Neural Network Forecast Algorithm Based on Iterated Unscented Kalman Filter

Liu Xiangheng
Department of navigation engineering
Naval University of Engineering
Wuhan, China, 430033
liuxiangheng240@163.com

Hu Baiqing
Department of navigation engineering
Naval University of Engineering
Wuhan, China, 430033
hubq@126.com

Chang Guobin
Department of navigation engineering
Naval University of Engineering
Wuhan, China, 430033
cgbin110@163.com

Abstract—**A novel algorithm based on the iterated unscented Kalman filter (IUKF) is proposed in this paper to train the weights and bias of the neural network. In the proposed algorithm, the weights and bias are considered as the states, and the outputs of the network are used as the measurements for the IUKF. In IUKF, the iteration concept is introduced into the unscented Kalman filter (UKF). By substituting the updated mean and covariance into the unscented transformation (UT), the total forecast precision is improved. Taking the Mackey-Glass chaos time sequences as an input of the net, the neural network is simulated with the IUKF, UKF and back propagation (BP) algorithms. The simulation results indicate that the IUKF algorithm has a faster training speed and higher forecast precision than the BP algorithm. Moreover, the IUKF algorithm avoids the network's convergence getting into the local minimum points. Compared with UKF algorithm, the proposed algorithm has a higher forecast precision.**

Keywords—Forecast, Iterated unscented Kalman filter, Neural network

I. INTRODUCTION

Artificial neural network is widely used due to its parallel computing, distributing storage, high fault-tolerant and nonlinear adaptive processing power for many obvious advantages. With the development and application of the neural network, rapid efficient learning algorithms have become the key in the research of neural network. The most common learning algorithms of neural network are back propagation (BP) algorithm and its various improved algorithms, but BP algorithm itself exists defects such as slow learning rate、easying to make the network into the local minimum points、slow convergence speed、being influenced by initialized weights and so on, and these seriously influences the BP algorithm in actual application. With the development of Kalman filter, a lot of theoretical algorithms have been developed in the neural network training [1]. The extended Kalman filter (EKF) algorithm for neural network learning has a lot of characteristics, such as fast convergence, high learning precision, etc, and it can solve the deficiency of the BP algorithm, but it has some obvious flaws. One is that the linearization may lead to unstability of filtering, the other is that the calculation of the jacobian matrix is a must in EKF, which results in a lot of restrictions in the network training. To solve these problems, Julier proposed a forward unscented Kalman filter (UKF) [2]. In the UKF, some sigma points are derived by sampling and propagated through the unscented transformation (UT) in order to approximate the probability density of nonlinear function. Compared with EKF algorithm, UKF algorithm could achieve higher precision in processing nonlinear problems, and avoids the calculation of the jacobian matrix. The UKF can also deal with the nonlinear function, which is discontinous or not exist display expression. Thus, it expands the scope of application. However, the UKF algorithm is a kind of approximated estimation and has some devitation unavoidabilities. Therefore, it is necessary to introduce the iteration theory into UKF algorithm to solve such a problem [3]. The proposed novel algorithm is iterated unscented Kalman filte (IUKF), which is used to train the neural network weights and bias with improved forecast precision of neural network.

II. NETWORK ARCHITECTURE、PROCESS EQUATION AND MEASUREMENT EQUATION

Neural network is composed by an input layer、an output layer and several hidden layers. A feedforward layered neural network, as shown in Fig.1, contains two hidden layers.

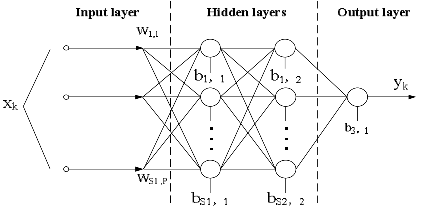

Figure 1. Feedforword layered neural network architecture

The IUKF algorithm used in trainning the neural network provides an optimized estimation of the weights and bias. The process equation and measurement equation of the the neural network can be described by the following nonlinear discrete-time system:

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

$$y_k = h(w_k, x_k) + r_k \tag{2}$$

The equation (1) is known as the process equation, and the equation (2) is the measurement equation. Where w_k is the state combined by weights and bias which is given by the neural network, x_k is the input of the network, y_k is the output, and q_k is the process noise which is characterized as zero-mean, Gaussian noise with covariance Q. Similarly, r_k is the measurement noise which is also characterized as zero-mean, Gaussian noise with covariance R. $h(\bullet)$ is the nonlinear excitation function of the neural network.

III. UNSCENTED TRANSFORMATION AND IUKF ALGORITHM

The unscented transformation (short of "UT") is a method for calculating the statistics of a random variable which undergoes a nonlinear transformation. Assuming that a state w (dimension L) is propagated through a nonlinear function, $y = h[w]$, and w has a mean $\hat{w}$ and covariance P_w, an UT is carried out in the follow expressions to calculate the mean $\hat{y}$ and covariance P_y:

$$\begin{cases} W_0 = \hat{w} \\ W_i = \hat{w} + (\sqrt{(L+\lambda)P_w})_i \quad i = 1, 2, \cdots, L \\ W_i = \hat{w} - (\sqrt{(L+\lambda)P_w})_i \quad i = L+1, \cdots, 2L \\ C_0^{(m)} = \lambda / (\lambda + L) \\ C_0^{(c)} = \lambda / (\lambda + L) + (1 - \alpha^2 + \beta) \\ C_i^{(m)} = C_i^{(c)} = 1 / \{2(L+\lambda)\} \quad i = 1, 2, \cdots 2L \end{cases} \tag{3}$$

where $\lambda = \alpha^2(L+k) - L$ is a scaling parameter. The constant α determins the spread of the sigma points around $\hat{w}$ and is usually set to a small positive value (e.g., 1e-4⩽ α ⩽1). The constant k is a secondary scaling parameter which is usually set to 0 or 3－L, and β is used to incorporate prior knowledge of the distribution of w (for Gaussian distributions, $\beta = 2$ is optimal). W_i is the sigma points and they are propagated through the nonlinear function,

$$Y_i = h[W_i] \quad i = 0, 1, \cdots, 2L \tag{4}$$

According to the equation (3) and (4), the mean $\hat{y}$ and covariance P_y for y can be approximated using a weighted sample mean and covariance of the posterior sigma points,

$$\hat{y} = \sum_{i=0}^{2L} C_i^{(m)} Y_i \tag{5}$$

$$P_y = \sum_{i=0}^{2L} C_i^{(c)} (Y_i - \hat{y})(Y_i - \hat{y})^T \tag{6}$$

The UKF algorithm is derived by introducing the UT into Kalman filter, and then the IUKF is derived by introducing the iteration concept into the UKF algorithm. Sine the process equation is linear, so the time update equations are the same as the strandord Kalman filter. Based on the UT and the process equation (equation (1)) and measurement equation (equation (2)), the step of the IUKF algorithm can be expressed as follows:

① Initialization

$$\hat{w}_0 = E[w]$$

$$P_{w_0} = E[(w - \hat{w}_0)(w - \hat{w}_0)^T]$$

② Time update

$$\hat{w}_k^- = \hat{w}_{k-1}$$

$$P_{w_k}^- = P_{w_{k-1}} + Q_{k-1}$$

③ State extend

In order to meet the condition of statistic orthogonality for estimate and measurement noise befor the iteration, it is necessary to extend the measurement noise R into the state. The extended variables are marked with a superscript of "α", and the extended mean and covariance of the state are expressed as follows:

$$\hat{w}_k^{a-} = \begin{bmatrix} \hat{w}_k^- \\ 0 \end{bmatrix} \tag{7}$$

$$P_{w_k}^{a-} = \begin{bmatrix} P_{w_k}^- & 0 \\ 0 & R \end{bmatrix} \tag{8}$$

④ Measurement update

$$W_{k|k-1}^a = [\hat{w}_k^{a-} \quad \hat{w}_k^{a-} + \gamma\sqrt{P_{w_k}^{a-}} \quad \hat{w}_k^{a-} - \gamma\sqrt{P_{w_k}^{a-}}] \tag{9}$$

$$Y_{k|k-1}^a = h(W_{k|k-1}^a, x_k) \tag{10}$$

$$\hat{y}_k^a = \sum_{i=0}^{2L} C_i^{(m)} Y_{k|k-1,i}^a \tag{11}$$

$$P_{\tilde{y}_k^a \tilde{y}_k^a}^a = \sum_{i=0}^{2L} C_i^{(c)} (Y_{k|k-1,i}^a - \hat{y}_k^a)(Y_{k|k-1,i}^a - \hat{y}_k^a)^T \tag{12}$$

$$P_{w_k^a y_k^a}^a = \sum_{i=0}^{2L} C_i^{(c)} (W_{k|k-1,i}^a - \hat{w}_k^{a-})(Y_{k|k-1,i}^a - \hat{y}_k^a)^T \tag{13}$$

$$K_k^a = P_{w_k^a y_k^a}^a (P_{\tilde{y}_k^a \tilde{y}_k^a}^a)^{-1} \tag{14}$$

$$\hat{w}_k^a = \hat{w}_k^{a-} + K_k^a (y_k - \hat{y}_k^a) \tag{15}$$

$$P_{w_k}^a = P_{w_k}^{a-} - K_k^a P_{\tilde{y}_k^a \tilde{y}_k^a}^a (K_k^a)^T \tag{16}$$

where $\gamma = \sqrt{L+\lambda}$

⑤ Iteration update

The process of the UKF algorithm is finished by step ① to ④. Compared with the standard UKF algorithm, step ③ is added in order to meet the condition of statistic orthogonality of the estimation and measurement noise befor the iteration. However, the UKF algorithm is a kind of approximated estimation and has some devitation unavoidably. Therefore, it is necessary to introduce a iterative computation for mean and covariance of the state after measurement update. Iteration update is to take $\hat{w}_k^a$ in equation (15) and $P_{w_k}^a$ in equation (16) as initial values, and then return to step ④. The mean and covariance of the state after iteration update can be expressed as follows:

$$\hat{w}_k^a(j+1) = \hat{w}_k^{a-}(j) + K_k^a(j)(y_k - \hat{y}_k^a(j)) \tag{17}$$

$$P_{w_k}^a(j+1) = P_{w_k}^{a-}(j) - K_k^a(j) P_{\tilde{y}_k^a \tilde{y}_k^a}^a(j)(K_k^a(j))^T \tag{18}$$

where j is the number of the iteration. When iteration endes [3], take the corresponding parts of the extended state estimation and covariance and assign them to the original state estimation and covariance for the next filter.

IV. PROCESS OF TRAINING NEURAL NETWORK BASED ON IUKF

Based on IUKF algorithm, the process of training neural network is show in Fig.2. Using the weights and bias trained by IUKF algorithm, the input training sample x_k is propagated through the neural network to produce a forecast output $\hat{y}_k^a$, Meanwhile, the forecast output $\hat{y}_k^a$ returns to the IUKF and participates in accomplishing measurement update of IUKF with the input training sample x_k, output training sample y_k, process noise Q, and measurement noise R, then a new weights and bias matrix $\hat{w}_k^a(j+1)$ is produced and transmitted to the neural network. The computation of the Sigma points, mean, covariance and the Kalman gain matrix are accompished in the frame of IUKF. The final weights and bias matrix $\hat{w}_k^a(j+1)$ for neural network forecast is computed when the iteration endes.

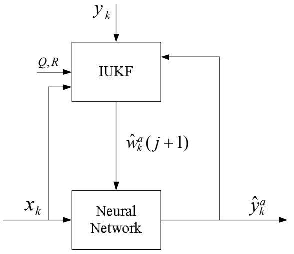

Figure 2. Process of training neural network based on IUKF

V. SIMULATION OF NEURAL NETWORK FORECAST

In this section, the forecast precision of the neural network based on IUKF algorithm is validated through being applicated in Mackey-Glass chaos time sequences, which is often used as a model of nonlinear neural network forecast for its high nonlinear. Mackey-Glass chaos time sequences can be produced by the equation as follows [4]:

$$\frac{dx}{dt} = \frac{0.2x(t-\tau)}{[1+x(t-\tau)^{10}]} - 0.1x(t) \tag{19}$$

where the equation (19) produces chaos phenomena when $\tau = 17$. In this similation, Mackey-Glass chaos time sequences is produced by fourth-order Runge-Kutta algorithm in Matlab. One thousand points of the sequences are taken as the sample and devided into 976 groups by the method as follows:

$$\{x(t), x(t+6), x(t+12), x(t+18), x(t+24); t = 1,2,\cdots 976\} \tag{20}$$

Where the first four sample points are used as the input sample and the last sample point is used as the output sample. The first 776 groups of the sample points are used as the training sample of neural network and the last 200 groups of the sample points are used as the forecast sample. According to the number of the input and output, A neural network of 4×5×1 is established (input number is 4, number of the single hidden layer neuron is 5, output number is 1). The neural network is trained by the IUKF, UKF and BP algorithms, and the results of the forecast with different algorithms are show in Fig.3~5 and Tab.1.

Where MSE is the mean-square error of each sample point, S_MSE is the sum of MSE, the number of the iteration is 3 in Fig.3. The results show that the IUKF algorithm has an obviously better forecast precision than UKF algorithm and BP algorithm. According to the Fig.5, S_MSE has reached steady when the number of the iteration more than 3. Generally speaking, 3 or 4 is enough in iteration.

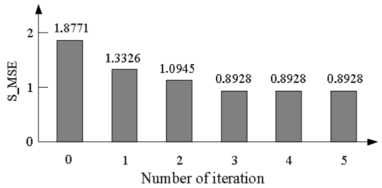

Figure 5. Iterative number of IUKF and S_MSE

TABLE I. S_MSE OF THREE ALGORITHMS

Algorithm	BP	UKF	IUKF
S_MSE	3.7619	1.8771	0.8928

VI. SUMMARIES

A novel algorithm based on the IUKF is proposed to train the weights and bias of the neural network in this paper, which can make full use of the valuable information from measurements to smooth state estimation and then a better weights and bias is derived to get a better forecast precision of the neural network. The results of the simulation indicate that the forecast precision of nerual network based on IUKF algorithm is obviously better than UKF algorithm and BP algorithm. The neural network based on IUKF algorithm has wide applications on modeling of nonlinear problems which only the input and output data are known.

Published by Atlantis Press, Paris, France.

REFERENCES

[1] BI Yan-li, SHANG Rui, LIU Yi-mei. A BP net analogue system based on the Kalman filtering [J]. Journal of Dalian Maritime University, 2006, 26(2): 86-88.

[2] Julier S J, Uhlmann J K. Unscented filtering and nonlinear estimation [J]. Proc IEEE, 2004, 92(3):401-422.

[3] Sibley G, Sukhatme G, and Matthies L .The iterated sigma point Kalman filter with applications to long range stereo [J]. Proceeding of Robotics: Science and Systems, Philadelphia, USA, August 2006.

[4] Farmer J D. Chaotic attractors of an infinite-dimensional dynamical system [J]. Physica, 1982, D4(3): 366-393.

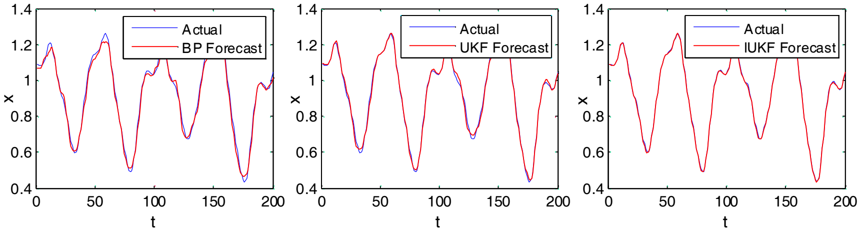

Figure 3. Neural network forecast of three algorithms

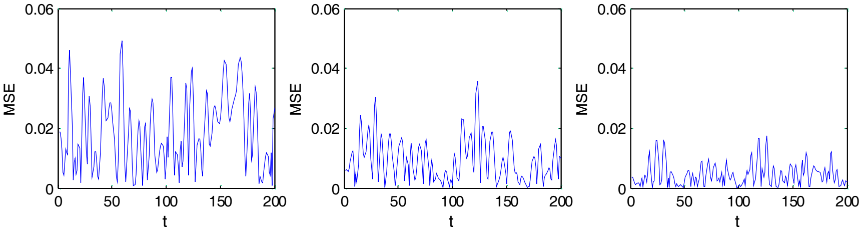

Figure 4. MSE of three algorithms

Published by Atlantis Press, Paris, France.