

Levenberg-Marquardt 算法在 T-S 型模糊 RBF 神经网络训练中的应用^①

徐奉友 张小刚 (湖南大学 电气与信息工程学院 湖南 长沙 410082)

摘要: 为了提高 T-S 型模糊 RBF 神经网络的训练效率,把 Levenberg-Marquardt 算法引入到 T-S 型模糊 RBF 神经网络的训练过程中,提高了网络训练的收敛速度,减小了训练过程陷入局部极小点的概率,然后基于这种算法推导出 T-S 型模糊 RBF 神经网络的快速训练算法,即混合学习算法。最后通过实验验证了这种算法的有效性和实用性。

关键词: T-S 模糊模型; RBF; Levenberg-Marquardt 算法; 混合学习算法

Application of Levenberg-Marquardt Algorithm to Training of T-S Fuzzy Model Based RBF Neural Network

XU Feng-You, ZHANG Xiao-Gang (Department of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

Abstract: To improve the efficiency of training the T-S fuzzy model based RBF neural network, the Levenberg-Marquardt algorithm is introduced into it, which speeds up the convergence and reduces the probability for the training to get into the local minimum point. Next, a kind of more efficient algorithm, named hybrid learning algorithm, is proposed. At last, the efficiency and practicability of the Levenberg-Marquardt algorithm for the training of the T-S fuzzy model based RBF neural network are tested through an experiment.

Keywords: T-S fuzzy model; RBF; levenberg-marquardt algorithm; hybrid learning algorithm

1 引言

1985 年 Takagi 和 Sugeno 提出了一种分段线性模糊模型,即 Takagi-Sugeno(T-S)模型^[1],因为它能够使用较少的模糊规则为一个复杂的非线性对象建模,所以得到的数学模型比一般模糊模型更为简洁,因而更加有利于数学分析,所以,与一般模糊模型相比,T-S 模型更适合为非线性系统建模。然而正如文献^[2]所言:如同一般模糊模型,T-S 模糊模型同样存在参数辨识困难的问题。神经网络与 T-S 模糊模型相结合,可以取长补短,组成 T-S 型模糊神经网络,使它在具备 T-S 模糊模型的可解释性、灵活性和简洁性

的前提下,可以使用神经网络的自学习能力优化 T-S 模糊模型的参数。T-S 型模糊 RBF 神经网络即是 T-S 型模糊神经网络中的一种,因其训练便捷易行、收敛速度快,而且容易避免局部极小点,所以表现出极大的魅力。虽然如此,找到一种简单且实用的 T-S 型模糊 RBF 神经网络的训练算法并不容易,文献^[3-6]在这方面做了一定的研究,并且提出了各自的学习算法。一般来说,T-S 型模糊神经网络经过粗调和细调两个过程,如文献^[4,5]:粗调过程用来优化网络结构,减小网络规模;细调过程用来优化参数,减小误差,细调过程使用梯度下降法。通常 T-S 模糊神经网络也存在

^① 基金项目:国家自然科学基金(No.60874096)(50704016)

收稿时间:2010-04-14;收到修改稿时间:2010-06-04

聚类 and 辨识两个过程, 如文献[3-4], 通过聚类寻找模糊规则前件的中心位置和展宽, 通过辨识寻找模糊规则后件函数的参数, 因为两个过程分开进行, 所以不利于自适应学习。文献[6]使用最速下降法训练网络, 并在训练过程中实施巧妙的干预措施, 有效地减小了局部极小点的形成, 从而得到最优网络结构和参数。因为 Levenberg-Marquardt 算法比最速下降法有更快的收敛速度, 和更好的寻优能力, 笔者把这种算法引入到 T-S 型模糊 RBF 神经网络的训练过程中, 这有助于加快网络的训练, 同时减小训练过程陷入局部极小点的概率, 然后为 T-S 型模糊 RBF 神经网络推导出它的混合训练算法。最后通过实验验证了 Levenberg-Marquardt 算法在 T-S 型模糊 RBF 神经网络训练中的有效性和实用性。

2 T-S型模糊RBF神经网络简述

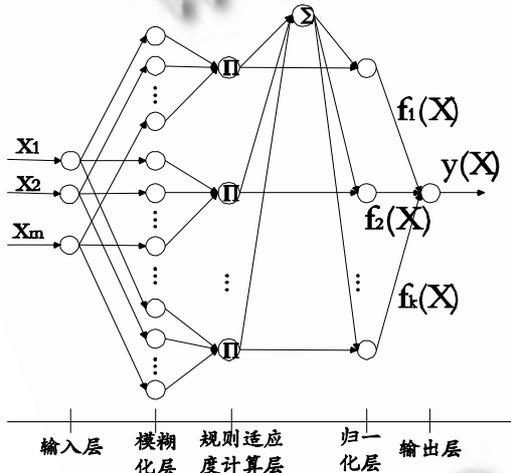


图 1 T-S 型模糊 RBF 神经网络的结构图

如图 1, 网络结构包括 5 层:

第一层, 输入层。该层把输入向量 X 直接传输给模糊化层, 其中 $X = [x_1 \ x_2 \ \dots \ x_m]^T \in R^n$

第二层, 模糊化层。该层基于高斯模型把输入向量的各分量空间模糊化为 k 个模糊子集, 隶属度函数为:

$$j_{ij}(x_j) = \exp\left[-\frac{(x_j - m_{ij})^2}{2s_{ij}^2}\right] \quad (1)$$

式中 $j_{ij}(x_j)$ 代表输入向量 X 中的第 j 个分量 x_j 隶属

于第 i 个模糊子集的隶属度。

第三层, 规则适应度计算层。该层计算输入向量 X 对各模糊规则的适应度, 计算公式如下:

$$z_i(X) = \prod_j j_{ij}(x_j) = \exp\left[-\frac{1}{2} \sum_j \frac{(x_j - m_{ij})^2}{s_{ij}^2}\right] \quad (2)$$

式中 $z_i(X)$ 代表输入向量 X 对第 i 条模糊规则的适应度。上式可以改写成如下公式:

$$z_i(X) = \exp\left[-\frac{1}{2} (X - m_i)^T Y_i (X - m_i)\right] \quad (3)$$

其中, $Y_i = \begin{bmatrix} y_{i(1,1)} & 0 & \mathbf{L} & 0 \\ 0 & y_{i(2,2)} & \mathbf{L} & 0 \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ 0 & 0 & \mathbf{L} & y_{i(m,m)} \end{bmatrix}$,

$$m_i = [m_{i1} \ m_{i2} \ \mathbf{L} \ m_{im}]^T \quad y_{ir} = \frac{1}{s_{ir}^2} \quad (4)$$

第四层, 归一化层。计算公式如下:

$$g_i(X) = \frac{z_i(X)}{\sum_{j=1}^k z_j(X)} \quad (5)$$

式中 $g_i(X)$ 代表输入向量 X 对第 i 条模糊规则的适应度经归一之后的结果。

第五层, 输出层。计算公式如下:

$$y(X) = \sum_{i=1}^k g_i(X) f_i(X) \quad (6)$$

$$f_i(X) = a_{i0} + a_{i1}x_1 + a_{i2}x_2 + \mathbf{L} + a_{im}x_m = A_i^T S \quad (7)$$

$$A_i = [a_{i0} \ a_{i1} \ \mathbf{L} \ a_{im}]^T, \quad S = [1 \ x_1 \ \mathbf{L} \ x_m]^T$$

其中 $f_i(X)$ 既是 T-S 模糊规则的后件表达式, 也是 T-S 型模糊 RBF 神经网络的权值函数。

3 T-S型模糊RBF神经网络的训练

3.1 Levenberg-Marquardt 优化算法在网络训练中的应用

因为 T-S 型模糊 RBF 神经网络可以看成一种前向人工神经网络, 在权值个数只有几百个或者更少的情况下, Levenberg-Marquardt 算法是训练前向神经网络最有效的一种算法之一, 如文献[7]所言, 在多种场合中, 共轭梯度法和可变速率梯度下

降法在不能收敛的情况下, Levenberg-Marquardt 算法可以使网络收敛到一个可以接受的极小值。文章按照文献[7]的思想推导 T-S 型模糊 RBF 神经网络的训练算法:

设 T-S 型模糊 RBF 神经网络的性能指标是:

$$V = \frac{1}{2} \sum_{n=1}^N e_n^2 = \frac{1}{2} \sum_{n=1}^N [(d_n - y_n)]^2 \quad (8)$$

式(8)中, N 为训练样本的个数, d_n 是对应第 n 个样本的期望输出, y_n 是对应于第 n 个样本的实际输出。若 T-S 型模糊 RBF 神经网络中的参数统一用 θ 表示, 则训练 T-S 型模糊 RBF 神经网络时使用的雅可比 (Jacobian) 矩阵为下式:

$$J = \begin{bmatrix} \frac{\partial e_1}{\partial q_1} & \frac{\partial e_1}{\partial q_2} & \mathbf{L} & \frac{\partial e_1}{\partial q_p} \\ \frac{\partial e_2}{\partial q_1} & \frac{\partial e_2}{\partial q_2} & \mathbf{L} & \frac{\partial e_2}{\partial q_p} \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ \frac{\partial e_N}{\partial q_1} & \frac{\partial e_N}{\partial q_2} & \mathbf{L} & \frac{\partial e_N}{\partial q_p} \end{bmatrix} \quad (9)$$

式(9)中 e_n 代表对应于第 n 个样本的误差, θ_p 代表模糊神经网络模型的第 p 个参数。

设 m_{ir} 代表第 i 个模糊规则中对应于第 r 个输入分量的模糊子集中心, x_{nr} 代表第 n 个样本的输入向量中的第 r 个分量, f_{in} 代表第 i 个模糊规则的后件对应于第 n 个样本的输出, g_{in} 代表 i 个模糊规则的前件对应于第 n 个样本的输出经归一化之后的结果。

由式(6), (7), (8)得到

$$\frac{\partial e_n}{\partial a_{ir}} = -g_{i,r} s_r \quad (10)$$

由式(3), (5), (6), (8)得到

$$\frac{\partial e_n}{\partial m_r} = -f_{in} g_{in} (1 - g_{in}) (x_{nr} - u_{ir}) y_{i(r,r)}^2 \quad (11)$$

$$\frac{\partial e_n}{\partial y_{i(r,r)}} = f_{in} g_{in} (1 - g_{in}) (x_{ni} - m_r)^2 y_{i(r,r)} \quad (12)$$

设 $E = [e_1 \ e_2 \ \mathbf{L} \ e_n]^T$, 文献 [7] 中介绍的 Levenberg-Marquardt 权值修正公式可以表示为下式:

$$q(n+1) = q(n) - [J^T(n)J(n) + m(n)I]^{-1} J^T(n)E \quad (13)$$

算法小结:

设 β 为一个大于 1 一个数, 用一个任意值初始化式(3), (7), (8)中的各参数, 并且使式(3)中的参数矩阵 ψ 满足这个条件: 它是一个对角阵并且它的各个元素的为一个足够小的正数。

① 把全部训练样本输入给 T-S 型模糊 RBF 神经网络, 分别计算它们的输出值, 计算公式为(3), (5), (6), (7), 根据式(8)计算 T-S 型模糊 RBF 神经网络的性能指标 V

② 使用式(9), (10), (11), (12), (3), (5), (6), (7)计算雅可比 (Jacobian) 矩阵 J , 然后计算 $\frac{\partial V}{\partial q} = J^T E$,

如果这个结果的范数小于一个设定的值, 或者 V 小于一个设定的值, 则退出循环, T-S 型模糊 RBF 神经网络训练结束。

③ 使用式(13)计算 T-S 型模糊 RBF 神经网络中的新参数 $\theta(n+1)$

④ 使用③步得到新参数 $\theta(n+1)$ 重新计算 T-S 型模糊 RBF 神经网络的性能指标 V , 如果这次得到的结果比①步得到结果小, 然后计算 $m(n+1) = m(n) / b$ 返回步骤①; 否则舍弃这次计算的得到新参数 $\theta(n+1)$, 计算 $m(n+1) = m(n)b$, 返回步骤③

3.2 网络的快速学习

参考文献[8], 在 ANFIS 的训练过程中, 可以使用所谓的快速学习算法, 即混合学习算法, 与之相似, T-S 型模糊 RBF 神经网络的训练过程也可以使用混合学习算法:

初始化网络的各个参数, 初始化内部循环最小误差变化率 ε_i , 最大循环步长 L_i , 例如 L_i 等于 1, 初始化外部循环最小误差变化率 ε_o , 最大循环步长 L_o 。

步骤一, 计算模糊神经网络性能指标式(8)。如果是初次训练, 继续下一步, 否则判断外部循环误差变化率是否小于一个合适的最小值 ε_o 或者达到最大训练步长 L_o , 如果达到这个条件, 则退出循环, 训练结束, 否则进入下一步。

步骤二, 模糊规则的后件参数固定, 辨识前件参数。此时雅可比矩阵可以剔除掉模糊规则后件的相关参数, 使用 Levenberg-Marquardt 优化算法更新前件参数, 直到内部循环误差变化率小于一个合适的最小值 ε_i 或者达到最大训练步长 L_i , 返回步骤二。

步骤三，模糊规则的前件参数固定，辨识后件参数。使用最小二乘法辨识后件参数：

设 $W = [A_1^T \ A_2^T \ \mathbf{L} \ A_k^T]^T$ 为模糊神经网络后件参数向量，另设：
 $U_n = [S_n^T g(1,n) \ S_n^T g(2,n) \ \mathbf{L} \ S_n^T g(k,n)]^T$

$U = [U_1 \ U_2 \ \mathbf{L} \ U_N]^T, D = [d_1 \ d_2 \ \mathbf{L} \ d_N]^T$ 则由最小二乘法得到

$$W = (U^T U)^{-1} U^T D \tag{14}$$

更新模糊神经网络后件参数，返回步骤一。

网络的快速训练算法有两个特点：模糊规则前件参数固定的情况下，辨识得到的后件参数是最优的；模糊规则的后件参数固定的情况下，网络的雅克比矩阵有较大幅度的降低，这两个特点使得网络训练过程非常快。

4 仿真实验

设实验模型为：

$$\begin{aligned} y(k) = & 0.5y(k-1) - 0.39y(k-2) \\ & + 0.75 \cos[y(k-1)z(k-1)] \\ & + 0.95 \sin[x(k-1)] + 0.2 \sin[x(k-2)] \\ & + 0.1 \sin[x(k-3)] + 0.7z(k-3) \end{aligned} \tag{15}$$

使用文章介绍的 T-S 型模糊 RBF 神经网络为它建模，并设置 15 个模糊规则，模糊规则的后件：线性函数表达式如下：

$$\begin{aligned} y(k) = & w_1 y(k-1) + w_2 y(k-2) + w_3 y(k-3) \\ & + w_4 x(k-1) + w_5 x(k-2) + w_6 x(k-3) \\ & + w_7 z(k-3) + w_8 z(k-1) + w_9 z(i-2) \end{aligned} \tag{16}$$

仿真结果如下：

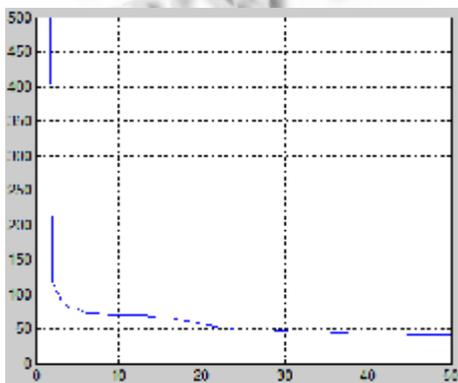


图 2 Levenberg-Marquardt 算法

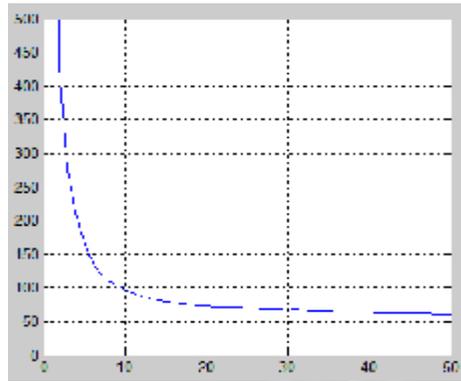


图 3 梯度下降法

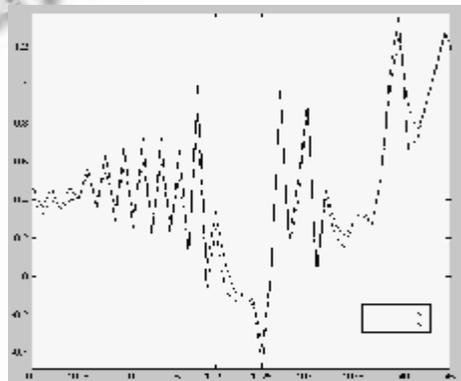


图 4 T-S 型模糊神经网络建模效果图

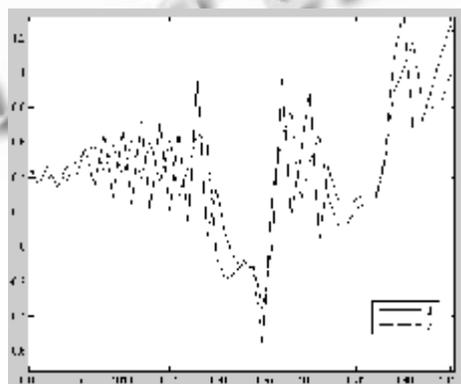


图 5 线性模型仿真效果图

图 2、3 为训练过程，训练模型都为模糊神经网络，其中前一种模型使用 Levenberg-Marquardt 训练算法，后一个模型使用梯度下降法。纵轴代表性能指标，由图 2、3 可以看出用 Levenberg-Marquardt 算法训练

T-S 型模糊 RBF 神经网络, 收敛更快, 训练后的模型精度更高; 如果使用单一线性模型为实验对象式(15)建模, 线性模型的表达式如式(16), 使用最小二乘法辨识它的相关参数, 这种线性模型的工作特性如图 5。图 4、5 中, 虚线 y 是数学模型的输出曲线, 实线 d 是目标值曲线。由图 2、3、4、5 可以看出, 当用 T-S 型模糊 RBF 神经网络为非线性、强耦合对象建模时, Levenberg-Marquardt 优化算法表现出良好的有效性和实用性。

5 小结

文章把 Levenberg-Marquardt 算法引入到 T-S 型模糊 RBF 神经网络的训练过程中, 提高了网络的训练质量: 提高了收敛速度并减小了陷入局部极小点的概率。如果 T-S 型模糊 RBF 神经网络的训练算法使用的是文献[4-5]讲述的训练算法, 在细调阶段引入 Levenberg-Marquardt 算法可以提高细调过程的训练质量。如果 T-S 型模糊 RBF 神经网络的训练算法使用的是文献[6]讲述的训练算法, 引入 Levenberg-Marquardt 算法, 可以大幅度减小网络训练的工作量, 降低网络的训练时间。也可以使用文章讲述的快速学习算法即混合学习算法, 自适应地快速训练 T-S 型模糊 RBF 神经网络, 如果结合文献[6]中讲述的训

练过程干预措施, 可以快速训练出近似最优的 T-S 型模糊 RBF 神经网络。

参考文献

- 1 Takagi T, Sugeno M. Fuzzy identification of system and its application to modeling and control. IEEE Trans Syst, Man Cyber. 1985, SMC-15(1):116 - 132.
- 2 丛爽. 几种模糊神经网络系统关系的对比研究. 信息与控制, 2001, 30(6) 486 - 491.
- 3 鲍鸿, 黄心汉, 李锡雄. T-S 型模糊 RBF 神经网络的结构研究. 华南理工大学学报, 1999, 27(1):11 - 13.
- 4 赵恒平, 俞金寿. 一种基于 T-S 模糊模型的自适应建模方法及其应用. 华东理工大学学报, 2004, 30(4):442 - 445.
- 5 李战明, 王君, 康爱红. 基于 T-S 模糊模型的 RBF 网络的自适应学习算法. 兰州理工大学学报, 2004, 30(2) 82 - 85.
- 6 程玉虎, 王雪松, 孙伟. 自适应 T-S 型模糊径向基函数网络系统仿真学报, 2007, 19(19):4440 - 4444.
- 7 Hagan MT, Menhaj MB. Training Feedforward Networks With the Marquardt Algorithm. IEEE Transactions on Neural Networks, 1994, 5(6):989 - 993.
- 8 Jang JSR. ANFIS: Adaptive-Neuro-based Fuzzy Inference Systems. IEEE Trans. Systems, Man, and Cybernetics, 1993, 23(3):665 - 685.