文章编号: 1001-0920(2001)01-0117-03

神经网络建模在热膨胀螺栓形变测量中的应用

陈增强, 袁著祉
(南开大学 自动化系, 天津 300071)

摘　要: 基于神经网络建立热膨胀螺栓形变的非线性数学模型。神经网络的辨识采用变尺度二阶快速学习算法,利用二阶插值法来优化搜索学习速率。新方法具有很快的收敛速度和良好的收敛精度,克服了BP算法在神经网络的权值训练中收敛速度过慢的缺点。热膨胀螺栓的受热形变测量结果表明,该学习算法适用于非线性系统的建模与辨识。

关键词: 神经网络; 非线性系统辨识; 快速二阶学习算法; 二阶插值; 热膨胀螺栓
中图分类号: TP 272　　文献标识码: A

Neural-net-based Modeling Used in the Protraction Measurement for Thermal Expansion Die-bolt

CHEN Zeng-qiang, YUAN Zhu-zhi
(Department of Automation, Nankai University, Tianjin 300071, China)

Abstract: Based upon neural networks, a nonlinear mathematical model of the protraction for the thermal expansion die-bolt is developed. A kind of variable metric fast second order learning algorithm was proposed for the identification of neural network. The second order interpolating method is used in the optimization of learning rate. The new algorithm has fast convergence rate and good precision. Therefore, it overcomes the drawback of BP algorithm which converge too slowly during the weights training of neural network. The method was applied into the protraction measurement for the thermal expansion die-bolt. The research result shows that the method is suitable for the modeling and identification of nonlinear system.

Key words: neural network; nonlinear system identification; fast second order learning algorithm; second order interpolating; thermal expansion die-bolt

1 引　　言

神经网络在控制工程领域的应用具有很大潜力[1, 2]。神经网络对非线性函数良好的逼近能力,使之成为非线性系统建模、辨识和控制中非常重要的技术与方法[3, 4],并逐渐得到广泛的使用。目前,多层前馈网络的权值学习多采用基于梯度下降方向搜索的反向传播(BP)算法。然而,BP算法的缺点是学习速度慢,要得到对系统满意的逼近往往需要许多步,甚至很难达到希望的精度。

收稿日期: 1999-12-03; 修回日期: 2000-02-28
基金项目: 国家863CIMS应用基础研究基金项目(863-511-945-010);天津市自然科学基金项目(983602011); 教育部骨干教师计划项目
作者简介: 陈增强(1964—),男,天津人,教授,博士,从事自适应控制、预测控制等方面的研究;袁著祉(1937—),男,山东青岛人,教授,博士生导师,从事自适应控制、智能控制等方面的研究。

本文将非线性规划的变尺度方法[5,6]应用于神经网络权值的学习,并采用二次插值法(抛物线法)对学习步长 η 进行一维搜索,提出变尺度二阶快速的神经网络学习算法。对非线性系统辨识的结果表明,该算法具有较快的收敛速度和较高的逼近精度。采用此方法建立了涤纶薄膜拉伸生产线中热膨胀螺栓的形变与加热时间关系的非线性数学模型。研究结果表明该方法收敛快,建立的模型精度高。该模型已用于生产线的厚度控制系统,取得了良好的效果。

2 前馈神经网络

假设由一个非线性动态系统得到 p 组输入、输出数据对 (X_j, Y_j), $j = 1, \cdots, p$。其中

$$X_j = [x_j^1, \cdots, x_j^K, 1]^{\mathrm{T}}, \quad Y_j = [y_j^1, \cdots, y_j^M]^{\mathrm{T}}$$

则可用这些数据对一个 3 层前馈神经网络进行学习,从而逼近该动态系统。其逼近模型为

$$\hat{Y}_j = f_{\mathrm{NN}}(X_j, W) \tag{1}$$

其中 W 为网络的权值向量,在网络的学习算法中,通过调节 W 使神经网络模型输出 $\hat{y}_j$ 逼近系统的实际输出 y_j。

用于辨识的 3 层前馈神经网络的输入层节点共 $K + 1$ 个,隐含层节点 L 个,输出层节点 M 个。当将第 j 个样本的原始数据 $X_j = [x_j^1, \cdots, x_j^K, 1]^{\mathrm{T}}$ 输入网络时,相应的隐单元净输入记为 $h_j^l (l = 1, 2, \cdots, L)$,隐单元输出状态记为 $H_j^l (l = 1, 2, \cdots, L)$,网络的最终输出即输出单元的输出记为 $y_j^m (m = 1, 2, \cdots, M)$。$\{w_{kl}^{\mathrm{IH}}: k = 1, 2, \cdots, K + 1; l = 1, 2, \cdots, L\}$ 为输入层节点到隐含层节点的连接权值,$\{w_{lm}^{\mathrm{HO}}: l = 1, 2, \cdots, L + 1; m = 1, 2, \cdots, M\}$ 为隐含层节点到输出层节点的连接权值。则

$$h_j^l = \sum_{k=1}^{K} w_{kl}^{\mathrm{IH}} x_j^k + w_{(K+1)l}^{\mathrm{IH}}, \quad l = 1, 2, \cdots, L \tag{2}$$

$$H_j^l = g(h_j^l), \quad l = 1, 2, \cdots, L \tag{3}$$

$$\hat{y}_j^m = \sum_{l=1}^{L} w_{lm}^{\mathrm{HO}} H_j^l + w_{(L+1)m}^{\mathrm{HO}}, \quad m = 1, 2, \cdots, M \tag{4}$$

其中 $g(\cdot)$ 为 Sigmoid 函数,一般可取 $g(x) = 1/(1 + \mathrm{e}^{-x})$。

3 变尺度二阶快速学习算法

变尺度法是求解无约束极值问题的一种有效方法。由于它既避免了计算二阶导数矩阵及其求逆过程,又比梯度法的收敛速度快,特别是对高维问题具有显著的优越性,因而被认为是求解无约束极值问题最有效的算法之一。多层前馈神经网络的权值学习实质上也是一个无约束极值问题,这就为变尺度方法应用于神经网络的权值学习提供了理论依据。令

$$E(W) = \frac{1}{2} \sum_{j=1}^{p} \sum_{m=1}^{M} (y_j^m - \hat{y}_j^m)^2 \tag{5}$$

则网络权值改变的目标就是要使能量函数 $E(W)$ 尽可能小。

设 $E(W)$ 对 W 的梯度为 $\nabla E(W)$,用变尺度二阶快速学习算法学习权值的计算步骤如下:

Step1: 初始化网络权值 $W^{(0)}$,给定梯度允许误差 $\epsilon > 0$。

Step2: 如果

$$\| \nabla E(W^{(0)}) \|^2 \le \epsilon \tag{6}$$

则 $W^{(0)}$ 即为优化的权值,停止迭代;否则,转向下一步。

Step3: 令

$$V^{(0)} = I(\text{单位阵}) \tag{7}$$

$$P^{(0)} = - V^{(0)} \nabla E(W^{(0)}) \tag{8}$$

在 $P^{(0)}$ 方向进行一维搜索,确定最佳步长 η_0

$$\min_{\eta} E(W^{(0)} + \eta P^{(0)}) = E(W^{(0)} + \eta_0 P^{(0)}) \tag{9}$$

这样便得到下一组权值

$$W^{(1)} = W^{(0)} + \eta_0 P^{(0)} \tag{10}$$

Step4: 设已得权值为 $W^{(k)}$,算出 $\nabla E(W^{(k)})$,若

$$\| \nabla E(W^{(k)}) \|^2 \le \epsilon \tag{11}$$

则 $W^{(k)}$ 即为所求的优化权值,停止迭代:否则,按下式计算 $V^{(k)}$

$$V^{(k)} = V^{(k-1)} + \left[1 + \frac{r_k^{\mathrm{T}} V^{(k-1)} r_k}{s_k^{\mathrm{T}} r_k}\right] \frac{s_k s_k^{\mathrm{T}}}{s_k^{\mathrm{T}} r_k} - \frac{s_k r_k^{\mathrm{T}} V^{(k-1)} + V^{(k-1)} r_k s_k^{\mathrm{T}}}{s_k^{\mathrm{T}} r_k} \tag{12}$$

其中

$$s_k = W^{(k)} - W^{(k-1)} \tag{13}$$

$$r_k = \nabla E(W^{(k)}) - \nabla E(W^{(k-1)}) \tag{14}$$

令

$$P^{(k)} = - V^{(k)} \nabla E(W^{(k)}) \tag{15}$$

在 $P^{(k)}$ 方向进行一维搜索,确定最佳步长 η_k

$$\min_{\eta} E(W^{(k)} + \eta P^{(k)}) = E(W^{(k)} + \eta_k P^{(k)}) \tag{16}$$

这样可得下一组权值

$$W^{(k+1)} = W^{(k)} + \eta_k P^{(k)} \tag{17}$$

Step5: 若 $W^{(k+1)}$ 满足精度要求,则停止迭代学习;否则,转 Step4,直到精度满足要求为止。如果迭代 n 次(n 为权值 $W^{(k)}$ 的维数)仍不收敛,则以 $W^{(n)}$ 为 $W^{(0)}$,以新的 $W^{(0)}$ 为起点重新进行一轮迭代。

上述算法中的一维搜索采用二次插值法(抛物线法),具体算法如下(以第 k 步为例):

令 $\varphi(\eta) = E(W^{(k)} + \eta P^{(k)})$，并假设 $P^{(k)}$ 为单位长向量。计算 $\varphi(1)$：

1) 如果 $\varphi(1) > \varphi(0)$，则对 $\eta = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \cdots$，计算 $\varphi(\eta)$，直到 $\varphi(\eta) < \varphi(0)$。令 $a = 0, b = \eta, c = 2\eta$，并计算 $\hat{\eta}$

$$\hat{\eta} = \frac{1}{2}\frac{[\varphi(a)(c^2 - b^2) + \varphi(b)(a^2 - c^2) + \varphi(c)(b^2 - a^2)]}{[\varphi(a)(c - b) + \varphi(b)(a - c) + \varphi(c)(b - a)]} \tag{18}$$

2) 如果 $\varphi(1) \leq \varphi(0)$，则对 $\eta = 2, 4, 8, \cdots, a, b, c$ 计算 $\varphi(\eta)$，这里 c 表示第一次 $\varphi(c) > \varphi(b)$ 的点；然后同样用式(18) 计算 $\hat{\eta}$。

由 1) 或 2) 得到 b 和 $\hat{\eta}$，如果 $\varphi(\hat{\eta}) < \varphi(b)$，则取 $\eta_k = \hat{\eta}$；如果 $\varphi(\hat{\eta}) \geq \varphi(b)$，则取 $\eta_k = b$。

4　在热膨胀螺栓建模中的应用研究

热膨胀螺栓是涤纶薄膜拉伸生产线中的重要执行机构。多个螺栓分布在生产线的唇口处，通过使其受热发生膨胀来调整涤纶片基的厚度分布。但由于每个螺栓的形变很难在线实时测量，因此只能通过固态继电器控制每个螺栓在一个控制间隔内的加热时间比，即占空度。这就需要建立螺栓形变与加热时间关系的数学模型，这是一个非线性的数学模型。我们用前馈网络来建模，并采用变尺度二阶快速学习算法进行辨识。控制间隔取为 10min，研究在这一间隔内加热时间与螺栓形变的关系。采用 2-4-1 网络结构，输入量为 $X = [x, 1]^T$，x 表示加热时间(单位 s)；输出量 y 表示膨胀形变(单位 μm)。

在离线情况下，用专用仪器测得关于 (x, y) 的 60 个样本。采用本文的学习算法进行训练，网络的初值取为[- 1, 1] 间的随机数，并取步长初值 $\eta = 0.98$。辨识后得到的模型为

$$y(x) = 25.22 - \frac{3.91}{1 + e^{4.07x + 12.25}} - \frac{25.29}{1 + e^{0.15x - 3.21}} + \frac{10.53}{1 + e^{57.58x + 1.34}} + \frac{38.12}{1 + e^{-0.25x + 13.67}}$$

该模型精度很高。模型输出拟合如图 1 所示，图中曲线 ① 为训练出的网络模型的输出；曲线 ② 为对应的实际测量数据。

这一数学模型已应用于薄膜的厚度控制系统，并取得了良好的效果。在该控制系统中，需要对厚片的横向剖面进行有效控制。通过贝塔射线测厚仪在线测得厚片在横向上多个点的厚度值，形成横向剖面。为使该剖面达到理想的剖面形状，通过一定的控制算法，计算出各个热膨胀螺栓应产生的形变位移，然后由该数学模型解算出每个热膨胀螺栓的加热时间，并驱动继电器给螺栓加热。

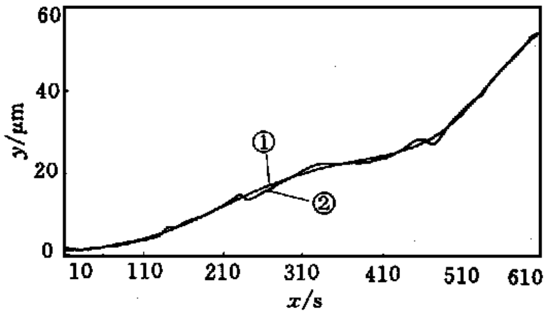

图 1　热膨胀螺栓模型拟合曲线

5　结　　语

由应用研究可以看出，变尺度二阶快速学习算法在神经网络辨识中具有极大的优越性。首先是收敛速度快，一般只需几十步即可达到满意效果，因此可在小样本的情况下训练网络。其次是不需人工选择步长 η，学习效果基本不受学习样本的数量、样本的输出是否归一化及网络初值等客观条件的影响，而且收敛精度高，在对精度要求较高的系统辨识问题中有较大的应用价值。虽然变尺度二阶快速方法每一步的计算量要比 BP 算法的计算量大，但要达到满意精度所需时间仍比 BP 节省许多。这种方法在非线性系统辨识领域具有广阔的应用前景。

参考文献:

[1] Narendra K S, Parthasarthy K. Identification and control of dynamical system using neural networks[J]. IEEE Trans on Neural Networks, 1990, 1(1): 4-27.

[2] Hunt K J, Sbarbaro D, Zbikowski R. Neural networks for control systems —— A survey [J]. Automatica, 1992, 28(2): 1083-1112.

[3] 林茂琼, 陈增强, 贺江峰, 等. 基于阻尼最小二乘法的神经网络自校正一步预测控制器[J]. 控制与决策, 1999, 14(2): 165-168.

[4] 王群仙, 陈增强, 袁著祉. 基于小波网络的非线性系统建模与控制[J]. 控制与决策, 1999, 14(4): 359-363.

[5] 甘应爱. 运筹学[M]. 北京: 清华大学出版社, 1990.

[6] Wismer D A, Chattergy R. 非线性最优化导论(问题求解) [M]. 马正午, 王佩玲译. 北京: 中国展望出版社, 1986.