

基于 Kalman 滤波的储备池多元 时间序列在线预报器

韩敏¹ 王亚楠¹

摘要 针对多元非线性时间序列, 结合回声状态网络和 Kalman 滤波提出一种新的在线自适应预报方法. 该方法将 Kalman 滤波应用于回声状态网络储备池高维状态空间中, 直接对网络的输出权值进行在线更新, 省去了传统递归网络扩展 Kalman 滤波中 Jacobian 矩阵的计算, 在提高预测精度的同时令算法的适用范围得到扩展. 在回声状态网络稳定时给出所提算法的收敛性证明. 仿真实例验证了所提方法的有效性.

关键词 卡尔曼滤波, 递归网络, 回声状态网络, 多变量序列, 预测

DOI 10.3724/SP.J.1004.2010.00169

Multivariate Time Series Online Predictor with Kalman Filter Trained Reservoir

HAN Min¹ WANG Ya-Nan¹

Abstract A novel online adaptive prediction method is proposed for multivariable nonlinear time series, which is based on echo state network (ESN) and Kalman filtering (KF) algorithm. The KF is adopted in the high-dimension “reservoir” state space to directly update the output weights of the ESN online. It is implemented without the computation of Jacobian matrices which is in the expanded KF (EKF) algorithm of traditional recurrent neural network (RNN), so as to improve the prediction accuracy and extend the applications. The convergence of the proposed method is proved when the ESN is steady. Simulation examples demonstrate the validity of the proposed method.

Key words Kalman filter (KF), recurrent neural network (RNN), echo state network (ESN), multivariate series, prediction

时间序列预测技术在复杂系统建模、数据流技术和故障诊断等领域都具有较为广泛的应用^[1], 是当前控制理论研究的热点之一. 而神经网络作为一种有效的非线性黑箱建模方法, 在当今时间序列研究中发挥着较大的作用. 特别是递归网络 (Recurrent neural network, RNN) 内部包含反馈回路, 可适应动态系统的时变特性, 比仅能反映系统输入-输出间静态映射关系的前馈网络更适合动态系统建模^[2].

尽管 RNN 已逐渐发展成为一种时间序列处理的基本方法, 但它始终存在训练复杂和稳定性分析困难的不足. 同时, 大多数预测方法都采用离线批处理算法. 对于时间序列数据来说, 其规模往往很大, 离线方法对存储资源和运算资源的要求高、成本大. 目前应用较广的在线 BP (Back propagation) 算法是一种梯度下降算法, 存在跟踪速度慢, 易陷入局部极小的缺陷. 扩展卡尔曼滤波 (Expanded Kalman filter, EKF)

作为卡尔曼 (Kalman filter, KF) 算法^[3] 的一种非线性推广形式, 近年来被用于网络在线学习中. EKF 属于最优状态估计算法, 用它训练神经网络时, 网络的权值被当作滤波器的待估状态^[4], 具有接近二次的收敛速度, 训练速率明显优于 BP. 但是, EKF 算法在计算中用到了误差协方差矩阵的近似值, 最终只得到最优解的一阶近似形式, 算法的精度和收敛性都无法保证, 仅在初始误差和噪声不大时有效. 且运算过程中含有 Jacobian 矩阵的求取, 需要通过 RTRL 或 BPTT 这样的梯度下降算法实现^[5]. 后来提出的 UKF 方法虽然能够提高预测的精度, 但它以牺牲运算量为代价, 计算复杂度远高于 EKF, 同时其 Unscented 变换参数的选取具有较大随机性^[6].

回声状态网络 (Echo state network, ESN) 是一种性能良好的新型递归网络, 具有较强的生物背景和非线性函数逼近能力^[7]. ESN 内部包含一个递归连接大且稀疏的储备池, 起到存储历史信息的“回声”作用, 储备池到输出之间呈线性映射. 与常规 RNN 网络不同, ESN 的输入权值和储备池内部连接权值在初始时刻即给定不变, 训练的目的在于得到储备池到输出间的最优输出权值, 可应用线性算法求取. 本文在 ESN 网络储备池学习机制基础之上, 将原始空间的非线性问题转化为储备池高维状态空间的线性问题后, 利用 KF 算法实时更新网络的输出权值. 由于储备池状态至输出值间为线性映射, KF 算法不需经任何扩展即可直接用于输出权值的寻优. 相比于传统 RNN 的 EKF 算法, 既消除了一阶线性近似产生的运算偏差, 解除了网络方程必为一阶可导且导数可求的限制, 又避免了 Jacobian 矩阵的求取运算. 在 ESN 网络稳定的前提下, 给出本文算法的收敛性证明, 并通过 Lorenz 多元时间序列预测仿真实验将本文方法与传统 EKF 方法进行比较.

1 多变量时间序列状态空间重构

对时间序列进行分析的第一步是相空间重构, 延迟坐标状态空间重构是最常用的方法. 假定多元离散时间序列为 $x_1(t), x_2(t), \dots, x_n(t)$, 借鉴单变量序列重构相空间的方法, 分别选取 m_i 和 τ_i ($i = 1, 2, \dots, n$) 为第 i 元混沌序列的嵌入维数和延迟, 得到多元变量嵌入延迟向量如式 (1) 所示.

$$\mathbf{x}(k) = [\mathbf{x}_1(k)^T, \mathbf{x}_2(k)^T, \dots, \mathbf{x}_n(k)^T]^T \quad (1)$$

其中

$$\mathbf{x}_i(k) = [x_i(k), x_i(k - \tau_i), \dots, x_i(k - (m_i - 1)\tau_i)]^T \quad (2)$$

对于 m_i 和 τ_i ($i = 1$) 的选取, 目前仍没有统一的方法. 本文采用 C-C 方法, 它具有计算量小、易操作和抗噪性能强的优点^[8]. 设预测时域为 h , 得到式 (1) 后, 建立 k 时刻嵌入延迟向量与预测值 $y_a(k+h)$ 间的定量关系. 由 Taken 定理可知, 一定存在 $F(\cdot)$, 使

$$\mathbf{x}(k+1) = F(\mathbf{x}(k))$$

假设 $y_a(k)$ 能够通过量测函数 $g(\cdot)$ 观测延迟坐标向量 (1) 得到

$$y_a(k) = g(\mathbf{x}(k))$$

相应地, 有下式成立

$$\mathbf{x}(k) = \mathbf{x}(k)$$

收稿日期 2008-12-26 录用日期 2009-03-24
Manuscript received December 26, 2008; accepted March 24, 2009
国家高技术研究发展计划 (863 计划) (2007AA04Z158), 国家自然科学基金 (60674073), 大连理工大学 2007 年“数学+X”交叉学科建设专项项目资助
Supported by National High Technology Research and Development Program of China (863 Program) (2007AA04Z158), National Natural Science Foundation of China (60674073), and 2007 “Math+X” Special Cross-disciplinary Construction of Dalian University of Technology

1. 大连理工大学电信学院 大连 116024
1. School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024

$$\begin{aligned} \mathbf{x}(k+1) &= F(\mathbf{x}(k)) \\ &\vdots \\ \mathbf{x}(k+h) &= F(\mathbf{x}(k+h-1)) \\ y_a(k+h) &= g(\mathbf{x}(k+h)) \end{aligned}$$

令 $F_0(\cdot) \equiv 0$, 将上面公式写作状态空间表达式形式

$$\begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k+1) \\ \vdots \\ \mathbf{x}(k+h) \end{bmatrix} = \begin{bmatrix} F_0(\mathbf{x}(k-1)) \\ F(\mathbf{x}(k)) \\ \vdots \\ F(\mathbf{x}(k+h-1)) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \mathbf{x}(k) \quad (3)$$

$$y_a(k+h) = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k+1) \\ \vdots \\ g(\mathbf{x}(k+h)) \end{bmatrix} \quad (4)$$

在 k 时刻, 状态空间表达式的输入为向量 $\mathbf{x}(k)$, 状态向量为 $[\mathbf{x}(k)^T, \mathbf{x}(k+1)^T, \dots, \mathbf{x}(k+h)^T]^T$, 输出信号为 $y_a(k+h)$. 预测原点和预测时域时间上的关系通过式 (3) 和式 (4) 转化为“空间”上的关系, 为使用递归网络直接预测多元时间序列奠定了基础.

2 储备池 KF 预测算法描述

本节将对储备池 KF 算法进行具体介绍, 给出算法的收敛性证明, 并指出测量误差协方差矩阵 $R(k)$ 对收敛速率具有较大影响. 同时, 从不同方面将本文方法与常规 RNN 的 EKF 方法进行比较分析.

2.1 储备池 KF 在线预测算法

首先, 对常规 RNN 网络的 EKF 算法进行简单介绍, 进而引出本文的储备池 KF 在线算法. 常规单输出 RNN 网络一步预测状态空间方程可表示为

$$\mathbf{x}(k+1) = f(\mathbf{u}(k), \mathbf{x}(k); \mathbf{w}_f) \quad (5)$$

$$y(k+1) = g(\mathbf{x}(k+1), y(k); \mathbf{w}_g) \quad (6)$$

其中, $\mathbf{u}(k)$ 为 k 时刻的嵌入延迟向量, $\mathbf{x}(k)$ 为隐含层状态, $y(k+1)$ 为网络输出. 用 EKF 算法训练 RNN 网络时, 网络中所有连接权值均为待估状态, 写成状态空间表达式形式为

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \boldsymbol{\omega}(k) \quad (7)$$

$$y_p(k+1) = h(\mathbf{w}(k), \mathbf{u}(k), \mathbf{x}(k)), y(k)) + v(k) \quad (8)$$

这里, $y_p(k+1)$ 为实际观测值, $\boldsymbol{\omega}(k)$ 和 $v(k)$ 为高斯白噪声. 对于常规 RNN 网络, 有

$$\mathbf{w}(k) = \begin{bmatrix} \mathbf{w}_f(k) \\ \mathbf{w}_g(k) \end{bmatrix}$$

取 $Q(k)$ 和 $R(k)$ 分别为过程噪声 $\boldsymbol{\omega}(k)$ 和测量噪声 $v(k)$ 的协方差矩阵

$$Q(k) = E[\boldsymbol{\omega}(k)\boldsymbol{\omega}(k)^T] \quad (9)$$

$$R(k) = E[v(k)v(k)^T] \quad (10)$$

Kalman 滤波器具有两个基本工作步骤. 第一步为时间更新, 通过前一刻的状态 $\hat{\mathbf{w}}(k-1)$ 和误差协方差矩阵 $P(k-1)$ 得到当前时刻状态和误差协方差矩阵的先验估计 $\hat{\mathbf{w}}^-(k)$ 和 $P^-(k)$

$$\hat{\mathbf{w}}^-(k) = F(k-1)\hat{\mathbf{w}}(k-1) \quad (11)$$

$$P^-(k) = F(k-1)P(k-1)F(k-1)^T + Q(k-1) \quad (12)$$

其中, $F(k)$ 为状态转移矩阵. 第二步为测量更新, 通过实际测量值对状态和误差协方差矩阵进行校正, 估计误差协方差矩阵和状态的更新式为

$$P(k) = (I - K(k)H(k)^T)P^-(k) \quad (13)$$

$$\hat{\mathbf{w}}(k) = \hat{\mathbf{w}}^-(k) + K(k)e(k) \quad (14)$$

其中, $e(k) = y_p(k) - y^-(k)$. $K(k)$ 为 Kalman 滤波增益, 计算如下

$$A(k) = (R(k) + H(k)^T P^-(k) H(k))^{-1} \quad (15)$$

$$K(k) = P^-(k) H(k) A(k) \quad (16)$$

在用 EKF 算法训练常规 RNN 网络时, 状态转移矩阵 $F(k) = I$, 相应的测量矩阵近似为

$$H(k) \approx \frac{\partial h(\mathbf{w}(k))}{\partial \mathbf{w}(k)} = \left[\frac{\partial h(k)}{\partial w_1}, \frac{\partial h(k)}{\partial w_2}, \dots, \frac{\partial h(k)}{\partial w_N} \right]^T \quad (17)$$

其中, N 为网络可调权值总数, $H(k)$ 为网络输出对可调权值的 Jacobian 矩阵. 由于 EKF 方法需计算 Jacobian 矩阵, 不能用于 $h(\cdot)$ 不可导或导数求解困难的系统. 同时, RNN 网络方程的形式一般比较复杂, 需应用 BPTT 或 RTRL 算法求得 Jacobian 矩阵, 令运算量增大. 因而, 从运算精度、复杂度和应用范围来说, 传统 RNN 的 EKF 算法均存在不足.

ESN 网络具有独特的非线性处理机制, 其网络方程为

$$\mathbf{x}(k+1) = \text{tansig}(W_x \cdot \mathbf{x}(k) + W_{in} \cdot \mathbf{u}(k)) \quad (18)$$

$$y(k) = \mathbf{w}(k)^T \mathbf{x}(k) \quad (19)$$

其中, $\mathbf{x}(k)$ 、 $\mathbf{u}(k)$ 代表 k 时刻储备池状态变量和输入向量, W_x 、 W_{in} 分别为储备池内部的连接矩阵和外部输入至储备池的连接矩阵. 通常, 储备池状态 \mathbf{x} 的维数较高 (一般取 100~1000 之间), 矩阵 W_x 保持 1%~5% 的稀疏连接, W_x 和 W_{in} 一经初始化后就保持不变, 当 W_x 的谱半径 $\rho(W_x) \leq 1$ 时^[9], ESN 网络为稳定运行, 学习的目的是求取最优输出向量 \mathbf{w} .

在最优估计理论和储备池机理的框架下, KF 算法可以直接用来在线估计 ESN 的输出权值, 相应的观测方程为

$$y_p(k) = \mathbf{w}(k)^T \mathbf{s}(k) + v(k) \quad (20)$$

需要注意的是, 与常规 RNN 不同, 这里 $\mathbf{w}(k)$ 仅表示输出权值. $\mathbf{s}(k)$ 为 k 时刻 ESN 储备池的内部状态变量 $\mathbf{x}(k)$, 它在储备池 KF 算法中充当了测量矩阵 $H(k)$. 对比式 (8) 和式 (20) 可知, 观测方程 (20) 较为简单. 这是因为在 ESN 中, 对于给定的网络输入, 可根据式 (18) 方便地求得储备池状态变量 $\mathbf{s}(k)$, 进而直接应用式 (11)~(16) 计算出 ESN 的输出权值.

具有渐消记忆功能的 Kalman 滤波算法只需将式 (11) 替换成下式即可

$$\hat{\boldsymbol{w}}^-(k) = \lambda^{-\frac{1}{2}} \hat{\boldsymbol{w}}^-(k-1), \quad 0 < \lambda \leq 1 \quad (21)$$

其中, λ 为遗忘因子.

2.2 储备池 KF 算法收敛性分析

本节将对储备池 KF 算法的收敛性进行证明. 为方便表达, 分析过程中用下标表示相应时刻的取值. 定义 $k+1$ 时刻状态估计误差

$$\tilde{\boldsymbol{w}}_{k+1} = \boldsymbol{w}^* - \hat{\boldsymbol{w}}_{k+1} \quad (22)$$

\boldsymbol{w}^* 为最优输出权值, 定义准 Lyapunov 函数为

$$V_{k+1} = \tilde{\boldsymbol{w}}_{k+1}^T P_{k+1}^{-1} \tilde{\boldsymbol{w}}_{k+1} \quad (23)$$

定理 1. 当 ESN 网络满足 $\rho(W_x) \leq 1$ 时, 通过式 (11) ~ (16) 得到 $\hat{\boldsymbol{w}}_k$. 当 $k \rightarrow \infty$ 时, 必有 $\tilde{\boldsymbol{w}}_k = 0$ 成立. 即当 ESN 稳定运行时, 储备池 KF 算法必收敛.

为了证明定理 1 成立, 我们首先证明 $\{V_k\}_{k=1, \dots, \infty}$ 为递减数列. 令 $F_k = I$, $Q_k = 0$, 由式 (11) 和 (14) 得

$$\hat{\boldsymbol{w}}_{k+1} = \hat{\boldsymbol{w}}_k + K_{k+1} e_{k+1} \quad (24)$$

$$\begin{aligned} e_{k+1} &= y_{k+1}^p - \hat{y}_{k+1} = H_{k+1}^T (\boldsymbol{w}^* - \hat{\boldsymbol{w}}_{k+1}) = \\ &= H_{k+1}^T (\boldsymbol{w}^* - \hat{\boldsymbol{w}}_k) = H_{k+1}^T \tilde{\boldsymbol{w}}_k \end{aligned} \quad (25)$$

在式 (24) 两边同时减 \boldsymbol{w}^* , 整理得到

$$\tilde{\boldsymbol{w}}_{k+1} = \tilde{\boldsymbol{w}}_k - K_{k+1} e_{k+1} \quad (26)$$

另一方面, 对式 (13)、(15) 和 (16) 进行等式变换, 可得

$$P_{k+1} H_{k+1} R_{k+1}^{-1} = P_k H_{k+1} A_{k+1} = K_{k+1} \quad (27)$$

$$P_{k+1}^{-1} = (I + H_{k+1} R_{k+1}^{-1} H_{k+1}^T P_k)^{-1} = P_k^{-1} + H_{k+1} R_{k+1}^{-1} H_{k+1}^T \quad (28)$$

将式 (25) ~ (28) 分别代入 (23), 整理得到

$$\begin{aligned} V_{k+1} &= (\tilde{\boldsymbol{w}}_k - P_{k+1} H_{k+1} R_{k+1}^{-1} e_{k+1})^T \times \\ &P_{k+1}^{-1} (\tilde{\boldsymbol{w}}_k - P_{k+1} H_{k+1} R_{k+1}^{-1} e_{k+1}) = \\ &V_k + e_{k+1}^T R_{k+1}^{-1} (H_{k+1}^T P_{k+1} H_{k+1} R_{k+1}^{-1} - I) e_{k+1} = \\ &V_k + e_{k+1}^T R_{k+1}^{-1} (H_{k+1}^T P_k H_{k+1} A_{k+1} - I) e_{k+1} \end{aligned}$$

由于 $F_k = I$, $Q_k = 0$, 将其代入式 (12) 可得 $P_{k+1}^- = P_k$, 结合式 (15), 得到

$$\begin{aligned} V_{k+1} &= V_k + e_{k+1}^T R_{k+1}^{-1} (H_{k+1}^T P_k H_{k+1} \\ &(R_{k+1} + H_{k+1}^T P_{k+1}^{-1} H_{k+1})^{-1} - I) e_{k+1} = \\ &V_k + e_{k+1}^T R_{k+1}^{-1} (H_{k+1}^T P_k H_{k+1} \\ &(R_{k+1} + H_{k+1}^T P_k H_{k+1})^{-1} - I) e_{k+1} \end{aligned} \quad (29)$$

通常, 取 $R_k = \sigma_0 I = R$ 为恒定的正对角阵. 要令 $V_{k+1} \leq V_k$, 只需保证

$$\Psi = H_{k+1}^T P_k H_{k+1} (R_{k+1} + H_{k+1}^T P_k H_{k+1})^{-1} - I \quad (30)$$

为负定即可, 该结论显然成立. 因此 $\{V_k\}_{k=1, \dots, \infty}$ 为递减数列. 同时, 由定义可知 $\forall V_k \geq 0$, 根据单调有界原理, 必有

$$\lim_{k \rightarrow \infty} V_k = V \quad (31)$$

另一方面, 由式 (28) 可得

$$P_{k+1}^{-1} = \Theta(1, k+1)^T \Gamma(1, k+1) \Theta(1, k+1) + P_0^{-1} \quad (32)$$

其中,

$$\Theta(1, k+1) = [H_1, H_2, \dots, H_{k+1}]^T$$

$$\Gamma(1, k+1) = \text{diag} \{R_1^{-1}, R_2^{-1}, \dots, R_{k+1}^{-1}\}$$

引理 1. 对于辨识算法 (11) ~ (16), 若存在正实数 σ_1 , σ_2 和正整数 n , 对于 $\forall k > n$ 均满足^[10]

$$0 < k\sigma_1 \leq \lambda(P_k^{-1} - P_0^{-1}) \leq k\sigma_2 \quad (33)$$

则有

$$\lim_{k \rightarrow \infty} \lambda_{\min}(P_k^{-1}) = \infty \quad (34)$$

$$\text{Sup} \lim_{k \rightarrow \infty} \frac{\lambda_{\max}(P_k^{-1})}{\lambda_{\min}(P_k^{-1})} < \infty \quad (35)$$

在 ESN 中, 当储备池谱半径 $\rho(W_x) \leq 1$ 时, 网络稳定运行. 由式 (18) 求得储备池内部状态变量充当 KF 算法的测量矩阵 H_k 时, 恒有

$$\|H_k\|_{\infty} = \|\boldsymbol{x}_k\|_{\infty} < 1 \quad (36)$$

此时, 通过不等式放大易得不等式 (33) 右侧成立. 同时, 当 n 足够大时, 对 $\forall k > n$ 一定有

$$\Omega = P_k^{-1} - P_0^{-1} = \Theta(1, k)^T \Gamma(1, k) \Theta(1, k)$$

为满秩. 又由式 (36) 可得等式 (33) 左侧成立. 进而由引理 1, 得到式 (34) 和 (35) 成立. 又因为

$$\frac{V_k}{\text{tr}(P_k^{-1})} \geq \frac{\lambda_{\min}(P_k^{-1}) \tilde{\boldsymbol{w}}_k^T \tilde{\boldsymbol{w}}_k}{d \lambda_{\max}(P_k^{-1})} \geq 0 \quad (37)$$

且 $\lim_{k \rightarrow \infty} \text{tr}(P_k^{-1}) > \lim_{k \rightarrow \infty} \lambda_{\min}(P_k^{-1}) = \infty$, 所以

$$\lim_{k \rightarrow \infty} \frac{V_k}{\text{tr}(P_k^{-1})} = \lim_{k \rightarrow \infty} \frac{\lambda_{\min}(P_k^{-1}) \tilde{\boldsymbol{w}}_k^T \tilde{\boldsymbol{w}}_k}{d \lambda_{\max}(P_k^{-1})} = 0 \quad (38)$$

由式 (35) 和 (38) 得到

$$\lim_{k \rightarrow \infty} \tilde{\boldsymbol{w}}_k = 0 \quad (39)$$

因此, 本文算法在 ESN 网络稳定时收敛. \square

对式 (29) 进一步化简, 可得

$$\begin{aligned} V_{k+1} &= V_k - e_{k+1}^T (H_{k+1}^T P_k H_{k+1} + R_{k+1})^{-1} e_{k+1} = \\ &V_k - \tilde{\boldsymbol{w}}_k^T H_{k+1} (H_{k+1}^T P_k H_{k+1} + R_{k+1})^{-1} H_{k+1}^T \tilde{\boldsymbol{w}}_k \end{aligned} \quad (40)$$

从式 (40) 中可以看出, R_k 取值对算法收敛速率有较大的影响, R_k 取值大时收敛较慢, R_k 取值小时收敛较快. 而在式 (15) 中, R_k 起到了正则项的作用, 取值过小时 $A(k)$ 易呈“病态”. 因此, R_k 的取值需要根据具体数据进行折衷选取.

3 仿真实例

为验证本文方法的有效性,将其应用于 Lorenz 混沌方程的三变量时间序列预测仿真中,并与 Elman 网络的 EKF 算法^[5]进行比较。

引入均方根误差 E_{RMSE} 来定量说明预测性能

$$E_{RMSE} = \left(\frac{1}{S-1} \sum_{i=1}^S (P_i - O_i)^2 \right)^{\frac{1}{2}} \quad (41)$$

其中, S 为测试样本个数, O_i 是某个实际观测值, P_i 为对应预测输出。

Lorenz 混沌方程如式 (42) 所示, 当取 $a = 10, b = 8/3, c = 28, x(0) = y(0) = z(0) = 1.0$ 时, 系统产生混沌。利用四阶 Runge-Kutta 法迭代产生混沌时间序列, 使用 $x(t), y(t)$ 和 $z(t)$ 序列共同预测 $x(t + \eta)$ 。

$$\begin{cases} \dot{x} = a(y - x) \\ \dot{y} = (c - z)x - y \\ \dot{z} = xy - bz \end{cases} \quad (42)$$

根据 C-C 方法计算得到 $\tau_1 = 19, \tau_2 = 13, \tau_3 = 12, m_1 = 3, m_2 = 5, m_3 = 7$, 相应的 15 维延迟坐标向量为

$$\mathbf{d}(t) = \{x(t), \dots, x(t - 2\tau_1), y(t), \dots, y(t - 4\tau_2), z(t), \dots, z(t - 6\tau_3)\}^T = \{\mathbf{x}^T, \mathbf{y}^T, \mathbf{z}^T\}^T$$

原序列长度为 2501, $\eta = 1$ 时经相空间重构生成 2428 组数据。令 $\mathbf{u}(k) = \mathbf{d}(k)$, 预测输出为 $x(t + \eta)$, 分别用 Elman 网

络结合 EKF 算法和本文的储备池 KF 算法对其进行预测建模。在 Elman 网络中, 隐含层节点数选为 5, 网络的可调权重总数为 111。Jacobian 矩阵通过截断 BPTT 算法求取, 截断长度 p 取 10。而 ESN 储备池的参数设置如表 1 所示。

表 1 ESN 的储备池参数设置

| 储备池参数 | 取值 |
|---------------|-----------------|
| 储备池维数 | 200 |
| 矩阵 W_x 稀疏度 | 2% |
| 矩阵 W_x 谱半径 | 0.98 |
| 输入权值 W_{in} | $[-1, +1]$ 均匀分布 |

两种方法自适应预测曲线和自适应误差曲线如图 1~4 所示。从图中可以看出递归网络受到初始暂态影响, 在训练前期产生较大偏差, 经一段时间以后逐渐趋于稳定。舍弃前面若干暂态点求取误差, 得到均方根误差结果如表 2 和表 3 所示。从图 1、2 和表 2 可以看出, Elman 网络结合 EKF 算法在第 1000 个预测点左右趋于稳定, 稳定后的预测误差在 0.4 左右。从图 3、4 和表 3 可以看出, 本文储备池 KF 方法具有较短的暂态过程, 预测结果第 400 个预测点左右已趋于稳定, 且稳定后预测误差在 0.04 左右, 优于传统方法。

对照两种方法 $x(t)$ 时间序列的预测结果可知, 本文方法的预测性能较常规方法有较大提升, 相应的暂态过程更短且预测精度更高。究其原因在于预测建模过程中, 非线性部分和动态部分均已在储备池内部完成, 储备池 KF 算法本质上是一种线性运算, 能够精确计算且不存在局部最优问题。而常规递归网络 EKF 方法需要线性化近似且存在局部最优, 对参数的初始值敏感, 因而实现效率比本文算法要低。

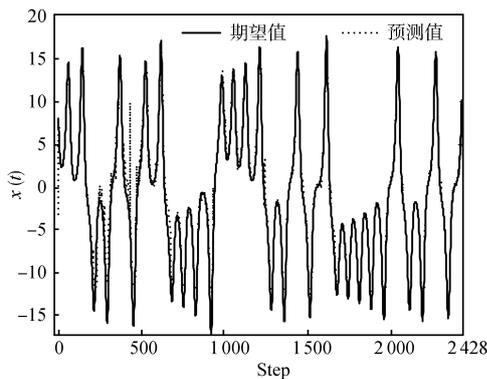


图 1 常规 RNN 在 EKF 算法下的预测曲线

Fig. 1 Prediction curves of EKF method in traditional RNN

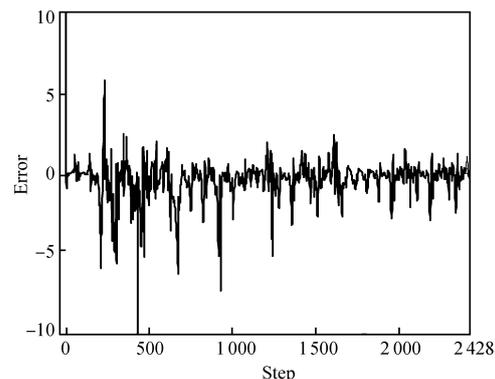


图 2 常规 RNN 在 EKF 算法下的预测误差曲线

Fig. 2 Prediction error curve of EKF method in traditional RNN

表 2 常规 RNN 网络 EKF 算法舍弃一定暂态后的预测均方根误差 (50 次随机试验取平均)

Table 2 Prediction E_{RMSE} of EKF method in traditional RNN excluded transient process with different lengths

| 舍弃暂态长度 | 200 | 400 | 600 | 800 | 1000 | 1200 | |
|------------|-------------|--------|--------|--------|--------|--------|--------|
| E_{RMSE} | $\eta = 1$ | 0.7449 | 0.6245 | 0.5829 | 0.5021 | 0.4493 | 0.4376 |
| | $\eta = 20$ | 2.0020 | 1.9496 | 1.8644 | 1.7743 | 1.6201 | 1.5725 |

表 3 储备池 KF 算法舍弃一定暂态后的预测均方根误差 (50 次随机试验取平均)

Table 3 Prediction E_{RMSE} of reservoir KF method excluded transient process with different lengths

| 舍弃暂态长度 | 100 | 200 | 400 | 600 | 800 | 1000 | |
|------------|-------------|--------|--------|--------|--------|--------|--------|
| E_{RMSE} | $\eta = 1$ | 0.1297 | 0.0967 | 0.0408 | 0.0388 | 0.0386 | 0.0342 |
| | $\eta = 20$ | 0.7305 | 0.7174 | 0.6875 | 0.6782 | 0.6761 | 0.6556 |

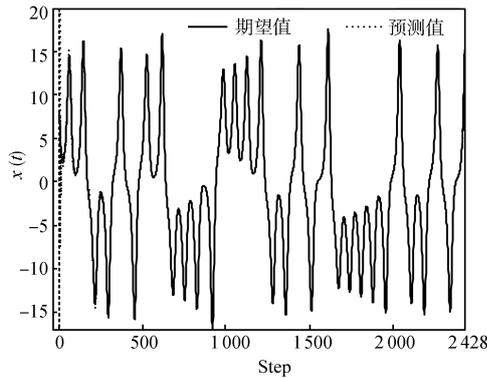


图3 储备池 KF 算法下的预测曲线

Fig. 3 Prediction curves of reservoir KF method

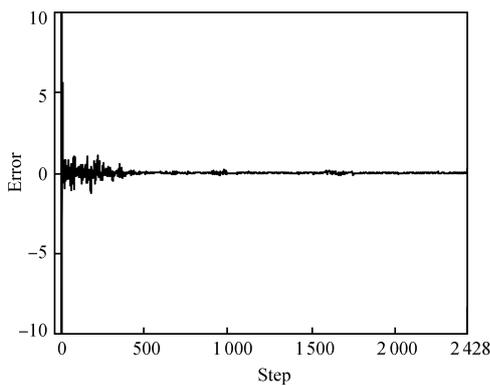


图4 储备池 KF 算法训练下的预测误差曲线

Fig. 4 Prediction err curve of reservoir KF method

4 结论

本文在储备池方法和 KF 算法基础之上, 针对复杂多元时间序列提出一种新的自适应预报方法. 重构多元相空间后, 通过 ESN 储备池将嵌入延迟向量至预测输出间的非线性动态映射转化为高维储备池状态至预测输出间的线性静态映射, 并运用 KF 算法随数据的逐步加入实时更新网络输出权重. 给出本文算法在 ESN 稳定时为收敛的证明, 并通过仿真, 比较 Elman 网络 EKF 算法与本文储备池 KF 算法的性能差异. 由于常规 EKF 需要对非线性部分进行线性化近似, 利用输出对可调参数导数的一阶 Taylor 展开来近似 KF 方程中的测量矩阵, 算法本身存在近似偏差, 效率较低. 而本文方法将非线性部分完全分离出去由储备池处理, KF 完成余下的线性参数估计问题, 因而算法效率较高. 实验结果表明, 本文的在线储备池 KF 算法为复杂系统的辨识提供了一条有效的途径.

References

- 1 Wang Yong-Li, Zhou Jing-Hua, Xu Hong-Bing, Dong Yi-Sheng, Liu Xue-Jun. An adaptive forecasting method for time-series data streams. *Acta Automatica Sinica*, 2007, **32**(2): 197–201
(王永利, 周景华, 徐宏炳, 董逸生, 刘学军. 时间序列数据流的自适应预测. *自动化学报*, 2007, **32**(2): 197–201)
- 2 Atiya A F, Parlos A G. New results on recurrent network training: unifying the algorithms and accelerating conver-

gences. *IEEE Transactions on Neural Networks*, 2000, **11**(3): 697–709

- 3 Sayed A H. A framework for state-space estimation with uncertain models. *IEEE Transactions on Automatic Control*, 2001, **46**(7): 998–1013
- 4 Sum J, Leung C S, Young G H, Kan W K. On the Kalman filtering method in neural network training and pruning. *IEEE Transactions on Neural Networks*, 1999, **10**(1): 161–166
- 5 Cernansky M, Benuskova L. Simple recurrent network trained by RTRL and extended Kalman filter algorithms. *Neural Network World*, 2003, **13**(3): 223–234
- 6 Choi J, Yeap T H, Bouchard M. Online state-space modeling using recurrent multilayer perceptrons with unscented Kalman filter. *Neural Processing Letters*, 2002, **22**(1): 69–84
- 7 Jaeger H, Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science*, 2004, **304**(5667): 78–80
- 8 Kim H S, Eykholt R, Salas J D. Nonlinear dynamics, delay times, and embedding windows. *Physica D*, 1999, **127**(1-2): 48–60
- 9 Buehner M, Young P. A tighter bound for the echo state property. *IEEE Transactions on Neural Networks*, 2006, **17**(3): 820–824
- 10 Boutayed M, Rafaralahy H, Darouach M. Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Transactions on Automatic Control*, 1997, **42**(4): 581–586

韩敏 大连理工大学教授. 主要研究方向为神经网络理论及其应用, 复杂系统建模及自适应控制. 本文通信作者.

E-mail: minhan@dlut.edu.cn

(HAN Min Professor at Dalian University of Technology. Her research interest covers neural networks theory and application, complex systems modeling, and adaptive control. Corresponding author of this paper.)

王亚楠 大连理工大学硕士研究生. 主要研究方向为神经网络应用和机器学习. E-mail: wyanan@student.dlut.edu.cn

(WANG Ya-Nan Master student at Dalian University of Technology. Her research interest covers neural networks application and machine learning.)