

滑动数据窗口驱动的贝叶斯-高斯网络及其 在非线性系统辨识中的应用

刘益剑^{1,2}, 方彦军², 马宝萍¹

(1. 南京师范大学 电气与自动化工程学院, 江苏 南京 210042; 2. 武汉大学 自动化系, 湖北 武汉 430072)

摘要: 工业控制场合中, 需要获取非线性被控对象的结构特性, 而系统动态响应的数据直接从外部特征上反映了非线性系统结构关系. 为了充分利用非线性动态系统响应过程中的数据, 本文提出了一种基于滑动数据窗口(sliding data window)的贝叶斯-高斯神经网络(SW-BGNN)模型. 该模型将数据融合于网络模型结构中, 借助于贝叶斯推理和高斯假设, 利用滑动窗口数据, 实现非线性动态系统的辨识和预测. 整个SW-BGNN本身需要确定的参数很少, 因此运算的时间很短, 适合于非线性动态系统的在线辨识. 将SW-BGNN应用于几个非线性动态系统的辨识和预测, 仿真试验结果表明了SW-BGNN模型的有效性.

关键词: 滑动窗口; 贝叶斯-高斯神经网络; 非线性; 辨识

中图分类号: TP18 **文献标识码:** A

Sliding-data-window-driven Bayesian-Gaussian neural network and its application to modeling of nonlinear system

LIU Yi-jian^{1,2}, FANG Yan-jung², MA Bao-ping¹

(1. School of Electrical & Automation Engineering, Nanjing Normal University, Nanjing Jiangsu 210042, China;

2. Department of Automation, Wuhan University, Wuhan Hubei 430072, China)

Abstract: In industrial control, the structure of the nonlinear dynamic system is determined by using the dynamic data of the controlled object. In order to make full use of the data obtained from the dynamic response process of the nonlinear dynamic system, a novel Bayesian-Gaussian neural network based on sliding-window(SW-BGNN) is proposed which combines the Bayesian reasoning formula with the Gaussian assumption. Based on the data in the sliding window, the operation process of SW-BGNN reasonably predicts the output of the nonlinear dynamic system in terms of a small number of parameters of the SW-BGNN. The SW-BGNN has limited computation time which makes it suitable to on-line identification applications. Examples of identification and prediction of nonlinear dynamic system are presented. Simulation results show the effectiveness of the SW-BGNN method.

Key words: sliding window; Bayesian-Gaussian neural network; nonlinear; identification

1 引言(Introduction)

工业控制中, 绝大多数被控对象都是非线性系统, 因此, 对于非线性系统的辨识和控制问题一直是研究的热点和难点. 通常, 非线性对象的辨识方法一般有机理建模、黑箱建模和灰箱建模3种, 对于系统结构复杂、具有滞后、时变等强非线性的系统, 一般很难通过机理建模方法得到系统精确的数学模型; 大部分的工业非线性动态系统, 人们往往基于系统已知的部分结构特性, 采用黑箱建模和灰箱建模方法, 建立合乎控制精度需要的对象模型.

系统辨识问题往往是由建立一个合适参数化的

辨识模型和通过最优化未知系统辨识和模型输出之间误差的性能函数来调整模型的参数两部分组成^[1]. 神经网络模型在非线性系统辨识中得到了许多应用^[2~4]. 虽然从理论上分析, 任何3层前向神经网络都能够逼近任意的非线性函数, 但是神经网络也存在隐层的确定没有系统方法、权值调整太多、易陷入局部最小等缺点. 此外, 面对结构变化的系统, 训练好的神经网络往往要重新训练才能适应, 因此, 单纯的前向神经网络不适合于时变的动态系统的辨识和预测.

而非线性系统在响应过程中, 大量的输入输出数

据已经从外部表达了非线性动态系统的特征. 从概率论的角度分析, 动态系统的结构特征应该包含于数据的推理关系中. 基于贝叶斯推理和高斯假设, 本文提出了一种基于SW-BGNN推理模型, 将滑动窗口数据融合于推理模型结构中. 只需离线训练个数有限的门槛矩阵参数, 利用滑动窗口中的历史数据, 就能够对系统输出进行在线预测. 当系统的结构发生变化时, 也能够实现良好的跟踪.

2 非线性系统辨识描述(Description of non-linear dynamic system identification)

考虑如下描述的稳定离散时间非线性动态系统

$$y(k) = f[y(k-i); u(k-j)]. \quad (1)$$

式中: $y(k)$ 为第 k 步时的系统输出, $y(k-i)$ ($i = 1, 2, \dots, n$) 代表前 n 步的系统输出; $u(k-j)$ ($j = 0, 1, \dots, m$) 代表当前第 k 步及前 m 步的控制输入; f 表示需要辨识的非线性动态系统的输入和输出函数关系.

3 滑动数据窗口贝叶斯-高斯神经网络

(Bayesian-Gaussian neural network based on sliding-data window)

3.1 贝叶斯-高斯推理模型推导(Derivation of Bayesian-Gaussian reasoning model)

设 (X_i, y_i) , $i = 1, 2, \dots, N$ 为训练的样本集, X_i 为第 i 步的采样输入, 表示为 m 行的列向量, $X_i = (X_{i1}, X_{i2}, \dots, X_{im})^T$, y_i 代表系统的输出, 利用贝叶斯推理和高斯假设, 在新的输入 X 下, 利用概率的方法来产生输出 y .

由高斯假设, Y 具有概率密度函数 $p(Y)$ 并且满足高斯正态分布 $Y \sim N(y_0, \sigma_0^2)$, 其中 y_0 为均值、 σ_0^2 为方差; 假设在 Y 给定的条件下, Y_i 满足正态分布 $Y_i \sim N(Y, \sigma_i^2)$ 并且具有概率密度函数 $p(y_i|Y = y)$, 根据贝叶斯定理, 得到在单一历史数据 (X_i, y_i) 已知时, 新输入 X 产生的输出 Y 的概率如式(2)所示:

$$p(Y|y_i) = c_1 \frac{1}{\sqrt{2\pi}\sigma_{0,i}} e^{-\frac{1}{2} \frac{(Y-y_{0,i})^2}{\sigma_{0,i}^2}}. \quad (2)$$

其中 c_1 为归一化参数, 均值参数 $y_{0,i}$ 和方差 $\sigma_{0,i}^2$ 参数表示如下:

$$\sigma_{0,i}^{-2} = \sigma_0^{-2} + \sigma_i^{-2}, \quad (3)$$

$$y_{0,i} = \sigma_{0,i}^2(\sigma_0^{-2}y_0 + \sigma_i^{-2}y_i). \quad (4)$$

当历史数据样本 (X_i, y_i) , $i = 1, 2, \dots, N$ 已知时, 假设 y_i 对 Y 的先验概率为 $p(Y|y_i)$, 并且 y_i 和 y_j ($i, j = 1, 2, \dots, N, i \neq j$) 在给定 Y 的条件下彼此独立, 则此 N 个数据样本输出对新输入 X 产生的输出 Y 的

条件概率为

$$p(Y|Y_1, Y_2, \dots, Y_N) = k \frac{\prod_{i=1}^N p(Y|Y_i)}{p^{N-1}(Y)}. \quad (5)$$

式中 k 为与 Y 无关的归一化常数, 上式详细证明过程请参考文献[5].

将式(2)代入式(5)中得到

$$p(Y|Y_1, \dots, Y_N) = \frac{c_2 \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_{0,i}} e^{-\frac{1}{2} \frac{(Y-y_{0,i})^2}{\sigma_{0,i}^2}}}{\left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^{N-1} e^{-\frac{N-1}{2} \frac{(Y-y_0)^2}{\sigma_0^2}}}. \quad (6)$$

c_2 是一个独立于 Y 的归一化常数, 由于先验概率分布近似为常数, 因此先验方差 σ_0^2 很大, 在此情况下, 式(3)和式(4)近似为 $\sigma_{0,i}^{-2} = \sigma_i^{-2}$, $y_{0,i} = y_i$, 由高斯假设, 化简式(6)得到

$$p(Y|Y_1, Y_2, \dots, Y_N) = c_3 \frac{1}{\sqrt{2\pi}} \frac{1}{s_i} e^{-\frac{1}{2} \frac{(Y-y_i)^2}{s_i^2}} = c_3 \frac{1}{\sqrt{2\pi}} \prod_{i=1}^N \frac{1}{\sigma_i} e^{-\frac{1}{2} \sum_{i=1}^N \frac{Y^2 - 2y_i Y + y_i^2}{\sigma_i^2}} = c_4 \frac{1}{\sqrt{2\pi}\sigma(N)} e^{-\frac{1}{2} \frac{(Y-y')^2}{\sigma(N)^2}}. \quad (7)$$

上式中 c_4 为与 Y 独立的归一化常数, 估计的均值 $y'(N)$ 和方差 $\sigma(N)$ 分别表示如下式

$$Y'(N) = \sigma(N)^2 \sum_{i=1}^N \sigma_i^{-2} y_i, \quad (8)$$

$$\sigma^{-2}(N) = \sum_{i=1}^N \sigma_i^{-2}. \quad (9)$$

假设方差满足式(10)

$$\sigma_i^2 = \sigma_0^2 e^{(X-X_i)^T D (X-X_i)}. \quad (10)$$

式(10)中 D 称之为门槛矩阵, 为对角矩阵,

$$D = \text{diag}\{d_{11}^{-2}, d_{22}^{-2}, \dots, d_{mm}^{-2}\}. \quad (11)$$

因此, 由式(8)、(9)和(10)构成了整个的贝叶斯-高斯推理模型, 而且整个推理模型的参数主要为门槛矩阵 D 和初始的估计方差 σ_0^2 , 门槛矩阵的维数与输入 X 变量个数相等, 因此由模型结构引入的需要确定的参数就很少, 从而可以大大节省推理模型的运算时间.

基于贝叶斯-高斯推理模型, 设计贝叶斯-高斯神经网络, 神经元节点与常规网络选择相同, 网络总共包含5层^[5], 值得注意的是第2层将门槛矩阵参数包含在激励函数中. 从实验过程分析, 这 N 组历史输入

数据样本对于系统的预测是至关重要的, 本文从减少贝叶斯-高斯神经网络的运算量和在线跟踪系统的动态响应角度出发, 对这 N 组历史数据采用滑动数据窗口的方法选择, 下文将详细描述.

3.2 SW-BGNN工作过程(Working procedure of SW-BGNN)

工作过程总体分为网络的离线训练和在线预测应用两个步骤, BGNN的训练主要是确定阈值矩阵参数 D ; 而在线预测时, 主要为利用 N 组历史样本, 预测系统当前输出, N 组预测样本集采用滑动窗口方法确定, 以上2个步骤分别描述如下.

3.2.1 BGNN的训练(Off-line training of BGNN)

首先得到 N_1 组训练样本 (X_i, y_i) , $i = 1, 2, \dots, N_1$, 利用下面的性能评价函数

$$V_N(D) = \frac{1}{2N_1} \sum_{i=1}^{N_1} (y_i - \hat{y}_i)^2. \quad (12)$$

式中: y_i 代表实际系统输出, $\hat{y}_i$ 为利用其它除 X_i 外的 $N_1 - 1$ 个训练样本, 用式(8)~(10)进行BGNN得到的预测值.

训练的目标是寻找合适的阈值矩阵 D , 使得系统实际输出与预测值达到很好的拟和, 使得式(12)最小或者满足工程应用精度.

上述过程是最小化式(12)的过程, 可以采用基于梯度的最小二乘法、单纯形法等^[6]优化算法, 此外也可以采用近年来得到深入研究和应用的随机进化优化算法^[7~9].

本文采用改进的E.Coli觅食优化算法进行阈值矩阵 D 的训练, 限于篇幅, 算法的具体内容和符号请参考文献[10], 其中对E.Coli个体的评价函数 eval_i^{n+1} 用式(12)的性能评价函数.

3.2.2 基于滑动窗口数据的 BGNN (BGNN based on sliding data window)

阈值矩阵 D 经过上述网络训练后, 可以进行系统在线预测, 虽然文献[5]采用的自适应调整方法, 但当评估的样本个数 N 很大时会带来极大的运算量. 本文利用滑动数据窗口方法来确定BGNN在线预测时的 N 组样本, 是基于这样的思路: 距离当前时刻近的数据对于当前系统输出贡献大, 即距离当前时刻近的数据样本对当前输出能进行较高精度的预测, 实验结果也验证了思路的正确性.

采用滑动数据窗口方法的目的也是维持在线预测估计输出 y 时, 维持BGNN的规模 N 不变, 具体的做法如图1所示:

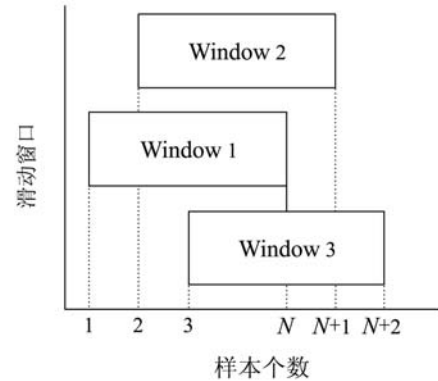


图 1 滑动窗口工作过程示意

Fig. 1 Illustration of sliding window working process

4 仿真试验(Simulation experiments)

采用以下两个非线性动态系统, 验证本文的SW-BGNN的辨识和预测效果.

4.1 非线性系统1(The first nonlinear system)

$$y(t) = \sin x + e. \quad (13)$$

式中 e 为随机白噪声, 试验中取 e 为零均值, 方差为0.2. 系统从320个采样时刻起结构发生变化, 变为 $y(t) = 2\sin x + e$, 到450采样时刻又变为原系统.

系统采样得到600个数据, 输入输出曲线如图2实线所示.

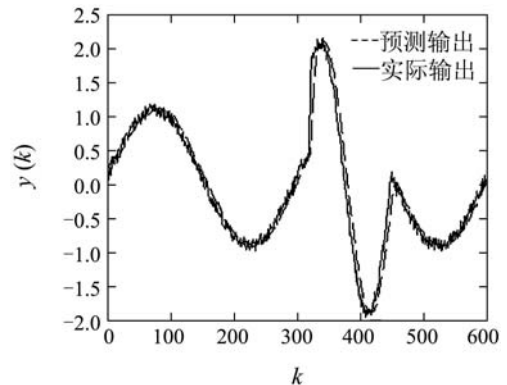


图 2 非线性系统1在线预测

Fig. 2 On-line prediction of the first nonlinear system

首先利用图2所示的前300个样本训练BGNN, 优化算法的初始参数为: $S = 10$, $N_e = 200$, $w_1 = 0.3$, $w_2 = 0.2$, 初始游动步长 $C(i)$ 设为0.02, BGNN的输入为 x , 所以阈值矩阵仅参数 d_{11}^{-1} 需要确定. 经过优化算法优化得到阈值矩阵 $D = [43.3729]$, 性能指标 $V_N(D) = 0.0128$.

利用训练好的阈值矩阵参数, 采用的滑动窗口宽度 N 为10, 进行系统在线预测, 得到如图2虚线所示的跟踪曲线.

4.2 非线性系统2(The second nonlinear system)

采用文献[5]中的非线性动态系统

$$y(t) = (0.8 - 0.5e^{-y^2(t-1)})y(t-1) - (0.3 + 0.9e^{(-y^2(t-1))})y(t-2) + u(t-1) + 0.2u(t-2) + 0.1u(t-1)u(t-2) + e(t). \quad (14)$$

BGNN的输入

$$X = [y(t-1), y(t-2), u(t-1), u(t-2)]^T,$$

因此对应的门槛对角矩阵 $D = \text{diag}\{d_{11}^{-1}, d_{22}^{-1}, d_{33}^{-1}, d_{44}^{-1}\}$ 需要确定4个参数. 控制量 u 为均值为2, 方差为1的随机序列; $e(t)$ 为零均值, 方差为0.04的高斯白噪声.

由式(14)产生100个采样数据作为训练样本, 优化算法的初始参数:

$$S = 20, N_c = 80, w_1 = 0.2, w_2 = 0.2,$$

初始游动步长 $C(i)$ 设为 0.01. 经优化后门槛矩阵 $D = \text{diag}\{28.6618, 45.5806, 36.7235, 27.7467\}$, 对应性能指标 $V_N(D) = 0.0247$.

从第101时刻起, 系统结构发生变化, 输入输出变为式(15):

$$y(t) = (0.8 + 0.5e^{-y^2(t-1)})y(t-1) - (0.3 - 0.9e^{(-y^2(t-1))})y(t-2) + u(t-1) + 0.2u(t-2) + 0.1u(t-1)u(t-2) + e(t). \quad (15)$$

利用训练得到的门槛矩阵 D , 选取滑动窗口宽度 N 为20, 进行系统在线预测, 得到如图3中虚线输出.

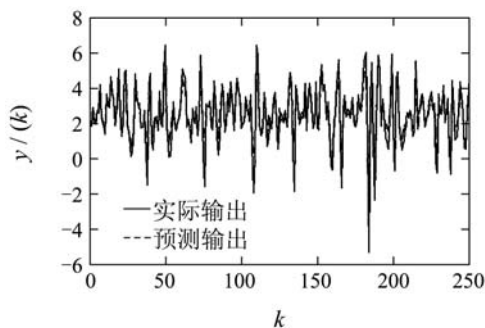


图3 非线性系统2在线预测

Fig. 3 On-line prediction of the second nonlinear system

5 分析与结论(Analysis and conclusion)

从非线性动态系统的仿真实验可出, SW-BGNN能够充分利用窗口数据, 对系统进行良好的在线预测. 从实验也看出, SW-BGNN对非线性动态系统的结构改变也能够很好的适应, 这主要是由

于SW-BGNN将窗口数据融合于其结构中, 通过窗口的不断滑动, 不断更新网络的结构, 从而能够较快的捕捉非线性系统结构的改变, 这个特点对于特性时变的动态系统是比较有吸引力的, 能够适合非线性动态系统的在线预测应用.

参考文献(References):

- [1] 张志华, 郑南宁, 赵振选. 基于模糊对向神经网络的非线性动态系统辨识器[J]. 控制理论与应用, 2000, 17(4): 566 – 568.
(ZHANG Zhihua, ZHENG Nanning, ZHAO Zhenxuan. Identification of nonlinear dynamic systems using fuzzy CP network[J]. *Control Theory & Applications*, 2000, 17(4): 566 – 568.)
- [2] 刘建昌, 陈莹莹, 张瑞友. 基于PSO-BP网络的板形智能控制器[J]. 控制理论与应用, 2007, 24(4): 674 – 678.
(LIU Jianchang, CHEN Yingying, ZHANG Ruiyou. Intelligent flatness-controller based on PSO-BP network[J]. *Control Theory & Applications*, 2007, 24(4): 674 – 678.)
- [3] ZENG Y H, HUAI W X. Application of artificial neural network to predict the friction factor of open channel flow[J]. *Communications in Nonlinear Science and Numerical Simulation*, 2009, 14(5): 2293 – 2300.
- [4] 从秋梅, 柴天佑, 于文. 污水处理过程的递阶神经网络建模[J]. 控制理论与应用, 2009, 26(1): 8 – 14.
(CONG Qiumei, CHAI Tianyou, YU Wen. Modeling wastewater treatment plant via hierarchical neural networks[J]. *Control Theory & Applications*, 2009, 26(1): 8 – 14.)
- [5] YE H W, NI W D. Nonlinear system identification using a Bayesian-Gaussian neural network for predictive control[J]. *Neurocomputing*, 1999, 28(1): 21 – 36.
- [6] DUAN H Y, GAO S Q, JIANG B N, et al. Analysis of a least-square finite element method for the thin plate problem[J]. *Applied Numerical Mathematics*, 2009, 59(5): 976 – 987.
- [7] GUO C Y, HE G, CHEN G J. Application of genetic algorithm in identifying nonlinear regressive model[J]. *Journal of Naval university of Engineering*, 2003, 15(2): 70 – 73.
- [8] ZECCHIN A C, SIMPSON A R, MAIER H R, et al. Parametric study for an ant algorithm applied to water distribution system optimization[J]. *IEEE Transactions on Evolutionary Computation*, 2005, 9(2): 175 – 191.
- [9] GE H W, SUN L, LIANG Y C, et al. An effective PSO and AIS-based hybrid intelligent algorithm for job-shop scheduling[J]. *IEEE Transactions on System, Man and Cybernetics*, 2008, 38(2): 358 – 368.
- [10] 刘益剑, 方彦军. 简化E.Coli觅食优化算法及其在非线形模型参数辨识中的应用[J]. 控制理论与应用, 2007, 24(6): 991 – 994.
(LIU Yijian, FANG Yanjun. Simplified E.Coli foraging optimization algorithm and its application to parameter identification of nonlinear system model[J]. *Control Theory & Applications*, 2007, 24(6): 991 – 994.)

作者简介:

刘益剑 (1978—), 男, 南京师范大学讲师, 武汉大学博士研究生, 目前研究方向为智能优化算法及嵌入式系统设计, E-mail: liuyijian_2002@163.com;

方彦军 (1957—), 男, 教授, 博士生导师, 主要研究方向为检测与控制方面的教学与研究, E-mail: yjfang@whu.edu.cn;

马宝萍 (1973—), 女, 博士, 讲师, 目前研究方向为广义系统、时滞系统和网络控制, E-mail: mabaoping@njnu.edu.cn.