Vol 13 No 12, December 2004
1009-1963/2004/13(12)/2045-08

Chinese Physics

© 2004 Chin. Phys. Soc.
and IOP Publishing Ltd

Modelling of chaotic systems based on modified weighted recurrent least squares support vector machines*

Sun Jian-Cheng(孙建成)†, Zhang Tai-Yi(张太镒), and Liu Feng(刘 枫)

Department of Information and Communication Eng., Xi'an Jiaotong University, Xi'an 710049, China

(Received 21 April 2004; revised manuscript received 22 September 2004)

Positive Lyapunov exponents cause the errors in modelling of the chaotic time series to grow exponentially. In this paper, we propose the modified version of the support vector machines (SVM) to deal with this problem. Based on recurrent least squares support vector machines (RLS-SVM), we introduce a weighted term to the cost function to compensate the prediction errors resulting from the positive global Lyapunov exponents. To demonstrate the effectiveness of our algorithm, we use the power spectrum and dynamic invariants involving the Lyapunov exponents and the correlation dimension as criterions, and then apply our method to the Santa Fe competition time series. The simulation results shows that the proposed method can capture the dynamics of the chaotic time series effectively.

Keywords: chaotic dynamics, dynamical invariants, support vector machines, least squares
PACC: 0545

1. Introduction

Building the model of a dynamical system by the time series analysis has been an important problem in science. For a long time the linear mode is the conventional tool exclusively used to describe the dynamical system. Since the linear equations can only lead to exponentially growing or periodically oscillating solutions, all irregular behaviour of the system has to be attributed to some random external input to the system. But, chaos theory shows that the random input is not the only possible source of irregularity in a system's output.[1] Chaotic systems constitute an important class of nonlinear dynamical systems in economics, meteorology, chemical processes, biology, hydrodynamics and many other situations. In common cases, the only information we can acquire about such systems is in the form of a time-ordered sequence of system state observations. Such time series do not necessarily contain observations for all states since some states may be hidden. The objective of time series analysis is then to build a model for the unknown dynamical system that generates the time series.

Various techniques for modelling and predicting nonlinear time series have been investigated in past years. There are many traditional methods associated with time series analysis such as linear regression and ARIMA models.[2] In addition, some new methods have been proposed to deal with the nonlinear time series analysis such as local linear mode,[3] chaotic attractor,[4] reconstructed embedding phase,[5] wavelets,[6] and modified neighbour regression.[7] However, in recent years, many researcher addressed the nonlinear time series analysis with the artificial neural networks.[8−10]

In this paper, a novel machine learning method, modified recurrent least squares support vector machines (MRLS-SVM), is developed for realizing the chaotic time series reconstruction. Support vector machines (SVM) have become a subject of intensive study and have been applied successfully to classification tasks as optical character recognition (OCR) and recently also to regression.[11] Least squares (LS) versions of SVM greatly simplify the problem since the solution is characterized by a linear system.[12] Some iterative operation is necessary for common prediction mission; to deal with this problem, recurrent least squares support vector machines (RLS-SVM) was deduced based on the SVM algorithm.[13] In our work,

*Project supported by the National Natural Science Foundation of China (Grant No 90207012).
†E-mail: sunjc@mailst.xjtu.edu.cn

we proposed the MRLS-SVM to deal with this problem. A component including the largest Lyapunov exponent is used in MRLS-SVM for weighting the sum squared error (SSE). With this weighting component, we expect that the errors resulting from the later iteration will have less influence since the largest Lyapunov exponent might introduce the exponential distortion in a chaotic time series.[8] On the other hand, the dynamic modelling is different from the prediction of chaotic time series since the former only addresses the long term behaviour, while the latter is interested in the instantaneous prediction error. For instance, two chaotic time series can be very different point to point but be produced from the same dynamical system. In this sense, the machine learning method also has better performance for dynamic modelling than does the conventional method.[11]

2. Embedding phase space of dynamical system

The deterministic dynamical system describes the time evolution of a system in some phase space $\Gamma \subset \boldsymbol{R}^m$; for simplicity we will assume that the phase space is a finite-dimensional vector space. A state is specified by a vector $\boldsymbol{x} \in \boldsymbol{R}^m$. Then we can describe the dynamics by an explicit system of m first-order ordinary differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t) = \boldsymbol{f}(t, \boldsymbol{x}(t)), \quad t \in \boldsymbol{R}, \tag{1}$$

or in discrete time $t = n\Delta t$ by maps of the form

$$\boldsymbol{x}_{n+1} = \boldsymbol{F}(\boldsymbol{x}_n). \tag{2}$$

A time series can then be thought of as a sequence of observations $\{s_n\} = h(\boldsymbol{x}_n)|_{n=1,2,\ldots,N_T}$ performed with some measurement function $h(\cdot)$, where N_T is the number of data points. Since the sequence $\{s_n\}$ (usually scalar) in itself does not properly represent the multidimensional phase space of the dynamical system, we have to employ some technique to unfold the multidimensional structure using the available data. Takens' embedding theorem[14] guarantees the reconstruction of a state space representation from a scalar signal alone. A delayed coordinate function Φ is defined by

$$\begin{aligned}\boldsymbol{s}_n &= \Phi(\boldsymbol{x}_n)\\ &= \left[h\left(\boldsymbol{x}_{n-(m-1)\tau_d}\right), \ldots, h\left(\boldsymbol{x}_{n-\tau_d}\right), h(\boldsymbol{x}_n)\right]\\ &= \left[s_{n-(m-1)\tau_d}, \ldots, s_{n-\tau_d}, s_n\right],\end{aligned} \tag{3}$$

where $\boldsymbol{s}_n$ are vectors in a new space, namely the embedding phase space, formed from time-delayed values of the scalar measurements. The number m of elements is called the embedding dimension, the time τ_d is generally referred to as the time delay or lag. So the trajectory matrix $\boldsymbol{S}$ can be constructed in m-embedding dimensions as

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{s}_n^T \\ \boldsymbol{s}_{n+1}^T \\ \vdots \\ \boldsymbol{s}_{n+N-1}^T \end{bmatrix} = \begin{bmatrix} s_{n-(m-1)\tau_d} & s_{n-(m-2)\tau_d} & \cdots & s_n \\ s_{n+1-(m-1)\tau_d} & s_{n+1-(m-2)\tau_d} & \cdots & s_{n+1} \\ \vdots & \vdots & & \vdots \\ s_{n+N-1-(m-1)\tau_d} & s_{n+N-1-(m-2)\tau_d} & \cdots & s_{n+N-1} \end{bmatrix}, \tag{4}$$

where $N = N_T - (m-1)$. Takens showed that if the sequence $\{s_n\}$ does indeed consists of scalar measurements of the state of a dynamical system, then under certain genericity assumptions, the time delay embedding provides a one-to-one image of the original set $\{\boldsymbol{x}\}$, provided m is large enough. For almost all τ_d and for some m, Takens' embedding theorem ensures that there is a smooth map $f : \boldsymbol{R}^m \to \boldsymbol{R}$ such that

$$s_{(n+1)\tau_d} = f(s_{n-(m-1)\tau_d}, \ldots, s_{n-\tau_d}, s_n) = f(\boldsymbol{s}_n). \tag{5}$$

The problem of remodelling then becomes equivalent to the problem of estimating the unknown function f in the embedding phase space. Many regression techniques can be used to solve problems of this type. In this paper, we concentrate on the MRLS-SVM algorithm, a novel regression-technique modified version of SVM.

For the sake of judging the success of dynamic modelling, some rules such as correlation dimension and Lyapunov exponents have been proposed since they can describe the global properties of the attractor. In addition, The power spectrum is a conventional method to analyse the time series and it can indicate if the dynamical system is periodic, quasiperiodic or chaotic. Chaotic and stochastic systems are easily distinguishable from periodic or quasiperiodic systems for they have rich broadband power spectra, as well as widely varying phase spectra. It was not possible

to distinguish the chaotic system from the stochastic system before the advent of the chaotic theory.

It is proven that the Lyapunov exponents are useful dynamic invariants to characterize the chaotic dynamic system. The Lyapunov exponents measure the rate at which nearby orbits converge or diverge. Roughly speaking, the Lyapunov exponent is the time constant that is the expression for the distance between two nearby orbits. If it is negative, then the orbits converge in time, and the dynamical system is insensitive to initial conditions. However, if it is positive, then the distance between nearby orbits grows exponentially in time, and the system exhibits sensitive dependence on initial conditions. Thus very different time series can be produced from the same dynamic system even if the initial conditions are different in a slight scale. There are as many Lyapunov exponents as there are dimensions in the state space of the system, but the largest is usually the most important since it indicates the chaotic degree of the dynamic system. Formally, the Lyapunov spectrum is defined by Wolf *et al*:[15] given an n-dimensional phase space, the long-term evolution of an infinitesimal sphere is monitored. As the sphere evolves, it will turn into an n-ellipsoid. The ith one-dimensional Lyapunov exponent is then defined in terms of the length of the resulting ellipsoid's principal axis

$$\lambda_i = \lim_{t\to\infty} \frac{1}{t} \log_2 \frac{p_i(t)}{p_i(0)}. \tag{6}$$

The Lyapunov spectrum is then formed by the set (λ_1, λ_2, ..., λ_n), where λ_i are arranged in decreasing order.

The difference between strange attractors and purely stochastic (random) processes is that the evolution of points in the phase space of a strange attractor has definite structure. The correlation integral provides a measure of the spatial organization of this structure, and is given by

$$C(r) = \lim_{N\to\infty} \frac{1}{N(N-1)} \sum_{i\neq j} \Theta(r - \|\boldsymbol{s}(i) - \boldsymbol{s}(j)\|), \tag{7}$$

where Θ is the Heaviside function. Grassberger and Procaccia found that, for a strange attractor, $C(r) \propto r^{\nu}$ for a limited range of r.[16] The power ν is called the correlation dimension of the attractor. Thus, we can plot the $\log C(r) - \log r$ graph to identify an attractor.

In addition to the above invariants, the Poincaré map is another parameter to characterize the chaotic system.[17] Since a chaotic system never revisits the same state, it will trace out contours on the Poincaré map. However, unlike a purely random process, these contours will have definite structure and will graphically indicate the presence of the responsible attractor.

3. Modified recurrent least squares support vector machines and learning algorithm

The foundations of SVM were developed by Vapnik and have gained popularity for their promising empirical performance.[11] Suykens *et al* proposed the RLS-SVM based on SSE to deal with some problems in prediction such as iterative operation.[13] However, with outliers on the data or the underlying assumption of a Gaussian distribution for the error, applying the SSE estimator for such data might result in completely erroneous estimates. In addition, since the chaotic character affects the prediction performance, using the SSE as criterion directly can reduce the prediction precision. In this work, we deal with this problem by considering the influence result from the Lyapunov exponents. In the following section, we formulate our algorithm based on the MRLS-SVM:

1. Choose optimal m, τ_d (scale with discrete sample units) and set $Q \subseteq \{1, ..., N\}$,

$$D = \{((s_{k-(m-1)\tau_d}, ..., s_{k-\tau_d}, s_k), s_{k+1}) \in \boldsymbol{R}^m \times \boldsymbol{R} | k \in Q\}.$$

2. Construct the trajectory matrix and build training data: given initial condition $\hat{s}_i = s_i$ for $i = 1, 2, ..., m$, the prediction problem is given by

$$\begin{aligned} \hat{s}_{(k+1)\tau_d} &= f(\hat{s}_{k-(m-1)\tau_d}, ..., \hat{s}_{k-\tau_d}, \hat{s}_k) \\ &= w^T \varphi \left(\begin{bmatrix} \hat{s}_{k-(m-1)_{\tau_d}} \\ \vdots \\ \hat{s}_{k-\tau_d} \\ \hat{s}_k \end{bmatrix} \right) + b, \end{aligned} \tag{8}$$

where $\varphi(\cdot) : \boldsymbol{R}^m \to \boldsymbol{R}^{n_h}$ is a nonlinear mapping in future space, $w \in \boldsymbol{R}^{n_h}$ is the output weight vector and $b \in \boldsymbol{R}$ is the bias term.

3. Choose the $\varphi(\cdot)$ and estimate the function $f(\cdot)$ using training data D.

4. Predict s by function $f(\cdot)$.

In Ref.[12], Suykens *et al* proposed that Eq.(8) can be converted into optimal problems which can be described as

$$\min_{w,e} J(w, e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=m+1}^{N+m} e_k^2 \tag{9}$$

subject to

$$s_{k+1} - e_{k+1} = w^T \varphi(\boldsymbol{s}_k - \boldsymbol{e}_k) + b, \tag{10}$$

$$k = m+1, ..., N+m,$$

where

$$e_k = s_k - \hat{s}_k, \boldsymbol{s}_k = [s_{k-(m-1)\tau_d}; ...; s_{k-\tau_d}, s_k],$$
$$\boldsymbol{e}_k = [e_{k-(m-1)\tau_d}; ...; e_{k-\tau_d}; e_k],$$

and γ is an adjustable constant. The basic idea of mapping function $\varphi(\cdot)$ is to map the data $\boldsymbol{s}$ into a high-dimensional feature space, and to do linear regression in this space.

In the above equation, we can see that the SSE is used as the cost function. According to the chaos theory, positive global Lyapunov exponents cause the errors in these predictions to grow exponentially, and it is conventionally assumed that the prediction error $E(t)$ will grow as

$$E(t) = E(0)\mathrm{e}^{\lambda_m t}, \tag{11}$$

where λ_m is the largest Lyapunov exponents. In this paper, we introduce a weighted component to compensate the cost function which is described as

$$P_k = (\mathrm{e}^{\lambda_m \Delta t})^{-(k-m-1)}, \tag{12}$$

where Δt is sampling interval. Rewrite the Eq.(9) as

$$\min_{w,e} J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=m+1}^{N+m} P_k e_k^2. \tag{13}$$

To resolve the optimal function Eq(11), we define the Lagrangian function

$$L(w,b,e;\alpha) = J(w,e) + \sum_{k=m+1}^{N+m} \alpha_{k-m} \times [h_{k+1} - e_{k+1} - w^T \varphi(\boldsymbol{s}_k - \boldsymbol{e}_k) - b], \tag{14}$$

where α_i are Lagrange multipliers. The conditions for optimality of the above function are

$$\begin{aligned}
\frac{\partial L}{\partial w} &= w - \sum_{k=m+1}^{N+m} \alpha_{k-m} \varphi(\boldsymbol{s}_k - \boldsymbol{e}_k) = 0, \\
\frac{\partial L}{\partial b} &= \sum_{k=m+1}^{N+m} \alpha_{k-m} = 0, \\
\frac{\partial L}{\partial e_k} &= \gamma P_k e_k - \alpha_{k-m} - \sum_{i=1}^{m} \alpha_{k-m+i} \\
&\quad \times \frac{\partial}{\partial e_{k-i}} [w^T \varphi(\boldsymbol{s}_k - \boldsymbol{e}_k)] = 0, \\
&\quad k = m+1, ..., N, \\
\frac{\partial L}{\partial \alpha_{k-m}} &= s_{k+1} - e_{k+1} - w^T \varphi(\boldsymbol{s}_k - \boldsymbol{e}_k) - b = 0, \\
&\quad k = m+1, ..., N+m.
\end{aligned} \tag{15}$$

After eliminating variables w and e one obtains the solution

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & \Omega + V_\gamma \end{bmatrix} \cdot \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \boldsymbol{s} \end{bmatrix}, \tag{16}$$

where

$$\boldsymbol{s} = [s_{m+1}; s_{m+2}; ...; s_{m+N}],$$
$$1_v = [1; ...; 1],$$
$$\boldsymbol{\alpha} = [\alpha_{m+1}; \alpha_{m+2}; ...; \alpha_{m+N}],$$
$$\Omega_{k,l} = \varphi(x_k)^T \varphi(x_l),$$

and V_γ is a diagonal matrix given by

$$V_\gamma = \mathrm{diag}\left\{ \frac{1}{\gamma P_{m+1}}, ..., \frac{1}{\gamma P_{m+N}} \right\}. \tag{17}$$

The resulting recurrent simulation model is described as follows:

$$s_{(k+1)\tau_d} = \sum_{l=m+1}^{N+m} \alpha_{l-m} K \left(\boldsymbol{z}_l, \begin{bmatrix} \hat{s}_{k-(m-1)\tau_d} \\ \vdots \\ \hat{s}_{k-\tau_d} \\ \hat{s}_k \end{bmatrix} \right) + b, \tag{18}$$

where $\boldsymbol{z}_l = \boldsymbol{s}_l - \boldsymbol{e}_l$.

The mapping function $\varphi(\cdot)$ can be paraphrased by a kernel function $K(\cdot,\cdot)$ because of the application of Mercer's theorem, which means that

$$K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j), \tag{19}$$

with radial basis function (RBF) kernels one employs

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|}{2\sigma^2} \right). \tag{20}$$

4. Simulation and results

In the following procedure, we use the laser data of the Santa Fe competition. The data set consists of laser intensity collected from a laboratory experiment.[15]

The reconstructed embedding phase space has been discussed in Section 2, the choice of the embedding dimension m and time delay τ_d is the first problem faced. Cao proposed a method to determine the minimal sufficient embedding dimension using the improved false nearest neighbour method.[16] The time-delayed mutual information was suggested by Fraser and Swinney[17] as effective tool to determine a reasonable τ_d. A certain value of τ corresponding to the minimum value of mutual information is a good candidate for a reasonable time delay. For the model structure, we estimated the embedding dimension of the Santa Fe data as being 4, by the Cao's algorithm. In our simulation, we have not controlled τ_d, the delay parameter of the Takens' embedding, so the sampling gives by default $\tau_d = 1$.

The kernel width σ and the regularization parameter γ must be decided when we use the RBF kernel with MRLS-SVM. In this paper, we adopt a simple method to determine these parameters based on the cross-validation idea. We define two data sets, namely the training set and the validation set from the observed time series, respectively. For each (σ, γ) combination, we perform cross-validation on the training and validation data sets. Then we choose optimal (σ, γ) based on the best cross-validation performance for each (σ, γ) combination.

The time series was divided into two subsets referred to as the training and test subset: the training subset consists of 1000 time entries and the test subset consists of 1500 entries. In order to solve the constrained nonlinear optimization problem (11) and (8), sequential quadratic programming (SQP) has been applied.

Figure 1 shows the prediction results of MRLS-SVM and RLS-SVM respectively. The prediction time series is able to follow with a very small error the Santa Fe time series up to 500 steps for MRLS-SVM and only up to 100 steps for RLS-SVM. However, the prediction error does not seem to help answer the original question of finding a criterion to establish if the MRLS-SVM was able to learn the dynamic model. Therefore we resorted to the estimation of the dynamical invariants of motion from the predicted model. The Poincaré maps for the Santa Fe time series are shown in Fig.2, where the original attractor (Fig.2(a)) is compared with the reconstructed ones using MRLS-SVM (Fig.2(b)) and RLS-SVM (Fig.2(c)), respectively. The similarity of the Fig.2(a) and Fig.2(b) suggests that the two time series represent two trajectories on the same attractor. However, a distinct difference exists between Fig.2(a) and Fig.2(c).

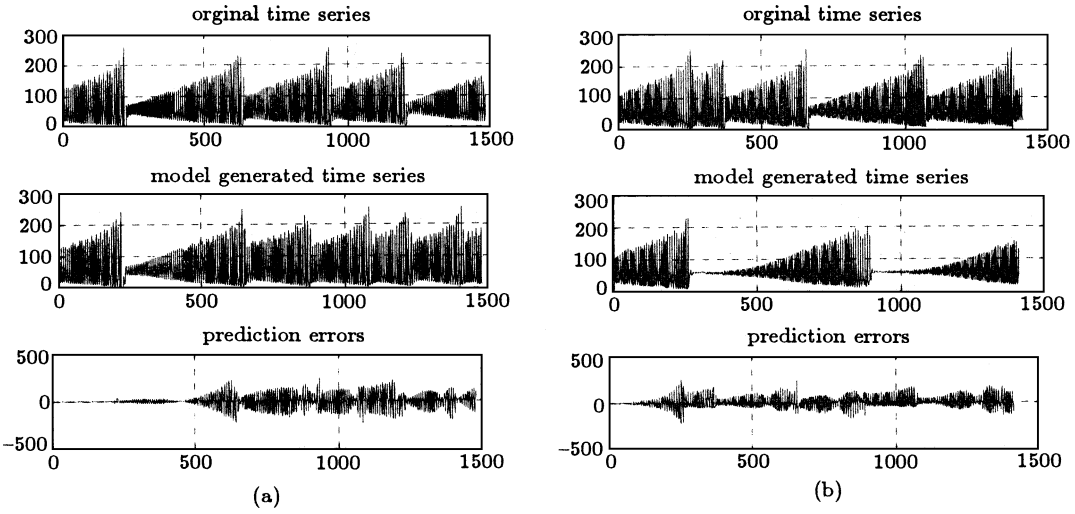


Fig.1. Prediction results for Santa Fe competition: (a) MRLS-SVM algorithm, (b) RLS-SVM algorithm.

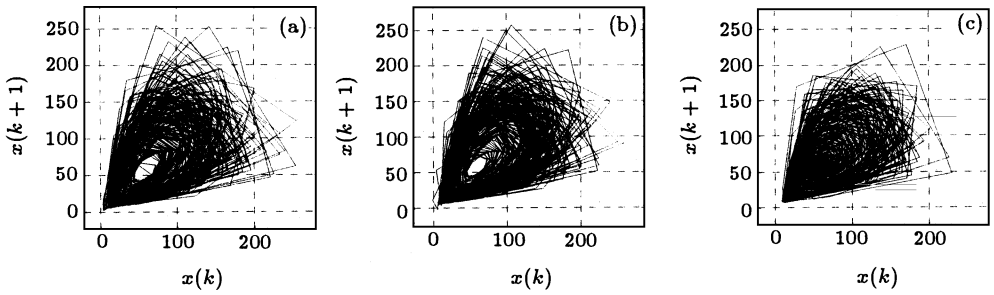


Fig.2. Poincaré maps: (a) original system, (b) MRLS-SVM reconstruction, (c) RLS-SVM reconstruction.

We have computed the spectrum of the original and reconstructed Santa Fe time series in Fig.3. The spectrum of the reconstructed time series of MRLS-SVM seems to follow very closely the spectrum of the original time series (Fig3.(a)), with only minor differences in the fine detail of the spectrum. However, a distinct difference exists between the original spectrum and the reconstructed time series of RLS-SVM (Fig3.(b)). Although these plots lead us to believe that the dynamics have been reasonably captured by the MRLS-SVM algorithm, for the chaotic time series case, the value of correlation dimension and the largest Lyapunov exponent is necessary. The correlation dimension was computed using the Grassberger and Proccacia algorithm.[13] In Fig.4 the correlation integral map (CIM) and its slope are depicted for both the original and reconstructed time series. The value of the correlation dimension is defined as the slope of the CIM curves for at least three consecutive embeddings. In this case the correlation dimension for the predicted time series is 2.147 for the region $3 < r < 5$. The correlation dimension for the original time series is 2.038 for $3 < r < 5$. In Fig.5 the largest Lyapunov exponent was also computed using Wolf's procedure for different embedding dimensions.[12] The Lyapunov exponent of the reconstructed time series seems to follow very closely the spectrum Lyapunov exponent of the original time series. In addition, for comparing the performance of RLS-SVM and MRLS-SVM, we computed the test error of the correlation dimension and the largest Lyapunov exponents of the time series in Fig.6. This plot shows that the proposed weighting component can improve the ability to learn the dynamics of the chaotic system effectively.

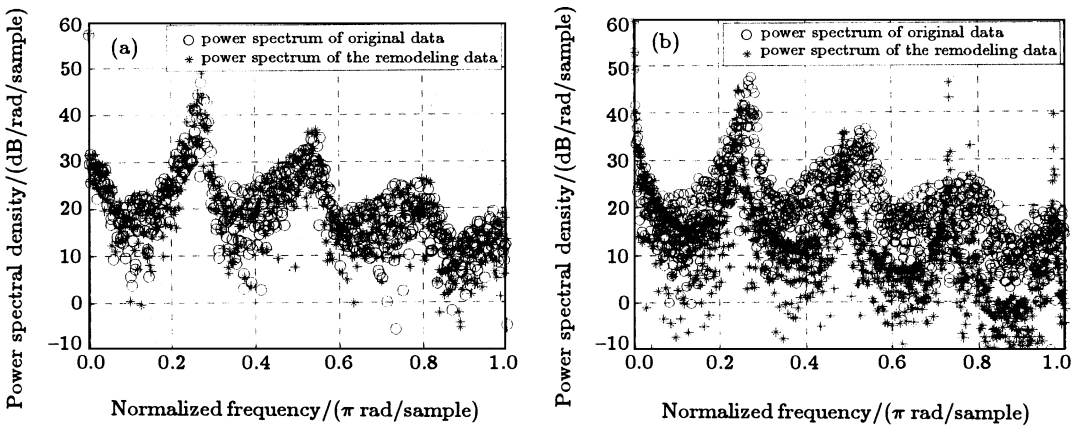


Fig.3. Power spectrum for the original data and MRLS-SVM remodelling data of Santa Fe competition: (a) MRLS-SVM algorithm (b) RLS-SVM algorithm.

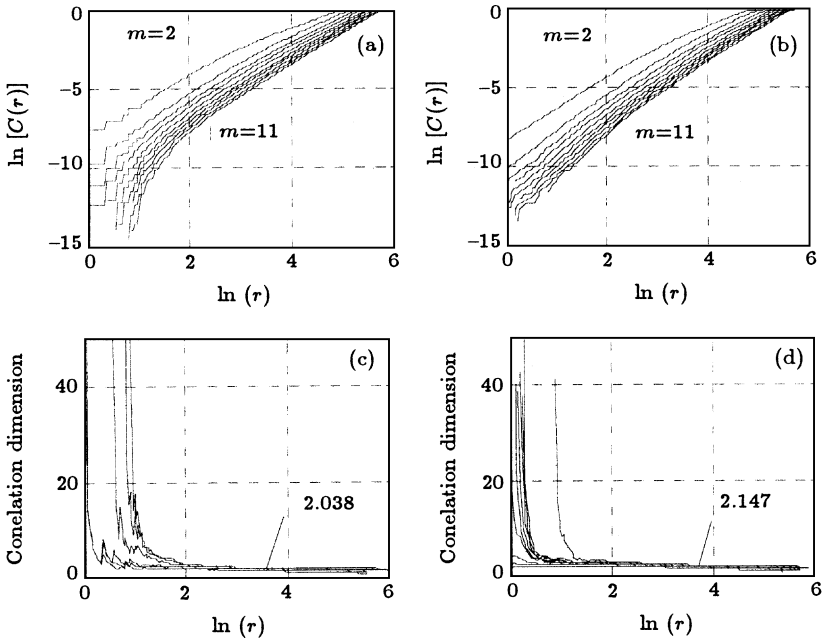


Fig.4. Correlation dimension estimates for embedding dimensions 2–11 (a) correlation integral map (CIM) for original time series, (b) CIM for series generated for MRLS-SVM, (c) slope (correlation dimension) estimate of CIM for original time series, (d) slope estimate of CIM for MRLS-SVM.

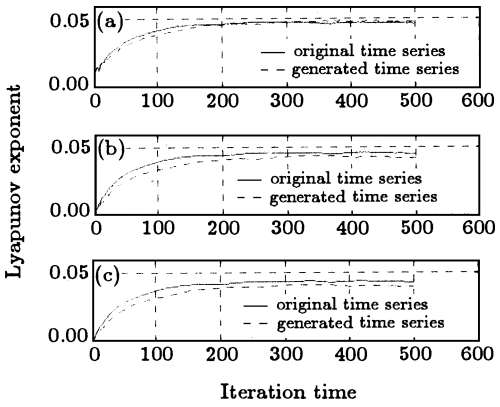


Fig.5. Estimation of largest Lyapunov exponent for original and generated time series. (a) m=2, (b) m=4, (c) m=6.

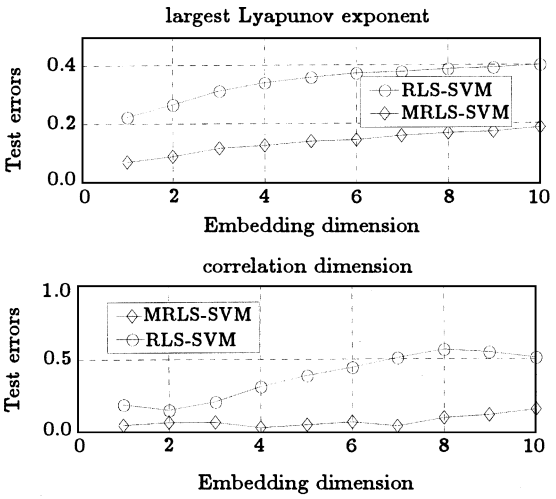


Fig.6. Test errors of dynamical invariants of Santa Fe.

Analysing the dynamical invariants and the state space trajectories, we can conclude that the dynamics of the reconstructed and original time series are very similar, and that according to our definition, the dynamics of the system that produced the original time series have been reasonably well captured in the MRLS-SVM. Notice that with the MRLS-SVM one obtains an explicit representation of the dynamics.

5. Conclusion and future work

From the simulation results, we can deduce the conclusion that MRLS-SVM have the ability to capture the dynamics of nonlinear dynamical systems as was demonstrated with the system of Santa Fe competition. This conclusion is based on the fact that the largest Lyapunov exponents and the correlation dimension of the original and generated time series are very similar. We also use additional dynamical invariants, power spectrum and the Poincaré maps to evaluate the remodelling performance. In addition, the comparison of the MRLS-SVM and the RLS-SVM shows that the proposed weighting term can improve the ability to learn the dynamical invariants of a chaotic dynamical system.

Our future work will be focused on applying the techniques to real world signals superposed with wideband noise. Because the noise distorts the trajectories in state space, and make the learning of dynamics more difficult and biased, we must remove the noise before carrying out the learning method. However, since the traditional linear filtering technique is unfit for the wideband noise case, we should develop new filtering techniques for successful dynamical modelling.

References

[1] Kaplan D and Glass L 1995 *Understanding Nonlinear Dynamics* (New York: Springer)

[2] Brillinger D R 1981 *Time series: Data Analysis and Theory* (New York: McGraw-Hill)

[3] Farmer J D and Sidorowich J J 1987 *Phys. Rev. Lett.* **59** 845.

[4] Li K P and Chen T L 2002 *Chin. Phys. Lett.* **19** 904

[5] Packard N H *et al* 1980 *Phys. Rev. Lett.* **45** 712

[6] Meyer Y and Ryan R D 1993 *Wavelets Algorithms and Applications* (Philadelphia, PA: SIAM)

[7] Li K P *et al* 2003 *Chin. Phys.* **12** 1213

[8] Tan W *et al* 2003 *Acta Phys. Sin.* **52** 795 (in Chinese)

[9] GanPcay R 1996 *Physica* D **89** 261

[10] Lapedes A and Farber R 1987 *Neural Information Processing System* (New York: AIP) p442

[11] Vapnik V 1995 *The Nature of Statistical Learning Theory* (New York: Springer)

[12] Suykens J A K *et al* 2002 *Least Squares Support Vector Machines* (World Scientific)

[13] Suykens J A K *et al* 2000 *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications* **47** 1109

[14] Takens F 1981 *Dynamical Systems and Turbulence* (Berlin: Springer) p366

[15] Wolf A *et al* 1985 *Physica* D **16** 285

[16] Grassberger P and Procaccia I 1983 *Phys. Rev. Lett.* **50** 346

[17] Thomsen J 1997 *Vibrations and Stability, Order and Chaos* (London: McGraw-Hill)