

文章编号: 1001-0920(2007)12-1425-04

一种时变输入输出过程神经网络及学习算法研究

许少华^{1,2}, 何新贵², 王 兵¹

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

摘 要: 针对输入/输出均为时变函数的非线性系统建模问题, 提出一种时变输入输出过程神经网络模型, 并给出了具体的学习算法. 过程神经网络的输入、输出均可为时变函数, 其空间、时间聚合算子分别取为空间加权求和及含时间变参积分, 聚合运算和激励能同时反映时变输入信号的空间聚合作用和输入过程中的阶段时间累积效应. 仿真实验结果验证了所提出模型和算法的有效性.

关键词: 时变输入输出系统; 过程神经网络; 聚合算子; 学习算法

中图分类号: TP183

文献标识码: A

Process neural networks with time-varying inputs and outputs and learning algorithm

XU Shao-hua^{1,2}, HE Xin-gui², WANG Bing¹

(1. School of Computer and Information Technology, Daqing Petroleum Institute, Daqing 163318, China; 2. School of Information Science and Technology, Peking University, Beijing 100871, China. Correspondent: XU Shao-hua, E-mail: xush6263@163.com)

Abstract: Aiming at the nonlinear system modeling problem that input and output are time-varying functions, a process neural network model with time-varying inputs and outputs is proposed, and the concrete learning algorithm is given. The inputs and outputs of the proposed process neural network are all time-varying functions. Its spatial-temporal aggregation operators are adopted as space weighted summation and integral depending on time parameter separately. Aggregation operation and activation can reflect the space aggregation function of the time-varying input signals and the stage time additive effect in the input process at the same time. The simulation experiment results show the effectiveness of the proposed model and algorithm.

Key words: System with time-varying inputs and outputs; Process neural networks; Aggregation operator; Learning algorithm

1 引 言

在非线性系统辨识和过程控制、多维过程信号合成、动态系统仿真等实际问题中, 许多系统的输入/输出均为随时间变化的函数, 系统输入/输出之间存在某种实时或若干时间单元延迟的对应关系. 对于这类时变系统的信息处理问题, 笔者在过程神经元基本模型^[1]的基础上, 对其时间聚合运算机制进行改进, 建立了一种输入输出均为时变函数的过程神经元和过程神经网络模型. 时变输入输出过程神经元的空间聚合算子仍为多输入时变信号的空间加权求和运算, 而时间聚合算子取为含时间变参积分或其他时间累积运算, 聚合运算和激励能同时

反映外部时变输入信号的空间聚合作用和输入过程中的阶段时间累积效应. 时变输入输出过程神经网络按照一定的结构关系和信息传递流程, 可构成输入输出均为时变函数的过程神经网络, 实现系统输入/输出之间的非线性实时映射关系.

时变输入输出过程神经网络的输入、输出和连接权均可为时变函数, 过程神经元增加了一个对于阶段时间效应的累积算子, 这使得这种网络模型与一般神经网络相比, 信息变换机制和学习算法有很大的不同. 作者针对网络输出为时变函数的情况, 构造了一种基于广义距离泛数的训练误差函数, 提出了一种基于权函数基展开的梯度下降结合数值计

收稿日期: 2006-08-31; 修回日期: 2006-12-25.

基金项目: 国家自然科学基金项目(60373102, 60473051); 教育部博士点基金项目(20030001701).

作者简介: 许少华(1962—), 男, 河北邢台人, 教授, 博士后, 从事模式识别、信息处理等研究; 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、进化计算等研究.

算的学习算法,采用时间粒度分割方法给出了具体的算法步骤,并以化工生产过程模拟问题为例验证了模型和算法的有效性.

2 过程神经元

2.1 过程神经元基本模型^[1]

过程神经元由时变输入信号加权,时间、空间二维聚合和阈值激励输出等运算组成,其结构如图 1 所示.图中: $x_1(t), x_2(t), \dots, x_n(t)$ 为过程神经元的时变输入函数; $w_1(t), w_2(t), \dots, w_n(t)$ 为相应的连接权函数; \oplus 为过程神经元的空间聚合算子, \otimes 为时间累积算子; $K(t)$ 为时间聚合核函数, $f(\cdot)$ 为连续的激励函数.

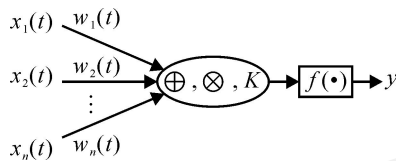


图 1 过程神经元

过程神经元输入与输出之间的映射关系为

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

如果 \oplus 取为空间加权和, \otimes 取对于时间的积分,且 $K(\cdot) = 1$,则式(1)可改写为

$$y = f\left(\int_0^T \left(\sum_{i=1}^n w_i(t) x_i(t)\right) dt - \right), \quad (2)$$

式中 $-$ 为过程神经元阈值.

2.2 时变输入输出过程神经元

时变输入输出过程神经元的输入、输出、连接权以及阈值均可时间为函数,其空间聚合算子仍取为多输入信号加权和,而时间累积算子取含时间变参积分,结构如图 2 所示.图中: Σ, \int_t 为过程神经元的空间聚合算子, \int_t 为含时间变参积分算子, $f(\cdot)$ 为激励函数, $y(t)$ 为过程神经元的输出函数.

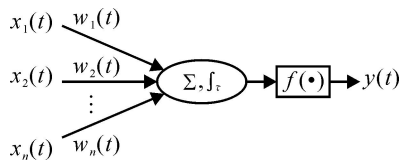


图 2 时变输入输出过程神经元

由图 2 可知,时变输入输出过程神经元输入与输出之间的映射关系为

$$y(t) = f\left(\int_0^t \left(\sum_{i=1}^n w_{ij}(\cdot) x_i(\cdot)\right) d - (t)\right), \quad (3)$$

其中 $- (t)$ 为过程神经元的阈值,也为时间函数.

3 时变输入输出过程神经网络

若干由式(3)定义的过程神经元和其他类型的神经元按照一定的结构关系可构成输入输出均为时变函数的过程神经网络模型.考虑含一个由式(3)

定义的过程神经元隐层、线性输出的时变输入输出过程神经网络模型,网络拓扑结构如图 3 所示.图中: $x_1(t), x_2(t), \dots, x_n(t)$ 为网络的时变输入函数; $w_{ij}(t) (i = 1, 2, \dots, n, j = 1, 2, \dots, m)$ 为输入层节点与隐层节点的连接权函数; $v_j(t) (j = 1, 2, \dots, m)$ 为隐层节点到输出节点的连接权函数; $y(t)$ 为网络输出函数.

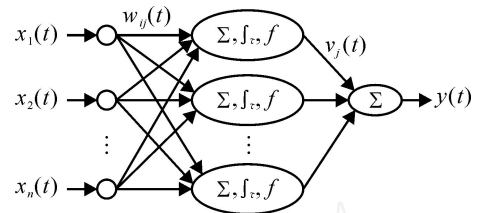


图 3 时变输入输出过程神经网络

由图 3 可知,网络输入输出之间的映射关系为

$$y(t) = \sum_{j=1}^m v_j(t) f\left(\int_0^t \left(\sum_{i=1}^n w_{ij}(\cdot) x_i(\cdot)\right) d - j(t)\right). \quad (4)$$

式中: $t \in [0, T], [0, T]$ 为时变信号输入过程区间; f 为激励函数; $- j(t)$ 为阈值函数.

4 学习算法

设时变输入输出过程神经网络的输入函数空间 $U = C[0, T]^n$; $b_1(t), b_2(t), \dots, b_L(t)$ 为 $C[0, T]$ 中一组满足输入函数拟合精度要求的基函数; $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$ 为 U 中的任意函数.则 $x_i(t)$ 的基函数展开形式可表示为

$$x_i(t) = \sum_{l=1}^L a_{il} b_l(t), \quad i = 1, 2, \dots, n. \quad (5)$$

将式(4)中的权函数 $w_{ij}(t), v_j(t)$ 和阈值函数 $- j(t)$ 也用基函数 $b_1(t), b_2(t), \dots, b_L(t)$ 的展开形式表示,即

$$w_{ij}(t) = \sum_{l=1}^L w_{ij}^{(l)} b_l(t), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m; \quad (6)$$

$$v_j(t) = \sum_{l=1}^L v_j^{(l)} b_l(t), \quad j = 1, 2, \dots, m; \quad (7)$$

$$- j(t) = \sum_{l=1}^L - j^{(l)} b_l(t), \quad j = 1, 2, \dots, m. \quad (8)$$

式中: $w_{ij}^{(l)}, v_j^{(l)}, - j^{(l)}$ 分别为 $w_{ij}(t), v_j(t), - j(t)$ 相对于 $b_l(t)$ 的基函数展开系数.

将式(5) ~ (8)代入(4),整理得

$$y(t) = \sum_{j=1}^m \left(\sum_{l=1}^L v_j^{(l)} b_l(t) \right) f\left(\sum_{i=1}^n \sum_{l=1}^L a_{il} w_{ij}^{(l)} \times \int_0^t \left(\sum_{s=1}^n b_s(\cdot) b_s(\cdot) \right) d - j^{(l)} b_l(t) \right). \quad (9)$$

记 $B_{ks}(t) = \int_0^t b_k(\cdot) b_s(\cdot) d -$,则式(9)可简写为

$$y(t) = \sum_{j=1}^m \left(\sum_{l=1}^L v_j^{(l)} b_l(t) \right) f \left(\sum_{i=1}^n \sum_{l=1}^L \sum_{s=1}^L a_{il} w_{ij}^{(l)} \times B_{ls}(t) - \sum_{l=1}^L \left(\sum_{j=1}^m v_j^{(l)} b_l(t) \right) \right). \quad (10)$$

给定 K 个学习样本 $(x_{k1}(t), x_{k2}(t), \dots, x_{kn}(t), d_k(t))$, $k = 1, 2, \dots, K$, 其中 $d_k(t)$ 为对应输入 $x_{k1}(t), x_{k2}(t), \dots, x_{kn}(t)$ 网络的期望输出函数. 设对应第 k 个样本输入网络的实际输出为 $y_k(t)$, 则网络误差函数定义为

$$E = \sum_{k=1}^K \int_0^T \left(y_k(t) - d_k(t) \right)^2 dt = \sum_{k=1}^K \int_0^T \left(\sum_{j=1}^m \left(\sum_{l=1}^L v_j^{(l)} b_l(t) \right) f \left(\sum_{i=1}^n \sum_{l=1}^L \sum_{s=1}^L a_{il}^{(k)} w_{ij}^{(l)} \times B_{ls}(t) - \sum_{l=1}^L \left(\sum_{j=1}^m v_j^{(l)} b_l(t) \right) - d_k(t) \right)^2 dt \right)^{1/2}, \quad (11)$$

式中 $a_{il}^{(k)}$ 为 $x_{ki}(t)$ 在基函数展开式中对应用于 $b_l(t)$ 的系数.

将过程区间 $[0, T]$ 进行适当的 P 等分, 记区间分割点为 t_0, t_1, \dots, t_p . 将式(5) ~ (8) 另写为

$$x_i(t_p) = \sum_{l=1}^L a_{il} b_l(t_p), \quad (12)$$

$$w_{ij}(t_p) = \sum_{l=1}^L w_{ij}^{(l)} b_l(t_p), \quad (13)$$

$$v_j(t_p) = \sum_{l=1}^L v_j^{(l)} b_l(t_p), \quad (14)$$

$$j(t_p) = \sum_{l=1}^L \left(\sum_{j=1}^m v_j^{(l)} b_l(t_p) \right). \quad (15)$$

将式(12) ~ (15) 代入(11), 有

$$E = \sum_{k=1}^K \sum_{p=0}^P \left(\sum_{j=1}^m \left(\sum_{l=1}^L v_j^{(l)} b_l(t_p) \right) f \left(\sum_{i=1}^n \sum_{l=1}^L \sum_{s=1}^L a_{il}^{(k)} w_{ij}^{(l)} \times B_{ls}(t_p) - \sum_{l=1}^L \left(\sum_{j=1}^m v_j^{(l)} b_l(t_p) \right) - d_k(t_p) \right)^2 \right)^{1/2}, \quad (16)$$

式中 $t = T/P$.

由误差梯度下降算法, 网络权值修正规则为

$$v_j^{(l)} = v_j^{(l)} + \Delta v_j^{(l)}, \quad (17)$$

$$w_{ij}^{(l)} = w_{ij}^{(l)} + \Delta w_{ij}^{(l)}, \quad (18)$$

$$j = j + \Delta j, \quad (19)$$

其中 η 为学习效率.

为表述简便, 记

$$\sum_{i=1}^n \sum_{l=1}^L \sum_{s=1}^L a_{il}^{(k)} w_{ij}^{(l)} B_{ls}(t_p) - \sum_{l=1}^L \left(\sum_{j=1}^m v_j^{(l)} b_l(t_p) \right) = u_{kj}(t_p),$$

$$\sum_{j=1}^m \left(\sum_{l=1}^L v_j^{(l)} b_l(t_p) \right) f \left(\sum_{i=1}^n \sum_{l=1}^L \sum_{s=1}^L a_{il}^{(k)} w_{ij}^{(l)} \times B_{ls}(t_p) - \sum_{l=1}^L \left(\sum_{j=1}^m v_j^{(l)} b_l(t_p) \right) - d_k(t_p) \right) = z_{kp},$$

则

$$\Delta v_j^{(l)} = - \frac{\partial E}{\partial v_j^{(l)}} = \sum_{k=1}^K \sum_{p=0}^P \left(\left(\sum_{p=0}^P (z_{kp})^2 \cdot t \right)^{-1/2} \times \left(z_{kp} \cdot t \cdot b_l(t_p) \cdot f(u_{kj}(t_p)) \right) \right), \quad (20)$$

$$\Delta w_{ij}^{(l)} = - \frac{\partial E}{\partial w_{ij}^{(l)}} = \sum_{k=1}^K \sum_{p=0}^P \left(\left(\sum_{p=0}^P (z_{kp})^2 \cdot t \right)^{-1/2} \cdot \left(\sum_{p=0}^P z_{kp} \cdot t \times \left(\sum_{l=1}^L v_j^{(l)} b_l(t_p) \right) \cdot f(u_{kj}(t_p)) \cdot a_{is}^{(k)} B_{ls}(t_p) \right) \right), \quad (21)$$

$$\Delta j = - \frac{\partial E}{\partial j} = \sum_{k=1}^K \sum_{p=0}^P \left(\left(\sum_{p=0}^P (z_{kp})^2 \cdot t \right)^{-1/2} \cdot \left(\sum_{p=1}^P z_{kp} \cdot t \times \left(\sum_{l=1}^L v_j^{(l)} b_l(t_p) \right) \cdot f(u_{kj}(t_p)) \cdot b_l(t_p) (-1) \right) \right). \quad (22)$$

5 在化工生产过程模拟中的应用

在化工生产中, 分馏塔反应器的过程预报控制对于提高化工产品质量具有重要的意义. 某化工装置反应器的过程输入变量有 4 个: 2FC1302_SP, 2FC1304_SP, 2FC1305_SP, 2FC1306_SP. 其中: 输入变量的 SP 为其 PID 回路的设定值, OP 为控制阀开度. 反应器过程输出变量有 3 个: 2TH1302, 2QH1306, 2QH1304.

测量记录 7 次生产中分馏塔反应器的过程输入值和输出值, 每次采样 361 个, 采样间隔为 1 min, 共获得 7 组采样样本. 其中一次采样部分数据见表 1 中第 1 ~ 7 列.

采用多输入 / 多输出时变过程神经网络进行反应器生产过程模拟, 网络拓扑结构确定为 4-21-3. 选择离散 Walsh 函数系^[21] 作为输入函数的拟合基函数. 对 7 组时变输入信号在精度 0.01 下进行拟合, 确定基函数项数为 256. 以 6 组测量数据组成训练集, 另一组数据为测试集. 网络训练误差精度取 0.05, 学习效率 η , 分别为 0.47, 0.55, 0.60, 网络迭代 10 322 后收敛. 测试样本 3 个输出变量过程预测结果见表 1 中第 8 ~ 10 列. 由表 1 可知, 过程神经网络预测模型达到了较高的模拟精度.

表 1 反应器部分测试数据和预测结果(部分)

2FC1302	2FC1304	2FC1305	2FC1306	2TII302	2QII306	2QII304	2TII302_ REAL	2QII306_ REAL	2QII304_ REAL
49.263 6	11 988.823	12 000.299	7.618 399	194.324 0	4.535 281	1.938 586	194.305 6	4.525 177	1.939 644
49.359 8	11 962.689	11 953.086	7.621 44	194.232 6	4.522 707	1.938 803	194.199 9	4.526 402	1.937 827
48.686 1	11 987.439	11 988.183	7.574 075	194.158 3	4.530 679	1.937 381	194.164 9	4.530 469	1.937 694
48.664 4	12 005.299	12 003.660	7.581 643	194.175 3	4.541 925	1.933 688	194.171 2	4.539 579	1.933 875
48.806 4	12 005.813	11 995.109	7.569 448	194.193 6	4.487 464	1.933 651	194.164 9	4.490 992	1.933 156
48.968 6	12 018.001	12 000.028	7.595 889	194.207 8	4.451 073	1.938 601	194.223 1	4.451 467	1.938 894
49.409 4	12 016.457	11 984.801	7.594 652	194.180 6	4.492 686	1.938 491	194.146 8	4.485 935	1.939 001

6 结 语

本文建立的时变输入输出过程神经网络的聚合运算和激励机制,能同时反映外部时变输入信号的空间聚合作用和输入过程中的阶段时间累积效应.所建立的网络模型对时变信息具有较强的非线性变换能力,可实现动态系统输入/输出之间的非线性实时映射关系,拓宽了人工神经网络^[3,4]的应用领域.

参考文献(References)

- [1] He Xin-gui, Liang Ji-rzhen. Process neural networks [C]. World Computer Congress 2000, Proc of Conf on Intelligent Information Processing. Beijing, 2000: 143-146.
- [2] 柳重堪. 正交函数及其应用[M]. 北京:国防工业出版社, 1982.
(Liu Chong-kan. Orthogonal functions and applications [M]. Beijing: National Defense Industry Press, 1982.)
- [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] Zhang Li I, Tao Huizhong W, Holt C E, et al. A critical window for cooperation and competition among developing retinotectal synapses[J]. Nature, 1998, 395(3): 37-44.
- [6] Charles J Corbett, Xavier de Groote. A supplier's optimal quantity discount policy under asymmetric information[J]. Management Science, 2000, 46(3): 444-450.
- [7] Apostolos Burnetas, Stephen M Gilbert, Craig Smith. Quantity discounts in single period supply contracts with asymmetric demand Information [DB/OL]. (2005-08-13). <http://ssrn.com/abstract/788109>.
- [8] Albert Y Ha. Supplier-buyer contracting: Asymmetric cost information and cutoff level policy for buyer participation[J]. Naval Research Logistics, 2001, 48: 41-64.
- [9] Cachon G, Lariviere M A. Contracting to assure supply: How to share demand forecasts in a supply chain[J]. Management Science, 2001, 47(5): 629-646.
- [10] 唐宏祥, 何建敏, 刘春林. 非对称需求信息条件下的供应链信息共享机制[J]. 系统工程学报, 2004, 19(6): 589-595.
(Tang Hong-xiang, He Jian-min, Liu Chun-lin. Supply chain information sharing mechanism under the condition of asymmetric demand information[J]. J of Systems Engineering, 2004, 19(6): 589-595.)
- [11] 郭琼, 杨德礼. 需求信息不对称下基于期权的供应链协作机制的研究[J]. 计算机集成制造系统, 2006, 12(9): 1466-1471.
(Guo Qiong, Yang De-li. Supply chain coordination mechanism with option under asymmetric demand information[J]. Computer Integrated Manufacturing Systems, 2006, 12(9): 1466-1471.)
- [12] Charles J Corbett, Deming Zhou, Christopher S Tang. Designing supply contracts: Contract type and information asymmetry [J]. Management Science, 2004, 50(4): 550-559.

(上接第 1424 页)