

基于填充函数算法的工业产品 小波神经网络质量模型¹⁾

李换琴¹ 万百五²

¹(西安交通大学理学院 西安 710049)

²(西安交通大学系统工程研究所 西安 710049)

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

摘要 为避免反传学习(BP)算法易于落入局部极小点,该文提出一种基于新填充函数的小波神经网络全局优化学习算法,用来解决连铸连轧过程的产品质量建模问题.该过程很复杂,影响其产品性能的因素很多,物理模型难以建立.该文以小波神经网络为模型,建立连铸连轧产品质量与其化学成分和轧制参数之间的复杂非线性模型.该模型用来对板材产品的断裂延伸率、屈服强度等质量性能指标进行预测.数值实验表明:所建立的模型拟合与校验命中率较高,能够较好地预测产品的物理性能.

关键词 小波神经网络,填充函数算法,全局优化,质量模型

中图分类号 TP321

A Wavelet Neural Network Model for Industrial Product Quality Based on Filled Function Algorithm

LI Huan-Qin¹ WAN Bai-Wu²

¹(Faculty of Science, Xi'an Jiaotong University, Xi'an 710049)

²(Systems Engineering Institute, Xi'an Jiaotong University, Xi'an 710049)

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

Abstract To avoid local minimum solutions in the back propagation learning, a new global optimization algorithm based on filled function is proposed to model the product quality of continuous casting furnace and hot rolling mill. The industrial process is very complicated and the number of parameters which determine the final properties can be quite large. It is extremely difficult to develop a physical model for predicting various properties like elongation and yield and tensile strengths. In the present work, a wavelet neural network has been employed to develop a quantitative method for estimating the elongation and yield and tensile strengths as a function of steel composition and rolling parameters. Experimental studies demonstrate that the predicted mechanical proper-

1) 国家“863”计划资助项目(863-51-945-011)、国家自然科学基金资助项目(60274055)和西安交通大学自然科学基金资助项目(0900-573024)

Supported by National “863” Program(863-51-945-011) and National Natural Science Foundation of P. R. China (60274055) and the Natural Science Foundation of Xi'an Jiaotong University(0900-573024)

收稿日期 2002-09-02 收修改稿日期 2003-08-11

Received September 2, 2002; in revised form August 11, 2003

ties have a good agreement with the measured data by using the developed wavelet network model.

Key words Wavelet neural network, filled function algorithm, global optimization, quality model

1 引言

大型加工生产线产品质量控制问题,一直缺乏有效的方法^[1].神经网络建模方法的出现和成功应用,给产品质量控制问题揭开新的一页^[2].小波网络由于将小波良好的时-频局部特性和神经网络的自学习、自适应性相结合,对于某些非线性函数的逼近,性能优于同等规模的其它神经网络^[3].但小波网络常用的学习算法是反向传播(BP)算法,由于BP算法是通过用梯度下降方法来最小化二次误差目标函数推导出来的,BP算法以及各种改进的BP算法均为局部优化算法,学习结果与初始权值的选择有关.近年来,各种随机型全局优化算法被相继提出,如遗传算法、模拟退火法、随机搜索方法和下山单纯形搜索方法^[4~7].理论上,这些方法的随机性保证了在给定计算时间内得到最优解的概率非零,然而实际上,为了得到给定问题的最优解,将要花费非常可观的计算时间.与此同时这些方法大部分是基于经验的,其解析解很难求得.文献[8]在最优化理论中提出了一种确定型全局优化的填充函数法.同随机型方法相比较,确定型法具有寻优速度快、优化效果好的特点,是全局优化理论的一个重要研究方向.作者首次将这种方法引入到神经网络学习中,提出一种训练小波神经网络全局优化新算法,并将其应用于建立产品质量模型.

2 高维输入小波神经网络

不失一般性,考虑具有一个输出的三层前向高维小波神经网络,其输出可表示为

$$\hat{y} = \hat{f}(x, \theta) = \sum_{i=1}^m w_i \psi(d_i x + t_i) + b$$

这里, $x = [x_1, x_2, \dots, x_n]^T$ 是网络的输入, m 为隐层节点数, w_i 是隐层与输出层神经元的连接权重, b 是输出层神经元的偏置权重, d_{ji} 是输入层与隐层神经元的连接权重, 它表示小波的膨胀系数, t_j 是隐层神经元的平移参数. $\theta \in R^S$ ($S = m \times p + m \times n + m + p$) 是所有权重组成的向量, $d_j = [d_{j1}, d_{j2}, \dots, d_{jn}]^T$, $\psi \in C^1: R \rightarrow R$ 是 4 阶 B 样条小波函数.

假设有 N 组训练样本, 第 k 组样本对应的网络输出和期望输出分别为 $\hat{y}(k, \theta)$ 和 $y(k)$, $e(k, \theta) = \hat{y}(k, \theta) - y(k)$ 表示误差. 定义小波网络的误差目标函数为

$$E(\theta) = \frac{1}{2} \sum_{k=1}^N e^2(k, \theta)$$

网络学习的目的就是寻找合适的 θ 使误差目标函数 $E(\theta)$ 最小.

3 填充函数法

填充函数法的基本思想是, 根据目标函数 $F(x)$ 的一个已知极小点 x_1 构造一个函数

$P(x)$, 称为填充函数. 它在 x_1 处取得极大值, 而在任何比 x_1 的盆 B_1 高的盆中没有极小点或鞍点, 但在某个比 x_1 的盆低的盆中有一个极小点或鞍点 x^* , 然后以 x^* 作为初始点去极小化 $F(x)$, 并找到 $F(x)$ 的一个新的极小点 x_2 , 使得 $F(x_2) \leq F(x_1)$. 用 x_2 代替 x_1 , 重复上述过程, 直到找到 $F(x)$ 的全局极小点.

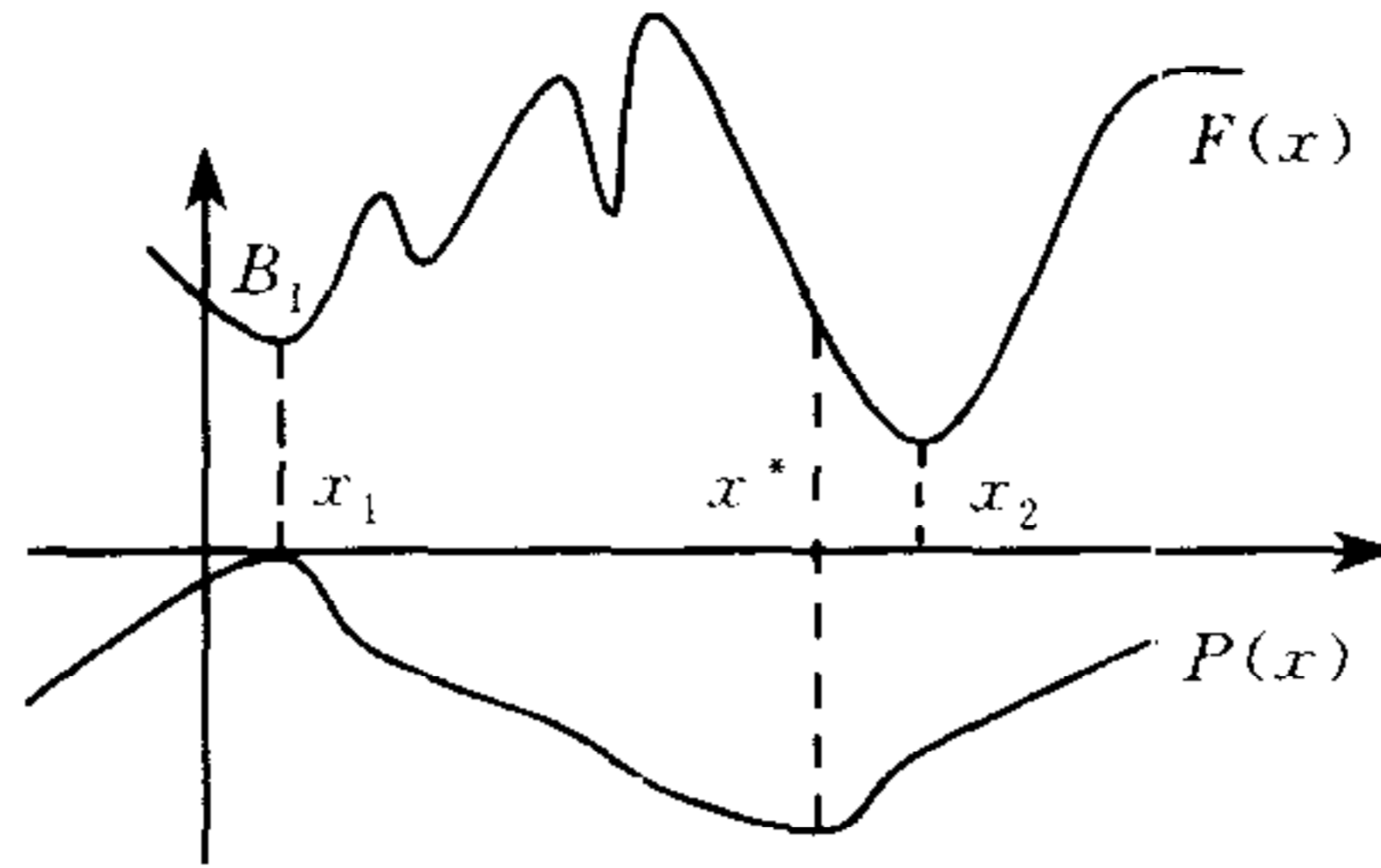


图 1 一维情况下填充函数示意图
Fig. 1 Graph of the filled function

在已有的填充函数中, 常用的有以下三种^[8]:

$$P(x, r, \rho) = \exp(-\|x - x_1\|^2 / \rho^2) / [r + F(x)]$$

$$G(x, r, \rho) = -\{\rho^2 \ln[r + F(x)] + \|x - x_1\|^\rho\}$$

$$Q(x, a) = -[F(x) - F(x_1)] \exp(a\|x - x_1\|^\rho)$$

这里, ρ 取 1 或 2, r, ρ 以及 $a > 0$ 是计算过程中需要调整的参数. 前两个为双参数, 调节起来比较麻烦、费时, 最后一个为单参数, 可是只有当 a 很大时, 才能成为 x_1 处的填充函数, 由于参数 a 在指数上, 当 a 很大时计算容易产生溢出, 计算性能不理想. 作者构造了一种新的填充函数:

$$L(x, a) = -(\sqrt[3]{F(x) - F(x_1)} + a\|x - x_1\|)$$

该填充函数为单参数, 且参数不在指数上, 易于调节和计算. 以下说明 $L(x, a)$ 满足填充函数的条件.

首先, 当 $a > 0$ 时, 对 $\forall x \in B_1$ 有 $L(x_1, a) = 0 > L(x, a)$ 即 x_1 是 $L(x, a)$ 的一个极大点.

其次, 若 $F(x) > F(x_1)$, d 是满足以下条件的方向

$$\begin{cases} d^T \nabla F(x) \geq 0 \\ d^T (x - x_1) > 0 \end{cases} \quad \text{或} \quad \begin{cases} d^T \nabla F(x) > 0 \\ d^T (x - x_1) \geq 0 \end{cases}$$

则有
$$d^T \nabla L(x, a) = -\left\{ \frac{d^T \nabla F(x)}{3 \sqrt[3]{[F(x) - F(x_1)]^2}} + a \frac{d^T (x - x_1)}{\|x - x_1\|} \right\} < 0$$

这就是说在 x_1 的盆以及高于 x_1 的盆中, 若在 x 处 $F(x)$ 沿方向 d 上升, 则 $L(x, a)$ 沿方向 d 下降.

再次, 若 $F(x) > F(x_1)$, d 是满足以下条件的方向

$$d^T \nabla F(x) < 0, \quad d^T (x - x_1) > 0$$

记 $a_x = -\frac{d^T \nabla F(x)}{3 \sqrt[3]{[F(x) - F(x_1)]^2} d^T (x - x_1)}$, 则当 $a > a_x$, $d^T \nabla L(x, a) < 0$, $L(x, a)$ 在 x 处沿方向 d 下降. 当 $a < a_x$, $d^T \nabla L(x, a) > 0$, $L(x, a)$ 在 x 处沿方向 d 上升.

也就是说, 在高于 x_1 的盆中, 如果 $a > a_x$, 则在 $F(x)$ 沿方向 d 下降的点 x 处, $L(x, a)$ 沿方向 d 仍然下降. 如果 $a < a_x$, 则在 $F(x)$ 沿方向 d 下降的点 x 处, $L(x, a)$ 沿方向 d 上升. 因

为 $F(x) > F(x_1)$, 且 $F(x) \rightarrow F(x_1)$ 时, $a_x \rightarrow +\infty$. 所以在比 x_1 的盆低的盆中, 一定存在点 x , 不论 a 取何确定的正数, 总有 $a < a_x$. 此时 $L(x, a)$ 沿方向 d 上升, 从而形成 $L(x, a)$ 的一个静止点.

利用填充函数法训练神经网络的算法归纳如下:

第 1 步: 给参数 a 取一个初始值(比如 $a=100$), $h=0.001$, 在 $[0, 1]$ 区间取很小的随机数作为初始权值, 利用 BP 算法得到 $E(\theta)$ 的一个局部极小点 θ_1 ;

第 2 步: 如果 $E(\theta_1) < h$, 结束训练, θ_1 为最优权值. 否则转第 3 步;

第 3 步: 构造 θ_1 处的填充函数:

$$L(\theta, a) = -[E(\theta) - E(\theta_1)]^{1/3} + a\|\theta - \theta_1\|$$

第 4 步: 在 θ_1 附近取一点 θ_0 , 比如 $\theta_0 = \theta_1 + \delta e_j$, 其中 δ 是很小的正数(例如 $\delta=0.05$), e_j 为第 j 分量为 1, 其余分量全为 0 的单位向量. 从 θ_0 出发, 极小化 $L(\theta, a)$, 得到 $L(\theta, a)$ 的极小点 $\hat{\theta}$, $\hat{\theta}$ 满足 $L(\hat{\theta}, a) \geq 0$ 或 $(\hat{\theta} - \theta_1)^T \nabla L(\hat{\theta}, a) \geq 0$ 或 $\|\nabla L(\hat{\theta}, a)\| \leq \epsilon$ 判据之一. 其中 ϵ 是事先给定的精度;

第 5 步: 以 $\hat{\theta}$ 为初始点, 极小化 $E(\theta)$, 得到 $E(\theta)$ 的极小点 θ_2 ;

第 6 步: 如果 $E(\theta_2) \leq E(\theta_1)$, 令 $\theta_1 \leftarrow \theta_2$, 返回第 2 步;

第 7 步: 如果 $E(\theta_2) > E(\theta_1)$, 放大 a , 例如令 $a \leftarrow 2a$, 返回第 3 步.

4 应用实例

连铸连轧过程是典型大工业生产过程. 成品钢的质量主要与炼钢过程中的 23 个化学元素、加热炉中 2 个参数及轧制过程中 7 个参数的变化有关. 在现场收集到的数据中, 经过数据挖掘选出 8600 组数据. 以这些数据为样本点, 随机选取其中 5000 组为建模样本, 其余 3600 组为校验样本, 以断裂延伸率为质量指标, 实测输入输出数据呈锯齿状, 以高维小波神经网络为模型是较为合理的选择. 本文利用填充函数法训练小波神经网络建立连铸连轧钢的质量模型.

为比较起见, 对 BP 网络、小波网络分别采用梯度算法、填充函数法进行了试验. 对于梯度算法, 随机选取 10 组初始权值, 多次训练, 选用最好的一组结果. 对于填充函数法, 初始权值取为上述 10 组初始权值中的一组. 拟合与校验结果如表 1.

表 1 4 种方法建立板材质量模型的结果比较
Table 1 Modeling results of four methods

	拟合均方误差	拟合命中率(%)	校验均方误差	校验命中率(%)
BP 网络梯度算法	0.0044	66.0	0.0051	58.0
小波网络梯度算法	0.0017	82.3	0.0020	80.5
BP 网络填充函数法	0.0008	84.4	0.0012	82.3
小波网络填充函数法	0.0006	87.1	0.0009	86.2

表中, 命中率指的是网络输出值与实际值相对误差小于 0.05 的样本占总样本的百分比. 可以看出, 就网络模型而言, 小波神经网络优于 BP 网络; 就学习算法而言, 梯度算法, 尽管选取多组初始权值, 多次训练, 但仍难跳出局部极小, 致使网络拟合误差较大, 而填充函数训练算法, 是一种全局优化算法, 它能找到误差函数的全局最小值, 所建立的板材神经网络质量模型的拟合命中率、校验命中率较高.

5 结束语

填充函数法通过构造新填充函数可使小波神经网络训练跳出局部极小,从而提高模型精度. 该法用于连铸连轧产品质量建模,取得较好的建模结果. 所提出的方法可推广用于其他大型加工生产线产品的质量建模,或高维非线性函数的逼近.

References

- 1 Schlang M. Current and future development in neural computation in steel processing. *Control Engineering Practice*, 2001, **9**(9):975~986
- 2 Wan Bai-Wu. Product quality model and quality control model for industries and their application. *Acta Automatic Sinica*, 2002, **28**(6):1019~1024(in Chinese)
- 3 Li Hua-Qin, Wan Bai-Wu. A product quality model for large-scale hot rolling mill based on high-dimension input wavelet neural network. *Systems Engineering*, 2002, **20**(5): 55~58 (in Chinese)
- 4 David J J, Frenzel J F. Training product unit neural networks with genetic algorithms. *IEEE Expert*, 1993, **8**(5): 26~33
- 5 Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by simulated annealing. *Science*, 1983, **220**(4598): 671~680
- 6 Babe N, Mogami Y, Kohzake M, Shiraishi Y, Yoshida Y. A hybrid algorithms for finding the global minimum of error function of neural networks and its applications. *Neural Networks*, 1994, **7**(8): 1253~1265
- 7 Nelder J A, Mead R. A simplex method for function minimization. *The Computer Journal*, 1965, **7**(2):308~313
- 8 Ge R, Qin Y. A class of filled functions for finding global minimizers of a function of several variables. *Journal of Optimization Theory and Applications*, 1987, **54**(2): 241~252

李换琴 西安交通大学理学院副教授. 分别于1986、1999年获西安交通大学计算数学专业学士、硕士学位,2003年获西安交通大学控制科学与工程博士学位. 现主要从事神经网络、小波神经网络及大工业过程产品质量控制研究.

(**LI Huan-Qin** Received her master and Ph. D. degree all from Xi'an Jiaotong University in 1999 and 2003, respectively. Now she is an associate professor in the School of Science at Xi'an Jiaotong University. Her research interests include neural network, wavelet neural network and their application to product quality control.)

万百五 西安交通大学系统工程研究所教授,博士生导师. 目前主要研究方向是大工业过程稳态递解优化理论与应用、智能控制、产品质量控制等.

(**WANG Bai-Wu** Professor in Systems Engineering Institute at Xi'an Jiaotong University. His research interests include theory and application of large-scale industrial processes, steady-state hierarchical optimization and intelligent control of large scale system, and product quality control.)