Tạp chí Tin học và Điều khiển học, T. 12, S. 4 (1996) (114–124)

VỀ MỘT THUẬT TOÁN CHẶT KHÚC CHUỖI THỜI GIAN THÀNH CÁC ĐOẠN TỰ HỒI QUY AR(p)

NGUYỄN HỒ QUỲNH, NGUYỄN TRUNG HÒA

Abstract. The problem of segmentationing time series is treated. The segments are considered as autoregressive models. The segmentation is based in the differency of the partiel autocorrelative function. A numeral result by Monte-Carlo simulation has been proposed in order to interprete the efficiency of the algorithm.

I - ĐẶT VẤN ĐỀ

Một chuỗi thời gian $\{x_t | t \in \mathbf{N}\}$ là một dãy các số thực phụ thuộc tham số thời gian rời rạc t. Cơ chế cảm sinh ra một chuỗi thời gian thường rất phức tạp. Với những quan điểm khác nhau người ta có thể đặt tương ứng mỗi chuỗi thời gian với một mô hình thống kê nào đó, hoặc cũng có thể gắn mỗi đoạn của chuỗi với một mô hình. Điều đó có nghĩa là: với một chuỗi thời gian giai đoạn này được đặt tương ứng với mô hình này, giai đoạn được đặt tương ứng với mô hình khác. Việc tách đoạn như vậy được gọi là *chặt khúc chuỗi thời gian* (segmentation of time series).

Một số phương pháp chặt khúc chuỗi thời gian như gắn nhãn cho các quan sát của Staley Sclove [9], dò tìm biên của từng khúc của Urich Appel và Achim Brandt [2], Michal Basseville và Albert Benvennist [3] đã được sử dụng. Tuy nhiên các phương pháp đã nêu đều phải sử dụng số các quan sát lớn và thường rất phức tạp.

Bài này đề xuất một phương pháp chặt khúc chuỗi thời gian bằng cách dò tìm biên đồng thời với việc ước lượng các tham số của nó thông qua hàm tự tương quan riêng của mỗi khúc, mà mỗi hàm tự tương quan riêng ấy đặc trưng cho một mô hình tự hồi quy, nghĩa là tách toàn bộ chuỗi đã cho thành một số đoạn tuân theo các mô hình tự hồi quy khác nhau.

II - CƠ SỞ CỦA THUẬT TOÁN

1. Mô hình tự hồi quy và các tự tương quan riêng

Ta nói rằng $\{x_t | t \in \mathbf{Z}\}$ là một *quá trình gauss, dừng cấp hai, quy tâm*, nghĩa là $\{x_t|t \in \mathbf{Z}\}$ là một họ các biến ngẫu nhiên thực, rời rạc, xác định trên một không gian xác suất $(\Omega, \mathcal{A}, P)$ thỏa mãn:

$$E(x_t) = 0\,, \quad E(x_t^2) = \sigma^2 < \infty\,, \quad E(x_t x_s) = \lambda_{t-s}\,, \tag{1}$$

trong đó: $E(.)$ là toán tử lấy kỳ vọng;

$t \in \mathbf{Z}$, $s \in \mathbf{Z}$ ($\mathbf{Z}$ là tập hợp các số nguyên).

Ta cũng nói rằng $\{x_t|t \in \mathbf{Z}\}$ là một quá trình *không kỳ dị* nếu không gian con $\mathcal{M}$ sinh bởi tập $\{x_t|t \in \mathbf{Z}\}$ của không gian $\mathbf{L}^2(\Omega, \mathcal{A}, P)$ có số chiều vô hạn.

Định nghĩa 1. Giả sử $\{x_t|t \in \mathbf{Z}\}$ là một quá trình gauss, không kỳ dị, dừng cấp hai, quy tâm có phương sai σ và tự hiệp phương sai giữa x_t và x_s là λ_{t-s}. Ta gọi *tự tương quan riêng cấp k* giữa x_t và x_{t+k} trong tập $\{x_t, x_{t+1}, ..., x_{t+k}\}$, ký hiệu là $\beta(k)$, là số thực thỏa mãn hệ thức truy hồi sau:

$$\sigma^2(0) = \sigma^2 \tag{2a}$$

Với mỗi $n > 0$:

$$\beta(n) = \frac{1}{\sigma^2(n-1)}\left[\lambda_n + \sum_{k=1}^{n-1} b(n-k,\, k)\lambda_{n-k}\right] \tag{2b}$$

$$\sigma^2(n) = [1 - \beta^2(n)]\sigma^2(n-1) \tag{2c}$$

$$b(n,\, n) = -\beta(n) \tag{2d}$$

$$b(n,\, k) = b(n-1,\, k) - b(n)\, b(n-1,\, n-k) \text{ với } 1 \le k \le n-1 \tag{2e}$$

Hàm $\beta(k)$ với $k \in \mathbf{Z}$ được gọi là *hàm tự tương quan riêng của quá trình* $\{x_t|t \in \mathbf{Z}\}$.

Định nghĩa 2. Một quá trình $\{x_t|t \in \mathbf{Z}\}$ được gọi là *quá trình tự hồi quy cấp p* ($p \in N^*$) ký hiệu $x_t \sim AR(p)$, nếu x_t thỏa mãn phương trình:

$$\epsilon_t = \sum_{i=0}^{p} a_i\, x_{t-i} \quad (t \in \mathbf{Z}) \tag{3a}$$

trong đó: ϵ_t là *ồn trắng*, nghĩa là

$$E(\epsilon_t) = 0\,; \;\; E(\epsilon_t^2) = \sigma_\epsilon^2 < \infty\,; \;\; E(\epsilon_t\, \epsilon_s) = 0\,, \text{ với } t \neq s \tag{3b}$$

và

$$a_0 = 1\,; \;\; a_p \neq 0 \tag{3c}$$

Với định nghĩa trên có thể nhận thấy rằng quá trình tự hồi quy là một quá trình dừng, quy tâm có phương sai hữu hạn

Đặt $\Phi^*(z) = \sum_{i=0}^{p} a_i z^i$ thì Φ^* được gọi là *đa thức tự hồi quy* tương ứng với quá trình tự hồi quy $\{x_t | t \in \mathbf{Z}\}$. Nếu quá trình đã cho là không kỳ dị thì đa thức tự hồi quy tương ứng có tất cả các nghiệm nằm ngoài đường tròn đơn vị U [4], và theo [10] nếu một quá trình $\{x_t | t \in \mathbf{Z}\}$ dừng, không kỳ dị thỏa mãn phương trình (3a) sao cho $\Phi^* = \sum_{i=0}^{p} a_i z^i$ có tất cả các nghiệm nằm ngoài đường tròn đơn vị U thì ϵ_t là ồn trắng, nghĩa là $\{x_t | t \in \mathbf{Z}\}$ là một quá trình tự hồi quy cấp p, hơn nữa ϵ_t không tương quan với $x_{t-1}, x_{t-2}, \ldots$

Mỗi quá trình tự hồi quy được đặc trưng bởi hàm tự tương quan riêng của nó. Điều đó được khẳng định bởi Dégerine [10].

Mệnh đề 1. *Quá trình dừng $\{x_t | t \in \mathbf{Z}\}$ là tự hồi quy cấp p khi và chỉ khi cá tự tương quan riêng của nó thỏa mãn các điều kiện:*

$$0 \leq |\beta(n)| < 1 \text{ với } 0 < n \leq p \tag{4a}$$

$$\beta(n) = 0 \text{ với } n > p \tag{4b}$$

$$\beta(b) \neq 0 \tag{4c}$$

2. Chặt khúc chuỗi thời gian

Định nghĩa 3. Giả sử $\{x_t | t \in \mathbf{Z}\}$ là một chuỗi thời gian dừng. Dãy bao gồm m quan sát liên tiếp $\{x_{i+1}, \ldots, x_{i+m}\}$ $(i \in \mathbf{N})$ được gọi là một *khúc tự hồi quy* của chuỗi đã cho nếu thỏa mãn các điều kiện sau:

(i) $\{x_{i+1}, \ldots, x_{i+m}\}$ tuân theo một quy một mô hình tự hồi quy cấp p nào đó, nghĩa là tồn tại $a_0, a_1, \ldots, a_p$ $(a_0 = 1,\ a_p \neq 0)$ sao cho $\epsilon_t = \sum_{k=0}^{p} a_k x_{t-k}$ là ồn trắng.

(ii) Các tập $\{x_i, \ldots, x_{i+m}$ và $\{x_{i+1}, \ldots, x_{i+m+1}\}$ không nghiệm đúng phương trình trên.

x_{i+1} và x_{i+m} được gọi lần lượt là *cận dưới* và *cận trên* của khúc đã cho.

Theo mệnh đề 1 có tương ứng một một giữa tập các quá trình tự hồi quy cấp p và tập các hàm $\beta(.)$ thỏa mãn các điều kiện (4a), (4b), (4c) của mệnh đề 1. Nói cách khác, hai chuỗi thời gian là các thể hiện của cùng một quá trình tự hồi quy khi và chỉ khi các hàm tự tương quang riêng của chúng trùng nhau.

Tuy nhiên, các quan sát x_t là ngẫu nhiên, nên $\beta(.)$ (sẽ được xác định theo các quan sát x_t) cũng là ngẫu nhiên. Vì vậy để làm cơ sở cho thuật toán chặt khúc ta thừa nhận giả thiết sau:

Ta có thể xem hai chuỗi thời gian $\{x_t|t \in \mathbf{N}\}$, $\{y_t|t \in \mathbf{N}\}$ không tuân theo cùng một mô hình tự hồi quy cấp p nào đó nếu như các tự tương quan riêng của mô hình này không thuộc khoảng tin cậy của tự tương quan riêng cùng cấp của mô hình kia. Hay nói một cách khác, chuỗi $\{y_t|t \in \mathbf{N}\}$ có thể xem như không tuân theo mô hình tự hồi quy cấp p tương ứng với chuỗi $\{x_t|t \in \mathbf{N}\}$ nếu trong số các tự tương quan riêng của chuỗi $\{y_t|t \in \mathbf{N}\}$ có ít nhất một tương quan riêng cấp k, chẳng hạn $\beta_y(k)$ không thuộc khoảng tin cậy $[\underline{\beta}_x(k), \overline{\beta}_x(k)]$ của sự tương quan riêng $\beta_x(k)$ của chuỗi $\{x_t|t \in \mathbf{N}\}$.

Với một chuỗi hữu hạn các quan sát $\{x_t|t \in \{1, 2, ..., m\}\}$ cho trước, ta sẽ ước lượng hàm tự tương quang riêng của đoạn đầu tiên bao gồm một số ít nhất (có thể được) các quan sát liên tiếp $x_1, x_2, ..., x_{m_0}$. Thêm quan sát x_{m_0+1} tiếp tục ước lượng hàm tự tương quan riêng của đoạn vừa bổ sung. Nếu hàm này thuộc khoảng tin cậy của hàm kia thì giả thuyết $\{x_1, x_2, ..., x_{m_0+1}\}$ và $\{x_1, x_2, ..., x_{m_0}\}$ *cùng thuộc một mô hình tự hồi quy* được chấp nhận, ngược lại, x_{m_0} được xem là cận trên của khúc đầu tiên.

Giả sử $\beta(k)$ là tự tương quan riêng cấp k của chuỗi các quan sát $\{x_t|t \in \{1, 2, ..., m\}\}$ (sẽ được xác định bởi thủ tục $MLE(x, m, n, \beta)$). Nếu sử dụng biến đổi Fisher, đặt $\alpha(k) = \frac{1}{2}\ln\frac{1+\beta(k)}{1-\beta(k)}$ thì theo [10], [11], $\alpha(k)$ có phân phối chuẩn với phương sai $\sigma_\alpha = \frac{1}{m-n-2}$, ta có thể xác định được khoảng tin cậy $[\underline{\alpha}(k), \overline{\alpha}(k)]$ của $\alpha(k)$. Khi đó khoảng tin cậy của $\beta(k)$ là $\underline{\beta}(k), \overline{\beta}(k)]$ được xác định bởi các công thức:

$$\underline{\beta}(k) = \frac{e^{2\underline{\alpha}(k)} - 1}{e^{2\underline{\alpha}(k)} + 1}, \quad \overline{\beta}(k) = \frac{e^{2\overline{\alpha}(k)} - 1}{e^{2\overline{\alpha}(k)} + 1} \tag{8}$$

Với một chuỗi thời gian cho trước, có nhiều phương pháp để ước lượng hàm tự tương quan riêng của nó, chẳng hạn xem [5, 6, 8]... Tuy nhiên chúng tôi chọn phương pháp ước lượng hợp lý cực đại lặp của các tác giả Phạm Đình Tuấn, vì phương pháp này có hiệu quả trong trường hợp có ít quan sát.

3. Ước lượng hợp lý cực đại lặp

Giả sử $\{x_t|t \in \mathbf{Z}\}$ là một quá trình gauss thỏa mãn định nghĩa 2. Ta gọi xích cấp n của quá trình là dãy các vectơ n chiều trong đó mỗi vectơ là n quan sát liên tiếp $(x^{(r)})^T = (x_1^{(r)}, x_2^{(r)}, ..., x_n^{(r)})$ $(r \in \mathbf{Z})$.

Hàm hợp lý loga dựa vào M phần tử của xích cấp n là:

$$L = -M[\ln(\det\pi\Lambda) + \operatorname{tr}(\Lambda^{-1}S)]$$

trong đó: Λ là ma trận covariance cấp n của quá trình và

$$S = (s_{ij}) = \frac{1}{M}\sum_{r=1}^{M}(x_1^{(r)}, x_2^{(r)}, ..., x_n^{(r)})^T\,(x_1^{(r)}, x_2^{(r)}, ..., x_n^{(r)}) \tag{5}$$

là ma trận convariance mẫu.

Nếu đặt $R = \frac{1}{\sigma^2}\Lambda$ thì

$$L = -M[\ln(\det\pi\sigma^2 R) + \frac{1}{\sigma^2}\mathrm{tr}(R^{-1}S)]$$

Theo [8], việc làm cực đại L tương đương với cực tiểu hóa L^* dưới đây:

$$L^* = \ln(\det R) + n\ln(\mathrm{tr}(R^{-1}S)) \tag{6a}$$

với

$$\ln(\det R) = -\sum_{k=1}^{n} k\ln[1 - \beta^2(k)] \tag{6b}$$

$$\mathrm{tr}(R^{-1}S) = \sum_{i,j=0}^{p} a_i\, q_{ij}\, a_j = a^T Q a \tag{6c}$$

trong đó: $\beta(k)$ là tự tương quan riêng cấp k, Q là ma trận vuông cấp $(p+1)$ thỏa mãn:

$$q_{ij} = q_{ji} = -q_{n-j,n-i} \tag{6d}$$

$$q_{ij} = \sum_{k=1}^{n-i-j} s_{k+i,k+j} \quad \text{với } i + j < n \tag{6e}$$

Với dãy có m quan sát, thống kê Q có thể tính được. Để cực tiểu hóa L^* ta cần xác định các $\beta(1)$, $\beta(2), \ldots, \beta(p)$. Phương pháp để tìm các $\beta(.)$ sao cho L^* đạt cực tiểu là phương pháp lặp lắc, nghĩa là với mỗi k thuộc tập $\{1,\ 2, \ldots,\ p\}$ ta xác định cự tiểu của L^* theo $\beta(k)$ trong khi giữ nguyên các $\beta(l)$ khác. Quá trình xác định các $\beta(k)$ đó được lặp đi lặp lại cho đến khi đạt được sự hội tụ cần thiết [8]. Trong thuật toán của Staley Sclove phương pháp này cũng được sử dụng dưới tên gọi là phương pháp giảm dư [8].

Ta ký hiệu thủ tục ước lượng hợp lý cực đại vừa nêu trên là $MLE(x,\ m,\ n,\ \beta)$ trong đó:

- x là chuỗi các quan sát liên tiếp,
- m là số các quan sát,
- n là cấp của xích,
- $\beta(.)$ là các tự tương quan riêng của chuỗi x.

Ma trận Q, được tính theo ma trận S, tương ứng với $M = m - n + 1$ phần tử của xích cấp n của chuỗi có m quan sát $\{x_t | t \in \{1,\ 2, \ldots,\ m\}\}$ sẽ được ký hiệu là $Q^{(m)}$. Nếu ký hiệu phần tử thứ r của xích cấp n của chuỗi $\{x_t | t \in \{1,\ 2, \ldots,\ m\}\}$

là $x^{(r)}$ thì $x^{(r)} = (x_{r+1}, ..., x_{r+n})$, $r = 0, 1, ..., M$. Khi đó $Q^{(m)}$ được tính trực tiếp như sau:

Bổ đề 1. *Với $i \geq 0, j \geq 0$, $i+j < n$, phần tử $q_{ij}^{(m)}$ của ma trận $Q^{(m)}$ trong công thức (6c) tương ứng với chuỗi $\{x_t | t \in \{1, 2, ..., m\}\}$ được tính bởi:*

$$q_{ij}^{(m)} = \frac{1}{m-n+1} \sum_{k=1}^{n-i-j} \sum_{r=0}^{m-n} x_{r+k+i}\, x_{r+k+j} \tag{7}$$

Chứng minh. Trong công thức (5) nếu chọn các vectơ mẫu là các phần tử liên tiếp của xích cấp n thì $(x^{(r)})^T = (x_{r+1}, x_{r+2}, ..., x_{r+n})$ $(r = 0, 1, ..., m-n+1)$. Khi đó:

$$s_{k+i}\, s_{k+j} = \frac{1}{m-n+1} \sum_{r=0}^{m-n} x_{r+k+i}\, x_{r+k+j}$$

do đó ta có:

$$\begin{aligned} q_{ij}^{(m)} &= \sum_{k=1}^{n-i-j} \frac{1}{m-n+1} \sum_{r=0}^{m-n} x_{r+k+i}\, x_{r+k+j} \\ &= \frac{1}{m-n+1} \sum_{k=1}^{n-i-j} \sum_{r=0}^{m-n} x_{r+k+i}\, x_{r+k+j} \end{aligned}$$

III - THUẬT TOÁN VÀ KẾT QUẢ SỐ

1. Thuật toán chặt khúc

Giả sử cần chặt khúc chuỗi gồm $D < \infty$ quan sát $\{x_t | t \in \{1, 2, ..., D\}\}$.

Bước chuẩn bị:

Cho trước cấp p của mô hình tự hồi quy: p có thể nhận được thông qua các thông tin tiên nghiệm về chuỗi thời gian, hoặc có thể nhận được dạng theo các phương pháp được đề cập trong [4], [7]....

Cho trước cấp n của xích: Theo [10], n phải thỏa mãn $n > [\frac{3}{2}p]$.

Cho trước điểm xuất phát m_0, $m_0 > n$.

Bước 1:

Ước lượng hàm tự tương quan riêng $\beta^{m_0)}$ của chuỗi $\{x_t | \{1, 2, ..., m_0\}\}$ bởi thủ tục $MLE(x, m_0, n, \beta^{(m_0)})$.

Đặt $m = m_0 + 1$.

Bước 2:

Nếu $m < D$ thì thực hiện thủ tục $MLE(x, m, n, \beta^{(m)})$, nếu ngược lại thì dừng.

Bước 3:

Nếu với mọi k thuộc tập 1, 2,..., p, $\beta^{(m)}(k)$ thuộc khoảng tin cậy của $\beta^{(m-1)}(k)$ thì tăng m và quay lại bước 2. Nếu ngược lại thì $m_1 = m - 1$ là cận trên của khúc tự hồi quy.

Quá trình được lặp đi lặp lại từ bước 1 (đối với mỗi chuỗi x_{m_1-n+1}, x_{m_1-n+2}, ..., x_D) nếu cận trên của đoạn vừa tìm được nhỏ hơn D.

2. Kết quả số:

Trong thuật toán trên phải nhiều lần thực hiện thủ tục $MLE(x, m, n, \beta)$ ở bước 1 và bước 2, trong đó có sự tham gia của ma trận $Q^{(m)}$ ứng với đoạn gồm m quan sát. Việc tính $Q^{(m)}$ theo (6e) hoặc theo Bổ đề 1 tốn rất nhiều thời gian (là thời gian tính chủ yếu của $MLE(x, m, n, \beta)$). Tuy nhiên để giảm đáng kể số lượng các phép tính $Q^{(m)}$ có thể được tính truy hồi theo công thức sau:

Mệnh đề 2. *Giả sử $Q^{(m)}$ là ma trận trong (6c) tương ứng với chuỗi $\{x_t | t \in \{1, 2, ..., m\}\}$ và $Q^{(m+1)}$ là ma trận tương ứng của chuỗi $\{x_t | t \in \{1, 2, ..., m+1\}\}$, khi đó với i $i \geq 0$, $j \geq 0$, $i + j < n$ ta có:*

$$q_{ij}^{(m+1)} = \frac{1}{m-n+2}\Big[(m-n+1)q_{ij}^{(m)} + \sum_{k=1}^{n-i-j} x_{m-n+1+k+i}\, x_{m-n+1+k+j}\Big] \quad (9)$$

Chứng minh. Thật vậy, theo Bổ đề 1, ta có:

$$q_{ij}^{(m+1)} = \frac{1}{m-n+2} \sum_{k=1}^{n-i-j} \sum_{r=0}^{m-n+1} x_{r+k+i}\, x_{r+k+j}$$

nên nhân cả hai vế với $(m - n + 2)$ thì:

$$\begin{aligned}(m-n+2)q_{ij}^{(m+1)} &= \sum_{k=1}^{n-i-j} \sum_{r=0}^{m-n+1} x_{r+k+i}\, x_{r+k+j} \\ &= \sum_{k=1}^{n-i-j} \Big(\sum_{r=0}^{m-n} x_{r+k+i}\, x_{r+k+j} + x_{m-n+1+k+i}\, x_{m-n+1+k+j} \Big) \\ &= \sum_{k=1}^{n-i-j} \sum_{r=0}^{m-n} x_{r+k+i}\, x_{r+k+j} + \sum_{k=1}^{n-i-j} x_{m-n+1+k+i}\, x_{m-n+1+k+j}\end{aligned}$$

do đó từ Bổ đề 1 ta có:

$$(m-n+2)q_{ij}^{(m+1)} = (m-n+1)q_{ij}^{(m)} + \sum_{k=1}^{n-i-j} x_{m-n+1+k+i}\, x_{m-n+1+k+j}$$

chia cả hai vế của đẳng thức cuối cùng cho $(m-n+2)$ ta có điều phải chứng minh.

Với mệnh đề trên ta chỉ cần phải sử dụng khoảng $2(n-i-j)+4$ phép tính cộng hoặc nhân để tính $q_{ij}^{(m+1)}$ thay vì phải sử dụng $(m-n+2)(n-i-j)$ phép tính (theo (7)). Hiển nhiên ở bước đầu tiên khi tính $Q^{(m_0)}$ vẫn phải sử dụng cách tính trực tiếp theo Bổ đề 1.

Thuật toán vừa nêu trên đã được chạy trên máy vi tính PC/AT 486. Chương trình được viết bằng ngôn ngữ Pascal. Các chuỗi đưa vào để chặt khúc là các chuỗi được nối liên tiếp bởi hai khúc tự hồi quy. Khúc đầu gồm 100 quan sát và khúc sau gồm 50 quan sát. Chúng được tạo nên bởi thuật toán mô phỏng [1]. Với các bộ hệ số tự hồi quy cố định cho trước, mỗi chuỗi ghép được mô phỏng nhiều lần, mỗi lần dò tìm vị trí cận trên $(m = 100)$ của khúc đầu đồng thời với việc ước lượng hệ số tự hồi quy của nó. Điểm xuất phát m_0 (nối ở bước chuẩn bị) để dò tìm cận trên của khúc đầu được chọn cố định là 30.

Mỗi bảng sau cho kết quả của việc dò tìm cận trên của khúc đầu qua 200 lần mô phỏng ước lượng các hệ số tự hồi quy của khúc đầu và dò tìm cận trên của một chuỗi ghép:

Bảng 1. Tín hiệu Narrow-band cấp 2

	a_1	a_2	m
Giá trị cho trước	-1,124	0,312	100
Giá trị trung bình	-1,064	0,292	96

Các cận trên của khúc đầu qua mỗi lần lặp:

101	101	72	102	88	102	102	72	92	102	102	101	102	115	101
102	102	102	101	101	102	101	102	108	50	101	54	102	102	81
101	101	101	102	101	102	102	102	101	91	102	101	102	81	101
108	67	102	57	101	102	101	102	102	102	101	101	61	101	102
83	101	101	101	101	101	87	102	101	61	102	102	102	101	
102	102	102	101	101	101	90	102	101	61	102	102	102	101	
101	101	74	102	102	102	101	62	101	101	102	102	71	102	
102	73	102	102	101	115	102	56	102	102	80	102	92	101	
102	102	101	102	102	102	102	102	102	102	102	88	102	101	
48	115	101	101	57	101	68	101	102	102	102	108	101	77	

84	102	50	101	101	102	102	101	101	86	102	94	102	102
101	102	108	81	97	102	88	101	101	115	102	53	101	101
102	102	91	102	102	53	102	102	96	96	102	43	77	115

Bảng 2. Tín hiệu Narrow-band cấp 3

	a_1	a_2	a_3	m
Giá trị cho trước	-0,771	-0,277	0,223	100
Giá trị trung bình	-0,743	-0,204	0,163	98

Các cận trên của khúc đầu qua mỗi lần lặp:

102	51	85	102	107	109	107	101	83	102	109	109	109	101	
99	109	89	93	107	66	109	107	48	109	109	109	98	109	102
101	109	107	102	102	107	107	99	109	101	103	85	107	68	109
109	107	93	109	101	102	51	109	109	107	86	101	96	102	102
109	62	102	87	101	107	102	101	102	107	50	109	107	109	
109	109	102	60	107	109	101	109	87	109	109	79	86	107	
102	70	88	107	109	101	102	78	109	107	107	101	101	103	
107	86	107	102	109	109	107	107	109	107	109	72	109	84	
109	101	109	109	66	107	96	102	107	91	109	109	107	101	
107	107	107	51	107	88	101	83	102	103	77	102	62	102	
98	107	103	107	107	107	109	107	109	109	97	101	109	46	
101	101	107	63	109	109	101	107	99	104	109	78	101	107	
91	109	101	100	103	77	109	104	107	109	109	103	102	83	
109	109	101	107	81	109	63	44	107	109	98	109	109	93	

Bảng 3. Tín hiệu Narrow-band cấp 5

	a_1	a_2	a_3	a_4	a_5	m
Giá trị cho trước	-0,636	-1,018	0,635	0,252	-0,154	100
Giá trị trung bình	-0,637	-0,916	0,572	0,227	-0,143	93

Các cận trên của khúc đầu qua mỗi lần lặp:

93	101	102	53	101	95	91	101	90	102	101	73	93	101	102
102	104	59	102	101	58	95	101	101	72	85	101	94	102	87
64	102	103	102	102	102	101	101	101	101	101	102	101	102	84
101	101	101	101	101	102	63	102	97	102	71	101	95	101	102

102	83	74	101	101	101	101	101	71	101	101	74	83	101
101	101	101	57	85	53	101	101	101	102	96	101	102	102
87	101	102	58	77	101	101	101	102	102	102	101	102	102
101	101	102	47	101	102	102	77	96	70	101	101	89	101
101	101	101	101	77	101	101	101	69	100	102	64	101	96
92	85	84	95	52	101	100	101	61	102	102	102	104	82
101	102	102	81	102	98	101	89	75	101	101	101	78	101
74	102	101	61	101	101	101	101	73	84	59	102	101	101
67	102	98	74	101	101	101	102	101	91	101	101	102	101
98	101	74	83	57	101	102	101	83	101	70	101	102	102

Bảng 4. Tín hiệu Wide-band cấp 5

	a_1	a_2	a_3	a_4	a_5	m
Giá trị cho trước	0,984	0,379	0,071	0,007	0,000	100
Giá trị trung bình	0,940	0,397	0,118	0,041	-0,008	100

Các cận trên của khúc đầu qua mỗi lần lặp:

109	109	109	105	105	109	106	92	109	109	109	105	109	106	71
103	105	65	103	109	109	101	107	109	109	59	109	106	105	107
109	83	109	106	105	105	47	91	109	109	109	78	72	106	59
90	106	106	94	87	105	101	50	109	109	107	103	103	106	109
89	107	90	109	69	105	105	109	106	106	109	55	84	109	
107	106	101	88	109	103	102	103	109	105	106	109	76	109	
103	109	103	105	109	107	106	67	54	109	105	109	107	106	
105	69	107	105	109	109	92	105	109	106	109	76	109	106	
105	69	107	105	109	109	92	105	109	106	109	76	109	90	
102	109	106	106	134	109	107	109	107	109	109	107	109	105	
65	106	109	105	105	101	95	107	105	107	109	103	70	109	
105	107	105	106	103	109	105	107	88	106	79	103	103	109	
101	109	94	109	107	107	109	109	103	103	109	67	76	79	
102	87	105	103	78	106	107	69	109	105	103	105	109	103	
109	109	105	103	109	102	103	103	103	109	94	107	109	107	

(phần tử dòng i cột j là kết quả cả vòng lặp thứ $i + (j-1)^*$ (tổng số dòng)

So sánh các giá trị cho trước với các giá trị trung bình tính được cho thấy thuật toán có hiệu quả hơn so với các chuỗi có đa thức tự hồi quy của nó có nghiệm nằm ngoài, gần đường tròn đơn vị (tín hiệu Narrow-band).

Các tác giả chân thành cảm ơn xemina toán ứng dụng thuộc liên trường ĐHTH Hà Nội, ĐHBK Hà Nội đã hỗ trợ, cổ vũ và đóng góp nhiều ý kiến để chúng tôi có thể hoàn thành công trình này.

TÀI LIỆU THAM KHẢO

1. Nguyễn Trung Hòa, *Về một thuật toán mô phỏng chuỗi thời gian tự hồi quy AR(p)*, Bài sẽ đăng ở tạp chí Tin học và Điều khiển học, (1996).
2. Appel U. and Brandt A. V., *Adaptive sequential segmentation of piecewise stationaray time series*, Information Sciences, **29** (1983), 27 - 56.
3. Basseville M. and Benveniste A., *Sequential segmentation of nonstationary digital signals using spectral analysis*, Information Sciences, **29** (1983), 57 - 73.
4. Box G. E. P and Jenkins G. M., *Time series analysis, forecasting and control*, Holden-day, 1970.
5. Kay S. M., *Recursive maximum likehood estimation of autoregressive processes*, IEEE Trans. ASSP, **31** (1983), 56 - 65.
6. Le Cadre J. P., *Maximum likehood estimation of an autoregressive model. Application to spatial signal processing*, Traitement du signal, **2-4** (1985), 291 - 303.
7. Matix Wax, *Order selection for AR models by predictive least squares*, IEEE trans. ASSP, **36** (1988), 581 - 588.
8. Pham D. T., *Maximum likehood estimation of aturegressive model by relaxation on the reflection coefficients*, IEEE Trans. ASSP, **365** (1988), 1363 - 1367
9. Sclove S., *Tie series segmetation a model and a method*, Information sciences, **29** (1983), 7 - 25.
10. Dégerine S., *Fonction d'autocorrelation partielle et estimation autoregressive dans le domaine temporel*, Thèse de doctorat ès science Mathématiques, Université de Joseph Fourier, Grenoble I, France, 1988.
11. Ванде Варден Б. Л., Математическая статистика, Иностранной литературы, Москва, 1960.

Khoa Toán - Đại học Bách khoa Hà Nội.

Khoa Toán - Đại học Sư phạm Vinh.

Nhận bài ngày 28-12-1995