Internal model control based on a novel least square support vector machines for MIMO nonlinear discrete systems

Chaoxu Mu · Changyin Sun · Xinghuo Yu

Abstract To improve the robustness of the traditional inverse system method, the internal model control based on a novel least square support vector machines (LS-SVM) is proposed. The novel LS-SVM considers general errors that include noises of input variables and output variables as empirical errors. The data of original MIMO discrete system is exploited to approximate its inverse model by the novel LS-SVM. By cascading the inverse model and the original system to constitute a decoupling pseudo-linear system, the internal model control strategy is carried out to the pseudo-linear system to realize the effective control. Simulation validates that the novel LS-SVM used in the inverse system identification is effective and shows that the internal model control of nonlinear discrete systems has better robustness of anti-interference and parameters varying than that of the open-loop system only based on inverse control.

Keywords Internal model control · Inverse system · Least square support vector machines · Robustness

1 Introduction

Internal model control has been extensively studied in the case of linear systems, and has been shown good robust properties against disturbances. Nonlinear systems widely exist in industrial processes, and the development of internal model control for nonlinear models has already been paid great attention [1, 2]. For nonlinear systems,

C. Mu · C. Sun (✉) · X. Yu
School of Automation, Southeast University,
Nanjing 210096, China
e-mail: cysun@seu.edu.cn

whether continuous systems or discrete systems, they are very difficult to establish internal models because of imperfect mathematic models of original systems. The internal model control cannot be implemented without inversion models, which hinders the internal model control method to solve control problems of nonlinear systems [3]. Hence, it is very insistent to find an approach to obtain inverse functions effectively by the information and knowledge of original systems.

In recent years, the birth of intelligent learning algorithms encourages the development of nonlinear control without precise models of processes. Neural networks, as early algorithms, have been discussed in the internal model control of nonlinear systems and have played an important role [4, 5]. Support vector machines (SVM) introduced by Vapnik are a new methodology in the area of nonlinear modeling after neural networks. While neural networks suffer from problems like the existence of many local minima and the choice of the number of hidden units [6], SVM are characterized by convex optimization problems based on sound theoretical principles, up to the determination of a few additional tuning parameters, and provide better generalization performance than that of neural networks [7, 8]. The convex quadratic programming problem is solved in dual space in order to determine the SVM model. In the optimization formulation, one works with equality instead of inequality constraints and the sum squared error instead of the epsilon-insensitive cost function, thus LS-SVM are proposed. This reformulation greatly simplifies the problem in such a way that the solution is a Karush-Kuhn-Tucker (KKT) system [9, 10]. Because errors in SVM and LS-SVM are only the noises of output variables, which is unreasonable since the input variables may be polluted by noises. Reference [11] proposed the total least squares method to deal with the

problem and the normal least square support vector machine in reference [12] considered that noises of the input variable based on the total least squares method.

In this paper, we consider a novel LS-SVM with general errors that include the noises of input variables and output variables and an iterative algorithm is introduced to solve the LS-SVM. Then we propose the internal model control of LS-SVM with general errors for MIMO nonlinear discrete systems. We use LS-SVM to approximate the inverse system based on input-output data from the original system. The internal model is a pseudo linear system by connecting with the inverse model and the original system. The internal model controller is then designed by cascading a filter and the inversion of pseudo linear system. We focus on robust properties of internal model control in the case of disturbing signals and parameters varying. Comparatively speaking, the inverse control is a simple open-loop method whose ability is limit for disturbances and parameters varying. Simulation shows that the internal model control strategy based on LS-SVM is effective and has good performance.

This paper is organized as follows. In Sect. 2, we give a description of the class of MIMO nonlinear discrete systems, which helps to model an inverse system, and the control method based on inverse system is also be introduced. In Sect. 3, LS-SVM with general errors are given and an iterative algorithm is used to solve the LS-SVM, then the inverse model is approached by the novel LS-SVM based on effective data. In Sect. 4, we present the approach of internal model control and analyze the ability of the close-loop system. In Sect. 5, simulated processes are conducted to illustrate robustness of the internal model control system based on LS-SVM, comparing to the inverse system method proposed in the past. Conclusion is in Sect. 6.

2 System description and the control based on inverse system

We are interested in reversible MIMO nonlinear discrete systems. This kind of system is described by the following discrete nonlinear input-output model, $\Sigma : \mathbf{u}(k) \rightarrow \mathbf{y}(k)$

$$
\begin{aligned}
F(\mathbf{y}(k+\alpha), &\mathbf{y}(k+\alpha-1), \ldots, \mathbf{y}(k+\alpha-p), \mathbf{u}(k), \\
&\mathbf{u}(k-1), \ldots, \mathbf{u}(k-q)) = 0,
\end{aligned}
\tag{1}
$$

where $\mathbf{y} = (y_1, \ldots, y_n) \in R^n$, $\mathbf{u} = (u_1, \ldots, u_m) \in R^m$, $\mathbf{y}(k + \alpha - p) = (y_1(k + \alpha_1 - p_1), y_2(k + \alpha_2 - p_2), \ldots, y_n(k + \alpha_n - p_n))$, $\mathbf{u}(k - q) = (u_1(k - q_1), u_2(k - q_2), \ldots, u_m(k - q_m))$, α expressed relative delays of outputs to inputs, q denoted the input delays and $max\{p_1, p_2, \ldots, p_n\}$ denoted the order of the system.

The inverse system of the described system Σ is expressed in the formula, $\Sigma' : \mathbf{y}(k) \rightarrow \mathbf{u}(k)$

$$
\begin{aligned}
\mathbf{u}(k) = F^{-1}(&\mathbf{y}(k+\alpha), \mathbf{y}(k+\alpha-1), \ldots, \mathbf{y}(k+\alpha-p), \\
&\mathbf{u}(k-1), \ldots, \mathbf{u}(k-q)).
\end{aligned}
\tag{2}
$$

Denote $\varphi(k) = \mathbf{y}(k + \alpha)$ and $\varphi_1(k) = y_1(k + \alpha_1), \ldots, \varphi_n(k) = y_n(k + \alpha_n)$. $\mathbf{y}(k + \alpha - 1) = z^{-1}\varphi(k)$, $\mathbf{y}(k + \alpha - 2) = z^{-2}\varphi(k), \ldots, \mathbf{y}(k + \alpha - p) = z^{-p}\varphi(k)$. For a reversible MIMO nonlinear discrete system, we express the α-th order inverse system [13] as follows, $\Sigma'' : \varphi(k) \rightarrow \mathbf{u}(k)$

$$
\begin{aligned}
\mathbf{u}(k) &= F^{-1}(\varphi(k), \mathbf{y}(k+\alpha-1), \ldots, \mathbf{y}(k+\alpha-p), \\
&\quad \mathbf{u}(k-1), \ldots, \mathbf{u}(k-q)) \\
&= F^{-1}(\varphi(k), z^{-1}\varphi(k), \ldots, z^{-p}\varphi(k), \mathbf{u}(k-1), \ldots, \\
&\quad \mathbf{u}(k-q)).
\end{aligned}
\tag{3}
$$

The formula with the input $\varphi(k)$ and the output $\mathbf{u}(k)$ is the inverse expression. We cascade the inverse system and the original system to rebuild a composite system as shown in the Fig. 1. The composite system is a pseudo-linear system with the following decoupling transfer function,

$$
G_{ij}(z) = \frac{y_i(z)}{\varphi_j(z)} = \begin{cases} z^{-\alpha_i}, & i = j. \\ 0, & i \neq j. \end{cases}
\tag{4}
$$

The relationship of nonlinear coupling still exists in the composite system, but it has the standard linear relationship in view of the transfer function, namely the original MIMO system is decoupled into independent single-input single-output pseudo-linear subsystems.

3 The inverse model based on LS-SVM

3.1 LS-SVM for nonlinear function estimation

Given a training data set of M points $\{\mathbf{x}_i, y_i\}_{i=1}^M$ with input data $\mathbf{x}_i \in R^{n_1}$ and output data $y_i \in R$, one considers the following optimization problem [14] in primal space,

$$
\begin{cases} min & J_1 = \frac{1}{2}\|\mathbf{w}\|^2 + \frac{\gamma}{2}\sum_{i=1}^M e_i^2 \\ \text{s.t.} & y_i = \mathbf{w}^T \phi(\mathbf{x}_i) + b + e_i, \quad i = 1, \ldots, M. \end{cases}
\tag{5}
$$

A function $\phi : R^{n_1} \rightarrow R^{n_2}$, it maps the input space into a high dimensional (possibly infinite dimensional) feature space. A weight vector $\mathbf{w}$, a error variable $e_i \in R$ and a bias term $b \in R$ are in primal space. The cost function J consists
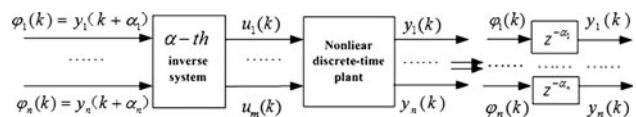

Fig. 1 A MIMO nonlinear system is decoupled into single-input single-output α-th order pseudo-linear systems

of a fitting error and a regularization term. The relative importance of these terms is determined by the positive real constant γ. In the case, a smaller γ value can avoid the over fitting of noise data. The regression model of LS-SVM is $f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}) + b$ in primal space. The weight vector may be infinite dimensional, which makes a calculation of $\mathbf{w}$ from (5) impossible in general. Therefore, one computes the model in the dual space instead of the primal space. The Lagrange function constructed for problem (5) is

$$L_1(\mathbf{w}, e_i, \beta_i, b) = J_1 + \sum_{i=1}^{M} \beta_i [y_i - \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) - b - e_i]. \quad (6)$$

$\beta_i, i = 1, \ldots, M$ is a Lagrange multiplier. According to KKT conditions [15], let the first order derivatives of L be zeros, namely

$$\frac{\partial L}{\partial \mathbf{w}} = 0; \quad \frac{\partial L}{\partial e_i} = 0; \quad \frac{\partial L}{\partial \beta_i} = 0; \quad \frac{\partial L}{\partial b} = 0. \quad (7)$$

The following equations can be acquired

$$\mathbf{w} = \sum_{i=1}^{M} \beta_i \phi(\mathbf{x}_i); \quad \sum_{i=1}^{M} \beta_i = 0;$$
$$\beta_i = \gamma e_i; \quad \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) + b + e_i - y_i = 0. \quad (8)$$

According to Mercer kernel conditions, one can choose a kernel $K(\cdot, \cdot)$, such that $K(\mathbf{x}_1, \mathbf{x}_2) = \phi(\mathbf{x}_1)^{\mathrm{T}}\phi(\mathbf{x}_2)$. After elimination of $\mathbf{w}$ and e_i, the optimization problem leads to the following linear system,

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & K(\mathbf{x}_1, \mathbf{x}_1) + \frac{1}{\gamma} & \cdots & K(\mathbf{x}_1, \mathbf{x}_M) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & K(\mathbf{x}_M, \mathbf{x}_1) & \cdots & K(\mathbf{x}_M, \mathbf{x}_M) + \frac{1}{\gamma} \end{bmatrix} \begin{bmatrix} b \\ \beta_1 \\ \vdots \\ \beta_M \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_M \end{bmatrix}. \quad (9)$$

The linear equation can be rewritten as:

$$\begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{I}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}, \quad (10)$$

where $\mathbf{A} = \mathbf{K} + \boldsymbol{\theta}, \mathbf{K} = (K_{ij})_{M \times M}, \boldsymbol{\theta} = diag(\frac{1}{\gamma}, \ldots, \frac{1}{\gamma}), \mathbf{I} = (1; 1; \ldots; 1), \boldsymbol{\beta} = (\beta_1; \beta_2; \ldots; \beta_M)$, and $\mathbf{y} = (y_1; y_2; \ldots; y_M)$. We focus on an RBF kernel with parameters γ and σ^2 in this paper. β_i and b are the solution to the linear equations. The LS-SVM model at $\mathbf{x}$ becomes

$$y(\mathbf{x}) = \sum_{i=1}^{M} \beta_i K(\mathbf{x}, \mathbf{x}_i) + b. \quad (11)$$

Errors in optimization problem of LS-SVM only measures noises of output variables which use errors

between expected outputs and predictions as the empirical errors, and then minimize the sum of square errors. But the noises of input variables still exist. So we describe the empirical errors [12] in the features pace as follows.

$$\frac{y_i - \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) + b}{\sqrt{1 + \| \mathbf{w} \|^2}} = e_i, \quad i = 1, \ldots, M. \quad (12)$$

The revised optimization problem in primal space is:

$$min J_2 = \frac{1}{2}\|\mathbf{w}\|^2 + \frac{\gamma}{2}\sum_{i=1}^{M} e_i^2; \quad (13)$$
$$\text{s.t.} (y_i - \mathbf{w}^{\mathrm{T}}\phi(\mathbf{x}_i) + b)/\sqrt{1 + \| \mathbf{w} \|^2} = e_i.$$

We discuss an iterative learning algorithm to solve the optimization problem in the formula (13) in the dual space. At the $t - th$ iteration, the Lagrange function is also rewritten as:

$$L_2^{(t)} = \frac{1}{2}\|\mathbf{w}^{(t)}\|^2 + \frac{\gamma}{2}\sum_{i=1}^{M} (e^{(t)})_i^2$$
$$+ \sum_{i=1}^{M} \beta_i^{(t)} \left(\frac{y_i - (\mathbf{w}^{(t)})^{\mathrm{T}}\phi(\mathbf{x}_i) - b^{(t)}}{\sqrt{1 + \| \mathbf{w}^{(t-1)} \|^2}} - e_i^{(t)} \right). \quad (14)$$

According to KKT conditions, partial differentiation of the Lagrange function is the following,

$$\frac{\partial L_2^{(t)}}{\partial \mathbf{w}^{(t)}} = 0 \Rightarrow \mathbf{w}^{(t)} = \sum_{i=1}^{M} \frac{\beta_i^{(t)}}{\sqrt{1 + \| \mathbf{w}^{(t-1)} \|^2}} \phi(\mathbf{x}_i);$$

$$\frac{\partial L_2^{(t)}}{\partial e_i^{(t)}} = 0 \Rightarrow e_i^{(t)} = \frac{1}{\gamma}\beta_i^{(t)};$$

$$\frac{\partial L_2^{(t)}}{\partial \beta_i^{(t)}} = 0 \Rightarrow \frac{y_i - (\mathbf{w}^{(t)})^{\mathrm{T}}\phi(\mathbf{x}_i) - b^{(t)}}{\sqrt{1 + \| \mathbf{w}^{(t-1)} \|^2}} = e_i^{(t)}; \quad (15)$$

$$\frac{\partial L_2^{(t)}}{\partial b^{(t)}} = 0 \Rightarrow \sum_{i=1}^{M} \frac{\beta_i^{(t)}}{\sqrt{1 + \| \mathbf{w}^{(t-1)} \|^2}} = 0.$$

We have the equation

$$\sum_{i=1}^{M} \frac{\beta_i^{(t)} K(\mathbf{x}_i, \mathbf{x}_j)}{\sqrt{1 + \| \mathbf{w}^{(t-1)} \|^2}} + \frac{1}{\gamma}\beta_i^{(t)}\sqrt{1 + \| \mathbf{w}^{(t-1)} \|^2} + b^{(t)} = y_i; \quad (16)$$

Define $\alpha_i^{(t)} = \frac{\beta_i^{(t)}}{\sqrt{1 + \|\mathbf{w}^{(t-1)}\|^2}}, i = 1, \ldots, M$, we have equations

$$\sum_{i=1}^{M} \alpha_i^{(t)} K(\mathbf{x}_i, \mathbf{x}_j) + \frac{1}{\gamma}\alpha_i^{(t)}(1 + \| \mathbf{w}^{(t-1)} \|^2) + b^{(t)} = y_i; \quad (17)$$

$$\sum_{i=1}^{M} \alpha_i^{(t)} = 0. \quad (18)$$

So we get the linear equation:

$$\begin{bmatrix} \mathbf{A}+ \parallel \mathbf{w}^{(t-1)} \parallel^2 \boldsymbol{\theta} & \mathbf{I} \\ \mathbf{I}^{\mathrm{T}} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}^{(t)} \\ b^{(t)} \end{bmatrix} = \begin{bmatrix} \mathbf{y} \\ 0 \end{bmatrix}. \tag{19}$$

where still $\mathbf{A} = \mathbf{K} + \boldsymbol{\theta}, \boldsymbol{\alpha}^{(t)} = (\alpha_1^{(t)}; \alpha_2^{(t)}; \ldots; \alpha_M^{(t)})$. In order to solve the Eq. (19), we need to update $\boldsymbol{\alpha}^{(t)}, b^{(t)}$ and $\parallel \mathbf{w}^{(t-1)} \parallel^2$ to find the solution of the optimization problem (13).

Lemma *Let* $\mathbf{A}$ *be an invertible matrix, for the given matrix* $\mathbf{A}$ *and* $\mathbf{U}, \mathbf{V}, \mathbf{D}$, *define* $\mathbf{B} = \mathbf{D} - \mathbf{V}\mathbf{A}^{-1}\mathbf{U}$, *then the inverse matrix*

$$\begin{bmatrix} \mathbf{A} & \mathbf{U} \\ \mathbf{V} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{U}\mathbf{B}^{-1}\mathbf{V}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{U}\mathbf{B}^{-1} \\ -\mathbf{B}^{-1}\mathbf{V}\mathbf{A}^{-1} & \mathbf{B}^{-1} \end{bmatrix} \tag{20}$$

Specially, if $\mathbf{U} = \mathbf{I}$ and $\mathbf{U} = \mathbf{V}^{\mathrm{T}}$ hold, it has

$$\begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{I}^{\mathrm{T}} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{I}\mathbf{I}^{\mathrm{T}}\mathbf{A}^{-1}}{\mathbf{I}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{I}} & \frac{\mathbf{A}^{-1}\mathbf{I}}{\mathbf{I}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{I}} \\ \frac{\mathbf{I}^{\mathrm{T}}\mathbf{A}^{-1}}{\mathbf{I}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{I}} & \frac{-1}{\mathbf{I}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{I}} \end{bmatrix} \tag{21}$$

The original matrix in LS-SVM linear equation is invertible and the revised matrix still is invertible due to the same as rank of the original matrix. We use special symbols respectively to express the original matrix and the revised matrix:

$$\Psi = \begin{bmatrix} \mathbf{A} & \mathbf{I} \\ \mathbf{I}^{\mathrm{T}} & 0 \end{bmatrix}, \tag{22}$$

$$\Psi^{(t)} = \begin{bmatrix} \mathbf{A}+ \parallel \mathbf{w}^{(t-1)} \parallel^2 \boldsymbol{\theta} & \mathbf{I} \\ \mathbf{I}^{\mathrm{T}} & 0 \end{bmatrix}. \tag{23}$$

According to the linear Eq. (19), our motivation is to get $(\Psi^{(t)})^{-1}$. From the lemma and the formula (21), if we hope to get $(\Psi^{(t)})^{-1}$, we need $(\mathbf{A}+ \parallel \mathbf{w}^{(t-1)} \parallel^2 \boldsymbol{\theta})^{-1}$. Since $\mathbf{A}$ is symmetric and positive definite, there exists an orthogonal matrix $\mathbf{P}$ and $\mathbf{P}^{\mathrm{T}} = \mathbf{P}^{-1}$, so that $\mathbf{A} = \mathbf{P}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{P}$, where $\mathbf{\Lambda} = diag(\lambda_1, \lambda_2, \ldots, \lambda_M)$ and $\lambda_1, \lambda_2, \ldots, \lambda_M$ are the positive eigenvalues of $\mathbf{A}$. So we have

$$\mathbf{A}+ \parallel \mathbf{w}^{(t-1)} \parallel^2 \boldsymbol{\theta}$$
$$= \mathbf{P}^{\mathrm{T}} diag\left\{ \lambda_1 + \frac{\parallel \mathbf{w}^{(t-1)} \parallel^2}{\gamma}, \ldots, \lambda_M + \frac{\parallel \mathbf{w}^{(t-1)} \parallel^2}{\gamma} \right\} \mathbf{P} \tag{24}$$

$$(\mathbf{A}+ \parallel \mathbf{w}^{(t-1)} \parallel^2 \boldsymbol{\theta})^{-1}$$
$$= \mathbf{P}^{\mathrm{T}} diag\left\{ \frac{1}{\lambda_1} + \frac{\gamma}{\parallel \mathbf{w}^{(t-1)} \parallel^2}, \ldots, \frac{1}{\lambda_M} + \frac{\gamma}{\parallel \mathbf{w}^{(t-1)} \parallel^2} \right\} \mathbf{P} \tag{25}$$

We derive the inverse of $\mathbf{A}+ \parallel \mathbf{w}^{(t-1)} \parallel^2 \boldsymbol{\theta}$, then update $(\Psi^{(t)})^{-1}$ by using the formula (21), and find the solution of (19). The iterative updating algorithm can be summarized as follows.

Algorithm 1 Iterative updating learning algorithm for solving the novel LS-SVM.

1. Set parameters of LS-SVM. Find the orthogonal matrix $\mathbf{P}$ and the diagonal matrix $\mathbf{\Lambda}$, so that $\mathbf{A} = \mathbf{P}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{P}^{-1}$. Store $\mathbf{P}$ and $\mathbf{\Lambda}$.

2. Computer the inverse of Ψ using the formula (21) and solve the problem (10). Set the solution of LS-SVM as $\boldsymbol{\beta}^0, b^0$. Let $t = 1$ and $\boldsymbol{\alpha}^0 = \boldsymbol{\beta}^0$ and computer $\parallel \mathbf{w}^{(0)} \parallel^2 = (\alpha^{(0)})^{\mathrm{T}}\mathbf{K}\alpha^{(0)}$.

3. Computer $(\mathbf{A}+ \parallel \mathbf{w}^{(t-1)} \parallel^2 \boldsymbol{\theta})^{-1}$ using the formula (25), then $(\Psi^{(t)})^{-1}$ can be computed using (21).

4. The solution of the novel LS-SVM in the formula (13) can be obtained by multiplying $(\Psi^{(t)})^{-1}$. Record the solution of the Eq. (13) as $\boldsymbol{\alpha}^{(t)}$ and $b^{(t)}$, and computer $\parallel \mathbf{w}^{(t)} \parallel^2 = (\boldsymbol{\alpha}^{(t)})^{\mathrm{T}}\mathbf{K}\boldsymbol{\alpha}^{(t)}$.

5. If the stop condition $\eta = \parallel \frac{\sqrt{\parallel \mathbf{w}^{(t)} \parallel^2}}{\parallel \mathbf{w}^{(t)} \parallel^2} - \frac{\sqrt{\parallel \mathbf{w}^{(t-1)} \parallel^2}}{\parallel \mathbf{w}^{(t-1)} \parallel^2} \parallel < \zeta$ holds for a positive number ζ, go to 6; Otherwise, set $t = t + 1$, go to 3.

6. Let $\beta_i = \frac{\alpha_i^{(t)}}{\sqrt{1 + \parallel \mathbf{w}^{(t-1)} \parallel^2}}$ and $b = b^{(t)}$. The output of the novel LS-SVM is $y(\mathbf{x}) = \sum_{i=1}^M \beta_i K(\mathbf{x}, \mathbf{x}_i) + b$.

3.2 α-th order inverse model based on LS-SVM

For the described nonlinear discrete system, the α-th order inverse system is expressed in the formula (3). Both the precise mathematic model of the original system and the explicit expression of $\mathbf{u}(k)$ cannot be obtained. For the imperfect model, we adapt the novel LS-SVM to approximate the inverse model based on the input-output data acquired from the original system. Because LS-SVM can only be used for the estimation of single output functions, in order to identify multiple output objects, it is necessary to learn respectively for each subsystem. The number of subsystems is equal to the number of output variables.

Algorithm 2 Inverse model approaching algorithm based on the novel LS-SVM.

1. Select a proper excitation signal, such as the white noise, etc.

2. Obtain input-output data of original system by using the excitation signal as the input. Sort data into training and testing samples in the form of $\{S_i, u_i\}_{i=1}^M$.

3. Use parameters γ and σ^2 and train the novel LS-SVM to acquire the inverse submodel of each decoupling subsystem.

4. Test the generalization ability of inverse models using testing data.

5. Assemble all inverse submodels to get the inverse

model of original system due to decoupling subsystems.

After acquiring the so-called inverse system of the original MIMO system, cascade the inverse model and the original system to constitute the pseudo-linear system which is a simple open loop control based on inverse system.

4 The internal model control of the MIMO nonlinear discrete system

The internal model control for discrete processes has the following properties [16].

Property 1   Stability Criterion. When the internal model is exact, stability of both controller and plant is sufficient for overall system stability.

Property 2   Perfect Controller. Under the assumption that the internal model is perfect and that the plant and the controller is stable, if there is no disturbance, the perfect control can be achieved when the controller is the inverse of internal model.

The pseudo-linear system acquired by connecting with the inverse model and the original system has basically linearization, the basic diagram of internal model control based on LS-SVM is shown in the Fig. 2. $G_m(z)$ represents the pulse transfer function matrix of internal model, $G(z)$ represents the controlled plant, $G_c(z)$ represents the pulse transfer function matrix of internal model controller, $D(z)$ is the disturbance function, $R(z)$ is the input function, $Y(z)$ is the output function. The internal model control strategy provides a feedback control for nonlinear systems. One usually just chooses a diagonal matrix constituted by relative orders of independent subsystems as the transfer function of the internal model, namely $G_m(z) = diag\{z^{-\alpha_1}, z^{-\alpha_2}, \ldots, z^{-\alpha_n}\}$. Considering that the actual composite system $G(z)$ maybe have an error in modeling, the pseudo-linear system can be assumed $G(z) = G_m(z)(1 + h_m(z))$, $h_m(z)$ expresses the unmodeled error function. We assume that $h_m(z)$ is linear and
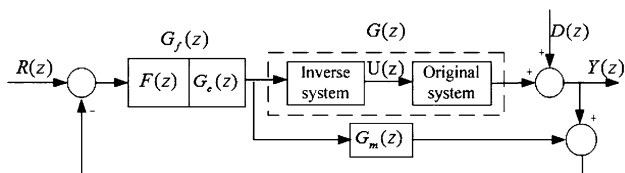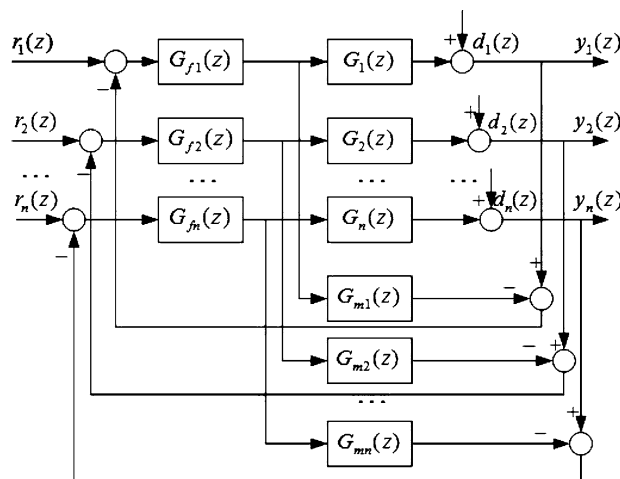

Fig. 3 The MIMO internal model control system

bounded. The MIMO internal model control is illustrated by the Fig. 3.

The internal model controller denotes as $G_f(z)$ that is the product of a robust filter $F(z)$ and $G_c(z)$. Still let $G_c(z) = G_m^{-1}(z)$. The robust filter $F(z)$ is usually to reduce the sensitivity of the internal model control system. Reference [16] offers a detailed introduction about how to design the robust filter. The internal model controller can be rewritten as $G_f(z) = F(z)G_m^{-1}(z)$. According to Property 1, which requires that the object and the controller are input-output stable to make sure the control system is stable. For a decoupling linear system, $G_m(z) = diag\{z^{-\alpha_1}, z^{-\alpha_2}, \ldots, z^{-\alpha_n}\}$, we ask $G_m^{-1}(z) = diag\{1, 1, \ldots, 1\}$ in order to keep the controller stable. Using the following filter

$$F(z) = diag\left\{\frac{1 - l_1}{1 - l_1 z^{-\alpha_1}}, \ldots, \frac{1 - l_n}{1 - l_n z^{-\alpha_n}}\right\}, \quad (26)$$

the output of the closed-loop system can be described as

$$Y(z) = \frac{G_f(z)G(z)R(z) + (1 - G_f(z)G_m(z))D(z)}{1 + G_f(z)(G(z) - G_m(z))}, \quad (27)$$

and the error is

$$\begin{aligned} E(z) &= Y(z) - R(z) \\ &= \frac{(G_f(z)G_m(z) - 1)R(z) + (1 - G_f(z)G_m(z))D(z)}{1 + G_f(z)(G(z) - G_m(z))}. \end{aligned} \quad (28)$$

5 Simulation

In this section, aiming at the multivariate, nonlinear and strong coupling plant, we illustrate the performance of internal model control based on LS-SVM. The discrete

Fig. 2 The diagram of internal model control based on LS-SVM

model in the simulation is as follows,

$$y_1(k) = \frac{0.6y_1(k-1)\sin(y_1(k-2))}{1+y_1^2(k-1)+y_2^2(k-2)}$$
$$+ 0.3u_1(k-2) + u_1(k-1) + 0.2u_2(k-2),$$
$$y_2(k) = \frac{0.5y_2(k-1)\cos(y_2(k-2))}{1+y_2^2(k-1)+y_1^2(k-2)}$$
$$+ 0.4u_2(k-2) + u_2(k-1) + 0.5u_1(k-2).$$
$$(29)$$

Suppose that the precise mathematic model of original system is unknown and it is reversible. $\alpha_1 = 1$, $\alpha_2 = 1$, $m = 2$, $n = 2, p_1 = 1, p_2 = 1, q_1 = 2, q_2 = 2$.

5.1 The internal model control and the open loop control based on inverse system

Give white noise sequences to the two input ends and the above model is used to produce data of 1000 groups. Utilize 500 groups to train and the other 500 groups to test. Fitting factors of every group are $S_1 = \{y_1(k), y_1(k-1),\ y_1(k-2), u_1(k-2), y_2(k-2), u_2(k-2)\}$ and $S_2 = \{y_2(k),\ y_2(k-1), y_2(k-2), u_2(k-2), y_1(k-2), u_1\ (k-2)\}$ respectively. With RBF kernel function, $\gamma = 1000$ and $\sigma^2 = 30$ are selected. We can obtain inverse submodels by the novel LS-SVM. The index of root mean square error is denoted as $RSME = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}$. RSMEs of inverse submodels for testing data equal 0.136 and 0.096. The Fig. 4 shows testing curves and error curves of the inverse models approximated.

Due to a decoupling linear system, assemble all inverse submodels to get the inverse model of original system.
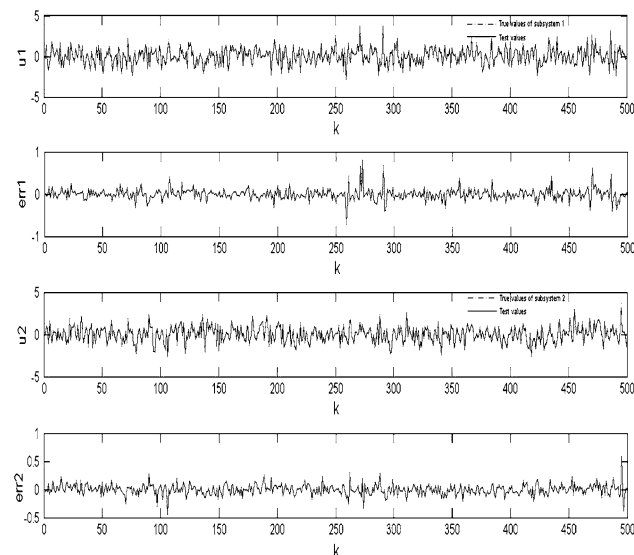

Fig. 4 *Testing curve* and *error curve* of the inverse model approximated by the novel LS-SVM

As the relative order of the system $\alpha_1 = \alpha_2 = 1$, the internal model $G_m(z)$ takes as $diag\{\frac{1}{z}, \frac{1}{z}\}$. The simple controller is $G_f(z) = G_m^{-1}(z) = diag\{1, \ldots, 1\}$. $G_f(z) = F(z)G_m^{-1}(z)$ is a internal model control with the filter $F(z) = diag\{\frac{1-l_i}{1-l_i z^{-1}}\}$, $0 \le l_i \le 1, i = 1, 2$. Set $l_1 = l_2 = 0.5$ in the simulation.

By cascading the inverse system, the multivariable coupling system has been decoupled into two pseudo-linear systems. The performance of open-loop system based on inverse control under a given square-wave reference is shown in the Fig. 5. Comparing with the open-loop system only based on inverse control under the same reference, the internal model control system achieves a good tracking to a square-wave. In order to reduce jitter, we introduced a filter into the simple internal model control system. The performance of open-loop system based on inverse control is bigger errors and jitter than that of the internal model control system according to tracking curves from trajectories shown in the figure. We use a mixture signal of sin waves with different frequency as a reference signal again. The tracking performances of the open-loop system and the internal model control system for the reference signal are shown in the Fig. 6, which illustrate a unit delay tracking for the reference signal. The internal model control system has better tracking performance than that of the open-loop system.
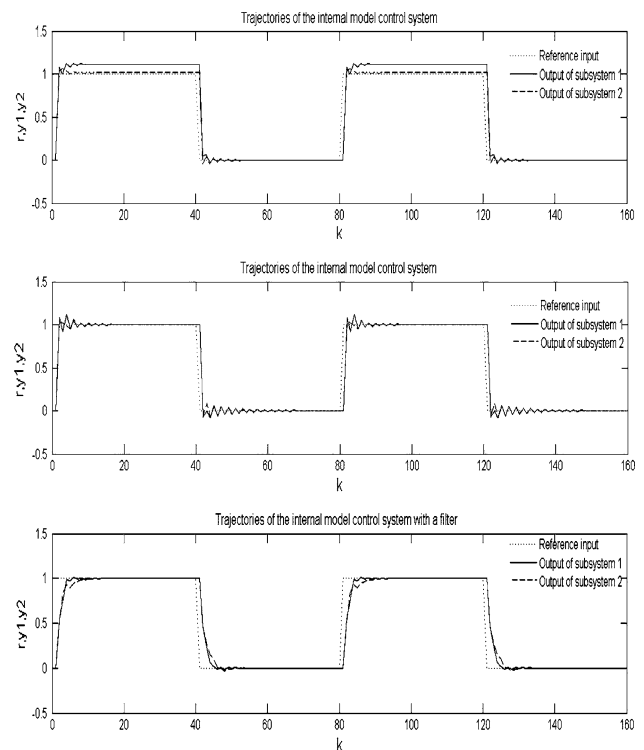

Fig. 5 Systems trajectories under a given *square-wave* reference. The first subfigure is the curves of open-loop control, the second one is the curves of simple internal model control, and the third one is the curves of internal model control with a filter
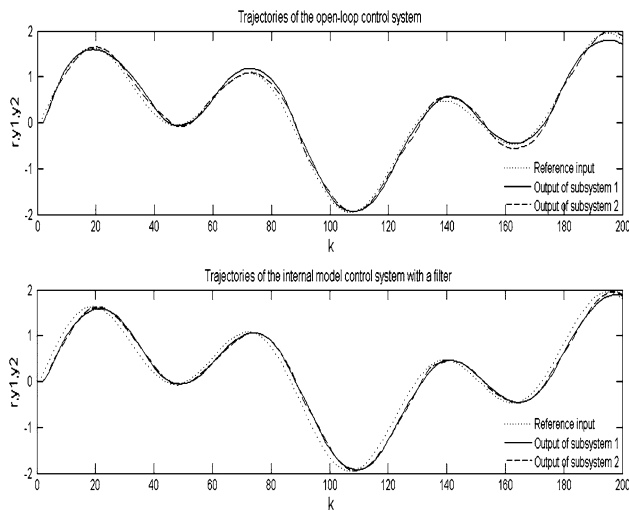
Fig. 6 Systems trajectories under a mixture reference signal. The first one is the curves of open-loop control, the second one is the curves of internal model control with a filter

5.2 Robustness of disturbance rejection

At $k = 100$ and $k = 110$, the two decoupling subsystems are interfered respectively by an external step signal with the amplitude of 0.1. The square-wave response of internal model control system can be observed from the Fig. 7, the strategy of
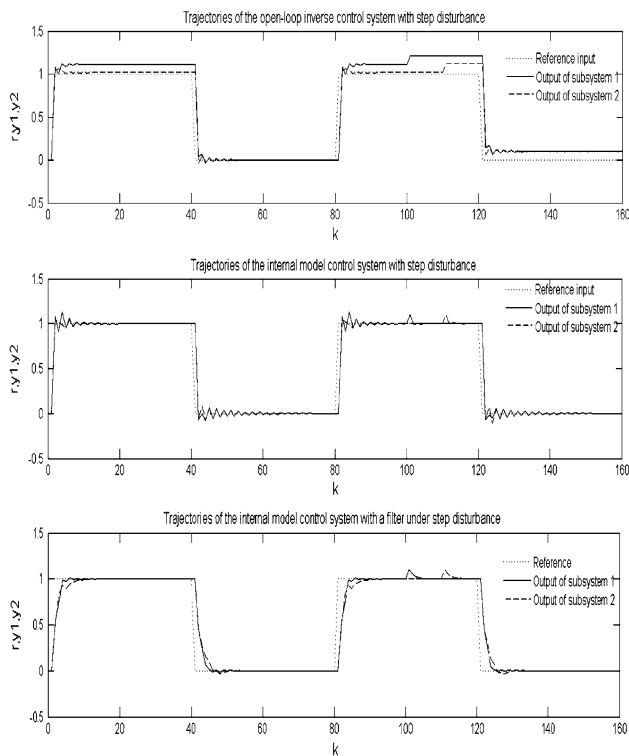
simple internal model control has good robustness to inhibit a step disturbance and keep following the reference signal after a little jitter, but the duration of the jitter is a litter longer than the time excepted. After adding a filter, the duration of the jitter is reduced. When the open loop system is disturbed by the same step signal, the system can not follow the reference signal and deviates from the reference signal greatly, which shows weak robustness. Disturbances lead to bad steady-state errors in the open-loop system.

5.3 Robustness of variable parameters

Parameters of the nonlinear system change, namely the original nonlinear system changes into the following formula,

$$
\begin{aligned}
y_1(k) &= \frac{0.8 y_1(k-1) \sin(y_1(k-2))}{1 + y_1^2(k-1) + y_2^2(k-2)} \\
&\quad + 0.6 u_1(k-2) + u_1(k-1) + 0.6 u_2(k-2), \\
y_2(k) &= \frac{0.6 y_2(k-1) \cos(y_2(k-2))}{1 + y_2^2(k-1) + y_1^2(k-1)} \\
&\quad + 0.6 u_2(k-2) + u_2(k-1) + 0.6 u_1(k-2).
\end{aligned}
$$
(30)

When parameters of nonlinear system vary as described in (30), it is equal to the mismatch between the object and the

Fig. 7 Systems trajectories with external step disturbance. The first subfigure is the curves of open-loop control, the second one is the curves of simple internal model control, and the third one is the curves of internal model control with a filter
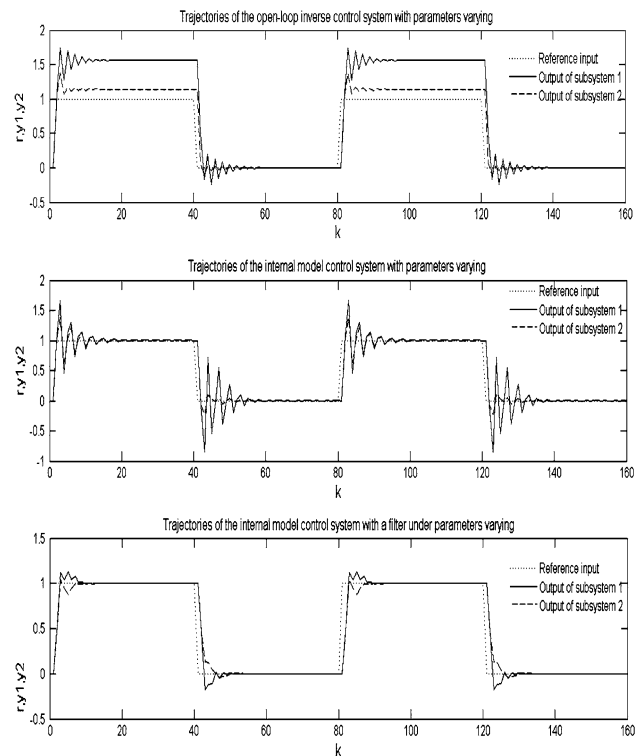

Fig. 8 Systems trajectories with parameters varying. The first subfigure is the curves of open-loop control, the second one is the curves of simple internal model control, and the third one is the curves of internal model control with a filter

model. Simulation results are shown in the Fig. 8. The simple internal model control can still achieve tracking the reference, but a large jitter exists. Under a filter, the jitter is reduced. But the disturbance of parameters varying induces a larger oscillations for the open loop system.

6 Conclusion

In this study, we firstly introduce the novel LS-SVM which considers noises of input variables and output variables, then the internal model control based on the novel LS-SVM for MIMO nonlinear discrete systems is presented. The proposed method overcomes the problem of the imperfect mathematic model of original system and identifies the inverse model accurately by way of input-output data. The advantage of internal model control is the excellent robustness with respect to a disturbance signal and a model mismatch, which is illustrated in the simulation.

References

1. Lightbody G, Irwin GW (1997) Nonlinear control structures based on embedded neural system models. IEEE Trans Neural Netw 8(3):553–567
2. Rivals I, Personnaz L (2000) Nonlinear internal model control using neural networks: application to processes with delay and design issues. IEEE Trans Neural Netw 11(1):80–90
3. Li M, Zhou ZK, Shi LL (2004) The neural network decoupling control based on the internal model control. In: Proceedings of the 5th world congress on intelligent control and automation, Hangzhou, China, pp 2643–2646
4. Haber RE, Alique JR (2004) Nonlinear internal model control using neural networks: an application for machining processes. Neural Comput Appl 13:47–55
5. Chidrawar SK, Patre BM (2008) Implementation of neural network for internal model control and adaptive control. In: Proceedings of the international conference on computer and communication engineering, Kuala Lumpur, Malaysia, pp 741–746
6. Cui WT, Yan XF (2009) Adaptive weighted least square support vector machine regression integrated with outlier detection and its application in QSAR. Chemom Intell Lab Syst 98:130–135
7. Vapnik VN (1998) Statistical learning theory. Wiley, NewYork
8. Evgeniou T, Pontil M, Poggio T (2000) Regularization networks and support vector machiness. Adv Comput Math 13(1):1–50
9. Suykens JAK, Vandewalle J (1999) Least squares support vector machines classifiers. Neural Process Lett 9(3):293–300
10. Suykens, JAK, Lukas L, Vandewalle J (2000) Sparse approximation using least-squares support vector machines. In: Proceedings of the IEEE international symposium on circuits and systems, Geneva, Switzerland, pp 11757–11760
11. Markovsky I, VanHuffel S (2007) Overview of total least squares methods. Signal Process 87:2283–2302
12. Peng XJ, Wang YF (2009) A normal least squares support vector machine (NLS-SVM) and its learning algorithm. Neurocomputing 72:3734–3741
13. Dai XZ (2005) Inverse control method based on neural networks for multi-variables nonlinear system (in Chinese). Science Press, Beijing
14. Sun CY, Mu CX, Li XM (2009) A weighted LS-SVM approach for the identification of a class of nonlinear inverse systems. Sci China Ser F Inf Sci 52(5):770–779
15. Suykens JAK, Brabanter JD, Lukas L, Vandewalle J (2002) Weighted least squares support vector machiness: robustness and sparse approximation. Neurocomputing 48:85–105
16. Garcia CE, Morari M (1982) Internal model control: a unifying review and some new results. Ind Eng Chem Process Des 21(2):308–323