

# 加动量项且变学习速率的 BP 算法 应用于变压器故障诊断

许 虎, 明志强, 侯 勇  
(南充电业局, 四川 南充 637000)

**摘 要:** BP 神经网络具有自学习和自适应能力, 非常适合于变压器故障诊断。分析了加动量项 BP 算法和变学习速率 BP 算法存在的不足, 给出了加动量项且变学习速率的 BP 算法的原理和优点, 减轻了网络训练过程中的振荡, 加速了网络的收敛。构造了适合于变压器油中溶解气体分析故障诊断的神经网络, 并将训练所得的神经网络用于变压器故障诊断, 结果表明了该算法的有效性。

**关键词:** 变压器; 故障诊断; 油色谱; 神经网络; BP 算法

**Abstract:** BP neural network has the self-learning and adaptive ability, so it is very suitable for fault diagnosis of transformer. BP algorithm with momentum can smooth out the oscillations in the network training trajectory, and BP algorithm with variable learning rate can change the learning rate adaptively according to network error surface. Every algorithm can speed up the convergence, but the principles of these two algorithms are different. The two algorithms are combined to become a BP algorithm with momentum and variable learning rate. The modified BP algorithm has the merits of the two algorithms, so it can smooth out the oscillations and speed up convergence of the network training. The artificial neural network for fault diagnosis of transformer is constructed, and is trained by the modified algorithm. The results verify the effectiveness of the proposed algorithm.

**Key words:** transformer; fault diagnosis; chromatogram of transformer oil; neural network; BP algorithm

**中图分类号:** TM855 **文献标志码:** A **文章编号:** 1003-6954(2010)01-0030-03

## 0 前 言

大型电力变压器是电力系统中至关重要的电气设备, 搞好变压器的故障诊断对于电力系统安全稳定运行具有重要意义。油中溶解气体分析方法是诊断电力变压器故障的有效手段<sup>[1]</sup>, 它对诊断变压器内部故障十分有效。BP 神经网络由于具有形成非线性映射、自学习、自适应、鲁棒性等优点, 在基于油中溶解气体分析的变压器故障诊断中得到了广泛的应用<sup>[2-3]</sup>。BP 神经网络广泛使用的标准 BP 算法是一种按照梯度规则下降的算法, 它存在着收敛速度慢、容易陷入局部极小点等缺点。

针对 BP 算法的这些缺点, 人们提出了各种改进算法, 加动量项 BP 算法和变学习速率 BP 算法就是其中的代表。加动量项 BP 算法能够平滑网络的学习曲线, 变学习速率 BP 算法能根据网络误差曲面的情况自适应调整网络的学习步长, 这两种算法都能加

速网络的收敛, 但是它们的原理是不同的。通过结合了这两种算法, 形成的 BP 算法不仅加入动量项而且学习速率自适应变化, 该算法在避免网络训练振荡的同时提高了网络的收敛速度。

## 1 加动量项 BP 算法和变学习速率 BP 算法

### 1.1 BP 算法的原理及存在的不足

BP 神经网络的结构如图 1 所示。

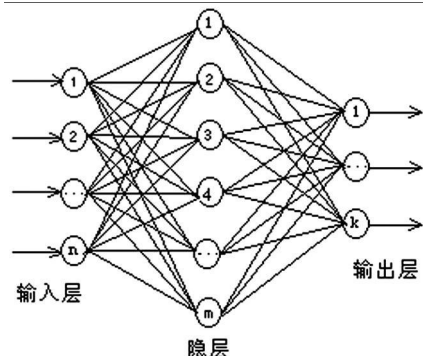


图 1 BP 神经网络的结构

BP神经网络由输入层、隐层和输出层组成。网络的隐层可以为一层或多层,但是含有一个隐层的BP网络已经能以任意精度完成任何连续函数的映射,故一般只采用一个隐含层<sup>[4]</sup>。网络可变参数由层间神经元的连接权值和神经元的偏置组成,网络的学习过程由信号的正向传播和误差的反向传播构成,学习的最终目的是通过修改网络的可变参数使网络能满足要求。在信号的正向传播过程中输入层的信号经过计算得到隐层的输出信号,此后信号逐层向前传播到达输出层;如果输出层的信号未能达到指定的误差要求,则输出信号的误差将沿网络的连通路经反向传播至输入层,沿途修改网络各层间的连接权值和神经元的偏置。BP算法每次迭代网络可变参数的调整量如下。

$$\Delta w_{ij}(k) = -\eta \cdot \frac{\partial e}{\partial w_{ij}}(k) \quad (1)$$

式(1)中,  $\eta$ 是学习步长;  $\frac{\partial e}{\partial w_{ij}}(k)$ 是可变参数对网络误差的一阶偏导数;  $k$ 是迭代次数。

BP神经网络的误差函数是一个以网络可变参数为自变量、网络误差为因变量的非线性多元函数,它是多维空间中的一个形状极为复杂的曲面。网络的误差曲面随可变参数变化十分剧烈,有些区域误差曲面非常平滑,此时应该选择较大的学习步长;有些区域误差曲面非常陡峭,此时应该选择较小的学习步长。但是标准BP算法在训练中的学习步长是恒定的,从而导致了学习步长选择困难,学习步长过大则网络容易振荡;学习步长过小则每次网络可变参数调整量过小,也会导致网络收敛速度变慢。而且网络误差曲面上存在着很多局部极小点,在局部极小点上网络误差距全局极小点上网络误差较大,但可变参数对网络误差的一阶偏导数接近于0。由于每次迭代网络可变参数的调整量是可变参数对网络误差的一阶偏导数与学习步长的乘积,而学习步长是某一较小的常数,它们的乘积仍然是一个较小的值。因此,在局部极小点上网络可变参数的调整量很小,这使网络虽然经过多次训练仍不能跳出局部极小点,训练不收敛<sup>[5]</sup>。

### 1.2 加动量 BP算法

加动量项的BP算法<sup>[6]</sup>是将网络可变参数上一次调整量乘以一个系数作为网络可变参数本次调整量的一部分,这个增加的部分人们通常将它称为动量项。此时网络可变参数的调整量变为

$$\Delta w_{ij}(k+1) = \beta \cdot \Delta w_{ij}(k) - \eta \cdot \frac{\partial e}{\partial w_{ij}}(k+1) \quad (2)$$

式(2)中,  $\beta$ 是动量项因子;  $w$ 是网络可变参数;  $0 < \beta < 1$ ,可取 0.9;  $k$ 是网络训练迭代次数。

当本次的  $\frac{\partial e}{\partial w_{ij}}(k)$ 与上一次同符号时,其加权求和值增大,使  $\Delta w_{ij}(k)$ 较大,从而在稳定调节的同时增加了  $w$ 的调节速度;当本次的  $\frac{\partial e}{\partial w_{ij}}(k)$ 与上一次符号相反时,说明有一定的振荡,此时加权和结果使  $\Delta w_{ij}(k)$ 较小,起到了稳定的作用。动量项的引入平滑了网络的训练路径,提高了网络的收敛速度,而且网络训练的计算量增加很少。

### 1.3 变学习速率 BP算法

标准BP算法在训练过程中学习步长始终保持不变,这未必是最好的选择。当神经网络误差曲面比较平滑时,网络的学习步长应该较大以使网络能尽快收敛;当网络的误差曲面较陡时,网络学习步长应该较小以使网络可变参数的调整量不会过大。变学习速率BP算法<sup>[7]</sup>网络可变参数的调整量如下。

$$\Delta w_{ij}(k) = -\eta(k) \cdot \frac{\partial e}{\partial w_{ij}}(k) \quad (3)$$

$$\eta(k+1) = \begin{cases} \alpha_+ \cdot \eta(k), & \text{当 } e(k+1) < e(k) \\ \alpha_- \cdot \eta(k), & \text{当 } e(k+1) > b \cdot e(k) \\ \eta(k), & \text{其他} \end{cases} \quad (4)$$

式(4)中,  $\alpha_+$ 取 1.05,  $\alpha_-$ 取 0.7,  $b$ 取 1.04。

当网络误差增加时说明调整量过大,网络学习步长变短;当网络误差减少时说明调整量过小,网络学习步长变长;当网络误差不变时,学习步长同样保持不变。所以该算法能根据网络误差曲面的情况实现学习步长的自适应调整,比BP算法具有更好的收敛性能,同时网络训练计算量的增加很少。

## 2 结合加动量项和变学习速率的BP算法

加动量项BP算法由于加入了动量项减轻了网络训练过程中的振荡现象,变学习速率BP算法根据网络误差曲面的情况自适应地调整网络的学习步长,这两种方法都能加速网络的收敛。但从上面的分析中可以看出这两种算法也存在各自的不足:加动量项BP算法不能根据误差曲面的情况自适应调整学习步长;变学习速率BP算法并没有利用动量项来减轻网

络学习曲线中的振荡。将这两种方法结合起来形成既加入动量项又变学习速率的 BP 算法,能结合这两种方法的优点,舍弃它们的缺点。新算法的网络可变参数调整公式如下。

$$w_{ij}(k) = w_{ij}(k-1) + \Delta w_{ij}(k-1) \quad (5)$$

$$\Delta w_{ij}(k) = \beta \cdot \Delta w_{ij}(k-1) - \eta(k) \cdot \frac{\partial e}{\partial w_{ij}}(k) \quad (6)$$

式中的各个变量的意义和取值同式(2)中相应变量相同。

该算法的计算步骤如下。

1)用较小的随机数对网络可变参数进行初始化,确定动量项因子  $\beta$  和初始学习步长  $\eta(0)$ ,  $k=0$ , 计算  $e(0)$ ;

2)计算网络可变参数对网络误差的一阶偏导数

$$\frac{\partial e}{\partial w_{ij}}(k);$$

3)根据式(6)计算网络可变参数的调整量  $\Delta w_{ij}(k)$ , 根据式(5)调整网络的可变参数;

4)  $k=k+1$ , 根据式(4)调整学习步长  $\eta(k)$ ;

5)计算网络误差,如果误差小于目标则训练结束;否则跳至步骤 2)继续训练。

### 3 神经网络应用于变压器故障诊断

#### 3.1 训练性能比较

采用结构为 5-20-4 的三层前馈神经网络对故障变压器油色谱样本进行训练,输入层 5 个神经元分别表示 5 种气体,即  $H_2$ 、 $CH_4$ 、 $C_2H_6$ 、 $C_2H_4$ 、 $C_2H_2$  占气体总浓度的比值;网络的输出层 4 个神经元分别表示网络能诊断的 4 种故障类型,即低温过热、高温过热、低能放电和高能放电。样本数目为 54 个,允许最多训练代数数为 40 000,误差目标为 4。分别采用加动量项 BP 算法、变学习速率 BP 算法和结合加动量项和变学习速率的 BP 算法训练,3 种算法的收敛速度如图 2 所示。

加动量项 BP 算法训练了 40 000 次仍不能收敛于误差目标,最终训练误差在 8 附近;变学习速率算法训练了 3 万多次网络收敛了,但学习曲线振荡严重;改进算法在训练进行到 10 000 次左右收敛到误差目标,而且学习曲线平滑,是 3 种方法中最好的。

#### 3.2 故障变压器实例诊断

将加动量项 BP 算法、变学习速率 BP 算法和动

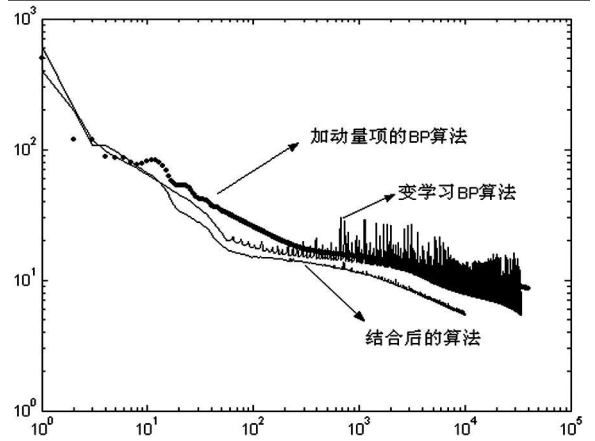


图 2 不同训练算法收敛速度比较

量项和变学习速率的 BP 算法训练所得的神经网络分别用于两个变压器油色谱故障诊断,色谱数据如表 1 所示。

表 1 变压器油中溶解气体数据

	$H_2$	$CH_4$	$C_2H_6$	$C_2H_4$	$C_2H_2$
1号	278.10	354.10	169.40	486.30	0.57
2号	271	95	22	162	183

3 种算法训练神经网络的诊断结果如表 2 所示。

表 2 神经网络的诊断结果

	低温过热	高温过热	低能放电	高能放电
1号	0.903 1	-0.097 9	0.070 4	0.107 9
2号	-0.117 0	0.023 2	0.236 3	0.843 0

1号变压器吊罩检查故障原因为变压器铁心下夹件工艺螺钉过热、高压 B 相组合导线一股导体开焊和油枕油囊破裂;神经网络诊断为低温过热。2号变压器吊罩检查故障原因为低压 10 kV 相引接片靠铁心侧一片未接在套管上,引起对上铁轭放电造成电弧闪络,单相对地电流较大,放电时被烧毁一片约 16 mm, 10 kV 侧线路另一相短路时造成较大电流;神经网络诊断为高能放电。两个故障变压器实例神经网络均诊断正确。

## 4 结 论

将加动量项和变学习速率的 BP 算法应用于基于油中溶解气体分析的变压器故障诊断,从训练和诊断情况可以看出,改进后的 BP 算法减轻了网络的振荡、加快了网络的收敛速度,可以准确诊断变压器油中溶解气体分析故障。

#### 参考文献

[1] 操敦奎. 变压器油中气体分析与诊断 [Z]. 北京: 中国水

利电力企业管理协会, 1987.

- [2] 王财胜, 孙才新, 廖瑞金. 变压器色谱监测中的 BPNN 故障诊断法 [J]. 中国电机工程学报, 1997, 17(5): 322-325.
- [3] 黄鞠名, 胡明华, 朱子述, 等. BP 网络在基于 DGA 变压器故障诊断中的应用 [J]. 高电压技术, 1996, 22(2): 21-23.
- [4] 史忠植. 神经网络计算 [M]. 北京: 电子工业出版社, 1993.

- [5] 张萍萍, 姚万业, 杨明玉, 等. 基于改进 BP 算法的发电机组振动故障诊断技术 [J]. 华北电力大学, 2003, 30(6): 34-37.
- [6] Rumelhart D E, et al Learning Representation by BP Errors [J]. Nature(London), 1986(7): 149-154.
- [7] R. A. Jacobs Increased Rates of Convergence through Learning Rate Adaptation [J]. Neural Networks 1998, 1(4): 295-308.

(收稿日期: 2009-10-22)

(上接第 24 页)

进行换相时, 因预计关断的阀关断后, 在反向电压期间未能恢复阻断能力, 当加在该阀上的电压为正时, 又重新导通, 即出现了换相失败现象, 直流电流激增到  $2 \text{ pu}$ 。  $t=0.8 \text{ s}$  时清除故障, VDCOL 将参考电流调节到  $0.3 \text{ pu}$ , 经  $0.35 \text{ s}$  后系统恢复。

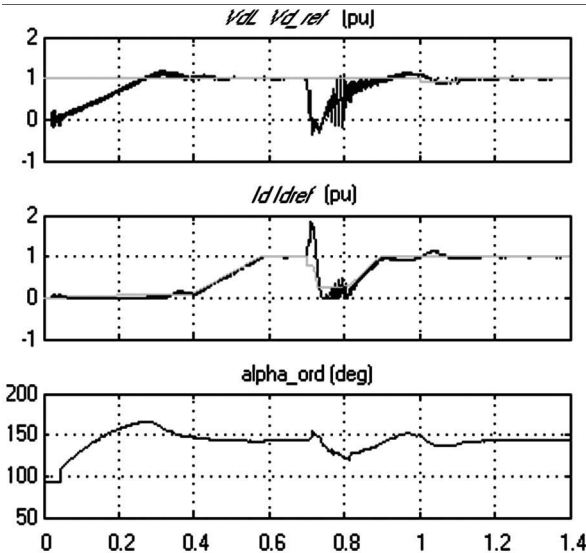


图 8 a 相接地故障波形

### 3.4 逆变侧 a b 两相接地故障

打开断路器模块, 使其在  $0.7 \text{ s}$  时导通,  $0.8 \text{ s}$  时断开, 整流侧交流电压和电流波形如图 9 所示。可见, 两相接地故障使整流侧交流电压和电流发生了一定的扰动, 交流电流的波动频率和幅度比较大。故障时, 交流电流值迅速增加, 之后逐渐下降为 0, 故障清除后经过小幅度波动恢复稳态。同样, 交流电压值在故障时有所增加, 但增加的幅度要小一些, 故障清除后也能够恢复稳态。

## 4 结论

直流输电已成为中国电网的重要组成部分, 所以

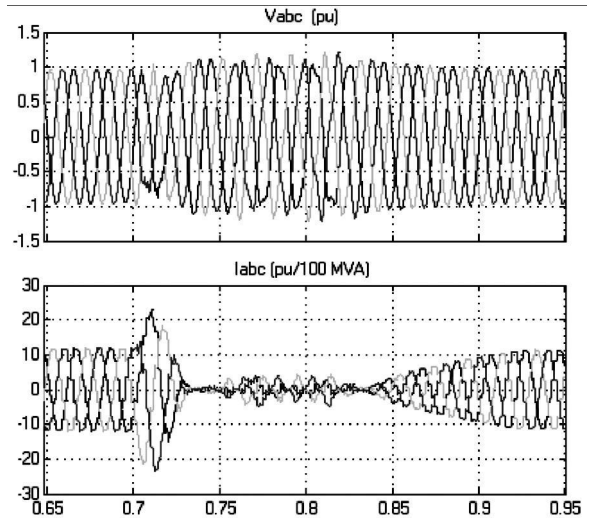


图 9 整流侧交流部分波形

HVDC 系统的暂稳态仿真对电力系统的研究、规划和运行等起着重要作用。前面利用 Matlab 的 Simulink 及 PSB 对高压直流输电系统建模, 并进行了稳态和一系列故障的仿真, 得出相应的波形。可见, 在 Matlab 环境下, 可以进行 HVDC 系统的建模与仿真等工作, 为今后实施德阳-宝鸡  $\pm 500 \text{ kV}$  直流输电工程换流站仿真系统工作奠定了基础。

## 参考文献

- [1] 赵畹君. 高压直流输电工程技术 [M]. 北京: 中国电力出版社, 2004.
- [2] 韩民晓, 文俊, 徐永海. 高压直流输电原理与运行 [M]. 北京: 机械工业出版社, 2008.
- [3] 谢小荣, 姜齐荣. 柔性交流输电系统的原理与应用 [M]. 北京: 清华大学出版社, 2006.
- [4] 卢强, 孙元章. 电力系统非线性控制 [M]. 北京: 科学出版社, 1993.
- [5] 王晶, 翁国庆, 张有兵. 电力系统的 MATLAB/SMULINK 仿真与应用 [M]. 西安: 西安电子科技大学出版社, 2008.

(收稿日期: 2009-10-28)