

基于模糊对向神经网络的非线性动态系统辨识器^{*}

张志华 郑南宁 赵振选

(西安交通大学人工智能与机器人研究所·西安 710049)

摘要: 模糊对向神经网络(FCP)在功能上同模糊逻辑系统的 TS 模型是等价的,它具有神经网络和模糊逻辑系统各自的优点,因而适宜作辨识模型.利用 FCP 网络为辨识模型,设计了一种新的非线性动态系统辨识器.辨识模型的参数调整采用无监督聚类法和误差回传算法相结合的混合学习方法进行的,同时文中还给出选取初始参数的方法.仿真实验表明提出的辨识器是十分有效的.

关键词: 模糊对向网络;系统辨识;学习算法

文献标识码: A

Identification of Nonlinear Dynamic Systems Using Fuzzy CP Network

ZHANG Zhihua, ZHENG Nanning and ZHAO Zhenxuan

(Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University · Xi'an, 710049, P.R. China)

Abstract: The fuzzy counterpropagation(FCP) network is functionally equivalent to the TS model of fuzzy logic systems, and has the characteristics of both neural networks and fuzzy logic systems. Therefore, it's applicable to use FCP network for identification model of nonlinear dynamic systems. Training the parameters of the model is to use a hybrid learning algorithm based on combining an unsupervised clustering method and the error back-propagation algorithm. A new identifier according to the above is developed. The simulation results show the identifier is efficient.

Key words: fuzzy counterpropagation network; system identification; learning algorithm

1 引言(Introduction)

非线性动态系统辨识问题是目前控制领域中的一个难点和焦点.通常,系统辨识问题由建立一个合适参数化的辨识模型和通过最优化未知系统辨识与模型输出之间误差的性能函数来调整模型的参数两部分组成.神经网络的出现为非线性系统辨识提供了一个新的工具.首先,神经网络具有丰富的非线性背景;其次,从逼近论角度来看,神经网络(比如多层感知机)能够逼近任一定义在 \mathbb{R}^n 中的紧子集上的连续函数,这为神经网络解决非线性系统辨识问题提供了物质基础.最后,神经网络具有学习能力,这对于辨识模型的参数调整是具有吸引力的. Narendra 和 Parthasarathy 基于反向传播神经网络提出了两种辨识模型^[1]: 并行模型和串-并行模型.

Wang 等讨论了使用模糊逻辑系统去辨识非线性系统^[2].模糊逻辑系统不仅具有函数的全局逼近能力,而且还有两个神经网络不具有的优点:它的参数具有确定的物理意义;它能够有效的接收人类语言期望的语义信息.模糊逻辑系统是适用于非线性

系统辨识的.

近几年来,模糊逻辑系统同神经网络的融合是一个流行的研究领域.通过结合这两者,一个具有上述两者各自的优点的混合系统被建立.模糊对向(FCP)网络^[3]正是基于这种思想提出的,FCP 网络在功能上等价于模糊逻辑系统的 TS 模型,所以它具有神经网络和模糊逻辑系统两者的优点,由此本文利用它设计出了一个非线性动态系统辨识器.

2 模糊 CP 网络和 TS 模型(Fuzzy CP network and TS-model)

这节我们首先简单介绍 FCP 网络和模糊逻辑系统中的 TS 模型,然后证明这两者之间的功能等价性.模糊 CP 网络是一个三层前向网络.设它的输入层、隐含层和输出层分别有 N_c 和 M 个神经元组成. p 个输入模式 $X = \{x_1, x_2, \dots, x_p\}$.对于输入向量 $x_k = (x_1^k, x_2^k, \dots, x_N^k)$,隐层神经元对应的输出向量为 $u_k = (u_1^k, u_2^k, \dots, u_c^k)$,输出层神经元对应的实际和期望输出向量分别为 $o_k = (o_1^k, o_2^k, \dots, o_M^k)$ 和 y_k

* 基金项目: 国家自然科学基金(69735010)重点资助项目和西安交通大学研究生院博士学位论文基金资助项目.

收稿日期: 1999-01-20; 收修改稿日期: 1999-09-30.

$= (y_1^k, y_2^k, \dots, y_M^k)$, 以上 $k = 1, 2, \dots, p$. 由输入层至隐含层连接权向量记为 $v_j = (v_{j1}, v_{j2}, \dots, v_{jN})$ ($j = 1, \dots, c$); 隐含层至输出层的连接权向量记为 $w_l = (w_{l1}, w_{l2}, \dots, w_{lc})$ ($l = 1, \dots, M$).

模糊 CP 网络是把竞争层的第 j 个神经元的输出函数看作是输入样本 x_k 属于该神经元代表的类的隶属度函数而导出的, 其中, 隶属度函数采用模糊 C-均值方法中的公式:

$$u_j(x_k) = u_j^k = \left(\sum_{r=1}^c \frac{\|x_k - v_j\|^{\frac{2}{m-1}}}{\|x_k - v_r\|^{\frac{2}{m-1}}} \right)^{-1}, \quad (1)$$

如果 $\|x_k - v_j\| = 0$, 则 $u_j^k = 1, u_i^k = 0 (\forall i \neq j)$, 这正是当 $\|x_k - v_j\| \rightarrow 0$ 时函数 u_j^k 的极限. 本文下面都假设 $\|x_k - v_j\| \neq 0 (j = 1, 2, \dots, c), m \in (1, \infty)$ 是模糊度因子.

网络输出层第 l 个神经元的实际输出函数 o_l^k 定义为隐含层神经元输出函数的加权线性组合 (由 (1) 式, 以正则化和非正则化方式计算的结果是相同的, 即都是下式)

$$o_l^k = \sum_{j=1}^c w_{lj} \cdot u_j^k. \quad (2)$$

对于第 k 个输入向量 x_k , 网络的误差函数 E_k 定义为:

$$E_k = \frac{1}{2} \sum_{l=1}^M (o_l^k - y_l^k)^2. \quad (3)$$

为了简便, 本文假设系统是单输出多输入的. 一个模糊逻辑系统有四个基本组成部分: 模糊规则库, 模糊推理机, 模糊产生器和模糊消除器. 模糊规则库由某些“IF-THEN”规则组成, 模糊推理机使用这些“IF-THEN”规则去决定输入域 $U \subset \mathbb{R}^n$ 上的模糊集和输出域 $V \subset \mathbb{R}$ 上的模糊集之间的对应关系. 由 Taka-gi 和 Sugeno^[4]在 1985 年提出的 TS 模型可以表示或建模一般静态或动态的非线性系统. 它的模糊规则具有下面的形式:

$$\begin{aligned} R^i: & \text{IF } x_1 \text{ is } A_1^i \wedge x_2 \text{ is } A_2^i \wedge \dots \wedge x_n \text{ is } A_n^i, \\ & \text{THEN } y^i = a_0^i + a_1^i x_1 + \dots + a_n^i x_n. \\ & (i = 1, 2, \dots, c) \end{aligned} \quad (4)$$

这里, $R^i (i = 1, 2, \dots, c)$ 记第 i 个模糊规则, $x_i (i = 1, 2, \dots, n)$ 和 y^i 分别是模糊规则 R^i 的输入和输出语义变量, c 是模糊“IF-THEN”规则的个数, \wedge 是一个模糊连接算子 (一般是 T 范数算子). 规则推理首先是通过计算前提部分的点火强度来实现的:

$$\omega^i = \mu_R(x) = \mu_{A_1^i}(x_1) \wedge \mu_{A_2^i}(x_2) \wedge \dots \wedge \mu_{A_n^i}(x_n), \quad (5)$$

其中 $\mu_{A_j^i}$ 是模糊集 A_j^i 的隶属度函数. TS 模型的输出通常定义为正则化求和的形式

$$y = f(x) = \frac{\sum_{i=1}^c \omega^i \cdot y^i}{\sum_{i=1}^c \omega^i}, \quad (6)$$

然而, 有时也简化为非正则化的形式:

$$y = f(x) = \sum_{i=1}^c \omega^i \cdot y^i. \quad (7)$$

CP 网络和 TS 模型的知识表示和推理形式^[5]两方面有严格的相似性. 这里我们从理论的角度来讨论 FCP 网络和 TS 模型之间在功能上的等价性.

径向基函数(RBF)网络和 TS 模型之间的功能等价性在参考文献[6]中证明. 比较 FCP 网络和 RBF 网络, 它们之间的主要区别是它们隐含层的输出函数. 前者的输出函数由 (1) 式定义, 而后的输出函数是 Gaussian 函数. 不幸的是 (1) 式不具有 Gaussian 函数的变量可分离性, 因而 FCP 网络和 TS 模型之间的功能等价性不能够直接得到. 但是, 对于第 k 个输入模式, 如果前提部分的点火强度 ω^j 用 $\mu_j^k (j = 1, 2, \dots, c)$ 代替, 即 (5) 式定义为 (1) 式, 则这两个系统之间知识表示的相似性是十分清楚的. 由于 ω^j 定义为点火强度, 其意义事实上就是输入模式属于第 j 个模糊规则 R^j 的隶属度值, 而 (1) 式正代表一个隶属度函数, 所以, 用它来定义 ω^j 是完全合理的.

由上文的分析, FCP 网络的输出层的输出函数的正则化形式和非正则化形式是相同的, 类似于 RBF 网络和 TS 模型的等价定理, 我们同样可以得到下面结论.

定理 1 如果下述三个条件成立, 则 FCP 网络和 TS 模型在功能上是等价的.

1) FCP 网络中竞争层的神经元的个数等于 TS 模型中的模糊“IF-THEN”规则数;

2) 每个模糊“IF-THEN”规则的输出是一个常数 (即, $a_1^j = a_2^j = \dots = a_n^j = 0, a_0^j = \text{常数 } j = 1, 2, \dots, c$);

3) 点火强度 ω^j 由 (1) 式定义.

3 基于 FCP 网络的系统辨识器 (System identification based on FCP network)

FCP 网络具有函数全局逼近能力, 另外, 它在功能上等价于 TS 模型, 因此 FCP 网络同样有模糊逻辑系统具有的两大优点, 使用它去辨识非线性系统是适宜的. 在本文, 我们仅仅考虑下面离散时间的非

线性系统：

$$y(k+1)=f(y(k),\dots,y(k-n+1);u(k),\dots,u(k-m+1)),$$
$$(8)$$

其中 f 是一个要被辨识的未知函数， u 和 y 分别是输入和输出， n 和 m 是正整数。被辨识的对象假设是稳定的，它具有一个已知参数表，但参数的值是未知的。未知系统的辨识问题是构造一合适辨识模型。对于与未知系统相同的输入 $u(k)$ ，它产生一个输出 $\hat{y}(k+1)$ 去逼近 $y_p(k+1)$ 。在本文，我们的目标是使用 FCP 网络去辨识未知函数。

3.1 辨识模型(Identification model)

使用 FCP 网络的非线性动态系统辨识模型同样有并行和串-并行两种模型。

· 并行模型。

$$\hat{y}(k+1)=\hat{f}(\hat{y}(k),\dots,\hat{y}(k-n+1);\hat{u}(k),\dots,\hat{u}(k-m+1))$$

· 串-并行模型。

$$\hat{y}(k+1)=\hat{f}(\hat{y}(k),\dots,\hat{y}(k-n+1);u(k),\dots,u(k-m+1))$$

其中， y 是由方程 (8) 定义的被辨识系统的输出。

3.2 辨识模型的参数学习(Learning for the parameters of identification model)

辨识模型的参数主要是 FCP 网络的权向量 $v_j(j=1,2,\dots,c)$ 和 $w_l(l=1,2,\dots,M)$ 。所以参数学习事实上就是网络权向量的训练，一般地，权向量 w_l 的学习方法采用梯度下降法，而权向量 v_j 具有不同的物理意义：一方面它是网络竞争层输出函数的中心向量，同时它又表示输入层到竞争层的连接，它的调整方法是网络学习算法的关键，根据它的不同意义我们分别用聚类法和梯度下降法来调整它。参数的学习主要包括三个步骤：

步骤 1 确定竞争层神经元的个数 c 和模糊度因子 m 。根据权向量的具体物理意义和模糊逻辑系统接收人类语言信息的能力来选取初始权向量。下面定理是选取初始权向量的一个依据。

定理 2 令 $v_k = x_{i_k}(k=1,2,\dots,c)$ ， $w_{jk} = y_j^{i_k}(j=1,2,\dots,m; k=1,2,\dots,c)$ ，其中 $\{i_1,i_2,\dots,i_c\} \subset \{1,2,\dots,p\}$ ，则下式成立

$$y_j^{i_k} = o_j^{i_k} \quad j=1,2,\dots,m.$$

其中， $y_j^{i_k}, o_j^{i_k}$ 分别是网络相应于输入样本 x_{i_k} 的实际和期望输出的第 j 个分量。

证 因为 $\|x_{i_k} - v_k\| = 0$ ，所以 $u_k^{i_k} = 1, u_j^{i_k} = 0(j=1,2,\dots,c; j \neq k)$ 。于是

$$o_j^{i_k} = \sum_{l=1}^c w_{jl} \cdot u_l^{i_k} = w_{jk} = y_j^{i_k}.$$

定理 2 表明具有上述初始参数的 FCP 网络能

够精确地拟合训练模式集中的任意 c 个输入-输出模式对。在本文实验中，我们把输入-输出样本集中前 c 对输入-输出模式作为权向量 v_j 和 w_l 的初始值，而在下面的学习中从第 $c+1$ 个训练模式开始。

步骤 2 用无监督聚类法和梯度下降法分别粗调权向量 v_j 和 w_l ，算法过程如下：

- 1) 对于 $k = c+1, c+2, \dots, p$ 。
- a) 由 (1) 式计算 u_j^k ，对于 $j = 1, 2, \dots, c$ 。
- b) 根据下式调整 v_j
$$v_{jk}(t+1) = v_{jk}(t) + \lambda(t)(x_{i_k}^k - v_{jk}(t))u_j^k.$$
- c) 由 (2) 式计算输出层第 l 个神经元的输出 o_l^k 。
- d) 按下式调整权向量 w_l ：

$$w_{ij}(t+1) = w_{ij}(t) - \gamma(t) \frac{\partial E_k}{\partial w_{ij}(t)}.$$

2) 计算误差 $E = \frac{1}{2} \sum_{k=1}^p \sum_{l=1}^M (o_l^k - y_l^k)^2$ ，若 $E \leq \epsilon$ 或 $t > T$ ，转入步骤 3，否则 $t = t+1$ ，转入 1)。其中 $\lambda(t), \gamma(t)$ 为学习率，它们的调整方法与自组织特征映射网络相同。

步骤 3 用误差回传 (BP) 学习算法细调 $v_j(j=1, \dots, c)$ 和 $w_l(l=1, 2, \dots, M)$ 。学习规则是基于求误差函数 E_k 最小的梯度下降法导出的：

$$\begin{cases} w_{ij}(t+1) = w_{ij}(t) - \alpha(t) \cdot \frac{\partial E_k}{\partial w_{ij}(t)}, \\ \quad l=1,2,\dots,M; \quad j=1,2,\dots,c, \\ v_j(t+1) = v_j(t) - \beta(t) \cdot \frac{\partial E_k}{\partial v_j(t)}, \quad j=1,2,\dots,c. \end{cases}$$

其中 $\alpha(t), \beta(t)$ 是学习率。

$$\frac{\partial E_k}{\partial w_{ij}} = (o_l^k - y_l^k) \cdot u_j^k,$$
$$\frac{\partial E_k}{\partial v_j} = \sum_{l=1}^M (o_l^k - y_l^k) \cdot w_{jl} \cdot \frac{\partial u_j^k}{\partial v_j},$$
$$\frac{\partial u_j^k}{\partial v_j} = -\frac{2}{m-1} \cdot \|x_k - v_j\|^{-2} \cdot u_j^k \cdot (1-u_j^k) \cdot (x_k - v_j)$$

上述学习算法充分利用了 FCP 网络的特性，首先根据模糊逻辑系统所具有的两大优点来选取初始权向量可以有效保证算法的收敛速度和全局收敛性，而利用无监督聚类算法和梯度下降法分别先粗调权向量 v_j 和 w_l ，然后利用 BP 学习算法细调它们，这样会使模型的推广能力增强，而且又较单纯的 BP 算法收敛得快。我们发现步骤 2 的学习过程正是标准 CP 网络学习算法的一种自然的推广，而 BP 学习算法却是 FCP 网络独有的，这是 FCP 网络较之 CP 网络的一大优点。限于篇幅，FCP 网络的学习算法的

细节另文讨论.

4 仿真实验(Simulation)

这节我们用一个例子来验证上节提出的辨识器的性能,辨识模型采用串-并行模型.我们还用标准 CP 网络和 FCP 网络在文献 [3] 中提出的学习算法去进行仿真.在下面实验中初始学习参数为 $\lambda(0) = 0.6$, $\gamma(0) = 0.6$, $\alpha(0) = 0.6$ 和 $\beta(0) = 0.6$. 在实验中取网络竞争层的节点数 $c = 25$ 和模糊度因子 $m = 2.0$. 标准 CP 网络和文献 [3] 的算法在实验中都采用序列格式.

被辨识的对象由二阶差分方程来描述

$$y_p(k+1) = \hat{f}[y_p(k), y_p(k-1)] + u(k),$$

其中

$$\hat{f}[y_p(k), y_p(k-1)] = \frac{y_p(k) \cdot y_p(k-1) [y_p(k) + 25]}{1 + y_p^2(k) + y_p^2(k-1)},$$

输入是

$$u(k) = \sin(2\pi k/25).$$

串-并行辨识模型由下述方程描述:

$$\hat{y}(k+1) = \hat{f}[\hat{y}(k), \hat{y}(k-1)] + u(k).$$

在仿真中, $\{\hat{y}(k), \hat{y}(k-1), u(k); \hat{y}(k+1)\} (k = 1, \dots, 200)$ 是网络的训练模式. 这里辨识的目标是在利用网络训练前 200 个模式之后,去追踪对象下 400 个输出. 实验结果在表 1 中给出. 上述仿真表明本文提出的辨识模型能够立即跟踪被辨识的系统,在学习精度和辨识泛化能力上较标准 CP 网络和文献 [3] 的算法都有极大的提高,而且比文献 [1] 和 [2] 中辨识效果也有明显的改善.

表 1 三种算法的学习和辨识结果

Table 1 The results of learning and identification of three algorithms

学习算法	迭代次数	总的学习误差 E	辨识的绝对误差和	平均辨识误差
CP 网络的算法	10000	15.883	133.696	0.334
文[3]的算法	5000	0.0113	2.25374	0.0056
本文的算法	$t_1 = 200, t_2 = 500$	0.0005	0.187	0.00047

5 结论(Conclusion)

本文基于 FCP 网络提出了一种非线性动态系统的辨识模型.该模型具有模糊逻辑系统和神经网络的各自优点.而且它的参数采用一种无监督和有监督相结合的混合学习算法进行训练,初始参数可以利用模糊逻辑系统易接收人类语言信息的能力和定理 2 为依据选取,这使得模型一方面收敛速度快,而且具有很强的泛化能力,因而适于用作非线性动态系统辨识.另外由于 FCP 网络中引入了模糊度因子 m ,网络及其学习算法具有十分丰富的性能,通过实验我们发现 m 在 1.1 和 5.0 之间取值,本文的算法性能都很好.对不同的辨识问题我们可以选取不同的 m 值或在学习过程中自适应调整它,这会使得 FCP 网络的辨识能力得到进一步的增强.

参考文献(References)

[1] Narendra K S and Parthasarathy K. Identification and control of dynamical systems using neural networks [J]. IEEE Trans. on Neural Networks, 1990, 1(1): 4-27
[2] Wang L X. Adaptive Fuzzy Systems and Control: Design and Stabili-

ty Analysis [M]. Englewood Cliffs, NJ: Prentice-Hall, 1996
[3] Zhang Zhihua, Zheng Nanning and Wang Tianshu. Fuzzy counterpropagation neural network, universal approximation, and application to time series prediction [A]. International Conference on Neural Network and Brain [C], Beijing, 1998
[4] Takagi T and Sugeno M. Fuzzy identification of systems and its applications to modeling and control [J]. IEEE Trans. on Systems, Man, Cybernetics, 1985, 15(1): 116-132
[5] Nie J and Linkens D A. Fast self-learning multivariable fuzzy controllers constructed from a modified fuzzy controllers constructed from a modified CPN network [J]. Int. J. Control, 1994, 60(3): 369-393
[6] Jang J S and Sun C T. Functional equivalence between radial basis function networks and fuzzy inference systems [J]. IEEE Trans. on Neural Networks, 1993, 4(1): 156-158

本文作者简介

张志华 1969 年生,博士生.主要从事图象处理,模式识别,神经网络及模糊逻辑系统等研究.
郑南宁 1952 年生,博士,中国工程院院士.主要从事图象处理,计算机视觉和模式识别等领域的理论及应用研究.现已在国内外各种重要期刊及会议上发表论文 100 余篇.
赵振选 1975 年生,硕士生.主要从事图象处理,数字视频技术等领域的研究.