Appl Intell (2011) 35:428–435
DOI 10.1007/s10489-010-0232-5

Fuzzy cost support vector regression on the fuzzy samples

**Abedin Vahedian · Mehri Sadoghi Yazdi ·
Sohrab Effati · Hadi Sadoghi Yazdi**

Published online: 21 May 2010

Abstract This paper presents a new version of support vector regression (SVR) named Fuzzy Cost SVR (FCSVR) with a unique property of operating on fuzzy data where fuzzy cost (fuzzy margin and fuzzy penalty) are maximized. This idea admits to have uncertainty in the penalty and margin terms jointly. Robustness against noise is shown to be superior in the experimental results as a property compared with conventional SVR.

Keywords Fuzzy samples · Support vector regression · Fuzzy input · Fuzzy cost

1 Introduction

The standard support vector machine works using crisp training samples. Chun-fu Lin in [1, 2] proposed fuzzy support vector machine (FSVM) by considering noise in the training samples. They used the membership function to express the membership value of a sample to positive or negative classes, with crisp training data. It, however, remains a conventional support vector machine from the view point of fuzzy theory. The degree of importance of training data is then modeled in the FSVM by inserting a membership value, μ_i as a penalty term of the cost function in the form of $\frac{1}{2}\|W\|^2 + C(\sum_{i=1}^{l} \mu_i \xi_i)$. The error term ξ_i is scaled by μ_i. The fuzzy membership values are used to weight the soft penalty term in the cost function of SVM. The weighted soft penalty term reflects the relative fidelity of the training samples during training. Important samples with large membership values will have more emphasis in the FSVM training procedure and more effect on the determination of hyperplanes. In [1] linear and quadratic functions are presented for μ_i in the FSVM, on which two main targets are followed, increasing margin and decreasing misclassification error.

Hong in [3] presented support vector fuzzy regression machines which introduces use of SVM for multivariate fuzzy linear and nonlinear regression models. The model presented in [3] for regression includes fuzzy input and output $(\tilde{x}, \tilde{y})$ in the form of:

$$\tilde{y} = w^T \tilde{x} + \tilde{b}. \tag{1}$$

A SVM model is, then, used for calculation of crisp w (weights). This model includes conventional fuzzy regression with new constraints in which upper and lower bounds of fuzzy input and output are used for generation of constraints. The effect of fuzzy variables (input and output) on the cost of SVR has not been considered though. Assuredly, uncertainty in input data affects margin and penalty maximization in the SVR, which has not been studied in the previous works.

A. Vahedian · H. Sadoghi Yazdi (✉)
Computer Department, Ferdowsi University of Mashhad, Mashhad, Iran
e-mail: h-sadoghi@um.ac.ir

M. Sadoghi Yazdi
Electrical and Computer Engineering Department, Shahid Beheshti University of Tehran, Tehran, Iran

S. Effati
Mathematics Department, Ferdowsi University of Mashhad, Mashhad, Iran

In [4] Ji studied the support vector machine with fuzzy chance constraints in the following form:

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|W\|^2 + C\sum_{i=1}^{l}\xi_i \\ \text{subject to} \quad & \text{Pos}\{y_i(W^T\tilde{X}_i + b) + \xi_i \geq 1\} \geq \lambda \\ & \xi_i \geq 0, \quad i = 1, 2, \ldots, l. \end{aligned} \tag{2}$$

They showed that $\text{Pos}\{\tilde{a} \leq 0\} \geq \lambda$ with triangular fuzzy number $\tilde{a} = (r_1, r_2, r_3)$ for any given level of λ $(0 < \lambda \leq 1)$ is equivalent to: $(1-\lambda)r_1 + r_2 \leq 0$. Thereupon, constraints in (2) are simplified.

In our previous work [5], probabilistic constraints were applied to reduce the impact of noisy samples in maximization of margin. A constraint is in the form of $\Pr(d_i(w^T x_i + b) \geq u_i) \geq \delta_i$ where u_i is an independent random variable with a known distribution function and $0 \leq \delta_i \leq 1$ is the value of effect of i samples in fixing the optimal hyperplane.

Liu in [6] presented total margin-based adaptive fuzzy support vector machines, TAF-SVM. TAF-SVM is a type of FSVM which also corrects the skew of the optimal separating hyperplane due to the very imbalanced data sets by using different cost algorithm. This work was performed by dividing training data into two categories with different levels of importance and results in dual problems in different boundaries for different Lagrange multipliers.

In [7] two new methods for calculation of membership function of μ_i are presented based on geometry of distribution of the training samples. Those samples are near to optimal hyperplane and have similar geometric properties. The main idea of FSVM is that if the input is detected as an outlier or noisy sample, membership function decreases so that total error decreases. In [8] a new method for μ_i of FSVM is presented which follows the same idea that one input is assigned with a low membership to the class if it is detected as an outlier. However, the method presented in [8] treats each input as an input of the opposite class with higher membership and makes full use of the data achieving better generalization ability. Also in two different works [9, 10], authors have tried to determine membership function in multi-category data classification.

The related works reviewed above can be categorized in the following form:

(I) Standard FSVM and its variants, which modify membership function μ_i.
(II) SVMs with special constraints for better performance against noisy samples.
(III) SVM as a method for finding optimal parameters of a regression model.

The main idea in this work is the presentation of a full fuzzy support vector machine according to a new fuzzy cost and fuzzy input signal. Undoubtedly, fuzzy input or fuzzy penalty cannot exist alone. If we assume that the input signal is a fuzzy number then fuzzyfication permeates into the output part of SVM, which includes margin and penalty terms.

This paper is organized as follows: The SVM and SVR are discussed in Sect. 2 in more detail. Section 3 is devoted to the proposed method, namely fuzzy cost SVR (FCSVR). Experimental results are discussed in Sect. 4. Final section incorporates conclusions and future work.

2 Support vector machine and regression

We first discuss Support vector machine and regression, prior to introducing our approach. The support vector machine (SVM) is a supervised learning method which generates input-output mapping functions from a set of labeled training data. The mapping function can be either a classification function, i.e., the category of the input data, or a regression function. Initially developed for solving classification problems, support vector techniques can be successfully applied to regression. The general regression learning problem is set as follow:

Suppose we are given the training data $\{(X_1, y_1), (X_2, y_2), \ldots, (X_l, y_l)\} \subset \bar{X} \times R$, where $\bar{X}$ denotes the space of the input patterns (e.g. $\bar{X} = R^D$). In ε-SV regression [Vapnik, 1995], the goal is to find a function $f(x)$ that has at most ε deviation from the actually obtained targets y_i for all the training data. The regressor must not only fit the given data well, but also make minimal error in predicting the values at any other arbitrary point in R^D. Nonlinear regression is accomplished by fitting a linear regressor in a higher dimensional feature space. A nonlinear transformation ϕ is used to transform data points from the input space with dimension D into a feature space having a higher dimension L. The nonlinear mapping is denoted by $\phi : R^D \to R^L$.

This problem can be written as a convex optimization problem; hence, we arrive at the formulation stated in [Vapnik, 1995].

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|W\|^2 + C\left(\sum_{i=1}^{l}(\xi_i + \xi_i^*)\right) \\ \text{subject to} \quad & y_i - W^T\phi(X_i) - b \leq \varepsilon + \xi_i \\ & -y_i + W^T\phi(X_i) + b \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0 \end{aligned} \tag{3}$$

where $W = [w_1, \ldots, w_d]^T$, $C > 0$ is a constant, ξ_i, ξ_i^* are slack variables for soft margin SVM, that allows to accept some deviation larger than the precision, ε. It turns out that in most cases the optimization problem (3) can be solved more easily in its dual formulation.

$$\begin{aligned}\text{Maximize}\quad & -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i-\alpha_i^*)(\alpha_j-\alpha_j^*)K(X_i,X_j)\\ & -\varepsilon\sum_{i=1}^{l}(\alpha_i+\alpha_i^*)+\sum_{i=1}^{l}y_i(\alpha_i-\alpha_i^*) \qquad (4)\\ \text{subject to}\quad & \sum_{i=1}^{l}(\alpha_i-\alpha_i^*)=0,\quad \alpha_i,\alpha_i^*\in[0,C]\end{aligned}$$

where α_i, α_i^* are Lagrange coefficients and matrix K is termed as a kernel matrix and its elements are given by: $K(X_i, X_j) = \phi(X_i)^T\phi(X_j)\ i, j = 1, 2, \ldots, M$.

By solving (4) we can find Lagrange coefficients and by replacing them, we have: $W = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\phi(X_i)$. Thus we can find the hyperplane function as:

$$f(X)=\sum_{i=1}^{l}(\alpha_i-\alpha_i^*)K(X_i,X)+b. \qquad (5)$$

3 The proposed fuzzy cost support vector regression (FCSVR)

We now discuss the proposed algorithm for support vector regression, termed as fuzzy cost support vector regression (FCSVR). Consider the fuzzy sample set $S = \{(\tilde{X}_1, y_1), (\tilde{X}_2, y_2), \ldots, (\tilde{X}_l, y_l)\}$, where $\tilde{X}_i = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_d)$ is a d-dimensional fuzzy input vector, y_i is the desired output, and l is the number of training samples for regression operation. The fuzzy input can be having different form of membership functions. Here we will consider the following linear membership function related to each fuzzy sample:

$$\mu_i(X)=\begin{cases}1 & X\le X_i\\ \frac{X_i+d_i-X}{d_i} & X_i\le X\le X_i+d_i\\ 0 & X\ge X_i+d_i\end{cases} \qquad (6)$$

d_i is the tolerance of ith input vector and $d_i \in (0, 1]$, $X \in R^D$, $i = 1, \ldots, l$. Some considerable are notable about the tolerance of data. (I) Without any further calculation, we enter the concept of noise in (6). (II) In many applications we cannot easily obtain prescience (prior knowledge) about Signal to Noise Ratio (SNR) so we consider SNR using data with tolerance. (III) Also (6) provides the ability of inserting samples with tolerance and degree of uncertainty in the training of learning data. (IV) Tolerance is an unwanted part, which is derived from unprecise nature of devices, and sensors in data acquisition, its inclusion in the estimation/regression, however, appears to be a challenging issue.

The support vector machine for fuzzy linear examples solves the following fuzzy quadratic equation:

$$\begin{aligned}\text{Minimize}\quad & \frac{1}{2}\|W\|^2+C\left(\sum_{i=1}^{l}(\xi_i+\xi_i^*)\right)\\ \text{subject to}\quad & y_i-W^T\tilde{X}_i-b\le\varepsilon+\xi_i \qquad (7)\\ & -y_i+W^T\tilde{X}_i+b\le\varepsilon+\xi_i^*\\ & \xi_i,\xi_i^*\ge 0.\end{aligned}$$

In (7) fuzzy input leads to fuzzy cost $Z = \frac{1}{2}\|W\|^2 + C(\sum_{i=1}^{l}(\xi_i + \xi_i^*))$. For entering the fuzzy concept in Z, we use the following algorithm which incorporates determining the upper and lower cost function (Z) and its fuzzyfication.

Step (I) Boundary calculation of the cost function

As the range of fuzzy samples is $[X_i, X_i + d_i]$, Z is therefore obtained in the bounds as the solution of two classical convex quadratic programming (QP) problems. It takes the form of (8) for the lower bound and that of (9) for the upper bound.

$$\begin{aligned}\text{Minimize}\quad & Z_1=\frac{1}{2}\|W\|^2+C\left(\sum_{i=1}^{l}(\xi_i+\xi_i^*)\right)\\ \text{subject to}\quad & y_i-W^TX_i-b\le\varepsilon+\xi_i \qquad (8)\\ & -y_i+W^TX_i+b\le\varepsilon+\xi_i^*\\ & \xi_i,\xi_i^*\ge 0\end{aligned}$$

and

$$\begin{aligned}\text{Minimize}\quad & Z_2=\frac{1}{2}\|W\|^2+C\left(\sum_{i=1}^{l}(\xi_i+\xi_i^*)\right)\\ \text{subject to}\quad & y_i-W^T(X_i+d_i)-b\le\varepsilon+\xi_i \qquad (9)\\ & -y_i+W^T(X_i+d_i)+b\le\varepsilon+\xi_i^*\\ & \xi_i,\xi_i^*\ge 0.\end{aligned}$$

Solving (8) and (9), results in Z_1 andZ_2 as the values for Z respectively. In the next step, cost function of Z is fuzzified for studying all states of the input signal in $[X_i, X_i + d_i]$.

Step (II) Fuzzyfication of the cost function

Lower and upper bounds of Z are:

$$\text{Min}\{Z_1,Z_2\}=Z_l \qquad (10)$$

$$\text{Max}\{Z_1,Z_2\}=Z_u \qquad (11)$$

where Z_u is the upper bound and Z_l is the lower bound of the object function of (8) and (9) respectively. Other optimum values are varying between the two values where inputs are varying between $[X_i, X_i + d_i]$. Now we consider the following linear membership function to determine the optimal grade for Z:

$$\mu_Z(W, \xi + \xi^*) = \begin{cases} 1, & Z \le Z_l \\ \frac{Z_u - Z}{Z_u - Z_l}, & Z_l \le Z \le Z_u \\ 0, & Z \ge Z_u \end{cases} \tag{12}$$

where, $W = [w_1, \ldots, w_d]^T$, $\xi = [\xi_1, \ldots, \xi_d]^T$, $\xi^* = [\xi_1^*, \ldots, \xi_d^*]^T$.

Step (III) Finding decision space

The membership function of the fuzzy set "decision" of fuzzy model is in the following form:

$$\mu_{ci}(W, \xi + \xi^*) = \begin{cases} 0 & y_i - W^T(X_i + d_i) - b \ge \varepsilon + \xi_i \\ \frac{\varepsilon + \xi_i - (y_i - W^T(X_i + d_i) - b)}{W^T d_i} & y_i - W^T(X_i + d_i) - b \le \varepsilon + \xi_i \le y_i - W^T X_i - b \\ 1 & y_i - W^T X_i - b \le \varepsilon + \xi_i \end{cases} \tag{13}$$

and

$$\mu_{ci}^*(W, \xi + \xi^*) = \begin{cases} 0 & -y_i + W^T X_i + b \ge \varepsilon + \xi_i^* \\ \frac{\varepsilon + \xi_i^* - (-y_i + W^T X_i + b)}{W^T d_i} & -y_i + W^T X_i + b \le \varepsilon + \xi_i^* \le -y_i + W^T(X_i + d_i) + b \\ 1 & -y_i - W^T(X_i + d_i) + b \le \varepsilon + \xi_i^* \end{cases} \tag{14}$$

where $W^T d_i \ne 0$.

The intersection of the membership function of objective function and the membership function of constraints are in fact the minimization of all membership functions. Therefore, we must maximize this minimum value. We have:

$$\text{Max Min}\{\mu_Z(W, \xi + \xi^*), \mu_{c1}(W, \xi + \xi^*), \ldots, \mu_{cl}(W, \xi + \xi^*), \mu_{c1}^*(W, \xi + \xi^*), \ldots, \mu_{cl}^*(W, \xi + \xi^*)\}. \tag{15}$$

Using α-cut method, we arrive at the following constraint:

$$\begin{aligned} &\text{Maximize} && \alpha \\ &\text{subject to} && \mu_Z(W, \xi + \xi^*) \ge \alpha \\ &&& \mu_{ci}(W, \xi + \xi^*) \ge \alpha \\ &&& \mu_{ci}^*(W, \xi + \xi^*) \ge \alpha \\ &&& 0 \le \alpha \le 1. \end{aligned} \tag{16}$$

Substituting in the above, assuming that $W^T d_i$ is nonzero, we have:

$$\begin{aligned} &\text{Maximize} && \alpha \\ &\text{subject to} && -\frac{1}{2}\|W\|^2 - C\left(\sum_{i=1}^{l}(\xi_i + \xi_i^*)\right) \ge \alpha(Z_u - Z_l) - Z_u \\ &&& \varepsilon + \xi_i - (y_i - W^T(X_i + d_i) - b) \ge \alpha W^T d_i \\ &&& \varepsilon + \xi_i^* - (-y_i + W^T X_i + b) \ge \alpha W^T d_i. \end{aligned} \tag{17}$$

Solving this, we find optimized W, b and maximum α.

4 Experimental results

We now demonstrate the effectiveness of our proposed model for linear function approximation. The experimental results pertaining to the model are compared to conventional support vector regression models. In this work, we have in fact studied the effect of measurement noise on the proposed method in estimation of the desired function. We used MATLAB for implementing and testing our method. Results are obtained from an average of 400 times of executing the program. Some definitions are mentioned before carrying out the experiments.

Triangular or trapezoidal form of fuzzy numbers is used for simulation of uncertain data in the operation of regression. They fall in the interval $[X_i, X_i + d_i]$. If $\tilde{X}$ is a fuzzy number then alpha-cut of $\tilde{X}$ is represented by $X_\alpha = \{X : \mu_{\tilde{X}} \ge \alpha\}$ that is a closed interval and is denoted by: $\tilde{X}_\alpha = [X_\alpha^L, X_\alpha^U]$, where $\alpha \in [0, 1]$.

An LR-type fuzzy number $\tilde{X}$ with its membership function $\mu_{\tilde{X}}(x)$ is defined as:

$$\mu_{\tilde{X}}(X) = \begin{cases} L(\frac{m_1 - X}{\alpha}) & \text{for } X \le m_1 \\ 1 & \text{for } m_1 \le X \le m_2 \\ R(\frac{X - m_2}{\beta}) & \text{for } X \ge m_2. \end{cases} \tag{18}$$

This is called an LR-type TFN (Trapezoidal Fuzzy Number) where m_1, m_2 are boundaries. The results are represented in Fig. 1(a), Fig. 2 and Fig. 6 respectively. α, β are slopes of

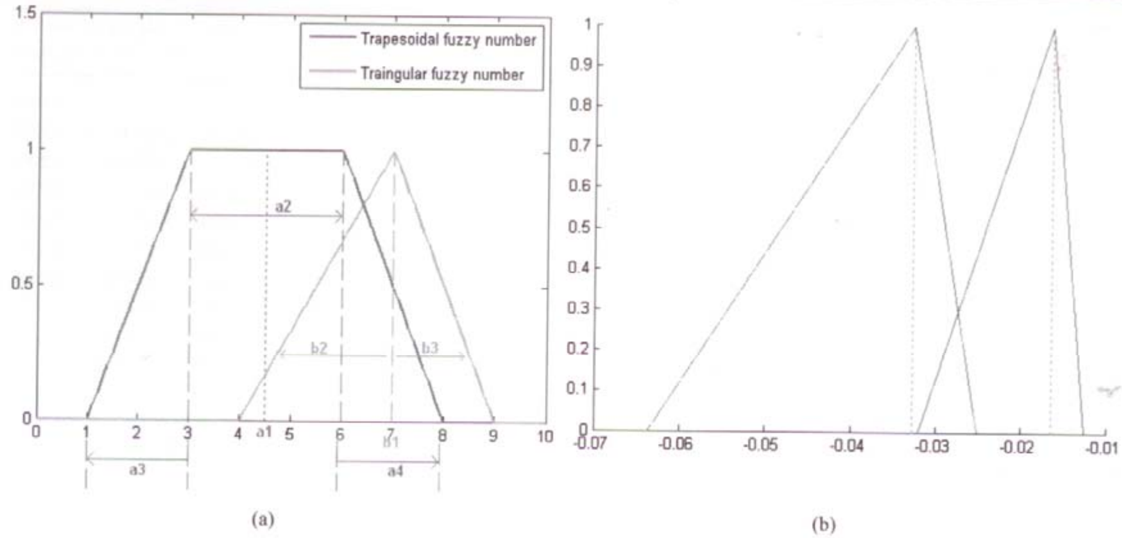

Fig. 1 (**a**) Two common samples of TFN, (**b**) LR-type triangular fuzzy number

right and left side of trapezoid. Two types of TFN are shown in Fig. 1(a).

In general, fuzzy number $\tilde{X}$ is a number in the interval $[X_i, X_i + d_i]$ with defined uncertainty degree. Noise may affect parameters of fuzzy numbers. This effect can be modeled as the following form using LR-type fuzzy numbers:

$$\mu_{\tilde{X}_n}(X) = \begin{cases} L(\frac{\hat{m}_1 - X}{\hat{\alpha}}) & \text{for } X \le \hat{m}_1 \\ 1 & \text{for } \hat{m}_1 \le X \le \hat{m}_2 \\ R(\frac{X - \hat{m}_2}{\hat{\beta}}) & \text{for } X \ge \hat{m}_2 \end{cases} \quad (19)$$

where $\hat{m}_1, \hat{m}_2, \hat{\alpha}, \hat{\beta}$ are noisy parameters of LR-type fuzzy numbers corrupted with uniform noise. Accurate study of noise effects and method of contamination is a new work in the field of fuzzy numbers. Signal to Noise Ratio (SNR) is also defined as $20 \log \frac{D_s}{D_n}$ where D_s is the main value of parameters and D_n is the domain of noise. Error is also defined in the form of $\frac{1}{N}\sum_{i=0}^{N}(\hat{y}_i - y_i)^2$ where $\hat{y}_i$ is the resulted output using SVR or FCSVR method and y_i is the desired output. N is the number of training samples.

Example Given the samples shown in Fig. 2, we want to estimate linear equation in the form of $y = wX + b$ to check the results. It is known that the given data are generated from $y = 3.73X + 4$ and noise is added to y which is modeled in the form of added noise into parameters of fuzzy numbers.

Obtained results using standard SVR and FCSVR are shown in Table 1. The optimum value of input tolerance

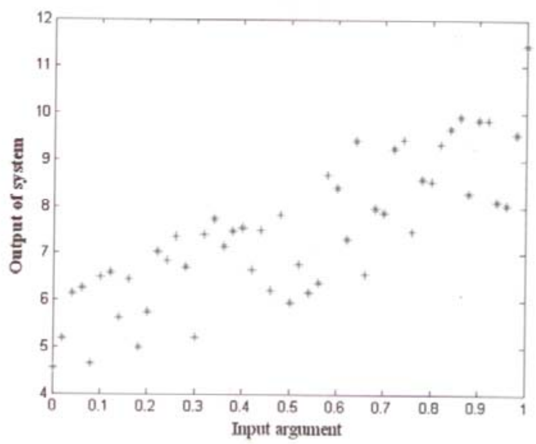

Fig. 2 Noisy captured data in signal to noise ratio equal 13.9 dB

(the parameter d_i, mentioned in (17)) is obtained using exhaustive search and is shown in second column of Table 1. Maximum value of membership degree α appears in the end column. Error indicates superiority of the proposed FCSVR relative to standard SVR. $\hat{w}_{\mathrm{SVR}}, \hat{b}_{\mathrm{SVR}}$ are estimated parameters by SVR and $\hat{w}_{\mathrm{FCSVR}}, \hat{b}_{\mathrm{FCSVR}}$ are estimated parameters by FCSVR. Also, error of SVR (e_{SVR}) and error of FCSVR (e_{FCSVR}) are shown in Fig. 3.

Figure 4 demonstrates the optimum tolerance of (d_i) in different SNRs. In the low noise condition or low SNR, to gain lower error, we need to decrease d_i. It means that decreasing the certainty degree (d_i) must be performed once the signal has been detected to be contaminated with noise.

Table 1 Result of estimation of $y = wX + b$ from captured noisy samples (Yn) (as shown in Fig. 2) in different SNR over 400 runs

SNR (per dB)	d_i	$\hat{w}_{SVR}$	$\hat{w}_{FCSVR}$	$\hat{b}_{SVR}$	$\hat{b}_{FCSVR}$	e_{SVR}	e_{FCSVR}	α
26.02	0.08	3.8211	3.6828	4.1015	4.0159	0.0412	0.0005	0.1865
20	0.14	3.9282	3.7527	4.2009	4.0046	0.1722	0.0006	0.1969
16.4782	0.2	4.0006	3.8351	4.2963	3.961	0.3549	0.0024	0.2086
13.9794	0.28	4.115	3.969	4.3966	3.861	0.6635	0.0112	0.2338
12.76	0.3	4.1039	3.9749	4.4608	3.8771	1.0394	0.02	0.2249
12.0412	0.34	4.1979	4.0595	4.5043	3.8268	1.4988	0.0409	0.2265
11.37	0.4	4.2109	4.1095	4.5595	3.7255	1.5443	0.0438	0.2491
10.4576	0.42	4.3304	4.1883	4.5838	3.8069	2.3889	0.0784	0.2164
9.1186	0.48	4.1579	4.0748	4.704	3.677	2.9048	0.0819	0.2381

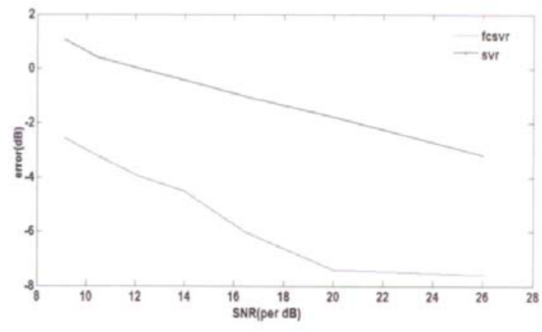

Fig. 3 Comparison of the proposed FCSVR and standard SVR

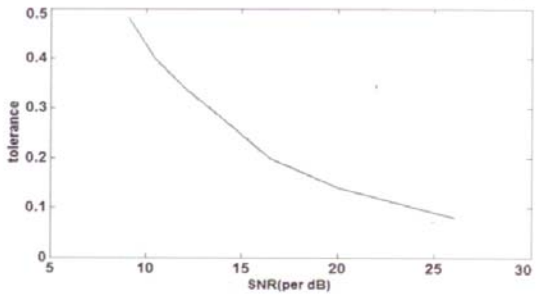

Fig. 4 Input tolerance (di) versus SNR

Therefore, if SNR decreases, to have lower error, d_i must increase. The main problem, however, would be to estimate the level of noise present in the signal. In the future work this issue will be addressed to complete regression system.

Maximum membership, (α) in (17) indicates its relation with SNR. An increase in the noise level, α increases about 10–15% for a 50% decrease in SNR. An increase in α means that fuzzy values are selected in a narrower range. In other

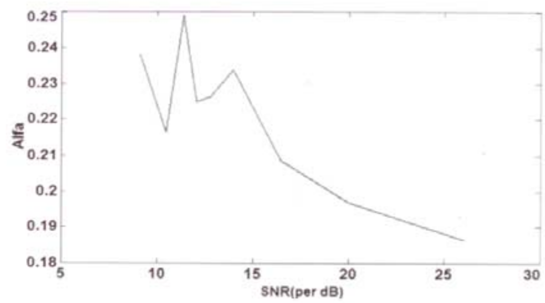

Fig. 5 α versus SNR

words, according to (17) constraints $\mu_Z(W, \xi + \xi^*) \geq \alpha$, $\mu_{ci}(W, \xi + \xi^*) \geq \alpha$, $\mu^*_{ci}(W, \xi + \xi^*) \geq \alpha$ are satisfied in higher certainty. For $\mu_g(x) \geq \alpha$ higher α means that the optimum Z in (9) moves towards Z_l or alternatively the cost function is spotted with higher degree of certainty. It is obvious that Z includes both the margin of SVR and the penalty term, therefore, decreasing uncertainty in the margin results in a medium to high level of SNR. Simultaneously, the penalty term has higher certainty. From $\mu_{ci}(W, \xi + \xi^*) \geq \alpha$, $\mu^*_{ci}(W, \xi + \xi^*) \geq \alpha$ we find that, constraints move toward standard SVR.

Alternatively, when SNR is high, the regression model moves toward SVR with high value of uncertainty due to uncertainty in modeling the input data. This lemma is correct only in medium to high level of SNR (more than 13 dB) according to Fig. 5. There is no basis for evaluating low value of SNR at the present, though.

Figure 6 indicates the function estimation using SVR and FCSVR in two different values (medium and high) of SNR. Robustness of FCSVR against noise is noticeable compared to SVR.

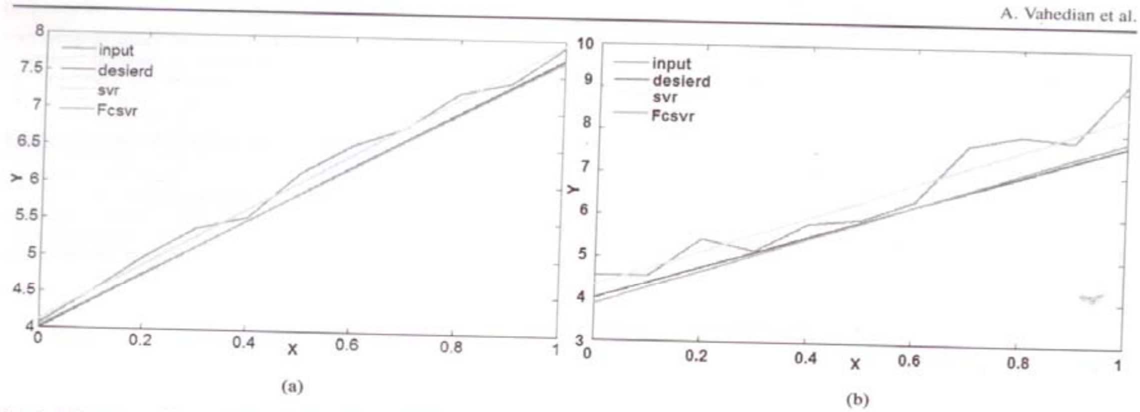

Fig. 6 Estimation of Y=3.73X+4 with two different SNR. (**a**) SNR = 30 dB, (**b**) SNR = 14 dB

5 Conclusion

Noisy samples cause performance decrease in the support vector regression method. Fuzzy margin with fuzzy penalty concept were introduced in this paper. The idea could result in decreasing the noise effect. Several experiments were performed and compared to standard SVR. The obtained results indicate superiority of the proposed method as opposed to conventional SVR.

References

1. Lin C-f, Wang S-d (2002) Fuzzy support vector machine. IEEE Trans Neur Netw 13(2):464–471
2. Lin C-f, Wang S-d (2004) Training algorithms for fuzzy support vector machines with noisy data. Pattern Recognit Lett 25:1647–1656
3. Hong DH, Hwang C (2003) Support vector fuzzy regression machines. Fuzzy Sets Syst 138:271–281
4. Ji A-B, Pang J-H, Li S-H, Sun J-P (2006) Support vector machine for classification based on fuzzy training data. In: Proceedings of the fifth int conf on machine learning and cybernetics, Dalian, 13–16 August 2006, pp 1609–1614
5. Sadoghi Yazdi H, Effati S, Saberi Z (2007) The probabilistic constraints in the support vector machine. Appl Math Comput 194(2):467–479
6. Liu Y-H, Chen Y-T (2007) Face recognition using total margin-based adaptive fuzzy support vector machines. IEEE Trans Neur Netw 18(1):178–192
7. Chu L, Wu C (2004) A fuzzy support vector machine based on geometric model. In: Proceedings of the fifth world congress on intelligent control and automation, Hangzhou, PR China, June 15–19 2004, pp 1843–1846
8. Wang Y, Wang S, Lai KK (2005) A new fuzzy support vector machine to evaluate credit risk. IEEE Trans Fuzzy Syst 13(6):820–831
9. Jayadeva J, Khemchandani R, Chandra S (2005) Fuzzy linear proximal support vector machines for multi-category data classification. Neurocomputing 67:426–435
10. Wang T-Y, Chiang H-M (2007) Fuzzy support vector machine for multi-class text categorization. Inf Process Manag 43:914–929

Abedin Vahedian received his Bsc. In Electronics in 1987 from Ferdowsi University of Mashhad, IRAN. He joined Iranian Research Org. for Science & Technology as applied research director in Electronics and Medical Engineering and worked for 5 years, when he was granted scholarship to continue his postgraduate in 1993. He completed his MEng. in DSP in 1995 and Phd in Digital Video Communications in 1996 from UNSW, Australia. He has been an academic board member since then and has been working on a number of applied research fields such as Video Conference, Video Compression Machine Vision, Intelligent Traffic Control and ICT. Dr Vahedian is currently Assistant Professor in Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad.

Mehri Sadoghi Yazdi was born in Iran in 1985. She received the BS degree in Computer Engineering from Azzahra University, Iran in 2007. She is currently pursuing her MS Program in computer engineering at Shahid Beheshti University of Tehran, Iran. Her research interests include artificial intelligence algorithm and image processing.

Sohrab Effati received the B.S. degree in Applied Mathematics from Birjand University, Birjand, Iran, the M.S. degree in Applied Mathematics from Institute of Mathematics at Tarbiat Moallem Tehran University, Tehran, Iran, in 1992 and 1995, respectively. He received the PhD degree in control systems from Ferdowsi University of Mashhad, Mashhad, Iran, in April 2000. He is an Associate Professor with the Department of Applied Mathematics at Ferdowsi University of Mashhad in Iran. His research interests include control systems, fuzzy theory, and neural network models and its applications in optimization problems.

Hadi Sadoghi Yazdi received the B.S. degree in electrical engineering from Ferdowsi University of Mashad, Iran in 1994, and then he received to the M.S. and Ph.D. degrees in electrical engineering from Tarbiat Modarres University, Iran, in 1996 and 2005, respectively, is currently Associate Professor in Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad. His research interests include pattern recognition, and optimization in signal processing.