

# CRITERE OPTIMISANT UN COMPROMIS BIAIS-VARIANCE POUR ESTIMER L'ORDRE D'UN PROCESSUS AUTOREGRESSIF MONO OU MULTIVARIABLE

Bernard LUMEAU\* et Jean-François BERCHER\*\*

\* LSS, CNRS-ESE, Plateau de Moulon, 91192 GIF-SUR-YVETTE Cedex

\*\* Service National, CELAR, 35170 BRUZ

## RESUME

Un critère est proposé pour l'estimation de l'ordre d'un processus autorégressif mono ou multivariable. L'ordre est estimé en modélisant les erreurs d'estimation comme l'effet d'un bruit blanc additif de variance connue. Cette modélisation est obtenue en exprimant le «vecteur de régression observé» sur une autre base. La base définie, l'ordre optimal est obtenu par minimisation d'un critère quadratique faisant apparaître un Compromis entre un terme de Biais et un terme de Variance (CBV). Cette approche s'étend aux processus ARMA à phase minimale. Elle se généralise au cas multivariable. Le critère proposé, noté CBV, est comparé aux critères classiques AIC et MDL. Il donne en particulier de meilleurs résultats pour des processus AR bruités ou des ARMA.

## I. INTRODUCTION

Dans le domaine de la prédiction linéaire [1] d'un processus autorégressif, dans les méthodes d'estimation spectrale paramétrique [2] et dans les méthodes dites à haute résolution [3] le choix de l'ordre du processus AR sous-jacent est un problème crucial. De nombreux critères ont été proposés pour estimer pratiquement cet ordre [4, 5] : AIC, FPE, MDL, CAT... Tous ces critères jouent sur le compromis entre une bonne modélisation du signal (ordre élevé) et une bonne stabilité statistique (ordre faible) : ils font intervenir une fonction de l'erreur de prédiction (décroissante avec l'ordre) et une fonction de l'ordre. La diversité et le nombre de critères utilisés montrent bien qu'il n'existe pas d'absolu en ce domaine. Les principaux reproches faits à ces critères sont de sous-estimer l'ordre lorsque le nombre de points est faible [6], ou encore (pour le critère AIC d'Akaike [4]) d'être inconsistant.

Nous avons cherché à définir un critère qui soit simple d'emploi, justifié théoriquement et efficace. Notre approche consiste à mettre en oeuvre un critère optimisant un compromis entre un terme de biais et un terme de variance afin de déterminer l'ordre optimal au sens de ce compromis. Ce critère s'inscrit dans une suite logique à la démarche qui nous a permis d'estimer la matrice de densité spectrale, au sens de ce même compromis [7]. Nous donnons, dans une première partie, une présentation générale de l'approche de ce problème d'optimisation. Nous montrons, dans la deuxième partie, par quelle transformation cette approche peut être appliquée aux processus autorégressifs. Nous étendons, dans une troisième partie, ces résultats au cas des processus autorégressifs à moyenne ajustée à phase minimale. La généralisation de cette approche au cas multivariable est envisagée dans la quatrième partie. Enfin, pour conclure, quelques comparaisons statistiques avec les critères classiques sont présentées.

## ABSTRACT

A criterion is proposed to estimate the order of a mono or multivariate autoregressive process. The order is estimated by a modelling of the estimation errors as the result of an additive white noise which has a known variance. This modelling is obtained from a transformation of "the observed regression vector" on another basis. After having defined the basis, the optimal order of the decomposition is obtained by a minimization of a quadratic criterion which shows a Compromise between the terms of Bias and the terms of Variance (CBV). This approach can be extended to a minimum phase ARMA model. It is generalized for the multivariate case. The proposed criterion, which is called CBV, is compared to the classic criteria AIC and MDL. In particular, it provides better results for noisy AR and ARMA models.

## II. APPROCHE GENERALE DU PROBLEME

L'approche générale du problème est introduite pour un modèle statistique d'une observation  $u(i)$  consistant en un signal  $s(i)$  à estimer, considéré comme déterministe mais inconnu, perturbé par un bruit blanc  $b(i)$  additif de variance supposée connue  $\sigma^2$ . Nous montrerons dans les paragraphes suivants comment on peut se ramener à un tel modèle à partir des coefficients du vecteur de régression estimé. La partie signal est estimée de manière paramétrique : nous représentons  $s(i)$  par un développement tronqué sur une base orthonormée, dont l'ordre est estimé à l'aide d'un critère d'erreur quadratique minimale. Ainsi, si nous considérons  $K$  observations, le signal observé s'écrit sous forme vectorielle dans un espace de dimension  $K$

$$(1) \quad \mathbf{x} = \mathbf{s} + \mathbf{b} \text{ où } \mathbf{x} = [x(1), \dots, x(K)]^T,$$

avec

$$(2) \quad E[\mathbf{x}] = \mathbf{s} \text{ et } E[|\mathbf{b}|^2] = \sigma^2.$$

Nous pouvons décomposer chacun des signaux composant (1) dans une base orthonormée particulière  $\{v_1, v_2, \dots, v_K\}$  :

$$(3) \quad \mathbf{x} = \sum_{i=1}^K \eta(i)v_i = \sum_{i=1}^K \delta(i)v_i + \sum_{i=1}^K \beta(i)v_i.$$

Nous en déduisons entre les projections

$$(4) \quad \eta(i) = \delta(i) + \beta(i).$$

La transformation entre les  $b(i)$  et les  $\beta(i)$  étant orthogonale, il s'en suit que  $\beta(i)$  est un bruit blanc :

$$(5) \quad E[|b(i)|^2] = E[|\beta(i)|^2] = \sigma^2.$$

Le choix de la base étant fait, considérons l'effet de l'estimation du



signal par un développement tronqué à l'ordre  $p$ , sous la forme

$$(6) \quad \hat{s} = \sum_{i=1}^p \eta(i)v_i = \sum_{i=1}^p (\delta(i) + \beta(i))v_i.$$

L'erreur quadratique moyenne totale d'estimation s'écrit :

$$(7) \quad Q(p) = E [s - \hat{s}]^2 = p\sigma^2 + \sum_{i=p+1}^N |\delta(i)|^2.$$

Elle fait apparaître un terme de variance qui croît avec  $p$  et un terme de biais qui décroît avec  $p$ ; la minimisation de cette erreur quadratique conduit alors à optimiser le compromis biais-variance. Cette erreur ne peut être calculée exactement à partir des observations car les  $\delta(i)$  sont inconnus. D'après (4) et (5), une estimation non biaisée  $\hat{Q}(p)$  de  $Q(p)$  nous est donnée en remplaçant dans (7)  $|\delta(i)|^2$  par  $(|\eta(i)|^2 - \sigma^2)$  qui a même espérance :

$$(8) \quad \hat{Q}(p) = (2p-N)\sigma^2 + \sum_{i=p+1}^N |\eta(i)|^2.$$

Le minimum de  $Q(p)$  sera atteint pour une valeur proche de celui de  $\hat{Q}(p)$  et sera donc utilisable comme critère d'estimation de l'ordre optimal de troncature. Le résultat fournit par (8) peut d'ailleurs être comparé au critère MAICE d'Akaike [4] dans le cas où les  $\{\beta(i)\}$  sont gaussiens.

Pour des raisons de simplicité, dans la mesure où ce n'est que l'ordre optimal de troncature que l'on veut déterminer, nous pouvons utiliser le critère équivalent

$$(9) \quad \hat{Q}'(p) = \sum_{i=1}^p (|\eta(i)|^2 - 2\sigma^2)$$

égal à  $-\hat{Q}(p)$  à une constante près. L'ordre optimal correspond cette fois au maximum de  $\hat{Q}'(p)$ . Ce critère est donc croissant pour  $|\eta(i)|^2 > 2\sigma^2$ , soit en moyenne, en «présence de signal», pour  $|\delta(i)|^2 > \sigma^2$ , et décroissant en «l'absence de signal», pour  $|\delta(i)|^2 < \sigma^2$ . Ainsi, la valeur de  $p$  qui maximise le critère quadratique,  $p_0$ , est donc telle que :

$$(10) \quad \begin{cases} |\eta(i)|^2 \geq \sigma^2 + \text{pour } i \leq p_0, \\ |\eta(i)|^2 \geq \sigma^2 + \text{pour } i > p_0. \end{cases}$$

### III. ORDRE OPTIMAL D'UN PROCESSUS AUTOREGRESSIF

Un signal autorégressif  $x(n)$  d'ordre  $p_0$  est représenté par un ensemble de  $(p_0+1)$  paramètres :  $\{a(1), \dots, a(p_0); \sigma^2\}$ , où  $\sigma^2$  est la puissance du bruit blanc générateur  $u(n)$  supposé centré et stationnaire.

$$(11) \quad x(n) = -a(1)x(n-1) - \dots - a(p_0)x(n-p_0) + u(n).$$

Nous pouvons interpréter le vecteur de régression  $\mathbf{a} = [a(1), \dots, a(p_0)]^T$  comme une représentation du signal sur une base particulière. Par suite, si l'ordre sélectionné  $K$  est inférieur à l'ordre réel  $p_0$ , la mauvaise adéquation du modèle se traduira par un effet de biais. Par contre, si l'ordre  $K$  est supérieur à  $p_0$ , les paramètres excédentaires qui en moyenne ont une contribution nulle vont avoir un effet sur l'accroissement de la variance. Ce comportement statistique a déjà été étudié [5], les résultats dont nous disposons sont des résultats asymptotiques (nombre de points tendant vers l'infini) utilisant la borne de Cramer-Rao :

1. l'estimée d'ordre  $K$ , avec  $K > p_0$ , du vecteur de régression est asymptotiquement sans biais :

$$(12) \quad E[\hat{\mathbf{a}}_k] = \mathbf{a}_k$$

2. la variance tend vers la borne de Cramer-Rao :

$$(13) \quad \text{Var}[\hat{\mathbf{a}}_k] = \sigma^2/N\Gamma_k^{-1}$$

avec  $N$  la taille de l'échantillon  $\{x(n)\}$  et  $\Gamma_k$  la matrice de corrélation d'ordre  $K$  du signal  $x(n)$ .

De ce fait, en posant

$$(14) \quad \begin{bmatrix} \hat{a}(1), \dots, \hat{a}(p_0), \dots, \hat{a}(K) \end{bmatrix}^T = \begin{bmatrix} a(1), \dots, a(p_0), 0, \dots, 0 \end{bmatrix}^T + \begin{bmatrix} \varepsilon(1), \dots, \varepsilon(K) \end{bmatrix}^T,$$

avec

$$(15) \quad E[\varepsilon] = 0,$$

nous constatons que si nous pouvons définir une transformation sur le «signal  $\hat{\mathbf{a}}_k$ » permettant de modéliser l'erreur d'estimation  $\varepsilon$  comme un bruit blanc additif de variance connue, alors nous serons ramenés au modèle statistique (1).

La matrice de corrélation  $\Gamma_k$  étant définie positive, nous pouvons la mettre sous la forme :

$$(16) \quad \Gamma_k = \mathbf{C}\mathbf{C}^+,$$

où  $\mathbf{C}$  est une matrice triangulaire inférieure (décomposition de Cholesky). Ainsi, nous avons en particulier l'égalité

$$(17) \quad \mathbf{C}\Gamma_k^{-1}\mathbf{C}^+ = \mathbf{1},$$

dont on déduit la transformation qui nous permet de nous ramener à un modèle à bruit additif de variance unité. En effet, si  $\hat{\alpha}_k$  désigne le vecteur résultant de la transformation de  $\hat{\mathbf{a}}_k$  :

$$(18) \quad \hat{\alpha}_k = (N^{1/2}/\sigma)\mathbf{C}^+\hat{\mathbf{a}}_k,$$

(expression dans laquelle  $\sigma$  et  $\mathbf{C}$  seront pratiquement des estimées), on peut facilement vérifier que

$$(19) \quad \text{Var}[\hat{\alpha}_k] = \mathbf{1}$$

Il est important de constater que la transformation (18) «conserve l'ordre», car la matrice  $\mathbf{C}^+$  est triangulaire supérieure, et de ce fait la transformation n'augmente pas la dimension de l'espace engendré par  $\hat{\mathbf{a}}_k$ . De plus, la transformation est inversible :

$$(20) \quad \hat{\mathbf{a}}_k = (\sigma/N^{1/2})\mathbf{C}^-\hat{\alpha}_k$$

et possède les mêmes propriétés (conservation de l'ordre). Ainsi, le vecteur  $\hat{\alpha}_k$  obtenu après transformation dans la base  $\{v_1, v_2, \dots, v_k\}$  peut s'écrire :

$$(21) \quad \hat{\mathbf{a}}_k = \delta + \beta \text{ où } \hat{\mathbf{a}}_k = \hat{a}(1)v_1 + \dots + \hat{a}(p_0)v_{p_0} + \dots + \hat{a}(K)v_K$$

Le vecteur  $\delta$  est de même dimension que  $\mathbf{a}$ , et :

$$(22) \quad \begin{cases} E[\hat{\alpha}_k] = \delta, \\ E[\beta] = 0, \\ E[\beta\beta^+] = \mathbf{1}. \end{cases}$$

Nous sommes ainsi ramenés au modèle statistique défini par (1).

L'ordre optimal maximisant le critère CBV qui d'après (9) s'écrit

$$(23) \quad \hat{Q}'(p) = \sum_{i=1}^p (|a_i|^2 - 2)$$

sera, dans le cas d'un processus autorégressif, l'ordre exact  $p_0$ . Le critère que nous proposons permet donc de déterminer l'ordre à adopter en modélisation autorégressive afin d'optimiser le compromis biais-variance sur le développement paramétrique. Comme nous le verrons, ce critère s'applique aussi bien aux AR «vrais»

qu'aux signaux non autorégressifs modélisables par un AR long, pourvu que la première estimation soit réalisée à un ordre assez grand.

La procédure à adopter est la suivante :

1. Calcul du vecteur  $\hat{a}_k$  pour un ordre K élevé (modélisation a priori par un AR long) et estimation de  $\sigma$  par  $\hat{\sigma}_k$
2. factorisation de Cholesky :  $\Gamma_k = CC^+$ ,
3. calcul du vecteur transformé :  $\hat{\alpha}_k = (N^{1/2}/\sigma)C^+ \hat{a}_k$ ,
4. sélection de l'ordre correspondant au maximum de  $\hat{Q}'(p)$ 

$$p_0 = \sup \left\{ \max_i \left( \hat{Q}'(p) = \sum_{i=1}^K |a_{ij}|^2 - 2 \right) \right\},$$
5. calcul du vecteur paramètre à l'ordre optimal.

Notons que pratiquement, l'ensemble des vecteurs paramètres (jusqu'à l'ordre K) pourra être déterminé par l'algorithme de Levinson, qui fournit de plus la factorisation de Cholesky de l'inverse de  $\Gamma_k$ . Il reste donc à calculer l'inverse d'une matrice triangulaire, à effectuer les étapes 3 et 4, et à sélectionner le vecteur d'ordre  $p_0$ .

Pour simplifier l'exposé, nous avons supposé implicitement que les valeurs des paramètres du vecteur de régression étaient concentrées près de l'origine. En réalité, on peut envisager d'obtenir plusieurs intervalles distincts pour lesquels  $|\delta(i)|^2 > 1$ . On devra alors définir plusieurs «fenêtres» correspondant à ces intervalles. Le vecteur  $\hat{\alpha}_k$  est par exemple de la forme :

$$(24) \quad \hat{\alpha}_k = [\hat{a}(1), \dots, \hat{a}(i), 0, \dots, 0, \hat{a}(j+1), \dots, \hat{a}(j+1), 0, \dots, 0].$$

Compte tenu de la transformation inverse, l'ordre effectif du processus est alors égal à l'indice p du dernier maximum retenu par le critère  $\hat{Q}'(p)$ . Nous devons noter que pratiquement, pour éviter de sélectionner un maximum secondaire lié à des valeurs anormalement élevées du bruit, il faut introduire des tests statistiques (voir [7]).

#### IV. EXTENSION AUX PROCESSUS ARMA

Si le signal que l'on cherche à modéliser peut être décrit par un ARMA(p,q) à phase minimale, nous pouvons le caractériser par un modèle AR infini (filtre stable causal et causalement inversible) :

$$(25) \quad x(n) = \sum_{m=1}^{+\infty} c(m)x(n-m) + u(n)$$

où  $u(n)$  représenterait le bruit générateur (l'innovation). En écrivant l'égalité des transformées en z, on obtient les relations suivantes entre les coefficients  $\{c(r)\}$  de l'AR long et les coefficients AR  $\{a(r)\}$  et MA  $\{b(r)\}$  du processus ARMA :

$$(26) \quad c(r) = -\sum_{i=1}^q b(i)c(r-i) + a(r) \quad \text{pour } 0 \leq r \leq p,$$

$$c(r) = -\sum_{i=1}^q b(i)c(r-i) \quad \text{pour } r > p.$$

Nous pouvons donc interpréter les coefficients  $c(r)$  comme la sortie d'un filtre autorégressif d'ordre q et de coefficients  $b(i)$  excité par la séquence  $\{a(r)\}$  de longueur p. Les  $\{b(i)\}$  sont les coefficients de la partie MA dont on a supposé la stabilité. Par suite, pour  $r > p$ , les  $c(r)$  constituent la réponse transitoire, évidemment décroissante, de la partie MA. Nous pouvons donc finalement approximer un ARMA(p,q) par un AR( $r_0$ ) en choisissant  $r_0$  tel que

$$(27) \quad c(r) \equiv 0 \text{ pour } r > r_0 > p \text{ avec } c(r+1) < c(r) \text{ pour } r > p.$$

Ainsi, il suffit de déterminer l'ordre qui permet d'optimiser le compromis biais-variance sur  $c = [c(1), \dots, c(p), \dots, c(r)]^T$ . Cette recherche s'effectue de la même transformation que (18) et en maximisant le critère (23) qui s'en déduira pour obtenir la valeur de  $r_0$ . Nous pouvons noter qu'une telle approche peut s'appliquer directement à un AR bruité par un bruit d'observation blanc : un tel processus étant caractérisé par un ARMA(p,p) [5].

#### V. GENERALISATION A UN PROCESSUS AR MULTIVARIABLE

Dans le cas d'un processus AR multivariable, le choix de l'ordre est au moins aussi important que dans le cas scalaire. On conçoit facilement que les ordres optimaux pour l'estimation des différentes composantes n'ont aucune raison d'être identiques entre eux. la meilleure estimation serait obtenue en utilisant pour chaque composante l'ordre optimal correspondant. La formulation multivariable des équations de Yule-Walker ne permet pas d'imposer un ordre différent pour l'estimation des modèles relatifs aux diverses composantes : une telle approche a toutefois été expérimentée, mais elle conduit à une perte de la symétrie hermitienne de la matrice spectrale qui peut en être déduite. nous sommes donc pour le moment «condamnés» à une approche globale. Marple rapporte [2] l'analogue du critère AIC dans le cas multivariable, mais il souligne cependant que ce critère n'est qu'approximatif. Nous allons généraliser le critère introduit précédemment en nous ramenant ici encore à un modèle statistique à bruit blanc additif.

Le signal vectoriel de dimension Q, supposé centré et stationnaire, est un processus vectoriel autorégressif s'il répond à un modèle du type

$$(28) \quad x(n) = \sum A_i x(n-i) + u(n),$$

où  $x$  et  $u$  sont des vecteurs de dimension Q et les  $A_i$  des matrices de dimension  $Q \times Q$ .

Une composante quelconque  $x_j$  s'exprime comme :

$$(29) \quad x_j(n) = -\sum_{i=1}^{p_{j1}} a_{j1} x_{j1}(n-i) - \dots - \sum_{i=1}^{p_{jQ}} a_{jQ} x_{jQ}(n-i) + u_j(n)$$

Soit  $p_j$  le max de  $\{p_{j1}, \dots, p_{jQ}\}$ , nous pouvons écrire (29) sous la forme :

$$(30) \quad x_1(n) = -\sum_{i=1}^{p_1} a_j^T x(n-i) + u_j(n) \quad \text{avec} \quad a_j(i) = [a_{j1}(i), \dots, a_{jQ}(i)]^T,$$

$$x(n-i) = [x_1(n-i), \dots, x_Q(n-i)]^T.$$

certaines des composantes du vecteur  $a_j(i)$  pouvant être nulles suivant les valeurs de i. L'ensemble des vecteurs  $a_j(i)$  forme un vecteur  $a_j$  de dimension  $p_j Q$  :

$$(31) \quad a_j = [a_j(1), \dots, a_j(p_j)]^T.$$

Comme dans la troisième partie, le vecteur de régression, ici caractérisé par  $a_j$ , sera estimé par un vecteur  $\hat{a}_j$  de dimension  $KQ$  avec  $K > p_j$  :

$$(32) \quad \hat{a}_j = [\hat{a}_j(1), \dots, \hat{a}_j(K)]^T.$$

On peut montrer que ce vecteur atteint la borne de Cramer-Rao :

$$(33) \quad \text{Var} [\hat{a}_j] = \sigma_u^2 N \Gamma_{KQ}^{-1}$$

où la matrice  $\Gamma_{KQ}$  est une matrice de Toeplitz dont chacun des éléments est une matrice de dimension  $Q \times Q$ .





Ainsi, la transformation

$$(34) \quad \hat{\alpha}_j = (N^{1/2}/\sigma_{uj})C^+ \hat{a}_j \text{ avec } \Gamma_{KQ} = CC^+$$

nous ramène à un modèle statistique à bruit blanc additif de variance connue, en supposant que  $\Gamma_{KQ}$  et  $\sigma_{uj}$  soient connus (pratiquement ils seront estimés), car :

$$(35) \quad \text{Var}[\hat{\alpha}_j] = 1.$$

Nous sommes donc ramenés au cas envisagé dans la troisième partie et dès lors nous pouvons déterminer l'ordre optimal sur chacune des composantes de  $x_Q(n)$ . La détermination des paramètres AR étant effectuée à un ordre unique, nous pourrions rechercher l'ordre optimal sur le vecteur :

$$(36) \quad \hat{\alpha}_j = \sum_{i=1}^Q \hat{a}_i \text{ tel que } \text{Var}[\hat{a}] = 1.$$

notons que ceci revient à sélectionner l'ordre

$$(37) \quad p_0 = \max_j \{ \text{ordre de } \hat{a}_j \}.$$

Nous disposons donc d'un critère global permettant de déterminer l'ordre optimal à adopter en estimation autorégressive multivariable afin d'optimiser le compromis biais-variance sur les paramètres du modèle. Ainsi, l'estimée du vecteur autorégressif multivariable peut se mettre sous la forme globale :

$$(38) \quad x(n) = -\sum_{i=1}^{p_0} A_i x(n-i) + u(n),$$

où chacune des lignes de rang  $j$  correspond au vecteur  $a_j(i)$  précédemment défini.

## VI. CONCLUSIONS

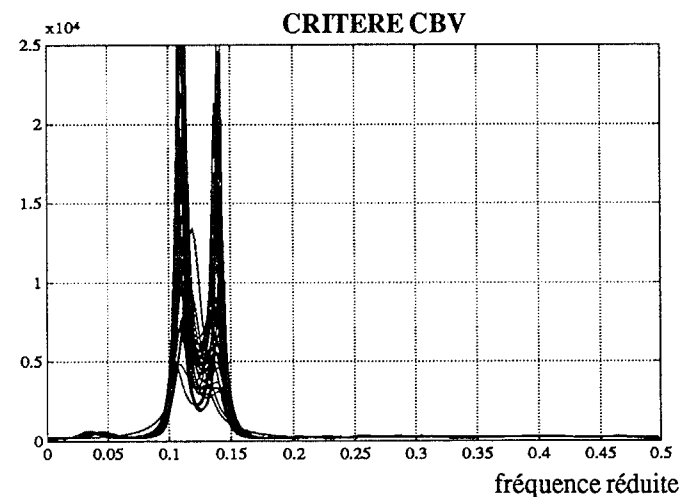
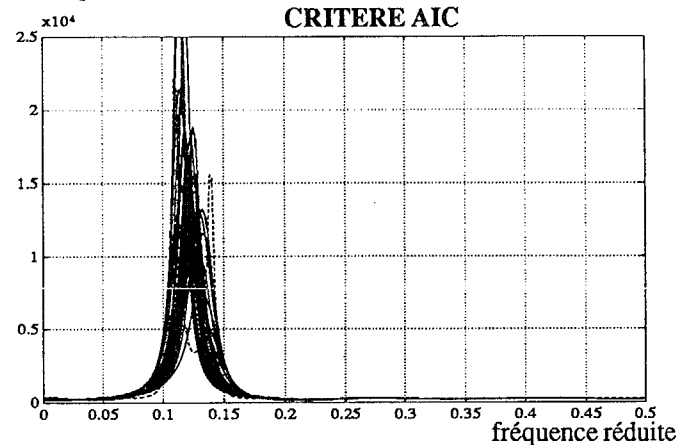
Le critère CBV proposé a été, en particulier, comparé dans le cas monodimensionnel aux critères AIC et MDL aussi bien sur des AR purs que sur des signaux non autorégressifs. Les résultats obtenus en fonction du nombre de points ont eux aussi été étudiés. Il est clair que ce type d'étude ne peut être exhaustif, puisque la qualité de la sélection de l'ordre dépend de nombreux facteurs : le nombre de points, l'emplacement des pôles dans le plan complexe, les effets de compensation entre pôles et zéros... Quoiqu'il en soit, les résultats obtenus peuvent se résumer comme indiqué ci-après.

Pour des signaux purement autorégressifs, l'ordre sélectionné est toujours très proche de l'ordre exact (même pour un nombre de points faible). Les performances et la tenue statistique ont été évaluées pour différents modèles par la Méthode de Monte Carlo. Nous pouvons alors établir la hiérarchie suivante suivant la tendance qu'ont ces critères à déterminer l'ordre le plus élevé : AIC > CBV > MDL, l'erreur sur l'ordre étant dans tous les cas très faible.

Pour des signaux autorégressifs bruités ou des ARMA, les critères classiques AIC et MDL sont très nettement mis en défaut. la hiérarchie est ici la suivante : CBV > AIC > MDL. On peut juger ici de la qualité des critères au regard de l'estimation des densités spectrales de puissance pour les ordres sélectionnés. Le critère CBV donne de (très?) bons résultats pour des AR bruités jusqu'à un rapport signal à bruit de +5 dB environ. Nous rapportons ci-après la comparaison de l'identification de la densité spectrale de puissance d'un AR(4) bruité à 10 dB présentant deux pics rapprochés, lorsque l'ordre du modèle est estimé par AIC et CBV (512 points, 30 réalisations). L'ordre détecté par AIC est en moyenne de 8 et est insuffisant pour faire ressortir les pics. l'ordre donné par CBV est en moyenne de 16, alors que les deux pics semblent se séparer pour un ordre de 14 en moyenne.

Les résultats que nous avons obtenus sont très satisfaisants dans la mesure où l'on peut maintenant envisager la modélisation de signaux ARMA ou AR bruités par la technique de l'AR long, sans

recourir à des essais successifs et à des choix toujours délicats. D'autre part, les résultats obtenus pour des modèles purement autorégressifs où l'on sait que les critères classiques sont très efficaces sont aussi très intéressants. Nous pouvons citer ici la réflexion de Dacunha-Castelle et Azencott [8] à propos du critère AIC : «Un tel choix, préconisé par Akaike, n'est pas satisfaisant car il conduit à une probabilité de surparamétrisation positive stricte». Il n'est donc pas anormal de sélectionner un ordre légèrement plus faible à celui de AIC lorsque le nombre de paramètres est grand.



## BIBLIOGRAPHIE

- [1] MAKHOUL, J., «Linear prediction : a tutorial review», Proc. IEEE, Vol. 63, pp. 561-580, 1975.
- [2] MARPLE, S.L., «Digital Spectral Analysis with Applications», Prentice Hall Signal Processing Series, New Jersey, 1987.
- [3] TUFTS, D.W., KUMARESAN, R., «Estimation of frequencies of Multiple Sinusoids : Making Linear Prediction Perform Like Maximum Likelihood», Proc. IEEE, Vol. 70, n°9, Sept. 82, pp. 975-989.
- [4] AKAIKE, H., «A new look at the statistical model identification», IEEE Transactions on Automatic Control, Vol. 19, pp. 716-723, 1974.
- [5] KAY, S.M., «Modern Spectral Estimation : Theory and Application», Prentice Hall Signal Processing Series, New Jersey, 1988.
- [6] LACOUPE J.L., «Nouvelles méthodes d'analyse spectrale », dans MAX J. et Col., «Méthodes et techniques de Traitement du Signal et Applications aux mesures physiques», chap. 15, Masson, Paris, 1987.
- [7] LUMEAU, B. et CLERGEOT, H., «Méthode d'estimation de la matrice de densité spectrale : optimisation du compromis biais-variance», Traitement du Signal, Vol. 7, pp 87-99, 1990.
- [8] DACUNHA-CASTELLE, D. et AZENCOTT, R., «Séries d'observations irrégulières : modélisation et prévision», Collection Techniques Stochastiques, Masson, 1984, pp. 175.