

ANNEXE 5. FONCTIONS SPHEROIDALES APLATIES.

Les fonctions sphéroïdales aplaties ont été étudiés pour leur relation avec la transformée de Fourier et les signaux à bande limitée dans une série de papiers par SLEPIAN, POLLAK, 1961, LANDAU, POLLAK, 1961, SLEPIAN, 1978, à temps continu d'abord, à temps discret dans le dernier papier. L'objet de cette annexe est un rappel de leur définition et leurs propriétés, et la présentation d'une méthode de calcul originale et rapide.

La définition des fonctions sphéroïdales aplaties à temps discret est donnée par SLEPIAN, 1978, à partir du concept de concentration sur un intervalle de temps $(0, T-1)$ de l'énergie d'un signal y_t supposé à bande limitée B , c'est-à-dire tel que $Y(\omega)=0$ dès que $|\omega| > 2\pi B$. Un indice de concentration est donné par $\lambda(B, T)$:

$$(2-215) \quad \lambda(B, T) = \left[\sum_{t=0}^{T-1} |y_t|^2 \right] \cdot \left[\sum_{t=-\infty}^{+\infty} |y_t|^2 \right]^{-1}$$

Parmi toutes les suites y_t qui sont à bande limitée, celle qui est la plus concentrée sur $(0, T-1)$ est celle qui maximise $\lambda(B, T)$. Il est montré par SLEPIAN, 1978, qu'elle vérifie une équation du type vecteur propre: pour tout t ,

$$(2-216) \quad \sum_{s=0}^{T-1} \frac{\sin 2\pi B(t-s)}{\pi(t-s)} f_{B, T}(s) = \lambda(B, T) f_{B, T}(t)$$

On pose dans (2-216) que pour $t-s=0$, la fraction vaut $2B$. Si on utilise cette équation pour $t \in (0, T-1)$, on obtient le vecteur de composantes $f(0) \dots f(T-1)$ comme symétrique. Cette remarque sera utilisée pour fournir un calcul rapide des fonctions f , qui sont au nombre de T , vue la non-

dégénérescence de la matrice construite sur le sinus cardinal de $2B(t-s)$. En négligeant dans l'écriture l'indice B, T , on obtient les fonctions $f_0(t) \dots f_{T-1}(t)$. Elles sont habituellement normalisées de façon que leur énergie sur $(0, T-1)$ soit 1, et que leur somme soit positive:

$$\sum_{t=0}^{T-1} (f_i(t))^2 = 1, \quad \sum_{t=0}^{T-1} f_i(t) \geq 0, \quad \sum_{t=0}^{T-1} (T-1-2t)f_i(t) \geq 0$$

Elles sont alors paires si j est pair et impaires si j est impair.

Elles sont de plus deux fois orthogonales, sur $(0, T-1)$ et sur $(-\infty, +\infty)$:

$$\sum_{t=0}^{T-1} f_i(t)f_j(t) = \lambda_i \delta_{i,j} \quad \sum_{t=-\infty}^{+\infty} f_i(t)f_j(t) = \delta_{i,j}$$

Une propriété très intéressante de ces fonctions est la notion de dimension d'un signal qu'elles induisent. En effet quand T est grand, l'espace des fonctions à bande limitée à B , qui sont concentrées sur l'intervalle $(0, T-1)$ a pour dimension approximative $2BT$. Ceci tient au fait que pour $T \rightarrow +\infty$, la limite de λ_i est 1 si $i = 2BT(1-\epsilon)$ et 0 si $i = 2BT(1+\epsilon)$, et ceci pour tout ϵ tel que $0 < \epsilon < 1$. Ainsi pour T grand, $2BT$ coefficients suffisent à décrire tout signal à bande limitée à B , et précisément ce sont les coefficients du développement de ce signal sur les fonctions sphéroïdales aplaties $f_0(t) \dots f_{2BT-1}(t)$.

L'emploi des fonctions sphéroïdales aplaties pour représenter les coefficients des modèles ARMA évolutifs se justifie à partir de la propriété précédente pour tout signal y_t dont les variations du modèle se font à bande limitée. Le produit $2BT$ fixera le nombre $m+1$ des fonctions nécessaires pour approximer correctement les variations limitées à la bande $(-B, +B)$ sur l'intervalle $(0, T-1)$. Dans la pratique, cela se décidera en sens inverse. On recherchera $m+1$ fonctions servant de base aux $a_i(t)$ et $b_i(t)$ du modèle.

La durée T étant connue, on choisira alors la largeur de bande B égale à $(m+1)/2T$, ce sera une mesure de la variation maximale avec t des $a_i(t)$. La matrice de Toeplitz R construite suivant (2-116) permet de calculer le vecteur x_i de composantes $f_i(t) \dots f_i(T-1)$ comme solution de $Rx_i = \lambda_i x_i$.

Le calcul de x_i présenté ici recherche le i -ème vecteur propre de R par la procédure décrite par GUEGUEN, GRENIER, GIANNELLA, 1980, pour le calcul de vecteurs propres de matrices de Toeplitz. Ceci s'oppose à la procédure habituelle qui détermine les fonctions sphéroïdales aplaties par une équation récurrente précédée de la diagonalisation d'une matrice tri-diagonale (SLEPIAN, 1978). Ici seuls nous intéressent les $m+1$ premiers vecteurs propres. On pourra donc utiliser un processus itératif faisant alterner le calcul du vecteur propre x_i par l'algorithme de Levinson sur $R - \lambda_i I$, et le calcul de la valeur propre λ_i par le coefficient de Rayleigh $x_i^T R x_i / (x_i^T x_i)^{-1}$ ou le coefficient modifié $x_i^T J R x_i / (x_i^T J x_i)$ où J est la matrice anti-diagonale qui renverse les colonnes ou les lignes d'une matrice. On améliorera la convergence de l'algorithme, qui est déjà rapide, en symétrisant les vecteurs x_i pairs ou en antisymétrisant les vecteurs x_i impairs, à chaque itération.

Pour les valeurs de T rencontrées dans la modélisation évolutive, on peut même faire plus vite en partant d'une valeur approchée de λ_k donnée par SLEPIAN, 1978, valable pour T grand et k faible:

$$(2-217) \quad \lambda_k = 1 - \pi^{1/2} (k!)^{-1} 2^{(14k+9)/4} \alpha^{(2k+1)/4} (2-\alpha)^{-(k+1)/2} T^{(k+1)/2} e^{-\gamma T}$$

avec $\alpha = 1 - \cos 2\pi B$

$$\gamma = \text{Log} \left(1 + \frac{2^{1/2}}{2^{1/2} - \alpha^{1/2}} \right)$$

Pour des valeurs de T comprises entre 400 et 1200, et des bases d'au

plus 10 fonctions, cette valeur approximative de λ_k était suffisamment bonne pour éviter les itérations, et la fonction $f_k(t)$ a pu être prise comme solution de $[R - \lambda_k I]x_k = 0$ avec comme seule correction la symétrisation des fonctions pour k pair, et l'anti-symétrisation de celles pour k impair. La figure 2.41 montre pour $m=3$ et $T=1200$ les fonctions calculées.

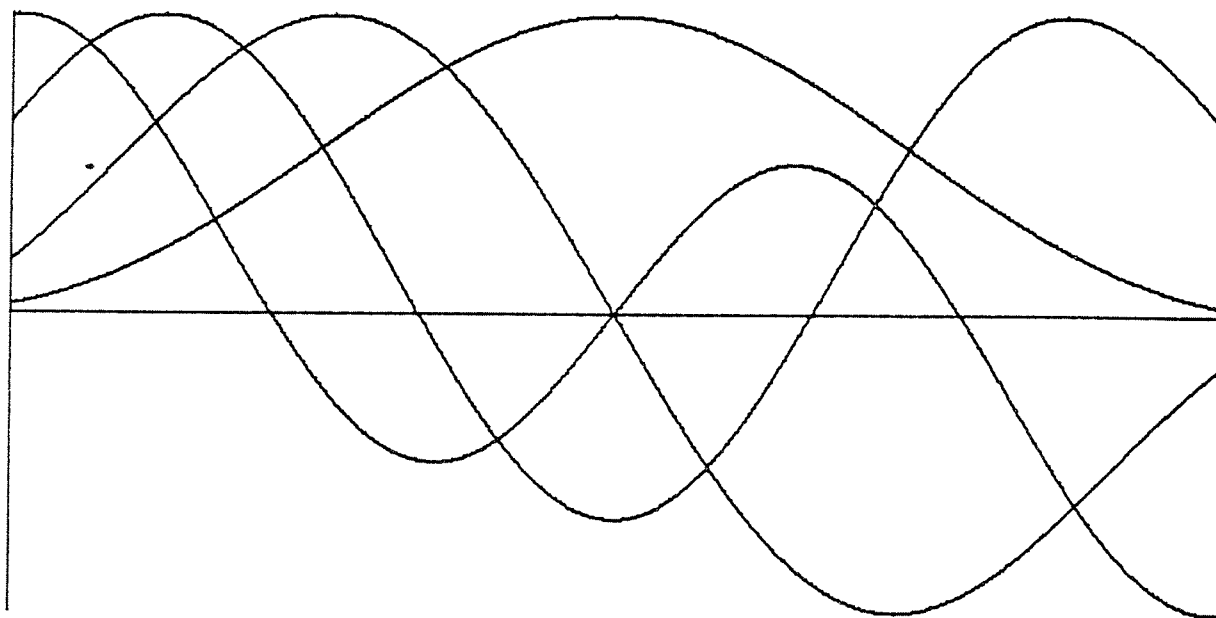
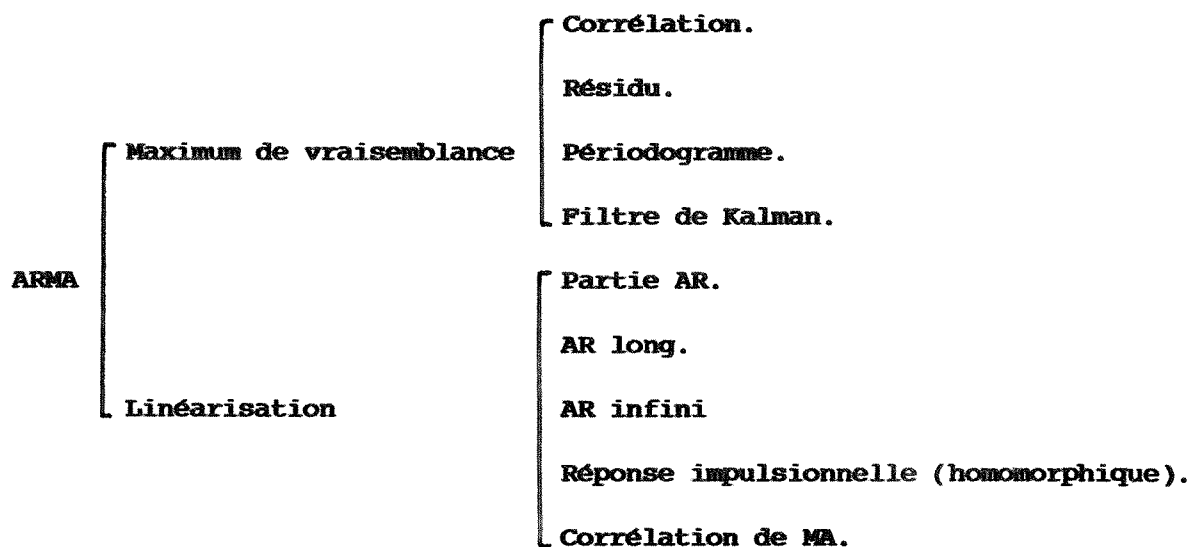


Figure 2.41. Quatre fonctions sphéroïdales aplaties pour $T=1200$.

ANNEXE 6. MODELE ARMA STATIONNAIRE.

Cette annexe est consacrée à une revue partielle des méthodes d'estimation de modèles ARMA stationnaires. La sélection réalisée parmi l'abondante littérature sur ce sujet est essentiellement subjective, et donc partielle. Les méthodes se répartissent en deux grands groupes, donnés dans le schéma ci-dessous.



1. Maximum de vraisemblance.

Toutes ces méthodes tentent de maximiser la vraisemblance soit $p(y_0 \dots y_T | \theta)$ où les mesures sont $y_0 \dots y_T$ et les paramètres du modèle ARMA sont regroupés dans le vecteur θ . Le problème étant non-linéaire en θ , la recherche de θ optimal se fait par une procédure itérative, le plus souvent Newton-Raphson, et parfois programmation non-linéaire, par exemple méthode de Powell. Les différences entre estimateurs jouent sur la façon de calculer la vraisemblance et ses dérivées, tantôt exactement, tantôt approxima-

tivement. L'évaluation peut aussi se faire sur le signal ou sa corrélation, sur le résidu du modèle ARMA, sur la transformée de Fourier du signal ou le périodogramme.

1.1. Evaluation sur la corrélation, ou le signal.

L'exemple le plus typique de cette approche est le travail d'ANSLEY, 1979, (et la tentative erronée d'ALI, 1977, dont il est inspiré). ANSLEY, 1979, propose un calcul exact de la vraisemblance du modèle MA, utilisable aussi pour le modèle ARMA après estimation de la partie AR. Le vecteur θ permet le calcul de Ω la covariance théorique du vecteur Y des observations sur $(0, T)$. La matrice Ω est de Toeplitz bande. Sa factorisation de Cholesky donne L tel que $\Omega = LL^T$. La vraisemblance est alors:

$$p(Y|\theta) = (2\pi)^{-(T+1)/2} \text{Det}(L)^{-1} e^{-\sum_{t=0}^T 1/2(L^{-1}Y)^T L^{-1}Y}$$

Un calcul analogue est fourni par RISSANEN, CAINES, 1979, qui réalisent la factorisation de Cholesky sous la forme de l'algorithme de BAUER, 1955. La même expression est utilisée par PHADKE, KEDEM, 1978, pour un calcul de la vraisemblance logarithmique, soit:

$$L(\theta) = -1/2(\text{LogDet}(\Omega) + Y^T \Omega^{-1} Y)$$

Ils proposent trois méthodes de calcul différentes de $L(\theta)$, mais pas des dérivées, ce qui rend plus délicate l'optimisation du critère $L(\theta)$.

GODOLPHIN, 1977, propose une méthode pour le modèle MA, que TUAN, 1979, étend au cas ARMA, et qui recherche le vecteur θ comme transformation linéaire du vecteur r des corrélations du signal: $\theta = Hr$, mais il s'agit d'une vraisemblance approchée. ZETTERBERG, 1969, donne aussi une expression approchée de la vraisemblance qui se linéarise, et lui permet dans une

procédure itérative d'estimer en alternance la partie AR connaissant la partie MA (problème linéaire) et la partie MA connaissant la partie AR (vraisemblance approchée, linéarisée).

1.2. Evaluation sur le résidu.

Le calcul de la vraisemblance peut se faire assez facilement quand le résidu du modèle ARMA est obtenu, par filtrage inverse. Ce résidu devrait être l'entrée du modèle et donc un bruit blanc. Cependant le filtrage inverse suppose connues les conditions initiales du filtre, et en particulier les entrées précédant le premier échantillon, qui ne sont pas accessibles. Aussi le filtrage suppose-t-il ces conditions initiales nulles. La vraisemblance qui s'en déduit ne peut être qu'approchée.

TRETTER, STEIGLITZ, 1967, donnent un calcul du gradient de $L(\theta)$ en fonction du résidu ϵ_t et de ses dérivées $\frac{d\epsilon_t}{d\theta_i}$. Celles-ci se calculent comme:

$$\frac{d\epsilon_t}{d\theta_i} = v_{t-i} \quad \text{avec} \quad v(z) = \frac{A(z)}{B(z)^2} Y(z)$$

Le calcul du Hessien peut aussi se faire avec les mêmes quantités comme le montrent WILSON, 1973, puis GUPTA, 1976, ce dernier donnant un moyen de calcul rapide des dérivées de ϵ_t . Si R est la variance de l'entrée du modèle ARMA, le critère s'écrit:

$$L(\theta) = -\frac{1}{2} \sum_{t=1}^T (\epsilon_t^T R^{-1} \epsilon_t + \text{Log} R)$$

Le gradient et le hessien s'écrivent:

$$\frac{\partial L}{\partial \theta_i} = \sum_t \frac{\partial \epsilon_t}{\partial \theta_i} R^{-1} \epsilon_t$$

et

$$\frac{\partial^2 L}{\partial \theta_i \partial \theta_j} = \sum_t \frac{\partial \epsilon_t^T}{\partial \theta_i} R^{-1} \frac{\partial \epsilon_t}{\partial \theta_j}$$

Des expressions plus complexes sont données par AKAIKE, 1973, ANDERSON, 1975, COOPER, WOOD, 1981, (ces derniers écrivent le modèle ARMA sous forme d'équation d'état). Chacune de ces expressions tente de réduire l'approximation introduite dans les précédentes approches, en particulier dans le terme $\text{Log}(R)$. Toutes souffrent cependant de l'erreur initialement faite en posant comme nulles les entrées précédant la première mesure.

Pour éviter cet écueil, NEWBOLD, 1974, et OSBORN, 1974, proposent des évaluations de la vraisemblance de manière exacte, par l'emploi de la covariance de ces entrées et sorties initiales, qui elle, est connue dès que l'est θ . Si e est le vecteur des résidus de $t=1$ à T , si e^* est le vecteur des entrées $\epsilon_{1-q} \dots \epsilon_0$ et des sorties $y_{1-p} \dots y_0$, les résidus vérifient:

$$\begin{bmatrix} e^* \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ L \end{bmatrix} Y + \begin{bmatrix} I \\ M \end{bmatrix} e^*$$

où L et M sont des matrices ne dépendant que de la réponse impulsionnelle du modèle inverse. La covariance de e^* sera Ω , elle aussi ne dépendant que du modèle inverse. L'idée essentielle est de diagonaliser Ω pour inclure e^* comme un bruit décorrélé dans le critère. En posant $M^* = M\Omega^{1/2}$, on obtient la valeur de e^* maximisant la vraisemblance:

$$e^* = -(M^{*T} M^*)^{-1} M^{*T} L Y$$

puis la vraisemblance:

$$p(Y|\theta) = (2\pi\sigma^2)^{-(T+1)/2} \text{Det}(M^* M^*)^{-1/2} e^{-\frac{1}{2} S^* S / \sigma^2}$$

avec $S = (I - M^* (M^* M^*)^{-1} M^*) LY$ et σ^2 , la variance de l'entrée du modèle. Le calcul du gradient et du Hessien deviennent très compliqués, et ne sont pas indiqués par ces auteurs.

1.3. Evaluation par le périodogramme.

Une autre manière de calculer la vraisemblance du modèle consiste à recourir au périodogramme, c'est-à-dire au carré du module de la transformée de Fourier du signal:

$$I(\omega) = \frac{1}{T} \left| \sum_{t=0}^{T-1} y_t e^{-j\omega t} \right|^2$$

La vraisemblance est donnée par PRADO, 1979, sous la forme:

$$p(I|\theta) = \frac{1}{S_\theta(\omega)} e^{-\frac{1}{2} \frac{I(\omega)}{S_\theta(\omega)}} \quad \text{avec} \quad S_\theta(\omega) = E(I(\omega))$$

Lorsque $S_\theta(\omega)$ est le spectre associé au modèle ARMA, un calcul du gradient et du Hessien est donnée, en vue d'une optimisation par Newton-Raphson. ANDERSON, 1977, avait présenté un calcul analogue, incluant le cas où la partie MA du modèle est remplacé dans θ par son autocorrélation. Il donne aussi une expression de l'espérance conditionnellement à θ du Hessien, afin d'utiliser une méthode de Gauss-Newton pour l'optimisation. Dans chaque cas, les expressions sont complexes.

Une simplification s'obtient, comme souvent dans les méthodes de vraisemblance en négligeant le terme hors de l'exponentielle. C'est ainsi que peuvent être vus les estimateurs proposés par HANNAN, 1969, 1970, 1975.

Il considère la transformée de Fourier de $I(\omega)S_{\theta}(\omega)^{-2}$ comme la corrélation du modèle AR inverse du modèle MA cherché. L'application de l'algorithme de Levinson à cette corrélation donne une estimation des b_i . Le vecteur θ obtenu est introduit dans la récurrence par $\theta_k = 2\theta - \theta_{k-1}$. Que ce soit dans cette version simplifiée ou dans la précédente, il faut souligner que le recours au périodogramme suppose que les données y_t sont périodiques, de période T, il ne saurait donc être question de vraisemblance exacte.

1.4. Evaluation par filtre de Kalman.

Une écriture du modèle ARMA sous forme d'équations d'état permet un calcul exact de la vraisemblance, comme l'a montré SCHWEPPE, 1965, dans le cas à temps continu comme dans celui à temps discret. Il se trouve de plus que le calcul se fait au moyen du filtre de Kalman, qui en extrayant à chaque instant une estimation de l'état du filtre fournit aussi une estimée de l'innovation et de sa variance R_t . La vraisemblance logarithmique s'écrit ainsi:

$$L(\theta) = -\frac{T}{2} \text{Log} 2\pi - \frac{1}{2} \sum_{t=1}^T \left(\frac{\epsilon_t^2}{R_t} + \text{Log} R_t \right)$$

Ce résultat a été étendu par HARVEY, PHILLIPS, 1979, à l'estimation d'une régression linéaire avec résidu ARMA, ou encore le modèle dit ARMAX (avec X comme exogène). Dans le cas stationnaire, le filtre de Kalman peut être remplacé par sa version rapide utilisant les équations de Chandrasekar (FRIEDLANDER, KAILATH, MORF, LJUNG, 1978). Ceci combiné à l'emploi d'une équation d'état d'une forme peu habituelle permet à PEARLMAN, 1980, d'écrire le calcul des résidus ϵ_t comme:

$$\epsilon_t = y_t - Cx_t$$

$$x_{t+1} = Ax_t + K_t \epsilon_t / R_t$$

$$\begin{bmatrix} K_{t+1} \\ L_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & \alpha_t \\ \alpha_t & 1 \end{bmatrix} \begin{bmatrix} K_t \\ AL_t \end{bmatrix} \quad \text{avec} \quad \alpha_t = -CL_t / R_t$$

$$R_{t+1} = R_t (1 - \alpha_t^2)$$

Les conditions initiales du filtre sont: $K_1 = L_1 = AP_1 C^T$ et $R_1 = CP_1 C^T$ où P_1 est solution de l'équation de Lyapunov: $P_1 = AP_1 A^T + BB^T$. Le modèle d'état est:

$$\begin{cases} x_t = Ax_{t-1} + Be_t \\ y_t = Cx_t \end{cases}$$

$$\text{avec} \quad A = \begin{bmatrix} -a_1 & 1 & & \\ & \cdot & \cdot & \\ & & \cdot & \\ & & & -a_n \end{bmatrix} \quad B = \begin{bmatrix} b_0 \\ \cdot \\ b_{n-1} \end{bmatrix} \quad C = [1 \ 0 \ \dots \ 0]$$

GUEGUEN, SCHARF, 1980, donnent une version similaire de ce calcul exact ainsi que deux versions approchées reposant essentiellement sur des choix différents des conditions initiales. Ils soulignent aussi que le filtre de Kalman-Chandrasekar, apparenté aux équations de Levinson s'identifie lorsque les conditions initiales sont issues des covariances expérimentales, avec la factorisation de Cholesky (rapide), ce qui établit un lien avec la méthode sur le résidu de RISSANEN, CAINES, 1979.

Le calcul du gradient et du Hessien n'est pas précisé par ces auteurs, mais il a été donné dans le cas non-stationnaire par GRENIER, 1981-a. Il utilise des filtres dérivés du filtre de Kalman, en parallèle avec lui, en nombre égal à la dimension de l'état pour le gradient et au carré de cette dimension pour le Hessien. Ces filtres se simplifieraient dans le cas stationnaire.

2. Linéarisation.

Les méthodes regroupées dans cette classe sont nombreuses, voire disparates. Leur point commun est de faire intervenir une linéarisation du problème, soit pour la partie MA seule, soit pour l'ensemble du modèle ARMA. Ce seront donc des méthodes approchées, sauf pour certaines d'entre elles qui proposent un schéma itératif.

2.1. Partie AR.

La linéarisation du problème est très simple pour la partie autorégressive, et consiste à décaler sur l'autocorrélation les équations de Yule-Walker du modèle AR, afin d'éviter les autocorrélatifs de R_{-q} à R_q qui sont affectées par la partie MA du modèle:

$$\begin{bmatrix} R_q & R_{q-p+1} \\ R_{q+p+1} & R_q \end{bmatrix} \begin{bmatrix} a_1 \\ \cdot \\ a_p \end{bmatrix} = \begin{bmatrix} R_{q+1} \\ \cdot \\ R_{q+p} \end{bmatrix}$$

Cet estimateur se trouve chez HSIA, LANDGREBE, 1967, ainsi que MASSEY, 1969, qui au lieu de résoudre le système en utilisant la structure de Toeplitz (algorithme de Levinson) permute les lignes pour avoir une structure de Hankel. GERSCH, 1970, a montré que cet estimateur n'est pas biaisé. L'ajout des équations au delà de $q+p$ donne un système sur-dimensionné qui est résolu par l'intermédiaire de la pseudo-inverse, ce qui équivaut à une méthode de moindres carrés sur la queue de la corrélation. Une méthode de moindres carrés pondérés est proposée par CADZOW, 1980-a et b et CADZOW, MOSES, 1981. Ces derniers suggèrent un choix optimal de la pondération.

2.2. Partie MA par modèle AR long.

Dans cette approche, on se sert de l'équivalence entre un modèle MA(q) et un modèle AR(∞). L'approximation consiste à remplacer le modèle AR(∞) par un modèle d'ordre L fini mais élevé. Les variantes jouent sur la manière de revenir du modèle AR(L) au modèle MA(q). L'idée originale est due à DURBIN, 1959, qui par un calcul de la vraisemblance des b_i étant donnés les a_i , $i=1 \dots L$ montre que la partie MA est solution d'une équation de Yule-Walker:

$$\begin{bmatrix} \alpha_0 & & & \\ & \alpha_{q-1} & & \\ & & \ddots & \\ & & & \alpha_0 \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_q \end{bmatrix} = - \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{bmatrix} \quad \text{avec} \quad \alpha_i = \sum_{j=0}^{L-i} a_j a_{j+i}$$

L'inversion AR-MA se fait donc par l'algorithme de Levinson. DURBIN, 1960, étend cette idée au cas ARMA: alternativement, la partie MA s'identifie comme ci-dessus mais après filtrage du modèle AR long par l'inverse de la partie AR estimée, puis la partie AR est estimée sur les corrélations du signal filtré par l'inverse de la partie MA. GRAUPE, KRAUSE, MOORE, 1975, simplifient cette procédure en identifiant la partie MA comme un modèle AR vérifié par les éléments $a_{p+1}^{(L)} \dots a_{p+q}^{(L)}$ du modèle AR long, puis la partie AR se détermine d'après les coefficients $a_1^{(L)} \dots a_p^{(L)}$. GRAUPE, CLINE, 1975, montrent que cet estimateur est non-biaisé. La méthode d'ATASHROO, 1977, est identique à celle de DURBIN, 1959. ATAL, SCHROEDER, 1978, procèdent en deux étapes, estimant un modèle ARMA dont les ordres sont supérieurs à ceux du modèle cherché, puis sur la réponse impulsionnelle de ce modèle les coefficients du modèle d'ordre correct.

La même méthode sert de base à trois variantes qui améliorent l'estimateur. Dans l'approche de MAYNE, FIROOZAN, 1978, le modèle AR long permet de reconstituer par filtrage inverse le résidu ϵ_t , et le modèle ARMA

est estimé avec ϵ_t comme son entrée, et y_t sa sortie, en minimisant l'erreur

w_t :

$$w_t = \sum_{i=0}^p a_i y_{t-i} - \sum_{j=0}^q b_j \epsilon_{t-j}$$

Une telle procédure donne des valeurs biaisées aussi une seconde étape filtre les signaux ϵ_t et y_t par l'inverse du modèle MA avant de reprendre la minimisation de l'erreur w_t par moindres carrés. La procédure peut s'itérer. La limitation essentielle reste due au modèle AR long initial. Pour la dépasser, KONVALINKA, MATAUSEK, 1979, suggèrent à chaque itération de leur procédure, d'identifier le modèle AR long d'où sera déduit le modèle ARMA, sur le signal d'origine filtré par l'inverse du modèle MA trouvé à l'itération précédente. La méthode est plus lourde mais bien plus précise, et converge assez rapidement sauf si des zéros ont un module très proche de 1. MILLS, MULLIS, ROBERTS, 1980, proposent la même procédure, mais en faisant croître l'ordre des parties AR et MA en commençant à 1, et en les augmentant de une unité à chaque itération jusqu'à l'ordre cherché où s'arrête le processus.

2.3. Modélisation AR d'ordre infini.

L'algorithme de décomposition donné par RISSANEN, 1973, pour la factorisation rapide d'une matrice de Toeplitz ou d'une matrice polynomiale positive est relié à l'algorithme de Levinson dans la version qu'en donnent LE ROUX, GUEGUEN, 1976. Appliquée à la corrélation d'un signal de type MA, cet algorithme converge vers le modèle MA (LE ROUX, GRENIER, 1980), et il s'avère n'être rien d'autre que la décomposition de SCHUR, 1917, dont une version vectorielle est donnée par DELSARTE, GENIN, KAMP, 1979, et YOULA, KAZANJIAN, 1978, qui en montrent la convergence (voir aussi GAMBOTTO, 1979).

Cet algorithme obéit à la récurrence suivante:

$$\begin{bmatrix} E_0(i) & \dots & E_q(i) \\ 0 & D_1(i) & \dots & D_q(i) \end{bmatrix} = (1 - k_i^2)^{-1/2} \begin{bmatrix} 1 & k_i \\ k_i & 1 \end{bmatrix} \begin{bmatrix} E_0(i-1) & \dots & E_q(i-1) \\ D_1(i-1) & \dots & D_q(i-1) & 0 \end{bmatrix}$$

avec $k_i = -D_1(i-1)(E_0(i-1))^{-1}$ et $E_j(0) = R_j$ $D_j(0) = R_{-j}$

L'algorithme outre sa simplicité possède une interprétation utile, puisqu'il équivaut à approximer le modèle MA par des modèles AR d'ordre croissant (LE ROUX, GRENIER, 1980) et permet donc au lieu de choisir arbitrairement l'ordre AR long comme dans les méthodes du groupe précédent, de ne s'arrêter qu'à la convergence à ϵ près. Cependant l'estimateur trouvé n'a pas de bonnes propriétés statistiques.

2.4. Réponse impulsionnelle (homomorphique).

Le principe de ces méthodes est donné par SHANKS, 1967, qui cherche le modèle ARMA en connaissant sa réponse impulsionnelle. La partie AR s'obtient par moindres carrés sur la queue de la réponse impulsionnelle, et la partie MA par filtrage de Wiener entre la réponse impulsionnelle cherchée, et celles décalées de la partie AR. Cette méthode, destinée à priori à la conception de filtres numériques peut être utilisée pour le modèle ARMA après calcul de la réponse impulsionnelle. Celle-ci peut s'obtenir par les techniques homomorphiques. Ainsi la partie d'indice positif du cepstre de y_t permet de factoriser le spectre de y_t et donne la réponse impulsionnelle h_t . En effet, le cepstre est la transformée de Fourier inverse du logarithme du spectre, ce qui en transformée en z s'écrit: $C(z) = \text{Log}R(z)$. La factorisation de $R(z) = H(z)H(z^{-1})$ en cepstre s'écrit $C(z) = \text{Log}H(z) + \text{Log}H(z^{-1})$. En ne conservant que les indices positifs du cepstre on extrait $\text{Log}H(z)$ et par l'opération inverse (exponentielle) on reconstruit $H(z)$ puis h_t (OPPENHEIM,

KOPEC, TRIBOLET, 1976).

Le calcul du modèle ARMA peut se faire sur h_t , soit par la méthode de Shanks (KOPEC, OPPENHEIM, TRIBOLET, 1979) ou par une méthode de moindres carrés itératifs (STEIGLITZ, 1977) minimisant à l'itération k le critère:

$$J = \sum_{t=0}^T \left[\frac{D_k v_t}{D_{k-1}} - \frac{N_k \delta_{t,0}}{D_{k-1}} \right]^2$$

où N_k et D_k sont la partie MA, et AR du modèle à l'itération k , v_t est la réponse impulsionnelle, les produits ou divisions doivent être lus comme des opérations de filtrage. Noter l'analogie avec la méthode que MAYNE, FIROOZAN, 1978, utilisent sur le modèle AR long qui équivaut à l'inverse de la réponse impulsionnelle. D'après DONE, RUSHFORTH, 1979, cette méthode donne de bons résultats sur des signaux du type parole voisée (entrée suite d'impulsions) mais de moins bons résultats sur un signal effectivement ARMA, avec entrée en bruit blanc.

Une variante est proposée par YEGNANARAYANA, 1981, qui utilise aussi le cepstre, mais en séparant directement les parties AR et MA du modèle. Pour cela il étudie la dérivée négative du spectre de phase, (DNSP) qui est la partie réelle de la transformée de Fourier de la suite des coefficients cepstraux multipliés par t , soit tc_t . On a alors approximativement une DNSP qui a des pics positifs pour chaque pôle et négatifs pour chaque zéro, et il suffit de séparer la DNSP en ses parties positive et négative pour reconstituer les cepstres des parties AR et MA. Deux possibilités pour revenir au modèle: soit reconstituer la réponse impulsionnelle de chaque partie et la tronquer à l'ordre voulu, soit tronquer les cepstres et reconstituer les parties du modèle par les équations écrites par GRAY, MARKEL, 1976. Il est clair que cette méthode n'est valable que si les pôles et zéros se manifestent à des fréquences différentes, sinon la séparation en parties

positives et négatives de la DNSP ne respectera pas la répartition des pôles et zéros.

2.5. Corrélations de la partie MA.

Les méthodes procédant par factorisation de la corrélation de la partie MA mènent à un estimateur qui est médiocre du point de vue statistique lorsque la corrélation est obtenue en tronquant de $-q$ à $+q$ la corrélation expérimentale du signal $R_k = \frac{1}{N-k} \sum_t y_t y_{t+k}$. Pour remédier au biais constaté, WALKER, 1961, propose de ne pas tronquer mais d'estimer les autocorrélations $\rho_0 \dots \rho_q$ du signal à partir des corrélations $R_0 \dots R_N$ avec $N > q$. En maximisant la vraisemblance des ρ_i il parvient à une procédure itérative sur les ρ_i qui par son recours aux puissances quatrièmes des ρ_i n'est pas sans évoquer la procédure de HANNAN, 1969, utilisant la vraisemblance du périodogramme. La procédure s'étend au modèle ARMA représenté par sa partie AR et les corrélations de sa partie MA (WALKER, 1962). Les liens qui apparaissent alors sont avec la procédure d'ANDERSON, 1977. Une version simplifiée est donnée dans le cas MA par NEWTON, 1980, qui ajuste au sens des moindres carrés le périodogramme du signal à la transformée de Fourier de l'autocorrélation ($R_{-q} \dots R_0 \dots R_q$). Il faut remarquer que si l'estimateur en est amélioré, la positivité de l'autocorrélation trouvée et donc sa factorisabilité ne sont pas pour autant assurées.

ANNEXE 7. REPRESENTATION GRAPHIQUE DES RELIEFS.

L'objet de cette brève annexe est de préciser les conventions permettant la lecture des divers exemples de reliefs présentés dans ce document. Dans chacune des figures proposées se trouve un ou plusieurs rectangles figurant la fonction $\rho(t,\omega)$ avec une présentation analogue à celle des spectrogrammes. L'axe horizontal représentera le temps, gradué en nombre d'échantillons. L'axe vertical figurera les pulsations ou les fréquences, graduées de 0 à 50 en pourcentage de la fréquence d'échantillonnage. A chaque point (t,ω) dans ce système d'axes, la troisième composante qu'est l'amplitude $\rho(t,\omega)$ est représentée à l'origine par une couleur, convertie ici pour les besoins de l'impression en un niveau de gris. Les niveaux sont en nombres limités (au plus 7) afin que leurs séparations, qui sont des courbes de niveaux, soient visibles. Leur code est lisible sur l'échelle figurant à gauche de chaque figure où se voit la correspondance entre niveau d'énergie de $\rho(t,\omega)$ converti en dB et niveau de gris. Le niveau 0 de référence est fixé arbitrairement à $\rho(t,\omega)=10^{-7}$. Quant aux valeurs négatives de certains reliefs (Wigner-Ville) elles sont éliminées avant la conversion en dB par une saturation à 10^{-7} de toutes les valeurs inférieures à ce niveau.

Sur les figures présentant un seul relief, on trouve sous ce relief deux signaux. Celui du haut est le signal qui a été analysé, et celui du bas est en général le résidu ou l'erreur de prédiction du modèle. L'exception concerne les méthodes paramétriques où ce second signal est la transformée de Hilbert du premier dans le cas du relief de Wigner-Ville, et le premier signal dupliqué dans le cas de la STFT.

Dans les images originales, les reliefs sont figurés en couleurs. Un choix satisfaisant s'est avéré être le code des cartes de géographie, du bleu sombre du fond des mers au rouge des sommets. Dans un système de couleurs RVB où chaque couleur est codée sur 4 éléments binaires (de 0 à 15), la liste retenue est la suivante:

Rouge	Vert	Bleu	
15	0	0	niveau haut
15	6	0	
15	11	0	
15	15	0	
12	15	0	
8	15	0	
0	15	0	
1	10	8	
1	5	13	
1	0	15	
1	0	11	
1	0	7	
1	0	3	niveau bas

ANNEXE 8. LA NOTION DE STABILITE LOCALE.

On a vu au chapitre 10 que la synthèse de la parole au moyen d'un modèle ARMA non-stationnaire, de type évolutif soulevait un problème de stabilité du modèle engendrant le signal. Il ne suffisait pas que le modèle soit stable au sens usuel, encore fallait-il qu'il ne se trouve aucun intervalle de temps sur lequel le modèle tangent au modèle utilisé soit instable. Cette annexe se propose de formaliser brièvement la notion qu'il est nécessaire d'introduire pour décrire cet effet. Il s'agira d'une stabilité locale qui s'oppose à la stabilité au sens usuel dans laquelle, rappelons-le le système est dit stable si et seulement si toute entrée bornée donne une sortie bornée. Ceci s'obtient si et seulement si la réponse impulsionnelle $h(t,s)$ du système est absolument sommable sur s , et pour tout t :

$$\sum_{s=-\infty}^t |h(t,s)| < \infty \quad \forall t$$

On introduit ici la notion plus restrictive de stabilité locale.

Définition.

Un modèle ARMA(p,q) non-stationnaire, à coefficients bornés sera dit localement stable si et seulement si pour tout t , le modèle tangent à l'instant t est ρ -contractant.

Dans cette définition, le modèle tangent est pris au sens de la partie I (chap.3), c'est le modèle obtenu en figeant à l'instant t le modèle d'état observable (A_t, B_t, C) équivalent au modèle ARMA(p,q). Quant à la notion de ρ -contraction, elle se définit comme suit:

Définition.

Un système sous forme d'équation d'état (A_t, B_t, C) sera dit ρ -contractant si et seulement si le rayon spectral (module de la plus grande valeur propre) de la matrice A_t est borné supérieurement pour tout t par ρ où ρ est réel et $0 < \rho < 1$.

Quels liens y-a-t'il entre la stabilité locale et la stabilité usuelle (entrée bornée/sortie bornée)? Pourquoi avoir privilégié le modèle d'état plutôt que le modèle ARMA? Pourquoi cette notion de ρ -contraction alors qu'il paraissait plus simple de poser simplement la stabilité asymptotique du modèle tangent? La réponse à ces trois questions tient en la proposition qui suit montrant que la stabilité est conséquence de la stabilité locale. Deux contre-exemples montrent ensuite que l'emploi du modèle ARMA au lieu du modèle d'état, comme l'emploi de la stabilité asymptotique du modèle tangent au lieu de la ρ -contraction auraient conduit à une notion de stabilité locale incompatible avec la stabilité.

Proposition.

Tout modèle ARMA non-stationnaire localement stable est stable.

Démonstration.

Il suffit de considérer l'équation d'état (A_t, B_t, C) équivalente au modèle ARMA, et la matrice de transition $\phi(t, s)$ permettant de calculer la réponse impulsionnelle $h(t, s)$. Si on a:

$$x_t = A_{t-1} x_{t-1} + B_t e_t$$

$$y_t = C x_t$$

alors

$$\phi(t, t-s) = A_{t-1} \dots A_{t-s}$$

$$h(t, t-s) = C\phi(t, t-s)B_{t-s}$$

Si pour tout t , le rayon spectral de A_t est borné par $\rho < 1$, alors celui de $\phi(t, t-s)$ est borné par ρ^s , et les éléments de B_t étant bornés, il existe K réel, positif tel que $|h(t, t-s)| < K\rho^s$ pour tous t et s . Il s'ensuit que le modèle est stable puisque la réponse impulsionnelle $h(t, s)$ est absolument sommable.

Contre-exemple 1.

Afin de montrer la nécessité de recourir au modèle d'état dans la définition de la stabilité locale, voici un exemple de modèle ARMA non-stationnaire qui si on le fige (dans sa forme ARMA) est stable pour tout t , ρ -contractant, et pourtant instable. Ce modèle est défini par:

$$y_{2t} = \alpha y_{2t-1} + e_{2t}$$

$$y_{2t+1} = -2\alpha y_{2t} - \alpha^2 y_{2t-1} + e_{2t+1}$$

On choisira α tel que $|\alpha| < 1$, ce qui assure la stabilité du modèle figé aux instants pairs comme aux instants impairs. Si la ρ -contraction était définie à partir du modèle ARMA, celui-ci serait ρ -contractant. Un calcul rapide montre que la réponse à l'entrée $e_t = \delta_{t,0}$ de ce modèle est:

$$y_{2t} = (-1)^t 2.3^{t-1} \alpha^{2t}$$

$$y_{2t+1} = (-1)^{t+1} 2.3^t \alpha^{2t+1}$$

Il est clair que cette réponse diverge dès que $|3\alpha^2| > 1$ et donc pour $3^{-1/2} < |\alpha| < 1$ le système est instable. Convertir ce modèle dans sa forme

d'état montrerait bien sûr qu'il n'est pas localement stable, du moins avec ces valeurs de α . On obtiendrait pour les matrices (A_t, B_t, C) :

$$A_{2t} = \begin{bmatrix} 2\alpha & 1 \\ 0 & 0 \end{bmatrix} \quad A_{2t+1} = \begin{bmatrix} \alpha & 1 \\ -\alpha^2 & 0 \end{bmatrix} \quad B_t = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad C = [1 \ 0]$$

d'où

$$\text{Det}(\lambda I - A_{2t}) = \lambda(\lambda - 2\alpha)$$

$$\text{Det}(\lambda I - A_{2t+1}) = \lambda^2 - \alpha\lambda + \alpha^2$$

Le modèle est donc localement stable dès que $|\alpha| < 1/2$.

Contre-exemple 2.

Voici maintenant un exemple de modèle mis sous forme d'état qui n'est pas ρ -contractant, mais dont pourtant le modèle tangent est stable quel que soit t . Il sera montré qu'il est instable. Le modèle a un état de dimension 2 avec:

$$A_t = \begin{bmatrix} 0 & 1 \\ a(t) & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad C = [1 \ 0] \quad a(t) = \begin{cases} (t+1)/(t-1) & \text{si } t < -1 \\ 1/2 & \text{si } t \geq -1 \end{cases}$$

Le calcul de la matrice $\Phi(1-n, 1-n-s)$ donne:

$$\text{si } s=2m, \quad \Phi(1-n, 1-n-s) = \begin{bmatrix} 0 & n/(n+s) \\ (n-1)/(n+s-1) & 0 \end{bmatrix}$$

$$\text{si } s=2m+1, \quad \Phi(1-n, 1-n-s) = \begin{bmatrix} (n-1)/(n+s-1) & 0 \\ 0 & (n-1)/(n+s) \end{bmatrix}$$

ce qui donne pour la réponse impulsionnelle:

$$h(1-n, 1-n-2m) = h(1-n, 1-n-2m+1) = n/(n+2m)$$

Prenons alors $n=0$, $e_t=1$ pour $t \leq 0$, alors:

$$y_0 = \sum_{s=-\infty}^0 h(0, s) e_s = 1 + \frac{1}{3} + \frac{1}{3} + \frac{1}{5} + \frac{1}{5} + \dots \rightarrow +\infty$$

Ainsi la valeur de la sortie à $t=0$ n'est pas bornée bien que l'entrée le soit, le système est instable. Ceci a été obtenu parce que le rayon spectral de A_t qui vaut $(1-\frac{2}{1-t})^{1/2}$ quand $t < -1$ n'a pas été borné par $\rho < 1$, et a pu tendre vers 1 pour $t \rightarrow -\infty$.

Imposer la stabilité par des contraintes instantanées sur le modèle ne peut se faire, ces exemples le montrent, que par une définition très restrictive de la stabilité locale passant par la ρ -contraction du modèle d'état.

