

多输出神经元模型的多层前向神经网络及其应用

汪秉文¹, 沈艳军¹, 何统洲²

(1. 华中科技大学 控制科学与工程系, 湖北 武汉 430074; 2. 陨阳师范高等专科学校 计算机系, 湖北 陨阳 442700)

摘要: 研究了一种新的多输出神经元模型. 首先, 给出这类模型的一般形式, 并将该模型应用于多层前向神经网络; 其次, 给出了其学习算法, 即递推最小二乘算法, 最后通过几个模拟实验表明, 采用多输出神经元模型的多层前向神经网络, 具有结构简单, 泛化能力强, 收敛速度快, 收敛精度高等特点, 其性能远远优于激活函数可调模型的多层前向神经网络.

关键词: 神经网络; 神经元模型; 递推最小二乘算法; 多输出神经元模型

中图分类号: TP271 **文献标识码:** A

Multilayer feedforward neural networks with multioutput neural model and its application

WANG Bing-wen¹, SHEN Yan-jun¹, HE Tong-zhou²

(1. Department of control science and engineering, Huazhong University of Science & Technology, Wuhan Hubei 430074, China;

2. Computer Department, Yunyang Teachers' College, Yunyang Hubei 442700, China)

Abstract: A multi-output neural model and its general form are presented. The recurrent least square, a learning algorithm, was used to train multilayer feedforward neural networks (MFNN) with this new model. Several simulations demonstrated that MO (multi-output neural)-MFNN has simple architecture, excellent generalization capacity, fast speed of convergence and improved accuracy. Its performance is superior to TAF (tunable activation function) MFNN.

Key words: neural networks; neural model; recurrent least square (RLS); multi-output neural model

1 引言 (Introduction)

神经元的 M-P 模型半个世纪以来在各种人工神经网络的研究中得到广泛的应用. 这种模型用连接权值和非线性激活函数分别模仿神经元的突触和细胞体的作用. Segee^[1]曾深入研究了一些常用的 BP 网络, 发现采用固定的 S-型激活函数神经元模型的网络来解决某些问题时, 很难获得满意的结果, Lee 和 Kil^[2]等人也曾对采用 S-型、Gauss 型和正弦型激活函数的神经元模型的映射能力分别进行了模拟实验, 结果表明, 在 3 种网络中采用 S-型激活函数的网络, 其映射性能最差. 吴佑寿等^[3]深入讨论了一种激活函数可调的神经元模型 (tunable activation function, TAF) 模型, 采用这种模型的多层前向神经网络的网络容量和性能, 优于采用通常 M-P 模型的网络. 但是其学习算法为 BP 算法. 本文作者^[4]曾将 TAF 网络做了一个适当的变形, 利用递推最小二乘

算法 (RLS) 训练网络, 大大提高网络学习的收敛速度及精度. 在此基础上, 本文将 TAF 模型再进行改进, 得出多输出神经元模型 (MO) 及其多层前向神经网络 (MO-MFNN), 其学习算法为 RLS 算法, 通过几个例子, 结果表明, 利用此神经元模型的网络具有结构简单, 泛化能力强, 学习算法待参数少等特点, 网络收敛速度快、精度高等优点.

2 MO 模型及其前向网络和学习算法 (Multi-output neural model and its feedforward networks and its learning algorithms)

MO-MFNN 示意图, 如图 1 所示, MO 一般形式为: $O = f(g(X, W), \alpha)$, 这里, W 是一个多维权向量, X 是多维输入向量, S 是突触输出信号, 形式同文^[3], 而 $f = [\varphi_1(s), \dots, \varphi_k(s)]^T, i = 1, \dots, k$, $\varphi_i(s)$ 是 f 基函数, k 为神经元的输出维数, 注意 f 维数不大于基函数的维数, 否则, 输出向量一定线性相

关,这里,取 $k = m$,此外,为方便计算, $\varphi_i(s)$ 应该是网络训练的副产品,即

$$\varphi_j(s) = \frac{d^{j-1}}{ds^{j-1}} \varphi_1(s), j = 1, \dots, m.$$

文中主要讨论突触模型具有内积形式,且 $f = [\varphi_1(s), \dots, \varphi_k(s)]^T$ 的神经元模型.具有 MO 模型的 MO-MFNN 的结构与一般 MFNN 的结构相同.这里,如果 k 太大,那么神经元将变得很复杂,这里取 $k = 4$, $\varphi(s) = 1/(1 + e^{-s})$ 作为激活函数.下面给出具有上述神经元模型的多层前向网络图,为方便起见只画了 3 层网络,如图 2 所示.

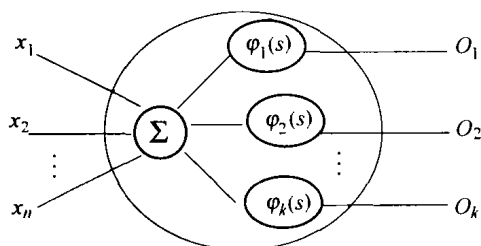


图 1 一般化的 MO 模型

Fig. 1 General MO model

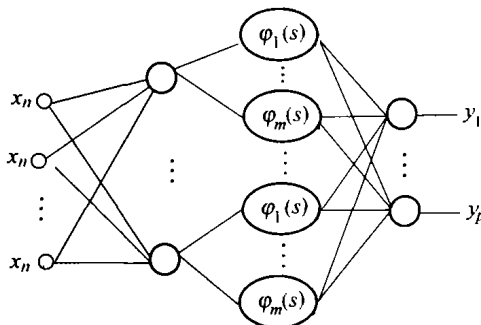


图 2 改进后的 MO 前向网络

Fig. 2 Modified MO-MFNN

定理 1 假设有 N 个训练样本 (x_k, y_k) , $\hat{y}_k$ 是希望输出向量,在输入层、隐含层、输出层节点相同的条件下,改进后的 MO-MFNN 网络,设每个隐含层的激活函数为 $\varphi(s) = [\varphi_1(s), \dots, \varphi_k(s)]^T$, 设它 TAF-MFNN 的激活函数为 $f(s) = \sum_{j=1}^k a_{ij} \varphi_j(s)$, 那么这两种网络具有相同的性能,即,如果两个网络关于式(1)存在最小值,则最小值相同.

$$J_0 = 1/2 \sum_{k=1}^N \sum_{j=1}^P |\hat{y}_k - y_k|^2. \quad (1)$$

证 令改进的网络相应权数值为 $a_{ik}w_{ij}$, $k = 1, 2, 3, 4$, 则改进后网络与原网络等价.

可见改进后的网络比原来网络少 q 个参数,但是其性能却保持不变,网络学习算法为带遗忘因子

的 RLS 算法^[5].

3 模拟实验与仿真结果 (Simulation and results)

3.1 非线性函数的逼近 (Approximation of nonlinear function)

对于非线性映射 Figenbaum 函数,其表达式为

$$x(t+1) = \gamma x(1)(1 - x(t)). \quad (2)$$

其中参数 $\gamma = 4$. 通过训练一个神经网络来逼近此函数,相对均方差用 RMSE 表示,网络结构 $(1, 4, 1)$, 激活函数维数 $k = 2$, 算法参数 $\gamma = 0.9999$, $P_1 = P_2 = 10,000I$, 计算结果见表 1.

表 1 仿真结果

Table 1 Simulation results

RMSE	-41	-50	-60	-71	-81	-95
迭代次数	5	7	12	14	16	46

3.2 一般化的 XOR 问题 (General XOR problem)

本文利用神经网络给出一般化 XOR 问题的判决面,随机产生 300 个样本来训练网络,训练完后,用随机产生的另外 1000 个样本来测试, TAF-MFNN 和 MO-MFNN 网络结构为 $(1, 2, 1)$, 然后,分别用不同的初始值做了 10 次实验,这两种网络都收敛,并得到其中一次实验的判决曲线(图 3、图 4), 表 2 给出了网络测试结果.

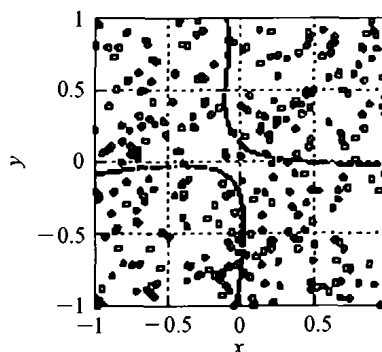


图 3 TAF-MFNN 判决曲线

Fig. 3 Decision curve of TAF-MFNN

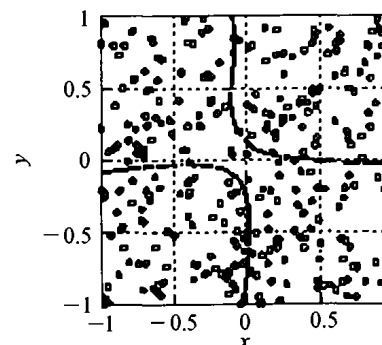


图 4 MO-MFNN 判决曲线

Fig. 4 Decision curve of MO-MFNN

表2 网络测试结果
Table 2 Network test result

网络结构	平均值	最大值	最小值	方差
TAF-MFNN	0.02	0.03	0.009	0.0058
MO-MFNN	0.0016	0.0035	0.0002	0.0002

3.3 非线性动态系统辨识 (Identification of nonlinear dynamic systems)

在控制系统的设计中一个关键步骤是建立系统的数学模型,设动态系统由下列方程给出:

$$y_p(k+1) = f[y_p(k), y_p(k-1), u(k), u(k-1), u(k-2)].$$

未知函数 $f(\cdot)$ 有下述形式

$$f[x_1, x_2, x_3, x_4, x_5] = \frac{x_1 x_2 x_3 x_5 (x_3 - 1) + x_4}{1 + x_2^2 + x_3^2}.$$

本文利用3层MO前向网络来辨识这个非线性系统,网络结构为(5,6,1),在训练过程中,将训练数据分成4个部分,即 $[-1, -0.5]$, $[-0.5, 0]$, $[0, 0.5]$, $[0.5, 1]$.这4个区间各产生50个数据,依次用这4个区间的数据训练网络的学习,在网络训练后对其性能进行测试,测试时输入为

$$u(k) = \sin(2k\pi/250), k \leq 500;$$

$$u(k) = 0.8\sin(2k\pi/250) + 0.2\sin(2k\pi/25), k > 500.$$

训练后网络对于测试信号的跟踪曲线如图5.

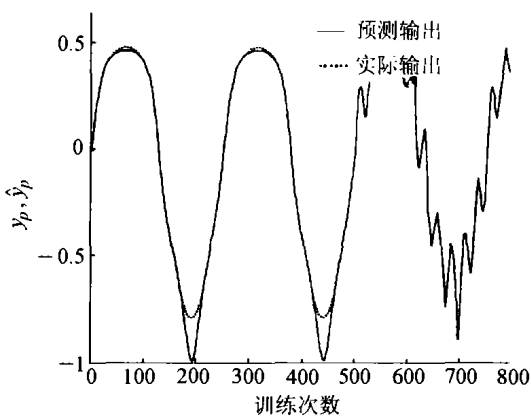


图5 测试曲线图

Fig. 5 Test curve

4 结论 (Conclusion)

本文提出一种新的神经元模型的多层前向神经网络,通过几个不同实验,结果表明,此神经元网络不但保持了TAF-MFNN的优点,而且简化网络结构,参数少,算法参数易于调整.

参考文献 (References):

- [1] SEGEE B E. Using spectral techniques for improved performance in artificial neural networks [C]// *IEEE Int Conf on Neural Networks*. Piscataway: NJ: IEEE Press, 1993, 28(1): 500-505.
- [2] LEE S, KIL R M. A Gaussian potential function network with hierarchically self organizing learning [J]. *Neural Networks*, 1991, 12(4): 207-224.
- [3] 吴佑寿, 赵明生. 激活函数可调的神经元模型及其有监督学习与应用[J]. *中国科学, E辑*, 2001, 31(3): 263-272.
(WU Youshou, ZHAO Mingsheng. An neural model with tunable activation function and its supervised learning algorithms and application [J]. *Science in China, Ser E*, 2001, 31(3): 263-272.)
- [4] 沈艳军, 汪秉文. 激活函数可调的神经网络的一种快速算法[J]. *中国科学, E辑*, 2003, 33(8): 733-740.
(SHEN Yanjun, WANG Bingwen. A fast learning algorithms of neural networks with tunable activation function [J]. *Science in China, Ser E*, 2003, 33(8): 733-740.)
- [5] ROBERT S S, NAZIF T. A fast new algorithm for training feedforward neural networks [J]. *IEEE Trans on Signal Processing*, 1992, 40(1): 203-210.

作者简介:

汪秉文 (1946—), 男, 教授, 博士生导师, 主要研究方向为控制理论与应用、计算机集成系统、人工智能及其应用, E-mail: wangbw@public.hb.wh.cn;

沈艳军 (1970—), 男, 博士研究生, 主要研究方向为人工智能及其应用;

何统洲 (1969—), 男, 讲师, 主要研究方向为计算机网络.