2005 International Symposium on Nonlinear Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Bootstrap Estimates for Nonlinear Predictors

Daisuke Haraki[†,1], Tomoya Suzuki[‡] and Tohru Ikeguchi[†]

[†]Graduate School of Science and Engineering, Saitama University, Japan,
255, Shimo-Okubo, Sakura-ku, Saitama City, Saitama, 338-8570, Japan
TEL/FAX: +81-48-858-9105
[1]E-mail: daisuke@nls.ics.saitama-u.ac.jp
[‡]Department of Electronic Engineering, Tokyo Denki University
2-2, Kanda-Nishiki-cho, Chiyoda-ku, Tokyo, 101-8457, Japan

Abstract

Estimating Jacobian matrices of observed data generated by a nonlinear dynamical system is one of the important steps for nonlinear prediction. The Jacobian matrices are estimated by using local information about divergences of nearby trajectories. Although the basic algorithm for estimating the Jacobian matrices generally works well, it often fails for noisy data. In this paper, we proposed a new scheme to select a better near neighbor set for more accurate estimation of the Jacobian matrix: making a bootstrap subset of nearest neighbors. As a result, our method much improves nonlinear predictability not only for mathematical models with observational noise but also for real time series.

1. Introduction

Many prediction methods are proposed for predicting real time series data, for example, seismic events, economic data and so on. Such real time series data are often generated by nonlinear dynamical systems. Thus, it is inevitable to involve nonlinear prediction even though its algorithms become more complicated.

In this paper, we adopt a local linear prediction for nonlinear dynamics by estimating Jacobian matrices[1, 2]. In the method, it is important to consider how to select appropriate local information, that is, a near neighbor set. If we observe noisy short time series, local information becomes poor, that is, if the data are corrupted by noise and the number of near neighbors is small, the estimated Jacobian matrices become unreliable, then it is almost impossible to achieve higher predictability. To solve the issue, we apply the bootstrap method[3] to statistically produce a wide variety of near neighbors sets. We combine a nonlinear prediction method using the approximated Jacobian matrices with the bootstrap method in order to realize higher prediction accuracy than conventional prediction methods. To confirm the validity, first we apply our method to the Ikeda map[4]. As a result, our method improved nonlinear predictability. Moreover, the method is also effective in the case that the time series data is disturbed by noises. In addition, we also apply our method to analyze nonlinearity of a Japanese vowel /a/[5].

2. The Jacobian matrix estimation method

Let us consider a nonlinear dynamical system:

$$\boldsymbol{x}(t+1) = \boldsymbol{F}(\boldsymbol{x}(t)), \tag{1}$$

where $\boldsymbol{F}$ is a k-dimensional nonlinear map, $\boldsymbol{x}(t)$ is a k-dimensional state at time t. To estimate the Jacobian matrix of $\boldsymbol{F}$, we linearize the Eq.(1) as follows:

$$\delta\boldsymbol{x}(t+1) = \boldsymbol{DF}(\boldsymbol{x}(t))\delta\boldsymbol{x}(t), \tag{2}$$

where $\boldsymbol{DF}(\boldsymbol{x}(t))$ is the Jacobian matrix at $\boldsymbol{x}(t)$. To evaluate $\boldsymbol{DF}(\boldsymbol{x}(t))$ only with the information of $\boldsymbol{x}(t)$, first, we extract a near neighbor set of $\boldsymbol{x}(t)$. Let us denote the i-th nearest neighbor by $\boldsymbol{x}(k_i)$, $i = 0, 1, 2, \ldots, M$, where $M+1$ is the number of near neighbors. After a short temporal evolution, $\boldsymbol{x}(t)$ and $\boldsymbol{x}(k_i)$ evolve $\boldsymbol{x}(t+1)$ and $\boldsymbol{x}(k_i+1)$, respectively. Then, we denote displacement vectors, $\boldsymbol{y}_i = \boldsymbol{x}(k_i) - \boldsymbol{x}(t)$ and $\boldsymbol{z}_i = \boldsymbol{x}(k_i+1) - \boldsymbol{x}(t+1)$. Here, $\boldsymbol{y}_i$ corresponds to $\delta\boldsymbol{x}(t)$ and $\boldsymbol{z}_i$ corresponds to $\delta\boldsymbol{x}(t+1)$ in Eq.(2). If the norms of $\boldsymbol{y}_i$ and $\boldsymbol{z}_i$ and the temporal evolution are small enough, we can describe the relation between $\boldsymbol{z}_i$ and $\boldsymbol{y}_i$ by the following equation:

$$\boldsymbol{z}_i = \boldsymbol{G}(t)\boldsymbol{y}_i, \tag{3}$$

where the matrix $\boldsymbol{G}(t)$ is the approximated Jacobian matrix $\boldsymbol{DF}(\boldsymbol{x}(t))$ in Eq.(2). Then, we estimate $\boldsymbol{G}(t)$ by using a least-square-error fitting, which minimizes the average square error S:

$$S = \frac{1}{M}\sum_{i=1}^{M} |\boldsymbol{z}_i - \boldsymbol{G}(t)\boldsymbol{y}_i| \to \min.$$

Then, we can calculate $\boldsymbol{G}(t)$ by the following equations:

$$\boldsymbol{G}(t)\boldsymbol{W} = \boldsymbol{C}, \quad (4)$$

where $\boldsymbol{W}$ is a variance matrix of $\boldsymbol{y}_i$ and $\boldsymbol{C}$ is a covariance matrix of $\boldsymbol{y}_i$ and $\boldsymbol{z}_i$. If $\boldsymbol{W}$ has its inverse matrix, we can estimate $\boldsymbol{G}(t)$ by $\boldsymbol{G}(t) = \boldsymbol{C}\boldsymbol{W}^{-1}$.

Let us consider a nonlinear prediction problem of $\boldsymbol{x}(T)$ on an attractor from the dynamical system of Eq.(1). Our issue is to predict its s step future $\boldsymbol{x}(T+s)$. First, we search the nearest neighbor $\boldsymbol{x}(k_0)$ of $\boldsymbol{x}(T)$. Then, we calculate a displacement vector

$$\boldsymbol{y}' = \boldsymbol{x}(T) - \boldsymbol{x}(k_0).$$

Next, we estimate the Jacobian matrix $\boldsymbol{G}(k_0)$ at $\boldsymbol{x}(k_0)$. Finally, we can estimate the predicted displacement vector $\hat{\boldsymbol{z}}'$as

$$\hat{\boldsymbol{z}}' = \boldsymbol{G}(k_0)\boldsymbol{y}',$$

where $\hat{\boldsymbol{z}}' = \hat{\boldsymbol{x}}(T+1) - \boldsymbol{x}(k_0+1)$. Then, we can predict $\hat{\boldsymbol{x}}(T+1)$ as follows:

$$\hat{\boldsymbol{x}}(T+1) = \boldsymbol{G}(k_0)(\boldsymbol{x}(T) - \boldsymbol{x}(k_0)) + \boldsymbol{x}(k_0+1).$$

Repeating above scheme for s times, we can predict the s step future of $\boldsymbol{x}(T)$.

3. Bootstrap estimates for Jacobian Matrices

In the conventional Jacobian matrix estimation method[1, 2], if the number of near neighbors of $\boldsymbol{x}(T)$ is small, the estimation of the Jacobian matrices becomes unreliable. Such an unreliable estimate would lead to undesirable results. To solve the issue, we introduced the bootstrap sampling scheme[3] to make a wide variety of nearest neighbor sets in order to obtain statistical reliability of the estimation of the Jacobian matrices.

In our proposed method, we apply the bootstrap method to the prediction method of the Jacobian matrices. First, we select the nearest neighbors on $\boldsymbol{x}(T)$ by the same way as the conventional method. The nearest neighbor on $\boldsymbol{x}(T)$ is denoted by $\boldsymbol{x}(k_0)$, and the other neighbors are denoted by $\mathrm{D}_T = \{\boldsymbol{x}(k_1), \boldsymbol{x}(k_2), \cdots, \boldsymbol{x}(k_M)\}$. We perform a sampling with replacement of D_T and obtain a new set of nearest neighbors $\mathrm{D}_1^* = \{\boldsymbol{x}_1^*(k_1), \boldsymbol{x}_1^*(k_2), \cdots, \boldsymbol{x}_1^*(k_M)\}$. Then, we estimate the Jacobian matrix $\boldsymbol{G}_1^*(k_0)$ at $\boldsymbol{x}(k_0)$ by using D_1^* as introduced in Sec.2, and we predict a future point of $\boldsymbol{x}(T)$ by

$$\hat{\boldsymbol{x}}_1^*(T+1) = \boldsymbol{G}_1^*(k_0)(\boldsymbol{x}(T) - \boldsymbol{x}(k_0)) + \boldsymbol{x}(k_0+1).$$

We repeat such bootstrap estimates for B times and the b-th bootstrap predicted point is obtained as $\hat{\boldsymbol{x}}_b^*(T+1) = \boldsymbol{G}_b^*(k_0)(\boldsymbol{x}(T) - \boldsymbol{x}(k_0)) + \boldsymbol{x}(k_0+1)(b = 1, 2, \ldots, B)$. Then, we calculate its mean value $\bar{\hat{\boldsymbol{x}}}_b^*(T+1) = \frac{1}{B}\sum_{b=1}^{B}\hat{\boldsymbol{x}}_b^*(T+1)$ as a final predicted point.

However, in this method, we also found that the matrix $\boldsymbol{W}$ in Eq.(4) dose not always have the inverse matrix, because it is possible that the matrix $\boldsymbol{W}$ is not full rank. In such a case, we retry a sampling with replacement of $\boldsymbol{X}(T)$ again, and obtain another set of nearest neighbors. We repeat this sampling until the rank of the matrix $\boldsymbol{W}$ is enough to estimate the stable Jacobian matrices and to perform a stable prediction.

4. Applying the proposed method

4.1. Simulations

We use two time series data for performing the proposed nonlinear prediction. First, we predict the Ikeda map [4] as a numerical model, which is described as follows:

$$\begin{cases} x(t+1) = p + b(x(t)\cos(\theta(t)) - y(t)\sin(\theta(t))) & (5) \\ y(t+1) = \quad b(x(t)\sin(\theta(t)) + y(t)\cos(\theta(t))), & (6) \end{cases}$$

$$\theta = \kappa - \alpha/(1 + x(t)^2 + y(t)^2),$$

where p, b, κ, and α are parameters. Then, we disturb both $x(t)$ and $y(t)$ by the observational noise, the strength of which is denoted by w.

We also apply our method to the Japanese vowel /a/[5].

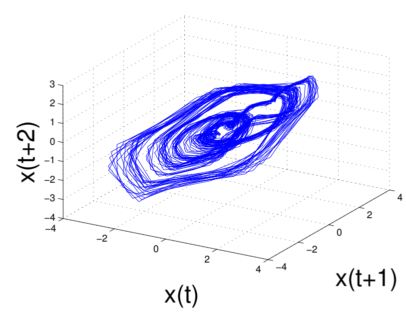


Figure 1: An attractor of the Japanese vowel /a/ reconstructed with $d = 3$ and $\tau = 5$.

Because the vowel is a single variable time series, we transform the time series in a d-dimensional state space with a delay τ by the Takens method[6] for performing nonlinear prediction.

We introduced a measure to compare the proposed method and conventional methods. First, we calculate the normalized root mean square error,

$$E = \frac{\sqrt{\sum_{t=1}^{N}|\boldsymbol{x}(t+s) - \hat{\boldsymbol{x}}(t+s)|^2}}{\sqrt{\sum_{t=1}^{N}|\boldsymbol{x}(t+s) - \bar{\boldsymbol{x}}(t+s)|^2}},$$

where s is a prediction step, $\hat{\boldsymbol{x}}(T+s)$ is a predicted point and N is the number of predicted data. We denote E of the conventional method and the proposed method as E_{c} and E_{p}, respectively. Then, the measure for the comparison is defined by

$$E_{\mathrm{r}} = \frac{E_{\mathrm{c}} - E_{\mathrm{p}}}{E_{\mathrm{c}}},$$

which is an improved ratio between E_{c} and E_{p}. The second measure evaluates how many times the proposed method improves prediction accuracies than the conventional method. We define this ratio by $I_r \in [0, 1]$.

4.2. Results and Discussion

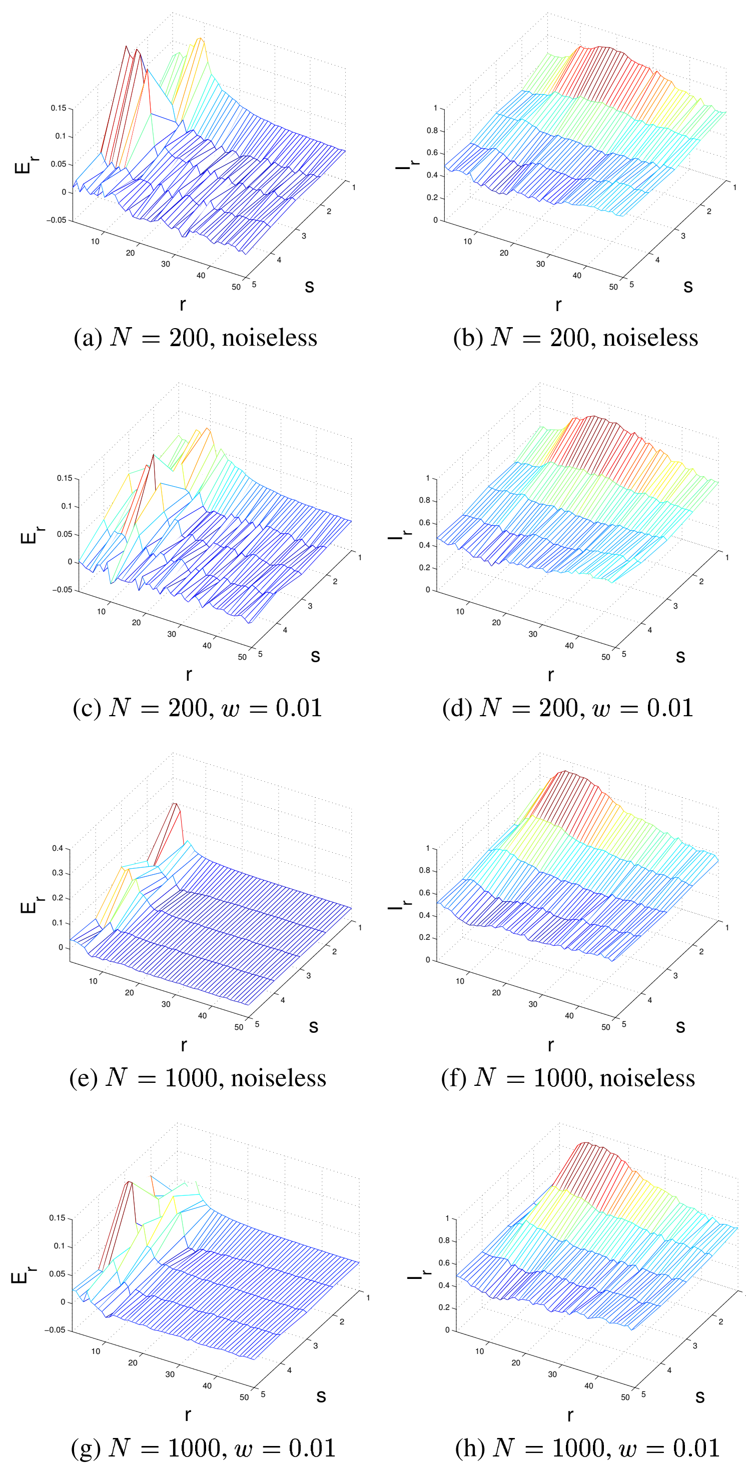
Figure 2 shows the results of the Ikeda map. In Figs.2(a), (c), (e) and (g), the horizontal plane shows the near neighbor size r(defined by $\frac{M+1}{N} \times 100$) used for local linear predictions and prediction step s, and the vertical axis shows the improved ratio of root mean square error E_{r}. These figures show that each E_{r} is almost positive, especially in the region of small r. Thus, the prediction accuracy of the proposed method is better than the conventional method. On the other hand, in the region of large r, the proposed method shows almost the same result as the conventional method. The reason is that the number of near neighbors used for prediction is large enough to evaluate Jacobian matrices. In such a case, the bootstrap method is not effective.

In Figs.2(b), (d), (f) and (h), the horizontal plane shows the near neighbor size r and delay time s, and the vertical axis shows I_r. These figures show that each I_r becomes larger than almost 50%. In particular, such tendencies are more clear as r is smaller. Then, if the delay time for embedding s is small, the ratio I_r is large. As a result, we confirm that our proposed method is effective even if time series is disturbed by the observational noise.

Next, we predict the Japanese vowel /a/ as a real time series data. Results are shown in Fig.3. Each axis is the same as Fig.2. Figures 3(b), (d), (f) and (h) show each ratio of improved cases that I_r is almost 50%. Thus, the proposed method has no advantage. However, Figs.3(a), (c), (e) and (g) show that the prediction accuracy of the proposed method is better than the conventional method at the region of small number of near neighbor r. From these results, we can confirm that the proposed method is also effective to predict real data.

5. Conclusions

We proposed a new nonlinear prediction method which combines a local linear prediction[1, 2] with the bootstrap method[3]. Then, we applied the proposed method to a mathematical model[4] and real data[5]. As a result, the proposed



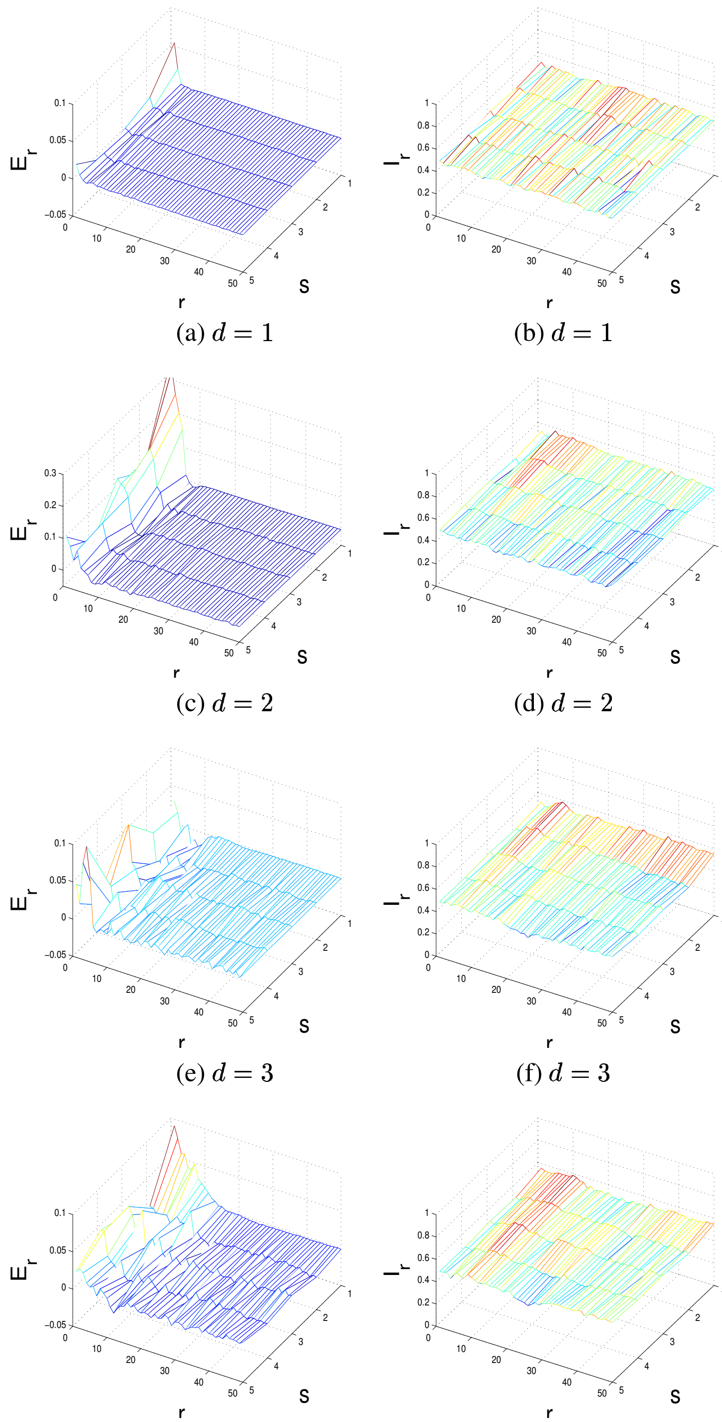
(a) $N = 200$, noiseless (b) $N = 200$, noiseless

(c) $N = 200$, $w = 0.01$ (d) $N = 200$, $w = 0.01$

(e) $N = 1000$, noiseless (f) $N = 1000$, noiseless

(g) $N = 1000$, $w = 0.01$ (h) $N = 1000$, $w = 0.01$

Figure 2: Simulation results of the Ikeda map. The parameters of the Ikeda map are $p = 1.0$, $b = 0.9$, $\kappa = 0.4$ and $\alpha = 6.0$. N is the data length and w is an additive noise level.



(a) $d = 1$ (b) $d = 1$

(c) $d = 2$ (d) $d = 2$

(e) $d = 3$ (f) $d = 3$

(g) $d = 4$ (h) $d = 4$

Figure 3: The simulation results of the Japanese vowel /a/. The sampling rate is 9.6[kHz], and the data length N is 1000.

method is effective even if the near neighbor size is small or time series data is corrupted by observational noise. That is, the bootstrap samples can compensate the lack of local information to estimate the Jacobian matrix.

Acknowledgments

The research of TI was partially supported by Grant-in-Aids for Scientific Research (c) (No.17500136) from JSPS and grant from the Research Grants Council of Hong Kong (Project No. PolyU 5235/03E).

References

[1] M. Sano and Y. Sawada, "Measurement of the Lyapunov Spectrum from a Chaotic Time Series," Physical Review Letters, Vol.55, No.10, pp.1082-1085, 1985.

[2] J. D. Farmer and J. J. Sidorowich, "Predicting Chaotic Time Series," Physical Review Letters, Vol.59, No.8, pp.845–848, August 1987.

[3] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, 1993.

[4] K. Ikeda, "Multiple-Valued Stationary State and Its Instability of the Transmitted Light by a Ring Cavity System," Optics Communications, Vol.30, No.2, pp.257–261, August 1979.

[5] I. Tokuda, R. Tokunaga, and K. Aihara, "A Simple Geometrical Structure Underlying Speech Signals of the Japanese Vowel /a/," International Journal of Bifurcation and Chaos, Vol.6, No.1, pp.149–160, 1996.

[6] F. Takens, *Detecting Strange Attractors in Turbulence*, In D. A. Rand and B. S. Young, editors, *Dynamical Systems of Turbulence*, Vol.898 of Lecture Notes in Mathematics, Berlin, Springer-Verlag, pp.366–381, 1981.