

文章编号: 1005-8893(2001)03-0013-04

BP 神经网络算法研究及其在燃煤锅炉中应用^{*}

肖立川, 薛国新

(江苏石油化工学院 机械工程系, 江苏 常州 213016)

摘要: 神经网络的种种特性使得它们在动力工程中得到广泛的应用。根据系统的特点研究使用合理的神经网络算法已成为广受关注的课题。BP神经网络是一种最基本的神经网络。国际国内学者围绕如何避免训练陷入局部极小和加速收敛, 提出了多种措施。但以往方法主要基于如何选择合理的学习率或加入适当的动量项, 未能充分利用控制系统的特点。通过分析具有不同极点分布的控制系统 BP神经网络的误差曲面形状, 提出在总体误差的下降方向和其相反方向上同时进行一维搜索的方法。算法较多地考虑了有关的细节性问题。此方法被用于某燃煤锅炉, 结果令人满意。

关键词: BP神经网络; 算法; 燃煤锅炉

中图分类号: TP 393

文献标识码: A

研究并使用智能化方法代表着系统优化控制的一个重要方向。近年来, 许多国际国内学者将专家系统和神经网络方法用于控制系统的动态建模和辨识上, 取得了一系列成果。其中以 BP神经网络的应用最为引人注目。BP神经网络能实现任何非线性连续映照^[1], 并且一旦反映实际问题的网络的权训练好了的话, 该网络的应用就只是较为简单的计算。另外当实际问题的样本发生变化时, 一般网络的结构不变, 只要改变其权即可, 即其权可通可动态学习不断调整。正是由于 BP网络的上述特性, 使得它在自动控制方面得到广泛应用。为了能将 BP神经网络用于在线计算以对系统行为不断作出动态调整, 快速而可靠的训练算法是关键。人们提出了多种改进 BP网络训练算法的措施, 但以往方法主要基于如何选择合理的学习率和加入适当的动量项。以往方法对如何克服误差曲面的非凸性, 尚有其不足之处。显然, 对此作进一步研究有着重要意义。

1 基于一维搜索的 BP神经网络改进训练算法

BP网络的基本算法是一种在梯度法的基础上推导出来的有教师的 δ 学习律, 或称广义 δ 法则。设输入学习样本为 P 个 $x^{(1)}, x^{(2)}, \dots, x^{(P)}$, 已知与其对应的教师为 $t^{(1)}, \dots, t^{(P)}$ 。学习算法是根据实际的输出 $y^{(1)}, \dots, y^{(P)}$ 与 $t^{(1)}, \dots, t^{(P)}$ 的误差 E (称作总体误差) 来修改由其连接权和阈值构成的权向量 w , 使得 $y^{(1)}, \dots, y^{(P)}$ 与 $t^{(1)}, \dots, t^{(P)}$ 尽可能的接近, 这里, 总体误差 E 被看作是权向量 w 的函数, 即 $E=E(w)$ 。每次迭代根据 $E(w)$ 的梯度方向来确定权向量 w 的修正量 Δw 。

为了加快 BP网络训练算法的速度和避免训练陷入局部极小^[2], 以往人们提出了各种措施。但在克服误差曲面非凸性方面, 尚有其不足之处^[3]。为此, 本文研究同时在 $E(w)$ 的下降方向及其相反方向进行一维搜索, 以改进原有训练算法。因为 $E(w)$ 存在局部极小, 所以在得到了 $E(w)$ 的一个下降方向后, 沿其作一维搜索能在一定程度上避免陷入局部极小。在计算中同时在 $E(w)$ 的下降方向及其相反方向进行一维搜索, 发现在好多情

^{*} 收稿日期: 2001-06-14

基金项目: 江苏省科技厅 2000 年度国际合作基金项目 (编号 BS2000730)

作者简介: 肖立川 (1946-), 男, 江苏靖江人, 教授, 硕士, 主要从事热能工程方面的研究。

形下在下降方向射线上 $E(w)$ 的最小值大于其相反方向射线上 $E(w)$ 的最小值。因为 $E(w)$ 不是凸函数, 所以如果假定只沿下降方向进行一维搜索, 那实际上已经用到了凸函数的性质, 故会使全局最小值漏网。即使能找到 $E(w)$ 的全局最小值, 如果在某次迭代中 $E(w)$ 在对应直线上的最小值发生在与其下降方向相反的射线上, 那末必将浪费大量的计算机时。因此, 展开对下降方向相反方向上的搜索有其重要意义。不过, 其具体算法应有别于在下降方向上的搜索算法。

现给出在 $E(w)$ 方向上的一维搜索算法。其基本思想是: 在大范围内先建立起一个网, 搜索出具有最小 $E(w)$ 值的网点, 然后在此网点附近用曲线拟合法估计 $E(w)$ 的最小值。具体算法如下。算法 P:

[P1] 将射线上的搜索范围限定在线段

$$w^{(0)} + t\Delta w \quad (0 \leq t \leq T_0) \quad (1)$$

上。这里 T_0 的选取要考虑到最近各次迭代中 $E(w)$ 在下降方向上的最小值位置;

[P2] 用一系列分点

$$w_0 = w^{(0)}, \quad w_1, \quad \dots, \quad w_n = w^{(0)} + T_0\Delta w \quad (2)$$

对 (1) 式所示线段进行等分, 将这种等分看作是一个搜索 $E(w)$ 在指定线段上最小值的一个网, 求得各分点上的 $E(w)$ 之值;

[P3] 求出 $E(w)$ 在上述各分点中的最小值, 设为 $E(w_{i0})$;

[P4] 分如下三种情形:

情形 (a) $w_{i0} = w_0$ 。因为 Δw 是 $E(w)$ 在 w_0 处的一个下降方向, 故造成此情形显然是由于在 w_0 处所作的分划过粗所致。这时, 应重新在以 w_0, w_1 为端点的线段上进行较细的划分, 求出各新分点上的 $E(w)$ 之值, 然后转 [P3];

情形 (b)

$$w_{i0} = w^{(0)} + T_0\Delta w \quad (3)$$

这时, 若将 (1) 式所示线段延长, 有望得到更小的 $E(w)$ 之值。故增大 T_0 之值, 例如令 $T_0 \leftarrow 1.2 T_0$ 。相应地由 (1) 式所表示的线段被延长, 在延长段上加入新的分点, 其中包括线段端点。计算出 $E(w)$ 在各新分点上的最小值, 转 [P3];

情形 (c) w_{i0} 不是 (1) 式所示线段的任一端点,

这时, 取出包括 w_{i0} 在内的至少三个分点。一般可取五个分点, 并使所取分点除 w_{i0} 外均衡地分布于 w_{i0} 的两侧。但是, 要保证对所取出的分点画出的 $E(w)$ 的离散图不再有除 $E(w_{i0})$ 而外的波峰。进行多项工曲线拟合, 对拟合曲线求出其最小值点 $w^{(*)}$; 如果 $E(w^{(*)}) \leq E(w_{i0})$, 且 $w^{(*)}$ 到与其最近分点的距离小于预定的小正数 δ_0 , 则将 $E(w^{(*)})$ 作为 $E(w)$ 在射线 $w^{(0)} + t\Delta w$ ($0 \leq t \leq T_0$) 上的最小值, 结束计算; 否则, 将 $w^{(*)}$ 作为一个新的分点, 转本情形的起点处。

在最初对 T_0 的选取无经验的情况下, 先取 T_0 为较大数, 将 (1) 式所示线段作较为粗略的等分, 求出 $E(w)$ 在各分点上的值, 求出其最小者, 然后缩小 T_0 , 进一步对 (1) 式所示线段实行分割, 但不再是等分, 而是在 $E(w)$ 取得较小值处有较大的分划密度。

在做完在下降方向上的一维搜索后, 再考虑在其相反方向上进行一维搜索。计算采用与算法 P 类似的思想。但是注意到, 在射线 $w^{(0)} - t\Delta w$ ($t > 0$) 充分接近于 $w^{(0)}$ 的范围内有 $E(w) > E(w^{(0)})$, 因此必须避免在此小范围内作不必要的计算。另外, 在射线 $w^{(0)} - t\Delta w$ ($t > 0$) 上作一维搜索之前, 必须确定其有一定价值, 为此必须存在可能条件使 $E(w)$ 在射线 $w^{(0)} - t\Delta w$ ($t > 0$) 上的最小值可能小于它在射线 $w^{(0)} + t\Delta w$ ($t > 0$) 的最小值。如果根据最近迭代结果看出这种可能性极小, 那末在本次迭代中也就没有必要进行这种搜索。还可假定, 在连续的两次迭代中, $E(w)$ 在射线 $w^{(0)} - t\Delta w$ ($t > 0$) 上的最小值位置表现出连续性, 并由此确定搜索范围。研究总结出算法如下。算法 N:

[N1] 将射线 $w^{(0)} - t\Delta w$ ($t > 0$) 上的搜索范围限定在线段

$$w^{(0)} - t\Delta w \quad (0.5\eta^* \leq t \leq T_1) \quad (4)$$

上。这里 T_1 的选取要考虑到最近各次迭代中 $E(w)$ 在下降方向上的最小位置。其中 η^* 的定义是 $w^{(0)} + \eta^*\Delta w$ 为 $E(w)$ 在射线 $w^{(0)} + t\Delta w$ ($t > 0$) 上的最小值;

[N2] 用一系列分点

$$w_0 = w^{(0)} - 0.5\eta^*\Delta w, \quad w_1, \quad \dots, \quad w_m = w^{(0)} - T_1\Delta w \quad (5)$$

将 (4) 式所表示的线段进行等分。将这种划分看作是一个搜索 $E(w)$ 在指定线段最小值的一个

网；求得各分点上的 $E(w)$ 之值；
[N3] 求出 $E(w)$ 在上述工各分点中的最小值，设为 $E(w_{j0})$ ；
[N4] 分如下三种情形：

情形 (a) $w_{j0} = w^{(0)} - 0.5 \eta^* \Delta w$ 。
这时，令 $\eta^* \leftarrow 0.618 \eta^*$ 。相应地由 (4) 式所表示的线段被延长；在其延长部分上再作进一步的划分，求出这些新分点上的 $E(w)$ 之值，转 [N3]；

情形 (b) $w_{j0} = w^{(0)} - T_1 \Delta w$ 。
这时，增大 T_1 之值，令 $T_1 \leftarrow 1.2 T_1$ 。同上，相应地由 (4) 式所表示的线段也被延长。在延长段上加入新的分点，其中包括由 (4) 式表示的新的线段端点，计算出 $E(w)$ 在各新分点上的最小值，转 [N3]；

情形 (c) w_{j0} 不是 (5) 式所示线段的任一端点。

这时，取出包括 w_{j0} 在内的至少三个分点（有可能则取五个以上分点，所取分点除 w_{j0} 外均衡地分布于 w_{j0} 两侧。但是，要保证对所取出的分点画出的 $E(w)$ 的离散图不再有除 $E(w_{j0})$ 而外的波峰。然后进行多项式曲线拟合，对拟合曲线求出其最小值点 $w^{(*)}$ ，最小值为 $E(w^{(*)})$ 。

如果 $E(w^{(*)}) >$ 线段 (4) 上的 $E(w)$ 的最小值，则放弃搜索，结束；如果 $E(w^{(*)}) \leq E(w_{j0})$ 且 $w^{(*)}$ 到与其最近分点的距离小于预先取定的小正数 δ_0 ，则将 $E(w^{(*)})$ 作为 $E(w)$ 在线段 (4) 上的最小值，结束计算；否则，将 $w^{(*)}$ 作为一个新的分点，转本情形的起点处。

自然地，取出线段 (1) 和线段 (5) 上的 $E(w)$ 的最小点作为下次迭代的起点。

2 应用实例

考虑某热电厂 410T/H 燃煤锅炉蒸汽温度的计算。这时，有关操作条件有：给水流量，给水温度，同操器开度，给煤机转速，引风机挡板开度，送风机挡板开度。根据测量得出运行过程中给水流量和给水温度随时间的变化规律如下：

$$F_{in} = F_{in}(t) \tag{6}$$

$$T_{in} = T_{in}(t) \tag{7}$$

由所有的操作条件和以上两式产生输入样本。采用一个隐层，使用本文方法进行训练，求得各权值。在主频为 133 兆的 PC 机上计算，训练所需机时为 65 秒。另外，使用常用方法对同一 BP 神经网络进行训练，所需机时为 259 秒。可见本文方法具有较快的训练速度。根据训练所得的 BP 神经网络的各权计算出各输出，计算结果列于表 1。表 1 还给出了运行资料值。由表 1 可见，本文的 BP 神经网络的训练方法是可靠的，且总体精度优于传统的训练方法。

表 1 运行温度和计算温度的比较		
运行资料值	蒸汽温度/℃	
	BP 网络计算值	
	传统方法	本文方法
551	549	550
549	546	548
542	540	543
536	533	537
530	528	530
524	521	525
518	516	518
512	510	511
508	509	508

3 结 论

为了克服 BP 神经网络误差曲面的非凸性所带来的训练上的困难，算法的处理细节十分重要。本文提出的一维搜索算法，充分考虑了各种可能情形，依据情形的不同采取不同的具体计算步骤，从而能节省计算机时。BP 神经网络是一种能用于自控系统的可靠方法。

参考文献:

[1] Rohani S, Haeri M, Wood H C. Modeling and Control of a Continuous Crystallization Process; Part 1. Linear and Non-linear Modeling [J]. Computers & Chemical Engineering, 1999, 23: 263—277.
[2] Voge T, Wangs I. Accelerating and Convergence of the Back Propagation Method [J]. Biol Cybernetics, 1998, 59: 257—261.
[3] Baldi P, Hornik K. Neural Networks and Principal Component Analysis: Learning from Examples Without Local Minimum [J]. Neural Networks, 1989, 2: 53—58.

A Study on the Algorithm of Neural Networks and its Application in a Coal—burning Boiler

XIAO Li—chuan, XUE Guo—xin

(Department of Mechanical Engineering, Jiangsu Institute of Petrochemical Technology, Changzhou 213016, China)

Abstract: Neural networks have found wide application in power engineering because of their special functions. The use of reasonable algorithms of neural networks for the control of a system according to its characteristics has been a subject for study attracting wide attention. A BP neural network is one of the basic type. Researchers at home and abroad have proposed a lot of methods to overcome the difficulty of local minimum points encountered in the training process of a BP neural network and to accelerate the training speed. These methods are based on the choice of a better learning rate or appending of a more appropriate momentum term. They didn't take advantage of the characteristics of control systems. Analyzing the appearances of error cambers of the BP neural networks of control systems with different limit points distributions, this paper presents an improved algorithm which performs one dimension searching in the descent direction and the opposite direction at the same time. The improved algorithm takes much consideration on related details. The application to a coal—burning boiler shows a great success.

Key words: BP neural networks; algorithm; coal—burning boiler

简 讯

会 议 信 息

2001 年 9 月 21 日 ~23 日, 江苏省化学化工学会高分子专业委员会和江苏省塑料加工工业协会, 在承办单位江苏石油化工学院召开了“纳米及超细技术在高聚物中的应用技术研讨交流会”, 会议邀请了中科院化学所、中国石油化工集团总公司、山东大学、南京大学、华东理工大学、南京理工大学、南京工业大学等国内高等院校科研院所的专家学者作了技术报告, 并和有关企业进行了纳米材料的制备与应用技术交流。

参加会议的 100 余位代表, 就纳米科技的研究现状, 纳米材料的特性, 主要的制造方法及应用范围等进行了热烈友好的研讨。与会专家一致认为纳米技术将对 21 世纪的信息科学、生命科学、分子生物学、新材料科学以及军事科学等领域的发展提供了一个新的技术基础, 并将推动 21 世纪的又一次产业革命。

材料科学与工程系