TELKOMNIKA, Vol. 11, No. 12, December 2013, pp. 7605~7610
e-ISSN: 2087-278X

A Novel Sensor Fault Detection Method

Chen Yongqi*[1], Zhao Yiming[2]
College of Science and Technology, Ningbo University, CHINA
*Corresponding author, e-mail: lingfen7781@163.com[1], zhaoyiming@nbu.edu.cn[2]

Abstract

To detect the sensor faults of the plant with min-disturbance, a weighted least squares interval regression model is proposed. The output of the proposed model is an interval band which can resist the disturbance influence and give correct sensor fault alarm. Additionally, the time complexity of this model is low because only a set of linear equations can determine the parameters. The experiments of fault instance demonstrate the feasibility and effectiveness of the interval regression model.

Keywords: *fault Detection, least squares, interval regression;*

1. Introduction

To detect the sensor faults, one can obtain the regression model of the plant without faults. The sensor faults can be detected by comparing the estimated outputs of the regression model and the observed outputs of the plant. Support vector regression (SVR), which owns high generalization performance, is an effective method to construct this regression model [1]. But, when the system is a plant with min-disturbance, outputs will locate in an interval band. The regression model based on traditional SVR only presents crisp outputs and can not describe the effect of min-disturbance in plant whose output is an interval band. For this reason, traditional SVR is not fit to detect the sensor faults of the plant with min-disturbance. It is likely to regard the disturbance as sensor faults and give a wrong alarm. In order to resolve this problem, interval regression model whose regression outputs is an interval band model must be proposed to detect the sensor faults of the plant with min-disturbance.

For these years, many interval regression models are proposed. Support vector interval regression networks (SVIRNs) is presented. This model utilizes two radial basis function networks to identify the upper and lower sides of the data interval [2]. Support vector interval regression machine (SVIRM) is designed for crisp input and output data [3]. SVIRM is robust in the sense that outliers do not affect the resulting interval regression. ν-support vector interval regression networks are proposed to evaluate interval linear and nonlinear regression models for crisp input and output data [4]. Experimental results manifest that ν-support vector interval regression networks is useful in practice, especially when noise is heteroscedastic. However, like SVM with inequality constraints, the weight vectors and the bias term are worked out by a complicated quadratic programming problem. Though by modifying the solution, the time complexity of SVMR based on sequential minimal optimization (SMO) algorithm is high[5]. Due to this, in this study, a novel weighted least squares interval regression (WLS_IR) is proposed by applying the fuzzy set principle to weight vectors for the purpose of estimating the interval of imprecise observations. Compared with the existing support vector fuzzy regression models, only a set of linear equations are needed to determine the weight vector and bias term of WLS_IR. Consequently, WLS_IR owns the advantage of low time complexity. Additionally, the bounds of the interval regression model are influenced by outliers in training data. In this paper, the reweighting scheme [6] is introduced to resist the influence of outliers.

The rest of this paper is organized as the following. least squares support vector machines is introduced in section 2. WLS_IR is studied in Section 3. Section 4 introduces the reweighting scheme to resist outliers influence. Fault instance Experiment is presented in Section 6. Section 7 puts forward the concluding remarks.

Received June 29, 2013; Revised August 17, 2013; Accepted September 4, 2013

2. Least Squares Support Vector Machines

In this section, least squares support vector machines（LS_SVM） is briefly introduced [7]. Given training data set $\{X_i, Y_i\},\ i = 1,\cdots,n$ LS-SVM supposes the Hyperplanes as the following:

$$y_i = \omega^T \phi(\mathbf{x}_i) + b \tag{1}$$

Where $\mathbf{x}_i$, y_i are input variable and output variable, $\phi(\mathbf{x}_i)$ is a nonlinear function which maps the feature space of input into a higher dimension feature space and can be reached by the kernel strategy. ω is a coefficient determining the margin of support vectors and b is a bias term. The coefficients (ω, b) are determined by minimizing the following regularized risk function and using the equality constraints.

$$\min_{\omega,b} J = \frac{1}{2}\omega^T\omega + \frac{1}{2}c\sum_{i=1}^{n} e_i^2 \qquad (c > 0) \tag{2}$$

$$\text{s.t. } y_i = w^T \varphi(x_i) + b + e_i\ ,\ i = 1\cdots n$$

Where e_i is the error variable and used to construct a soft margin hyper plane. In Equation (2), the first term, measures the inverse of the margin distance. In order to obtain the minimum structural risk, the first term should be minimized. c is the regularization parameter determining the fitting error minimization and smoothness.

Finally, the decision function of the classifier of LS_SVM can be expressed as following:

$$f(\mathbf{x}) = \sum_{i=1}^{n} a_i \phi(\mathbf{x}_i)\phi(\mathbf{x}) + b \tag{3}$$

The functional form of $\phi(\mathbf{x}_i)$ need not to be known since it is defined by the kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^t \phi(\mathbf{x}_j),\ i = 1\cdots n,\ j = 1\cdots n$. Different kernel functions present different mappings from the input space to the high dimension feature space. The commonly used kernels for regression problem are given as follows:

Linear kernel: $K(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t\mathbf{y}$

Polynomial kernel: $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^t\mathbf{y} + 1)^d$

RBF kernel: $K(\mathbf{x}, \mathbf{y}) = \exp(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\sigma^2})$

Due to the equality constraints in the formulation, LS_SVM is solved by a set of linear equalities instead of a complicated quadratic programming problem. For this reason, LS_SVM is a low computational complexity method. But the regression output of LS_SVM in only a crisp data. When Available information is uncertain and imprecise, LS_SVM can not solve this problem. For this reason, Weighted Least Squares interval Regression is proposed based the theory of LS_SVM and the interval regression of Tanaka. This interval regression model is presented as follows:

3. The Design of WLS_IR

In this section, a novel WLS_IR model is proposed based on LS_SVM and the interval regression theory. This model can be resolved by a set of linear equations as against the complicated quadratic programming. the interval regression problem is to find the parameters $(\boldsymbol{\omega}, \mathbf{c}, b)$ which is the solution of the objection function as follows:

$$\min J(w,\xi)=\frac{1}{2}[w^T w+(c^T c+b^2)]$$
$$\frac{\gamma}{2}\sum_{i=1}^{n}(v_{1i}\xi_{1i}^2+v_{2i}\xi_{2i}^2) \tag{4}$$

Subject to:

$$\begin{cases} w\phi(x_i)+b+(1-H)(c\phi(|x_i|))+\xi_{1i} \\ =y_i+(1-H)e_i \\ -(w\phi(x_i)+b)+(1-H)(c\phi(|x_i|))+\xi_{2i} \\ =-y_i-(1-H)e_i\ ,\quad i=1,2,\cdots n, \end{cases} \tag{5}$$

Where ξ_{1i} and ξ_{2i} are slack variables. This optimization problem, including the constraints, can be solved by the Lagrange function as follows:

$$\begin{aligned} &L(w,\alpha_{1i},\alpha_{2i},b,c,\xi_{1i},\xi_{2i}) \\ &=\frac{1}{2}[w^T w+(c^T c+b^2)]+\frac{\gamma}{2}\sum_{i=1}^{n}(\xi_{1i}^2+\xi_{2i}^2) \\ &-\sum_{i=1}^{n}\alpha_{1i}[(w\phi(x_i)+b)+(1-H)((c\phi|x_i|)+b) \\ &-y_i-(1-H)e_i+\xi_{1i}]-\sum_{i=1}^{n}\alpha_{2i}[-(w\phi(x_i)+b) \\ &+(1-H)(c\phi(|x_i|)+b)+y_i+(1-H)e_i+\xi_{2i}] \end{aligned} \tag{5}$$

Where α_{1i},α_{2i} are Lagrange multipliers. Computing the partial derivatives of (5), one can derive:

$$\begin{aligned} &\frac{\partial L}{\partial w}=0\Rightarrow w=\sum_{i=1}^{n}(\alpha_{1i}-\alpha_{2i})\phi(x_i) \\ &\frac{\partial L}{\partial c}=0\Rightarrow c=\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i})(1-H)\phi(|x_i|) \\ &\frac{\partial L}{\partial b}=0\Rightarrow -b=(2-H)\sum_{i=1}^{n}\alpha_{1i}-H\sum_{i=1}^{n}\alpha_{2i} \\ &\frac{\partial L}{\partial \alpha_{1i}}=0\Rightarrow [w\phi(x_i)+b]+(1-H)(c\phi|x_i|+b) \\ &-y_i-(1-H)e_i+\xi_{1i}=0 \\ &\frac{\partial L}{\partial \alpha_{2i}}=0\Rightarrow -[w\phi(x_i)+b]+(1-H)(c\phi|x_i|+b) \\ &+y_i+(1-H)e_i+\xi_{2i}=0 \end{aligned} \tag{6}$$

As mentioned by Vapnic [1], the map function ϕ does not need to be known since it is defined by the choice of kernel function. For this reason, two kernel functions, $k(x_i,x_j)=\phi(x_i)^T\phi(x_j)$ and $k(|x_i|,|x_j|)=\phi(|x_i|)^T\phi(|x_j|)$, are used to replace ϕ. Lagrange multipliers α_{1i},α_{2i} and bias term b can be obtained. Then, the upper bound and lower bound of LS_SVFR are derived as follows:

$$
\begin{aligned}
&F\_up(x) = \langle w \cdot \phi(\bar{x}) \rangle + \langle c \cdot \phi(|\bar{x}|) \rangle \\
&= \sum_{i=1}^{n} (\alpha_{1i} - \alpha_{2i}) k(x_i, x) + (\alpha_{1i} + \alpha_{2i}) \\
&- \frac{(4 - 2 \cdot H)}{\lambda} \sum_{i=1}^{n} \alpha_{1i} + \frac{2 \cdot H}{\lambda} \sum_{i=1}^{n} \alpha_{2i} \\
&+ \sum_{i=1}^{n} (\alpha_{1i} + \alpha_{2i})(1 - H) k(|x_i|, |x|) \\
&- 2[(2 - H) \sum_{i=1}^{n} \alpha_{1i} - H \sum_{i=1}^{n} \alpha_{2i}] \\
&F\_down(x) = \langle w \cdot \phi(\bar{x}) \rangle - \langle c \cdot \phi(|\bar{x}|) \rangle \\
&= \sum_{i=1}^{n} (\alpha_{1i} - \alpha_{2i}) k(x_i, x) + (\alpha_{1i} + \alpha_{2i}) \\
&- \sum_{i=1}^{n} (\alpha_{1i} + \alpha_{2i})(1 - H) k(|x_i|, |x|)
\end{aligned} \tag{7}
$$

From the conditions for optimality, this regression problem can be solved by the matrix Equation (5): The choice of the weights v_{1i} and v_{2i} is determined based upon the error variables $\xi_{1i} = \alpha_{1i} / \gamma$, $\xi_{2i} = \alpha_{2i} / \gamma$.

4. Reweighting Scheme to Resist Outliers Influence

Robust estimates are obtained by using the same iteratively reweighting approach in reference [6]. The iterative approach is summarized as follows:

1) Set $v_{1i}=1, v_{2i}=1, i=1\cdots$ The proposed interval regression is used to obtain the estimated outputs. Then, the regression errors $\xi_{1i} = \alpha_{1i} / \gamma, \xi_{2i} = \alpha_{2i} / \gamma$ are calculated.

2) Repeat:

3) $s = 1.483 * MAD([\text{errors ot the ith} \xi_{1i}, \xi_{2i}])$

4) Determine the weights $(v_{1k}^{(i)}, v_{2k}^{(i)})$ based upon $r_{1k}^{(i)} = \xi_{1k}^{(i)} / s, r_{2k}^{Ii)} = \xi 2_k^{(i)} / s,$ and the logistic weights function $v_{1k}(r) = \frac{\tanh((r_{1k}^{(i)})}{r_{1k}^{(i)}}$, $v_{2k}(r) = \frac{\tanh((r_{2k}^{(i)})}{r_{2k}^{(i)}}$.

5) Solve the weighted interval regression model with weight v_{1k} and v_{2k} .

6) Set $i = i + 1$

7) Until Lagrange multipliers $\alpha_{1k}^{(i-1)}, \alpha_{2k}^{(i-1)}$ and $\alpha_{1k}^{(i)}, \alpha_{2k}^{(i)}$, $k = 0,1,\cdots m$ are sufficiently close to each other.

5. Experiments

In the first example, We apply the unified WLS_IR to the data set of crisp inputs and interval outputs shown in Table 1.

To illustrate the proposed method, the second example [22] are presented. Because this function is not affected by outliers, the weight parameters v_{1i} v_{2i} is assumed From Figure 1, LS_SVFR performs fairly well for this function.

Now, LS_SVFR is applied in sensor faults detection for the plant with min-disturbance. Sensor is important for the plant to achieve its optimal performance. All sensor faults must be detected accurately and rapidly to prevent serious accidents. Consider the plant:

Table 1. Crisp Inputs and Interval Outputs

No(i)	Crisp input x_i	Fuzzy output (y_i, e_i)
1	0.1	(2.25,0.785)
2	0.2	(2.875,0.875)
3	0.3	(2.5,1)
4	0.4	(4.25,1.75)
5	0.5	(4,1.5)
6	0.6	(5.25,1.25)
7	0.7	(7.5,2)
8	0.8	(8.5,1.5)

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$$
$$y(t) = \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} d(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} f(t) \quad (8)$$

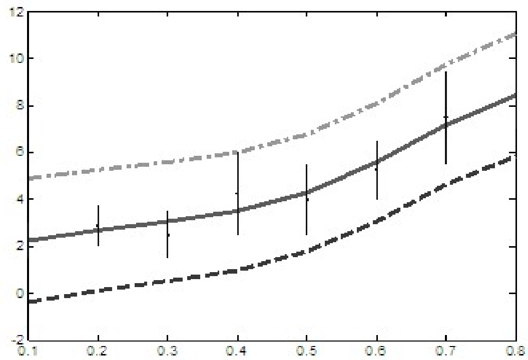

Figure 1. Simulation Result for the Second Example

Where $x \in R^n$ is the state of the plant, $u \in R^m$ is the control input, $y \in R^p$ is the measurable output of the plant, $d \in R^l$ is the unknown min-disturbance, $f \in R^l$ is sensor fault. f and d are the uncoupled forms.

The unknown min-disturbance is assumed as following:

$$d(t) = 0.1 - 0.2 * 1 * \text{rand}(\,) \quad (23)$$

It denotes noises generated in the interval [-0.1, 0.1] at random. The sensor fault is given as:

$$f(t) = \begin{cases} 0.3\sin(t) + 0.1 & 4s \le t \le 8s \\ 0 & t < 4s, t > 8s \end{cases} \quad (24)$$

In this experiment, $\Delta t = 0.2s$, m=4, n=4. Simulation time is chosen as 10 second. TWhen sensor faults occur, output of the plant will beyond the interval. Figure 2-3 show estimates of the interval regression and output of plant with sensor faults. As shown in Figure 2-3, because there are sensor faults in the plant between 4s and 8s, the outputs of plant also is beyond the interval between 4s and 8s. LS_SVFR is successful in detecting the sensor faults as early as possible.

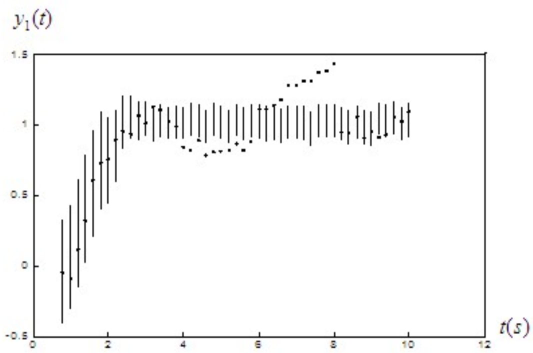

Figure 2. Estimates of Interval Bounds and the First Output Parameter of the Plant with Sensor Faults and Disturbance

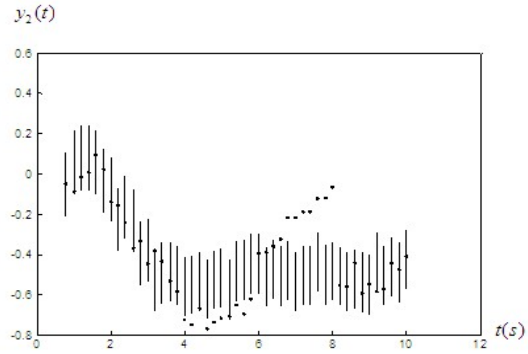

Figure 3. Estimates of Interval Bounds and the Second Output Parameter of the Plant with Sensor Faults and Disturbance

6. Conclusion

In order to preserve the advantages of LS_SVM and fuzzy regression, WLS_IR is presented by incorporating the concept of fuzzy set theory. By choosing different kernel functions, WLS_IR can denote different type nonlinear regression model to adapt different data sets. The experiments of fault instance demonstrate the feasibility and effectiveness of the interval regression model.

Acknowledgement

This work was supported by the Twelfth Five-Year Plan of Zhejiang province key discipline, Scientific Research Fund of Zhejiang Provincial Education Department (Y201119567), The Zhejiang Excellent Yong Teachers in University and Colleges Funding program (No. 20101215).

References

[1] C Cortes, V Vapnik. Support Vector Network. *Machine Learning*. 1995; 20(3): 273-297.
[2] J Jeng, C Chuang, Shun-Feng Su. Support Vector Interval Regression Networks For Interval Regression Analysis. *Fuzzy Sets and Systems*. 2003; 138(2): 283–300.
[3] C Hwang, D Hong, K Seok. Support Vector Interval Regression Machine For Crisp Input And Output Data. *Fuzzy Sets and Systems*. 2006; 157(8): 1114-1125.
[4] P Hao. Interval Regression Analysis Using Support Vector Networks. *Fuzzy sets and systems.* 2009; 160(7): 2466-2485.
[5] De Brabanter K, Pelckmans K, De Brabanter J, Debruyne M, Suykens JAK, Hubert M, De Moor B. *Robustness Of Kernel Based Regression: A Comparison Of Iterative Weighting Schemes.* Proceedings of the 19th International Conference on Artificial Neural Networks (ICANN), Limassol. Cyprus. 2009: 100-110.
[6] LWS Sun, J Li. BISVM: Block-Based Incremental Training Algorithm Of SVM For Very Large Dataset. *Application Research of Computers.* 2008; 25(1): 98-100.
[7] JAK Suykens, J Vandewalle. Least Squares Support Vector Machine Classifiers. *Neural Processing Letters*. 1999; 9(3): 293-300.
[8] L Hu, K Cao, H Xu. Fault Diagnosis For Hydraulic Actuator On Support Vector Regression. *Journal of system simulation*. 2007; 19(23); 5508-5512.
[9] H Tanaka. Fuzzy data analysis by possibilistic linear models. *Fuzzy Sets and Systems*. 1987; 24(1): 363–375.