QUATORZIEME COLLOQUE GRETSI - JUAN-LES-PINS - DU 13 AU 16 SEPTEMBRE 1993

SUR L'ESTIMATION AUTOREGRESSIVE DANS LE CAS MULTIVARIE

Serge DEGERINE

Laboratoire de Modélisation et Calcul, C.N.R.S., Université Joseph Fourier, B.P. 53, 38041 Grenoble Cedex 9.

RÉSUMÉ

Nous reprenons brièvement la définition du coefficient d'autocorrélation partielle, ou coefficient de réflexion, dans le cas multivarié pour considérer ensuite les méthodes d'estimation autorégressive directement liées à ce paramètre. Nous introduisons un coefficient matriciel qui mesure la liaison linéaire symétrique entre deux vecteurs aléatoires et dont les valeurs singulières sont inférieures ou égales à un. Cela permet de définir un cadre général en trois étapes pour l'estimation du coefficient de réflection. La stabilité du filtre autorégressif associé est alors garantie. Nous retrouvons ainsi certaines généralisations de la méthode de Burg et proposons de nouvelles approches.

ABSTRACT

The definition of the partial autocorrelation coefficient, or reflection coefficient, in the multivariate case is briefly revisited in order to investigate the autoregressive estimation methods based on this parameter. A matrix coefficient, which measures the symmetric linear link between two random vectors, is introduced and its singular values are less than or equal to one. This gives rise to a general three steps procedure for the estimation of the reflection coefficient. Then the associated autoregressive filter is guaranteed to be stable. Some generalizations of the Burg's technique fit in this framework and new methods are given.

1. INTRODUCTION ET NOTATIONS

La notion de coefficient d'autocorrélation partielle, ou coefficient de réflexion, est au cœur de bien des méthodes d'estimation autorégressive. Pour une série scalaire on peut citer la célèbre méthode de Burg [1] ou celle de Dickinson [2] dans le cadre des critères de moindres carrés, mais ces coefficients sont aussi à l'origine des méthodes de maximum de vraisemblance approché (Kay, [3]) ou exact (Pham, [4]). Nous avons également proposé une estimation empirique naturelle des autocorrélations partielles (Dégerine, [5]).

L'extension de ces méthodes au cas multivarié présente de nombreuses difficultés. Ainsi plusieurs généralisations de la méthode de Burg ont été proposées (Nutall, [6], Strand, [7]) mais seule celle de Morf et coll. [8] est directement basée sur la notion d'autocorrélation partielle. C'est aussi le cas de l'extension de la méthode de Dickinson [9] et ces deux méthodes utilisent l'autocorrélation partielle liée à la décomposition de Cholesky des matrices de covariance.

En effet l'une des difficultés essentielles dans la définition d'une corrélation (partielle ou non) pour le cas multivarié est le choix de la racine carrée d'une matrice symétrique positive dans la phase de normalisation. Ceci nous a conduit à proposer la notion d'autocorrélation partielle canonique (Dégerine, [10]).

Soit $X(.) = \{X(t),\ t\in \mathbb{Z}\}$ une série vectorielle réelle de dimension m, stationnaire au second ordre et centrée. La fonction d'autocovariance

$$R(n) = E\{X(t+n)X(t)^T\},\quad n\in \mathbb{Z}\ ,$$

caractérise la structure au second ordre de $X(.)$ mais doit satisfaire les contraintes de positivité. Elle peut être remplacée par la fonction d'autocovariance partielle, $\delta(0) = R(0)$ et

$$\delta(n) = E\{\varepsilon(t;n-1)\varepsilon^*(t-n;n-1)^T\},\quad n\in \mathbb{N}^*,$$

où $\varepsilon(t;n)$ et $\varepsilon^*(t;n)$ désignent les erreurs de prédiction progressive et rétrograde d'ordre n ($\varepsilon(t;0) = \varepsilon^*(t;0) = X(t)$). Cependant $\delta(.)$ doit à son tour satisfaire les conditions garantissant la positivité des matrices de covariance des erreurs d'ordre n, $n\in \mathbb{N}^*$, construites selon la récurrence,

$$\sigma^2(n) = \sigma^2(n-1) - \delta(n)\sigma^{*2}(n-1)^-\delta(n)^T,$$
$$\sigma^{*2}(n) = \sigma^{*2}(n-1) - \delta(n)^T\sigma^2(n-1)^-\delta(n),$$

où A^- représente un inverse généralisé quelconque de A.

La fonction d'autocorrélation partielle, $\beta(0) = \delta(0)$ et

$$\beta(n) = \sigma(n-1)^-\delta(n)\sigma^*(n-1)^{-T},\quad n\in \mathbb{N}^*, \tag{1}$$

dans laquelle σ est une racine carrée de σ^2 ($\sigma\sigma^T = \sigma^2$) et $\sigma^{-T} = [\sigma^-]^T$, permet d'éliminer ce genre de contraintes.

Cette notion est précisée dans le second paragraphe. Le troisième paragraphe est consacré à l'extension au cas vectoriel de l'estimation empirique des autocorrélations partielles. Les démonstrations des résultats présentés ici ainsi que des compléments figurent dans Dégerine [11].

2. AUTOCORRELATIONS PARTIELLES

2.1. Procédés d'orthogonalisation

Un *procédé d'orthogonalisation standard* désignera ici tout moyen qui permet d'associer de façon unique, à une matrice de covariance σ^2, une racine carrée σ ($\sigma\sigma^T = \sigma^2$) et un inverse généralisé σ^- de σ ($\sigma\sigma^-\sigma = \sigma$) satisfaisant $\sigma^-\sigma = I_r$ (matrice identité d'ordre r) où r est le rang de σ^2.

a) *Orthogonalisation triangulaire.* Soit TD^2T^T la décomposition de Cholesky de σ^2 dans laquelle D^2 est une matrice carrée diagonale à éléments positifs ou nuls et T une matrice triangulaire inférieure unitaire dont l'unicité est obtenue en annulant les termes non diagonaux des colonnes de T correspondant aux colonnes nulles de TD^2. On note $\sqrt{D^2}$ la matrice carrée diagonale définie par la racine carrée des termes de D^2 et $\sqrt{D^2}^+$ celle définie par l'inversion des termes non nuls de $\sqrt{D^2}$. Alors σ est la matrice $m{\times}r$ obtenue en supprimant les colonnes nulles de $T\sqrt{D^2}$ et σ^- la matrice $r{\times}m$ obtenue en supprimant les lignes nulles de $\sqrt{D^2}^+T^{-1}$. Le procédé peut être utilisé à partir d'une décomposition de Cholesky basée sur les matrices triangulaires supérieures.

b) *Orthogonalisation selon les composantes principales.* Soit $V\Delta^2V^T$ une décomposition spectrale de σ^2 dans laquelle Δ^2 est la matrice carrée diagonale formée par les valeurs propres de σ^2, avec leur multiplicité et rangées par ordre décroissant. Les vecteurs propres associés, colonnes de V, sont rendus uniques en imposant que la première composante non nulle soit positive. Lorsqu'une valeur propre est multiple, le premier vecteur propre associé est celui pour lequel le nombre de composantes nulles successives, à partir de la dernière, est maximum; les suivants sont construits selon le même principe tout en conservant les conditions d'orthogonalité. Alors σ et σ^- sont obtenus comme ci-dessus à partir de $V\sqrt{\Delta^2}$ et $\sqrt{\Delta^2}^+V^T$.

2.2. Fonctions d'autocorrélation partielle standards

Une *fonction d'autocorrélation partielle standard* $\beta(.)$ est associée par les relations (1) à la fonction d'autocovariance $\delta(.)$ à partir de n'importe quel procédé d'orthogonalisation. Le domaine de variation de $\beta(.)$ se présente ainsi: la matrice $\beta(0)$ est symétrique positive d'ordre m; la suite $\{\beta(n), n = 1, 2, ...\}$ est constituée de matrices carrées dont les valeurs singulières sont inférieures où égales à un et telle que l'ordre de $\beta(n+1)$ soit égal au rang de $I-\beta(n)\beta(n)^T$, celui de $\beta(1)$ étant égal au rang de $\beta(0)$.

a) *Forme directe.* On appelle ainsi toute définition de $\beta(.)$ qui résulte d'un procédé dans lequel σ et σ^- sont définis uniquement en fonction de σ^2.

b) *Forme récursive.* Un procédé, dit de base, étant fixé on construit les matrices par la récurrence, $n \in \mathbb{N}^*$,

$$\begin{aligned} \sigma(n) &= \sigma(n-1)[I - \beta(n)\beta(n)^T]^{1/2} \\ \sigma(n)^- &= [I - \beta(n)\beta(n)^T]^{-1/2}\sigma(n-1)^- \end{aligned} \qquad (2)$$

à partir de valeurs initiales $\sigma(0)$ et $\sigma(0)^-$, où $[.]^{1/2}$ et $[.]^{-1/2}$ sont associés à $[.]$ selon le procédé de base retenu. Le même principe est adopté dans le sens rétrograde, éventuellement à l'aide d'un autre procédé de base.

Parmi les définitions possibles de $\beta(.)$, celles basées sur les décompositions triangulaires présentent de nombreux avantages:

– Les matrices σ et σ^- résultent d'un calcul algébrique simple (non itératif).

– Les formes directes et récursives coïncident.

Par la suite nous identifierons par l'indice t (triangulaire) les éléments $\beta(.)_t$, $\sigma(.)_t$, etc, associés au procédé utilisant la forme triangulaire supérieure dans le sens progressif et la forme triangulaire inférieure dans le sens rétrograde.

2.3. Fonctions d'autocorrélation partielle canoniques

La fonction d'autocorrélation partielle canonique définie dans [10] correspond à la forme standard obtenue par le procédé récursif ayant pour base le procédé b). Nous utiliserons l'indice cr (canonique récursive) pour identifier les éléments de cette version. Le terme canonique signifie que les composantes de

$$\eta(t;n)_{cr} = \sigma(n)_{cr}^-\varepsilon(t;n),$$
$$\eta^*(t-n;n)_{cr} = \sigma^*(n)_{cr}^-\varepsilon^*(t-n;n),$$

sont les variables canoniques dans l'analyse canonique du couple $\{\varepsilon(t;n), \varepsilon^*(t-n;n)\}$. L'unicité obtenue de façon récursive semble cependant peu adaptée aux problèmes d'estimation.

Soit $\beta = L^T\Delta\Theta M = L^T\Theta\Delta M$, la décomposition en valeurs singulières d'une matrice réelle carrée, dont les valeurs singulières sont inférieures ou égales à un, réalisée de façon unique comme suit. Les matrices orthogonales L et M sont celles des décompositions spectrales $I-\beta\beta^T = L^T(I-\Delta^2)L$ et $I-\beta^T\beta = M^T(I-\Delta^2)M$, dont l'unicité est obtenue selon les conventions du procédé b). La matrice carrée diagonale Δ contient les valeurs singulières de β, avec leur multiplicité et rangées par ordre croissant. La matrice Θ est définie par $L\beta M^T = \Delta\Theta$ en convenant de poser égaux à zéro les éléments des lignes et colonnes correspondant aux valeurs nulles de Δ. D'autre part les matrices diagonales rectangulaires $[I - \Delta^2]^{1/2}$ et $[I - \Delta^2]^{-1/2}$ sont associées à $(I - \Delta^2)$ selon le procédé b) de même que Δ_0 et Δ_0^- à partir de $\Delta^2(0)$ dans le cas particulier de $\sigma^2(0) = L(0)^T\Delta^2(0)L(0)$.

Proposition 1.– *Soit $\beta(.)$ une fonction d'autocorrélation partielle standard quelconque. A partir des décompositions $\beta(n) = L(n)^T\Delta_n\Theta(n)M(n)$, $n \geq 1$, on définit les nouvelles variables $\eta(t;n)_c$ et $\eta^*(t-n;n)_c$ par*

$$[I-\Delta^2(n)]^{-1/2}L(n)\{\eta(t;n-1) - \beta(n)\eta^*(t-n;n-1)\},$$
$$[I-\Delta^2(n)]^{-1/2}M(n)\{\eta^*(t-n;n-1) - \beta^T(n)\eta(t;n-1)\},$$

où $\eta(t;n-1)$ et $\eta^*(t-n;n-1)$ sont les erreurs normalisées selon le procédé de définition de $\beta(.)$. Alors les variables ainsi définies sont normalisées et leur corrélation est égale à

$$-[I - \Delta^2(n)]^{-1/2}\Delta_n\Theta(n)[I - \Delta^2(n)]^{1/2}.$$

La matrice de corrélation ci-dessus est celle obtenue en supprimant dans $-\Delta_n\Theta(n)$ les lignes et colonnes correspondant aux éléments de Δ_n égaux à un. Cette proposition montre que les composantes de $\eta(t;n)_c$ et de $\eta^*(t-n;n)_c$ sont un choix de variables canoniques dans l'analyse canonique du couple $\{\varepsilon(t;n), \varepsilon^*(t-n;n)\}$. Il est alors possible de définir une fonction d'autocorrélation partielle canonique en posant $\beta(n+1)_c = E\{\eta(t;n)_c\eta^*(t-n-1;n)_c^T\}$, $n \geq 1$.

Cette forme de définition de l'autocorrélation partielle canonique permet d'échapper à la définition récursive initiale.

Elle dépend cependant du choix de la fonction standard $\beta(.)$ utilisée. Ici seule la simplicité doit motiver ce choix, nous utiliserons donc la fonction $\beta(.)_t$ et nous identifierons cette version par l'indice *ct* (canonique triangulaire). De façon précise on pose $\beta(0)_{ct} = \delta(0)$, $\beta(1)_{ct} = \Delta_0^{-} L(0)_{cr} \delta(1) L(0)_{cr}^T \Delta_0^{-T}$ et $\beta(n+1)_{ct} = P(n)\beta(n+1)_t Q(n)^T$, $n \geq 1$, où $P(n)$ et $Q(n)$ sont les matrices orthogonales,

$$P(n) = [I - \beta(n)_t \beta(n)_t^T]_{cp}^{-1/2} [I - \beta(n)_t \beta(n)_t^T]_{ts}^{1/2},$$
$$Q(n) = [I - \beta(n)_t^T \beta(n)_t]_{cp}^{-1/2} [I - \beta(n)_t^T \beta(n)_t]_{ti}^{1/2},$$

les indices *cp*, *ts* ou *ti* signifiant que l'on utilise le procédé *composantes principales*, *triangulaire supérieure* ou *inférieur*.

3. ESTIMATIONS EMPIRIQUES

3.1. Position du problème

Soit $x(1), \ldots, x(N)$ l'observation d'une séquence de longueur N de $X(.)$. Considérons, pour $n = 0, \ldots, p < N/2$, l'espace vectoriel engendré par les matrices d'ordre $m \times (N-n)$:

$$x_n(t) = [x(t), \ldots, x(t+N-n-1)], \quad t = 1, \ldots, n+1.$$

Posant $[x_n(t), x_n(t-k)] = \{x_n(t) x_n(t-k)^T\}/(N-n)$ on a $E\{[x_n(t), x_n(t-k)]\} = R(k)$ et le produit scalaire est défini par $\langle x_n(t), x_n(t-k)\rangle = tr\{[x_n(t), x_n(t-k)]\}$.

L'erreur de prédiction progressive d'ordre k

$$\varepsilon_n(t;k) = \sum_{j=0}^{k} a_n(k,j) x_n(t-j), \quad a_n(k,0) = I_m,$$

se caractérise par $[\varepsilon_n(t;k), x_n(t-j)] = 0$, $j = 1, \ldots, k$. Son homologue dans le sens rétrograde est notée $\varepsilon_n^*(t;k)$. Ainsi on définit sans ambiguïté les éléments empiriques suivants pour $n = 0, \ldots, p$ ($\varepsilon_n(t;-1) = \varepsilon_n^*(t;-1) = \varepsilon_n(t;0) = \varepsilon_n^*(t;0) = x_n(t)$),

$$\delta_e(n) = [\varepsilon_n(n+1;n-1), \varepsilon_n^*(1;n-1)],$$
$$\sigma_e^2(n) = [\varepsilon_n(n+1;n), \varepsilon_n(n+1;n)],$$
$$\sigma^{*2}_e(n) = [\varepsilon_n^*(1;n), \varepsilon_n^*(1;n)].$$

On utilisera également les variances résiduelles, $n = 1, \ldots, p$,

$$\tilde{\sigma}_e^2(n-1) = [\varepsilon_n(n+1;n-1), \varepsilon_n(n+1;n-1)],$$
$$\tilde{\sigma}^{*2}_e(n-1) = [\varepsilon_n^*(1;n-1), \varepsilon_n^*(1;n-1)].$$

Proposition 2.– *Pour $n = 1, \ldots, p$, les erreurs de prédiction empiriques satisfont:*

$$\varepsilon_n(n+1;n) = \varepsilon_n(n+1;n-1) - \delta_e(n) \tilde{\sigma}^{*2}_e(n-1)^{-} \varepsilon_n^*(1;n-1),$$
$$\varepsilon_n^*(1;n) = \varepsilon_n^*(1;n-1) - \delta_e(n)^T \tilde{\sigma}_e^2(n-1)^{-} \varepsilon_n(n+1;n-1).$$

Les relations entre les variances résiduelles sont:

$$\sigma_e^2(n) = \tilde{\sigma}_e^2(n-1) - \delta_e(n) \tilde{\sigma}^{*2}_e(n-1)^{-} \delta_e(n)^T,$$
$$\sigma^{*2}_e(n) = \tilde{\sigma}^{*2}_e(n-1) - \delta_e(n)^T \tilde{\sigma}_e^2(n-1)^{-} \delta_e(n).$$

La suite des valeurs $\delta_e(n)$, $n = 0, \ldots, p$, ne constitue pas nécessairement une suite d'autocovariances partielles. La méthode de Dickinson [9] consiste à utiliser la séquence des autocovariances partielles définies pour $n = 0, \ldots, p$ par $[\varepsilon_p(n+1;n-1), \varepsilon_p^*(1;n-1)]$, c'est-à-dire qu'elles sont toutes obtenues à partir de $N-p$ séquences.

3.2. Coefficient de liaison linéaire

Nous introduisons dans cette section un coefficient matriciel mesurant la liaison linéaire symétrique existant entre deux vecteurs aléatoires réels X et Y, centrés, de dimensions respectives p et q, à partir du critère quadratique défini sur l'ensemble des matrices C de dimensions $p \times q$ par

$$Q(C;X,Y) = tr\{Var(X - CY)\} + tr\{Var(Y - C^T X)\}.$$

Proposition 3.– *Le critère $Q(C;X,Y)$ est convexe par rapport à C et son minimum est réalisé par toute matrice satisfaisant l'équation de Lyapunov*

$$Var(X)C + CVar(Y) = 2Cov(X,Y).$$

Celle-ci admet une seule solution satisfaisant la contrainte

$$u^T Var(X) u = 0 \ \text{ et } \ v^T Var(Y) v = 0 \ \Rightarrow \ u^T C v = 0.$$

Cette solution, notée C_{XY}, est appelée coefficient de liaison linéaire entre X et Y.

Le coefficient matriciel C_{XY} est égal à 0 si et seulement si les vecteurs X et Y sont non corrélés ($Cov(X,Y) = 0$). Il coïncide avec le coefficient de corrélation ρ_{XY} lorsque ces vecteurs sont préalablement normés. La valeur minimum du critère est égale à zéro si et seulement si les deux vecteurs X et Y se déduisent l'un de l'autre par une transformation orthogonale, $X = C_{XY} Y$ avec $C_{XY} C_{XY}^T = I$. Dans ce cas on a aussi $\rho_{XY} \rho_{XY}^T = I$ mais la réciproque n'est pas vraie. Cependant dans tous les cas les valeurs singulières de C_{XY} sont inférieures ou égales à un.

3.3. Corrélation partielle empirique

Il est naturel, au vu de la Proposition 2, d'estimer la suite des autocorrélations partielles par $\hat{\beta}(0) = \delta_e(0)$ et pour $n = 1, \ldots, p$,

$$\hat{\beta}(n) = \tilde{\sigma}_e(n-1)^{-} \delta_e(n) \tilde{\sigma}_e^*(n-1)^{-T}, \tag{3}$$

après s'être fixé un choix de définition des éléments σ et σ^{-}. Cela équivaut à $\hat{\beta}(n) = [\eta_n(n+1;n-1), \eta_n^*(1;n-1)]$, où $\eta_n(n+1;n-1) = \tilde{\sigma}_e(n-1)^{-} \varepsilon_n(n+1;n-1)$ et $\eta_n^*(1;n-1) = \tilde{\sigma}_e^*(n-1)^{-} \varepsilon_n^*(1;n-1)$ sont normalisées. Les valeurs singulières de $\hat{\beta}(n)$ sont donc inférieures ou égales à un.

On notera $\hat{\beta}(n)_t$ l'estimation obtenue selon (3) en utilisant la forme triangulaire supérieure dans le sens progressif et la forme triangulaire inférieure dans le sens rétrograde pour les définitions de σ et σ^{-}. Les valeurs singulières de $\hat{\beta}(n)$, $n = 1, \ldots, p$, sont un invariant des estimations obtenues selon (3).

Plus généralement on pose $\hat{\beta}(0) = \delta_e(0)$ et pour $n = 1, \ldots, p$, on effectue un choix au cours des trois étapes suivantes.

(i) Déterminer les coefficients $\bar{a}_n(n-1,k)$ et $\bar{a}_n^*(n-1,k)$, $k = 0, \ldots, n-1$, des filtres définissant les erreurs d'ordre $n-1$,

$$\bar{\varepsilon}_n(n+1;n-1) = \sum_{k=0}^{n-1} \bar{a}_n(n-1,k) x_n(n+1-k),$$
$$\bar{\varepsilon}_n^*(1;n-1) = \sum_{k=0}^{n-1} \bar{a}_n^*(n-1,k) x_n(1+k),$$

auxquelles sont alors associés les éléments empiriques

$$\bar{\sigma}_e^2(n-1) = [\bar{\varepsilon}_n(n+1;n-1), \bar{\varepsilon}_n(n+1;n-1)],$$
$$\bar{\sigma}_e^{*2}(n-1) = [\bar{\varepsilon}_n^*(1;n-1), \bar{\varepsilon}_n^*(1;n-1)],$$
$$\bar{\delta}_e(n) = [\bar{\varepsilon}_n(n+1;n-1), \bar{\varepsilon}_n^*(1;n-1)].$$

(ii) Choisir les estimateurs $\hat{\sigma}^2(n-1)$ et $\hat{\sigma}^{*2}(n-1)$.

(iii) Opter pour un procédé de normalisation pour définir $\hat{\sigma}(n-1)^{-}$ et $\hat{\sigma}^*(n-1)^{-}$.

L'estimation $\hat{\beta}(n)$ est alors le coefficient de liaison linéaire entre

$\widehat{\eta}_n = \widehat{\sigma}(n-1)^{-}\overline{\varepsilon}_n(n+1;n-1)$ et $\widehat{\eta}_n^* = \widehat{\sigma}^*(n-1)^{-}\overline{\varepsilon}_n^*(1;n-1)$. En d'autres termes $\widehat{\beta}(n)$ est solution de l'équation de Lyapunov (on omet les arguments $(n-1)$)

$$\widehat{\sigma}^{-}\overline{\sigma}_e^2\widehat{\sigma}^{-T}\beta(n) + \beta(n)\widehat{\sigma}^{*-}\overline{\sigma}_e^{*2}\widehat{\sigma}^{*-T} = 2\widehat{\sigma}^{-}\overline{\delta}_e(n)\widehat{\sigma}^{*-T}, \qquad (4)$$

associée au problème de minimisation par rapport à $\beta(n)$ de

$$\|\widehat{\eta}_n - \beta(n)\widehat{\eta}_n^*\|^2 + \|\widehat{\eta}_n^* - \beta(n)^T\widehat{\eta}_n\|^2.$$

Le principe des généralisations de la méthode de Burg consiste à utiliser, dans la phase (i), les filtres définis par les estimations $\widehat{\beta}(k)$, $k = 0, \ldots, n-1$, obtenues aux étapes précédentes. Dans la phase (ii) Nutall [6] et Strand [7] retiennent en plus les estimations $\widehat{\sigma}^2(n-1)$ et $\widehat{\sigma}^{*2}(n-1)$ qui en résultent et alors la méthode est indépendante de la phase (iii). Par contre Morf et coll. [7] utilisent, en phase (ii), les éléments empiriques $\overline{\sigma}_e^2(n-1)$ et $\overline{\sigma}_e^{*2}(n-1)$ comme estimateurs (ce qui évite l'équation de Lyapunov) et le procédé triangulaire inférieur dans les sens progressif et rétrograde pour la normalisation de la phase (iii).

Les méthodes visant à estimer les autocorrélations partielles canoniques peuvent aussi entrer dans le cadre général ci-dessus. Cependant la particularité du procédé de normalisation conduit à une présentation plus spécifique. Une estimation possible de $\beta(.)_{cr}$ consiste à utiliser, dans la phase (i), les erreurs empiriques $\varepsilon_n(n+1;n-1)$ et $\varepsilon_n^*(1;n-1)$ définies par le critère des moindres carrés. Mais les deux autres phases sont réalisées simultanément sur la base de la récurrence (2); par exemple dans le sens progressif $\widehat{\sigma}(n-1)_{cr}^{-}$ est défini en fonction de l'étape précédente par

$$\widehat{\sigma}(n-1)_{cr}^{-} = [I - \widehat{\Delta}^2(n-1)_{cr}]^{-1/2}\widehat{L}(n-1)_{cr}\widehat{\sigma}(n-2)_{cr}^{-}.$$

L'estimation $\widehat{\beta}(n)_{cr}$ est alors solution de l'équation de Lyapunov (4) correspondante. On aurait pu aussi utiliser les erreurs définies par les filtres associés aux étapes précédentes et obtenir ainsi une généralisation de la méthode de Burg basée sur les autocorrélations partielles canoniques. Les différentes contraintes introduites dans ces deux premières approches risquent d'affaiblir les qualités de résolution de la méthode; en particulier les valeurs singulières de $\widehat{\beta}(n)_{cr}$ diffèrent de l'invariant obtenu dans (3).

L'estimation empirique naturelle de $\beta(.)_{ct}$ se présente de la façon suivante. $\widehat{\beta}(0)_{ct} = \delta_e(0)$, $\widehat{\beta}(1)_{ct} = \widetilde{\sigma}_e(0)_{cp}^{-}\delta_e(1)\widetilde{\sigma}_e(0)_{cp}^{*-T}$, où $\widetilde{\sigma}_e(0)_{cp}^{-}$ et $\widetilde{\sigma}_e(0)_{cp}^{*-T}$ sont définis à partir de $\widetilde{\sigma}_e^2(0)$ et $\widetilde{\sigma}_e^{*2}(0)$ selon le procédé b). Pour $n \geq 2$ on utilise sur le plan empirique le principe de la Proposition 1 à partir du procédé triangulaire. Dans le sens progressif on introduit

$$\delta_n(n+1;2) = [\varepsilon_n(n+1;n-2), \varepsilon_n^*(2;n-2)],$$
$$\sigma_n^{*2}(2;n-2) = [\varepsilon_n^*(2;n-2), \varepsilon_n^*(2;n-2)],$$
$$\sigma_n^2(n+1;n-2) = [\varepsilon_n(n+1;n-2), \varepsilon_n(n+1;n-2)],$$

$$\beta_n(n+1;2)_t = \sigma_n(n+1;n-2)_{ts}^{-}\delta_n(n+1;2)\sigma_n^*(2;n-2)_{ti}^{-},$$

pour obtenir l'innovation canonique en fonction de l'innovation triangulaire $\eta_n(n+1;n-1)_{ct} = \widehat{P}_n(n-1)\eta_n(n+1;n-1)_t$ par la transformation orthogonale,

$$\widehat{P}_n(n-1) = [I - \beta\beta^T]_{cp}^{-1/2}[I - \beta\beta^T]_{ts}^{1/2}, \beta = \beta_n(n+1;2)_t.$$

Dans le sens rétrograde on utilise la corrélation partielle triangulaire $\beta = \beta_n(n;1)_t$ entre $\varepsilon_n^*(1;n-2)$ et $\varepsilon_n(n-1;n-2)$ pour aboutir à $\eta_n^*(1;n-1)_{ct} = \widehat{Q}_n(n-1)\eta_n^*(1;n-1)_t$ par la transformation orthogonale,

$$\widehat{Q}_n(n-1) = [I - \beta\beta^T]_{cp}^{-1/2}[I - \beta\beta^T]_{ti}^{1/2}, \beta = \beta_n(n;1)_t.$$

L'estimation $\widehat{\beta}(n)_{ct}$ de $\beta(n)_{ct}$ est alors donnée par

$$[\eta_n(n+1;n-1)_{ct}, \eta_n^*(1;n-1)_{ct}] = \widehat{P}_n(n-1)\widehat{\beta}(n)_t\widehat{Q}_n(n-1)^T.$$

Notons que les valeurs singulières de $\widehat{\beta}(n)_{ct}$ sont égales à celles de $\widehat{\beta}(n)_t$ et que l'on retrouve ainsi l'invariant des méthodes issues de (3). C'est en effet une méthode de même nature mais dont le procédé de normalisation passe par une estimation préalable.

La comparaison entre toutes ces méthodes devrait être menée à travers un important travail de simulation. Cependant le minimum de contraintes qu'utilisent les méthodes directes associées à (3), y compris $\widehat{\beta}(.)_{ct}$ ci-dessus, plaide en leur faveur. Ces dernières sont identiques dans le cas scalaire, puisque le procédé de normalisation ne se pose pas, et coïncident avec l'autocorrélation partielle empirique qui, dans ce cas, est préférable à la méthode de Burg (cf. [5]). D'autre part les valeurs singulières de module 1 sont estimées de façon presque sûre, pour N suffisamment grand, alors que ce n'est pas le cas dans les généralisations de la méthode de Burg.

REFERENCES

[1] BURG, J.P. (1968) A new analysis technique for time series data. Presented at NATO Advanced Study Institute on Signal Processing, Enschede, Netherlands.

[2] DICKINSON, B.W. (1978) Autoregressive estimation using residual energies ratios. IEEE Trans. on Information Theory, Vol. IT-24, n° 2, 503-506.

[3] KAY, S. M. (1983) Recursive maximum likelihood estimation of autoregressive processes. IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-31, n° 1, 56-65.

[4] PHAM, D.T. (1988) Maximum likelihood estimation of autoregressive model by relaxation on the reflection coefficients. IEEE, Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP 38, n° 1, 175-177.

[5] DEGERINE, S. (1993) Sample Partial Autocorrelation Function. IEEE, Trans. on Signal Processing, Vol. SP 41, n° 1, 403-407.

[6] NUTALL, A.H. (1976) Positive definite spectral estimate and stable correlation recursion for multivariate linear predictive spectral analysis. Naval Underwater System Center, NUSC Tech. Doc. 5729, New London, CT, Nov. 14.

[7] STRAND, O.N. (1977) Multichannel complex maximum entropy (autoregressive) analysis. IEEE Trans. Automat. Contr., vol. AC-22, 634-640.

[8] MORF, M., VIEIRA, A., LEE, D. and KAILATH, T. (1978) Recursive multichannel maximum entropy spectral estimation. IEEE Trans. on Geoscience Electronics, Vol. GE-16, n° 2, 85-94.

[9] DICKINSON, B.W. (1979) Estimation of partial correlation matrices using Cholesky decomposition. IEEE Trans. on Automatic Control, Vol. AC-24, n° 2, 302–305.

[10] DEGERINE, S. (1990) Canonical Partial Autocorrelation Function of a Multivariate Time Series. The Annals of Statistics, Vol. 18, N° 2, 961-971.

[11] DEGERINE, S. (1992) Autocorrélation partielle empirique d'une série chronologique vectorielle. RR n° 896-M- Lab. LMC, Univ. J. Fourier, Grenoble.