

文章编号: 1000-8055(2003) 05-0705-04

前向神经网络快速学习算法在发动机模型 辨识中的应用

王立峰, 李正熙
(北方工业大学 机电工程学院, 北京 100041)

摘要: 介绍一种采用前向神经网络辨识发动机非线性模型的快速学习方法, 与现有同类方法相比, 该方法采用矩阵的 QR 分解求解线性方程组, 解决了维数过高的矩阵直接求逆带来的解失真问题, 并把这种方法应用于航空发动机的非线性模型辨识。算例仿真验证了算法的可行性。

关键词: 航空、航天推进系统; 神经网络; 学习算法; 发动机模型; 模型辨识
中图分类号: V 233. 7 **文献标识码:** A

A Rapid Learning Algorithm of Feedforward Neural Network for Identification of Aeroengine Model

WANG Li-feng, LI Zheng-xi
(North China University of Technology, Beijing 100041, China)

Abstract: To build a precise model is a key technique to aeroengine adaptive control. A rapid learning algorithm of feedforward neural network for identification of aeroengine model is presented in the paper. By contrast with similar algorithms, the present approach employs the decomposition of Matrix QR to solve a set of linear equations. As a result, inversing of huge matrix inverse and fuzzy solutions can be avoided. The algorithm was applied to the aeroengine nonlinear model identification, and validated by the simulation results of the calculation samples.

Key words: aerospace propulsion system; neural networks; learning algorithm;
engine model; model identification

1 引言

航空发动机的自适应控制需要精确的发动机模型。实际飞行过程中, 发动机的部件磨损, 飞行姿态的变化(攻角, 侧滑角), 外界气流扰动等, 都直接影响发动机性能, 故建立良好的发动机模型成为自适应控制的关键技术。本文介绍一种采用前向神经网络辨识发动机非线性模型的方法。

前向神经网络的 BP 算法^[1]存在收敛慢和局部极小点问题, 迄今已经提出很多改进算法, 其中

采用增广卡尔曼滤波的学习算法^[2]引起广泛关注。这类算法特点是: 收敛速度快, 精度高。其基本思想是把网络权值作为一个相应动态系统的状态, 运用增广卡尔曼滤波估计网络权值。然而这种方法由于维数过高且需矩阵求逆, 限制了网络的规模。局部化算法^[3, 4]的提出, 降低了维数, 递推算法^[5, 6]的提出, 避免了矩阵求逆, 这种算法的收敛速度是 BP 算法的100倍以上, 个别情况可达1000倍。本文提出一种基于最小二乘的快速学习算法, 采用矩阵的 QR 分解求解线性方程组, 解决了维

数过高的矩阵直接求逆带来的解失真问题, 并把这种方法应用于航空发动机的非线性模型辨识。

2 前向神经网络和相应的动态系统

考虑一个任意的具有一个输入层、一个输出层和若干隐层的前向网络, 设此网络包含一个由全部连接权构成的权向量 $\theta = [\theta_1 \ \theta_2 \ \cdots \ \theta_m] \in R^M$, 并认为一个单元的阈值是由单位强度输入到连接该单元的连接权。假定网络中单元的激活函数是 S 形函数 $\varphi(x) = K / (1 + e^{-gx})$, 其中 x 为单元的纯输入, 它等于输入到该单元信号的和。现在的问题是要确定 θ 使得对于一组给定的输入向量 $i(k) \in R^n (k = 1, \cdots, p)$, 网络产生的实际输出向量 $o(k) \in R^m (k = 1, \cdots, p)$ 应尽可能接近期望的输出向量 $d(k) \in R^m (k = 1, \cdots, p)$ 。这里考虑在线学习算法, 即输入模式依次输入至输入层, 权值则递推地改进。将 θ 视为如下非线性动态系统的状态:

$$\theta(k+1) \approx \theta(k) \approx \theta_0 \quad (1)$$

$$d(k) = f(\theta, i(k)) + e(k) \quad (2)$$

这里 $f(\theta, i(k)) = o(k, \theta_0)$ 表示网络的输入、权值和输出关系的非线性函数, $e(k)$ 为模型误差。 $f(\theta, i(k))$ 可在当前估计值 $\theta(k-1)$ 附近作展开, 略去高阶余项, 故 (1), (2) 式可改写为:

$$\theta(k+1) \approx \theta(k) \approx \theta_0 \quad (3)$$

$$d(k) = f(\hat{\theta}(k-1), i(k)) + G^T(k) \cdot (\theta - \hat{\theta}(k-1)) + e(k) \quad (4)$$

这里 $\hat{\theta}(k-1)$ 为估计值, $G(k)$ 是 $M \times m$ 维梯度矩阵:

$$G(k) = \left. \frac{\partial f(\theta, i(k))}{\partial \theta} \right|_{\theta = \hat{\theta}(k-1)} \quad (5)$$

展开后可写为:

$$G^T(k) = \begin{bmatrix} \left[\frac{\partial o_1}{\partial \theta_1} \right] & \left[\frac{\partial o_1}{\partial \theta_2} \right] & \cdots & \left[\frac{\partial o_1}{\partial \theta_m} \right] \\ \left[\frac{\partial o_2}{\partial \theta_1} \right] & \left[\frac{\partial o_2}{\partial \theta_2} \right] & \cdots & \left[\frac{\partial o_2}{\partial \theta_m} \right] \\ \vdots & \vdots & \ddots & \vdots \\ \left[\frac{\partial o_m}{\partial \theta_1} \right] & \left[\frac{\partial o_m}{\partial \theta_2} \right] & \cdots & \left[\frac{\partial o_m}{\partial \theta_m} \right] \end{bmatrix} \quad (6)$$

$G(k)$ 中的偏导数可通过反传网络 $\alpha(k)$ 求得。

若设:

$$\xi(k) = f(\hat{\theta}(k-1), i(k)) - G^T(k) \hat{\theta}(k-1) \quad (7)$$

则 (4) 式可简写为:

$$d(k) \approx G^T(k) \theta + \xi(k) + e(k) \quad (8)$$

3 学习算法

基于模型 (8) 给定一个输入/输出序列对 $(i(j), d(j)) (j = 1, \cdots, k)$, 选择如下的目标函数:

$$\begin{aligned} \epsilon(\hat{\theta}(k)) &= \sum_{j=1}^k \|d(j) - f(\hat{\theta}(j), i(j))\|^2 \lambda^{k-j} \\ &= \sum_{j=1}^k \|e(j)\|^2 \lambda^{k-j} = \sum_{j=1}^k \sum_{i=1}^m e_i^2(j) \lambda^{k-j} \\ &= \lambda \epsilon(\hat{\theta}(k-1)) + \|e(k)\|^2 \end{aligned} \quad (9)$$

这里 $\|\cdot\|$ 表示向量的欧几里得范数, $e_i(j)$ 是 (8) 式中第 i 个建模误差, λ 为遗忘因子 $0 < \lambda \leq 1$, 估计值 $\hat{\theta}(k)$ 是由使 $\epsilon(k)$ 达到极小而得的。此解可从下列正则方程推导:

$$\begin{aligned} \nabla_{\theta_0} \epsilon(k) &= -2 \sum_{j=1}^k \sum_{i=1}^m g_i(j) [d_i(j) - \xi_i(j)] \lambda^{k-j} = 0 \end{aligned} \quad (10)$$

其中, $1 \times M$ 维向量 $g_i^T(j)$ 是矩阵 $G^T(j)$ 的第 i 行。由 (10) 式可得:

$$S(k) \hat{\theta}(k) = T(k) \quad (11)$$

其中:

$$\begin{aligned} S(k) &= \sum_{j=1}^k \sum_{i=1}^m g_i(j) g_i^T(j) \lambda^{k-j} = \\ &\lambda S(k-1) + \sum_{i=1}^m g_i(k) g_i^T(k) \end{aligned} \quad (12)$$

$$T(k) = \sum_{j=1}^k \sum_{i=1}^m g_i(j) [d_i(j) - \xi_i(j)] \lambda^{k-j} =$$

$$\lambda T(k-1) + \sum_{i=1}^m g_i(k) [d_i(k) - \xi_i(k)] \quad (13)$$

(11) 式中, 如果 $M \times M$ 维矩阵 $S(k)$ 线性无关, 可用 Householder 变换, 将其进行 QR 分解^[7], $S(k) = QR$ 。其中, 其 Q 阵为正交矩阵, R 阵为上三角矩阵。可推得 QR 方法求解线性方程组如下:

$$R \hat{\theta}(k) = Q^T T(k) \quad (14)$$

由于 R 阵为上三角矩阵, 可以直接求解 $\hat{\theta}(k)$ 。算法初值可取 $S(0) = I$, I 为单位阵 $T(0) = \hat{\theta}(0)$, $\hat{\theta}(0)$ 可为非零的随机向量。使用求解线性方程组, 算法非常稳定^[7], 并解决矩阵病态问题, 也避免了大型矩阵直接求逆。

4 航空发动机模型辨识问题

航空发动机模型具有时变的、非线性特点。动力学模型由转子运动的微分方程构成, 描述如下:

$$\dot{x} = f(x, y) \quad (15)$$

$$\mathbf{y} = \mathbf{y}(\mathbf{x}) \tag{16}$$

其中, $\mathbf{x} = [n_l, n_h]^T$ 状态向量 代表低压转子转速和高压转子转速。输出向量 $\mathbf{y}$ 代表发动机推力、耗油率、转子转速和各截面温度压力等, 这里 $\mathbf{y} = [n_l, n_h, pr_6, tr_6]^T$ 分别代表转子转速、涡轮后压力、温度分别与发动机进口压力、温度之比值。在非加力情况下控制向量 $\mathbf{u} = [w_f, A_8]^T$ 。对于连续系统方程(15)代表两个转子的运动方程, 反映发动机动态特性, 方程(16)代表发动机模型的输出特性。将上述连续系统模型按50 Hz 的采样周期(0.02 s)进行离散化, 模型转化为:

$$\mathbf{x}(k+1) = \mathbf{x}(k) + f(\mathbf{x}(k), \mathbf{u}(k))h \tag{17}$$

$$\mathbf{y}(k+1) = \mathbf{y}(\mathbf{x}(k+1)) \tag{18}$$

那么我们要辨识的发动机模型是下面的形式:

$$\mathbf{y}(k+1) = \mathbf{y}(\mathbf{x}(k)) + f(\mathbf{x}(k), \mathbf{u}(k), \mathbf{u}(k)) \tag{19}$$

模型中4个输入量为:

$$[n_l(k), n_h(k), w_f(k), A_8(k)]^T$$

4个输出量:

$$[n_l(k+1), n_h(k+1), pr_6(k+1), tr_6(k+1)]^T$$

5 仿真算例验证

利用文献[8]建立的发动机部件级的非线性模型进行发动机的加速模拟, 不失一般性, 选择高度0 km, 飞行马赫数 $Ma=0$, 用3 s 左右的时间把发动机从低速状态加速到中间状态, 得到时间长度为6 s 的基本辨识数据, 所有数据都按照方程(19)要求的格式输出, 数据共得到300组, 辨识模型数只取259组, 所有数据都是发动机设计点数据的相对量。

人工神经网络模型取4- 8- 4结构, 既输入层节点为4, 隐含层节点为8, 输出层节点为4; 模型辨识时每前进一步都进行训练, 训练样本数取最近的4个, 即取样本 $(k, k-1, k-2, k-3)$ 时刻的样本, 训练后预报 $k+1$ 时刻输出数据; 训练停止误差准则为 $\epsilon(\hat{\theta}(k)) < 0.0002$, S 函数增益取为 $K=1.15$ 和 $g=1$, 是为了适应输出样本大于1的情况; 遗忘因子取 $\lambda=0.97$, 网络权重取 $[0, 1]$ 之间随机分布值。

在模型辨识启动时, 训练初始的4个数据样本, 图1为训练的误差收敛曲线, 可以看出在达到25个迭代次数后, 满足训练停止的误差准则, 这样

的迭代次数是令人满意的, 其它曲线是输出层的每个节点与目标值的误差随迭代次数的收敛曲线; 以后各个时刻的训练都以当前权重为初始值, 图2为每步训练的迭代次数, 在前进100步左右 ($0.02 \times 100 = 2$ s 情况下), 模型出现迭代次数突然增加后又恢复低迭代次数的水平, 经过多次仿

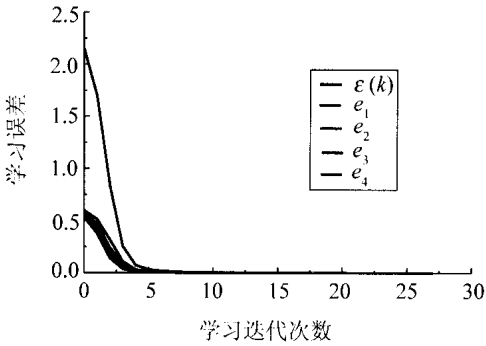


图1 学习误差 vs 学习迭代次数

Fig. 1 Error vs iteration number

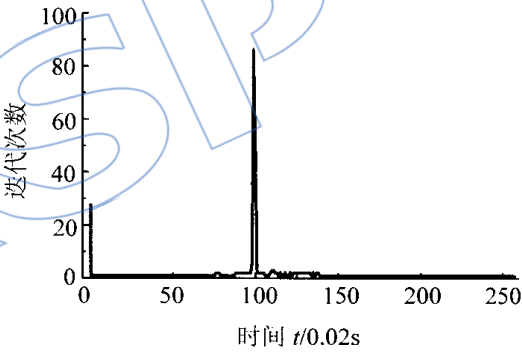


图2 迭代次数 vs 时间进程

Fig. 2 Iteration number vs time

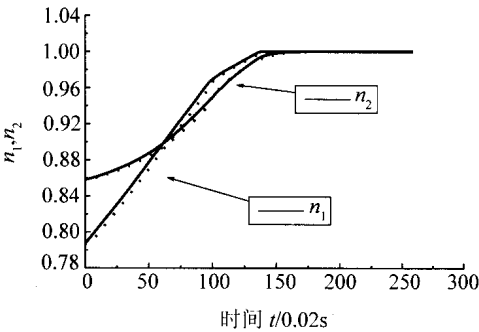


图3 预测值和样本 n_1, n_2 vs 时间 t

Fig. 3 Prediction and Train n_1, n_2 vs time

真试验表明这跟辨识数据变化以及要求的训练误差准则有关; 图3为发动机高低压转子转速的模型预测值与实际的样本数据对比图, 虚线为使用 $(k, k-1, k-2, k-3)$ 时刻数据预报的 $k+1$ 时刻模型

的预测值,实线是样本数据值,看出模型很好的跟踪发动机转速;图4为发动机涡轮后压力、温度与发动机进口压力、温度之比值的模型预测值与实际的样本数据对比图,虚线为使用 $(k, k-1, k-2, k-3)$ 时刻数据预报的 $k+1$ 时刻模型的预测值,实线是样本数据值,看出模型很好的跟踪发动机截面的压力和温度参数。

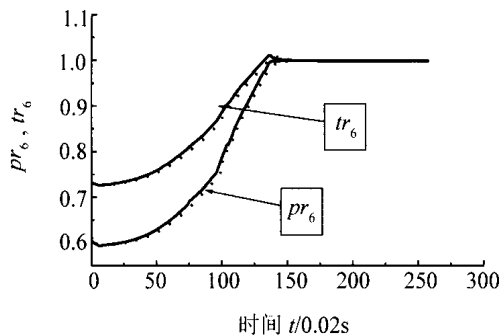


图4 预测值和样本 pr_6, tr_6 vs 时间 t

Fig. 4 Prediction and Real pr_6, tr_6 vs time

仿真的经验表明:样本多少与模型预测精度关系不大,样本多反而带来计算量增大;神经网络的结构简单,则训练速度就快,而且对样本的训练精度影响不大。

6 结 论

文中介绍一种采用前向神经网络辨识发动机非线性模型的快速方法,与现有同类方法相比,该

方法采用矩阵的QR分解求解线性方程组,解决了维数过高的矩阵直接求逆带来的解失真问题,并把这种方法应用于航空发动机的模型辨识。仿真结果表明该方法训练速度快,能够很好的辨识发动机模型,但达到实时性目标还需进一步研究。

参考文献:

- [1] Rumelhart D E, Hinon D E, Williams R J. Learning Internal Representations by Error Propagation, Parallel Distributed Processing: Explorations in the Microstructure of Cognition [M]. Cambridge MA: MIT Press, 1986, 318-362.
- [2] Singhal S, Wu L. Training Feed-Forward Networks with the Extended Karman Algorithm [R]. In: Proc. of the IEEE int. Conf. on Acoustics, Speech and Singal Processing, Glasgow, 1989, 1187-1190.
- [3] Kollas S, Amastasson D. An Adaptive Least Squares Algorithm for the Efficient Training of Artificial Neural Networks [J]. IEEE Trans. Circuits and Systems, 1990, 36: 1092-1101.
- [4] Shah S, Palmieri F, Datum M. Optimal Filtering Algorithms for Fast Learning in Feed-Forward Neural Networks [J]. Neural Networks, 1992, 5: 779-787.
- [5] 王正欧, 林晨. 一种前向一种前向神经网络快速学习算法及其在系统辨识中的应用 [J]. 自动化学报, 1997, 23(6): 728-735.
- [6] 李冬梅, 王正欧. 提高前向神经网络泛化性能和实时性能的新算法 [J]. 电机与控制学报, 2002, 6(3): 241-244.
- [7] 徐士良编. C 常用算法程序集 [M]. 北京: 清华大学出版社, 1996.
- [8] 王立峰, 张津. 战斗机过失速机动飞行条件下综合飞行/推进控制研究 [D]. 北京: 北京航空航天大学, 2000.