

ESTIMATING CORRELATION DIMENSION IN CHAOTIC TIME SERIES

Kawaguchi, Atsushi
Graduate School of Mathematics, Kyushu University

Yanagawa, Takashi
Graduate School of Mathematics, Kyushu University

<https://doi.org/10.5109/13504>

出版情報 : Bulletin of informatics and cybernetics. 33 (1/2), pp.63-71, 2001-12. Research
Association of Statistical Sciences

バージョン :

権利関係 :



ESTIMATING CORRELATION DIMENSION IN CHAOTIC TIME SERIES

By

Atsushi KAWAGUCHI* and Takashi YANAGAWA†

Abstract

This paper is concerned with estimating the correlation dimension from chaotic time series. First, we show in simulation that confidence intervals of correlation dimension constructed by Smith(1992) and Judd(1992) fail to take into account the variability due to the sensitivity to the initial conditions, the characteristic feature of chaos. Next, we propose a new estimator with weighted least squares in the formulation of Cutler(1990). It is shown that the estimator has smaller dependence on scaling region and variance than Cutler's and Smith's estimator.

Key Words and Phrases: Chaotic time series, correlation dimension, weighted least squares

1. Introduction

We consider chaotic time series $\{X_t\}$ satisfying

$$X_t = F(X_{t-1}, X_{t-2}, \dots, X_{t-d})$$

for some known integer d and known F . Putting $Y_t = (X_t, X_{t-1}, \dots, X_{t-(d-1)})$ and

$$C_N(r) = \binom{N}{2}^{-1} \sum_{i < j} I(\|Y_i - Y_j\| \leq r), \quad (1.1)$$

where I denotes the indicator function and $\|\cdot\|$ is a norm, Grassberger and Procaccia (abridged by G-P) (1983a,b) called $C(r) := \lim_{N \rightarrow \infty} C_N(r)$ the correlation integral and introduced the correlation dimension as

$$p = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r}$$

if the limit exists. It was introduced as a measure of representing the fractal dimension of the attractor of $\{Y_t\}$. Estimating the dimension of an attractor of a chaotic time series can provide useful, even vital information (see for example, Abraham et al. 1989).

The G-P proposed a procedure for estimating the correlation dimension. It essentially consists of plotting $\log C_N(r)$ against $\log r$ and looking for a proportion over

* Graduate School of Mathematics, Kyushu University 33 Hakozaki 6-10-1 Higashi-Ku Fukuoka 812-8581 Japan. Email: atsushi@math.kyushu-u.ac.jp

† Graduate School of Mathematics, Kyushu University 33 Hakozaki 6-10-1 Higashi-Ku Fukuoka 812-8581 Japan. Email: yanagawa@math.kyushu-u.ac.jp

which the plot is approximately linear, the slope over that portion is the estimator of dimension p .

Typically such a graph for a finite length time series looks like Figure 1. The graph shows that for large r the graph flattens, at moderate r the graph is quite linear and for small r the graph jumps irregularly. The irregularity is a result of having only a finite amount of data. The linear part of data is often called the scaling region.

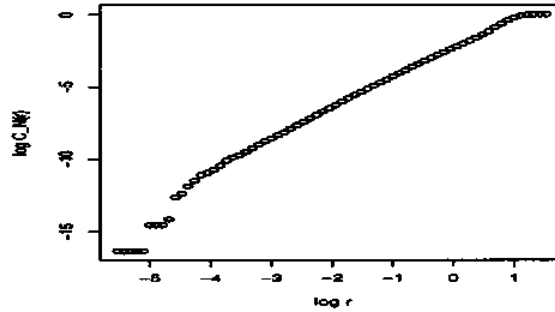


Figure 1: log-log plot

The idea of G-P was mathematically realized by Cutler(1990) who proposed a least squares estimator of p over the scaling region taking into account the intercept and by setting

$$r = r_0, r_1, \dots, r_M \quad (1.2)$$

where $r_m = r_0 \phi^m$ ($m = 0, 1, \dots, M$) for some r_0 , M and $0 < \phi < 1$.

It is well known (see, for example, Judd, 1992) that the estimator depends sensitively on the selection of the scaling region. In order to avoid this problem for ordinary least squares, Takens(1984) introduced maximum-likelihood approach. Smith(1992) developed a statistical theory by assuming that interpoint distances which are less than r are independent conditioned on distance which are less than ε for some ε , and proposed an estimator which is essentially equal to Takens' estimator.

Judd(1992) improved the estimator of Smith(1992) in the same framework as Smith(1992), but with different mathematical manifestation. A target of the statistical theory developed by Smith(1992) and Judd(1992) was the construction of confidence intervals of the correlation dimension. One purpose of this article is to show numerically that the interval by Smith(1992), and thus Judd(1992), fails to take into account the variability due to the sensitivity to the initial condition, the characteristic feature of chaos.

The other purpose of this article is to suggest the use of a weighted least squares estimator in the formulation of Cutler(1990). The estimator is compared with Cutler's and Smith's estimators. It is shown by simulation that the impact of the scaling region on the proposed estimator is smaller and that the estimator has smaller variance than the others; the decrease of the variance is substantial, in particular, when the embedding dimension is large.

2. A criticism on Smith's estimator

Put $b_m = \sum_{i < j}^N I(\|Y_i - Y_j\| \leq r_m)$ for r_m ($m = 0, 1, \dots, M$) in (1.2). Assuming that $C(r) = \alpha r^p$ and that the conditional distribution of b_m given b_{m-1} is binomial with parameter b_{m-1} and $C(r_m)/C(r_{m-1})$, Smith(1992) led to maximum-likelihood estimator of p

$$\hat{p}_S = \frac{1}{\log \phi} \left(\log \sum_{m=1}^M b_m - \log \sum_{m=1}^M b_{m-1} \right)$$

based on the conditional joint distribution of b_1, b_2, \dots, b_M given b_0 . Smith(1992) also obtained the confidence interval of p from the likelihood function by computing the information matrix. We show in the following example that Smith's confidence intervals fail to take into account the variability due to the sensitivity of the initial values in a chaotic series.

EXAMPLE. Smith used the part of data whose interpoint distances are less than ε . We set constants r_0 and M in (1.2) as $r_0 = \max_{i,j} \{\|Y_i - Y_j\|\}$, and $M = \max\{m; b_m \neq 0\}$, and delaminated the behavior of the estimator when $\varepsilon = r_{\gamma M}$ ($\gamma = 0.75, 0.7, 0.65, 0.6, 0.55, 0.5$), where $r_m = r_0 \phi^m$. To get possibly large M , we set $\phi = 0.9$. The data of size 5000 are generated from the Henon map $X_t = 1 - 1.4X_{t-1}^2 + 0.3X_{t-2}$, and estimate \hat{p}_S of p is obtained for each ε . The process is repeated 20 times using randomly selected initial values, X_0 and X_1 , from Henon's domain of convergence (see below). Also Smith's 95% confidence interval is obtained based on one randomly selected trial for each ε . The estimator and confidence interval are plotted in Figure 2 taking the values of γ on horizontal line and values of the estimated correlation dimension on vertical line. The figure shows that the substantial number of estimates than expected are outside the confidence interval. More precisely, the percentage of estimates outside the confidence interval is exhibited in Table 1 for each ε , showing the percentages are surprisingly far from 5%. We would like to put it as an open problem to construct confidence intervals of the correlation dimension that may take into account the sensitivity of the initial values.

ε	$r_{0.75M}$	$r_{0.7M}$	$r_{0.65M}$	$r_{0.6M}$	$r_{0.55M}$	$r_{0.5M}$
outside	20%	50%	50%	55%	45%	55%

Table 1: The percentage of estimates outside the confidence intervals

3. Weighted least squares estimator

3.1. Introduction of the estimator

In this section, we consider a weighted least squares estimator in the framework of Cutler(1990). It will be indicated by simulation that the new estimator has less variance than Smith's estimator. The estimated variance of $\log C_N(r)$ is used for the weight. Instead of the conditional i.i.d. assumption employed in Smith(1992) and Judd(1992) we use the i.i.d. assumption only for obtaining the weight. To begin with we show the following proposition.

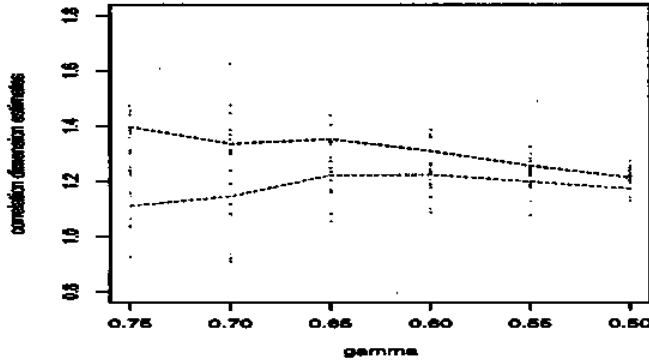


Figure 2: The binomial estimates and confidence intervals

PROPOSITION. Assume that Y_1, Y_2, \dots, Y_N are independent random variables with same marginal distribution, and let $\log C_N(r)$ be the statistic given in (1.1). Then the variance of $\log C_N(r)$ is approximated by

$$V[\log C_N(r)] \sim 6 \binom{N}{3} \left\{ \frac{P(\|Y_1 - Y_2\| \leq r, \|Y_1 - Y_3\| \leq r)}{\{P(\|Y_1 - Y_2\| \leq r)\}^2} - 1 \right\} + \binom{N}{2} \left\{ \frac{1}{P(\|Y_1 - Y_2\| \leq r)} - 1 \right\}. \quad (3.1)$$

The proof is given in Appendix.

Next we consider the estimation of $V[\log C_N(r_m)]$. Put

$$b_m^* = \sum_{i \neq j, i \neq k, j \neq k}^N I(\|Y_i - Y_j\| \leq r_m, \|Y_k - Y_j\| \leq r_m),$$

for $m = i_0 = \lfloor \lambda M^*/2 \rfloor, i_0 + 1, \dots, i_L = \lfloor \lambda M^* \rfloor$, where $M^* = \max\{m; b_m^* \neq 0\}$ and λ is a prespecified constant. Now it is clear under the assumption of Proposition that $C_N(r_m)$ is an unbiased estimator of $P(\|Y_1 - Y_2\| \leq r_m)$ and that

$$C_{N2}(r_m) = \frac{1}{N(N-1)(N-2)} b_m^*$$

is an unbiased estimator of $P(\|Y_1 - Y_2\| \leq r_m, \|Y_1 - Y_3\| \leq r_m)$. Substituting these estimators into the right hand side of the equation (3.1), we estimate $V[\log C_N(r_m)]$ by

$$\hat{V}[\log C_N(r_m)] = 6 \binom{N}{3} \left\{ \frac{C_{N2}(r_m)}{(C_N(r_m))^2} - 1 \right\} + \binom{N}{2} \left\{ \frac{1}{C_N(r_m)} - 1 \right\}$$

Using $\hat{w}_m = \{\hat{V}[\log C_N(r_m)]\}^{-1}$ as a weight, we propose to estimate the correlation dimension p by minimizing

$$Q = \sum_{m=i_0}^{i_L} \hat{w}_m (v_m - \alpha - p u_m)^2$$

over all possible choices of α and p , where $u_m = \log r_m$ and $v_m = \log C_N(r_m)$. Put $\bar{u}_w = \hat{w}^{-1} \sum_{m=i_0}^{i_L} \hat{w}_m u_m$, $\bar{v}_w = \hat{w}^{-1} \sum_{m=i_0}^{i_L} \hat{w}_m v_m$, and $\hat{w} = \sum_{m=i_0}^{i_L} \hat{w}_m$. Then the weighted least squares estimator of p is explicitly given by

$$\hat{p} = \left(\sum_{m=i_0}^{i_L} \hat{w}_m (u_m - \bar{u}_w)(v_m - \bar{v}_w) \right) / \left(\sum_{m=i_0}^{i_L} \hat{w}_m (u_m - \bar{u}_w)^2 \right).$$

Note that the scaling region of the weighted least squares estimator is $(r_{i_0}, r_{i_0+1}, \dots, r_{i_L})$, where $i_0 = \lfloor \lambda M^*/2 \rfloor$ and $i_L = \lfloor \lambda M^* \rfloor$.

3.2. Numerical comparison

Simulation is conducted to compare the weighted least squares estimator with Cutler's and Smith's estimator. The scaling region for computing Cutler's and Smith's estimator is $(r_{\lfloor \lambda M/2 \rfloor}, r_{\lfloor \lambda M/2 \rfloor+1}, \dots, r_{\lfloor \lambda M \rfloor})$, whereas that used for the weighted least squares estimator is $(r_{\lfloor \lambda M^*/2 \rfloor}, r_{\lfloor \lambda M^*/2 \rfloor+1}, \dots, r_{\lfloor \lambda M^* \rfloor})$. For given initial value $(X_0, X_1, \dots, X_{d-1})$ the data of size 6000 are generated from each of the following dynamics, the first 1000 data are abandoned and 5000 data are used to compute the estimates, i.e. $N = 5000$.

(D1) $X_t = 1 - 1.4X_{t-1}^2 + 0.3X_{t-2}$ (Henon map)

(D2) $X_t = 0.01X_{t-1} + 4X_{t-4} \exp(-X_{t-4}^2)$

(D3) $X_t = 0.5X_{t-1} + (-0.5 + 2 \exp(-X_{t-1}^2))X_{t-7}$

Note that the embedding dimension of (D1), (D2) and (D3) are $d = 2, 4$, and 7 , respectively. The time for computing $C_{N2}(r_m)$ in the weighted least squares estimator is enormous. In order to reduce it we used the sub data of size $N^* = 50N^{1/3} = 855$ selected as follows: first select randomly one element, say Y_{i_0} from Y_1, Y_2, \dots, Y_N and put $Y_1^* = Y_{i_0}$, $Y_2^* = Y_{i_0+1}, \dots, Y_{N^*}^* = Y_{i_0+(N^*-1)}$. It was shown that the impact of the selection was negligible. Twenty trials are conducted using the initial values which are selected randomly from Henon's domain of convergence (Henon, 1976) for (D1), that is, the quadrilateral ABCD defined by $A = (-1.33, 1.4)$, $B = (1.32, 0.443)$, $C = (1.245, -0.466)$, $D = (-1.06, -1.666)$. The same number of trials is conducted with initial values from interval $[0, 2]$ for (D2) and from interval $[-1, 1]$ for (D3). Note that dynamics (D3) shows three different patterns of trajectory depending on initial values; one remains positive, one remains negative and the other takes positive and negative values as if randomly. We only employed the last pattern in the simulation. For (D1) the value of M is assessed in the region $120 \sim 130$ and the constant λ is set as $\lambda = 1, 0.95, 0.9, \dots, 0.5$; for (D2) it is assessed $55 \sim 65$ and λ is set as $\lambda = 1, 0.95, 0.9, \dots, 0.75$; and for (D3) it is assessed to be $60 \sim 70$ and λ is set as $\lambda = 1, 0.95, 0.9, \dots, 0.75$.

Figures 3, 4, and 5 exhibit the results of computation for (D1), (D2) and (D3), respectively. In each figure, the left panel shows the average of the twenty estimates and the right panel shows the standard deviation of the estimates computed from the twenty estimates. The horizontal line in each figure shows the value of λ . The numbers "1", "2", and "3" in the figures show Cutler's estimator, the weighted least squares estimator and Smith's estimates, respectively.

Figures show that Cutler's estimator depends on the scaling region; that it overestimates the correlation dimension, in particular, when $\lambda = 1$ and the embedding

dimension is large; and that the estimator is sensitive to initial values. Figures also show that Smith's estimates and the weighted least squares estimates are essentially equal over the whole scaling region; that the standard deviation of Smith's estimates is larger than that of the weighted least squares estimator; and that the difference of the standard deviation increases as the increase of the embedding dimension.

4. Discussion

A weighted least squares estimator is introduced in this article and its superiority to Cutler's and Smith's estimator is shown by simulation. Also it is shown that Smith's confidence interval fails to take into account the variability due to the sensitivity of the initial values, the characteristic feature of chaos, although it has been long considered that an advantage of Smith's estimator is in the confidence interval constructed from the estimator.

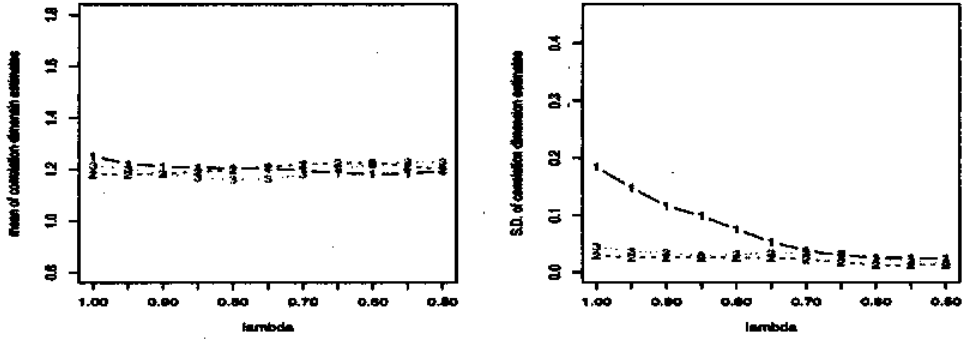
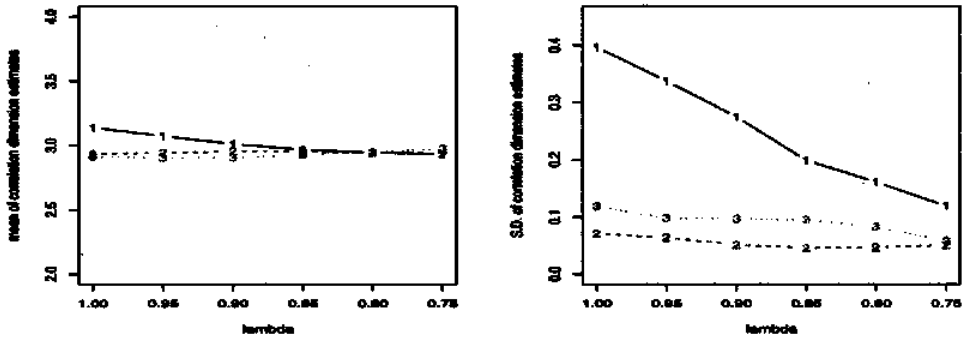
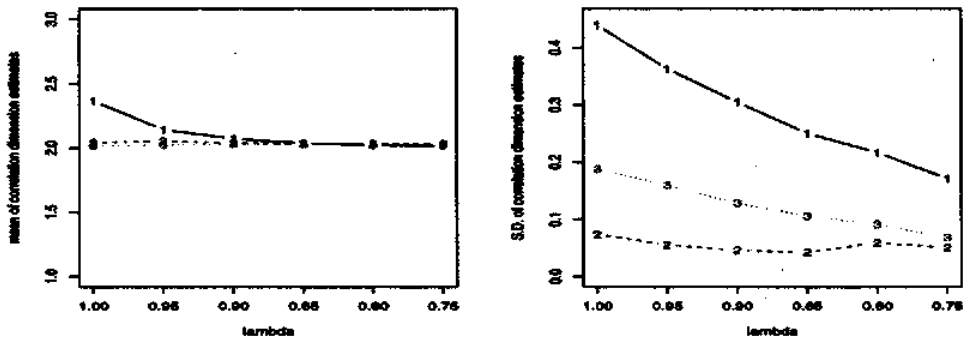
It is assumed in this paper that F and d are known. The assumption is, of course, not acceptable in practice. Also noises are involved in practical data. Including the open problem stated in the text, there are a lot of challenging problems in the subject. This article is a first step of our research on the interesting subject.

Acknowledgement

The authors would like to thank the referees for their helpful comments.

References

- Abraham, N. B. Albano, A. M. Passamante, A. and Rapp, P. E. (1989). *Measure of Complexity and Chaos*. New York: Plenum.
- Cutler, C. D. (1990). Some Results on the Behavior and Estimation of the Fractal Dimensions of Distributions on Attractor, *Journal of Statistical Physics* **62**, 651-708.
- Grassberger, P. and Procaccia, I. (1983a). Characterization for Strange Attractors, *Physical Review Letters*, **50**, No.5, 346-349.
- Grassberger, P. and Procaccia, I. (1983b). Measuring the strangeness of strange attractors, *Physica*, **9D**, No.5, 189-208.
- Henon, M. (1976). A Two-dimensional Mapping with a Strange Attractor, *Communications in Mathematical Physics*, 69-77.
- Judd, K. (1992). An improved estimator of dimension and some comments on providing confidence intervals, *Physica D*, **56**, 216-228.
- Judd, K. (1994). Estimating dimension from small samples, *Physica D*, **71**, 421-429.
- Smith, R. L. (1992). Estimation Dimension in Noisy Chaotic Time Series, *Journal of the Royal Statistical Society*, **B54**, No.2, 329-351.
- Takens, F. (1984). On the numerical determination of the dimension of an attractor, *Lecture Notes in Mathematics*, **1125**, 99-106.


 Figure 3: Dynamics (D1) $X_t = 1 - 1.4X_{t-1}^2 + 0.3X_{t-2}$

 Figure 4: Dynamics (D2) $X_t = 0.01X_{t-1} + 4X_{t-4} \exp(-X_{t-4}^2)$

 Figure 5: Dynamics (D3) $X_t = 0.5X_{t-1} + (-0.5 + 2 \exp(-X_{t-1}^2))X_{t-7}$

Appendix

Proof of Proposition. First,

$$\begin{aligned}\log C_N(r) &= \log E[C_N(r)] + \left(\log C_N(r) - \log E[C_N(r)] \right) \\ &= \log E[C_N(r)] + \left[\log \left(1 + \frac{C_N(r) - E[C_N(r)]}{E[C_N(r)]} \right) \right].\end{aligned}$$

Using the Taylor expansion in the second term, it follows that

$$\log C_N(r) \sim \log E[C_N(r)] + \frac{C_N(r) - E[C_N(r)]}{E[C_N(r)]}.$$

Thus,

$$V[\log C_N(r)] \sim \frac{1}{(E[C_N(r)])^2} V[C_N(r)]. \quad (\text{A.1})$$

Under the assumption that Y 's have the same marginal distribution, we have for the denominator of (A.1)

$$E[C_N(r)] = \binom{N}{2}^{-1} \sum_{i < j}^N P(\|Y_i - Y_j\| \leq r) = P(\|Y_1 - Y_2\| \leq r). \quad (\text{A.2})$$

Put $g(y_i, y_j) = I(\|y_i - y_j\| \leq r) - E[I(\|y_i - y_j\| \leq r)]$, the numerator of (A.1) can be written as follows:

$$V[C_N(r)] = \binom{N}{2}^{-2} \sum_{i < j}^N \sum_{k < l}^N E[g(Y_i, Y_j)g(Y_k, Y_l)].$$

Since we assumed that Y_1, Y_2, \dots, Y_N are independent, it follows that

$$V[C_N(r)] = \binom{N}{2}^{-2} \left\{ \sum_{i \neq j, i \neq k, j \neq k}^N E[g(Y_i, Y_j)g(Y_k, Y_l)] + \sum_{i < j}^N E[(g(Y_i, Y_j))^2] \right\}. \quad (\text{A.3})$$

Under the assumption that Y 's have the same marginal distribution, the first term of (A.3) tends to be

$$\begin{aligned}& \sum_{i \neq j, i \neq k, j \neq k}^N E[g(Y_i, Y_j)g(Y_k, Y_l)] \\ &= 6 \sum_{i < j < k}^N \left\{ E[I(\|Y_i - Y_j\| \leq r)I(\|Y_i - Y_k\| \leq r)] \right. \\ & \quad \left. - E[I(\|Y_i - Y_j\| \leq r)]E[I(\|Y_i - Y_k\| \leq r)] \right\} \\ &= 6 \binom{N}{3} \left\{ P(\|Y_1 - Y_2\| \leq r, \|Y_1 - Y_3\| \leq r) - (P(\|Y_1 - Y_2\| \leq r))^2 \right\} \quad (\text{A.4})\end{aligned}$$

Again, under the assumption that Y 's have the same marginal distribution, we also have for the second term of (A.3)

$$\begin{aligned} \sum_{i < j}^N E[(g(Y_i, Y_j))^2] &= \sum_{i < j}^N \left\{ E[(I(\|Y_i - Y_j\| \leq r))^2] - (E[I(\|Y_i - Y_j\| \leq r)])^2 \right\} \\ &= \binom{N}{2} \left\{ P(\|Y_1 - Y_2\| \leq r) - (P(\|Y_1 - Y_2\| \leq r))^2 \right\} \quad (\text{A.5}) \end{aligned}$$

Combining the equations (A.1)-(A.5), we can get required form as (3.1).

Received February 14, 2002

Revised February 27, 2002