1991
7

BIBLIOTHÈQUE UNIVERSITAIRE
SECTION DE SCIENCES
LILLE

1991
7

N^O d'ordre : 762

THÈSE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE FLANDRES ARTOIS

pour obtenir

LE TITRE DE DOCTEUR DE L'UNIVERSITÉ

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

par

ACCARDO Jérôme

VALEURS PROPRES DES MATRICES DE TOEPLITZ ET MATRICES DE COVARIANCE DE PROCESSUS.

Soutenue le 9 Octobre 1991 devant la Commission d'Examen :

B.U. LILLE I
D 030 093402 2

Président : P. JACOB, Université de Lille

Rapporteurs : M.C. VIANO, Université de Lille

G. MÉLARD, Université Libre de Bruxelles

Examinateur : C. GOURIÉROUX, E.N.S.A.E., C.R.E.S.T., Paris

Je tiens à remercier

M. Pierre JACOB, qui m'a initié à la théorie des probabilités et qui m'a fait l'honneur de présider ce jury.

Mme Marie Claude VIANO et M. Guy MELARD, rapporteurs de cette thèse pour leur lecture approfondie et leurs judicieuses critiques.

Je veux aussi témoigner ma gratitude à Françoise GABARROCHE et Evelyne COUTANT qui ont dactylographié le manuscrit dans une bonne humeur jamais démentie et, pour leurs conseils et l'attention qu'ils ont bien voulu m'accorder, aux membres du laboratoire de probabilités et statistique de Lille I, en particulier M. Raymond MOCHE, ainsi qu'aux membres du CREST ENSAE, particulièrement Emmanuel GUERRE, Jean Michel ZAKOIAN, Philippe BARBE et Patrice BERTAIL.

Je remercie de tout coeur Vivette et Alain ACCARDO et Juliette BAILLOT pour leur soutien et leur patience.

Enfin je veux souligner ici ma dette envers M. Christian GOURIEROUX, qui a dirigé ce travail. Je lui suis reconnaissant certes de ses intuitions, de son enthousiasme, de sa disponibilité, bref de son inestimable soutien théorique et moral mais plus encore, peut-être, de m'avoir appris à chercher la plage des idées sous le pavé des calculs.

à Martine "BAB",

Le ministre : "Here is the report Mr President, it is rather complex and.. "

Groucho : "Complex? Ah,ah,ah! You're kiddin'! A four years old boy would easily get it! ... Well, hum, ...could you find me a four years old boy?

Les MARX BROTHERS, "Duck Soup".

CONTENU :

-INTRODUCTION p 2

-PARTIE 1 : Processus stationnaires unidimensionnels p 8

I Introduction p 9

II Etude des valeurs propres dans le cas ARMA p 11

a) Processus moyenne mobile d'ordre 1 p 11

b) Extension au cas ARMA p 14

c) Encadrement des valeurs propres dans le cas de processus moyennes mobiles ou autorégressifs purs. p 19

d) Vitesse de convergence et distribution empirique p 25

III Distribution asymptotique des valeurs propres p 28

a) Le théorème de Szegö p 29

b) Distribution asymptotique des valeurs propres, lorsque la densité spectrale n'est pas bornée p 34

c) Le cas des processus singuliers p 38

Annexe 1 p 42

Annexe 2 p 47

Annexe 3 p 52

-PARTIE 2 : Processus multivariés p 58

I Approximation des valeurs propres d'un processus MA(1) multivarié p 59

II Approximation des valeurs propres d'un AR(1) multivarié p 66

a) Inversion de la matrice de covariance d'un processus autorégressif multivarié p 67

b) Valeurs propres d'un AR(1) multivarie P 70

c) Etude de l'erreur d'approximation p 71

III Processus multivariés et théorème de Szegö p 73

-PARTIE 3 : Processus non stationnaires p 78

I Valeurs propres et marche aléatoire p 79

1) Calcul des valeurs propres p 80

2) Comportement asymptotique p 86
3) Résultats p 87
4) Théorème de Szegö p 88
5) Comparaison stationnaire-non stationnaire p 89
II Processus autorégressifs non stationnaires p 90
1) Approximation des valeurs propres p 90
2) Théorème de Szegö p 92
III Etude des valeurs propres maximales p 94
1) Ordre de la valeur propre maximale p 94
2) Degré de non stationnarité et distribution des "grandes" valeurs propres p 102
Annexe p 109

-BIBLIOGRAPHIE p 115

VALEURS PROPRES des MATRICES de TOEPLITZ
et
MATRICES de COVARIANCE de PROCESSUS

par Jérôme ACCARDO

INTRODUCTION

Les matrices de Toeplitz jouent un rôle tout à fait important dans la théorie des processus stochastiques, où elles apparaissent comme les matrices d'autocovariance de processus stationnaires à temps discret (pour un survey voir par exemple Makhoul(1975)).

Rappelons que la particularité de telles matrices est d'avoir ses éléments constants le long des parallèles à la diagonale principale. En d'autres termes, si $c_{i,j}$ désignent les éléments de la matrice, on a :

$$c_{i,j} = c_{i',j'} \quad , \quad \text{dès que} \quad i-j = i'-j'.$$

De sorte que leur forme typique est :

$$\begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \cdots & c_{-1} & c_0 & c_1 & c_2 & c_3 & \cdots \\ \cdots & c_{-2} & c_{-1} & c_0 & c_1 & c_2 & c_3 \cdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & \cdots & \dot{c}_{-2} & \dot{c}_{-1} & \dot{c}_0 & \dot{c}_1 & \dot{c}_2 \cdots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}$$

Les matrices de Toeplitz ont une structure spectrale trés simple à décrire dès qu'elles sont infinies, autrement dit quand elles s'identifient à des opérateurs T sur les series formelles $\sum_{-\infty}^{+\infty} a_h X^h$, du type

$$T\left(\sum_{-\infty}^{+\infty} a_h X^h\right) = \sum_{h=-\infty}^{+\infty} \left(\sum_{j=-\infty}^{+\infty} c_j \, a_{j+h}\right) X^h \quad) \quad ;$$

dans ce cas en effet, Toeplitz a montré que les valeurs propres sont données par les valeurs de la fonction

$$f(z) = \sum_{-\infty}^{+\infty} c_h z^h$$

sur le cercle unité ($|z| = 1$) (Toeplitz 1911).

En revanche le cas de matrices finies $(c_h)_{-n\leq h\leq n}$ est beaucoup plus complexe et donne lieu à deux types d'approche :

- la détermination des valeurs propres elles même (ou d'approximations intéressantes) d'une part,
- l'étude de leur distribution asymptotique (i.e. quand $n\to\infty$), d'autre part.

1) DETERMINATION

En fait on connait mal la structure des valeurs propres des matrices de Toeplitz finies (voir par exemple Delsarte, Genin (1983) ou Cybenko (1984)). Si, dans la décomposition en éléments propres de ces matrices (étape essentielle dans les méthodes haute résolution d'analyse spectrale ,voir Golub (1984)), on s'appuie effectivement sur leur forme trés particulière quand on calcule les vecteurs propres (par exemple en utilisantl'algorithme de Levinson, voir Guerin (1983) ou Delsarte, Genin (1986)), la recherche des valeurs propres reste, elle, un problème ouvert ; les algorithmes existants n'utilisent en effet que trés faiblement la forme spécifique de ce type de matrices (essentiellement le fait qu'elles sont centro-symétriques, Cadzow (1984)) ou Zhang (1989). Si le calcul des éléments propres paraît ainsi hors de portée dans le cas de matrices de Toeplitz quelconques, on peut néanmoins chercher à obtenir des résultats en se bornant aux matrices associées à certaines classes de processus. De fait, les seuls exemples relativement généraux où le spectre de la matrice est connu (ou "bien" approché) sont fournis par les matrices de covariance de processus unidimensionnel stationnaire vérifiant une équation moyenne mobile d'ordre 1 (MA(1)) ou autorégressive d'ordre 1 (AR(1)), c'est à dire si le processus $\{ x_t ; t \in \mathbb{Z} \}$ vérifie :

$$X_t = \varepsilon_t - \rho\varepsilon_{t-1} \quad \text{ou} \quad X_t - \rho X_{t-1} = \varepsilon_t \quad \text{avec } -1 < \rho < 1 \ ,$$

où $\{ \varepsilon_t ; t \in \mathbb{Z} \}$ est un bruit blanc

Dans ces deux cas en effet la matrice de covariance est suffisamment simple pour que l'on puisse développer le déterminant caractéristique de la matrice $\Gamma_n = \text{var}\{ (X_t, X_{t-1}, \ldots, X_{t-n+1})' \}$. Rappelons que dans le cas d'un

processus moyenne mobile les valeurs propres sont déterminées exactement alors que dans le cas autorégressif on doit se contenter d'approximations dont la précision augmente avec n la taille de la matrice (voir Grenander,Szegö (1958) ou encore Stroeker (1984) pour un essai d'amélioration de l'approximation pour les AR(1)).

On se propose dans une première partie de ce travail d'étendre ces résultats à des processus plus généraux, à savoir d'abord aux processus stationnaires unidimensionnels autorégressifs ou moyenne mobile d'ordre quelconque (AR(p) ou MA(q)), ensuite aux processus multivariés.

La difficulté tient ici à ce que les calculs menés pour les processus MA(1) et AR(1) unidimensionnels reposent sur des développements de déterminants caractéristiques qui ne sont praticables qu'en raison de l'extrême simplicité des matrices considérées.C'est pourquoi on abordera la détermination des valeurs propres de la matrice de covariance d'un processus stationnaire en utilisant une approche alternative fondée sur l'approximation de ces matrices par d'autres matrices dont la structure spectrale est connue. Cette méthode permet de résoudre les cas "purs", autrement dit les cas où le processus est purement autorégressif ou purement moyenne mobile. En revanche elle échoue devant la complexité des cas "mixtes" (i.e. ARMA(p,q)).

Les problèmes que posent la détermination, exacte ou approchée, des valeurs propres d'une matrice de Toeplitz générale ont motivé historiquement un autre type de description de la stucture spectrale de telles matrices : l'étude de la distribution asymptotique de ces valeurs propres quand la taille n de la matrice tend vers l'infini.

2) DISTRIBUTION ASYMPTOTIQUE

Cette approche a été initiée par G.Szegö qui a montré le résultat fondamental suivant :si Γ_n est la matrice de Toeplitz (finie)

$$\Gamma_n = \begin{bmatrix} c_0 & c_{-1} & c_{-2} & \cdots & c_{n-2} & c_{n-1} \\ c_{-1} & c_0 & c_{-1} & c_{-2} & \cdots & c_{n-3} & c_{n-2} \\ c_{-2} & c_{-1} & c_0 & c_{-1} & & c_{n-4} & c_{n-3} \\ & & \ddots & \ddots & & \vdots & \vdots \\ \vdots & & & \ddots & & & \\ & & & & & c_0 & c_{-1} \\ c_{-n+1} & & \cdots & & & c_{-1} & c_0 \end{bmatrix}$$

et si on note $\lambda_\nu^{(n)}$ $\nu = 1,2,\ldots,n$ ses valeurs propres, alors pour toute fonction continue h on a

$$(1) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} h\,[\lambda_\nu^{(n)}] = \frac{1}{2\pi} \int_{-\pi}^{+\pi} h\,[f(\omega)]\, d\omega$$

où f est la fonction

$$f(z) = \sum_{-\infty}^{+\infty} c_k\, z^k \qquad .$$

[Voir Szegö (1920)].

Ainsi dans un cadre fini les valeurs propres d'une matrice de Toeplitz ne sont plus nécessairement égales aux valeurs prises par f, mais définissent une suite de mesures ponctuelles $\frac{1}{n} \sum_{\nu=1}^{n} \delta_{\lambda_\nu^{(n)}}$, (δ_x désignant la mesure de Dirac en x) qui converge en loi vers la transportée par f de la loi uniforme sur $[-\pi,+\pi]$.

On dira, heuristiquement, que les $\lambda_\nu^{(n)}$ sont asymptotiquement répartis comme f. Ce résultat est plus connu dans un de ses cas particuliers; pour h(x)=log(x), on retrouve l'égalité de Kolmogorov (voir par exemple Gouriéroux,Monfort(1990)).

Le point essentiel est que le théorème de Szegö n'a été établi que pour des processus unidimensionnels, stationnaires et *admettant une densité spectrale bornée.*

Or d'une part les processus étudiés sont généralement multivariés, d'autre part une classe de processus apparue récemment dans la littérature, les processus fractionnaires (voir Gonçalves (1987)par exemple), et dont l'intérêt est de faire le lien entre stationnarité et non stationnarité,

fournit des exemples de processus stationnaires à densité spectrale non bornée.

On s'est donc attaché, dans une seconde partie de cette étude, à étendre le théorème de Szegö ; on montre, de fait, que le théorème se généralise aisément aux processus multidimensionnels et surtout que les processus à densité non bornée intégrable vérifient encore le théorème.

Ce point est tout à fait intéressant dans la mesure où il suggère de poursuivre la généralisation au cas non stationnaire.

3) CAS NON STATIONNAIRE.

L'analyse des valeurs propres des matrices de covariances d'un processus non stationnaire (convenablement définies) s'appuie sur les résultats et les méthodes introduits dans les parties précédentes et fournit quelques conclusions remarquables.

En particulier, l'étude détaillée d'une marche aléatoire

$$X_t = X_{t-1} + \varepsilon_t$$

fait apparaître les valeurs propres et leur distribution asymptotique, comme la limite quand ρ tend vers 1 des quantités homologues calculées pour un processus autorégressif d'ordre 1 (AR(1)) stationnaire

$$X_t = \rho X_{t-1} + \varepsilon_t \qquad , -1 < \rho < 1$$

ce qui est d'autant plus frappant que les matrices de covariance de la marche aléatoire ne sont d'aucune manière limite des matrices correspondantes du processus autorégressif stationnaire ; elles ne présentent d'ailleurs pas une structure de matrice de Toeplitz.

En fait, et c'est sans doute le résultat le plus intéressant, le théorème de Szegö qu'on associe toujours aux matrices de Toeplitz est beaucoup plus général et concerne une classe plus vaste de matrices.

Il s'étend ainsi aux processus non stationnaires vérifiant une équation autorégressive ; à cette occasion on peut mettre en évidence le lien entre la

fournit des exemples de processus stationnaires à densité spectrale non bornée.
On s'est donc attaché, dans une seconde partie de cette étude, à étendre le théorème de Szegö ; on montre, de fait, que le théorème se généralise aisément aux processus multidimensionnels et surtout que les processus à densité non bornée intégrable vérifient encore le théorème.
Ce point est tout à fait intéressant dans la mesure où il suggère de poursuivre la généralisation au cas non stationnaire.

3) CAS NON STATIONNAIRE.

L'analyse des valeurs propres des matrices de covariances d'un processus non stationnaire (convenablement définies) s'appuie sur les résultats et les méthodes introduits dans les parties précédentes et fournit quelques conclusions remarquables.
En particulier, l'étude détaillée d'une marche aléatoire

$$X_t = X_{t-1} + \varepsilon_t$$

fait apparaître les valeurs propres et leur distribution asymptotique, comme la limite quand ρ tend vers 1 des quantités homologues calculées pour un processus autorégressif d'ordre 1 (AR(1)) stationnaire

$$X_t = \rho X_{t-1} + \varepsilon_t \qquad , -1 < \rho < 1$$

ce qui est d'autant plus frappant que les matrices de covariance de la marche aléatoire ne sont d'aucune manière limite des matrices correspondantes du processus autorégressif stationnaire ; elles ne présentent d'ailleurs pas une structure de matrice de Toeplitz.
En fait, et c'est sans doute le résultat le plus intéressant, le théorème de Szegö qu'on associe toujours aux matrices de Toeplitz est beaucoup plus général et concerne une classe plus vaste de matrices.
Il s'étend ainsi aux processus non stationnaires vérifiant une équation autorégressive ; à cette occasion on peut mettre en évidence le lien entre la

distribution des valeurs propres et le degré de non stationnarité du processus.

4) PLAN

Le travail s'organise de la façon suivante :

- Dans la première partie (" Processus stationnaires unidimensionnels "), on présente une méthode d'approximation des matrices de covariance des processus stationnaires satisfaisant un modèle ARMA. Ces approximations fournissent à leur tour des valeurs approchées des valeurs propres de ces matrices quand le processus considéré est moyenne mobile (MA) ou autorégressif (AR) pur. On montre, au passage, comment cette méthode permet d'obtenir dans le cas des processus ARMA une démonstration "élémentaire" du théorème de Szegö.On étend ensuite ce théorème aux processus fractionnaires stationnaires et aux processus stationnaires dont la mesure spectrale admet une partie discrète.

-Dans la seconde partie on reprend les questions abordées précédemment mais cette fois dans un cadre multidimensionnel(" Processus Multivariés ") ; on fournit des approximations des valeurs propres dans le cas des processus AR(1) stationnaires multivariés et on démontre une version multidimensionnelle du théorème de Szegö.

-La partie 3 aborde alors, d'un point de vue théorique le cas des processus non stationnaires. On dérive en premier lieu les valeurs propres d'une marche aléatoire dont on étudie ensuite la distribution asymptotique. Les résultats obtenus sont alors généralisés aux processus non stationnaires vérifiant une équation autorégressive. On montre enfin que le théorème de Szegö persiste dans ce cadre non stationnaire et on en déduit une caractérisation du degré de non stationnarité

Remarque: La première partie a fait l'objet d'une publication dans la série des Documents de Travail de l'INSEE. De sorte que dans les parties 2 et 3 la mention : "Accardo (1990)" fait référence à cette partie 1.

PARTIE 1 : PROCESSUS STATIONNAIRES UNIDIMENSIONNELS

RESUME

Dans cette partie, nous présentons divers résultats concernant les valeurs propres de la matrice de covariance d'un processus stationnaire : détermination exacte ou approchée de ces valeurs propres, étude asymptotique de leur distribution empirique, propriétés limites concernant la valeur propre maximale.

ABSTRACT

In this part, we describe several results on eigenvalues of the covariance matrix of a stationary process : exact or approximate determination of the eigenvalues, asymptotic study of their empirical distribution, limit properties of the maximal eigenvalue.

I - INTRODUCTION

Les séries chronologiques qu'elles soient univariées ou multivariées, stationnaires ou non sont habituellement analysées à l'aide de modèles et de techniques linéaires. On commence en général par retirer l'évolution moyenne : ceci est réalisé en retirant la moyenne temporelle de la série dans le cas stationnaire, en différenciant la série ou en retranchant une tendance polynômiale estimée dans le cas non stationnaire. L'étude porte alors sur les covariances entre les séries résiduelles. Si $X_1, \ldots, X_T$ désignent les observations de cette série résiduelle, on détermine une matrice de covariance empirique temporelle, en se donnant une valeur n et en calculant :

$$\hat{\Gamma}_n = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} X_t^2 & \cdots & \cdot & \frac{1}{T}\sum_{t=1}^{T} X_t\, X_{t-n+1} \\ \frac{1}{T}\sum_{t=1}^{T} X_t\, X_{t-1} & & & \cdot \\ & & & \cdot \\ & & & \cdot \\ & & & \cdot \\ \frac{1}{T}\sum_{t=1}^{T} X_t\, X_{t-n+1} & \cdots & \frac{1}{T}\sum_{t=1}^{T} X_t\, X_{t-1} & \frac{1}{T}\sum_{t=1}^{T} X_t^2 \end{bmatrix},$$

Dans les cas suffisamment réguliers cette matrice devrait constituer un estimateur convergent de la matrice de covariance théorique du processus, qui dans le cas où X est stationnaire est donnée par :

$$\Gamma_n = \begin{bmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \gamma(n-1) & \cdots & \gamma(0) \end{bmatrix},$$

avec $\gamma(h) = E(X_t X_{t-h}) = \mathrm{Cov}(X_t, X_{t-h})$.

La matrice $\hat{\Gamma}_n$ sert de base au calcul des principales statistiques : autocorrélations partielles estimées, périodogramme... intervenant dans l'étude inferentielle de la série. Ces diverses statistiques sont en fait très liées aux valeurs propres de la matrice $\hat{\Gamma}_n$. Comme la matrice est symétrique positive, ces valeurs propres $\hat{\lambda}_\upsilon^{(n)}$, $\upsilon = 1, \ldots n$, $n \geq 1$ sont toutes réelles positives et devraient constituer des approximations des valeurs propres $\lambda_\upsilon^{(n)}$, $\upsilon = 1, \ldots, n$, $n \geq 1$ de la matrice Γ_n. De plus si le nombre de valeurs propres augmente avec la taille de la matrice n, on peut penser que la distribution de ces valeurs propres présente une certaine

"stabilité, lorsque la taille n augmente. Dans cet article nous nous proposons de faire une étude détaillée de ces valeurs propres dans le cas d'un processus X stationnaire.

Dans le paragraphe II, nous commençons par étudier le cas d'un processus moyenne mobile d'ordre 1, seul cas où les valeurs propres $\lambda_\upsilon^{(n)}$ de la matrice de covariance théorique peuvent être explicitées sous forme analytique. Nous donnons leur expression, explicitons les vecteurs propres et étudions la distribution empirique de ces valeurs propres.

De plus une étude des valeurs propres peut avoir un autre intérêt, cette fois numérique. Ainsi l'utilisation de la méthode d'estimation par le maximum de vraisemblance ou le calcul de certaines prévisions peuvent nécessiter l'inversion de matrices de covariance de grande taille (souvent égale au nombre d'observations). Divers algorithmes ont été proposés dans la littérature pour approcher une telle matrice inverse [voir e.g. Akaïke (1973) ou Ali (1977)]. Déterminer des approximations des valeurs propres $\lambda_\upsilon^{(n)}$ peut permettre de construire des matrices approchées de la matrice de covariance, matrices approchées dont certaines sont peut être facilement inversibles.

Le cas particulier du processus moyenne mobile d'ordre 1, nous permet alors un calcul approché de la distribution empirique des valeurs propres d'un processus ARMA (autorégressif-moyenne mobile) quelconque. On montre alors que cette distribution empirique tend en loi vers la transformée de la loi uniforme sur $(-\pi, \pi)$ par l'application densité spectrale. On déduit comme corollaire l'égalité de *Kolmogorov* pour les processus ARMA.

Dans le cas des processus moyennes mobiles purs ou autorégressifs purs, on propose alors des encadrements pour les valeurs propres elles-mêmes et pas seulement pour le résumé que constitue la distribution empirique.Divers résultats concernant la vitesse de convergence de la distribution empirique vers sa limite sont ensuite décrits.

Dans le paragraphe III, nous proposons une analyse directe du comportement asymptotique de la distribution empirique des valeurs propres $\lambda_\upsilon^{(n)}$, $\upsilon = 1,\ldots, n$, lorsque n tend vers l'infini. C'est Toeplitz qui le premier [Toeplitz (1911)] a montré que les "valeurs propres" d'une matrice de Toeplitz infinie $T = (c_{\upsilon-\mu},\ \upsilon, \mu \in \mathbb{Z})$ coïncidaient avec les valeurs de la série $\sum_{-\infty}^{+\infty} c_n \exp inw$, $-\pi \leq w \leq \pi$ (quand elle converge). En d'autres termes, il établissait une relation étroite entre les "valeurs propres" d'une matrice de Toeplitz et la "densité spectrale" associée. Quelques années plus tard, Szegö (1921) prouvait le théorème correspondant pour les matrices finies. Plus précisément si f est une fonction bornée sur $(-\pi, \pi)$,

si $M_n(f)$ désigne la matrice de carrée de taille n, d'élément $m_{kl} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp[i(k-l)w] f(w)dw$, si $\lambda_\upsilon^{(n)}(f)$, $\upsilon = 1,\dots n$ sont les valeurs propres de cette matrice, il établit que :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} h\ [\lambda_\upsilon^{(n)}(f)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} h[f(w)]\ dw,$$

pour toute fonction h continue. Autrement dit, il montre la convergence en loi de la distribution empirique des valeurs propres vers la transformée par f de la loi uniforme sur l'intervalle $(-\pi, \pi)$. Nous commençons par rappeler la démonstration de ce résultat, puis nous l'étendons à divers cas interessants pour l'application aux séries temporelles : cas où la densité spectrale n'est pas bornée, ce qui permet d'inclure des catégories de processus comme les processus fractionnaires et constitue une première étape vers l'analyse de la non stationnarité. De façon analogue, nous considérons le cas de processus stationnaires admettant une mesure spectrale avec une partie discrète et vérifions que la loi limite est modifiée et comporte alors une "masse à l'infini".

II - ETUDE DES VALEURS PROPRES DANS LE CAS DE PROCESSUS ARMA

II.a - Processus Moyenne Mobile d'ordre 1

Le cas fondamental est celui des processus moyenne-mobile d'ordre 1 ; en effet les valeurs propres sont alors explicitement calculables et elles serviront pour approcher les valeurs propres associées à des représentations ARMA plus complexes.

Considérons donc un processus MA(1) : $X = (X_t,\ t \in \mathbb{Z})$ défini par :

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}, \tag{2.1}$$

où $\varepsilon = (\varepsilon_t,\ t \in \mathbb{Z})$ est un bruit blanc de variance 1 et où le paramètre θ est un réel non nul, $\theta \neq 0$. La densité spectrale du processus est $f(w) = |1 - \theta\exp iw|^2$, $-\pi \leq w \leq \pi$ [on prendra garde que nous retenons comme définition de la densité spectrale $f(w) = \sum_k \gamma(h) \exp(ihw)$, où $\gamma(h)$ désigne l'autocovariance d'ordre h ; il y a donc un rapport de 2π avec certaines définitions souvent utilisées en séries temporelles].

La matrice de covariance d'ordre n est :

$$(2.2) \qquad \Gamma_n(\theta) = \begin{bmatrix} 1+\theta^2 & -\theta & \cdots & 0 \\ -\theta & 1+\theta^2 & & \vdots \\ \vdots & & & -\theta \\ 0 & \cdots & -\theta & 1+\theta^2 \end{bmatrix}.$$

Proposition 2.3 : Les valeurs propres $\lambda_\upsilon^{(n)}$, $\upsilon = 1,\ldots, n$ de la matrice $\Gamma_n(\theta)$ sont :

$$\lambda_\upsilon^{(n)} = 1 + \theta^2 - 2\theta \cos \frac{\upsilon\pi}{n+1}, \quad \upsilon = 1,\ldots, n.$$

Des vecteurs propres correspondants sont :

$$\ell_\upsilon^{(n)} = \sqrt{\frac{2}{n+1}} \left[\sin \frac{\mu\upsilon\pi}{n+1}\right]_{\mu=1,\ldots,n}, \quad \upsilon = 1,\ldots, n.$$

Preuve :

i) Nous présentons le calcul pour $\theta > 0$, le cas $\theta < 0$ s'en déduisant facilement. D'après l'annexe 1.A, on sait que les valeurs propres se trouvent dans l'intervalle [Min f, Max f] = $[(1 - \theta)^2, (1 + \theta)^2]$. On peut donc poser : $\lambda = 1 + \theta^2 + 2\theta \cos \alpha$, $-\pi \le \alpha \le \pi$.

Dans ces conditions, on a :

$$\det (\Gamma_n(\theta) - \lambda \mathrm{Id}_n] = (-1)^n \theta^n \begin{bmatrix} 2\cos\alpha & 1 & 0 \\ \vdots & & \\ 1 & 2\cos\alpha & 1 \\ 0 & 1 & 2\cos\alpha \end{bmatrix}$$

$$= (-1)^n \theta^n \mathcal{U}_n(\cos \alpha),$$

où $\mathcal{U}_n$ désigne le polynôme de Tchebychev normalisé par $\mathcal{U}_n(1) = n + 1$ [voir Abramovitz-Stegun (1985)].

Une relation de récurrence linéaire sur le polynôme de Tchebychev est facilement obtenue en développant le déterminant.

On a : $\mathcal{U}_n(\cos\alpha) = (2\cos\alpha)\, \mathcal{U}_{n-1}(\cos\alpha) - \mathcal{U}_{n-2}(\cos\alpha)$, , $n \ge 2$
$\mathcal{U}_1(\cos\alpha) = 2\cos\alpha$, $\mathcal{U}_0(\cos\alpha) = 1$.

Résolvant l'équation caractéristique associée, on en déduit :

$$\mathcal{U}_n(\cos\alpha) = \frac{\sin(n + 1)\alpha}{\sin\alpha}, \quad -\pi \le \alpha \le \pi.$$

Par conséquent $\lambda = 1 + \theta^2 + 2\theta \cos \alpha$ est valeur propre, si $\alpha = \frac{\upsilon\pi}{n+1}$, $\upsilon = 1, \ldots, n$.

ii) On vérifie par ailleurs facilement que le vecteur $\ell_\upsilon^{(n)}$ donné dans la proposition 2.3 est bien vecteur propre.

QED

Remarque 2.4 : On peut noter que les vecteurs propres sont indépendants de la valeur du paramètre θ. En particulier les matrices $(\Gamma_n(\theta),\ \theta \in]-1, 1[)$ commutent entre elles. Ceci peut d'ailleurs se voir directement en remarquant que :

$$\Gamma_n(\theta) = (1 + \theta^2)\ Id_n - \theta J_n,$$

$$\text{où : } J_n = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ 1 & & & \vdots \\ & & & \vdots \\ & & & \vdots \\ \vdots & & & 1 \\ \vdots & & & \\ \vdots & & & \\ 0 & \cdots & 1 & 0 \end{bmatrix}$$

Lorsque θ est positif, les valeurs propres $\lambda_\upsilon^{(n)} = 1 + \theta^2 - 2\theta \cos \frac{\upsilon\pi}{n+1}$ sont rangées par ordre croissant. On peut alors déterminer facilement la fonction de répartition empirique de ces valeurs propres. Elle est définie par :

$$(2.5) \qquad \mathcal{F}_n(\xi) = \frac{\text{Card } \{\upsilon : \lambda_\upsilon^{(n)} < \xi\}}{n}.$$

Pour $0 \leq \theta \leq 1$, on remarque que :

si $\xi \leq \underset{w}{\text{Min}}\ |f(w)| = (1 - \theta)^2$, alors $\mathcal{F}_n(\xi) = 0$,

si $\xi \geq \underset{w}{\text{Max}}\ |f(w)| = (1 + \theta)^2$, alors $\mathcal{F}_n(\xi) = 1$.

Pour les valeurs intermédiaires : $(1 - \theta)^2 \leq \xi \leq (1 + \theta)^2$, on a :

$$1 + \theta^2 - 2\theta \cos \frac{\upsilon\pi}{n+1} < \xi \iff 1 \leq \upsilon < \frac{n+1}{\pi} \text{Arc cos} \left(\frac{1 + \theta^2 - \xi}{2\theta}\right) \text{ et donc :}$$

$$\mathcal{F}_n(\xi) = \frac{n+1}{n} \frac{1}{\pi} \text{Arc cos} \left(\frac{1 + \theta^2 - \xi}{2\theta}\right).$$

Cette distribution empirique converge en loi, lorsque n tend vers l'infini, vers la fonction de répartition limite :

$$\mathcal{F}_\infty(\xi) = \begin{cases} 0, \text{ si } \xi \leq (1-\theta), \\ \frac{1}{\pi} \text{Arc cos} \left(\frac{1 + \theta^2 - \xi}{2\theta} \right), \text{ si } (1-\theta)^2 \leq \xi \leq (1+\theta)^2, \\ 1, \text{ si } \xi \geq (1+\theta)^2. \end{cases}$$

On peut noter que cette fonction de répartition n'est autre que celle de la loi transformée de la loi uniforme sur $(-\pi, \pi)$ par l'application densité spectrale : $f(w) = |1 - \theta \exp iw|^2 = 1 + \theta^2 - 2\theta \cos w$. C'est un cas particulier d'une situation générale étudiée dans le paragraphe suivant.

II.b - Extension aux processus ARMA

Le but est ici d'étendre le résultat précédent à tous les processus ARMA stationnaires, autrement dit de relier la fonction de répartition empirique des valeurs propres et la densité spectrale. On commence par introduire une approximation de la matrice de covariance d'un processus ARMA au moyen de matrices de covariance de processus MA(1). La qualité de l'approximation sera évaluée en utilisant deux types de normes. Si A_n est une matrice carrée de taille n, nous notons : $\|A_n\| = \sup_x \frac{\|A_n x\|}{\|x\|}$ la norme usuelle ; nous introduisons également une norme prenant en compte la taille n :

$$|A_n| = \sqrt{\frac{1}{n} \sum_{i,j} (a_{ij})^2}.$$

Propriété 2.6 : Soit $X = (X_t, t \in \mathbb{Z})$ un processus stationnaire admettant la représentation ARMA (p, q) :

$$\prod_{i=1}^{p} (1 - \varphi_i L) X_t = \prod_{j=1}^{q} (1 - \theta_j L) \varepsilon_t, \; V\varepsilon_t = 1,$$

où L désigne l'opérateur retard et où les coefficients φ_i sont tels que $|\varphi_i| < 1$, $i = 1, \ldots, p$.

Si Γ_n^X désigne la matrice de covariance d'ordre n du processus X, on a :

$$\Gamma_n^X = \prod_{i=1}^{p} [\Gamma_n(\varphi_i)]^{-1} \prod_{j=1}^{q} \Gamma_n(\theta_j) + A_n,$$

où $\Gamma_n(\theta)$ désigne la matrice de covariance d'un processus MA(1) de paramètre θ et où la suite A_n est telle que : $\|A_n\| \leq K < +\infty$, $\forall n$ et

$\lim_n |A_n| = 0.$

Preuve :

i) Nous commençons par regarder comment est transformée la matrice de covariance d'un processus stationnaire, lorsqu'on transforme celui-ci par un opérateur du type $1 - \varphi L$.

Soit donc $X = (X_t, t \in \mathbb{Z})$ un processus stationnaire, de fonction d'autocovariance $\gamma_X(h)$, $h \in \mathbb{Z}$. Sa matrice de covariance d'ordre n est :

$$\Gamma_n^X = \begin{bmatrix} \gamma_X(0) & \cdots & \gamma_X(n-1) \\ \vdots & & \vdots \\ \gamma_X(n-1) & & \gamma_X(0) \end{bmatrix}.$$

Le processus transformé $Y = [Y_t, t \in \mathbb{Z}]$ défini par :

$$Y_t = (1 - \varphi L) X_t = X_t - \varphi X_{t-1},$$

a pour fonction d'autocovariance :

$$\gamma_Y(h) = -\varphi \gamma_X(h-1) + (1+\varphi^2) \gamma_X(h) - \varphi \gamma_X(h+1).$$

La matrice de covariance d'ordre n du processus transformé Y est alors liée à celle du processus initial x par :

$$\Gamma_n(\varphi) \Gamma_n^X = \Gamma_n^Y + \varphi E_n^X, \tag{2.7}$$

où $\Gamma_n(\varphi)$ désigne la matrice de covariance d'un processus MA(1) de paramètre φ et où E_n^X est la matrice :

$$E_n^X = \begin{bmatrix} \gamma_X(1) & \cdots & \gamma_X(n) \\ 0 & \cdots & 0 \\ \vdots & & \\ 0 & & 0 \\ \gamma_X(n) & \cdots & \gamma_X(1) \end{bmatrix}.$$

ii) Nous pouvons alors appliquer récursivement la formule (2.7) pour obtenir une approximation de la matrice de covariance d'un processus ARMA général. Considérons donc un tel processus $X = (X_t, t \in \mathbb{Z})$ satisfaisant l'équation de récurrence :

$$\prod_{i=1}^{p} (1 - \varphi_i L)\, X_t = \prod_{j=1}^{q} (1 - \theta_j L)\, \varepsilon_t .$$

Posons alors :

$$z_t^j = \prod_{i=j+1}^{p} (1 - \varphi_i L) X_t,\ j = 1, \ldots, p-1,\ z_t^p = X_t,$$

$$w_t^\ell = \prod_{j=\ell+1}^{q} (1 - \theta_j L)\varepsilon_t,\ \ell = 1, \ldots, q-1,\ w_t^q = \varepsilon_t ;$$ appliquant récursivement la formule (2.7), on obtient :

$$(2.8) \qquad \left[\prod_{i=1}^{p} \Gamma_n(\varphi_i) \right] \Gamma_n^x - \sum_{k=1}^{p} \varphi_k \left[\prod_{j=1}^{k-1} \Gamma_n(\varphi_j) \right] E_n^{z^k}$$

$$= \prod_{j=1}^{q} \Gamma_n(\theta_j) - \sum_{\ell=1}^{q} \theta_\ell \left[\prod_{m=1}^{\ell-1} \Gamma_n(\theta_m) \right] E_n^{w^\ell}.$$

Cette formule fait clairement apparaître la forme de la matrice A_n.

iii) Il nous reste donc seulement à vérifier que cette matrice satisfait les conditions énoncées dans la propriété. Pour cela nous allons établir plusieurs majorations uniformes.

a) Toute valeur propre de $\Gamma_n(\varphi_i)$ appartient (voir annexe 1.A) à l'intervalle $[(1 - |\varphi_i|)^2, (1 + |\varphi_i|)^2]$. On a donc :

$\|\Gamma_n(\theta_j)\| \leq \operatorname{Max}_j (1 + |\theta_j|)^2$ et $\|\Gamma_n(\varphi_i)^{-1}\| \leq \operatorname{Max}_i \dfrac{1}{(1 - |\varphi_i|)^2}$ ainsi les matrices $\Gamma_n(\theta_j)$, $j = 1 \ldots q$ et les matrices $\Gamma_n(\varphi_i)^{-1}$ $i = 1, \ldots p$, sont toutes uniformément bornées pour la norme $\|.\|$.

b) De plus les processus intermédiaires z^d et w^ℓ sont stationnaires. Posons alors :

$$K = \sup \left\{ \sup_j \sum_{k=0}^{\infty} (\gamma_{z^j}(h))^2 \ , \sup_\ell \sum_{h=0}^{\infty} (\gamma_{w^\ell}(h))^2 \right\}.$$

Cette valeur K est finie. De plus, on a :

$$|E_n^{z^j}| = \left\{ \frac{1}{n}\, 2 \sum_{h=1}^{n} (\gamma_{z^j}(h))^2 \right\}^{\frac{1}{2}} \leq \frac{\sqrt{2K}}{\sqrt{n}},$$

et de même $|E_n^{w^\ell}| \leq \dfrac{\sqrt{2K}}{\sqrt{n}}$.

On en déduit que : $\lim_{n\to\infty} |E_n^{z^j}| = 0$ et $\lim_{n\to\infty} |E_n^{w^\ell}| = 0$.

c) Enfin, on a, d'après l'inégalité de Schwarz :

$$\|E_n^{z^j}\| \leq \left\{ \sum_{h=1}^{n} (\gamma_{z^j})(h))^2 \right\}^{1/2} \leq K ;$$

de même $\|E_n^{w^\ell}\| \leq K$.

L'erreur d'approximation est donnée par :

$$A_n = \left[\prod_{i=1}^{p} \Gamma_n(\varphi_i)^{-1} \right] \left\{ \sum_{k=1}^{p} \varphi_k \left[\prod_{j=1}^{k-1} \Gamma_n(\varphi_j) \right] E_h^{z^k} - \sum_{\ell=1}^{p} \theta_\ell \left[\prod_{m=1}^{\ell-1} \Gamma_n(\theta_m) \right] E_n^{w^\ell} \right\}$$

Utilisant les propriétés $\|A + B\| \leq \|A\| + \|B\|$,
$\|AB\| \leq \|A\| \, \|B\|$,

et le fait que les diverses matrices intervenant dans le membre de droite sont uniformément bornées en norme $\|.\|$, on en déduit que $\|A_n\|$ est également uniformément borné en n.

Par ailleurs, il est montré dans l'annexe 1.B. Proposition 6 que :

$$|AB| \leq \|A\| \, |B|.$$

On obtient alors immédiatement que $\lim_n |A_n| = 0$.

QED

Comme les matrices de covariance Γ_n^x du processus ARMA peuvent être approchées par les matrices $\left[\prod_{i=1}^{p} \Gamma_n(\varphi_i) \right]^{-1} \left[\prod_{j=1}^{q} \Gamma_n(\theta_j) \right]$, on peut raisonnablement penser que les valeurs propres de Γ_n^x peuvent l'être par celle de l'approximation. Ces dernières sont évidemment faciles à déterminer, puisque nous avons remarqué que les matrices $\Gamma_n(\theta)$ admettent des vecteurs propres indépendants de la valeur du paramètre θ. C'est cette idée qui va permettre de montrer le théorème ci-dessous.

Propriété 2.9 : Soient λ_υ^n, $\upsilon = 1, \ldots, n$, les valeurs propres de la matrice de covariance d'un processus ARMA et f la densité spectrale de ce processus. Alors pour toute fonction h continue sur (Min f, Max f) :

$$\lim_n \frac{1}{n} \sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(f(w))dw.$$

Preuve :

i) Notons β_υ^n, $\upsilon = 1,\ldots, n$, les valeurs propres de l'approximation $\left[\prod_{i=1}^{p} \Gamma_n(\varphi_i) \right]^{-1} \left[\prod_{j=1}^{q} \Gamma_n(\theta_j) \right]$. Comme les matrices $\Gamma_n(\theta)$ ont même vecteurs propres, nous avons :

$$\beta_\upsilon^n = \left[\prod_{i=1}^{p} (1 + \varphi_i^2 - 2\varphi_i \cos \frac{\upsilon\pi}{n+1}) \right]^{-1} \prod_{j=1}^{q} (1 + \theta_j^2 - 2\theta_j \cos \frac{\upsilon\pi}{n+1}),$$

ou encore : $\beta_\upsilon^n = f\left(\frac{\upsilon\pi}{n+1}\right)$, $\upsilon = 1\ldots n$.

Par conséquent si h est continue sur [Min f, Max f], $\frac{1}{n} \sum_{\upsilon=1}^{n} h(\beta_\upsilon^n) = \frac{1}{n} \sum_{\upsilon=1}^{n} h\left[f\left(\frac{\upsilon\pi}{n+1}\right) \right]$ apparaît sous la forme d'une somme de Riemann. Comme h(f) est continue, il y a donc convergence vers l'intégrale correspondante :

$$\lim_n \frac{1}{n} \sum_{\upsilon=1}^{n} h(\beta_\upsilon^n) = \frac{1}{\pi} \int_0^{\pi} h[f(w)]dw$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} h[f(w)]dw,$$

car la densité spectrale est paire.

ii) Il reste maintenant à montrer que cette limite est aussi celle obtenue pour $\frac{1}{n} \sum_{\upsilon=1}^{n} h(\lambda_\upsilon^n)$. Ce résultat est une conséquence directe de la proposition 7 de l'annexe 1. La matrice $T_n = \Gamma_n^x$ est en effet de Toeplitz, l'approximation $L_n = \left[\prod_{i=1}^{p} \Gamma_n(\varphi_i) \right]^{-1} \prod_{j=1}^{q} \Gamma_n(\theta_j)$ est hermitienne, L_n et T_n sont uniformément bornées en norme $\|.\|$ et l'écart $A_n = T_n - L_n$ est tel que $\lim_n |A_n| = 0$. Toutes les hypothèses de la proposition 7 sont donc réunies et la distribution empirique des valeurs propres de λ_υ^n a même limite en loi que la distribution empirique des valeurs propres β_υ^n.

Q.E.D.

La démonstration que nous venons de présenter peut être vue comme une preuve constructive du théorème de Szegö dans le cas particulier des

processus ARMA. Une démonstration très différente de ce théorème sera décrite dans le paragraphe III. En particulier dans notre approche nous n'utilisons pas l'étape constituée par l'égalité de Kolmogorov ; cette dernière apparait comme le cas particulier correspondant à la fonction logarithme ; on a :

$$\lim_n \frac{1}{n} \sum_{\upsilon=1}^{n} \text{Log } \lambda_\upsilon^n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{log} f(w) dw$$

$$\Longleftrightarrow \lim_n \frac{1}{n} \text{ Log det } \Gamma_n^x = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log} f(w) dw.$$

On sait par ailleurs (voir Azencott-Dacunha-Castelle (1983) que le premier membre $\lim_n \frac{1}{n}$ Log det Γ_n^x est égal à la variance de l'innovation et on en déduit donc l'égalité de Kolmogorov : Log $\sigma^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi}$ Log f(w) dw.

La propriété 2.9 fournit la limite en loi de la distribution empirique des valeurs propres λ_υ^n, $\upsilon = 1 \ldots n$. Pour exhiber cette limite, il suffit d'appliquer la formule de transfert. On a :

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} h[f(w)] dw = \int_{(-\pi,\pi)} h[f(w)] \; d\mathcal{U}(w)$$

$$= \int_{[-\pi,\ \pi]} h(x) \; d\mathcal{U}^f(w)$$

où $\mathcal{U}_{[-\pi,\ \pi]}$ désigne la loi uniforme sur $[-\pi, \pi]$ et $\mathcal{U}^f_{[-\pi,\ \pi]}$ sa transformée par la densité spectrale. Ainsi la propriété 2.9 équivaut à la convergence en loi de la distribution empirique vers la loi $\mathcal{U}^f_{[-\pi,\ \pi]}$.

II.c - Encadrement des valeurs propres dans le cas de processus moyennes mobiles ou autorégressifs purs.

Dans le paragraphe précédent, nous avons fait une étude globale de l'ensemble des valeurs propres, étude essentiellement asymptotique. Dans ce paragraphe nous nous proposons de donner des propriétés des valeurs propres prises séparément, propriétés valables pour toutes les valeurs de n.

En dehors du cas MA(1) présenté en II.a, le calcul analytique des valeurs propres parait impossible. Dans le cas AR(1), Grenander et Szegö (1958) et Stroeker (1978), (1983) obtiennent par des méthodes sensiblement différentes des approximations des valeurs propres, mais leurs calculs ne sont praticables qu'à cause de la relative simplicité des matrices de covariance des AR(1).

Par une méthode distincte de celles introduites précédemment dans la littérature, on donne ici des approximations des valeurs propres de

processus MA(q), $q \geq 1$ et AR(p), $p \geq 1$. On établit au préalable un résultat général qui fournit des encadrements des valeurs propres à partir d'approximations d'opérateurs [voir GLAZMAN et LIUBITCH (1974)].

Proposition 2.10 : Soient P un projecteur orthogonal de L_{n+1} (dim $L_{n+1} = n + 1$) sur un sous-espace L_n (dim $L_n = n$) et S un opérateur autoadjoint de L_{n+1}.

On définit la restriction $\tilde{S}$ de PS au sous-espace L_n. $\tilde{S}$ est un opérateur autoadjoint de L_n. De plus notant λ_k^n, $k = 0,\ldots, n$ et $\tilde{\lambda}_k^n$, $k = 0,\ldots n-1$ les valeurs propres de S et $\tilde{S}$ rangées par ordre croissant, on a :

$$\lambda_0 \leq \tilde{\lambda}_0 \leq \lambda_1 \leq \tilde{\lambda}_1 \leq \ldots\ldots \leq \lambda_{n-1} \leq \tilde{\lambda}_{n-1} \leq \lambda_n .$$

Preuve :

1) Les restrictions de PS et de PSP au sous-espace L_n sont clairement confondues. Cette remarque montre que $\tilde{S}$ est autoadjoint.

2) Commençons la comparaison des valeurs propres par celles des plus petites λ_0 et $\tilde{\lambda}_0$. On a (voir annexe 1) :

$$\lambda_0 = \min_{x \in L_{n+1},\, \|x\| = 1} x'Sx \leq \min_{x \in L_n,\, \|x\|=1} x'Sx = \min_{y \in L_n;\, \|y\|=1} y'PSPy = \tilde{\lambda}_0 .$$

3) Plus généralement, notons $\mathcal{V}_k$ l'ensemble des sous espaces V_k de L_{n+1} tels que dim $V_k \geq n + 1 - k$ et $\mathcal{U}_k$ l'ensemble des sous-espaces U_k de L_n tels que dim $U_k \geq n - k$.

D'après la caractérisation Max-Min des valeurs propres proposée par Weyl-Courant [voir Kobilinsky (1979)], on a :

$$\lambda_k = \max_{\mathcal{V}_k} \min_{x \in V_k} x'Sx .$$

$$\text{On a alors : } \lambda_k \leq \max_{\mathcal{V}_k} \min_{x \in V_k \cap L_n} x' S x .$$

Comme dim $V_k \cap L_n \geq n - k$, on voit que : $V_k \cap L_n \in \mathcal{U}_k$. Introduisant le maximum sur l'ensemble $\mathcal{U}_k$, nous avons :

$$\lambda_k \leq \max_{\mathcal{V}_k} \min_{x \in V_k \cap L_n} x'Sx \leq \max_{\mathcal{U}_k} \min_{y \in \mathcal{U}_k} y' S y = \tilde{\lambda}_k ,\ 1 \leq k \leq n$$

4) Par ailleurs : $\lambda_{k+1} = \max_{\mathcal{V}_{k+1}} \min_{x \in V_{k+1}} x'Sx$, et donc en se restreignant aux sous-espaces V_{k+1} contenus dans L_n :

$$\lambda_{k+1} \quad \underset{V_{k+1} \in \mathcal{V}_{k+1}}{\text{Max}} \quad \underset{x \in V_{k+1}}{\text{Min}} \quad x' S x.$$
$$V_{k+1} \subset L_n$$

Comme : $\dim V_{k+1} \geq n - (k+1) + 1 = n - k$, il y a identité entre les sous-espaces V_{k+1} inclus dans L_n et les éléments de $\mathcal{U}_k$. Par conséquent nous avons :

$$\lambda_{k+1} \geq \underset{\mathcal{U}_k}{\text{Max}} \ \underset{y \in \mathcal{U}_k}{\text{Min}} \ y'S\,y = \tilde{\lambda}_k, \quad k = 1, \ldots, n-1.$$

5) Mettant ensemble les majorations obtenues en 2), 3), 4), nous obtenons : $\tilde{\lambda}_0 \leq \tilde{\lambda}_0 \leq \lambda_1 \leq \tilde{\lambda}_1 \leq \ldots \leq \tilde{\lambda}_{n-1} \leq \lambda_n$.

Q.E.D.

On obtient alors un corollaire qui relie l'approximation des matrices et celles des valeurs propres.

Proposition 2.11 : Soient $(B_n)_{n\geq 1}$ et $(C_n)_{n\geq 1}$ deux suites de matrices symétriques, où les matrices C_n sont de la forme :

$$C_n = \begin{bmatrix} C_1^{(n)} & & 0 \\ & 0 & \\ 0 & & C_2^{(n)} \end{bmatrix}$$

avec $C_1^{(n)}$, $C_2^{(n)}$ de dimension r x r indépendante de n.

Posons : $A_n = B_n + C_n$, $n \geq 1$ et désignons par $\lambda_\upsilon^{(n)}(A)$, $\upsilon = 1, \ldots, n$ [resp $\lambda_\upsilon^{(n)}(B)$, $\upsilon = 1, \ldots, n$] les valeurs propres rangées par ordre croissant de A_n [resp de B_n]. On a alors pour $2r+1 \leq \upsilon \leq n - 2\upsilon$, les inégalités :

$$\lambda_{\upsilon-2r}^{(n)}(B) \leq \lambda_\upsilon^{(n)}(A) \leq \lambda_{\upsilon+2r}^{(n)}(B),$$

$$\lambda_{\upsilon-2r}^{(n)}(A) \leq \lambda_\upsilon^{(n)}(B) \leq \lambda_{\upsilon+2r}^{(n)}(A).$$

Preuve : Introduisons le sous-espace de $\mathbb{R}^n$ défini par :

$L = \{0\}^r \times \mathbb{R}^{n-2r} \times \{0\}^r$ et p le projecteur orthogonal sur L. Les restrictions de PA_n et de PB_n au sous-espace L sont confondues. Soient alors $\alpha_1 \leq \alpha_2 \ldots \leq \alpha_{n-2r}$ les n - 2r valeurs propres de cette restriction appliquant 2r fois consécutives la proposition 2.10, on voit que :

$$\lambda_\upsilon^{(n)}(A) \leq \alpha_\upsilon \leq \lambda_{\upsilon+2r}^{(n)}(A)$$

$$\lambda_\upsilon^{(n)}(B) \leq \alpha_\upsilon \leq \lambda_{\upsilon+2r}^{(n)}(B), \quad \upsilon = 1, \ldots, n-2r.$$

Combinant ces deux séries d'inégalités, on obtient celles annoncées dans la proposition 2.11.

Q.E.D.

On peut maintenant utiliser la proposition précédente pour approcher les valeurs propres de la matrice covariance d'un MA(q), $q \geq 1$.

Propriété 2.12 : Soient $X_t = \prod_{j=1}^{q} (1 - \theta_j L)\varepsilon_t$ un processus MA(q), Γ_n^X sa matrice de covariance d'ordre n, $\lambda_1^{(n)} \leq \ldots \leq \lambda_n^{(n)}$ ses valeurs propres rangées par ordre croissant. Soient $\mu_1^{(n)} \leq \ldots \leq \mu_n^{(n)}$ les valeurs propres de la matrice (symétrique) $\prod_{j=1}^{q} \Gamma_n(\theta_j)$.

Alors : $\mu_{\upsilon-2q}^{(n)} \leq \lambda_\upsilon^{(n)} \leq \mu_{\upsilon+2q}^{(n)}$, $\quad \upsilon = 2q + 1, \ldots, n - 2q.$

Preuve :

i) On sait que :

$$\Gamma_n^X = \prod_{j=1}^{q} \Gamma_n(\theta_j) - \sum_{\ell=1}^{q} \theta_\ell \left[\prod_{m=1}^{\ell-1} \Gamma_n(\theta_m)\right] E_n^{W\,\ell} = \prod_{j=1}^{q} \Gamma_n(\theta_j) + A_n.$$

Nous allons montrer en utilisant deux lemmes que la matrice A_n est de la forme : $A_n = \begin{bmatrix} C_n & 0 \\ 0 & D_n \end{bmatrix}$, où C_n et D_n sont carrées de tailles q x q.

ii) La fonction d'autocovariance d'un processus moyenne mobile pur s'annule à partir d'un certain rang, égal à q + 1 où q est l'ordre de la moyenne mobile. Les matrices de covariance associées ont donc des éléments non nuls sur un nombre fini de paradiagonales parallèles à la diagonale principale. Ceci nous conduit à introduire la notion de matrice M à bande diagonale de largeur H. Il s'agit d'une matrice d'élément général m_{ij} tel que : $m_{ij} = 0$, si $|i-j| > H$ et m_{ij} non tous nuls, si $|i-j| = H$.

Lemme 2.13 : Si M_1 et M_2 sont deux matrices à bandes diagonales de largeurs respectives H_1 et H_2, la matrice $M_1 M_2$ est à bande diagonale de largeur au plus $H_1 + H_2$.

Preuve immédiate.

Lemme 2.14 : Si M est une matrice à bande diagonale de largeur H, si W est

un processus MA(q) et E_n^W la matrice associée : $E_n^W = \begin{bmatrix} \gamma_W(1) & \cdots & \gamma_W(n) \\ & 0 & \\ \gamma_W(n) & \cdots & \gamma_W(1) \end{bmatrix}$,

alors le produit ME_n^W est de la forme $\begin{bmatrix} C_n & 0 \\ 0 & D_n \end{bmatrix}$, où les matrices C_n et D_n sont de tailles H x q.

La preuve de ce lemme est immédiate dès que l'on a noté que la matrice E_n^W a la forme spécifique :

$$E_n^W = \begin{bmatrix} \gamma_W(1) & \cdots & \gamma_W(q) & 0 & \cdots & 0 \\ & & 0 & & & \\ 0 & \cdots & & 0\gamma_W(q) & \cdots & \gamma_W(1) \end{bmatrix}$$

iii) Utilisant le lemme 2.13, nous voyons que les matrices : $\prod_{m=1}^{\ell-1} \Gamma_n(\theta_m)$, $\ell = 1,\ldots, q$ sont des matrices à bandes diagonales de largeurs au plus q. Utilisant alors le lemme 2.14 nous en déduisons que l'erreur d'approximation :

$$A_m = \sum_{\ell=1}^{q} \theta_\ell \left(\prod_{m=1}^{\ell-1} \Gamma_n(\theta_n) \right) \text{ est de la forme } \begin{bmatrix} C_n & 0 \\ 0 & D_n \end{bmatrix},$$

où C_n et D_n sont carrées de taille q. La fin de la démonstration est alors une conséquence directe de la propriété 2.11.

Q.E.D.

Exemple 2.15 : A titre d'illustration prenons le cas d'un processus MA(2) défini par : $X_t = (1 - \theta_1 L)(1 - \theta_2 L)\varepsilon_t$, $V\varepsilon_t = 1$, avec : $0 \leq \theta_1, \theta_2 \leq 1$. Les valeurs propres de $\Gamma_n(\theta_1)\,\Gamma_n(\theta_2)$ rangées par ordre croissant sont :

$$\left[1 + \theta_1^2 - 2\theta_1 \cos \frac{\upsilon\Pi}{n+1}\right] \left[1 + \theta_2^2 - 2\theta_2 \cos \frac{\upsilon\Pi}{n+1}\right], \quad \upsilon = 1,\ldots, n$$

Elles encadrent les valeurs propres $\lambda_\upsilon^{(n)}$ de la matrice de covariance Γ_n^*. On notera que l'amplitude des encadrements est de l'ordre de $\frac{1}{n}$ et diminue, lorsque n tend vers l'infini.

Le cas des processus autorégressifs purs se déduit assez facilement du cas moyenne mobile pur. L'idée est d'utiliser le fait qu'une représentation autorégressive est "l'inverse" d'une représentation moyenne mobile. Plutôt que d'approcher directement la matrice de covariance de l'autorégressif, on cherchera donc à approcher son inverse. Soit $(X = X_t,\ t \in \mathbb{Z})$ un processus AR(p) satisfaisant la représentation :

$$\prod_{j=1}^{p} (1-\varphi_j L)X_t = \varepsilon_t,\ V\varepsilon_t = 1,\ |\varphi_j| < 1.$$

La matrice de covariance d'ordre n du processus X est alors reliée à celle du processus moyenne mobile : $Y_t = \prod_{j=1}^{p} (1 - \varphi_j L)\varepsilon_t$ par :

$$(\Gamma_n^X)^{-1} = \Gamma_n^Y + \begin{bmatrix} C_n & & 0 \\ & 0 & \\ 0 & & D_n \end{bmatrix}, \tag{2.16}$$

où C_n, D_n sont deux matrices carrées de taille p x p [voir ALI (1977)].

La formule (2.16) permet ainsi de se ramener au cas précédent.

Proposition 2.17 : Soit $X = (X_t,\ t \in \mathbb{Z})$ un processus AR(p) de représentation $\prod_{j=1}^{p} (1 - \varphi_j L)\ X_t = \varepsilon_t$, $V\varepsilon_t = 1$, $|\varphi_j| < 1$; soient Γ_n^X sa matrice de covariance d'ordre n, $\lambda_1^{(n)} \le \ldots \le \lambda_n^{(n)}$ les valeurs propres de cette matrice rangées par ordre croissant et $\mu_1^{(n)} \le \ldots \le \mu_n^{(n)}$ les valeurs propres de $\prod_{j=1}^{p} \Gamma_n(\varphi_j)$. Alors on a :

$$\mu_{\upsilon-4p}^{(n)} \le \frac{1}{\lambda_{n-\upsilon+1}^{(n)}} \le \mu_{\upsilon+4p}^{(n)},\ \upsilon = 4p + 1, \ldots, n - 4p$$

Preuve : D'après la propriété 2.12 les valeurs propres $\alpha_1^{(n)} \le \ldots \le \alpha_n^{(n)}$ de Γ_n^Y vérifient : $\mu_{\upsilon-2p}^{(n)} \le \alpha_\upsilon^{(n)} \le \mu_{\upsilon+2p}^{(n)}$; appliquant alors la propriété 2.11, on a aussi :

$$\alpha_{\upsilon-2p}^{(n)} \le \frac{1}{\lambda_{n-\upsilon+1}^{(n)}} \le \alpha_{\upsilon+2p}^{(n)}.$$

Comparant les deux encadrements, nous obtenons ceux de la propriété.

Q.E.D.

II.d - Vitesse de convergence et distributions empiriques

Nous avons vu dans le paragraphe II.c que si X est un processus ARMA la distribution empirique des valeurs propres de la matrice de covariance tend vers la loi limite $\mathcal{U}^{f}_{[-\pi,\ \pi]}$. Nous allons ici nous intéresser à la vitesse à laquelle se produit cette convergence.

Nous avions introduit diverses valeurs propres : celles $\lambda_{\upsilon}^{(n)}$ $\upsilon = 1,\ldots$ n de la matrice Γ_n^{*}, celles $\beta_{\upsilon}^{(n)}$, $\upsilon = 1,\ldots,$ n de la matrice $\left[\prod_{i=1}^{p} \Gamma_n(\varphi_i)\right]^{-1} \prod_{j=1}^{q} \Gamma_n(\theta_j)$. Les distributions empiriques qui leurs sont associées sont désignées par :

$$\mathcal{F}_n(x) = \frac{1}{n}\,\mathrm{Card}\{\upsilon : \lambda_{\upsilon}^{(n)} < x\},\quad \mathcal{G}_n(x) = \frac{1}{n}\,\mathrm{Card}\,\{\upsilon : \beta_{\upsilon}^{(n)} < x\}.$$

Nous établissons successivement des résultats concernant l'écart entre $\mathcal{G}_n$ et la distribution limite $\mathcal{F} = \mathcal{U}^{f}_{[-\pi,\ \pi]}$ puis l'écart entre $\mathcal{F}_n$ et $\mathcal{G}_n$.

Proposition 2.18 : Pour tout réel x, il existe M(x) dépendant de x, mais pas de n tel que :

$$|\mathcal{G}_n(x) - \mathcal{F}(x)| \le \frac{M(x)}{n},\quad n \ge 1.$$

De plus : $\sup_x M(x) \le 4(p + q) - 3.$

Preuve :

i) La densité spectrale f du processus ARMA est une fraction rationnelle en cos w, sin w. Pour tout réel $x \ge 0$, l'image inverse $f^{-1}([o,\ x[)$ s'exprime donc comme une réunion finie d'intervalles disjoints :

$$f^{-1}([O,\ x[) = \bigcup_{i=1}^{r(x)} (a_i(x),\ b_i(x)).$$

Le nombre r(x) dépend de la valeur x, mais est de toutes façons inférieur au nombre de variations de la fonction f sur $(-\pi,\ \pi)$, c'est-à-dire au plus à $2(p + q - 1)$.

ii) Par ailleurs, on sait que $\beta_{\upsilon}^{(n)} = f\left(\frac{\upsilon\pi}{n+1}\right)$. On en déduit que :

$$\beta_{\upsilon}^{(n)} < x \iff \upsilon \in \mathcal{J}_n(x) = \bigcup_{i=1}^{r(x)} \left(\frac{n+1}{\Pi}\, a_i(x),\ \frac{n+1}{\Pi}\, b_i(x)\right)$$

Puisque le nombre d'entiers contenu dans l'intervalle (a, b) est compris entre b - a et b - a + 2, nous avons :

$$\mathrm{Card}\,[\upsilon : \upsilon \in \mathcal{J}_n(x)] - \sum_{i=1}^{r(x)} \frac{n+1}{\Pi}\,[b_i(x) - a_i(x)] \le 2r(x)$$

$$\Longrightarrow \quad \left| \frac{1}{n} \text{ Card } \{\upsilon : \beta_\upsilon^{(n)} < x\} - \frac{1}{\Pi} \sum_{i=1}^{r(x)} [b_i(x) - a_i(x)] \right| \leq \frac{2r(x) + 1}{n}$$

Remarquant que : $\mathcal{F}(x) = \int \mathbb{1}_{[0,x]}(z) \, d \, \mathcal{U}^f_{[-\pi, \pi]}(z)$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbb{1}_{[f(w) \leq z]} \, dw$$

$$= \frac{1}{\pi} \int_0^{\pi} \mathbb{1}_{w \in f^{-1}([0, z[)} \, dw$$

$$= \frac{1}{\pi} \sum_{i=1}^{r(x)} [b_i(x) - a_i(x)] ,$$

nous voyons que l'inégalité obtenue s'écrit :

$$|\mathcal{G}_n(x) - \mathcal{F}(x)| \leq \frac{M(x)}{n} ,$$

où $M(x) = 2r(x) + 1 \leq 2[2(p + q - 1] + 1 = 4(p + q) - 3$.

Proposition 2.19 : Pour toute fonction h dérivable à dérivée bornée sur l'intervalle $\mathcal{J} = [0, \text{Max } f]$, on a :

$$\frac{1}{n} \sum_{\upsilon=1}^{n} |h(\lambda_\upsilon^{(n)}) - h(\beta_\upsilon^{(n)})| \leq \frac{K_0}{n} \sup_{\mathcal{J}} \left| \frac{dh}{dz}(z) \right| ,$$

où K_0 est une constante dépendant de f, mais pas de h et n.

Preuve : Nous savons que la matrice de covariance Γ_n^x s'écrit :

$$\Gamma_n^x = \left[\prod_{i=1}^{p} \Gamma_n(\varphi_i) \right]^{-1} \prod_{j=1}^{q} \Gamma_n(\theta_j) + \left[\prod_{i=1}^{p} \Gamma_n(\varphi_i) \right]^{-1} \sum_{k=1}^{p} \varphi_k \left[\prod_{j=1}^{k-1} \Gamma_n(\varphi_j) \right] E_n^{Z^k} - \sum_{\ell=1}^{q} \theta_\ell \left[\prod_{m=1}^{\ell-1} \Gamma_n(\theta_m) \right] E_n^{W^\ell} ,$$

où $Z_t^k = \prod_{j=k+1}^{p} (1 - \varphi_j L) \, X_t$, $W_t^\ell = \prod_{j=\ell+1}^{q} (1 - \theta_j L) \, \varepsilon_t$.

i) Le rang des matrices $E_n^{W^\ell}$ ou $E_n^{Z^k}$ est inférieur ou égal à 2. On en déduit plus généralement que pour deux suites de matrices A_j, $E_n^{W^j}$ $j = 1, \ldots s$, la combinaison $\sum_{j=1}^{s} A_j \, E_n^{W^j}$ a un rang inférieur ou égal à 2s.

Par conséquent l'erreur d'approximation :

$$\Gamma_n^x - \left[\prod_{i=1}^{p} \Gamma_n(\varphi_i) \right]^{-1} \prod_{j=1}^{q} \Gamma_n(\theta_j) ,$$

est symétrique de rang r inférieur à 2(p + q).

On peut alors trouver une matrice orthogonale Q_n telle que :

$$Q_n \Gamma_n^{*} Q_n' - Q_n \left[\prod_{i=1}^{p} \Gamma_n(\varphi_i)\right]^{-1} \prod_{j=1}^{q} \Gamma_n(\theta_j) Q_n' = \begin{bmatrix} \alpha_1^{(n)} & & & & 0 \\ & \ddots & & & \\ & & \alpha_r^{(n)} & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & 0 \end{bmatrix}.$$

ii) Cette égalité permet alors d'utiliser le lemme suivant :

Lemme 2.20 : Soient M et N deux matrices symétriques, dont les valeurs propres rangées par ordre croissant sont respectivement : λ_υ^M, $\upsilon = 1, \ldots, n$ et λ_υ^N, $\upsilon = 1, \ldots, n$. On suppose que $M - N = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Alors nous avons : $\sum_{\upsilon=1}^{n} |\lambda_\upsilon^M - \lambda_\upsilon^N| \leq \sum_{\upsilon=1}^{n} |\alpha_\upsilon|$.

Preuve du lemme :

i) Comme $(\alpha_1, \alpha_2 \ldots, \alpha_n) = (\alpha_1, 0, \ldots 0) + (0, \alpha_2, \ldots 0) + \ldots + (0 \ldots 0, \alpha_n)$, il suffit de faire le raisonnement dans le cas où : $(\alpha_1, \alpha_2, \ldots \alpha_n) = (\alpha_1, 0 \ldots 0)$, le cas général s'en déduisant par récurrence et application de l'inégalité triangulaire.

ii) Supposons donc $M - N = \mathrm{diag}(\alpha_1, 0, \ldots, 0)$ avec par exemple $\alpha_1 > 0$. Les deux matrices symétriques sont alors ordonnées $M \gg N$, ce qui entraîne (voir annexe 1) : $\lambda_\upsilon^M \geq \lambda_\upsilon^N$, $\forall \upsilon$.

Par ailleurs, on a : $\mathrm{Tr}\, M - \mathrm{Tr}\, N = \mathrm{Tr}\, \mathrm{diag}(\alpha_1, 0, \ldots, 0)$

$$\Leftrightarrow \sum_{\upsilon=1}^{n} \lambda_\upsilon^M - \sum_{\upsilon=1}^{n} \lambda_\upsilon^N = \alpha_1$$

$$\Leftrightarrow \sum_{\upsilon=1}^{n} |\lambda_\upsilon^M - \lambda_\upsilon^N| = |\alpha_1|.$$

Le résultat est analogue pour $\alpha_1 < 0$ et l'inégalité apparaissant dans le lemme s'introduit au moment de la récurrence et de l'application de l'inégalité triangulaire.

Q.E.D.

iii) Revenant alors à la partie centrale de la preuve, on note que la norme spectrale :

$$\|\Gamma_n^x - \left[\prod_{i=1}^{p}\Gamma_n(\varphi_i)\right]^{-1} \prod_{j=1}^{q}\Gamma_n(\theta_j)\| \leq \sup_w \frac{\prod_{j=1}^{q}|1 - \theta_j \exp iw|^2}{\prod_{i=1}^{p}|1-\varphi_i \exp iw|^2} + \frac{\prod_{j=1}^{q}(1 + |\theta_j|)^2}{\prod_{i=1}^{p}(1 - |\varphi_i|)^2}$$

c'est-à-dire par une constante K_0 ne dépendant que de la densité spectrale f.

Appliquant le lemme, on en déduit alors que :

$\sum_{\upsilon=1}^{n} |\lambda_\upsilon^{(n)} - \beta_\upsilon^{(n)}| \leq rK_0 \leq 2(p+q)\ K_0$, et donc pour une fonction h de dérivée bornée :

$$\frac{1}{n}\sum_{\upsilon=1}^{n} |h(\lambda_\upsilon^{(n)}) - h(\beta_\upsilon^{(n)})| \leq \frac{2(p+q)K_0}{n} \sup_{z\in\mathcal{J}} \left|\frac{dh}{dz}(z)\right|$$

Q.E.D.

III - DISTRIBUTION ASYMPTOTIQUE DES VALEURS PROPRES (CAS DE PROCESSUS STATIONNAIRES)

Dans le paragraphe précédent, nous avons considéré un processus stationnaire $X = (X_t, t \in \mathbb{Z})$ admettant une représentation autorégressive moyenne mobile. Nous avons déterminé certaines approximations des valeurs propres $\{\lambda_\upsilon^{(n)}\ \upsilon = 1,\ldots, n\}$ de la matrice de variance covariance de $[X_t, X_{t+1},\ldots X_{t+n-1}]$. Ceci nous a permis en particulier d'approcher la distribution empirique de ces valeurs propres et de montrer de façon directe que cette distribution converge vers la loi image de la loi uniforme sur $[-\pi, \pi]$ par l'application densité spectrale. Ce résultat est en fait un cas particulier d'un théorème général établi par Szegö (1921) et valable pour tous les processus stationnaires admettant une densité spectrale bornée. Dans le premier sous-paragraphe nous commençons par rappeler ce résultat et son principe de démonstration. Celui-ci repose de façon fondamentale sur l'égalité de Kolmogorov. Dans le second sous-paragraphe nous étendons ce résultat au cas de processus stationnaires à densité spectrale non bornée. Ceci permet donc d'inclure les processus dits à mémoire longue qui tout en étant stationnaires admettent une densité spectrale infinie à l'origine. Le cas de processus stationnaires admettant une mesure spectrale mélange de mesure discrète et de mesure continue s'en déduit alors aisément.

III.a - Le théorème de Szegö

III.a.1 - Matrices de Toeplitz

Le résultat a été historiquement présenté dans le cadre des matrices de Toeplitz. Rappelons qu'une matrice M_n carrée de taille n, d'élément général m_{kl} $k = 1,\ldots, n$, $l = 1,\ldots, n$ est dite de Toeplitz si $m_{kl} = m_{k',l'}$ pour $k - l = k' - l'$.

Il s'agit donc d'une matrice admettant les mêmes éléments sur les paradiagonales principales :

$$M_n = \begin{bmatrix} a_0 & a_1 & \cdot & \cdot & a_{n-1} \\ a_{-1} & a_0 & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & a_1 \\ a_{1-n} & \cdot & \cdot & a_{-1} & a_0 \end{bmatrix}, \quad a_\upsilon \in \mathbb{C}, \ \upsilon = 1 - n,\ldots, n - 1.$$

Des matrices de Toeplitz sont généralement obtenues à partir de la transformation de Fourier. Supposons que f désigne une fonction réelle intégrable sur l'intervalle $[-\pi, \pi]$, on peut introduire la matrice de Toeplitz :

$$M_n(f) = (m_{k,l})_{1\le k,\ l\le n}, \tag{3.1}$$

avec : $m_{k,l} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp[i(k-l)w] f(w)\, dw.$

Cette matrice est de plus hermitienne.

Remarquons immédiatement que des résultats obtenus pour de telles matrices $M_n(F)$ construites à partir d'une fonction intégrable f s'appliqueront en particulier aux matrices de variance-covariance associées à un processus stationnaire. En effet si $X = (X_t, t \in \mathbb{Z})$ est un tel processus, admettant une densité spectrale (intégrable) f. La fonction d'autocovariance se déduit de la densité spectrale par :

$$\gamma(h) = \int_{-\pi}^{\pi} \exp(ihw)\, f(w)\, dw.$$

La matrice de variance-covariance de $(X_t, X_{t+1},\ldots X_{t+n-1})$ définie par :

$$\Gamma_X^{(n)} = (\gamma(k-l))_{1\le k,\ l\le n},$$

coïncide alors la matrice de Toeplitz M_n $(2\pi f)$.

III.a.2 - Distribution asymptotique des valeurs propres, lorsque la densité spectrale est bornée

Théorème 3.2 : Soit f une fonction mesurable réelle définie sur $[-\pi, \pi]$. On suppose que cette fonction est bornée (presque partout) : $-\infty < m \leq f \leq M < +\infty$ (p.p), et on désigne par $\{\lambda_\upsilon^{(n)}, \upsilon = 1,\ldots, n\}$ les valeurs propres de la matrice de Toeplitz $M_n(f)$. Alors pour toute fonction réelle h, continue sur l'intervalle [m, M], on a :

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h[f(w)]\, dw.$$

Preuve : Nous rappelons brièvement le principe de démonstration de ce résultat.

i) La preuve s'appuie sur l'égalité de Kolmogorov [voir par exemple Whittle (1963), Azencott-Dacunha-Castelle (1983)].

Si f désigne une fonction bornée (p.p) positive, on a :

$$\lim_{n\to\infty} \frac{1}{n} \text{Log det } M_n(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log } f(w)\, dw. \tag{3.3}$$

Nous présentons ici le résultat sous une forme peut être peu usuelle pour les spécialistes de séries temporelles, mais forme qui est directement utilisable dans notre cas.

De façon plus classique si X est un processus stationnaire de densité spectacle f bornée (nécessairement réelle positive), on introduit l'innovation $\varepsilon_t = X_t - {}_{t-1}\hat{X}_t$, où ${}_{t-1}\hat{X}_t$ désigne la meilleure prévision linéaire de X_t fondée sur le passé du processus X et l'erreur quadratique moyenne de prévision $\sigma^2 = V\varepsilon_t = V(X_t - {}_{t-1}\hat{X}_t)$. Il existe alors deux façons de déterminer cette variance :

- dans le domaine des temps : $\sigma^2 = \lim_{n\to\infty} \frac{1}{n} \text{Log det } M_n(f)$.
- dans le domaine des fréquences : $\sigma^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log } f(w)\, dw$.

L'égalité est obtenue en comparant ces deux modes de calcul.

l'égalité peut aussi être écrite en introduisant les valeurs propres $\lambda_\upsilon^{(n)}$ de la matrice $M_n(f)$. On obtient :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} \text{Log } \lambda_\upsilon^{(n)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log } f(w)\, dw. \tag{3.4}$$

Avec cette écriture, on note que l'égalité de Kolmogorov n'est

autre que l'égalité apparaissant dans le théorème de Szegö et appliquée à la fonction $h(x) = \log x$. L'idée de la démonstration est alors d'atteindre l'ensemble des fonctions continues à partir de transformées bien choisies de la fonction Log.

ii) Considérons alors un réel z tel que : $|z| < \frac{1}{2M}$. La fonction : $w \longrightarrow 1 + z\, f(w)$ est positive bornée (p.p). De plus on vérifie facilement que la matrice $M_n(1 + zf)$ admet comme valeurs propres $1 + z\,\lambda_\nu^{(n)}$, $\upsilon = 1, \ldots, n$,valeurs propres qui sont d'ailleurs toutes strictement positives [voir Propriété3 de l'annexe]. On peut donc aussi appliquer l'égalité de Kolmogorov à cette fonction, ce qui conduit aux relations :

$$(3.5) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{Log}\,[1 + z\,\lambda_\upsilon^{(n)}] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Log}(1 + z\, f(w))\, dw,$$

$$\forall z,\ |z| < \frac{1}{2M}.$$

L'égalité du théorème apparait maintenant vérifiée pour toutes les fonctions $h(x) = \mathrm{Log}(1 + zx)$, avec z réel assez petit. La suite de la démonstration consiste à montrer qu'elle l'est aussi pour les fonctions puissances $h(x) = x^p$, qui forment une famille dense dans l'ensemble des fonctions continues. Pour cela il est naturel d'identifier les développements en série entière en z des deux membres de l'égalité (3.5). Ceci demande cependant dans une première étape, d'étendre l'égalité (3.5) à des valeurs complexes de l'argument z.

iii) Considérons donc la suite $(g_n,\ n \geq 1)$ de fonctions de la variable complexe z, suite définie par :

$$g_n(z) = \frac{1}{n} \sum_{\upsilon=1}^{n} \{\mathrm{Log}\,[1 + z\,\lambda_\upsilon^{(n)}] - \mathrm{Log}\,[1 + zf(-\pi + \frac{2\upsilon\pi}{n+1})]\}.$$

$(g_n,\ n \geq 1)$ est une suite de fonctions analytiques dans le disque $\{z\ ; |z| < \frac{1}{2M}\}$ et est uniformément bornée en n, z dans ce disque (par exemple par $2\ \mathrm{Log}\ \frac{3}{2}$). En se restreignant aux valeurs réelles de z, on déduit de (3.5) que :

$$\lim_{n\to\infty} g_n(z) = 0 \qquad , \quad \forall z \in \mathbb{R},\ |z| < \frac{1}{2M}$$

Appliquant le théorème de Vitali [voir Valiron (1951)], ceci implique que $(g_n,\ n \geq 1)$ tend uniformément vers 0 sur le disque $\{z : |z| < \frac{1}{2M}\}$.

iv) Les coefficients $(a_p^{(n)})$ $p \geq 0$ apparaissant dans le développement en série entière de g_n tendent donc vers zéro. Ceci permet d'obtenir les égalités :

$$(3.6) \qquad \forall p \geq 0 : \lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} [\lambda_\upsilon^{(n)}]^p = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(w)]^p \, dw.$$

La démonstration se termine en remarquant que la famille des combinaisons de fonctions puissance est dense pour la convergence uniforme dans l'ensemble des fonctions continues sur [m, M] (théorème de Stone-Weierstrass).

Q.E.D.

Le théorème 3.2 a une interprétation directe en terme de distribution empirique des valeurs propres. En effet si μ_n désigne cette distribution empirique, l'égalité du théorème s'écrit :

$$\lim_{n\to\infty} \int h(u) \, d\mu_n(u) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h[f(w)] \, dw$$

$$= \int_{[-\pi,\pi]} h[f(w)] \, d\mathcal{U}(w) \quad (\text{où } \mathcal{U}_{[-\pi,\pi]} \text{ désigne la loi uniforme sur } [-\pi, \pi])$$

$$= \int h(u) \, d\mathcal{U}^f_{[-\pi,\ \pi]}(u) \quad (\text{où } \mathcal{U}^f_{[-\pi,\pi]} \text{ désigne la loi image par l'application f de la loi uniforme sur } [-\pi, \pi]).$$

Le théorème fournit donc un résultat de convergence faible de la loi empirique des valeurs propres $\lambda_\upsilon^{(n)}$, $\upsilon = 1, \ldots, n$ vers la loi $\mathcal{U}^f_{[-\pi,\ \pi]}$.

De cette convergence, il est naturel de déduire certains résultats concernant les supports des mesures μ_n et $\mathcal{U}^f_{[-\pi,\ \pi]}$. En terme de valeurs propres, ceci permet d'étudier le comportement de valeurs propres extrêmes, c'est-à-dire des plus petites (resp. des plus grandes). La propriété ci-dessous généralise un résultat figurant dans Grenander-Szegö (1958) (voir aussi Wilf (1970)).

Propriété 3.7 : Soit M (resp m) la borne supérieure (resp inférieure) essentielle de la fonction f, avec $-\infty < m \leq M < +\infty$. Pour toute suite $(p_n)_{n\geq 1}$ d'entiers vérifiant $\lim_{n\to\infty} \frac{p_n}{n} = 0$, on a :

$$\lim_{n\to+\infty} \lambda_{p_n}^{(n)} = m \text{ et } \lim_{n\to+\infty} \lambda_{n-p_n}^{(n)} = M.$$

Preuve : D'après les propriétés relatives au classement des valeurs propres (voir annexe), nous savons que : $\lambda_\upsilon^{(n)} \to m$, $\upsilon = 1, \ldots n$. Soient alors ε un

réel strictement positif et F une fonction réelle continue, nulle sur ^{C} [m, n + ε[et strictement positive sur]m, m + ε[. On peut par exemple prendre pour fonction F celle définie par :

$$\begin{cases} F = 0, \text{ sur }]m,\ m + \varepsilon[, \\ F = |x - m||x - m - \varepsilon|, \text{ sur }]m,\ m + \varepsilon[. \end{cases}$$

Par définition de la borne inférieure essentielle, il existe un ensemble J de $[-\pi, \pi]$, de mesure strictement positive et tel que :
$m \leq f < m + \varepsilon$ sur J.

On peut alors appliquer le théorème 3.2 à la fonction F.
On obtient :

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{i=1}^{n} F(\lambda_{\upsilon}^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F[f(w)]\ dw > 0.$$

Comme la fonction F est nulle à l'extérieur de]m, m + ε[, ceci implique : $\lim_{n\to+\infty} \frac{1}{n} \sum_{\lambda_{\upsilon}^{(n)} \in\ ]m,\ m+\varepsilon[} F[\lambda_{\upsilon}^{(n)}] > 0.$

Donc le nombre de valeurs inférieures à m + ε est d'ordre n. On en déduit que si : $\lim_{n\to\infty} \frac{p_n}{n} = 0$, on a pour n assez grand $\lambda_{p_n}^{(n)} < m + \varepsilon$. Ceci étant valable quelque soit ε, on aboutit au résultat cherché $\lim_{n\to\infty} \lambda_{p_n}^{(n)} = m$.

Q.E.D.

Ce résultat reste d'ailleurs valable, si la fonction f n'est pas bornée.

Propriété 3.8 : L'énoncé de la propriété 3.7 reste vrai, si $m = -\infty$, $M < +\infty$ ou si $m > -\infty$, $M = +\infty$.

Preuve : supposons $m = -\infty$; on peut se ramener au cas précédent en tronquant la fonction f, c'est-à-dire en posant :

$$f_K = \begin{cases} f, \text{ si } f \geq -K, \\ -K, \text{ si } f \leq -K. \end{cases}$$

On a alors : $\lim_{K\to\infty} \inf_{w} f_K(w) = -\infty$.

Introduisons les valeurs propres $\lambda_{\upsilon}^{(n)}$ associées à $M_n(f_K)$, valeurs rangées dans l'ordre croissant. D'après l'appendice nous avons $f_K \geq f$, ce qui entraine $\lambda_{\upsilon}^{(n)}(K) \geq \lambda_{\upsilon}^{(n)}$, $\forall \upsilon, n, K$. Donc :

$\limsup_{n\to\infty} \lambda_{p_n}^{(n)} \leq \lim_{n\to\infty} \lambda_{p_n}^{(n)}(K) = \inf f_K = -K$, d'après la propopition 3.7.

La valeur de troncature K étant arbitraire, on en déduit que :

$$\lim_{n\to+\infty} \sup \lambda_{p_n}^{(n)} = -\infty.$$

Q.E.D.

III.b - Distribution asymptotique des valeurs propres, lorsque la densité spectrale n'est pas bornée.

Le théorème de Szegö (3.2) s'applique aux processus stationnaires, dont la densité spectrale est bornée (presque partout). Il est intéressant d'examiner, si le résultat est encore vérifié, quand on considère des processus stationnaires, mais dont la densité spectrale, tout en étant intégrable n'est pas bornée. De tels processus dits à mémoire longue apparaissent utiles pour étudier comment s'effectue la transition entre le cas stationnaire ARMA classique et le cas non stationnaire. Des exemples de tels processus sont fournis par les processus fractionnaires (voir Mandelbrot-Van Ness (1968), Granger (1978) Hosking(1981) ou Gonçalves (1987) pour un survey de ce sujet). Dans le cas le plus simple, ils sont définis par l'équation :

$$(1 - L)^d X_t = \varepsilon_t, \quad t \in \mathbb{Z}$$

où L désigne l'opérateur retard, et ε un bruit blanc. La densité spectrale est nécessairement de la forme :

$$f(w) = \frac{\sigma^2}{2\pi} \frac{1}{|1 - \exp iw|^{2d}}.$$

Elle peut ne pas exister pour la fréquence nulle. En fait comme son développement limité au voisinage de ces fréquences est :

$$f(w) \sim \frac{\sigma^2}{2\pi} \frac{1}{w^{2d}},$$

on voit que : si $d \geq \frac{1}{2}$, la densité n'est pas intégrable et le processus n'est pas stationnaire ;

si $0 < d < \frac{1}{2}$, la densité est intégrable [processus stationnaire], mais elle est infinie à l'origine;

si $d \leq 0$, le processus est stationnaire, de densité spectrale bornée.

Nous allons voir que la généralisation du théorème 3.2 est obtenue en imposant des conditions différentes aux fonctions h pour lesquelles l'égalité :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h[f(w)]\, dw,$$

est satisfaite.

Dans une première étape nous introduisons une métrique Δ sur les suites de suites partielles, métrique adaptée au problème. Nous étudions alors l'ensemble de fonctions réelles continues vis à vis de cette métrique, fonctions dites Δ - continues et examinons les liens entre Δ continuité et continuité uniforme. Dans le second sous-paragraphe nous établissons un théorème sur la distribution limite de valeurs propres pour les fonctions h Δ - continues.

III.b.1 - Fonctions Δ continues

Soient $\mathcal{J}$ un intervalle de $\mathbb{R}$ et $\prod_{n=1}^{\infty} \mathcal{J}^n = \mathcal{E}(\mathcal{J})$ l'ensemble des suites de n-uplets : $[x] = \{x_\upsilon^{(n)},\ \upsilon = 1...n\}_{n\geq 1}$, dont les termes sont tous pris dans $\mathcal{J}$.

Pour $[x] = \{x_\upsilon^{(n)},\ \upsilon = 1,\ldots,\ n\}_n \geq 1$, $[y] = \{y_\upsilon^{(n)},\ \upsilon = 1,\ldots,\ n\}_n \geq 1$, éléments de $\mathcal{E}(\mathcal{J})$, on pose :

$$\Delta([x]\ ;\ [y]) = \sup_n \frac{\frac{1}{n} \sum_{\upsilon=1}^{n} |x_\upsilon^{(n)} - y_\upsilon^{(n)}|}{1 + \frac{1}{n} \sum_{\upsilon=1}^{n} |x_\upsilon^{(n)} - y_\upsilon^{(n)}|}. \tag{3.9}$$

L'application Δ est clairement une distance sur $\mathcal{E}(\mathcal{J})$.

Soit alors h une fonction réelle définie sur $\mathcal{J}$ à valeurs dans $\mathcal{J}$, on peut immédiatement lui associer la fonction $\tilde{h}$ définie sur $\mathcal{E}(\mathcal{J})$ par :

$$\tilde{h}([x]) = \{h(x_\upsilon^{(n)}),\ \upsilon = 1,\ldots,\ n\}_n \geq 1. \tag{3.10}$$

Dans la suite on identifie h et $\tilde{h}$, quand il n'y a pas de confusion possible.

On pose alors la définition suivante :

Définition 3.11 : h est dite Δ-continue, si et seulement si l'application associée $\tilde{h}$ est continue pour la métrique Δ.

La condition de Δ - continuité peut être explicitée de la façon suivante.

Soit $[x_p]_{p\geq 1}$ une suite de suites de n-uplets convergeant vers [x], c'est-à-dire telle que :

$$\lim_p \sup_n \frac{1}{n} \sum_{\upsilon=1}^{n} |x_\upsilon^{(n)}(p) - x_\upsilon(p)| = 0 ;$$

alors si h est Δ - continue, on a :

$$\lim_p \sup_n \frac{1}{n} \sum_{\upsilon=1}^{n} |h(x_\upsilon^{(n)}(p)) - h(x_\upsilon(p))| = 0.$$

Il s'agit donc d'une condition assurant que la fonction h admet un bon comportement vis à vis de moyennes de Cesaro.

La proposition suivante, montrée dans l'annexe 2, fournit des renseignements sur les liens entre Δ-continuité et continuité uniforme.

Proposition 3.12 :

i) Toute fonction Δ-continue est uniformément continue.

ii) Toute fonction uniformément continue, bornée est Δ-continue.

iii) Si h est uniformément continue sur $\mathcal{J} = [m, +\infty[$ et k-lipschitzienne sur $[M, +\infty[$, $(M \geq m)$, alors h est Δ -continue.

Il est clair que la classe des fonctions Δ-continues excède celle des fonctions uniformément continues bornées sur $\mathcal{J}$, sauf évidemment dans le cas d'un intervalle $\mathcal{J}$ compact où ces deux familles coïncident. Par exemple l'application identité $h(x) = x$ est Δ - continue non bornée. La partie iii) de la proposition prend en compte de telles fonctions uniformément continues, non bornées à l'infini. Remarquons que la restriction $m \geq 0$ intervenant dans cette partie iii) est peu gênante ; en effet dans la suite, nous appliquerons cette proposition à des applications h définie sur l'image de la densité spectrale et cette dernière est positive.

III-b.2 - Etude asymptotique pour des densités spectrales non bornées.

En utilisant la notion de Δ-continuité, on obtient une extension du théorème 3.2 aux cas de densités spectrales non bornées.

Théorème 3.13 : Soit f une densité spectrale, fonction intégrable sur $[-\pi,\pi]$; soient $\lambda_\upsilon^{(n)}$, $\upsilon = 1 \ldots n$, les valeurs propres de $M_n(f)$, $n \geq 1$. Alors pour

toute fonction h Δ-continue sur un intervalle $\mathcal{J}$ contenant$f([-\pi, \pi])$ et telle que $h[f]$ soit intégrable, on a :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} h[\lambda_\upsilon^{(n)}] = \frac{1}{2\pi} \int_{-\pi}^{\pi} h[f(w)]dw.$$

L'ensemble des fonctions Δ-continues étant plus large que l'ensemble des fonctions continues à support compact, le théorème 3.13 implique comme en 3.2 la convergence en loi de la distribution empirique des valeurs propres vers la loi transformée : $\mathcal{U}^f_{[-\pi, \pi]}$.

Preuve du théorème 3.13 :

Nous allons nous ramener au théorème 3.2 en bornant la densité spectrale f. Définissons :

$$f_K = \begin{cases} f, \text{ si } f \leq K, \\ K, \text{ sinon.} \end{cases}$$

i) f_K est intégrable, bornée par K et inférieure à la densité spectrale f. Notons $\mu_\upsilon^{(n)}(K)$, $\upsilon = 1 \ldots n$ les valeurs propres de $M_n(f_K)$ rangées dans l'ordre croissant. On peut appliquer le théorème 3.2 à la fonction f_K :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} h(\mu_\upsilon^{(n)}(K)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h[f_K(w)]\ dw,$$

pour toute fonction h Δ-continue (ou de façon équivalente continue sur [0, K]).

ii) Par ailleurs d'après l'annexe 1, nous savons que $f_K \leq f$ entraîne $\mu_\upsilon^{(n)}(K) \leq \lambda_\upsilon^{(n)}$, $\upsilon = 1, \ldots n$, $n \geq 1$ et par conséquent :

$$\frac{1}{n} \sum_{\upsilon=1}^{n} |\lambda_\upsilon^{(n)} - \mu_\upsilon^{(n)}(K)| = \frac{1}{n} \sum_{\upsilon=1}^{n} [\lambda_\upsilon^{(n)} - \mu_\upsilon^{(n)}(K)]$$

$$= \frac{1}{n} (\mathrm{Tr}\ M_n(f) - \mathrm{Tr}\ M_n(f_K)] \text{ (où Tr désigne la trace de la matrice)}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \{f(w) - f_K(w)\}\ dw$$

$$= \frac{1}{2\pi} \int_{(f \geq K)} [f(w) - K]\ dw.$$

Cette dernière quantité, indépendante de n, tend vers zéro, si K tend vers l'infini à cause de l'intégrabilité de la densité spectrale f.

Autrement dit : $\lim_{K\to\infty} \Delta([\mu_K], [\lambda]) = 0$, et par Δ - continuité de la

fonction h, nous avons : $\lim_{K\to\infty} \Delta(h([\mu_K]), h([\lambda])) = 0$.

iii) Finalement pour obtenir le résultat, il suffit d'écrire la décomposition :

$$\left|\frac{1}{n}\sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) - \frac{1}{2\pi}\int_{-\pi}^{\pi} h[f(w)]dw\right|$$

$$\leq \left|\frac{1}{n}\sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) - \frac{1}{n}\sum_{\upsilon=1}^{n} h(\mu_\upsilon^{(n)}(K))\right|$$

$$+ \left|\frac{1}{n}\sum_{\upsilon=1}^{n} h(\mu_\upsilon^{(n)}(K)) - \frac{1}{2\pi}\int_{-\pi}^{\pi} h[f_K(w)]\right|$$

$$+ \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} h[f_K(w)]\, dw - \frac{1}{2\pi}\int_{-\pi}^{\pi} h[f(w)]\, dw\right|$$

Comme h[f] est supposée intégrable, on a pour $\varepsilon > 0$ fixé et K assez grand, le dernier terme : $\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} h[f_K(w)]\, dw - \frac{1}{2\pi}\int_{-\pi}^{\pi} h[f(w)]\, dw\right|$ qui est inférieur à $\varepsilon/3$.

Pour K assez grand, on a aussi d'après ce qui précède que : $\left|\frac{1}{n}\sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) - \frac{1}{n}\sum_{\upsilon=1}^{n} h(\mu_\upsilon^{(n)}(K))\right| < \frac{\varepsilon}{3}$, uniformément en n. Choisissant n assez grand et utilisant le théorème 3.2, on a alors :

$$\left|\frac{1}{\mu}\sum_{\upsilon=1}^{n} h(\mu_\upsilon^{(n)}(K)) - \frac{1}{2\pi}\int_{-\pi}^{\pi} h[f_K(w)]\, dw\right| < \frac{\varepsilon}{3}.$$

En résumé : $\lim_{n\to\infty} \frac{1}{n}\sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} h[f(w)]\, dw.$

Q.E.D.

III.c - Le cas des processus singuliers

III.c.1 - Mesure spectrale

On sait, (voir [Doob (1953]) que tout processus stationnaire $(X_t, t \in \mathbb{Z})$ admet une mesure spectrale, dont on note F la fonction de répartition ; dans ces conditions, F est monotone non décroissante et, si $\gamma(h) = \text{cov}(X_t, X_{t-h})$, $h \in \mathbb{Z}$ désigne la covariance d'ordre h alors $\gamma(h) = \int_{-\pi}^{\pi} e^{-iwh} dF(w)$. Par ailleurs, F se décompose suivant le lemme de Lebesgue (voir [Priestley (1981)] en une somme de F_a, fonction absolument continue, F_j fonction étagée avec des sauts d'amplitude p_r en

w_r, r = 1, 2, 3... et enfin F_s une fonction "singulière" continue à dérivée nulle presque partout : $F = F_a + F_j + F_s$.

Ainsi, dans les chapitres précédents on a considéré exclusivement le cas où F s'identifie à sa partie absolument continue ; en particulier : $dF(w) = f(w)dw$, où $f(w)$ est la densité spectrale du processus.

On se propose, dans ce paragraphe d'étudier en quoi ce qui précède en particulier les valeurs propres des matrices de covariance et le théorème de Szegö, est modifié quand on se place dans la situation plus générale d'une mesure spectrale admettant une partie absolument continue et une partie discrète.

On peut remarquer que Grenander et Szegö dans une note marginale (Appendice, note 8-5 p 234) de leur ouvrage ([Grenander, Szegö (1958)] abordent rapidement la question et obtiennent (par des procédés d'ailleurs plus "heuristiques" que rigoureux) les résultats les plus immédiats, à savoir, essentiellement, que la présence d'une partie discrète engendre des valeurs propres d'ordre n. On établira rigoureusement ce résultat et on l'utilisera pour étendre le théorème de Szegö à des mesures spectrales avec partie discrète.

III.c.2 - Encadrement des valeurs propres

On suppose que $(X_t, t \in \mathbb{Z})$ est un processus stationnaire dont la mesure spectrale admet une fonction de répartition $F = F_a + F_j$, avec la partie continue F_a de densité φ bornée sur $[-\pi, \pi]$: $dF_a(w) = \varphi(w)\, dw$ et la partie discrète du type $F_j = \sum_{k=1}^{r} \theta_k \delta_{w_k}$, où δ_{w_k} désigne la masse ponctuelle en w_k. On suppose que les valeurs w_k sont toutes distinctes dans $[-\pi, \pi]$ et les poids sont rangés par ordre décroissant : $\theta_1 \geq \theta_2 \ldots \geq \theta_r$. On a alors la proposition :

Propriété 3.14 : Soient $\lambda_\upsilon^{(n)}$, $\upsilon = 1, \ldots, n$ les valeurs propres rangées dans l'ordre décroissant de la matrice de covariance Γ_n d'ordre n du processus. Soient d'autre part Γ_n^a la matrice de Toeplitz associée à la fonction φ : $\Gamma_n^a = M_n(\varphi)$, et $\mu_\upsilon^{(n)}$, $\upsilon = 1, \ldots, n$ les valeurs propres de cette matrice rangées par ordre décroissant. Il existe une constante M_0 indépendante de n telle que : $n\,\theta_\upsilon - M_0 \leq \lambda_\upsilon^{(n)} \leq n\,\theta_\upsilon + M_0$, $\upsilon = 1, \ldots, r$,

$$0 \leq \lambda_\upsilon^{(n)} \leq M_0 \quad \upsilon = r+1, \ldots, n,$$

$$0 \leq \mu_\upsilon^{(n)} \leq M_0 \quad \upsilon = 1, \ldots, n$$

Preuve : voir Annexe 3

III.c.3 - Distribution limite des valeurs propres

L'existence d'une partie discrète dans la mesure spectrale

introduit donc des valeurs propres $\lambda_\upsilon^{(n)}$ d'ordre n. Il est alors clair que le théorème de Szegö n'est pas applicable à des fonctions h quelconques. Un contrôle du comportement à l'infini de telles fonctions est alors nécessaire.

Propriété 3.15 : Soit h une fonction continue sur $\mathbb{R}^+$ telle que $\lim_{n\to\infty} \frac{h(x)}{x} = \ell < +\infty$. Alors :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\varphi(w))\, dw + \ell \left(\sum_{k=1}^{r} \theta_k \right)$$

Preuve :

a) Nous donnons essentiellement une idée de la démonstration qui est décrite de façon complète dans l'annexe 3. Elle repose sur le lemme ci-dessous.

Lemme 3.16 : Si inf ess $\varphi = m > 0$, on a :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} \text{Log } \lambda_\upsilon^{(n)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log } \varphi(w)\, dw.$$

b) Soit alors z un réel positif et $Y_t = \varepsilon_t + \sqrt{z}\, X_t$, $t \in \mathbb{Z}$, où (ε_t) désigne un bruit blanc, de variance 1, non corrélé avec (X_t). Les valeurs propres de niveau n associées au processus (Y_t) sont $1 + z\lambda_\upsilon^{(n)}$, $\upsilon = 1, \ldots n$ et la mesure spectrale correspondant à ce processus admet une partie continue de densité $1 + z\varphi$. Appliquant le lemme, on obtient :

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} \text{Log } (1 + z\lambda_\upsilon^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log}[1 + z\varphi(w)]dw, \text{ si } z \geq 0.$$

c) D'après la proposition 3.14, $\lambda_1^{(n)}, \ldots, \lambda_r^{(n)}$ sont d'ordre n et puisque $\frac{\text{Log } n}{n}$ est de limite nulle, il vient $\lim_{n\to+\infty} \frac{\text{Log } \lambda_\upsilon^{(n)}}{n} = 0$ $\upsilon = 1, \ldots r$. On a donc aussi l'égalité :

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{\upsilon=r+1}^{n} \text{Log } [1 + z\, \lambda_\upsilon^{(n)}] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log}[1 + z\, \varphi(w)]\, dw, \; z \geq 0.$$

d) On peut alors en déduire (voir annexe 3) que pour toute fonction continue h définie sur un intervalle borné, où les valeurs propres $\lambda_\upsilon^{(n)}$ $\upsilon = r + 1, \ldots n$, $n \geq 1$ prennent leurs valeurs, on a :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=r+1}^{n} h(\lambda_\upsilon^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\varphi(w))\ dw.$$

e) Finalement si la fonction h est continue sur $\mathbb{R}^+$, telle que $\lim_{x\to\infty} \frac{h(x)}{x} = \ell < \infty$, on a :

$$\lim_{n\to+\infty} \frac{h(\lambda_\upsilon^{(n)})}{n} = \lim_{n\to+\infty} \frac{h(\lambda_\upsilon^{(n)})}{\lambda_\upsilon^{(n)}} \frac{\lambda_\upsilon^{(n)}}{n} = \theta_\upsilon\ \ell \text{ si } \upsilon \leq r,$$

et donc :

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} h(\lambda_\upsilon^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\varphi(w))\ dw + \left(\sum_{k=1}^{r} \theta_k \right) \ell.$$

Q.E.D.

ANNEXE 1

Résultats Préliminaires

On rappelle ici quelques résultats classiques et on montre quelques propositions utilisées dans le texte :

a) Valeurs propres et ordre sur les formes quadratiques

On rappelle d'abord la caractérisation de Weyl-Courant [Chatelin (1988)].

Proposition 1 :

Soit A une matrice hermitienne $(n+1) \times (n+1)$ et $\lambda_0 \leq \lambda_1 \leq \ldots \lambda_n$ ses valeurs propres ; alors $\lambda_\upsilon = \underset{a_1, \ldots a_\upsilon}{\text{Max}} \{ \underset{\substack{\|u\|=1 \\ \langle u, a_s \rangle = 0 \\ 1 \leq 1 \leq \upsilon}}{\min} \bar{u}' A u \}$

pour $\upsilon = 1, \ldots n$.

Preuve : Soient $\xi_0, \ldots, \xi_n$, $n+1$ vecteurs propres unitaires respectivement associés à $\lambda_0, \ldots, \lambda_n$; pour $\upsilon \geq 1$, on considère un sous espace $\mathcal{V}$ de dimension supérieure à $n + 1 - \upsilon$. En notant E_υ le sous-espace engendré par les $\{\xi_s\}_{0 \leq s \leq \upsilon}$ on a alors :
$\dim E_\upsilon \cap \mathcal{V} = \dim E_\upsilon + \dim \mathcal{V} - \dim(E_\upsilon + \mathcal{V}) \geq n + 1 + \upsilon + 1 - \upsilon - (n+1)$, donc $\dim E_\upsilon \cap \mathcal{V} \geq 1$ et il existe un vecteur u de $\mathcal{V}$ tel que $u = \sum_{k=0}^{\upsilon} \alpha_k \xi_k$ avec $\sum_{k=0}^{\upsilon} |\alpha_k|^2 = 1$.

Il vient alors : $\bar{u}' A u = \sum_{k=0}^{\upsilon} |\alpha_k|^2 \lambda_k \leq \lambda_\upsilon$. Ce raisonnement pouvant être mené avec tous les sous-espaces $\mathcal{V}$ de dimension supérieure à $n + 1 - \upsilon$, il vient :

$$\underset{a_1, \ldots a_\upsilon}{\text{Max}} \{ \underset{\substack{\|u\|=1 \\ \langle u, a_s \rangle = 0 \\ 1 \leq s \leq \upsilon}}{\min} \bar{u}' A u \} \leq \lambda_\upsilon \}.$$

Pour $\upsilon \geq 1$, posons alors $a_i = \xi_{i-1}$, $i = 1, \ldots, \upsilon$.

Tout vecteur u, orthogonal aux vecteurs a_i s'écrit :

$u = \sum_{k=\upsilon}^{n} \alpha_k \xi_k$, avec $\sum_{k=\upsilon}^{n} |\alpha_k|^2 = 1$, si u est pris unitaire.

Alors : $\bar{u}'\ A\ u = \sum_{k=\upsilon}^{n} |\alpha_k|^2\ \lambda_k \geq \lambda_\upsilon$ d'où :

$$\underset{a_1,\ldots,\ a_\upsilon}{\text{Max}} \left\{ \min_{\substack{\|u\|=1 \\ <u-a_\ell>=0 \\ 1\leq 1\leq \upsilon}} \bar{u}'\ Au \right\} \geq \lambda_\upsilon$$

et donc finalement l'égalité.

Q.E.D.

Les propositions qui suivent relient ordre sur les valeurs propres et ordres sur les formes quadratiques.

Rappelons que, pour deux formes quadratiques A et B, $A \ll B$ si $A - B$ est semi définie négative. ($x'(A - B)\ x \leq 0$, pour tout x).

Proposition 2 : [Kobilinsky (1979)] :

Soient A (resp B) une forme quadratique et $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ (resp $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n$) ses valeurs propres ; alors
$A \ll B \Longrightarrow \lambda_i \leq \mu_i$ $i = 1,\ldots, n$.

Preuve : Notons $(a_1,\ldots, a_n)$ une base orthonormée de vecteurs propres de A, a_i associé à λ_i, $i = 1,\ldots, n$ et de même $(b_1,\ldots b_n)$ une base orthonormée de vecteurs propres de B, b_i associé à μ_i, $i = 1,\ldots, n$.

L'espace engendré par $b_1,\ldots, b_i$ de dimension i et l'espace engendré par $a_i,\ldots, a_n$ de dimension $n - i + 1$ ont en commun un vecteur u, non nul, qu'on suppose unitaire et qu'on écrit simultanément :

$u = \sum_{k=1}^{i} \alpha_k\ b_k$ et $u = \sum_{h=i}^{n} \beta_h\ a_h$, avec $\sum_{k=1}^{i} \alpha_k^2 = \sum_{h=i}^{n} \beta_h^2 = 1$. Alors :

$$u'Au = \beta_i^2\ \lambda_i + \ldots + \beta_n^2\ \lambda_n \geq \lambda_i,$$

$$u'Bu = \alpha_1^2\ \mu_1 + \ldots + \alpha_i^2\ \mu_i \leq \mu_i$$

et donc $\lambda_i \leq u'Au \leq u'Bu \leq \mu_i$ d'où la proposition.

Q.E.D.

Proposition 3 :

Soient f et g deux fonctions réelles de $L^1[-\pi, \pi]$;
$(f \leq g$ p.p.$) \Longrightarrow (M_n(f) \leq M_n(g),\ n \geq 1)$

Preuve : Soit $X = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$ un vecteur de $\mathbb{C}^n$; pour w dans $[-\pi ; \pi]$ on pose :

$u_X(w) = \sum_{s,t=1}^{n} x_s \bar{x}_t e^{i(s-t)w}$; Il vient alors : $\bar{X}' M_n(f) X = \int_{-\pi}^{\pi} u_X(w) f(w) dw$

et $\bar{X}' M_n(g) X = \int_{-\pi}^{\pi} u_X(w) g(w) dw$.

Or $u_X(w)$ s'écrit : $u_X(w) = \bar{X}' M_n(w) X$, où on a posé : $M_n(w) = [e^{i(s-t)w}]_{1 \le s,t \le n}$. $M_n(w)$ est hermitienne, de rang 1 ; sa seule valeur propre non nulle est donc égale à : $Tr\, M_n(w) = n$; on en déduit que $u_X(w) \ge 0$ $\forall X$, $\forall w$.

Par conséquent : $f \le g$ p.p. $\Longrightarrow u_X(w) f(w) \le u_X(w) g(w)$ p.p. et, par intégration : $\bar{X}' M_n(f) X \le \bar{X}' M_n(g) X$, $\forall X \in \mathbb{C}^n$; d'où $M_n(f) \le M_n(g)$, $n \ge 1$.

Q.E.D.

Corollaire 4 : Soient f et g, deux fonctions de $L^1([-\pi, \pi])$ à valeurs réelles et paires ; alors $f \le g$ p.p. $\Longrightarrow \lambda_\upsilon^{(n)}(f) \le \lambda_\upsilon^{(n)}(g)$, $\upsilon = 1, \ldots, n$; $n \ge 1$.

Preuve : f et g étant réelles, paires, $M_n(f)$ et $M_n(g)$ sont symétriques. Si $f \le g$ on a, par la proposition 3 : $M_n(f) \le M_n(g)$ et donc, par la proposition 2 : $\lambda_\upsilon^{(n)}(f) \le \lambda_\upsilon^{(n)}(g)$, $n \ge 1$, $\upsilon = 1, \ldots, n$.

Corollaire 5 : Si f est paire réelle et $m \le f \le M$ alors $m \le \lambda_\upsilon^{(n)}(f) \le M$.

Preuve : Soit $g = M$ alors $M_n(g) = M\, Id_n$ et $\lambda_\upsilon^{(n)}(g) = M$ pour tout n et υ. On en déduit par le corollaire 1 : $\lambda_\upsilon^{(n)}(f) \le M$, $\upsilon = 1, \ldots, n$; $n \ge 1$. On montre de même que : $m \le \lambda_\upsilon^{(n)}(f)$.

B - Approximation des matrices de Toeplitz

On désigne par $\|.\|$ la norme usuelle de matrice, c'est-à-dire le sup sur la boule unité. On définit de plus une norme, notée $|.|$, plus adaptée à l'étude des distributions de valeurs propres : $|(m_{ij})_{1 \le i,j \le n}| = \sqrt{\frac{1}{n} \sum_{i,j} |m_{ij}|^2}$.

Proposition 6 : Si A et B sont deux matrices à coefficients complexes, n x n :

$|AB| \le \|A\|\, |B|$.

Preuve : $|AB|^2 = \frac{1}{n} \sum_{i,j} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|^2$ avec $A = (a_{ij})_{i \leq i,\ j \leq n}$, $B = (b_{ij})$.

Notons b_j, le jième vecteur colonne de B : $b_j = \begin{bmatrix} b_{1j} \\ b_{2j} \\ \cdot \\ \cdot \\ \cdot \\ b_{nj} \end{bmatrix}$, $j = 1, \ldots, n$.

Alors $\sum_{i=1}^{n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|^2 = \|Ab_j\|_2^2$, $\|.\|_2$ étant la norme euclidienne usuelle dans $\mathbb{C}^n$. Puisque $\|Ab_j\|_2 \leq \|A\| \ \|b_j\|_2$, il vient $|AB|^2 \leq \sum_{j=1}^{n} \|A\|^2 \ \|b_j\|_2^2$, d'où $|AB|^2 \leq \|A\|^2 \ |B|^2$.

Q.E.D.

On peut, dans ces conditions, établir un théorème d'approximation :

Remarquons d'abord que $|A|^2 = \frac{1}{n} \mathrm{Tr}[A\bar{A}'] = |\bar{A}'|^2$; de plus $|AB|^2 = |\bar{B}'\bar{A}'|^2$; enfin si A est hermitienne, de valeurs propres $\lambda_1, \ldots, \lambda_n$: $|A|^2 = \frac{1}{n} \sum_{\upsilon=1}^{n} \lambda_\upsilon^2$.

Proposition 7 : [Grenander et Szegö], (1958)].

Soient $(T_n)_{n \geq 1}$ et $(L_n)_{n \geq 1}$, T_n matrices de Toeplitz hermitiennes n x n, L_n matrices hermitiennes n x n ; on suppose de plus que :

1) la distribution empirique des valeurs propres de L_n converge en loi vers une fonction de répartition $\mathcal{F}$;

2) $(L_n)_{n \geq 1}$, $(T_n)_{n \geq 1}$ sont uniformément bornées en norme $\|.\|$ et $|.|$ par $K < +\infty$;

3) $\lim_{n \to +\infty} |T_n - L_n| = 0$.

Alors la distribution empirique des valeurs propres de T_n converge en loi vers la répartition $\mathcal{F}$.

Preuve : Posons $T_n = L_n + \Delta_n$, $n \geq 1$; $(\Delta_n)_{n \geq 1}$ est donc une suite de

matrices hermitiennes. Pour un entier s, $s \geq 1$ on a : $T_n^s = L_n^s + R_n$, où R_n est somme de $2^s - 2$ termes de la forme $A_1 A_2 \cdots A_q$, $q \leq s$ avec $A_i = \Delta_n$ ou L_n. Il y a au moins une fois Δ_n dans chacun de ces termes. Comme $|AB| = |\bar{B}'\bar{A}'|$ (évident) on a : $|A_1 \cdots A_{r-1}\ \Delta_n\ A_{r+1} \cdots A_q| \leq \|A_1 \cdots A_{r-1}\|\ |\Delta_n\ A_{r+1} \cdots A_q|$, d'après la proposition 6.

$$\text{Or } |\Delta_n\ A_{r+1} \cdots A_q| = |\bar{A}_q' \cdots \bar{A}_{r+1}'\ \bar{\Delta}_n'| \leq \|\bar{A}_q' \cdots \bar{A}_{r+1}'\|\ |\Delta_n|.$$

$(T_n)_{n\geq 1}$ et $(L_n)_{n\geq 1}$ étant uniformément majorées par K pour les deux normes, il vient :

$|A_1 \cdots A_q| \leq K^{q-1}\ |\Delta_n|$ et donc : $|T_n^s - L_n^s| \leq K^s\ (2^s - 2)\ |\Delta_n|$, qui tend vers 0 avec n. Notant $\lambda_\upsilon^{(n)}$ (resp $\mu_\upsilon^{(n)}$). Les valeurs propres de T_n (resp L_n), on obtient donc :

$$\lim \frac{1}{n} \sum_{\upsilon=1}^{n} [\lambda_\upsilon^{(n)}]^s = \lim \frac{1}{n} \sum_{\upsilon=1}^{n} [\mu_\upsilon^{(n)}]^s,\ s \geq 1$$

et les deux suites ont donc même fonction de répartition $\mathcal{F}$.

Q.E.D.

ANNEXE 2

Δ-CONTINUITE ET UNIFORME CONTINUITE

1) Si h est Δ-continue, alors h est uniformément continue

Il nous faut montrer que si h est Δ - continue, si $(x_n)_{n\geq 1}$ et $(y_n)_{n\geq 1}$ sont deux suites dans $\mathcal{J}$ vérifiant $\lim_{n\to\infty} |x_n - y_n| = 0$, alors $\lim_{n\to\infty} |h(x_n) - h(y_n)| = 0$.

Considérons donc deux telles suites $(x_n)_{n\geq 1}$ et $(y_n)_{n\geq 1}$. On peut alors introduire la suite d'éléments $[x_K]_{K\geq 1}$ de $\mathcal{E}(\mathcal{J})$ définie par :

$$[x_K] = \{x_\upsilon^{(n)}(K),\ \upsilon = 1, \ldots, n\}_{n\geq 1},$$

où : $$x_\upsilon^{(n)}(K) = \begin{cases} y_n, & \text{si } n < K, \\ x_n, & \text{si } n \geq K. \end{cases}$$

On pose d'autre part :

$$[y] = \{(y_\upsilon^{(n)})\ \upsilon = 1, \ldots n\}_{n\geq 1}, \text{ avec } y_\upsilon^{(n)} = y_n\ \forall \upsilon, n.$$

Dans ces conditions, on a :

$$\Delta([x_K], [y]) = \sup_{n\geq K} \frac{|x_n - y_n|}{1 + |x_n - y_n|}$$

Comme $\lim_{n\to\infty} |x_n - y_n| = 0$, on voit que la suite $([x_k])_{K\geq 1}$ converge vers $[y]$ pour la distance Δ. Comme la fonction h est Δ-continue, il en résulte que :

$$\lim_{K\to\infty} \Delta[\tilde{h}([x_K]), \tilde{h}([y])] = 0,$$

ou encore que :

$$\lim_{n\to\infty} |h(x_n) - h(y_n)| = 0.$$

Q.E.D.

2) Si h est uniformément continue bornée, alors h est Δ - continue.

On a en fait montré une propriété un peu plus forte qui servira aussi pour établir la proposition suivante. De façon plus précise, nous considérons deux suites de $\mathcal{E}(\mathcal{J})$, notée $([x_p])_{p\geq 1}$, $([y_p])_{p\geq 1}$ telles que

$\lim_{p\to\infty} \Delta([x_p], [y_p]) = 0$, et nous établissons que :
$\lim_{p\to\infty} \Delta(\tilde{h}([x_p]), \tilde{h}([y_p]) = 0.$

Considérons donc deux telles suites et un nombre ε strictement positif.

i) Comme h est uniformément continue, il existe $\alpha > 0$ tel que : $x, y \in \mathcal{J}$, $|x-y| < \alpha D \Longrightarrow |h(x) - h(y)| < \frac{\varepsilon}{3}$, où D est la constante définie par : $D = \frac{4 \sup |h|}{\varepsilon}$.

ii) De plus, il existe un entier K, tel que :

$$\forall p \geq K : \Delta([x_p], [y_p]) < \alpha ;$$

ceci entraîne : $\frac{1}{n} \sum_{\upsilon=1}^{n} |x_\upsilon^{(n)}(p) - y_\upsilon^{(n)}(p)| < \alpha$, $n \geq 1$, $p \geq K$.

Introduisons alors l'ensemble d'indices :

$$H_p^{(n)} = \{\upsilon \in \{1,\ldots, n\}, |x_\upsilon^{(n)}(p) - y_\upsilon^{(n)}(p)| > \alpha D\}.$$

Il vient : $\alpha > \frac{1}{n} \sum_{\upsilon=1}^{n} |x_\upsilon^{(n)}(p) - y_\upsilon^{(n)}(p)|$

$$\geq \frac{1}{n} \sum_{\upsilon \in H_p^{(n)}} |x_\upsilon^{(n)}(p) - y_\upsilon^{(n)}(p)| > \alpha D \frac{\text{Card } H_p^{(n)}}{n}.$$

Ceci implique : $\frac{\text{Card } H_p^{(n)}}{n} < \frac{1}{D}$, si $p \geq K$.

Par conséquent :

$$\frac{1}{n} \sum_{\upsilon \in H_p^n} |h(x_\upsilon^{(n)}(p)) - h(y_\upsilon^{(n)})| \leq \frac{2 \sup |h|}{n} \text{Card } H_p^{(n)} \leq \frac{\varepsilon}{2},$$

pour $p \geq K$; puisque par ailleurs, on a :

$$\frac{1}{n} \sum_{\upsilon \notin H_p^n} |h(x_\upsilon^{(n)}(p)) - h(y_\upsilon^{(n)}(p)| \leq \frac{\varepsilon}{2}$$

on en déduit que $\Delta\{\tilde{h}([x_p]), \tilde{h}([y_p])\} < \varepsilon$, c'est-à-dire la Δ-continuité.

Q.E.D.

3) <u>Si h est uniformément continue sur $\mathcal{J} = [m, +\infty[$ et si elle est k - lipschitzienne sur $[M, +\infty[$, alors h est Δ - continue.</u>

Soit $([x_p])_{p\geq 1}$ une suite de $\mathcal{E}(\mathcal{J})$ convergeant vers $[x]$ pour la métrique Δ. On a donc : $\sup_n \frac{1}{n}\sum_{\upsilon=1}^{n} |x_\upsilon^{(n)}(p) - x_\upsilon^{(n)}| \xrightarrow[p\to\infty]{} 0.$

On décompose cette somme de la manière suivante :

$$\frac{1}{n}\sum_{\upsilon=1}^{n} |x_\upsilon^{(n)}(p) - x_\upsilon^{(n)}| = \frac{1}{n}\sum_{\substack{x_\upsilon^{(n)}(p)\leq M \\ x_\upsilon^{(n)} \leq M}} |x_\upsilon^{(n)}(p) - x_\upsilon^{(n)}| + \frac{1}{n}\sum_{\substack{x_\upsilon^{(n)}(p)\leq M \\ x_\upsilon^{(n)} > M}} (x_\upsilon^{(n)} - M)$$

$$+ \frac{1}{n}\sum_{\substack{x_\upsilon^{(n)}(p)\leq M \\ x_\upsilon^{(n)} > M}} (M - x_\upsilon^{(n)}(p)) + \frac{1}{n}\sum_{\substack{x_\upsilon^{(n)}(p)>M \\ x_\upsilon^{(n)} \leq M}} (M - x_\upsilon^{(n)}) + \frac{1}{n}\sum_{\substack{x_\upsilon^{(n)}(p)>M \\ x_\upsilon^{(n)} \leq M}} (x_\upsilon^{(n)}(p) - M)$$

$$+ \frac{1}{n}\sum_{\substack{x_\upsilon^{(n)}(p)>M \\ x_\upsilon^{n} > M}} |x_\upsilon^{(n)}(p) - x_\upsilon^{(n)}|$$

Dans cette écriture tous les termes sont positifs et il y a donc convergence vers 0 uniformément en n de chacun des six termes, lorsque p tend vers l'infini.

Il y a deux types de sommes :

a) Les termes sommés sont différence de deux réels se trouvant dans l'intervalle [m, M]. On mène alors le calcul de la manière suivante.

Pour la première somme $A = \frac{1}{n}\sum_{\substack{x_\upsilon^{(n)}(p) \leq M \\ x_\upsilon^{(n)} \leq M}} |x_\upsilon^{(n)}(p) - x_\upsilon^{(n)}|$ par exemple.

On définit $[a_p]$ et $[\tilde{a}_p]$ par :

$$\begin{cases} a_\upsilon^{(n)}(p) = x_\upsilon^{(n)}(p) \text{ et } \tilde{a}_\upsilon^{(n)}(p) = x_\upsilon^{(n)}, \text{ si } x_\upsilon^{(n)}(p) \leq M,\ x_\upsilon^{(n)} \leq M, \\ a_\upsilon^{(n)}(p) = \tilde{a}_\upsilon^{(n)}(p) = M, \text{ sinon.} \end{cases}$$

On a alors : $A = \frac{1}{n} \sum_{\upsilon=1}^{n} |a_\upsilon^{(n)}(p) - \tilde{a}_\upsilon^{(n)}(p)|$ et

$\lim_{p\to\infty} \Delta([a_p], [\tilde{a}_p]) = 0.$

Si on pose : $h^* = \begin{cases} h, \text{ sur } [m, M], \\ h(M), \text{ sur } [M, +\infty[\end{cases}$ on obtient une fonction h^* uniformément continue <u>bornée</u> sur $\mathcal{J} = [m, +\infty[$. D'après la preuve de 2), on en déduit que :

$$\lim_{p\to\infty} \Delta(\tilde{h}^*([a_p]), \tilde{h}^*([\tilde{a}_p]) = 0,$$

c'est-à-dire : $\lim_{p\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{n} (h^*[a_\upsilon^{(n)}(p)] - h^*[\tilde{a}_\upsilon^{(n)}(p)]) = 0.$

Remplaçant en fonction des suites $x_\upsilon^{(n)}(p)$ et $x_\upsilon^{(n)}$ initiales, il vient : $\lim_{p\to+\infty} \sup_n \frac{1}{n} \sum_{\substack{x_\upsilon^{(n)}(p) \le M \\ x_\upsilon^{(n)} \le M}} [h(x_\upsilon^{(n)}(p)) - h(x_\upsilon^{(n)})] = 0.$

b) Pour les termes qui s'obtiennent en sommant la différence de deux réels appartenant à $[M, +\infty[$, on utilise le fait que la fonction h est k-lipschitzienne :

$$\forall x, y \in [M, +\infty[: |h(x) - h(y)| \le k|x - y|.$$

Considérant par exemple le dernier terme, ceci permet immédiatement d'écrire la majoration :

$$\frac{1}{n} \sum_{\substack{x_\upsilon^{(n)}(p) > M \\ x_\upsilon^{(n)} > M}} |h(x_\upsilon^{(n)}(p)) - h(x_\upsilon^{(n)})| \le \frac{k}{n} \sum_{\substack{x_\upsilon^{(n)}(p) > M \\ x_\upsilon^{(n)} > M}} |x_\upsilon^{(n)}(p) - x_\upsilon^{(n)}|,$$

et d'en déduire que $\sup_n \frac{1}{n} \sum_{\substack{x_\upsilon^{(n)}(p) > M \\ x_\upsilon^{(n)} > M}} |h(x_\upsilon^{(n)}(p)) - h(x_\upsilon^{(n)})|$ tend vers zéro, lorsque p tend vers l'infini.

Rassemblant les résultats obtenus pour les diverses sommes, on en déduit que :

$$\lim_{p\to\infty} \sup_n \frac{1}{n} \sum_{\upsilon=1}^{n} |h(x_\upsilon^{(n)}(p)) - h(x_\upsilon^{(n)})| = 0$$

c'est-à-dire la Δ-continuité de la fonction h.

Q.E.D.

ANNEXE 3

ETUDE DES VALEURS PROPRES DANS LE CAS SINGULIER

i) Preuve de la propriété 3.14

a) Les autocovariances du processus $(X_t,\ t \in \mathbb{Z})$ s'écrivent :

$$\gamma(h) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(iwh)\ \varphi(w)\ dw + \sum_{k=1}^{r}\theta_k\ \exp(iw_k h).$$

Introduisant la matrice (n, n) définie par $E_n(w) = [\exp iw(s-r)]_{1\le s,r\le n}$, on obtient la décomposition ci-dessous de la matrice Γ_n :

$$\Gamma_n = \Gamma_n^a + \sum_{k=1}^{r}\theta_k\ E_n(w_k).$$

b) Remarquons d'abord que Γ_n^a est une matrice de Toeplitz associée à une fonction bornée positive. Ses valeurs propres sont alors bornées supérieurement par supess φ [et inférieurement par infess φ].

c) Considérons par ailleurs la matrice $\xi_n = \sum_{k=1}^{r}\theta_k\ E_n(w_k)$ associée à la partie discrète et introduisons les vecteurs :

$$d_n'(w_j) = \sqrt{\theta_j}\ (1,\ \exp iw_j, \ldots,\ \exp i(n-1)\ w_j),\ j = 1, \ldots,\ r$$

Ces vecteurs sont indépendants entre eux. En effet il est classique que le déterminant de Vandermonde :

$$\Delta_n = \det \begin{bmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ 1 & x_2 & \ldots & x_2^{n-1} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 1 & x_n & \ldots & x_n^{n-1} \end{bmatrix} = \prod_{1\le i<j\le n}(x_i - x_j),$$

(voir Lelong-Ferrand (1974)). Prenant $x_j = \exp iw_j$ avec des valeurs w_j, j=1,...n distinctes, on voit que le déterminant de Vandermonde associé est non nul ce qui prouve bien l'indépendance des vecteurs $d_n(w_j)$, $j = 1, \ldots, r$.

On peut alors introduire une base de $\mathbb{C}^n$ du type :
$\hat{\mathcal{B}} = \{d_n(w_1), \ldots, d_n(w_r), g_{r+1}^{(n)}, \ldots g_k^{(n)}, \ldots g_n^{(n)}\}$, où les vecteurs $g_k^{(n)}$ $k = r+1, \ldots n$ engendrent l'orthogonal de $\{d_n(w_j),\ j = 1, \ldots r\}$ dans $\mathbb{C}^n$.

Par définition de ces vecteurs complétant la base, on a :

$$E_n(w_k)\ g_s^{(n)} = 0 \quad \forall k = 1,\dots r,\ s = r+1,\dots, n.$$

Par ailleurs, nous avons :

$$E_n(w_k)\ d_n(w_j) = a_n(w_k - w_j)\ d_n(w_k),$$

avec : $a_n(w) = \begin{cases} n, \text{ si } w = 0, \\ \dfrac{1 - \exp(-iwn)}{1 - \exp(-iw)}, \text{ sinon.} \end{cases}$

Par conséquent exprimée dans la base $\hat{\mathcal{B}}$ la matrice ξ_n devient :

$$\hat{\xi}_n = \begin{pmatrix} \sqrt{\theta_j}\ \sqrt{\theta_k}\ a_n(w_k - w_j) & 0 \\ 0 & 0 \end{pmatrix}.$$

Comme pour $j \neq k$, on a l'inégalité : $|a_n(w_k - w_j)| \leq \dfrac{2}{|1 - \exp - i(w_k - w_j)|}$, et que les valeurs w_k sont toutes distinctes, il existe une constante K uniforme en n, k, j et majorant $|a_n(w_k - w_j)|$.

Soient alors A_n la matrice nord-ouest de taille (r, r) de $\hat{\xi}_n$, $\alpha_i^{(n)}$ $i = 1,\dots r$ les valeurs propres de cette matrice rangées par ordre décroissant et $D_n = \text{diag}[n\theta_1,\dots, n\theta_r]$.

La matrice A_n est hermitienne et donc par le théorème de Fisher-Courant [voir Châtelin (1988)], on a :

$$\max_{1\leq i\leq r} |\alpha_i^{(n)} - n\theta_i| \leq \|A_n - B_n\|.$$

$$\text{De plus : } \|A_n - B_n\| \leq \sup_k \left[\sum_{j\neq k} (\sqrt{\theta_j}\ \sqrt{\theta_k}\ |a_n(w_k - w_j)|)^2\right]^{\frac{1}{2}} \leq r\theta_1 K.$$

On voit donc que les valeurs propres de la matrice $\hat{\xi}_n$ (et donc de ξ_n) sont toutes nulles, sauf r d'entre elles qui vérifient :

$$n\theta_i - r\theta_i K \leq \alpha_i^{(n)} \leq n\theta_i + r\theta_i K, \qquad i = 1,\dots, r.$$

d) D'autre part $\|\Gamma_n - \xi_n\| = \|\Gamma_n^e\|$ et appliquant de nouveau le théorème de Fisher-Courant avec la convention $\alpha_\upsilon^{(n)} = 0$, si $\upsilon \geq r+1$, on a : $|\lambda_\upsilon^{(n)} - \alpha_\upsilon^{(n)}| \leq \|\Gamma_n^e\|$ $\qquad \upsilon = 1,\dots, n.$

Le résultat en découle en prenant comme constante :

$$M_0 = \sup\,[r\theta_1 K\ ;\ \operatorname{supess}\,\varphi].$$

ii) Preuve du lemme 3.16

On démontre ici le lemme en utilisant divers résultats énoncés dans [Grenander, Szegö (1958)] ; quand leur démonstration n'y est qu'esquissée, on fournira une preuve complète.

Soit α une distribution dont la partie absolument continue admet une densité φ bornée avec infess $\varphi = m > 0$; soit de plus : $C_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{inw} d\alpha(w)$ pour $n \geq 0$; on pose alors $D_n = \det\,(C_{\upsilon-\mu})_{0\leq\upsilon,\mu\leq n}$. $C_{\upsilon-\mu}$ s'interprète comme le produit hermitien de $\exp i\upsilon w$ et $\exp i\mu w$ dans $L^2([-\pi\ ;\ \pi])$ muni de sa structure d'espace hilbertien : $C_{\upsilon-\mu} = (\exp i\upsilon w, \exp i\mu w)$, avec : $(f, g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f\bar{g}\, d\alpha(w)$.

Proposition 1 : Pour tout $n \geq 0$, le déterminant D_n est strictement positif.

Preuve : Soit x un vecteur de $\mathbb{C}^{n+1}$. On a : $X'(C_{\upsilon-\mu})_{0\leq\upsilon,\mu\leq n}\, X = \int_{-\pi}^{\pi} \left|\sum_{h=0}^{n} x_h e^{-ihw}\right|^2 d\alpha(w)$.

Puisque $\varphi \geq m > 0$ p.p., cette intégrale n'est nulle que si $X = 0$; $(C_{\upsilon-\mu})_{0\leq\upsilon,\mu\leq n}$ est donc une forme hermitienne définie positive. D'où $D_n > 0$, pour $n \geq 0$.

On construit alors un système $\Phi_0, \ldots, \Phi_n, \ldots$ de polynômes en $z = e^{iw}$, orthonormal pour le produit $(.,.)$ défini ci-dessus.

Proposition 2 : Soit $\Phi_n(z) = (D_{n-1}D_n)^{-\frac{1}{2}} \det \begin{Vmatrix} C_0 & C_1 & \cdots & C_n \\ C_{-1} & C_0 & \cdots & C_{n-1} \\ \vdots & & & \\ C_{n+1} & C_{-n+2} & \cdots & C_0 \\ 1 & z & \cdots & z^n \end{Vmatrix}$, $n \geq 1$

$\vdots$

$\Phi_0(z) = 1$

alors : $\frac{1}{2\pi}\int_{-\pi}^{\pi} \Phi_n \bar{\Phi}_m\, d\alpha(w) = \begin{cases} 0, & \text{si } m \neq n \\ 1, & \text{si } m = n \end{cases}$ $m, n \geq 0$.

Preuve : Notons $z^h = f_h(w)$, $h \geq 0$; d'après l'interprétation donnée aux coefficients C_n, $n \geq 0$, il vient :

$$\Phi_n(z) = \frac{1}{\sqrt{D_{n-1} D_n}} \det \begin{bmatrix} (f_0, f_0) & \dots & (f_n, f_0) \\ (f_0, f_1) & \dots & (f_n, f_1) \\ \vdots & & \\ (f_0, f_{n-1}) & \dots & (f_n, f_{n-1}) \\ 1 & \dots & z^n \end{bmatrix}.$$

On notera, pour simplifier $\Phi_n(z) = \frac{1}{\sqrt{D_n D_{n-1}}} \det \begin{bmatrix} A \\ 1 \dots z^n \end{bmatrix}$.

Il est immédiat que, pour toute fonction g :

$$(\Phi_n, g) = (D_n D_{n-1})^{-\frac{1}{2}} \det \begin{bmatrix} A \\ (1,g) \ (z,g) \dots (z^n,g) \end{bmatrix}.$$

De sorte que $(\Phi_n, f_k) = 0$, si $k \leq n-1$, puisqu'alors la dernière ligne du déterminant est $[(f_0, f_k) \dots (f_n, f_k)]$ et donc est égale à la $(k+1)$ième ligne.

Φ_p étant, pour tout $p \geq 0$, combinaison linéaire de $f_0, f_1, \dots, f_p$, on a $(\Phi_n, \Phi_m) = 0$ si $m \leq n-1$.

D'autre part $(\Phi_n, \Phi_n) = (D_n D_{n-1})^{-\frac{1}{2}} \det \begin{bmatrix} A \\ (f_0,\Phi_n) \dots (f_n, \Phi_n) \end{bmatrix}$ $(n \geq 1)$

D'après ce qui précède :

$$(\Phi_n, \Phi_n) = (D_n D_{n-1})^{-1/2} \det \begin{bmatrix} A \\ 0, \dots, 0, (f_n, \Phi_n) \end{bmatrix},$$

$$(\Phi_n, \Phi_n) = \frac{(f_n, \Phi_n) D_{n-1}}{\sqrt{D_{n-1} D_n}} \text{ ; or } (f_n, \Phi_n) = (\overline{\Phi_n, f_n}) = D_n.$$

Q.E.D.

Ces polynômes $\Phi_1, \dots, \Phi_n \dots$ permettent de calculer le minimum d'une intégrale :

Proposition 3 : $\sqrt{\frac{D_n}{D_{n-1}}} \Phi_n(z)$, minimise l'intégrale $\frac{1}{2\pi} \int_{-\pi}^{\pi} |g(z)|^2 d\alpha(w)$, quand g décrit l'ensemble des polynômes unitaires de degré n. Le minimum lui-même

vaut $\frac{D_n}{D_{n-1}}$.

Preuve : $\Phi_0, \ldots, \Phi_n$ engendrent clairement l'espace des polynômes de degré n et le coefficient de z^n dans Φ_n est $k_n = \sqrt{\frac{D_{n-1}}{D_n}}$. Tout polynôme unitaire de degré n, g(z) s'écrit alors :

$$g(z) = \sum_{j=0}^{n} v_j \, \Phi_j(z), \quad v_j \in \mathbb{C} \quad \text{et} \quad v_n k_n = 1.$$

On en déduit : $\frac{1}{2\pi} \int_{-\pi}^{\pi} |g|^2 \, d\alpha(w) = \sum_{j=0}^{n} |v_j|^2 \geq |v_n| = \frac{1}{k_n^2} = \frac{D_n}{D_{n-1}}$.

Q.E.D.

La limite de ce minimum est fournie par le théorème suivant pour la démonstration duquel on renvoie à [Grenander, Szegö (1958)] (p. 44 et suiv.).

Théorème : Soit α une distribution de type infini, dont la partie absolument continue est notée $\beta(w)\, dw$, $-\pi \leq w \leq \pi$.

On considère $u_n(u_0, \ldots, v_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{j=0}^{n} u_j \, z^j \right|^2 d\alpha(w)$, $u_j \in \mathbb{C}$ et on pose : $\mu_n = \min_{\substack{u_0, \ldots, u_n \\ u_0 = 1}} T_n$.

Alors : $\lim_{n \to +\infty} \mu_n = \exp \left\{ \int_{-\pi}^{\pi} \text{Log}\beta(w) \, \frac{dw}{2\pi} \right\}$. Si $\text{Log}\beta(w)$ n'est pas intégrable, alors $\mu_n = 0$.

Par la proposition 3 et le théorème, on a donc (puisqu'ici Log φ est intégrable, φ est bornée et inf ess $\varphi = m > 0$) :

$$\lim \mu_n = \lim_{n \to +\infty} \frac{D_n}{D_{n-1}} = \exp \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log } \varphi(w) \, dw \right].$$

Ainsi, pour $n \geq n_0$, n_0 assez grand, on a en posant $J = \exp \int_{-\pi}^{\pi} \text{Log } \varphi(w) \, \frac{dw}{2\pi}$.

$\forall p \geq 0$, $(J - \varepsilon)^p \leq \frac{D_{n_0}}{D_{n_0 - 1}} \frac{D_{n_0+1}}{D_{n_0}} \ldots \frac{D_{n_0+p}}{D_{n_0+p-1}} \leq (J + \varepsilon)^p$, ε réel positif fixé.

Par conséquent :

$$\lim_{p\to+\infty} \frac{p}{p} \text{Log } (J - \varepsilon) \leq \lim_{p\to+\infty} \frac{1}{p} \log \frac{D_{n_0+p}}{D_{n_0-1}} \leq \lim_{p\to+\infty} \frac{p}{p} \text{Log } (J + \varepsilon),$$ si $\varepsilon > 0$ assez petit.

D'où : $\lim_{n\to+\infty} \frac{1}{n} \text{Log } D_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Log } \varphi(w)\, dw$ d'où le lemme. (On se reportera à [Szegö (1939)] pour la preuve originale de cette dernière égalité).

PARTIE 2 : PROCESSUS MULTIVARIES

RESUME :

Cette partie applique les résultats et les méthodes de la partie 1 au cas des processus stationnaires multidimensionnels. On calcule des approximations des valeurs propres quand le processus est AR(1) et MA(1) et l'on établit une version multidimensionnelle du théorème de Szegö.

ABSTRACT :

Using the results and the methods developped in the first part, we deal here with the multivariate case. We compute approximation of the eigenvalues when the process is AR(1) and MA(1) and we prove a multidimensional version of Szegö's theorem.

On se propose d'aborder l'étude des valeurs propres des matrices de variance-covariance de processus stationnaires multidimensionnels.

On considère ainsi des processus $\{X_t ; t \in \mathbb{Z}\}$, où X_t est un vecteur de $\mathbb{R}^p$, $p \geqslant 1$; aux autocovariances $\gamma(h)$ du cas univarié ($p = 1$), (cf. Accardo (1989)), correspondent alors des matrices $p \times p$ notées $\Gamma(h)$ et définies par : $\Gamma(h) = \text{cov}(X_t, X_{t-h})$. La matrice de variance-covariance du vecteur $np \times 1$: $(X_1', X_2', \ldots, X_n')'$ est alors une matrice de Toeplitz par blocs. On la notera Γ_n

$$\Gamma_n = \begin{bmatrix} \Gamma(0) & \Gamma(1) & \ldots & \ldots & \Gamma(n-1) \\ \Gamma'(1) & \Gamma(0) & \Gamma(1) & \ldots & \Gamma(n-2) \\ \vdots & \Gamma'(1) & \Gamma(0) & . & \vdots \\ \vdots & & . & . & \Gamma(1) \\ \Gamma'(n-1) & \Gamma'(n-2) & \ldots & \Gamma'(1) & \Gamma(0) \end{bmatrix} = (\Gamma(i-j))_{1 \leqslant i,\, j \leqslant n}.$$

Il en résulte, par rapport au cas univarié, des difficultés supplémentaires aussi bien du point de vue du calcul (le développement du moindre déterminant devient vite impraticable) que d'un point de vue plus structurel (les matrices de covariance des processus moyenne mobile d'ordre 1 (MA (1)) ne commutent pas entre elles). De fait il n'existe pas, à notre connaissance, de travaux consacrés aux valeurs propres des matrices de covariance de processus multivariés.

Ces remarques expliquent pourquoi l'étude qui suit, tout en restant fidèle à la démarche adoptée lors de l'étude des processus univariés, aboutit parfois à des résultats moins complets. Ainsi des approximations explicites des valeurs propres ne sont proposées que dans le cas de processus multivariés moyenne mobile et auto-régressif d'ordre 1 (MA(1) et AR(1)) (paragraphes 1 et 2).

L'intérêt particulier des AR(1) multivariés réside dans ce que l'étude des valeurs propres de leur matrice de variance-covariance va conduire à des procédures d'inversion des matrices de variance covariance de processus auto-régressifs d'ordre supérieur. De plus on obtiendra des approximations des valeurs propres extrêmes à partir du cas AR(1). On pourra alors, en utilisant les approximations obtenues, vérifier le théorème de Szegö (voir Accardo (1989)) pour les processus stationnaires MA(1) et AR(1) multivariés. Le paragraphe 3 est consacré à l'extension de ce théorème aux processus multivariés stationnaires généraux.

I APPROXIMATIONS DES VALEURS PROPRES D'UN PROCESSUS MA(1) MULTIVARIE :

Soit $\{X_t ; t \in \mathbb{Z}\}$ un processus multivarié, de dimension p, vérifiant un modèle MA(1) : $X_t = \varepsilon_t - \Phi \varepsilon_{t-1}$ où $\{\varepsilon_t ; t \in \mathbb{Z}\}$ est un bruit blanc et Φ une matrice p x p.

On note Ω la matrice de variance de ε_t :

$$(\Omega = V(\varepsilon_t))$$

La matrice Γ_n de variance-covariance du processus s'écrit alors :

$$\Gamma_n = \begin{bmatrix} \Omega + \Phi\Omega\Phi' & -\Phi\Omega & & 0 \\ -\Omega\Phi' & \Omega + \Phi\Omega\Phi' & -\Phi\Omega & \\ & & & -\Phi\Omega \\ 0 & & -\Omega\Phi' & \Omega + \Phi\Omega\Phi' \end{bmatrix}.$$

Comme il est noté dans l'introduction, la matrice Γ_n , n'est toeplitz que par l'intermédiaire des blocs, mais ne l'est pas élément par élément ; ceci interdit d'employer, à la différence du cas univarié, des méthodes fondées sur le développement de déterminants bien choisis (voir Grenander, Szegö (1958) ou Stroeker (1983) par exemple). On va alors revenir à l'idée de base de l'étude menée en univarié, qui consistait à "approcher" Γ_n par une matrice convenable. On choisira ici une "matrice cyclique-blocs", c'est-à-dire une matrice dont la décomposition blocs est telle que :

$$C = \begin{bmatrix} C_o & C_1 & C_2 & C_3 & \cdots & C_{n-2} & C_{n-1} \\ C_{n-1} & C_o & C_1 & C_2 & \cdots & C_{n-3} & C_{n-2} \\ C_{n-2} & C_{n-1} & C_o & C_1 & \cdots & \cdots & C_{n-3} \\ & & & & & & \\ \vdots & \vdots & \vdots & & & & C_1 \\ C_1 & C_2 & C_3 & \cdots & \cdots & C_{n-1} & C_o \end{bmatrix}$$

En d'autres termes : $C = (C_{\nu\mu})_{1 \leq \nu, \mu \leq n}$, où $C_{\nu\mu}$ est un bloc pxp et $C_{\nu\mu} = C_{\nu', \mu'} = C_{\nu-\mu}$ si $\nu-\mu = \nu'-\mu' (\mathrm{mod}\ n)$.

On remarquera que la matrice Γ_n d'un processus MA(1) n'est pas cyclique, au sens précis du terme, mais peut néanmoins être "approchée" par une telle matrice ; en effet la matrice C_n définie par :

$$C_n = \Gamma_n + E_n, \text{ avec } E_n = \begin{bmatrix} 0 & \dots & 0 & -\Omega\,\Phi' \\ \vdots & & & 0 \\ \vdots & & 0 & \vdots \\ 0 & & & \vdots \\ -\Phi\,\Omega & 0 & \dots\ \dots & 0 \end{bmatrix},$$

vérifie, elle, la condition de "cyclicité-blocs". On remarque de plus que C_n et Γ_n sont intuitivement "peu différentes".

L'utilisation de telles matrices cyclique-bloc est fréquente dans l'étude des processus unidimensionnels : que ce soit pour étudier l'inverse de Γ_n (voir Wise (1955), Whittle (1963)) ou pour établir la distribution d'un coefficient de corrélation (voir Dixon (1944) et Koopmans (1941)). On trouvera par ailleurs dans Grenander, Szegö (1958) une preuve du théorème de Szegö fondée sur l'usage de matrices cycliques.

L'importance des matrices cycliques et leur simplicité d'emploi tient fondamentalement au fait qu'elles sont reliées au groupe cyclique fini de la même façon que les matrices de Toeplitz infinies sont associées au groupe cyclique infini. (voir Kac, Murdoch, Szegö (1953)).

A notre connaissance de telles matrices n'ont cependant jamais été utilisées dans un contexte multivarié. On a alors la proposition suivante qui précise l'idée intuitive de "proximité" des matrices C_n et Γ_n.

Proposition 1 : On note $\lambda_\nu^{(n)}$ $\left(\text{resp. } \alpha_\nu^{(n)}\right)$, $\nu = 1,\dots,np$ les valeurs propres rangées dans l'ordre croissant de Γ_n (resp C_n). On a alors :

$$\alpha_\nu^{(n)} \leqslant \lambda_\nu^{(n)} \leqslant \alpha_{\nu+2p}^{(n)} \quad \nu = 1,2,\dots np - 2p.$$

Preuve : Soit Op (resp.Ip) la matrice nulle (resp.identité) pxp . On considère l'orthoprojecteur P, de rang (n-2)p défini par sa matrice dans la base canonique de $\mathbb{R}^{np}$

$$P = \begin{bmatrix} 0_p & & & \\ & I_p & & 0 \\ & \ddots & & \\ & 0 & I_p & \\ & & & 0_p \end{bmatrix}$$

P projette donc $\mathbb{R}^{np}$ sur le sous-espace L_{n-2} de dimension (n-2)p tel que $L_{n-2} = \{0\}^p \times \mathbb{R}^p \times \dots \times \mathbb{R}^p \times \{0\}^p$.

. On a clairement $P\Gamma_n = PC_n$ (puisque $PE_n = 0$)

D'autre part la restriction de $P\Gamma_n$ à L_{n-2} définit un opérateur autoadjoint de l'espace L_{n-2}. On rappelle le lemme (voir Accardo (1990)) :

Lemme : Si P est un projecteur orthogonal d'un espace L_{n+1} de dimension n+1 sur un sous espace L_n de dimension n on a, pour tout opérateur autoadjoint S de L_{n+1} :

a) La restriction $\tilde{S}$ de PS au sous espace L_n $\left(\tilde{S} = PS/L_n\right)$ définit un opérateur autoadjoint de L_n.

b) Si $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n \leqslant \lambda_{n+1}$ (resp. $\tilde{\lambda}_1 \leqslant \tilde{\lambda}_2 \leqslant \ldots \leqslant \tilde{\lambda}_n$) désignent les valeurs propres de S (resp. de $\tilde{S}$) ou a :

$$\lambda_1 \leqslant \tilde{\lambda}_1 \leqslant \lambda_2 \leqslant \tilde{\lambda}_2 \leqslant \ldots \leqslant \lambda_n \leqslant \tilde{\lambda}_n \leqslant \lambda_{n+1}.$$

. En vertu de ce lemme, appliqué 2p fois successives, les valeurs propres de $PM_{n/L_{n-2}}$ (donc celles de PC_n / L_{n-2}) sont séparées par les valeurs propres de C_n :

$\alpha_\nu^{(n)} \leqslant$ (val. propre d'indice ν de $P\Gamma_{n/L_{n-2}}$) $\leqslant \alpha_{\nu+2p}^{(n)}$ $\nu = 1, \ldots,$ np-2p.

.Si on considère ξ, vecteur de L_{n-2}, ξ s'écrit

$$\xi' = (0, \ldots 0, \hat{\xi}', 0, \ldots 0) \text{ où } \hat{\xi} \in \mathbb{R}^{(n-2)\,p}$$

et on a : $(P\Gamma_n)\,\xi = \begin{pmatrix} 0 \\ \cdots\cdots \\ \Gamma_{n-2}\,\hat{\xi} \\ \cdots\cdots \\ 0 \end{pmatrix}$.

Les valeurs propres de $P\Gamma_{n/L_{n-2}}$ coïncident donc avec celles de Γ_{n-2}, d'où la proposition.

Q.E.D.

L'intérêt de cette proposition tient à ce que les valeurs propres de C_n sont explicitement calculables en fonction de Φ et Ω

Proposition 2 : Les valeurs propres $\alpha_\nu^{(n)}$, $\nu = 1, \ldots, np$ de C_n sont les valeurs propres, rangées dans l'ordre croissant, des matrices :

$$A_k^{(n)} = \Omega + \Phi\,\Omega\,\Phi' - \exp\left[-\frac{2ik\pi}{n}\right]\Phi\,\Omega - \exp\left[\frac{2ik\pi}{n}\right]\Omega\,\Phi',$$

pour $o \leqslant k \leqslant n-1$.

Preuve : Posons $\mu_k^{(n)} = \exp\left(\frac{2ik\pi}{n}\right)$, $o \leqslant k \leqslant n-1$.

1) Chaque matrice $A_k^{(n)}$ est hermitienne et peut donc être diagonalisée sur une base orthonormée de vecteurs propres.

On notera $\theta_{i,k}^{(n)}$ $i = 1,\ldots,p$, les valeurs propres (réelles) de $A_k^{(n)}$ et $e_{i,k}^{(n)}$, $i = 1,\ldots p$ des vecteurs propres orthonormaux associés.

On considère alors le vecteur $Y_{i,k}^{(n)}$, np x 1, défini par :

$$\left[Y_{i,k}^{(n)}\right]' = \left[1,\ \mu_k^{(n)},\ \left(\mu_k^{(n)}\right)^2,\ldots,\ \left(\mu_k^{(n)}\right)^{n-1}\ \right] \otimes \left[e_{i,k}^{(n)}\right]'$$

On va montrer que les vecteurs $Y_{i,k}^{(n)}$ forment une base de vecteurs propres de C_n ;

2) Vérifions d'abord qu'ils sont vecteurs propres :

$C_n\ Y_{i,k}^{(n)}$ s'écrit matriciellement

$$\begin{bmatrix} \Omega + \Phi\,\Omega\,\Phi' & -\Phi\,\Omega & 0 & \ldots & 0 & -\Omega\,\Phi' \\ -\Omega\,\Phi' & \Omega + \Phi\,\Omega\,\Phi' & -\Phi\,\Omega & 0 & \ldots & 0 \\ 0 & -\Omega\,\Phi' & \cdot & \cdot & & \\ \vdots & & & \cdot & \cdot & \\ 0 & & & & \cdot & -\Phi\,\Omega \\ -\Phi\,\Omega & 0 & \ldots\ldots & 0 & -\Omega\,\Phi' & \Omega + \Phi\,\Omega\,\Phi' \end{bmatrix} \begin{bmatrix} e_{i,k}^{(n)} \\ \mu_k^{(n)}\ e_{i,k}^{(n)} \\ \left(\mu_k^{(n)}\right)^2\ e_{i,k}^{(n)} \\ \vdots \\ \left(\mu_k^{(n)}\right)^{n-2}\ e_{i,k}^{(n)} \\ \left(\mu_k^{(n)}\right)^{n-1}\ e_{i,k}^{(n)} \end{bmatrix} ;$$

tenant compte de ce que $\mu_k^{-(n)} = \frac{1}{\mu_k^{(n)}} = \left(\mu_k^{(n)}\right)^{n-1}$, on voit immédiatement que la $t^{ième}$ composante bloc pxl de $C_n\ Y_{i,k}^{(n)}$ s'écrit :

$$\left[(\Omega + \Phi\,\Omega\,\Phi')\ \left(\mu_k^{(n)}\right)^{t-1} - \Phi\,\Omega\ \left(\mu_k^{(n)}\right)^{t} - \Omega\,\Phi'\ \left(\mu_k^{(n)}\right)^{t-2}\right] e_{i,k}^{(n)} \text{ pour}$$

$t = 1,\ldots n$.

On retrouve donc en factorisant $\left(\mu_k^{(n)}\right)^{t-1}$:

$$\left(\mu_k^{(n)}\right)^{t-1} A_k^{(n)} e_{i,k}^{(n)}$$

et donc, d'après la définition de $e_{i,k}^{(n)}$ la $t^{ième}$ composante bloc pxl de C_n $Y_{i,k}^{(n)}$ est : $\theta_{i,k}^{(n)} Y_{i,k}^{(n)}$, $1 \leqslant i \leqslant p$, $o \leqslant k \leqslant n-1$

3) Vérifions qu'on a bien une base :

. D'abord les $e_{i,k}^{(n)}$ étant normés par hypothèse et puisque $\left|\mu_k^{(n)}\right| = 1$ pour tout n et k, on a facilement : $\left\|Y_{i,k}^{(n)}\right\| = \sqrt{n}$ ($\|.\|$ désigne ici la norme euclidienne sur $\mathbb{C}^{np}$)
$\frac{1}{\sqrt{n}} Y_{i,k}^{(n)}$ est donc un vecteur propre unitaire de C_n, associé à la valeur propre $\theta_{i,k}^{(n)}$ $1 \leqslant i \leqslant p$, $o \leqslant k \leqslant n-1$.

. On a alors,

$$\overline{\left[Y_{i,k}^{(n)}\right]'} \cdot \left[Y_{j,l}^{(n)}\right] = \begin{bmatrix} 1 \\ \bar{\mu}_k^{(n)} \\ \vdots \\ \vdots \\ \left(\bar{\mu}_k^{(n)}\right)^{n-1} \end{bmatrix} \otimes \bar{e}_{i,k}^{(n)\prime} \cdot \begin{bmatrix} 1 \\ \mu_l^{(n)} \\ \vdots \\ \vdots \\ \left(\mu_l^{(n)}\right)^{n-1} \end{bmatrix} \otimes e_{j,l}^{(n)}$$

$$= \left(\sum_{h=o}^{n-1} \left(\bar{\mu}_k^{(n)}\right)^h \right) \left[\bar{e}_{i,k}^{(n)\prime} \cdot e_{j,l}^{(n)} \right].$$

Si $k \neq l$ on a : $\sum_{h=o}^{n-1} \left(\bar{\mu}_k^{(n)}\right)^h \left(\mu_l^{(n)}\right)^h = \sum_{h=o}^{n-1} \exp \frac{2i\pi h\,(k-l)}{n} = o$;

si $k = l$ et si $i \neq j$: $\bar{e}_{i,k}^{(n)\prime} \cdot e_{j,l}^{(n)} = o$.

Donc $Y_{i,k}^{(n)}$ et $Y_{j,l}^{(n)}$ sont orthogonaux dès que $(i,k) \neq (j,l)$ {avec $1 \leqslant i \leqslant p$, $o \leqslant k \leqslant n-1$ } et $\left\{\frac{1}{\sqrt{n}} Y_{i,k}^{(n)} ;\ 1 \leqslant i \leqslant p,\ 0 \leqslant k \leqslant n-1\right\}$ constitue bien une base orthonormée de vecteurs propres de C_n associés aux valeurs propres $\theta_{i,k}^{(n)}$.

Q.E.D.

Dans le cas des processus moyenne mobile MA(1) univariés d'ordre 1, on sait (voir Accardo (1990)) que les valeurs propres de Γ_n coïncident avec

les valeurs $f\left(\frac{k\pi}{n+1}\right)$, $k = 1,\ldots,n$ prises par la densité spectrale f du processus en n points équirépartis sur $[o, \pi]$.

Les deux propositions précédentes ont montré, de leur côté, qu'on pouvait encadrer les valeurs propres de Γ_{n-2} par celles de C_n, qui sont, en particulier, les valeurs propres des matrices $A_k^{(n)}$; or il est immédiat de constater que si S désigne la matrice densité spectrale du processus $X_t = \epsilon_t - \Phi\epsilon_{t-1}$. on a :

$$S(\omega) = \Omega + \Phi\,\Omega\,\Phi - e^{-i\omega}\,\Phi\,\Omega - e^{i\omega}\,\Omega\,\Phi',$$

et donc que :

$$A_k^{(n)} = S\left(\frac{2k\pi}{n}\right)$$

On peut alors se demander si les valeurs propres des matrices Γ_n ne sont pas exactement les valeurs propres des n matrices $S\left(\frac{k\pi}{n+1}\right)$, $k = 1,\ldots n$ ou de toute autre suite de matrices correspondant aux valeurs de S en des points de $[o, \pi]$.

On va voir sur un exemple qu'il n'en est rien ; en effet, considérons le cas où $\Omega = Id_2$ et $\Phi = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

La densité spectrale $S(\omega) = I_2 + \Phi\,\Phi' - e^{-i\omega}\,\Phi' - e^{i\omega}\Phi$, $o \leqslant \omega \leqslant 2\pi$ s'écrit matriciellement :

$$S(\omega) = \begin{bmatrix} 2 & -e^{-i\omega} \\ -e^{i\omega} & 1 \end{bmatrix}.$$

Il est clair que ses valeurs propres $\alpha_1 = \frac{3-\sqrt{5}}{2}$, $\alpha_2 = \frac{3+\sqrt{5}}{2}$ sont indépendantes de ω

La matrice Γ_n du processus s'écrit, quant à elle :

$$\Gamma_n = \begin{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} & & & o \\ \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} & \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} & \ddots & & \\ & \ddots & & \ddots & \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \\ o & & \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} & \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} & \end{bmatrix}$$

Il est alors immédiat que det $(\Gamma_n - I_n)$ (resp.det $(\Gamma_n - 2I_n)$) a sa $2^{ème}$ (resp. $(2n-1)^{ième}$) colonne nulle ; (ici I_n = l'identité 2n x 2n) donc Γ_n admet $\lambda = 1$ et $\lambda' = 2$ comme valeurs propres.

Cet exemple appelle deux remarques :

1) Si on applique la proposition 1, on a :

$\alpha_\nu^{(n)} \leqslant \lambda_\nu^{(n-2)} \leqslant \alpha_{\nu+2p}^{(n)}$ $\nu = 1,\ldots,$ np - 2p, avec ici p = 2

or $\alpha_\nu^{(n)} = \alpha_1 = \dfrac{3 - \sqrt{5}}{2}$ dès que $\nu \leqslant n - 4$,

et $\alpha_\nu^{(n)} = \alpha_2 = \dfrac{3 - \sqrt{5}}{2}$ dès que $\nu \geqslant n + 1$.

On voit donc, au passage, que Γ_{n-2} admet n- 4 valeurs propres égales à $\alpha_1 = \dfrac{3 - \sqrt{5}}{2}$, n - 4 autres égales à $\alpha_2 = \dfrac{3 + \sqrt{5}}{2}$ les 4 restantes étant respectivement $\lambda = 1$ et $\lambda' = 2$ comptées avec leur ordre de multiplicité.

La méthode d'approximation proposée est donc particulièrement efficace dès que la matrice de densité spectrale a ses valeurs propres constantes.

2) En revanche il est clair qu'à la différence du cas univarié, les valeurs propres de Γ_n ne sont pas dans le cas général exactement les valeurs propres des matrices S (ω_k), avec $o \leqslant \omega_k \leqslant 2\pi$, k = 1,..., n (par exemple $\omega_k = \dfrac{k\pi}{n+1}$)

II APPROXIMATIONS DES VALEURS PROPRES D'UN AR(1) MULTIVARIE

Ici encore, les résultats obtenus dans le cadre des processus moyenne mobile vont servir de point de départ à l'analyse d'autres types de processus.

Ainsi l'étude du cas des processus AR (1) s'appuie sur la dualité processus autorégressif-processus moyenne mobile déjà mentionnée dans l'analyse du cas univarié (voir Accardo (1990) ; Cleveland (1972)).

On va montrer en particulier que si $\{X_t ; t \in \mathbb{Z}\}$ est un processus AR (1) : $X_t = \Phi X_{t-1} + \varepsilon_t$; $V(\varepsilon_t) = \Omega$

et si $\{Y_t ; t \in \mathbb{Z}\}$ est un processus MA (1) : $Y_t = \eta_t - \Phi'\eta_{t-1}$; $V(\eta_t) = \Omega^{-1}$

alors leurs matrices de covariance respectives $\Gamma_{n,x}$ et $\Gamma_{n,y}$ sont

"approximativement" inverses l'une de l'autre : $\Gamma_{n,x} \simeq \Gamma_{n,y}^{-1}$; on en déduira des valeurs approchées des valeurs propres de $\Gamma_{n,x}$.

En fait on établira un résultat un peu plus complet, en donnant l'expression exacte de l'inverse de Γ_n, pour tout processus auto-régressif d'ordre r.

a) Inversion de la matrice de covariance d'un processus auto-régressif multivarié :

L'inversion de la matrice de covariance d'un processus ARMA est un problème classique que l'on rencontre dans la méthode d'estimation par le maximum de vraisemblance, par exemple, ou encore dans le calcul des conséquences de certains chocs et dans la détermination de valeurs de mesure de causalité (voir Geweke (1982)).

Dans le cas des processus ARMA généraux, il n'existe pas d'expression analytique de la matrice inverse et l'inversion est souvent menée par des algorithmes (voir par exemple Akaïke (1973)).

Pour les cas les plus simples, c'est-à-dire pour des processus MA(2) ou AR(2), on connaît cependant l'expression générale exacte de l'inverse (voir Ali (1977) pour un résumé des différents résultats et aussi Shaman (1975), Shaman (1976)) mais, à notre connaissance, la question a toujours été abordée dans le cas univarié. Or, pour des processus autorégressifs "purs", c'est-à-dire sans partie moyenne mobile (AR(r)), on dispose d'une méthode simple proposée par Siddiqui (1958) (voir aussi Box, Jenkins (1976) ou Verbyla (1985)), qui se généralise directement au cas multivarié.

On considère donc un processus p-dimensionnel $\{X_t \ ; \ t \in \mathbb{Z}\}$ satisfaisant la formule autorégressive :

$$X_t = \sum_{i=1}^{r} \Phi_i \ X_{t-1} + \varepsilon_t ,$$

(Rappelons que, comme précédemment, les Φ_i sont des matrices pxp et que $\{\epsilon_t \ ; \ t \in \mathbb{Z}\}$ est un bruit blanc non corrélé, de variance Ω).

On fait ici l'hypothèse que Ω est inversible ; la non corrélation de ϵ_t avec x_{t-j} , $j \geqslant 1$ entraîne alors immédiatement l'inversibilité de la matrice de covariance np x np

$$G_n = \text{Var} \left\{ (X_n', \ X_{n-1}', \ldots, \ X_2', \ X_1')' \right\}$$

Remarquons que G_n n'est autre que la matrice de covariance d'ordre n du processus X, notée précédemment Γ_n, dans le cas où X est stationnaire.

On a alors la proposition :

Proposition 3 : La matrice inverse de G_n admet comme $(\alpha,\beta)^{ième}$ bloc de dimension p :

$$1)\ A_{\alpha\beta} = \sum_{k=o}^{n-1-r} \Phi'_{\alpha-k-1}\ \Omega^{-1}\ \Phi_{\beta-k-1} \quad \text{si } \alpha \leqslant n-r \text{ ou } \beta \leqslant n-r$$

$$2)\ A_{\alpha\beta} = \sum_{k=o}^{n-1-r} \Phi'_{\alpha-k-1}\ \Omega^{-1}\ \Phi_{\beta-k-1} + T_{\alpha-n+r,\ \beta-n+r}$$

$$\text{si } \alpha \geqslant n-r+1 \text{ et } \beta \geqslant n-r+1$$

On a désigné par T, la matrice inverse de G_r ($G_r = \text{Var}\left\{(X'_r, X'_{r-1}, \ldots, X'_2, X'_1)'\right\}$ et on a posé $\Phi_o = -Id_p$ et $\Phi_s = o$ si $s \notin \{0, 1, \ldots, r\}$

Preuve : Si le bruit blanc ϵ était gaussien, il en serait de même pour $(X'_n, X'_{n-1}, \ldots, X'_2, X'_1)'$. L'inverse de G_n apparaîtrait alors sous l'exponentielle dans l'expression de la densité $l(X_n, X_{n-1}, \ldots, X_2, X_1)$.

Faire comme si le bruit blanc était gaussien fournit donc un moyen simple de déterminer G_n^{-1}.

Sous l'hypothèse de normalité et en utilisant la structure autorégressive, nous pouvons écrire la densité :

$$l(X_n, X_{n-1}, \ldots, X_1) = \prod_{k=o}^{n-1-r} l(X_{n-k} / X_{n-k-1}, \ldots, X_1)\ l(X_r, \ldots, X_1)$$

$$= \prod_{k=o}^{n-1-r} l(X_{n-k} / X_{n-k-1}, \ldots, X_{n-k-r})\, l(X_r, \ldots, X_1)$$

$$= \prod_{k=o}^{n-r-1} \frac{1}{(2\Pi)^{p/2}\ (\det\Omega)^{1/2}} \left\{\exp -\frac{1}{2}\left[X_{n-k} - \sum_{i=i}^{r} \Phi_i X_{t-k-i}\right]' \Omega^{-1} \left[X_{n-k} - \sum_{i=i}^{r} \Phi_i X_{n-k-i}\right]\right\} \cdots$$

$$\cdots \frac{1}{(2\Pi)^{rp/2}\ (\det T)^{1/2}} \left\{\exp -\frac{1}{2}(X'_r, \ldots, X'_1)\ T \begin{bmatrix} X_r \\ X_{r-1} \\ \cdot \\ \cdot \\ \cdot \\ X_1 \end{bmatrix}\right\}.$$

Comme d'autre part on a :

$$l\,(X_n,\ X_{n-1},\ldots,\ X_1) = \frac{1}{(2\,\pi)^{nP/2}\ (\det G_n)^{1/2}} \exp -\frac{1}{2}\,(X_n',\ldots,\ X_1')\ G_n^{-1} \begin{pmatrix} X_n \\ X_{n-1} \\ \cdot \\ \cdot \\ \cdot \\ X_1 \end{pmatrix},$$

on obtient alors directement la forme des blocs $A_{\alpha\beta}$ en identifiant les formes quadratiques figurant sous l'exponentielle.

Q.E.D.

On voit donc que c'est la structure autorégressive qui donne facilement le résultat, que le processus soit ou non stationnaire.

D'un point de vue pratique la proposition permet de ramener l'inversion d'une matrice np x np à l'inversion de deux matrices de tailles plus petites, p x p pour Ω et pr x pr pour G_r. La formule d'inversion donne, dans le cas d'un processus autorégressif d'ordre 1 :

Proposition 4 :
Si $X_t = \Phi X_{t-1} + \varepsilon_t$ et si $\Gamma(0)$ désigne la variance de X_1, on a :

$$G_n^{-1} = \begin{bmatrix} \Omega^{-1} & -\Omega^{-1}\Phi & & & 0 \\ -\Phi'\Omega^{-1} & \Omega^{-1}+\Phi'\Omega^{-1}\Phi & -\Omega^{-1}\Phi & & \\ & -\Phi'\Omega^{-1} & \Omega^{-1}+\Phi'\Omega^{-1}\Phi & & \\ & & & \Omega^{-1}+\Phi'\Omega^{-1}\Phi & -\Omega^{-1}\Phi \\ & 0 & & -\Phi'\Omega^{-1} & \Gamma_{(0)}^{-1}+\Phi'\Omega^{-1}\Phi \end{bmatrix}$$

Preuve : immédiate.

Remarque 1 : l'égalité $X_t = \Phi X_{t-1} + \epsilon_t$ entraine, dans le cas stationnaire,

que $X_t = \sum_{h \geqslant 0} \Phi^h\ \epsilon_{t-h}$.

Si Φ, Ω et Φ' commutent, on a immédiatement :

$$\Gamma(0) = \Omega \sum_{h \geqslant 0} \Phi^h \Phi'^h = \Omega(I-\Phi\,\Phi')^{-1}$$

D'où $\Gamma(0)^{-1} = \Omega^{-1} - \Phi\,\Omega^{-1}\Phi'$ et donc, dans l'expression de G_n^{-1}, les termes diagonaux extrêmes Ω^{-1} et $\Gamma(0)^{-1} + \Phi'\Omega^{-1}\Phi$ sont égaux.

Un cas particulier est le cas univarié ($p = 1$) ; en effet on a alors : $\Omega = \sigma^2 \in \mathbb{R}_+$, $\Phi = \varphi \in \mathbb{R}$ ($|\varphi| < 1$) et donc $\Gamma(0) = \dfrac{\sigma^2}{1-\varphi^2}$ d'où $\Gamma(0)^{-1} + \Phi'\Omega^{-1}\Phi = \dfrac{1}{\sigma^2} = \Omega^{-1}$

$$\text{et } G_n^{-1} = \frac{1}{\sigma^2}\begin{bmatrix} 1 & -\varphi & & & \\ -\varphi & 1+\varphi^2 & -\varphi & 0 & \\ & -\varphi & 1+\varphi^2 & & \\ & & & 1+\varphi^2 & -\varphi \\ & 0 & & -\varphi & 1 \end{bmatrix}.$$

On retrouve la formule classique (voir par exemple Gourieroux-Monfort)(1990).

Ce résultat n'est cependant pas général, prenons en effet

$$\Omega = I_3 \ , \ \Phi = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} ; \ \Phi \text{ est nilpotente } (\Phi^3 = 0).$$

On a alors l'expression de $\Gamma(0)$:

$$\Gamma(0) = I_3 + \Phi\Phi' + \Phi^2\Phi'^2 \text{ c'est-à-dire } \Gamma(0) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} ;$$

$$\text{d'où } \Gamma(0)^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ et, clairement : } \Gamma(0)^{-1} + \Phi'\Phi = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & 2 \end{bmatrix} \neq \Omega^{-1} = I_3$$

Remarque 2 : Il suffit de comparer l'expression de G_n^{-1} dans la proposition, à l'expression de la matrice Γ_n de covariance d'un processus MA(1) : $Y_t = \eta_t - \Phi'\eta_{t-1}$, où $\{\eta_t \ ; \ t \in \mathbb{Z}\}$ est un bruit blanc, $V(n_t) = \Omega^{-1}$, pour voir immédiatement que I_n et G_n^{-1} coïncident, à l'exception des blocs p x p, nord-ouest et sud-est. C'est ce résultat qui est utilisé dans la suite pour déterminer les valeurs propres de G_n.

b) Valeurs propres d'un AR(1) multivarié.

Compte tenu de la proposition 4 et de la définition de la matrice C_n introduite dans le paragraphe I , pour tout processus autorégressif X, d'ordre 1, vérifiant $X_t = \Phi X_{t-1} + \epsilon_t$, $V(\epsilon_t) = \Omega$, on a :

$$G_n^{-1} = \tilde{C}_n + \begin{bmatrix} -\Phi'\Omega^{-1}\Phi & \vdots & \Omega^{-1}\Phi' \\ \vdots & \vdots & \vdots \\ 0 & \cdots\ 0\ \cdots & 0 \\ \vdots & \vdots & \vdots \\ \Phi\,\Omega^{-1} & & -\Omega^{-1}+\Gamma(0)^{-1} \end{bmatrix},$$

où $\tilde{C}_n$ est la matrice cyclique par blocs contruite comme indiqué au I, à partir du MA(1) p dimensionnel $Y_t = \eta_t - \Phi'\eta_{t-1}$ où $[\eta_t ; t \in \mathbb{Z}]$ est un bruit blanc (non corrélé) de matrice de variance Ω^{-1}. Il suffit alors de reprendre la preuve de la proposition 1, en remplaçant simplement la matrice Γ_n par G_n^{-1} et C_n par $\tilde{C}_n$ pour obtenir:

Proposition 5 : Soit $\lambda_1^{(n)} \leqslant \ldots \leqslant \lambda_{np}^{(n)}$ les valeurs propres de G_n et $\alpha_1^{(n)} \leqslant \ldots \leqslant \alpha_{np}^{(n)}$ celles des matrices

$$\tilde{A}_k^{(n)} = \Omega^{-1} + \Phi'\Omega^{-1}\Phi - e^{\frac{2ik\pi}{n}}\,\Omega^{-1}\,\Phi - e^{\frac{-2ik\pi}{n}}\,\Phi'\Omega^{-1}$$

(pour $k = 0,\ldots,n-1$) ; on a alors $\dfrac{1}{\alpha_{\nu+2p}^{(n)}} \leqslant \lambda_\nu^{(n-2)} \leqslant \dfrac{1}{\alpha_{\nu-2p+1}^{(n)}}$, $2p \leqslant \nu \leqslant np - 2p$.

Remarque : a proposition précédente n'est valable que si, bien sûr, les valeurs propres de $\tilde{C}_n$ sont non nulles ; autrement dit il faut que, pour tout k, $\tilde{A}_k^{(n)}$ soit inversible.

Or : $\tilde{A}_k^{(n)} = (I - e^{\frac{2ik\pi}{n}}\,\Phi)\,\Omega^{-1}\,(I - e^{\frac{2ik\pi}{n}}\,\Phi)$, $k = o,\ldots,\ n-1$.

Une condition suffisante est donc clairement : les valeurs propres de Φ sont strictement inférieures à 1, en module.

On sait, en particulier, que cette condition est équivalente à la stationnarité du processus X (voir Gouriéroux, Monfort (1990)).

c) Etude de l'erreur d'approximation

Dans la proposition 5 ci-dessus, les valeurs propres "extrêmes" de G_n (c'est-à-dire $\lambda_\nu^{(n)}$, $\nu \leqslant 2p-1$) ne sont pas encadrées. Le résultat suivant fournit quelques précisions sur la distribution de ces valeurs propres.

Proposition 6 : Soit M_n l'ensemble des valeurs propres de G_n^{-1} ; il existe une constante H, indépendante de n, telle que :
pour tout $j = 1,\ldots,\ pn$,

$$\min_{\mu \in M_n} \left|\mu - \alpha_j^{(n)}\right| \leqslant \frac{H}{\sqrt{n}}.$$

Preuve : D'après un résultat classique (voir Fox (1964)) on a pour tout vecteur unitaire et tout réel a :

$$\operatorname*{Min}_{\mu \in M_n} |\mu - a| \leqslant \left\| G_n^{-1} Y - a Y \right\|.$$

Appliquons ce résultat à $a = \alpha_j^{(n)}$ et $Y = Y_j^{(n)}$ un vecteur propre unitaire associé à la valeur propre $\alpha_j^{(n)}$;

On obtient : $$\operatorname*{Min}_{\mu \in M_n} \left|\mu - \alpha_j^{(n)}\right| \leqslant \left\|G_n^{-1}\ Y_j^{(n)} - \alpha_j^{(n)}\ Y_j^{(n)}\right\|$$
$$= \left\|\left(G_n^{-1} - \tilde{C}_n\right) Y_j^{(n)}\right\|.$$

On sait que :

$$G_n^{-1} = \tilde{C}_n + \begin{bmatrix} -\Phi'\Omega^{-1}\Phi & \vdots & \Omega^{-1}\Phi' \\ \vdots & \vdots & \vdots \\ 0 & \cdots\ 0\ \cdots & 0 \\ \vdots & \vdots & \vdots \\ \Phi\ \Omega^{-1} & & -\Omega^{-1}+\Gamma(0)^{-1} \end{bmatrix},$$

et d'autre part que $Y_j^{(n)}$ est de la forme (voir proposition 2)
$Y_j^{(n)} = \dfrac{1}{\sqrt{n}}\ (e', \xi e', \xi^2 e', \ldots, \xi^{n-1} e')'$, avec ξ racine $n^{ième}$ de l'unité et e vecteur propre unitaire d'une matrice $\tilde{A}_k^{(n)}$.

On obtient donc

$$\left\|\left(G_n^{-1} - \tilde{C}_n\right) Y_j^{(n)}\right\| \leqslant \frac{\|\Phi'\Omega^{-1}\Phi\| + \|\Omega^{-1}\Phi'\| + \|\Phi\Omega^{-1}\| + \|\Omega^{-1}-\Gamma(0)^{-1}\|}{\sqrt{n}}.$$

Il suffit donc de poser H égal au numérateur ci-dessus pour obtenir le résultat.

Q.E.D.

Remarque 4 : Une autre manière d'énoncer la proposition 6 est de dire que tout intervalle de la forme

$$\left[\frac{1}{\alpha_j^{(n)} + \dfrac{H}{\sqrt{n}}}\ ;\ \frac{1}{\alpha_j^{(n)} - \dfrac{H}{\sqrt{n}}}\right]$$ contient au moins une valeur propre de G_n. On obtient ainsi des informations d'autant plus précises que n est grand, sur les valeurs propres de G_n, en particulier sur les 2p plus petites, et sur l'ordre de grandeur de l'erreur commise en approchant les $\lambda_\nu^{(n)}$ par les $\dfrac{1}{\alpha_\nu^{(n+2)}}$.

II PROCESSUS MULTIVARIES ET THEOREME DE SZEGO

Rappelons le théorème de Szegö (voir Grenander, Szegö (1958) ou Accardo (1989)) :

Soit $\{X_t , t \in \mathbb{Z}\}$ un processus stationnaire univarié admettant une densité spectrale f intégrable. Alors si $o < m \leqslant f(\omega) \leqslant M$, $\omega \in [-\pi ; \pi]$ on a, pour toute fonction h continue sur [m,M] :

$$\lim \frac{1}{n} \sum_{\nu=1}^{n} h\left(\lambda_\nu^{(n)}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h\,(f\,(\omega))\, d\omega \quad ,$$

où $\lambda_\nu^{(n)}$ désigne les valeurs propres de la matrice de covariance Γ_n du processus X.

Il est alors naturel d'examiner ce résultat dans un cadre multivarié ; on donnera d'abord une idée de sa généralisation sur un exemple simple.

Soit $X_t = \varepsilon_t - \Phi\, \varepsilon_{t-1}$ un processus MA(1) de dimension p.
Sa densité spectrale s'écrit : $S(\omega) = (I - e^{-i\omega}\Phi)\ \Omega\ (I - e^{i\omega}\Phi')$ pour $\omega \in [-\pi , \pi]$ et $\Omega = \mathrm{Var}\,(\varepsilon_t)$.

Pour chaque ω, posons $\mu_1(\omega) \leqslant \mu_2(\omega) \leqslant \ldots \leqslant \mu_p(\omega)$ les valeurs propres (positives de manière évidente) de $S(\omega)$. On va voir, dans un cadre plus général, que $\omega \mapsto \mu_j(\omega)$ définit une fonction continue sur $[-\pi , \pi]$. Soit alors $m = \min_{\omega} \mu_1(\omega)$, $M = \max_{\omega} \mu_p(\omega)$. On montrera également que les valeurs propres $\lambda_\nu^{(n)}$ $\nu = 1,\ldots,np$ de M_n vérifient : $m \leqslant \lambda_\nu^{(n)} \leqslant M$.

Dans ces conditions on utilise les propositions 1 et 2 du paragraphe I pour approcher les $\lambda_\nu^{(n)}$ par les valeurs propres des matrices $A_k^{(n)} = \left(I - e^{\frac{2ik\pi}{n}} \Phi\right) \Omega \left(I - e^{\frac{2ik\pi}{n}} \Phi'\right)$, c'est-à-dire précisément les $\mu_j\left(\frac{2k\pi}{n}\right)$ $j = 1,\ldots, p$, $k = o,\ldots, n-1$.

De façon heuristique on a pour toute fonction h continue sur [m,M]

$$\frac{1}{n} \sum_{\nu=1}^{np} h\left(\lambda_\nu^{(n)}\right) \simeq \frac{1}{n} \sum_{k=o}^{n-1} \sum_{j=1}^{p} h\left(\mu_j\left(\frac{2k\pi}{n}\right)\right)$$

(pour n assez grand)

Le second membre est clairement une somme de Riemann et, les fonctions considérées étant toutes continues :

$$\lim \frac{1}{n} \sum_{\nu=1}^{np} h\left(\lambda_\nu^{(n)}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\sum_{j=1}^{p} h\left(\mu(\omega)\right) \right] d\omega.$$

Ainsi , moyennant le remplacement sous l'intégrale de la fonction densité spectrale par les valeurs propres de la matrice densité spectrale, on obtient l'analogue multivarié du théorème de Szegö.

C'est ce résultat qu'on va établir en toute généralité.

On considèrera dans tout ce qui suit un processus $X = \{X_t \ ; \ t \in \mathbb{Z}\}$ de dimension p., stationnaire admettant une matrice densité spectrale continue (en particulier X est donc représentable comme une moyenne mobile infinie $X_t = \sum_{-\infty}^{+\infty} \Psi_j \ \eta_{t-j}$ avec $\sum_{-\infty}^{+\infty} \|\Psi_j\|^2 < +\infty$, voir Doob (1953)).

On a alors les deux lemmes :

lemme 1 : (voir Rellich (1969)) Pour tout ω dans $[-\pi, \pi]$, $S(\omega)$ admet p valeurs propres positives $0 \leqslant \mu_1(\omega) \leqslant \mu_2(\omega) \leqslant \ldots \leqslant \mu_p(\omega)$ et pour tout j, $\omega \mapsto \mu_j(\omega)$ est continue sur $[-\pi, \pi]$.

Preuve : L'existence et la positivité des $\mu_j(\omega)$ découlent directement de la définition de $S(\omega)$ qui est hermitienne positive.

Montrons la continuité des μ_j.

Par l'hypothèse de continuité de S, on a la continuité des coefficients S_{ij} , $1 \leqslant i,j \leqslant p$, de la matrice densité spectrale.

Ces coefficients sont donc uniformément bornés sur $[-\pi, \pi]$ et on en déduit classiquement l'existence d'une constante K_0 telle que :

$$\sup_{j,\omega} \mu_j(\omega) \leqslant K_0 < +\infty$$

Soit alors ω_0 réel fixé de $]-\pi ; \pi[$ et $(h_n)_{n \geqslant 1}$ une suite de réels de limite nulle. Pour n_0 assez grand on a :

$$-\pi \leqslant \omega_0 + h_n \leqslant \pi, \quad n \geqslant n_0.$$

Dans ces conditions : $0 \leqslant \mu_j(\omega_0 + h_n) \leqslant K_0$ $j = 1, \ldots, p$, $n \geqslant n_0$; de chaque suite $(\mu_j(\omega_0 + h_n))_{n \leqslant n_0}$ on peut alors extraire une sous suite $\left(\mu_j(\omega_0 + h_{n_m})\right)_{m \geqslant 1}$ convergeant vers une limite notée μ_j.

On a d'autre part :

$$\det \left[S\ (\omega_o + h_{n_m}) - \lambda\ I_p \right] = (-1)^p \prod_{j=1}^{p} \left(\lambda - \mu_j\ (\omega_o + h_{n_m}) \right).$$

Faisant tendre m vers l'infini, on obtient, S étant continue

$$\det \left[S\ (\omega_o) - \lambda\ I_p \right] = \prod_{j=1}^{p} (-1)^p\ (\lambda - \mu_j)\ ;$$

l'ensemble des μ_j, $j = 1, \ldots, p$ coïncide donc (par identification des deux polynômes en λ) avec l'ensemble des valeurs propres $\mu_j\ (\omega_o)$ $j = 1, \ldots, p$ de $S\ (\omega_o)$.

On remarque alors que, par hypothèse :

$$\mu_j\ (\omega_o + h_{n_m}) \leqslant \mu_{j+1}\ (\omega_o + h_{n_m}),\ \forall_j,\ \forall m$$

On en déduit que $\mu_1 \leqslant \mu_2 \leqslant \ldots \leqslant \mu_p$ et donc, puisque on a aussi $\mu_1\ (\omega_o) \leqslant \mu_2\ (\omega_o) \leqslant \ldots \leqslant \mu_p\ (\omega_o)$, il vient $\mu_j = \lim_{m \to \infty} \mu_j\ (\omega_o + h_{n_m})$ $= \mu_j\ (\omega_o)$ $j = 1, \ldots, p$; d'où la continuité.

Q.E.D.

<u>lemme 2</u> : Les valeurs propres $\lambda_\nu^{(n)}$, $\nu = 1, \ldots, np$ de Γ_n, la matrice de covariance d'ordre n du processus X, vérifient :

$$\min_\omega \mu_1\ (\omega) \leqslant \lambda_\nu^{(n)} \leqslant \max_\omega \mu_p\ (\omega),\ \nu = 1, \ldots, np.$$

<u>Preuve</u> : X s'écrivant comme une moyenne mobile, on a la formule de représentation pour les autocovariances ($\Gamma\ (h) = \text{cov}\ (X_t, X_{t-h})$ $h \in \mathbb{Z}$) :

$$\Gamma\ (h) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-i\omega h}\ S\ (\omega)\ d\omega.$$

Soit $\xi' = (\xi_1', \xi_2', \ldots, \xi_n')$ un vecteur propre unitaire de Γ_n associé à une valeur propre λ (en particulier, ξ_i est un vecteur $p \times 1$ de $\mathbb{R}^p$, $i = 1, \ldots, n$).

On vérifie alors que :

$$\lambda = \xi'\ M_n\ \xi = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{1 \leqslant h,\ k \leqslant n} \left(\xi_h'\ S\ (\omega)\ \xi_k \right) e^{-i\ (h-k)\omega}\ d\omega\ ;$$

soit encore que :

$$\lambda = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \left(\sum_{h=1}^{n} \xi_h' e^{i\omega h} \right) S\ (\omega) \left(\sum_{k=1}^{n} \xi_k e^{-i\omega k} \right) d\omega.$$

Or on a, pour tout ω et tout vecteur Y de $\mathbb{R}^p$

$$\min_{\omega} \mu_1 (\omega) \quad \|Y\|_2 \leq Y' S (\omega) Y \leq \max_{\omega} \mu_p (\omega) \|Y\|_2 .$$

Notant M (resp m) ce maximum (resp ce minimum) il vient :

$$\frac{m}{2\pi} \int_{-\pi}^{\pi} \left\| \sum_{h=1}^{n} \xi_h e^{-i\omega h} \right\|_2^2 d\omega \leq \lambda \leq \frac{M}{2\pi} \int_{-\pi}^{\pi} \left\| \sum_{k=1}^{n} \xi_k e^{-i\omega k} \right\|_2^2 d\omega ;$$

l'égalité de Parseval et le fait que ξ est unitaire donnent le résultat :

$$\int_{-\pi}^{\pi} \left\| \sum_{h=1}^{n} \xi_h e^{-i\omega h} \right\|_2^2 d\omega = \sum_{h=1}^{n} \|\xi_h\|_2^2 \; 2\pi = 2\pi.$$

Q.E.D.

On peut alors montrer le résultat :

Proposition 7 : Soit $X = \{X_t ; t \in \mathbb{Z}\}$ un processus stationnaire de dimension p, admettant une matrice de densité spectrale continue $S : \omega \mapsto S(\omega)$, $\omega \in [-\pi, \pi]$.

On note $\mu_j : \omega \mapsto \mu_j (\omega)$, $\omega \in [-\pi, \pi]$ les valeurs propres de S rangées dans l'ordre croissant.

On pose $M = \max_{\omega} \mu_p (\omega)$ et $m = \min_{\omega} \mu_1 (\omega)$ et on note $\lambda_\nu^{(n)}$, $\nu = 1, \ldots,$ np les valeurs propres de Γ_n, la matrice de covariance d'ordre n de X.

Alors pour toute fonction h continue sur [m, M],

$$\text{on a : } \lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{np} h(\lambda_\nu^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=1}^{p} h(\mu_j (\omega)) \, d\omega .$$

Preuve : 1) La démonstration est analogue à celle du cas unidimensionnel (voir Accardo (1990)). Elle s'appuie en particulier sur l'égalité de Kolmogorov qui s'écrit, dans le cas multivarié, en supposant m > 0 :

$$\lim_{n\to\infty} \log \det V [X_t / X_{t-1}, \ldots, X_{t-n}] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \det S (\omega) \, d\omega.$$

($V(X_t / X_{t-1}, \ldots, X_{t-n})$ désigne la matrice de covariance du résidu dans la régression linéaire de X_t sur $\{X_{t-1}, \ldots, X_{t-n}\}$).
On a donc aussi convergence des moyennes de Cesaro :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\log \det V [X_t / X_{t-1}, \ldots, X_{t-k}]) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \det S(\omega)\, d\omega.$$

On exprime alors les matrices V en fonction de M_n

$$\det V [X_t / X_{t-1}, \ldots, X_{t-n}] = \frac{\det \Gamma_n}{\det \Gamma_{n-1}}$$

on a alors :

$\lim \frac{1}{n} \log \det \Gamma_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \det S(\omega) d\omega$ et utilisant les valeurs propres $\lambda_\upsilon^{(n)}$ et $\mu_j(\omega)$, on obtient : $\lim \frac{1}{n} \sum_{\upsilon=1}^{np} \log \lambda_\upsilon^{(n)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=1}^{p} \log \mu_j(\omega) d\omega$ qui est un cas particulier de l'égalité de la proposition, pour la fonction $h(x) = \log x$.

2) On raisonne alors tout à fait comme dans le cas unidimensionnel pour obtenir $\lim \frac{1}{n} \sum_{\upsilon=1}^{np} \log \left(1 + z \lambda_\upsilon^{(n)}\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=1}^{p} \log (1 + z \mu_j(\omega)) d\omega$ uniformément dans $\left\{|z| < \frac{1}{2M}\right\}$. On peut alors développer en série les deux membres de l'égalité et identifier les coefficients des séries ;

$\lim \frac{1}{n} \sum_{\upsilon=1}^{np} \left\{\lambda_\upsilon^{(n)}\right\}^s = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=1}^{p} \{\mu_j(\omega)\}^s dw$ pour tout $s \geqslant 1$ étant vérifiée pour toutes les fonctions puissances x^s , $s \geqslant 1$, la relation est vraie pour toute fonction h continue sur [m, M]

3) Si $m = 0$, il suffit d'appliquer la démarche précédente au processus $Y_t = X_t + \epsilon_t$, où ϵ_t est un bruit blanc non correlé avec X_t, avec $V(\epsilon_t) = I_p$; la matrice densité spectrale S_y vérifie en effet $S_y(\omega) = I_p + S(\omega)$, $-\pi \leqslant \omega \leqslant \pi$; en particulier ses valeurs propres sont minorées par $m + 1 = 1 > 0$.

On obtient alors pour toute fonction continue h sur [1, M+1]

$\lim_{n\to\infty} \frac{1}{n} \sum_{\upsilon=1}^{np} h \left[\lambda_\upsilon^{(n)} + 1\right] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=1}^{p} h (\mu_j(\omega)+1) dw$ et, par translation, l'égalité annoncée.

Q.E.D.

PARTIE 4 : PROCESSUS NON STATIONNAIRES

RESUME:

On étudie dans ce travail les valeurs propres des matrices de covariance de processus non stationnaires. On montre d'abord, sur l'exemple de la marche aléatoire, que ces valeurs propres vérifient de nombreuses propriétés qu'on croit généralement satisfaites par les seuls processus stationnaires.

On étend alors ces résultats aux processus vérifiant une équation autorégressive; on prouve en particulier une généralisation d'un théorème de Szegö concernant la distribution asymptotiques de ces valeurs propres.

On établit enfin deux caractérisations du degré de non stationnarité d'un processus vérifiant un modèle ARIMA, l'une en s'appuyant sur l'examen de la valeur propre maximale, l'autre en se fondant sur la distribution asymptotique des grandes valeurs propres.

ABSTRACT :

We study the eigenvalues of the covariance matrices of non stationnary processes. We analyse the case of a random walk, then the asymptotic distribution of these eigenvalues. A generalized Szegö theorem is proved. We finally establish two characterizations of the degree of non stationnarity of an ARIMA process .

L'étude des valeurs propres des matrices de covariance d'un processus stationnaire montre que ces valeurs propres sont fondamentalement liées à la densité spectrale du processus. Rappelons en effet que pour un tel processus $\{ X_t ; t \in \mathbb{Z} \}$ les valeurs propres $\lambda_\nu^{(n)}$ $(\nu = 1,\ldots,n)$ de la matrice de covariance de taille n, autrement dit de $\mathrm{Var}\{ (X_t, X_{t-1}, \ldots, X_{t-n+1})' \}$, vérifient pour toute fonction continue h l'égalité :

$$\mathrm{Lim} \ \frac{1}{n} \sum_{\nu=1}^{n} h(\lambda_\nu^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} h[f(\omega)] \, d\omega \quad ,$$

où f désigne la densité spectrale du processus.
Initialement prouvée par G. Szegö dans le cas d'une densité spectrale bornée (voir Grenander-Szegö (1958)) cette relation est susceptible de généralisation (voir Accardo (1990)), mais l'hypothèse de stationnarité reste cruciale puisque les démonstrations reposent largement sur la représentation spectrale des autocorrélations du processus considéré.
On se propose ici de relâcher l'hypothèse de stationnarité et d'examiner en quoi est alors affecté le spectre de la matrice de covariance du processus, matrice qu'il faut au préalable redéfinir.
On posera ainsi pour tout processus $\{ X_t ; t \in \mathbb{Z} \}$:

$$\Gamma_n \stackrel{def}{=} \mathrm{Var} \{ (X_1, X_2, \ldots, X_n)' \}$$

Il est clair que si le processus est stationnaire on retrouve la matrice de covariance usuelle.
Si au contraire le processus est non stationnaire, la matrice de covariance n'est, en général, plus Toeplitz et dépend des conditions initiales.
Dans un premier temps, on aborde l'étude par un cas simple de processus non stationnaire : les marches aléatoires.

I VALEURS PROPRES ET MARCHE ALEATOIRE.

Soit la marche aléatoire (dans $\mathbb{R}$) :

$$\begin{cases} X_t = X_{t-1} + \varepsilon_t \\ X_t = 0 \quad t \leq -p \end{cases} \qquad \text{avec } (\varepsilon_t) \text{ bruit blanc et } \sigma_\varepsilon^2 = 1,\ p \in \mathbb{N}$$

Il est immédiat de vérifier que :
pour tout t et t'≥ 1, Cov ($X_t, X_{t'}$) = Min (t,t') + p
On en déduit la matrice de covariance :

$$\Gamma_{n,p} = \begin{bmatrix} p+1 & p+1 & \cdots & p+1 \\ p+1 & p+2 & \cdots & p+2 \\ p+1 & p+2 & & p+3 \\ \vdots & \vdots & \ddots & \vdots \\ p+1 & p+2 & \cdots & p+n \end{bmatrix}.$$

On note qu'à la différence du cas stationnaire la matrice dépend directement des conditions initiales, ici la date $t_0 = -p$ de démarrage du processus.
C'est ce qui motive l'étude du spectre de la matrice de deux points de vue complémentaires :
- On fixe p et on étudie les valeurs propres de $\Gamma_{n,p}$, notamment leur distribution asymptotique quand la taille n de l'échantillon grandit.
- On fixe la taille n de l'échantillon et on étudie le comportement des valeurs propres quand le processus démarre de plus en plus tôt ($p \to +\infty$).

1) Calcul des valeurs propres.

Le calcul des valeurs propres des $\Gamma_{n,p}$ est effectué en développant le déterminant caractéristique :

$$D_{n,p}(\lambda) = \det\{ \Gamma_{n,p} - \lambda\ Id_n \}$$

ce qui va fournir une relation de récurrence linéaire grâce à laquelle on peut obtenir explicitement les $D_{n,p}$.
On a en effet :

Proposition 1 :

(1) $D_{n,p}(\lambda) = [(p+1)-\lambda(p+2)]\, D_{n-1,0}(\lambda) - \lambda^2(p+1)\, D_{n-2,0}$, $n \geq 3$,

(2) $D_{n,0}(\lambda) = (1-2\lambda)\, D_{n-1,0}(\lambda) - \lambda^2\, D_{n-2,0}(\lambda)$, $n \geq 3$.

Preuve : Voir Annexe.

Ces deux relations sont, en l'état, peu maniables, puisqu'elles définissent une récurrence sur les *deux* indices n et p.
On va, en fait, montrer que les $D_{n,p}$ vérifient (2). Une remarque permet de simplifier notablement les calculs ultérieurs :
considérons la suite des $D_{n,0}(\lambda)$. Elle n'est définie qu'à partir de n = 1. Mais grâce à (2) il est possible de définir récursivement :

$$D_{0,0}(\lambda) = 1 \ , \quad D_{-1,0}(\lambda) = -1/\lambda \ , \quad D_{-2,0}(\lambda) = \frac{1}{\lambda^2}\,(1-1/\lambda) \ ,$$

de sorte que :

(3) $D_{n,0}(\lambda) = (1-2\lambda)\, D_{n-1,0}(\lambda) - \lambda^2\, D_{n-2,0}(\lambda)$, $n \geq 0$.

On peut alors définir, grâce à (1) :

$$D_{2,p} = \lambda^2 - (2p+3)\lambda + (p+1) \ , \qquad D_{1,p} = (p+1) - \lambda, \qquad D_{0,p} = 1$$

et l'on obtient donc :

(4) $D_{n,p}(\lambda) = [(p+1)-\lambda(p+2)]\, D_{n-1,0}(\lambda) - \lambda^2(p+1)\, D_{n-2,o}$, $n \geq 0$

Dans ces conditions, on a le lemme :

Lemme 1 :
$$\begin{cases} D_{n,p}(\lambda) = (1-2\lambda)\, D_{n-1,p}(\lambda) - \lambda^2\, D_{n-2,p}(\lambda) & , n \geq 2 \\ D_{1,p}(\lambda) = p+1-\lambda \\ D_{0,p}(\lambda) = 1 \end{cases}$$

Preuve : Voir annexe.

De la récurrence énoncée par le lemme, on tire l'équation caractéristique associée :

$$(5) \qquad r^2 - (1-2\lambda)r + \lambda^2 = 0$$

Pour mener à bien le calcul on fait le changement de variable :

$$(6) \qquad \lambda = \frac{1}{2(1+\cos 2x)} \qquad \left(\text{pour } \quad 0 < x < \pi \quad \text{et } x \neq \frac{\pi}{2} \right)$$

Ceci suppose donc que :

$$(7) \qquad \lambda > \frac{1}{4}$$

On va donc exprimer $D_{n,p}(\lambda)$ pour λ vérifiant (7). On constatera ultérieurement que l'hypothèse (7) n'élimine pas de solution et qu'on trouve bien les n racines de l'équation $D_{n,p}(\lambda) = 0$.
Moyennant le changement de variable (6) l'équation (5) devient :

$$[2(1+\cos 2x)r]^2 - 2\cos 2x\ [2(1+\cos 2x)r] + 1 = 0$$

qui admet deux racines complexes conjuguées :

$$r_1 = \frac{e^{2ix}}{2(1+\cos 2x)} \quad , \quad r_2 = \frac{e^{-2ix}}{2(1+\cos 2x)} \quad .$$

La solution générale de l'équation de récurrence

$$D_{n,p} = (1-2\lambda)\ D_{n-1,p} - \lambda^2\ D_{n-2,p}$$

est donc :

$$D_{n,p}(\lambda) = A\ r_1^n + B\ r_2^n \quad ,$$

où A et B sont des constantes déterminées par :

$$(8)\qquad \begin{cases} D_{1,p} = A\, r_1 + B\, r_2 = p+1-\lambda \\ D_{0,p} = A + B = 1 \end{cases},$$

De (8) on tire :

$$A = \frac{1}{2} - i(2p+1)\,\frac{\cos x}{2\sin x} \quad \text{et } B = \bar{A} = \frac{1}{2} + i(2p+1)\,\frac{\cos x}{2\sin x},$$

de sorte qu'il vient :

$$(9)\qquad \begin{cases} D_{n,p}(\lambda) = \dfrac{1}{[2(1+\cos 2x)]^n \sin x}\{\, \sin x.\cos 2nx + (2p+1)\sin 2nx.\cos x \,\} \\ \lambda = \dfrac{1}{2(1+\cos 2x)}\ , \text{ avec } \lambda > \dfrac{1}{4} \qquad (\text{i.e. } x \in \left]0, \pi\right[\setminus \{\tfrac{\pi}{2}\}) \end{cases}$$

La résolution de $D_{n,p}(\lambda) = 0$, sous l'hypothèse que $\lambda > \frac{1}{4}$ se ramène donc par (9) à la résolution de :

$$(10)\qquad \begin{cases} \sin x.\cos 2nx + (2p+1)\sin 2nx.\cos x = 0 \\ \text{et } x \in \left]0, \pi\right[\setminus \{\tfrac{\pi}{2}\} \end{cases}$$

Remarque : de la formule élémentaire $\sin a.\cos b = \frac{1}{2}[\sin(a+b)+\sin(a-b)]$ on conclut que la première équation dans (10) peut s'écrire de façon équivalente :

$$(p+1)\sin(2n+1)x + p\sin(2n-1)x = 0$$

(10) ne semble pas résoluble analytiquement et simplement pour p quelconque; cependant on vérifie facilement

- d'une part que la résolution est immédiate si p = 0
- d'autre part que les solutions obtenues dans le cas p = 0 séparent les solutions de l'équation générale (p quelconque).

α) En effet si p = 0 , (10) devient :

$$(11)\quad \begin{cases} \sin(2n+1)x = 0 \\ x \in \left]\, 0\,,\, \pi \,\right[\setminus \{ \frac{\pi}{2} \} \end{cases}$$

Les solutions de (11) sont :

$$x = \frac{\nu\pi}{2n+1} \quad , \ \nu = 1,\ldots,2n$$

qui définissent :

$$\lambda_{\nu}^{(n,0)} = \frac{1}{2(1+\cos\frac{2\nu\pi}{2n+1})} \quad , \ \nu = 1,\ldots,n$$

c'est à dire n solutions *distinctes* de l'équation $D_{n,0}(\lambda) = 0$, toutes de la forme (6) retenue pour λ.

β) si $p \geq 1$, (10) n'est plus résoluble exactement. Mais les solutions $\lambda_{\nu}^{(n,0)}$ obtenues pour $p = 0$ permettent d'encadrer les solutions $\lambda_{\nu}^{(n,p)}$ de l'équation $D_{n,p}(\lambda) = 0$.

En effet , il suffit de remarquer que $D_{n,p}(\lambda_{\nu}^{(n,0)})$ est du signe de :

$$\sin\frac{\nu\pi}{2n+1} \cdot \cos\frac{2n\nu\pi}{2n+1} + (2p+1)\sin\frac{2n\nu\pi}{2n+1} \cdot \cos\frac{\nu\pi}{2n+1}$$

Tenant alors compte de :

$$\frac{2n\nu\pi}{2n+1} = \nu\pi - \frac{\nu\pi}{2n+1} \ , \ \sin(\nu\pi-\omega) = (-1)^{\nu+1}\sin\omega \quad \text{et} \quad \cos(\nu\pi-\omega) = (-1)^{\nu}\cos\omega$$

on voit que $D_{n,p}(\lambda_{\nu}^{(n,0)})$ est du signe de :

$$(-1)^{\nu+1}\left(2p \sin\frac{\nu\pi}{2n+1} \cos\frac{\nu\pi}{2n+1} \right)$$

donc du signe de $(-1)^{\nu+1}$ pour $\nu = 1,\ldots,n$

En particulier $D_{n,p}(\lambda_{\nu}^{(n,0)})$ est positif ; $D_{n,p}(\lambda_{\nu}^{(n,0)})$ est positif si n est impair et négatif si n est pair.

Or il est immédiat de constater que, pour λ assez grand (i.e. pour x tendant vers $\frac{\pi}{2}$ par valeurs inférieures), $D_{n,p}(\lambda)$ est du signe de $\cos n\pi$ et donc est

positif si n est pair et négatif sinon. De sorte que :

$$D_{n,p}(\lambda_n).D_{n,p}(\lambda) < 0 \quad \text{pour } \lambda \text{ assez grand}$$

Ainsi, si $\lambda_\nu^{(n,p)}$, $\nu = 1,\ldots,n$ désigne les solutions, rangées dans l'ordre croissant, de : $D_{n,p}(\lambda) = 0$
il vient :

$$(13) \qquad \lambda_1^{(n,0)} < \lambda_1^{(n,p)} < \lambda_2^{(n,0)} < \lambda_2^{(n,p)} < \ldots < \lambda_{n-1}^{(n,p)} < \lambda_n^{(n,0)} < \lambda_n^{(n,p)}$$

Conséquences : 1) (13) montre que les valeurs propres de $\Gamma_{n,0}$ encadrent celles de $\Gamma_{n,p}$.

2) Les valeurs propres de $\Gamma_{n,p}$, comme celles de $\Gamma_{n,0}$ sont supérieures à $\frac{1}{4}$ ce qui justifie le changement de variables (6).

3) Enfin on peut déduire de (13) un résultat à la fois plus précis et plus général : rappelons d'abord que si A et B désignent 2 matrices réelles symétriques de taille n, on peut définir une relation d'ordre

$$A \ll B \quad \Leftrightarrow \quad (\; x'Ax \leq x'Bx \; , \; \forall \; x \in \mathbb{R}^n \;)$$

Dans ces conditions, on peut montrer que :

$$(14) \qquad A \ll B \Rightarrow (\; \lambda_i^A \leq \lambda_i^B \; , \; i = 1,2, \ldots,n \;)$$

λ_i^A et λ_i^B désignant les valeurs propres respectivement de A et B *rangées dans le même ordre.*(Voir Kobilinski (1979)).
De la relation immédiate :

$$\Gamma_{n,p} - \Gamma_{n,q} = (p-q)\, e.e' \quad \text{avec } e = (1,1,\ldots,1)' \in \mathbb{R}^n$$

on déduit :

$$\Gamma_{n,p} \ll \Gamma_{n,q} \quad \text{dès que } q \geq p \; .$$

La proposition (14) s'applique donc et on obtient :

$$(15) \qquad p \leq q \quad \Rightarrow \quad \lambda_{\nu}^{(n,p)} \leq \lambda_{\nu}^{(n,q)} , \qquad \nu = 1,\ldots,n$$

Ainsi ,

$$(16) \quad \lambda_{\nu}^{(n,p-1)} \leq \lambda_{\nu}^{(n,p)} \leq \lambda_{\nu+1}^{(n,0)} \leq \lambda_{\nu+1}^{(n,p)} , \quad p \geq 1 \text{ et } \nu = 1,\ldots,n-1$$

Au total on peut résumer (13) et (16) dans une formule unique (17) :

$$\lambda_1^{(n,p-1)} \leq \lambda_1^{(n,p)} \leq \lambda_2^{(n,p-1)} \leq \lambda_2^{(n,p)} \leq \ldots \leq \lambda_{n-1}^{(n,p-1)} \leq \lambda_{n-1}^{(n,p)} \leq \lambda_n^{(n,p-1)} \leq \lambda_n^{(n,p)}$$

On voit donc que les valeurs propres de $\Gamma_{n,p-1}$ encadrent celles de $\Gamma_{n,p}$ (à l'exception de la plus grande $\lambda_n^{(n,p)}$) .

Remarque : les inégalités précédentes prouvent par ailleurs que les valeurs propres sont approchées par les valeurs que prend la pseudo densité spectrale (c'est à dire la fonction $f(\omega) = \frac{1}{|1-e^{i\omega}|^2}$, $0 < \omega < 2\pi$) en n points équirépartis de l'intervalle $[0,2\pi]$, résultat déjà établi pour des processus AR(p) et MA(q) stationnaires (Voir Accardo (1990)).

2) Comportement asymptotique.

De (17) on déduit le comportement des valeurs propres $\lambda_{\nu}^{(n,p)}$, $\nu = 1,\ldots,n$ quand n ou p tendent vers 0.

- Soit d'abord n fixé. On a : $\mathrm{Tr}\ \Gamma_{n,p} = np + \frac{n(n+1)}{2}$. Par (13) on obtient donc que : $\mathrm{Tr}\ \Gamma_{n,p} \leq (n-1)\lambda_n^{(n,0)} + \lambda_n^{(n,p)}$.

Par conséquent, si $p \longrightarrow \infty$, on voit que toutes les valeurs propres $\lambda_{n,p}^{\nu}$ sont bornées (par $\lambda_n^{(n,0)}$) sauf la plus grande $\lambda_n^{(n,p)}$ qui est un $O(p)$. Plus

précisément,

$$\lim_{p\to\infty} \frac{\lambda_n^{(n,p)}}{p} = \lim_{p\to\infty} \frac{\mathrm{Tr}\ \Gamma_{n,p}}{p} = n$$

- Soit maintenant p fixé. On a par un calcul élémentaire :

$$\lambda_{\nu+1}^{(n,p)} - \lambda_{\nu}^{(n,p)} = \frac{\sin\frac{(2\nu+1)\pi}{2n+1}\ \sin\frac{\pi}{2n+1}}{(1+\cos\frac{2\nu\pi}{2n+1})(1+\cos\frac{(2\nu+2)\pi}{2n+1})}$$

et donc, à ν et p fixés, $\lambda_{\nu}^{(n,0)}$ constitue une approximation en $\frac{1}{n}$ de $\lambda_{\nu}^{(n,p)}$.
D'autre part pour n grand on a :

$$\lambda_{n-\nu}^{(n,0)} \simeq \frac{1}{2(\frac{2\nu+1}{2n+1})^2\ \pi^2}$$

autrement dit $\lambda_{n-\nu}^{(n,0)}$ est, à ν fixé, un $O(n^2)$. Par (17) il en est de même de $\lambda_{n-\nu}^{(n,p)}$.

3) Résultats:

En résumé,
-l'étude précédente montre que les valeurs propres d'une marche aléatoire sont approchées par les valeurs prises par

$$f(\omega) = \frac{1}{|1-e^{i\omega}|^2}$$

en n points équirépartis de l'intervalle $[0,\pi]$:

$$\lambda_{\nu}^{(n,0)} = f(\frac{\nu\pi}{n+1/2}) \quad , \ \nu = 1,\ldots,n$$

ou encore :

$$\lambda_\nu^{(n,0)} = \frac{1}{2\left(1+\cos\frac{\nu\pi}{(n+\frac{1}{2})}\right)} .$$

-On voit alors que les plus grandes de ces valeurs propres sont d'ordre n^2.
-Les valeurs propres $\lambda_\nu^{(n,p)}$ sont, à ν fixé, approchées par les $\lambda_\nu^{(n,0)}$ en n^{-2}.
Ceci traduit le fait que les valeurs propres s'accumulent en $\frac{1}{4}$, pour un ν donné, quand n tend vers $+\infty$.

4) Théorème de Szegö.

On peut maintenant établir un résultat qui étend aux marches aléatoires le théorème de Szegö, suivant lequel les valeurs propres de la matrice de covariance $\Gamma_{n,p}$ d'un processus stationnaire sont distribuées, asymptotiquement (quand $n \to \infty$), comme la densité spectrale du processus.

Proposition 2 : soit $\Gamma_{n,p}$, la matrice précédemment définie, et $\lambda_\nu^{(n,p)}$ (pour $\nu = 1,\ldots,n$) ses valeurs propres ; soit f la pseudo-densité spectrale de la marche aléatoire associée à $\Gamma_{n,p}$:

$$f(\omega) = \frac{1}{|1-e^{i\omega}|^2}, \quad 0 < \omega < 2\pi .$$

Alors pour toute fonction h définie sur $\mathbb{R}^+$, continue et bornée, on a :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} h\,(\lambda_\nu^{(n,p)}) = \frac{1}{2\pi}\int_0^{2\pi} h\,[f(\omega)]\,d\omega .$$

Preuve :

Soit ζ un réel positif. Définissons $\Phi_{n,p}$ par :

$$\Phi_{n,p}(\zeta) = \frac{1}{n} \text{ card } \{\lambda_\nu^{(n,p)} < \zeta\}$$

Par (13) on a :

$$\left|\text{card}\{\ \lambda_\nu^{(n,p)} < \zeta\ \} - \text{card}\{\lambda_\nu^{(n,0)} < \zeta\ \}\ \right| \leq 1$$

d'où

$$\lim_{n\to\infty} \left|\ \Phi_{n,p}(\zeta) - \Phi_{n,0}(\zeta)\right| = 0 \qquad (18)$$

Comme d'autre part:

$$\text{card}\{\ \lambda_\nu^{(n,0)} < \zeta\ \} = \text{card}\{\ \frac{1}{2(1+\cos\frac{2\nu+1}{2n+1}\pi)} < \zeta\ \}$$

$$= \text{card}\{\ \nu \in \mathbb{N}\ ;\ \nu < \tfrac{1}{2}[\tfrac{1}{\pi}(2n+1)\text{Arccos}(\tfrac{1}{2\zeta} - 1)-1]\ \}\ ,\ \text{si}$$

ζ est infèrieur ou égal à $\frac{1}{4}$,

$$= 0 \quad \text{sinon} ,\quad \text{il vient finalement :}$$

$$\lim_{n\to\infty} \Phi_{n,0}(\zeta) = \frac{1}{\pi}\ \text{Arccos}(\frac{1}{2\zeta} - 1) \qquad \text{si } \zeta \geq \frac{1}{4}$$

$$= 0 \qquad \text{sinon .}$$

On note Φ cette fonction limite.

Il suffit maintenant de remarquer que $f(\omega) < \zeta$ dès que $\cos\omega < 1 - \frac{1}{2\zeta}$, c'est à dire dès que : $\text{Arccos}(1 - \frac{1}{2\zeta}) < \omega < 2\pi - \text{Arccos}(1 - \frac{1}{2\zeta})$.

Ainsi on a : $\frac{1}{2\pi}\int_0^{2\pi} \chi_{[0,\zeta]}(\ f(\omega))\ d\omega = \frac{1}{\pi}\ \text{Arccos}(\frac{1}{2\zeta} - 1)$.

On voit donc que la proposition est vérifié pour toute fonction h du type indicatrice $\chi_{[0,\pi]}$; donc pour toute fonction continue bornée.

Q.E.D.

5) Comparaison Stationnaire - Non Stationnaire.

On peut, à ce stade de l'étude, comparer les résultats obtenus pour une marche aléatoire avec ceux que l'on dérive usuellement lors de l'analyse des processus autorégressifs d'ordre 1 ; pour un processus stationnaire vérifiant la relation : $X_t = \rho X_{t-1} + \varepsilon_t$, avec $-1 < \rho < 1$, on sait que les

valeurs propres sont données par :

$$\lambda_{\nu}^{(n)} = \frac{1}{1+\rho^2-2\rho\cos\frac{\nu\pi}{n+1}}$$

On a de façon immédiate :

$$\lim_{\rho\to 1} \lambda_{\nu}^{(n)} = \frac{1}{2(1-\cos\frac{\nu\pi}{n+1})}$$

Cette limite est clairement équivalente (pour n grand, ν étant fixé) à la valeur propre $\lambda_{\nu}^{(n,p)}$ calculée dans le cas de la marche aléatoire (de fait, on a pu vérifier dans les paragraphes 3) et 4) que les valeurs propres correspondant à la marche aléatoire présentent sous plusieurs rapports les mêmes propriétés que celles obtenues dans le cas stationnaire). En ce sens les valeurs propres de la matrice de covariance $\Gamma_{n,p}$ d'une marche aléatoire apparaissent *comme limite* (*quand* ρ tend vers 1) des valeurs propres des processus AR(1) stationnaires alors que, du point de vue des matrices de covariance, processus stationnaires et processus non stationnaires sont totalement différents.

Cette stabilité des éléments propres des matrices de covariance suggère que les résultats établis pour une marche aléatoire persistent dans un cadre plus général. Ce point est abordé dans la partie suivante.

II PROCESSUS AUTOREGRESSIFS NON STATIONNAIRES.

A partir d'une approximation des valeurs propres de la matrice de covariance d'un processus vérifiant une équation autorégressive, on montre une généralisation au cas non stationnaire du théorème de Szegö.

1) Approximation des valeurs propres.

Soit X_t un processus défini par la donnée de $X_1, X_2, \ldots, X_p$ (de moyenne nulle) et de l'équation autorégressive :

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \varepsilon_t \quad , \quad \text{pour } t \geq p \ .$$

On ne fait aucune hypothèse sur les racines (notées $\theta_1, \theta_2, \ldots, \theta_p$) du polynôme autorégressif $P(z) = z^p - \sum_{1}^{p} a_j z^{p-j}$.

On a alors la proposition :

Proposition 3 : Soit Γ_n la matrice de covariance (de taille n) du processus $\{X_t ;\ t \geq 1\}$ (autrement dit $\Gamma_n = \mathrm{Var}[(X_1,\ldots,X_n)]$) et $\lambda_\nu^{(n)}$, $\nu = 1,\ldots,n$ ses valeurs propres rangées dans l'ordre croissant.

Soit aussi $\mu_\nu^{(n)}$ les quantités $\prod_{j=1}^{p} \left|1-\theta_j \exp\left(i\frac{k\pi}{n+1}\right)\right|^2$, pour $k = 1,\ldots,n$, réordonnées dans le sens croissant. On a alors :

$$(20) \qquad \mu_{\nu-4p}^{(n)} \leq \frac{1}{\lambda_{n+1-\nu}^{(n)}} \leq \mu_{\nu+4p}^{(n)} \quad , \ \nu = 4p+1,\ldots,n-4p.$$

Preuve : Voir annexe.

Remarques :

1) L'inégalité (20) exprime d'abord que presque toutes les valeurs propres $\lambda_\nu^{(n)}$ de Γ_n (c'est à dire toutes sauf éventuellement les 8p valeurs extrêmes) sont encadrées par les quantités $\prod_{j=1}^{p} \frac{1}{\left|1-\theta_j \exp i\frac{k}{n+1}\right|^2}$, autrement dit par les valeurs prises aux points $\omega = \frac{k\pi}{n+1}$ par la pseudo densité spectrale du processus ce qui généralise un résultat déjà vérifié par les marches aléatoires (et, rappelons le, par les processus stationnaires vérifiant une équation autorégressive ou moyenne mobile, cf Accardo (1990)).

2) La proposition permet de préciser 3 cas de figure selon la position des racines θ_j du polynôme autorégressif P(z) :

a) Soit toutes les racines sont de module strictement inférieur à 1. Le processus est alors stationnaire et les valeurs propres sont toutes dans

l'intervalle (indépendant de n) [$\min_{\omega} f(\omega)$, $\max_{\omega} f(\omega)$] (Voir Accardo(1990)). Il n'y a donc pas de valeurs explosives.

b) Soit certaines racines sont de module strictement supérieur à 1, les autres étant de module strictement inférieur à 1. Le processus est alors non stationnaire, mais les $\mu_\nu^{(n)}$ étant bornées inférieurement par la quantité strictement positive $\min_j(1-|\theta_j|)^{2p}$ (vérification immédiate), l'inégalité (20) assure que les valeurs propres $\lambda_\nu^{(n)}$ "explosives" éventuelles sont nécessairement parmi les 4p plus grandes.

c) Soit il existe des racines de module 1. Dans ce cas les $\mu_\nu^{(n)}$ n'admettent pas de borne inférieure strictement positive. Le nombre de valeurs propres "explosives" croît avec n comme le montre l'exemple de la marche aléatoire.

3) Une autre conséquence de l'inégalité (20) est d'assurer que les valeurs propres $\lambda_\nu^{(n)}$ de Γ_n, ou du moins leurs inverses se comportent presque toutes (i.e. sauf éventuellement les 8p extrêmes) comme les $\mu_\nu^{(n)}$. Or ces dernières définissent, asymptotiquement, une distribution égale à la loi image par la fonction $g(\omega) = \prod_{j=1}^{p} |1-\theta_j e^{i\omega}|^2$ de la loi uniforme sur $[0,2\pi]$ (voir Accardo (1990)). On conçoit ainsi intuitivement que les valeurs propres $\lambda_\nu^{(n)}$ sont distribuées asymptotiquement suivant la pseudo-densité spectrale du processus, ici $f(\omega) = \frac{1}{g(\omega)}$. C'est dire que le processus $\{ X_t ;\ t \geq 1 \}$ bien qu'éventuellement non (asymptotiquement) stationnaire vérifie le théorème de Szegö, ce qu'établit rigoureusement la proposition suivante :

Proposition 4 : Soit $\{ X_t ;\ t \geq 0 \}$ un processus vérifiant une équation autorégressive,

$$X_t = \sum_{j=1}^{p} a_j X_{t-j} + \varepsilon_t \quad , t \geq p \quad \text{et } \mathrm{Var}(\varepsilon_t) = \sigma^2.$$

Soit $\Gamma_n, \lambda_\nu^{(n)}$, $\mu_\nu^{(n)}$ définis comme dans la proposition précédente et P(z) le polynôme autorégressif associé,

$$P(z) = z^p - \sum_{j=1}^{p} a_j z^{p-j}.$$

Alors pour toute fonction h continue et bornée sur $\mathbb{R}^+$, on

a :

$$(21) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} h(\lambda_\nu^{(n)}) = \frac{1}{2\pi}\int_0^{2\pi} h\left(\frac{\sigma^2}{|P(e^{i\omega})|^2}\right) d\omega.$$

Preuve : D'après (20) on a, pour tout réel positif ζ,

$$(22) \qquad \left| \text{card}\{ \lambda_\nu^{(n)}; \frac{1}{\lambda_\nu^{(n)}} < \zeta \} - \text{card}\{ \mu_\nu^{(n)} ; \mu_\nu^{(n)} < \zeta \} \right| \leq 4p.$$

Or on sait d'après l'étude menée dans le cas stationnaire (Voir Accardo (1990)) que les fonctions $\Phi_n(\zeta) = \frac{1}{n}$ card $\{ \mu_\nu^{(n)}; \mu_\nu^{(n)} < \zeta \}$ convergent vers la fonction Φ telle que :

$$\Phi(\zeta) = \frac{1}{2\pi} \int_0^{2\pi} \chi_{[0,\zeta]}(|P(e^{i\omega})|^2)\, d\omega$$

(où $\chi_{[0,\zeta]}$ désigne la fonction caractéristique de $[0,\zeta]$) qui est la fonction de répartition de la loi image par $|P(e^{i\omega})|^2$ de la loi uniforme sur $[0,2\pi]$.
En réalité l'égalité précédente n'est rien d'autre que le théorème de Szegö appliqué au processus moyenne mobile "dual" du processus autorégressif considéré:

$$Y_t = \sum_{j=1}^{p} a_j \varepsilon_{t-j} + \varepsilon_t \quad , t \geq p \quad \text{et } \text{Var}(\varepsilon_t) = \sigma^2,$$

qui est évidemment stationnaire et dont les valeurs propres peuvent être approchées par les $\mu_\nu^{(n)}$

On obtient donc :

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\nu=1}^{n} \chi_{[0,\zeta]}\left(\frac{1}{\lambda_\nu^{n}} \right) = \frac{1}{2\pi} \int_0^{2\pi} \chi_{[0,\zeta]}(|P(e^{i\omega})|^2)\, d\omega .$$

Cette égalité étant vérifiée pour tout réel positif ζ, elle l'est aussi pour toute fonction g continue et bornée sur $\mathbb{R}^+$.On obtient alors l'égalité (21)

en posant $h(x) = g(\frac{1}{x})$ pour $x > 0$.

Q.E.D.

Remarque :ce résultat montre donc que les valeurs propres se distribuent de la même façon que dans le cas stationnaire, dès que sont éliminées (par l'utilisation de fonctions bornées) les valeurs explosives. On est alors naturellement conduit à analyser plus précisément la "queue" de la distribution, c'est à dire les plus grandes valeurs propres de Γ_n quand n tend vers $+\infty$. C'est l'objet de la partie suivante.

III Etude des valeurs propres maximales (cas ARMA).

On a vu, à l'occasion de l'étude des processus stationnaires que si la densité spectrale est bornée, alors la valeur propre maximale $\lambda_\nu^{(n)}$ tend vers Sup $f(\omega)$ quand n tend vers $+\infty$; en fait on a plus précisément le résultat suivant, classique dans la théorie spectrale des matrices de Toeplitz,

$$\lambda_{n-k}^{(n)} - \sup f(\omega) = O(1/n^2) \quad , \text{ pour k fixé.}$$

(Voir Wilf (1970) ou Widom (1958)).

On a établi d'autre part que si la densité spectrale n'est pas bornée alors les "grandes" valeurs propres $\lambda_{n-k}^{(n)}$ tendent vers $+\infty$ avec n (à k fixé) (voir Accardo(1990)).

C'est ce résultat qu'on se propose de préciser ici en étudiant en particulier la relation entre le degré de non stationnarité du processus (en se limitant à ceux qui vérifient un modèle ARMA), et la vitesse avec laquelle la valeur propre maximale $\lambda_n^{(n)}$ de la matrice de covariance Γ_n tend vers l'infini avec la taille n de l'échantillon.

On montrera ainsi que cette valeur propre est d'ordre n^{2d} si d désigne le degré de non stationnarité.

Remmarque : ce résultat correspond bien à l'ordre n^2 obtenu dans le cas de la marche aléatoire (voir I 2)) où le degré de non stationnarité est d = 1.

1) Ordre de la valeur propre maximale.

On considère un processus univarié $\{ X_t ; t \in \mathbb{Z} \}$ vérifiant un modèle ARIMA

(condition (α)) : $$(1-L)^d \Phi(L) X_t = \Theta(L)\varepsilon_t.$$

$\Phi(z)$ est supposé avoir toutes ses racines de module strictement plus petit que 1 de sorte que $\frac{\Theta(L)}{\Phi(L)} \varepsilon_t$ est un processus stationnaire.
Un tel processus admet une pseudo densité spectrale donnée par :

$$f(\omega) = \frac{|\Theta(e^{i\omega})|^2}{|\Phi(e^{i\omega})|^2 |1-e^{i\omega}|^{2d}} \quad , \ 0 \leq \omega \leq 2\pi$$

On remarque que :

$$f(\omega) \simeq \frac{1}{\omega^{2d}}$$

dans un voisinage de zéro.

Si $d \geq \frac{1}{2}$, X_t n'est plus stationnaire et on pose alors

condition (β): $$\begin{cases} X_t = 0 \\ \varepsilon_t = 0 \end{cases} \quad , \quad \text{pour } t \leq 0.$$

On a alors la proposition:

Proposition 5 : Soit un processus vérifiant les conditions (α) et β), $d \geq 0$.
Si $d \neq \frac{1}{2}$, la valeur propre maximale de Γ_n est d'ordre n^{2d}.
Pour $d = \frac{1}{2}$, l'ordre est compris entre n et nLogn.

Preuve : On distingue deux cas selon que $d < \frac{1}{2}$ (le processus est alors stationnaire) ou que $d \geq \frac{1}{2}$ (processus non stationnaire).

A) $d < \frac{1}{2}$.

les autocovariances vérifient (asymptotiquement) pour tout entier h :

$$\gamma(h) = \int_{-\pi}^{+\pi} e^{i\omega h} f(\omega)\, d\omega \ .$$

De sorte que la valeur propre maximale de Γ_n est donnée par :

$$\lambda_n^{(n)} = \text{Max } X' \left[\int_{-\pi}^{+\pi} \left(e^{i(k-l)\omega} \right)_{1\le k,l\le n} f(\omega)\, d\omega \right] X$$

avec X parcourant la sphère unité de $\mathbb{C}^n$. On peut encore écrire :

$$\lambda_n^{(n)} = \text{Max} \int_{-\pi}^{+\pi} \left| x_1 e^{i\omega} + x_2 e^{i2\omega} + \ldots + x_n e^{in\omega} \right|^2 f(\omega) d\omega$$

avec $|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 = 1$.

a) Il suffit alors de remarquer que :

- il existe une constante K telle que

$$0 \le f(\omega) \le \frac{K}{\omega^{2d}} \quad , \quad \text{pour } -\pi < \omega < +\pi \quad \text{et } \omega \ne 0 \ .$$

- d'après la contrainte

$$|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 = 1,$$

on a la majoration :

$$\left| x_1 e^{i\omega} + x_2 e^{i2\omega} + \ldots + x_n e^{in\omega} \right|^2 \le n \ .$$

Il vient alors :

$$\int_0^{1/n} \left| x_1 e^{i\omega} + \ldots + x_n e^{in\omega} \right|^2 f(\omega) d\omega \le n \int_0^{1/n} f(\omega)\, d\omega \le nK \left[\frac{1}{(-2d+1)\ \omega^{2d-1}} \right]_0^{1/n},$$

autrement dit

$$\int_0^{1/n} |x_1 e^{i\omega} + x_2 e^{i2\omega} + \ldots + x_n e^{in\omega}|^2 f(\omega) d\omega \leq \frac{K}{1-2d} n^{2d}.$$

De même on a :

$$\int_{1/n}^{\pi} |x_1 e^{i\omega} + x_2 e^{i2\omega} + \ldots + x_n e^{in\omega}|^2 f(\omega) d\omega \leq \ldots$$

$$\ldots \leq \frac{K}{(\frac{1}{n})^{2d}} \int_{1/n}^{\pi} |x_1 e^{i\omega} + x_2 e^{i2\omega} + \ldots + x_n e^{in\omega}|^2 \, d\omega \leq \ldots$$

$$\ldots \leq K\, n^{2d} \int_{1/n}^{\pi} |x_1 e^{i\omega} + x_2 e^{i2\omega} + \ldots + x_n e^{in\omega}|^2 \, d\omega \leq \ldots$$

$$\ldots \leq K\, n^{2d} \int_{-\pi}^{+\pi} |x_1 e^{i\omega} + x_2 e^{i2\omega} + \ldots + x_n e^{in\omega}|^2 \, d\omega = K\, n^{2d}$$

par l'égalité de Parseval.
On voit donc que $\lambda_n^{(n)}$, la valeur propre maximale, est d'ordre inférieur à n^{2d}.

b) Considérons maintenant le choix particulier de X :

$$x_i = \frac{1}{\sqrt{n}} \quad , \text{ pour } i = 1,\ldots,n.$$

On a alors la minoration :

$$\lambda_n^{(n)} \geq \frac{1}{n} \int_{-\pi}^{+\pi} |e^{i\omega} + e^{2i\omega} + \ldots + e^{ni\omega}|^2 \, f(\omega) \, d\omega \, ,$$

c'est à dire,

$$\lambda_n^{(n)} \geq \frac{1}{n} \int_{-\pi}^{+\pi} \left| \frac{e^{in\omega}-1}{e^{i\omega}-1} \right|^2 f(\omega) \, d\omega. \qquad (*)$$

Du développement limité en zéro de e^{ix}, on déduit aisément l'existence d'une constante L telle que :

$$|x| < \frac{1}{L} \quad \Rightarrow \quad \frac{x^2}{2} < |1-e^{ix}|^2 < 2x^2.$$

Ceci permet d'obtenir :

$$\left|\frac{e^{in\omega}-1}{e^{i\omega}-1}\right|^2 \geq \left(\frac{n^2\omega^2}{2}\right)/(2\omega^2) = \frac{n^2}{4} \quad , \text{ dès que } \quad |\omega| < \frac{1}{nL} .$$

Reportant dans (*) on trouve :

$$\lambda_n^{(n)} \geq \frac{1}{n} \int_{-1/nL}^{1/nL} \left|\frac{e^{in\omega}-1}{e^{i\omega}-1}\right|^2 f(\omega)\, d\omega \geq \frac{2}{n} \int_0^{1/nL} \frac{n^2}{4} f(\omega)\, d\omega.$$

Puisque $f(\omega) \simeq \omega^{-2d}$ au voisinage de zéro, on sait qu'il existe une constante $M > 0$ telle que pour n assez grand,

$$f(\omega) \geq \frac{M}{\omega^{2d}}, \text{ pour } |\omega| < \frac{1}{n}.$$

Ainsi on a finalement,

$$\lambda_n^{(n)} \geq \frac{2}{n}\frac{n^2}{4} M \left[\frac{\omega^{1-2d}}{1-2d}\right]_0^{1/nL} = O(n^{2d}).$$

De a) et b) on déduit que la valeur propre maximale est d'ordre n^{2d}.

B) $d \geq \frac{1}{2}$.

La suite de la démonstration se fait en deux temps ; en s'appuyant sur une approximation des coefficients moyenne mobile, on prouve d'abord que la valeur propre maximale de la matrice de covariance Γ_n est d'ordre au moins n^{2d}, ensuite qu'elle est majorée par la trace de cette matrice elle même d'ordre n^{2d}.

a) $\lambda_n^{(n)} \geq O(n^{2d})$.

BU LILLE

Lemme 2 : on a $X_t = \sum_{k=1}^{t} \Pi_{t-k} \varepsilon_k$, $t \geq 1$ et , pour tout $k \geq 1$,

$$\left| \Pi_k - \frac{\Theta(1)}{\Phi(1)\Gamma(d)} k^{d-1} \right| < C\, k^{d-2}$$

avec Γ la fonction Gamma usuelle et C une constante positive.

Démonstration du lemme 2 : Voir Gouriéroux, Monfort [1990], p 527 .

Soit alors $(X_n^1)' = (X_1, X_2, \ldots, X_n)$; on peut écrire :

$$X_n^1 = L_n \varepsilon_n^1 \qquad (23)$$

où L_n est une matrice nxn triangulaire inférieure:

$$L_n = \begin{bmatrix} \Pi_0 & & & \\ \Pi_1 & \Pi_0 & & 0 \\ \Pi_2 & \Pi_1 & \Pi_0 & \\ \vdots & & \ddots & \\ \Pi_n & \cdots & \Pi_1 & \Pi_0 \end{bmatrix} .$$

Prenant les covariances dans (23), on obtient : $\Gamma_n = \sigma^2 L_n L_n'$.
Par conséquent

$$\lambda_n^{(n)} = \max_{1 \leq \nu \leq n} \lambda_n^{\nu} = \sigma^2 \max_{\zeta} \zeta' L_n L_n' \zeta \quad ,$$

où ζ décrit la sphère unité de $\mathbb{R}^n$ (muni de la norme euclidienne $\| \cdot \|_2$).
Pour le vecteur unitaire $w_n' = n^{-1/2}(1,1,\ldots,1)$ de $\mathbb{R}^n$, il vient donc :

$$\frac{1}{\sigma^2} \max \lambda_n^{\nu} \geq \| L_n' w_n \|^2 = \frac{1}{n} \sum_{j=1}^{n} \left(\sum_{k=0}^{n-j} \Pi_k \right)^2 \qquad (24)$$

Posons $A = \Theta(1)/\Phi(1)\Gamma(d)$. Quitte à changer ε_t en $-\varepsilon_t$, on peut supposer $A \geq 0$. Le lemme donne alors l'encadrement (25) :

$$A \sum_{k=1}^{n-j} k^{d-1} - C \sum_{k=1}^{n-j} k^{d-2} \leq \sum_{k=1}^{n-j} \Pi_k \leq A \sum_{k=1}^{n-j} k^{d-1} + C \sum_{k=1}^{n-j} k^{d-2} \text{ , pour } j < n.$$

Il est classique que : $\sum_{1}^{m} \int_{k-1}^{k} x^s \, dx \leq \sum_{1}^{m} k^s \leq \sum_{1}^{m} \int_{k}^{k+1} x^s \, dx$

On en déduit que, pour tout $m \geq 1$

$$H_m \, m^d < \sum_{k=1}^{m} \Pi_k \quad ,$$

avec $H_m = \frac{A}{d} - \frac{C}{d-1} \left[\frac{1}{m} \left(1 + \frac{1}{m}\right)^{d-1} - \frac{1}{m^d} \right]$, si $d \neq 1$ et $H_m = A$ si $d = 1$.

Clairement H_m tend vers $\frac{A}{d}$ quand m tend vers $+\infty$; par conséquent il existe un entier m_0 et une constante positive K_1 tels que :

$$m \geq m_0 \quad \Rightarrow \quad 0 < K_1 m^d < \sum_{1}^{m} \Pi_k < \sum_{0}^{m} \Pi_k \quad ,$$

si on suppose $\Pi_0 > 0$ (sinon on considère $|\Pi_0|$).

Il vient donc

$$\sum_{m=1}^{n-1} \left(\sum_{k=0}^{m} \Pi_k \right)^2 = \sum_{m=1}^{m_0} \left(\sum_{k=0}^{m} \Pi_k \right)^2 + \sum_{m_0+1}^{n-1} \left(\sum_{k=0}^{m} \Pi_k \right)^2 > K_2 + K_1^2 \frac{(n-2)^{2d+1}}{2d+1} \quad ,$$

où K_2 est une constante positive. On en déduit, d'après (24), que pour une constante K, $K > 0$:

$$\operatorname{Max}_{1\le\nu\le n} \lambda_\nu^{(n)} \ge \sigma^2 K n^{2d} , \quad \text{pour n assez grand.}$$

b) $\lambda_n^{(n)} \le O(n^{2d})$.

Il suffit maintenant d'évaluer la trace de Γ_n. Il est facile de voir que le $i^{ème}$ terme diagonal de la matrice $\Gamma_n = \sigma^2 L_n L_n'$ vaut :

$$\sum_{h=0}^{i-1} \Pi_h^2 ,$$

et donc la trace est donnée par :

$$\operatorname{Tr} \Gamma_n = \sum_{i=1}^{n} \Big(\sum_{h=0}^{i-1} \Pi_h^2 \Big) = n \Pi_0^2 + \sum_{m=1}^{n-1} \sum_{k=1}^{m} \Pi_k^2 .$$

Par l'encadrement (25) on obtient :

$$0 \le \Pi_k^2 \le (A k^{d-1} + C k^{d-2})^2 , \quad k \ge 1 .$$

Il existe alors une constante K_3 telle que :

$$0 \le \Pi_k^2 \le K_3 k^{2d-2}$$

Il faut ici distinguer deux cas :

- Si $d = \frac{1}{2}$, on a $\Pi_k^2 \le K_3 k^{-1}$, de sorte que

$$\sum_{k=1}^{m} \Pi_k^2 \le K_3 \log(m+1)$$

et donc

$$\operatorname{Tr} \Gamma_n \le n + K_3 \sum_{m=1}^{n-1} \log(m+1) \le n + K_3 \, [\, x\log(x+1) \,]_1^{n-1} \le K_4 n\log n$$

et donc $\lambda_n^{(n)} = \lambda_{Max} \le \operatorname{Tr} \Gamma_n \le O(n\log n)$

- Si $d > \frac{1}{2}$, on a $\sum_{k=1}^{m} \Pi_k^2 \leq K_5 \frac{m^{2d-1}}{2d-1}$

et donc

$$\sum \sum_{k=1}^{m} \Pi_k^2 \leq K_6 \frac{(n-1)^{2d}}{2d(2d-1)}$$

au total on a ici aussi :

$\mathrm{Tr}\ \Gamma_n \leq n\ \Pi_0^2 + K_6 \frac{(n-1)^{2d}}{2d(2d-1)}$ qui est un $O(n^{2d})$, ce qui prouve b).

Q.E.D.

Remarque : l'argument de la démonstration précédente qui repose sur les propriétés des sommes de puissances d'entiers ne s'applique plus, on le voit, dès que le processus est différencié avec un ordre trop faible (i.e. $d \leq \frac{1}{2}$).
En fait la démonstration précédente montre que seul le cas $d = \frac{1}{2}$, c'est à dire précisément la première valeur du degré de différenciation pour laquelle le processus n'est plus stationnaire, pose problème et constitue une "discontinuité" dans la relation simple entre ce degré d et l'ordre de grandeur de la valeur propre maximale.

2) degré de non stationnarité et distribution des "grandes" valeurs propres.

La proposition précédente donne l'ordre de grandeur de la valeur propre maximale et le relie au degré de non stationnarité.
On peut aussi raisonner non plus sur une seule valeur propre (en l'occurence la plus grande) mais sur une partie d'entre elles (on retrouve ainsi la problématique du théorème de Szegö) ; on va voir qu'en effet la distribution asymptotique des "grandes" valeurs propres (la "queue" de la distribution) fournit des indications sur le degré de non stationnarité du processus. La démonstration de ce résultat nécessite que le processus

(toujours supposé ARMA) considéré vérifie le théorème de Szegö ; il ne sera donc pas, en toute rigueur, établi pour tous les processus ARMA non stationnaires mais seulement pour une classe de processus incluant les processus autorégressifs (et aussi les processus fractionnaires stationnaires, puisqu'ils vérifient le théorème de Szegö, voir Accardo(1990)).

Soit $\{ X_t ; t \in \mathbb{Z} \}$ un processus ARMA défini par les conditions (α) et (β) précédentes. On suppose de plus que le polynôme moyenne mobile $\Theta(z)$ n'admet aucune racine dans le cercle unité, de sorte que la densité spectrale

$$f(\omega) = \frac{|\Theta(e^{i\omega})|^2}{|\Phi(e^{i\omega})|^2 |1-e^{i\omega}|^{2d}} \quad , \ 0 \leq \omega \leq 2\pi$$

est minorée par un réel strictement positif m. (Condition (γ)).
On suppose enfin que m minore les valeurs propres : $m \leq \lambda_\nu^{(n)}$, pour tout n et tout ν. (Condition (δ)).

On Pose, pour $M > 0$, $L(M) = \lim_{n\to\infty} \frac{1}{n} \sum_{\lambda>M} \log\lambda$, où on a simplifié les notations en écrivant λ pour $\lambda_\nu^{(n)}$. On a alors :

Proposition 6 :pour tout processus vérifiant d'une part les conditions(α) (β), (γ), (δ) et d'autre part le théorème de Szegö on a :
L(M) est d'ordre $\frac{LogM}{M^{1/d}}$, pour M tendant vers $+\infty$.

Preuve : On va d'abord montrer que

$$(*) \qquad \lim_{n\to\infty} \frac{1}{n} \sum_{\lambda>M} \log\lambda = \frac{1}{2\pi} \int_{f>M} \log[(f(\omega)] \, d\omega \quad , \text{ pour M assez grand.}$$

On évaluera ensuite le second membre de cette égalité, ce qui donnera le résultat.

1) montrons (*)
De la relation $\Gamma_n = \sigma^2 L_n L'_n$ obtenue dans la démonstration

précédente, on déduit l'expression du déterminant de Γ_n :

$$\det\Gamma_n = \sigma^{2n}\Pi_0^{2n}$$

Par conséquent, $\frac{1}{n}\sum_{\nu=1}^{n}\log\lambda_\nu^n = \log\sigma^2 + \log\Pi_0^2$, pour tout n .

Comme par ailleurs, pour tout polynôme complexe à racines de module supérieur ou égal à 1, P(z), on a

$$\int_0^{2\pi} \log|P(z)|^2 d\omega = 0 \ , \ (\text{ avec } z = e^{i\omega}\ ),$$

on voit que le processus X_t vérifie l'égalité de Kolmogorov :

$$\lim \frac{1}{n}\sum_{\nu=1}^{n}\log\lambda_\nu^{(n)} = \frac{1}{2\pi}\int_0^{2\pi} \log\,[f(\omega)]\,d\omega \quad , \tag{26}$$

avec f la pseudo densité spectrale du processus (compte tenu des conditions portant sur les racines des polynômes $\Theta(z)$ et $\Phi(z)$ il est facile de voir que $\Pi_0 = 1$).

On définit alors $\log_M$ comme la fonction $\chi_{[m,M]}\log$, où $\chi_{[m,M]}$note la fonction caractéristique de l'intervalle [m,M].

Les fonctions $\log_M$ sont clairement bornées sur $\mathbb{R}$. Elles ne sont cependant continues que pour $x \neq m$ et $x \neq M$. On ne peut donc appliquer directement le théorème de Szegö (supposé vérifié par le processus). Il faut utiliser le lemme suivant :

Lemme 3 : Si pour toute fonction continue et bornée h on a :

$$\lim \frac{1}{n}\sum_{\nu=1}^{n} h(\lambda_\nu^{(n)}) = \frac{1}{2\pi}\int_0^{2\pi} h[f(\omega)]\,d\omega \ . \tag{27}$$

alors les fonctions $\log_M$ vérifient :

$$\lim \frac{1}{n}\sum_{\nu=1}^{n} \log_M(\lambda_\nu^{(n)}) = \frac{1}{2\pi}\int_0^{2\pi} \log_M\,[f(\omega)]\,d\omega \ . \tag{28}$$

Preuve : voir annexe.

Dans ces conditions il suffit de soustraire (26) et (28) pour avoir, puisque on a supposé toutes les valeurs propres dans [m , +∞ [(condition (δ)):

$$(29) \qquad \lim \frac{1}{n} \sum_{\lambda>M} \log\lambda = \frac{1}{2\pi} \int_{f>M} \log[f(\omega)]d\omega \ .$$

Evaluons alors le second membre $\int_{f>M} \log[f(\omega)]\, d\omega$ de la relation précédente.

2) Evaluation de $\int_{f>M} \log f$

Φ et Θ n'admettant pas, par hypothèse, de racines dans le cercle unité, on sait qu'il existe deux réels r et R , strictement positifs tels que :

$$(30) \quad 0 < \frac{r}{|1 - e^{i\omega}|^{2d}} \leq f(\omega) \leq \frac{R}{|1 - e^{i\omega}|^{2d}} \quad , \quad \text{pour tout } \omega \text{ de }]-\pi,+\pi[\setminus\{0\}.$$

On a donc :

$$(31) \quad \int_{f>M} \log \frac{r}{|1-e^{i\omega}|^{2d}} d\omega \leq \int_{f>M} \log[f(\omega)]\, d\omega \leq \int_{f>M} \log \frac{R}{|1-e^{i\omega}|^{2d}} d\omega$$

On en déduit :

$$(32) \quad (\log r)\ [\ \text{mes}\{\omega\ ; f(\omega)>M\}\ ] \leq \int_{f>M} \log[f(\omega)]d\omega - \int_{f>M} \log \frac{1}{|1-e^{i\omega}|^{2d}} d\omega \leq$$

$$\leq (\log R)\ [\ \text{mes}\{\omega; f(\omega)>M\}\ ]$$

(mesX désignant ici la mesure de Lebesgue de l'ensemble X).

De (30) on tire :

$$(33)\quad \{\ \omega\ ;\frac{r}{|1-e^{i\omega}|^{2d}} > M\ \} \subseteq \{\omega\ ;f(\omega)>M\} \subseteq \{\ \omega\ ;\ \frac{R}{|1-e^{i\omega}|^{2d}} > M\ \}$$

La fonction $\log \frac{1}{|1-e^{i\omega}|^{2d}}$ est positive dans un voisinage de $\omega = 0$; d'autre part :

$$\frac{R}{|1-e^{i\omega}|^{2d}} > M \quad \text{implique que} \quad |\omega| \leq \text{Arccos}\ [\ 1-\frac{1}{2}(\frac{R}{M})^{1/d}\ ] \simeq \frac{1}{2}(\frac{R}{M})^{1/d}$$

de sorte que pour M assez grand on peut déduire de (33) l'encadrement :

$$(34)\quad \int_{A_r} \log \frac{1}{|1-e^{i\omega}|^{2d}} d\omega \leq \int_{f>M} \log \frac{1}{|1-e^{i\omega}|^{2d}} d\omega \leq \int_{A_R} \log \frac{1}{|1-e^{i\omega}|^{2d}} d\omega$$

où A_r (resp A_R) désigne l'ensemble $\{\ \omega\ ;\ |\omega| \leq \text{Arccos}\ [\ 1-\frac{1}{2}(\frac{r}{M})^{1/d}\ ]\}$ (resp l'ensemble $\{\ \omega\ ;\ |\omega| \leq \text{Arccos}\ [\ 1-\frac{1}{2}(\frac{R}{M})^{1/d}\ ]\}$). On peut alors trouver $a_r(M)$ (resp.$a_R(M)$) tel que :

$$(35)\qquad a_r(M) \simeq \frac{1}{2}(\frac{r}{M})^{1/d} \quad (\text{resp.}\ a_R(M) \simeq \frac{1}{2}(\frac{R}{M})^{1/d})\ ,$$

et vérifiant :

$$(36)\quad \int_0^{a_r(M)} \log \frac{1}{|1-e^{i\omega}|^{2d}} d\omega \leq \int_{f>M} \log \frac{1}{|1-e^{i\omega}|^{2d}} d\omega \leq \int_0^{a_R(M)} \log \frac{1}{|1-e^{i\omega}|^{2d}} d\omega$$

Si on développe $\cos\omega$ au voisinage de zéro, autrement dit si on écrit

$$1-\cos\omega = \frac{\omega^2}{2} + \omega^2\varepsilon(\omega),\ \text{avec}\ \varepsilon(\omega) \to 0\ \text{quand}\ \omega \to 0,$$

on trouve alors :

$$d\int_0^{a_r(M)} \log \frac{1}{|1-e^{i\omega}|^{2}}\, d\omega = -2d\int_0^{a_r(M)} \log\omega\ d\omega\ -d\int_0^{a_r(M)} \log(1+2\varepsilon(\omega))\ d\omega,$$

$$= -2d\ a_r(M)\ \log[a_r(M)]\ + 2d\ a_r(M) - d\ a_r(M)\ \eta_r(M),$$

avec $|\eta_r(M)| \leq \text{Max}\ \{\ |\log(1+2\varepsilon(\omega))|;\ 0 \leq \omega \leq a_r(M)\ \}$, de sorte que $\eta_r(M)$ tend vers 0 quand M tend vers $+\infty$.
On a de même :

$$d\int_0^{a_R(M)} \log\frac{1}{|1 - e^{i\omega}|^2}\, d\omega = -2d\ a_R(M)\ \log[a_R(M)]\ + 2d\ a_R(M) - d\ a_R(M)\ \eta_R(M),$$

avec $\eta_R(M)$ tendant vers 0 quand M tend vers $+\infty$.
Reportant ces deux égalités dans (36), on obtient :

$$(37)\quad -2d\ a_r(M)\ \log[a_r(M)]\ + 2d\ a_r(M) - d\ a_r(M)\ \eta_r(M) \leq \int_{f>M} \log\frac{1}{|1 - e^{i\omega}|^{2d}}\, d\omega$$

$$\leq -2d\ a_R(M)\ \log[a_R(M)] + 2d\ a_R(M) - d\ a_R(M)\ \eta_R(M).$$

D'autre part, d'aprés (33), on a :

$$\text{mes}\ A_r \leq \text{mes}\ \{\ f > M\ \} \leq \text{mes}\ A_R,$$

ou encore,

$$2\ a_r(M) \leq \text{mes}\ \{\ f > M\ \} \leq 2\ a_R(M).$$

De sorte que (32) entraine que :

$$2\ a_r(M)\ \log r \leq \int_{f>M} \log[f(\omega)]\ d\omega - \int_{f>M} \log\frac{1}{|1-e^{i\omega}|^{2d}}\, d\omega \leq 2\ a_R(M)\ \log R.$$

Utilisant (36) et (37) on obtient :

$$(38)\quad -2d\ a_r(M)\ \log[a_r(M)] + 2d\ a_r(M) - d\ a_r(M)\ \eta_r(M) + 2\ a_r(M)\ \log r \leq$$

$$\leq \int_{f>M} \log f \leq -2d\ a_R(M)\ \log[a_R(M)] + 2d\ a_R(M) - d\ a_R(M)\ \eta_R(M) + 2\ a_R(M)\ \log R.$$

Etant donné les équivalents de $a_r(M)$ et de $a_R(M)$ fournis par (35) pour M tendant vers $+\infty$, il est facile de calculer que le premier membre de (38) est équivalent pour M grand à :

$$(\frac{r}{M})^{\frac{1}{d}} \log M.$$

De même, le dernier membre est équivalent à :

$$(\frac{R}{M})^{\frac{1}{d}} \log M.$$

Ainsi,

$$\limsup_{M\to\infty} \frac{M^{\frac{1}{d}}}{\log M} \frac{1}{2\pi} \int_{f>M} \log[f(\omega)]\, d\omega \leq R^{\frac{1}{d}},$$

et de même,

$$r^{\frac{1}{d}} \leq \liminf_{M\to\infty} \frac{M^{\frac{1}{d}}}{\log M} \frac{1}{2\pi} \int_{f>M} \log[f(\omega)]\, d\omega ,$$

ce qui donne le résultat.

Q.E.D.

ANNEXE :

1) Preuve de la proposition 1 :

On montre la relation

$$D_{n,p}(\lambda) = [(p+1) - \lambda(p+2)]\, D_{n-1,0}(\lambda) - (p+1)\lambda^2\, D_{n-2,0}(\lambda)$$

pour $n \geq 3$.

Soit donc

$$D_{n,p}(\lambda) = \begin{bmatrix} p+1-\lambda & p+1 & p+1 & \dots & p+1 \\ p+1 & p+2-\lambda & p+2 & \dots & p+2 \\ \vdots & \vdots & & & \\ p+1 & & \dots & & p+1-\lambda \end{bmatrix}$$

Soustrayant la première colonne aux suivantes, on obtient

$$D_{n,p}(\lambda) = \begin{bmatrix} p+1-\lambda & \lambda & \lambda & \dots & \lambda \\ p+1 & 1-\lambda & 1 & \dots & 1 \\ \vdots & 1 & 2-\lambda & & 2 \\ \vdots & \vdots & \vdots & & \\ p+1 & 1 & 2 & & n-1-\lambda \end{bmatrix}$$

La seconde colonne est alors multipliée par p+1 puis soustraite à la première colonne ; on trouve :

$$D_{n,p}(\lambda) = \begin{bmatrix} p+1-\lambda(p+2) & \lambda & \lambda & \dots & \lambda \\ (p+1)\lambda & 1-\lambda & 1 & & 1 \\ 0 & 1 & 2-\lambda & & 2 \\ 0 & 1 & 2 & & \\ \vdots & \vdots & \vdots & & \\ 0 & 1 & 2 & & n-1-\lambda \end{bmatrix}$$

Il suffit alors de développer ce déterminant par rapport à la première colonne pour obtenir :

$$D_{n,p}(\lambda) = [(p+1) - \lambda(p+2)]\, D_{n-1,0}(\lambda) - (p+1)\lambda\, \hat{D}$$

avec :

$$\hat{D} = \begin{bmatrix} \lambda & \lambda & \cdots & \lambda \\ 1 & 2-\lambda & & 2 \\ 1 & 2 & & 3 \\ \vdots & \vdots & & \vdots \\ 1 & 2 & & n-1-\lambda \end{bmatrix}$$

Pour finir on soustrait dans $\hat{D}$ la première colonne à toutes les autres pour avoir :

$$\hat{D} = \begin{bmatrix} \lambda & 0 & 0 & \cdots & 0 \\ 1 & 1-\lambda & 1 & \cdots & 1 \\ 1 & 1 & 2-\lambda & & 2 \\ \vdots & \vdots & & & \\ 1 & 1 & & & n-2-\lambda \end{bmatrix}$$

ce qui donne immédiatement :

$$\hat{D} = \lambda\, D_{n-2,0}(\lambda)$$

et donc la relation (1)

Conséquence :la relation (2) est du même coup démontrée puisqu'elle correspond à (1) dans le cas p = 0.

Q.E.D.

2) Preuve du lemme 1:

Reportant dans (4) l'expression de $D_{n-1,0}$ et $D_{n-2,0}$ tirée de (3) on obtient pour $n \geq 2$

$$D_{n,p} = [(p+1) - \lambda(p+2)]\, \{(1-2\lambda)\, D_{n-2,0} - \lambda^2\, D_{n-3,0}\} - \lambda^2(p+1)\, \{(1-2\lambda)\, D_{n-3,0} - \lambda^2\, D_{n-4,0}\}$$

c'est à dire

$$D_{n,p} = (1-2\lambda)\, \{\, [(p+1)-\lambda(p+2)]\, D_{n-2,0} - \lambda^2(p+1)\, D_{n-3,0}\} - \lambda^2\, \{\, [(p+1)-\lambda(p+2)]\, D_{n-3,0} - \lambda^2(p+1)\, D_{n-4,0}\}$$

soit finalement

$$D_{n,p} = (1-2\lambda)\, D_{n-1,p} - \lambda^2 D_{n-2,p}$$

pour $n \geq 2$ Q.E.D.

3) Preuve de la proposition 3 :

Pour $n \geq p$ le vecteur $(X_n^1) = (X_1, X_2 \ldots, X_n)'$ se déduit du vecteur V_n^1 , avec $V_n^1 = (X_1, \ldots, X_p, \varepsilon_{p+1}, \ldots, \varepsilon_n)'$ par la relation :

$$X_n^1 = QV_n^1$$

où Q est une matrice triangulaire inférieure dont la diagonale principale est composée uniquement de 1.On en déduit, en prenant les covariances:

$$\Gamma_n = QMQ'$$

où M est la matrice nxn :

$$M = \begin{bmatrix} \mathrm{Var}(X_p^1) & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$

Si on suppose $\det(\mathrm{var} X_p^1)$ non nul, alors $\det(\Gamma)_n$ est non nul et Γ_n est inversible, pour tout n. L'inversion de Γ_n peut s'effectuer comme on l'a vu dans le cadre plus général des processus multivariés, en utilisant la forme autorégressive du processus. Rappelons que dans cette inversion la stationnarité (ou la non stationnarité) ne joue aucun rôle. On sait alors que Γ_n^{-1}est de la forme :

$$\Gamma_n^{-1} = \Gamma_{n,y} + E_n$$

où $\Gamma_{n,y}$ est la matrice de covariance d'ordre n du processus $\{ y_t ;\ t \in \mathbb{Z} \}$,

avec $y_t = \varepsilon_t - \sum_{j=1}^{p} a_j \varepsilon_{t-j}$, qui est un processus MA(p) stationnaire (mais éventuellement non inversible) "dual" du processus X_t.

E_n de son côté désigne une matrice nulle partout sauf dans les coins Nord-ouest et Sud-est, occupés par deux matrices de taille p, A et B.

$$E_n = \begin{bmatrix} A & & \\ & 0 & \\ & & B \end{bmatrix}$$

On voit ainsi que Γ_n^{-1}et $\Gamma_{n,y}$ coïncident sur un sous-espace de dimension n-2p. On en déduit (Voir Accardo (1990), propriété 2.12) que :

$$(*) \qquad m_{\nu-2p}^{(n)} \leq \frac{1}{\lambda_{n+1-\nu}^{(n)}} \leq m_{\nu+2p}^{(n)} , \qquad \nu = 2p+1,\ldots,n-2p.$$

où les $\lambda_\nu^{(n)}$ sont rangées dans l'ordre croissant. Le processus $\{ y_t ; t \in \mathbb{Z}\}$ étant stationnaire on peut approcher ses valeurs propres $m_\nu^{(n)}$par :

$$(**) \qquad \mu_{\nu-2p}^{(n)} \leq m_\nu^{(n)} \leq \mu_{\nu+2p}^{(n)} , \quad \nu = 2p+1,\ldots,n-2p .$$

où $\mu_\nu^{(n)}$ désigne les quantités

$$\prod_{j=1}^{p} \left| 1-\theta_j \exp(i\frac{k\pi}{n+1}) \right|^2 , \quad k = 1,\ldots n.$$

(Voir Accardo(1990)).

De (*) et (**) on déduit immédiatement la relation (20).

Q.E.D.

4) Preuve du lemme 3 :

Soit $\varepsilon > 0$. Pour tout réel $t > 0$,on définit :

$I_t = [t-\varepsilon,t]$, $S_t = [t,t+\varepsilon]$, $W_t = [t-\varepsilon,t+\varepsilon]$.

A) On supose d'abord que $m \geq 1$.

1) la fonction $\log_M$ est donc positive ou nulle. On définit alors la fonction h par :

$$h(x) = \frac{1}{\varepsilon} \log m\,(x-m+\varepsilon) \qquad \text{si } x \in I_m$$
$$h(x) = \log x \qquad \text{si } x \in [m,M]$$
$$h(x) = \frac{1}{\varepsilon} \log M\,(-x+M+\varepsilon) \qquad \text{si } x \in S_M$$
$$h(x) = 0 \qquad \text{sinon.}$$

h est donc continue, positive ou nulle et bornée. D'où, par hypothèse, h vérifie la relation (27).

D'autre part, h majore clairement $\log_M$ et on a ainsi :

$$0 \leq \frac{1}{n} \sum_{\nu=1}^{n} \{\, h(\lambda_\nu^{(n)}) - \log_M (\lambda_\nu^{(n)}) \,\} = \frac{1}{n} \sum_{\lambda \in I_m \cup S_M} h(\lambda_\nu^{(n)}).$$

2) On considère alors la fonction k définie par :

$$k(x) = \frac{1}{\varepsilon} \log m\,(-x+m+\varepsilon) \qquad \text{si } x \in S_m$$
$$k(x) = h(x) \qquad \text{si } x \in S_M \cup I_m$$
$$k(x) = \frac{1}{\varepsilon} \log M\,(x-M+\varepsilon) \qquad \text{si } x \in I_M$$
$$k(x) = 0 \qquad \text{sinon.}$$

On a alors une fonction continue, positive ou nulle et bornée qui vérifie donc la relation (27).

Puisque $k = h$ sur $S_M \cup I_m$, on a :

$$\frac{1}{n} \sum_{\lambda \in I_m \cup S_M} h(\lambda_\nu^{(n)}) \leq \frac{1}{n} \sum_{\nu=1}^{n} k(\lambda_\nu^{(n)}).$$

On en déduit la majoration :

$$0 \leq \limsup_{n \to \infty} \frac{1}{n} \sum_{\lambda \in I_m \cup S_M} h(\lambda_\nu^{(n)}) \leq \frac{1}{2\pi} \int_{-\pi}^{+\pi} k[f(\omega)]\,d\omega\,,$$

le dernier terme étant clairement majoré par :

$$\frac{1}{2\pi}(\log M)\ \mathrm{mes}[\omega; f(\omega)\in W_m \cup W_M]\ ,\ \text{qu'on note } K\ ,$$

(mesX désignant ici la mesure de Lebesgue de l'ensemble X).

Il est de même immédiat de constater que :

$$\frac{1}{2\pi}\int_{-\pi}^{+\pi}(h-\log_M)[f(\omega)]d\omega \le K\ . \qquad (\alpha)$$

3) Au total on a :

$$\limsup_n \frac{1}{n}\sum_{\nu=1}^{n}\{\ h(\lambda_\nu^{(n)}) - \log_M(\lambda_\nu^{(n)})\ \} \le K\ . \qquad (\beta)$$

(α) et (β) sont vérifiées pour tout $\varepsilon > 0$. Par ailleurs, la forme de la fonction f implique que $\mathrm{mes}[\omega; f(\omega)\in W_m \cup W_M]$ tend vers 0 quand ε tend vers 0 . On obtient donc :

$$\lim_n \frac{1}{n}\sum_{\nu=1}^{n}\log_M(\lambda_\nu^{(n)}) = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\log_M[f(\omega)]\ d\omega . \quad (27).$$

B) Si $0 < m < 1$, il suffit de remarquer que :

$$\log_M = \chi_{[m,1]}\log + \chi_{[1,M]}\log \overset{\text{déf}}{=} a + b\ .$$

-a et b sont des fonctions <u>positives</u>, bornées et continues sauf en un point (m ou M). le raisonnement de A) s'applique alors; a et b vérifient donc (27) et donc leur somme $\log_M$ aussi.

Q.E.D.

BIBLIOGRAPHIE

ABRAMOWITZ M. et STEGUN I.A. (1985) : "Handbook of Mathematical Functions", Dover.

ACCARDO J. (1990) : "Approximation des Valeurs Propres de la Matrice de Covariance d'un Processus Stationnaire" - D.T. n° 9003 INSEE.

AKAIKE H. (1973) : "Block Toeplitz Matrix inversion", SIAM J. Appl. Math., Vol 23, n°2,234-241.

ALI M. (1977) : "Analysis of ARMA Models : Estimation and Prediction", Biometrika, 65, 535-545.

AZENCOTT R. et DACUNHA-CASTELLE D. (1983) : "Series d'Observations Irrégulières : Modélisation et Prévision", Masson, Paris.

BOX G. et JENKINS G. (1970) : "Time Series Analysis : Forecasting and Control", Holdenday.

CADZOW J.A. (1984) : "SVD Representation of Unitary Invariant Matrices", IEEE TASSP, vol ASSP-32, n° 3.

CHATELIN F. (1988) : "Valeurs Propres de Matrices", Masson, Paris.

CLEVELAND W. (1972) : "The Inverse Autocorrelations of Time Series and their Application", Technometrics, 14, 277-293.

CYBENKO G. (1984) : "On the Eigenstructure of Toeplitz Matrices", IEEE TASSP, vol ASSP-32, n° 4, pp 918-925.

DELSARTE P. et GENIN Y. (1983) : "Spectral Properties of Finite Toeplitz Matrices", in Proc. 1983 Int. Symp. Math. Theory of Networks and Syst., Beer-Sheva, Israel.

DELSARTE P. et GENIN Y. (1986) :"The Split Levinson Algorithm", IEEE TASSP, Vol ASSP-34, n° 3, pp 472-478.

DIXON W. J. (1944) : "Further Contributions to the Problem of Serial Correlation", Ann. Maths. Statistics, 15, 119-144.

DOOB J. L. (1953) : "Stochasic Process", Wiley.

FOX L. (1964) : "An Introduction to Numerical linear Algebra", Oxford University Press.

GLAZMAN I. et LIUBITCH Y. (1974) : " Analyse Linéaire dans les Espaces de Dimensions Finies", Editions de Moscou.

GOLUB G. et VAN LOAN C. (1984) : "Matrix Computations", Baltimore MD:John Hopkins University Press.

GONCALVES E. (1987) : "Une Généralisation des Processus ARMA", Annales d'Economie et de Statistique 5, 109-146.

GOURIEROUX C. et MONFORT A (1990) : "Series Temporelles et Modèles Dynamiques", Economica.

GRANGER C. (1978) : "New Classes of Time Series Models", Statistician 3-4, 27, 237-253.

GRENANDER U. et SZEGO G. (1958) : "Toeplitz Forms and their Applications", University of California Press.

HOSKING J. (1981) : "Fractional Differencing", Biometrika 68, 165-176.

KAC M., MURDOCK W.L. et SZEGO G. (1953) : "On the Eigenvalues of Certain Hermitian Forms", Journal of Rational Mechanics and Analysis, 2, 767-800.

KOBILINSKY A. (1979) : "Ordre entre Formes Quadratiques. Application à l'Optimalité des Sous-Espaces en Analyse des Données", Revue de Statistique Appliquée, XXVII, 45-54.

LELONG-FERRAND J. et ARNAUDIES J.M. (1974) : "Cours de mathématiques, Algèbre",

MAKHOUL J. (1975) : "Linear Prediction : A Tutoral Review", Proc IEEE, vol 63, pp 561-580.

MANDELBROT B. et VAN NESS J. (1968) : "Fractional Brownian Motions, Fractional Noises and Applications", SIAM Review 10, 422-437.

METIVIER M. (1979) : "Notions Fondamentales de la Théorie des Probabilités", Dunod.

PARTER S. V. (1961) : "On the Extreme Eigenvalues of Toeplitz Matrices", Trans. Math. Soc. 100, 263-276.

PRIESTLEY M. (1981) : "Spectral analysis and Time Series, Vol 1 : Unvariate Series", Academic Press.

RELLICH F. (1969) : "Perturbation Theory of Eigenvalue Problems", Gordon and Breach.

SHAMAN P. (1975) : "An Approximate Inverse for the Covariance Matrix of Moving Average and Autoregressive Process", Am. Statist., 3, 532-538.

SHAMAN P. (1976) : "Approximation for Stationnary Covariance Matrix and their Inverses with Application to ARMA Models", Am. Statist., 4, 292-301.

SIDDIQUI M.M. (1958) : "On the Inversion of the Sample Covariance Matrix in a Stationnary Autoregressive Process", Ann. Math.Statist.,29, 585-588.

STROEKER R. J. (1978) : "Note on the Eigenvalues of the Covariance Matrix of Disturbances in the General Linear Model", Report 7818/M, Econometric Institute, Erasmus University, Rotterdam.

STROEKER R. J. (1983) : "Approximation of the Eigenvalues of the Covariance Matrix of an AR(1) Process", Journal of Econometrics, 22, 269-279.

SZEGO G. (1921) : "Beitrage zur Theorie der Toeplitzschen Formen I, II", Mathematische Zeitschrift, Vol 6, 196-202 et Vol 9, 167-190.

SZEGO G. (1939) : "Orthogonal Polynomials", Amer. Math. Soc., Colloquium Publications 3, New-York.

TOEPLITZ O. (1911) : "Zur Theorie der Quadratischen und Bilinearen Formen von Wendlichvielen Veränderlichen, I Teil : Theorie der L Formen", Mathematischen Annalen, Vol 70, 351-376.

VALIRON G. (1951) : "Théorie des fonctions", Masson, Paris.

VERBYLA P. (1985) : "A Note on the Inverse Covariance Matrix of the Autoregressive Process", Australian Journal of Statistics 27, 2, 221-224.

WHITTLE P. (1963) : "Prediction and Regulation by Linear Least Square Methods", the English Universities Press.

WIDOM H. (1958) : "On the Eigenvalues of Certain Hermitian Operators", Trans. Amer. Math. Soc. 88, 491-522.

WILF H.S. (1970) : "Finite Sections of Some Classical Inequalities", Springer-Verlag.

WILKINSON J. M. (1965) : "The Algebraic Eigenvalue Problem", Oxford University Press.

WISE J. (1955) : "The Autocorrelation Function and the Spectral Density Function", Biometrika, Vol 42.

ZHANG H.M. (1989) : "Algorithmes Rapides et Matrices de Toeplitz", Thèse de Doctorat, ENS Telecom, Paris.