

输入输出均为时变函数的过程神经网络及应用*

何新贵¹⁺, 许少华²

¹(北京大学 信息科学技术学院 视觉与听觉信息处理国家重点实验室,北京 100871)

²(北京航空航天大学 计算机科学与工程系,北京 100083)

Process Neural Network with Time-Varied Input and Output Functions and Its Applications

HE Xin-Gui¹⁺, XU Shao-Hua²

¹(National Laboratory on Machine Perception, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

²(Department of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

+ Corresponding author: Phn: 86-10-66356595, Fax: 86-10-66356595, E-mail: hexingui@pku.edu.cn

<http://www.cs.pku.edu.cn>

Received 2002-09-13; Accepted 2002-11-12

He XG, Xu SH. Process neural network with time-varied input and output functions and its applications. *Journal of Software*, 2003,14(4):764~769.

Abstract: In order to solve the problems in real systems where inputs and outputs are time-varied continuous functions, a process neural network model based on expansion of base functions is proposed in this paper. In this model, the continuous input-output mapping of the system is realized by nonlinear mapping capability to the time variable of process neural networks. A learning algorithm is also given in this paper. In order to simplify the algorithm, orthogonal functions are selected as base functions, and the effectiveness of the model and the algorithm is proved by simulation of oil reservoir exploitation.

Key words: process neural network; approximation of function; time-varied function; orthogonal base function; learning algorithm

摘要: 为了解决实际系统中输入、输出经常是时变连续函数的问题,提出了一类基于基函数展开的过程神经网络模型.该模型利用过程神经网络所具有的对时间变量的非线性映射能力,实现系统的输入、输出之间的连续映射关系.另外,还给出了一种学习算法.为了简化计算,选择正交函数作为基函数,并以油藏开发仿真为例,验证了模型和算法的有效性.

关键词: 过程神经网络;函数逼近;时变函数;正交基函数;学习算法

中图法分类号: TP18 文献标识码: A

* Supported by the Natural Science Foundation of Heilongjiang Province of China under Grant No.F01-20 (黑龙江省自然科学基金)

第一作者简介: 何新贵(1938—),男,浙江浦江人,教授,博士生导师,中国工程院院士,主要研究领域为模糊逻辑,神经网络,进化计算,数据库理论.

在现实世界中,许多系统的输入是依赖于时间变化的连续函数,而其输出值既依赖于输入函数的空间聚合,又与时间的累计效应密切相关.在文献[1,2]中针对此类问题提出了过程神经网络的概念,把这种神经网络从数学上解释为一个泛函,解决了过程神经网络的连续性、可逼近性、计算能力等理论问题,并给出了学习算法和初步的应用.过程神经元的结构与传统神经元^[3]的不同之处在于,过程神经元的输入和权都可以是时变的函数(或称过程).过程神经网络是人工神经网络在时间域上的一种扩展,从而使问题更为一般化.对于求解与过程有关的众多实际问题有较好的适应性,拓宽了神经网络的应用领域.

在系统控制和过程仿真等问题中,我们经常遇到许多系统的输入、输出均为与时间有关的函数.要解决这类问题,传统方法一般需要建立和求解较为复杂的数学模型或建立经验统计公式.但这些问题往往是影响因素较多的非线性系统,传统方法建模困难、精度低、适应性差且求解难度大.在文献[1]中曾提出过一类基于基函数展开的过程神经网络模型.该模型利用过程神经网络所具有的对时变输入的非线性映射能力,能够实现系统的连续输入、输出映射关系.本文将在此基础上提出一种输入、输出均为时变函数的过程神经网络.

由于过程神经网络包括对时间的聚合运算,因此在学习和运算过程中,计算较为复杂.为了解决这个问题,基函数不妨选择标准正交基函数.应用情况表明,这种算法简化了过程神经网络的计算,同时还可以提高网络学习的稳定性和收敛率.本文给出了具体的学习算法,并以石油开采三次采油过程仿真为例,证明了该模型和算法的有效性.

1 过程神经网络模型

设过程神经网络的输入为 $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$, 期望输出为 $y(t)$, $x_i(t), y(t) \in C[0, T]$. $b_1(t), b_2(t), \dots, b_L(t), \dots$ 为 $C[0, T]$ 空间的一组基函数,则只要基函数选择合适, $y(t)$ 就可以在该基函数下展开,并可按任意给定的精度进行有限项展开^[4].设其展开式为

$$y(t) = \sum_{l=1}^L c_l b_l(t). \tag{1}$$

1.1 网络拓扑结构

我们提出的用于函数过程逼近问题的神经网络是一类 4 层前馈过程神经网络,拓扑结构为 $n - m - L - 1$. 输入层有 n 个节点,用于完成 n 个时变函数 $x_1(t), x_2(t), \dots, x_n(t)$ 向网络的输入;第 1 隐层为过程神经元隐层,由 m 个节点单元组成,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,以及对样本过程模式特征的提取;第 2 隐层为非时变一般神经元隐层,有 L 个节点单元;第 4 层为输出层,输出为 $\sum_{l=1}^L \mu_l b_l(t)$. 网络拓扑结构如图 1 所示.

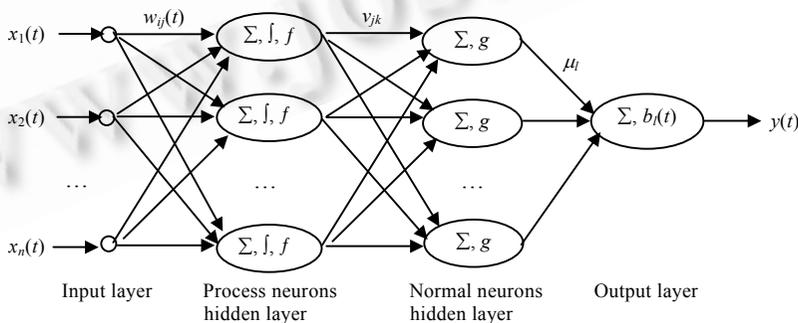


Fig.1 Process neural network
图 1 过程神经网络

1.2 网络各层之间的输入、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:

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

在式(2)中, $y_j^{(1)}$ 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 个神经元的输出阈值, $[0, T]$ 为时间采样区间, f 为第 1 隐层的激励函数.

第 2 隐层输出:

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

在式(3)中, $y_k^{(2)}$ 为第 2 隐层第 k 个神经元的输出, v_{jk} 为第 1 隐层神经元 j 到第 2 隐层神经元 k 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:

设基函数为 $b_l(t)$, $l=1,2,\dots,L$, 则系统输出为

$$y(t) = \sum_{l=1}^L y_l^{(2)} b_l(t). \quad (4)$$

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. 其实, 我们可以根据实际情况选择其他合适的聚合算子. 因此, 上述模型其实是给出了一类过程神经网络, 可以应用于各种不同场合.

2 学习算法

网络学习分成两个阶段: 第 1 阶段为预处理阶段, 选用一种合适的展开方法, 根据需要的精度, 把样本中的输出函数按基函数展开, 从而把每个输出函数变换为一组展开式系数; 第 2 阶段是进行过程神经网络的多输出学习阶段.

2.1 预处理阶段

根据不同的基函数, 要把一个函数按基函数展开已有很多可供选择的方法. 特别是具有周期性的函数, 已有不少快速算法可供选择, 这里不再赘述. 此外, 采用各种小波基作为基函数自然是很值得推荐的.

2.2 学习阶段

给定 K 个学习样本函数:

$$(x_{11}(t), x_{12}(t), \dots, x_{1n}(t), y_1(t)),$$

$$(x_{21}(t), x_{22}(t), \dots, x_{2n}(t), y_2(t)),$$

...

$$(x_{K1}(t), x_{K2}(t), \dots, x_{Kn}(t), y_K(t)),$$

其中, $x_{ij}(t)$ 的第 1 下标表示学习样本序号, 第 2 下标表示输入函数向量分量序号. $y_k(t)$ 为对应输入 $x_{k1}(t)$, $x_{k2}(t)$, ..., $x_{kn}(t)$ 网络的期望输出函数.

在预处理阶段, $y_k(t)$ 已被基函数 $b_l(t)$ ($l=1,2,\dots,L$) 展开, 设相应展开式的系数向量为 $C_k = (c_{k1}, c_{k2}, \dots, c_{kL})$.

首先, 将权函数 $w_{ij}(t)$ 表示为基函数展开形式:

$$w_{ij}(t) = \sum_{l=1}^L w_{ij}^{(l)} b_l(t). \quad (5)$$

其中 $w_{ij}^{(l)}$ 为相对于 $b_l(t)$ 的输入层与第 1 隐层的连接权值.

由式(2)可得, 基于基函数展开的过程神经网络第 1 隐层输出为

$$y_j^{(1)} = f \left(\sum_{i=1}^n \int_0^T \left(\sum_{l=1}^L w_{ij}^{(l)} b_l(t) x_i(t) \right) dt - \theta_j^{(1)} \right), \quad j = 1, 2, \dots, m. \quad (6)$$

由式(3)可得,第2隐层的输出为

$$y_l^{(2)} = g \left(\sum_{j=1}^m v_{jl} f \left(\sum_{i=1}^n \int_0^T \left(\sum_{l=1}^L w_{ij}^{(l)} b_l(t) \right) x_i(t) \right) dt - \theta_j^{(1)} \right) - \theta_l^{(2)}, \quad l = 1, 2, \dots, L.$$

设对应第 k 个学习样本输入的第2隐层第 l 个神经元的实际输出为 $y_{kl}^{(2)}$,期望输出为 c_{kl} , $l = 1, 2, \dots, L$,则网络误差函数可定义为

$$E = \sum_{k=1}^K \sum_{l=1}^L (y_{kl}^{(2)} - c_{kl})^2 = \sum_{k=1}^K \left[\sum_{l=1}^L \left(g \left(\sum_{j=1}^m v_{jl} f \left(\sum_{i=1}^n \int_0^T \left(\sum_{l=1}^L w_{ij}^{(l)} b_l(t) x_{ki}(t) \right) dt - \theta_j^{(1)} \right) - \theta_l^{(2)} \right) - c_{kl} \right)^2 \right]. \quad (7)$$

按照梯度下降算法,网络权值学习规则如下:

$$v_{jl} = v_{jl} + \alpha \Delta v_{jl}, \quad j=1, 2, \dots, m; l=1, 2, \dots, L, \quad (8)$$

$$w_{ij}^{(l)} = w_{ij}^{(l)} + \beta \Delta w_{ij}^{(l)}, \quad i=1, 2, \dots, n; j=1, 2, \dots, m; l=1, 2, \dots, L, \quad (9)$$

$$\theta_j^{(1)} = \theta_j^{(1)} + \gamma \Delta \theta_j^{(1)}, \quad j=1, 2, \dots, m, \quad (10)$$

$$\theta_l^{(2)} = \theta_l^{(2)} + \mu \Delta \theta_l^{(2)}, \quad l=1, 2, \dots, L, \quad (11)$$

其中 $\alpha, \beta, \gamma, \mu$ 为学习速度.

为了表述方便,记:

$$\sum_{i=1}^n \int_0^T \left(\sum_{l=1}^L w_{ij}^{(l)} b_l(t) \right) x_{ki}(t) dt - \theta_j^{(1)} = u_{kj}, \quad (12)$$

$$\sum_{j=1}^m v_{jl} f \left(\sum_{i=1}^n \int_0^T \left(\sum_{l=1}^L w_{ij}^{(l)} b_l(t) \right) x_{ki}(t) dt - \theta_j^{(1)} \right) - \theta_l^{(2)} = z_{kl}, \quad (13)$$

则

$$\Delta v_{jl} = -\frac{\partial E}{\partial v_{jl}} = -2 \sum_{k=1}^K \left((g(z_{kl}) - c_{kl}) g'(z_{kl}) f(u_{kj}) \right), \quad (14)$$

$$\Delta w_{ij}^{(l)} = -\frac{\partial E}{\partial w_{ij}^{(l)}} = -2 \sum_{k=1}^K \left((g(z_{kl}) - c_{kl}) g'(z_{kl}) f'(u_{kj}) \int_0^T b_l(t) x_{ki}(t) dt \right), \quad (15)$$

$$\Delta \theta_j^{(1)} = -\frac{\partial E}{\partial \theta_j^{(1)}} = -2 \sum_{k=1}^K \left(\sum_{l=1}^L (g(z_{kl}) - c_{kl}) g'(z_{kl}) f'(u_{kl}) (-1) \right), \quad (16)$$

$$\Delta \theta_l^{(2)} = -\frac{\partial E}{\partial \theta_l^{(2)}} = -2 \sum_{k=1}^K (g(z_{kl}) - c_{kl}) g'(z_{kl}) (-1). \quad (17)$$

这里,如果网络各层神经元的激励函数均取为 S 函数,即 $f(u) = g(u) = \frac{1}{1+e^{-u}}$,则 $f'(u) = f(u)(1-f(u))$.

学习算法步骤如下:

步骤 1. 给定误差精度 $\varepsilon > 0$; 累计学习迭代次数 $s=0$; 学习最大迭代次数为 M ; 选取标准正交基函数 $b_1(t), b_2(t), \dots, b_L(t)$.

步骤 2. 初始化权值和阈值 $w_{ij}^{(l)}, v_{jl}, \theta_j^{(1)}, \theta_l^{(2)}$ ($i=1, 2, \dots, n; j=1, 2, \dots, m; l=1, 2, \dots, L$).

步骤 3. 由式(7)计算误差函数 E , 如果 $E < \varepsilon$ 或者 $s > M$, 转步骤 5; 否则转步骤 4.

步骤 4. 按照式(8)~(17)修正权值和阈值; $s+1 \rightarrow s$; 转步骤 3.

步骤 5. 输出结果; 结束.

3 应用实例

例:油藏开发三次采油过程仿真.

油藏采出程度是衡量一个油田开发水平的重要指标.油藏开发过程主要分为 4 个阶段,即依靠原始地层能量的自然开采(一次采油)、注水驱油开采(二次采油)、注聚合物开采(三次采油)以及再次注水驱油开采.实验室仿真实验是确定注采方式与采出程度之间关系的重要依据.实验室岩心驱替实验分为 3 个阶段:原始含油饱和度下的注水驱油、注水驱油不见效后的注聚合物驱油以及最后阶段的注水驱油.注采系统有两个输入变量:驱替物(指注入的水或聚合物)的注入体积和注入压力增量,系统输出为采出程度.这 3 个变量都是随时间变化的函数.

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(即不同的开发方式)下进行实验,采样间隔为 0.5 小时,整个实验共得到 11 组记录.其中第 1 号岩心样品实验的真实记录数据见表 1.

Table 1 Experimental records of No.1 core sample
表 1 第 1 号岩心样品实验记录

Number	Time (h)	Injected volume (PV)	Injected pressure increment (Mpa)	Oil recovery (%)
1	0.5	0.15	0.015	19.75
2	1.0	0.30	0.035	26.30
3	1.5	0.45	0.065	30.61
4	2.0	0.86	0.115	36.00
5	2.5	1.20	0.065	37.88
6	3.0	1.86	0.030	37.88
7	3.5	1.93	0.015	37.88
8	4.0	1.99	0.055	38.42
9	4.5	2.06	0.090	41.02
10	5.0	2.12	0.140	45.77
11	5.5	2.16	0.185	48.91
12	6.0	2.25	0.230	52.77
13	6.5	2.36	0.165	57.25
14	7.0	2.43	0.095	57.79
15	7.5	2.50	0.035	57.88
16	8.0	2.57	0.025	57.88
17	8.5	2.63	0.025	57.88

以其中 8 块岩心样品组成训练集,3 块岩心样品组成测试集.网络结构参数选择如下:网络正交基函数选择三角函数,基函数个数为 50 个;2 个输入节点,30 个过程神经元隐层节点,50 个非时变神经元隐层节点,1 个输出节点.学习速度 $\sigma=0.5$, $\beta=0.65$, $\gamma=0.6$, $\mu=0.7$,最大学习次数 $N=30000$,学习精度 $\varepsilon=0.05$.网络模型中各神经元的激励函数均取为 S 函数,即

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

实际网络学习 7513 轮后收敛.对测试集样本进行采出程度预测,其中一块岩心样品预测结果如图 2 所示.在图 2 中,实线表示岩心实际测量值曲线,虚线为网络输出函数曲线,可见两者非常接近.

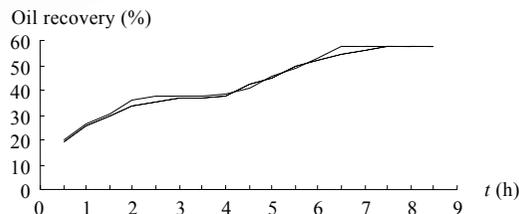


Fig.2 Oil recovery prediction curve
图 2 采出程度预测曲线

3 块岩心样品在采样点处的采出程度实验测量值与过程神经网络计算值平均绝对误差为 2.71, 平均相对误差为 5.63%,拟合精度能够满足实际问题分析的要求.因此,采用过程神经网络建模,可以较大地减少仿真实验

次数,预测出各种不同开发方式下采收率的变化情况.这样,不但可以大量节约实验经费,而且与根据实验数据由人来进行预测相比,可以提高预测效率和精度.

本文提出的输入、输出均为时变过程的过程神经网络模型应用很广,很多实际问题都可以用该模型来描述,具有较大的实用价值.另一方面,我们也认为,关于该模型的一些理论问题,如连续性、可逼近性^[5]、泛化误差估计以及更好的学习算法等问题,都需要进一步进行研究.关于本文提出的学习算法在各种实际问题中应用的有效性以及各种方法的比较等,也还可以做进一步的工作,限于篇幅,这里不再赘述.

References:

- [1] He XG, Liang JZ. Some theoretical issues on process neural networks. *Engineering Science*, 2000,2(12):40~44 (in Chinese with English abstract).
- [2] He XG, Liang JZ. Process neural network. In: Shi ZZ, Faltings B, Musen M, eds. *The 16th World Computer Congress 2000, Proceedings of the Conference on Intelligent Information Processing*. Beijing: Publishing House of Electronics Industry, 2000, 143~146.
- [3] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in neuron activity. *Bulletin Mathematical Biophysics*, 1943,5(1): 115~133.
- [4] Liu CK. *Orthogonal Functions and Applications*. Beijing: National Defence Industry Press, 1982. 7~16 (in Chinese).
- [5] Hornik K, Stinchcombe M, White H. Multi-Layer feed-forward networks are universal approximators. *Neural Networks*, 1989, 2(3):359~366.

附中文参考文献:

- [1] 何新贵,梁久祯.过程神经网络的若干理论问题. *中国工程科学*, 2000,2(12):40~44.
- [4] 柳重堪. *正交函数及其应用*.北京:国防工业出版社,1982.7~16.

中国计算机学会 2003 年部分学术活动计划 (国际部分)

会议名称及内容	地点	时间	主办单位	联系人及地址
第 4 届 Web 时代信息管理 (WAIM'2003) 国际会议	成都	8 月 17 日~19 日	数据库专委, ACM SIGMOD	唐常杰, (610064) 四川大学计算机学院; Tel: 028-85412404, E-mail: chjtang2002@sohu.com; 大会网址: 主节点(美国) http://www.cs.wright.edu/~waim03/ 副节点(成都) http://www.scu.edu.cn/waim03/index.htm
第 5 届高级并行处理技术国际会议 (APPT2003)	厦门	9 月 17 日~19 日	体系结构专委, 国防科学技术大学, 哈尔滨工业大学, 厦门软件园	徐明, (410073) 国防科学技术大学计算机学院训练部; Tel: 0731-4573613, 4533819; E-mail: xuming64@cs.hn.cn
International Conference on CAD & CG (CAD/CG 2003)	香港/澳门	10 月底	中国计算机学会, 国家自然科学基金委, 澳门基金会	唐泽圣, Macau University of Science and Technology, Avenida Wai Long, Taipa, Macau; Tel: +853-8972189 传真: +853-880022; E-mail: ztang@must.edu.mo
2003 国际计算机网络及可移动计算会议	上海	10 月 20 日~23 日	中国计算机学会, IEEE CS	
The IEEE Twelfth Asian Test Symposium (ATS2003)	西安	11 月 17 日~19 日	IEEE Computer Society, Test Technology Technical Council	李晓维, Tel: 010-62654541, E-mail: ats03@ict.ac.cn; http://ats03.ict.ac.cn
4th IEEE Workshop on RTL and High Level Testing (WRTL'03)	西安	11 月 20 日~21 日	IEEE Computer Society, Test Technology Technical Council	李华伟, Tel: 010-62565533-5673, E-mail: wrtlt03@ict.ac.cn; http://wrtlt03.ict.ac.cn