

回声状态网络的递推训练算法

雷晓义 曹柳林* 余 晋

(北京化工大学 信息科学与技术学院, 北京 100029)

摘要: 针对回声状态网络(ESN)传统的训练方法无法解决高维矩阵不可逆时的训练,以及无法应用于需要在线训练的建模当中等问题,提出了两种新的递推训练算法。分别将含遗忘因子递推最小二乘算法(FRRLS)和无先导卡尔曼滤波算法(UKF)应用到回声状态网络输出神经元为线性函数和非线性函数的权值训练中,进而直接对网络的输出权值进行递推更新。与传统的训练方法相比,所提新方法不仅具有在线更新、精度高的优点,而且还可以解决传统训练方法中批量数据构成的向量矩阵不可逆及输出神经元为非线性函数且其反函数不可求的问题。通过对连续搅拌釜式反应器(CSTR)浓度和温度的预测仿真,结果证明了所提新方法的有效性。

关键词: 回声状态网络(ESN); 遗忘因子递推最小二乘算法(FRRLS); 无先导卡尔曼滤波算法(UKF)

中图分类号: TP183

引 言

神经网络技术作为一种有效的非线性黑箱建模方法,在复杂系统建模、非线性系统预测等领域得到了广泛的应用。特别是递归神经网络由于其可以适应动态系统的时变特性,与前馈网络相比更适合动态系统建模。回声状态网络(ESN)是由 Jaeger^[1]提出的一种新型的递归神经网络,并且在非线性时间序列预测等领域取得了成功应用。但是 ESN 传统训练方法采用离线批处理算法,虽然方法简单、有效,但是训练数据规模、存储量以及运算量较大,并且不能应用于在线辨识和建模的场合。另外,输出权值的计算需要对矩阵直接求逆,这对于高维数据矩阵由于共线性而导致病态方程无法求解,或某些系统数据组成的矩阵无法求逆时不能使用。

针对上述问题葛谦等^[2]利用粒子群算法训练 ESN 的输出权值,虽然能在有限的学习样本下有效的提高预测精度,但是该方法由于演化粒子和迭代次数较多,使得模型训练时间较长。王建民等^[3]将 ESN 的储备池参数以及输出权矩阵作为目标参数,利用扩展卡尔曼滤波(EKF)对其进行联合递归训练,但是 EKF 算法在计算中用到了误差协方差矩阵

的近似,最终只得到最优解的一阶近似。鉴于此,本文提出了两种 ESN 权值的递推训练算法,将 FRRLS 和 UKF 算法分别应用到 ESN 输出神经元为线性函数和非线性函数情况下权值的训练中。所提方法不但具有训练样本少、运算量小、可在线更新以及精度高的优点,而且拓展了 ESN 建模的应用范围。

1 回声状态网络

ESN 是一种离散型神经网络,结构如图 1 所示。

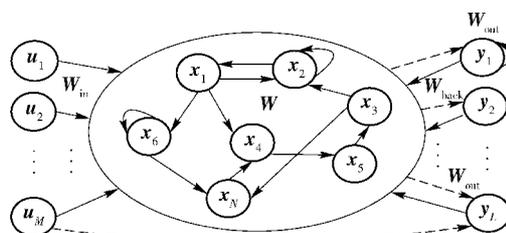


图 1 回声状态网络结构

Fig. 1 Structure of ESN

状态方程和输出方程分别如式(1)~(2)所示。

$$\mathbf{x}(k+1) = \mathbf{f}(\mathbf{W}_{in} \mathbf{u}(k+1) + \mathbf{W} \mathbf{x}(k) + \mathbf{W}_{back} \mathbf{y}(k)) \quad (1)$$

$$\mathbf{y}(k+1) = \mathbf{f}_{out}(\mathbf{W}_{out}(\mathbf{u}(k+1), \mathbf{x}(k+1), \mathbf{y}(k))) \quad (2)$$

其中 $\mathbf{u}(k) = (u_1(k) \cdots u_M(k))^T \in \mathbf{R}^M$, $\mathbf{x}(k) = (x_1(k) \cdots x_N(k))^T \in \mathbf{R}^N$, $\mathbf{y}(k) = (y_1(k) \cdots y_L(k))^T \in \mathbf{R}^L$, 分别表示 k 时刻的输入向量、状态向量和输出向量。 $\mathbf{W}_{in} \in \mathbf{R}^{N \times M}$, $\mathbf{W} \in \mathbf{R}^{N \times N}$, $\mathbf{W}_{back} \in \mathbf{R}^{N \times L}$ 分别表

收稿日期: 2012-08-16

基金项目: 国家自然科学基金(60974031/61174128)

第一作者: 男, 1987 年生, 硕士生

* 通讯联系人

E-mail: caoll@mail.buct.edu.cn

示输入、内部状态和反馈连接权值矩阵, 这些权值在初始化时随机生成并且固定不变。 $\mathbf{W}_{out} \in \mathbf{R}^{L \times (N+L+M)}$ 表示输出连接权矩阵, 需要通过网络训练得到。 $\mathbf{f}(\cdot) = (f_1(\cdot) \cdots f_N(\cdot))$ 表示神经元激活函数一般选为 sigmoid 函数。 $\mathbf{f}_{out}(\cdot) = (f_{out,1}(\cdot) \cdots f_{out,L}(\cdot))$ 表示输出神经元函数, 根据问题的复杂程度不同, 可以选为线性或非线性函数。

ESN 的训练过程就是根据输入、输出训练样本对 $(\mathbf{u}(k), \mathbf{y}(k), k=1, 2 \cdots T)$ 确定输出权值矩阵 \mathbf{W}_{out} 的过程, 可作如下描述。

1) 对 \mathbf{W}_{in} 、 \mathbf{W} 、 $\mathbf{x}(0)$ 等参数进行初始化。

2) 依次将教师信号 $(k=0, 1 \cdots T)$ 分别送入输入单元与输出单元, 根据式(1) ~ (2) 进行网络更新。

3) 将每个时间下标的状态向量 $\mathbf{x}(k)$ 作为单独一行组成规模为 $(T - T_0 + 1) \times (M + N + L)$ 的状态矩阵 \mathbf{B} 。同样收集对应的 $(T - T_0 + 1) \times L$ 教师信号 \mathbf{S} , 其中 T_0 为需要删除的开始不准确数据的个数。

4) 输出权值矩阵 \mathbf{W}_{out} 由式(3)得到

$$\mathbf{W}_{out} = \mathbf{B}^{-1} \mathbf{S} \quad (3)$$

至此, ESN 的训练已经完成, 对于训练好的 ESN, 可

$$\mathbf{Q}_F = \begin{bmatrix} \mathbf{q}^T(1) \\ \mathbf{q}^T(2) \\ \vdots \\ \mathbf{q}^T(F) \end{bmatrix} = \begin{bmatrix} u_1(1) & \cdots & u_M(1) & x_1(1) & \cdots & x_N(1) & y_1(0) & \cdots & y_L(0) \\ u_1(2) & \cdots & u_M(2) & x_1(2) & \cdots & x_N(2) & y_1(1) & \cdots & y_L(1) \\ \vdots & \vdots \\ u_1(F) & \cdots & u_M(F) & x_1(F) & \cdots & x_N(F) & y_1(F-1) & \cdots & y_L(F-1) \end{bmatrix} \quad (6)$$

$\mathbf{E} \in \mathbf{R}^{F \times L}$ 为白噪声矩阵, 而把回声状态网络所需要训练的权值 \mathbf{W}_{out} 当作参数估计值如式(7)所示

$$\mathbf{W}_{out} = [a_1 \ a_2 \ \cdots \ a_M \ b_1 \ b_2 \ \cdots \ b_N \ c_1 \ c_2 \ \cdots \ c_L]^T \quad (7)$$

这样通过式(5) ~ (6) 就推导出批量的最小二乘公式, 进而将推导中增广矩阵的 k 时刻与遗忘因子递推最小二乘算法和 UKF 算法相结合就可以推导出回声状态网络权值的递推训练算法。

2.2 基于 FFRLS 的 ESN 权值训练算法

本文将基于遗忘因子的最小二乘算法应用在回声状态网络输出神经元为线性函数权值的递推训练中, 此时式(6)变为式(8)所示的形式。其中 $\beta(0 < \beta < 1)$ 为衰减因子。由于式(5)中输出函数为线性函数, 所以通过让式(9)所示的目标函数最小, 就可以推导出用于回声状态网络权值训练的遗忘因子递推最小二乘算法, 如式(10) ~ (12) 所示。

$$\mathbf{Q}_F = [\beta^{F-1} \mathbf{q}^T(1) \ \beta^{F-2} \mathbf{q}^T(2) \ \cdots \ \mathbf{q}^T(F)]^T \quad (8)$$

$$J(\mathbf{W}_{out}(k)) = \sum_{k=1}^F [\mathbf{y}(k) - \mathbf{q}^T(k) \mathbf{W}_{out}(k)]^2 =$$

以通过新的输入值 $\mathbf{u}(k)$, 利用式(1) ~ (2) 得到对应的输出预测值 $\mathbf{y}(k)$ 。

2 回声状态网络的递推训练算法

2.1 算法推导

为了实现回声状态网络权值的递推训练, 必须推导出权值训练的递推公式。考虑到 \mathbf{W}_{back} 可选择(这里不考虑此部分), 将回声状态网络更新方程(1)代入输出方程(2)中得到 ESN 的综合方程(4)。

$$\mathbf{y}(k+1) = \mathbf{f}_{out}(\mathbf{W}_{out}[\mathbf{u}(k+1), \mathbf{f}(\mathbf{W}_{in} \mathbf{u}(k+1) + \mathbf{W} \mathbf{x}(k))], \mathbf{y}(k)) \quad (4)$$

设采样个数为 F , 式(4)可用式(5)所示的矩阵形式表达。此表达式的构造为回声状态网络权值从批量训练方法转变为递推训练方法提供了依据。

$$\mathbf{Y} = \mathbf{F}_{out}(\mathbf{Q}_F \mathbf{W}_{out}) + \mathbf{E} \quad (5)$$

其中 $\mathbf{Y} \in \mathbf{R}^{F \times L}$ 为输出神经元构成的输出矩阵, \mathbf{F}_{out} 为输出神经元函数构成的矩阵, $\mathbf{Q}_F \in \mathbf{R}^{F \times (M+N+L)}$ 为由当前时刻的输入状态以及上一时刻的输出所组成的增广矩阵, 可以用式(6)表示

$$\sum_{k=1}^F [\mathbf{e}(k)]^2 \quad (9)$$

$$\mathbf{W}_{out}^+(k) = \mathbf{W}_{out}^+(k-1) + \mathbf{K}(k) [\mathbf{y}(k) - \mathbf{q}^T(k) \mathbf{W}_{out}^+(k-1)] \quad (10)$$

$$\mathbf{K}(k) = \mathbf{P}(k-1) \mathbf{q}(k) [\mathbf{q}^T(k) \mathbf{P}(k-1) \mathbf{q}(k) + \mu]^{-1} \quad (11)$$

$$\mathbf{P}(k) = (1/\mu) [\mathbf{I} - \mathbf{K}(k) \mathbf{q}^T(k)] \mathbf{P}(k-1) \quad (12)$$

其中, $\mu = \beta^2, 0 < \mu < 1, \mu$ 为遗忘因子。对于递推过程中所需的初值 $\mathbf{P}(0)$ 和 $\mathbf{W}_{out}^+(0)$, 一般选为 $\mathbf{P}(0) = \alpha \mathbf{I}, \mathbf{W}_{out}^+(0) = \boldsymbol{\varepsilon}$, 而 α 为充分大的正实数 ($10^4 \sim 10^6$), $\boldsymbol{\varepsilon}$ 为随机生成的实向量。

上述推导过程可以将基于遗忘因子最小二乘的 ESN 权值递推训练算法作如下概括。

1) 对输入维数 M 、状态池规模 N 、输出向量维数 L 、ESN 状态池谱半径 ρ 以及 FFRLS 中初值 $\mathbf{P}(0)$ 和 $\mathbf{W}_{out}^+(0)$ 等进行初始化。

2) 随机产生 ESN 的权值 \mathbf{W}_{in} 和 \mathbf{W} 。

3) 采集当前输入 $\mathbf{u}(k)$ 和输出 $\mathbf{y}(k)$ 。

4)根据式(10)~(12)计算 $\mathbf{K}(k)$ 、 $\mathbf{P}(k)$ 、 $\mathbf{W}_{out}^+(k)$,进行 ESN 网络权值的训练。

5) $k \rightarrow k+1$,返回 3),直到训练误差达到要求。

2.3 基于 UKF 的 ESN 权值训练算法

本文采用 UKF 算法对 ESN 权值进行训练。UKF 是由 Julier 等根据卡尔曼滤波的基本原理,从非线性均值和方差传播的角度提出的一种方法^[4]。该算法的核心思想和基础为 UT 变换,有关 UT 变换的原理参见文献[5]。

为了实现 UKF 用于 ESN 权值的递推训练,首先要构建合适的 UKF 状态空间表达式,进而将 ESN 的网络方程与所构造的状态空间表达式进行对照,推导出用于 ESN 权值训练的递推算法。不失一般性,在这里将参数估计的动态系统状态空间表达式(13)和(14)所示的形式

$$\mathbf{W}_{out}(k+1) = \mathbf{W}_{out}(k) + \mathbf{r}(k) \quad (13)$$

$$\mathbf{d}(k) = \mathbf{G}(\mathbf{q}^T(k), \mathbf{W}_{out}(k)) + \mathbf{e}(k) \quad (14)$$

其中式(13)为卡尔曼滤波的状态方程,本文将 ESN 需要训练的输出连接权值 \mathbf{W}_{out} 作为待估计的状态。式(14)为滤波器的输出观测方程,在此将其视为网络的输出方程。式(6)中的 $\mathbf{q}^T(k)$ 作为输入,将 ESN 输出神经元的非线性函数 f_{out} 记为 \mathbf{G} ,而 $\mathbf{r}(k)$ 、 $\mathbf{e}(k)$ 、 $\mathbf{d}(k)$ 分别为过程噪声、观测噪声和输出。

将利用 UKF 算法训练 ESN 权值的递推过程作如下概括。

1)对 ESN 输入维数 M 、状态池规模 N 、输出向量维数 L 以及递推所需量等进行初始化。

$$\tilde{\mathbf{W}}_{out}(0) = E[\mathbf{W}_{out}] \quad (15)$$

$$\mathbf{P}_{\mathbf{W}_{out}(0)} = E[(\mathbf{W}_{out} - \tilde{\mathbf{W}}_{out}(0))(\mathbf{W}_{out} - \tilde{\mathbf{W}}_{out}(0))^T] \quad (16)$$

2)时间更新和 sigma 点的计算如式(17)~(21)所示

$$\tilde{\mathbf{W}}_{out}^\wedge(k) = \tilde{\mathbf{W}}_{out}^\wedge(k-1) \quad (17)$$

$$\mathbf{P}_{\mathbf{W}_{out}(k)}^- = \mathbf{P}_{\mathbf{W}_{out}(k-1)} + \mathbf{R}^r(k-1) \quad (18)$$

$$\mathbf{J}^1(k|k-1) = [\tilde{\mathbf{W}}_{out}^\wedge(k), \tilde{\mathbf{W}}_{out}^\wedge(k) + \sqrt{(V+\lambda_1)\mathbf{P}_{\mathbf{W}_{out}(k)}^-}, \tilde{\mathbf{W}}_{out}^\wedge(k) - \sqrt{(V+\lambda_1)\mathbf{P}_{\mathbf{W}_{out}(k)}^-}] \quad (19)$$

$$\mathbf{D}(k|k-1) = \mathbf{G}(\mathbf{q}^T(k), \mathbf{J}^1(k|k-1)) \quad (20)$$

$$\hat{\mathbf{d}}(k) = \sum_{i=0}^{2V} \mathbf{W}_i^m \mathbf{D}_i(k|k-1) \quad (21)$$

3)根据式(22)~(26)进行测量更新

$$\mathbf{P}_{\hat{\mathbf{d}}(k)\hat{\mathbf{d}}(k)} = \sum_{i=0}^{2V} \mathbf{W}_i^c (\mathbf{D}_i(k|k-1) - \hat{\mathbf{d}}(k)) (\mathbf{D}_i(k|k-1) - \hat{\mathbf{d}}(k))^T + \mathbf{R}^e(k) \quad (22)$$

$$\mathbf{P}_{\mathbf{W}_{out}(k)\mathbf{d}(k)} = \sum_{i=0}^{2V} \mathbf{W}_i^c (\mathbf{J}_i^1(k|k-1) - \tilde{\mathbf{W}}_{out}^\wedge(k)) (\mathbf{D}_i(k|k-1) - \hat{\mathbf{d}}(k))^T \quad (23)$$

$$\boldsymbol{\kappa}(k) = \mathbf{P}_{\mathbf{W}_{out}(k)\mathbf{d}(k)} \mathbf{P}_{\hat{\mathbf{d}}(k)\hat{\mathbf{d}}(k)}^{-1} \quad (24)$$

$$\tilde{\mathbf{W}}_{out}^\wedge(k) = \tilde{\mathbf{W}}_{out}^\wedge(k) + \boldsymbol{\kappa}(k) (\mathbf{d}(k) - \hat{\mathbf{d}}(k)) \quad (25)$$

$$\mathbf{P}_{\mathbf{W}_{out}(k)} = \mathbf{P}_{\mathbf{W}_{out}(k)}^- - \boldsymbol{\kappa}(k) \mathbf{P}_{\hat{\mathbf{d}}(k)\hat{\mathbf{d}}(k)} \boldsymbol{\kappa}^T(k) \quad (26)$$

其中 \mathbf{R}^r 为过程噪声协方差, \mathbf{R}^e 为测量噪声协方差。 \mathbf{W}^m 为一阶统计特性权系数,也即均值的权系数。 \mathbf{W}^c 为二阶统计特性权系数,也即方差的权系数。

4)若满足要求则停止,否则返回 2)继续执行。

3 实验仿真结果

3.1 仿真对象

为了验证本文所提方法的有效性,将两种方法分别应用于某 CSTR 浓度和温度的建模中,并与经典的回声状态网络算法进行比较。

CSTR 物料平衡和能量平衡方程如式(27)

$$\begin{cases} \frac{dc_A}{dt} = \frac{q}{V}c_{A,f} - \frac{q}{V}c_A - R \\ \frac{dc_B}{dt} = \frac{q}{V}c_{B,f} - \frac{q}{V}c_B - R \\ \frac{dc_C}{dt} = -\frac{q}{V}c_C + R \\ \frac{dT}{dt} = \frac{q}{V}(T_f - T) + \frac{-\Delta H}{\rho C_p}R + \frac{UA}{V\rho C_p}(T_{c,f} - T) \\ R = Kc_Ac_B \\ K = k_0 \exp\left(-\frac{E_0}{R'T}\right) \end{cases} \quad (27)$$

式(27)中, R 为反应速率, c_A 、 c_B 为反应物 A、B 的浓度, c_C 是生成物的浓度, T 是反应温度, $c_{A,f}$ 、 $c_{B,f}$ 分别为反应物 A、B 的进料浓度, T_f 是进料温度, $T_{c,f}$ 是冷却剂温度。其他过程参数的含义和数值参考文献[6],其中除反应器体积、密度之外均为未知量。

为简化模型的状态空间表达,可选择 $c_{A,f}$ 、 $c_{B,f}$ 、 T_f 、 $T_{c,f}$ 作为模型的输入变量(操纵变量和扰动变量),即 $\mathbf{u} = [c_{A,f} \quad c_{B,f} \quad T_f \quad T_{c,f}]^T$ 。但是在实际过程中,一般进料流量 q 和进料温度 T_f 都作为控制变量引入,所以选择 q 和 T_f 作为输入变量更符合实际

工艺过程,即 $\mathbf{u} = [q \ T_f]^T$ 。选择反应物浓度 c_A 、 c_B , 生成物浓度 c_C , 反应温度 T 为输出变量, 即 $\mathbf{y} = [c_A \ c_B \ c_C \ T]^T$ 。对于 FFRLS 算法为了训练方便以及考虑到 c_C 及 T 更重要的实际情况, 在训练时将采用 4 个输出的加权平均相对误差作为训练停止的判断标准。此外为了试验的公平性, 在利用经典 ESN 批量算法、FFRLS、UKF 递推算法训练回声状态网络权值时, 将 ESN 网络的相关参数设置为一致。为了定量的体现预测的性能, 引入式 (28) 所示的平均相对误差 E 。

$$E = \frac{1}{S-1} \sum_{i=1}^S (|y_i^- - y_i|/y_i) \quad (28)$$

其中, S 为样本个数, y_i 为实际观测值, y_i^- 为预测输出值。

3.2 仿真结果及分析

为了充分说明所提方法的有效性, 本文对以下 3 种情形进行了实验仿真。情形 1: ESN 输出神经元函数为线性函数情况; 情形 2: ESN 输出神经元函数为非线性函数并且其反函数可以求得的情况; 情形 3: ESN 输出神经元函数为非线性函数并且其反函数不可求得的情况。这里仅给出了生成物浓度 c 的泛化误差曲线和分布情况, 如图 2 ~ 3 所示。

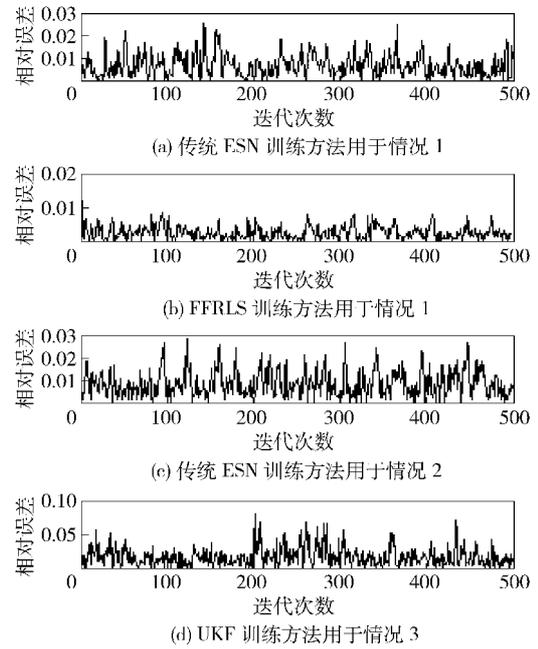


图 2 浓度 c_C 的泛化相对误差

Fig. 2 Tests of relative errors in the concentration c_C

从图 2 ~ 3 中可以看出与传统的 ESN 训练方法相比 FFRLS 方法泛化精度有明显提高; 对于传统的训练方法, 情形 1 和情形 2 的泛化平均相对误差相差不多; 对于情形 3 传统的 ESN 训练方法和 FFRLS

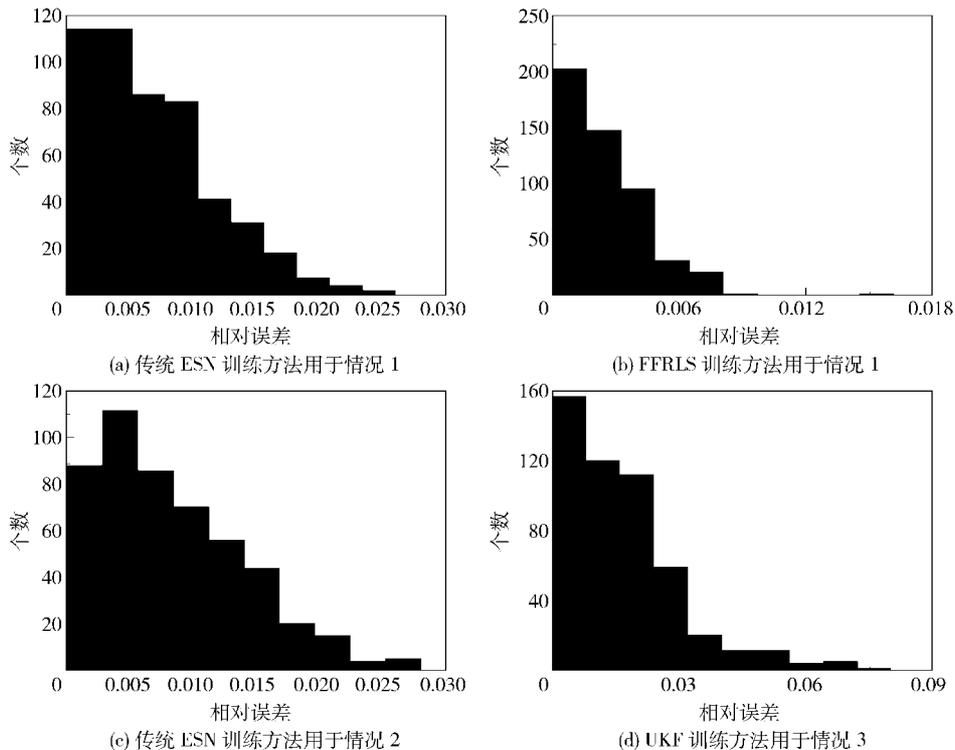


图 3 浓度 c_C 泛化相对误差分布情况

Fig. 3 The distribution of the test of relative errors in the concentration c_C

无法解决,而 UKF 训练方法可得到一个不错的结果。究其原因在于 FFRLS 算法是一种递推的训练算法,其可防止数据饱和,而 ESN 是一种批量训练方法,其对新老数据不加区分同样对待,并且通过线性求解,所以 FFRLS 的训练精度更高。传统的训练方法,情形 1 和情形 2 的泛化平均相对误差相差不多,说明针对本文的这个对象,情形 1 可以比较充分的体现对象的动态特性,所以即使输出神经元为非线性函数对网络的改善作用也不是很明显。UKF 算法其需要设置的参数比较多,且其需根据不同对象选择合适值,可能由于参数设置不恰当造成误差相对较大,但是泛化误差在可接受范围内。

4 结束语

本文在对 ESN、FFRLS、UKF 分析的基础上提出了两种不同的 ESN 权值递推训练的算法。FFRLS 算法不但避免了繁琐的矩阵求逆过程,而且提高了运算速度、精度以及实现了在线训练。而 UKF 算法解决了传统 ESN 训练方法不能解决的一些问题。仿真结果也证明了所提方法的有效性。此外,所提方法还为充分利用系统先验知识,构建结构逼近的系统模型提供条件,有关研究正在进行中。

参考文献:

[1] Jaeger H. The “echo state” approach to analyzing and

training recurrent neural networks [R] // GMD-German National Research Center for Information Technology, Bremen, 2001.

[2] 葛谦, 蔚承建. 基于 PSO 的回声状态网训练算法[J]. 计算机工程与设计, 2009, 30(8): 1947-1949.

Ge Q, Wei C J. Approach for optimizing echo state network training based on PSO [J]. Computer Engineering and Design, 2009, 30(8): 1947-1949. (in Chinese)

[3] 王建民, 彭宇, 彭喜元, 等. 基于扩展卡尔曼滤波的回声状态网络在线训练算法[J]. 仪器仪表学报, 2011, 32(7): 1514-1520.

Wang J M, Peng Y, Peng X Y, et al. An on-line training algorithm for echo state networks with extended kalman filter [J]. Chinese Journal of Scientific Instrument, 2011, 32(7): 1514-1520. (in Chinese)

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

[5] 刘铮. UKF 算法及其改进算法的研究[D]. 长沙: 中南大学, 2009: 1-65.

Liu Z. UKF algorithm and the research of the improved algorithm [D]. Changsha: Central South University, 2009: 1-65. (in Chinese)

[6] 刘华. 利用结构分解和参数分解的灰箱建模方法建立化学反应器模型[D]. 北京: 北京化工大学, 2011: 1-83.

Liu H. Simulation of Chemical processes using hybrid modeling method [D]. Beijing: Beijing University of Chemical Technology, 2011: 1-83. (in Chinese)

A recursive training algorithm for echo state networks (ESN)

LEI XiaoYi CAO LiuLin YU Jin

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: Two new recursive algorithms are proposed in the light of the problems of the irreversibility of the high dimensional matrix, and the inability to apply online training, associated with a traditional echo state network (ESN). We put forward a forgetting factor recursive least square (FFRLS) algorithm and an unscented Kalman filter (UKF) algorithm for the training of the connecting weights in association with the linear and nonlinear output neuron functions, which can directly and recursively update the output connecting weights. The proposed methods have the advantages of higher precision and can be updated online and, in addition, can solve problems associated with the traditional echo state network training methods, such as the batch data based matrix inversion being difficult to perform, and the inability to solve the inverse of the nonlinear output function. Simulations of the concentration and temperature in a continuous stirred tank reactor (CSTR) demonstrate the viability and effectiveness of our proposed methods.

Key words: echo state network (ESN); forgetting factor recursive least square (FFRLS); unscented Kalman filter (UKF)