INFORMATICA, 1992, Vol.3, No.1, 37–46

ESTIMATION OF MANY CHANGE-POINTS IN LONG AUTOREGRESSIVE SEQUENCES

Antanas LIPEIKA and Joana LIPEIKIENĖ

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

Abstract. The dynamic programming method for estimation of many change-points in univariate autoregressive (AR) sequences with known AR parameters between change-points is investigated. A problem how to use this method for long autoregressive sequences is solved and a constructive solution is given. A simulation experiment illustrates the advantages of the solution obtained.

Key words: autoregressive sequence, change-points in properties, dynamic programming, segmentation of speech.

1. Introduction. The problem of estimation of many change-points in random sequences is rather complicated because of a large amount of computations required. There are many papers (e.g. Ozaki, Tong, 1975; Kitagawa, Akaike, 1978; Praetorius, Bodenstein, 1977) in which suboptimal methods to solve this problem are proposed. But these methods do not enable us to reach a global maximum (or minimum) of the objective function. For this reason we discuss here optimal methods.

As far as we know, the problem of estimation of many change-points in random processes has been investigated for the first time by Telksnys in 1970. Later on, under the assumption that the parameters between change-points are known, the problem was formulated for autoregressive random sequences (Lipeika, 1975; 1977) and important properties of the likelihood function of change-points

have been established. Using specific properties of the likelihood function of change points a method for global maximization of the likelihood function was constructed (Lipeika, 1979; 1987a). This method enabled us to solve the problem by computer. The application of a dynamic programming method to solve the maximization problem (Lipeika, Lipeikienė, 1987b; 1990) allowed to reduce significantly the computation amount. However, the usage of the developed methods for estimation of change-points in the segmentation of speech signals proved to be inconvenient for estimation of change-points in long AR sequences. (Long sequences are assumed to have 5000 or more points). Thus there arose a necessity to apply the developed methods to long sequences. We present here an algorithm which enables us to estimate the change-points in long sequences without complicated manipulatious with the main and external memory of a computer. An example illustrates the performance of the constructed algorithm. All computations were carried out by IBM PC AT. We used C language for programming.

2. Statement of the problem. We consider the output X_t, $t = \ldots, 1, 2, \ldots$ of a linear discrete nonstationary dynamic system, whose input is a sequence of Goussian independent random variables V_t, $t = \ldots, 1, 2, \ldots, N, \ldots$ with zero mean and variance 1. The system structure is described by an autogressive model

$$X_t = -a_1(t)X_{t-1} - a_2X_{t-2} - \ldots - a_{p_t}(t)X_{t-p_t} + b(t)V_t + \mu(t). \quad (1)$$

At the unknown points of time $u_1, \ldots, u_n$ the system parameters $A_t = (a_1(t), \ldots, a_{p_t}(t), b(t), \mu(t))$ abruptly change their values:

$$A_t = \begin{cases} Q_1, & t = \ldots, 1, 2, \ldots, u_1 \\ Q_2, & t = u_1 + 1, \ldots, u_2 \\ - - - - - - - - - - - - - \\ Q_i, & t = u_{i-1} + 1, \ldots, u_i \\ - - - - - - - - - - - - - \\ Q_n, & t = u_{n-1} + 1, \ldots, u_n \\ Q_{n+1}, & t = u_n + 1, \ldots, N, \ldots, \end{cases}$$

where $Q_i = (a_1^{(i)}, \ldots, a_{p_i}^{(i)}, b_i, \mu_i)$, $i = 1, \ldots, n+1$ are known. The problem is to obtain the maximum likelihood estimates $\hat{u} = (\hat{u}_1, \ldots, \hat{u}_n)$ of change-points $u = (u_1, \ldots, u_n)$ using the realization $x = (x_1, \ldots, x_N)$ of the random sequence X_t. It is convenient to denote $u_0 = p_{\max} + 1$, $u_{n+1} = N$, where $p_{\max} = \max(p_1, \ldots, p_{n+1})$.

3. Solution of the problem. The maximum likelihood estimate of change-points has the following form:

$$\hat{u} = \underset{p_{\max} < u_1 < \ldots < u_n < N}{\arg\max_{u}} L(u/x), \tag{2}$$

where the logarithmic function $L(u/x)$ of change-points (if we neglect the constants not depending on u) is (Lipeika, 1977):

$$L(u/x) = -\sum_{i=1}^{n+1} (u_i - u_{i-1}) \ln b_i - \frac{1}{2} \sum_{i=1}^{n+1} \frac{1}{b_i^2} \sum_{t=u_{i-1}+1}^{u_i} (x_t - \mu_i + a_1^{(i)} x_{t-1} + \ldots + a_{p_i}^{(i)} x_{t-p_i})^2. \tag{3}$$

For the logarithmic likelihood function $L(u/x)$ the equality

$$\hat{u} = \underset{p_{\max} < u_1 < \ldots < u_n < N}{\arg\max_{u}} L(u/x) = \underset{p_{\max} < u_1 < \ldots < u_n < N}{\arg\max_{u}} \theta(u/x) \tag{4}$$

is valid, where the function $\theta(u/x)$ consists of the sum of partial functions $L_i(u_i/x)$:

$$\theta(u/x) = L_1(u_1/x) + L_2(u_2/x) + \ldots + L_n(u_n/x). \tag{5}$$

The functions $L_i(u_i/x)$ are calculated recurrently:

$$L_i(u_i/x) = L_i(u_i - 1/x) + \ln \frac{b_{i+1}}{b_i} + \frac{1}{2b_{i+1}^2} (x_{u_i} - \mu_{i+1} + a_1^{(i+1)} x_{u_i - 1} + \ldots + a_{p_{i+1}}^{(i+1)} x_{u_i - p_{i+1}})^2 - \frac{1}{2b_i^2} (x_{u_i} - \mu_i + a_1^{(i)} x_{u_i - 1} + \ldots + a_{p_i}^{(i)} x_{u_i - p_i})^2 \tag{6}$$

with the initial conditions

$$L_i(p_{\max} + 1/x) = 0, \quad i = 1, 2, \ldots, n.$$

Since the function $\theta(u/x)$ consists of the sum of partial functions and each of these functions depends only on one variable, we may use the dynamic programming method to determine the global maximum of this function.

According to the dynamic programming method let us define the Bellmon functions

$$g_1(u_2/x) = \max_{\substack{u_1 \\ p_{\max} < u_1 < u_2}} L_1(u_1/x), \qquad u_2 = p_{\max} + i + 1, \ldots, N$$

$$g_i(u_{i+1}/x) = \max_{\substack{u_i \\ p_{\max}+i-1 < u_i < u_{i+1}}} \left[L_i(u_i/x) + g_{i-1}(u_i/x)\right]. \tag{8}$$

The value of $g_n(N|x)$ is the global maximum of $\theta(u/x)$ and simultaneously we find the exact global maximum of the function $L(u/x)$.

For further reduction of computation amount one may compute the functions $g_i(u_{i+1}/x)$, $i = 1, \ldots, n$ recurrently. Then

$$g_1(u_2/x) = \max\left[g_1(u_2 - 1/x),\ L_1(u_2 - 1/x)\right], \tag{9}$$

$u_2 = p_{\max} + 3, \ldots, N$ with the initial conditions

$$g_1(p_{\max} + 2/x) = L_1(p_{\max} + 1/x).$$

For $i = 2, 3, \ldots, n$

$$g_i(u_{i+1}/x) = \max\Big\{g_i(u_{i+1} - 1/x),\ \left[L_i(u_{i+1} - 1/x) + g_{i-1}(u_{i+1} - 1/x)\right]\Big\}, \qquad u_{i+1} = p_{\max} + i + 2, \ldots, N$$

with the initial conditions

$$g_i(p_{\max} + i + 1/x) = L_i(p_{\max} + i/x) + g_{i-1}(p_{\max} + i/x). \tag{11}$$

Thus we get the following estimates of change-points:

$$\hat{u}_n = \min\Big[\ \operatorname*{arg\,max}_{\substack{u_{n+1} \\ p_{\max}+n \leqslant u_{n+1} \leqslant N}} g_n(u_{n+1}/x)\Big], \tag{12}$$

$$\hat{u}_{n-1} = \min\Big[\ \operatorname*{arg\,max}_{\substack{u_n \\ p_{\max}+n-1 \leqslant u_n < \hat{u}_n}} g_{n-1}(u_n/x)\Big], \tag{13}$$

- - - - - - - - - - - - - - - - - - - -

$$\widehat{u}_1 = \min\Big[\underset{\substack{u_2 \\ p_{\max}+1 \leqslant u_2 < \widehat{u}_2}}{\arg\max} g_1(u_2/x)\Big], \tag{14}$$

4. Minimization of the main memory of a computer. Expressions (6), (9), (10) show that in order to realize the algorithm by computer one must store at least four arrays of length N (array $x(N)$, $g_i(N)$, $g_{i-1}(N)$ and $L_i(N)$). We have modified the algorithm for determination of maximum of the function $L(u/x)$ in such a way that we have to store only two arrays of length N in the main memory of a computer, where we save $X(N)$ and the meanings of $g_{i-1}(\ )$. We store the calculated functions $g_i(\ )$ in the external memory. In this case the functions $g_i(\ )$ are calculated as follows. Let us denote

$$\begin{aligned} r(i,u_i) =& L_i(u_i/x) - L_i(u_i-1)/x) \\ =& \ln\frac{b_{i+1}}{b_i} + \frac{1}{2b_{i+1}^2}\big(x_{u_i} - \mu_{i+1} + a_1^{(i+1)}x_{u_i-1} + \dots \\ & + a_{p_{i+1}}^{(i+1)}x_{u_{i+1}-p_{i+1}}\big)^2 - \frac{1}{2b_i^2}\big(x_{u_i} - \mu_i + a_1^{(i)}x_{u_i-1} \\ & + \dots + a_{p_i}^{(i)}x_{u_i-p_i}\big)^2. \end{aligned} \tag{15}$$

When calculating the Bellmon functions, we do not compute $r(i,u_i)$ in advance but calculate them right away when it is necessary to use them. Then $g_i(u_2/x)$ are calculated in such a way:

$$\begin{aligned} & g_1(p_{\max}+2/x) = r(1,p_{\max}+1); \\ & l = r(1,p_{\max}+1) + r(1,p_{\max}+2); \\ & g_1(u_2/x) = \max\big[g_1(u_2-1/x), l\big]; \\ & l = l + r(1,u_2); \\ & u_2 = p_{\max}+3,\dots,N. \end{aligned} \tag{16}$$

For $i = 2,3,\dots,n$, the calculating procedure of $g_i(u_{i+1}/x)$ is the following. We put the values of the function $g_{i-1}(u_i/x)$ into the auxiliary array $h(\ )$ of length N :

$$h(u_i) = g_{i-1}(u_i/x), \qquad u_i = 1,\dots,N.$$

Afterwards we calculate the initial values:

$$\begin{cases} l = \sum\limits_{t=p_{\max}+1}^{t=p_{\max}+i} r(i,t); \\ \alpha = h(p_{\max}+i+1); \\ h(p_{\max}+i+1) = h(p_{\max}+i)+l; \\ l = l + r(i, p_{\max}+i+1). \end{cases} \tag{17}$$

Then, for $u_{i+1} = p_{\max}+i+2,\ldots,N$,

$$\begin{cases} \beta = h(u_{i+1}); \\ h(u_{i+1}) = \max\left[h(u_{i+1}-1), (l+\alpha)\right]; \\ l = l + r(i, u_{i+1}); \\ \alpha = \beta. \end{cases} \tag{18}$$

After the calculations the array $h(\ )$ contains the values of the function $g_i(u_{i+1}/x)$, stored in the external memory for determination of change-points according to (12) – (14).

5. Example. We have solved the following simulation problem. We generated a realization of the random second-order autoregressive sequence with 10 change-points (values in Table 2). The length of the realization was $n = 8000$. The parameters of the autoregressive model between change-points are presented in Table 1

Table 1. The parameters of the autoregressive model.

$\begin{matrix} i \\ Q_i \end{matrix}$	1	2	3	4	5	6	7	8	9	10	11
$a_1^{(i)}$	0.9	0.7	0.5	0.3	0.1	0	−0.1	−0.3	−0.5	−0.7	−0.9
$a_2^{(i)}$	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9
b_i	1	1	1	1	1	1	1	1	1	1	1

Fig. 1 illustrates the generated realization (as one can see, it is long enough to be considered long) and the Bellman functions $g_i(u_{i+1}/x)$, $i = 1,\ldots,10$. Fig. 2 displays the spectral densities which

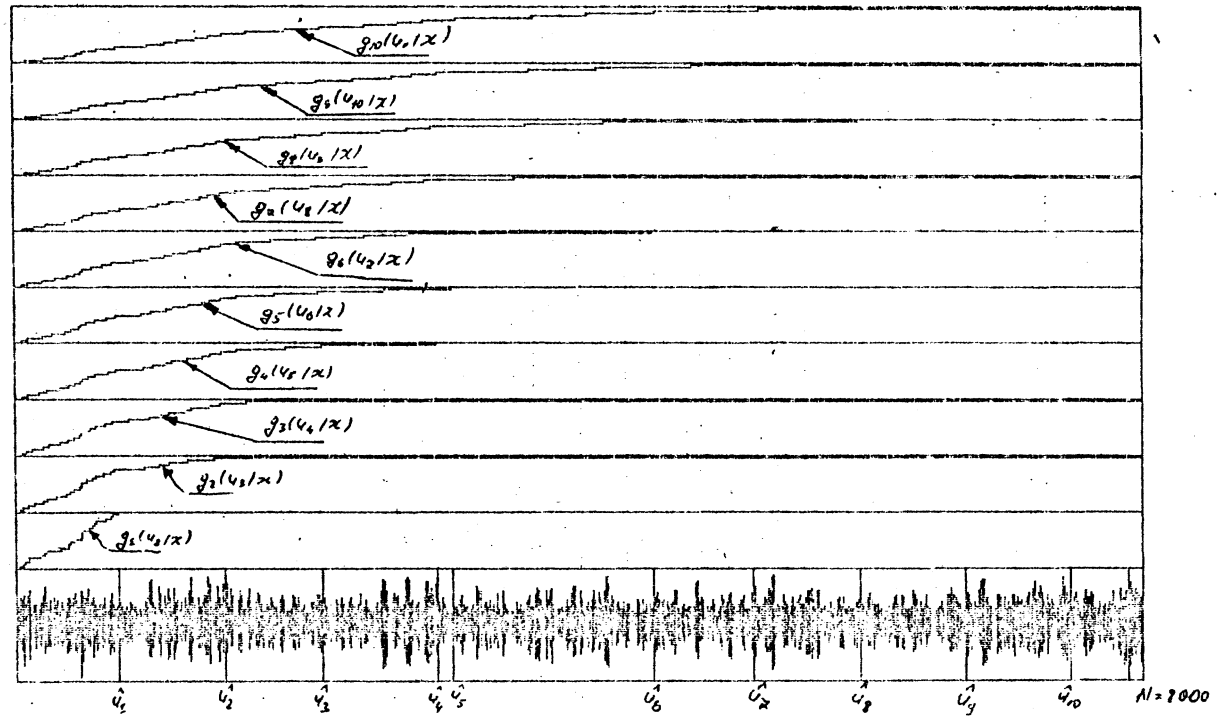

Fig. 1. The generated realization x and Bellmon functions $g_i(u_{i+1}/x)$, $i = 1, \ldots, 10$.

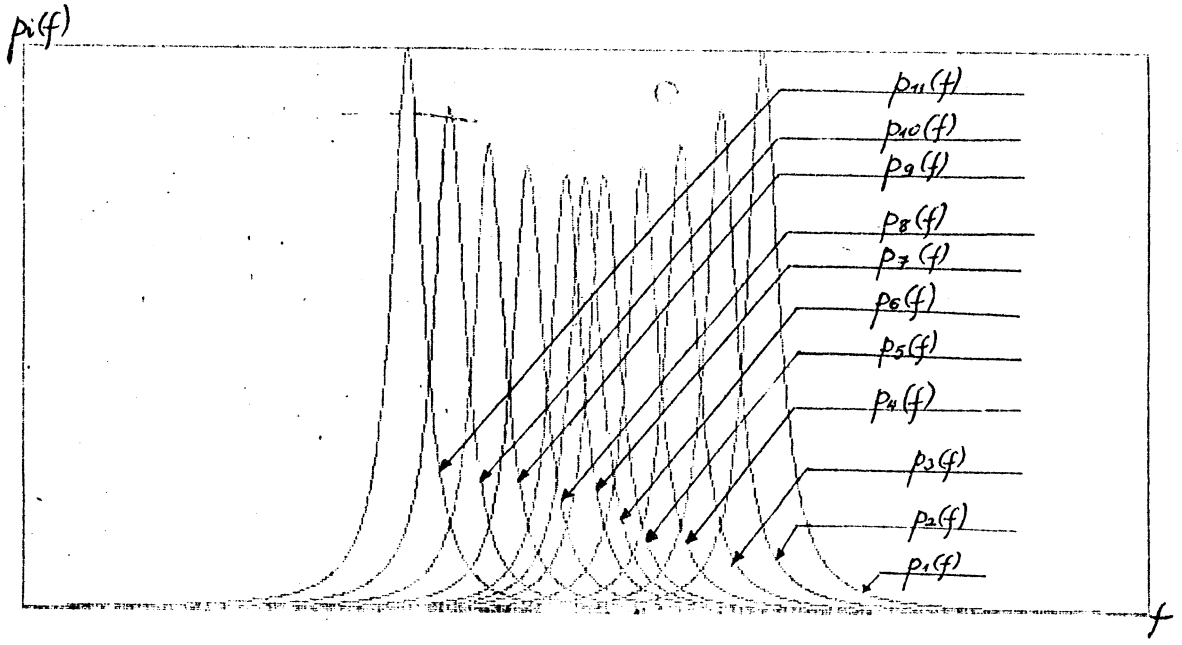

Fig. 2. Spectral densities $p_i(f)$ corresponding to parameters of AR model Q_i, $i = 1, \ldots, 11$.

correspond to the parameters of AR between the change-points. These functions show how much the parameters differ. We have got estimates of change-points which are presented in Table 2.

Table 2. Change-points and their estimates

i	1	2	3	4	5	6	7	8	9	10
u_i	750	1500	2250	3000	3750	4500	5250	6000	6750	7500
$\widehat{u}_i$	749	1497	2195	3007	3108	4541	5251	6002	6751	7494

6. Conclusions. The presented change-points detection algorithm enables us to solve the problem for comparatively long AR sequences and extends the sphere of applications. It can be used for the analysis of real random sequences (e.g., for segmentation of speech signals). For the use of the algorithm one must store two arrays of length N in the main menory of a computer instead of four such arrays as before.

REFERENCES

Kitagawa, G., and Akaike, H. (1978). A procedure for the modeling of nonstationary time series. *Ann. Inst. Statist. Math.*, **30**(B), 351–363.

Lipeika, A. (1975). Determination of changes in the properties of autoregressive sequences. A two-dimensional case. In L. Telksnys (Ed.), *Statisticheskie Problemy Upravleniya*, Vol.12, pp. 27–41 (in Russian).

Lipeika, A. (1977). Determination of changes in the properties of an autoregressive sequence. In L. Telksnys (Ed.), *Statisticheskie Problemy Upravleniya*, Vol.24, pp. 27–71 (in Russian).

Lipeika, A. (1979). On the determination of changes in the properties of an autoregressive sequence. In L. Telksnys (Ed.), *Statisticheskie Problemy Upravleniya*, Vol.39, pp. 9–23 (in Russian).

Lipeika, A. (1987a). On the determination of change-points in multivariate autoregressive sequences. In N.K. Sinha, and L. Telksnys (Eds.), *Stochastic Control*, IFAC Proceedings series, part II, Vilnius, pp. 267–269.

Lipeika, A., and Lipeikienė, J. (1987b). Application of the dynamic programming method to change detection in the properties of random autoregressive sequences. In L. Telksnys (Ed.), *Statisticheskie Problemy Upravleniya*, Vol.80, pp. 9–23 (in Russian).

Lipeika, A., and Lipeikienė, J. (1980). Main results on the estimation of change-points in multivariate autoregressive sequences. In L. Telksnys (Ed.), *Statisticheskie Problemy Upravleniya*, Vol.90, pp. 17–35.

Ozaki, T., and Tong, H. (1975). On fitting of nonstationary autoregressive models in time series analysis. *Proc. of the 8th Hawaii Internat. Conf. on Systems Science*, Western Periodical C^0, 224–226.

Praetorius, H., Bodenstein, G. and Creutzfelld, O. (1977). Adaptive segmentation of EEG records: A new approach to automatic EEG analysis. *Electrocncephalography and Clnical Neurophysiology*, **42**, 84–94.

Telksnys, L. (1970). Determination of maximum likelihood estimates of change-points in the properties of random signals. *Automatika i vychislitelnaja technika*, **1**, 24–27 (in Russian).

Received December 1991

A. Lipeika is a Candidate of Technical Sciences, a senior researcher of the Recognition Processes Departament at the Institute of Mathematics and Informatics (Lith. Acad. Sci). Scientific interests include: processing and recognition of random processes, detection of changes in the properties of random processes.

J. Lipeikienė is a Candidate of Technical Sciences, a senior researcher of the Recognition Processes Department at the Institute of Mathematics and Informatics (Lith. Acad. Sci). Scientific interests include: processing of random signals, robust methods for determination of change-points in the properties of random processes.