Fifth IFSA World Congress (1993), 857-860

Dimension Analysis of Chaotic Time Series Using Self Generating Neuro Fuzzy Model

Ryu Katayama, Kaihei Kuwata, Yuji Kajitani, Masahide Watanabe, Yukiteru Nishida

Information & Communication Systems Research Center, Sanyo Electric Co.,Ltd.
1-18-13, Hashiridani, Hirakata, Osaka, 573, Japan
E-mail: katayama@rd.sanyo.co.jp

Abstract

In this paper, we apply the self generating neuro fuzzy model (SGNFM) to the dimension analysis of the chaotic time series. Firstly, we formulate a nonlinear time series identification problem with nonlinear autoregressive (NARMAX) model.

Secondly, we propose an identification algorithm using SGNFM. We apply this method to the estimation of embedding dimension for chaotic time series, since the embedding dimension plays an essential role for the identification and the prediction of chaotic time series. In this estimation method, identification problems with gradually increasing embedding dimension are solved, and the identified result is used for computing correlation coefficients between the predicted time series and the observed one. We apply this method to the dimension estimation of a chaotic pulsation in a finger's capillary vessels.

1. Introduction

The identification and prediction of a system is a fundamental task, not limited only to the field of engineering, but in all kinds of other fields as well. So far, studies on identification and prediction of time series using neural networks and radial basis function have been reported [2,6]. For the identfication and prediction of these kinds of nonlinear dynamical systems, the time series which we can actually measure are often time series of only one variable. Therefore, to identify the original dynamical system, we need to reconstruct the system state space based on the embedding theorem[3] of Takens. In doing that, we have to appropriately determine the embedding dimension and delay time. In regard to the dimension estimation, Grassberger and Procaccia [4] suggested a method of estimating the original system's dimension in seeking by use of the correlation dimension. Also Sugihara and May [5] proposed a technique for estimating the original system's dimension from the correlation coefficient between the predicted and observed data for a model determined by using a reconstructd state space vector. Grassberger's method with the correlation dimension requires a time series of 10^4 order. But in contrast with the Grassberger's method, the advantage of Sugihara's method is that it requires fewer data values to estimate the dimension.

In this paper, we apply the self generating neuro fuzzy model (SGNFM) to the dimension analysis of the chaotic time series. By using SGNFM, we can automatically determine the smallest number of basis functions required to achieve a model with the desired degree of accuracy. We apply this method to the dimension estimation of a chaotic pulsation in a finger's capillary vessels.

2. Nonlinear Autoregressive Model Using Self Generating Neuro Fuzzy Model

For a nonlinear autoregressive model, if we do not consider the control input, and if the input vector prior to time t is taken to be $\underline{y}=(y(t-1), y(t-2), \cdots, y(t-s))$, then the output y(t) at time t can be expressed by (1);

$$y(t) = f(y(t-1), y(t-2), \cdots, y(t-s)) \qquad (1)$$

To identify the function f in (1), several methods using neural networks or fuzzy model are reported. In this paper, we use self generating neuro fuzzy model (SGNFM) with gaussian type radial basis function. In SGNFM, we can automatically determine the least number of basis function numbers to accomplish a given degree of model error. Representing the nonlinear autoregressive model using a three-layered neural network with one output, then in the case where s inputs are given for the model, the unit number of the hidden layer can be determined automatically by the self generating algorithm [6,7] (Fig.1).

(a), (b), and (c) in Fig.1, are nonlinear autoregressive models for identifying the output y(t) from the previous 1 input, two inputs, and s inputs, respectively. The identification of the models is carried out by approximating the value of the functions f_1, f_2, $\cdots$, f_s in (2);

$$y(t) = f_1(y(t-1))$$
$$y(t) = f_2(y(t-1), y(t-2))$$
$$\cdots$$
$$y(t) = f_s(y(t-1), y(t-2), \cdots, y(t-s)) \quad (2)$$

The functions f_1, f_2, $\cdots$, f_s of (2) are determined by neuro fuzzy model with radial basis function as shown in (3);

$$f_1 = \sum_{k=1}^{n1} w_k \mu_k(x_1,\underline{a}_k,\underline{b}_k), \quad x_1=y(t-1)$$

$$f_2 = \sum_{k=1}^{n2} w_k \mu_k(\underline{x}_2,\underline{a}_k,\underline{b}_k), \quad \underline{x}_2=(y(t-1),y(t-2))$$

$$\cdots$$

$$f_s = \sum_{k=1}^{ns} w_k \mu_k(\underline{x}_s,\underline{a}_k,\underline{b}_k), \quad \underline{x}_s=(y(t-1),y(t-2),\ldots,y(t-s)) \quad (3)$$

where n_1, n_2, $\cdots$, n_s are basis function numbers, $\underline{a}_k \in R^s$, $\underline{b}_k \in R^s$, $k=1,\cdots,n_s$ are center value vector and width value vector of RBF, respectively, and w_k, $k=1,\cdots,n_s$ are conclusion part of k-th basis.

In addition,

$$\mu_k(\underline{x}_s,\underline{a}_k,\underline{b}_k) = \prod_{i=1}^{s} A_{ik}(y(t-i),a_{ik},b_{ik}) \quad (4)$$

$$A_{ik}(y(t-i),a_{ik},b_{ik}) = \exp(-(y(t-i) - a_{ik})^2 / b_{ik}) \quad (5)$$

A_{ik} is the radial basis function for the input component y(t-i) which has b_{ik} as its width of the distribution, and a_{ik} at its center. Various function forms for A_{ik} can be taken [7], but here we use Gaussian basis functions as in (5). Let us introduce the following parameter vectors $\underline{a}^n \in R^{sn}$, $\underline{b}^n \in R^{sn}$, $\underline{w}^n \in R^n$ defined by (6) when the number of RBF is n;

$$\underline{a}^n=(\underline{a}_1,\cdots,\underline{a}_k,\cdots,\underline{a}_n) \in R^{sn}$$
$$\underline{b}^n=(\underline{b}_1,\cdots,\underline{b}_k,\cdots,\underline{b}_n) \in R^{sn}$$
$$\underline{w}^n=(w_1,\cdots,w_k,\cdots,w_n) \in R^{sn} \quad (6)$$

Then, the nonlinear identification problem by NFM is to determine the minimal number of basis n and optimal model parameter values which satisfy

$$E_n(*\underline{a}^n,*\underline{b}^n,*\underline{w}^n) < \varepsilon \quad (7)$$

where

$$E_n(\underline{a}^n,\underline{b}^n,\underline{w}^n) = \frac{1}{2}\sum_{p=1}^{N} (y^p - *y^p)^2 \quad (8)$$

and $(*\underline{a}^n,*\underline{b}^n,*\underline{w}^n)$ are optimal solution which minimizes E_n with n fuzzy rules or n hidden units in hidden layer, and y^p, $*y^p$ are p-th output data and inference output, respectively.

To solve the above problem, we proposed a self generating algorithm called MAE (Maximum Absolute Error) selection method [6,7]. By using MAE method, we can automatically determine the minimum number of basis functions to satisfy the inequality (7), where ε is a any desired model accuracy specified by the designer. This method is very effective especially for this kind of dynamical system identification problem, since so many model structure identification problems and model parameter optimization problems with different inputs' dimension must be solved.

3. Estimating the Dimension of Nonlinear Dynamical System

Even in a dynamical system that can be expressed by multi-dimensional state variables, what we can actually measure in our expriments is often time series data of just one variable. Takens [3] showed that it is possible for the attractor of the original multi-dimensional dynamical system to be restructured in a DE dimensional state space (Fig.2), using the state space vector represented in (9);

$$\underline{y}_1 = \{y(1), y(1-Td), \cdots, y(1-(DE-1)Td)\}$$
$$\cdots$$
$$\underline{y}_j = \{y(j), y(j-Td), \cdots, y(j-(DE-1)Td)\}$$
$$\cdots$$
$$\underline{y}_N = \{y(N), y(N-Td), \cdots, y(N-(DE-1)Td)\} \quad (9)$$

where DE is an embedding dimension and Td is a appropriate delay time, which is determined by the autocorrelation of the original time series y(t). Suitable value of Td is such a time at which the value of autocorrelation of the y(t) becomes 0 initially.

Now for the dimension D of the original dynamical system, if the embedding dimension DE equals 2D+1 or more, then the attractor of the original dynamical system can be embedded in the reconstructed state space.

Sugihara and May suggest a technique for estimating the dimension of the original nonlinear dynamical system using identification and prediction procedure. In this estimation method, identification problems with gradually increasing embedding dimension are solved, and the identified result is used for computing correlation coefficients between the predicted time series and the observed one. The correlation coefficien ρ is defined by (10) between original time series y(t) and predicted time series y'(t), t=1,···,M;

$$\rho = (1/M)\sum_t (y(t)-y_m)(y'(t)-y'_m) \Big/ \sqrt{(1/M)\sum_t (y(t)-y_m)^2}\sqrt{(1/M)\sum_t (y'(t)-y'_m)^2} \quad (10)$$

where y_m, y'_m are mean value of y(t), y'(t), t=1,···,M, respectively. The most suitable embedding dimenston D_E is such that it gives the maximal correlation coefficient value.

We applied above method to estimate the dimension of a human pulse wave. In Fig.3, we show a normalized time series for a human pulse measured during a thirty-second interval, with a sampling time of 19.5 msec, and with the amplitude pulse values normalized between a range of -1.0 to 1.0. The number of samples is 1792. In identifying the model, 600 samples were taken from the first half of the pulse wave time series as learning data. In this learning data, the delay time Td is taken as 234 msec. (19.5 msec. x 12, see Fig.4). By modifying the embedding dimension from 1 to 15, we used a reconstructed state space vector, identified models with different dimensions. Total learning iterations is 5,000.

By using thus identified models, the predicted time series y'(t) for different embedding dimension D_E are computed. Then, we calculate correlation coefficient ρ between the actually measured and predicted values obtained by using six hundred sets of data with the latter half of the model. We show the relation between ρ and D_E in Fig.5. As shown in Fig.5, since the correlation is highest when the embedding dimension D_E is in the 4-6 range, it can be said that the dimension of the the pulse wave's dimension is in the 2-3 range based upon the Takens' Theorem.

4. Conclusion

In this research, we have discussed a method for calculating the dimension of a nonlinear dynamical system, based on a time series from one measured variable, using the correlation coefficient between predicted and measured values of a nonlinear autoregressive model. We applied this method to estimate the a chaotic pulsation in a finger's capillary vessels.

For a time-varying nonlinear dynamical system, it can be said that the statistical characteristics of a measured time series are constantly changing. In this kind of a situation, it is necessary to calculate the dimension by varying the structure of the model to adapt to that structure change of the system. One of the future research direction is to develop a method for identifying a time-varying nonlinear dynamical system while considering the fractal dimension of time series and simultaneously modifying the structure of the model accordingly.

[References]

[1] M.J.D.Powell: "Radial basis function approximations to polynomials", Proc. 12th Biennial Numerical Analysis Conf. (Dundee), pp.223-241, 1987.

[2] S.Chen, S.A.Billings, C.F.N.Cowan, P.M.Grant: "Practical identification of NARMAX models using radial basis functions", Int.J.Control, vol.52, No.6, pp.1327-1350, 1990.

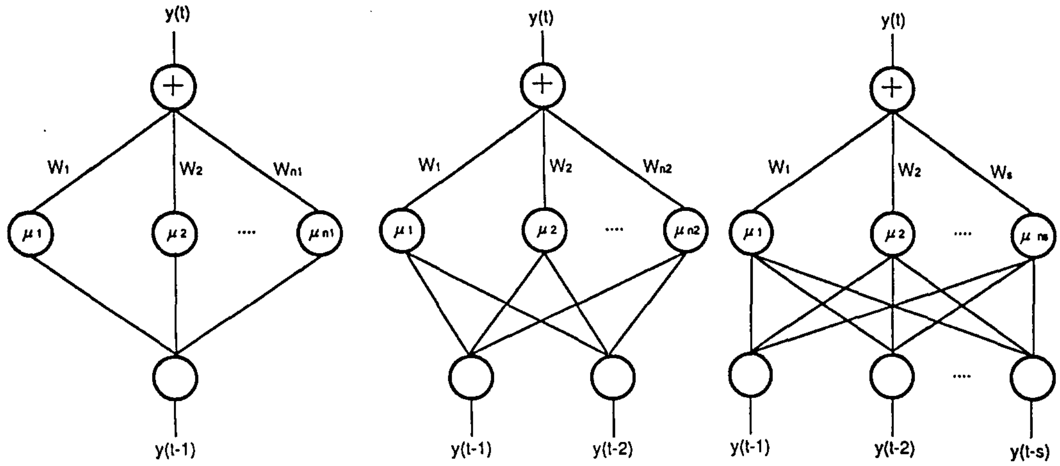
[3] F.Takens:"Detecting Strange Attractors in Turbulence", Lect. Notes in Mathematics, 898, pp.366-381, 1991.

[4] P.Grassberger, I.Procaccia:"Characterization of Strange Attractor", Phys. Rev. Lett., 50, pp.346-349, 1983.

[5] G.Sugihara, R.May:"Nonlinear Forcasting as a Way of Distinguishing Chaos from Measurement Error In Time Series", Nature, Vol.344, No.19, pp.734-741, 1990.

[6] R.Katayama,Y.Kajitani,K.Kuwata,Y.Nishida: "Self Generating Radial Basis Function as Neuro-Fuzzy Model and its Application to Nonlinear Prediction of Chaotic Time series" , Proc. of the Second IEEE International Conference on Fuzzy Systems, San Francisco, pp.407-414, 1993.

[7] R.Katayama,Y.Kajitani,K.Kuwata,Y.Nishida: "Comparative Analysis on Learning Ability of Self Generating Algorithm for Neuro Fuzzy Model with Various Radial Basis Functions", Proc. of the Fifth IFSA, Seoul, Korea, 1993.



(a) 1 input number (b) 2 input numbers (c) s input numbers

Fig.1 Nonlinear Autoregressive Models
Using Self Generating Neuro Fuzzy Model

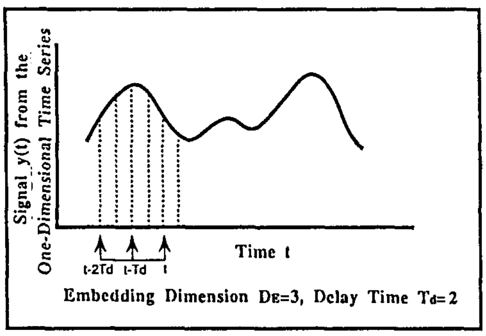


Fig.2 Reconstruction of the State Space from the Time Series

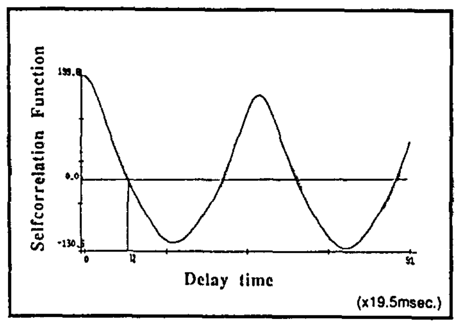


Fig.4 Autocorrelation Function of a Pulse Wave Time Series

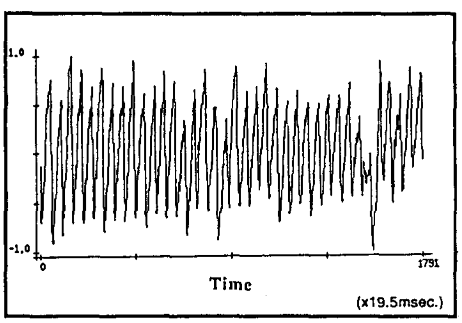


Fig.3 Time Series of a Pulse Wave

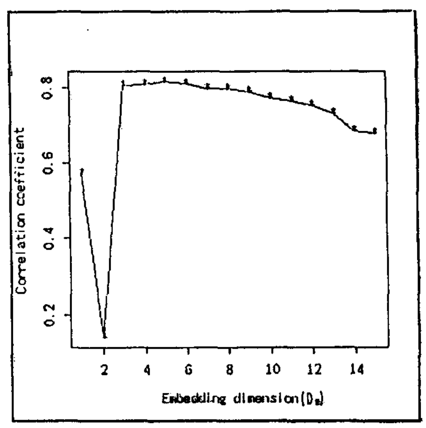


Fig.5 Correlation Coefficient vs Embedding Dimension between Observed and Predicted Time Series