Vol 12 No 6, June 2003
1009-1963/2003/12(06)/0594-05

Chinese Physics

© 2003 Chin. Phys. Soc.
and IOP Publishing Ltd

Determining the input dimension of a neural network for nonlinear time series prediction*

Zhang Sheng(张 胜)[a)b)], Liu Hong-Xing(刘红星)[a)],
Gao Dun-Tang(高敦堂)[a)], and Du Si-Dan(都思丹)[a)]

[a)] *Department of Electronic Science and Engineering, Nanjing University, Nanjing* 210093, *China*

[b)] *Department of Physics, Nanjing Normal University, Nanjing* 210097, *China*

(Received 18 November 2002; revised manuscript received 9 February 2003)

Determining the input dimension of a feed-forward neural network for nonlinear time series prediction plays an important role in the modelling. The paper first summarizes the current methods for determining the input dimension of the neural network. Then inspired by the fact that the correlation dimension of a nonlinear dynamic system is the most important feature of it, the paper presents a new idea that the input dimension of the neural network for nonlinear time series prediction can be taken as an integer just greater than or equal to the correlation dimension. Finally, some validation examples and results are given.

Keywords: nonlinear time series, prediction, phase space reconstruction, neural network, input dimension

PACC: 0545, 0555, 9260X

1. Introduction

Time series prediction plays an important role in areas of weather forecasting, economy and engineering predicting, and so on. Most traditional time series models, such as autoregressive, moving averaging and autoregressive moving averaging, are of linear types, but there are also many nonlinear time series prediction problems in practical situations. The artificial neural network (ANN), as a powerful tool of nonlinear time series modelling, has already been used in solving nonlinear prediction problems.[1−3] However, deep and systematic research in this area is still necessary and wanted.

ANN modelling involves several elements such as neuron model, network topological structure and learning arithmetic. The determining of an ANN input dimension is an important part of the network topological structure, and depends on the actual situation. The input dimension must be carefully selected when modelling a nonlinear time series. Too small input dimension will lead to divergence of the network, and too large a value will result in increasing the computation load, and even more, lead to over-fitting of ANN. An optimum input dimension should be chosen, such that while making enough modelling capability available, it is kept as small as possible.

In this paper a summary is first made on the existing methods of determining ANN input dimension. Then by analysing the relationship between the correlation dimension of a nonlinear dynamic system and the determining of input dimension in nonlinear time series ANN prediction, a new input dimension determining method for ANN model is proposed.

2. Current methods for determining the input dimension

Time series prediction is to predict a value at certain future time according to its historic values. The key for doing this is to obtain a mathematical model to describe the relationship between time series values. Mark a nonlinear time series as $\{x_n\}$, its mathematical model can be described as

$$x_{n+k} = f(x_n, x_{n-\tau}, \cdots, x_{n-(m-1)\tau}). \qquad (1)$$

Here k=1 means a single step prediction model, and $k > 1$ means a multi-step prediction model; τ is the time delay. In this paper, only the single step model of k=1, τ=1 is analysed, that is,

$$x_{n+1} = f(x_n, x_{n-1}, \cdots, x_{n-(m-1)}). \qquad (2)$$

*Project supported by the National Natural Science Foundation of China (Grant Nos 59905011 and 60275041).

To model a nonlinear time series with the widely used feed-forward neural network means to fit the above nonlinear self-regression function $f(^*)$. The network can be illustrated as Fig.1.

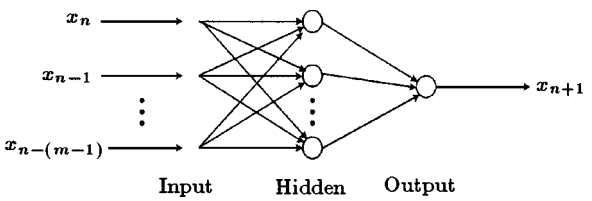


Fig.1. Structure of a feed-forward ANN for nonlinear time series prediction. Here m is the input dimension of the ANN, x_{n+1} is the output of the network as the value to be predicted, $x_n, x_{n-1}, \cdots, x_{n-(m-1)}$ are known historic values. The ANN should first be trained using sufficiently many known input and output values before it can be used to predict.

2.1. Determining the ANN input dimension by trials

Some instructive discussion has been done in determining the input dimension of predicting ANN, and the intuitive method is the method of trial and error. That is, the optimum network input dimension is determined by trying a variety of ANN input dimensions and choosing the least one meeting the requirement of training the ANN.[4] This method is simple and direct, but obviously less efficient.

2.2. Determining the ANN input dimension based on Takens' theory

Reference [5] indicates that, because the nonlinear time series prediction using ANN is based on the correlation of the nonlinear time series in time-delay phase space, the input dimension of ANN should be at least equal or greater than the embedding dimension of time series phase space reconstruction, so that certain intrinsic properties, such as some geometrical invariants, attractor dimension, measure entropy and positive Lyapunov exponents, can be retained. According to the Takens theory,[6] the original dynamic properties can be retained under the topological transform, as long as the embedding dimension $d \geq 2D_2 + 1$ (D_2 is the correlation dimension of system attractor and can be obtained by using GP arithmetic[7]). Therefore, the input dimension m of ANN can be taken as an integer not smaller than $2D_2$+1. For instance, in the case of using X component of the Lorenz equation as a nonlinear time series to predict, the correlation dimension D_2 is 2.06, so by Takens theory, the input dimension of this ANN model can be taken as 6.

2.3. Determining the ANN input dimension as minimum embedding dimension

The embedding dimension determined by Takens theory in section 2.2 above is enough for phase space reconstruction, but it is always too big and redundant. This dimension is here called the full embedding dimension d_{E} for distinguishing purpose. Much nonlinear theory research[8−11] indicate that the minimum embedding dimension $d_{\min\ \mathrm{E}}$ for phase space reconstruction does not necessarily satisfy Takens theory, that is, the condition $d_{\min\ \mathrm{E}} \geq 2D_2$+1 is not met. For example, in the case of X component of the Lorenz series, the embedding dimension is 6 by Takens theory, but actually only 3 by the method of false neighbours[12] and Cao arithmetic[13] is satisfactory. Hence, according to the Ref.[5], the input dimension of the ANN can be taken as the minimum embedding dimension of the phase space reconstruction, that is, for the X component of the Lorenz series, the input dimension of ANN model should be 3.[14]

3. The author's ideas on determining the input dimension

In current methods of determining the ANN input dimension, the minimum embedding dimension is smaller than that required by the Takens theory, since it is the smallest Euclidean space dimension containing the system attractor. However, there is another problem in phase space reconstructions, that is the nonlinear systems having the same correlation dimension often do not have the same minimum embedding dimension.

Take a simple example to illustrate the problem. In general, the correlation dimension of a system attractor is close to the topological dimension of the attractor, with a difference of only a fraction of a dimension. Approximately, it is reasonable to take the topological dimension of the attractor as the correlation dimension of the system. When the attractor is a straight line, its topological (or correlation) dimension is 1, the same as that of the minimum embedding dimension that can contain itself. But in the case that the attractor is a 3-D helix, its topological (or correlation) dimension remains 1, even though its minimum embedding dimension is 3 in this case.

In some cases, the minimum embedding dimension of a system's attractor is probably much bigger

than its correlation dimension. Facing the two different dimensions, a natural idea is to determine the input dimension of the ANN model directly from the correlation dimension, so as to reduce the input dimension of the model. This is also reasonable from the point of nonlinear system theory analysis, because it is the correlation dimension rather than the minimum embedding dimension that truly reflects the essence of the nonlinear system.

From the above discussion, it is proposed that a proper and practical way is to give the input dimension a value a little bit greater than the system's correlation dimension, which should enable the network to fit the intrinsic properties of nonlinear system. Since the input dimension must be an integer, the input dimension of ANN model should be the upper integer of the correlation dimension.

4. Experimental verification

4.1. Benchmark signals

Several chaos signals, such as Lorenz,[15] Roessler,[16] Logist,[17] Henon[18] and two machine-fault signals[19] have been sampled as nonlinear time series. Parameters, such as full embedding dimension d_{E}, minimum embedding dimension $d_{\min\ \mathrm{E}}$ and correlation dimension D_2 have been computed for all signals. Prediction experiments have been applied on the signals with ANN. In order to verify the optimum input dimension for the network models, the input dimension is varied and corresponding prediction MSE (mean squared error) is calculated.

1) The Lorenz attractor equation:

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= \sigma(y-x),\\ \frac{\mathrm{d}y}{\mathrm{d}t} &= \gamma x - xz - y,\\ \frac{\mathrm{d}z}{\mathrm{d}t} &= xy - bz.\end{aligned} \tag{3}$$

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

2) The Roessler attractor equation:

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= -y-z,\\ \frac{\mathrm{d}y}{\mathrm{d}t} &= x+\alpha y,\\ \frac{\mathrm{d}z}{\mathrm{d}t} &= \alpha + xz - \mu z.\end{aligned} \tag{4}$$

Take its X component, where α=0.2, μ=5.7; $x(0)$=0, $y(0)$=0.01, $z(0)$=–0.01.

3) Logist mapping:

$$x(n+1) = 4^{*}x(n)^{*}(1-x(n)). \tag{5}$$

where

$$x(1) = 0.1.$$

4) Henon mapping:

$$\begin{aligned}x(n+1) &= 1 + y(n) - \alpha x(n),\\ y(n+1) &= \beta x(n).\end{aligned} \tag{6}$$

Take its X component, where $\alpha = -1.4$, β=0.3; $x(1)$=0.01, $y(1)$=0.04.

5) A real axial vibration signal of a rotating machine under oil whirl fault[19]

As oil whirl fault happens, the complex oil turbulent flow may occur, so the axial vibration signal is complex nonlinear.

6) A real axial vibration signal of a rotating machine under base loose fault[19]

In the case of a base loose fault mode of mechanical rotor system, the base will jump up and down due to the unbalanced force on the rotor. This jumping leads to variation of the system's stiffness, accompanying certain striking phenomena, resulting in pseudo-cycle or chaotic motion in some particular parameter range.

Signals 1 and 2 above are sampled at a proper sampling frequency in solving the differential equations. Signals 3 and 4 are discrete series obtained from the iterative equation. Signals 5 and 6 are vibration signals recorded from real large machine sets. All signals mentioned above have complex nonlinear dynamic properties.

4.2. Experiments

1. Apply the method of Kennel[12] to obtain the time delay τ for phase space reconstruction for series 1,2,5,6, the resulting values are 7,11,2,3, respectively, signals 3 and 4 are discrete series, so τ=1.

2. Apply GP arithmetic[7] to obtain the correlation dimensions of all series (see Table 1).

3. Apply Cao arithmetic[13] to obtain the minimum embedding dimension of $d_{\min\ \mathrm{E}}$ for every series (see Table 1).

4. Increase ANN input dimension m from 1. Take hidden neuron number big enough as 20, that is, take m–20–1 as the ANN structure. Combine Trainbr arithmetic and Early stopping methods in Matlab to train the feed-forward ANN to prevent over-fitting. Results are listed in Table 1.

Table 1. Predicting results under different input dimensions.

Series	Signal name	Correlation dim. D_2	Correlation dim. upper integer	Min. embedding dim. $d_{\min\,\mathrm{E}}$	Full embedding dim. d_{E}	Network input dim. m	Predicting MSE
1	Lorenz	2.02	3	3	6	1	0.79
						2	0.022
						3	0.0036
						4	0.0034
						5	0.0031
2	Rossler	2.10	3	3	6	1	0.27
						2	0.0037
						3	0.00088
						4	0.00098
						5	0.00083
3	Logist	0.97	1	2	3	1	0.0000412
						2	0.000192
						3	0.0000723
4	Henon	1.70	2	2	5	1	0.19
						2	0.0022
						3	0.0012
						4	0.0011
						5	0.0046
5	Oil whirl fault signal	1.24	2	4	4	1	19.03
						2	2.33
						3	2.25
						4	1.96
						5	2.04
6	Base loose fault signal	1.98	2	4	5	1	42.31
						2	16.39
						3	16.56
						4	15.79
						5	15.90

4.3. Results

Table 1 shows that the accuracy of prediction is stabilized at certain level as the input dimension of ANN model increases. The accuracy reaches or approaches the optimum value when taking the input dimension as the correlation dimension or its upper-integer. In the case of signals 1 and 2, the correlation dimensions are both just a little bigger than 2, and their accuracies have increased markedly to a very high level as their input dimensions reach 2. It reveals that the correlation dimension is the true reflection of a system, and that determining the ANN input dimension based on the correlation dimension is reasonable. Signals 5 and 6 are real and contain noise, so the resultant final MSEs of the networks are big.

5. Conclusion

Current methods of determining input dimension for nonlinear time series ANN prediction model have been summarized and analysed. Based on further analysing the nonlinear dynamic theories, it has been proposed that the input dimension of the predicting ANN can be taken as the upper integer of the system correlation dimension. Experimental results show that the proposed idea is correct. The correlation dimension reflects the system's essence, better than the embedding dimension of phase space reconstruction does.

References

[1] Yuan J and Xiao X C 1997 *Acta Phys. Sin.* **46** 2095 (in Chinese)

[2] Guan X P *et al* 2001 *Acta Phys. Sin.* **50** 2112 (in Chinese)

[3] Thiesing F M and Vornberger O 1997 IEEE (ed) *Proc. ICNN'97* 2125

[4] Li Q and Zheng D L 1999 *J. Univ. Sci. Technol. Beijing* **21** 90 (in Chinese)

[5] Wang D S and Cao L 1995 *Chaos Fractal and Their Applications* (Hefei: University of Science and Technology of China Press) 403 (in Chinese)

[6] Takens F 1981 *Lecture Notes in Mathematics* **898** 366

[7] Grassberger P and Procaccia I 1983 *Physica* D **9** 189

[8] Mees A I, Rapp P E and Jennings L S 1987 *Phys. Rev.* A **36** 340

[9] Schroer C G *et al* 1998 *Phys. Rev. Lett.* **80** 1410

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

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

[12] Kennel M B, Brown R and Abarbanel H D I 1992 *Phys. Rev.* A **45** 3403

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

[14] Oliveira K A, Alvaro V and Silva E C 2000 *Physica* A **284** 393

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

[16] Feng G L *et al* 2001 *Acta Phys. Sin.* **50** 606 (in Chinese)

[17] Yuan J and Xiao X C 1998 *Acta Phys. Sin.* **47** 897 (in Chinese)

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

[19] Jiang J D and Qu L S 1998 *J. Xi'an Jiaotong University* **32** 27 (in Chinese)