SUPPORT VECTOR REGRESSION WITH FADING MEMORY FOR ARMA IDENTIFICATION

Waseda University Member Tadanobu Sato
Kyoto University Student Member ○Jian Zhang

1. Introduction

Support vector machine (SVM) has received many researchers' attentions especially in these years for it's potential capability as one of machine learning methods. Structural dynamical response can be written in ARMA form, therefore ARMA identification is the first step to make structural identification. In this paper, support vector regression (SVR) first is used to identify the stationary ARMA model. Next, an improved SVM by using the forgetting factor is given, which can be used to identify the time varying problems.

２. SVR formulation for ARMA identification

The ARMA model is defined as

$$y_t = \sum_{k=1}^{p} a_k y_{t-k} + \sum_{k=1}^{q} b_k x_{t-k+1} + e_t \tag{1}$$

Eq. (1) can be rewritten as

$$y_t = < w, x_t > + e_t \tag{2}$$

where, $<\cdot,\cdot>$ denotes the dot product, $w = \{a_1, \cdots, a_p, b_1, \cdots, b_q\}$ and $x_t = \{y_{t-p}, \cdots, y_{t-1}, y_t, x_t, x_{t-1}, \cdots, x_{t-q+1}\}$. The Support Vector Regression will search a optimum hyper-plane (defined by Eq.(2)) to separate these training data into two subsets by minimizing the sum error in Eq.(3).

$$R_{reg}[f] = \frac{1}{2} \| w \|^2 + C \sum_{i=1} (\xi_i + \xi_i^*) \tag{3}$$

Eq.(3) is constraint by three boundaries defined by

$$\begin{cases} y_i - < w, x_i > - b \le \varepsilon + \xi_i \\ -y_i + < w, x_i > + b \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \tag{4}$$

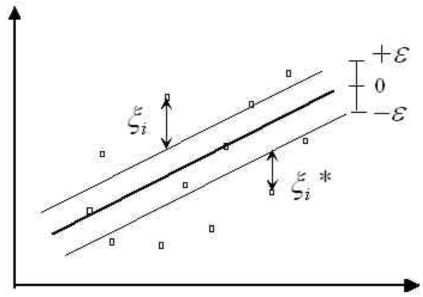

Fig. 1 ε-Insensitive loss function in SVR

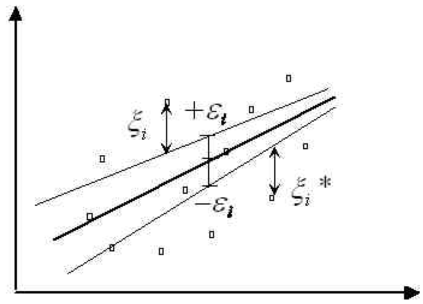

Fig. 2 ε_i-Insensitive loss function in SVR

Optimum problem Eq. (3) and constraint Eq. (4) is obtained based on the ε-Insensitive loss function (Fig. 1). After solving the optimum problem, w consisting the ARMA model parameters is identified.

３. ARMA numerical example

ARMA model as defined by Eq. (5) is used o to check the SVR identification effectiveness.

$$y_t = 0.2y_{t-3} + 0.5y_{t-2} - 0.3y_{t-1} + 2.0x_t - 0.6x_{t-1} + 1.5x_{t-2} - 0.8x_{t-3} + e_t \tag{5}$$

The noise added to the ARMA model is 5%, 20% Gauess White noise, where 5% means the standard variance of noise is 5% of that of the simulated training data.

Table 1. Identified result for the ARMA model

True value	0.2	0.5	-0.3	2.0	-0.6	1.5	-0.8	Mean error
5% noise	0.2152	0.5123	-0.2912	1.9936	-0.6106	1.5029	-0.8339	2.79%
20% noise	0.2689	0.5448	-0.2820	1.9838	-0.6062	1.5036	-0.9154	9.41%

After data preparation, the SVR training is carried out in the Matlab environment. The parameters of SVR for this example are chose as C=0.46 and ε equaling the noise level of input data. The training data number is 100. The computing time is 114.1 seconds. The identification results under different level noise are shown in Table 1.

〒611-0011 Disaster Prevention Research Institute, Kyoto University, Kyoto. TEL 0774-38-4071

４．Non-stationary ARMA identification by SVR

Support vector machine so far is efficiency to deal with stationary problem but can't track non-stationary variation of the state variable, because each training data has same weight to influence the identified objective function in the training procedure of SVM. Sato has developed adaptive Kalman filter and H_∞ filter with fading memory to identify non-stationary properties of structure. Cooper made a study in detail how to choose adaptive forgetting factor in his paper. The similar idea is adopt to improve the SVR, which is easy to understand that nearer observed data has larger influence on the structural parameter, therefore it is better to put larger weight to the near training data.

The further problem is how to add the forgetting factor in the support vector regression. As described before, the basic idea of SVR is looking for an optimal hyper-plane maximizing the margin. In the time varying case, the nearest observed data reflects the true properties of the studied problem, and the observed data in the past have a large bias, which can be taken as the "noise". That is, the longer time for the observed time, the larger noise is considered. It is well known that ε in Eq.(3) reflects the noise level in the observed data, therefore a varying ε instead of the constant ε can be adopt to reflect the different noise in each training data as shown in Fig. 2. That is, one ε_i is designed for each observed data, and larger values should be added for older observed data to reflect the larger "noise". Let's check Eq. (3).

$$C\sum_{i=1}^{n}\xi_i = C\sum_{i=1}^{n}(f(x_i)-\varepsilon_i) = C\sum_{i=1}^{n}\beta_i(f(x_i)-\varepsilon) \tag{6}$$

The forgetting factor can be defined as

$$\beta_i = \lambda_i\beta_{i-1} = \prod_{j=0}^{j=i}\lambda_j \tag{7}$$

where λ_j is smaller that 1 and larger than 0.

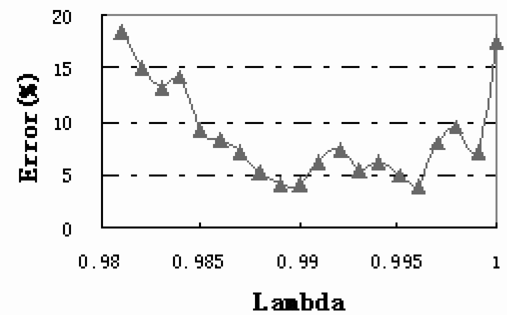

Fig. 3 Results under different Lambda

５．Non-stationary ARMA example

An ARMA model with time varying coefficients is defined as

$$y_t = 0.2f_ty_{t-3} + 0.5f_ty_{t-2} - 0.3f_ty_{t-1} + 2.0f_tx_t - 0.6f_tx_{t-1} + 1.5f_tx_{t-2} - 0.8f_tx_{t-3} + e_t \tag{8}$$

where $f_t = 1-0.005t$. Table 2 shows the identified result before and after using the forgetting factor. It is easy to find the result error is very large (17.32%) when their is no forgetting factor used, and the error can be reduced to 4% when using the forgetting factor 0.989. Figure 3 shows the identified results under different forgetting factor.

Table 2. Identified result for the non-stationary ARMA model

True value		$0.2f_t$	$0.5f_t$	$-0.3f_t$	$2.0f_t$	$-0.6f_t$	$1.5f_t$	$-0.8f_t$	Mean error
C=0.6	λ=0.989	2.52%	1.60%	1.37%	1.61%	4.76%	10.65%	5.56%	4.01%
	λ=1	11.84%	13.59%	26.70%	1.46%	29.14%	6.08%	32.42%	17.32%

６．Conclusion

Stationary and non-stationary ARMA model parameters identification by improved support vector regression were carried out in this paper. The numerical results showed that the improved SVR was one powerful, robust, and fast algorithm to make system identification. Therefore structural identification by SVR with fading memory is possible which will be studied in next stage work. How to choose optimum forgetting factor should be studied in the future.

7. Reference

A. J. Smola , B. Scholkopf. (1998). A tutoral on support vector regression. NeuroCOLT2 Reports Series. Berlin, Germany.

J. E. Cooper. K. Worden. (2000). On-line physical parameter estimation with adaptive forgetting factors. Mechanical Systems and Signal Processing 14(5), 705-730.

Tadanobu Sato, Kai Qi. (1998). Adaptive H_∞ filter: its application to structural identification. Journal of Engineering Mechanics. 1244(11),1233-1240.