ハイパフォーマンス
コンピューティング 78－4
（１９９９．１０．７）

Neural predictions for chaos represented by multi descriptors

Tomoo AOYAMA and Hanxi ZHU
The Faculty of Engineering, Miyazaki University
Gakuen Kibanadai-nishi 1-1, Miyazaki 889-2192, Japan
E-mail: aoyama@esl.miyazaki-u.ac.jp

Abstract

We discuss long range forecasting for the chaos that is represented by multi descriptors. The discussion is based on an extended embedding theory and multi layer neural networks. Next, we investigate forecasting under insufficient observations. Where the future movements of phenomena are predicted by a hypothesis for objects.

1. Introduction

We find many studies for forecasting on use of neural networks. In the studies, making use of Takens' embedding theory [1], time-series data for one dimension descriptor are adopted. The theory is useful certainly, however it secures correct predictions under the one step futures. Rossler's chaos is described by three variables, but by using the theory, only one descriptor makes movements of the chaos. It is strange, but certain, because of the one step prediction.

We must study new forecasting that is not based on the embedding theory for more practical use. We introduce one extension method that we published in a previous paper.

Usually we must forecast future movements of phenomena under insufficient sampling data. We face such situations always. For the fuzzy predictions, we must study another forecasting method. This way is extremely difficult and requires elaborate operations. The way easily branch into arbitrariness and subjective. We adopted multi value logic to avoid the problems. The logic is mathematical and objective.

2. Discussion for multi layer neural network

2.1 Cross terms

We write a three-layer neural network as following,

$\{x1,x2,....xn\} \equiv x$,

$yj = \Sigma i[Vji*xi]$, $pj=f(yj)$,

Where a vector {x1,x2,...xn} has an individual information, and its suffix corresponds to numbered neurons in the first layer. We put one bias neuron in the layer, and we substitute the bias neuron for threshold-values of actions of neurons. Even if the bias is one, since it has many connections, the substitution is realized. Suffices "i" and "j" are used for 1st and 2nd layer, respectively. Vij is a matrix of connection weights between neurons in 1st and 2nd layer. The "yj" is a variable as temporary defined one. And f() is a function simulated a neuron, which is a differential function. We name the function as "neuron-function." The "pj" is a vector for output of neurons in 2nd layer.

We get following relations for informed propagations between 2nd and 3rd layers.

$\{p1,p2,....,pm\} \equiv p$,

$zk = \Sigma j[Wjk*pj]$, $ok=g(zk)$,

$\{o1,o2,...ok\} \equiv o$.

Where Wjk is a matrix of connection weights, and g() is a neuron-function. Suffix "k" is used for 3rd layer, so a vector {o1,o2,...ok} is output for neurons in 3rd layer.

If g() is a linear function, the three-layer neural network is equivalent to,

$$ok = \Sigma j[Wjk*f(\ \Sigma i[Vji*xi]\ )], \quad (1)$$
$$= \Sigma j[Wjk*F(xi)].$$

That is well known as ANN whose approximate expansion is similar to the ortho-normal expansion. In the case, as the function F is represented by sigmoid functions and their thresholds, the Fs are not ortho-normal set exactly. In order to close the Fs to the set, radial function's approach is often tried. It has been believed that methods similar to the ANN are useful. But they are not complete. If a phenomenon is described by multi variables that is {x,y,z}, we will derivate following expressions,

$$ok = \Sigma j[Wjk*F(xi) + W'jk*F(x'i)$$
$$+ W''jk*F(x''zi)].$$

Where descriptors {xi,x'i,x"i} are correspond to variables {x,y,z}. In the expression (1), cross terms such as xi*x'i are not found. We are sure that the expression is insufficient. We are reluctant to put the cross terms in first layer. That method adds new neurons that don't connect to outer region and don't reflect observations directly. But cross terms are necessary, we assume them to be {(x**a)*(y**b)*(z**c), a=0,1,2,..., b=0,1,2,..., c=0,1,2,...}. The bias neuron exists in the assumption, that is a=b=c=0.

2.2 Window width

We must define a window width in order to part observation data into many fragments. The window is not for one dimension, but for the multi descriptors. But, we need to adopt a definition that includes one dimension domain. We resulted following correspondence, which are:

For x-coordinate, {{x1,x2,...xk}<-->xk+1, {x2,x3,...xk+1}<-->xk+2,....},

For y-coordinate, {{y1,y2,...yk}<-->yk+1, {y2,y3,...yk+1}<-->yk+2,....},

For z-coordinate, {{z1,z2,...zk}<-->zk+1, {z2,z3,...zk+1}<-->zk+2,....},

For any cross-term, {{r1,r2,...rk}<-->rk+1, {r2,r3,...rk+1}<-->rk+2,....},

where the letter "r" is a genetic symbol of (x**a)*(y**b)*(z**c)-term. Sampling for x, y, and z are at same time. We neglected (xi**a) (yj**b)*(zk*c) terms. The terms don't appear in the original differential equation. Using above correspondences, the prediction scheme in the previous paper is expanded straightly.

3. Numerical calculations

3.1 Rossler's chaos

A definition of the chaos is,

dx/dt=-y-z,

dy/dt=x+a*y,

dz/dt=b*x-c*z+x*z,

where a, b, and c are constant that we set 0.36, 0.4, and 0.5. Initial values x0, y0, and z0 were 0.5, 0.0, and 0.0, respectively. In practice, we used a relation: dt=(xi+1-xi)/160000. Then, we got three kinds of learning sets that were {x0,x1,...,x39}, {y0,y1,...,y39}, and {z0,z1,...,z39}. The reference sets were {x40,....x159}, {y40,...y159}, and {z40,...z159}.

3.2 One dimensional prediction

At first we used forecasting based on x-coordinate data. The window width is 8. The calculated results on long range are listed in figure 1. Those of short range are done in figure 2.

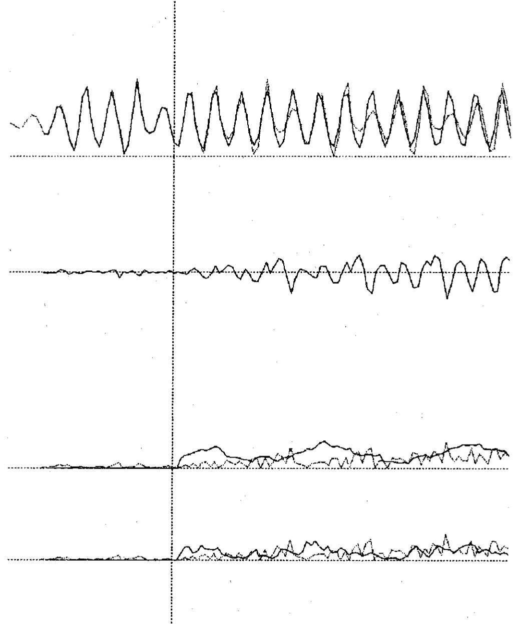
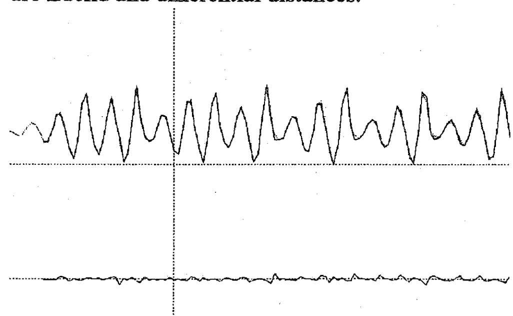


Figure 1. Long range forecasting, its error, Euclid and differential distances for Rossler's chaos. The left side of a dotted vertical line means a learning term. The right side is forecasting. Number of the learning data is 32, and forecasting points are 120. In the first graph, the solid curve is results from a neural network, and the plotted curve is true values of the chaos. In the second graph, the solid curve is differences between forecasting values and true ones. In the third and fourth graphs, the plotted lines are square of the differences. The solid curves are Euclid and differential distances.



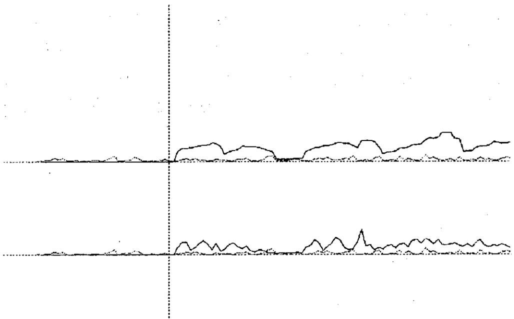


Figure 2. Short range forecasting, its error, Euclid and differential distances for Rossler's chaos. In these graphs, drawing methods are same as upper figure 1.

Figure 1 shows clearly that long range forecasting for Rossler's chaos is very difficult. We notice that the error between forecasting and true values is increasing as time passing. It is a grave problem.

At second we calculated a short range forecasting under the same conditions.

Figure 2 shows that short range forecasting is possible certainly. Takens' embedding theory is enabled. Some researchers satisfy the results, however, the Euclid and differential distances show a tendency. The forecasting is usually failing against true values of the chaos. If there is no attraction power, the forecasting is failed by the tendency. One-step prediction is the attractor. This is a reason that long range forecasting is difficult.

3.3 Multi dimensional prediction

Rossler's chaos is primarily three-dimensional phenomenon. Then, it should be expressed by multi descriptors.

We did the method mentioned in the section 2 as an elementary test. In practice, we used three kinds of learning data and one product data: {x0*z0,x1*z1,....,x39*z39}. The product is correspond to the cross terms in section 2. We must select properly all products (x**p)*(y**q)*(z**r), where p=q=r={0,1,2,...}. But the selection is difficult to execute because of limitation of computer resources. As an elementary trial, we used one x*z term only. We set that window width was 8 for each descriptors. Therefore, the neuron number of the first layer was 8*4+1(bias)=33. The number of the third layer must be equal to number of descriptors, i.e. 3. The number of the second layer was set as 30. This value was not optimized.

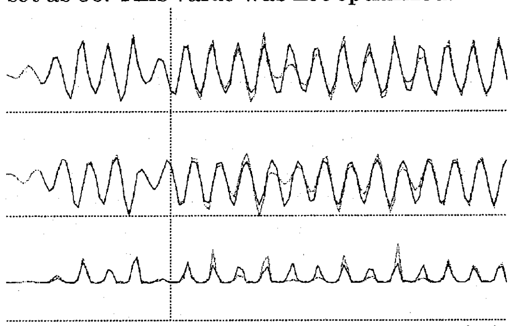


Figure 3. Long range forecasting for Rossler's chaos. Values for x, y, and z coordinates are plotted in order.

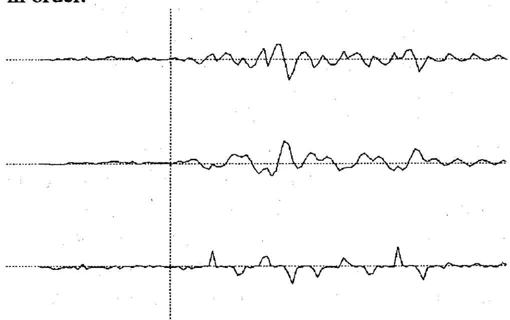


Figure 4. Differences between forecasting and Rossler's chaos.

The figures 3 and 4 showed pulses of errors during the forecasting. But they didn't indicate increasing error as time passing. It is clearly an improvement for the grave problem.

4. Introducing hypothesis

Above mentioned the forecasting by using neural networks is a reasonable extension based on Takens' embedding theory. But there is a limitation on the prediction, because only observed data are used and is not done an implicit property of phenomena. Usually introductions of the property have rejected in the traditional way. We find that arbitrariness is arisen in the introductions, and have avoided it. However, we must convince that we already used the implicit property without consciousness. We have used neural networks. That is forecasting is limited to functions of neural networks.

When we select a scheme for a solution, the arbitrariness exists already. This is a problem we cannot avoid, that is why we will discuss them mathematically. We try to diminish the problem, which we introduce a hypothesis under the multi value logic [2].

4.1 One-body operators for multi value logic

There are many operators to calculate multi value logic. In them, we consider the one-body operators at first, which are not-operator (~) and rotation operator (R).

The not-operator operates the fragments of input vector, and generates new fragments as, {xi,xi+1,...,xi+k-1}={ 1-xi,1-xi+1,...,1-xi+k-1}. Similarly the scalar value xi+k is translated into 1-xi+k. The operations are accepted when observations are symmetric for the mean value.

The rotation operator operates as,
R{xi,xi+1,...,xi+k-1}
={S+xi,S+xi+1,...,S+xi+k-1},
where "S+" means the addition of a constant "S" to the x-elements. The results seem to be equal to the value-shift, however the operation is based on Post's not-operation whose character is a kind of rotation. The operations are accepted when similar phenomena are observed on the different bias conditions.

4.2 Numerical calculations

In order to examine the one-body operator "~", we forecasted a sine function under discrete data sampled on a quarter period. Naturally they are not enough for predictions. But they have partial information to predict. This is the reason why we select it. Results we calculated showed their insufficiency.

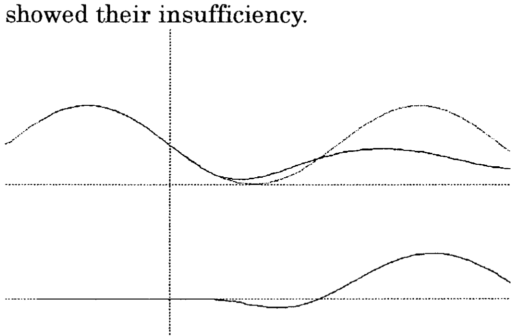


Figure 5. Long range forecasting for sine function under insufficient sampling.

The inappropriate forecasting is due to insufficient sampling. Some increasing fragments data in the sine are not found in sampling. They can be easily introduced by operating not-operator "~". We operated it to the learning set, and got following forecasting.

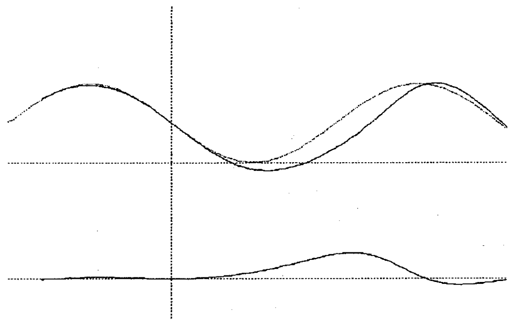


Figure 6. Long range forecasting for sine function under insufficient sampling data that are operated by the not-operator

The not-operator improved explicitly an insufficiency of the sampling data. Moreover, no secondary effect was arisen. Then we got an acceptable forecasting.

However, we should not satisfy with these results. Since we know the sine function, we can judge the reasonableness for the forecasting. It is uncertain whether the learning data calculated from the sine represent the original function or not. Therefore, we need an index that relates with a precision of forecasting. We believe the index based on the principle "nothing can't generate significance". Euclid and differential distances satisfy the principle.
We calculated them for case of figure 5 and 6. These results are following.

Figures 7 and 8 are drawn on same scale. The figures clearly suggest that the precision by using not-operator is higher than the original.

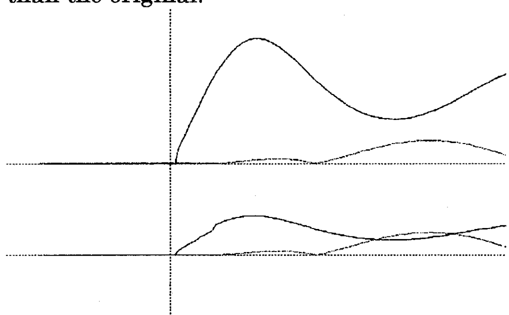


Figure 7. Euclid and differential distances for long range forecasting under the conditions of figure 5.

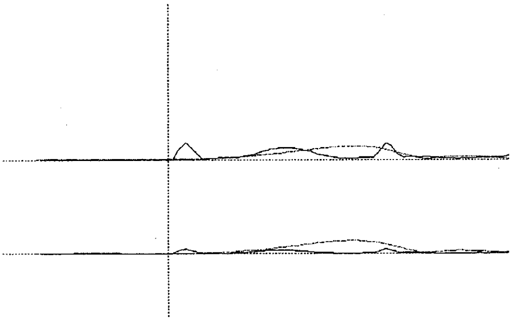


Figure 8. . Euclid and differential distances for long range forecasting under the conditions of figure 6.

But we must consider that the operated learning data are more than the original. It is natural consequence that the distances are shorter. Figure 8 indicates that the operation ought not to be a negation. It doesn't indicate positively that the introduction is correct. To revise the defects, we recalculated the distances between prediction input data and the restricted learning data that include the operated data.

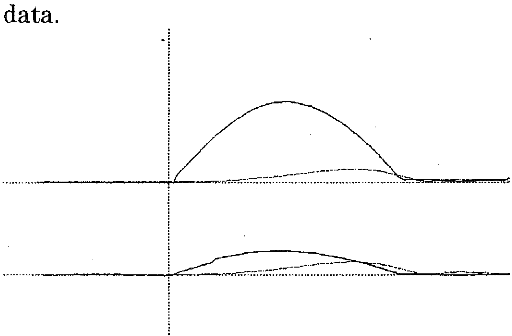


Figure 9. . Euclid and differential distances for long range forecasting under the restricted learning data set.

The restricted distances show uncertain predictions at first. But, the predictions become more and more certain. This implies that initial predictions are executed by the operated terms, and as passing time, they are done by the original terms. That is likely that a lack of data was supplemented by that not-operator. We must determine that such supplementation is reasonable or not. It is not determined by mathematics; therefore we call the method as "introducing hypothesis".

If we use the method for short range forecasting, we will get more accurate result. So we tried it, and made sure that.

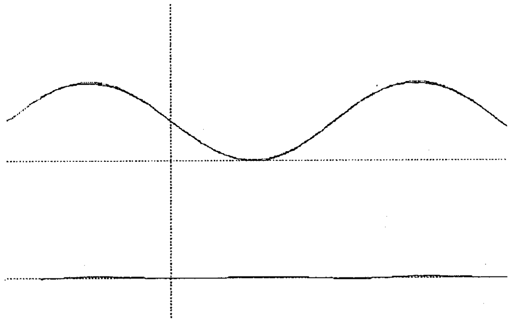


Figure 10. Short range forecasting for sine function under insufficient sampling data that are operated by the not-operator

4.3 Two-body operators

We consider the two-body operators now. They are "or-" (|) and "and-" operators (&). The "or-" and "and-" operators operate the fragments as following,

{xi,xi+1,...,xi+k-1} | {xj,xj+1,...,xj+k-1}
={max(xi,xj),max(xi+1,xj+1)...,max(xi+k-1,xj+k-1)},

xi+k | xj+k=max(xi+k,xj+k),

{xi,xi+1,...,xi+k-1}&{xj,xj+1,...,xj+k-1}
={min(xi,xj),min(xi+1,xj+1)...,min(xi+k-1, xj+k-1)},

xi+k & xj+k=min(xi+k,xj+k).

These operations are accepted when turndown or turn-up is not observed yet, but they are expected.

5. Conclusion.

We investigated followings; A fragmentation technique for chaos that was represented by multi descriptors, Forecasting of the chaos by using the fragmentation and neural networks. The technique is a multi-version of the embedding theory. We examined it for Rossler's. But we should test it in case of much other chaos. We introduced a working hypothesis in order to supplement a lack of sampling data, and discussed that the introduction was acceptable or not by using multi value logic.

references

[1] Kazuyuki AIHARA and Ryuji TOKUNAGA, "Practical Strategy for Chaos", (Ohm Co. Ltd., Tokyo, 1993.10), Japanese.

[2] Masao MUKAIDONO, "Fuzzy Logic", (Daily Industrial News Co. Ltd., 1993.3), Japanese.