

基于UKF和神经网络的一类非线性系统状态估计

刘 济, 高丽君

(华东理工大学 信息科学与工程学院, 上海 200237)

摘要: 在模型未知的情况下, 估计过程的重要变量尤为重要. 鉴于此, 采用不敏卡尔曼滤波(UKF)与神经网络相结合的方法, 解决一类未知模型非线性系统的状态估计问题. 采用动态神经网络对非线性系统进行建模, 利用UKF对状态和权值进行同时更新, 从而达到神经网络逼近真实模型, 估计值跟随真实值的目的. 通过两个仿真实例表明了所提出的方法具有良好的估计效果, 并且状态在输出中的比重越大, 其估计精度越高.

关键词: 模型未知; 神经网络; 不敏卡尔曼滤波

中图分类号: TP183

文献标志码: A

State estimation of a class of nonlinear system based on UKF and neural network

LIU Ji, GAO Li-jun

(School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China. Correspondent: Gao Li-jun, E-mail: junjun1028@yeah.net)

Abstract: It is significant to estimate the important process variables when process models are unknown. Therefore, the method of combining unscented Kalman filter(UKF) with neural network is used to solve the state estimation problems for a class of nonlinear systems whose models are unknown. The dynamic neural network is used to model for the nonlinear system, and the state and weights are updated at the same time by using UKF, which can achieve the purposes that the neural network approximate the real model, and the estimated values follow the real values. Two simulation examples are given to verify that the proposed approach gets good effects of estimation, and the greater the proportion of state in the output, the higher the estimation precision.

Key words: unknown model; neural network; unscented Kalman filter

0 引 言

对于复杂非线性系统, 其中间变量的控制和调整对实现整个非线性系统的监控十分重要. 然而, 大多数情况下这些中间变量难以直接测量, 只能通过输出的测量量对其进行估计, 因此, 为了能够以较高精度和较快的计算速度处理非线性状态估计问题, Julier等^[1]提出了基于UT变换的卡尔曼滤波(UKF). UKF适用于非线性高斯系统的状态估计问题, 尤其对于强非线性系统, 其精度和稳定性较EKF均有明显提高^[2-3]. UKF基于模型已知的情况对状态进行估计, 但对于一个复杂的非线性过程, 由于其多变量、强非线性、高不确定性, 导致建立精确的机理模型非常困难, 且一般情况下, 这些机理模型的建立是在一定的实验背景下实现的, 不能完全适应偏离实验背景的所

有环境. 因此, 对于未知模型的非线性过程, 仅靠UKF不能完成对中间变量的估计, 还需要对其未知模型进行精确的建模.

当对系统的机理不甚了解或不能用数学模型表示时, 人工神经网络往往是最有利的工具, 其方法实质上是一种非线性函数的逼近. BP神经网络结构简单、参数数量少, 是目前使用最为广泛的神经网络之一, 在很多非线性问题的解决中显示出了一定的优势. 神经网络根据实际数据集确定网络逼近的非线性函数, 但当实际过程的状态变量不完全可测, 或者可测的仅是状态的线性函数时, 单独采用神经网络将无法建立过程的模型. 为此, 本文基于UKF和BP神经网络各自的特点, 将UKF与BP神经网络相结合, 以解决一类模型未知且可测输出为状态线性组合的非线

收稿日期: 2013-07-15; 修回日期: 2013-10-11.

基金项目: 上海市自然科学基金项目(11ZR1409800).

作者简介: 刘济(1977-), 女, 副教授, 从事复杂非线性系统的建模与仿真等研究; 高丽君(1989-), 女, 硕士生, 从事复杂非线性系统的建模与仿真的研究.

性系统状态估计问题.

对于未知模型非线性系统的状态估计问题, 已有不少学者将 UKF 算法和神经网络结合起来解决实际问题. 文献 [5-7] 利用神经网络建立一个一维非线性时间序列模型, 模型的输入和输出分别是时间序列在当前时刻的值和下一时刻的值, 用 UKF 算法同时对网络权值和时间序列进行实时更新, 并将该方法与标准神经网络学习算法和单独的 UKF 估计算法进行比较, 以证明该方法的估计效果更好. 文献 [8-9] 利用神经网络建立一个多维的空间模型, 根据已知数据集用 UKF 先对神经网络进行训练, 得到满足精度要求的神经网络, 再用此神经网络和 UKF 算法对状态进行估计, 将此方法用于化工过程的状态估计, 达到了较好的估计效果. 此类文献研究中, 神经网络的输入和输出都可以直接测量.

如上所述, 很多实际系统的状态变量并不都可测, 即可以测量的仅是状态变量的某个函数关系 (输出方程). 这时, 单纯采用神经网络将无法获得此类系统的状态估计. 本文所要解决的是在模型未知且可测量输出是状态的线性函数关系情况下的状态估计问题. 根据神经元激励函数的非线性特点, 利用 UKF 实现神经网络权值系数的自适应调整, 从而自适应地模拟复杂的非线性系统, 同时对状态进行实时更新来达到估计的目的.

1 UKF 状态估计算法

UKF 结构与卡尔曼滤波相同, 也分为预测 (时间更新) 和更新 (测量更新) 两个步骤, 其预测过程是 UT 变换的过程, 更新过程则沿用卡尔曼滤波的测量过程. 考虑如下的非线性离散系统:

$$x(k+1) = f[x(k)] + w(k), \quad (1)$$

$$z(k) = h[x(k)] + v(k). \quad (2)$$

其中: $x(k)$ 为 n 维的状态向量; $f(\cdot)$ 为非线性系统状态转移函数; $w(k)$ 和 $v(k)$ 分别为系统的过程噪声和测量噪声, 且均为零均值的高斯白噪声. 对于非线性系统 (1) 和 (2) 的状态估计问题, 其关键在于如何由已知状态 $x(k)$ 的均值 $\bar{x}(k)$ 和方差阵 $P_x(k)$ 计算 $y(k) = f[x(k)]$ 的均值 $\bar{y}(k)$ 和方差阵 $P_{yy}(k)$. 这里采用 UT 变换技术来近似非线性函数的均值和方差.

一个 n 维随机变量 $x(k)$ 的均值和方差阵分别为 \bar{x} 和 P_x , 其一、二阶矩信息可根据以下 $2n+1$ 个 Sigma 点获得:

$$\chi_i = \begin{cases} \bar{x}, & i = 0; \\ \bar{x} + (\sqrt{(n+\lambda)P_x})_i, & i = 1, 2, \dots, n; \\ \bar{x} - (\sqrt{(n+\lambda)P_x})_i, & i = n+1, n+2, \dots, 2n. \end{cases} \quad (3)$$

其中: λ 为常数, $(\sqrt{(n+\lambda)P_x})_i$ 为矩阵 $(n+\lambda)P_x$ 方根矩阵的第 i 行或第 i 列. 对于 x 的非线性变换

$$y = f(x), \quad (4)$$

因变量 y 的一、二阶矩信息, 即均值 \bar{y} 和方差阵 P_{yy} 为

$$Y_i = f(\chi_i), \quad i = 0, 1, \dots, 2n; \quad (5)$$

$$\bar{y} = \sum_{i=0}^{2n} W_i^{(m)} Y_i; \quad (6)$$

$$P_{yy} = \sum_{i=0}^{2n} W_i^{(c)} (Y_i - \bar{y})(Y_i - \bar{y})^T. \quad (7)$$

其中: $W_i^{(m)}$ 和 $W_i^{(c)}$ 分别为计算均值和方差的第 i 个 Sigma 点的权值, 定义为

$$W_i^{(m)} = W_i^{(c)} = \begin{cases} \lambda/(n+\lambda), & i = 0; \\ 1/[2(n+\lambda)], & i = 1, 2, \dots, 2n. \end{cases} \quad (8)$$

UKF 状态估计方法只是在式 (1) 中状态转移函数已知的前提下才能达到估计的目的, 对于模型未知的情况, UKF 不能单独解决状态估计的问题.

2 基于 NN 的 UKF 的状态估计系统

为了能够在模型未知的情况下对中间变量进行估计, 利用神经网络建模, UKF 对状态和神经网络的权值同时估计, 其系统原理如图 1 所示.

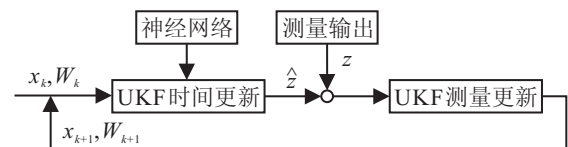


图 1 状态估计系统

图 1 中, 利用神经网络构建状态转移函数, 基于此模型, 以 k 时刻状态 x_k 和神经网络的权值 W_k 作为输入, 利用 UKF 对其进行时间更新, 得到先验估计和先验输出 \hat{z} . 通过与测量输出 z 的比较得到误差, 再利用 UKF 对先验估计进行测量更新, 得到后验状态 x_{k+1} 和后验权值 W_{k+1} , 返回到输入端作为下一次系统状态估计的输入.

神经网络模型利用一个输入输出节点数与状态个数相等、只含一个隐含层的 BP 神经网络建立, 以 k 时刻的状态 $x_{1,k}, x_{2,k}, \dots, x_{N,k}$ 作为输入, 以 $k+1$ 时刻的状态 $x_{1,k+1}, \dots, x_{N,k+1}$ 作为输出, 从输入层到隐含层和从隐含层到输出层的连接权值分别为 W_{nl}^1 和 W_{ln}^2 . 由于本文中 BP 神经网络的权值未知, 也将它们视为被估计的状态, 与目标状态进行联合估计, 形成增广的估计状态 $X^a = [X \ W]^T$. 对于神经元激励函数的选择, 应根据应用不同而异. 事实上, 研究表明, 隐含层采用 Sigmoid 激励函数、输出层采用线性激励函数的两层神经网络, 几乎可以以任意的精度逼近任

何函数^[10], 因此, 本文也采用此激励函数进行建模.

3 具体算法

如前面所述, 将所估计状态和神经网络权值联合起来, 形成增广的估计状态 $X^a = [X \ W]^T$, 因此非线性系统为

$$X_{k+1}^a = \begin{bmatrix} X_{k+1} \\ W_{k+1} \end{bmatrix} = \begin{bmatrix} f^1(X_k) \\ f^2(X_k) \\ \vdots \\ W_k \end{bmatrix} + n_k = \begin{bmatrix} f(X_k) \\ W_k \end{bmatrix} + n_k = f^a(X_k^a) + n_k, \quad (9)$$

$$z_k = h(X_k^a) + v(k). \quad (10)$$

其中

$$f^j(X_k) = \sum_{l=i}^L \left(W_{j,l}^2 \left(g \left(\sum_{i=1}^N (x_{i,k} W_{i,l}^1) \right) \right) \right), \quad j = 1, 2, \dots, N; \quad (11)$$

$g(x) = 1/(1 + e^{-x})$ 为 Sigmoid 函数; W 为神经网络的连接权值; $h(X_k^a)$ 为一个线性函数, 本文中为被估计状态的线性组合, 如 $h(X_k^a) = ax_{1,k} + bx_{2,k}$, a, b 为非零常数; n_k 为满足分布 $n_k \sim N(0, Q_k)$ 的系统过程噪声; v_k 为满足分布 $v_k \sim N(0, R_k)$ 的测量噪声.

假设已知状态在上一时刻的状态估计值和方差阵分别为 \hat{X}_{k-1}^a 和 $P_{X^a, k-1}$, 则对上述非线性系统采用 UKF 进行状态估计的具体步骤如下.

Step 1: 根据经验设定初值

$$\hat{X}_0^a = E(X_0^a), \quad P_0 = E[(X_0^a - \hat{X}_0^a)(X_0^a - \hat{X}_0^a)^T]. \quad (12)$$

Step 2: 时间更新. 当 $k > 1$ 时, 按照式 (3) 构造 $2n + 1$ 个 Sigma 点, 即

$$\chi_{k-1}^a = \begin{cases} \hat{X}_{k-1}^a, & i = 0; \\ \hat{X}_{k-1}^a + (\sqrt{(n + \lambda)P_{X^a, k-1}})_i, & i = 1, 2, \dots, n; \\ \hat{X}_{k-1}^a - (\sqrt{(n + \lambda)P_{X^a, k-1}})_i, & i = n + 1, n + 2, \dots, 2n. \end{cases} \quad (13)$$

计算预测 Sigma 点, 即

$$\chi_{i,k}^{a-} = f^a(\chi_{i,k-1}^a), \quad i = 0, 1, \dots, 2n. \quad (14)$$

计算预测 Sigma 点的均值和方差, 即

$$\hat{X}_k^{a-} = \sum_{i=0}^{2n} W_i^{(m)} \chi_{i,k}^{a-}, \quad (15)$$

$$P_{X^a, k}^- = \sum_{i=0}^{2n} W_i^{(c)} (\chi_{i,k}^{a-} - \hat{X}_k^{a-})(\chi_{i,k}^{a-} - \hat{X}_k^{a-})^T + Q_k. \quad (16)$$

Step 3: 测量更新. 当获得新的测量值 z_k 后, 可对

状态均值和方差进行更新. 具体公式如下:

$$\hat{X}_k^a = \hat{X}_k^{a-} + K[z_k - \hat{z}_k^-], \quad (17)$$

$$P_{X^a, k} = P_{X^a, k}^- - KP_{z, k}K^T, \quad (18)$$

$$K = P_{X^a, z, k}P_{z, k}^{-1}. \quad (19)$$

其中

$$\hat{z}_k^- = \sum_{i=0}^{2n} W_i^{(m)} h(\chi_{i,k}^{a-}), \quad (20)$$

$$P_{z, k} = \sum_{i=0}^{2n} W_i^{(c)} [h(\chi_{i,k}^{a-}) - \hat{z}_k^-][h(\chi_{i,k}^{a-}) - \hat{z}_k^-]^T + R_k, \quad (21)$$

$$P_{X^a, z, k} = \sum_{i=0}^{2n} W_i^{(c)} [\chi_{i,k}^{a-} - \hat{X}_k^{a-}][h(\chi_{i,k}^{a-}) - \hat{z}_k^-]^T. \quad (22)$$

4 仿真实验和结果分析

为了验证该估计方法的有效性, 通过两个实例的仿真进行证实.

4.1 仿真实例 1

考虑一个普遍使用的非线性系统离散模型^[11]

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} x_1(k) + \tau x_2(k) \\ -x_1(k) + (x_1^2(k) + x_2^2(k) - 1)x_2(k) \end{bmatrix} + n_k, \quad (23)$$

其中 $\tau = 0.001$ 为模型参数. 选取初始条件为 $x_{1,0} = 0.8, x_{2,0} = 0.2$, 过程噪声方差和测量噪声方差为 $Q_k = \text{diag}\{0.003^2, 0.005^2\}, R_k = 0.03^2$.

根据该非线性系统建立包含一个隐含层的神经网络模型 (2 个输入节点和输出节点, 3 个隐含节点), 权值个数为 12 个, 初始估计值为方差是 0.2 的随机噪声, 漂移方差阵为 $Q_w = 0.05^2 \times I_{12 \times 12}$, 状态的初始估计值为 $\hat{x}_{1,0} = 0.1, \hat{x}_{2,0} = 0.6$, 增广状态方差初始值为 $P_{x,0} = I_{14 \times 14}$; 调节参数 $\alpha = 0.1, \lambda = 1$. 取输出方程分别为

$$\begin{aligned} z(k) &= x_1(k) + x_2(k) + v(k), \\ z(k) &= x_1(k) + 3x_2(k) + v(k), \\ z(k) &= x_1(k) + v(k) \end{aligned}$$

3 种情况进行 20 次仿真, 得到两个状态 MSE (误差均方差) 值的平均值如表 1 所示, 其中一次的仿真结果如图 2 所示.

表 1 3 种情况 2 个状态的 MSE 值

MSE	状态	
	x_1	x_2
$z = x_1 + x_2$	0.047 252 9	0.047 195 8
$z = x_1 + 3x_2$	0.087 187 1	0.009 687 715
$z = x_1$	9.87e-07	0.056 760 15

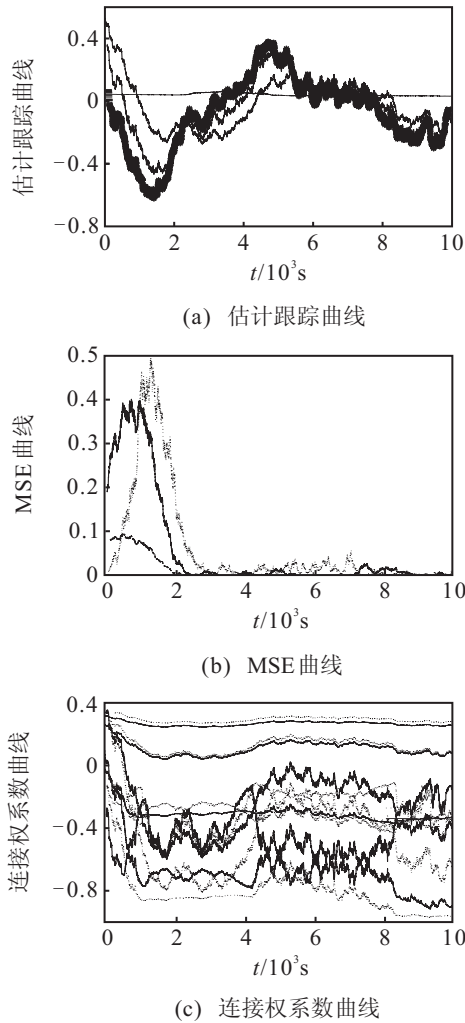


图 2 状态 x_2 在 3 种输出情况下的曲线

由表 1 可见: 当输出为两个状态的线性组合时, 两个状态的 MSE 值均较小, 达到了较高的精度要求; 当输出只为 1 个变量 x_1 时, 状态 x_1 的 MSE 较 x_2 小得多, 即状态 x_1 的估计精度比状态 x_2 的估计精度高得多。

由图 2(a) 和图 2(b) 可见, 当输出为两个状态的线性组合时, 状态 x_2 的估计曲线在约 3 000 步后能够较好地跟踪真实曲线, MSE 曲线也逐渐趋于 0, 且状态在输出方程中的系数权重越大, 其跟踪能力越高, 这与表 1 中的 MSE 值相对应, 即 MSE 值越小, 精度越高。由图 2(c) 可见, 神经网络的连接权值在整个学习过程中是一个自适应的调整过程。

4.2 仿真实例 2

考虑一个非等温化学搅拌塔反应器^[8,12], 其反应物浓度 C_A 和温度 T 的动态微分方程为^[8]

$$\begin{aligned} \dot{C}_A(t) &= -k_0 C_A e^{-\frac{E}{RT}} + \frac{q}{V} (C_{in} - C_A), \\ \dot{T}(t) &= -\frac{\delta H}{\rho C_p} k_0 C_A e^{-\frac{E}{RT}} + \frac{q}{V} (T_{in} - T) + \frac{h_A}{\rho C_p V} (T_w - T). \end{aligned} \quad (24)$$

对两个状态变量进行归一化处理

$$C'_A(k) = C_A(k) / \max(C_A), T'(k) = T(k) / \max(T).$$

采用与实例 1 中同样的网络结构和参数, 采样时间为 0.5 s. 分两种输出情况进行 20 次仿真, 得到两个物理量 MSE 值 (误差均方差) 的平均值如表 2 所示, 取其中一次的仿真结果如图 3 和图 4 所示。

表 2 化学反应器中反应物浓度与温度的 MSE 值

MSE	状态	
	C'_A	T'
$z = C'_A + T'$	5.36e-04	0.000 548 761
$z = T'$	0.972 841 5	6.66e-07

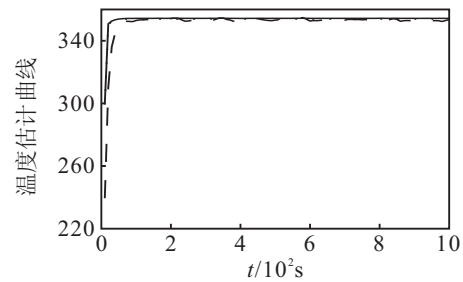


图 3 化学搅拌反应器温度估计曲线

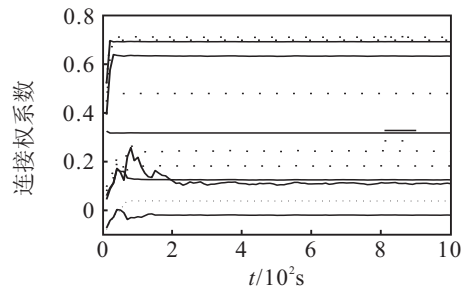


图 4 化学搅拌反应器神经网络连接权系数

由表 2 可见: 当输出为两个物理量的线性组合时, 其 MSE 值均较小且相当, 即两个物理量的估计精度都比较高; 当输出只为温度时, 温度的 MSE 值比浓度的 MSE 值小得多, 即温度的估计精度比浓度高得多。

由图 3 可见: 当输出为反应物浓度与温度的线性组合时, 温度的估计值与真实值除开始几步外几乎重合, 跟踪得较好; 当输出只为温度时, 反应物温度精确地跟踪了真实值, 这与表 2 中的 MSE 值相对应, 即 MSE 值越小, 跟踪越准确。由图 4 可见, 在整个过程中, 网络权值不断调整, 是一个自适应的动态学习过程。

4.3 结果分析

由两个实例的仿真结果可见, 当模型未知时, 利用神经网络建模, UKF 进行状态和神经网络权值的同时更新, 达到了神经网络连接权值自适应调整的目的, 且输出为不同的线性组合时, 每个状态的估计精度也有所不同。这是因为不同线性组合情况下, 各个状态在输出中的比重不同, 比重越大, 其实际测量的信息量越大, 精度更高一些。可见, 实例 2 的仿真效果较实

例 1 好,这是因为实例 1 的动态性和随机性较强,是一个强非线性的过程,实例 2 是一个较平稳的化工过程,非线性较弱.在相同的神经网络结构下,实例 2 的自适应过程较实例 1 快一些.对于实例 1 这种具有强非线性的过程,采用本文方法能够达到较好的估计效果,因此,对于大部分的模型未知且输出为状态线性组合的非线性状态估计问题,本文提出的估计方法是可以解决的.

5 结 论

本文利用动态神经网络,对一类模型未知且输出为状态的线性组合的非线性系统进行建模,利用 UKF 对状态和权值同时更新,达到神经网络逼近真实模型、估计值跟随真实值的目的.通过两个实例可以看出,在相同神经网络结构的情况下,系统的非线性强度较高时,估计精度相对较弱,但也能达到一定的精度要求.同时可以看出,当输出为不同的线性组合时,在输出中所占比重较大的状态其估计精度相对较高.

本文所解决的是模型未知且输出为状态的线性组合的情况下,非线性系统的状态估计问题.相对于估计状态为可测量对象的系统,更具有现实意义.但是对于输出为状态的非线性关系的状态估计问题,由于非线性函数中状态与状态的分离和状态的可视性难以确定,是一个更复杂的问题,还需进一步的研究.下一步的研究方向是考虑当状态维数较高时,如何选择合适的逼近网络建立状态转移方程,以降低估计问题的维数,获得较好的估计结果.

参考文献(References)

- [1] Julier S, Uhlmann J. Unscented filtering and nonlinear estimation[J]. Proc of the IEEE, 2004, 92(3): 401-422.
- [2] Wan E A, Merwe R. The unscented Kalman filter for nonlinear estimation[C]. Proc of Symposium 2000 on Adaptive Systems for Signal Processing, Communication and Control. Lake Louise: IEEE, 2000: 153-158.
- [3] Luo Zhen, Fang Hua-jing. Modified state prediction algorithm based on UKF[J]. J of Systems Engineering and Electronics, 2013, 24(1): 135-140.
- [4] 毕晓君. 信息智能处理技术[M]. 北京: 电子工业出版社, 2010: 117-148.
(Bi X J. Information intelligent processing technology[M]. Beijing: Electronic Industry Press, 2010: 117-148.)
- [5] Wu Xue-dong, Wang Yao-nan. Extended and unscented Kalman filtering based feedforward neural networks for time series prediction[J]. Applied Mathematical Modelling, 2012, 36(3): 1123-1131.
- [6] Li Hong-li, Wang Jiang, Chen Yan-qiu, et al. On neural network training algorithm based on the unscented Kalman filter[C]. The 29th Chinese Control Conf. Beijing, 2010, 29-31: 1447-1450.
- [7] Ronghui Z, Wan J. Neural network-aided adaptive unscented Kalman filter for nonlinear state estimation[J]. IEee Signal Processing Letters, 2006, 13(7): 445-448.
- [8] Karim Salahshoor, Amin Sabet Kamalabady. On-line multivariable identification by adaptive rbf neural networks based on ukf learning algorithm[C]. Proc of Chinese Control and Decision Conference. Yantai: IEEE, 2008: 4754-4759.
- [9] 李太福, 侯杰, 易军, 等. 基于 UKF 神经网络子空间逼近的非线性动态过程演化建模[J]. 应用基础与工程科学学报, 2013, 21(1): 186-194.
(Li T F, Hou J, Yi J, et al. Evolutional modeling for complex nonlinear dynamiprocess based on UKF & NN in subspace approximation[J]. J of Basic Science and Engineering, 2013, 21(1): 186-194.)
- [10] 魏海坤. 神经网络结构设计的理论与方法[M]. 北京: 国防工业出版社, 2005: 73-79.
(Wei H K. Theory and method of the neural network structure design[M]. Beijing: National Defence Industry Press, 2005: 73-79.)
- [11] Xiong K, Zhang H Y, Chan C W. Performance evaluation of UKF-based nonlinear filtering[J]. Automatica, 2006, 42(2): 261-270.
- [12] Denis Dochain. State and parameter estimation in chemical and biochemical processes: A tutorial[J]. J of Process Control, 2003, 13(8): 801-818.

(责任编辑: 郑晓蕾)