

A prediction comparison between univariate and multivariate chaotic time series

Wang Haiyan¹ Zhu Mei²

(¹College of Economics and Management, Southeast University, Nanjing 210096, China)

(²Department of Mathematics, Southeast University, Nanjing 210096, China)

Abstract: The methods to determine time delays and embedding dimensions in the phase-space delay reconstruction of multivariate chaotic time series are proposed. Three nonlinear prediction methods of multivariate chaotic time series including local mean prediction, local linear prediction and BP neural networks prediction are considered. The simulation results obtained by the Lorenz system show that no matter what nonlinear prediction method is used, the prediction error of multivariate chaotic time series is much smaller than the prediction error of univariate time series, even if half of the data of univariate time series are used in multivariate time series. The results also verify that methods to determine the time delays and the embedding dimensions are correct from the view of minimizing the prediction error.

Key words: multivariate chaotic time series; phase space reconstruction; prediction; neural networks

In practical problems, complex systems exist everywhere. They usually contain many variables. In many realistic situations, it is hard to build up exact analytic models for complex systems because their constructions are very intricate and the information available is also incomplete and inaccurate. Complete systems are usually analyzed by time series observed or measured from the systems^[1-3]. Meanwhile, multivariate time series are available in many practical situations, for example, physiological data and economical data. Porporato and Reick, et al. studied local predictions of multivariate time series in hydrological and biological system^[4,5], respectively. However, they didn't take the selection of embedding parameters for multivariate time series into account. Based on the selection methods of embedding parameters for multivariate time series^[1-3], we systematically study three nonlinear prediction methods for multivariate chaotic time series, including local mean prediction, local linear prediction and BP neural networks prediction. The prediction errors of univariate time series and multivariate ones are compared by the Lorenz system with three prediction methods. The results also verify that the methods to determine the time delays and the embedding dimensions are correct from the view of minimizing the prediction error.

1 Phase Space Reconstruction^[1]

Suppose that we have an M -dimensional time

series $\{\mathbf{x}_n\}_{n=1}^N = \{(x_{1,n}, x_{2,n}, \dots, x_{M,n})\}_{n=1}^N$. As in the case of univariate time series (where $M = 1$), we make the time delay reconstruction:

$$\begin{aligned} \mathbf{V}_n &= \{x_{1,n}, x_{1,n-\tau_1}, \dots, x_{1,n-(m_1-1)\tau_1}; x_{2,n}, x_{2,n-\tau_2}, \\ &\dots, x_{2,n-(m_2-1)\tau_2}; \dots; x_{M,n}, x_{M,n-\tau_M}, \dots, x_{M,n-(m_M-1)\tau_M}\}^T \\ n &= J_0, J_0 + 1, \dots, N; J_0 = \max_{1 \leq i \leq M} (m_i - 1)\tau_i + 1 \end{aligned} \quad (1)$$

where $\tau_i, m_i (i = 1, 2, \dots, M)$ are the time-delays and the embedding dimensions, respectively. Following

Takens's delay embedding theorem, if $m = \sum_{i=1}^M m_i$ or each m_i is sufficiently large, there exists a mapping $\mathbf{F}: \mathbf{R}^m \rightarrow \mathbf{R}^m$, such that

$$\mathbf{V}_{n+1} = \mathbf{F}(\mathbf{V}_n) \quad (2)$$

Thus the evolution from state \mathbf{V}_n to state \mathbf{V}_{n+1} reflects the motion of the original unknown dynamics. This means that the geometrical characteristics of the attractor in the reconstructed space are equivalent to the original one. So any differential or topological invariant quantities computed for the reconstructed attractor are identical to those in the original state space.

We use the same method for choosing τ_i as Liangyue Cao's^[2], that is, we find the time delays τ_i with mutual information method separately for each univariate time series $\{x_{i,n}\}_{n=1}^N, i = 1, 2, \dots, M$. In univariate time series, a popular method that is used for choosing the embedding dimensions m_i is the so-called false nearest-neighbor method^[6]. Here we extend this method to the multivariate case.

For each V_n we find its nearest neighbor $V_{\eta(n)}$, i.e.,

$$\|V_{\eta(n)} - V_n\|^{(m_1, \dots, m_i, \dots, m_M)} = \min_{j=J_0, \dots, N; j \neq n} \|V_j - V_n\| \quad (3)$$

When the dimension is increased from m to $m+1$, for example from m_i to m_i+1 , the distance between V_n and $V_{\eta(n)}$ becomes $\|V_{\eta(n)} - V_n\|^{(m_1, \dots, m_i+1, \dots, m_M)}$. If the distance $\|V_{\eta(n)} - V_n\|^{(m_1, \dots, m_i+1, \dots, m_M)}$ at dimension $m+1$ is greatly larger than the distance $\|V_{\eta(n)} - V_n\|^{(m_1, \dots, m_i, \dots, m_M)}$ at dimension m , this may be the result of non-near points in high dimensional attractor becoming near points in a lower dimensional space. So the point $V_{\eta(n)}$ is regarded as a false neighbor. In other words, if

$$\frac{\|V_{\eta(n)} - V_n\|^{(m_1, \dots, m_i+1, \dots, m_M)} - \|V_{\eta(n)} - V_n\|^{(m_1, \dots, m_i, \dots, m_M)}}{\|V_{\eta(n)} - V_n\|^{(m_1, \dots, m_i, \dots, m_M)}} \geq R_T \quad (4)$$

where $V_{\eta(n)}$ is the false nearest neighbor of V_n ; R_T is a threshold and its value is selected according to practical situations.

For measured time series, we start at dimension m realized with a choice of an initial set of subspace dimension $\{m_i\}$ (usually one begins with $m_i = 1$, $i = 1, 2, \dots, M$) and increase some dimension m_i by one each time. When false nearest neighbors are very few or no longer decrease with dimension m increasing, it can be thought that the underlying geometrical structure has been completely unfolded. At this time, dimensions $(m_1, \dots, m_i, \dots, m_M)$ are optimum. As $M = 1$, this is just the case of false nearest-neighbor method for choosing the minimum embedding dimensions to univariate time series in Ref. [6].

2 Nonlinear Prediction

Nonlinear prediction is to construct an approximate functional form F based on time series $\{x_n\}_{n=1}^N$. Of course, F must be nonlinear, because the underlying system is chaotic. There are two main kinds of methods for constructing prediction function F . One is a local prediction method, which only makes full of near neighbors of the predicted points. And the other is a global prediction method, which uses all data to fit the prediction function.

We generalize nonlinear prediction methods from univariate time series to the multivariate ones; including local mean prediction, local linear prediction and back propagation (BP) neural networks.

2.1 Local mean prediction

Suppose that the state vector at time T is V_T and

the K nearest neighbors of state vector V_T are $V_{a_1}, V_{a_2}, \dots, V_{a_K}$. If the underlying system is determined, the future state V_{T+1} will also be close to $V_{a_{k+1}}$ ($k = 1, 2, \dots, K$) when V_T is close to V_{a_k} ($k = 1, 2, \dots, K$). Therefore, the finally accepted prediction of $x_{1, T+1}$ is then the average of all these individual predictions $x_{1, a_1+1}, x_{1, a_2+1}, \dots, x_{1, a_K+1}$, that is

$$\hat{x}_{1, T+1} = \frac{1}{K} \sum_{k=1}^K x_{1, a_k+1} \quad (5)$$

2.2 Local linear prediction

Suppose that the state vector at time T is V_T and the prediction model is $x_{1, T+1} = c_0 + c_1 x_{1, T} + c_2 x_{1, T-\tau_1} + \dots + c_{m_1} x_{1, T-(m_1-1)\tau_1} + \dots + c_M x_{M, T-(m_M-1)\tau_M}$, where c_0, c_1, \dots, c_m are the coefficients that need to be determined. To estimate the coefficients, we define K nearest neighbors $V_{a_1}, V_{a_2}, \dots, V_{a_K}$ of V_T as

$$V_{a_k} = \{x_{1, a_k}, x_{1, a_k-\tau_1}, \dots, x_{1, a_k-(m_1-1)\tau_1}; x_{2, a_k}, x_{2, a_k-\tau_2}, \dots, x_{2, a_k-(m_2-1)\tau_2}; \dots; x_{M, a_k}, x_{M, a_k-\tau_M}, \dots, x_{M, a_k-(m_M-1)\tau_M}\}^T \quad (6)$$

$k = 1, 2, \dots, K$

So the current problem that needs to be solved is how to obtain the coefficients c_0, c_1, \dots, c_m which make $\sum_{k=1}^N |x_{1, a_k+1} - (c_0 + c_1 x_{1, a_k} + \dots + c_M x_{M, a_k-(m_M-1)\tau_M})|^2$ minimal. The estimation of c_0, c_1, \dots, c_m is obtained by the least squares method. So we get the prediction of $x_{1, T+1}$ via the following evolutionary relationship:

$$\hat{x}_{1, T+1} = c_0 + c_1 x_{1, T} + c_2 x_{1, T-\tau} + \dots + c_{m_1} x_{1, T-(m_1-1)\tau} + \dots + c_M x_{M, T-(m_M-1)\tau_M} \quad (7)$$

2.3 Neural networks prediction

One particular sort of network that has been used for prediction study is the so-called feed-forward network. A three-layer feed-forward network has the capability of approximating any complicated continuous function.

Suppose that the state vector at time T is V_T and the inputs of a network are the state vectors $V_n, n = J_0, J_0+1, \dots, T, J_0 = \max_{1 \leq i \leq M} (m_i - 1)\tau_i + 1$. Suppose that the hidden layer consists of q neurons. In the case of scalar prediction, that is, finding a function \hat{f}_1 such that $x_{1, n+1} = \hat{f}_1(V_n)$, the output layer only comprises a single unit, namely $x_{1, n+1}$. After the $T - J_0 + 1$ data have been trained, a net will be obtained. Then work out the prediction $\hat{x}_{1, T+1}$ of $x_{1, T+1}$ based on the trained

net. The algorithm in which the weights of neural networks are adjusted is the Levenberg-Marquardt training method^[7].

2.4 Prediction error

Give an M -dimensional time series $\{\mathbf{x}_n\}_{n=1}^N = \{(x_{1,n}, x_{2,n}, \dots, x_{M,n})\}_{n=1}^N$, divide N into two parts T and L . The former T data are used to construct a model and the latter T data are used to make predictions, which are called the prediction set. Let the prediction of $x_{1,T+1}$ be $\hat{x}_{1,T+1}$, the prediction of $x_{2,T+1}$ be $\hat{x}_{2,T+1}$, \dots , and the prediction of $x_{M,T+1}$ be $\hat{x}_{M,T+1}$. This prediction is defined as one-step prediction. Continuing the above process, the predictions $\hat{x}_{i,T+2}$, $\hat{x}_{i,T+3}$, \dots , $\hat{x}_{i,T+L}$ of $x_{i,T+2}$, $x_{i,T+3}$, \dots , $x_{i,T+L}$ are obtained, $i = 1, 2, \dots, M$. For the purpose of simplification, we only predict the first variable.

In order to compare univariate prediction and multivariate prediction, we use the following indices, namely root mean squared prediction error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{L} \sum_{l=1}^L (x_{1,T+l} - \hat{x}_{1,T+l})^2} \quad (8)$$

and normalized mean squared prediction error (NMSE)

$$\text{NMSE} = \frac{\text{RMSE}^2}{\sigma^2} \quad (9)$$

where σ^2 is the variance of observed data $\{x_{1,T+l}\}_{l=1}^L$, and correlation coefficient r .

3 Numerical Simulation

Consider Lorenz system $\frac{dx}{dt} = \sigma(y - x)$, $\frac{dy}{dt} = x(R - z) - y$, $\frac{dz}{dt} = xy - bz$; where $\sigma = 10$, $R = 28$, $b = \frac{8}{3}$. The step length of integral is $h = 0.04$ and initial points are $x_0 = 15.34$, $y_0 = 13.68$, $z_0 = 37.91$. We use four-order Runge-Kute integral method and obtain two simulated time series of x and y individually with 12 100 data. To reduce the influence of transition, we abandon first 10 000 data and only keep the last 2 100 data. The former 2 000 data are used to train samples and the latter 100 data are used as prediction set.

For the time series $\{x_n\}_{n=1}^{2000}$, we obtain the optimal time lag $\tau = 4$ with mutual information method and the minimum embedding dimension $m = 3$ with false nearest-neighbor method. We reconstruct phase space with (x_n, x_{n-4}, y_n) . For the multivariate time series $\{(x_n, y_n)\}_{n=1}^{2000}$, we obtain $\tau_1 = \tau_2 = 4$, $m_1 = 2$,

$m_2 = 1$ ^[1]. We reconstruct phase space with (x_n, x_{n-4}, y_n) . Root mean squared prediction errors, normalized mean squared prediction errors and correlation coefficients r are shown in Tab.1. Multivariate 1 is defined as predicting x_n with $\{(x_n, y_n)\}_{n=1}^{2000}$, while multivariate 2 is defined as predicting x_n with the last 1 000 pairs of $\{(x_n, y_n)\}_{n=1}^{2000}$. The number of nearest neighbors is 12 (i.e. $K = 12$) both in local mean prediction and local linear prediction. The hidden layer consists of 8 neurons and the training times are 1 000 in neural networks prediction.

Tab.1 RMSE, NMSE and correlation coefficient r

		RMSE	NMSE	r
Local mean prediction	Univariate	0.229 7	$8.513 3 \times 10^{-4}$	0.989 6
	Multivariate 1	0.126 0	$2.562 0 \times 10^{-4}$	0.989 9
	Multivariate 2	0.179 2	$5.180 7 \times 10^{-4}$	0.989 8
Local linear prediction	Univariate	0.016 6	$4.693 1 \times 10^{-6}$	0.990 0
	Multivariate 1	0.002 7	$1.136 4 \times 10^{-7}$	0.990 0
	Multivariate 2	0.004 0	$2.587 4 \times 10^{-7}$	0.990 0
Neural networks prediction	Univariate	0.053 6	$4.634 6 \times 10^{-5}$	0.990 0
	Multivariate 1	0.002 7	$1.189 7 \times 10^{-7}$	0.990 1
	Multivariate 2	0.003 7	$2.149 0 \times 10^{-7}$	0.990 1

According to Tab.1, we can conclude that whatever nonlinear prediction method, the prediction results with multivariate reconstruction are significantly better than those with univariate reconstruction under the circumstances of the same data, even if half of the data of univariate time series are used to multivariate time series. Therefore, if the data available are not sufficient, the prediction results with multivariate reconstruction can greatly improve the prediction accuracy.

In order to further discuss the influence of reconstructing parameters to prediction, we reconstruct the state vector $(x_n, x_{n-\tau_1}, \dots, x_{n-(m_1-1)\tau_1}, y_n)$ with bivariate time series $\{(x_n, y_n)\}_{n=1}^{2000}$. Using a neural networks prediction method, fix τ_1 as 4 and change m_1 . The normalized mean squared prediction errors are shown in Tab.2. The NMSE is minimum in the case of $m_1 = 2$. By the same argument, fix m_1 as 2 and change τ_1 . The normalized mean squared prediction errors are shown in Tab.3. The NMSE is minimum in the case of $\tau_1 = 4$. All

Tab.2 NMSE as m_1 is changed

m_1	NMSE
1	$2.131 1 \times 10^{-5}$
2	$1.189 7 \times 10^{-7}$
3	$7.264 7 \times 10^{-5}$
4	$9.507 9 \times 10^{-6}$
5	$1.237 6 \times 10^{-7}$
6	$5.774 5 \times 10^{-7}$

these show that the parameters of multivariate reconstruction are proper.

Tab.3 NMSE as τ_1 is changed

τ_1	NMSE
3	$1.201\ 3 \times 10^{-7}$
4	$1.189\ 7 \times 10^{-7}$
5	$7.303\ 6 \times 10^{-7}$
6	$1.639\ 0 \times 10^{-7}$
7	$1.493\ 6 \times 10^{-6}$

4 Conclusion

In this paper, we discuss three nonlinear prediction methods to multivariate chaotic time series and compare the prediction results of univariate reconstruction with those of multivariate reconstruction. The results show that prediction errors with multivariate reconstruction are still much smaller than those of univariate reconstruction, even if half of the data of univariate time series are used in the multivariate time series. Therefore, in practical applications, when multivariate time series data are available, better prediction results can be obtained from multivariate reconstruction, even if the data are not sufficient.

References

[1] Wang Haiyan, Sheng Zhaohan, Zhang Jin. Phase space reconstruction of complex systems based on multivariate time series [J]. *Journal of Southeast University (Natural Science Edition)*, 2003, **33**(1): 115 – 118. (in Chinese)

[2] Cao Liangyue, Mees A, Judd K. Dynamics from multivariate time series [J]. *Physica D*, 1998, **121**(1,2):75 – 88.

[3] Boccaletti S, Valladares D L, Pecora L M, et al. Reconstructing embedding spaces of coupled dynamical systems from multivariate data [J]. *Physical Review E*, 2002, **65**(1):1 – 4.

[4] Porporato A, Ridolfi L. Multivariate nonlinear prediction of river flows [J]. *Journal of Hydrology*, 2001, **248**(1 – 4):109 – 122.

[5] Reick C H, Page B. Time series prediction by multivariate next neighbor methods with application to zooplankton forecasts [J]. *Mathematics and Computers in Simulation*, 2000, **52**(3, 4):289 – 310.

[6] Kennel M B, Brown R, Abarbanel H D I. Determining embedding dimension for phase space reconstruction using a geometric construction [J]. *Physical Review A*, 1990, **151**(5): 225 – 233.

[7] Cong Shuang. *Neural networks, fuzzy systems and their applications to motion control* [M]. Hefei: University of Science and Technology of China Press, 2001. 26 – 27. (in Chinese)

混沌时间序列单变量和多变量重构的预测比较

王海燕¹ 朱 梅²

(¹ 东南大学经济管理学院, 南京 210096)

(² 东南大学数学系, 南京 210096)

摘 要 提出了多变量混沌时间序列相空间延迟重构中延迟时间间隔和嵌入维数的选取方法,给出了多变量混沌时间序列的局部平均预测法,局部线性预测法和 BP 神经网络预测法等 3 种非线性预测方法.通过 Lorenz 系统的仿真计算表明,无论用 3 种非线性预测方法中的哪一种,多变量混沌时间序列要比单变量混沌时间序列的预测误差小得多,即使前者的数据长度只有后者的一半,前者的预测误差也要小很多.另外从预测误差最小的角度验证了多变量混沌时间序列相空间延迟重构中延迟时间间隔和嵌入维数选取方法的有效性.

关键词 多变量混沌时间序列; 相空间重构; 预测; 神经网络

中图分类号 O175; O241