

一种基于数值积分的过程神经网络训练算法

许少华^{1,2} 王颖¹ 王皓¹ 何新贵²

(大庆石油学院计算机与信息技术学院 大庆 163318)¹ (北京大学信息科学技术学院 北京 100871)²

摘要 针对过程神经网络的训练问题,提出了一种基于数值积分的学习算法。直接采用数值积分进行网络中动态样本与连接权函数的时域加权聚合运算,采用梯度下降法实现连接权函数特征参数及网络性质参数的调整。设计了基于梯形积分、辛普森积分、柯特斯积分等3种过程神经网络数值积分训练方法,以太阳黑子数据预测为例进行仿真实验,结果表明,基于数值积分的过程神经网络训练算法是有效的,其中辛普森积分算法的性能最优。

关键词 过程神经网络,学习算法,数值积分,时域聚合运算

中图分类号 TP18 **文献标识码** A

Training Algorithm of Process Neural Networks Based on Numerical Integration

XU Shao-hua^{1,2} WANG Ying¹ WANG Hao¹ HE Xin-gui²

(School of Computer & Information Technology, Daqing Petroleum Institute, Daqing 163318, China)¹

(School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)²

Abstract Aiming at the training problem of process neural networks, a training algorithm based on numerical integration was proposed. In proposed algorithm, the numerical integration was directly applied to deal with the weighted aggregation of dynamic samples and weight functions in time-domain, and the gradient descent method was used to adjust the weight function characteristic parameters and network property parameters. Three kinds of numerical integration methods of Trapezoidal, Simpson, and Cotes were designed. Taking the prediction of sunspot data as an example, the simulation results show that the training algorithms based on numerical integration are efficient, and the approximation performance of Simpson integration is optimal.

Keywords Process neural networks, Learning algorithm, Numerical integration, Time-domain aggregation operation

1 引言

自从1989年多层神经网络被证明是连续函数的一致逼近器^[1]以来,神经网络的研究得到了长足发展,并被广泛应用于各种实际领域^[2-5]。在实际应用中,许多系统的输入都是依赖于时间变化的函数,一些控制信号的输出既依赖于多输入函数的空间加权聚合,又与时间效应的累积密切相关,为此文献^[6]提出了过程神经网络(Process Neural Networks, PNN)的概念和模型。PNN的结构与传统神经元MP模型^[7]的结构相类似,不同之处在于过程神经元的输入/输出和权值可以是时变函数,其聚合运算既有对空间的多输入加权聚合,也有对时间过程效应的积累,可对时空二维信息直接进行处理。PNN将传统神经网络的函数映像关系推广为泛函映像,拓展了神经网络的应用领域。

对于神经网络的训练问题,已有许多成熟的学习算法,其中应用最为广泛和有效的是基于梯度下降的误差反传播学习算法(BP算法)及其各种改进形式。已有学者指出,可将神经网络的学习问题从动力学系统的角度来加以考察^[8,9],神经

网络学习算法与微分动力系统有着许多的相似性^[10]。因此,借助于微分动力系统的求解方法研究神经网络学习算法是一种有价值的尝试,也为寻找替代神经网络传统算法的新的学习算法提供了一条途径。文献^[10]证明了在一定条件下,BP学习算法的数值解与微分动力系统的解在渐近性方面是一致的,并通过研究微分动力系统解的存在性、唯一性和稳定性分析了BP算法解序列的特征。因此,神经网络的学习在一定条件下可以借助微分动力系统常用的数值解法,从而开拓出更广泛的神经网络学习方法。

过程神经网时域聚合运算的有效处理是过程神经网络训练的核心问题。对于PNN的训练,已有多种有效的学习方法^[11-13]。笔者在前面研究工作的基础上,借助于微分动力系统的数值解法,提出了一种基于数值积分的过程神经网络训练算法(Numerical Integrate-based PNN, NIPNN)。该算法直接采用数值积分处理过程神经元中动态样本与连接权函数的时域聚合运算,采用梯度下降算法进行网络参数的调整,从而可有效降低算法复杂度,增强算法的适用性。理论分析和仿真实验结果表明,该算法是可行有效的。

到稿日期:2009-12-15 返修日期:2010-03-11 本文受国家自然科学基金项目(60572174),黑龙江省科技攻关项目(GZ07A103)资助。

许少华(1962-),男,博士,教授,博士生导师,CCF会员,主要研究方向为神经网络、数据挖掘、智能信息处理,E-mail: xush62@163.com;

王颖(1986-),女,硕士生,主要研究方向为数据挖掘、智能信息处理;王皓(1986-),女,硕士生,主要研究方向为神经网络、智能信息

处理;何新贵(1938-),男,教授,博士生导师,中国工程院院士,CCF会员,主要研究方向为模糊逻辑、神经网络、进化计算、数据库理论。

2 过程神经网络模型

2.1 过程神经元

过程神经元是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,拓扑结构如图1所示。

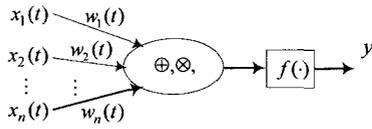


图1 过程神经元

图1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, “ \oplus ”和“ \otimes ”构成了过程神经元时空二维聚合算子对; $K(t)$ 为过程神经元的时间聚合核函数, $f(\cdot)$ 为激励函数。

如果“ \oplus ”取为空间加权和运算, “ \otimes ”取对时间的积分, 且 $K(\cdot)=1$, 则过程神经元输入输出之间的映射关系可写为:

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

2.2 过程神经网络模型

过程神经网络是由若干个过程神经元按照一定的拓扑结构组成的网络。考虑包含1个隐层、多输入/单输出的系统, 网络结构如图2所示。

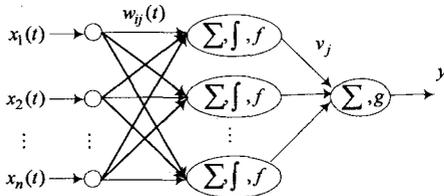


图2 过程神经网络

图2中, 输入层有 n 个单元, 中间层(过程神经元隐层)有 m 个单元, 输出层为线性关系。网络输入与输出之间的关系为:

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

式中, $w_{ij}(t)$ 为输入层与隐层的连接权函数, v_j 为隐层过程神经元到输出层的连接权值, θ_j 为隐层的输出阈值, $[0, T]$ 为时间过程区间。

3 NIPNN 训练算法

由于过程神经网络的连接权可以是任意形式的时间函数, 如果不将其限定在某一函数类中, 则很难通过对函数样本集的训练来确定。实际应用中, 可根据时变系统的模态变化特征或动态样本的统计规律, 选择统计特性与实际系统变化特征相接近且拟合能力强、柔韧性好的函数系作为权函数空间, 例如二次或三次 B-样条函数、三角函数系、Walsh 函数系等。这样, 在网络训练中只需对连接权函数中的性质参数进行调整即可。

3.1 数值积分基本公式^[14]

$$\text{记: } x_k = a + kh, k=0, 1, \dots, N, h = \frac{b-a}{N}.$$

梯形积分公式:

$$I_1(f(x)) = \frac{h}{2} (f(a) + 2 \sum_{k=1}^{N-1} f(x_k) + f(b)) \quad (3)$$

辛普森积分公式:

$$I_2(f(x)) = \frac{h}{6} (f(a) + 4 \sum_{k=1}^N f(x_{2k-1}) + 2 \sum_{k=1}^{N-1} f(x_{2k}) + f(b)) \quad (4)$$

柯特斯公式:

$$I_3(f(x)) = \frac{h}{90} [7f(a) + 32 \sum_{k=1}^N f(x_{4k-3}) + 12 \sum_{k=1}^N f(x_{4k-2}) + 32 \sum_{k=1}^N f(x_{4k-1}) + 14 \sum_{k=1}^{N-1} f(x_{4k}) + 7f(b)] \quad (5)$$

3.2 NIPNN 训练算法

过程神经网络的输入可以是连续时间函数或离散时间采样数据序列。对于连续时间函数, 在满足采样定理的条件下, 可按照一定的时间粒度在过程输入区间上离散为时间序列形式, 以方便数值积分的运算。

设网络输入过程区间为 $[0, T]$, 将 $[0, T]$ 进行适当的 P 分割, 记区间分割点为 t_0, t_1, \dots, t_P ; 其中, $t_0=0, t_P=T$ 。

给定 K 个序列长度为 P 的学习样本: $(x_{1j}^{(k)}(t_p), x_{2j}^{(k)}(t_p), \dots, x_{mj}^{(k)}(t_p) d_k)$; 其中, $j=1, 2, \dots, m; p=1, 2, \dots, P; k=1, 2, \dots, K; d_k$ 为期望输出。

对于图2描述的过程神经网络, 对应于第 k 个样本输入, 基于数值积分时域聚合运算的 PNN 输入/输出关系为(为表述方便, 令式(3)一式(5)中, $h=1$):

$$y_k = \sum_{j=1}^m v_j f\left(\sum_{p=1}^P \sum_{i=1}^n I_s(x_{ij}^{(k)}(t_p) w_{ij}(t_p)) \Delta t_p - \theta_j\right) \quad (6)$$

式中, $s=1, 2, 3; \Delta t_p = t_p - t_{p-1}$ 。训练误差函数可取为:

$$E = \sum_{k=1}^K (y_k - d_k)^2 = \sum_{k=1}^K \left(\sum_{j=1}^m v_j f\left(\sum_{p=1}^P \sum_{i=1}^n I_s(x_{ij}^{(k)}(t_p) w_{ij}(t_p)) \Delta t_p - \theta_j\right) - d_k\right)^2 \quad (7)$$

参数调整规则为:

$$v_j = v_j + \alpha \Delta v_j \quad (8)$$

$$w_{ij}(t) = w_{ij}(t) + \beta \Delta w_{ij}(t) \quad (9)$$

$$\theta_j = \theta_j + \gamma \Delta \theta_j \quad (10)$$

式(9)是对 $w_{ij}(t)$ 中的性质参数进行调整; α, β, γ 为学习速率。

为表述方便, 令:

$$\sum_{j=1}^m v_j f\left(\sum_{p=1}^P \sum_{i=1}^n I_s(x_{ij}^{(k)}(t_p) w_{ij}(t_p)) \Delta t_p - \theta_j\right) = u_{kj}, \text{ 则:}$$

$$\Delta v_j = -\frac{\partial E}{\partial v_j} = -2 \sum_{k=1}^K \left(\sum_{j=1}^m v_j f(u_{kj}) - d_k\right) f(u_{kj}) \quad (11)$$

$$\begin{aligned} \Delta w_{ij}(t) &= -\frac{\partial E}{\partial w_{ij}(t)} \\ &= -2 \sum_{k=1}^K \left(\sum_{j=1}^m v_j f(u_{kj}) - d_k\right) v_j f'(u_{kj}) c x_{ij}^{(k)}(t) \end{aligned} \quad (12)$$

其中, 当 $s=1, t=1, p$ 时, $c=1/2$;

当 $s=1, t=2, 3, \dots, p-1$ 时, $c=1$;

当 $s=2, t=1, p$ 时, $c=1/6$;

当 $s=2, t=2, 4, \dots, p-1$ 时, $c=4/6$;

当 $s=2, t=3, 5, \dots, p-2$ 时, $c=2/6$;

当 $s=3, t=1, p$ 时, $c=7/90$;

当 $s=3, t=2, 6, \dots, p-3$ 时, $c=32/90$;

当 $s=3, t=3, 7, \dots, p-2$ 时, $c=12/90$;

当 $s=3, t=4, 8, \dots, p-1$ 时, $c=32/90$ 。

$$\Delta \theta_j = -\frac{\partial E}{\partial \theta_j} = -2 \sum_{k=1}^K \left(\sum_{j=1}^m v_j f(u_{kj}) - d_k\right) v_j f'(u_{kj}) \quad (13)$$

若激励函数 f 取为 Sigmoid 函数, 则:

$$f'(u) = f(u)(1-f(u)) \quad (14)$$

3.3 NIPNN 算法描述

- Step1 PNN 结构初始化;
 Step2 给定误差精度 ϵ ; 最大迭代次数 M ; 置当前步数 $s=0$;
 Step3 初始化权值和阈值 $v_j, w_{ij}(t), \theta_j (i=1, 2, \dots, n; j=1, 2, \dots, m; t=1, 2, \dots, p)$;
 Step4 由式(7)计算网络输出误差 E , 若 $E < \epsilon$ 或 $s > M$ 转 Step6;
 Step5 按式(8)一式(13)修正权值和阈值; $s=s+1$; 转到 Step4;
 Step6 保存学习结果, 结束。

4 实际应用

太阳黑子活动对人类的生活环境有着较大的影响。人们在观测太阳的漫长过程中积累了很多关于太阳黑子的观测数据, 这些数据构成了一个典型的时变过程记录。利用过程神经网络可根据这些观测数据较方便地建立太阳黑子的预测模型。

4.1 预测方案

本实验采用连续 259 年(1749 年—2007 年)太阳黑子季数据和年数据^[15]进行建模。季数据指一个季度的太阳黑子平均数, 年数据指全年 4 个季度太阳黑子数的平均值。采用的预测方案是: 用前 n 年的季数据预测第 $n+1$ 年的年数据。例如, 用 1850 年、1851 年、1852 年的季数据, 预测 1853 年的年太阳黑子数平均值, 此时 $n=3$ 。

4.2 网络模型

根据预测方案, 在过程神经网络中, 系统输入为 n 个太阳黑子的季数据序列, 每个输入的时间序列长度 $P=4$; 输出为一个单值(第 $n+1$ 年的年数据)。实验中, 网络输入节点个数取 $n=5$, 隐层节点个数取 $2n=10$, 输出层节点为 1 个。

4.3 训练结果

选取前 180 年(1749 年—1928 年)的数据用于网络训练, 使其逼近太阳黑子序列在不同时段之间的复杂映射关系, 后 59 年(1929 年—2007 年)数据用于测试网络的泛化推广能力。网络权函数的类型取为 Walsh 函数, 基函数个数为 8。记网络输出的最大误差、最小误差、均方误差分别为 E_{\max} , E_{\min} , E_{avg} ; 训练步数为 10^4 。3 种数值积分算法的训练结果如表 1 所列。

表 1 3 种算法的训练结果对比

积分类型	E_{\max}	E_{\min}	E_{avg}
梯形	40.3556	0.0545	13.1223
辛普森	20.0500	0.0358	5.1450
柯特斯	30.2768	0.1521	8.7296

由表 1 可知, 时域聚合运算采用辛普森积分时网络逼近性能最佳, 训练结果对比如图 3 所示。

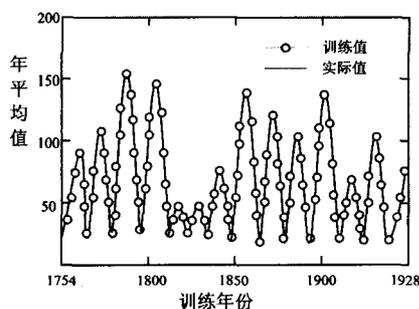


图 3 1754 年—1928 年训练结果与实际结果对比

4.4 预测结果

用 1929 年—2007 年的样本数据, 对辛普森积分算法训练好的网络进行测试, 预测结果: $E_{\max} = 22.1614$; $E_{\min} = 0.3481$; $E_{\text{avg}} = 10.0583$ 。预测结果与实际结果对比如图 4 所示。

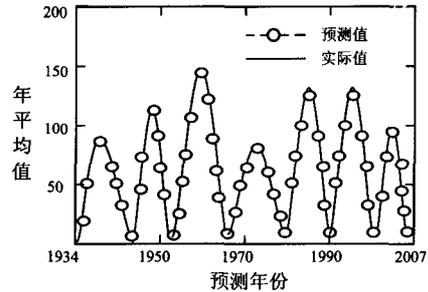


图 4 1934 年—2007 年预测结果与实际结果对比

结束语 针对过程神经网络训练中时域聚合运算的有效处理问题, 本文提出了一种基于数值积分的 PNN 训练方法。由于过程神经网络的聚合运算机制和计算过程较传统神经网络复杂, 笔者提出的方法为过程神经网络学习算法的研究提供了一种新的途径。

参考文献

- [1] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2: 359-366
- [2] Matuki T, Kudo T, Kondo T. Three dimensional medical images of the lungs and brain recognized by artificial neural networks [C]//SICE, Annual Conference. 2007; 1117-1121
- [3] Swain A K, Morris A S. A novel hybrid evolutionary programming method for function optimization[C]//Proceedings of the 2000 Congress on Evolutionary Computation. 2000, 1: 699-705
- [4] Dehghan M, Faez K, Ahmadi M. A hybrid handwritten word recognition using self-organizing feature map, discrete HMM, and evolutionary programming[C]//Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. 2000, 5: 515-520
- [5] Lieslehto J. PID controller tuning using evolutionary programming[C]//Proceedings of the 2001 American Control Conference. 2001, 4: 2828-2833
- [6] 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
- [7] 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
- [8] Roychowdhury P, Singh Y P, Chansarkar R A. Dynamic tunneling technique for efficient training of multilayer perceptron[J]. IEEE Trans on Neural Networks, 1999, 10(1): 48-55
- [9] Fukuoka Y, et al. A modified back-propagation method to avoid false local minim[J]. Neural Networks, 1998, 11(8): 1059-1072
- [10] 梁久祯. 模糊神经网络逼近理论及学习算法研究[D]. 北京: 北京航空航天大学, 2001
- [11] 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
- [12] 许少华, 何新贵. 过程神经网络一种基于函数正交基展开的学习算法[J]. 计算机学报, 2004, 27(5): 645-649
- [13] 葛利. 一种基于混合遗传算法学习的过程神经网络[J]. 哈尔滨工业大学学报, 2005, 37(7): 986-988
- [14] 关治, 陆金普. 数值分析基础[M]. 北京: 高等教育出版社, 1998
- [15] 本文所引用数据来自 SIDC(solar influences data analysis center)