*Tạp chí Tin học và Điều khiển học T 12, S 1 (1996) (21 - 28)*

# VỀ MỘT THUẬT TOÁN MÔ PHỎNG MÔ HÌNH TỰ HỒI QUY AR(p)

NGUYỄN TRUNG HOÀ

An algorithm of simulating the regular stationary gaussian autoregressive model is proposed. Numerical examples have been issued in order to interpete the efficiency of the algorithm .

## I. ĐẶT VẤN ĐỀ

Mô hình tự hồi quy AR là một trong những mô hình có nhiều ứng dụng trong kinh tế kỹ thuật, đặc biệt là trong lý thuyết xử lý tín hiệu. Bài toán nhận dạng mô hình $AR(p)$ là bài toán ước lượng bậc hồi quy $p$ và các hệ số tự hồi quy $a_k$, $(k = 1, ..., p)$ đã được nhiều tác giả quan tâm và đã có nhiều thủ tục khác nhau. Chẳng hạn các thủ tục xác định cấp $p$ được đề cập trong [1], [2], 3[],... và các thủ tục tục ước lượng $a_k$ có thể tìm thấy trong [2], [4], [5], [6],... Trên thực tế, để kiểm tra tính ưu việt của mỗi thủ tục nhận dạng người ta phải cho chạy mô hình đó trên rất nhiều dãy số liệu cụ thể khác nhau. Tuy nhiên không dễ gì để có được nhiều dãy số liệu như vậy. Bài này đề nghị một thủ tục để mô phỏng các dãy số liệu tuân theo mô hình $AR(p)$ với $p$ cho trước và các $a_k$, $k = 1, ..., p$ cho trước.

## II. CƠ SỞ THUẬT TOÁN

### 2.1. Hàm tự tương quan riêng của một quá trình dừng, không kỳ dị

Trong mục này, nếu không nói gì thêm, ta luân giả thiết $\{x_t |\ t \in \mathbf{Z}\}$ là một quá trình dừng cấp hai, quy tâm, nghiã 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

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

trong đó

- $E(.)$ là một toán tử kỳ vọng

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

Giả sử $\mathcal{M}$ là không gian con của $L^2(\Omega, \mathcal{A}, P)$ sinh bởi các $x_t$. Khi đó phần tử của $\mathcal{M}$ là tổ hợp tuyến tính có dạng $\sum \alpha_t x_t$, ($\alpha_t$ là các số thực). Tích vô hướng trong $\mathcal{M}$ được cảm sinh bởi tích vô hướng trong $L^2(\Omega, \mathcal{A}, P)$ là toán tử hiệp phương sai, nghiã là $\langle x_t, x_s \rangle = E(x_t, x_s)$. Chuẩn trong $\mathcal{M}$ là phương sai của biến ngẫu nhiên, $\|x_t\|^2 = E(x_t^2)$.

Không gian $\mathcal{M}$ được gọi là *chính quy* nếu số chiều của $\mathcal{M}$ là vô hạn. Khi đó quá trình $\{x_t |\ t \in \mathbf{Z}\}$ được gọi là quá trình *không kỳ dị* (non singulier) [7].

Giả sử $\{x_t |\ t \in \mathbf{Z}\}$ là quá trình không kỳ dị, gọi $\mathcal{M}(t,n)$, $(n > 0)$ là không gian con sinh bởi tập $\{x_t, x_{t-1}, ..., x_{t-n+1}\}$, $\mathcal{M}(t, 0) = \{0\}$. Vì không gian này có số chiều là $n$ nên hình chiếu trực giao của $x_t$ lên $\mathcal{M}(t-1,n)$ được biểu diễn duy nhất dưới dạng: [7]

$$x(t,n) = -\sum_{k=1}^{n} b(n,k) x_{t-k} \tag{2}$$

Đặt

$$\varepsilon(t,n) = x_t - x(t,n) = \sum_{k=0}^{n} b(n,k) x_{t-k}, \ \ (b(n,0) = 1) \tag{3}$$

và ta sẽ gọi $\varepsilon(t,n)$ thỏa mãn (3) là *thặng dư tiến riêng cấp n* của $x_t$.

Hoàn toàn tương tự, ta gọi $\varepsilon^*(t,n)$ là *thặng dư lùi riêng cấp n* của $x_t$:

$$\varepsilon^*(t,n) = x_t - x^*(t,n) = \sum_{k=0}^{n} b(n,k) x_{t+k}, \tag{4}$$

vì giả thiết về tính dừng của $\{x_t |\ t \in \mathbf{Z}\}$ nên phương sai của $\varepsilon(t,n)$ và phương sai của $\varepsilon^*(t,n)$ bằng nhau và được ký hiệu bởi

$$\sigma^2(n) = \|\varepsilon(t,n)\|^2 = \|\varepsilon^*(t,n)\|^2. \tag{5}$$

**Định nghĩa 2.1.1.** *Hàm tương quan riêng* của quá trình $\{x_t |\ t \in \mathbf{Z}\}$, kí hiệu $\beta(.)$, là hàm xác định trên tập $\mathbf{Z}$ các số nguyên, được cho bởi biểu thức

$$\beta(0) = 1, \tag{6.a}$$

$$\beta(n) = \langle \varepsilon(t, n-1), \varepsilon^*(t-n, n-1) \rangle / \sigma^2(n-1), \ \ (n > 0) \tag{6.b}$$

$$\beta(n) = \beta(-n), \ \ (n < 0), \tag{6.c}$$

vì $\{x_t |\ t \in \mathbf{Z}\}$ dừng nên (6.c) là hợp lý.

Với $k$ cố định thuộc $N^*$, $\beta(k)$ được gọi là *tự tương quan riêng cấp k* giữa $x_t$ và $x_{t+n}$ trong tập $\{x_t, ..., x_{t+n}\}$.

Với giả thiết Gauss, hàm tự tương quan riêng sẽ được tính qua phương sai và các hiệp phương sai của quá trình theo hệ thức truy hồi sau:

**Mệnh đề 2.1.2** *(Thuật toán Levison-Durbin)* [7].

Giả sử $\{x_t|\ t \in \mathbf{Z}\}$ là qwuá trình Gauss, dừng cấp hai, quy tâm, không kỳ dị có $E(x_t^2) = \sigma^2$ và $E(x_t, x_k) = \lambda_{t-k}$. Khi đó các tự tương quan riêng $\beta(n)$ thỏa mãn hệ thức truy hồi sau

$$\sigma^2(0) = \sigma^2, \text{ với mỗi } n > 0: \tag{7.a}$$

$$\beta(n) = \overline{1\sigma^2(n-1)}\left[\lambda_n + \sum_{k=1}^{n-1} b(n-k,k)\lambda_{n-k}\right] \tag{7.b}$$

$$\sigma^2(n) = \left[1 - \beta^2(n)\right]\sigma^2(n-1) \tag{7.c}$$

$$b(n,n) = -\beta(n) \tag{7.d}$$

$$b(n,k) = b(n-1,k) - \beta(n)b(n-1,n-k), \text{ với } 1 \le k \le n-1. \tag{7.e}$$

Để xây dựng một chuỗi thời gian dừng $\{x_t|\ t \in \mathbf{Z}\}$ có hm tự tương quan riêng cho trước ta sử dụng mệnh đề sau:

**Mệnh đề 2.1.3.** *(Dégerine)* [7]

Giả sử $\{x_t|\ t \in \mathbf{Z}\}$ là một dãy trực giao của $L^2(\Omega, \mathcal{A}, P)$ thỏa mãn

$$\|y_0\|^2 = \sigma^2(0) = \sigma^2, \ \|y_n\|^2 = \sigma^2(n) = [1 - \beta^2(n)]\sigma^2(n-1), \ n \in N^*, \tag{8a}$$

trong đó

$$|\beta(n)| \le 1(8.b)$$

$$\text{Nếu } |\beta(n)| = 1, \text{ thì } \beta(n) = \beta(n+1). \tag{8.c}$$

Khi đó chuỗi thời gian $\{x_t|\ t \in \mathbf{Z}\}$, sao cho

$$x_0 = y_0 \tag{9.a}$$

$$x + n = x(n, n-1) + \beta(n)\varepsilon^*(0, n-1) + y_n, \ n \in N^* \tag{9.b}$$

là một dãy dừng có các hàm tự tương quan riêng là $\beta(.)$.

### 2.2. Mô hình tự hồi quy và các tự tương quan riêng

Một quá trình $\{x_t|\ t \in \mathbf{Z}\}$ được gọi là *quá trình tự hồi quy cấp* $(p \in N^*)$, kí hiệu $x_t \sim \mathbf{A}R(p)$, nếu $x_t$ thỏa mãn phương trình

$$\varepsilon_t = \sum_{i=0}^{p} a_i x_{t-i} \ (t \in \mathbf{Z}), \tag{10.a}$$

trong đó: $\varepsilon_t$ là *ồn trắng*, nghĩa là

$$E(\varepsilon_t) = 0,\ E(\varepsilon_t^2) = \sigma_\varepsilon^2 < \infty;\ E(\varepsilon_t \varepsilon_s) = 0,\ \text{với } t \neq s \tag{10.b}$$

$$a_0 = 1,\ a_p \neq 0. \tag{10.c}$$

Với định nghiã 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 $\Phi^*$ 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 $U$ [2], và theo [7] nếu một quá trình $\{x_t|\ t \in \mathbf{Z}\}$ dừng, không kì dị thỏa mãn phương tình (10.a) sao cho $\Phi^* = \sum_{i=0}^{p} a_i z^i$ có tất cả các nghiệm nằm ngòa đường tròn đơn vị $U$ thì $\varepsilon_t$ là ồn trắng, nghiã là $\{x_t|\ t \in \mathbf{Z}\}$ là một quá trình tự hồi quy cấp $p$, hơn nữa $\varepsilon_t$ 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 [7]:

**Mệnh đề 2.2.2.**

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

$$0 \leq |\beta(n)| < 1 \text{ với } 0 < n \leq p \tag{11.a}$$

$$\beta(n) = 0 \text{ với } n > p \tag{11.b}$$

$$\beta(n) \neq 0 \tag{11.c}$$

Định lý sau đây cho phép tính các tự tương quan riêng của một quá trình không kỳ dị, tự hồi quy cấp $p$ khi biết các hệ số tự hồi quy của nó.

**Định lý 2.2.3.** *Giả sử $\{x_t|\ t \in \mathbf{Z}\}$ là quá trình không kìi dị tự hồi quy cấp p thỏa mãn (10.1), (10.b), (10.c). khi đó các tự tương quan riêng tương ứng với quá trình đó được tính truy hồi theo công thức*

$$\text{Với } k = 0, 1, \ldots, p;\ b(p,k) = a_k \tag{12.a}$$

$$\beta_p = -b(p,p) \tag{12.b}$$

*Với $n = p, \ldots, 2$:*

$$\text{Với } k = 1, \ldots, n-1:\ b(n-1,k) = [b(n,k) + \beta_n b(n, n-k)]/[1 - \beta^2(n)] \tag{12.c}$$

$$\beta(n-1) = -b(n-1, n-1). \tag{12.d}$$

*Chứng minh.* Giả sử $\beta(n)$ là tự tương quan riêng cấp $n$ và $b(n,k)$ là các hệ số biểu diễn hình chiếu trực giao của $x_t$ lên $\mathcal{M}(t-1,n)$. Từ định nghiã 2.2.1 ta có

$$x_t = -\sum_{i=1}^{p} a_i x_{t-i} + \varepsilon_t$$

với $\varepsilon_t$ là ồn trắng, và vì quá trình được giả thiết là không kỳ dị nên

$$E(\varepsilon_t(-\sum_{i=1}^{p} a_i x_{t-i})) = E(-\sum_{i=1}^{p} a_i x_{t-i}\varepsilon_t) = 0$$

suy ra $-\sum_{i=1}^{p} a_i x_{t-i})$ là hình chiếu trực giao của $x_t$ lên $\mathcal{M}(t-1,p)$, do đó $b(p,k) = a_k$ (với $k = 0,1,...,p$) và $\beta(p) = -b(p,p)$.

Theo mệnh đề 2.1.2 thì $\beta(n-1) = -b(n-1,n-1)$ và

$$\begin{cases} b(n,k) = b(n-1,k) - \beta(n)b(n-1,k-1) \\ b(n,n-k) = b(n-1,n-k) - \beta(n)b(n-1,k) \end{cases}$$

và từ hệ trên dễ dàng có được

$$b(n-1,k) = [b(n,k) + \beta_n b(n,n-k)]/[1-\beta^2(n)].$$

**Định lý 2.2.4.** *Giả sử $\{\xi_t|\ (t = 0,1,...)\}$ là dãy các ồn trắng và $\{\beta(k)|\ (n = 0,...)\}$ là dãy các số thực thỏa mãn: $0 \le \beta(k) < 1$ với $k = 1,...,p$; $\beta(p) \neq 0$; $\beta(0) = 1$; $\beta(k) = 0$ với $k > p$. Khi đó chuỗi thời gian $\{x_t|\ t \in \mathbf{Z}\}$ được tạo nên bởi hệ thức truy hồi sau là chuỗi tự hồi quy cấp p có các tự tương quan riêng là $\beta(k)$ :*

$$x(0) = \xi_0$$

*Với* $t = 1,...,p$

$$x_t = -\sum_{k=1}^{t} b(t,k)x_{t-k} + \sqrt{\Pi_{k=1}^{t}[1-\beta^2(k)]\xi_t}$$

*Với* $t = p+1, p+2, ...$

$$x_t = -\sum_{k=1}^{t} b(p,k)x_{t-k} + \sqrt{\Pi_{k=1}^{p}[1-\beta^2(k)]\xi_t}$$

*Chứng minh.* Đặt $y_0 = \xi_0$ và $y_t = \sqrt{\Pi_{k=1}^{t}[1-\beta^2(k)]\xi_t}$ (với $t = 1,2,...$) nên $\|y_0\|^2 = \sigma_\xi^2$; $\|y_t\|^2 = \sigma^2(t) = \Pi_{k=1}^{t}[1-\beta^2(k)]\sigma_\xi^2$. Vì $\beta(p+1) = \beta(p+2) = ... = 0$ nên $\sigma^2(p) = \sigma^2(p+1) = ... = \sigma^2$ không đổi, và $y_t$ là dãy trực giao do $\xi_t$ là dãy các ồn trắng.

Theo mệnh đề 2.1.3, chuỗi thời gian $\{x_t|\ t \in \mathbf{Z}\}$ thỏa mãn

$$x_0 = y_0$$

$$x_t = x(t,t-1) + \beta(t)\varepsilon^*(0,t-1) + y_t$$

là chuỗi dừng có các tự tương quan riêng cấp $k$ là $\beta(k)$ $(k \in N^*)$.

Thay $x(t, t-1)$ và $\varepsilon^*(0, t-1)$ ở đẳng thức trên bởi các hệ thức (2), (4) ta có

$$x_t = -\sum_{k=1}^{t-1} b(t-1,k)x_{t-k} + \beta(t)\sum_{k=0}^{t-1} b(t-1,k)x_k + y_t$$

$$x_t = -\sum_{k=1}^{t-1} b(t-1,k)x_{t-k} + \beta(t)\sum_{k=1}^{t-1} b(t-1,t-k)x_{t-k} + \beta(t)x_0 + y_t$$

$$x_t = -\sum_{k=1}^{t-1} [b(t-1,k)x_{t-k} + \beta(t)b(t-1,t-k)]x_{t-k} - b(t,t)x_0 + y_t.$$

Theo mệnh đề 2.1.2, từ đẳng thức cuối cùng ta đạt được

$$x_t = -\sum_{k=1}^{t} b(t,k)x_{t-k} + y_t.$$

Vì $\beta(t) = 0$ với $t > p$ nên

$$x_t = -\sum_{k=1}^{p} b(t,k)x_{t-k} + y_t, \text{ với } t > p$$

và nếu đặt $a_k = b(p,k)$ và $\varepsilon_t = y_t$ thì

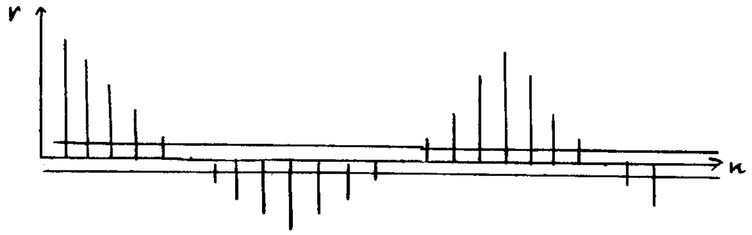
$$x_t = -\sum_{k=1}^{t} a_k x_{t-k} + \varepsilon_n$$

trong đó $\varepsilon_t$ (với $t \geq p$) là các ồn trắng có phương sai $\sigma_\varepsilon^2 = \Pi_{k=1}^t[1-\beta^2(k)]\sigma_\xi^2$. Vì vậy, theo mệnh đề 2.2.2 thì $\{x_t|\ t \in \mathbf{Z}\}$ là chuỗi tự hồi quy cấp $p$.

Như vậy, theo định lý 2.2.3, nếu cho trước $p$ hệ số tự hồi quy của quá trình không kỳ dị ta sẽ tính được các tự tương quan riêng, và theo định lý 2.2.4 ta sẽ xây dựng được chuỗi tự hồi quy cấp $p$ khi biết $p$ tự tương quan riêng đầu tiên của chuỗi. Để đảm bảo rằng chuỗi được tạo nên là không kỳ dị ta phải chọn các hệ số tự hồi quy sao cho đa thức hồi quy có $p$ nghiệm phân biệt nằm ngoài đường tròn đơn vị $U$. Ta có thuật toán sau.

## III. THUẬT TOÁN VÀ KẾT QUẢ SỐ

*Bước 1.*




Hình 1. Tự tương quan mẫu của chuỗi Narrow-band, Gauss AR(3)

Cho trước $p$ nghiệm của $\Phi^*(z)$ nằm ngoài đường tròn đơn vị $U$ (có thể có các cặp nghiệm phức liên hợp): $z_1, z_2, ..., z_p$.

Khai triển chính tắc

$$\Phi^*(z) = (\Pi_{i=1}^{p} z_i)^{-1} \Pi_{i=1}^{p} (z - z_i)$$

dưới dạng

$$\Phi^*(z) = \sum_{i=0}^{p} a_i z^i$$

để có $a_0 = 1$.

*Bước 2*

Tính các $\beta(k)$, $k = 1, ..., p$ và ma trận $b(n, k)$ cấp $p \times p$ theo định lý 2.2.3.

*Bước 3.*

Tạo chuỗi tự hồi quy có phân phối cho trước từ dãy các ồn trắng cho trước theo định lý 2.2.4.

Thuật toán đã được thể hiện trên ngôn ngữ chương trình Pascal, với các ồn trắng tuân theo phân phối chuẩn $\mathcal{N}(0,1)$ do công thức $\xi_t = \sqrt{-2log\varsigma_{t_1}}cos(2\pi\varsigma_{t_1+1}$, [8] trong đó $\varsigma_{t_1}$ và $\varsigma_{t_1+1}$ là các biến ngẫu nhiên liên tiếp sẵn có của Turbo Pascal.

Các đa thức tự hồi quy được chọn theo hai loại: Loại thứ nhất, đặc trưng cho tín hiệu narrow-band, có tất cả các nghiệm nằm ngoài, sát đường tròn đơn vị. Loại thứ hai, đặc trưng cho tín hiệu wide-band, có tất cả các nghiệm nằm ngoài, xa đường tròn đơn vị.

Trong mỗi trường hợp, các tự tương quan mẫu của chuỗi tạo nên đã được tính lại theo [2] và được thể hiện trên các hình 1,2. Qua các đồ thị nhận được, ta nhận thấy rằng dáng điệu của các hàm tự tương quan mẫu đều tuân theo quy luật sin/mũ và tắt dần tại những trễ bậc cao, theo [2] điều đó chứng tỏ các chuỗi mô phỏng thực sự là các chuỗi tự hồi quy.


Hình 2. Tự tương quan mẫu của chuỗi Wide-band, Gauss AR(3)

Tác giả tỏ lòng biết ơn GS Nguyễn Hồ Quỳnh đã tận tình giúp đỡ để có thể hoàn thành được bài báo này. Đồng thời tác giả cũng chân thành cám ơn xemina Toán ứng dụng thuộc liên trường ĐHTH, ĐHBK,... và các đồng nghiệp đã hỗ trợ, cổ vũ nhiều khi tác giả thực hiện bài báo này.

## TÀI LIỆU THAM KHẢO

1. Vũ Chấn Hưng, *Về một thuật toán đánh giá hệ động lực tuyến tính*, Tin học và Điều khiển học, Tập 10, N. 1. 1994, 24-29.
2. Box G.E.P & Jenkins G.M., Time series analysis, forecasting and control, Holden-day 1970.
3. Dickie J.R. & Nandi A.K., *A comparative study of AR order selection methods*, Signal processing, 40 (2-3) (1994), 239-255.
4. Kay S.M., *Recursive maximum likehood estimation of autoregressive processes, IEEE Trans. ASSP Vol. 31, 1983, 56-65.*
5. *Le Cadre J.P., Maximum likehood estimation of an autoregressive model. Application to spatial signal processing, Traitement du signal, 2-4, 1985, 291-303.*
6. *Pham D.T., Maximum likehood estimation of an autoregressive model by relaxation on the reflection coeeficients, IEEE Trans. ASSP V. 36, 1988, 1363-1367.*
7. *Dégerine S., Fonction d'autocorrelation et partielle et estimation autoregressive dans le domaine temporel, Thése de doctorat ès sciences "Mathématiques, Université de Joseph Fourier, Grenoble I., France 1988.*
8. *Ermakov C.M. Montecarlo method and relative problems M.Nauka, 1975 (in Russian).*

*Khoa toán, Đại học sư phạm Vinh*

*Nhận ngày 3, 9 1995*