A SEQUENTIAL NONLINEAR MAPPING FOR DATA ANALYSIS

Algirdas Mykolas MONTVILAS

Institute of Mathematics and Informatics,
2600 Vilnius, Akademijos St.4, Lithuania

Abstract. An algorithm for the sequential analysis of multivariate data is presented along with some experimental results. The algorithm is based upon the sequential nonlinear mapping of L-dimensional vectors from the L-hiperspace into a lower-dimensional (two-dimensional) vectors such that the inner structure of distances between the vectors is preserved. Expressions for the sequential nonlinear mapping are obtained. The sequential nonlinear mapping is applied to sequential clusterization of random processes and creation of an essentially new method for sequential detection of many abrupt or slow changes in several unknown states of dynamic systems.

Key words: dimensionality reduction, sequential nonlinear mapping, sequential clustering, state of system, sequential detection of changes.

Introduction. The purpose of this paper is to describe the sequential nonlinear mapping algorithm which has been found to be highly effective in the sequential analysis of multivariate data. Random processes often have rather a complicated structure. Sometimes .nonstationary processes consist of stationary segments with typical properties of a certain class. While forming classifiers for those processes we need to build clusters and to reject the segments of a random process untypical for the given class. Therefore, a task turns up to define typical properties of the segments by any means and to compare them. It is very convenient to do that denoting stationary segments of random processes by some mark on the PC screen. Then the marks of the same properties of segments find themselves in a certain place of the screen and that of different properties - in another place. It is achieved using the

nonlinear mapping algorithm in (Sammon, 1969). The stationary segments of random processes are described by a proper mathematical model, and L estimates of the model's parameters make up the vectors in the L-hyperspace. These L-dimensional vectors are non-linearly mapped into two-dimensional vectors preserving the inner structure of distances among them. However, it can be done only after having got the whole random process. In such a case when we are working in a real time and segments of a random process are received sequentially, we need to create a sequential nonlinear mapping algorithm.

Another task. While watching the states of dynamic objects or technological processes and their changes one has to keep sequential observations as well. Objects can change their states abruptly or slowly. When the state of an object changes the L parameters describing the state change as well. If the object state is described by a random process generated by this object, then the state is described by the L parameters characterizing the random process. Thus, in all cases we can decide about the object state or its abrupt or slow change according to the same data or their changes. The object can have several unknown states and we need to watch the states and detect their changes sequentially and independently of the history. It is convenient to watch the object state and changes marking it by some mark on the PC screen. According to the mark position we can make a decision of the object state and its change if the mark position changes.

For solution of these problems it is necessary to have a method of sequential detection of many changes in several unknown properties of random processes. There are many methods of detection of changes in the properties of random processes in scientific publications (Kligienė and Telksnys, 1984; Basseville and Benveniste, 1986; Nikiforov, 1983), but there are no methods to solve the above mentioned problems.

In this paper a sequential nonlinear mapping is considered. The expressions for sequential nonlinear mapping of vectors from L-hyperspace onto the plane are obtained. The sequential non-

linear mapping is applied to sequential clusterization of random processes and creation of an essentially new method for sequential detection of many abrupt or slow changes in several unknown states of dynamic systems. Examples are given.

Statement of the problem. Let us have in general a non-stationary random process Z_t. However separate segments of the process have their own cinstant parameters, i.e., we have a locally-stationary process. The stationary segments may be described by a proper mathematical model, e.g., an autoregressive (AR) sequence:

$$z_t - m = -\sum_{k=1}^{p} a_k(z_{t-k} - m) + bv_t, \qquad (1)$$

where m is the segment average, p is AR order, $a_k(k = 1,\ldots,p)$ are AR parameters, b is amplification factor of input excitation, v_t is discrete white noise distributed by $\mathcal{N}(0, 1)$. The assumption of stationarity involves a condition that all the roots of the characteristic equation are less that 1 in absolute value. The parameters of AR are estimated using the Yule-Walker equations (Box and Jenkins, 1970). Then, we have $L = p + 2$ parameters: m, b, a_k $(k = 1,\ldots,p)$, and the L-dimensional vectors represent stationary segments of a random process. In (Sammon, 1969) the algorithm of simultaneous nonlinear mapping of multidimensional vectors onto the plane is presented. Let us denote these vectors by X_i, $i = 1,\ldots,M$. The L-dimensional vectors X_i are mapped onto the plane into two-dimensional vectors Y_i, $i = 1,\ldots,M$. The main requirement of mapping the L-dimensional vectors into two-dimensional vectors is to preserve the inner structure of distances between the vectors. This is achieved using a nonlinear mapping procedure.

In our case it is necessary to map the L-dimensional vectors into two-dimensional vectors nonlinearly and sequentially in order to process the sequentially received data.

Solution of the problem. While realizing sequential nonlinear mapping first of all we have to nonlinearly map M vectors $(M \geqslant 2)$ simultaneously. Afterwards we need to map sequentially

and nonlinearly the received parameters vectors and, in such a way, to clusterize the stationary segments of a random process or to watch the present state of a dynamic object, to detect its changes and deviations from it for a practically unlimited time. In order to formalize the method we denote by N this practically unlimited number of arriving vectors.

Thus, let us have $M + N$ vectors in the L-hyperspace. We denote them X_i, $i = 1, \ldots, M$; X_j, $j = M + 1, \ldots, M + N$. M vectors are already simultaneously mapped into two-dimensional vectors Y_i, $i = 1, \ldots, M$. Now we need to sequentially map the L-dimensional vectors X_j into two-dimensional vectors Y_j, $j = M + 1, \ldots, M + N$. Here the nonlinear mapping expressions will change into sequential nonlinear mapping expressions respectively. First, before performing iterations it is expedient to put the two-dimensional vectors being mapped in the same initial conditions, i.e., $y_{jk} = c_k$, $j = M + 1, \ldots, M + N$; $k = 1, 2$. Note that in the case of simultaneous mapping of the first M vectors, the initial conditions must be chosen in a random way (Sammon, 1969). Let the distance between the vectors X_i and X_j in the L-hyperspace be defined by d_{ij}^x and on the plane – by d_{ij}^y respectively. This algorithm uses the Euclidean distance measure, because if we have no a priori knowledge concerning the data, we would have no reason to prefer any metric over the Euclidean metric (Sammon, 1969).

Next, we compute the normalized error of distances E.

$$E_{j\cdot} = \left(\sum_{i=1}^{M} d_{ij}^x \right)^{-1} \sum_{i=1}^{M} (d_{ij}^x - d_{ij}^y)^2 / d_{ij}^x,$$

$$j = M + 1, \ldots, M + N.$$

(2)

For correct mapping we have to change the positions of vectors Y_j, $j = M + 1, \ldots, M + N$, on the plane in such a way that the error E_j be minimal.

This is achieved by using the steepest descent procedure. After

the r-th iteration the error of distances will be

$$E_j(r) = \left(\sum_{i=1}^{M} d_{ij}^x \right)^{-1} \sum_{i=1}^{M} [d_{ij}^x - d_{ij}^y(r)]^2 / d_{ij}^x,$$

$$j = M + 1, \ldots, M + N. \tag{3}$$

here

$$d_{ij}^y(r) = \left(\sum_{k=1}^{2} [y_{ik} - y_{jk}(r)]^2 \right)^{1/2} \tag{4}$$

$$i = 1, \ldots, M; \quad j = M + 1, \ldots, M + N.$$

During the $r + 1$ iteration the coordinates of the mapped vectors Y_j will be

$$y_{jk}(r + 1) = y_{jk}(r) - F\Delta_{jk}(r),$$

$$j = M + 1, \ldots, M + N; \quad k = 1, 2; \tag{5}$$

where

$$\Delta_{jk}(r) = \frac{\partial E_j(r)}{\partial y_{jk}(r)} \bigg/ \left| \frac{\partial^2 E_j(r)}{\partial y_{jk}^2(r)} \right|, \tag{6}$$

F is the coefficient for correction of the coordinates, and it is defined empirically to be $F = 0.35$;

$$\frac{\partial E_j}{\partial y_{jk}} = H \sum_{i=1}^{M} (d_{ij}^x \cdot d_{ij}^y)^{-1} \cdot D \cdot C, \tag{7}$$

$$\frac{\partial^2 E_j}{\partial y_{jk}^2} = H \sum_{i=1}^{M} (d_{ij}^x \cdot d_{ij}^y)^{-1} \{ D - C^2 (d_{ij}^y)^{-1} [1 + D(d_{ij}^y)^{-1}] \}, \tag{8}$$

where

$$H = -2 \left(\sum_{i=1}^{M} d_{ij}^x \right)^{-1}; \quad D = d_{ij}^x - d_{ij}^y; \quad C = y_{jk} - y_{ik}.$$

When $E_j < \epsilon$, where ϵ is chosen under concrete conditions, the iteration process is over and the result is shown on the PC screen. In fact it is enough $\epsilon = 0.01$. In order to have equal computing time for each mapping we can execute constant number of iterations. In practice it is enough $I = 30$.

Sequential clusterization of random processes. The sequential nonlinear mapping can be applied to sequential clusterization of random processes or formation of clusters. A real acoustic signal has been received, which was in fact a locally-stationary one, and in discrete segments 1024 long it was stationary. The signal was approximated in the segments by the sixth order AR model. For each segment the AR parameters a_k $(k = 1,\ldots,6)$, the amplification factor of input excitation b, and segment average m were estimated. Then, $L = 6 + 2 = 8$ – dimensional vectors, representing the locally-stationary segments of the acoustic signal, were mapped onto the plane.

20 vectors, as if belonging to two acoustic signal classes, were involved in the experiment, besides, in all cases the first class was represented by the vectors numbered from 1 to 10, and the second class – from 11 to 20, respectively. Note that the first M vectors, mapped simultaneously, are denoted in Fig. 1 and Fig. 2 by mark $\times$ with the index which means the vector's number. The later received vectors, mapped sequentially, are denoted by mark + with the respective index. Besides, in all cases there was 30 iterations.

We present two cases of mapped vectors, which are taken from (Montvilas, 1990). In Fig. 1 the results of mapping of 20 vectors are presented, when after simultaneous mapping of $M = 19$ vectors (10 from the first class and 9 vectors from the second one) the vector number 20, as if received later, was mapped, i.e., $N = 1$.

In Fig. 2 the case of real sequential clusterisation is presented, i.e., after simultaneous mapping of the minimal amount $M = 2$ of initial vectors, belonging to the first class, the remaining $N = 18$ vectors (8 from the first class and 10 from the second one) were mapped sequentially.

While analyzing Fig. 1 and Fig. 2 where different cases of nonlinear mapping of the set of the same vectors are presented, we see, that vector number 2, as if belonging to the first class, is distant from other vectors of the same class, i.e., the properties of this acoustic signal segment are different from the properties typical of the first class. Thus, we can reject the vector number 2 while
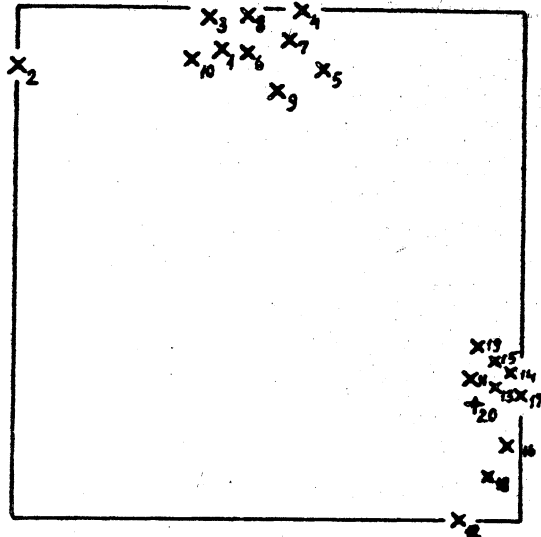
Fig. 1. The view of the mapped vectors of two classes (1 ÷ 10, 11 ÷ 20), when $M = 19$, $N = 1$.
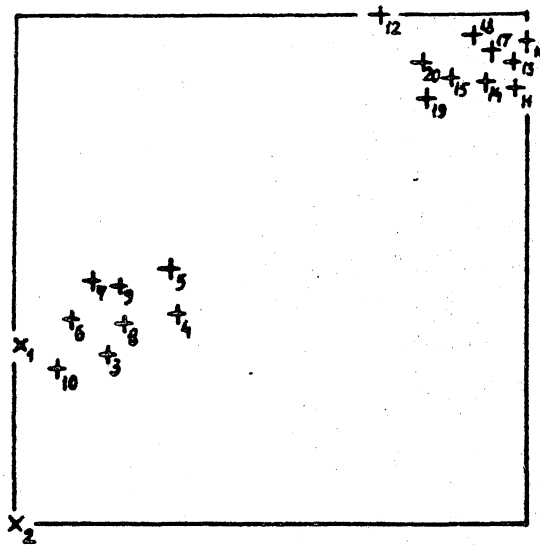

Fig. 2. The view of the same vectors mapped sequentially, when $M = 2$, $N = 18$

forming the classifier. As it turned out later, the acoustic signal segment number 2 was really spoiled by the sensor.

Sequential detection of many abrupt or slow changes in several unknown states of dynamic systems. We can apply the sequential nonlinear mapping in creation of an essentially new method for sequential detection of many abrupt or slow changes in several unknown properties of random processes, as well as for watching the states of dynamic systems or technological processes and detecting their various changes.

Let a dynamic system (DS) be in any state s_i of the set of possible states: $s_i \in S$. We can watch L parameters at the output of DS. Those parameters can be of any physical nature (then we must introduce the scale coefficient for each parameter). A DS can be described by some proper mathematical model, too, e.g., AR (1). Then we have $L = p + 2$ parameters: m, b, a_k $(k = 1, \ldots, p)$ as well, estimated from locally-stationary segments 256 long of the observed discrete process. DS states may be unknown. A DS may change its states abruptly or slowly. We need to map sequentially nonlinearly the L-dimensional vectors, representing the states, into two-dimensional vectors in order to reflect the present system state by some mark on the PC screen and, having in mind the existence of particular states, to identify the current state, a deviation from it or a transfer to other state when the mark changes its position.

We present some simulated experiments. At first we have a case when the DS has three states: $S = 3$. The DS(3) is described by the 3-rd-order AR with the parameters (see Table 1).

Table 1. The AR parameters of the states $1 \div 3$ of DS(3)

STATE	a_1	a_2	a_3	b
1	0.30	0.10	0.10	1.0
2	0.10	0.60	0.40	1.0
4	−0.50	−0.10	0.20	1.0

We detect the states of DS(3) at $M + N = 12$ time moments. First we take such a case when the number of initial simultaneous

mapping of state vectors is equal to the number of stationary states of the DS(3): $M = S = 3$, and during the time moments $M = 1 \div 3$ the DS(3) passes through all its possible states S. Then the view on the screen is "fixed" from the very beginning because of the automatic scale of coordinates. After that we detect the states of DS(3) at the time moments $N = 4 \div 12$ sequentially. According to the conditions of the experiment a priori the states of DS(3) are known at the time moments (see Table 2).

Table 2. The states of DS(3) at the time moments
$$M + N = 3 + 9 = 12$$

MAPPING	SIMULTAN (i)			SEQUENTIAL (j)								
MARK	x			+								
TIME MOMENT	1	2	3	4	5	6	7	8	9	10	11	12
STATE	1	2	3	1	3	2	2	1	1	3	3	2

In Fig. 3 the results of mapping are presented, where at the first $M = 3$ time moments state vectors, mapped simultaneously, are denoted by mark x with the index which means the time moment number, and the state vectors, mapped sequentially, are denoted by mark + with the respective index.

In Fig. 4 the situation similar to that of Fig. 3 is presented, but after time moment number 6 the DS(3) changed its state slowly, and at time moment number 7 it was between the states number 2 and number 1.

Next, let us take another case. It is taken from (Montvilas, 1992). Now the DS has $S = 4$ stationary states. The DS(4) is described by 3-rd-order AR with the parameters (see Table 3).

In this case the number of initial simultaneously mapped state vectors will be $M = 2$ and not all the possible states of DS(4) are involved. We detect the states of DS(4) at the time moments $M + N = 2 + 14 = 16$. According to the experiment conditions the states of DS(4) are known at the time moments (see Table 4).
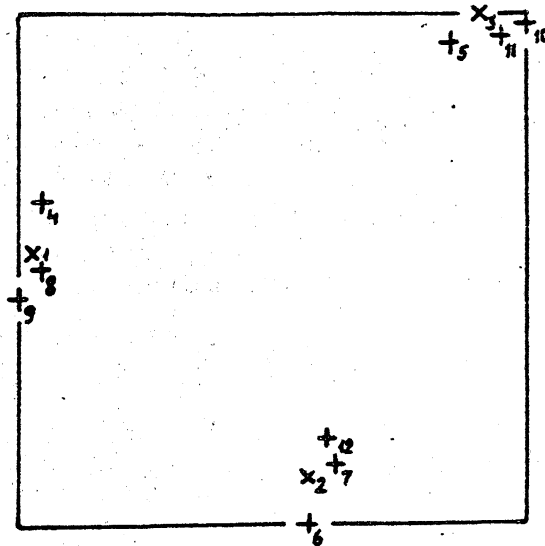
Fig. 3. The view on the PC screen of the mapped vectors of
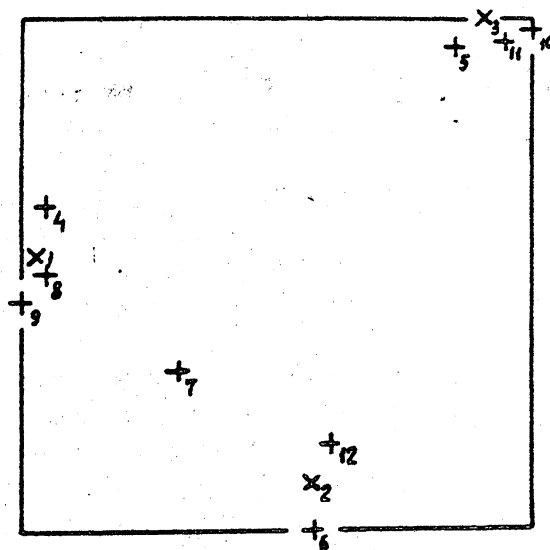DS(3) states for the first situation.

Fig. 4. The view on the PC screen of the mapped vectors of
DS(3) states for the second situation.

Table 3. The AR parameters of the states $1 \div 4$ of DS(4)

STATE	a_1	a_2	a_3	b
1	0.30	0.10	0.10	1.0
2	0.20	0.65	0.40	1.0
3	−0.50	0.30	0.40	1.0
4	−0.30	−0.10	0.15	1.0

Table 4. The states of DS(4) at the time moments
$M + N = 2 + 14 = 16$

MAPPING	SIMULTAN (i)		SEQUENTIAL (j)													
MARK	x		+													
TIME MOMENT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
STATE	1	2	3	4	1	4	2	4	1	3	3	2	1	3	4	2

The mapping results are presented in Fig. 5.

In Fig. 6 the situation similar to that of Fig. 5 is presented. The difference is at time moment number 11, where a slow change of the state of DS(4) took place, and DS(4) was between the states number 3 and number 2. In Fig. 6 this situation is clear.

Conclusions. The considered sequential nonlinear mapping of vectors from the L-hyperspace onto the plane can be applied to sequential clusterization of random processes and visual formation of clusters rejecting the spoiled elements.

The sequential nonlinear mapping enabled us to create the essentially new method which allows us to sequentially detect many abrupt or slow changes in several unknown properties of random processes and to watch dynamic system or technological process states, their abrupt or slow changes on the PC screen in fact without time limitation.

At the very beginning, before sequential nonlinear mapping, it suffices to map simultaneously only $M = 2$ vectors.
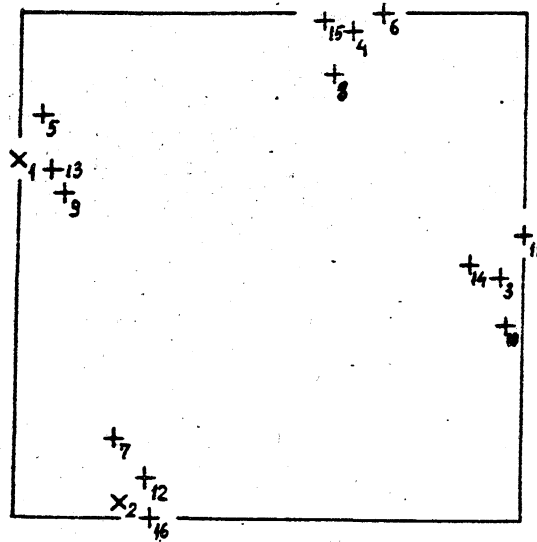
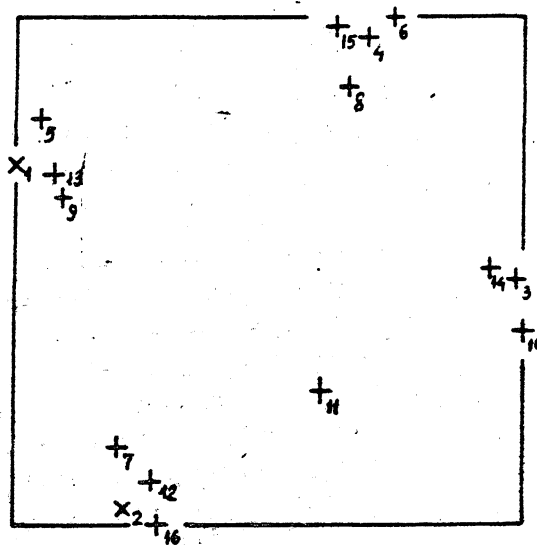**Fig. 5.** The view on the PC screen of the mapped vectors of DS(4) states for the first situation.

Fig. 6. The view on the PC screen of the mapped vectors of DS(4) states for the second situation.

REFERENCES

Basseville, M., and A.Benveniste (1986). *Detection of Abrupt Changes in Signals and Dynamical Systems*. Springer-Verlag Berlin, Heidelberg, New York, Tokyo.

Box, G., and G.Jenkins (1970). *Time Series Analysis, Forecasting and Control*. San Francisco, Cambridge, London, Amsterdam.

Kligiene, N., and L.Telksnys (1984). Methods of detecting instants of change of random process properties.-A survey. *Automation and Remote Control,* part 1, 44(10), 1241–1283. (translated from Russian original article in Automatika i Telemechanika, 10, 1983, pp. 5–56).

Montvilas, A-M. (1990). Sequential clusterization of multiparametric nonstationary processes.. In L.Telksnys (Ed.), *Statistical Problems of Control*, Vol.90. Inst. Math. Inform. Lithuanian Acad. Sci., Vilnius. pp. 62–72 (in Russian).

Montvilas, A-M. (1992). Sequential detection of many changes in several unknown states of dynamic systems. *Informatica*, 3(1), 72–79.

Nikiforov, I.V. (1983). *Sequential Detection of a Change in the Properties of Time Series*. Nauka, Moscow. (in Russian).

Sammon, J.W (1969). A nonlinear mapping for data structure analysis. *IEEE Trans. on Computers*, 18(5), 401–409.

A.-M. Montvilas received the degree of Candidate of Technical Sciences from the Kaunas Technological University, Kaunas, Lithuania, in 1974. He is a senior researcher at the Process Recognition Department of the Institute of Mathematics and Informatics of the Lithuanian Academy of Sciences. Scientific interests include: processing of random processes, classification, detection of a change in the properties of random processes, simulation.