Vol 13 No 5, May 2004
1009-1963/2004/13(05)/0633-04

Chinese Physics

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

Determining the minimum embedding dimension of nonlinear time series based on prediction method

Bian Chun-Hua(卞春华)† and Ning Xin-Bao(宁新宝)

State Key Laboratory of Modern Acoustics, Institute for Biomedical Electronic Engineering, Nanjing University, Nanjing 210093, *China*

(Received 12 August 2003; revised manuscript received 24 September 2003)

Determining the embedding dimension of nonlinear time series plays an important role in the reconstruction of nonlinear dynamics. The paper first summarizes the current methods for determining the embedding dimension. Then, inspired by the fact that the optimum modelling dimension of nonlinear autoregressive (NAR) prediction model can characterize the embedding feature of the dynamics, the paper presents a new idea that the optimum modelling dimension of the NAR model can be taken as the minimum embedding dimension. Some validation examples and results are given and the present method shows its advantage for short data series.

Keywords: nonlinear time series, embedding dimension, NAR model, prediction
PACC: 0545, 0547,9260X

1. Introduction

The embedding dimension plays an important role in nonlinear analysis.[1] It concerns the phase space reconstruction. According to Takens' theory, the original dynamic properties can be retained under the topological transform, as long as the embedding dimension $d_E \geq 2d_2 + 1$ (d_2 is the correlation dimension of the attractor).[2] But this is only a sufficient but not necessary condition. Take the X component of the Lorenz equation for instance, d_2 is 2.06, according to Takens' theory $d_E = 6$, but actually $d_E = 3$ can satisfy the reconstruction of phase space. In realistic application, when the embedding dimension is chosen too small the phase space cannot be fully unfolded, but when it is set too large, noise will occupy a large part of the embedding space.

In this paper a summary is first made of the existing methods of determining the minimum embedding dimension. Then by analysing the relationship between the optimum nonlinear autoregressive (NAR) modelling dimension and the determination of minimum embedding dimension, a new minimum embedding dimension determining method is proposed.

2. Current methods for determining the embedding dimension

There are three basic methods that are usually used to choose minimum embedding dimension:

(1) Computing an invariant on the attractor. By increasing the embedding dimension used for the computation one notes when the value of the invariant stops changing.[3]

(2) Singular value decomposition. The procedure identifies orthogonal directions in the embedding space that may be ordered according to the magnitude of the variance of the trajectory's projection on them. The ordering is done using the singular values of the embedding. The number of these directions visited by the reconstructed trajectory, and indicated by large singular values, is an estimate of the dimension of the smallest space that contains the trajectory.[4]

(3) The method of false nearest neighbours (FNN). This was developed based on the fact that choosing too low an embedding dimension results in points that are far apart in the original phase space being moved closer together in the reconstruction space.[5,6]

The above three methods start from different aspects but reach the same aim of calculating the minimum embedding dimension. Among the above methods, FNN method is now commonly used. Certainly this method is a good approach. But it has some shortcomings. In this paper we provide an-

†E-mail: mail_abcde@yahoo.com.cn

other method to get the value of minimum embedding dimension and make a comparison with the FNN method.

3. The authors' ideas on determining the minimum embedding dimension

Our method is based on the nonlinear modelling prediction method. First let us introduce the NAR model and embedding theory.

3.1. NAR model

NAR model is a special case of the NARMA (nonlinear autoregressive moving average) model that was first introduced by Haber and Keviczky:[7]

$$y_n = G[y_{n-1}, y_{n-2}, \ldots, y_{n-k}], \tag{1}$$

where G can be any nonlinear function. G specifies the dependence of the output on the past history. Equation (1) could be used to model chaos. When G is some chaotic map (such as Lorentz, Logistic map, ...), the output will be a chaotic series. In this case we can say the system is deterministic chaotic. The NAR modelling method has been widely used in detecting nonlinearity.[8]

3.2. Time delay embedding

The most important phase space reconstruction technique is the method of time delay embedding.[2,9] Vectors in a new space, the embedding space, are formed from time delayed values of the scalar measurements:

$$y(n) = (x(n), x(n+\tau), \ldots, x(n+(d-1)\tau)), \tag{2}$$

where d is called the embedding dimension, τ is generally referred to as the delay or lag.

From Eq.(2) we can see that multivariate vectors in d-dimensional space are used to trace out the orbit of the system. The output y is determined by the past history $x(n), x(n+\tau), \ldots, x(n+(d-1)\tau)$.

3.3. The authors' prediction model

According to the above description, the output of both time the delay embedding dynamics and the NAR model are determined by the past history. The time delay embedding space is combined by multivariate vectors in d-dimensional space. So we can make a NAR model consisting of all distinct combinations of the past history of y_n up to d dimensions.

The original time series is $\{y_n\}, n = 1, 2, \ldots, N$. We use a global nonlinear function with increasing exponents to model it:

$$\begin{aligned} y_n^{\text{cal}} =& G(1, (y_{n-1} + y_{n-2} + \cdots + y_{n-k}), \\ & (y_{n-1} + y_{n-2} + \cdots + y_{n-k})^2, \\ & (y_{n-1} + y_{n-2} + \cdots + y_{n-k})^3 \cdots \\ & (y_{n-1} + y_{n-2} + \cdots + y_{n-k})^{d-1}, \\ & (y_{n-1} + y_{n-2} + \cdots + y_{n-k})^d), \end{aligned} \tag{3}$$

$n = k+1, k+2, \ldots, N$.

Spreading Eq.(3) out, we get a polynomial series. The order of its terms is: constant term, linear terms, quadratic terms, cubic terms, ..., dth-order terms. The coefficient of each term can be calculated through least squares curve-fitting. k is the memory of the system. d stands for the nonlinearity degree of the model and correspondingly it is the dimension of embedding space. Through Eq.(3) we can calculate the prediction values y_n^{cal}.

The prediction error between prediction values and original values is given by the normalized error:

$$\varepsilon = \left(\left(\sum_{n=k+1}^{N} (y_n^{\text{cal}} - y_n)^2 \right) \Big/ \left(\sum_{n=k+1}^{N} (y_n - \overline{y})^2 \right) \right)^{\frac{1}{2}},$$

where

$$\overline{y} = \left(\sum_{k+1}^{N} y_n \right) / (N-k). \tag{4}$$

The modelling procedure is the following: for each data series, first we create a model like Eq.(3) with parameters k and d, then use the least squares method to get the coefficient of each term and the prediction value (y_n^{cal}). Finally, we calculate the prediction error (ε) and figure the relation between the number of polynomial terms (r) and the error. When ε reaches the minimum and its value is small enough, then we say this r is the optimum term number (r_{optM}), and d is the optimum modelling dimension (d_{optM}).

4. Experimental verification

4.1. Common chaotic signal

We take several chaos signals, such as Lorentz,[10] Roessler,[11] logistic map,[12] Henon map [13] and two physiological systems [14] as examples.

1. The Lorentz attractor equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sigma(y - x),$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \gamma x - xz - y,$$
$$\frac{\mathrm{d}z}{\mathrm{d}t} = xy - bz.$$

Take its X component, where σ=10, γ=28, b=8/3.

2. The Roessler attractor equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -y - z,$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = x + \alpha y,$$
$$\frac{\mathrm{d}z}{\mathrm{d}t} = \alpha + xz - \mu z.$$

Take its X component, where α=0.2, μ=5.7.

3. Logistic map:

$$y(i+1) = r * y(i) * (1 - y(i))$$
$$r = 4.0$$
$$y(1) = 0.1$$

4. Henon map:

$$x(n+1) = 1 - a * x(n) * x(n) + b * y(n)$$
$$y(n+1) = x(n)$$
$$a = 1.4, b = 0.3$$
$$x(1) = 0.01, y(1) = 0.04$$

5. ECG (Electrocardiogram) data of a healthy person. The subject is a healthy 77-year-old female.

6. Heartbeat interval time series. Usually we call this the HRV (Heart rate variability) signal. The subject is a healthy 77-year-old female.

Signals 1 and 2 of the above are sampled at a proper sampling frequency in solving the differential equations. Signals 2, 4 and 6 are discrete series. Signal 5, the ECG signal, is sampled at 250Hz.

4.2. Experiments and results

1. Apply the mutual information method [15] to obtain the time delay τ. When N=500, 1000, 3000, 7000, Lorentz τ=6, 9, 11, 11; Roessler τ=14, 19, 24, 23; ECG τ=3, 2, 2, 2. The other signals are discrete series, τ=1.

2. Apply the FNN method [5] to get the minimum embedding dimension. The parameter τ for this calculation has been obtained in Procedure 1.

3. Apply the NAR modelling method to each series to get the optimum modelling dimension. Figure 1 shows the $\ln(\varepsilon) \sim r$ curve of Lorentz-X as an example. Table 1 is the results of all data series. In order to verify the modelling ability, the corresponding prediction error is recorded. For the prediction error we calculated the natural logarithm of ε and k=4.

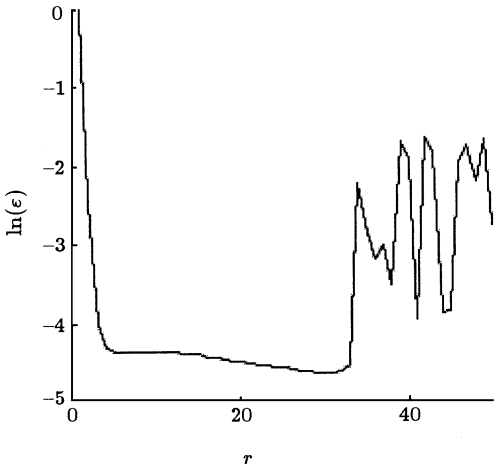


Fig.1. The prediction error with different terms numbers. k=4.

Figure 1 shows that the prediction error reaches the minimum when term number is equal to 31. When k=4 and d=1 the total term number of the NAR model is 5; when d=2 the total term number of the NAR model is 15; when d=3 the total term number of the NAR model is 35. $15 < 31 < 35$, so d_{optM}=3. We can see that the prediction error is obviously smaller in all the area of d=3 than in other areas. We think this is because when the model dimension is set too small the original dynamic information cannot be fully included, and when it is set too large, noise will occupy a large part. This is consistent with the embedding theory as we described earlier.

Table 1 shows that our method provides more stabilized results when the signal length is changing and our method can determine $d_{\mathrm{min}E}$ from short data sets while the FNN method cannot. This has realistic meaning. In application, sometimes long data sets cannot be easily obtained. Take the HRV signal for instance. For the HRV signal in Table 1, calculating $d_{\mathrm{min}E}$ through the FNN method needs at least 3000-point data length. The sampling is inconvenient, for the sampling time would be at least 30 minutes.

Table 1. Results of prediction method and comparison with that of the FNN method

Series	Signal name	N	$d_{\mathrm{min}E}$ (FNN method)	Optimum terms number	Optimum modelling dim d_{optM}	Prediction error $\ln(\varepsilon)$
1	Lorentz	500	2	32	3	–4.9413
		1000	3	31	3	–4.6442
		3000	3	32	3	–4.6636
		7000	3	32	3	–4.6554
2	Roessler	500	2	29	3	–4.7918
		1000	2	31	3	–4.7337
		3000	3	32	3	–4.7707
		7000	3	32	3	–4.6853
3	Logistic	500	2	9	2	–12.9165
		1000	2	9	2	–12.9158
		3000	2	9	2	–12.9268
		7000	2	9	2	–12.9178
4	Henon	500	2	9	2	–13.8899
		1000	2	9	2	–13.8669
		3000	2	9	2	–13.8519
		7000	2	9	2	–13.8604
5	ECG signal	500	4	32	3	–1.9241
		1000	5	41	4	–1.9319
		3000	5	34	3	–1.7962
		7000	6	41	4	–1.7123
6	HRV signal	500	/	37	4	–0.9852
		1000	/	37	4	–0.9550
		3000	5	38	4	–1.2042
		7000	6	43	4	–1.1961

5. Conclusion

This paper summarizes the current methods of determining the minimum embedding dimension and proposes that the optimum modelling dimension of the NAR model can be taken as the minimum embedding dimension. Experimental results have shown that the proposed idea is correct. The proposed method shows its advantage for short time series which could have realistic meaning.

References

[1] Yang S Q and Jia C Y 2002 *Acta Phys. Sin.* **51** 2452 (in Chinese)

[2] Takens F 1981 *Lecture Notes in Mathematics* (Berlin: Springer) vol 898 p366

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

[4] Broomhead D S and King G P 1986 *Physica* D **20** 217

[5] Kennel M, Brown R and Abarbanel H 1992 *Phys. Rev.* A **45** 340

[6] Cao L Y 1997 *Physica* D **110** 43

[7] Haber R and Keviczky L 1976 Identification of nonlinear dynamic systems *IFAC Symp. Ident. Syst. Param. Est. (Tbilisi, Georgia)* p62

[8] Chon K H *et al* 1997 *Physica* D **99** 471

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

[10] Guan X P *et al* 2001 *Acta Phys. Sin.* **50** 26 (in Chinese)

[11] Chen S H *et al* 2002 *Acta Phys. Sin.* **51** 749 (in Chinese)

[12] Gordon W B 1996 *Math. Mag.* **69** 118

[13] Li L X *et al* 2001 *Acta Phys. Sin.* **50** 629 (in Chinese)

[14] Iyengar N *et al* 1996 *Am. J. Physiol.* **271** 1078

[15] Fraser A M and Swinney H L 1986 *Phys. Rev.* A **33** 1134