

基于改进的 PRPE 算法的神经网络建模 及其在工业电加热炉中的应用

古 勇

胡协和

(浙江大学工控所·杭州, 310027) (浙江大学电机系·杭州, 310027)

褚 健

(浙江大学工控所·杭州, 310027)

摘要: 本文改进了多层神经网络的并行递推预报误差(PRPE)算法,极大提高了算法的运行速度,并采用此算法对一非线性多变量耦合的电加热炉的工业实际对象,建立了对象的动态离散神经网络模型,取得了满意的模型拟合效果。

关键词: 非线性系统; 多层神经网络; 建模; 辨识; 递推预报误差法; 电加热炉

1 引 言

实际工业过程具有复杂性、多样性和随机性等特点,使许多过程无法利用机理分析法建立其基本的数学模型结构.特别是实际过程都带有一定的非线性特性,随着对实际系统预测和控制要求的提高,基于工作点上线性化模型已不能满足高标准的要求.因此非线性系统的建模和辨识受到越来越多的重视.现在被应用领域所接受的非线性系统辨识模型,主要是NARMAX多项式模型,但它的缺点是表达非线性特征的能力不强,对于高维、高阶模型存在项数组合爆炸现象.而多层前馈神经网络已经在理论上被证明,即使仅含单隐层,也能任意渐近近似任何连续函数,因此在复杂非线性系统的建模中有很好的前景.

基于预报误差原理,Billings等人提出了多层前馈神经网络的递推预报误差(RPE)算法^[1],取得了比反向传播(BP)算法收敛速度快、预报精度高的效果.但针对多变量系统,神经网络输入输出维数较高,隐层结点数亦较高的情况下,RPE算法涉及高维矩阵的计算,计算量激增,使相对BP算法收敛速度快的优势丧失殆尽,同时内存开销亦激增,限制了隐层结点数的选取,使网络模型与系统固有误差无法进一步减少直至消除,预报精度难以保证.因此,为使预报误差原理得到实际应用,Billings等将网络按隐含节点和输出节点分成若干个子系统,得到了并行递推预报误差(PRPE)算法,大大降低了内存开销和计算量^[2].

本文改进了多层前馈神经网络的PRPE算法,降低了运算维数,而且完全避免了矩阵求逆,进一步降低了计算量.本文采用此算法学习训练工业电加热炉的多层神经网络模型,与BP算法和原PRPE算法比较,具有预报精度高、收敛速度快的双重优势.

2 改进的多层神经网络 PRPE 算法

假设系统可以由以下多变量NARMAX模型来表示

$$Y(t) = f(Y(t-1), \dots, Y(t-n_y), U(t-1), \dots, U(t-n_u)) + e(t). \quad (1)$$

其中 $f(\cdot)$ 是非线性函数. 我们采用多层前馈神经网络来逼近(1)式所描述的非线性系统. 设多层神经网络扩展的时变模型为:

$$\begin{bmatrix} \hat{Y}(t) \\ \Psi(t) \end{bmatrix} = \begin{bmatrix} f(t, \Theta(t-1)) \\ G(t, \Theta(t-1)) \end{bmatrix}. \quad (2)$$

其中 $\Psi(t)$ 为 $\hat{Y}(t)$ 关于参数向量 Θ 的梯度矩阵, 极小化如下的预报误差准则:

$$J_N(\Theta) = \frac{1}{2N} \sum_{t=1}^N \varepsilon^T(t, \Theta) \varepsilon(t, \Theta). \quad (3)$$

其中 $\varepsilon(t)$ 为预报误差向量, N 为数据长度. 完整的递推预报误差算法如下:

$$\varepsilon(t) = Y(t) - \hat{Y}(t), \quad (4)$$

$$\Delta(t) = \alpha_n \Delta(t-1) + \alpha_g \Psi^T(t) \varepsilon(t), \quad (5)$$

$$P(t) = \frac{1}{\lambda} [P(t-1) - P(t-1) \Psi^T(t) (\lambda I + \Psi^T(t) P(t-1) \Psi(t))^{-1} \Psi^T(t) P(t-1)], \quad (6)$$

$$\hat{\Theta}(t) = \hat{\Theta}(t-1) + P(t) \Delta(t). \quad (7)$$

其中 $\Delta(t)$ 为平滑随机梯度.

假设多层神经网络的输出层及隐层共有 q 个节点, 则参数向量 Θ 及梯度矩阵 $\Psi(t, \Theta)$ 可安排成分块形式:

$$\Theta = [\Theta_1 \cdots \Theta_q]^T, \quad \Psi(t, \Theta) = [\Psi_1(t, \Theta) \cdots \Psi_q(t, \Theta)]^T. \quad (8)$$

其中 Θ_i 是第 i 个节点与前层各节点的联系权值及节点神经元阈值组成的子参数向量, $\Psi_i(t, \Theta)$ 是 $n_{\theta_i} \times r$ 维矩阵, 是 $\hat{Y}(t, \Theta)$ 关于 Θ_i 的梯度. 由此针对(8)式决定的各子系统, 分别采用 RPE 算法, 就构成了并行递推预报误差 PRPE 算法.

本文根据 PRPE 算法中梯度矩阵的特定形式, 即 $\Psi_i = uv^T$, u, v 皆是列向量, 而改进了 PRPE 算法, 将矩阵相乘及矩阵与向量相乘分别降维为矩阵与向量相乘及向量与向量相乘, 而且完全避免了矩阵求逆. 实际上, 多层神经前馈网络各节点子系统都可看做是多输入单输出的简单系统. 隐含层节点的活化函数虽然为非线性函数, 但从以下推导知, 其 PRPE 算法形式也与多输入单输出线性系统几乎完全一致. 以单隐层网络为例:

设系统 m 维输入, r 维输出, 令 $n = rn_y + mn_u$.

$$X(t) = [x_1(t) \cdots x_n(t)]^T = [Y^T(t-1) \cdots Y^T(t-n_y), U^T(t-1) \cdots U^T(t-n_u)]^T. \quad (9)$$

为单隐层网络的输入向量. 网络所有权值和阈值参数排列成 n_{θ} 维参数向量:

$$\Theta = [\theta_1 \cdots \theta_{n_{\theta}}]^T, \quad \text{其中 } n_{\theta} = (n+1)n_h + (n_h+1)r.$$

则单隐层网络的扩展模型的具体形式为:

$$\hat{y}_k(t, \Theta) = \sum_{i=1}^{n_h} w_{ki}^o a \left[\sum_{j=1}^{n_h} w_{ij}^h x_j(t) + \mu_i^h \right] + \mu_k^h, \quad 1 \leq k \leq r, \quad (10)$$

$$\Psi_{ij}(t, \Theta) = \frac{d\hat{y}_j(t, \Theta)}{d\theta_i} = \begin{cases} o_{hk}(t), & \theta_i = w_{jk}^o, & 1 \leq k \leq n_h, \\ w_{jk}^o d_{hk}(t), & \theta_i = \mu_k^h, & 1 \leq k \leq n_h, \\ w_{jk}^o d_{hk}(t) x_i(t), & \theta_i = w_{ki}^h, & 1 \leq k \leq n_h, \quad 1 \leq l \leq n, \\ 1, & \theta_i = \mu_k^o, & 1 \leq k \leq r, \\ 0, & \text{Others.} \end{cases} \quad (11)$$

其中隐层节点活化函数取

$$a(z) = \tanh(z) = \frac{1 - \exp(-2z)}{1 + \exp(-2z)}. \quad (12)$$

引入的各标识为: n_h , 为隐节点数; μ_i^h , 为隐含层第 i 个节点的阈值; w_{ij}^h , 为第 j 输入 $x_j(t)$ 与第 i 个隐节点的联系权值; $o_{hi}(t)$, 为第 i 个隐节点的输出; w_{ki}^o , 为第 i 个隐节点与第 k 个输出层节点联系权值; $d_{hi}(t)$, 为第 i 个隐节点活化函数的导数值; μ_i^o , 为输出层第 i 个节点阈值。

根据(10), (11)式, 对应于隐含层第 i 个节点子系统的 Ψ_i 矩阵为:

$$\begin{aligned} \Psi_i(t) &= \begin{bmatrix} w_{i1}^o d_{hi}(t) x_1(t) & w_{i2}^o d_{hi}(t) x_1(t) & \cdots & w_{in}^o d_{hi}(t) x_1(t) \\ \vdots & \vdots & \vdots & \vdots \\ w_{i1}^o d_{hi}(t) x_n(t) & \cdots & \cdots & w_{in}^o d_{hi}(t) x_n(t) \\ w_{i1}^o d_{hi}(t) & \cdots & \cdots & w_{in}^o d_{hi}(t) \end{bmatrix} \\ &= d_{hi}(t) \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \\ 1 \end{bmatrix} \begin{bmatrix} w_{i1}^o & w_{i2}^o & \cdots & w_{in}^o \end{bmatrix} \\ &= d_{hi}(t) X(t) W_i^{(o)T}. \end{aligned} \quad (13)$$

$$\text{令} \quad S_i(t) = \lambda I + \Psi_i^T(t) P_i(t-1) \Psi_i(t). \quad (14)$$

将(13)式代入(14)式得 $S_i(t) = \lambda I + \beta W_i^{(o)T} W_i^{(o)}$, 其中 $\beta = d_{hi}^2(t) X^T(t) P_i(t-1) X(t)$. 由矩阵求逆公式 $(A + BC)^{-1} = A^{-1} - A^{-1} B (I + CA^{-1} B)^{-1} CA^{-1}$ 得

$$S_i^{-1}(t) = \frac{1}{\lambda} I - \gamma W_i^{(o)T} W_i^{(o)}. \quad (15)$$

$$\text{其中} \quad \gamma = \frac{\beta}{\lambda(\lambda + \beta W_i^{(o)T} W_i^{(o)})}.$$

将(15)式代入(6)式得

$$P_i = \frac{1}{\lambda} \left[P_i(t-1) - \frac{P_i(t-1) X(t) X^T(t) P_i(t-1)}{\lambda' + X^T(t) P_i(t-1) X(t)} \right]. \quad (16)$$

其中 $\lambda' = \frac{\lambda}{d_{hi}^2(t) W_i^{(o)T} W_i^{(o)}}$. 显然(16)式的计算相对(6)式大为简洁。

对于输出层第 i 个节点的子系统的 Ψ_i 矩阵为:

$$\Psi_i = \begin{bmatrix} 0 & \cdots & o_{h1} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ \vdots & & o_{hn_h} & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} o_{h1} \\ \vdots \\ o_{hn_h} \\ 1 \end{bmatrix} [0 \quad \cdots \quad 1 \quad \cdots \quad 0] = O_h e_i^T. \quad (17)$$

则(6)式即为:

$$P_i(t) = \frac{1}{\lambda} \left[P_i(t-1) - \frac{P_i(t-1) O_h(t) O_h^T(t) P_i(t-1)}{\lambda + O_h^T(t) P_i(t-1) O_h(t)} \right]. \quad (18)$$

将 PRPE 算法中的(6)式相应换成(16)式或(18)式就构成了改进的 PRPE 算法。

为增强算法的数值稳定性, 对以上算法中 $P_i(t)$ 矩阵的计算引入定迹算法:

$$P_i(t) = \frac{K_0}{\text{trace}[\bar{P}_i(t)]}, \quad K_0 > 0. \quad (19)$$

其中 $\bar{P}_i(t)$ 的计算仍按(16), (18)式.

3 电加热炉建模实例

本文所描述的系统是仙居某厂无纺针织热处理淬火炉^[3], 炉膛形式近似矩形, 由五段加热区组成, 每个区采用一组由可控硅控制的电阻丝加热, 以达到炉内温度分布均匀、稳定的目的, 系统如图 1 所示.

由于该加热炉炉膛结构分布不均匀, 加热电阻丝的功率各有差异及电阻丝加热过程具有死区和非线性特性等, 从而使得对象表现出较强的非线性特性. 同时由

于针的规格不同和操作人员排针的不均匀性等而引起热处理负荷发生较大且频繁的变化, 另外由于工件的连续传送, 引起炉外空气不断地抽入炉内, 对炉内(特别是第一区)温度产生频繁影响, 从而使对象表现较强的随机性. 由于可以测得电加热炉五个区的炉膛温度, 故可将电加热炉看作一个五输入(五个可控硅控制电流信号)/五输出(五个区的炉膛温度)的系统. 因此我们针对的系统是一个大惯性的非线性多变量耦合的随机过程.

令系统的输出输入向量为:

$$Y = [T_1 - \bar{T}_1, T_2 - \bar{T}_2, T_3 - \bar{T}_3, T_4 - \bar{T}_4, T_5 - \bar{T}_5]^T,$$

$$U = [u_1 - \bar{u}_1, u_2 - \bar{u}_2, u_3 - \bar{u}_3, u_4 - \bar{u}_4, u_5 - \bar{u}_5]^T.$$

其中带上划线“—”表示相应变量的稳态工作点. 根据仿真得出经验, 多层神经网络的输入向量取 $X(t) = [Y^T(t-1), Y^T(t-2), U^T(t-1)]^T$ 且隐含层节点选取为 40 时, 能得到收敛速度及拟合精度的较佳折衷. 则电加热炉的多层神经网络模型结构为 15 个输入节点, 40 个隐含节点, 5 个输出节点, 共含有 $n_0 = (15 + 1) \times 40 + (40 + 1) \times 5 = 845$ 个参数. 在一般 486/66 微机上, 辨识如此多参数的 RPE 算法不可能实现, 而 PRPE 算法, 将网络分成了 $q = 40 + 5 = 45$ 个子系统, 极大降低了算法中涉及到的矩阵的维数, 使算法可以在 486/66 微机上实现, 改进 PRPE 算法递推计算 1000 组数据约需 1500 秒, 每步递推运算时间不到 1.5 秒, 比原 PRPE 算法提高速度约 5 倍. BP 算法虽然较 PRPE 算法远为简单, 但其收敛速度及精度都比不上 PRPE 算法, 对于含噪声的实际过程, 模型甚至可能发散.

我们引入误差指标 R^2 作为表征模型拟合精度的指标:

$$R^2 = \frac{\sum_{t=1}^N \epsilon^T(t)\epsilon(t)}{\sum_{t=1}^N Y^T(t)Y(t)}. \quad \text{其中 } N \text{ 为数据长度.}$$

根据炉子实际情况, 数据采样时间间隔为 10 秒, 利用计算机实时采样的 1000 组数据做为神经网络的训练样本. 分别采用 BP 算法、PRPE 算法和改进 PRPE 算法做为学习算法, 比较了各算法的收敛速度及的拟合精度, 网络初始权值及阈值参数取为 -0.25 到 0.25 间的随机数. BP 算法分别取两组 $\alpha_g = 0.001, \alpha_m = 0.8$ (实线) 和 $\alpha_g = 0.1, \alpha_m = 0.8$ (点划线); 改进 PRPE 算法取 $\alpha_g = 0.2, \alpha_m = 0.8$, 并采用鲁棒性强的定迹算法(19)式计算 $P_i(t)$ 矩阵, 仿真结果表明, 在计算隐含层子系统 $P_i(t)$ 矩阵和输出层子系统 $P_i(t)$ 矩阵时, 采用不同参数 K_0 , 更有希望提高收敛速度, 也分别取两组, $K_{0h} = K_{0o} = 10$ (点划线) 和 $K_{0h} = 2, K_{0o} =$

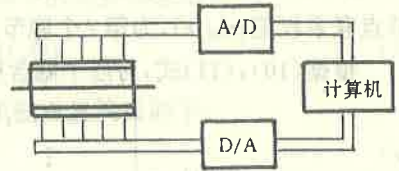


图 1 淬火炉计算机控制系统

20(实线), 结果如图 2 所示.

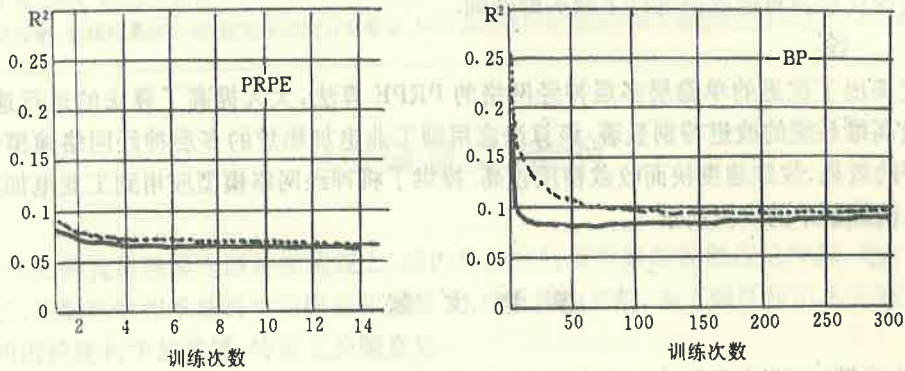


图 2 各算法收敛性比较

从图 2 可以明显看出 PRPE 算法的收敛性质比 BP 算法优越得多. BP 算法在数百次反复学习后趋于发散, 而 PRPE 算法经几次反复学习后就趋于较低的误差限了. 而计算输出层子系统 $P_i(t)$ 矩阵的参数 K_{oi} . 取得较隐含层子系统的 K_{oh} 大一些会提高收敛性质.

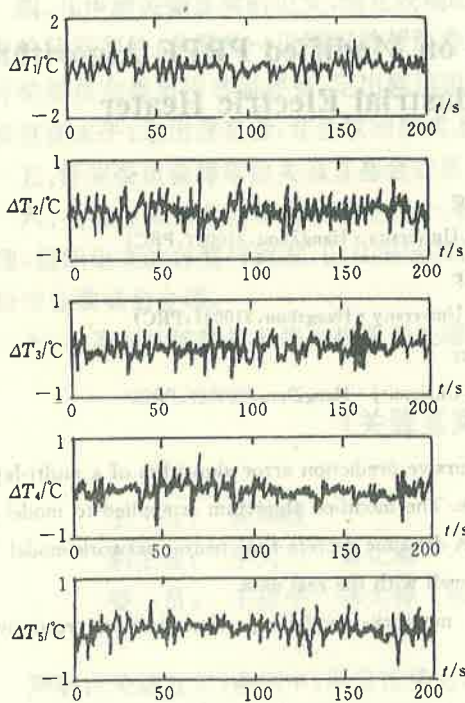


图 3 训练样本拟合误差曲线

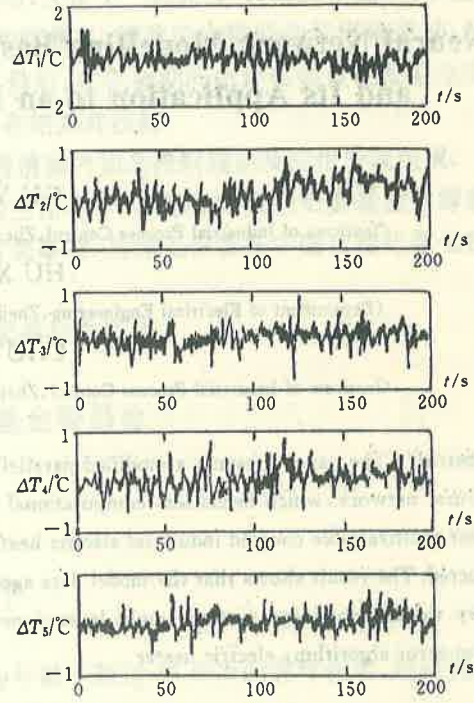


图 4 检验数据拟合误差曲线

我们还利用 200 组未做为训练样本的数据, 检验了多层神经网络模型的有效性; 可扩展性. 图 4 显示了各区检验温度数据的拟合后的误差曲线, 图 3 显示的是训练样本中部分数据的拟合后的误差曲线. 经相关分析, 图 3 所示的误差接近于白噪声特性. 因此, 多层神经网络能建立电加热炉的满意的动态离散模型. 从图 4 可以看出检验数据误差有一定的直流偏差,

但改进的 PRPE 算法也有望实现对模型的在线学习修正,以适应系统随机性的特征,为实现生产过程优化及自适应控制打下坚实的基础。

4 结 论

本文提出了改进的单隐层多层神经网络的 PRPE 算法,大大提高了算法的运行速度,对多变量高维系统的改进特别显著。该算法应用到工业电加热炉的多层神经网络建模中取得了良好的效果,收敛速度快而收敛精度较高,提供了将神经网络模型应用到工业电加热炉多变量系统控制中的广阔前景。

参 考 文 献

- [1] Chen, S. and Billings, S. A. . Non-linear System Identification using Neural Networks. *Int. J. Control*, 1990, 51(6): 1191-1214
- [2] Chen, S. and Billings, S. A. . Neural Networks for Nonlinear Dynamic System Modelling and Identification. *Int. J. Control*, 1992, 56(2): 319-346
- [3] 苏宏业, 褚健, 王骥程. 工业电加热炉非线性模型辨识. *控制与决策*, 1994, 9(4): 281-285

A Neural Network Modelling Based on Modified PRPE Algorithm and Its Application to an Industrial Electric Heater

GU Yong

(Institute of Industrial Process Control, Zhejiang University • HangZhou, 310027, PRC)

HU Xiehe

(Department of Electrical Engineering, Zhejiang University • Hangzhou, 310027, PRC)

CHU Jian

(Institute of Industrial Process Control, Zhejiang University • HangZhou, 310027, PRC)

Abstract: The paper presents a modified parallel recursive prediction error algorithm of a multi-layered neural network, which needs less computational time. The modified algorithm is applied to model a nonlinear multivariable coupled industrial electric heater. A dynamic discrete-time neural network model is constructed. The result shows that the model data agrees well with the real data.

Key words: nonlinear system; multi-layered neural network; modelling; identification; recursive prediction error algorithm; electric heater

本文作者简介

古 勇 1972年生, 1993年毕业于浙江大学, 1996年获浙江大学工学硕士学位, 现于浙江大学工控所攻读博士学位。主要研究兴趣是非线性系统辨识技术、神经网络在工业控制中的应用、非线性系统预测控制。

胡协和 1958年生, 1982年毕业于浙江大学, 1986年获浙江大学工学硕士学位, 后留浙江大学电机系工作, 1992年被聘为浙江大学副教授, 主要从事非线性控制、智能仪表等研究。