

文章编号: 1001-0920(2004)01-0036-04

## 基于基函数展开的双隐层过程神经网络及其应用

许少华<sup>1</sup>, 何新贵<sup>2</sup>, 尚福华<sup>1</sup>

(1. 大庆石油学院 计算机科学系, 黑龙江 大庆 163318; 2 北京大学 信息科学技术学院, 北京 100871)

**摘要:** 提出一类基于基函数展开的双隐层过程神经网络模型。过程神经元隐层完成对输入信息过程模式特征的提取和对时间的聚合运算, 非时变一般神经元隐层用于提高网络对系统输入输出之间复杂关系的映射能力。在输入空间中引入一组函数正交基, 将输入函数和网络权函数表示为该组正交基的展开形式, 利用基函数的正交性简化过程神经元聚合运算。以旋转机械故障诊断和油藏开发过程采收率的模拟为例, 验证了模型和算法的有效性。

**关键词:** 过程神经网络; 双隐层; 函数正交基; 学习算法

中图分类号: TP183

文献标识码: A

## Research and application of process neural network with two hidden-layer based on expansion of basis function

XU Shao-hua<sup>1</sup>, HE Xin-gui<sup>2</sup>, SHANG Fu-hua<sup>1</sup>

(1. Department of Computer Science, Daqing Petroleum Institute, Daqing 163318, China; 2 School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China Correspondent: XU Shao-hua, E-mail: xush62@163.com)

**Abstract:** A class of process neural network model with two hidden-layer based on expansion of basis function is brought forward. The hidden layer of process neuron performs the aggregation operation of time, while the hidden layer of generic neuron raises the mapping capability of network to the complex relation between the system input and output. By introducing a group of function orthogonal basis into the input space, the input functions and the network weight functions are expressed in the expansion form. Using the orthogonality of basis function simplifies the aggregation operation of process neuron. The machinery fault diagnosis and process simulation of oil reservoir development show the effectiveness of the algorithm and model.

**Key words:** process neural network; two hidden-layer; function orthogonal basis; learning algorithm

### 1 引言

在实际问题中, 许多系统的输入依赖于时间变化, 一些控制信号的输出既依赖于输入信号的空间聚合, 又与输入信号的时间累计效应密切相关。文献 [1, 2] 提出了过程神经网络理论和模型。过程神经元的结构与传统的 MP 模型<sup>[3]</sup>相似, 不同之处在于其输入和权值可以是时变的, 其聚合运算既有对空

间的多输入聚合, 也有对时间过程的积累。具有过程式输入的神经网络是人工神经网络在时间域上的扩展, 它放宽了传统神经网络模型对输入的同步瞬时限制, 使问题更具一般性。

本文考虑过程神经网络对时间变量的非线性映射能力和计算复杂性, 提出一类基于基函数展开的双隐层过程神经网络模型。该模型为 4 层结构, 由

收稿日期: 2002-10-25; 修回日期: 2003-02-17

基金项目: 黑龙江省自然科学基金资助项目 (F01-20)

作者简介: 许少华 (1962—), 男, 河北邢台人, 教授, 博士, 从事模式识别、神经网络等研究; 何新贵 (1938—), 男, 浙江浦江人, 教授, 博士生导师, 中国工程院院士, 从事模糊逻辑、神经网络等研究

输入层、过程神经元隐层、非时变一般神经元隐层和输出层组成。过程神经元隐层完成对输入信息过程模式特征的提取和对时间的聚合运算; 一般神经元隐层用于提高系统输入输出之间复杂关系的映射能力, 以增强网络的柔韧性和知识存储能力。由于过程神经网络在训练和计算中对时间的聚合运算十分复杂, 在输入空间中引入一组函数正交基, 将输入函数和网络权函数表示为同一组正交基的展开形式, 利用基函数的正交性简化聚合运算过程。文中给出了学习算法, 并以旋转机械故障诊断和油藏开发过程采收率的模拟为例, 验证了模型和算法的有效性。

## 2 过程神经元

过程神经元由加权、聚合和激励运算三部分组成, 其输入输出为与时间有关的函数。聚合运算既包括对空间的多输入聚集, 也包括对时间过程的累积。过程神经元的结构如图 1 所示。

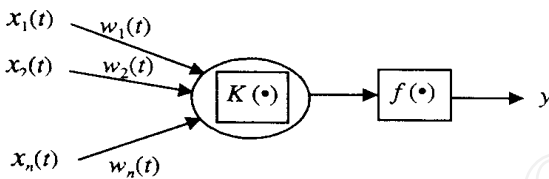


图 1 过程神经元模型

在图 1 中,  $x_1(t), x_2(t), \dots, x_n(t)$  为过程神经元输入函数向量,  $x_i(t) \in C[0, T], [0, T]$  为输入过程区间;  $w_1(t), w_2(t), \dots, w_n(t)$  为相应的权函数;  $K(t)$  为过程神经元的时间聚合基函数;  $f(\bullet)$  为激励函数, 可取线性函数、Sigmoid 函数等。过程神经元的输入与输出之间的关系为

$$y = f((W(t) \oplus X(t)) \otimes K(\bullet) - \theta). \quad (1)$$

其中:  $\theta$  为过程神经元输出阈值,  $y$  为过程神经元的输出值,  $\oplus$  表示某种空间聚合运算,  $\otimes$  表示某种时间(过程)聚合运算。

## 3 双隐层过程神经网络模型

### 3.1 网络拓扑结构

本文提出的双隐层过程神经网络是一类 4 层前馈网络, 拓扑结构为  $n \times m \times K \times 1$ 。输入层有  $n$  个节点, 用于完成  $n$  个时变函数  $x_1(t), x_2(t), \dots, x_n(t)$  对网络的输入; 第 1 隐层由  $m$  个过程神经元节点组成, 完成对  $n$  个输入函数在空间上的加权聚合和对时间过程的聚合运算; 第 2 隐层由  $K$  个非时变一般神经元节点组成, 用于提高网络对系统输入输出之间复杂关系的映射能力; 第 4 层为输出层, 完成系统输出。网络拓扑结构如图 2 所示。

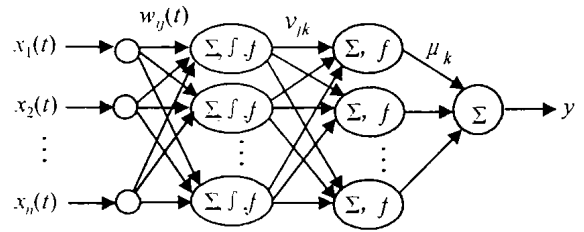


图 2 双隐层过程神经网络拓扑结构

### 3.2 网络各层间的输入输出关系

系统输入

$$X(t) = (x_1(t), x_2(t), \dots, x_n(t)).$$

第 1 隐层输出

$$y_j^{(1)} = f \left[ \int_0^T \left( \sum_{i=1}^n w_{ij}(t) x_i(t) \right) dt - \theta_j^{(1)} \right]. \quad (2)$$

其中:  $y_j^{(1)}$  ( $j = 1, 2, \dots, m$ ) 为第 1 隐层第  $j$  个神经元的输出,  $w_{ij}(t)$  为输入层与第 1 隐层的连接权函数,  $\theta_j^{(1)}$  为第 1 隐层第  $j$  个过程神经元的输出阈值,  $[0, T]$  为系统输入过程区间,  $f$  为第 1 隐层的激励函数。

第 2 隐层输出

$$y_k^{(2)} = g \left[ \sum_{j=1}^m v_{jk} y_j^{(1)} - \theta_k^{(2)} \right]. \quad (3)$$

其中:  $y_k^{(2)}$  ( $k = 1, 2, \dots, K$ ) 为第 2 隐层第  $k$  个神经元的输出,  $v_{jk}$  为第 1 隐层与第 2 隐层的连接权值,  $\theta_k^{(2)}$  为第 2 隐层第  $k$  个神经元的输出阈值,  $g$  为第 2 隐层的激励函数。

输出层: 系统输出为

$$y = \sum_{k=1}^K \mu_k y_k^{(2)}. \quad (4)$$

其中:  $y$  为过程神经网络输出,  $\mu_k$  为第 2 隐层到输出层的连接权值。

在上述模型中, 空间聚合采用了加权求和, 时间聚合采用了对时间的积分。实际应用中可根据情况, 选择其他合适的聚合算子。

综合式 (2) ~ (4), 系统的输入输出关系为

$$y = \sum_{k=1}^K \mu_k g \left( \sum_{j=1}^m v_{jk} f \left( \int_0^T \left( \sum_{i=1}^n w_{ij}(t) x_i(t) \right) dt - \theta_j^{(1)} \right) - \theta_k^{(2)} \right). \quad (5)$$

## 4 基于函数正交基展开的网络模型及学习算法

### 4.1 输入函数和网络权函数的正交基展开

设  $b_1(t), b_2(t), \dots, b_k(t), \dots$  为  $C[0, T]$  中的一组标准正交基函数(例如三角基函数、小波基函数等),  $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$  为输入空间中的任意函数。则对于  $\forall \epsilon > 0, \exists L$ , 在给定精度  $\epsilon > 0$  下, 有<sup>[4]</sup>

$$x_i(t) = \sum_{l=1}^L a_{il} b_l(t), \quad i = 1, 2, \dots, n \quad (6)$$

其中  $a_{il}$  为  $x_i(t)$  展开式中相对于基函数  $b_l(t)$  的系数

将权函数  $w_{ij}(t)$  用  $b_1(t), b_2(t), \dots, b_L(t)$  的展开形式表示, 并记

$$w_{ij}(t) = \sum_{l=1}^L w_{ij}^{(l)} b_l(t), \quad (7)$$

其中  $w_{ij}^{(l)}$  ( $i = 1, 2, \dots, n, j = 1, 2, \dots, m$ ) 为相对于  $b_l(t)$  的连接权系数

将式(6)和(7)代入式(5), 并由基函数的正交性, 得

$$y = \sum_{k=1}^K \mu_k g \left[ \sum_{j=1}^m v_{jk} f \left( \sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} b_l(t) \right) \times \left( \sum_{l=1}^L a_{il} b_l(t) \right) dt - \theta^{(1)} - \theta^{(2)} \right] = \sum_{k=1}^K \mu_k g \left[ \sum_{j=1}^m v_{jk} f \left( \sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} a_{il} - \theta^{(1)} \right) - \theta^{(2)} \right]. \quad (8)$$

#### 4.2 学习算法

为便于讨论, 设各层的激励函数均为  $S$  函数, 即

$$f(u) = g(u) = \frac{1}{1 + e^{-u}}$$

给定  $P$  个学习样本函数  $(x_{p1}(t), x_{p2}(t), \dots, x_{pn}(t), d_p), p = 1, 2, \dots, P$ . 其中:  $x_{ij}(t)$  的第 1 个下标表示学习样本序号, 第 2 个下标表示输入函数向量分量序号;  $d_p$  为对应于输入  $(x_{p1}(t), x_{p2}(t), \dots, x_{pn}(t))$  的网络期望输出

网络误差能量函数定义为

$$E = \sum_{p=1}^P (y_p - d_p)^2 = \sum_{p=1}^P \left( \sum_{k=1}^K \mu_k f \left( \sum_{j=1}^m v_{jk} f \left( \sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} \times a_{il}^{(p)} - \theta^{(1)} \right) - \theta^{(2)} \right) - d_p \right)^2. \quad (9)$$

其中  $a_{il}^{(p)}$  为函数  $x_{pi}(t)$  展开式中相对于基函数  $b_l(t)$  的系数

采用梯度下降法, 过程神经网络的权值修正公式为

$$\mu_k = \mu_k + \alpha \Delta \mu_k, \quad (10)$$

$$v_{jk} = v_{jk} + \beta \Delta v_{jk}, \quad (11)$$

$$w_{ij}^{(l)} = w_{ij}^{(l)} + \gamma \Delta w_{ij}^{(l)}, \quad (12)$$

$$\theta^{(1)} = \theta^{(1)} + \eta \Delta \theta^{(1)}, \quad (13)$$

$$\theta^{(2)} = \theta^{(2)} + \lambda \Delta \theta^{(2)}. \quad (14)$$

其中  $\alpha, \beta, \gamma, \eta, \lambda$  为网络学习速度

为表述方便, 记

$$\sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} a_{il}^{(p)} - \theta^{(1)} = u_{jp},$$

$$\sum_{j=1}^m v_{jk} f \left( \sum_{i=1}^n \sum_{l=1}^L w_{ij}^{(l)} a_{il}^{(p)} - \theta^{(1)} \right) - \theta^{(2)} = z_{kp}.$$

则

$$\Delta \mu_k = - \frac{\partial E}{\partial \mu_k} = - 2 \sum_{p=1}^P (f(z_{kp}) - d_p) f(z_{kp}), \quad (15)$$

$$\Delta v_{jk} = - \frac{\partial E}{\partial v_{jk}} = - 2 \sum_{p=1}^P (f(z_{kp}) - d_p) f(z_{kp}) f(u_{jp}), \quad (16)$$

$$\Delta w_{ij}^{(l)} = - \frac{\partial E}{\partial w_{ij}^{(l)}} = - 2 \sum_{p=1}^P \left[ \sum_{k=1}^K (f(z_{kp}) - d_p) f(z_{kp}) f(u_{jp}) a_{il}^{(p)} \right], \quad (17)$$

$$\Delta \theta^{(1)} = - \frac{\partial E}{\partial \theta^{(1)}} = - 2 \sum_{p=1}^P \left[ \sum_{k=1}^K (f(z_{kp}) - d_p) f(z_{kp}) f(u_{jp}) (-1) \right], \quad (18)$$

$$\Delta \theta^{(2)} = - \frac{\partial E}{\partial \theta^{(2)}} = - 2 \sum_{p=1}^P ((f(z_{kp}) - d_p) f(z_{kp})) (-1).$$

其中

$$f(u) = f(u) (1 - f(u)).$$

#### 4.3 学习算法描述

第 1 步: 给定误差精度  $\epsilon$ , 学习累计次数  $s = 0$ , 最大学习次数  $M$ , 选取输入空间标准正交基函数  $b_1(t), b_2(t), \dots, b_L(t)$ ;

第 2 步: 初始化权值和阈值  $w_{ij}^{(l)}, v_{jk}, \mu_k, \theta^{(1)}, \theta^{(2)}$ ;

第 3 步: 由式(9)计算误差函数  $E$ , 如果  $E < \epsilon$  或  $s > M$ , 则转第 5 步;

第 4 步: 按式(10)~(19)修正权值和阈值,  $s + 1$ , 并转第 3 步;

第 5 步: 输出学习结果, 结束

### 5 应用举例

#### 5.1 旋转机械故障诊断

旋转机械是以转子及其他回转设备部件作为工作主体的机械, 是机械的重要组成部分. 故障诊断是保障旋转机械正常运行和维护的一项重要工作<sup>[5]</sup>. 对于旋转机械故障诊断问题, 不同故障类型对



应于不同的时域波形信号, 因此可通过对采样区间内连续信号波形的识别来实现故障诊断。旋转机械运动呈周期性, 故可将机械的一个旋转周期作为采样区间, 将一个周期内的信号连续变化过程作为一个样本。旋转机械的典型故障主要分为偏心、不同轴、磨损加上正常共 4 种情况。典型信号曲线如图 3 所示。

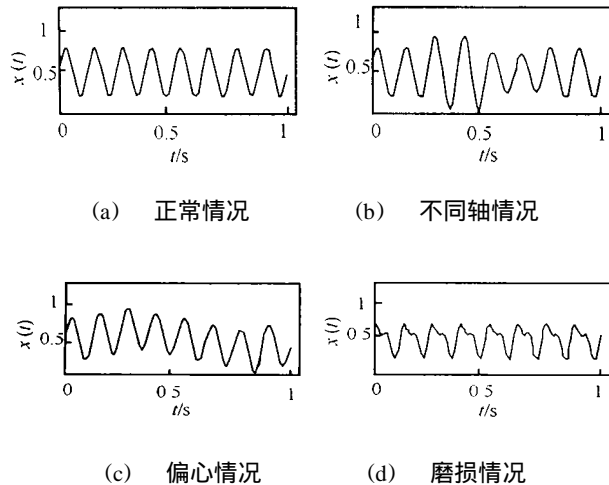


图 3 旋转机械运动 4 种典型曲线

针对上述 4 种情况, 在实际测量信号中, 分别选取了 5 条不同轴、6 条偏心、5 条磨损和 3 条正常共 19 条曲线组成样本训练集。网络结构参数选择如下: 输入节点 1 个, 过程神经元隐层节点 30 个, 非时变一般神经元隐层节点 20 个, 输出节点 1 个; 正交基函数选用三角函数, 基函数为 50 个。学习速率  $\alpha = 0.5$ ,  $\beta = 0.70$ ,  $\gamma = 0.65$ ,  $\eta = 0.55$ ,  $\lambda = 0.55$ ; 最大学习次数  $M = 30\,000$ , 学习精度  $\epsilon = 0.05$ , 实际网络学习 5 337 次后收敛。对 11 个测试样本进行识别, 判对 10 个。这在旋转机械故障自动诊断问题中是一个较好的结果。

## 5.2 油藏开发过程模拟

油藏采收率是衡量油田开发水平和经济效益的重要标志。开发方式不同, 采收率也不同。模拟地层条件下的驱替采油实验是制定油田合理开发方案的重要依据。实验过程分为 3 个阶段: 原始含油饱和度下的注水驱油, 注水驱油效果变差后的注聚合物驱油以及最后阶段的注水驱油。注采系统输入变量有两个: 驱替物(指水或聚合物)注入体积和注入压力增量, 这两个变量都是随时间变化的函数; 系统输出为采收率。

选取 11 块体积尺寸、岩性、物性、含油性十分接近的人造岩心样品, 在不同的注入速度、注入压力增量下(即不同的开发方式)进行实验, 采样间隔为 30

min。整个实验共得到 11 组记录, 其中 1# 岩心样品实验数据列于表 1。

表 1 1# 岩心样品实验记录

序号	注入体积 /PV	注入压力增量 /MPa
1	0.15	0.015
2	0.30	0.035
3	0.45	0.065
4	0.86	0.115
5	1.20	0.065
6	1.86	0.030
7	1.93	0.015
8	1.99	0.055
9	2.06	0.090
10	2.12	0.140
11	2.16	0.185
12	2.25	0.230
13	2.36	0.165
14	2.43	0.095
15	2.50	0.035
16	2.57	0.025
17	2.63	0.025

采用快速傅立叶变换将离散实验数据拟合为  $[0, 8.5]$  区间上的连续函数, 以其中 8 块岩心样品实验结果组成训练样本集, 3 块岩心样品组成测试集。网络结构参数选择如下: 输入节点 2 个, 过程神经元隐层节点 40 个, 一般神经元隐层节点 30 个, 输出节点 1 个; 正交基函数选择三角函数, 基函数为 50 个。学习速率  $\alpha = 0.60$ ,  $\beta = 0.50$ ,  $\gamma = 0.65$ ,  $\eta = 0.75$ ,  $\lambda = 0.70$ ; 最大学习次数  $M = 30\,000$ , 学习精度  $\epsilon = 0.05$ , 网络学习 3 765 次后收敛。对测试集样本进行采收率预测, 结果列于表 2。预测结果可以满足实际问题分析的需要。

表 2 测试样本预测结果

测试集样本序号	9#	10#	11#
实验采收率 /%	52.23	56.11	50.02
预测采收率 /%	50.11	58.09	48.01
绝对误差	2.12	1.98	2.01
相对误差 /%	4.06	3.53	4.01

## 6 结 语

本文提出的基于函数正交基展开的双隐层过程神经网络模型和学习算法, 提高了网络对系统输入输出之间复杂关系的映射能力, 较好地解决了过程神经网络在训练过程中对时间和空间开销较大的问题, 简化了聚合运算过程。基于函数正交基展开的学习算法对于过程神经网络的训练问题具有一般性, 可提高网络学习效率和解决实际问题的适应性。

(下转第 48 页)

表中数据的水平,但由于变异的作用,进一步下降的趋势被抑制 另一方面,由于OSC独立地重组各基因座上的等位基因,在一定程度上破坏了优良积木块的作用,从而难以产生有效的搜索

## 5 结 论

本文分析了广义交叉算子的遗传漂移现象,证明了SDC,ADC,USC等算子不会导致基因级的遗传漂移,但OSC可引起强烈的遗传漂移,其程度随重组规模的增加而加剧 仿真优化实验表明,OSC的遗传漂移不能提高局部搜索速度,仅产生降低种群多样性的消极效果

提高优良个体的选择概率有可能抵消OSC所引起的基因漂移,但增加了选择自身的遗传漂移 OSC对独立的基因座进行重组,不同基因座的多数基因频率相互独立,对每个基因座采用不同的变异概率并自适应调整变异概率,可能是克服基因漂移的有效方法,但这种措施并不能减少OSC对优良积木块的破坏性 因此,OSC所引起的遗传漂移和破坏力对进化搜索的不利影响是难以克服的

## 参考文献(References):

- [1] Rechenberg I. *Evolutionsstrategie 94* [M]. Stuttgart: Frommann-Holzboog, 1994

- [2] Beyer H G. Toward a theory of evolution strategies: On the benefits of sex — The  $(\mu/\mu, \lambda)$  theory [J]. *Evolutionary Computation*, 1995, 3(1): 81-111.
- [3] 章珂, 刘贵忠. 交叉位置非等概率选取的遗传算法[J]. *信息与控制*, 1997, 26(1): 53-60  
(Zhang Ke, Liu Gui-zhong. Selecting crossover site with unequal probability in genetic algorithms [J]. *Information and Control*, 1997, 26(1): 53-60.)
- [4] 张文修, 梁怡. *遗传算法的数学基础* [M]. 西安: 西安交通大学出版社, 2000
- [5] Mahfoud S W. Genetic drift in sharing methods [A]. *Proc of the First IEEE Conf on Evolutionary Computation* [C]. Piscataway: IEEE Service Center, 1994 67-72
- [6] Schmitt L M. Theory of genetic algorithms [J]. *Theoretical Computer Science*, 2001, 259(1-2): 1-61.
- [7] Eiben A E. Multiparent recombination [A]. *Evolutionary Computation 1: Basic Algorithms and Operators* [C]. New York: Institute of Physics Publishing, 2000 289-307.
- [8] Back T. *Evolutionary Algorithms in Theory and Practice* [M]. New York: Oxford University Press, 1996

(上接第39页)

## 参考文献(References):

- [1] He Xin-gui, Liang Jiu-zhen. Process neural network [A]. *Proc of Conf on Intelligent Information Processing* [C]. Beijing, 2000 143-146
- [2] 何新贵, 梁久祯, 许少华. 过程神经网络的训练及其应用[J]. *中国工程科学*, 2001, 3(4): 31-35  
(He X G, Liang J Z, Xu S H. Learning for process neural networks and its applications [J]. *Engineering*

*Science*, 2001, 3(4): 31-35.)

- [3] McCulloch W S, Pitts W H. A logical calculus of the ideas immanent in neuron activity [J]. *Bulletin Mathematical Biophysics*, 1943, 5(1): 115-133
- [4] 柳重堪. *正交函数及其应用* [M]. 北京: 国防工业出版社, 1982 7-16
- [5] 韩捷, 张瑞林. *旋转机械故障机理及诊断技术* [M]. 北京: 机械工业出版社, 1997. 91-114

(上接第43页)

为在减少计算量的同时,保证机器人实际行走过程中障碍检测的准确性,根据地面坐标系与摄像机坐标系之间相对位姿的不断变化,系统加入了在线实时校正映射变换矩阵模块和空间障碍物投影图模块,减少了外部参数和系统内部参数变化对障碍检测的影响,增加了系统运行的可靠性

## 参考文献(References):

- [1] Chow Y H, Ronald C. Obstacle avoidance of legged robot without 3D reconstruction of the surroundings [R]. Hong Kong: The Chinese University of Hong

Kong, 1999.

- [2] Zheng Y, Jones D G, Billings S A, et al. SW ITCJER: A stereo algorithm for ground plane obstacle detection [J]. *Image and Vision Computing*, 1990, 28(8): 57-62
- [3] Badal S, Ravela S. A practical obstacle detection and avoidance system [A]. *Proc of the Second IEEE Workshop* [C]. 1994 97-104
- [4] 高文, 陈熙霖. *计算机视觉* [M]. 北京: 清华大学出版社, 1999