

符号函数激励的 WASD 神经网络与 XOR 应用*

张雨浓¹, 王 茹¹, 劳稳超¹, 邓健豪²

(1. 中山大学信息科学与技术学院, 广东 广州 510006;

2. 中山大学软件学院, 广东 广州 510006)

摘 要: 基于权值与结构确定 (WASD) 算法, 提出和构建了一种以非连续符号函数为隐层神经元激励函数的 WASD 神经网络模型。通过 WASD 算法, 能有效地确定所构建网络的权值及网络的最优结构。该文也将此网络模型应用于 XOR (异或) 上, 并详细讨论了在带噪类型不同时网络在此应用上的性能。计算机数值实验结果验证了所提出的权值与结构确定法能够有效地确定出网络的最优权值与结构, 所构建的 WASD 网络在 XOR 应用上具有优秀的抗噪性能。另外, 通过对比符号函数激励的 WASD 神经网络与幂函数激励的 WASD 神经网络在高维 XOR 应用方面的性能差异, 证实了所提出的符号函数激励的 WASD 神经网络及算法在解决非线性问题时的优越性。

关键词: 权值与结构确定 (WASD) 算法; 非连续符号函数; 神经网络; XOR (异或); 噪声; 高维

中图分类号: TP183 **文献标志码:** A **文章编号:** 0529-6579 (2014) 01-0001-08

Signum-Function-Activated WASD Neuronet and Its XOR Application

ZHANG Yunong¹, WANG Ru¹, LAO Wenchao¹, DENG Jianhao²

(1. School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, China;

2. School of Software, Sun Yat-sen University, Guangzhou 510006, China)

Abstract: A discontinuous signum-function-activated (SFA) weights-and-structure-determination (WASD) neuronet model is presented and constructed based on the WASD algorithm. By this algorithm, the optimal weights and structure can be determined effectively. We apply the SFA-WASD neuronet model to XOR (i. e., exclusive or), and detail its performance in the XOR application with various types of disturbance noise considered. Numerical verification results substantiate the validity of the WASD algorithm in determining the optimal weights and structure, as well as the good anti-noise ability of the SFA-WASD neuronet in the XOR application. Moreover, for high-dimension XOR application, the performance comparison is made between the power-function-activated (PFA) WASD neuronet and the SFA-WASD neuronet. The numerical results verify the superiority of the SFA-WASD neuronet in terms of solving nonlinear problems.

Key words: weights-and-structure-determination (WASD) algorithm; discontinuous signum function; neuronet; noise; XOR; high-dimension

人工神经网络具有并行处理、分布式储存、自适应学习和高容错能力等特性^[1], 因此, 被众多研究者广泛而深入地研究并应用于多个领域, 如模

式识别、系统辨识、信号处理和机器人控制等^[2-5]。其中 BP (Back Propagation, 误差回传) 神经网络模型是目前最为成熟的网络模型之一^[1],

* 收稿日期: 2013-07-02

基金项目: 国家自然科学基金资助项目 (61075121, 60935001); 教育部高等学校博士学科点专项科研基金博导类课题资助项目 (20100171110045)

作者简介: 张雨浓 (1973 年生), 男; 研究方向: 人工神经网络、冗余机器人; E-mail: zhyong@mail.sysu.edu.cn

但是该模型存在着一些固有的缺点和局限,如收敛速度慢、易陷入局部最小点、网络结构难以确定、激励函数须可导等^[1,6]。虽然有很多研究者在算法上对其进行了优化^[1,6],但是由于 BP 算法本质上仍是迭代算法,所以难以完全克服其缺点。

为避免 BP 冗长的迭代过程以及克服其固有的缺点,我们提出了基于矩阵伪逆的权值直接确定(WDD)方法^[7-9],可一步确定出网络的隐层神经元与输出层神经元之间的最优权值。此外,我们前期研究的工作表明网络学习误差与隐层神经元数目并非存在着严格的递减关系^[8-10],即,网络的性能不一定随着隐层神经元数目增加而提高。因此,一种基于 WDD 方法的权值与结构确定(WASD)算法被提出。该算法可以根据网络的学习误差变化,自动确定出网络的最优隐层神经元数目(即最优结构)。

与以往以连续光滑函数(如幂函数)作为神经网络的激励函数不同^[10],本文基于 WASD 算法,创新地提出和构造了一种以不连续的符号函数为隐层神经元激励函数的神经网络(即,符号函数激励的 WASD 神经网络),并将该网络应用于非线性分类问题求解,其中本文讨论的是 XOR 问题^[11]。考虑到在模式分类、数据挖掘与图像处理等过程中,所获取的数据可能存在着一定的噪声干扰,且噪声强度对后续实验的结果及处理质量有着很大的影响^[12-13]。本文因此详细地探讨了所构建的符号函数激励的 WASD 神经网络在数据受到不同类型噪声干扰时(也即,带噪 XOR 应用方面)的优越的抗噪性能。数值实验结果验证了这一优越性能。值得指出的是,相对于只能使用连续光滑激励函数的 BP 神经网络,使用非连续激励函数的 WASD 神经网络更容易以电路的形式实现。

另外,本文进一步将 XOR 推广到高维情况,使得分类非线性程度更高更难。通过对比符号函数激励的 WASD 神经网络与幂函数激励的 WASD 神经网络在高维 XOR 应用方面的性能差异,证实了符号函数激励的 WASD 神经网络在非线程度增加时学习效果更加优越且明显优于幂函数激励的神经网络。

1 WASD 神经网络模型

本节首先给出了旨在保证符号函数激励的神经网络学习能力的理论基础,然后根据其理论基础,构建出该神经网络模型。

1.1 理论基础

定理 1^[13] 设 F 是一个线性赋范空间, $f(t) \in F$, 函数系 $\{p_k(t); k = 1, 2, \dots, n\}$ 是 F 中的线性无关系, 则存在待定的权值 $\{w_k; k = 1, 2, \dots, n\}$ 使得 $f(t)$ 关于 $p_k(t)$ 的最佳逼近多项式 $p(t)$ 存在, 其中 $p(t)$ 可表示为

$$p(t) = w_1 p_1(t) + w_2 p_2(t) + \dots + w_k p_k(t) + \dots + w_n p_n(t)$$

定理 2^[15] 函数 $\theta_1(t), \theta_2(t), \dots, \theta_n(t)$ 在区间 $[a, b]$ 上线性无关的充要条件为存在 $t_i \in [a, b]$ (其中 $i = 1, 2, \dots, n$) 使得向量组 $\{[\theta_1(t_i), \theta_2(t_i), \dots, \theta_n(t_i)]\}_{i=1}^n$ 线性无关。

对于函数 $\theta_i(t) = \text{sgn}(t - a_i)$, 其中 $i = 1, 2, \dots, n, a_0 = a, a_n = b$ 且 $a_0 < a_1 < \dots < a_n$, 则任取 $t_i \in (a_{i-1}, a_i)$ 可得矩阵 Θ :

$$\Theta = [\theta_1 \quad \theta_2 \quad \dots \quad \theta_n] = \begin{bmatrix} \theta_1(t_1) & \theta_2(t_1) & \dots & \theta_n(t_1) \\ \theta_1(t_2) & \theta_2(t_2) & \dots & \theta_n(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \theta_1(t_n) & \theta_2(t_n) & \dots & \theta_n(t_n) \end{bmatrix} = \begin{bmatrix} -1 & -1 & \dots & -1 \\ 1 & -1 & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

其中 \sim 表示矩阵初等变换。显然矩阵 Θ 的秩 $\text{Rank}(\Theta) = n$, 即向量组 $\theta_1, \theta_2, \dots, \theta_n$ 线性无关。因此由定理 1 和定理 2 可推知, 目标函数 $f(t)$ 的最佳逼近多项式可以表示为

$$\theta(t) = w_1 \theta_1(t) + w_2 \theta_2(t) + \dots + w_i \theta_i(t) + \dots + w_n \theta_n(t)$$

其中 $\theta_i(t) = \text{sgn}(t - a_i)$ ($i = 1, 2, \dots, n$)。即, $f(t) \approx \theta(t) = w_1 \text{sgn}(t - a_1) + w_2 \text{sgn}(t - a_2) + \dots + w_i \text{sgn}(t - a_i) + \dots + w_n \text{sgn}(t - a_n)$ 由上述理论分析可得, 通过选取适当的权值 $\{w_i\}_{i=1}^n$, 可使 $\theta(t)$ 逼近目标函数 $f(t)$ 。

1.2 网络模型

由上述理论基础分析推广, 构造出符号函数激励的 WASD 前向神经网络模型如图 1 所示。该模型有三层神经元, 分别是线性激励输入层、符号函数激励隐层和线性激励输出层。具体来说, 输入层和输出层神经元的激励函数都是线性恒等函数, 且神经元的阈值都设为 0; 隐层神经元的激励函数则为符号函数, 神经元的阈值为 a_i 。值得说明的是, 实验中我们采用了试错法来确定 a_i 的最佳变化范

围，即从不同的范围中随机生成 a_i 进行实验，通过对比分析得出使得实验结果最好的 a_i 的取值范围。因此设输入层有 m 个神经元，隐层有 n 个神经元，输出层有 1 个神经元，则隐层神经元的激励函数为 $\theta_i = \text{sgn}(\sum_{j=1}^m x_j - a_i), i = 1, 2, \dots, n$ 。

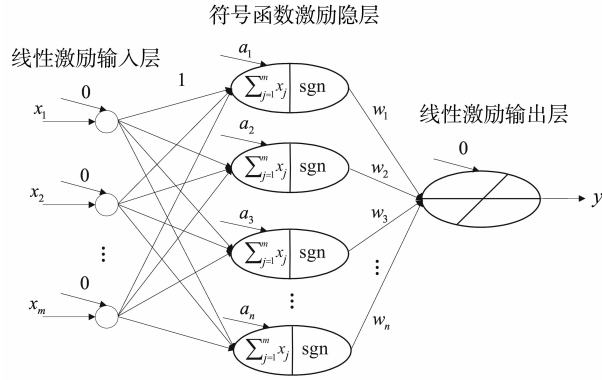


图 1 符号函数激励 (SFA) 的 WASD 神经网络 (SFA - WASD) 模型

Fig. 1 Model of signum function activated (SFA) WASD neuronet, i. e., SFA - WASD neuronet

而且，设输入层与隐层神经元间的连接权值均为 1，隐层与输出层神经元间的连接权值向量为 $W = [w_1, w_2, \dots, w_n]^T \in \mathbf{R}^n$ 。该网络的输入与输出关系故可以表示为 $y = \sum_{i=1}^n \text{sgn}(\sum_{j=1}^m x_j - a_i) w_i$ 。

2 权值与结构确定 (WASD) 算法

本节首先给出了权值直接确定法，然后在此基础上提出权值与结构确定算法并给出其具体实现过程。

如前所述，利用权值直接确定法可以一步得到隐层神经元与输出层神经元之间的最优连接权值，从而避免了传统 BP 算法的冗长迭代。在定理 3 中，我们给出权值直接确定法的理论基础与实现。

定理 3^[7-10] 如图 1 所示的前向神经网络，若已知目标输出向量 $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_l, \dots, \gamma_N]^T \in \mathbf{R}^N$ ，其中， γ_l 表示第 l 个样本对应的目标输出；以及输入受激励矩阵

$$J = \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \cdots & \theta_{1,n} \\ \theta_{2,1} & \theta_{2,2} & \cdots & \theta_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{N,1} & \theta_{N,2} & \cdots & \theta_{N,n} \end{bmatrix} \in \mathbf{R}^{N \times n}$$

其中， $\theta_{l,i} = \text{sgn}(\sum_{j=1}^m x_{l,j} - a_i)$ 表示第 i 个隐层神经元对第 l 个样本 $x_l = [x_{l,1}, x_{l,2}, \dots, x_{l,m}]^T$ 的激励响应值， N 表示训练样本个数。则隐层与输出层神经元

间的最优连接权值 $W = [w_1, w_2, \dots, w_n]^T$ 可直接确定为

$$W = (J^T J)^{-1} J^T \gamma = J^+ \gamma$$

其中， J^+ 表示矩阵 J 的伪逆，也可写作 $\text{pinv}(J)$ 。

然后，在权值直接确定法的基础上，我们进一步研究了网络误差与隐层神经元个数间的关系^[8-10]，发现随着神经元数的增加，网络输出误差不是单调递减的。换言之，当神经元数增长到一定值时，输出误差不但不减小甚至会有增大的趋势。值得指出的是，本文使用公式 $E = \|\gamma - JW\|_2^2 / N$ 来计算网络平均误差并将其作为衡量网络性能的标准，其中 $\|\cdot\|_2$ 表示向量的二范数。因此，为了优化网络结构，我们提出并研究了权值与结构确定算法，该算法在权值直接确定法的基础上，根据网络学习误差的变化，自动确定出网络的最优结构（即，最优的隐层神经元数目）。

WASD 算法的基本思想是：逐个增加隐层神经元数 n ，并且每次都计算出当前结构时的网络平均误差 E_{tmp} ；同时将 E_{tmp} 与之前结构中的最小网络平均误差 E_{min} 相比较，如果满足 $E_{\text{tmp}} \leq E_{\text{min}}$ ，则将此次的网络结构与网络误差保存作为当前的最优网络结构 N_{opt} 及最小误差 E_{min} ；否则不保存并继续增加隐层神经元（简称，隐神经元）。重复以上过程，如果向前连续增加 n_f 个隐神经元都不能满足 $E_{\text{tmp}} \leq E_{\text{min}}$ ，那么隐层神经元数为 $n - n_f$ 的结构就为网络的最优结构，具体流程如图 2 所示。

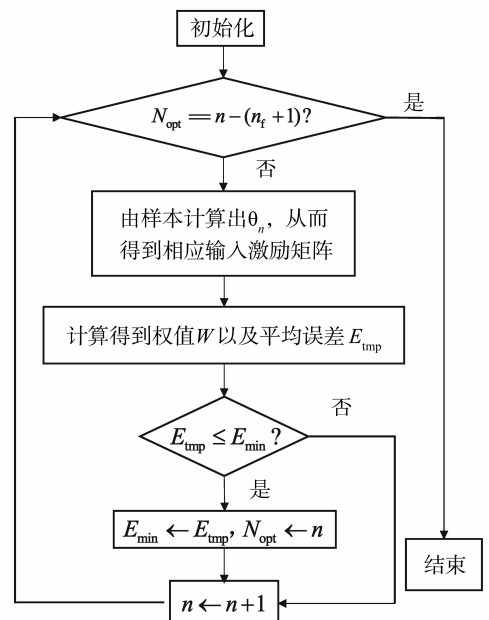


图 2 权值与结构确定算法流程图

Fig. 2 Flowchart of WASD algorithm

3 WASD 神经网络在 XOR 中的应用

本节使用一个两输入的符号函数激励的 WASD 神经网络以实现 XOR 应用, 具体而言, 当输入为 $(1, -1)$ 或 $(-1, 1)$ 时, 输出为 1; 输入为 $(1, 1)$ 或 $(-1, -1)$ 时, 输出为 -1。然而, 在实际应用中学习样本的输入输出值可能存在着各种噪声的干扰, 如本文所讨论的均匀噪声和高斯噪声。为了验证所提出的网络在 XOR 应用上具有优越的抗噪性能, 且能有效地自动确定最优权值与结构, 本文分别讨论了无噪学习、输入带噪学习以及输出带噪学习这三种不同情况的学习过程, 并分别对学习点以外的带噪声的点进行测试。值得指出的是本节实验中隐层神经元阈值 a_i 的取值范围为 $[-10, 10]$, 向前连续增加的隐层神经元数目 $n_i = 6$ 。

3.1 噪声类型

本节实验中所有的学习和测试样本受到的干扰噪声类型有两种, 均匀噪声和高斯噪声。假设无噪情况时向量 $z = [z_1, z_2, \dots, z_q]^T$, 若其受均匀噪声干扰且受扰幅值范围为 D , 则所有带噪向量的各个分量都服从均匀分布, 分布区间为 $[z_j - D, z_j + D]$, 其中 $j = 1, 2, \dots, q$ 。若其受高斯噪声干扰且受扰标准差为 σ , 则各个分量均服从高斯分布, 期望分别为 z_1, z_2, \dots, z_q ; 也值得指出的是, 根据 3σ 准则^[16], 受高斯噪声干扰的向量分量分布在范围 $[z_j - 3\sigma, z_j + 3\sigma]$ 内的概率都为 99.7%, 其中 $j = 1, 2, \dots, q$ 。因此为了增加不同噪声类型下实验结果的可比性, 本文设定 $3\sigma = D$ 。

3.2 无噪学习与带噪测试

学习过程中, 样本的输入为 $(1, 1)$, $(1, -1)$, $(-1, 1)$ 和 $(-1, -1)$, 对应的目标输出分别为 -1, 1, 1 和 -1。网络学习完成后, 分别在 4 个样本输入附近各选取 100 个带噪数据, 对学习所得的网络进行测试, 实验结果如表 1 所示。需要指出的是, 表中的噪声大小对于均匀噪声指的是分布范围 D , 对于高斯噪声指的是 3σ ; 表 1 中所有数据均为实验 100 次后的平均结果。为使结果更加直观, 我们在图 3 和图 4 (a) 中分别展示了测试噪声为高斯噪声 $3\sigma = 0.4$ 的一次测试效果及网络结构自确定过程。从图表中可以看出算法能够有效地自主确定网络结构, 而且其学习时间较短, 网络测试的正确率也较高。同时值得指出的是, 网络对均匀噪声和高斯噪声均具有较为优秀的抗噪性能, 其中对高斯噪声的抗噪性更强。

表 1 无噪学习带噪测试结果

测试噪声		噪声大小	最优隐层神经元数	正确率/%	学习时间/s
均匀噪声	0.2	15	90.935 0	0.035 9	
	0.4	17	86.017 5	0.041 9	
	0.6	17	81.265 0	0.044 2	
高斯噪声	0.2	17	92.785 0	0.029 9	
	0.4	17	89.945 0	0.025 8	
	0.6	16	89.770 0	0.022 0	

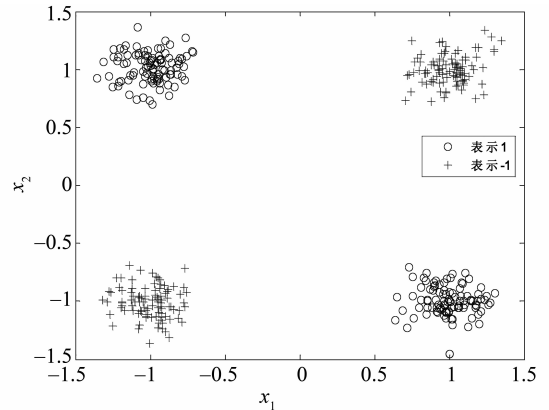


图 3 无噪学习的网络进行带噪测试的一次结果

Fig. 3 One noisy - test result after learning without noise

3.3 输入带噪学习与测试

接着, 我们讨论输入带噪的学习和测试情况。也即, 分别在 $(1, 1)$, $(1, -1)$, $(-1, 1)$ 和 $(-1, -1)$ 附近各选取 100 个带噪输入进行样本学习, 其相应目标输出仍分别为 -1, 1, 1 和 -1; 网络学习完成后, 重新在这四个无噪输入周围各选取 100 个带噪数据对所得的网络进行测试。学习中的输入噪声与测试中的噪声均使用不同大小的两类噪声。为了充分地显示网络的优越性能, 实验中对比了在学习噪声与测试噪声相同以及不同的情况时网络的抗噪性能。实验结果如表 2 和图 5 所示。另外, 图 4 (b) 中展示了输入带噪下网络结构的一次自确定过程。

表 2 中数据对应的测试噪声为 $D = 0.3$ 的均匀噪声和 $3\sigma = 0.3$ 的高斯噪声, 且所有数据都为实验 100 次后的均值。分析表中的数据可得, 当输入带噪时, 网络也能够有效地实现结构和权值的最优自确定, 而且算法的正确率很高, 并高于不带噪学习时的情况。另外, 学习中输入带噪类型与测试中输入带噪类型的异同对结果影响不大, 正确率均较高。此外, 图 5 直观地展示了学习点受 $D = 0.2$ 的均匀噪声干扰时, 测试正确率随两种测试噪声大小

的变化情况。从图中可以看出：测试噪声在一定范围时，网络能够保证具有很高的正确率；且测试数据受高斯噪声干扰时的网络抗噪效果好于受均匀噪

声干扰时的效果。总而言之，表 2 和图 5 的实验结果充分地显示了网络在输入带噪学习时具有优秀的抗噪性能。

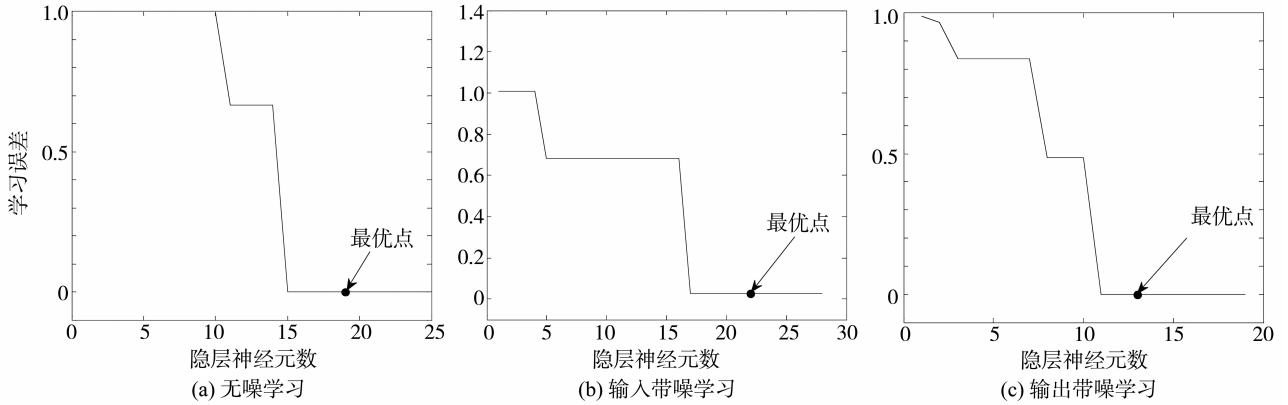


图 4 学习误差与隐神经元数关系曲线

Fig. 4 Curves of the relationship between the learning error and the number of hidden - layer neurons

表 2 输入带噪的学习与测试结果

Table 2 Results of learning and testing both with input noise

学习样本		学习样本与测试样本噪声类型相同			学习样本与测试样本噪声类型不同		
噪声类型	噪声大小	WASD 算法最优 隐神经元数	正确率/%	学习时间/s	WASD 算法最优 隐神经元数	正确率/%	学习时间/s
均匀 噪声	0.2	19	92.372 5	0.032 0	19	93.260 0	0.032 9
	0.4	30	94.875 0	0.105 9	31	94.510 0	0.140 4
	0.6	39	94.547 5	0.230 4	39	94.160 0	0.399 2
	0.8	46	93.602 5	0.326 4	50	97.270 0	0.295 7
高斯 噪声	0.2	19	94.012 5	0.030 2	20	93.130 0	0.035 2
	0.4	21	94.485 0	0.042 4	21	92.670 0	0.043 0
	0.6	28	96.772 5	0.112 4	30	94.662 5	0.145 9
	0.8	43	94.485 0	0.552 5	47	94.307 5	2.769 0

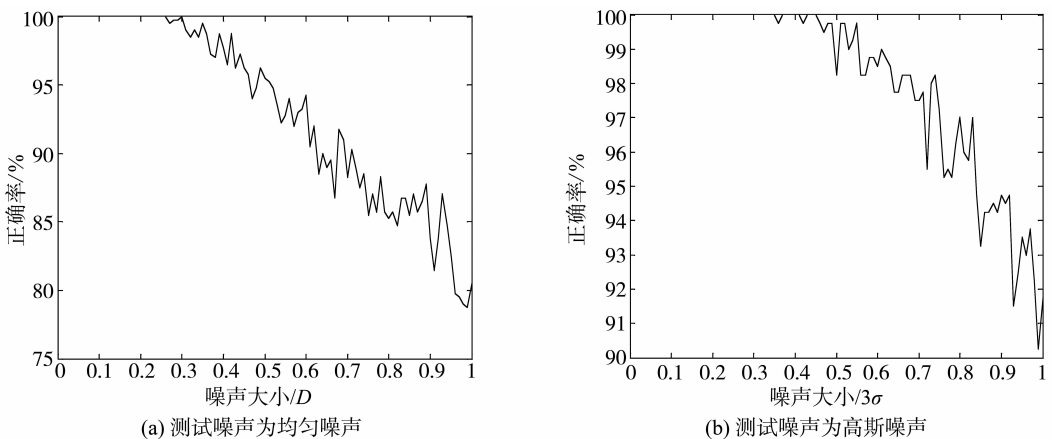


图 5 学习输入带均匀噪声的网络正确率随不同测试噪声的变化曲线

Fig. 5 Variation curves of network accuracy over different test noise after learning with uniformly noisy input

3.4 输出带噪学习与测试

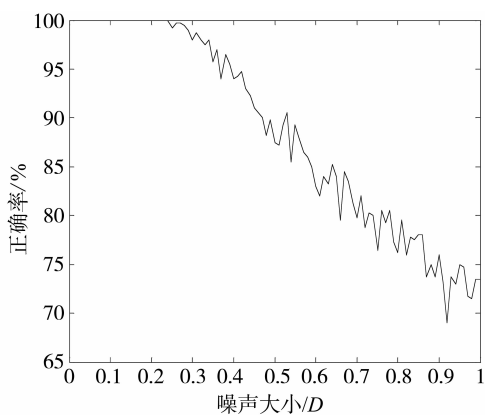
本节最后探讨输出带噪学习与测试的情况。也即, 平均分布 400 个样本输入在 $(1, 1)$, $(1, -1)$, $(-1, 1)$ 和 $(-1, -1)$ 处, 其输出分别对应为在 $-1, 1, 1$ 和 -1 周围的带噪值。网络学习完成后, 重新在这四个无噪输入周围选取 100 个带噪数据对所得的网络进行测试。学习中输出噪声与测试中的噪声均使用不同大小的两类噪声, 实验同样比较在学习噪声和测试噪声异同时的学习效果, 结果如表 3 与图 6 所示; 并在图 4 (c) 展示了输出带噪学习时网络结构的一次自确定过程。与表 2 做法一致, 表 3 中数据对应的测试噪声为 $D = 0.3$ 的均匀噪声和 $3\sigma = 0.3$ 的高斯噪声, 且所有数

据都为实验 100 次后的均值。图 6 中学习样本的输出噪声为 $D = 0.2$ 的均匀噪声。分析图表可以得出: 当学习样本输出带噪时, 网络能够有效地实现结构和权值的最优自确定, 并具有很高的正确率。另外, 输出噪声类型与测试噪声类型在相同与不同时网络均有很好的测试效果, 而且输出噪声大小与测试正确率及学习时间的关联性不强。此外, 由图 6 可看出: 测试噪声在一定范围时网络均可以保证较高的正确率; 且测试数据受高斯噪声干扰时的网络抗噪效果优于受均匀噪声干扰时的效果。实验结果充分显示了网络在输出带噪学习的情况时具有优越的抗噪性能。

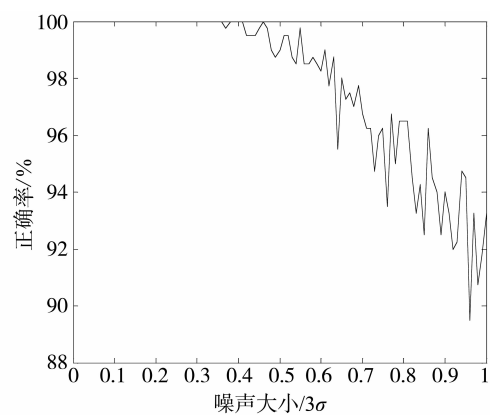
表 3 输出带噪的学习与测试结果

Table 3 Results of learning and testing both with output noise

学习样本		学习样本与测试样本噪声类型相同			学习样本与测试样本噪声类型不同		
噪声类型	噪声大小	WASD 算法最优 隐神经元数	正确率/%	学习时间/s	WASD 算法最优 隐神经元数	正确率/%	学习时间/s
均匀 噪声	0.2	33	91.947 5	0.954 8	25	93.507 5	0.140 3
	0.4	28	92.862 5	0.106 3	29	93.125 0	0.219 0
	0.6	23	90.832 5	0.075 1	24	93.267 5	0.063 0
	0.8	29	94.097 5	0.109 6	34	94.392 5	1.799 3
高斯 噪声	0.2	30	94.680 0	0.494 4	33	93.505 0	0.678 7
	0.4	34	94.507 5	1.350 6	36	92.662 5	1.412 4
	0.6	35	94.970 0	0.854 2	25	91.967 5	0.100 8
	0.8	37	94.757 5	1.037 2	37	91.560 0	1.768 8



(a) 测试噪声为均匀噪声



(b) 测试噪声为高斯噪声

图 6 学习输出带均匀噪声的网络正确率随不同测试噪声的变化曲线

Fig. 6 Variation curves of network accuracy over different test noise after learning with uniformly noisy output

4 高维 XOR 提出、WASD 应用与对比

上节探讨了传统的 XOR 问题, 即判断两个输入是否相同。为了进一步验证本文所提出的不连续

符号函数激励的 WASD 神经网络在解决非线性程度更高的分类问题中的有效性与优越性, 本节以作者新近提出的唯一性逻辑为基础将传统 XOR 推广到高维情况并给出了如下定义^[17]。

定义 1 对于双极二进制向量 $[B_1, B_2, \dots, B_p]^T \in \{-1, 1\}^p$ (其中 $p > 2$)，若找到且只找到 (换言之，有且只有) 一个分量 $B_k \in \{-1, 1\}$ ，使得 $B_k \neq B_i, \forall i \in \{1, 2, \dots, p\}$ ，且 $i \neq k$ ，则此 p 维双极二进制向量的该逻辑运算称为异或 (XOR) 逻辑运算且其结果为 1，即此 p 维双极二进制向量的异或 (XOR) 逻辑运算结果为 1。否则，此 p 维双极二进制向量的异或 (XOR) 逻辑运算结果为 -1。

定义 2 对于单极二进制向量 $[B_1, B_2, \dots, B_p]^T \in \{0, 1\}^p$ (其中 $p > 2$)，若找到且只找到一个分量 $B_k \in \{0, 1\}$ ，使得 $B_k \neq B_i, \forall i \in \{1, 2, \dots, p\}$ ，且 $i \neq k$ ，则此 p 维单极二进制向量的异或 (XOR) 逻辑运算结果为 1。否则，此 p 维单极二进制向量的异或逻辑运算结果为 0。

根据上述定义，可以得到高维 XOR 的输入输出数据表，其中表 4 示例性地给出了 5 维双极二进制向量的 XOR 逻辑运算的部分输入输出数据表。

表 4 五维双极二进制向量的 XOR 运算的部分输入输出数据表

Table 4 Partial input-and-output data of XOR operation for 5-dimension bipolar binary vectors

输入向量	输出	输入向量	输出
-1, -1, -1, -1, -1	-1	-1, 1, -1, 1, -1	-1
-1, -1, -1, -1, 1	1	-1, 1, -1, 1, 1	-1
-1, -1, -1, 1, -1	1	-1, 1, 1, -1, -1	-1
-1, -1, -1, 1, 1	-1	-1, 1, 1, -1, 1	-1
-1, -1, 1, -1, -1	1	-1, 1, 1, 1, -1	-1
-1, -1, 1, -1, 1	-1	-1, 1, 1, 1, 1	1
-1, -1, 1, 1, -1	-1	1, -1, -1, -1, -1	1
-1, -1, 1, 1, 1	-1	1, -1, -1, -1, 1	-1
-1, 1, -1, -1, -1	1	1, -1, -1, 1, -1	-1
-1, 1, -1, -1, 1	-1	1, -1, -1, 1, 1	-1

此外，我们将本文提出的符号函数激励的 WASD 神经网络应用于该高维 XOR 逻辑问题求解 (具体为 3 至 6 维 XOR 情况)，实验结果如表 5 所示。从表中可知，符号函数激励的 WASD 神经网络能够有效地实现高维 XOR 问题求解 (正确率为 100%)，这展示了所提出的符号函数激励的 WASD 神经网络具有优越的学习性能。值得指出的是，在本节的计算机数值实验里，隐层神经元阈值的取值范围均为 $[-70, 70]$ ，向前连续增加的隐层神经元数为 49。

表 5 符号函数激励的与幂函数激励的 WASD 神经网络的高维 XOR 学习正确率

Table 5 Learning accuracy of SFA and PFA (power function activated) WASD neuronets on high-dimension XOR

向量维数	符号函数激励神经网络正确率/%	幂函数激励神经网络正确率/%
3	100.00	100.00
4	100.00	87.50
5	100.00	81.25
6	100.00	87.50

最后，我们也对比了所提出的符号函数激励的 WASD 神经网络与幂函数激励的 WASD 神经网络在 3 至 6 维 XOR 应用方面的性能差异，其对比性的实验结果也如表 5 (尤其是表 5 最右列) 所示。由该表中可以看出，当向量的维数较低时，即学习样本非线性程度较小时，符号函数激励的和幂函数激励的 WASD 神经网络均能够实现正确率为 100% 的学习。随着学习样本非线性程度的增加，即 XOR 的维数增加，符号激励函数的神经网络仍能保持 100% 的学习正确率，而幂激励函数神经网络的学习性能却变得相对较差 (其学习正确率低于 90%)。这说明不连续的符号函数激励的 WASD 神经网络在学习非线性程度较大的 XOR 方面优于以往常用的连续光滑型的幂函数激励的 WASD 神经网络。以上对比结果进一步验证了所提出的符号函数激励的 WASD 神经网络在解决非线性程度更高的高维 XOR 问题的有效性与优越性。

5 结 论

为了探讨并研究非连续函数作为隐层激励函数的可行性及其学习问题，本文构造了一种以非连续符号函数为激励函数的 WASD 神经网络模型，其通过权值与结构确定 (WASD) 算法得到网络的最优权值及最优结构，然后将其应用于非线性分类中的 XOR 问题求解。计算机数值实验结果表明：构建的 WASD 神经网络能有效地实现其最优权值与最优结构的自确定，且此神经网络在 XOR 应用中具有十分优越的抗噪性能。另外，本文进一步将 XOR 推广到高维，通过对比符号函数激励的神经网络与幂函数激励的神经网络在高维 XOR 应用上的性能差异，证实了符号函数激励的神经网络在非线程度增加时具备更为优越的学习性能。这也展现了符号函数激励的 WASD 神经网络在模式识别和数据挖掘等方面的应用潜力和前景。

ceedings of STOC 2005, 2005: 523 – 532.

- [5] CANETTI RAN, VARIA MAYANK. Non-malleable obfuscation [C] // TCC 2009, LNCS, 2009, 5444: 73 – 90.
- [6] LYNN BENJAMIN, PRABHAKARAN MANOJ, SAHAI AMIT. Positive results and techniques for Obfuscation [C] // EUROCRYPT 2004, LNCS, 2004, 3027: 20 – 39.
- [7] GOLDWASSER S, KALAI YAEL TAUMAN. On the impossibility of obfuscation with auxiliary input [C] // FOCS 2005, 2005: 60.
- [8] HOHENBERGER SUSAN, ROTHBLUM GUY N, SHE-LAT ABHI, et al. Securely obfuscating re-encryption [C] // TCC 2007, LNCS, 2007, 4392: 233 – 252.
- [9] SHACHAM HOVAV, WATERS BRENT. Efficient ring signatures without random oracles [C] // PKC 2007, 2007, 4450: 166 – 180.
- [10] BONEH DAN, BOYEN XAVIER, SHACHAM HOVAV. Short group signatures [C]. // CRYPTO 2004, LNCS, 2004, 3152: 227 – 242.
- [11] BELLARE MIHIR, DESAI ANAND, POINTCHEVAL DAVID, et al. Relations among notions of security for public-key encryption schemes [C] // CRYPTO 98, LNCS, 1998, 1462: 26 – 45.

(上接第 7 页)

参考文献:

- [1] 张雨浓, 杨逸文, 李巍. 神经网络权值直接确定法 [M]. 广州: 中山大学出版社, 2010.
- [2] 韩红桂, 甄博然, 乔俊飞. 动态结构优化神经网络及其在溶解氧控制中的应用 [J]. 信息与控制, 2010, 39 (3): 354 – 360.
- [3] 常春, 陈怡群, 肖宏儒, 等. 基于神经网络图像分析的智能鲜茶叶分选机 [J]. 中国农机化学报, 2013, 34 (1): 137 – 141.
- [4] 贾鹤鸣, 张利军, 齐雪, 等. 基于神经网络的水下机器人三维航迹跟踪控制 [J]. 控制理论与应用, 2012, 29(7): 877 – 883.
- [5] 习会峰, 汤立群, 何庭惠, 等. 基于神经网络和遗传算法的桥梁参数优化方法与分析 [J]. 中山大学学报: 自然科学版, 2008, 47(增刊 2): 46 – 49.
- [6] 张月琴, 刘翔, 孙先洋. 一种改进的 BP 神经网络算法与应用 [J]. 计算机技术与发展, 2012, 22(8): 163 – 166.
- [7] 杨文光. 权值直接确定的三角型模糊前向神经网络 [J]. 中山大学学报: 自然科学版, 2013, 52(2): 33 – 37.
- [8] ZHANG Y N, CHEN J W, FU S B, et al. Weights and structure determination of multiple-input Hermit orthogonal polynomials neural network [C] // Control and Decision Conference (CCDC), IEEE, 2012: 1106 – 1111.
- [9] 张雨浓, 李钧, 张智军, 等. SIMO 傅里叶三角基神经网络的权值直接确定法和结构自确定算法 [J]. 信息与控制, 2011, 40(4): 507 – 513.
- [10] 张雨浓, 劳稳超, 余晓填, 等. 两输入幂激励前向神经网络权值与结构确定 [J]. 计算机工程与应用, 2012, 48(15): 102 – 106.
- [11] 王进, 烤贾雨, 熊超, 等. 基于神经网络异或算法的误差演算 [J]. 电子世界, 2012, (11): 87.
- [12] YIN H, DONG H B, HU Y P. A new view on noise cleaning for classification [C] // Conference on Information Management and Engineering (ICIME), IEEE, 2011: 597 – 601.
- [13] RAFAEL C G, RICHARD E W, STEVEN L E. 数字图像处理 (MATLAB 版) [M]. 北京: 电子工业出版社, 2009: 103 – 116.
- [14] 邹阿金, 张雨浓. 基函数神经网络及应用 [M]. 广州: 中山大学出版社, 2009: 18 – 19.
- [15] 张小勇, 徐香勤, 贾利新. 函数组线性无关性的研究 [J]. 河南科学, 2013, 31(1): 25 – 27.
- [16] 盛骤, 谢式千, 潘承毅. 概率论与数理统计 [M]. 第四版. 北京: 高等教育出版社, 2008: 49.
- [17] 张雨浓, 邓健豪, 金龙, 等. 唯一性逻辑及其 BP 神经网络侦测 [J]. 中山大学学报: 自然科学版, 2013, 52(3): 1 – 5.