第28卷 第5期
2002年9月
自 动 化 学 报
ACTA AUTOMATICA SINICA
Vol. 28, No. 5
Sept., 2002

研究简报

神经网络输出两阶段优化及其应用[1)]

邢进生[1]　万百五[2]　冯祖仁[2]

[1](西安交通大学管理学院　西安　710071)

[2](西安交通大学系统工程研究所　西安　710049)

(E-mail: jsxing@xjtu.edu.cn)

关键词　神经网络,优化,热轧产品,质量模型

THE TWO-STAGES OPTIMIZATION OF NEURAL NETWORK OUTPUT AND ITS APPLICATION

XING Jin-Sheng[1]　WAN Bai-Wu[2]　FENG Zu-Ren[2]

[1](*College of Management, Xi'an Jiaotong University, Xi'an*　710071)

[2](*Institute of Systems Engineering, Xi'an Jiaotong University, Xi'an*　710049)

(E-mail: jsxing@xjtu.edu.cn)

Key words　Neural network, optimization, hot rolling mill production, quality model

1　引言

随着计算机在生产控制中的广泛应用,钢铁企业在其生产过程中必然搜集了大量的热轧数据. 这些数据实际上是企业的经验和知识积累,当其积累到一定程度时,必然会反映出生产和决策的规律性. 同时,在大多数情况下用户对热轧产品质量的要求往往越高越好. 这就要求钢铁企业借助于已有的热轧产品质量的数据通过数据挖掘得出一种生产方案,使其热轧产品的质量达到最优. 解决这一问题的思路是,首先利用已有的热轧产品质量的数据建立神经网络质量模型[1],其输出是产品质量指标(如断裂延伸率、曲服强度等),然后对所建的神经网络模型输出进行优化,即找出在某一区域中使神经网络输出达到极值的输入变量.

2　神经网络输出优化

2.1　梯度法寻优

以 $y=N(\boldsymbol{x})$ 表示已经建立的 BP 神经网络模型,其中 $\boldsymbol{x}\in[0,1]^n$ 是 n 维输入向量,y 为

1) 国家"863"高技术研究发展计划(863-51-945-011)资助

收稿日期　2000-05-09　　收修改稿日期　2001-04-10

标量输出. 用 $x_j^{(k)}$ 表示输入向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 中的第 j 个分量 x_j 在第 k 次迭代时的取值，此时梯度方向为 $(\delta_1(k),\delta_2(k),\cdots,\delta_n(k))\triangleq\left(-\frac{\partial y}{\partial x_1},-\frac{\partial y}{\partial x_2},\cdots,-\frac{\partial y}{\partial x_n}\right)\Bigg|_{(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)})}$，$\alpha$ 表示学习率. 假如在第 $k+1$ 次迭代后得到的输入向量是 $\boldsymbol{x}^{(k+1)}=(x_1^{(k+1)},x_2^{(k+1)},\cdots,x_n^{(k+1)})\notin[0,1]^n$，由实数的稠密性必可找到某一自然数 N_k，使得 $\left(x_1^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_1(k),x_2^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_2(k),\cdots,x_{n_1}^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_n(k)\right)\in[0,1]^n$，此时用 $\frac{\alpha}{2^{N_k}}$ 代替学习率 α 作为新的学习率. 用梯度算法寻找神经网络输出的最优值的步骤如下：

1) 置 $k=0$，在输入向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 的允许范围内，依上述标准随机选取其初值 $\boldsymbol{x}^{(k)}=(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)})$；

2) 计算 $\delta_i(k)$ 和 $x_i^{(k)}+\alpha\cdot\delta_i(k),i=1,2,\cdots,n$；

3) 若 $\left(x_1^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_1(k),x_2^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_2(k),\cdots,x_{n_1}^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_n(k)\right)\in[0,1]^n$，令 $x_i^{(k+1)}=x_i^{(k)}+\alpha\cdot\delta_i(k)(i=1,2,\cdots,n)$；否则，取 N_k 使得 $\left(x_1^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_1(k),x_2^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_2(k),\cdots,x_{n_1}^{(k)}+\frac{\alpha}{2^{N_k}}\cdot\delta_n(k)\right)\in[0,1]^n$，并置用 $\alpha=\frac{\alpha}{2^{N_k}}$，$x_i^{(k+1)}=x_i^{(k)}+\alpha\cdot\delta_i(k)$，$i=1,2,\cdots,n$；

4) 置 $k=k+1$，并由 $\boldsymbol{x}^{(k)}=(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)})$ 计算神经网络的输出值 $y^{(k)}$；

5) 若 $y^{(k)}$ 达到质量要求，转 6)；否则转 2)；

6) 输出 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 的终值. 停止.

2.2 随机寻优

随机寻优算法的步骤如下：

1) 选梯度算法所得到的最后的输入向量为初始的输入向量 $\boldsymbol{x}^{(0)}=(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)})\in[0,1]^n$；

2) 随机取 $\boldsymbol{x}^{(1)}=(x_1^{(1)},x_2^{(1)},\cdots,x_{n_1}^{(1)})\in[0,1]^n$，计算神经网络的输出值 $y^{(1)}=N(\boldsymbol{x}^{(1)})$；

3) 确定 $\boldsymbol{x}'$ 使得 $N(\boldsymbol{x}')=\max(N(\boldsymbol{x}^{(0)}),N(\boldsymbol{x}^{(1)}))$，置 $k=2$；

4) 选取输入向量 $\boldsymbol{x}^{(k)}\in[0,1]^n\cap o(\boldsymbol{x}',r)$，计算神经网络的输出值 $y^{(k)}=N(\boldsymbol{x}^{(k)})$；

5) 确定 $\boldsymbol{x}'$ 使得 $N(\boldsymbol{x}')=\max(N(x'),N(\boldsymbol{x}^{(k)}))$；

6) 若 $N(\boldsymbol{x}')$ 已达到质量指标的要求，转 7)；否则置 $k=k+1$，然后转 4)；

7) 输出 $\boldsymbol{x}'$. 停止.

下面定理保证了上述随机寻优法一定能找到极值.

定理. 随机寻优算法中，序列 $\{N(\boldsymbol{x}^{(k)})\}$ 是收敛的，且收敛到神经网络输出函数的极值点.

证明. 若 $\{N(\boldsymbol{x}^{(k)})\}$ 中仅有有限个互不相同的数，则由其构造知从某一项，不妨设从第 K 项起，以后所有的项都相同. 由算法知，在 $\boldsymbol{x}^{(K)}$ 的半径为 r 的邻域中经无限次的随机搜索仍找不到比 $N(\boldsymbol{x}^{(K)})$ 更大的值，因此集合

$$\Delta=\{\boldsymbol{x}\in[0,1]^{n_1}\cap o(\boldsymbol{x}^{(K)},r)\,|\,N(\boldsymbol{x})>N(\boldsymbol{x}^{(K)})\}$$

的勒贝格测度为 0. 由神经网络函数的连续性 $o(\boldsymbol{x}^{(K)},r)$，在不可能有比 $N(\boldsymbol{x}^{(K)})$ 更大的神经网络函数值(否则在该点的某一邻域中的神经网络函数值均大于 $N(\boldsymbol{x}^{(K)})$). 于是 $\{N(\boldsymbol{x}^{(k)})\}$

收敛到极值点.

若$\{N(\boldsymbol{x}^{(k)})\}$是由无限多个不同的数构成的序列,由算法知,$\{N(\boldsymbol{x}^{(k)})\}$是单调递增序列,而神经网络作为闭区域$[0,1]^n$内的连续函数,必然可以达到其最大值(不妨设为$M$),于是有对于任何$k=1,2,\cdots$,总有$N(\boldsymbol{x}^{(k)})\leq M$. 因此$\{N(\boldsymbol{x}^{(k)})\}$是单调有界序列,从而必然收敛. 另外,如果$\{N(\boldsymbol{x}^{(k)})\}$的极限值不是神经网络函数的极值,迭代将继续进行,因此$\{N(\boldsymbol{x}^{(k)})\}$必然收敛到神经网络函数的极值点. 证毕.

3 在多辊热连轧产品质量控制中的应用

某大型钢铁企业生产的一种热轧产品板材,其质量板材断裂延伸率$el1$可由个炼钢及轧制等参数确定[2]. 取代表这种产品质量的实测样本4928个,样本中$el1$的最大值为55,最小值为13. 将各参数规一化在[0,1],将质量$el1$"不足归一"在[0.2, 0.8]内,并利用这些规一化或"不足归一"化处理的数据建立BP神经网络质量模型. 模型的结构采用$32\times60\times1$型. 经过10 000步的迭代可获得神经网络的权值和阈值,从而确定了热轧产品质量的BP神经网络质量模型.

欲将热轧产品质量指标的最大值55和最小值13分别"不足归一"为0.8和0.2,设所求的变换式为$y=ax+b$,其中a,b为待定系数,y为原来的热轧产品质量指标值,x为其"不足归一"值,可得出变换式$y=70x-1$.

采用上述的神经网络输出的梯度算法,经6 000步迭代得到了神经网络输入向量,其相应的神经网络输出值为0.87,相应的热轧产品质量指标值为$el1=70\times0.87-1=59.9$.

在上述结果的基础上利用随机搜索算法继续进行寻优,经4 000步的迭代,神经网络输出值达到0.94,相应的热轧产品质量指标值为$el1=70\times0.94-1=64.8$.

可见,采用随机搜索算法寻优可使热轧产品质量指标值得到进一步改善.

4 结论

本文所给出的产品质量神经网络模型输出的优化可近似等同于生产方案的优化,而优化算法由于在第二阶段采用了随机寻优算法,其收敛性已得到保证,因此两阶段优化算法的收敛是自然的. 从文中例子中可以看出,这种算法对产品质量神经网络模型输出的优化的确具有很好的效果,有提高产品质量的作用.

参 考 文 献

1 万维汉,万百五,杨金义. 闪速炉的神经网络冰镍质量模型与稳态优化研究. 自动化学报, 1999, **25**(6):800~803

2 刘 介,孙一康. 带钢热连轧计算机控制. 北京:机械工业出版社, 1997

邢进生 西安交通大学系统工程研究所博士生. 主要研究领域为智能控制、数据挖掘等.

万百五 简介见本刊第25卷第1期.