

以高维输入神经网络作为生产线产品质量模型*

贾磊 万百五 冯祖仁
(西安交通大学系统工程研究所 710049)

摘要 探索用高维输入的神经网络对复杂工业生产过程的建模方法。针对网络输入变量维数较高的特点,提出一种 BP 网络各权重独立训练的分散训练方法。该方法用附加大惯性项来协调各个权重的优化训练,运用非线性优化方法调节步长。与用普通的 BP 训练方法相比,用该方法训练高维输入的 BP 网络具有较快的收敛速度和较高的模型精度,较好地解决了实际生产过程的产品质量模型问题。

关键词 小波网络,径向基网络, BP 网络,最小二乘法,梯度法,类牛顿法

分类号 TP 18

Neural Network Quality Model with High-dimension Inputs for Production-line Products

Jia Lei, Wan Baiwu, Feng Zuren
(Xi'an Jiaotong University)

Abstract An approach is discussed to establish quality model for complicated industrial process. It presents a learning algorithm that each weight of BP neural network is trained individually. Large inertia term is used in the training of each weight in case against the divergence of the algorithm. Some crucial optimization methods are used to adjust the training step length of each weight. Compared with usual BP training algorithm, this algorithm provides a better convergence rate and the obtained BP neural network has a higher accuracy. An actual neural quality network model is established for a complicated industrial production line.

Key words wavelet neural network, radial basis function neural network, BP neural network, least mean square error method, gradient method, Newton-like method

1 引言

神经网络作为一种有效的建模方法,已广泛应用于生产过程。但在实际生产过程中,由于相关的输入变量和干扰变量过多,使得所建立的神经网络模型输入变量维数过高,网络规模过于复杂,从而导致网络训练精度下降,建模效果变差。

基于上述原因,针对热连轧机的质量控制问题,本文提出一种大惯性各权重独立训练的 BP 网络分散训练方法。将该训练方法与传统的 BP 网络训练方法进行比较,可以看出该训练方法在计算速度和

计算精度上都有明显提高。用该训练方法建立的网络模型较成功地解决了连轧机的质量控制建模问题。

2 问题描述

某厂要求建立一个大型热连轧机的板材稳态质量模型,以进行质量预测和质量控制,使输入钢板坯在轧制前就能根据模型大致估计出轧制结果,对不合格的工艺进行质量调节,提高成品率。通过分析,可将大型连轧机的轧制过程分为板坯加热、粗轧、精轧、卷曲和精整等 5 个阶段。经过模型抽象和变量筛选,把整个生产过程抽象为 32 个输入变量和 4 个输出变量的非线性关系,筛选出 32 个输入变量

* 国家 863 计划项目(863-51-945-011)

1999-08-03 收稿,1999-10-25 修回

(包括 23 个化学元素变量和 9 个其它类型变量) 和 4 个板材物理性能输出变量。在现场收集到的数据中, 通过数据挖掘选出 8 600 组数据, 利用这些数据建立连轧机质量的稳态模型, 进行质量预测。

一个 n 输入 m 输出的多输入多输出问题可以分成 m 个 n 输入单输出问题。限于篇幅, 本文只给出一个输出质量指标(钢板断裂延伸率)的神经网络建模过程和结果。

3 径向基网络和小波网络的神经网络建模

径向基神经网络是近年来使用较多的一种单前向网络^[1], 该网络具有简单的网络结构和较快的学习速度, 已成功应用于低维输入的工业生产过程的建模问题^[2]。本文采用文献[2]的方法建立径向基神经网络模型。小波网络也是近年来应用较多的一种网络, 也有成功的工业实际应用^[3]。本文采用文献[3]的方法建立小波网络模型。对上述两种网络选取 5 000 组数据训练, 3 600 组数据测试。经训练后, 拟合不合格率(厂方定义)分别为 64% 和 72%, 两种网络的拟合精度都远远达不到厂方要求(不合格率上限为 20%)。

4 BP 网络的神经网络建模

BP 网络是神经网络中较为成熟的一种网络^[4], 但用其解决输入变量维数较多的建模问题还存在一定困难。本文针对普通 BP 网络训练算法在处理高维输入问题时存在的不足, 引入各权重独立训练并附加大惯性项的 BP 网络分散训练方法对原算法进行改进, 以提高训练速度和模型精度。

建立一个 4 层 BP 神经网络模型对原生产过程进行质量控制。第 1 层是输入层, 由 32 维输入变量组成; 中间层由两个隐层组成: 隐层 1 由 40 个神经元构成, 隐层 2 由 20 个神经元构成; 输出层由一个神经元构成, 输出结果表示断裂延伸率。

在网络训练中采用整体训练原则, 即在每次训练时, 将所有训练样本一次性输入, 计算出网络在现有权重条件下的所有训练样本的误差平方和, 然后根据该均方误差之和来修正各个权系数。网络的初始权系数采用随机赋值的原则。

4.1 普通 BP 的神经网络建模

由普通的 BP 算法表达式可以看出^[4], 该算法在每次修正权系数时, 首先求出所有训练样本的误差

平方和对权系数的负梯度, 然后对所有权系数按统一一步长 α 在负梯度方向上一次性调整。本文对上述基本算法配合以步长的自适应算法, 以提高训练速度。该自适应算法使用了非线性规化中“成功-失败”法的思想, 即

If (本次迭代误差 < 上次迭代误差)

步长 $\alpha =$ 步长 $\alpha + \alpha \times \eta$

Else 步长 $\alpha =$ 步长 $\alpha - \alpha \times \eta$

这里 $\eta = 0.01 / (\text{迭代次数} / 1000 + 1)$ 。

取 8 600 组数据中的 5 000 组训练网络, 3 600 组数据实际测试。经在奔腾 266 微机上 37 个小时 20 000 次迭代, 模型收敛。下面给出用上述普通 BP 算法建模的建模结果:

图 1(a) 和图 1(b) 是采用普通 BP 算法所建立的网络模型的实际拟合结果和实际测试结果。其中, 横坐标表示输入的样本点的序号, 纵坐标表示每输入一个样本点时输出的断裂延伸率值。图 1(a) 网络拟合结果的均方误差为 0.004 4, 拟合不合格率为 34%; 图 1(b) 网络测试结果的均方误差为 0.005 1, 测试不合格率为 42%。显然不能满足工厂提出的模型精度指标。

4.2 适合高维输入 BP 网络的改进 BP 算法

本文提出一种各权重独立训练并附加大惯性项的 BP 网络分散训练方法, 对高维输入的 BP 网络进行训练, 以提高训练精度。高维输入的 BP 神经网络的主要特点是权重过多, 给各种优化工作带来很大困难。而 BP 网络分散训练方法对各权重分别独立训练, 训练时给各权重和各神经元阈值都配有各自独立的步长, 在计算出所有训练样本的误差平方和对各个权重和阈值的导数后, 利用各个权重和神经元阈值各自的导数信息和各自的步长信息, 独立地优化调节各权重和各阈值。在每一权重和阈值的优化调节式中, 采用自适应步长方法和类牛顿法来调节优化步长, 以达到提高网络收敛速度和收敛精度的目的。

因为算法中网络的各个权重和各个阈值是独立调节的, 容易出现某个方向上优化的变化幅度忽然加大, 影响其它方向上的优化, 甚至导致算法发散。为防止这种情况的发生, 在对每个权重和阈值调优时, 在本次权重增量和阈值增量的基础上都附加一个大惯性项, 这个惯性项是由惯性系数乘以上一次的权重增量和阈值增量。

步长的调节是决定网络收敛速度和收敛精度的主要因素, 是算法的核心。网络训练初期采用梯度法

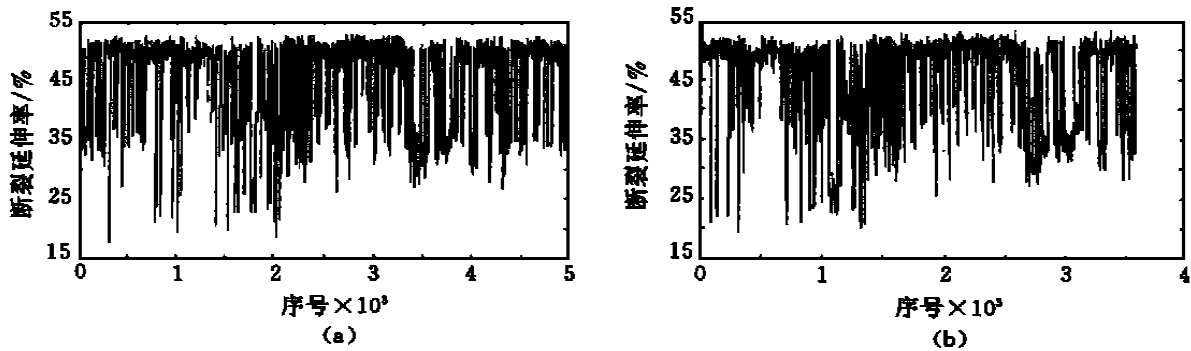


图 1 建模结果

(a) 普通 BP 的拟合结果 (b) 普通 BP 的测试结果

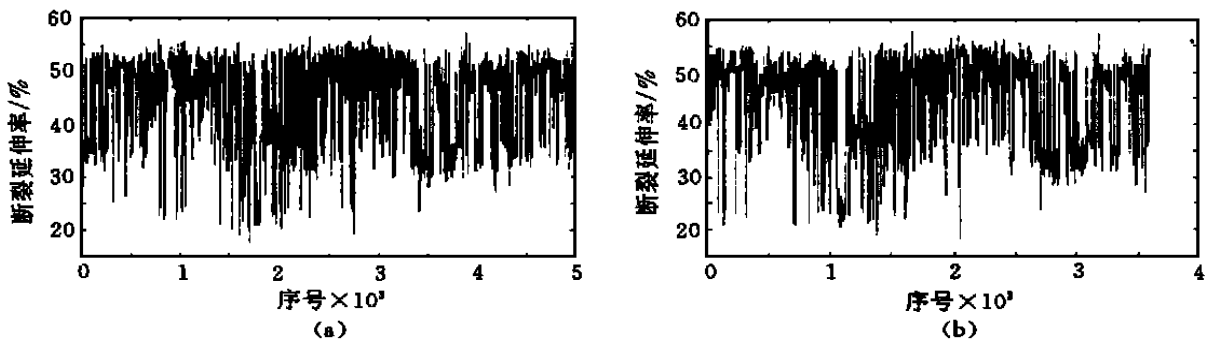


图 2 拟合与测试结果

(a) 改进 BP 的拟合结果 (b) 改进 BP 的测试结果

来调节权重和阈值,梯度法的步长采用“成功-失败”法进行调节。为了避免梯度法在接近收敛点时收敛速度缓慢对网络精度和速度的影响,在网络接近收敛时采用线性插值的方法,以加速网络收敛^[5],提高网络精度。

上述神经网络的权值和阈值的迭代公式可以写成

$$\Delta\omega_j^{(l)}(t) = \alpha_j^{(l)}(t) \times p_{ij}^{(l)}(t) + \eta\Delta\omega_j^{(l)}(t-1) \quad (1)$$

$$\Delta\theta_i^{(l)}(t) = \beta_i^{(l)}(t) \times p_i^{(l)}(t) + \eta\Delta\theta_i^{(l)}(t-1) \quad (2)$$

式中, ω_j 是第 i 个神经元到第 j 个神经元的连接权重; θ 表示第 i 个神经元的阈值; $p_{ij}^{(l)}(t)$ 和 $p_i^{(l)}(t)$ 表示网络在第 t 步修正某个权值和阈值时的迭代方向,此方向是网络误差平方和对该权值和阈值的负梯度方向, $p_{ij}^{(l)}$ 和 $p_i^{(l)}$ 的具体计算可采用一般 BP 网络的训练算法求出; η 是惯性系数,通常是接近于 1 的小数; $\eta\Delta\omega_j^{(l)}(t-1)$ 和 $\eta\Delta\theta_i^{(l)}(t-1)$ 就是附加的大惯性项。

网络训练初期的步长调节算法为

$$\text{if } p_{ij}^{(l)}(t) \times p_{ij}^{(l)}(t-1) > 0$$

$$\begin{aligned} \alpha_j^{(l)}(t) &= \alpha_j^{(l)}(t-1) + \gamma\alpha_j^{(l)}(t-1) \\ \text{Else } \alpha_j^{(l)}(t) &= \alpha_j^{(l)}(t-1) - \gamma\alpha_j^{(l)}(t-1) \\ \text{If } p_i^{(l)}(t) \times p_i^{(l)}(t-1) &> 0 \\ \beta_i^{(l)}(t) &= \beta_i^{(l)}(t-1) + \gamma\beta_i^{(l)}(t-1) \\ \text{Else } \beta_i^{(l)}(t) &= \beta_i^{(l)}(t-1) - \gamma\beta_i^{(l)}(t-1) \end{aligned}$$

其中 γ 是步长变化幅度。

网络接近收敛时的步长调节算法为

$$\begin{aligned} \text{If } (p_{ij}^{(l)}(t) \times p_{ij}^{(l)}(t-1) < 0, \\ |p_{ij}^{(l)}(t)| < 0.001) \\ \omega_j^{(l)}(t+1) &= \\ \omega_j^{(l)}(t) - (\omega_j^{(l)}(t) - \omega_j^{(l)}(t-1)) \times \\ p_{ij}^{(l)}(t) / (p_{ij}^{(l)}(t) - p_{ij}^{(l)}(t-1)) \end{aligned}$$

Else 调用网络初期的自适应步长调节算法

$$\begin{aligned} \text{If } (p_i^{(l)}(t) \times p_i^{(l)}(t-1) < 0, \\ |p_i^{(l)}(t)| < 0.001) \\ \theta_i^{(l)}(t+1) &= \\ \theta_i^{(l)}(t) - (\theta_i^{(l)}(t) - \theta_i^{(l)}(t-1)) \times \\ p_i^{(l)}(t) / (p_i^{(l)}(t) - p_i^{(l)}(t-1)) \end{aligned}$$

Else 调用网络初期的自适应步长调节算法

在网络的基本结构、初始权重、初始步长、训练数据、测试数据都与第 4.1 节相同的情况下,给出建

模结果如下:

经过 33 000 次迭代, 模型收敛. 图 2(a) 和图 2(b) 是网络拟合结果和网络测试结果, 它们较好地逼近了期望拟合结果和期望测试结果. 由此可见, 模型精度有很大提高. 此时, 网络拟合的均方误差为 0.001 9, 拟合不合格率为 14%, 网络测试的均方误差为 0.002 1, 测试不合格率为 16%. 显然, 采用改进 BP 算法训练的网络精度大大高于采用基本 BP 算法训练的网络精度, 且已满足工厂对模型精度指标的要求.

5 结 论

从建模结果可以看出, 采用改进的 BP 算法训练高维输入的神经网络, 建立工业过程的产品质量稳态模型是非常有效的. 与普通的 BP 算法相比, 改进的 BP 算法具有较高的精度和较快的收敛速度. 本文采用这种方法建立的连轧机稳态产品质量模型具有较高的精度, 能起到指导生产的作用.

参 考 文 献

1 Chen S, S A Billings, C F N Cowan *et al*. Practical

identification of NARMAX models using radial basis function networks. *Int J Control*, 1990, 52(6): 1327 ~ 1350

2 王朝利, 曲丽萍, 薛海波, 等. 用 RBF 神经网络建立石墨化炉过程模型. *控制与决策*, 1997, 12(5): 576 ~ 580

3 吕柏权, 李天铎. 基于小波网络的动态系统辨识方法及应用. *控制理论与应用*, 1998, 15(5): 495 ~ 499

4 杨行峻, 郑君里. *神经网络*. 北京: 高等教育出版社, 1992

5 蔡宣三. *最优化与最优控制*. 北京: 清华大学出版社, 1982

作 者 简 介

贾 磊 男, 1975 年生. 1997 年在西北工业大学自动控制专业获学士学位, 现为西安交通大学系统工程专业硕士研究生. 研究方向为大系统的建模和控制.

万百五 男, 1928 年生. 1951 年毕业于交通大学电信研究所, 现为西安交通大学系统工程研究所教授, 博士生导师. 主要研究方向为大系统模型简化, 递阶控制, 智能控制等.

冯祖仁 男, 1953 年生. 1988 年在西安交通大学系统工程专业获博士学位, 现为西安交通大学系统工程研究所所长, 教授, 博士生导师. 目前研究方向为并行结构机器人, 工业生产过程建模及优化控制等.

下 期 要 目

多阶段 CONWIP 系统流通卡分布的确定方法	黄 敏 王兴伟 等
基于模糊目标和模糊约束的满意控制	李少远 席裕庚
基于不变集的一类混合系统的稳定性	翟长连 吴智铭
存贷利率不等下具有随机跳跃收入的消费与投资策略	黄 薇 曹长修 等
一种利用感应电机的鲁棒变结构观测器	凌 强 徐文立 等
时变调参控制系统的稳定裕度	王 勇 高金源
变加权系数减小变结构系统抖振的设计方法	李 涛 冯 勇 等
基于遗传算法的自学习模糊逻辑系统	王宏伦 吕庆凤 等
自适应模糊变结构控制器的一种新设计	张天平
一类非线性不确定组合系统基于观测器的鲁棒分散镇定	魏永德
基于 MTO 管理系统的钢厂合同计划方法	张 涛 王梦光 等
基于逆系统方法的感应电机调速控制系统	张兴华 戴先中
多区域重叠互联电力系统的 H 控制	李小华 陈雪波
基于 GA 和 SA 的制造单元成组方法	李 岩 吴智铭
基于遗传算法的非线性系统时变时滞的在线估计方法	张泳健 周东华
基于 PGA 的全天自主星图识别算法研究	李立宏 宋申民 等
全局指数收敛的机器人 PD 自适应轨迹跟踪	陈启军 王月娟 等