

BP 学习算法在 CNN 发电机建模中应用

施伟锋

(上海海事大学 电气自动化系, 上海 200135)

摘要: 在应用混沌神经网络(CNN)进行同步发电机的建模过程中,对于 CNN 的学习,网络训练过程的收敛性很难控制。在研究了 BP 学习算法及其一些改进方法进行人工神经网络训练的轨迹收敛特性后,观测到运用梯度下降动量与自适应学习速率相结合的 BP 学习算法的神经网络训练轨迹的收敛特性良好。在用基于 Aihara 混沌神经元构成的 3 层反馈 CNN 进行同步发电机建模的应用中,用该 BP 学习算法对 CNN 进行了训练。结果表明:用该 BP 算法进行 CNN 发电机建模具有学习速度快和均方误差曲线轨迹收敛性好的特点,而且所建立的 CNN 同步发电机模型运行的动态过程误差小。

关键词: BP 算法; 混沌神经网络; 发电机; 建模

中图分类号: TM 31

文献标识码: A

文章编号: 1006-6047(2007)07-0009-04

0 引言

混沌神经网络 CNN(Chaotic Neural Networks)的学习问题是一个值得研究的课题^[1],CNN 训练轨迹的收敛特性与学习方法等因素有关^[2-3]。BP 学习算法是用于人工神经网络 ANN(Artificial Neural Network)训练的常用算法之一^[4],通过设置 ANN 的目标函数,运用权值的链式调整算法,可得到 ANN 权值和阈值的 BP 学习算法的修正递推计算式。BP 学习算法也存在一些公认的缺点,为使训练的 CNN 具有良好的函数逼近特性,需要优化 BP 学习算法的学习速度与收敛特性^[5]。目前,有一些对 BP 算法进行改进的方法;在实际的针对性问题的应用中,采用何种改进 BP 学习方法进行学习可以使 CNN 的学习收敛特性好还难以预料与控制,因为实际问题不仅与使用的算法本身有关,还与问题的复杂性、训练样本的数量、CNN 权值和阈值的数量、误差目标值、CNN 的用途等诸多因素有关;对于具体问题的结论需通过对实际问题的求解后得出。

1 BP 学习算法的改进

梯度下降 BP 学习算法是 ANN 训练的一种方法,下面分析该算法的计算过程^[6]。

1.1 网络的目标函数

$$E = \frac{1}{2} \sum_{j=t_0+1}^l (d_j - y)^2 = \frac{1}{2} \sum_{j=t_0+1}^l e_j^2 \quad (1)$$

式中 d_j 和 y 分别为输出样本参数和神经网络的输出参数。

1.2 权值调整算法

若隐层至输出层的连接权为 $\mathbf{W}^{(2)}$,有

$$\frac{\partial E(n)}{\partial w_i^{(2)}} = \frac{\partial E(n)}{\partial y(n)} \frac{\partial y(n)}{\partial x(n)} \frac{\partial x(n)}{\partial w_i^{(2)}} = \delta(n) o_i(n) \quad (2)$$

式中 $\delta(n) = \partial E(n) / \partial x(n) = -e(n) f_2'(x(n))$ 。

隐层连接权为 $\mathbf{W}^{(H)}$,有

$$\frac{\partial E(n)}{\partial w_{ij}^{(H)}} = \frac{\partial E(n)}{\partial o_i(n)} \frac{\partial o_i(n)}{\partial x_{o,i}(n)} \frac{\partial x_{o,i}(n)}{\partial w_{ij}^{(H)}} \quad (3)$$

对于第 q 个输入,输入层至隐层的连接权为 $\mathbf{W}^{(1)}$,有

$$\frac{\partial E(n)}{\partial w_{iq}^{(1)}} = \frac{\partial E(n)}{\partial x_{o,i}(n)} \frac{\partial x_{o,i}(n)}{\partial w_{iq}^{(1)}} = \delta_{o,i}(n) u_q(n-1) \quad (4)$$

得到的动态 BP 学习算法的权值调整式为

$$\Delta w_i^{(2)} = -\eta e(n) f_2'(x(n)) o_i(n) = \eta \delta(n) o_i(n)$$

$$\Delta w_{ij}^{(H)} = -\eta \delta_{o,i}(n) \frac{\partial x_{o,i}(n)}{\partial w_{ij}^{(H)}} \quad (5)$$

$$\Delta w_{iq}^{(1)} = -\eta \delta_{o,i}(n) u_q(n-1)$$

为寻找具有良好收敛特性的 CNN 学习方法,先研究各种 BP 改进学习算法对常规 ANN 网络函数逼近问题训练过程的影响。采用相同的神经元算法、网络结构与参数进行 ANN 训练的稳定性与快速性研究。研究中运用的学习算法有^[7-8]:最速下降 SD(Steepest Descent) BP 算法、共轭梯度 CG(Conjugate Gradient) BP 算法、LM(Levenberg Marquardt) BP 算法、动量 MO(MOMentum) BP 算法、自适应学习速率 VL(adaptive Variable Learning rate)BP 算法。ANN 函数逼近训练过程中得到学习过程在均方误差 MSE(Mean Square Error)性能曲面中的轨迹如图 1 所示。由于训练过程初始点(2.5,3)和终点(10,0.5)是相同的,从而可以根据训练的轨迹判断学习过程的稳定性与快速性的优劣。

观察图 1(a)(b),最速下降算法收敛的速度慢、训练需要的步数多,而且当学习速率系数过大时,在学习过程还未到达终点,就发生了振荡(图 1(b)中大块黑点密集区域),发生发散现象;观察图 1(c)(d),

共轭梯度算法与 LMBP 算法收敛较快,但步长过大,轨迹的平滑性不好。

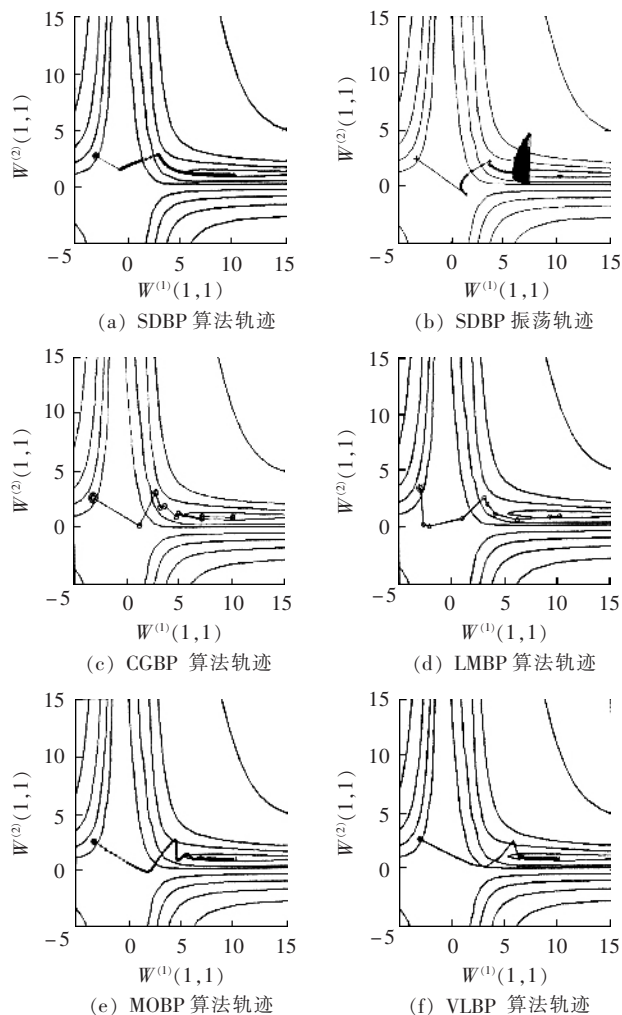


图 1 MSE 曲面中 ANN 训练的轨迹

Fig.1 Trajectories of ANN training on MSE surface

对于动量算法,算法在原 BP 迭代式的基础上,权值与阈值的增量计算的迭代为

$$\Delta W^{(m)}(n) = \gamma \Delta W^{(m)}(n-1) - (1-\gamma) \alpha S^{(m)}[o^{(m-1)}(n)]^T \quad (6)$$

$$\Delta \theta^{(m)}(n) = \gamma \Delta \theta^{(m)}(n-1) - (1-\gamma) \alpha S^{(m)} \quad (7)$$

式中 $\Delta W^{(m)}(n)$ 和 $\Delta \theta^{(m)}(n)$ 分别为第 m 层第 n 次迭代权值和阈值动量项的增量; γ 为动量系数; α 为学习速率; $S^{(m)}$ 为灵敏度向量,是近似均方误差对第 m 层的输入元素变化的敏感性; $o^{(m-1)}(n)$ 为输出动量项。

动量系数的加入,增加了一种滤波器的效果,动量系数 γ 增加时,滤波后输出的振荡减少,而平均输入与输出相同,仍然保持平均值,这体现出动量项对学习过程的类似低通滤波的作用,从而改善学习过程,达到使学习轨迹维持平滑稳定的前提下加速收敛的目的。观察图 1(e),动量算法收敛轨迹平滑,稳定性好。

自适应学习速率算法的神经网络权值处理原理为

$$\alpha(k+1) = \begin{cases} k_{inc} \alpha(k) & E(k+1) < E(k) \\ k_{dec} \alpha(k) & E(k+1) > E(k) \end{cases} \quad (7)$$

式中 k_{inc} 为增量因子; k_{dec} 为减量因子; $\alpha(k+1)$ 为进

行处理的权值或阈值。

该算法在曲面非常平坦的区域用大的学习速率,在曲率很高的区域用小的学习速率,改善学习过程,达到学习过程稳定与快速的目的。观察图 1(f),自适应学习速率算法收敛轨迹的步长可以自动调整得很恰当,且收敛速度快。

从结果分析得出,动量 BP 算法、自适应学习速率 BP 算法 MSE 轨迹收敛特性较好;结合动量与自适应学习速率产生动量自适应学习速率 BP 算法,将增强神经网络在学习过程中 MSE 轨迹的稳定性与收敛的快速性。这里将这两者结合的 BP 算法用于 CNN 发电机建模的学习。

2 基于 Aihara 神经元的 CNN

基于 Aihara 混沌神经元^[9]建立反馈型 CNN, Aihara 混沌神经元算法为

$$\begin{aligned} x(k+1) &= lx(k) - \gamma g\{f[x(k)]\} + a(k) \\ o(k+1) &= f[x(k+1)] \\ f[x(k)] &= 1 / \{1 + \exp[-x(k)/\varepsilon]\} \\ a(k) &= [A(k) - \theta](1-l) \end{aligned} \quad (8)$$

式中 x 为神经元内部状态; o 为神经元输出; l 为不应性衰减率, $0 \leq l \leq 1$; γ 为比例系数, $\gamma > 0$; f 为神经元传输函数; $g(\cdot)$ 为不应性函数,通常简化为 $g(x) = x$; θ 为神经元的输入阈值; ε 为输出函数的陡度参数; $A(k)$ 为在离散时刻 k 的神经元外加的输入强度。

将 Aihara 混沌神经元作为网络的隐层神经元,形成含有 1 个输入层、1 个带有反馈的隐层和 1 个输出层的 CNN,网络结构如图 2 所示。

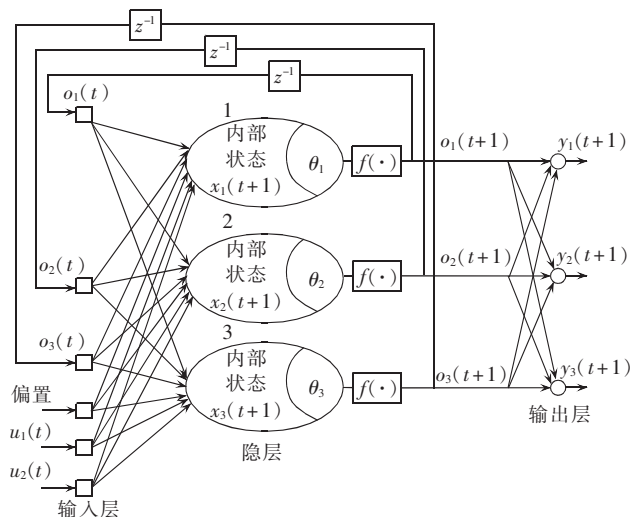


图 2 CNN 的网络结构

Fig.2 Network configuration of CNN

设网络外部的输入时间序列为 $u(t)$, 隐层的输出为 $o(t)$, 网络的输出为 $y(t)$, 则网络结构描述为

$$\begin{aligned} x_o(t+1) &= W^{(H)} o(t) + W^{(1)} u(t) + \theta^{(1)} \\ o(t) &= f_1(x_o(t)) \\ y(t) &= f_2(x_o(t)) = f_2(W^{(2)} o(t) + \theta^{(2)}) \end{aligned} \quad (9)$$

式中 $f_1(\cdot)$ 是非线性 Logistic 函数; $f_2(\cdot)$ 是连续输出函数; $W^{(1)}$ 、 $W^{(2)}$ 和 $W^{(H)}$ 分别为输入层至隐层、隐

层至输出层以及隐层节点之间(包括节点自身)的连接权矩阵。

神经网络的混沌特性体现在神经元传输函数是陡度参数 $\varepsilon=0.04$ 的 Logistic 函数,神经元传输函数混沌特性如图 3 所示。分析函数的分岔图,可以清楚地看到 $\varepsilon=0.04$ 时神经元传输函数的倍周期分岔过程如图 3(a)所示;进一步,求出混沌神经元传输函数的 Lyapunov 指数,曲线如图 3(b)所示。当 Lyapunov 指数大于 0 时,神经元表现出混沌特性。混沌算法具有强烈的遍历性,将混沌算法引入到神经元算法中形成混沌神经元,混沌神经元将具有强烈的遍历性^[10],使基于该 CNN 建模时参数的寻优范围增大;反馈型网络对信息具有密集的处理能力,在同步发电机建模中,将对发电机的模型产生极强的辨识能力。运用神经网络对非线性函数极强的逼近能力^[11-12],考虑运用该 CNN 进行同步发电机的建模。因此,采用 Lyapunov 指数大于 0 时的关于参数 $a(t)$ 的神经元算法完成发电机建模工作。

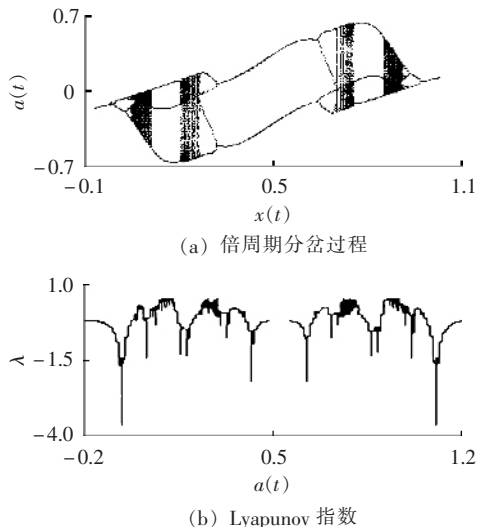


图 3 Aihara 神经元传输函数混沌特性图

Fig.3 Chaotic characteristics of Aihara neuron transfer function

3 基于 CNN 的同步发电机建模

由于同步发电机的输入机械功率与同步发电机的有功功率相关,即与发电机的输出转速(产生电网的频率)有关;同步发电机的输入励磁电流与同步发电机的无功功率相关,即与发电机的输出电压有关^[13],因此,选择同步发电机的输入机械功率和励磁电流作为 CNN 的输入,选择同步发电机的频率和发电机端电压作为 CNN 的输出^[14]。系统建模采用二输入、二输出串并联模型辨识结构,如图 4 所示。

为探索 BP 学习算法对 CNN 训练的有效性,分别采用各种改进的 BP 学习算法进行训练,观察与测量建模过程中的 MSE 曲线的收敛性,研究 CNN 的训练方法,分析与判断 BP 学习算法在 CNN 建模应用中的优劣。

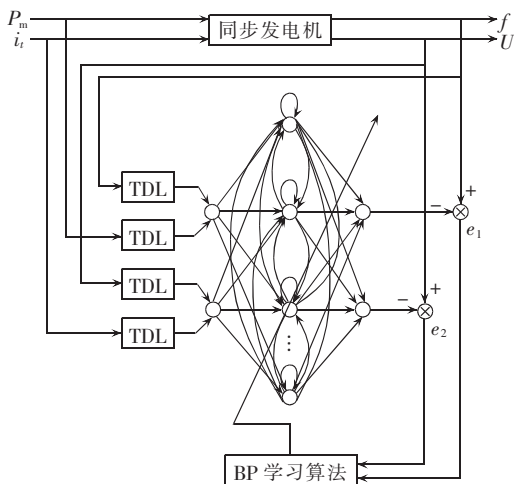


图 4 CNN 发电机建模结构

Fig.4 Configuration of CNN generator modeling

建模时同步发电机的样本参数需要充分地反映同步发电机工作的动态范围^[15],选择发电机启动工况、电网有小负载工况、电网有大负载工况、电网有短路故障 4 种工况的动态响应的样本数据用于网络学习;将 4 种工况下输入机械功率与励磁电流和同步发电机的频率与发电机端电压输出的对应数据作为神经网络学习的样本数据。CNN 采用图 4 所示结构,输入层有 2 个神经元(2 个输入信号)、隐层有 16 个 Aihara 混沌神经元(建模中发现,随着隐层神经元数的增加,MSE 的收敛性会得到改善,当隐层神经元数大于 16,MSE 的收敛性不再有明显的改善,训练速度会减慢)、输出层有 2 个神经元(2 个输出信号);CNN 输出与发电机输出样本参数的比较采用 MSE 形式。设 CNN 学习时 MSE 的目标值为 0.5×10^{-3} ,训练过程中记录 MSE 曲线,进行学习收敛的稳定性与快速性比较。几种 BP 改进算法在 CNN 发电机建模训练运用中,关于发电机频率参数训练过程的 MSE 曲线的收敛情况见图 5(N 为训练周期数),可测量到 MSE 值减小到一定的值后减小的速度下降,甚至不减小(图 5(c))。

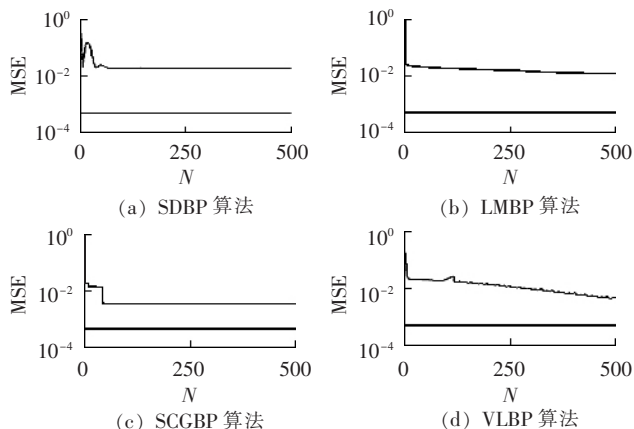


图 5 BP 算法用于 CNN 训练时 MSE 特性

Fig.5 MSE characteristics of BP algorithm for CNN training

图 6 是建模中动量与自适应学习速率 BP 算法的 CNN 训练的 MSE 曲线的收敛情况,图 6(a)是训练时发电机输出频率参数的 MSE 曲线,图 6(b)是训练时发电机输出电压参数的 MSE 曲线,与图 5 相比较可看出动量与自适应学习速率 BP 算法的 MSE 曲线收敛特性最好。而图 5 的几种 BP 算法训练的 MSE 由于减小到一定程度后就不能继续以较快速度减小,无法达到目标 MSE 值(0.5×10^{-3}),这些算法对于 CNN 的训练意义不大。在 CNN 同步发电机建模中,可观察到运用动量与自适应学习速率 BP 算法训练的模型与实际样本数据的拟合良好。图 7 是选择船舶电力系统常见的 30% 单台发电机额定负载运行工况下发发电机输出电流与样本数据的拟合结果,可看到发电机 CNN 模型在有负载时输出电流的变化与实际电流变化的数值很接近。图 7(a)反映了加负载过程 CNN 发电机模型与样本数据的拟合情况,其中虚线是发电机组输出端电流样本数据值,实线是 CNN 发电机模型输出电流值;图 7(b)是 CNN 发电机模型输出与样本数据之间的误差值,可看到误差为 10^{-3} 数量级。

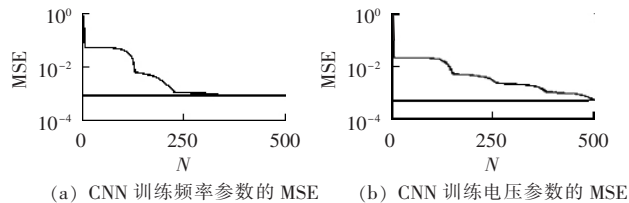


图 6 动量与自适应学习速率 BP 算法训练时的 MSE

Fig.6 MSE characteristics of proposed BP algorithm for CNN training

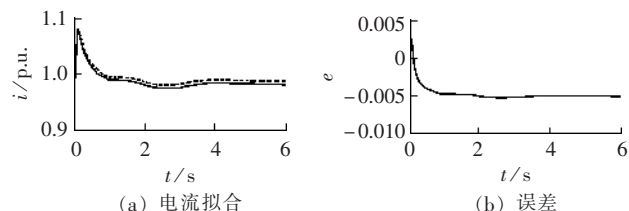


图 7 发电机 CNN 模型的数据拟合情况

Fig.7 Data fitting of generator CNN model

4 结论

从各种 BP 改进算法在 ANN 的训练中,可观察到动量 BP 算法与自适应学习速率 BP 算法的学习轨迹的收敛性比较理想。在同步发电机 CNN 建模中,对于基于 Aihara 混沌神经元建立的反馈型 CNN 的学习,建模的 MSE 收敛特性与采用不同改进方法的 BP 算法有很大的关系。这里将 2 种 BP 算法的改进方法结合起来,形成动量与自适应学习速率的 BP 学习算法,对于 CNN 进行训练的 MSE 曲线的收敛性比较理想,其他 BP 学习方法均不理想。动量与自适应学习速率 BP 算法在训练中,随着 CNN 的混沌神经元的权值和阈值数量的增加,收敛的目标 MSE 值可以减小,但权值和阈值数量超过一定数量后 MSE 值不再减小。对于训练样本的数量,数量过多时训练的时间会加大,过多的样本、过长时间的学习还会使 CNN 对模型产生过拟合,反而使 CNN 的泛化能力下降,使建

立的 CNN 发电机模型的动态性能下降。

参考文献:

- [1] 阮炯,顾凡及,蔡志杰. 神经动力学模型方法和应用[M]. 北京: 科学出版社,2002.
- [2] ADACHI M, AIHARA K. Associative dynamics in chaotic neural network[J]. Neural Networks, 1997(10): 83-98.
- [3] 李广,邹德忠,谈顺涛. 基于混沌神经网络理论的小电网短期电力负荷预测[J]. 电力自动化设备, 2006, 26(2): 50-52. LI Guang, ZOU De-zhong, TAN Shun-tao. Short-term load forecast for small power net based on chaos-artificial neural network theory[J]. Electric Power Automation Equipment, 2006, 26(2): 50-52.
- [4] CHEN D S, JAIN R C. A robust back propagation learning algorithm for function approximation[J]. IEEE Trans on Neural Networks, 1994, 5(3): 467-479.
- [5] HAYKIN S. Neural networks—a comprehensive foundation[M]. 2nd ed. New York: Prentice Hall, 1999.
- [6] HAGAN M T, DEMUTH H B, BEALE M H. Neural network design[M]. Boston: PWS Publishing Company, 1996.
- [7] DU L M, HOU Z Q, LI Q H. Optimum block-adaptive learning algorithm for error back-propagation networks[J]. IEEE Trans on Signal Processing, 1992, 40(12): 3032-3042.
- [8] YU X H, CHEN G A, CHENG S X. Dynamic learning rate optimization of the back propagation algorithm[J]. IEEE Trans on Neural Networks, 1995, 6(3): 669-677.
- [9] AIHARA K, TAKABE T, TOYODA M. Chaotic neural networks[J]. Physics Letter A, 1990, 144(6): 333-340.
- [10] CHEN L N, AIHARA K. Global searching ability of chaotic neural networks[J]. IEEE Trans on Circuits and Systems, 1999, 46(8): 974-993.
- [11] 唐蕾,陈维荣. 电力系统负荷建模中的小波神经网络新技术[J]. 电力自动化设备, 2003, 23(3): 72-77. TANG Lei, CHEN Wei-rong. New technology of wavelet neural network in power system load modeling[J]. Electric Power Automation Equipment, 2003, 23(3): 72-77.
- [12] 石敏,吴正国,徐袭. 基于概率神经网络和双小波的电能质量扰动自动识别[J]. 电力自动化设备, 2006, 26(3): 5-8. SHI Min, WU Zheng-guo, XU Xi. Auto recognition of power quality disturbance based on probabilistic neural networks and double wavelet[J]. Electric Power Automation Equipment, 2006, 26(3): 5-8.
- [13] YEAGER K E, WILLIS J R. Modeling of emergency diesel generators in an 800 megawatt nuclear power plant[J]. IEEE Trans on Energy Conversion, 1993, 8(3): 433-441.
- [14] OURARI M L, LOUIS A D, DO Van-que. Dynamic equivalent modeling of large power systems using structure preservation technique[J]. IEEE Trans on Power Systems, 2006, 21(3): 1284-1295.
- [15] 施伟锋,聂益文. 船舶大功率发电机混沌神经网络建模[J]. 中国电机工程学报, 2005, 25(21): 156-162. SHI Wei-feng, NIE Yi-wen. Marine high power generator modeling based on chaotic neural network[J]. Proceedings of the CSEE, 2005, 25(21): 156-162.

(责任编辑:李育燕)

作者简介:



施伟锋(1963-),男,浙江宁波人,教授,博士,主要研究方向为电力系统建模、分析与控制,控制理论与应用(E-mail: wfshi@shmtu.edu.cn)。

BP learning algorithm for CNN generator modeling

SHI Wei-feng

(Department of Electrical Engineering and Automation,
Shanghai Maritime University, Shanghai 200135, China)

Abstract: In the application of CNN (Chaotic Neural Networks) to synchronous generator modeling, the training process convergence of network learning is difficult to control. Based on the characteristic study of trajectory convergence in artificial neural network training using BP (Back Propagation) learning algorithm and its improved algorithms, it is found that the improved BP algorithm combining gradient descent momentum and adaptive learning rate is better, which is then used to train the CNN in the synchronous generator modeling based on a three-layer feed back CNN with Aihara chaotic neuron. Results indicate that, the learning speed is fast, the trajectory convergence of mean square error is better and the dynamic process error of synchronous generator operation is small.

This project is supported by the Fund of Science and Technology of Shanghai Municipal Education Committee(05FZ09) and the Key Discipline of Shanghai Maritime University(XL0105).

Key words: BP algorithm; chaotic neural network; generator; modeling