

# STATISTIQUE ET ANALYSE DES DONNÉES

ROSE-MARIE KESTEMONT

## **Sur l'évaluation de la fonction de vraisemblance du modèle ARMA stationnaire**

*Statistique et analyse des données*, tome 9, n° 1 (1984), p. 60-75

[http://www.numdam.org/item?id=SAD\\_1984\\_\\_9\\_1\\_60\\_0](http://www.numdam.org/item?id=SAD_1984__9_1_60_0)

© Association pour la statistique et ses utilisations, 1984, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

*Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques*

<http://www.numdam.org/>

Statistique et Analyse des Données  
mai 1984 - Vol. 9 n° 1 - pp 60-75

SUR L'EVALUATION DE LA FONCTION DE  
VRAISEMBLANCE DU MODELE ARMA STATIONNAIRE

Rose-Marie KESTEMONT

Unité PROB  
Université Catholique de Louvain  
Chemin du Cyclotron 2  
B 1348 Louvain-la-Neuve

Résumé : Nous donnons une présentation directe et simplifiée de l'évaluation de la fonction de vraisemblance du modèle ARMA stationnaire laquelle permet non seulement de retrouver aisément les résultats connus mais permet également de rectifier simplement un résultat erroné.

Abstract : We give a direct and simplified presentation for the evaluation of the likelihood function of a stationary ARMA model which not only lead easily to known results but also permit to correct an erroneous one.

Mots clés : fonction de vraisemblance, modèle ARMA.

0 - INTRODUCTION

Supposons avoir affaire à un modèle ARMA (p,q) univarié. Nous ferons toujours l'hypothèse que le processus stochastique qui génère la série chronologique est un processus gaussien, écrit habituellement sous la forme

$$\phi(B)\omega_t = \theta(B)a_t \quad (0.1)$$

où 1) B est l'opérateur de retard :  $B\omega_t = \omega_{t-1}$

$$2) \phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p \quad (0.2)$$

$$3) \theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q \quad (0.3)$$

4)  $a_t$  est un bruit blanc, c'est-à-dire que les  $a_t$  sont des v.a. indépendantes normales de moyenne nulle et de variance  $\sigma^2$ .  
Nous ferons en plus l'hypothèse que le processus est stationnaire et inversible, c'est-à-dire que les zéros des polynômes  $\phi$  et  $\theta$  sont extérieurs au cercle unité.

Considérons la série chronologique  $\omega = (\omega_1, \dots, \omega_n)'$  générée par ce processus ARMA (p,q). Le problème statistique posé est celui de l'estimation des paramètres  $\phi = (\phi_1, \dots, \phi_p)'$ ,  $\theta = (\theta_1, \dots, \theta_q)'$  et  $\sigma^2$  du modèle.

Puisque le processus stochastique qui génère la série chronologique est gaussien, la loi de probabilité du vecteur  $\omega$  est normale multivariée, de moyenne nulle et de matrice de covariance  $\sigma^2 \Sigma$  où

$$\Sigma = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & & . \\ \gamma_2 & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & \gamma_1 \\ \gamma_{n-1} & . & . & . & \gamma_1 & \gamma_0 \end{pmatrix} \quad (0.4)$$

car, à cause de la stationnarité,

$$\text{cov}(\omega_i, \omega_j) = \sigma^2 \gamma_{|i-j|}.$$

La fonction de vraisemblance du modèle est

$$L(\phi, \theta, \sigma^2) = (2\pi\sigma^2)^{-n/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2\sigma^2} \omega' \Sigma^{-1} \omega\right). \quad (0.5)$$

Comme il semble utopique d'espérer obtenir analytiquement les estimateurs du maximum de vraisemblance des paramètres, on est obligé de recourir à des procédures d'approximation qui toutes demandent l'évaluation répétée de la fonction de vraisemblance  $L$ . La difficulté majeure dans l'évaluation de  $L$  (ou de manière préférable de  $\ln L$ ) réside dans la nécessité d'inverser la matrice  $\Sigma$ , de taille  $n \times n$ , et d'en calculer son déter-

minant. Ceci n'est pas un problème simple même avec l'aide d'un ordinateur de grande dimension.

Par un choix judicieux d'une transformation des données de départ, nous viserons à remplacer l'inversion de la matrice  $\Sigma$  par celle d'une autre matrice dont l'ordre sera réduit au maximum. Ceci nous permettra non seulement de retrouver de manière directe, synthétique et élégante des résultats obtenus par divers auteurs mais encore de rectifier un point d'histoire. En effet, l'article de Ali (1977) contient dès le départ une erreur, relevée par Ansley (1979), qui rend toutes les formules inexactes. Il est pourtant cité par Priestley (1981). La correction proposée par Ali semble bien mystérieuse. Pourtant tout s'éclaire quand on remarque que la correction n'est qu'un cas particulier de l'article de Dent (1977) non cité par Priestley.

Les notations utilisées sont habituelles dans le problème. Les quantités introduites gardent la même signification tout au long du travail. Les matrices sont notées par des lettres majuscules et les vecteurs, par des lettres minuscules. Nous affectons du signe \* le vecteur des conditions initiales.

Remarque : Les méthodes examinées dans la suite ne sont pas les seules. Il resterait au moins les méthodes par factorisation et les méthodes ré-cursives. Le problème du temps de calcul et de la place en mémoire n'a pas été examiné. Pour ces deux aspects, on pourrait consulter les articles de Mèlard (1982) et (1983).

## 1 - PREMIERE REDUCTION

Dans ce paragraphe, nous nous proposons d'indiquer comment on peut ramener à  $p+q$  l'ordre de la matrice à inverser dans l'évaluation de la fonction de vraisemblance.

Le point de départ consiste à traduire les relations



non corrélées et par la relation (1.1), nous déduisons la matrice de covariance du vecteur  $\omega$ , soit  $\sigma^2 \Sigma$  où

$$\begin{aligned}\Sigma &= L_1^{-1} [L_2 L_2' + V \Sigma_{u^*} V'] L_1'^{-1} \\ &= L_1^{-1} L_2 [I + L_2^{-1} V \Sigma_{u^*} V' L_2'^{-1}] L_2' L_1'^{-1}.\end{aligned}\tag{1.2}$$

En posant

$$Z = L_2^{-1} L_1 \quad \text{et} \quad X = L_2^{-1} V$$

on obtient encore

$$\Sigma = Z^{-1} [I + X \Sigma_{u^*} X'] Z'^{-1}.$$

D'après les propriétés rappelées en annexe, il vient

$$\Sigma^{-1} = Z' [I - X \{\Sigma_{u^*}^{-1} + X'X\}^{-1} X'] Z\tag{1.3}$$

$$|\Sigma| = |\Sigma_{u^*}| |\Sigma_{u^*}^{-1} + X'X|.\tag{1.4}$$

Les expressions (1.3) et (1.4) ont été obtenues par un raisonnement analogue à celui suivi par Dent (1977) et Ali (1977). Elles montrent que le calcul de l'inverse de la matrice  $\Sigma$  et celui de son déterminant sont ramenés à l'inversion de deux matrices d'ordre  $p+q$ .

#### Remarques :

1) Galbraith & Galbraith (1974), Newbold (1974), Ljung & Box (1979) étendent aux processus ARMA l'approche de Box & Jenkins développée pour les processus MA, laquelle en faisant intervenir la fonction de densité de  $\omega$  et  $u^*$  allonge les calculs qui conduisent aux formules (1.3) et (1.4) contrairement à l'approche développée ici qui est plus directe.

2) Newbold (1974) introduit un artifice supplémentaire qui est en fait la décomposition de Cholesky de  $\Sigma_{u^*}$  sous la forme

.65.

$$T\Sigma_{u^*}T' = I_{p+q} \quad (1.5)$$

laquelle est encore équivalente à l'introduction du vecteur

$$e^* = Tu^* \quad (1.6)$$

dont la matrice de covariance est  $\sigma^2 I_{p+q}$ .

Cette dernière transformation facilite à nouveau l'évaluation des formules (1.3) et (1.4) car l'inversion de la matrice  $\Sigma_{u^*}$  se ramène maintenant à celle de la matrice triangulaire  $T$ .

3) Ceci termine les techniques développées pour ramener à  $p+q$  l'ordre des matrices à inverser dans la relation (1.3). C'est une transformation du type (1.6) qui conduira à ramener à  $m = \max(p,q)$  l'ordre des matrices à inverser dans l'évaluation de  $\Sigma^{-1}$  et  $|\Sigma|$ .

## 2 - DEUXIEME REDUCTION

Partant de la relation (0.1), on peut écrire suivant Dent (1977)

$$L_1 w = L_2 a + \Delta u \quad (2.1)$$

où 1° les matrices  $L_1$  et  $L_2$  sont définies comme plus haut

2°  $u = (u_1, \dots, u_m)'$  est un vecteur lié à  $u^*$  par la relation

$$u = V_1 u^* \quad (2.2)$$

ce qui revient à introduire la matrice  $\Delta$  telle que

$$\Delta u = Vu^* \quad (2.3)$$

et lui donne la forme particulière d'une matrice  $n \times m$ , de rang  $m$ , où  $m = \max(p,q)$

$$\Delta = \begin{pmatrix} I_m \\ 0_{n-m,m} \end{pmatrix}.$$

Il vient

$$\Delta \Sigma_U \Delta' = V \Sigma_{U*} V'$$

et de plus, la matrice  $\Sigma_U$  est régulière.

En suivant pas à pas la démarche utilisée au paragraphe précédent et en conservant les mêmes notations, on obtient

$$\Sigma^{-1} = Z' \{ I_n + M \Sigma_U M' \}^{-1} Z \quad (2.4)$$

dans laquelle

$$M = L_2^{-1} \Delta.$$

Suivant la formule rappelée en annexe, on a encore

$$\Sigma^{-1} = Z' \{ I_n - M [\Sigma_U^{-1} + M'M]^{-1} M' \} Z. \quad (2.5)$$

De même, on trouve

$$|\Sigma| = |\Sigma_U| |\Sigma_U^{-1} + M'M|. \quad (2.6)$$

Les expressions (2.4) et (2.5) ont été trouvées par Dent (1977) et mystérieusement par Ali (1977) puisque sa transformation de base était erronée. Elles demandent l'inversion de deux matrices d'ordre  $m = \max(p,q)$ .

#### Remarques.

1° Du point de vue historique, la relation (2.2) est importante car en l'explicitant, on montre aisément que l'expression



$$u_i = \sum_{k=1}^m (\phi_{i+k-1} \omega_{1-k} - \theta_{i+k-1} a_{1-k}) \quad i=1, \dots, m \quad (2.7)$$

dans laquelle

$$\begin{aligned} \phi_j &= 0 \text{ si } j > p \\ \theta_j &= 0 \text{ si } j > q \end{aligned}$$

est précisément la variable définie par Ali (1977) dans la correction apportée à son travail.

2° La relation (2.3) peut encore s'écrire

$$u = \Delta^{-} V u^* \quad (2.8)$$

où  $\Delta^{-} = \begin{pmatrix} I_m & \\ & 0_{m, n-m} \end{pmatrix}$  est l'inverse généralisée de  $\Delta$ .

3° Par rapport au paragraphe précédent, la méthode développée ici n'apporte en fait une amélioration que pour les processus ARMA, puisque  $\max(p, q) < p+q$ . Pour les processus AR(p) ou pour les processus MA(q), les résultats coïncident avec ceux obtenus par Newbold.

### 3 - LE POINT DE VUE NUMERIQUE

Du point de vue numérique, le point capital est l'évaluation de la forme quadratique  $\omega' \Sigma^{-1} \omega$  qui figure dans la fonction de vraisemblance. En posant

$$\begin{aligned} x &= Z\omega = L_2^{-1} L_1 \omega \\ y &= M'x = \Delta' L_2^{-1} x \end{aligned}$$

nous obtenons l'expression déduite de (2.5)

$$\omega' \Sigma^{-1} \omega = x'x - y'(\Sigma_u^{-1} + M'M)^{-1} y. \quad (3.1)$$

évaluation du vecteur  $x = (x_1, \dots, x_n)'$

De la relation

$$L_2 x = L_1 \omega$$

et par la définition des matrices  $L_1$  et  $L_2$ , on voit que les éléments du vecteur  $x$  se calculent par itération à partir du modèle général

$$\theta(B)x_t = \phi(B)\omega_t \quad t=1, \dots, n$$

autrement dit  $x_t = \theta_1 x_{t-1} + \dots + \theta_1 x_{t-q} + \omega_t - \phi_1 \omega_{t-1} \dots - \phi_p \omega_{t-p}$   
 $t=1, \dots, n$  dans lequel  $x_t = \omega_t = 0 \quad t \leq 0$ .

Remarque : Le  $x_t$  ainsi obtenu coïncide avec le  $\tilde{a}_t$  obtenu par la méthode conditionnelle (Box-Jenkins) avec des valeurs initiales nulles. De ce fait, le second terme de (2.5) peut s'interpréter, comme l'ont fait Ljung & Box, en tant que terme correctif dû à l'absence de conditions initiales.

évaluation du vecteur  $y = (y_1, \dots, y_m)'$

Le vecteur

$$z = L_2^{-1} x$$

est facilement calculable puisque la matrice  $L_2$  est facilement inversible. De plus, la relation

$$y = \Delta' z = (I_m \mid 0_{n-m, m}) z$$

montre que

$$y_i = z_i \quad i=1, 2, \dots, m.$$

évaluation de la matrice  $\Sigma_u$

Pour montrer que l'évaluation de (2.6) est facile, il reste à montrer comment on peut obtenir aisément les termes de la matrice  $\Sigma_u$  car il suffira alors d'appliquer un algorithme d'inversion.

A partir de la formule (2.7), on déduit

$$E(u_i u_j) = \sigma^2 \sum_{k=1}^m \sum_{l=1}^m \{ \phi_{i+k-1} \phi_{j+l-1} \gamma_{|k-l|} - \phi_{i+k-1} \theta_{j+l-1} \psi_{l-k} - \phi_{j+l-1} \theta_{i+k-1} \psi_{k-l} + \theta_{i+k-1} \theta_{j+l-1} \delta_{kl} \}$$

où 1°)  $\sigma^2 \gamma_{|h|} = \text{cov}(\omega_t, \omega_{t+h})$

2°)  $\sigma^2 \psi_h = \text{cov}(a_t, \omega_{t+h})$

avec  $\psi_0 = 1$  et  $\psi_h = 0$  si  $h < 0$

3°)  $\delta_{kl}$  est le symbole de Kronecker.

Les différents coefficients  $\gamma_h, \psi_h$  contenus dans cette formule s'obtiennent en résolvant les systèmes

1°)  $\psi_h = \sum_{i=1}^p \phi_i \psi_{h-i} - \theta_h \quad (h=1,2,\dots)$

avec  $\theta_i = 0$  si  $i > q$

puisque

$\text{cov}(a_t, \phi(B)\omega_{t+h}) = \text{cov}(a_t, \theta(B)a_{t+h}) \quad (h=1,2,\dots)$

2°)  $\psi_{|h|} = \sum_{i=1}^p \phi_i \gamma_{|h+i|} + \psi_h - \sum_{i=1}^q \theta_i \psi_{h+i}$

puisque

$\text{cov}(\omega_{t+h}, \phi(B)\omega_t) = \text{cov}(\omega_{t+h}, \theta(B)a_t).$

#### 4 - AMELIORATION DE DENT

Le résultat obtenu au paragraphe 2 nécessite encore l'inversion de deux

matrices et, au paragraphe 3, nous avons montré comment évaluer la forme quadratique  $\omega' \Sigma^{-1} \omega$ .

Dent (1977) propose une technique qui évite l'inversion de ces matrices. Il procède en deux étapes. Tout d'abord, il effectue la décomposition spectrale de la matrice  $\Sigma_U$  sous la forme

$$\Sigma_U = ADA' \tag{4.1}$$

où  $D = \text{diag}(\lambda_1, \dots, \lambda_m)$   
 $A = [a_1, a_2, \dots, a_m]$ .

$\lambda_i$  étant une valeur propre de  $\Sigma_U$  et  $a_i$  le vecteur propre correspondant.

En définissant la matrice T de type nxm et de rang  $\leq m$  par

$$T = L_2^{-1} \Delta AD^{1/2} \tag{4.2}$$

on peut écrire l'expression (2.4) sous la forme

$$\Sigma^{-1} = Z'(I_n + TT')^{-1}Z. \tag{4.3}$$

il ne reste plus maintenant qu'à effectuer la décomposition en valeurs singulières de T. Celle-ci se fait sous l'une des deux formes suivantes

$$a) T = U^\circ S^\circ W' \tag{4.4}$$

où 1°  $U^\circ$  est une matrice orthogonale nxn formée des vecteurs propres ortho-normalisés (en colonnes) de  $TT'$ .

$$2^\circ S^\circ = \begin{bmatrix} S \\ \dots \\ 0 \end{bmatrix}$$

est une matrice nxm dans laquelle S est une matrice diagonale positive d'ordre m dont les éléments diagonaux sont les racines carrées positives des valeurs propres de  $T'T$ .

3° W est une matrice formée des vecteurs propres orthonormalisés de T'T.

$$b) T = U S W' \quad (4.5)$$

où U est la matrice formée des m premières colonnes de U° correspondant aux m valeurs propres non nulles de TT'.

Il vient encore (cf. annexe)

$$(I_n + TT')^{-1} = I_n - (US)(I_m + S^2)^{-1}(S'U') \quad (4.6)$$

et finalement, par (4.3),

$$\omega' \Sigma^{-1} \omega = x'x - z'S(I_m + S^2)^{-1}S'z \quad (4.7)$$

ou encore

$$\omega' \Sigma^{-1} \omega = x'x - \sum_{i=1}^m \delta_i z_i^2 \quad (4.8)$$

dans laquelle

$$1^\circ x = Z\omega = L_2 L_1^{-1} \omega$$

$$2^\circ z = U'x$$

$$3^\circ \delta_i = s_{ii}^2 / (1 + s_{ii}^2).$$

Pour le déterminant, on trouve (cf. annexe)

$$\begin{aligned} |\Sigma| &= |I_m + T'T| \\ &= |I_m + S^2| \end{aligned}$$

et finalement,

$$|\Sigma| = \prod_{i=1}^m (1 + s_{ii}^2). \quad (4.9)$$

Il est important de remarquer que la procédure développée par Dent n'oblige à aucune inversion de matrice et de plus elle ne nécessite de gar-

der en mémoire que les termes diagonaux de la matrice S, le calcul du vecteur z se faisant à l'intérieur de l'algorithme de décomposition en valeurs singulières.

### 5 - DEUX CAS PARTICULIERS

Le cas des processus AR(p) conduit directement aux relations suivantes

$$\begin{aligned}
 L_2 &= I_n \\
 x &= L_1 \omega \\
 H &= I_n + \Delta \Sigma_u \Delta' \\
 T &= \Delta A D^{1/2} = \begin{bmatrix} A D^{1/2} \\ \dots \\ 0 \end{bmatrix} = \begin{pmatrix} A & 0 \\ 0 & I_{n-p} \end{pmatrix} \begin{bmatrix} D^{1/2} \\ 0 \end{bmatrix} I_p.
 \end{aligned}$$

Cette dernière expression fournit immédiatement la décomposition en valeurs singulières de T.

Le cas des processus MA(q) conduit directement aux relations suivantes

$$\begin{aligned}
 L_1 &= I_n \\
 u &= R a^* \quad \text{où} \quad R = \begin{pmatrix} -\theta_q & \dots & -\theta_1 \\ & \ddots & \\ 0 & & -\theta_q \end{pmatrix} \\
 a^* &= (a_{1-q}, \dots, a_0)'.
 \end{aligned}$$

Il en résulte que

$$\Sigma_u = R R',$$

ce qui donne immédiatement la décomposition spectrale de  $\Sigma_u$ .

Annexe

Soit A une matrice définie positive d'ordre s et B une matrice quelconque de dimension rxs, ( $r > s$ ), alors :

- 1)  $I_r + BAB'$  est une matrice définie positive d'ordre r
- 2)  $A^{-1} + B'B$  est définie positive d'ordre s
- 3)  $(I_r + BAB')^{-1} = I_r - B(A^{-1} + B'B)^{-1}B'$
- 4)  $|I_r + BAB'| = |A| |A^{-1} + B'B|$ .

Démonstration : Les trois premières propriétés sont évidentes. En posant

$$U = BA$$

$$N = B'$$

la quatrième résulte de la relation

$$\begin{vmatrix} I_r & -U \\ N & I_s \end{vmatrix} = |I_r + UN| = |I_s + NU|.$$

En particulier, si  $A = I_s$ , on a :

$$(I_r + BB')^{-1} = I_r - B(I_s + B'B)^{-1}B'.$$

Bibliographie

- ALI, M.M. (1977) : Analysis of autoregressive-moving average models : estimation and precision. *Biometrika* 64, 535-545.
- ALI, M.M. (1978) : Corrections. *Biometrika* 65, 677.
- ANSLEY, C.F. (1979) : An algorithm for the exact likelihood of a mixed autoregressive-moving average process. *Biometrika* 66, 59-65.
- BOX, G.E.P. & JENKINS, G.M. (1975) : Time series analysis, Forecasting and Control. Holden-Day, San Francisco (2nd edition).
- DENT, W.T. (1977) : Computation of the exact likelihood function for an ARIMA process. *J. Statist. Comput. Simul.* 5, 193-206.
- GALBRAITH, R.F. & GALBRAITH, J.I. (1974) : On the inverse of some patterned matrices arising in the theory of stationary time series. *J. Appl. Prob.* 11, 63-71.
- KESTEMONT, R.M. (1980) : La fonction de vraisemblance du modèle ARMA stationnaire. Mémoire de licence en Sciences Mathématiques, U.C.L.
- LJUNG, G.M. & BOX, G.E.P. (1979) : The likelihood function of stationary autoregressive-moving average models. *Biometrika* 66, 265-270.
- MELARD, G. (1982) : L'application de la méthode du maximum de vraisemblance dans des modèles de séries chronologiques. *Cahiers du C.E.R.O.*, vol. 24, n<sup>os</sup> 2-3-4.
- MELARD, G. (1983) : A fast algorithm for the exact likelihood of autoregressive-moving average models. Soumis pour publication.
- NEWBOLD, P. (1974) : The exact likelihood function for a mixed autoregressive-moving average process. *Biometrika* 61, 423-426.



.75.

PRIESTLEY, M.B. (1981) : Spectral analysis and time series (vol. I). Academic Press, New York.