

# 一类乘性有理样条权函数神经网络灵敏度分析

张代远<sup>1,2,3</sup>, 王雷雷<sup>1</sup>

- (1. 南京邮电大学 计算机学院, 江苏 南京 210003;  
2. 江苏省无线传感网高技术研究重点实验室, 江苏 南京 210003;  
3. 南京邮电大学 计算机技术研究所, 江苏 南京 210003)

**摘要:**样条权函数神经网络是一种新型的神经网络,它克服了传统神经网络收敛速度慢、初值敏感、局部极小的问题。因其能精确学习给定的样本,并且结构简单、训练速度快,因此被广泛关注。结合分子三次、分母一次的有理样条函数和样条权函数神经网络的优势,研究了分子三次、分母一次乘性有理样条权函数神经网络,并对其灵敏度进行了理论分析和实验仿真。通过理论分析和仿真可以看出,该神经网络具有分子三次、分母一次的有理样条和样条权函数神经网络的优越特性,在一定扰动范围内,该样条权函数神经网络的灵敏度稳定,具有很强的抗干扰能力。

**关键词:**样条权函数;样条插值;神经网络;灵敏度分析

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2016)10-0050-05

doi:10.3969/j.issn.1673-629X.2016.10.011

## Sensitivity Analysis of Neural Network with Rational Spline Weight Functions Using Multiplicative Neurons

ZHANG Dai-yuan<sup>1,2,3</sup>, WANG Lei-lei<sup>1</sup>

- (1. College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;  
2. Key Laboratory of Jiangsu High Technology Research for Wireless Sensor Networks, Nanjing 210003, China;  
3. Institute of Computer Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** The neural network with spline weight function is a new kind of neural network, which overcomes many problems such as slow convergence speed, sensitive to initial value and local minima. It is widely concerned because of its accurate learning approach to given patterns, simple network topology, fast training speed and so on. Based on the advantage of neural network with spline weight function, the sensitivity of neural network with cubic numerator and linear denominator of rational spline weight functions using multiplicative neurons is discussed, and the accuracy of analytical results is verified by simulation. Both the theoretical analysis and simulation results show that when the disturbance is in a certain range, the sensitivity of this kind of spline weight function neural network is very stable, and is featured with strong noise resistance.

**Key words:** spline weight function; spline interpolation; neural network; sensitivity analysis

## 0 引言

随着云计算、机器学习和大数据计算的蓬勃兴起,人工神经网络理论和方法得到了更加广泛的关注和应用。但是传统的BP算法往往存在局部极小、不收敛或者收敛速度慢、初值敏感等缺陷。这些缺陷极大限制了神经网络在实际工程中的应用。许多学者对传统的BP算法进行了多方面的改进,比如采用遗传算法

调整网络参数<sup>[1]</sup>,优化初始权值或者其他参数<sup>[2]</sup>等。但是这些改进无法从根本上克服梯度下降类算法的缺点。文献[3]提出了新的神经网络(样条权函数神经网络)及其训练算法,该算法能够精确学习给定的样本,克服了传统算法(如BP算法等)存在的局部极小、收敛速度慢、初值敏感等缺点。由于样条插值在曲线和平面设计中的重要作用,许多学者对其进行了大量

收稿日期:2016-01-05

修回日期:2016-05-12

网络出版时间:2016-09-19

基金项目:江苏高校优势学科建设工程资助项目(yx002001)

作者简介:张代远(1957-),男,教授,博士,硕士生导师,研究方向为人工智能、计算机体系结构、计算机应用等;王雷雷(1991-),男,硕士研究生,研究方向为智能计算技术与应用。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160919.0841.038.html>

研究<sup>[4-5]</sup>,发现分子三次、分母一次的有理样条<sup>[6]</sup>能够很好地逼近目标函数,而且通过选取适当参数可以使其获得很好的形状控制能力<sup>[7]</sup>。

当输入样本或者权值存在扰动时,样条权函数神经网络的输出会发生变化。从理论上分析这些变化对于整个网络系统产生的影响是很重要的。灵敏度作为神经网络的一个重要指标,可以分析这些变化给系统带来的影响,而且有不少学者对灵敏度或者神经网络灵敏度做过分析<sup>[8-11]</sup>。文中结合分子三次、分母一次的有理样条函数和样条权函数神经网络的优势,研究分子三次、分母一次乘性有理样条权函数神经网络,利用文献[10]中对灵敏度的定义,对其灵敏度进行了分析。该神经网络具有分子三次、分母一次的有理样条和样条权函数神经网络的优越特性,可以被广泛应用到工程实践中。

## 1 分子三次、分母一次乘性有理样条权函数神经网络

文献[3]提出的样条权函数神经网络,只有两层。为了简单起见,这里引用多输入单输出的乘性样条权函数神经网络结构,如图 1 所示。其中输入样本向量是  $m$  维,因此输入端有  $m$  个节点,每个输入端节点通过连接权前馈连接到所有乘法单元。对于多输入多神经输出网络,可以得到类似的网络结构<sup>[3]</sup>。

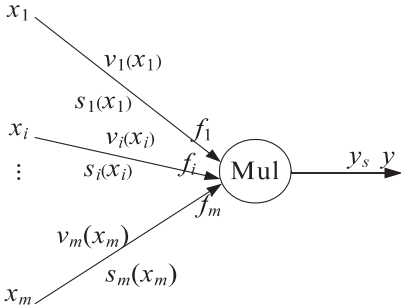


图 1 多输入单输出乘性样条权函数神经网络结构

从图中可以看出,输入向量  $\mathbf{x}$  有  $m$  维,其中  $\mathbf{x}_i, i = 1, 2, \dots, m$  为  $m$  维输入向量的第  $i$  个分量。假设有  $N + 2$  个训练样本,则  $\mathbf{x}_i$  的向量表示为:

$$\mathbf{x}_i = (x_{i0}, x_{i1}, \dots, x_{i(N+1)}) \quad (1)$$

那么对应输出节点的向量表示为:

$$\mathbf{y} = (y_0, y_1, \dots, y_{N+1}) \quad (2)$$

为了求得近似乘性权函数  $s_i(x_i)$  ( $v_i(x_i)$  为理论乘性权函数),除了知道  $\mathbf{x}_i$  的取值之外,还要知道函数  $s_i(x_i)$  在对应样本点函数值  $\mathbf{f}_i = (f_{i0}, f_{i1}, \dots, f_{i(N+1)})$ :

$$\mathbf{f}_i = (f_{i0}^p, f_{i1}^p, \dots, f_{i(N+1)}^p) \quad (3)$$

乘积因子  $\rho_i$  表示每一个输入端节点对相应输出的贡献大小,平方数据说明输入节点在网络训练中对输

出的影响越大。近似权函数  $s_i(x_i)$  由输入向量  $\mathbf{x}_i$  和输出向量  $\mathbf{f}_i$  决定,所以对应的插值点为:

$$\text{Ip}_i = \{(x_{i0}, f_{i0}), \dots, (x_{i(N+1)}, f_{i(N+1)})\} \quad (4)$$

既然已经得到了插值点,那么就可以通过数学上的插值方法精确通过这些插值点(样本点),所以说权函数神经网络是精确学习算法,就是说经过训练的网络能够精确“回想”起已经学过的样本。

插值理论方法丰富多彩,考虑到实际样本可能会在极点附近,或者一些应用中需要考虑形状控制时,有理函数具有一定的优势。因此文中采用分子三次、分母一次有理样条函数作为权函数,这与文献[3]所采用的插值多项式是不同的。

根据文献[6-7],对分子三次、分母一次有理样条函数给出如下定义:

令  $f_p \in R^n, p = 0, 1, \dots, N + 1$  为给定的数据点,而  $x_0 < x_1 < \dots < x_N < x_{N+1}$  为节点的划分。给定插值点  $\{(x_p, f_p)\}, p = 0, 1, \dots, n + 1$ , 且  $f_p$  为被插函数  $f(x)$  在点  $x_p$  处的函数值,即  $f(x_p) = f_p$ , 记  $h_p = x_{p+1} - x_p, \theta = (x - x_p)/h_p$ , 因此分子三次、分母一次的有理样条插值函数在  $[x_p, x_{p+1}]$  上的表达式为:

$$s(x) |_{[x_p, x_{p+1}]} = \frac{P_p(x)}{Q_p(x)} \quad (5)$$

$$P_p(x) = (1 - \theta)^3 \alpha_p f_p + \theta (1 - \theta)^2 V_p + \theta^2 (1 - \theta) W_p + \theta^3 \beta_p f_{p+1} \quad (6)$$

$$Q_p(x) = (1 - \theta) \alpha_p + \theta \beta_p \quad (7)$$

其中

$$V_p = (\alpha_p + \beta_p) f_p + \alpha_p f_{p+1} \quad (8)$$

$$W_p = (\alpha_p + 2\beta_p) f_{p+1} - h_p \beta_p \Delta_{p+1} \quad (9)$$

$\Delta_p = (f_{p+1} - f_p)/h_p, \alpha_p, \beta_p > 0$ , 为形状控制参数。

## 2 分子三次、分母一次乘性有理样条权函数神经网络灵敏度分析

### 2.1 灵敏度的统计学定义

对图 1 所示的神经网络,统计灵敏度定义<sup>[10]</sup>为:

$$S_p = \lim_{\sigma_p \rightarrow 0} \frac{[\text{var}[\Delta \mathbf{Y}_p]]^{1/2}}{\sigma_p} \quad (10)$$

其中,  $\Delta \mathbf{Y}_p$  是输入层第  $p$  个样本的扰动引起的输出层误差;  $\text{var}[\Delta \mathbf{Y}_p]$  为  $\Delta \mathbf{Y}_p$  的方差;  $\sigma_p$  为输入层第  $p$  个样本扰动的标准差。

对于多个输入样本,假设第  $p$  个输入样本的灵敏度为  $S_p, p = 0, 1, \dots, N + 1$ , 通常,可以采用以下两种方式来评价神经网络的统计灵敏度。

采用所有样本中的最大灵敏度:

$$S = \max\{S_p, p \in [0, N + 1]\} \quad (11)$$

采用所有样本的平均灵敏度:

$$S = \frac{1}{N+2} \sum_{p=0}^{N+1} S_p \quad (12)$$

## 2.2 模型逼近误差和逼近噪声误差分析

假设输入层有  $m$  个节点,输出层有  $n$  个节点,如果第  $p$  个输入样本存在噪声扰动,假设输入扰动向量为  $\Delta \mathbf{x}_p = (\Delta x_{1p}, \Delta x_{2p}, \dots, \Delta x_{mp})^T$ ,则由第  $p$  个输入样本扰动而造成的网络输出值的理论噪声误差  $\Delta y_p$  可有如下表示形式:

$$\Delta y_p = (\Delta y_{1p}, \Delta y_{2p}, \dots, \Delta y_{np})^T \quad (13)$$

根据上述的神经网络理论噪声误差输出值的运算方法,对网络输出值的理论噪声误差进行具体分析。

先考虑输出层第  $j$  个神经元理论噪声输出误差。同样地,假设第  $p$  个不含噪声的标准输入样本为  $\mathbf{x}_p^* = (x_{1p}^*, x_{2p}^*, \dots, x_{mp}^*)^T$  ( $p = 0, 1, \dots, N+1$ ),在这种输入情况下,第  $j$  个神经元的理论输出值为:

$$y_{jp}^* = \sum_{i=1}^m v_{ji}(x_{ip}^*) \quad (j = 1, 2, \dots, n) \quad (14)$$

假设第  $p$  个样本的噪声扰动向量为  $\Delta \mathbf{x}_p = (\Delta x_{1p}, \Delta x_{2p}, \dots, \Delta x_{mp})^T$ ,那么对应的噪声样本可以表示为  $\mathbf{x}_p = (x_{1p}^* + \Delta x_{1p}, x_{2p}^* + \Delta x_{2p}, \dots, x_{mp}^* + \Delta x_{mp})^T$ ,那么理论噪声引起的输出值就是:

$$y_{jp} = \sum_{i=1}^m v_{ji}(x_{ip}) = \sum_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) \quad (j = 1, 2, \dots, n) \quad (15)$$

则第  $j$  个神经元的理论噪声输出误差为:

$$\begin{aligned} \Delta y_{jp} &= y_{jp} - y_{jp}^* = \sum_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m v_{ji}(x_{ip}^*) = \\ &= \sum_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) + \\ &+ \sum_{i=1}^m s_{ji}(x_{ip}^*) - \sum_{i=1}^m v_{ji}(x_{ip}^*) + \\ &= \sum_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_{ji}(x_{ip}^*) = \\ &= \sum_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) + \\ &+ \sum_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_{ji}(x_{ip}^*) \end{aligned} \quad (16)$$

由于分子三次、分母一次乘性有理样条权函数仍然采用样条权函数来逼近,因此经过样本的差值点  $x_{ip}^*$ ,其样本代价函数等于 0,即  $\sum_{i=1}^m s_{ji}(x_{ip}^*) -$

$\sum_{i=1}^m v_{ji}(x_{ip}^*) = 0$ 。由式(16)可知,分子三次、分母一次乘性有理样条权函数神经网络灵敏度需要分析模型逼近误差  $\sum_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip})$  和逼近噪声误差  $\sum_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_{ji}(x_{ip}^*)$ 。

先求逼近噪声误差,首先对多输入单输出分子三

次、分母一次乘性有理样条权函数逼近噪声误差进行分析,为了简单起见,这里先去掉下标  $j$ ,即  $\sum_{i=1}^m s_i(x_{ip}^* +$

$\Delta x_{ip}) - \sum_{i=1}^m s_i(x_{ip}^*)$ 。易得  $\sum_{i=1}^m s_i(x_{ip}^*) = \sum_{i=1}^m v_i(x_{ip}^*) = \sum_{i=1}^m f_{ip}$ 。对于  $\sum_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip})$ ,由式(5)可知,  $\theta = (x - x_i)/h_i, x_i \in [x_{ip}^*, x_{i(p+1)}^*]$ ,在区间  $[x_{ip}^*, x_{i(p+1)}^*]$  上,  $h_{ip} = h_i = x_{i(p+1)}^* - x_{ip}^*, \theta = (x - x_{ip}^*)/h_{ip}$ ,因此当  $\Delta x_{ip} \rightarrow 0^+, x = x_{ip}^* + \Delta x_{ip}$  时,  $\theta = (x - x_{ip}^*)/h_{ip} = \Delta x_{ip}/h_{ip}$ ,代入  $\sum_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip})$  可得:

$$\sum_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip}) = \sum_{i=1}^m \left[ \frac{[A_{ip} \delta_{ip}^3 + B_{ip} \delta_{ip}^2 + C_{ip} \delta_{ip} + D_{ip}]}{E_{ip} \delta_{ip} + F_{ip}} \right] \quad (17)$$

其中,  $A_{ip} = (\beta_{ip} f_{i(p+1)} - \alpha_{ip} f_{ip} + V_{ip} - W_{ip}), B_{ip} = (3\alpha_{ip} f_{ip} - 2V_{ip} + W_{ip}), C_{ip} = (V_{ip} - 3\alpha_{ip} f_{ip}), D_{ip} = a_{ip} f_{ip}, \delta_{ip} = \Delta x_{ip}/h_{ip}, E_{ip} = (\beta_{ip} - \alpha_{ip}), F_{ip} = a_{ip}$ 。

因为  $\delta_{ip} \rightarrow 0$ ,所以

$$\begin{aligned} \sum_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip}) &= \sum_{i=1}^m \frac{A_{ip} \delta_{ip}^3 + B_{ip} \delta_{ip}^2 + C_{ip} \delta_{ip} + D_{ip}}{E_{ip} \delta_{ip} + F_{ip}} \approx \\ &= \sum_{i=1}^m \frac{C_{ip} \delta_{ip} + D_{ip}}{E_{ip} \delta_{ip} + F_{ip}} \end{aligned} \quad (18)$$

因此可得:

$$\begin{aligned} \sum_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_i(x_{ip}^*) &\approx \sum_{i=1}^m \frac{C_{ip} \delta_{ip} + D_{ip}}{E_{ip} \delta_{ip} + F_{ip}} - \sum_{i=1}^m f_{ip} = \\ &= \sum_{i=1}^m \frac{C_{ip} \delta_{ip} + D_{ip}}{E_{ip} \delta_{ip} + F_{ip}} - \sum_{i=1}^m f_{ip} = \\ &= \sum_{i=1}^m \left( f_{ip} \frac{(C_{ip} \delta_{ip} + D_{ip})/f_{ip}}{E_{ip} \delta_{ip} + F_{ip}} \right) - \sum_{i=1}^m f_{ip} = \\ &= \sum_{i=1}^m f_{ip} \left( \prod_{i=1}^m \left( \frac{(C_{ip} \delta_{ip} + D_{ip})/f_{ip}}{E_{ip} \delta_{ip} + F_{ip}} \right) - 1 \right) \end{aligned} \quad (19)$$

将  $D_{ip}, F_{ip}, \delta_{ip}$  代入式(19)可得:

$$\begin{aligned} \sum_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip}) - \sum_{i=1}^m s_i(x_{ip}^*) &= \sum_{i=1}^m f_{ip} \left( \prod_{i=1}^m \frac{(C_{ip} \delta_{ip}/f_{ip} + a_{ip})}{E_{ip} \delta_{ip} + a_{ip}} - 1 \right) = \\ &= \sum_{i=1}^m f_{ip} \left( \prod_{i=1}^m \left( \frac{C_{ip} \delta_{ip}/f_{ip} - E_{ip} \delta_{ip}}{E_{ip} \delta_{ip} + a_{ip}} + 1 \right) - 1 \right) \approx \\ &= \sum_{i=1}^m f_{ip} \left( \prod_{i=1}^m \left( \frac{C_{ip}/f_{ip} h_{ip} - E_{ip}/h_{ip} \Delta x_{ip}}{a_{ip}} + 1 \right) - 1 \right) \end{aligned} \quad (20)$$

令  $Q_{ip} = (C_{ip}/f_{ip}h_{ip} - E_{ip}/h_{ip})/a_{ip}$ , 且  $Q_{ip}\Delta x_{ip} \rightarrow 0$ , 因此可将式(20)化为:

$$\begin{aligned} \prod_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip}) - \prod_{i=1}^m s_i(x_{ip}^*) &= \\ \prod_{i=1}^m f_{ip} \left( \prod_{i=1}^m (Q_{ip}\Delta x_{ip} + 1) - 1 \right) &\approx \\ \prod_{i=1}^m f_{ip} \left( \sum_i Q_{ip}\Delta x_{ip} \right) \end{aligned} \quad (21)$$

显然,只要考虑区间  $[x_{p-1}, x_p]$ , 就可以求得  $\Delta x_{ip} \rightarrow 0^-$  时的情况。

而对于多输入多输出乘性样条权函数逼近噪声误差的表达式,可以由式(21)增加反映输出节点序号的加下标  $j$  表示。

对于模型逼近误差  $\prod_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) - \prod_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip})$ , 可以根据文献[12]进行分析,得到分子三次、分母一次乘性有理样条权函数神经网络的模型误差为:

$$\begin{aligned} \prod_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) - \prod_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) &\leq \\ M \parallel \prod_{i=1, i \neq N}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) \parallel \parallel v'_{ji}(x_{ip}^* + \Delta x_{ip}) \parallel &c_N h \end{aligned} \quad (22)$$

因此可得:

$$\begin{aligned} \Delta y_{jp} = y_{jp} - y_{jp}^* &= \prod_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) - \prod_{i=1}^m s_{ji}(x_{ip}^*) + \\ \prod_{i=1}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) - \prod_{i=1}^m s_{ji}(x_{ip}^* + \Delta x_{ip}) &\approx \\ \prod_{i=1}^m f_{jip} \left( \sum_i Q_{jip}\Delta x_{ip} \right) + M \parallel \prod_{i=1, i \neq N}^m v_{ji}(x_{ip}^* + \Delta x_{ip}) \parallel &\parallel v'_{ji}(x_{ip}^* + \Delta x_{ip}) \parallel c_N h \end{aligned} \quad (23)$$

显然,根据式(13)可知,多输入多输出的输出层误差可以表示成如下的矩阵形式:

$$\Delta Y = (\Delta y_0, \Delta y_1, \cdots, \Delta y_p, \cdots, \Delta y_N, \Delta y_{N+1}) \quad (24)$$

此时可以将式(11)或式(12)中增加反映式(24)中行、列的下标即可。这样可以得到多维输出神经网络灵敏度,因此根据式(10)、(11)、(12)、(24)就可以计算出分子三次、分母一次乘性有理样条权函数神经网络的统计灵敏度。

3 仿真实验

下面进行数值仿真实验,验证分子三次、分母一次乘性有理样条权函数的灵敏度稳定性。

实验 1:构建 2-1 的分子三次、分母一次的乘性有理样条权函数神经网络的拓扑结构、训练样本和权函数。其中目标函数为  $e^{(-0.5x_1-0.5x_2)}$ , 学习曲线为  $x_1 = t, x_2 = 2t, t \in [1, 2]$  时,平均相对误差  $\beta_p = 100, \beta_{ip} = 2$ 。取 100 个训练样本

点,样本扰动数组为  $[10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^0]$ 。对每一个  $x_{ip}$ , 在  $[-\tau, \tau]$  中插入 100 个均匀随机数,其中  $\tau = 10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^0$ 。然后根据式(25)计算相对误差,再求其平均值。

$$R = (Y^* - Y)/Y \quad (25)$$

其中,  $Y^*$  表示没有扰动时的输出;  $Y$  表示有扰动时的目标输出。

灵敏度的实际仿真值采用数理统计<sup>[13]</sup>的方法求得,对不同组的扰动利用式(21)计算网络输出层误差,然后再统计出扰动的标准差和网络输出层误差的方差,按照式(10)计算灵敏度仿真值。平均相对误差曲线和灵敏度曲线分别如图 2 和图 3 所示。

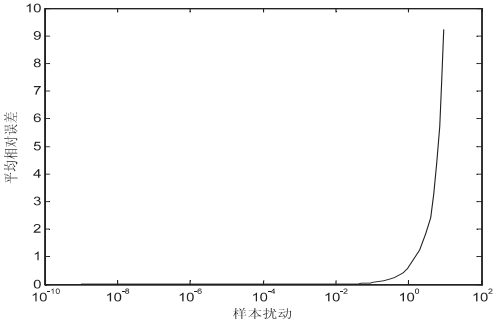


图 2 平均相对误差曲线(学习曲线为  $x_1 = t, x_2 = 2t$ )

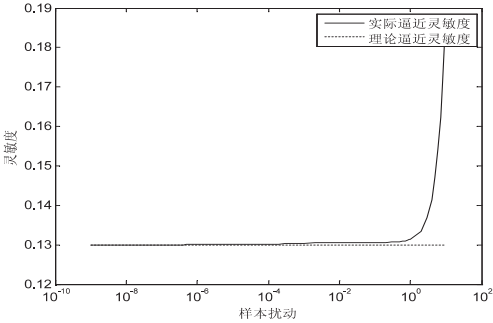


图 3 灵敏度曲线(学习曲线为  $x_1 = t, x_2 = 2t$ )

实验 2:采用实验 1 中构建的分子三次、分母一次乘性有理样条权函数神经网络的拓扑结构、训练样本、目标函数以及扰动数组,学习曲线为  $x_1 = t^2, x_2 = \sin 0.5t, t \in [1, 2]$  时,平均相对误差曲线和灵敏度曲线分别如图 4 和图 5 所示。

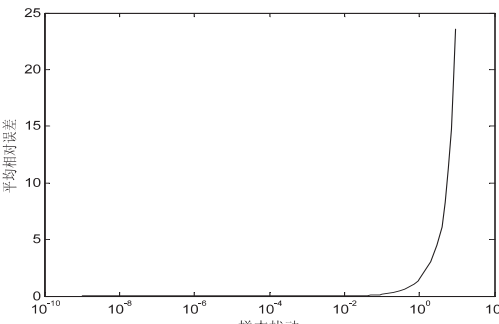


图 4 平均相对误差曲线(学习曲线为  $x_1 = t^2, x_2 = \sin 0.5t$ )



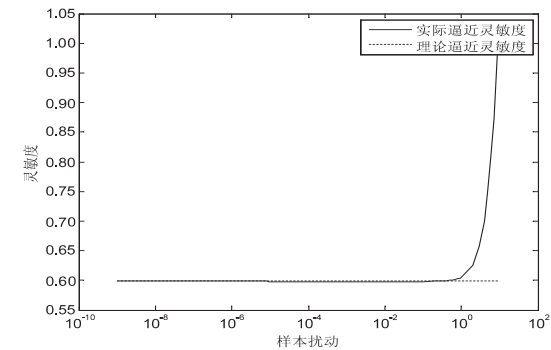


图 5 灵敏度曲线(学习曲线为  $x_1 = t^2, x_2 = \sin 0.5t$ )

从实验可以看出,当输入维数为 2,扰动在  $10^{-9} \rightarrow$

$10^{-3}$  时,  $Z^* = \prod_{i=1}^m f_{ip}(\sum_i Q_{ip} \Delta x_{ip})$  可很好地逼近  $Z =$

$\prod_{i=1}^m s_i(x_{ip}^* + \Delta x_{ip}) - \prod_{i=1}^m s_i(x_{ip}^*)$ , 分子三次、分母一次的乘性有理样条权函数神经网络灵敏度非常稳定。

实验 3:构建  $m-1$  的分子三次、分母一次的乘性有理样条权函数神经网络的拓扑结构,其中输入维数  $m$  分别为 2,5,50,100,200,400,500,600,800,1 000。取 100 个训练样本点,其目标函数为  $0.1e^{(0.5x_1+0.5x_2+\cdots+0.5x_m)}$ ,学习曲线组为  $[t, 2t, t^2, t^3, \cos t, \sin 2t, (\sin 2t)^2, (\cos 2t)^3, e^t, 2^t]$ ,  $t \in [1, 2]$ 。当  $m$  大于 10 时,  $x_{11} \cdots x_m$  的学习曲线从学习曲线组中依次重复选取。对每一个  $x_{ip}$ ,在  $[-10^{-3}, 10^{-3}]$  之间插入 100 个均匀分布的随机数,计算平均相对误差,计算相对误差的公式为  $(Z - Z^*)/Z$ 。

输入维数对平均相对误差的影响如表 1 所示。

	2	5	50	100	200
平均相对误差	0.001	0.002	0.006	0.008	0.011
	400	500	600	800	1 000
平均相对误差	0.019	0.023	0.027	0.035	0.044

从表 1 可以看出,当输入维数  $m \leq 1\,000$  时,平均相对误差非常小,即  $Z^*$  非常接近  $Z$ 。而根据 UC Irvine Machine Learning Repository 的数据,92% 的数据集输入属性个数小于 1 000,因此式(21)具有实用性和代表性。

4 结束语

文中将样条权函数神经网络和分子三次、分母一次的有理样条相结合,研究分子三次、分母一次乘性有理样条权函数神经网络灵敏度。通过理论分析和仿真实验,表明当扰动在较宽的变化范围内,灵敏度值非常稳定。在此扰动范围内神经网络的仿真输出与目标样本的理论输出偏差很小,能与目标样本较好匹配。说明当样本扰动在较宽范围内时,分子三次、分母一次乘

性有理样条权函数神经网络的网络泛化能力<sup>[14]</sup>不会受到外界噪声的影响,即具有较强的抗干扰能力,可以广泛应用于工程实践的各个领域。

参考文献:

[1] Khashman A. A modified backpropagation learning algorithm with added emotional coefficients [J]. IEEE Transactions on Neural Networks, 2008, 19(11): 1896-1909.

[2] Wedge D, Ingram D, McLean D, et al. On global-local artificial neural networks for function approximation [J]. IEEE Transactions on Neural Networks, 2006, 17(4): 942-952.

[3] 张代远. 神经网络新理论与方法[M]. 北京:清华大学出版社, 2006.

[4] Sun Q, Bao F, Duan Q. Shape-preserving weighted rational cubic interpolation [J]. Journal of Computational Information Systems, 2012, 8(18): 7721-7728.

[5] Bao F, Sun Q, Duan Q. Point control of the interpolating curve with a rational cubic spline [J]. Journal of Visual Communication & Image Representation, 2009, 20(4): 275-280.

[6] Duan Q, Djidjeli K, Price W G, et al. A rational cubic spline based on function values [J]. Computers & Graphics, 1998, 22(4): 479-486.

[7] Duan Q, Djidjeli K, Price W G, et al. Weighted rational cubic spline interpolation and its application [J]. Journal of Computational & Applied Mathematics, 2000, 117(2): 121-135.

[8] 张代远. 零代价函数条件下的前馈神经网络统计灵敏度分析 [J]. 模式识别与人工智能, 2004, 17(4): 510-515.

[9] Zhang D. A new algorithm of training neural networks by orthogonal weight functions and sensitivity analysis [C]//Proc of 2012 IEEE 5th international conference on advanced computational intelligence. [s. l.]: IEEE, 2012: 324-327.

[10] Choi J Y, Choi C H. Sensitivity analysis of multilayer perceptron with differentiable activation functions [J]. IEEE Transactions on Neural Networks, 1992, 3(1): 101-107.

[11] Li C G, Li H F, Zhang Y F, et al. Sensitivity analysis of multilayer perceptron based on elastic function [C]//Proc of international conference on machine learning and cybernetics. [s. l.]: IEEE, 2010: 1490-1493.

[12] Zhang D Y, Wang L L. Error analysis of neural network with multiply neurons using rational spline weight functions [J]. Applied Mechanics and Materials, 2015, 713-715: 1716-1720.

[13] Ng W W Y, Yeung D S, Ran Q, et al. Statistical output sensitivity to input and weight perturbations of radial basis function neural networks [C]//Proc of IEEE international conference on systems, man and cybernetics. [s. l.]: IEEE, 2002: 503-508.

[14] Zhang D Y. Analysis of generalization for a new kind of neural network [J]. Advanced Materials Research, 2013, 760-762: 2023-2027.