2015 International Symposium on Nonlinear Theory and its Applications
NOLTA2015, Kowloon, Hong Kong, China, December 1-4, 2015

Barycentric coordinates revisited: relaxation with linear programming and its evaluations

Yoshito Hirata[†,‡,*], Hideyuki Suzuki[‡,*], and Kazuyuki Aihara[†,‡,*]

†Institute of Industrial Science, The University of Tokyo,
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
‡Graduate School of Information Science and Technology, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
*CREST, JST, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan
Email: yoshito@sat.t.u-tokyo.ac.jp

Abstract– Barycentric coordinates are first used in Mees, Int. J. Bifurcat. Chaos (1991) to model a short nonlinear time series faithfully, while his formulation is restricted to low-dimensional dynamics because it employs triangulation. We recently relaxed his formulation by using linear programming (Hirata et al., Chaos (2015)). Using this relaxation, we can generate prediction and freeruns from a high dimensional time series. In this talk, we will review these recent advancements on barycentric coordinates and discuss some indices that evaluate locally the modelling accuracy for barycentric coordinates.

1. Introduction

Modelling time series data is an essential step to understand the underlying dynamics. In this context, Mees [1] provided a great contribution. He divided the phase space by tessellations and constructed barycentric coordinates, within which the current point is expressed as a linear combination of data points. The coefficients of the linear combination take values between 0 and 1 and their total is exactly 1. When one predicts p steps ahead, one can take the average of p steps ahead of data points weighted by the same coefficients. Mees [1] demonstrated that the Hénon map can be modelled faithfully only by using a time series of length 50.

Although the barycentric coordinates proposed by Mees [1] was fantastic, they also had a limitation, which is that we cannot model high-dimensional dynamics easily because obtaining tessellations in a high-dimensional space is very expensive.

To overcome this limitation, we recently proposed to construct barycentric coordinates using linear programming [2]. In this relaxation, we allowed some error for approximating the current point by neighboring points in the phase space. We also enforced the conditions of the coefficients that they are between 0 and 1 and their sum is equal to 1. Then, we minimized the approximation error. This minimization problem is a linear program and we can use a variety of commercial software to solve it. In [2], we demonstrated that the Rössler model and the Lorenz'63 models can be modelled such that their reconstructed models also reproduce the original shapes of the attractors. In addition, we modelled violin sounds, which cannot be distinguished from the original sounds easily.

In this presentation, we introduce our relaxation of barycentric coordinates [2] with more details and discuss that the modelling accuracy is evaluated locally using two indices related to this relaxation.

2. Barycentric Coordinates Defined by Linear Programming

2.1. Definition

Suppose that we have a dynamical system $f: M \to M$ ($x_{t+1} = f(x_t)$ for $t = 1,2,...$) we are interested in. We observe a scalar time series $s(t) = g(x_t)$ ($t = 1,2,...$) through the observation function $g: M \to R$. Using delay coordinates [3,4] denoted by

$$v(t) = \left(s(t-d+1), s(t-d+2),..., s(t)\right) \equiv G(x_{t-d+1})$$

for $t \geq d$, we reconstruct the state x_t of the underlying dynamics from this scalar time series. Namely, due to the following diagram, we can reconstruct the dynamics $\tilde{f}: R^d \to R^d$ on the delay coordinates that is equivalent to the original dynamics f :

$$\begin{array}{ccc} x_{t-d+1} & \xrightarrow{f} & x_{t-d+2} \\ \downarrow G & & \downarrow G \\ v(t) & \xrightarrow{\tilde{f}} & v(t+1). \end{array}$$

We denote the k th component of $v(t)$ by $v_k(t)$.

Suppose that we may generate prediction up to P steps ahead. In addition, suppose that $t \geq T$ and that we predict p steps ahead ($p = 1,2,..., P$). First, we find K neighboring points to $v(t)$ from $\left\{v(i) \mid i = d, d+1,..., T-P\right\}$ in the sense of the

Euclidean norm, and denote by I_t a set of time indices for the neighboring points. Then, we approximate $v(t)$ by a linear combination of $\{v(i) \mid i \in I_t\}$ using the coefficients λ_i satisfying $0 \le \lambda_i \le 1$ for $i \in I_t$ and $\sum_{i \in I_t} \lambda_i = 1$, namely,

$$v(t) \approx \sum_{i \in I_t} \lambda_i v(i). \qquad (1)$$

In our formulation in [2], we allow the error $\varepsilon \ge 0$ for the approximation of Eq. (1). Thus, we have

$$-\varepsilon \le v_k(t) - \sum_{i \in I_t} \lambda_i v_k(i) \le \varepsilon, \text{ for k=1,2,...,d.}$$

Thus, finding a set of the coefficients for barycentric coordinates can be formulated as the following linear program:

$$\min_{\{\lambda_i \mid i \in I_t\}} \varepsilon \qquad (2)$$

subject to

$$\varepsilon \ge 0,$$

$$-\varepsilon \le v_k(t) - \sum_{i \in I_t} \lambda_i v_k(i) \le \varepsilon, \text{ for } k = 1,2,...,d,$$

$0 \le \lambda_i \le 1$ for $i \in I_t$ and

$$\sum_{i \in I_t} \lambda_i = 1.$$

Then p steps ahead prediction is given by

$$\hat{v}(t+p) \approx \tilde{f}^p(v(t)) = \sum_{i \in I_t} \lambda_i v(i+p).$$

We used Matlab's function "linprog" to solve this linear program.

2.2. Formula for Evaulating Modelling Accuracy Locally

By using Taylor's expansion, the following relation can be obtained [2] in terms of the approximated dynamics $\hat{f}$ by barycentric coordinates:

$$\hat{f}(v(t)) = \tilde{f}(v(t)) + \tilde{f}'(v(t))\left(\sum_{i \in I_t} \lambda_i v(i) - v(t)\right) + O(\delta^2), \qquad (3)$$

where δ shows the size of convex hull constructed by the neighboring points. The second term is directly related to the quantity obtained by Eq. (2). For example, if this second term vanishes due to $\varepsilon = 0$, then the approximation by barycentric coordinates becomes the first order approximation as similarly to [1]. The optimal ε is related to the distance between the current point and the data manifold [5] spanned by the past points. Therefore, the value defined by Eq. (2) can be used as an index for evaluating the approximation error. Thus we call the index of Eq. (2) as the approximation error.

2.3. Quantity for Evaulating the Goodness for a Set of Neighboring Points

Another quantity for evaluating our approximation is related to how good the space spanned by neighboring points is. For this sake, we use the Gram matrix constructed by the neighboring points. Let us define a matrix V as

$$V = \big(v(i_2(t)) - v(i_1(t)), v(i_3(t)) - v(i_1(t)), \ldots, v(i_K(t)) - v(i_1(t))\big),$$

where $i_l(t)$ denotes the lth component of I_t. We use $\det$ for representing the determinant of a matrix followed. Then, we employ the next quantity for evaluating the goodness of prediction:

$$\det(V'V). \qquad (4)$$

This quantity, the Gram determinant, is expected to be close to 0 when the neighboring points are degenerated and not linearly independent. In addition, Eq. (4) becomes large when the convex hull spanned by the neighboring points is large and the approximation by Eq. (3) becomes rough. Thus, we can use Eq. (4) for evaluating the goodness of our prediction using barycentric coordinates.

3. Example

3.1. Lorenz'96II model

We evaluated the two quantities for evaluating locally the modelling accuracy using a time series generated from the Lorenz'96II model [6,7]. The Lorenz'96II model is a toy model of the atmosphere. In this model, two types of variables are connected in the structure of double rings, between which are connected locally. The outer ring corresponds to the upper sky of the atmosphere. The inner ring corresponds to the air close to the surface of the earth. The variables corresponding to the outer ring are denoted by y_i ($i = 1,2,\ldots,I$), while the variables corresponding to the inner ring are denoted by $z_{i,j}$ ($i = 1,2,\ldots,I$ and $j = 1,2,\ldots,J$). The equations for the Lorenz'96II model are defined as follows:

$$\dot{y}_i = -y_{i-2}y_{i-1} + y_{i-1}y_{i+1} - y_i + F - \frac{h_y c}{b}\sum_j z_{i,j},$$

$$\dot{z}_{i,j} = -cbz_{i,j+1}z_{i,j+2} + cbz_{i,j-1}z_{i,j+1} - cz_{i,j} + \frac{h_z c}{b} y_i,$$

$$y_i = y_{i+I},$$

$$z_{i,j+J} = z_{i+1,j},$$

where we set the parameters as follows: $F = 8$, $b = 10$, $c = 10$, $h_y = 1$, $h_z = 1$, $I = 40$, and $J = 5$. We generated a scalar time series of length $6 \times 24 \times 365 \times 2$ by observing $z_{1,1}$ every 0.01 unit times.

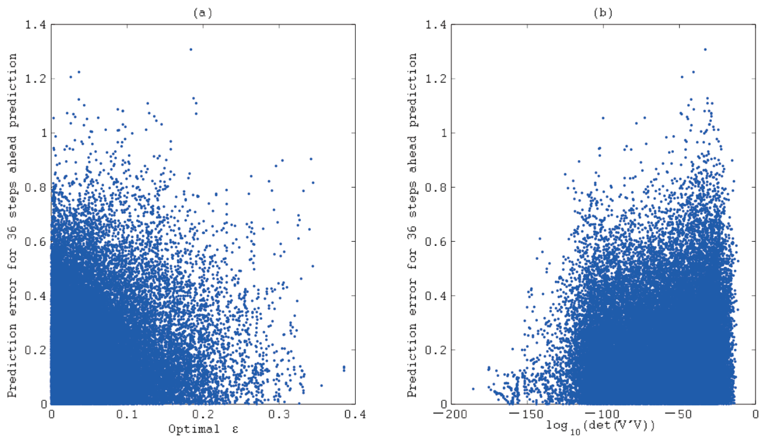

Fig. 1. Prediction error vs optimal ε using Eq. (2) (panel (a)) and $\log_{10} \det(V'V)$ using Eq. (4) (panel (b)), for the example of Lorenz'96II model.

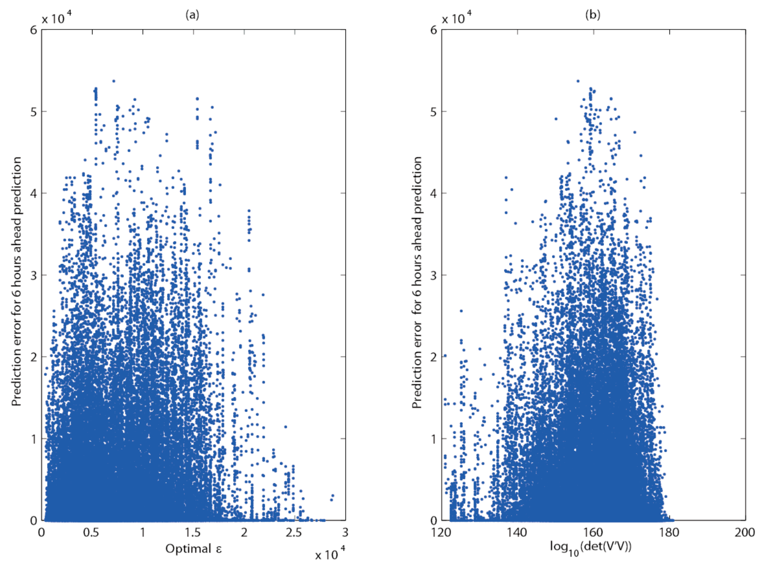

Fig. 2. Prediction error vs optimal ε using Eq. (2) (panel (a)) and $\log_{10} \det(V'V)$ using Eq. (4) (panel (b)), for the example of solar irradiance at Wakkanai, Japan.

We used the first half of the dataset for modelling and evaluated the prediction using the second half. We set $d = 36$.

The results are shown in Fig. 1. We found that the prediction errors for 36 steps ahead prediction were correlated with the optimal ε (Fig. 1(a)) and the logarithm for the Gram determinant (Fig. 1(b)). The correlation coefficients were 0.0467 (p-value: 9.43×10^{-27}) and 0.0461 (p-value: 4.09×10^{-26}), respectively.

3.2 Solar irradiance at Wakkanai, Japan

We applied the credibility measures of the approximation error and the Gram determinant to the dataset of solar irradiance for evaluating their relation to the prediction error by barycentric coordinates. The dataset was provided by the Japan Meteorological Agency. The dataset recorded the solar irradiance at Wakkanai, Japan every 10 minutes between 1 January 2010 and 31 December 2011. We used the dataset of year 2010 to predict the dataset of year 2011. We set $d = 180$ to take into account the temporal changes for solar irradiance for the previous days.

The results are shown in Fig. 2. The approximation error has a positive correlation with the prediction error (correlation coefficient: 0.0398, p-value: 6.98×10^{-20}), while the Gram determinant has a negative correlation with the prediction error (correlation coefficient: -0.0564, p-value: 2.82×10^{-38}).

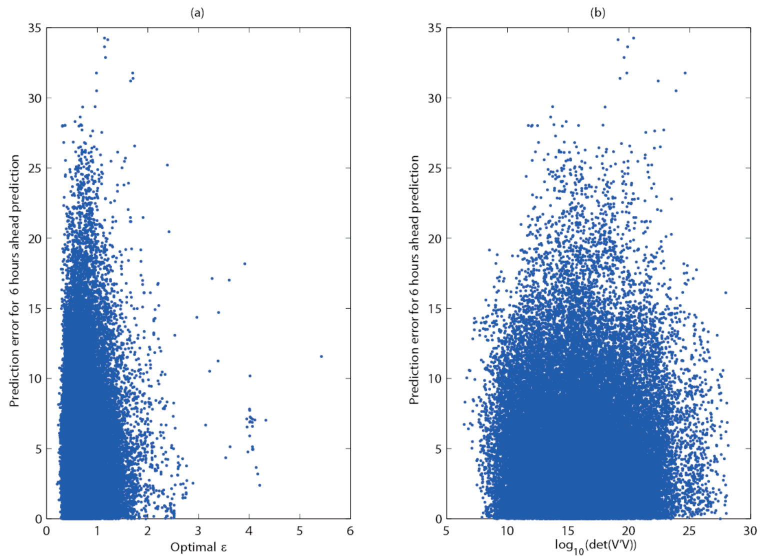

Fig. 3. Prediction error vs optimal ε using Eq. (2) (panel (a)) and $\log_{10}\det(V'V)$ using Eq. (4) (panel (b)), for the example of mean wind speed over Japan.

3.3. Mean wind speed over Japan

We also evaluated the performance for the approximation error and the Gram determinant using the mean wind speed over Japan. The dataset was provided by the Japan Meteorological Agency. We used the dataset of year 2010 to predict the dataset of year 2011. We had the mean wind speed every 10 minutes. We set $d = 36$ and thus the delay coordinates correspond to the time window of 6 hours.

The results are shown in Fig. 3. Both the approximation error and the Gram determinant have small positive correlations (0.0563 and 0.0109, the corresponding p-values 3.57×10^{-38} and 1.26×10^{-2}, respectively) with the prediction error by barycentric coordinates.

4. Discussions

Modelling a time series by barycentric coordinates was introduced. We found that barycentric cooridnates for a high-dimensional space can be constructed easily using linear programming. In addition, the prediction produced by barycentric coordinates can be evaluated locally using the two quantities of the approximation error and the Gram determinant. In this sense, the time series prediction using barycentric coordinates is sophisticated. We hope that barycentric coordinates enable us to introduce more renewable energy resources to power grid systems.

Acknowledgments

We thank the Japan Meteorological Agency for providing the datasets of solar irradiation and wind speeds used in this work. This research is partially supported by Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Agency (JST). This manuscript is also partially based on results obtained from a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

[1] A. I. Mees, "Dynamical systems and tessellations: detecting determinism in data," *Int. J. Bifurcat. Chaos*, vol.1, pp.777-794, 1991.

[2] Y. Hirata, M. Shiro, N. Takahashi, K. Aihara, H. Suzuki, and P. Mas, "Approximating high-dimensional dynamics by barycentric coordinates with linear programming," *Chaos*, vol.25, art.no.013114, 2015.

[3] F. Takens, "Detecting strange attractors in turbulence," *Lect. Notes Math.*, vol.898, p..366-381, 1981.

[4] T. Sauer, J. A. Yorke, and M. Casdagli, "Embeddology," *J. Stat. Phys.,* vol.65, pp.579-616, 1991.

[5] Y. Hirata, "Fast time-series prediction using high-dimesional data: Evaluating confidence interval credibility," *Phys. Rev. E*, vol.89, art.no.052916, 2014.

[6] E. N. Lorenz, "Predictability: A problem partly solved," in *Proceedings of the Seminar on Predictability,* vol.1, pp.1-18, ECMWF, 1996.

[7] J. A. Hansen and L. A. Smith, "The role of operational constraints in selecting supplementary observations," *J. Atmos. Sci.* vol.57, pp.2859-2871, 2000.