

文章编号: 1001-0920(2006)09-1037-05

基于过程神经网络的时间序列预测及其应用研究

丁刚, 钟诗胜

(哈尔滨工业大学 机电工程学院, 哈尔滨 150001)

摘要: 针对传统时间序列预测方法难以表达时间序列中时间累积效应的缺陷, 提出一种基于过程神经网络的时间序列预测方法。就时间序列的短期预测和长期预测问题分别应用该方法建立了两种预测模型, 并给出了相应的学习算法。以航空发动机状态监控中滑油铁金属含量预测为例验证了两种预测模型及其学习算法的有效性, 并得到了满意的结果。

关键词: 时间序列预测; 过程神经网络; 航空发动机状态监控; 学习算法

中图分类号: TP183 **文献标识码:** A

Time Series Prediction Based on Process Neural Networks and Its Applications

DING Gang, ZHONG Shi-sheng

(School of Mechanical and Electrical Engineering, Harbin Institute of Technology, Harbin 150001, China
Correspondent: DING Gang, E-mail: dingganghit@163.com)

Abstract: To the difficulty of expression of the temporal accumulation in the time series using conventional time series prediction methods, a time series prediction method based on process neural network is proposed. Time series short-term prediction model and long-term prediction model based on the proposed method are developed respectively, and the corresponding learning algorithms are given. The effectiveness of this two models and their learning algorithms are proved by the lubricating oil iron concentration prediction in the aircraft engine condition monitoring, and the test results are satisfactory.

Key words: Time series prediction; Process neural network; Aircraft engine condition monitoring; Learning algorithm

1 引言

时间序列预测问题在工程实践中具有重要意义。传统的时间序列预测方法主要以回归分析为主, 在理论上已十分成熟, 但精度不高, 容错性差。自从多层人工神经网络被证明是连续函数的一致逼近器^[1,2]以来, 应用人工神经网络进行时间序列预测引起了广泛关注。人工神经网络作为一种新的时间序列预测方法, 以其良好的非线性性质、并行分布式的存储结构和高容错性等特点, 在很多实际应用领域中都取得了成功^[3,4]。但受输入的同步瞬时限制, 传统人工神经网络难以表达时间序列中实际存在的时

间累积效应, 同时, 传统人工神经网络难以解决较大样本的学习和泛化问题, 因此传统人工神经网络在解决复杂非线性时间序列预测问题时还存在一定的不适应性。为解决此类问题, 文献[5]提出了一种人工神经网络模型——过程神经网络。它在结构上与传统人工神经网络类似, 不同之处在于过程神经网络的输入和相应的连接权可以是时变函数。过程神经网络由于增加了一个时间聚合运算算子, 能够充分反映时间序列中存在的时间累积效应。

本文在文献[5]的基础上就时间序列的短期预测和长期预测问题分别构建了两种基于过程神经网络

收稿日期: 2005-06-20; 修回日期: 2005-08-22

基金项目: 国家自然科学基金项目(60373102, 60572174)。

作者简介: 丁刚(1976—), 男, 山东青岛人, 博士生, 从事人工神经网络及其应用的研究; 钟诗胜(1964—), 男, 江西龙南人, 教授, 博士生导师, 从事人工智能、人工神经网络等研究。

络的时间序列预测模型,并给出了各自的学习算法最后以航空发动机健康状态监控中滑油铁金属含量的预测为例验证了这两种时间序列预测模型及其学习算法的有效性,并取得了满意的结果

2 过程神经元

过程神经元由加权、聚合和激励3部分组成^[5]。聚合运算既有对空间的多输入聚合,也有对时间过程的累积聚合,从而使过程神经元同时具有时空二维信息处理能力。过程神经元的结构如图1所示

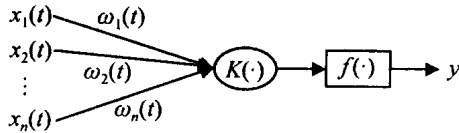


图1 过程神经元结构

过程神经元输出与输入之间的关系为

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

其中: $x_i(t)$ ($i = 1, 2, \dots, n$) 为过程神经元的输入函数; $[0, T]$ 为样本采样区间; $\omega_i(t)$ 为相应的连接权函数; θ 为过程神经元阈值; $f(\cdot)$ 为激活函数, 一般可以取为 Sigmoid 函数

文献[6]中假定 $x_i(t)$ 和 $\omega_i(t)$ 均为 $C[0, T]$ 空间上的函数, 然而在工程实际中, $x_i(t)$ 和 $\omega_i(t)$ 往往是 $L^p[0, T]$ ($p > 0$) 空间上的函数。本文遵循后者

若 $x_i(t) \in L^p[0, T]$, 由 Weierstrass 逼近定理在 $L^p[0, T]$ 空间上的推广^[7]可知, 对 $\forall \epsilon > 0$, 必存在多项式 $P(t)$, 使 $|P(t) - x_i(t)| < \epsilon$ 即 $x_i(t)$ 可由 $C[0, T]$ 空间上的多项式函数系 $P_k(t)$ ($k = 1, 2, \dots, K$) 以任意精度逼近, 从而 $x_i(t)$ 可表示为

$$x_i(t) = \sum_{k=1}^K c_{ik} P_k(t), \quad c_{ik} \in R.$$

设多项式函数系 $P_k(t)$ 在 $C[0, T]$ 空间上是相互独立的, 则可以通过 Gram-Schmidt 正交化步骤^[8]使之正交化, 于是得到 $C[0, T]$ 空间上的一组标准正交基函数 $b_k(t)$, 则 $x_i(t)$ 和 $\omega_i(t)$ 可分别表示为

$$x_i(t) = \sum_{k=1}^K x_{ik} b_k(t),$$

$$\omega_i(t) = \sum_{k=1}^K \omega_{ik} b_k(t).$$

其中 $x_{ik}, \omega_{ik} \in R$ 为展开式系数。根据正交函数的性质有

$$\int_0^T b_k(t) b_l(t) dt = \begin{cases} 1, & k = l; \\ 0, & k \neq l, \end{cases}$$

因此, 式(1)可简化为

$$y = f \left(\sum_{i=1}^n \sum_{k=1}^K \omega_{ik} x_{ik} - \theta \right) \quad (2)$$

3 基于过程神经网络的时间序列短期预测模型

应用传统方法进行时间序列预测在本质上是根据时间序列 $\{x_m\}$ 的历史观测数据 x_m, x_{m-1}, \dots 对 x_{m+h} ($h > 0$) 进行估计, 即认为 x_{m+h} 与其前面的数据 x_m, x_{m-1}, \dots 之间存在某种函数映射关系, 描述为

$$x_{m+h} = G(x_m, x_{m-1}, \dots) \quad (3)$$

此时, 时间序列预测问题就转化为对函数 $G(\cdot)$ 的逼近问题。当 $h = 1$ 时称为一步预测, $h > 1$ 时称为多步预测。显然, 应用传统人工神经网络对 x_{m+h} 进行预测时, x_{m-1}, x_{m-2}, \dots 对 x_m 的影响难以表达, 即难以表达时间序列中实际存在的时间累积效应, 这必然会影响到预测精度。为解决此问题, 本文分别就时间序列的短期预测和长期预测问题构建了两种基于过程神经网络的时间序列预测模型。首先介绍基于过程神经网络的时间序列短期预测模型

将 x_m, x_{m-1}, \dots 进行函数拟合, 产生一个时变函数 $x_m(t)$ 。可以认为时间序列在 $m + 1$ 时刻的取值 x_{m+1} 与函数 $x_m(t)$ 之间存在某种泛函映射关系, 即

$$x_{m+1} = F(x_m(t)) \quad (4)$$

此时, 时间序列短期预测问题转化为对泛函 $F(\cdot)$ 的逼近问题。文献[5]已经证明过程神经网络能够逼近任意连续泛函, 因此, 利用过程神经网络解决时间序列短期预测问题具有较强的适应性

3.1 模型拓扑结构

本文所提出的基于过程神经网络的时间序列短期预测模型是一种3层前向过程神经网络, 其输入层只有1个单元, 隐层由 n 个过程神经元构成, 输出层只有1个单元。模型拓扑结构如图2所示

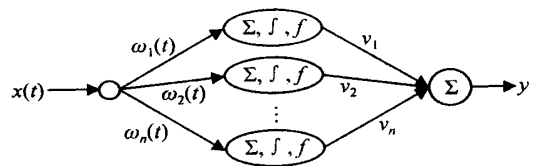


图2 时间序列短期预测模型

令 $x(t) = x_m(t)$, 则基于过程神经网络的时间序列短期预测模型对 x_{m+1} 的预测为

$$y = \sum_{i=1}^n v_i f \left(\int_0^T \omega_i(t) x(t) dt - \theta \right) - \theta \quad (5)$$

由前述可知 $x(t)$ 可展开为 $x(t) = \sum_{k=1}^K a_k b_k(t)$, $a_k \in R$, 相应的 $\omega_i(t)$ 可展开为 $\omega_i(t) = \sum_{k=1}^K \omega_{ik} b_k(t)$, $\omega_{ik} \in R$ 。因此, 式(5)可简化为

$$y = \sum_{i=1}^n v_i f \left(\sum_{k=1}^K \omega_{ik} a_k - \theta \right) - \theta \quad (6)$$

3.2 学习算法

给定 L 组学习样本 $\{x_l(t), d_l\}, l = 1, 2, \dots, L, d_l$ 为对应于样本 $x_l(t)$ 的网络期望输出, 设 y_l 为相应的网络实际输出, 则网络模型的误差函数可定义为

$$E = \frac{1}{2} \sum_{l=1}^L (y_l - d_l)^2 = \frac{1}{2} \sum_{l=1}^L \left(\sum_{i=1}^n v_{if} \left(\sum_{k=1}^K \omega_{kaki} - \theta \right) - \theta - d_l \right)^2 \quad (7)$$

为便于分析, 令 $Z_{il} = \sum_{k=1}^K \omega_{kaki} - \theta$ 根据最速梯度下降算法, 网络模型的待训练参数调整规则为

$$\begin{cases} v_i(s+1) = v_i(s) + \alpha \Delta v_i(s), \\ \omega_k(s+1) = \omega_k(s) + \beta \Delta \omega_k(s), \\ \theta(s+1) = \theta(s) + \gamma \Delta \theta(s), \\ \theta(s+1) = \theta(s) + \lambda \Delta \theta(s). \end{cases} \quad (8)$$

其中: $\alpha, \beta, \gamma, \lambda$ 为学习速率; s 为学习迭代次数

在每一次学习迭代中, $\Delta v_i, \Delta \omega_k, \Delta \theta$ 和 $\Delta \theta$ 可分别表示为

$$\begin{cases} \Delta v_i = - \frac{\partial E}{\partial v_i} = - \sum_{l=1}^L (y_l - d_l) f(Z_{il}), \\ \Delta \omega_k = - \frac{\partial E}{\partial \omega_k} = - \sum_{l=1}^L (y_l - d_l) v_{if}(Z_{il}) a_{ki}, \\ \Delta \theta = - \frac{\partial E}{\partial \theta} = \sum_{l=1}^L (y_l - d_l) v_{if}(Z_{il}), \\ \Delta \theta = - \frac{\partial E}{\partial \theta} = \sum_{l=1}^L (y_l - d_l). \end{cases} \quad (9)$$

学习过程可完整描述如下:

- Step 1: 选取合适的正交基函数将网络模型的输入函数和连接权函数同时展开;
- Step 2: 给定学习误差精度 ϵ , 学习迭代次数 $s = 0$, 最大学习迭代次数 M 和学习速率;
- Step 3: 初始化网络模型待训练参数;
- Step 4: 根据式 (7) 计算误差函数 E , 如果 $E < \epsilon$ 或 $s > M$ 则转至 Step 6, 否则转到 Step 5;
- Step 5: 根据式 (8) 和式 (9) 的规则调整待训练参数, $s + 1 \rightarrow s$, 转至 Step 4;
- Step 6: 输出学习结果, 结束

4 基于过程神经网络的时间序列长期预测模型

当 $h > 1$ 时, 将 $x_{m+1}, x_{m+2}, \dots, x_{m+h}$ 拟合成一个时变函数 $x_{m+h}(t)$, 可以认为 $x_{m+h}(t)$ 与 $x_m(t)$ 之间存在某种算子映射关系, 描述为

$$x_{m+h}(t) = T(x_m(t)). \quad (10)$$

此时, 时间序列长期预测问题转为对算子 $T(\bullet)$ 的逼

近问题 文献 [9, 10] 已证明过程神经网络能逼近任意连续过程, 即过程神经网络能实现对任意连续算子的逼近 因此, 采用输入输出均为时变函数的过程神经网络模型能够实现对时间序列的长期预测

4.1 模型拓扑结构

本文所提出的基于过程神经网络的时间序列长期预测模型是一种 4 层前向过程神经网络 其输入层只有 1 个单元, 第 1 隐层由 n 个过程神经元构成, 第 2 隐层由 K 个传统神经元构成, 输出层只有 1 个单元 模型拓扑结构如图 3 所示

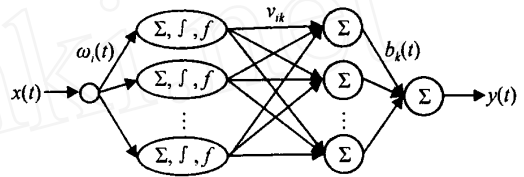


图 3 时间序列长期预测模型

令 $x(t) = x_m(t)$, 则基于过程神经网络的时间序列长期预测模型对 $x_{m+h}(t)$ 的预测可表示为

$$y(t) = \sum_{k=1}^K \left(\sum_{i=1}^n v_{ik} f \left(\int_0^T \omega_k(t) x(t) dt - \theta \right) - \theta \right) b_k(t). \quad (11)$$

由前述可知 $x(t)$ 和 $\omega(t)$ 均可由正交基函数展开, 则式 (11) 可简化为

$$y(t) = \sum_{k=1}^K \left(\sum_{i=1}^n v_{ik} f \left(\sum_{j=1}^K \omega_j a_{ji} - \theta \right) - \theta \right) b_k(t). \quad (12)$$

4.2 学习算法

给定 L 组学习样本 $\{x_l(t), d_l(t)\}, l = 1, 2, \dots, L, d_l(t)$ 为对应于样本 $x_l(t)$ 的网络期望输出, 设 $y_l(t)$ 为相应的网络实际输出, 将 $d_l(t)$ 展开为 $d_l(t) = \sum_{k=1}^K d_{lk} b_k(t)$. 为便于分析, 记 $U_{il} = \sum_{j=1}^K \omega_j a_{ji} - \theta$, $y_{kl} = \sum_{i=1}^n v_{ik} f(U_{il}) - \theta$. 则网络模型的误差函数可定义为

$$E = \frac{1}{2} \sum_{l=1}^L \sum_{k=1}^K (y_{kl} - d_{kl})^2 = \frac{1}{2} \sum_{l=1}^L \sum_{k=1}^K \left(\sum_{i=1}^n v_{ik} f \left(\sum_{j=1}^K \omega_j a_{ji} - \theta \right) - \theta - d_l \right)^2 \quad (13)$$

根据最速梯度下降算法, 网络模型的待训练参数调整规则为

$$\begin{cases} v_{ik}(s+1) = v_{ik}(s) + \alpha \Delta v_{ik}(s), \\ \omega_k(s+1) = \omega_k(s) + \beta \Delta \omega_k(s), \end{cases}$$

$$\begin{aligned}\theta(s+1) &= \theta(s) + \gamma \Delta \theta(s), \\ \theta(s+1) &= \theta(s) + \lambda \Delta \theta(s),\end{aligned}\quad (14)$$

其中: $\alpha, \beta, \gamma, \lambda$ 为学习速率; s 为学习迭代次数

在每一次学习迭代中, $\Delta v_{ik}, \Delta \omega_{ij}, \Delta \theta$ 和 $\Delta \theta$ 可分别表示为

$$\left\{ \begin{aligned}\Delta v_{ik} &= - \frac{\partial E}{\partial v_{ik}} = \\ & - \sum_{l=1}^L \sum_{k=1}^K (y_{kl} - d_{kl}) f'(U_{il}), \\ \Delta \omega_{ij} &= - \frac{\partial E}{\partial \omega_{ij}} = \\ & - \sum_{l=1}^L \sum_{k=1}^K (y_{kl} - d_{kl}) v_{ik} f'(U_{il}) a_{jl}, \\ \Delta \theta &= - \frac{\partial E}{\partial \theta} = \\ & \sum_{l=1}^L \sum_{k=1}^K (y_{kl} - d_{kl}) v_{ik} f'(U_{il}), \\ \Delta \theta &= - \frac{\partial E}{\partial \theta} = \sum_{l=1}^L \sum_{k=1}^K (y_{kl} - d_{kl}).\end{aligned}\right.\quad (15)$$

学习过程可完整描述如下:

Step 1: 选取合适的正交基函数将网络模型的期望输出函数展开;

Step 2: 用与 Step 1 中相同的正交基函数将网络模型的输入函数和连接权函数同时展开;

Step 3: 给定学习误差精度 ϵ , 学习迭代次数 $s=0$, 最大学习迭代次数 M 和学习速率;

Step 4: 初始化网络模型待训练参数;

Step 5: 根据式 (13) 计算误差函数 E , 如果 $E < \epsilon$ 或 $s > M$ 则转至 Step 7, 否则转至 Step 6;

Step 6: 根据式 (14) 和式 (15) 的规则调整待训练参数, $s+1=s$, 转至 Step 5;

Step 7: 输出学习结果, 结束

5 应用实例

航空发动机作为飞机的主动力装置, 由于其复杂的结构和恶劣的工作环境, 成为飞行故障的主要来源。其中, 因关键部件磨损所造成的故障在发动机故障中占有相当大的比例。由于关键部件之间存在摩擦, 总会有因过度磨损而产生的金属颗粒, 这些金属颗粒悬浮于循环的润滑油中, 当含量达到一定量级就会造成磨损类机械故障。通过对航空发动机润滑油金属含量进行预测可以监控滑油润滑部件的早期损伤, 以便及时采取相应措施, 消除事故隐患。

航空公司定期对航空发动机的滑油取样, 并进行滑油光谱分析。通过滑油光谱分析可以得到航空发动机滑油中的金属 (通常为 Fe, Mg, Cu, Al) 含量。根据现有的滑油金属含量历史数据, 建立时间序

列模型来预测滑油金属含量的未来变化趋势。由于影响航空发动机滑油金属含量的复杂因素很多, 很难用确定的数学模型来描述滑油金属含量的变化趋势。本文尝试用文中提出的两种时间序列预测模型来解决这一问题。

航空发动机中摩擦副材料大多是铁磁性金属, 因此在滑油磨粒分析技术中主要是对铁磁性磨粒的分析。图 4 是某型航空发动机滑油中 Fe 金属含量光谱分析的历史数据 (共计 155 个), 通过对该数据的自相关分析, 认为该数据属于平稳的随机序列。

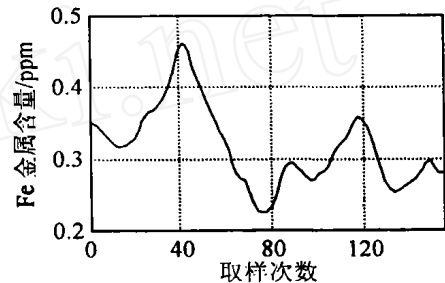


图4 Fe金属含量时间序列

首先采用基于过程神经网络的时间序列短期预测模型进行 Fe 金属含量的短期预测。将 Fe 金属含量时间序列中的连续 5 个数据进行拟合, 构成一个时变函数作为短期预测模型的输入, 以第 6 个数据作为短期预测模型的相应理想输出, 这样一共得到 150 组样本, 用前 100 组样本作为学习样本, 后 50 组样本作为检测样本。

短期预测模型的拓扑结构定为 1-10-1, 将其输入函数和连接权函数用 Legendre 正交基函数展开。学习误差精度设定为 10^{-4} , 学习速率为 0.01, 最大迭代次数为 2000 次, 网络模型经 122 次学习迭代后收敛。为测试学习完成后网络模型的泛化能力, 用后 50 组非学习样本进行测试。测试结果如图 5 所示。平均相对误差为 1.65%, 表明基于过程神经网络的时间序列短期预测模型对航空发动机滑油中 Fe 金属含量的预测满足实际工程要求, 能够达到对航空发动机健康状况进行实时监控的目的。图 5 同时给出拓扑结构为 5-10-1 的传统人工神经网络在相同条件下的一步预测结果, 该网络经 599 次学习迭代后收敛, 平均相对误差为 3.39%。通过对比可以看出, 基于过程神经网络的时间序列短期预测模型不仅具有较高的预测精度, 而且学习收敛速度较快。

采用基于过程神经网络的时间序列长期预测模型进行 Fe 金属含量的长期预测。将 Fe 金属含量时间序列中的连续 5 个数据进行拟合, 构成一个时变函数作为长期预测模型的输入, 将后连续 5 个数据进行拟合构成长期预测模型的相应理想输出函数。

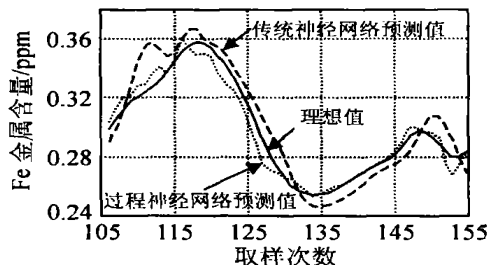


图5 Fe金属含量短期预测结果

这样一共得到30组样本,用前20组样本作为学习样本,后10组样本作为检测样本

长期预测模型的拓扑结构定为1-10-5-1,将模型的输入函数、连接权函数和样本中的理想输出用Legendre正交基函数展开。学习误差精度设定为 10^{-4} ,学习速率为0.01,最大迭代次数为10000次,网络模型经233次学习迭代后收敛。为测试学习完成后网络模型的泛化能力,用后10组非学习样本进行测试。测试结果如图6所示,平均相对误差为2.69%,表明基于过程神经网络的时间序列长期预测模型对航空发动机滑油中Fe金属含量的预测也能够满足实际工程要求。图6中同时给出了拓扑结构为5-10-5的传统人工神经网络在相同条件下的多步预测结果,该网络经6547次学习迭代后收敛,平均相对误差为4.26%。通过对比可以看出,基于过程神经网络的时间序列长期预测模型具有较高的预测精度,而且其学习收敛速度远远快于基于传统神经网络的时间序列长期预测模型。

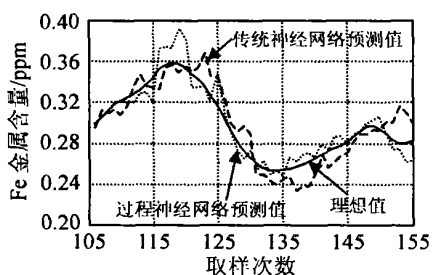


图6 Fe金属含量长期预测结果

文中提出的基于过程神经网络的两种时间序列预测模型都能满足实际工程要求,但通过图5与图6对比可以看出,短期预测模型的预测精度要比长期预测模型的预测精度高,且学习收敛速度较快。

6 结 语

本文针对时间序列中的短期预测和长期预测问题,分别建立了两种不同的基于过程神经网络的时间序列预测模型,并给出了相应的学习算法。由于过程神经网络能够方便地实现对泛函或算子的逼近,因而在处理时间序列预测问题时能够充分反映时间

序列中实际存在的时间累积效应,其预测精度与传统时间序列预测方法相比有所提高。以航空发动机健康状态监控中滑油Fe金属含量的预测为例进行实验,结果表明基于过程神经网络的时间序列预测模型与基于传统人工神经网络的时间序列预测模型相比具有更快的学习收敛速度。

参考文献 (References)

- [1] Hornik K, Stinchcombe M, White H. Multilayer Feedforward Networks Are Universal Approximators [J]. *Neural Networks*, 1989, 2(5): 359-366
- [2] Funahashi K. On The Approximate Realization of Continuous Mappings by Neural Networks [J]. *Neural Networks*, 1989, 2(3): 183-192
- [3] Frank R J, Davey N, Hunt S P. Time Series Prediction and Neural Networks [J]. *J of Intelligent and Robotic Systems: Theory and Applications*, 2001, 31(3): 91-103
- [4] Zhang G P, Qi M. Neural Network Forecasting for Seasonal and Trend Time Series [J]. *European J of Operational Research*, 2005, 160(2): 501-514
- [5] 何新贵, 梁久祯. 过程神经网络的若干理论问题 [J]. *中国工程科学*, 2000, 2(12): 40-44
(He X G, Liang J Z. Some Theoretical Issues on Process Neural Networks [J]. *Engineering Science*, 2000, 2(12): 40-44)
- [6] 许少华, 何新贵. 基于函数正交基展开的过程神经网络学习算法 [J]. *计算机学报*, 2004, 27(5): 645-650
(Xu S H, He X G. Learning Algorithms of Process Neural Networks Based on Orthogonal Function Basis Expansion [J]. *Chinese J of Computers*, 2004, 27(5): 645-650)
- [7] Jeffreys H, Jeffreys B S. *Methods of Mathematical Physics* [M]. 3rd ed. Cambridge: Cambridge University Press, 1988
- [8] Cheney E W. *Introduction to Approximation Theory* [M]. New York: McGraw-Hill, 1966
- [9] 何新贵, 许少华. 输入输出均为时变函数的过程神经网络及应用 [J]. *软件学报*, 2003, 14(4): 164-169
(He X G, Xu S H. Process Neural Network with Time-varied Input and Output Functions and Its Applications [J]. *J of Software*, 2003, 14(4): 164-169)
- [10] 许少华, 何新贵, 李盼池. 一类用于连续过程逼近的过程神经网络及其应用 [J]. *信息与控制*, 2004, 33(1): 116-119
(Xu S H, He X G, Li P C. A Process Neural Network for Continuous Process Approximation and Its Applications [J]. *Information and Control*, 2004, 33(1): 116-119)