

文章编号: 1000-8152(2008)03-0393-05

Hammerstein-Wiener模型最小二乘向量机辨识及其应用

桂卫华, 宋海鹰, 阳春华

(中南大学 信息科学与工程学院, 湖南 长沙 410083)

摘要: 借鉴最小二乘支持向量机求解的思路, 文中提出了辨识多输入-单输出Hammerstein-Wiener模型的方法。引入共线性约束假设, 将辨识问题转换为有约束的优化问题, 从而辨识出Hammerstein-Wiener模型的参数。基于Hammerstein-Wiener模型, 我们建立了一个多输入-单输出的滚动预测模型, 对铜转炉造渣S2期吹炼所需总氧量进行了预测, 其相对均方根误差为12.1%。仿真结果表明, 该模型预测准确、具有较好的应用价值。

关键词: Hammerstein-Wiener模型; 最小二乘向量机; 系统辨识; 智能建模; 铜转炉吹炼预测

中图分类号: TP18 文献标识码: A

Hammerstein-Wiener model identified by least-squares-support-vector machine and its application

GUI Wei-hua, SONG Hai-ying, YANG Chun-hua

(School of Information Science and Engineering, Central South University, Changsha Hunan 410083, China)

Abstract: The identification method for a multi-input single-output Hammerstein-Wiener model is proposed by using the solving method of the least-squares-support-vector machine. The identification problem is converted into a constrained optimization problem by assuming collinear constraints so that the parameters of Hammerstein-Wiener model can be identified. Based on the Hammerstein-Wiener model, a multi-input single-output receding-horizon prediction model is developed for predicting the total oxygen quantity required by a copper converter in slag making S2 stage. The relative root-mean-square error (RRMSE) is 12.1%. The simulation research shows that this model provides accurate prediction and is with desirable application value.

Key words: Hammerstein-Wiener model; least squares vector machine; system identification; intelligence model; copper converting prediction

1 引言(Introduction)

通常化学、冶金反应过程具有强烈的非线性特点, 而线性模型对非线性过程的描述能力十分有限, 但非线性模型, 即使是最简单的非线性模型也比线性模型能更好地描述整个生产过程中的特征。因此, 非线性动态模型的辨识与建立, 是基于模型的控制方法中的研究重点。

目前, 在非线性动态建模研究中, 一类新颖的块结构非线性动态模型则是其中的一个研究热点。该模型结构同时结合了线性动态模型和静态(无记忆)非线性函数模型, 具有较易辨识、计算量少、能较好地反映过程特征的特点, 适合作为过程控制模型使用。按其具体连接形式, 块结构非线性动态模型又可分为: Hammerstein模型、Wiener模型, 以及其各种组合形式Wiener-Hammerstein(L-N-L)模型与Hammerstein-Wiener(N-L-N)模型。其中, Hammerstein模型由Narendra和Gallman(1966)提出, 该模

型由一个非线性块再串联一个线性动态模块构造而成, 而由Masry和Cambanis(1980)所建立的Wiener模型的连接关系与此正好相反。由于Hammerstein模型中非线性部分无需以前的输入、输出信息, 其计算量要少于Wiener模型, 因此Hammerstein模型得到了更广泛的应用。

Bai^[1,2]、以及Chen等^[3]相继提出迭代法、将线性部分和非线性部分分别辨识的blind方法, 以及递推最小二乘法辨识Hammerstein系统。Krzyzak等^[4]将协相关法和递推核回归法用于对Hammerstein进行辨识。陈宗海等^[5]将非线性静态部分和线性动态部分分别用非线性基和LagueRe级数表示, 从而辨识出两部分参数。而万百五等^[6]采用神经网络对Hammerstein系统进行辨识, Goethals等^[7]则将最小二乘支持向量机(least squares support vector machine, LS SVM)应用在Hammerstein系统的辨识中, 通过增加约束及冗余参数法, 辨识出线性部分和非线性部分

收稿日期: 2007-03-14; 收修改稿日期: 2007-09-12。

基金项目: 国家自然科学基金资助项目(60634020, 60574030); 国家重点基础研究发展计划资助项目(2002cb312200); 博士点基金资助项目(20050533016)。

的系数。而对于Wiener模型, Chen^[8]利用了递推的方法进行辨识。Voros^[9]也进行了相应的研究。

由于 Hammerstein-Wiener(N-L-N) 模型包含了 Hammerstein 和 Wiener 两个非线性模块, 因此, 该模型结构比上述两者中的任一模型更接近于非线性系统。Bai^[10]针对基于非线性函数基的N-L-N模型, 提出了两段辨识法, 首先采用递推最小二乘辨识出模型的回归系数, 再采用奇异值分解法辨识出线性和非线性部分的系数。Zhu^[11]采用松弛迭代法, 通过最小化辨识误差对N-L-N模型进行参数估计。Crama等^[12]首先通过傅立叶变换, 将问题由时域空间转到频域空间, 再采用迭代法对N-L-N模型进行辨识。而Park等^[13]研究了采用激励信号辨识N-L-N模型的工作。虽然已对N-L-N模型进行了一些研究, 但是与其它块结构模型的研究工作相比, 目前的研究工作还远远不够。

通常在采用迭代法辨识N-L-N模型时, 局部渐进收敛是该类算法的缺点。而最小二乘向量机本身能较好地表示非线性过程, 并且其参数具有冗余性。因此, 提出了一种利用最小二乘支持向量机辨识N-L-N模型的方法, 假设满足共线性约束, 同时结合奇异值分解方法(SVD)辨识出N-L-N模型的系数。

2 Hammerstein-Wiener 模型 (Hammerstein-Wiener model)

Hammerstein-Wiener模型由一个静态的输入非线性模块 f , 一个动态的线性模块 G 和一个静态的输出非线性模块 h 组成, 如图1所示。



图1 Hammerstein-Wiener模型

Fig. 1 Hammerstein-Wiener model

图1中: $f(x), h(v)$ 表示静态非线性函数, 而 $G(u)$ 表示线性动态系统。对于SISO型N-L-N模型各模块分别表示为:

$$u = f(x), G(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_m z^{-m}}{1 - a_1 z^{-1} - \cdots - a_n z^{-n}},$$

其中: $(a_1, \dots, a_n)^T$ 和 $(b_0, \dots, b_m)^T$ 为参数向量, 而 $y = h(v)$ 。

由 $G(z)$ 可得:

$$v(k) = \sum_{i=1}^n a_i v(k-i) + \sum_{j=0}^m b_j u(k-j), \quad (1)$$

若函数 h 具有一对一映射关系时, 则存在逆映射 $v = h^{-1}(y) = g(y)$, 因此, 可以建立预测模型, 表示如下:

$$y(k) = \sum_{i=1}^n a_i g(y(k-i)) + \sum_{j=0}^m b_j f(x(k-j)) + e(k). \quad (2)$$

3 Hammerstein-Wiener模型的最小二乘向量机辨识(Hammerstein-Wiener model identified by LS SVM)

对于给定的训练样本集 $S = (\mathbf{x}_i, y_i), (i = 1, \dots, N)$, 其中: 输入向量 $\mathbf{x} \in \mathbb{R}^{n_x}$, ($\mathbf{x} = [\dots, x_i, \dots], x_i \in \mathbb{R}^m$), 一维输出 $y_i \in \mathbb{R}$. 因此, MISO型 Hammerstein-Wiener模型可表示为:

$$\begin{aligned} y(k) = & \sum_{i=1}^n a_i g(y_{k-i}) + \sum_{j=0}^m \mathbf{B}_j f(\mathbf{x}_{k-j}) + e(k) = \\ & \sum_{i=1}^n a_i g(y_{k-i}) + \sum_{j=0}^m \sum_{l=1}^{n_x} \mathbf{b}_l f(x_{l,k-j}) + e(k), \end{aligned} \quad (3)$$

其中: $e(k) \in \mathbb{R}$ 表示系统的预测误差, g, f 为非线性函数

$$\begin{aligned} g : \mathbb{R} &\rightarrow \mathbb{R}, f : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}, \\ (f(\mathbf{x})) &= [f(x_1), \dots, f(x_l), \dots, f(x_{n_x})]^T, \end{aligned}$$

而

$$\begin{aligned} \mathbf{a} &= (a_1, \dots, a_n)^T, \mathbf{B}_j \in \mathbb{R}^{n_x}, \\ \mathbf{b}_j &= (\mathbf{b}_1, \dots, \mathbf{b}_l, \dots, \mathbf{b}_{n_x}), \end{aligned}$$

其中: $\mathbf{b}_l = (b_0, \dots, b_j, \dots, b_m)^T$ 分别表示回归参数。

在支持向量机回归法中, 可以利用非线性映射函数 $\varphi(\mathbf{x})$ 将样本映射到高维特征空间, 这样就将原样本空间中的非线性函数估计问题转化成高维特征空间中的线性函数估计问题。因此, 如果用最小二乘支持向量机表示公式(3)中的非线性函数, 可以令:

$$\begin{aligned} a_i g(y_{k-i}) &= a_i (\omega^T \phi(y_{k-i}) + d_1) = \\ & \omega_i^T \phi(y_{k-i}) + a_i d_1, \end{aligned} \quad (4)$$

$$\begin{aligned} \sum_{l=1}^{n_x} \mathbf{b}_l f(x_{l,k-j}) &= \sum_{l=1}^{n_x} \mathbf{b}_l (\omega^T \varphi(x_{l,k-j}) + d_0) = \\ & \mathbf{w}_j^T \varphi(\mathbf{x}_{k-j}) + \mathbf{B}_j d_0. \end{aligned} \quad (5)$$

因此, 可以得到:

$$\begin{aligned} y_k &= \sum_{l=1}^n (\omega_l^T \phi(y_{k-l}) + a_l d_1) + \\ & \sum_{j=0}^m (\mathbf{w}_j^T \varphi(\mathbf{x}_{k-j}) + \mathbf{B}_j d_0) + e(k). \end{aligned} \quad (6)$$

其中:

$$\omega_i = a_i \omega, ,$$

$$\mathbf{w}_j = [\mathbf{b}_1 w, \dots, \mathbf{b}_l w, \dots, \mathbf{b}_{n_x} w]^T,$$

$$\varphi(\mathbf{x}) = [\varphi(x_1), \dots, \varphi(x_l), \dots, \varphi(x_{n_x})]^T,$$

而 d_0 和 d_1 表示回归参数, 再令

$$\sum_{i=1}^n a_i d_1 = c_1, \sum_{j=1}^m \mathbf{B}_j d_0 = \mathbf{c}_0.$$

为辨识出 Hammerstein-Wiener模型中的系数, 假设式(6)满足共线性约束条件:

$$\begin{aligned} (\omega_i^T \omega_k)^2 &= \sqrt{\omega_i^T \omega_i} \cdot \sqrt{\omega_k^T \omega_k}, j, k = 1, \dots, n, \\ (\mathbf{w}_j^T \mathbf{w}_k)^2 &= \sqrt{\mathbf{w}_j^T \mathbf{w}_j} \cdot \sqrt{\mathbf{w}_k^T \mathbf{w}_k}, j, k = 0, \dots, m, \end{aligned}$$

即系数 ω, \mathbf{w} 唯一性条件成立.

因此, 定义优化问题如下:

$$\begin{aligned} \min_{\mathbf{w}_j, \mathbf{c}_0, \omega_i, c_1, e} J(\mathbf{w}_j, \omega_i, e) = \\ \frac{1}{2} \sum_{j=0}^m \mathbf{w}_j^T \mathbf{w}_j + \frac{1}{2} \sum_{i=1}^n \omega_i^T \omega_i + \frac{C}{2} \sum_{k=r}^N e_k^2, \end{aligned} \quad (7)$$

满足约束条件:

$$\left\{ \begin{array}{l} \sum_{i=1}^n \omega_i^T \phi(y_{k-i}) + c_1 + \sum_{j=0}^m \mathbf{w}_j^T \varphi(\mathbf{x}_{k-j}) + \\ \mathbf{c}_0 + e_k - y_k = 0, \\ \sum_{k=1}^N \mathbf{w}_j^T \varphi(\mathbf{x}_k) = 0, \\ \sum_{k=1}^N \omega_j^T \phi(y_k) = 0. \end{array} \right. \quad (8)$$

公式(7)(8)中($k = r, \dots, N$), $r = \max(m, n) + 1$, 而公式(8)中的第2,3项, 使投影在特征空间中的非线性函数的值, 都位于整个训练样本集的平均值左右, 从而消除了由误差项所造成的不确定性, 继而可以辨识出模型的回归系数 a_i, b_j .

为了求解公式(7)和(8)表示的优化问题, 构造拉格朗日函数:

$$\begin{aligned} L(\mathbf{w}_j, \mathbf{c}_0, \omega_i, c_1, e, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, C) = \\ J(\mathbf{w}_j, \omega_i, e) - \sum_{j=0}^m \beta_j \sum_{k=1}^N \mathbf{w}_j^T \varphi(\mathbf{x}_k) - \\ \sum_{i=1}^n \gamma_i \sum_{k=1}^N \omega_i^T \phi(y_k) - \sum_{k=r}^N \alpha_k [\sum_{i=1}^n \omega_i^T \phi(y_{k-i}) + c_1 + \\ \sum_{j=0}^m \mathbf{w}_j^T \varphi(\mathbf{x}_{k-j}) + \mathbf{c}_0 + e_k - y_k]. \end{aligned} \quad (9)$$

$$\begin{bmatrix} 0 & 0 & \mathbf{1}_{N-r+1}^T & 0 & 0 \\ 0 & 0 & \mathbf{1}_{N-r+1}^T & 0 & 0 \\ \mathbf{1}_{N-r+1} \mathbf{1}_{N-r+1}^T \mathbf{K} + \mathbf{R} + \mathbf{I}/C & \Delta & \mathbf{\Omega} & \mathbf{c}_1 \\ 0 & 0 & \Delta^T & \mathbf{1}_N^T \boldsymbol{\kappa}_x \mathbf{1}_N \cdot \mathbf{I}_{m+1} & 0 \\ 0 & 0 & \mathbf{\Omega}^T & 0 & \mathbf{1}_N^T \boldsymbol{\kappa}_y \mathbf{1}_N \cdot \mathbf{I}_n \end{bmatrix} \cdot \begin{bmatrix} \mathbf{c}_0 \\ \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mathbf{Y}_f \\ 0 \\ 0 \end{bmatrix}, \quad (12)$$

其中:

$$\boldsymbol{\alpha} = [\alpha_r, \dots, \alpha_N]^T, \boldsymbol{\beta} = [\beta_0, \dots, \beta_m]^T,$$

$$\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_n]^T,$$

$$\mathbf{K}_{(N-r+1) \times (N-r+1)}(k_1, k_2) = \sum_{j=0}^m \varphi_{k_1-j}^T \varphi_{k_2-j},$$

公式(9)中, $\alpha_k, \beta_j, \gamma_i$ 为拉格朗日乘子.

又根据KKT条件得到下列方程:

$$\left\{ \begin{array}{l} \frac{\partial L}{\partial \mathbf{w}_j} = 0 \rightarrow \mathbf{w}_j = \beta_j \sum_{k=1}^N \varphi(\mathbf{x}_k) + \sum_{k=r}^N \alpha_k \varphi(\mathbf{x}_{k-j}), \\ j = 0, \dots, m, \\ \frac{\partial L}{\partial \omega_i} = 0 \rightarrow \omega_i = \gamma_i \sum_{k=1}^N \phi(y_k) + \sum_{k=r}^N \alpha_k \phi(y_{k-i}), \\ i = 1, \dots, n, \\ \frac{\partial L}{\partial \mathbf{c}_0} = 0 \rightarrow -\sum_{k=r}^N \alpha_k = 0, \frac{\partial L}{\partial c_1} = 0 \rightarrow -\sum_{k=r}^N \alpha_k = 0, \\ \frac{\partial L}{\partial \beta_j} = 0 \rightarrow \sum_{k=1}^N \mathbf{w}_j^T \varphi(\mathbf{x}_k) = 0, \\ \frac{\partial L}{\partial \gamma_i} = 0 \rightarrow \sum_{k=1}^N \omega_i^T \phi(y_k) = 0, \\ \frac{\partial L}{\partial \alpha_k} = 0 \rightarrow \sum_{i=1}^n \omega_i^T \phi(y_{k-i}) + c_1 + \\ \sum_{j=0}^m \mathbf{w}_j^T \varphi(\mathbf{x}_{k-j}) + \mathbf{c}_0 + e_k = y_k, \\ \frac{\partial L}{\partial e_k} = 0 \rightarrow C e_k = \alpha_k, k = r, \dots, N. \end{array} \right. \quad (10)$$

由方程组(10)可得:

$$\begin{aligned} y_{k_2} = & \sum_{i=1}^n \gamma_i \sum_{k_1=1}^N \phi^T(y_{k_1}) \phi(y_{k_2-i}) + \\ & \sum_{i=1}^n \sum_{k_1=r}^N \alpha_k \phi^T(y_{k_1-i}) \phi(y_{k_2-i}) + \\ & \sum_{j=0}^m \beta_j \sum_{k_1=1}^N \varphi^T(\mathbf{x}_{k_1}) \varphi(\mathbf{x}_{k_2-j}) + \\ & \sum_{j=0}^m \sum_{k_1=r}^N \alpha_k \varphi^T(\mathbf{x}_{k_1-j}) \varphi(\mathbf{x}_{k_2-j}) + \\ & c_1 + \mathbf{c}_0 + \frac{\alpha_{k_2}}{C}, \end{aligned} \quad (11)$$

公式(11)中: $k_2 = r, \dots, N$. 由公式(10)可得:

$$\begin{aligned} \mathbf{R}_{(N-r+1) \times (N-r+1)}(k_1, k_2) &= \sum_{i=1}^n \phi_{k_1-i}^T \phi_{k_2-i}, \\ (k_1, k_2 = r, \dots, N), \\ \Delta_{(N-r+1) \times (m+1)}(k_2, j) &= \sum_{k_1=1}^N \varphi_{k_1}^T \varphi_{k_2-j}, \end{aligned}$$

$$(k_2 = r, \dots, N; j = 0, \dots, m),$$

$$\boldsymbol{\Omega}_{(N-r+1) \times n}(k_2, i) = \sum_{k_1=1}^N \phi_{k_1}^T \phi_{k_2-i},$$

$$(k_2 = r, \dots, N; i = 1, \dots, n),$$

$$\mathbf{k}_{x(N \times N)}(k_1, k_2) = \varphi(\mathbf{x}_{k_1})^T \varphi(\mathbf{x}_{k_2}),$$

$$\mathbf{k}_{y(N \times N)}(k_1, k_2) = \phi(y_{k_1})^T \phi(y_{k_2}),$$

$$(k_1, k_2 = 1, \dots, N),$$

$$\begin{bmatrix} \mathbf{B}_0 \\ \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_m \end{bmatrix} \begin{bmatrix} \widehat{f}(\mathbf{x}_1) \\ \widehat{f}(\mathbf{x}_2) \\ \vdots \\ \widehat{f}(\mathbf{x}_N) \end{bmatrix}^T = \begin{bmatrix} \alpha_N \cdots \alpha_r & 0 \\ \alpha_N \cdots \alpha_r & \ddots \\ \ddots & \ddots \\ 0 & \alpha_N \cdots \alpha_r \end{bmatrix} \times \begin{bmatrix} \mathbf{K}_{N,1} & \mathbf{K}_{N,2} & \cdots & \mathbf{K}_{N,N} \\ \mathbf{K}_{N-1,1} & \mathbf{K}_{N-1,2} & \cdots & \mathbf{K}_{N-1,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{K}_{r-m,1} & \mathbf{K}_{r-m,2} & \cdots & \mathbf{K}_{r-m,N} \end{bmatrix} + \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \sum_{k=1}^N \begin{bmatrix} \mathbf{k}_{\mathbf{x}_{k,1}} \\ \mathbf{k}_{\mathbf{x}_{k,2}} \\ \vdots \\ \mathbf{k}_{\mathbf{x}_{k,N}} \end{bmatrix}. \quad (13)$$

公式(13)中 $\widehat{f}(\mathbf{x}_k)$ 是 $\underline{f}(\mathbf{x}_k) = f(\mathbf{x}_k) - \mathbf{c}_0$ 的估计值, 从而可推出: $f(\mathbf{x}) = \underline{f}(\mathbf{x}) + \mathbf{c}_0 / \sum_{j=0}^m \mathbf{B}_j$, 由于公

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \begin{bmatrix} \widehat{g}(y_1) \\ \widehat{g}(y_2) \\ \vdots \\ \widehat{g}(y_N) \end{bmatrix}^T = \begin{bmatrix} \alpha_N \cdots \alpha_r & 0 \\ \alpha_N \cdots \alpha_r & \ddots \\ \ddots & \ddots \\ 0 & \alpha_N \cdots \alpha_r \end{bmatrix} \times \begin{bmatrix} \Re_{N,1} & \Re_{N,2} & \cdots & \Re_{N,N} \\ \Re_{N-1,1} & \Re_{N-1,2} & \cdots & \Re_{N-1,N} \\ \vdots & \vdots & \ddots & \vdots \\ \Re_{r-(n-1),1} & \Re_{r-(n-1),2} & \cdots & \Re_{r-(n-1),N} \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} \sum_{k=1}^N \begin{bmatrix} k_{y_{k,1}} \\ k_{y_{k,2}} \\ \vdots \\ k_{y_{k,N}} \end{bmatrix}. \quad (14)$$

同理, 公式(14)中 $\widehat{g}(y_k)$ 是 $g(y_k) = g(y_k) - c_1$ 的估计值. 因为, 公式(14)等号右边为已知项, 因此, 对其进行奇异值分解可求得系数 α_j .

4 应用研究(Application research)

P-S铜转炉的吹炼过程, 是具有强烈非线性特征的冶金生产过程. 目前国内外P-S铜转炉的生产方式均以人工操作方式为主, 生产质量依赖于工人的技术水平. 若能对吹炼过程做出准确预报, 则有利于实现冶金工业过程的优化操作, 本文将建立基于最小二乘向量机的Hammerstein-Wiener(N-L-N)动态预测模型, 对造渣期S2期吹炼所需总氧量进行预测. 研究铜转炉造渣期吹炼机理可知, 影响造渣S2期吹炼总氧量(y_1)的变量包括: 造渣S1期和S2期的铜硫添加量(u_1, u_2)、造渣期熔剂添加量(u_3)、造渣期多种冷料的添加量(u_4, u_5)、造渣S1期的渣量(u_6)、造渣S1期的渣成分(u_7)、造渣S1期所用的时间(u_8)、造渣S1期所用的风量(u_9)、造渣S1期所用的氧量(u_{10}). 因此, 选择上述各变量在不同时刻的数据值组成输入矩阵(\mathbf{x}), 利用基于最小二乘支持向量机的N-L-N模型建立了造渣S2期吹炼所需总氧量(输出向量)的动态预测模型.

论文使用了某铜冶炼厂的5号P-S铜转炉的一个生产周期的共173炉数据, 在采用改进型肖维

$$\mathbf{1} = [1, \dots, 1]^T, \mathbf{Y}_f = [y_r, \dots, y_N]^T,$$

$$\mathbf{I} = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \end{bmatrix}.$$

由公式(12)可以直接求出待定参数 c_1, \mathbf{c}_0 , 而回归系数 $\alpha_i (i = 1, \dots, n)$ 和 $\mathbf{B}_j (j = 0, \dots, m)$ 需通过奇异值分解法确定, 因此建立等式:

式(13)等号右边为已知项, 因此, 对其进行奇异值分解可求得系数: \mathbf{B}_j .

耐(Chauvenet)法^[14]剔除部分异常样本后. 在应用滑动窗技术后, 每次利用100炉生产数据构建造渣S2期吹炼所需总氧量的动态预测模型. 通过所构建的动态模型, 对58炉样本数据进行滚动预测.

仿真研究采用未平滑的原始数据, 分别对基于最小二乘向量机的N-L-N模型和基于NARX结构的最小二乘向量机动态模型^[15]进行了对比. 上述两模型中最小二乘向量机均采用参数值相同的RBF核函数(核宽度 $\sigma^2 = 40$, 调整系数 C 取为 1000), 这是因为仿真数据是工人的经验操作值, 具有波动大的特点, 而建立本模型的目的是对操作量进行准确预测, 因此选择的调整系数(即经验误差项的罚系数)偏大, 从而使得模型的非线性逼近能力增加, 仿真实验的结果如图2所示.

曲线Hammerstein-Wiener表示采用最小二乘向量机进行参数辨识的N-L-N模型的预测值, 而曲线NARX表示输入矩阵为ARX形式的最小二乘支持向量机的预测值. 由上图对比分析可知, N-L-N模型的预测能力比采用ARX输入结构的最小二乘支持向量机的预测能力更好. 为评价预测模型的性能, 本文分别使用了: 最大相对误差 $E_{\max} = \max\left(\frac{|y_i - \hat{y}_i|}{y_i}\right)$,

均方根误差RMSE = $\sqrt{\frac{1}{l} \sum_{i=1}^l (y_i - \hat{y}_i)^2}$, 和相对均方根误差RRMSE = $\sqrt{\frac{1}{l} \sum_{i=1}^l \frac{(y_i - \hat{y}_i)^2}{y_i}}$, 其中: y_i 为实际数据向量中的第*i*个数据, \hat{y}_i 为预测所得数据向量中的第*i*个数据。根据上述公式对模型的预测准确性和跟踪变化趋势的能力进行了分析, 计算得到的各误差如下所示: N-L-N的 $E_{max} = 23.6\%$, RMSE = $10.3 \text{N} \cdot \text{m}^3$, RRMSE = 12.1%; 而NARX的 $E_{max} = 27.7\%$, RMSE = $15.4 \text{N} \cdot \text{m}^3$, RRMSE = 18.4%。

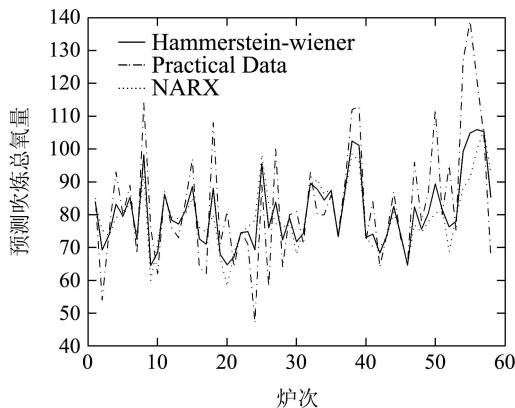


图2 N-L-N模型和NARX模型的预测对比

Fig. 2 Prediction comparison between N-L-N and NARX

由此可见, 基于LS SVM的N-L-N模型的预测性能要优于NARX结构的最小二乘支持向量机动态模型。基于LS SVM的N-L-N模型的预测吹炼总氧量的均方根误差为10.3标准牛立方米, 符合工业操作要求, 预测准确度较高; 其预测结果的相对均方根误差为12.1%, 表明该模型的对过程变化趋势跟踪性能较好, 能较好地预测过程的变化趋势; 但是由于采用了未经平滑滤波处理的数据进行建模, 导致预测结果的最大相对误差较大(为23.6%), 然而这表明该模型具有较高的抗噪性。

5 结论(Conclusion)

本文提出了一种利用最小二乘支持向量机辨识Hammerstein-Wiener模型的方法, 通过采用该方法, 可以避免过去为辨识Hammerstein-Wiener模型所采用的迭代计算, 从而有效地提高了Hammerstein-Wiener模型的使用效率, 使得Hammerstein-Wiener模型更适合作为工业过程的控制模型。仿真研究表明, 采用该方法构造的预测模型具有预测精度高, 计算效率快的特点, 具有较好的工业应用价值。

参考文献(References):

- [1] BAI E W. A blind approach to Hammerstein model identification[J]. *IEEE Transactions on Signal Process*, 2002, 50(7): 1610 – 1619.
- [2] LIU Y, BAI E W. Iterative identification of Hammerstein systems[J]. *Automatica*, 2007, 43(2): 346 – 354.
- [3] CHEN H F. Pathwise convergence of recursive identification algorithms for Hammerstein systems[J]. *IEEE Transactions on Automatic Control*, 2004, 49(10): 1641 – 1649.
- [4] KRZYZAK A, PARTYKA M A. Global identification of nonlinear Hammerstein systems by recursive kernel approach[J]. *Nonlinear Analysis, Theory, Methods and Applications*, 2005, 63(5/7): 1263 – 1272.
- [5] 向微, 陈宗海. 基于Hammerstein模型描述的非线性系统辨识新方法[J]. 控制理论与应用, 2007, 24(1): 143 – 147.
(XIANG Wei, CHEN Zonghai. New identification method of nonlinear systems based on Hammerstein models[J]. *Control Theory & Applications*, 2007, 24(1): 143 – 147.)
- [6] 孔金生, 万百五. 一种多输入单输出Hammerstein系统的集成辨识方法[J]. 控制理论与应用, 2005, 22(4): 517 – 519.
(KONG Jinseng, WAN Baiwu. Integrated identification method for multi-input/single-output Hammerstein system[J]. *Control Theory & Applications*, 2005, 22(4): 517 – 519.)
- [7] GOETHALS I, PELCKMANS K, SUYKENS J A K, et al. Identification of MIMO Hammerstein Models using least squares support vector machines[J]. *Automatica*, 2005, 41(7): 1263 – 1272.
- [8] CHEN H F. Recursive identification for Wiener model with discontinuous piece-wise linear function[J]. *IEEE Transactions on Automatic Control*, 2006, 51(3): 390 – 400.
- [9] VOROS, J. Parameter identification of Wiener systems with multisegment piecewise-linear nonlinearities[J]. *Systems and Control Letters*, 2007, 56(2): 99 – 105.
- [10] BAI E W. An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems[J]. *Automatica*, 1998, 34(3): 333 – 338.
- [11] ZHU Y C. Estimation of an N-L-N Hammerstein-Wiener model[J]. *Automatica*, 2002, 38(9): 1607 – 1614.
- [12] CRAMA P, SCHOUKENS J. Hammerstein-Wiener system estimator initialization[J]. *Automatica*, 2004, 40(9): 1543 – 1550.
- [13] PARK H C, SUNG S W, LEE J. Modeling of Hammerstein-Wiener processes with special input test signals[J]. *Industrial & Engineering Chemistry Research*, 2006, 45(3): 1029 – 1038.
- [14] 宋海鹰, 桂卫华, 阳春华, 等. 基于核偏最小二乘的动态预测模型在铜转炉吹炼中的应用[J]. 中国有色金属学报, 2007, 17(7): 1201 – 1205.
(SONG haiying, GUI weihua, YANG Chunhua, et al. Application of dynamical prediction model based on kernel partial least squares for copper converting[J]. *The Chinese Journal of Nonferrous Metals*, 2007, 17(7): 1201 – 1205.)
- [15] 宋海鹰, 桂卫华, 阳春华. 基于核偏最小二乘的简约最小二乘支持向量机及其应用研究[C] // 第26届中国控制会议. 张家界: 北京航空航天大学出版社, 2007, 3: 207 – 211.
(SONG haiying, GUI weihua, YANG Chunhua. Reduced least squares support vector based on kernel partial least squares and its application research[C] // *Proceedings of the 26th Chinese Control Conference*. Zhang Jiajie, Hunan: Beihang University Press, 2007, 3: 207 – 211.)

作者简介:

桂卫华 (1958—), 男, 中南大学教授, 博士生导师, 目前研究方向为复杂工业过程建模与优化、大系统及分散鲁棒控制, E-mail: gwh@mail.csu.edu.cn;

宋海鹰 (1975—), 男, 中南大学博士研究生, 目前研究方向为复杂冶炼过程建模与优化, E-mail: songhaiying1975@163.com;

阳春华 (1965—), 女, 中南大学教授, 博士生导师, 目前研究方向为复杂工业过程建模与优化, E-mail: ychh@mail.csu.edu.cn.