

MISO 多元广义多项式神经网络及其权值直接求解^{*}

肖秀春^{1,2}, 张雨浓¹, 姜孝华¹

(1. 中山大学信息科学与技术学院, 广东广州 510275; 2. 广东海洋大学信息学院, 广东湛江 524025)

摘 要: 基于多元函数逼近理论, 构建一种 MISO (Multiple-Input, Single-Output) 多元广义多项式神经网络。依据最小二乘原理, 推导出基于伪逆的最优权值一步计算公式——简称为权值直接确定法; 在此基础上, 提出基于指数增长和折半删减搜索策略的隐神经元数自适应增删搜索算法。该新型神经网络具有结构简单的优点, 其权值直接确定法、隐神经元增删算法可以避免冗长的迭代计算、局部极小点和学习率难选取等问题, 同时解决了传统 BP 神经网络难以确定隐神经元数这一难题。仿真实验显示其具有训练速度快、逼近精度高和良好的去噪特性等特点。

关键词: 多元广义多项式; 权值直接确定; 结构自适应确定; 指数增长; 折半删减

中图分类号: TP183 **文献标识码:** A **文章编号:** 0529-6579 (2009) 04-0042-06

MISO Multivariate Generalized Polynomials Neural Network and its Weights-Direct-Determination

XIAO Xiuchun^{1,2}, ZHANG Yunong¹, JIANG Xiaohua¹

(1. School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510275
China; 2. College of Information, Guangdong Ocean University, Zhanjiang 524025, China)

Abstract: A new type of MISO (Multiple-Input, Single-Output) multivariate generalized polynomials neural network is constructed based on multivariate function approximation theory. According to least square theorem, a pseudoinverse-based weights-direct-determination method is further presented to determine the neural-weights just in one step. Moreover, on the basis of this weights-direct-determination, a hidden-layer evolution algorithm is proposed based on exponential-growth and binary-delete-search strategy. Theoretical analysis demonstrates that, since the weights-direct-determination method and the hidden-layer evolution algorithm could obtain the optimal weights directly without lengthy iterative BP-training, the constructed neural network could remedy the weakness of conventional BP neural networks, such as the existence of local-minima, choosing of learning-rate as well as the determination of the hidden-layer neurons. Computer simulation results substantiate the advantages of weights-direct-determination method and hidden-layer evolution algorithm for the constructed neural network, in the sense of training speed and high approximation precision.

Key words: multivariate generalized polynomials; weights-direct-determination; structure-adaptive-determination; exponential growth; binary search

人工神经网络 (Artificial Neural Network, 以下简称神经网络) 是模拟人的神经系统之结构、原

理及功能的一种仿生数学模型, 由于其并行计算、分布式存储、高度容错、自组织自学习等特性而受

^{*} 收稿日期: 2008-12-11

基金项目: 国家自然科学基金资助项目 (60643004、60775050); 中山大学科研启动费、后备重点资助项目
作者简介: 肖秀春 (1976 年生), 男, 讲师, 博士生; 通讯作者: 姜孝华; E-mail: springxc@163.com

到人们的广泛关注^[1-4]。

BP (Back Propagation) 神经网络是由 Rumelhart 和 McClelland 等人于 1986 年提出的一种前向神经网络, 是目前应用最为广泛的神经网络模型及算法之一^[1,5], 由于采用基于梯度下降思想的误差回传 (即 BP) 学习算法而得名。

然而, 传统 BP 神经网络及其学习算法却存在一些局限^[6], 如: 迭代过程冗长, 学习过程易发生震荡; 难以确保搜索到全局最优值; 网络隐神经元数过多时, 算法时间及空间代价高; 不容易确定网络最优结构等。尽管一些改进的 BP 算法, 如加动量项的 BP 算法、拟牛顿 BP 算法、与遗传算法结合的 BP 算法及与模拟退火相结合的 BP 算法等均可从一定程度上克服上述缺陷, 但这些改进并没有从根本上解决上述问题^[7]。值得指出的是, BP 神经网络及其训练算法存在的上述局限性, 是其自身所固有的, 一般而言, 难以找到有效、彻底的解决办法^[6-7]。

基于多元函数逼近理论, 本文构造了一种 MISO 多元广义多项式神经网络。依据最小二乘原理, 推导出基于伪逆的最优权值一步计算公式。考虑其最优权值可一步计算确定, 该新型神经网络结构可以自适应地确定。仿真实验显示其还具有训练速度快、逼近精度高和对随机噪声可有效去除等特点。

1 神经网络拓扑结构

如图 1 所示, 本文构造的 MISO 多元广义多项式神经网络为三层前向结构。输入层 m 个神经元和输出层的单神经元可采用线性激励函数; 隐层有 n 个神经元, 使用一组多元广义多项式 $\{p_i: = p_i(x_1, x_2, \dots, x_m)\}_{i=1}^n$ 作为其激励函数。另外, 所有

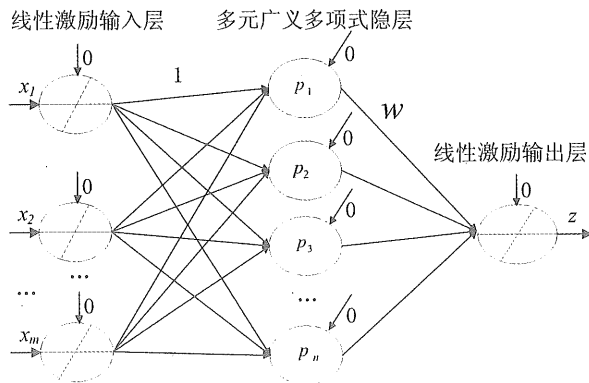


图1 MISO 多元广义多项式神经网络结构

Tab. 1 Structure of MISO multivariate generalized polynomials neural network

神经元阈值固定为 0; 输入层与隐神经元之间的连接权值恒为 1; 第 i 个隐神经元与输出神经元的连接权值为 w_i ($i = 1, 2, \dots, n$); 神经网络输出为 z 。值得指出的是, 此神经网络模型中, 仅隐层到输出层权值向量 w 需要调整 [或确定], 具有结构简单的优点。

对于上述结构的神经网络, 其输入与输出的关系显然可以简单描述为下式

$$z = \sum_{i=1}^n w_i p_i(x_1, x_2, \dots, x_m) \quad (1)$$

2 最优权值直接求解及性能验证

本节基于最小二乘法原理, 讨论 MISO 多元广义多项式神经网络最优权值直接求解方法, 且通过仿真实例、与 RBF 神经网络的性能比较, 对本文神经网络的可行性、有效性进行验证。不失一般性 (同时, 也为简单起见), 这里仅以二元幂单项式簇 $\{p_i(u, v)\}_{i=1}^n$ 作为隐层神经元激励函数进行讨论。

2.1 二元幂单项式序列的表达式

当二元幂单项式簇 $\{p_i(u, v)\}_{i=1}^n$ 按照 Graded lexicographic 次序^[8]排列时, 则序列中第 i 个二元幂单项式 $p_i(u, v)$ 的定义可如下式:

$$p_i(u, v) = u^{j_1} v^{j_2}, i = 1, 2, \dots, n, \sigma = \lceil (-3 + \sqrt{1 + 8i})/2 \rceil, j_1 = (\sigma + 1)(\sigma + 2)/2 - i, j_2 = \sigma - j_1 \quad (2)$$

于是, 二元幂多项式神经网络的表达式为,

$$z = \sum_{i=1}^n w_i p_i(u, v) \quad (3)$$

式 (2) 和 (3) 中, $p_i(u, v) = u^{j_1} v^{j_2}, i = 1, 2, \dots, n$ 表示二元幂单项式序列, j_1, j_2 分别表示变量 u 和 v 的幂次。 σ 表示变量 u 和 v 总幂次, 即 j_1, j_2 之和, 且 j_1, j_2, σ 的表达式可如下推导得出。

如果二元幂单项式簇中单项式按照 Graded lexicographic 次序排序^[8], 则其第 i 个二元幂单项式的变量 u, v 的总幂次 σ 与其序号 i 存在如下不等式关系, 即

$$\sigma(\sigma + 1)/2 < i \leq (\sigma + 1)(\sigma + 2)/2$$

解此不等式, 可得

$$(-1 + \sqrt{1 + 8i})/2 < \sigma \leq (-3 + \sqrt{1 + 8i})/2$$

从而有, $\sigma = \lceil (-3 + \sqrt{1 + 8i})/2 \rceil$;

进而有,

$$j_1 = (\sigma + 1)(\sigma + 2)/2 - i \text{ 和 } j_2 = \sigma - j_1$$

值得指出的是, 对于三元幂单项式簇亦存在类似式 (2) 的直接表达的序列, 但其表达形式要复杂一些, 而对于三元以上的多元幂单项式簇, 不存

在类似式(2)的直接表达式。但此时多元单项式簇中第*i*个单项式仍可以采用递推的方法递推求得。因此,这不影响本文所提出的神经网络的推广应用。

2.2 最优权值直接求解

对于欲逼近的二元连续目标函数 $f(u, v)$, 定义样本集 $\{(u_{s_1}, v_{s_2}), f_s\}, s_1 = 1, 2, \dots, l; s_2 = 1, 2, \dots, r, s = s_1 + (s_2 - 1)l$ [其中采样函数值 $f_s = f(u_{s_1}, v_{s_2})$, 样本总量为 $d = lr$], 且定义样本集中的输出向量 $\gamma = [f_1 \ f_2 \ \dots \ f_d \ 1]^T \in \mathbb{R}^d$; 定义神经网络权值向量 $w = [w_1 \ w_2 \ \dots \ w_n]^T \in \mathbb{R}^n$; 定义隐层神经元激励函数如式(2); 定义第*i*个隐神经元对第*s*个样本的激励响应为 $q_i^{(s)} = p_i(u_{s_1}, v_{s_2}), s_1 = 1, 2, \dots, l, s_2 = 1, 2, \dots, r, s = s_1 + (s_2 - 1)l, i = 1, 2, \dots, n$; 定义隐层对第*s*个样本的激励响应向量为 $q^{(s)} = [q_1^{(s)} \ q_2^{(s)} \ \dots \ q_n^{(s)}]^T \in \mathbb{R}^n$; 定义隐层对全体样本的激励响应矩阵 Q 如下式

$$Q = \begin{bmatrix} q_1^{(1)} & q_2^{(1)} & \dots & q_n^{(1)} \\ q_1^{(2)} & q_2^{(2)} & \dots & q_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ q_1^{(d)} & q_2^{(d)} & \dots & q_n^{(d)} \end{bmatrix} \in \mathbb{R}^{d \times n}$$

则对于如下定义的均方误差函数 $E(w)$,

$$E(w) = \frac{1}{2} \sum_{s=1}^d \left(\sum_{i=1}^n w_i q_i^{(s)} - f_s \right)^2$$

显然有,

$$\begin{aligned} E(w) &= \frac{1}{2} \sum_{s=1}^d \left(\sum_{i=1}^n w_i q_i^{(s)} - f_s \right)^2 = \\ &= \frac{1}{2} (Qw - \gamma)^T (Qw - \gamma) = \\ &= \frac{1}{2} \|Qw - \gamma\|_2^2 \end{aligned}$$

依据最小二乘原理,要优化误差函数 $E(w)$, 可令 $E(w)$ 的梯度值 $\nabla E(w) = 0$, 也即,

$$\nabla E(w) = \frac{\partial E(w)}{\partial w} = Q^T(Qw - \gamma) = 0$$

即, $Q^T Qw - Q^T \gamma = 0$,

上式等价于, $Q^T Qw = Q^T \gamma$,

当激励响应矩阵 Q 列满秩时,有

$$w = (Q^T Q)^{-1} Q^T \gamma = Q^+ \gamma$$

2.3 性能验证及与RBF神经网络性能对比

2.3.1 最优权值直接求解的可行性和去噪性能验证 为验证最优权值直接求解的可行性、有效性及去噪性能,如前所述,我们选用二元幂单项式序列 $p_i(u, v), i = 1, 2, \dots, n$ 作为隐神经元激励函数,对如下实例函数I和II进行计算机仿真验证(仿真环境为Intel P4 CPU 2.4GHz, 512MB内存,下同)。

$$\text{实例 I } f_1(u, v) = \frac{0.015 \times \sin(\sqrt{(10u)^2 + (10v)^2})}{\sqrt{(10u)^2 + (10v)^2}} + 20$$

$$\text{实例 II } f_2(u, v) = 2uve^{-2u^2-2v^2} + 20$$

我们对上述两例目标函数在 $u \in [-1, 1], v \in [-1, 1]$ 上均匀采样 51×51 组数据组成训练样本集,采用本文提出的MISO广义多项式神经网络的权值求解法(注:隐神经元数为63个)对样本集进行逼近实验。上述实例I和实例II的仿真逼近结果如图2和图3所示。

在图2和图3中,子图(a)为欲逼近的目标函数曲面,子图(b)和(c)分别为网络逼近输出及其相对误差曲面。子图(d)为实例目标函数带高斯噪声时的情况[噪声用来模拟数据采集时的测量或观测误差,其幅值为原始数据最大幅值的10%],子图(e)和(f)为网络对噪声的去除结果及其对应的相对误差。

观察图2和图3中仿真逼近结果及相对误差曲面,可看出本文算法能很好地学习样本集数据,以逼近上述目标函数实例I和实例II,且对于带有较大幅值随机高斯噪声的样本集,也能较有效地去除噪声的影响。在实际应用中,这可用于有效消除数据采集时所引入到样本集的测量或观测误差。另外,与文献[10-12]仿真实验结果相比较可知,本文算法在逼近精度方面也具有优势。

2.3.2 本文神经网络与RBF神经网络性能比较

为了进一步考查本文算法的性能,我们对本文神经网络与RBF神经网络进行了性能对比。为此,应用Matlab神经网络工具箱实现了RBF神经网络对于本文实例的逼近仿真实验,其中训练样本和神经元数目与本文神经网络相同,学习率为0.1,最大学习次数为10 000次。其训练时间和逼近精度与本文算法数据对比如表1所示。

表1 本文神经网络与RBF神经网络性能比较

Table 1 Comparison on performance of the proposed neural network and RBF neural network

实例	训练时间/s	逼近均方误差
	(本文神经网络/RBF)	(本文神经网络/RBF)
实例 I	11.31 / 2.22×10^3	5.83×10^{-11} / 2.73×10^{-7}
实例 II	11.38 / 2.11×10^3	6.83×10^{-10} / 2.40×10^{-4}

考察表1中对比数据可知,本文神经网络在训练时间及逼近精度上都优于RBF神经网络。

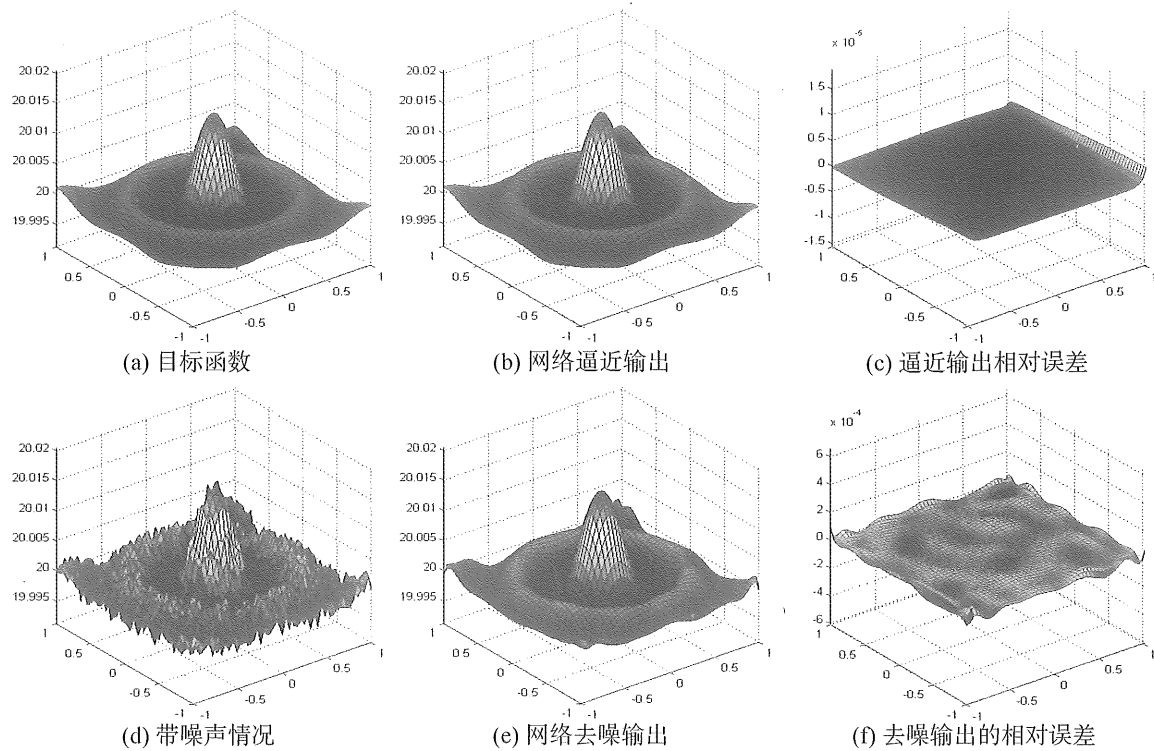


图 2 MISO 多元广义多项式神经网络对实例 I 的逼近效果及去噪性能

Fig. 2 Approximation and denoise results of MISO multivariate generalized polynomials neural network to instance I

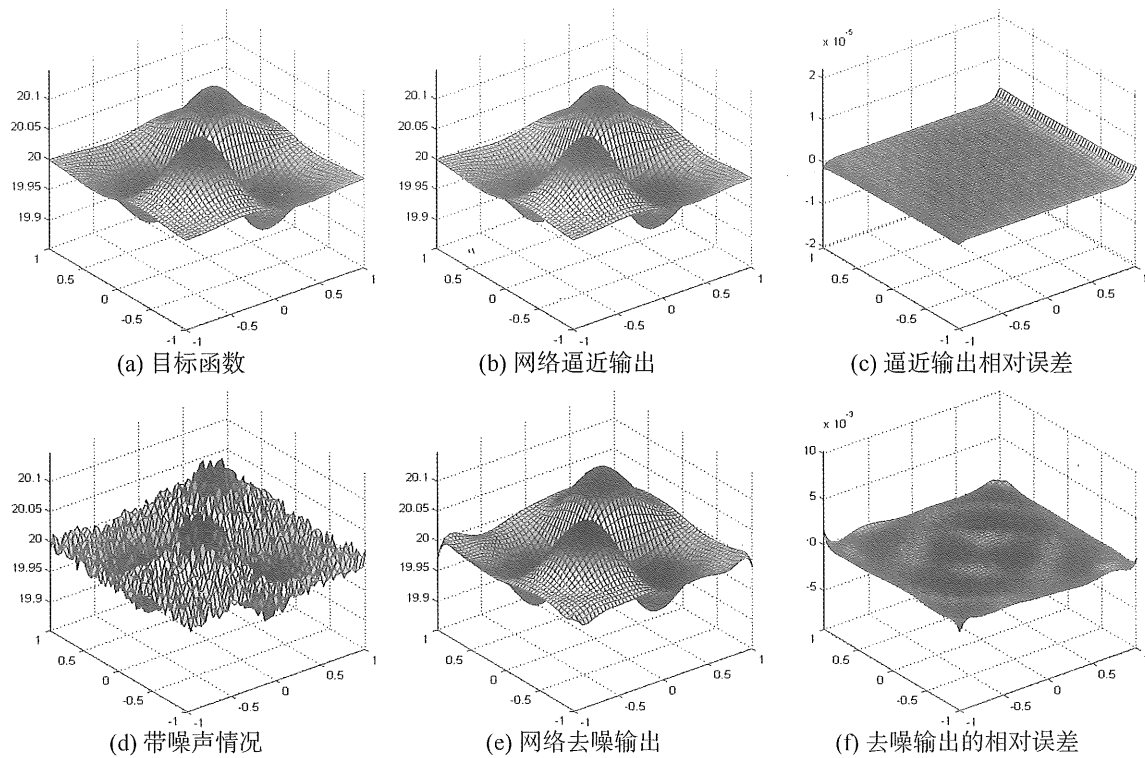


图 3 MISO 多元广义多项式神经网络对实例 II 的逼近效果及去噪性能

Fig. 3 Approximation and denoise results of MISO multivariate generalized polynomials neural network to instance II

3 隐神经元增删算法

一般来说,神经网络隐神经元数对于网络性能有非常重要的影响^[13]。因此,确定合适的隐神经元数是神经网络理论研究的一个重要方面。本文针对 MISO 多元广义多项式神经网络隐神经元数目越多,其能够达到的逼近精度越高的特点,提出了一套基于指数增长和折半删减搜索策略的隐神经元增删算法,它与自然数增长方式的搜索策略相比^[14],可以较大幅度地减少搜索次数,即,算法搜索次数由 $O(n)$ 降低到 $O(\log_2(n))$ 。

本文算法主要思想可简述为:首先按照指数增长的方式增加神经网络隐神经元数目,直到神经网络满足精度要求,但由于隐神经元的生长采用的是指数增长方式,因此,最后一次增长可能增加了过多的隐神经元,此时的神经网络未必是满足要求的最优结构;为此,我们需要删减过多的隐神经元,为了删减过多的隐神经元,采用折半搜索策略以减少搜索次数。算法流程如图 4 所示,其中 c 为网络当前隐神经元数且 $c = 2^h$, h 为控制指数增长的幂, a 和 b 分别为折半搜索时搜索区间的上界和下界, ε 和 e 分别为神经网络的目标误差和当前误差。

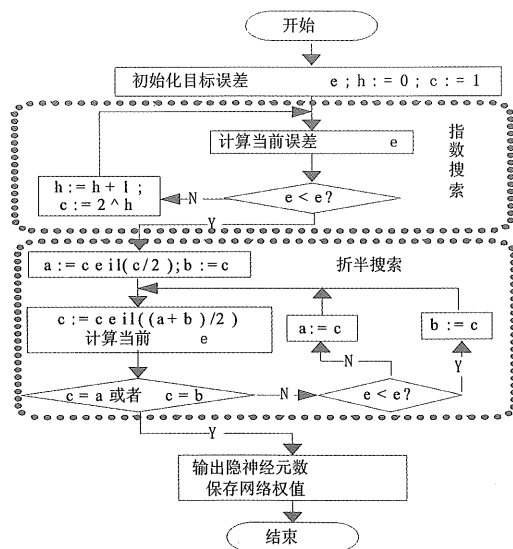


图 4 MISO 多元广义多项式神经网络
隐神经元增删算法

Fig. 4 Flow chart of the proposed evolution algorithm for MISO multivariate generalized polynomials neural network

上述算法与隐神经元自然增长方式的方法仿真性能对比数据如表 2 所示。由表 2 可知,当需要隐神经元较多时[尤其当需要 100 个以上的隐神经元

时],本文算法可有效地减少搜索次数和搜索时间。

表 2 增删算法/自然数增长算法性能对比

Table 2 Comparison on performance of the proposed evolution algorithm and natural evolution algorithm

目标 MSE 精度	搜索时间/s (增删算法/自然 增长搜索法)	搜索次数(增删 算法/自然增长 搜索法)
10^{-5}	1.500 / 5.656	12 / 63
10^{-6}	5.063 / 16.172	14 / 88
10^{-7}	7.079 / 39.188	14 / 115
10^{-8}	20.734 / 72.719	16 / 146
10^{-9}	21.469 / 136.062	16 / 183
10^{-10}	25.422 / 237.594	16 / 220

4 结 论

本文构建了 MISO 多元广义多项式神经网络。对于该神经网络参数求解方法及网络结构的确定算法等方面进行了数学上的推导。仿真实验显示本文神经网络具有良好的去除噪声性能,且相比 RBF 神经网络具有训练速度快、逼近精度高等优势。

参考文献:

- [1] 张雨浓,刘巍,易称福,等. Legendre 正交基前向神经网络的权值直接确定法[J]. 大连海事大学学报, 2008, 34(1):32-36.
ZHANG Yunong, LIU Wei, YI Chengfu, et al. Weights immediate determination for feed-forward neural network with Legendre orthogonal basis function [J]. Journal of Dalian Maritime University, 2008, 34(1):32-36.
- [2] ZHANG Yunong, WANG Jun. Global exponential stability of recurrent neural networks for synthesizing linear feedback control systems via pole assignment [J]. IEEE Transactions on Neural Networks, 2002, 13(3):633-644.
- [3] 杨国良,吴旷怀. 利用 BP 神经网络反算沥青路面结构层弹性模量的研究[J]. 中山大学学报:自然科学版, 2008, 47(5):44-48.
YANG Guoliang, WU Kuanghuai. Back calculation of layered elastic moduli of asphalt pavement using BP neural network [J]. Acta Scientiarum Naturalium Universitatis Sunyatsen, 2008, 47(5):44-48.

(下转第 56 页)

- Rostellularia procumbens* [J]. Chin J Appl Environ Biol, 2006, 12(2): 170-175.
- [11] 杨鸣华, 孔令义. 露珠杜鹃的化学成分[J]. 药学与临床研究, 2007, 15(3): 199-201.
- YANG M H, KONG L Y. Chemical constituents of *Rhododendron irroratum* Franch [J]. Pharmaceutical and Clinical Research, 2007, 15(3): 199-201.
- [12] 王金龙, 李灵娜, 贺礼东, 等. 牵牛全草的化学成分研究[J]. 天然产物研究与开发, 2007, 19(3): 427-429.
- WANG J L, LI L N, HE L D, et al. Studies on the Chemical Constituents of *Pharbitisnil chois* [J]. Nat Prod Res Dev, 2007, 19: 427-429.
- [13] SUN S, LIU J, KONG L Y, et al. 骨缘当归的化学成分研究[J]. 中国药学杂志, 2006, 41(19): 1455-1457.
- SUN S, LIU J, KONG L Y, et al. Studies on chemical constituents of *Angelica cartilagino-marginatavar foliata* [J]. Chin Pharm J, 2006, 41(19): 1455-1457.
- [14] 金京玉, 刘向前, 陆昌洙, 等. 小前胡中香豆素类成分研究[J]. 天然产物研究与开发, 2006, 18: 7-10.
- JIN J Y, LIU X Q, LU C S, et al. Coumarins from *Angelica czernaevia* Kita Forma Dentate Yook [J]. Nat Prod Res Dev, 2006, 18: 7-10.
-
- (上接第 46 页)
- [4] 张雨浓, 张禹珩, 陈轲, 等. 线性矩阵方程的梯度法神经网络求解及其仿真验证 [J]. 中山大学学报: 自然科学版, 2008, 47(3): 26-32.
- ZHANG Yunong, ZHANG Yuheng, CHEN Ke, et al. Gradient-based neural network for solving linear matrix equations and its MATLAB simulative verification [J]. Acta Scientiarum Naturalium Universitatis Sunyatsen, 2008, 47(3): 26-32.
- [5] RUMELHART David, McCLELLAND James. Parallel distributed processing: Explorations in the microstructure of cognition [M]. Cambridge: MIT Press, 1986.
- [6] YANG Xuhua, DAI Huaping, SUN Youxian. SIMO Fourier neural networks research [C]. Proceedings of IEEE Intelligent Transportation Systems, 2003, 2: 1606-1609.
- [7] 谢宏, 程浩忠, 牛东晓, 等. 前向神经网络的神经元分层逐个线性优化快速学习算法 [J]. 电子学报, 2005, 33(1): 111-114.
- XIE Hong, CHENG Haozhong, NIU Dongxiao, et al. A fast learning algorithm for feedforward neural network based on the layer-by-layer and neuron-by-neuron optimizing procedure [J]. Acta Electronica Sinica, 2005, 33(1): 111-114.
- [8] DUNKL Charles F, YUAN Xu. Orthogonal polynomials of several variables [M]. Cambridge University Press, 2001.
- [9] LIANG Xuebin, TSO Shukit. An improved upper bound on step-size parameters of discrete-time recurrent neural networks for linear inequality and equation system [J]. IEEE Transactions on Circuits and Systems, 2002, 49(5): 695-698.
- [10] LIU Puyin. Mamdani fuzzy system: universal approximator to a class of random processes [J]. IEEE Transactions on Fuzzy Systems, 2002, 10(6): 756-766.
- [11] TOH Karann, YAU Weiyun, JIANG Xudong. A reduced multivariate polynomial model for multimodal biometrics and classifiers fusion [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14(2): 224-233.
- [12] 周永权, 赵斌, 焦李成. 基于泛函网络的多维函数逼近理论及学习算法 [J]. 系统工程与电子技术, 2005, 27(5): 906-909.
- ZHOU Yongquan, ZHAO Bin, JIAO Licheng. Theory and learning algorithm of multi-dimensional function approximation based on functional networks [J]. Systems Engineering and Electronics, 2005, 27(5): 906-909.
- [13] 邹阿金, 成继勋, 沈建中. 基于 Chebyshev 神经网络的衍生算法 [J]. 湘潭矿业学院学报, 2001, 16(1): 76-78.
- ZOU Ajin, CHENG Jixun, SHEN Jianzhong. A deriving arithmetic based on Chebyshev neural network [J]. Journal of Xiangtan Mining University, 2001, 16(1): 76-78.
- [14] ZHANG Yunong, ZHONG Tongke, LI Wei, et al. Growing algorithm of Laguerre orthogonal basis neural network with weights directly determined [C] // In Proc. ICIC08. Berlin: Springer Press, 2008: 60-67.