

文章编号: 1001-0920(2012)08-1206-05

## 限定记忆极端学习机及其应用

张弦, 王宏力

(第二炮兵工程大学 自动控制工程系, 西安 710025)

**摘要:** 为了实现极端学习机(ELM)的在线训练, 提出一种限定记忆极端学习机(FM-ELM). FM-ELM以逐次增加新训练样本与删除旧训练样本的方式, 提高其对于系统动态变化特性的自适应性, 并根据矩阵求逆引理实现了网络输出权值的递推求解, 减小了在线训练过程的计算代价. 应用于具有动态变化特性的非线性系统在线状态预测表明, FM-ELM是一种有效的ELM在线训练模式, 相比于在线贯序极端学习机, FM-ELM具有更快的调节速度和更高的预测精度.

**关键词:** 神经网络; 极端学习机; 在线训练; 非线性系统

**中图分类号:** TP183

**文献标识码:** A

## Fixed-memory extreme learning machine and its applications

ZHANG Xian, WANG Hong-li

(Department of Automatic Control Engineering, The Second Artillery Engineering University, Xi'an 710025, China.

Correspondent: ZHANG Xian, E-mail: sltecas@163.com)

**Abstract:** To solve the problem of extreme learning machine(ELM) on-line training, an algorithm, fixed-memory extreme learning machine(FM-ELM), is proposed. FM-ELM adopts the latest training sample and abandons the oldest training sample iteratively to enhance its adaptive capacity. The output weights of FM-ELM are determined recursively based on Sherman-Morrison formula. Thus, the computational cost of FM-ELM training procedure is effectively reduced. Numerical experiments on nonlinear system on-line condition prediction show that FM-ELM has better performance in adjusting speed and prediction accuracy in comparison with on-line sequential extreme learning machine(OS-ELM).

**Key words:** neural networks; extreme learning machine; on-line training; nonlinear systems

### 1 引言

神经网络是一种由大量神经元互连而成的复杂非线性系统, 在现代神经生物学与认知科学对人脑信息处理能力研究的基础上提出, 是对人脑生理结构的数学模拟形式, 具有较强的自适应性、自学习能力和非线性映射能力, 广泛应用于解决工程领域中的建模与控制问题<sup>[1-2]</sup>.

极端学习机(ELM)<sup>[3]</sup>是一种新颖的单隐层前馈神经网络. 它将传统的神经网络参数训练的迭代调整过程转化为求解线性方程组, 根据 Moore-Penrose 广义逆矩阵理论, 解析求得具有最小范数的最小二乘解作为网络权值, 整个训练过程可一次完成无需迭代. 相比于传统神经网络训练算法需经多次迭代调整才可最终确定网络权值, ELM的训练速度获得了显著提升, 这一优势使得 ELM 在模式识别与回归估计等

问题中获得了成功应用<sup>[4-11]</sup>. 然而, 传统 ELM 训练算法是一种针对固定训练样本集的离线训练算法, 当有独立于训练样本集之外的新训练样本加入时, 需要重新求解其输出权值, 因此不适宜用于 ELM 在线训练.

针对 ELM 在线训练问题, 本文提出一种 ELM 的限定记忆训练模式, 通过不断学习新训练样本来提高 ELM 的动态自适应性, 并逐步删除旧训练样本, 以摆脱其对于 ELM 的影响. 两个典型非线性系统在线状态预测实例表明了该方法的有效性.

### 2 极端学习机

基于训练样本集  $S_k = \{(\mathbf{x}_i, t_i)\}_{i=1}^k$ ,  $\mathbf{x}_i \in \mathbf{R}^n$  为  $n$  维输入,  $t_i \in \mathbf{R}$  为 1 维输出, 包含  $L$  个隐层神经元函数  $f(\cdot)$  的 ELM 回归模型可表示为

$$\sum_{j=1}^L w_j f(\mathbf{a}_j \mathbf{x}_1 + b_j) = t_1,$$

收稿日期: 2011-01-26; 修回日期: 2011-05-31.

基金项目: 国家部委预先研究基金项目(51309060302).

作者简介: 张弦(1982-), 男, 博士生, 从事工业过程故障监控的研究; 王宏力(1965-), 男, 教授, 博士生导师, 从事控制系统的故障检测与诊断等研究.

$$\begin{aligned} \sum_{j=1}^L w_j f(\mathbf{a}_j \mathbf{x}_2 + b_j) &= t_2, \\ &\vdots \\ \sum_{j=1}^L w_j f(\mathbf{a}_j \mathbf{x}_k + b_j) &= t_k. \end{aligned} \quad (1)$$

其中:  $k$  为训练样本的数量,  $\mathbf{a}_j = [a_{j1} \ a_{j2} \ \cdots \ a_{jn}]$  为连接第  $j$  个神经元的输入权值,  $w_j$  为连接第  $j$  个神经元的输出权值,  $b_j$  为第  $j$  个神经元的偏差. 将式 (1) 写成矩阵形式为

$$\mathbf{H}_k \mathbf{W}_k = \mathbf{T}_k. \quad (2)$$

其中:  $\mathbf{H}_k = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \cdots \ \mathbf{h}_k^T]^T$  为神经元矩阵;  $\mathbf{h}_i = [f(\mathbf{a}_1 \mathbf{x}_i + b_1) \ f(\mathbf{a}_2 \mathbf{x}_i + b_2) \ \cdots \ f(\mathbf{a}_L \mathbf{x}_i + b_L)]$  为与样本  $(\mathbf{x}_i, t_i)$  对应的神经元输入向量;  $\mathbf{W}_k = [w_1 \ w_2 \ \cdots \ w_L]^T$  为输出权值向量;  $\mathbf{T}_k = [t_1 \ t_2 \ \cdots \ t_k]^T$  为输出向量. 多数情况下  $k$  远大于  $L$ , 所以求解式 (2) 可得输出权值<sup>[4]</sup>

$$\mathbf{W}_k = (\mathbf{H}_k^T \mathbf{H}_k)^{-1} \mathbf{H}_k^T \mathbf{T}_k. \quad (3)$$

最终得到训练后的 ELM 回归模型为

$$t = \sum_{j=1}^L w_j f(\mathbf{a}_j \mathbf{x} + b_j). \quad (4)$$

其中:  $\mathbf{x}$  为模型输入,  $t$  为模型输出.

为了使 ELM 具备在线训练能力, 文献 [4] 提出了在线贯序极端学习机 (OS-ELM). OS-ELM 使得新训练样本加入训练样本集时, ELM 的输出权值可利用之前已获得的输出权值递推而得, 因此有效提高了 ELM 应用于在线训练时的计算效率<sup>[4]</sup>.

在实际工程实践中, 多数系统的输入与输出之间表现出复杂的非线性特性, 且系统中往往还包含多个对系统输出有显著影响的参数. 一旦系统输入或者参数发生较大变化时, 会直接造成系统输出和基于系统输入、输出或参数的训练样本发生较大变化. 因此在上述条件下, 新、旧训练样本之间会存在较大差异. 然而, OS-ELM 只是不断地将新训练样本加入训练样本集, 而并未对旧训练样本作任何处理. 因此, OS-ELM 这种同时兼顾对于新、旧训练样本拟合能力的回归策略, 必然会导致其对于新训练样本的学习能力有限, 从而难以准确刻画出变化后的系统特性.

### 3 限定记忆极端学习机

为了避免 OS-ELM 中存在的缺陷, 本文提出一种具有固定训练样本数量的限定记忆极端学习机 (FM-ELM). 其训练模式为: 在训练样本集中, 每加入一个新训练样本的同时删除一个距离当前时刻最远的旧训练样本, 以逐步摆脱旧训练样本对于当前回归模型的影响.

假设当前共有  $k+1$  个训练样本  $(\mathbf{x}_i, t_i), (\mathbf{x}_{i+1}, t_{i+1}), \cdots, (\mathbf{x}_{i+k}, t_{i+k})$ , 则基于这  $k+1$  个样本的神经元矩阵为  $\mathbf{H}_i = [\mathbf{h}_i^T \ \mathbf{h}_{i+1}^T \ \cdots \ \mathbf{h}_{i+k}^T]^T$ , 输出向量为  $\mathbf{T}_i = [t_i \ t_{i+1} \ \cdots \ t_{i+k}]^T$ , 输出权值为

$$\mathbf{W}_i = (\mathbf{H}_i^T \mathbf{H}_i)^{-1} \mathbf{H}_i^T \mathbf{T}_i. \quad (5)$$

将新训练样本  $(\mathbf{x}_{i+k+1}, t_{i+k+1})$  加入训练样本集后, 神经元矩阵变化为  $\mathbf{H}_{i+1} = [\mathbf{H}_i^T \ \mathbf{h}_{i+k+1}^T]^T$ , 输出向量变化为  $\mathbf{T}_{i+1} = [\mathbf{T}_i \ t_{i+k+1}]^T$ , 输出权值变化为

$$\mathbf{W}_{i+1} = (\mathbf{H}_{i+1}^T \mathbf{H}_{i+1})^{-1} \mathbf{H}_{i+1}^T \mathbf{T}_{i+1}. \quad (6)$$

由于

$$\mathbf{H}_{i+1}^T \mathbf{H}_{i+1} = \mathbf{H}_i^T \mathbf{H}_i + \mathbf{h}_{i+k+1}^T \mathbf{h}_{i+k+1}, \quad (7)$$

$$\mathbf{H}_{i+1}^T \mathbf{T}_{i+1} = \mathbf{H}_i^T \mathbf{T}_i + \mathbf{h}_{i+k+1}^T t_{i+k+1}, \quad (8)$$

令  $\mathbf{P}_i = (\mathbf{H}_i^T \mathbf{H}_i)^{-1}$ ,  $\mathbf{P}_{i+1} = (\mathbf{H}_{i+1}^T \mathbf{H}_{i+1})^{-1}$ , 将式 (7) 重新表示为

$$\mathbf{P}_{i+1}^{-1} = \mathbf{P}_i^{-1} + \mathbf{h}_{i+k+1}^T \mathbf{h}_{i+k+1}. \quad (9)$$

对式 (9) 两端求逆, 根据 Sherman-Morrison 矩阵求逆引理<sup>[12]</sup>可得  $\mathbf{P}_{i+1}$  的递推表达式为

$$\begin{aligned} \mathbf{P}_{i+1} &= (\mathbf{P}_i^{-1} + \mathbf{h}_{i+k+1}^T \mathbf{h}_{i+k+1})^{-1} = \\ &= \mathbf{P}_i - \frac{\mathbf{P}_i \mathbf{h}_{i+k+1}^T \mathbf{h}_{i+k+1} \mathbf{P}_i}{1 + \mathbf{h}_{i+k+1}^T \mathbf{P}_i \mathbf{h}_{i+k+1}}. \end{aligned} \quad (10)$$

将式 (5), (8) 和 (9) 代入 (6), 可得

$$\begin{aligned} \mathbf{W}_{i+1} &= \mathbf{P}_{i+1} (\mathbf{P}_i^{-1} \mathbf{P}_i \mathbf{H}_i^T \mathbf{T}_i + \mathbf{h}_{i+k+1}^T t_{i+k+1}) = \\ &= \mathbf{P}_{i+1} (\mathbf{P}_i^{-1} \mathbf{W}_i + \mathbf{h}_{i+k+1}^T t_{i+k+1}) = \\ &= \mathbf{W}_i + \mathbf{P}_{i+1} \mathbf{h}_{i+k+1}^T (t_{i+k+1} - \mathbf{h}_{i+k+1} \mathbf{W}_i). \end{aligned} \quad (11)$$

在加入了新训练样本  $(\mathbf{x}_{i+k+1}, t_{i+k+1})$  后, 为了消除旧训练样本对当前回归模型的影响, 对  $(\mathbf{x}_{i+1}, t_{i+1})$  进行剔除. 剔除  $(\mathbf{x}_{i+1}, t_{i+1})$  后的神经元矩阵变化为  $\mathbf{H}_{i+2} = [\mathbf{h}_{i+1}^T \ \mathbf{h}_{i+2}^T \ \cdots \ \mathbf{h}_{i+k+1}^T]^T$ , 输出向量变化为  $\mathbf{T}_{i+2} = [t_{i+1} \ t_{i+2} \ \cdots \ t_{i+k+1}]^T$ , 输出权值变化为

$$\mathbf{W}_{i+2} = (\mathbf{H}_{i+2}^T \mathbf{H}_{i+2})^{-1} \mathbf{H}_{i+2}^T \mathbf{T}_{i+2}. \quad (12)$$

由于

$$\mathbf{H}_{i+2}^T \mathbf{H}_{i+2} = \mathbf{P}_{i+1}^{-1} - \mathbf{h}_i^T \mathbf{h}_i, \quad (13)$$

$$\mathbf{H}_{i+2}^T \mathbf{T}_{i+2} = \mathbf{H}_{i+1}^T \mathbf{T}_{i+1} - \mathbf{h}_i^T t_i, \quad (14)$$

令  $\mathbf{P}_{i+2} = (\mathbf{H}_{i+2}^T \mathbf{H}_{i+2})^{-1}$ , 将式 (13) 重新表示为

$$\mathbf{P}_{i+2}^{-1} = \mathbf{P}_{i+1}^{-1} - \mathbf{h}_i^T \mathbf{h}_i. \quad (15)$$

对式 (15) 两端求逆, 根据 Sherman-Morrison 矩阵求逆引理<sup>[12]</sup>可得  $\mathbf{P}_{i+2}$  的递推表达式为

$$\mathbf{P}_{i+2} = (\mathbf{P}_{i+1}^{-1} - \mathbf{h}_i^T \mathbf{h}_i)^{-1} = \mathbf{P}_{i+1} + \frac{\mathbf{P}_{i+1} \mathbf{h}_i^T \mathbf{h}_i \mathbf{P}_{i+1}}{1 - \mathbf{h}_i^T \mathbf{P}_{i+1} \mathbf{h}_i}. \quad (16)$$

将式 (6), (14) 和 (15) 代入 (12), 可得

$$\mathbf{W}_{i+2} = \mathbf{P}_{i+2} (\mathbf{P}_{i+1}^{-1} \mathbf{P}_{i+1} \mathbf{H}_{i+1}^T \mathbf{T}_{i+1} - \mathbf{h}_i^T t_i) =$$

$$\begin{aligned} P_{i+2}(P_{i+1}^{-1}W_{i+1} - h_i^T t_i) = \\ W_{i+1} + P_{i+2}h_i^T(h_i W_{i+1} - t_i). \end{aligned} \quad (17)$$

因此, 经过增加新样本与删减旧样本后的输出权值  $W_{i+2}$  可利用增减样本之前的输出权值  $W_i$  递推而得, 避免了根据式 (3) 以矩阵求逆的方式重新计算. FM-ELM 的训练过程可描述如下:

Step 1: 令  $i = 1$ , 利用初始的  $k+1$  个训练样本  $(x_i, t_i), (x_{i+1}, t_{i+1}), \dots, (x_{i+k}, t_{i+k})$  训练 FM-ELM, 获得初始  $P_i$  和初始输出权值  $W_i$ .

Step 2: 将新训练样本  $(x_{i+k+1}, t_{i+k+1})$  加入训练样本集, 并按照式 (18) 对  $P_i$  和  $W_i$  进行递推更新, 以获得增加新样本后的  $P_{i+1}$  和  $W_{i+1}$ , 即

$$\begin{cases} P_{i+1} = P_i - \frac{P_i h_{i+k+1}^T h_{i+k+1} P_i}{1 + h_{i+k+1} P_i h_{i+k+1}^T}, \\ W_{i+1} = W_i + P_{i+1} h_{i+k+1}^T (t_{i+k+1} - h_{i+k+1} W_i). \end{cases} \quad (18)$$

Step 3: 将旧训练样本  $(x_i, t_i)$  剔除出训练样本集, 按照式 (19) 对  $P_{i+1}$  和  $W_{i+1}$  进行递推更新, 以获得剔除旧样本后的  $P_{i+2}$  和  $W_{i+2}$ , 即

$$\begin{cases} P_{i+2} = P_{i+1} + \frac{P_{i+1} h_i^T h_i P_{i+1}}{1 - h_i P_{i+1} h_i^T}, \\ W_{i+2} = W_{i+1} + P_{i+2} h_i^T (h_i W_{i+1} - t_i). \end{cases} \quad (19)$$

Step 4: 利用  $W_{i+2}$  建立式 (4) 所示的回归模型.

Step 5: 令  $i = i + 1$  后转至 Step 2.

#### 4 仿真测试

以一个由式 (20) 描述的典型非线性系统为例验证 FM-ELM 的有效性, 并将其测试结果与 OS-ELM 进行比较, 有

$$y(n+1) = \frac{y(n)}{1 + y(n)^2} + u(n)^3. \quad (20)$$

其中:  $u(n)$  为系统输入,  $y(n)$  为系统输出. 以  $x(n) = [y(n) \ u(n)]^T$  为输入,  $t(n) = y(n+1)$  为输出构成训练与测试样本, 利用  $\{(x(n), t(n))\}_{n=1}^{100}$  训练初始的 FM-ELM 和 OS-ELM. 从  $n > 100$  开始, 不断利用新训练样本对 FM-ELM 和 OS-ELM 进行在线训练, 并利用训练后的 FM-ELM 和 OS-ELM 预测未来时刻的系统输出  $y(n+1)$ . 选择 FM-ELM, OS-ELM 的隐层神经元函数为 Sigmoid 函数<sup>[3]</sup>(数量为 20), 分别测试 FM-ELM 和 OS-ELM 在以下 3 种变化模式下的在线预测性能:

1) 变化模式 1. 系统输入  $u(n)$  随  $n$  而变, 变化规律为

$$u(n) = \begin{cases} r[0, 1], & n \leq 100; \\ \sin(2\pi n/50), & 100 < n \leq 300; \\ 1, & 300 + m_i < n \leq 350 + m_i; \\ -1, & 350 + m_i < n \leq 400 + m_i. \end{cases} \quad (21)$$

其中:  $r[0, 1]$  为 0~1 间的随机数;  $i = 1, 2, 3, 4$ ;  $m_1 = 0$ ,  $m_2 = 100$ ,  $m_3 = 200$ ,  $m_4 = 300$ . 系统结构如式 (20) 所示, 保持不变.

2) 变化模式 2. 系统输入  $u(n)$  保持为  $r[0, 1]$ , 系统结构随  $n$  而变. 当  $n \leq 100$  时, 系统结构如式 (20) 所示. 当  $n > 100$  时, 系统结构如下所示:

$$y(n+1) = \begin{cases} \frac{2y(n)}{(2 + y(n)^2)} + u(n)^3, & 100 < n \leq 300; \\ \frac{y(n)}{(1 + 2y(n)^2)} + 2u(n)^3, & 300 < n \leq 700. \end{cases} \quad (22)$$

3) 变化模式 3. 系统输入  $u(n)$  与系统结构均随  $n$  而变,  $u(n)$  根据式 (21) 变化, 系统结构根据式 (22) 变化. 最终基于 FM-ELM 和 OS-ELM 的系统输出预测结果如图 1~图 4 和表 1 所示.

由表 1 可见, 在 3 种变化模式下, FM-ELM 对于系统的预测误差均明显小于 OS-ELM. 由图 1~图 4 可

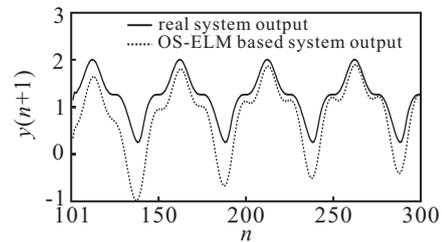


图 1 100 < n ≤ 300 时基于 OS-ELM 的预测结果

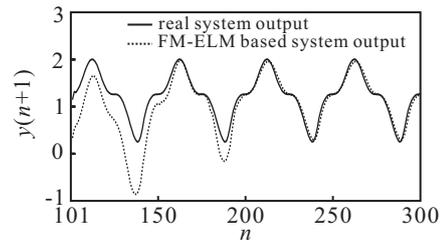


图 2 100 < n ≤ 300 时基于 FM-ELM 的预测结果

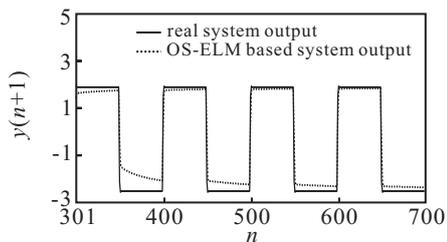


图 3 300 < n ≤ 700 时基于 OS-ELM 的预测结果

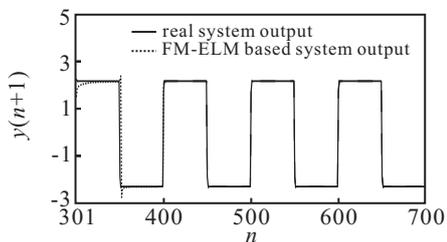


图 4 300 < n ≤ 700 时基于 FM-ELM 的预测结果

表 1 FM-ELM, OS-ELM 对系统的预测均方根误差

预测模型	变化模式 1	变化模式 2	变化模式 3
FM-ELM	0.004 0	0.023 2	0.010 7
OS-ELM	0.005 5	0.039 2	0.015 5

见, 基于 FM-ELM 的预测值比基于 OS-ELM 的预测值更加接近系统的真实输出. 其原因在于: 系统输入或结构从初始训练阶段结束后开始发生变化, 因此训练样本也随之发生变化, 系统在 3 种变化模式下均表现出明显的阶段变化特性, 在各阶段的系统输入、输出与结构均明显不同于其余阶段. 在初始训练阶段结束后的任意时刻, OS-ELM 用于模型训练的样本中, 不仅包含反映当前时刻系统特性的新样本, 还包含许多反映过去时刻系统特性的旧样本. 由于系统特性在当前时刻与过去时刻之间的显著差异, 利用这种模式训练出的 OS-ELM 模型对于当前时刻系统的逼近能力必然有限. FM-ELM 由于具备了抛弃旧样本的能力, 可随系统特性的变化而变化, 使得用于模型训练的样本中, 始终都是反映系统当前特性的新样本, 比 OS-ELM 更适用于具有动态变化特性的系统在线建模与预测.

### 5 工程应用

典型的连续搅拌釜式反应器可由如下所示的非线性微分方程描述:

$$\begin{aligned} \frac{dA}{dt} &= \frac{Q}{V}(A_1 - A) - \psi A e^{(-E/T)}, \\ \frac{dT}{dt} &= \frac{Q}{V}(T_1 - T) - \frac{\Delta\varphi\psi A}{\rho f} e^{(-E/T)} + \frac{U}{\rho f V}(T_2 - T). \end{aligned} \quad (23)$$

其中:  $A$  为反应浓度,  $T$  为反应温度,  $A_1$  为进料浓度,  $T_1$  为进料温度,  $T_2$  为冷却剂温度,  $Q$  为进料流速,  $V$  为反应器体积,  $E$  为活化能,  $\psi$  为预指数因子,  $\Delta\varphi$  为反应热,  $\rho$  为密度,  $f$  为热容量.

设定  $A_1 = 1 \text{ mol/L}$ ,  $T_1 = 400 \text{ K}$ ,  $T_2 = 420 \text{ K}$ ,  $Q = 100 \text{ L/min}$ ,  $V = 100 \text{ L}$ ,  $E = 5000 \text{ K}$ ,  $\psi = e^{10}$ ,  $\Delta\varphi = 200 \text{ kJ/mol}$ ,  $\rho = 1 \text{ kg/L}$ ,  $f = 5 \text{ J/(g} \cdot \text{K)}$ ,  $U = 12 \text{ kJ/(min} \cdot \text{K)}$ . 以  $t = 0.2 \text{ min}$  为间隔对式 (23) 进行离散采样, 在任意时刻  $n$ , 以

$$\mathbf{x}_A(n) = [A(n) \ T(n) \ \leftarrow A_1(n) \ T_1(n)]^T$$

为输入,  $t_A(n) = A(n+1)$  为输出, 构成反应浓度  $A$  的训练与测试样本. 利用  $\{\mathbf{x}_A(n), t_A(n)\}_{n=1}^{100}$  训练初始的 FM-ELM 和 OS-ELM, 并在初始训练阶段结束后, 逐次利用新训练样本对 FM-ELM 和 OS-ELM 进行在线训练, 并利用训练后的 FM-ELM 和 OS-ELM 预测未来时刻的反应浓度  $A$ .

在以下 4 种模式下测试具有 30 个隐层神经元函数的 FM-ELM 和 OS-ELM 对于反应浓度的预测精度:

1) 突变模式 1. 进料浓度  $A_1$  在初始训练阶段结束后, 从  $1 \text{ mol/L}$  跳变为  $2 \text{ mol/L}$ , 其余参数保持不变.

2) 突变模式 2. 进料温度  $T_1$  在初始训练阶段结束后, 从  $400 \text{ K}$  跳变为  $300 \text{ K}$ , 其余参数保持不变.

3) 缓变模式 1. 进料浓度  $A_1$  在初始训练阶段结束后, 从  $1 \text{ mol/L}$  开始以每分钟  $0.005 \text{ mol/L}$  的速度缓慢增加, 其余参数保持不变.

4) 缓变模式 2. 进料温度  $T_1$  在初始训练阶段结束后, 从  $400 \text{ K}$  开始以每分钟  $0.05 \text{ K}$  的速度缓慢减少, 其余参数保持不变. 最终, 基于 FM-ELM 和 OS-ELM 的反应浓度  $A$  的预测结果如图 5~图 8 和表 2 所示.

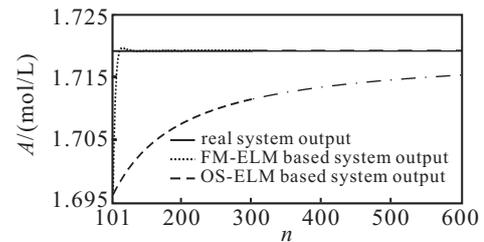


图 5 突变模式 1 的预测结果

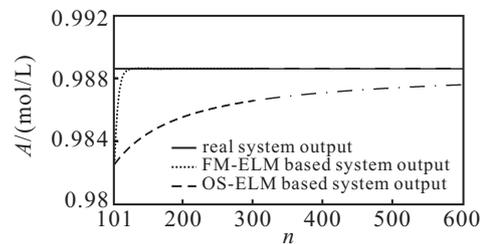


图 6 突变模式 2 的预测结果

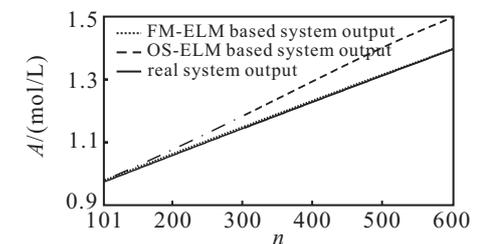


图 7 缓变模式 1 的预测结果

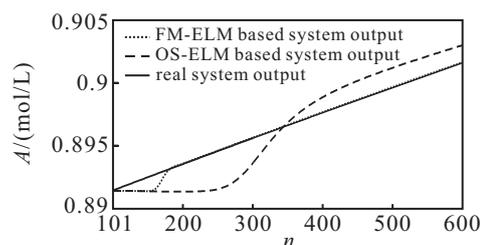


图 8 缓变模式 2 的预测结果

表 2 FM-ELM, OS-ELM 对反应浓度的预测均方根误差

预测模型	突变模式 1	突变模式 2	缓变模式 1	缓变模式 2
FM-ELM	6.4723e-5	2.0113e-4	4.9864e-5	1.4023e-5
OS-ELM	4.1704e-4	3.6815e-4	5.6825e-4	7.8708e-5

由表 2 可见, FM-ELM 在 4 种模式下对反应浓度的预测均方根误差均小于 OS-ELM, 再次验证了 FM-ELM 对旧训练样本的抛弃能力. 在系统特性发生变化时, 可有效摆脱 FM-ELM 对系统过去时刻特性的记忆力, 因此, 对于系统当前时刻特性的模拟能力更为精准. 由图 5 和图 6 可见, 当系统输入发生跳变时, 基于 FM-ELM 的预测值迅速作出反应, 在极短的时间内即可调节到与系统真实输出几乎完全重合的状态. 基于 OS-ELM 的预测值由于受到过去时刻系统输入的影响, 在很长一段时间内仍与系统真实输出之间存在较大差距. 由图 7 和图 8 可见, 当系统缓慢变化时, 基于 FM-ELM 的预测值的调节速度依然很快, 基于 OS-ELM 的预测值则与系统真实输出之间存在较大差距. 因此相比于 OS-ELM, 基于 FM-ELM 的预测结果更加接近系统真实特性, 且反应速度更快.

## 6 结 论

不同于传统神经网络训练算法中的网络权值迭代调整策略, ELM 以求解线性方程组的新颖方式完成网络训练过程, 极大地提升了神经网络的训练速度, 因此具有较高的学术研究与工程应用价值. 传统 ELM 的训练算法主要适用于离线训练, 不适于在线训练. FM-ELM 采用递推模式求解输出权值, 使 ELM 具备了在线训练能力. 典型非线性系统在线状态预测实例表明, 相比 OS-ELM, FM-ELM 中对旧训练样本的抛弃能力使其更适用于具有动态变化特性的非线性系统在线建模与状态预测.

## 参考文献(References)

- [1] Song Q S, Feng Z R. Effects of connectivity structure of complex echo state network on its prediction performance for nonlinear time series[J]. *Neurocomputing*, 2010, 73(10-12): 2177-2185.
- [2] Muhammad A F, Saeed Z. Chaotic time series prediction with residual analysis method using hybrid Elman-NARX neural networks[J]. *Neurocomputing*, 2010, 73(13-15): 2540-2553.
- [3] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: Theory and applications[J]. *Neurocomputing*, 2006, 70(1-3): 489-501.
- [4] Liang N Y, Huang G B, Saratchandran N, et al. A fast and accurate on-line sequential learning algorithm for feedforward networks[J]. *IEEE Trans on Neural Networks*, 2006, 17(6): 1411-1423.
- [5] Miche Y, Soriamaa A, Bas P, et al. OP-ELM: Optimally pruned extreme learning machine[J]. *IEEE Trans on Neural Networks*, 2010, 21(1): 158-162.
- [6] Liu N, Wang H. Ensemble based extreme learning machine[J]. *IEEE Signal Processing Letters*, 2010, 17(8): 754-757.
- [7] Lan Y, Soh C Y, Huang G B. Constructive hidden nodes selection of extreme learning machine for regression[J]. *Neurocomputing*, 2010, 73(16/17/18): 3191-3199.
- [8] Lan Y, Soh C Y, Huang G B. Two-stage extreme learning machine for regression[J]. *Neurocomputing*, 2010, 73(16/17/18): 3028-3038.
- [9] Malathi V, Marimuthu N S, Baskar S. Intelligent approaches using support vector machine and extreme learning machine for transmission line protection[J]. *Neurocomputing*, 2010, 73(10-12): 2160-2167.
- [10] 张弦, 王宏力. 局域极端学习机及其在状态在线监测中的应用[J]. *上海交通大学学报*, 2011, 45(2): 236-240. (Zhang X, Wang H L. Local extreme learning machine and its application to condition on-line monitoring[J]. *J of Shanghai Jiaotong University*, 2011, 45(2): 236-240.)
- [11] 韩敏, 王新迎. 基于信赖域 Newton 算法的 ELM 网络[J]. *控制与决策*, 2011, 26(5): 757-760. (Han M, Wang X Y. ELM based on trust region Newton method[J]. *Control and Decision*, 2011, 26(5): 757-760.)
- [12] 张贤达. 矩阵分析与应用[M]. 北京: 清华大学出版社, 2005. (Zhang X D. *Matrix analysis and applications*[M]. Beijing: Tsinghua University Press, 2005.)
- [10] Stutters L, Liu H, Tiltman C, et al. Navigation technologies for autonomous underwater vehicles[J]. *IEEE Trans on Systems, Man and Cybernetics*, 2008, 38(3): 581-589.
- [11] Kohonen T. Analysis of a simple self-organizing process[J]. *Biological Cybernetics*, 1982, 44(2): 145-140.
- [12] Kohonen T. Self-organization of a massive document collection[J]. *IEEE Trans on Neural Networks*, 2000, 11(3): 574-585.
- [13] Zhu A, Yang S X. Self-organizing behavior of a multi-robot system by a neural network approach[C]. *Proc IEEE Int Conf Intelligent Robotics and Systems and Signal Processing*. Nevada, 2003: 1204-1209.
- [14] Yang S X, Meng O H M. Real-time collision-free motion planning of mobile robots using neural dynamics based approaches[J]. *IEEE Trans on Neural Network*, 2003, 14(5): 1541-1552.

(上接第 1205 页)