文章编号：1001-0920(2009)08-1243-04

改进块式递推偏最小二乘建模方法及应用

常玉清[a,b]，袁 勇[a]，王福利[a,b]
（东北大学 a. 信息科学与工程学院，b. 教育部流程工业综合自动化重点实验室，沈阳 110004）

摘 要：针对传统偏最小二乘(PLS)模型的在线更新问题，提出了带有自适应遗忘因子的块式递推 PLS 建模方法. 通过 Hotelling-T^2 和 Q 统计量确定遗忘因子的大小，并且进行模型递推更新，确保模型跟踪过程特性的变化. 将所提出的方法应用于管坯斜轧穿孔能耗过程，表现出较强的模型在线更新能力. 测试结果表明，带有自适应遗忘因子的块式递推 PLS 方法的性能优于传统的迭代偏最小二乘方法的性能.

关键词：偏最小二乘；自适应遗忘因子；T^2 统计量；Q 统计量；模型更新
中图分类号：TP273 **文献标识码**：A

Improved block-wise recursive partial least square modeling method and its application

CHANG Yu-qing[a,b], *YUAN Yong*[a], *WANG Fu-li*[a,b]
(a. College of Information Science and Engineering, b. Key Laboratory of Integrated Automation of Process Industry, Ministry of Education, Northeastern University, Shenyang 110004, China. Correspondent: CHANG Yu-qing, E-mail: changyuqing@mail.neu.edu.cn)

Abstract: An improved block-wise recursive partial least square(PLS) modeling method with adaptive forgetting factor is proposed to overcome the problem of the traditional PLS model which can not be updated effectively. The model is updated recursively with a forgetting factor calculated by T^2 statistic and Q statistic, which are in accordance with the process changes. The proposed method is applied to the online update of piercing energy consumption soft sensing model. The ability for updating the model online is presented. The testing results show that the performance of the improved RPLS with adaptive is better than that of the traditionl RPLS model.

Key words: Partial least square; Adaptive forgetting factor; T^2 statistic; Q statistic; Model updating

1 引 言

作为一种基于数据回归的软测量建模方法，偏最小二乘(PLS)得到了广泛的重视[1]. 然而，当过程特性或操作条件变化时，传统的 PLS 不能及时进行模型在线更新. 为此，文献[2]提出一种块式递推偏最小二乘法(PLS)，根据新数据和原有的 PLS 模型参数调整模型，大大提高了模型的在线训练速度，但该方法确定的遗忘因子是固定的，无法根据生产过程的变化快速更新数据的可信度.

本文在分析块式递推 PLS 的基础上，基于 Hotelling-T^2 和 Q 统计量的过程监测的思想，提出了带有自适应遗忘因子的块式递推 PLS 建模方法. 该方法通过 Hotelling-T^2 和 Q 统计量监测生产状况的变化，并用其更新遗忘因子. 将所提出的方法应用于管坯斜轧穿孔能耗软测量模型的在线更新，取得了良好的效果.

2 块式递推 PLS 算法

2.1 PLS 算法

给定两个数据阵 $\boldsymbol{X}\in\boldsymbol{R}^{n\times m}$ 和 $\boldsymbol{Y}\in\boldsymbol{R}^{n\times p}$. 其中：$n$ 为样本个数，m 和 p 分别为自变量和因变量的维数. 则 $\boldsymbol{Y}$ 与 $\boldsymbol{X}$ 之间的偏最小二乘关系可表示为

$$\boldsymbol{Y}=\boldsymbol{X}\boldsymbol{C}^{\mathrm{PLS}}+\boldsymbol{V}. \tag{1}$$

其中：$\boldsymbol{C}^{\mathrm{PLS}}$ 为偏最小二乘回归系数矩阵，$\boldsymbol{V}$ 为残差矩阵.

在 PLS 回归建模时，对 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 进行如下分解[3]：

$$\boldsymbol{X}=\sum_{a=1}^{A}\boldsymbol{t}_a\boldsymbol{p}_a^{\mathrm{T}}+\boldsymbol{E}=\boldsymbol{T}\boldsymbol{P}^{\mathrm{T}}+\boldsymbol{E}, \tag{2a}$$

$$\boldsymbol{Y}=\sum_{a=1}^{A}\boldsymbol{u}_a\boldsymbol{q}_a^{\mathrm{T}}+\boldsymbol{F}=\boldsymbol{U}\boldsymbol{Q}^{\mathrm{T}}+\boldsymbol{F}. \tag{2b}$$

收稿日期：2008-07-15；**修回日期**：2008-10-28.
基金项目：国家自然科学基金项目(60774068).
作者简介：常玉清(1973—)，女，沈阳人，副教授，博士，从事工业过程建模与控制的研究；王福利(1957—)，男，辽宁辽阳人，教授，博士生导师，从事工业过程建模与控制、故障诊断等研究.

其中：$\boldsymbol{U}=\boldsymbol{TB}$，$\boldsymbol{T}=\boldsymbol{XW}^*$；$A$ 是保留的潜变量的个数；$\boldsymbol{T}=[\boldsymbol{t}_1,\boldsymbol{t}_2,\cdots,\boldsymbol{t}_A]$，$\boldsymbol{U}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_A]$，$\boldsymbol{t}_a(n\times 1)$ 和 $\boldsymbol{u}_a(n\times 1)$ 分别为 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的第 a 个得分向量；$\boldsymbol{P}=[\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_A]$，$\boldsymbol{Q}=[\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_A]$，$\boldsymbol{p}_a(m\times 1)$ 和 $\boldsymbol{q}_a(p\times 1)$ 分别为 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 的负载向量；$\boldsymbol{E}$ 和 $\boldsymbol{F}$ 为拟合残差矩阵；$\boldsymbol{W}^*=[\boldsymbol{w}_1^*,\boldsymbol{w}_2^*,\cdots,\boldsymbol{w}_A^*]$，$\boldsymbol{w}_a^*=\prod_{h=1}^{a-1}(\boldsymbol{I}_m-\boldsymbol{w}_h\boldsymbol{p}_h^{\mathrm{T}})\boldsymbol{w}_a$，$\boldsymbol{w}_a$ 是得分向量 $\boldsymbol{t}_a$ 的权值向量.

令 $\boldsymbol{W}=[\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_A]$；$\boldsymbol{B}=\mathrm{diag}\{\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_A\}$，$\boldsymbol{b}_a=\boldsymbol{t}_a^{\mathrm{T}}\boldsymbol{u}_a/(\boldsymbol{t}_a^{\mathrm{T}}\boldsymbol{t}_a)$ 是 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 空间潜变量 $\boldsymbol{t}$ 和 $\boldsymbol{u}$ 的内部回归系数. 则 $\{\boldsymbol{X},\boldsymbol{Y}\}$ 的 PLS 回归模型[4] 可表示为

$$\{\boldsymbol{X},\boldsymbol{Y}\}\xrightarrow{\mathrm{PLS}}\{\boldsymbol{T},\boldsymbol{W},\boldsymbol{P},\boldsymbol{B},\boldsymbol{Q}\}. \tag{3}$$

PLS 的回归系数为

$$\boldsymbol{C}^{\mathrm{PLS}}=(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{+}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}, \tag{4}$$

其中算子 $(\cdot)^+$ 为 PLS 意义上的广义逆矩阵.

文献[5] 给出了回归系数的另一种表达形式，即

$$\boldsymbol{C}^{\mathrm{PLS}}=\boldsymbol{W}^*\boldsymbol{BQ}^{\mathrm{T}}. \tag{5}$$

2.2 块式递推 PLS 算法

为了实现递推求解，文献[6] 对 PLS 回归算法进行改进. 对自变量矩阵 $\boldsymbol{X}$ 的得分矩阵 $\boldsymbol{T}$ 进行归一化，有

$$\boldsymbol{T}^{\mathrm{T}}\boldsymbol{T}=\boldsymbol{I}. \tag{6}$$

块式递推 PLS 算法只有在新的可用数据积累到 w 时（w 称为数据块宽度），才对模型进行更新. 与一般的递推 PLS 相比，该算法避免了不必要的重复计算，并且增强了模型的鲁棒性.

设新积累的数据块为 $\{\boldsymbol{X}_l,\boldsymbol{Y}_l\}$，所对应的 PLS 模型为 $\{\boldsymbol{X}_l,\boldsymbol{Y}_l\}\xrightarrow{\mathrm{PLS}}\{\boldsymbol{T}_l,\boldsymbol{W}_l,\boldsymbol{P}_l,\boldsymbol{B}_l,\boldsymbol{Q}_l\}$；旧数据块为 $\{\boldsymbol{X},\boldsymbol{Y}\}$，所对应的 PLS 模型为 $\{\boldsymbol{X},\boldsymbol{Y}\}\xrightarrow{\mathrm{PLS}}\{\boldsymbol{T},\boldsymbol{W},\boldsymbol{P},\boldsymbol{B},\boldsymbol{Q}\}$. 则有

$$\boldsymbol{C}_{\mathrm{new}}^{\mathrm{PLS}}=\left(\begin{bmatrix}\boldsymbol{X}\\ \boldsymbol{X}_l\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{X}\\ \boldsymbol{X}_l\end{bmatrix}\right)^{+}\begin{bmatrix}\boldsymbol{X}\\ \boldsymbol{X}_l\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{Y}\\ \boldsymbol{Y}_l\end{bmatrix}=$$
$$\left(\begin{bmatrix}\boldsymbol{P}^{\mathrm{T}}\\ \boldsymbol{P}_l^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{P}^{\mathrm{T}}\\ \boldsymbol{P}_l^{\mathrm{T}}\end{bmatrix}\right)^{+}\begin{bmatrix}\boldsymbol{P}^{\mathrm{T}}\\ \boldsymbol{P}_l^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{BQ}^{\mathrm{T}}\\ \boldsymbol{B}_l\boldsymbol{Q}_l^{\mathrm{T}}\end{bmatrix}. \tag{7}$$

可以看出，对全部数据进行回归等价于分别对新老数据块的 PLS 模型参数进行回归.

为避免因采集数据不断增加而产生数据饱和的现象，以及建模过程中新老数据的可信度相同的问题，Qin[2] 提出了两种自适应算法：移动窗口法和遗忘因子法. 移动窗口法根据设定的数据窗口长度限制进入算法的信息量，从而防止数据饱和；遗忘因子法是对新老数据分配不同的可信度，降低老数据提供的信息量，增加新数据提供的信息量.

汪晓勇等将移动窗口法与遗忘因子法相结合[7]，提出了基于块式递推 PLS 的限定记忆法，其原理如下式所示：

$$\begin{array}{ccc}\text{序号} & \text{数据块} & \text{遗忘因子}\\ \left[\begin{array}{ccc} l & \{\boldsymbol{X}_l,\boldsymbol{Y}_l\} & 1\\ l-1 & \{\boldsymbol{X}_{l-1},\boldsymbol{Y}_{l-1}\} & \lambda\\ \vdots & \vdots & \vdots\\ 2 & \{\boldsymbol{X}_2,\boldsymbol{Y}_2\} & \lambda^{l-2}\\ 1 & \{\boldsymbol{X}_1,\boldsymbol{Y}_1\} & \lambda^{l-1}\end{array}\right] & \begin{array}{c}\text{新}\\ \Big\downarrow\\ \text{旧}\end{array}\end{array}. \tag{8}$$

该算法应用如下数据队列计算回归系数 $\boldsymbol{C}^{\mathrm{PLS}}$：

$$\begin{bmatrix}\boldsymbol{P}_1^{\mathrm{T}}\\ \lambda\boldsymbol{P}_{l-1}^{\mathrm{T}}\\ \vdots\\ \lambda^{l-1}\boldsymbol{P}_1^{\mathrm{T}}\end{bmatrix},\begin{bmatrix}\boldsymbol{B}_l\boldsymbol{Q}_l^{\mathrm{T}}\\ \lambda\boldsymbol{B}_{l-1}\boldsymbol{Q}_{l-1}^{\mathrm{T}}\\ \vdots\\ \lambda^{l-1}\boldsymbol{B}_1\boldsymbol{Q}_1^{\mathrm{T}}\end{bmatrix}. \tag{9}$$

其中：l 为数据块队列长度，λ 为遗忘因子.

3 自适应遗忘因子块式递推 PLS 建模方法

块式递推 PLS 算法中有 3 个重要参数：数据块宽度 w，数据块队列长度 l 和遗忘因子 λ. 遗忘因子决定了队列中各个数据块对模型所起的作用，对预测误差和模型跟随速度将产生重要影响. 遗忘因子太大，模型更新的速度慢；遗忘因子太小，模型更新的速度快，但会产生较大的误差[8].

传统的 RPLS 算法一般将遗忘因子固定在某一特定值，或采用特定函数为不同的数据块分配不同的可信度[7]. 这些方法所确定的子数据块的可信度都是固定的，缺乏针对性和灵活性，无法根据生产过程的变化快速更新不同数据块的可信度. 在选取遗忘因子 λ 时，应根据批次间过程的变化程度来确定 λ 的大小. 过程变化不大，λ 可取 0.9 ～ 1；过程变化较大，λ 可取较小的值，使模型尽快适应新的生产过程.

Hotelling-T^2 统计量和 Q 统计量是过程监控中常用的两个多元统计量. Hotelling-T^2 统计量用来表征过程变量关系的变化情况，Q 统计量用来衡量过程测量信息偏离模型的程度. 当系统处于某一稳定工作状态，即系统的稳态工作点不变时，如果 T^2 和 Q 统计量已超过控制限，则表明此时的生产过程出现异常；如果生产过程一直处于正常状态，而 T^2 和 Q 统计量也已超过各自的控制限，则表明系统的稳态工作点发生了变化，即系统从一个正常工作状态转换到另一个正常工作状态.

T^2 和 Q 统计量的变化程度从两个不同的侧面表明了过程特性的变化程度，因此可用它们的相关函数来确定遗忘因子，保证具有新过程特性的数据信息占据重要的位置，进而得到带有自适应遗忘因子的块式递推 PLS 建模方法.

新测量数据 $\boldsymbol{x}_{\text{new}}$ 的得分和残差量由下式获得：

$$\begin{cases}\boldsymbol{t}=\boldsymbol{x}_{\text{new}}\boldsymbol{P},\\ \hat{\boldsymbol{x}}_{\text{new}}=\boldsymbol{t}\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{x}_{\text{new}}\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}},\\ \boldsymbol{e}=\boldsymbol{x}_{\text{new}}-\hat{\boldsymbol{x}}_{\text{new}}=\boldsymbol{x}_{\text{new}}(\boldsymbol{I}-\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}).\end{cases}\tag{10}$$

Hotelling-T^2 统计量定义为

$$T^2=\boldsymbol{t}^{\mathrm{T}}\boldsymbol{S}_A^{-1}\boldsymbol{t}.\tag{11}$$

其中：$\boldsymbol{t}$ 为由式(10)计算的得分，对角矩阵 $\boldsymbol{S}_A=\text{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_A\}$ 由 $\boldsymbol{X}$ 的协方差矩阵 $\boldsymbol{\Sigma}=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 的前 A 个特征值所组成.

Q 统计量也称预测误差平方和指标(SPE)，定义为

$$\text{SPE}=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}.\tag{12}$$

为了客观地判断系统稳态工作点是否发生变化，需要确定其变化前 T^2 和 Q 统计量的统计控制限. T^2 统计量的控制限利用 F 分布按下式计算[4]：

$$T_\alpha^2=\frac{A(n-1)}{n-A}F_{A,n-A,\alpha}.\tag{13}$$

其中：n 为建模样本的个数，A 为模型中保留的特征向量的个数，α 为显著性水平. $n-A$ 条件下的 F 分布临界值可从统计表中查到.

Q 统计量的控制限由下式计算[4]：

$$Q_\alpha=\theta_1\left(\frac{C_\alpha\sqrt{2\theta_2h_0^2}}{\theta_1}+1+\frac{\theta_2h_0(h_0-1)}{\theta_1^2}\right)^{1/h_0}.\tag{14}$$

其中：$\theta_i=\sum_{j=A+1}^{m}\lambda_j^i(i=1,2,3)$，$h_0=1-\frac{2\theta_1\theta_3}{3\theta_2^2}$，$C_\alpha$ 是正态分布在显著性水平 α 下的临界值.

随着新数据的增加，T^2 和 Q 统计量的控制限也应作相应的调整. 本文假设系统一直处于正常工作状态，即系统不出现故障，由测量仪表所测得的过程变量数据都是真实有效的.

在进行块式递推 PLS 模型更新时，当得到一个新的可用数据块 $\{\boldsymbol{X}_l,\boldsymbol{Y}_l\}$ 后，按下式计算该数据块的平均统计量 T_l^2 和 Q_l：

$$T_l^2=\frac{1}{w}\sum_{i=1}^{w}T_i^2,Q_l=\frac{1}{w}\sum_{i=1}^{w}Q_i.\tag{15}$$

当过程处于正常操作工况时，如果该数据块的平均统计量 T_1^2 和 Q_1 超限，则表明发生了较为明显的操作工况变化. 这时需要进行模型更新，数据块遗忘因子 λ 由统计量的变化情况决定，即

$$\lambda=\begin{cases}\lambda_0\min\left(\frac{T_\alpha^2}{T_1^2},\frac{Q_\alpha}{Q_1}\right),T_1^2>T_\alpha^2,Q_1>Q_\alpha;\\ \lambda_0,\text{otherwise}.\end{cases}\tag{16}$$

改进的块式递推 PLS 算法建模步骤如下：

1) 采集正常操作工况下的过程数据并进行有效性检查，由合格数据组成原始数据块 $\{\boldsymbol{X},\boldsymbol{Y}\}$，并对数据进行标准化处理.

2) 使用改进的 PLS 算法计算模型参数 $\{\boldsymbol{X},\boldsymbol{Y}\}\xrightarrow{\text{PLS}}\{\boldsymbol{T},\boldsymbol{W},\boldsymbol{P},\boldsymbol{B},\boldsymbol{Q}\}$ 和回归系数 $\boldsymbol{C}^{\text{PLS}}$.

3) 计算 T^2 和 Q 统计量的控制限 T_α^2 和 Q_α.

4) 当得到一个新的过程测量值 $\boldsymbol{x}_{\text{new}}$ 时，计算模型输出 $\hat{\boldsymbol{y}}=\boldsymbol{x}_{\text{new}}\boldsymbol{C}^{\text{PLS}}$，并计算 $\boldsymbol{x}_{\text{new}}$ 的 T^2 和 Q 统计量.

5) 用宽度为 w 的新数据块 $\{\boldsymbol{X}_l,\boldsymbol{Y}_l\}$(其中包含当前时刻开始的 w 个历史数据信息)计算平均统计量，并用PLS算法求取新数据块的模型参数 $\{\boldsymbol{X}_l,\boldsymbol{Y}_l\}\xrightarrow{\text{PLS}}\{\boldsymbol{T}_l,\boldsymbol{W}_l,\boldsymbol{P}_l,\boldsymbol{B}_l,\boldsymbol{Q}_l\}$.

6) 将新的模型参数矩阵 $[\boldsymbol{P}_l^{\mathrm{T}},\boldsymbol{B}_l\boldsymbol{Q}_l^{\mathrm{T}}]$ 加入数据块队列. 如果队列已满，则剔除队列中最老的数据块.

7) 由式(16)计算遗忘因子 λ，更新队列中各数据块的可信度，并重新计算统计量的控制限 T_α^2 和 Q_α.

8) 利用式(5)重新计算回归系数 $\boldsymbol{C}^{\text{PLS}}$，并返回 4).

4　应用仿真研究

将本文提出的方法应用于某钢厂 SWW 斜轧穿孔机穿孔过程，实现穿孔能耗的软测量建模及在线更新. 根据现场工艺选择上下轧辊转速、上下轧辊倾角值、左右导盘位置等 16 个过程变量作为输入变量，管坯的穿孔能耗为输出变量，建立基于 PLS 的穿孔能耗软测量模型. 取SWW斜轧穿孔过程200根钢管的穿孔数据，前 120 个数据作为训练样本建立初始模型，剩余的 80 个数据作为测试样本.

假设系统在第 20 个数据生产另一批具有不同特性的产品，并发生了操作点变化. 在递推建模过程中，数据块宽度 $w=10$，数据块队列长度 $l=5$，$\lambda_0=0.9$，$\alpha=0.95$.

模型更新过程中测试样本的 T^2 和 Q 统计量的变化情况如图 1 和图 2 所示. 可以看出，当操作点发生变化时，T^2 和 Q 统计量均大大超出了各自的控制限；随着模型的不断更新，T^2 和 Q 统计量又逐渐回到各自的控制限以下. 这种情况表明，通过 T^2 和 Q 统计量能准确地检测出生产状况的变化.

改进的自适应遗忘因子 RPLS 模型与传统的 PLS 模型的预测效果对比如图 3 所示. 可以看出，当生产过程特性没有发生较大变化时，两种模型的预测性能较为接近；当生产过程特性发生变化时，固定模型的预测值出现了很大误差，而改进的 RPLS 模型却表现出很强的过程特性变化适应能力，经过 12 个批次后便可达到良好的预测效果.

具有固定遗忘因子和自适应遗忘因子的 RPLS

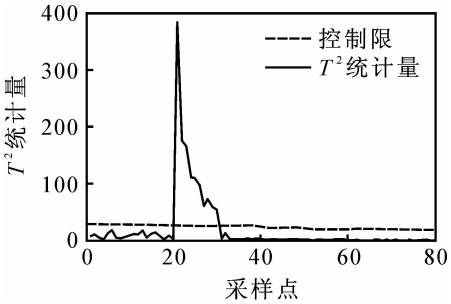

图 1 T^2 统计量变化曲线

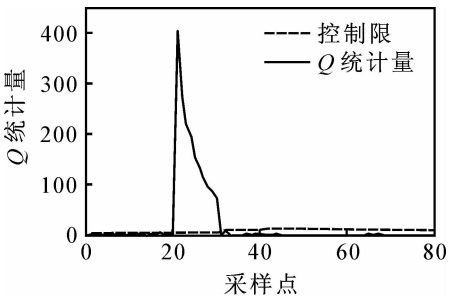

图 2 Q 统计量变化曲线

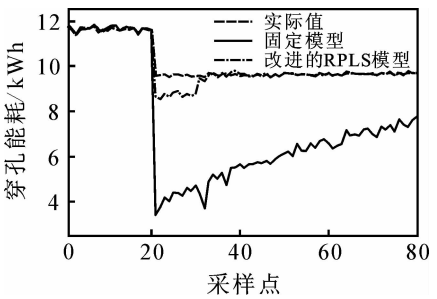

图 3 两种模型的预测效果仿真对比

模型更新性能如图 4 所示. 与传统固定遗忘因子的RPLS模型相比,自适应遗忘因子的RPLS模型具有更强的模型更新能力.

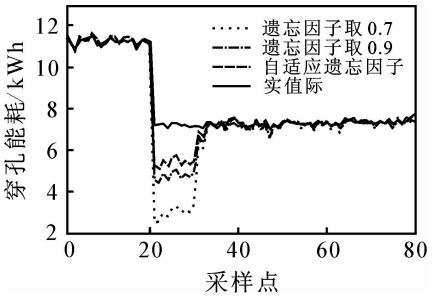

图 4 不同遗忘因子的预测值与实际值比较

取不同的遗忘因子,模型的均方误差(RMSE)如表 1 所示. RMSE 按下式计算:

$$\mathrm{RMSE}=\sqrt{\frac{1}{p}\sum_{i=1}^{p}(\hat{y}_i-y_i)^2}. \tag{17}$$

其中:$\hat{y}_i$ 和 y_i 分别为第 i 个样本的模型预测值和实际值,p 为样本个数.

表 1 不同遗忘因子的模型预测精度比较

	$\lambda=0.7$	$\lambda=0.9$	自适应遗忘因子
RMSE	1.2366	0.7234	0.5278

5 结 论

本文基于 T^2 和 Q 统计量监测生产过程稳态操作点的变化,将其与块式递推 PLS 建模方法相结合,提出了带有自适应遗忘因子的块式 RPLS 建模方法. 该算法根据 T^2 和 Q 统计量的变化程度自适应更新遗忘因子,使得模型对过程特性变化具有更强的自适应能力. 测试结果表明,自适应遗忘因子的RPLS 建模方法不仅保留了传统递推 PLS 建模方法结构简单、建模速度块、抗噪声的优势,而且具有更强的模型跟踪性能和较好的预测效果.

参考文献(References)

[1] Liang J, Qian J X. Multivariate statistical process monitioring and control: Recent developments and applications to chamical industry [J]. Chinese J of Chemical Engineering, 2003, 11(2): 191-203.

[2] Qin S J. Recursive PLS algorithms for adaptive data modeling [J]. Computers & Chemical Engineering, 1998, 22(45): 503-514.

[3] 张杰, 阳宪惠. 多变量统计过程控制[M]. 北京:化学工业出版社, 2000.
(Zhang J, Yang X H. Multi-variable statistic process control[M]. Beijing: Chemical Industry Press, 2000.)

[4] 王惠文. 偏最小二乘回归方法及其应用[M]. 北京: 国防工业出版社, 1999.
(Wang H W. Partial least squares regression method and its application [M]. Beijing: National Defence Industrial Press, 1999.)

[5] Hoskuldsson A. PLS regression methods [J]. J of Chemometrics, 1988, 2(3): 211-228.

[6] Hellend K, Berntsen H E, Borgen O S, et al. Recursive algorithm for partial least squares regression [J]. Chemometrics Intelligence Laboratory System, 1992, 14(1): 129-137.

[7] 汪小勇, 梁军, 刘育明. 基于递推 PLS 的自适应软测量模型及其应用[J]. 浙江大学学报, 2005, 39 (5): 676-680.
(Wang X Y, Liang J, Liu Y M. Recursive PLS based adaptive soft-sensor model and its application[J]. J of Zhejiang University, 2005, 39(5): 676-680.)

[8] 成忠. PLSR 用于化学化工建模的几个关键问题的研究[D]. 杭州: 浙江大学,2005.
(Cheng Z. Research on some key problems for chemical modeling[D]. Hangzhou: Zhejiang University, 2005.)