR列主元计算,其中, \tilde{Q} 是正交矩阵, $\tilde{\Pi}$ 是交换矩阵, \tilde{R} 是上三角矩阵,设:

$$H \tilde{\Pi} = \begin{bmatrix} \tilde{H}^b & \tilde{H}^{(v-b)} \\ b & v-b \end{bmatrix} \quad (10)$$

交换矩阵 $\tilde{\Pi}$ 对矩阵 H 的列排序使得 b 个最重要的列进入式(10)中的 \tilde{H}^b ,也等价于用下式定义的测量指数来确定:

$$I_{TLS} = [I_{TLS_1}, \dots, I_{TLS_v}]^T = [\|\tilde{s}_1\|_2, \dots, \|\tilde{s}_v\|_2]^T \quad (11)$$

步骤4 基于 b 个最重要的模糊规则构造归一化矩阵 \tilde{H}^b 。

b 个最重要的模糊规则的结论部分可以通过解如下的简化方程得到:

$$\tilde{H}^b \tilde{\theta}^b = D \quad (12)$$

步骤5 由式(14)计算SVD并保存 \tilde{S} 的 $(b+1)$ 列,由式(13)计算求解优化的参数 $\tilde{\theta}^b$ 。

其中 \tilde{H}^b 是通过选择 b 个最重要的模糊规则而重新构造的, Van Huffel 和 Vandewalle 在文献[7]中提出 $\tilde{\theta}^b$ 的SVD-TLS解为:

$$\tilde{\theta}^b = \frac{[\hat{S}_{1,b+1}, \dots, \hat{S}_{b,b+1}]^T}{\hat{S}_{b+1,b+1}} \quad (13)$$

式中, $[\hat{S}_{1,b+1}, \dots, \hat{S}_{b,b+1}]$ 为矩阵 \hat{S} 的最后一列元素, \hat{S} 通过对 $[\tilde{H}^b \quad D]$ 做SVD得到,即:

$$[\tilde{H}^b \quad D] = \hat{U} \hat{\Lambda} \hat{S}^T \quad (14)$$

SVD-TLS方法考虑了所有变量的误差,同时通过利用观察向量 D 提供的信息,使选择有关子集的准确性得到改善。

4 仿真实验

我们将通过实例来验证SVD-TLS及EKF学习算法相结合的动态自组织模糊神经网络(STE_DSFNN)的性能,将STE_DSFNN与ANFIS,DFNN及UKF_DFNN做了比较。

本文用STE_DSFNN进行Mackey-Glass混沌时间序列预测,逼近混沌时间序列模型定义如下:

$$x(t+1) = (1-a)x(t) + \frac{bx(t-k)}{1+x^{10}(t-k)} \quad (15)$$

为了能在相同的条件下进行性能比较,选择参数为: $a=0.1, b=0.2, k=17$,则其预测模型可表示如下:

$$x(t+6) = f[x(t), x(t-6), x(t-12), x(t-18)] \quad (16)$$

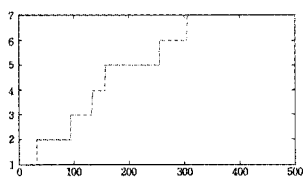


图3 模糊规则的产生

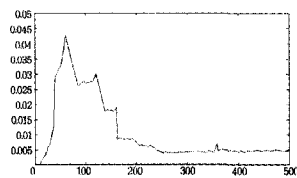


图4 均方根误差(RMSE)

初始条件选为: $x(0)=1.2, x(t)=0, t < 0$,在 $118 \leq t \leq 617$ 之间选择1500对样本,其中前500对作为训练样本,后1000

对作为测试样本来验证被辨识的模型。图3显示了STE_DSFNN生成的规则数;图4、图5显示了STE_DSFNN良好的预测性能。

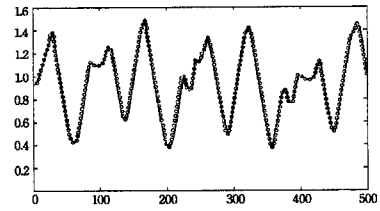


图5 期望和辨识的输出结果(期望值(—)和实际值(·))

为了方便比较,用归一化的均方根误差(RMSE)来评估STE_DSFNN相对于DFNN,ANFIS及UKF_DFNN的泛化能力,表1列出了STE_DSFNN与其它模糊神经网络的比较结果。

表1 STE_DSFNN与其它方案的性能比较

方案	规则数	参数量	训练的RMSE	测试的RMSE
STE_DSFNN	7	89	0.0049	0.0051
DFNN	10	100	0.0082	0.0083
UKF-DFNN	8	98	0.0057	0.0061
ANFIS	16	104	0.0016	0.0015

比较结果显示,STE_DSFNN具有最小的规则数,同UKF_DFNN以及DFNN一样,相对于ANFIS来说,只能获得次优解。但是,STE_DSFNN采用了SVD-TLS规则修剪策略,因此STE_DSFNN的RMSE比其它结构都小得多,说明STE_DSFNN网络的结构更紧凑,预测精度有所提高。

结束语 本文提出了SVD-TLS与EKF相结合的动态自组织模糊神经网络STE_DSFNN,模糊规则修剪与参数学习同时进行。其中,用SVD-TLS和EKF交替学习线性和非线性参数,SVD-TLS在学习线性参数的同时提出重要模糊规则。通过典型Machey-Glass时间序列预测实例,验证了本文得到的STE_DSFNN网络结构更紧凑,泛化性能也有所提高。

参考文献

- [1] Wu S Q, Er M J. Dynamic fuzzy neural networks; a novel approach to function approximation [J]. IEEE Trans Systems, Man and Cybernetics, Part B, 2000, 30(4): 358-364
- [2] Wu S Q, Er M J. A fast approach for automatic generation of fuzzy rules by generalized dynamic fuzzy neural networks [J]. IEEE Trans Fuzzy Systems, 2001, 9(5): 578-594
- [3] Chu K C, Gang F, Jian M. An adaptive fuzzy neural network for MIMO system model approximation in high-dimensional spaces [J]. IEEE Trans System, Man and Cybernetics, Part B, 1998, 28(3): 436-446
- [4] Er M J, Wong W M, Wu S Q. A comparative study of different methods for realizing D-FNN algorithm [C]// Proc of the 38th IEEE Conference on Decision and Control. 1999; 652-661
- [5] Lin C T, Lee C S G. Neural-Network-Based Fuzzy Logic Control and Decision System [J]. IEEE Trans Computers, 1991, 40: 1320-1336
- [6] 李庆良,雷虎民.基于UKF的自组织模糊神经网络训练算法[J].系统工程与电子技术,2010,32(5):1029-1034
- [7] Van Huffel S. Subset Selection Using the Total Least Squares Approach in Collinearity Problems with Errors in the Variables [J]. Linear Algebr. Application, 1987, 88(89): 695-714