

平方根 UKF 神经网络及其在预测中的应用

黄冬民

(西北工业大学 自动化学院, 陕西 西安 710072)

摘要: 提出一种利用平方根无迹卡尔曼滤波 (SR-UKF) 进行神经网络权值估计的算法, 该算法可以克服 BP 算法存在的学习速率缓慢、计算量大、容易使学习陷入局部极小等缺点. 以 Mackey-Grass 混沌时间序列作为神经网络输入, 运用 SR-UKF 算法、UKF 算法、BP 算法仿真神经网络. 结果表明, SR-UKF 算法较之 BP 算法具有更快的训练速度和更高的预测精度, 且可以避免网络学习陷入局部极小, 而相对于 UKF 算法又具有计算量小且能保证状态方差半正定的特点.

关键词: 神经网络; 卡尔曼滤波

中图分类号: TP183 **文献标识码:** A **文章编号:** 1007-855X(2007)03-0060-04

Neural Network and Its Application in Prediction Based on Square Root Unscented Kalman Filter (SR-UKF)

HUANG Dong-min

(College of Automation, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract An algorithm based on the square-root unscented Kalman filter (SR-UKF) is proposed for weight estimation of the neural network, which can overcome the BP algorithm's drawbacks of slow learning speed, high computational complexity and the likeliness to make the study fall into the local minimum points. Taking the Mackey-Grass chaos time sequences as its input, the neural network is simulated with the SR-UKF, UKF and BP algorithm, with the result showing that the SR-UKF algorithm has the faster training speed and higher prediction precision than the BP algorithm, and it can avoid the network's convergence to the local minimum points. In addition, compared to the UKF algorithm, the SR-UKF algorithm is less complex in computation and it can guarantee positive semi-definite of the state covariance.

Key words neural network; Kalman filtering

0 引言

人工神经网络 (ANN) 因具有分布式存储与联想记忆功能, 以及较强的自适应性、自组织性和容错性与鲁棒性好等特点而被广泛应用. 其较强的学习能力和在一定条件下对任意非线性函数的精确近似, 使之成为强有力的非线性动态系统预测、建模与控制的工具. 随着神经网络的广泛应用, 要求有一种快速、高效的学习算法, 以满足神经网络运用的需要. 最普遍的神经网络学习方法是 BP 算法及其各种改进算法, 但是 BP 算法存在学习速率缓慢、计算量大、容易使网络学习进入局部极小点、出现收敛速度慢或不收敛等诸多缺陷, 极大地影响了 BP 算法的实际运用. 近年来, Kalman 滤波的广泛应用, 为神经网络的训练提供了理论支持. 其中运用扩展 Kalman 滤波 (EKF) 进行网络学习具有收敛速度快、学习精度高等特点^[1], 可以弥补 BP 算法的不足. 然而因为 EKF 是通过截取系统方程和观测方程泰勒展开式的一次项, 使非线性系统线性化, 所以该算法对系统状态量的估计仅可达到一次精度, 精度相对较低, 而且在计算步长不是充分小

收稿日期: 2006-10-20

作者简介: 黄冬民 (1977-), 男, 在读博士研究生. 主要研究方向: 复杂系统建模与仿真.

E-mail: googk0613@163.com

的情况下, 线性化会造成滤波性能不稳定; 另外, 对于复杂的非线性系统, 线性化过程中雅可比矩阵的推导比较困难. 对此, Julier 和 Uhlman 等人提出了无迹 Kalman 滤波 (Unscented Kalman Filter(UKF))^[5]. 由于采用了近似分布而非逼近非线性函数的方法, 因而不需要知道非线性函数细节, 可将非线性函数当作黑盒来处理. 相对于 EKF, 运用 UKF 进行神经网络训练的精度更高, 而计算复杂度相当, 且不要求导计算 Jacobian 矩阵^[2], 同时可以避免学习陷入局部极小^[3]. 但 UKF 算法在计算过程中状态方差的半正定性难以保证且计算量较大, 而平方根 UKF (SR-UKF) 算法却能很好地克服这一点^[8].

1 神经网络的状态空间表示

一个多层神经网络由一个输出层、一个输入层和若干隐含层组成. 如图 1 所示:

对网络进行训练即是对网络的连接权值进行最优估计, 寻求最佳的网络连接权值. 因此, 可将网络连接权值作为网络系统的状态变量, 而其输出作为网络的测量方程. 神经网络的状态空间模型可表述为^[4]:

$$w_{k+1} = w_k + \omega_k \tag{1}$$

$$y_k = f_k(w_k, u_k) + v_k \tag{2}$$

其中, w_k 为权值向量, u_k 为网络输入向量, y_k 为网络输出, f_k 为神经网络的激励函数. ω_k, v_k 分别为网络系统的过程和测量噪声, 服从均值为 0 方差分别为 R^w, R^v 的高斯分布.

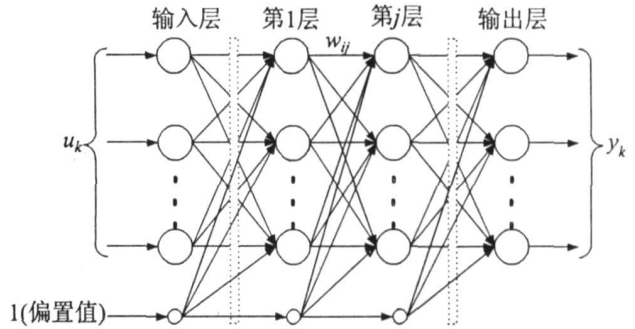


图 1 神经网络结构图
Fig.1 Structure of the neural network

2 UT 变换与 SR - UKF 原理

UT 变换是计算经非线性变换的随机变量之统计特性的一种方法^[5]. 设 n 维随机变量 x , 其均值和方差阵分别为 \bar{x}, P_x , 且满足非线性方程 $y = f[x]$. 为了确定 y 的均值 \bar{y} 和方差 P_y , 做 UT 变换:

$$Y_i = f[X_i], i = 0, 1, \dots, 2n \tag{3}$$

$$\begin{cases} X_0 = \bar{x} \\ X_i = \bar{x} + (\sqrt{(n+\lambda)P_x})_i & i = 1, 2, \dots, n \\ X_i = \bar{x} - (\sqrt{(n+\lambda)P_x})_i & i = n+1, \dots, 2n \\ W_0^{(m)} = \lambda/(n+\lambda) \\ W_0^{(c)} = \lambda/(n+\lambda) + (1-\alpha^2 + \beta) \\ W_i^{(m)} = W_i^{(c)} = 1/\{2(n+\lambda)\} & i = 1, 2, \dots, 2n \end{cases} \tag{4}$$

式中, $\lambda = \alpha^2(n+k) - L$ 为一比例参数, α 为一常数 (如 $10^{-4} \leq \alpha \leq 1$), 决定 Sigma 点围绕均值的延伸程度, k 为冗余量, 通常设为 $3-L$; β 为用于引入先验状态概率分布信息的参数 (高斯分布时, $\beta = 2$ 最优)^[6]. 由 (3), (4) 式可得的均值 \bar{y} 和方差 P_y 的近似为:

$$\bar{y} \approx \sum_{i=0}^{2n} W_i^{(m)} Y_i \tag{5}$$

$$P_y \approx \sum_{i=0}^{2n} W_i^{(c)} (Y_i - \bar{y})(Y_i - \bar{y})^T$$

根据 UT 变换和 Kalman 滤波原理, 在式 (1), (2) 定义的神经网络系统模型的基础上, 可得出基于 SR - UKF 的神经网络学习算法步骤如下:

1) 初始化

w-hat_0 = E[w]; S_0 = chol{[(w - w-hat_0)(w - w-hat_0)^T]}

2) Sigma点计算与时间更新

w-hat_k = w-hat_{k-1}; S-w_k = lambda_RL S_{k-1}

w_{k|k-1} = [w-hat_k w-hat_k + sqrt(S-w_k) w-hat_k - sqrt(S-w_k)]

Y_{k|k-1} = f(u_{k-1}, w_{k|k-1}); y-hat_k = sum_{i=0}^{2n} W_i^{(m)} Y_{i|k-1}

3) 测量更新

S_{y_k} = qr{[sqrt(W_1^c)[D_{1,2n,k} - y-hat_k sqrt(R^v)]]; P_{w_k y_k} = sum_{i=0}^{2n} W_i^{(c)} (w_{i|k-1} - y-hat_k)(Y_{i|k-1} - y-hat_k)^T

K_k = (P_{w_k y_k} S_{y_k}^{-1}) / S_{y_k}; w-hat_k = w-hat_k + K_k (y_k - y-hat_k); U = K_k S_{y_k}; S-w_k = cholupdate{S-w_k, U, -1}

其中, k in {1, ..., infinity}, sqrt(L + lambda) R^omega, R^v 分别为过程噪声和测量噪声. 运用 SR - UKF 进行神经网络权值估计, 当权值方差 P_w_k 趋于 0 时 (此时 Kalman 增益 K_k 亦趋于 0) 绝对收敛.

3 学习算法原理

运用 SR - UKF 进行神经网络学习的步骤如图 2 所示.

初始化权值及其方差, 进行时间更新, 在 UT 变换产生 Sigma 点后进入神经网络并向前传播, 计算方差、协方差和 KALMAN 增益矩阵, 最后进行权值和误差方差的测量更新, 循环往复直至达到滤波收敛条件为止.

4 仿真实验及结论

以 Mackey - Glass 混沌时间序列预测作为仿真实例, 以验证神经网络的性能. Mackey - Glass 混沌时间序列可由时滞微分方程 [7]:

x-dot(t) = (0.2x(t-t) / (1 + x^10(t-t))) - 0.1x(t) (6)

得到, t 为唯一可调参数. 当 t > 17 时, 方程 (6) 产生混沌现象.

对 Mackey - Glass 序列进行预测主要是根据 t 时刻以前的一组数据 x(•) 去预测 x(t + Delta t), 其中 Delta t 为预测时间步幅. 预测方法是令 Delta t 为时滞参数, 取 D 个点, 即 [x(t - (d - 1) Delta t) ... x(t - Delta t) x(t)] 去预测将来时刻 x(t + Delta t).

取 t = 20, Delta t = 6 选取 1000 组样本数据, 即 [x(t - 18) x(t - 12) x(t - 6) x(t); x(t + 6)]; t = 21, 22, ..., 1020 其中前 4 个变量数据作为网络的输入, 最后一个变量数据作为网络的输出. 前 800 组 (t = 21, 22, ..., 820) 当做神经网络训练数据, 其余 200 组 (t = 821, 822, ..., 1020) 作为测试数据以验证算法的有效性. 选取单隐层神经网络且隐含层神经元数为 5 从而构造一个 4 x 5 x 1 的神经网络. 分别运用 SR - UKF 算法、UKF 算法和 BP 算法对网络进行仿真, 其仿真预测结果如图 3~ 8 所示.

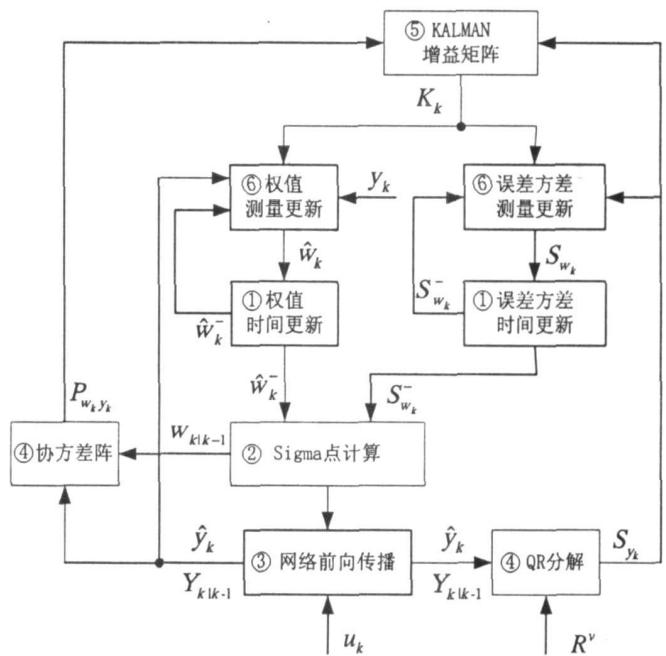


图 2 运用 SR-UKF 进行神经网络训练流程图 Fig.2 Flow chart of neural network training based on SR-UKF

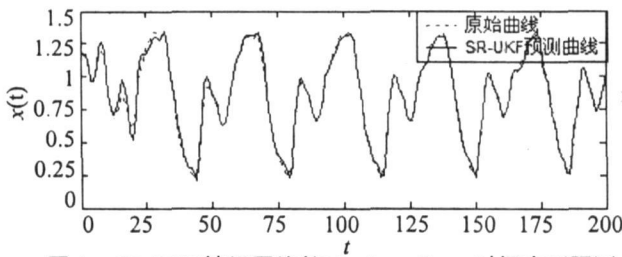


图 3 SR-UKF神经网络的Machey-Grass时间序列预测
Fig.3 Machey-Grass time series Prediction of SR-UKF neural network

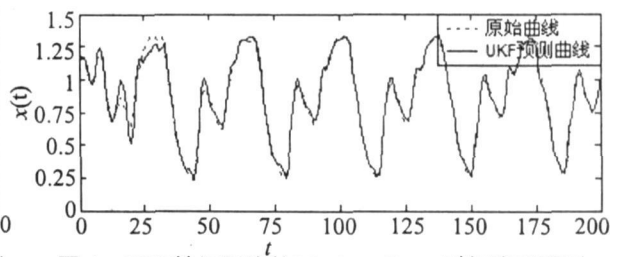


图 4 UKF神经网络的Machey-Grass时间序列预测
Fig.4 Machey-Grass time series Prediction of UKF neural network

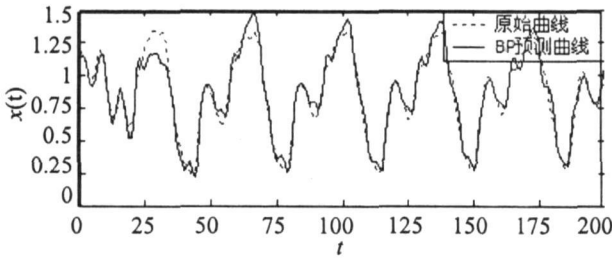


图 5 BP神经网络的Machey-Grass时间序列预测
Fig.6 Machey-Grass time series Prediction of BP neural network

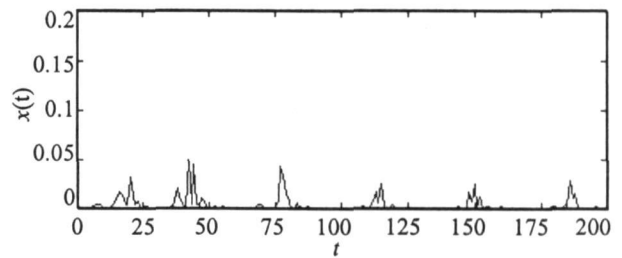


图 6 SR-UKF神经网络的Machey-Grass时间序列预测误差
Fig.6 Machey-Grass time series Prediction errors of SR-UKF neural network

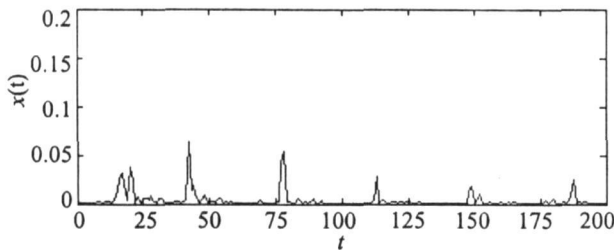


图 7 UKF神经网络的Machey-Grass时间序列预测误差
Fig.7 Machey-Grass time series Prediction errors of UKF neural network

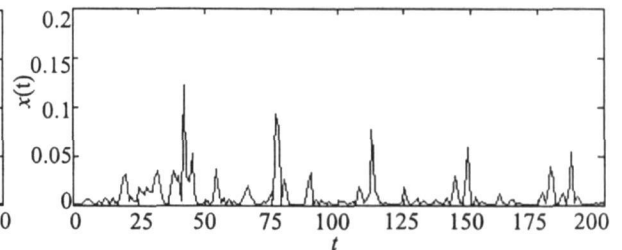


图 8 BP神经网络的Machey-Grass时间序列预测误差
Fig.8 Machey-Grass time series Prediction errors of BP neural network

仿真过程中, SR - UKF 算法可以保证状态方差的半正定性且计算量小, 结果显示利用其训练的神经网络对 Machey - Grass 时间序列的预测曲线与真实曲线吻合程度较好, 且与 UKF 算法近似, 而 BP 神经网络的预测曲线与真实曲线的则相去较远. 图 6 ~ 8 的预测误差曲线表明 SR - UKF 神经网络与 UKF 神经网络预测误差相近, 较之 BP 神经网络的预测精度高, 从而说明 SR - UKF 算法运用于神经网络训练的有效性.

(下转第 67 页)

49.4%, 可见地震峰值最大加速度对桥梁地震响应影响非常大. 文献 [5-6] 指出, 对于高墩连续刚构, 其稳定系数均非常高. 因此在地震作用下, 桥墩不会出现失稳.

2) 当地震峰值最大加速度为 $0.3g$ 时, 墩身主筋采用 HRB335 级 $\phi 32$ 钢筋, 横、纵桥向均布置两排, 主筋间距为 10 cm . 此时钢筋已非常密集, 如地震峰值最大加速度再提高, 则需要增大桥墩的截面才能配筋. 而桥墩截面增大后, 由于刚度也随之增大, 地震响应又将加大. 这两者成为矛盾. 故如再增大地震峰值最大加速度, 桥墩的配筋能否通过抗震尚需要进一步研究.

3) 对于高墩连续刚构, 通常其横向振型贡献较大, 故一般以墩底的横向弯矩作为控制桥墩的设计.

4) 查国家地震动峰值加速度图, 大于 $0.3g$ 的地区非常少. 故此, 可以认为连续刚构桥梁在高烈度地区具有较好的适应性.

参考文献:

- [1] 张利. 预应力混凝土连续刚构桥梁箱梁顶板纵向裂缝分析 [J]. 公路, 2006 (8): 262-264
- [2] 王克海, 李茜. 高墩桥梁地震响应分析 [J]. 世界桥梁, 2006 (1): 41-43
- [3] 李权编. ANSYS 在土木工程中的应用 [M]. 北京: 人民邮电出版社, 2005
- [4] 黄侨, 王永平. 桥梁混凝土结构设计原理计算示例 [M]. 北京: 人民交通出版社, 2006
- [5] 郭梅. 高墩大跨连续刚构稳定性分析 [J]. 西安公路交通大学学报, 1999 (3): 31-35
- [6] 彭元诚, 方秦汉, 李黎. 超高墩连续刚构设计中的关键技术 [J]. 桥梁建设, 2006 (4): 30-33

(上接第 63 页)

参考文献:

- [1] Iguni Y, Sakai H, Tokumaru H. A Real-Time Learning Algorithm for a Multilayered Neural Network Based on the Extended Kalman filter [J]. IEEE Trans Signal Processing, 1992, 40(4): 959-966
- [2] J J LaViola Jr. A Comparison of Unscented and Extended Kalman Filtering for Estimating Quaternion Motion [R]. Proc. 2003 Am. Control Conf. 2003: 2435-2440
- [3] Simon Haykin. Kalman Filtering and Neural Networks [M]. John Wiley & Sons, 2001.
- [4] R van derMerwe and E. Wan. Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models [C]. in Proc. Workshop Adv. in Machine Learning, Montreal, QC, Canada, 2003.
- [5] Julier S J, Uhlmann J K. A New Extension of the Kalman Filter to Nonlinear Systems [A]. In Proc. of AeroSense: The 11th Int. Symp. On Aerospace/Defence Sensing, Simulation and Controls, 1997.
- [6] Julier S J, Uhlmann J K. The scaled Unscented Transformation [A]. In IEEE American Control Conference, Anchorage, AK, USA, IEEE (2002) 4555-4559.
- [7] Farmer J D. Chaotic Attractors of an Infinite-Dimensional Dynamical System [J]. Physica D, 1982, 4: 366-393.
- [8] R Van derMerwe, Wan E A. The Square-root Unscented Kalman Filter for state and Parameter-estimation [A]. in International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, Utah, 2001.