

DOI: 10.3724/SP.J.1219.2012.00123

基于振动频谱的磨机负荷在线软测量建模

汤 健¹, 赵立杰^{1,2}, 柴天佑^{1,3}, 岳 恒³

(1. 东北大学流程工业综合自动化教育部重点实验室, 辽宁 沈阳 110004; 2. 沈阳化工大学信息工程学院, 辽宁 沈阳 110142;
3. 东北大学自动化研究中心, 辽宁 沈阳 110004)

摘 要: 针对磨机负荷 (ML) 软测量模型难以适应磨矿过程的时变特性, 模型需要依据工况实时在线更新的问题, 基于磨机筒体振动频谱, 通过递归主元分析 (RPCA) 和在线最小二乘支持向量回归机 (LSSVR) 的集成, 提出了 ML 参数 (料球比、矿浆浓度、充填率) 在线软测量方法. 首先, 针对训练样本, 采用主元分析 (PCA) 分别提取振动频谱在低、中、高频段的谱主元; 然后以串行组合后的谱主元为输入, 采用 LSSVR 方法构造 ML 参数离线软测量模型; 最后, 采用旧模型完成预测后, 应用 RPCA 及在线 LSSVR 算法分别递归更新模型的输入和模型的回归参数, 从而实现了 ML 软测量模型的在线更新. 实验结果表明, 该软测量方法与其它常规方法相比具有较高的精度和更好的预测性能.

关键词: 在线软测量; 递归主元分析; 最小二乘支持向量回归机; 磨机负荷; 振动频谱

中图分类号: TP29

文献标识码: A

文章编号: 1002-0411(2012)-01-0123-06

On-line Soft-sensing Modelling of Mill Load Based on Vibration Spectrum

TANG Jian¹, ZHAO Lijie^{1,2}, CHAI Tianyou^{1,3}, YUE Heng³

(1. Key Laboratory of Integrated Automation of Process Industry, Ministry of Education, Northeastern University, Shenyang 110004, China;
2. College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, China;
3. Research Center of Automation, Northeastern University, Shenyang 110004, China)

Abstract: The soft-sensing model for mill load (ML) is difficult to adapt to the time-varying characters of the mineral process, and it needs to be updated online in real-time according to the changes of condition. Aiming at these problems, based on the vibration spectrum of the mill shell, an on-line soft-sensing approach is proposed to measure the ML parameters, such as material to ball volume ratio (MBVR), pulp density (PD) and charge volume ratio (CVR) inside the mill. The method is realized by the integration of recursive principal component analysis (RPCA) and on-line least square support vector regression (LSSVR). At first, for the training samples, spectral principal components (PCs) at low, medium and high frequency bands of the shell vibration spectrum are extracted through PCA. Then, the spectral PCs of serial combination with different bands are used to construct ML parameters off-line soft sensing models based on LSSVR. At last, when a new sample is given, after predicted with the older models, the inputs and regression parameters of the soft sensing models are updated by RPCA and on-line LSSVR algorithm respectively. Therefore, the on-line updating of the soft-sensing models for ML parameters are implemented. Experiment result shows that the proposed approach has higher accuracy and better predictive performance than other normal approaches.

Keywords: on-line soft sensing; recursive principal component analysis; least square support vector regression; mill load; vibration spectrum

1 引言 (Introduction)

磨机负荷 (ML) 是指磨机中球、料、水的总和, 是选矿行业磨矿过程^[1]的重要参数. 准确检测 ML 是实现磨矿过程优化控制、节能降耗及提高磨矿效率、磨矿产品质量的关键技术之一. 但是到目前为止, 磨机负荷的在线检测仍未得到完全解决^[2]. 磨矿过程的湿式球磨机负荷的检测主要结合现场专

家的经验进行 ML 状态识别^[3]. 针对能够准确表征 ML 的磨机内部参数 (料球比 MBVR、矿浆浓度 PD、充填率 CVR) 的检测研究逐渐成为热点^[4]. 文 [5-6] 基于振动及振声的频谱, 分别采用偏最小二乘 (PLS)、主元回归 (PCR)、遗传算法-偏最小二乘 (GA-PLS) 方法建立磨矿过程操作参数、ML 参数与频谱变量间的线性回归模型. 但是, 采用全部频谱

基金项目: 国家自然科学基金资助项目 (61020106003, 60874057); 中国博士后科学基金资助项目 (20100471464); 国家科技支撑计划资助项目 (2008BAB31B03).

通讯作者: 汤健, tjian001@126.com 收稿/录用/修回: 2010-10-21/2010-12-27/2011-05-23

变量的回归模型会导致过拟合、泛化能力差等问题；而采用特征选择后的频谱建模则会导致部分频谱信息丢失，降低模型的普适性。主元分析（PCA）—支持向量机（SVM）相结合的建模方法对分类问题具有快速、鲁棒和高识别率的优点^[7-8]。因此，依据振动频谱可以被分为自然频率段、主冲击频率段和次冲击频率段的特性，汤健等人提出了一种先分频段提取谱特征，再串行组合谱特征，最后采用 SVM 预测 ML 参数的软测量方法^[9]。考虑到筒体振动、振声、电流等信号与磨机负荷参数存在相关性、信息互补与冗余现象，文^[10]则提出了基于多源数据特征融合的球磨机负荷参数软测量方法。

但是，工业过程具有时变特性，如设备磨损、传感器和过程漂移、预防性的维护和清洗等。而且，构建软测量模型时，有时难以得到足够的训练样本，如磨矿过程中机理复杂且不能检测的钢球及衬板的磨损，进料的硬度、粒度分布的波动等均会导致该过程的时变。而且，磨矿过程运行的连续性使建模初期难以获得不同运行工况下的筒体振动信号。递归 PCA（RPCA）可以适应工业过程的这些变化^[11]，能够在保持模型原有信息的基础上递归更新模型。SVM 需要解决二次规划（QP）问题，难以实现在线建模。最小二乘支持向量回归机（LSSVR）简化 QP 问题为解线性方程组^[12]，并且在线 LSSVR 方法在机械结构系统辨识和化工过程关键指标的软测量中取得了较好的应用效果^[13-14]。因此，本文将 RPCA 和在线 LSSVR 相结合，提出了 ML 软测量模型的在线更新方法。首先，采用训练样本离线构建基于 PCA-LSSVR 的 ML 参数模型；然后，在线使用时，先采用 RPCA 递归更新 PCA 模型，计算更新后的谱主元矩阵及新样本的谱主元，再对 LSSVR 的模型进行在线递归更新；最后，更新后的模型准备用于下一样本的预测和再次更新。

2 基于 PCA 和 RPCA 的谱主元提取 (Spectral PCs extraction based on PCA and RPCA)

2.1 初始谱主元矩阵的提取

采用 PCA 处理振动频谱变量的共线性，其中心思想是通过建立谱变量间的线性合并，同时保持原始频谱变量中尽可能多的变化信息。这样，便实现了频谱数据的维数约简和特征提取^[15]。

假设原始频谱数据 $\mathbf{X}_k^0 \in \mathbf{R}^{k \times p}$ 由 k 个样本（行）和 p 个变量组成（列），则 \mathbf{X}_k^0 首先被标准化为均值

为 0、方差为 1 的 \mathbf{X}_k 。 \mathbf{X}_k 按下式分解：

$$\mathbf{X}_k = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{t}_2 \mathbf{p}_2^T + \cdots + \mathbf{t}_h \mathbf{p}_h^T + \mathbf{t}_{h+1} \mathbf{p}_{h+1}^T + \cdots + \mathbf{t}_p \mathbf{p}_p^T \quad (1)$$

其中， \mathbf{t}_i 和 \mathbf{p}_i ($i = 1, \dots, p$) 分别称为得分向量和负载向量。 \mathbf{p}_i 是如下式所示的相关系数阵 $\mathbf{R}_k \in \mathbf{R}^{p \times p}$ 的第 i 个特征向量：

$$\begin{cases} \mathbf{R}_k \approx \frac{1}{k-1} \mathbf{X}_k^T \cdot \mathbf{X}_k \\ (\mathbf{R}_k - \lambda_k) \mathbf{p}_k = 0 \end{cases} \quad (2)$$

其中， λ_k 是 \mathbf{R}_k 的特征值。由于 $\mathbf{T}_k \in \mathbf{R}^{k \times p}$ 是 \mathbf{X}_k 在 \mathbf{P}_k 上的正交投影，有：

$$\mathbf{T}_k = \mathbf{X}_k \mathbf{P}_k \quad (3)$$

通过分解频谱 \mathbf{X}_k 实现维数约简：

$$\mathbf{X}_k = \hat{\mathbf{X}}_k + \tilde{\mathbf{X}}_k = \hat{\mathbf{T}}_k \hat{\mathbf{P}}_k^T + \tilde{\mathbf{T}}_k \tilde{\mathbf{P}}_k^T \quad (4)$$

其中， $\hat{\mathbf{X}}_k$ 和 $\tilde{\mathbf{X}}_k$ 分别是建模部分和残差部分。 $\hat{\mathbf{P}}_k \in \mathbf{R}^{p \times h}$ 称为负荷矩阵，在过程监视中称为 PCA 模型； $\hat{\mathbf{T}}_k \in \mathbf{R}^{n \times h}$ 称为得分矩阵，在过程建模中用于构建过程模型，本文中称为谱主元矩阵； $\tilde{\mathbf{P}}_k^T \in \mathbf{R}^{p \times (p-h)}$ 和 $\tilde{\mathbf{T}}_k \in \mathbf{R}^{n \times (p-h)}$ 称为残差的负荷和得分矩阵。 $\hat{\mathbf{T}}_k$ 写为

$$\hat{\mathbf{T}}_k = \mathbf{X}_k \hat{\mathbf{P}}_k \quad (5)$$

根据上述定义，低、中、高不同频段的振动频谱可表示为

$$\mathbf{X}_{L_k} = \hat{\mathbf{T}}_{L_k} \hat{\mathbf{P}}_{L_k}^T + \tilde{\mathbf{T}}_{L_k} \tilde{\mathbf{P}}_{L_k}^T \quad (6)$$

$$\mathbf{X}_{M_k} = \hat{\mathbf{T}}_{M_k} \hat{\mathbf{P}}_{M_k}^T + \tilde{\mathbf{T}}_{M_k} \tilde{\mathbf{P}}_{M_k}^T \quad (7)$$

$$\mathbf{X}_{H_k} = \hat{\mathbf{T}}_{H_k} \hat{\mathbf{P}}_{H_k}^T + \tilde{\mathbf{T}}_{H_k} \tilde{\mathbf{P}}_{H_k}^T \quad (8)$$

其中，L、M 和 H 分别表示低、中和高频段。因此训练样本中提取的谱主元矩阵可表为

$$\mathbf{X}_{s_k}^0 = [\hat{\mathbf{T}}_{L_k}, \hat{\mathbf{T}}_{M_k}, \hat{\mathbf{T}}_{H_k}] \quad (9)$$

$\mathbf{X}_{s_k}^0$ 通过下式实现标准化：

$$\mathbf{X}_{s_k} = (\mathbf{X}_{s_k}^0 - \mathbf{1}_k \mathbf{u}_{s_k}^T) \cdot \boldsymbol{\Sigma}_{s_k}^{-1} \quad (10)$$

其中， $\mathbf{1}_k = [1, \dots, 1]^T \in \mathbf{R}^k$ ， \mathbf{u}_{s_k} 为 $\mathbf{X}_{s_k}^0$ 的均值， $\boldsymbol{\Sigma}_{s_k} = \text{diag}(\sigma_{s_{k1}}, \dots, \sigma_{s_{kp}})$ ， $\sigma_{s_{ki}}$ 是谱主元矩阵中第 i 变量的标准偏差。

2.2 谱主元的在线提取

当新的振动频谱样本 \mathbf{x}_{k+1}^0 可用时，在线更新 $\hat{\mathbf{P}}_k$ ，需要初始频谱矩阵 \mathbf{X}_k^0 的均值和方差。 \mathbf{X}_k^0 的均值可表示为

$$\mathbf{u}_k = \frac{1}{n} (\mathbf{X}_k^0)^T \cdot \mathbf{1}_k, \quad \mathbf{u}_k \in \mathbf{R}^{p \times 1} \quad (11)$$

\mathbf{X}_k^0 的标准化公式为

$$\mathbf{X}_k = (\mathbf{X}_k^0 - \mathbf{1}_k \mathbf{u}_k^T) \cdot \boldsymbol{\Sigma}_k^{-1} \quad (12)$$

其中 $\boldsymbol{\Sigma}_k = \text{diag}(\sigma_{k1}, \dots, \sigma_{kp})$, σ_{ki} 是 i 谱变量的标准偏差. \mathbf{u}_{k+1} 、 $\sigma_{(k+1)i}$ 和 \mathbf{R}_{k+1} 的递归计算如下^[10]:

$$\mathbf{u}_{k+1} = \frac{k}{k+1} \mathbf{u}_k + \frac{1}{k+1} (\mathbf{x}_{k+1}^0)^T \quad (13)$$

$$\sigma_{(k+1)i}^2 = \frac{k-1}{k} \sigma_{ki}^2 + \Delta \mathbf{u}_{k+1}^2(i) + \frac{1}{k} \|\mathbf{x}_{k+1}^0(i) - \mathbf{u}_{k+1}(i)\|^2 \quad (14)$$

$$\mathbf{R}_{k+1} = \frac{k-1}{k} \cdot \boldsymbol{\Sigma}_{k+1}^{-1} \cdot \boldsymbol{\Sigma}_k \cdot \mathbf{R}_k \cdot \boldsymbol{\Sigma}_k \cdot \boldsymbol{\Sigma}_{k+1}^{-1} + \boldsymbol{\Sigma}_{k+1}^{-1} \cdot \Delta \mathbf{u}_{k+1} \Delta \mathbf{u}_{k+1}^T \cdot \boldsymbol{\Sigma}_{k+1}^{-1} + \frac{1}{k} \cdot \mathbf{x}_{k+1}^T \cdot \mathbf{x}_{k+1} \quad (15)$$

其中, $\mathbf{x}_{k+1} = (\mathbf{x}_{k+1}^0 - \mathbf{1} \cdot \mathbf{u}_{k+1}^T) \cdot \boldsymbol{\Sigma}_{k+1}^{-1}$, $\boldsymbol{\Sigma}_j = \text{diag}(\sigma_{j1}, \dots, \sigma_{jp})$ ($j = k, k+1$).

通过对 \mathbf{R}_{k+1} 进行奇异值分解 (SVD), 求得 \mathbf{R}_{k+1} 的特征向量 $\mathbf{P}_{\mathbf{R}_{k+1}}$. 假设选择的最大主元个数为 h , 则新 PCA 模型 $\hat{\mathbf{P}}_{k+1}$ 为

$$\hat{\mathbf{P}}_{k+1} = \mathbf{P}_{\mathbf{R}_{k+1}}(:, 1:h) \quad (16)$$

利用公式 (5), 新样本的谱主元向量和新的谱主元矩阵可用下式计算

$$\begin{cases} \hat{\mathbf{i}}_{k+1} = \mathbf{x}_{k+1} \cdot \hat{\mathbf{P}}_{k+1} \\ \hat{\mathbf{T}}_{k+1} = \begin{bmatrix} \hat{\mathbf{T}}_k \cdot \hat{\mathbf{P}}_k^T \cdot \boldsymbol{\Sigma}_k \cdot \boldsymbol{\Sigma}_{k+1}^{-1} - \mathbf{1}_k \cdot \Delta \mathbf{u}_{k+1}^T \cdot \boldsymbol{\Sigma}_{k+1}^{-1} \\ \mathbf{x}_{k+1} \end{bmatrix} \cdot \hat{\mathbf{P}}_{k+1} \end{cases} \quad (17)$$

根据式 (11) ~ (17), 计算新样本不同频段的谱主元向量及不同频段的谱主元矩阵, 并串行组合得到:

$$\mathbf{x}_{s_{k+1}}^0 = [\hat{\mathbf{i}}_{L_{k+1}}, \hat{\mathbf{i}}_{M_{k+1}}, \hat{\mathbf{i}}_{H_{k+1}}] \quad (18)$$

$$\mathbf{X}_{s_{k+1}}^0 = [\hat{\mathbf{T}}_{L_{k+1}}, \hat{\mathbf{T}}_{M_{k+1}}, \hat{\mathbf{T}}_{H_{k+1}}] \quad (19)$$

并分别进行标准化

$$\mathbf{x}_{s_{k+1}} = (\mathbf{x}_{s_{k+1}}^0 - \mathbf{1} \cdot \mathbf{u}_{s_{k+1}}^T) \cdot \boldsymbol{\Sigma}_{s_{k+1}}^{-1} \quad (20)$$

$$\mathbf{X}_{s_{k+1}} = (\mathbf{X}_{s_{k+1}}^0 - \mathbf{1}_{k+1} \cdot \mathbf{u}_{s_{k+1}}^T) \cdot \boldsymbol{\Sigma}_{s_{k+1}}^{-1} \quad (21)$$

其中, $\mathbf{u}_{s_{k+1}}$ 和 $\boldsymbol{\Sigma}_{s_{k+1}}$ 是分别是 $\mathbf{X}_{s_{k+1}}$ 的均值和以其方差为对角线元素的对角阵. $\mathbf{X}_{s_{k+1}}$ 和 $\mathbf{x}_{s_{k+1}}$ 用于更新 ML 参数软测量模型.

3 基于 LSSVR 和在线 LSSVR 的 ML 参数软测量 (ML parameters soft-sensing based on LSSVR and on-line LSSVR algorithm)

3.1 初始软测量模型的建立

记采用 PCA 提取的训练数据的谱主元矩阵 \mathbf{X}_{s_k}

与 ML 参数的真值 \mathbf{Y}_k 组成的训练数据为 $\{(\mathbf{x}_{s_l}, y_l)\}$, $l = 1, \dots, k$. 采用如下的回归模型构造 ML 参数软测量模型:

$$y(\mathbf{x}_s) = \mathbf{W}^T \boldsymbol{\Phi}(\mathbf{x}_s) + b_k \quad (22)$$

其中 $\boldsymbol{\Phi}(\cdot)$ 将 PCA 提取的谱主元向量 \mathbf{x}_s 映射到高维特征空间; \mathbf{W} 是权值向量, b_k 是偏置项. 基于 LSSVR 估计 ML 参数, 需要考虑解决如下的优化问题:

$$\begin{cases} \min_{\mathbf{W}, b} J = \frac{1}{2} \mathbf{W}^T \mathbf{W} + \frac{1}{2} \gamma \sum_{l=1}^k \xi_l^2 \\ \text{s.t. } y_l = \mathbf{W}^T \boldsymbol{\Phi}(\mathbf{x}_{s_l}) + b_k + \xi_l \end{cases} \quad (23)$$

其中, J 是目标函数, γ 是在模型的复杂性和近似精度之间均衡的惩罚参数, ξ_k 是近似误差. 上式的拉格朗日形式为

$$L(\mathbf{W}, b_k, \boldsymbol{\xi}, \boldsymbol{\beta}) = \frac{1}{2} \mathbf{W}^T \mathbf{W} + \frac{1}{2} \gamma \sum_{l=1}^k \xi_l^2 - \sum_{l=1}^k \beta_l (\mathbf{W}^T \boldsymbol{\Phi}(\mathbf{x}_{s_l}) + b_k + \xi_l - y_l) \quad (24)$$

其中, β_l 是拉格朗日乘子. 通过求解如下 $(k+1) \times (k+1)$ 的线性等式, 可得式 (24) 的解.

$$\mathbf{A}_k \boldsymbol{\Theta}_k = \mathbf{Y}'_k \quad (25)$$

其中, 核相关矩阵 $\mathbf{A}_k = \begin{bmatrix} 0 & \tilde{\mathbf{1}}^T \\ \tilde{\mathbf{1}} & \boldsymbol{\Omega} + \frac{1}{\gamma} \mathbf{I} \end{bmatrix}$, 回归参数

$\boldsymbol{\Theta}_k = \begin{bmatrix} b_k \\ \mathbf{B}_k \end{bmatrix}$, $\mathbf{Y}'_k = \begin{bmatrix} 0 \\ \mathbf{Y}_k \end{bmatrix}$, $\tilde{\mathbf{1}} = [1, 1, \dots, 1]^T$, $\mathbf{B}_k = [\beta_1, \beta_2, \dots, \beta_k]^T$, $\mathbf{Y}_k = [y_1, y_2, \dots, y_k]^T$, \mathbf{I} 是 $k \times k$ 的单位阵, $\boldsymbol{\Omega}$ 满足如下的 Mercer 条件:

$$\boldsymbol{\Omega}_{lm} = K(\mathbf{x}_{s_l}, \mathbf{x}_{s_m}) = \boldsymbol{\Phi}^T(\mathbf{x}_{s_l}) \boldsymbol{\Phi}(\mathbf{x}_{s_m}), \quad l, m = 1, 2, \dots, k \quad (26)$$

其中 $\boldsymbol{\Omega}_{lm}$ 代表 $\boldsymbol{\Omega}$ 的第 l 行 m 列. 本文采用径向基函数 (RBF) $K(\mathbf{x}_{s_l}, \mathbf{x}_{s_m}) = \exp(-\|\mathbf{x}_{s_l} - \mathbf{x}_{s_m}\|^2 / \lambda^2)$ 作为 LSSVR 模型的核函数, λ 为 RBF 的核半径. 最终, ML 参数的 LSSVR 模型可表示为

$$y(\mathbf{x}_s) = \sum_{l=1}^k \beta_l K(\mathbf{x}_s, \mathbf{x}_{s_l}) + b_k \quad (27)$$

3.2 软测量模型的预测及在线更新

LSSVR 模型预测及更新需要全部训练样本即谱主元矩阵 \mathbf{X}_{s_k} . 采用类似式 (6) ~ (9) 的方法得到新样本的谱主元向量 $\mathbf{x}_{s_{k+1}}^{\text{pr}0}$ 后, 采用下式进行标准化:

$$\mathbf{x}_{s_{k+1}}^{\text{pr}} = (\mathbf{x}_{s_{k+1}}^{\text{pr}0} - \mathbf{1} \cdot \mathbf{u}_{s_k}^T) \cdot \boldsymbol{\Sigma}_{s_k}^{-1} \quad (28)$$

并采用式 (27) 表示的旧模型进行预测, 即:

$$\hat{y}(\mathbf{x}_{s_{k+1}}^{\text{pr}}) = \sum_{l=1}^k \beta_l K(\mathbf{x}_{s_{k+1}}^{\text{pr}}, \mathbf{x}_{s_l}) + b_k \quad (29)$$

采用 2.2 节的方法得到 $\mathbf{x}_{s_{k+1}}$, 记 $(\mathbf{x}_{s_{k+1}}, y_{k+1})$ 为新样本. 依据式 (25), 当前模型与新模型间的关系可表示为

$$\mathbf{A}_{k+1} \boldsymbol{\Theta}_{k+1} = \mathbf{Y}'_{k+1} \quad (30)$$

$$\text{其中, } \mathbf{A}_{k+1} = \begin{bmatrix} \mathbf{A}_k & \mathbf{a}_{k+1}^T \\ \mathbf{a}_{k+1} & \frac{1}{\gamma} + \boldsymbol{\Phi}^T(\mathbf{x}_{s_{k+1}}) \boldsymbol{\Phi}(\mathbf{x}_{s_{k+1}}) \end{bmatrix}, \boldsymbol{\Theta}_{k+1}$$

$$= \begin{bmatrix} b_{k+1} \\ \mathbf{B}_{k+1} \end{bmatrix}, \mathbf{B}_{k+1} = [\beta_1^u, \beta_2^u, \dots, \beta_k^u, \beta_{k+1}^u]^T, \text{上标 } u$$

表示该回归参数为更新后的回归参数; $\mathbf{Y}'_{k+1} = [0, \mathbf{Y}_{k+1}]^T = [\mathbf{Y}'_k, y_{k+1}]^T$. 新模型的回归参数 $\boldsymbol{\Theta}_{k+1}$ 的递归计算如下 [13-14]:

$$\boldsymbol{\Theta}_{k+1} = \begin{bmatrix} \boldsymbol{\Theta}_k + \eta_{k+1} \cdot \mathbf{A}_k^{-1} \cdot \mathbf{a}_{k+1}^T \cdot (\mathbf{a}_{k+1} \cdot \boldsymbol{\Theta}_k - y_{k+1}) \\ -\eta_{k+1} \cdot \mathbf{a}_{k+1} \cdot \boldsymbol{\Theta}_k + \eta_{k+1} \cdot y_{k+1} + \eta_{k+1} \cdot y_{k+1} \end{bmatrix} \quad (31)$$

$$\mathbf{A}_{k+1}^{-1} = \begin{bmatrix} \mathbf{A}_k^{-1} + \mathbf{A}_k^{-1} \cdot \mathbf{a}_{k+1}^T \cdot \eta_{k+1} \cdot \mathbf{a}_{k+1} \cdot \mathbf{A}_k^{-1} & -\mathbf{A}_k^{-1} \cdot \mathbf{a}_{k+1}^T \cdot \eta_{k+1} \\ -\eta_{k+1} \cdot \mathbf{a}_{k+1} \cdot \mathbf{A}_k^{-1} & \eta_{k+1} \end{bmatrix} \quad (32)$$

其中,

$$\eta_{k+1} = (\zeta_{k+1}^{-1} - \mathbf{a}_{k+1} \cdot \mathbf{A}_k^{-1} \cdot \mathbf{a}_{k+1}^T)^{-1} \quad (33)$$

$$\zeta_{k+1} = \left(\frac{1}{\gamma} + \boldsymbol{\Phi}^T(\mathbf{x}_{s_{k+1}}) \boldsymbol{\Phi}(\mathbf{x}_{s_{k+1}}) \right)^{-1} \quad (34)$$

$$\mathbf{a}_{k+1} = [1, \boldsymbol{\Phi}^T(\mathbf{x}_{s_1}) \boldsymbol{\Phi}(\mathbf{x}_{s_{k+1}}), \dots, \boldsymbol{\Phi}^T(\mathbf{x}_{s_k}) \boldsymbol{\Phi}(\mathbf{x}_{s_{k+1}})] \quad (35)$$

依据以上推导, 更新后的模型如下:

$$y(\mathbf{x}_s) = \sum_{l=1}^{k+1} \beta_l^u K(\mathbf{x}_s, \mathbf{x}_{s_l}) + b_{k+1} \quad (36)$$

4 在线软测量的步骤 (Steps of on-line soft sensing)

4.1 离线训练步骤

- (1) 采集作为训练数据的磨机筒体振动信号;
- (2) 采用带通滤波器对振动信号进行滤波;
- (3) 采用 Welch 方法计算振动频谱;
- (4) 将振动频谱分割为低、中、高三个频段;
- (5) 对三个频段进行标准化并存储各频段的均值与方差;

(6) 采用 PCA 提取不同频段的谱主元并存储三个频段的相关系数阵、PCA 模型、谱主元矩阵及其均值和方差;

(7) 对低、中、高频段分别选择 h_z 个谱主元 ($h_z \in \{h_L, h_M, h_H\}$), 利用式 (9) 和 (10) 计算模型的输入矩阵;

(8) 训练基于 LSSVR 的 ML 参数软测量模型;

(9) 采用测试数据测试 PCA-LSSVR 模型, 并计算其均方根相对误差 (RMSSE), 这里将其记为 $RMSSE(h_z)$;

(10) 如果 $RMSSE(h_z) < RMSSE(h_z - 1)$, 则 $h_z = h_z + 1$, 返回步骤 (7), 否则转步骤 (11);

(11) 存储选定的主元个数、LSSVR 的模型参数及回归系数.

4.2 模型预测及在线更新步骤

(1) 采集磨机筒体振动信号.

(2) 振动信号滤波.

(3) 计算振动频谱.

(4) 划分振动频谱为低、中、高三个频段.

(5) 预测: 对三个频段的频谱进行标准化, 并采用类似式 (6) ~ (9) 的方法计算振动频谱的不同频段的谱主元向量, 依式 (28) 和 (29) 进行 ML 参数预测.

(6) 更新:

1) 依据式 (13) ~ (16) 递归更新三个频段 PCA 模型;

2) 依据式 (17) ~ (20) 计算 LSSVR 模型的输入谱主元向量和谱主元矩阵;

3) 依据式 (31) ~ (35) 递归更新 LSSVR 模型;

4) 更新替换离线训练阶段存储的相关参数.

5 应用验证 (Application validation)

采用型号为 XMQL-420×450 的小型实验球磨机进行实验数据采集. 为了保证磨机内负荷的恒定, 实验时将磨机筒体的入口和进口封闭. 数据采集系统安装在磨机筒体上并随磨机旋转. 实验的详细描述见文 [6]. 为研究不同研磨条件下振动频谱的不同, 多数实验在异常工况下进行, 如高料球比、低浓度、高充填率等. 采用 90 个样本建立 ML 参数模型. 将频谱划分为 LF、MF、HF 频段后, PCA 提取谱主元, 各频段的贡献率如图 1 所示.

PCA 结果显示, 相同数量的不同频段谱主元的贡献率是不同的. 如 98% 的贡献率对于 LF、MF 和 HF 对应的主元个数分别为 5、3 和 1 个谱 PC. 尽管低贡献率的主元对于建模可能具有较好的预测能

力，但是这类主元建模通常会导致过程模型的不稳定^[15]。因此，设定 LF、MF 和 HF 的主元贡献率为 98%。

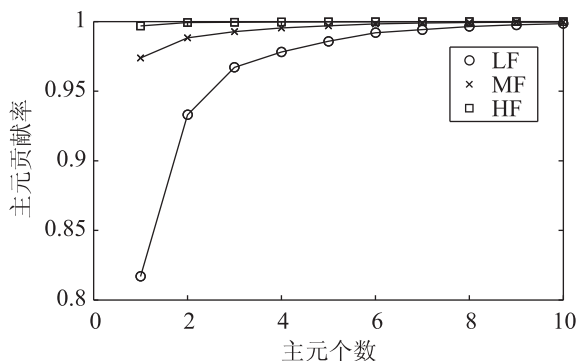


图 1 不同频段谱主元的贡献率

Fig.1 Contribution ratio of PCs for different frequency bands

测试数据采用 100 个样本，其中，前 50 个样本的工况与训练样本相同，后 50 个样本则来自不同的工况。也就是说，初始训练模型针对后 50 个样本的预测精度通常会很差，但采用模型在线更新算法克服该问题。本文所提的方法与 PCA-LSSVR、PLS 和 RPLS 的预测性能比较见表 1，预测曲线如图 2 ~ 4 所示。

表 1 不同软测量模型的性能比较

Tab.1 Performance comparison among different soft sensing models

方法	RMSSE (MBVR)	RMSSE (PD)	RMSSE (CVR)	RMSSE (average)
PLS	0.859 3	1.170 3	0.235 9	0.755 1
RPLS	0.059 75	0.132 2	0.070 11	0.087 35
PCA-LSSVR	0.222 9	0.086 83	0.134 5	0.148 1
RPCA-LSSVR	0.073 29	0.055 72	0.074 86	0.067 95

结果表明在线 RPCA-LSSVR 方法具有较高的预测精度。对料球比、矿浆浓度和充填率模型的预测精度分别是 0.073 29、0.055 72 和 0.074 86。最后 50 个测试样本的变化主要是物料负荷的递增。结果同时表明，某些样本点精度较低，原因是 LSSVR 模型的惩罚参数 γ 和核半径 λ 没有更新。与 RPLS 比较表明，基于 RPLS 算法的 MBVR 和 CVR 的精度稍高于 RPCA-LSSVR，原因主要是 RPLS 适合于处理高维谱数据。但 RPLS 本质上是线性模型，针对 PD 则是 PCA-LSSVR 具有更高的精度。

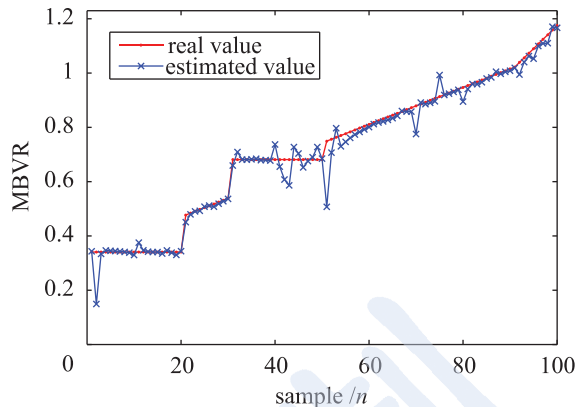


图 2 料球比软测量结果

Fig.2 Soft sensing results of MBVR

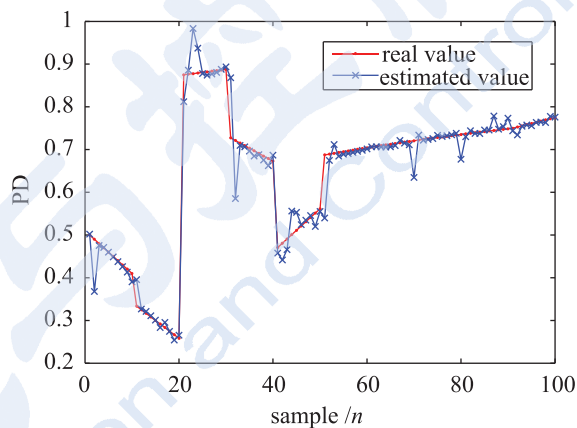


图 3 矿浆浓度软测量结果

Fig.3 Soft sensing results of PD

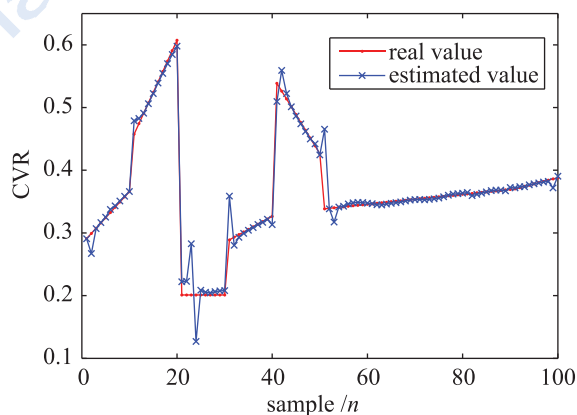


图 4 充填率测量结果

Fig.4 Soft sensing results of CVR

6 结论 (Conclusion)

本文提出了基于振动频谱的 ML 参数在线软测量方法。该方法有效地结合了 RPCA 和在线 LSSVR 算法，前者实现高维谱数据的在线特征提取，同时更新 PCA 模型和谱主元矩阵，后者利用谱主元矩阵和新样本的谱主元向量完成 ML 参数模型的在线更新，从而实现了特征提取和软测量模型的同步更

新. 实验结果表明该方法可以显著提高模型的预测精度. 进一步的研究方向是如何提高该方法的计算效率并同时提高 LSSVR 模型的稀疏性.

参考文献 (References)

- [1] Zhou P, Chai T Y. Intelligent optimal-setting control for grinding circuits of mineral processing[J]. IEEE Transactions on Automation Science and Engineering, 2009, 6(4): 730-743.
- [2] Behera B, Mishra B K, Murty C V R. Experimental analysis of charge dynamics in tumbling mills by vibration signature technique[J]. Minerals Engineering, 2007, 20(1): 84-91.
- [3] 周平, 柴天佑. 磨矿过程磨机负荷的智能监测与控制 [J]. 控制理论与应用, 2008, 25(6): 1095-1099.
- Zhou P, Chai T Y. Intelligent monitoring and control of mill load for grinding processes[J]. Control Theory & Applications, 2008, 25(6): 1095-1099.
- [4] 李勇, 邵诚. 灰色软测量在介质填充率检测中的应用研究 [J]. 中国矿业大学学报, 2006, 35(4): 549-555.
- Li Y, Shao C. Application research of grey soft sensor for charge ratio of media[J]. Journal of China University of Mining & Technology, 2006, 35(4): 549-555.
- [5] Zeng Y, Forssberg E. Monitoring grinding parameters by vibration signal measurement - A primary application[J]. Minerals Engineering, 1994, 7(4): 495-501.
- [6] Tang J, Zhao L J, Zhou J W, et al. Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell[J]. Minerals Engineering, 2010, 23(9): 720-730.
- [7] Li S, Fevens T, Krzyzak A, et al. Automatic clinical image segmentation using pathological modeling, PCA and SVM[J]. Engineering Applications of Artificial Intelligence, 2006, 19(4): 403-410.
- [8] Gumus E, Kilic N, Sertbas A, et al. Evaluation of face recognition techniques using PCA, wavelets and SVM[J]. Expert Systems with Applications, 2010, 37(9): 6404-6408.
- [9] Tang J, Zhao L J, Yu W, et al. Soft sensor modeling of ball mill load via principal component analysis and support vector machines[M]//Lecture Notes in Electrical Engineering: vol.67. Berlin, Germany: Springer-Verlag, 2010: 803-810.
- [10] 汤健, 赵立杰, 岳恒, 等. 基于多源数据特征融合的球磨机负荷软测量方法 [J]. 浙江大学学报: 工学版, 2010, 44(7): 1406-1413.
- Tang J, Zhao L J, Yue H, et al. Soft sensor for ball mill load based on multi-source data feature fusion[J]. Journal of Zhejiang University: Engineering Science, 2010, 44(7): 1406-1413.
- [11] Li W H, Yue H H, Valle-Cervantes S, et al. Recursive PCA for adaptive process monitoring[J]. Journal of Process Control, 2000, 10(5): 471-486.
- [12] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.
- [13] Tang H S, Xue S T, Chen R, et al. Online weighted LS-SVM for hysteretic structural system identification[J]. Engineering Structures, 2006, 28(12): 1728-1735.
- [14] Li L J, Su H Y, Chu J. Modeling of isomerization of C8 aromatics by online least squares support vector machine[J]. Chinese Journal of Chemical Engineering, 2009, 17(3): 437-444.
- [15] Jolliffe I T. Principal component analysis[M]. Berlin, Germany: Springer, 2002.

作者简介:

- 汤健 (1974-), 男, 博士生, 工程师. 研究领域为综合自动化系统, 基于数据的软测量建模等.
- 赵立杰 (1972-), 女, 博士, 副教授. 研究领域为多元统计分析方法, 工业过程性能监视和故障诊断.
- 柴天佑 (1947-), 男, 博士, 教授. 研究领域为自适应控制, 多变量智能解耦, 工业过程综合自动化等.

(上接第 122 页)

- [9] Zhang Q Z, Zhang W D. Priority scheduling in switched industrial Ethernet[C]//Proceedings of the 2005 American Control Conference. Piscataway, NJ, USA: IEEE, 2005: 3366-3370.
- [10] OPNET Model[EB/OL]. (2010-10-23)[2010-12-20]. <http://www.opnet.com>.

作者简介:

田中大 (1978-), 男, 博士生. 研究领域为网络控制系统,

复杂工业建模与控制, 智能控制理论.

高宪文 (1955-), 男, 博士生导师. 研究领域为复杂工业建模与控制, 智能控制理论应用与开发.

李焜 (1983-), 男, 博士生. 研究领域为故障诊断与优化, 复杂工业建模与控制.