

STATISTIQUE ET ANALYSE DES DONNÉES

JEAN-CLAUDE DEVILLE

Décomposition des processus aléatoires, une comparaison entre les méthodes factorielles et l'analyse spectrale

Statistique et analyse des données, tome 6, n° 3 (1981), p. 5-33.

http://www.numdam.org/item?id=SAD_1981__6_3_5_0

© Association pour la statistique et ses utilisations, 1981, tous droits réservés.

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

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

Statistiques et Analyse des données
1981 - 3 - pp. 5-33

DECOMPOSITION DES PROCESSUS ALEATOIRES
UNE COMPARAISON ENTRE LES METHODES FACTORIELLES
ET L'ANALYSE SPECTRALE

Jean-Claude DEVILLE

INSEE

18, Bd Adolphe Pinard - 75675 PARIS Cédex 14

Résumé : On examine les rapports entre l'analyse factorielle des processus du second ordre (développement de Karuhnen-Loève) et l'analyse spectrale des processus stationnaires. Dans un cas l'ensemble des temps doit être compact, dans le second, ce doit être un groupe. On s'intéresse donc aux processus définis sur des groupes compacts ; dans le cas stationnaire on montre que l'analyse factorielle se déduit de la représentation spectrale après "transfert" par la densité spectrale. On généralise ensuite la décomposition de Karuhnen-Loève à des cas où l'espace des temps est \mathbb{Z} puis \mathbb{R} et on montre que, si on ajoute l'hypothèse de stationnarité, cette représentation peut encore s'obtenir par transport de la densité spectrale.

On montre, enfin, que cette propriété ne peut pas être vérifiée dans le cas de champs aléatoires.

Summary : The relations between factor analysis of second order random processes (Karuhnen-Loève expansion) and spectral analysis are examined. In the first case, the set of times has to be compact ; in the second it has a group structure. Therefore we look at processes defined on a compact group and show that, in the stationary case, factor analysis results of the "transportation" of the spectral representation by the spectral density fonction. An extension of the Karuhnen - Loève expansion is then constructed for some processes on \mathbb{Z} or \mathbb{R} . In the stationary case, this new representation is once more given by transportation of the spectral representation.

Unfortunately, this property cannot hold for random fields.

Mots clés : *Analyse harmonique, Représentation spectrale, Processus du second ordre, Processus stationnaires, Décomposition de Karuhnen-Loève.*

Toutes les données que manipulent les statisticiens sont munies de structures de complexité variable : suite de nombres, tableaux à entrées multiples, schémas comptables très élaborés etc. On arrive souvent à les décrire de façon satisfaisante en s'appuyant sur certains éléments pertinents des structures ainsi mises en jeu ; l'information se transforme, se simplifie, devient plus parlante par le simple fait de calculer des éléments dont l'existence est assurée par la structure initiale des données.

Un exemple à la fois général et banal est celui du fichier organisé par individus (repérés par un indice i variant de 1 à N) et par variables (repérées par l'indice j variant de 1 à P). La structure de ces données est alors celle d'une matrice X de nombres réels à N lignes et P colonnes. On peut alors calculer d'autres éléments dans cette structure comme, par exemple, la matrice des covariances $\frac{1}{N} XX^t$ qui est symétrique définie positive. Le spectre et les vecteurs propres de cette matrice constituent une information déduite de la structure qui permet une description des données : c'est l'analyse en composante principale.

Dans cet exemple, on s'intéressait à une structure assez formelle de tableau à double entrée. La "structure" comporte aussi des éléments "sémantiques" : sur quels individus portent les mesures ? Comment celles-ci ont-elles été réalisées ? Quelles transformations ont été faites avant la mise en tableau ? Quel est le mode d'échantillonnage ? Les réponses à toutes ces questions font aussi partie de la structuration des données et il vaut mieux en tenir compte.

L'information qu'on obtient est très particulière quand les données sont obtenues par des mesures à des instants successifs de la même variable sur les mêmes individus ou qu'elles permettent de reconstituer des calendriers d'événements. Le langage naturel pour décrire ces données est alors celui des processus aléatoires. Certaines hypothèses, plus ou moins techniques, deviennent nécessaires pour pouvoir en dire quelque chose, pour les organiser et les décrire. L'analyse des données sur processus a pour but le calcul d'éléments liés à la structure "processus" au vu de données considérées comme des réalisations de celui-ci.

Dans cet article on s'efforcera de comparer deux techniques de représentation des processus : l'analyse harmonique (ou factorielle) et l'analyse spectrale. Elles s'appliquent en principe à des classes différentes de processus et donc à des structures de données différentes. Dans certains cas particulier, il y a néanmoins coïncidence et on montrera qu'alors les représentations qu'on obtient sont équivalentes. Pour une classe assez grande de processus en temps discret on exhibera une analyse factorielle un peu plus générale qui se ramène dans certains cas à l'analyse factorielle habituelle, dans d'autres à l'analyse spectrale. Une construction analogue est pensable pour les processus en temps continu au prix d'un appareillage mathématique plus compliqué ; elle ne sera entreprise que de façon très allusive à la fin de cet article.

*
* *

Dans toute la suite X_t désignera un processus du second ordre à valeurs complexes, séparable et mesurable de moyenne nulle défini sur un ensemble de "temps" T muni d'une mesure σ -finie qu'on notera toujours dt . La fonction de covariance sur $T \times T$ est notée $C(t,s) (=EX_t \overline{X_s})$. Dans les cas où T est un groupe abélien localement compact, dt sera toujours une mesure de Haar invariante par translation. Si X_t est alors stationnaire, on conviendra que pour tout $t \in T$ $E|X_t|^2 = 1$; la fonction de covariance ne dépendra que de la différence entre les temps et on écrira $C(t,s) = \rho(t-s)$.

Enfin $H(X)$ désignera l'espace de Hilbert des variables engendrées par le processus, c'est-à-dire le plus petit espace de Hilbert contenant toutes les variables X_t pour t variant dans T .

1 - ANALYSE HARMONIQUE (OU FACTORIELLE DES PROCESSUS DU SECOND ORDRE)

La variance totale du processus est supposée finie en ce sens que :

$$\text{Var}(X) = \int_T \text{Var}(X_t) dt < +\infty. \quad (1-1)$$

Cette condition est réalisée, en particulier, dès que T est un intervalle compact de \mathbb{R} et que le processus est continu en moyenne quadratique ; ce qui constitue le cas pratique le plus habituel.

On peut, dès lors, affirmer les résultats suivants [1].

L'opérateur linéaire : $f \xrightarrow{C} g$ de $L^2(T, dt)$ défini par :

$$g(t) = \int_T C(t,s) f(s) ds \quad (1-2)$$

est hermitien compact positif. Son spectre est constitué d'une suite λ_i décroissante vers zéro de valeurs propres positives. Chaque sous-espace propre est de dimension finie et on a :

$$\text{Var}(X) = \sum_{i=1}^{\infty} \lambda_i \quad (1-3)$$

à condition de répéter λ_i autant de fois que la dimension du sous-espace propre qui lui est associé. Il existe alors une famille $f_i(t)$ de fonctions propres de C telles que :

$$\int_T f_i(t) \overline{f_j(t)} dt = \delta_{ij} \quad (\text{Symbole de Kronecker}) \quad (1-4)$$

formant une base hilbertienne de $\overline{\text{Im}C}$, et une famille ξ_i de variables scalaires données par :

$$\xi_i = \lambda_i^{-1/2} \int_T X(t) \overline{f_i(t)} dt \quad (1-5)$$

telles que :

$$E\xi_i = 0 \quad E\xi_i \overline{\xi_j} = \delta_{ij} \quad (1-6)$$

forment une base hilbertienne de $H(X)$.

Le processus admet alors la représentation :

$$X_t = \sum_{i=1}^{\infty} \lambda_i^{1/2} \xi_i f_i(t) \quad (1-7)$$

Cette série converge en m.q. pour tout t ainsi qu'en variance totale.

On définit également l'opérateur d'Escoufier V qui opère sur les variables par :

$$V\xi = \int_T E(\xi \overline{X_t}) X_t dt \quad (1-8)$$

L'image de V n'est autre que $H(X)$, son spectre est le même que celui de C et les variables ξ_i en sont vecteurs propres associés aux λ_i . On a enfin les représentations suivantes, convergentes en normes d'opérateurs :

$$C = \sum_{i=1}^{\infty} \lambda_i f_i \otimes f_i \quad (1-9)$$

$$V = \sum_{i=1}^{\infty} \lambda_i \xi_i \otimes \xi_i \quad (1-10)$$

Le symbole $x \otimes y$ désigne l'opérateur linéaire de rang 1 :

$$a \longmapsto (a|x)y \quad (1-11)$$

Notons enfin J l'opérateur linéaire suivant qui transforme les variables en fonctions du temps :

$$J\xi(t) = E(\xi \bar{X}_t). \quad (1-12)$$

L'adjoint J^* (qui transforme des fonctions de $L^2(T)$ en variables) est donné par :

$$J^*f = \int_T f(t) \bar{X}_t dt \quad (1-13)$$

On a alors les relations suivantes (qui constituent le schéma de dualité) :

$$C = J J^* \quad (1-14)$$

$$V = J^* J \quad (1-15)$$

Soit W l'isomorphisme d'espaces de Hilbert entre $H(X)$ et $\overline{\text{Im}C}$ défini par :

$$W\xi_i = f_i \quad (i=1, \dots) \quad (1-16)$$

Autrement dit pour tout $\xi \in H(X)$ on aura :

$$W\xi = \sum_i E(\xi \bar{\xi}_i) f_i \quad \text{et} \quad W^{-1}f = \sum_i (f|f_i) \xi_i \quad (1-17)$$

On a alors la congruence suivante entre les opérateurs V et C :

$$C = W V W^{-1} \quad (1-18)$$

$$V = W C W^{-1} \quad (1-19)$$

L'analyse harmonique est l'utilisation de ces éléments dérivés de la structure du processus en vue de sa description. On s'aide alors de l'allure des fonctions $f_i(t)$ et de la signification concrète des variables X_t . Les composantes ξ_i s'interprètent nécessairement par des mesures de liaison (corrélations en général) avec d'autres variables disponibles au niveau individuel ou au niveau de groupes d'individus pour lesquels on dispose de trajectoires du processus.

2 - ANALYSE SPECTRALE (OU TRANSFORMATION DE FOURIER DES PROCESSUS)

La structure mathématique qui sous-tend l'analyse spectrale est généra-

lement très mal présentée dans la littérature relative à ce sujet*. Celle-ci aide cependant à comprendre la portée des hypothèses qu'on fait et en quoi ce mode de description des données se fait à partir d'un modèle plus complexe que celui du numéro 1.

2.a - Groupes

T a maintenant une structure de groupe abélien localement compact. La mesure dt est une mesure de Haar (normalisée à 1 si T est compact). Le processus X_t est supposé stationnaire au sens large et continu en moyenne quadratique (ce qui n'engage à rien dans le cas où T est un groupe discret).

Un caractère sur T est un homomorphisme dans le groupe multiplicatif \mathbb{T} des nombres complexes de module 1. Pour laisser aux notations un sens assez intuitif nous noterons un tel caractère ω par :

$$t \longmapsto e^{i(\omega, t)}$$

On a donc, par définition, la loi de T étant notée additivement :

$$\left. \begin{aligned} \cdot e^{i(\omega, t+s)} &= e^{i(\omega, t)} e^{i(\omega, s)} \\ \cdot e^{i(\omega, 0)} &= 1 \\ \cdot e^{i(\omega, -t)} &= e^{-i(\omega, t)} = \overline{e^{i(\omega, t)}} \end{aligned} \right\} \quad (2-1)$$

Le produit de deux caractères est encore un caractère et on peut, au prix de menues précautions, avoir la notation suivante :

$$e^{i(\omega, t)} e^{i(\omega', t)} = e^{i[(\omega, t) + (\omega', t)]} = e^{i(\omega + \omega', t)} \quad (2-2)$$

L'application $e^{i(0, t)} = 1$ est élément neutre pour cette loi de composition et l'ensemble des caractères se trouve muni d'une structure de groupe abélien localement compact si on le munit de la topologie de la convergence uniforme sur toute partie compacte de T. On note \hat{T} ce groupe dual et on montre (Th. de PONTRJAGIN) que $\hat{\hat{T}} = T$ à un vague isomorphisme près, ce qui justifie une notation de dualité. On peut alors choisir sur \hat{T} une mesure de Haar $d\omega$ telle que la transformation de Fourier de $L^2(T, dt)$ dans $L^2(\hat{T}, d\omega)$ soit une isométrie :

$$\cdot \hat{f}(\omega) = \int_T e^{i(\omega, t)} f(t) dt \quad (\text{transformation de Fourier}) \quad (2-3)$$

$$\cdot \int_{\hat{T}} |\hat{f}(\omega)|^2 d\omega = \int_T |f(t)|^2 dt \quad (\text{identité de Parseval Plancherel}) \quad (2-4)$$

Pour plus de détail on peut se reporter à [2] [3] ou [4] entre autres.

* Il s'agit là d'une affaire de goût, bien entendu, mais quand même !

2.b - Représentation unitaire

On définit dans $H(X)$ une famille d'opérateurs linéaires U_h ($h \in T$; "opérateurs de décalage") tels que :

$$\forall t \in T : U_h X_t = X_{t+h} \quad (2-5)$$

Par linéarité ces opérateurs sont définis pour toute combinaison linéaire des X_t . Du fait de la stationnarité du processus, les U_h sont des isométries sur cette partie partout dense de $H(X)$ et s'étendent donc à des isométries de l'espace tout entier. On obtient de façon immédiate que :

$$\left. \begin{aligned} \cdot U_{h+k} &= U_h U_k = U_k U_h \\ \cdot U_0 &= 1 \\ \cdot U_{-h} &= U_h^* = U_h^{-1} \end{aligned} \right\} \quad (2-6)$$

L'application $h \longmapsto U_h$ est un homomorphisme de T dans $L(H)$ c'est-à-dire une représentation unitaire de T . Cette représentation est fortement continue c'est-à-dire que pour tout $X \in H$ l'application $h \longmapsto U_h X$ est continue de T dans H . Ceci résulte directement que la continuité m.q. de X_t .

Le théorème de Stone ([4],[5],[6]) affirme alors l'existence d'une mesure spectrale ⁽¹⁾ $P(d\omega)$ sur \hat{T} à valeurs dans l'ensemble du projecteur de H , définie de façon unique et telle que l'on ait :

$$\forall h \in T \quad U_h = \int_{\hat{T}} e^{i(\omega, h)} P(d\omega) \quad (2-7)$$

(1) Voir note page suivante.

Note de la page précédente

Rappelons la définition et les principales propriétés des mesures spectrales. Etant donnés un espace de Hilbert (séparable) H et un espace mesurable (A, \mathcal{Q}) , une mesure (ou famille) spectrale est une fonction qui à toute partie $B \in \mathcal{Q}$ associe un projecteur orthogonal (et donc un sous-espace) de H , $P(B)$ et qui vérifie :

- . $B_1 \cap B_2 = \emptyset \Rightarrow \text{Im } P(B_1) \perp \text{Im } P(B_2) \Leftrightarrow P(B_1) P(B_2) = P(B_2) P(B_1) = 0$
- . $P(\emptyset) = 0$ et $P(A) = 1$ (identité de H)
- . Si les B_n sont disjoints ($n=1,2,\dots$) $P(\cup_n B_n) = \sum_n P(B_n)$

Ce qui équivaut à $\text{Im } P(\cup_n B_n) = \bigvee_n \text{Im } P(B_n) = \sum_n \text{Im } P(B_n)$.

Si x est un vecteur quelconque de H , $m_x(B) = (x|P(B)x) = ||P(B)x||^2$ est une mesure positive sur A de masse totale $||x||^2$. Soit \mathcal{K} la famille des parties N de A telles que $\exists B \in \mathcal{Q} : P(B) = 0$ et $N \subset B$. \mathcal{K} est une classe héréditaire ($N_1 \subset N_2$ et $N_2 \in \mathcal{K} \Rightarrow N_1 \in \mathcal{K}$) qu'on peut supposer contenue dans a (complétion de la tribu). On a alors :

$$N \in \mathcal{K} \Leftrightarrow P(N) = 0 \Leftrightarrow \forall x \in H \quad m_x(N) = 0$$

Une fonction complexe sur A est P -essentiellement bornée si $\exists r \geq 0 : \{|f| \leq r\} \in \mathcal{K}$. La borne inférieure de ces r est notée $\text{sup.ess } |f|$ ou $||f||_\infty$. Si on identifie deux fonctions essentiellement bornées ne différant que sur un ensemble de \mathcal{K} on construit une algèbre de Banach de (classe de) fonctions pour les opérations habituelles d'addition et de multiplication des fonctions qu'on note $L^\infty(P)$.

Si $f \in L^\infty(P)$ les quantités $\int_A f(u) m_x(du)$ ne dépendent que de la classe de f . A f fixée, on obtient une forme quadratique de x ; il existe donc un opérateur borné qu'on notera $P(f)$ ou encore $\int_A f(u) P(du)$ tel que :

$$\forall x \in H \quad \int_A f(u) m_x(du) = (x|P(f)x)$$

Si f est en escalier, c'est-à-dire de la forme $\sum_{i=1}^n f_i 1_{B_i}$ où $B_1 \dots B_n$ est une partition de A il est immédiat que l'on a :

$$\int_A f(u) P(du) = \sum_{i=1}^n f_i P(B_i)$$

Les propriétés suivantes, trivialement vérifiées pour les fonctions en escaliers, restent vraies par passage à la limite dans tous les cas :

$$\cdot \left\| \int f dP \right\| = \|f\|_{\infty}$$

• Linéarité ; involutivité : $\int \bar{f} dP = \left(\int f dP \right)^*$; en particulier si f est réelle (resp. positive), $P(f)$ est hermitien (resp. positif). Autrement dit si $f \geq g$, $\int f dP \geq \int g dP$ pour l'ordre des opérateurs hermitiens.

$$\begin{aligned} \cdot \int f(u) P(du) \int g(v) P(dv) &= \int f(u) g(u) P(du) P(dv) \\ &= \int f(u) g(u) P(du) \end{aligned}$$

ce qui s'écrit aussi $P(f) P(g) = P(fg)$

On a donc un morphisme d'algèbre de Banach de $L^{\infty}(P)$ dans $L(H)$.

On a enfin que :

$$S_P(P(f)) = \text{Im ess } f = \cap \{K ; K \text{ compact et } (\text{Im } f \cap (A-K)) \in \mathcal{A}\}.$$

Dans le cas particulier où A est topologique compact, f continue, et où $P(B) \neq 0$ pour tout B ouvert on a tout bonnement :

$$S_P(P(f)) = \text{Im } f = \{f(u) ; u \in A\}.$$

On montre, enfin, que tout opérateur normal Q définit de façon unique une mesure spectrale P sur son spectre tel que l'on ait :

$$Q = \int \lambda dP(\lambda)$$

Lorsque Q est compact le spectre est discret formant une suite convergent vers zéro. Chaque valeur spectrale non nulle est valeur propre pour un sous-espace de dimension finie. En choisissant une base orthonormée dans chacun d'eux et en remarquant qu'un projecteur de rang 1 s'écrit $x \otimes x$ pour un x unitaire, on obtient la représentation de $Q = \sum_i \lambda_i x_i \otimes x_i$ déjà donné p. 4 (1-9).

2.c - Représentation spectrale

On s'intéresse maintenant à la fonction d'ensemble définie sur les parties mesurables de \hat{T} à valeurs dans $H(X)$ définie par : (X_0 est la valeur pour $t=0$ au processus).

$$Z(A) = P(A) X_0 \tag{2-8}$$

Cette fonction est une mesure stochastique orthogonale (m.s.o. [7]) car elle vérifie :

$$\left. \begin{array}{l} \cdot Z(A) \text{ centrée } E|Z(A)|^2 = ||P(A) X_0||^2 = m(A) < \infty \\ \cdot A \cap B = \emptyset \quad E Z(A) \overline{Z(B)} = 0 \\ \cdot A_n \text{ disjoints } \quad Z(\bigcup_n A_n) = \sum_n Z(A_n) \text{ au sens de la} \end{array} \right\} (2-9)$$

convergence en moyenne quadratique.

La fonction d'ensemble m est une mesure de probabilité sur \hat{T} dite mesure structurelle. Si m est absolument continue par rapport à $d\omega$ on a :

$$m(d\omega) = f(\omega) d\omega \quad (2.10)$$

et on dit que f est la densité spectrale du processus X_t . Si f admet une version continue on la privilégie systématiquement dans l'écriture précédente. On montre alors que la fonction de covariance est la transformée de Fourier de la densité spectrale soit :

$$\phi(h) = \int_{\hat{T}} e^{-i(\omega, h)} dm(\omega) \quad (2.11)$$

et on obtient la représentation spectrale du processus par l'intégrale stochastique :

$$X_t = \int_{\hat{T}} e^{i(\omega, t)} dZ(\omega) \quad (2.12)$$

Cette formule ne fait que traduire l'isométrie entre les espaces $L^2(\hat{T}, m)$ et $H(X)$ associant les fonctions $\omega \mapsto e^{i(\omega, t)}$ aux X_t et la multiplication par $e^{i(\omega, h)}$ à l'opérateur unitaire U_h .

2,d - Exemples courants

- Si $T = \mathbb{R}$, $\hat{T} = \mathbb{R}$ et $e^{i(\omega, t)} = e^{i\omega t}$.
- Si $T = \mathbb{Z}$, $\hat{T} = \mathbb{T}$ et $e^{i(\omega, t)} = e^{i\omega t}$ ou $\omega \in \mathbb{R} | 2\pi$ et t entier ; pratiquement on prend ω dans $[0, 2\pi[$ ou dans $[-\pi, +\pi[$.
- Si $T = \mathbb{Z}_p$ groupe des entiers modulo p , $\hat{T} = \mathbb{Z}_p$
et $e^{i(\omega, t)} = e^{i \frac{2\pi}{p} \omega t}$

• Pour tout produit $T = T_1 \times T_2$ on a $\hat{T} = \hat{T}_1 \times \hat{T}_2$, ce qui permet de traiter tous les cas courants de groupes abéliens.

2.e - Commentaire

Dès que l'ensemble des temps est un groupe, on dispose de "beaucoup de structure" sur les données. Si de plus on suppose le processus stationnaire la richesse de la structure devient digne de Crésus et la description des données peut devenir, en principe, très fine. Une fois calculée $f(\omega)$, l'interprétation des données se fait en terme de cycles, et, dans la littérature des "électriciens" la densité spectrale représente une énergie transportée par bande de fréquence.

L'ennui, sur le plan statistique, c'est qu'en général on suppose que $T = \mathbb{R}$ ou \mathbb{Z} alors qu'on ne dispose que d'observation partout sur une durée limitée. On s'en tire généralement en disant que la série doit être très longue ... et qu'on lui fait subir une transformation (ad hoc) pour la rendre stationnaire. Assez souvent aussi, lorsqu'on dispose de plusieurs échantillons de trajectoires de X_t et que la stationnarité est de plus en plus difficile à assumer on utilise des méthodes (du genre cospectral) qui "permettent" de se ramener à une seule série vectorielle (ah mais !) sur laquelle on puisse en toute impunité faire une hypothèse de stationnarité.

L'étude des processus stationnaires est néanmoins utile pour plusieurs raisons : techniques d'abord car l'analyse "théorique" peut être poussée très loin ; pratiques aussi car les processus stationnaires apparaissent asymptotiquement dans les phénomènes non explosifs régis par des équations différentielles ou des équations récurrentes du genre ARMA. Si ces équations varient un peu au cours du temps on obtient une série qui évolue vers un processus stationnaire variable au cours du temps et qui apparaît donc comme une sorte de "mélange" de processus stationnaires.

D'où l'idée de s'intéresser à des processus relativement "voisins" de processus stationnaires admettant une représentation "canonique" qui se ramène à la décomposition spectrale dans le cas stationnaire. Il faudrait, de plus, pouvoir faire la statistique de cette représentation, ce qui est une autre paire de manche !

(Certaines idées dans cette optique peuvent se trouver dans [8],[9] et [9 bis], ou [10]).

3 - COMPARAISON ENTRE LES DEUX METHODES : GENERALITES

Pour comparer les deux techniques qui viennent d'être décrites on peut procéder soit par "réduction" soit par "extension".

La "réduction" consiste à étudier les hypothèses (ou structures) permettant de les appliquer toutes les deux. On doit donc supposer que T est un groupe, X_t un processus stationnaire, et la mesure de Haar sur T . La variance totale doit en outre être finie et donc la mesure totale du groupe. Cette condition est suffisante (et nécessaire) pour que T soit compact.

On étudiera donc deux archétypes qui, d'ailleurs, [11] permettent de construire tous les groupes abéliens compacts :

- . $T = \mathbb{Z}_p$ (ce cas permet de traiter tous les groupes finis).
- . $T = \mathbb{T}$

L'extension consiste à trouver une généralisation commune des deux formes de représentation. En fait on trouvera une extension de l'analyse factorielle pour certains processus en temps discret de variance totale infinie comprenant une classe assez large de processus stationnaires. La représentation factorielle, se déduira, pour ces derniers, de la représentation spectrale.

On pourrait, très vraisemblablement, arriver à une extension presque universelle, englobant tous les processus stationnaires. Ce serait en des termes mathématiques assez abstraits et vraisemblablement inaccessibles à la statistique numérique. Nous n'en parlerons pas ici car la mathématique statistique est, à notre avis, un discours sur le calculable.

4 - PROCESSUS DU SECOND ORDRE SUR \mathbb{Z}_p .

Sans précisions supplémentaires, il s'agit d'une famille de p variables du second ordre centrées, la mesure sur \mathbb{Z}_p est uniforme en masse totale égale à 1. Sur $\hat{\mathbb{Z}}_p$ ce sera la mesure de comptage, de masse totale égale à p . L'analyse harmonique d'un tel processus n'est autre, évidemment, que l'A.C.P. des p variables en question.

Précisons un peu le cas où X_t est stationnaire. Soit $U = U_1$ l'opérateur de $H(X)$ qui envoie X_t sur X_{t+1} . On a alors $U_h = U^h$ ($h=0,1,\dots,p-1$) et $U^p = 1$.

Introduisons dans $L^2(\mathbb{Z}_p)$ l'opérateur S de décalage d'indice qui associe à $(x_0, x_1, \dots, x_{p-1}) = x$ la suite $(x_1, x_2, \dots, x_{p-1}, x_0) = Sx$. On calcule alors l'opérateur de covariance :

$$(Cx)_t = \frac{1}{p} \sum_s C(t,s) x_s = \frac{1}{p} \sum_s \rho(t-s) x_s = \frac{1}{p} \sum_h \rho(h) x_{t-h} = \frac{1}{p} \sum_h \rho(h) (S^{-h}x)_t$$

soit
$$C = \frac{1}{p} \sum_h \overline{\rho(h)} S^h \tag{4-1}$$

Le même genre de calcul montre que l'opérateur d'Escoufier s'écrit :

$$V = \frac{1}{p} \sum_h \overline{\rho(A)} U^h \tag{4-2}$$

L'opérateur S admet les vecteurs propres unitaires :

$$f_\omega = \{e^{i \frac{2\pi}{p} \omega t} ; t = 0, 1, \dots, p-1\} \text{ pour } \omega = 0, 1, \dots, p-1 \tag{4-3}$$

associés aux valeurs propres simples $\lambda_\omega = e^{i\omega \frac{2\pi}{p}}$.

De ce fait C admet les mêmes vecteurs propres pour les valeurs propres

$$\phi(\omega) = \frac{1}{p} \sum_h \rho(h) e^{-i\omega h \frac{2\pi}{p}} \tag{4-4}$$

Les composantes de l'A.C.P. sont faciles à déterminer. Posons :

$$z_\omega = \frac{1}{p} \sum_t e^{i \frac{2\pi}{p} \omega t} X_t. \text{ On a } \text{Var}(z_\omega) = \phi(\omega) \text{ de sorte qu'en posant}$$

$\xi_\omega = (\phi(\omega))^{-\frac{1}{2}} z_\omega$ on obtient la représentation factorielle (à l'ordre près des termes) :

$$X_t = \sum_\omega \sqrt{\phi(\omega)} f_\omega \xi_\omega = \sum_\omega e^{i\omega t \frac{2\pi}{p}} z_\omega \tag{4-5}$$

Ceci n'est pas autre chose que la représentation spectrale du processus, et la fonction $\omega \mapsto \phi(\omega)$ se révèle être la densité spectrale. Par ailleurs l'opérateur U admet les λ_ω comme vecteurs propres unitaires associés aux λ_ω ; V admet ces mêmes vecteurs propres associés au $\phi(\omega)$.

Cette identification présente néanmoins une difficulté. Dès que X_t est réel - par exemple ! - la fonction ϕ est symétrique de sorte que, à l'exception de $\phi(0)$ et $\phi(p/2)$ si p est pair, toutes les valeurs propres de C et V sont doubles. La formation de C (resp. V) comme polynôme sur S (resp. U) a transporté de spectre de S (les racines $p^{\text{ième}}$ de 1) sur \mathbb{R}^+ et l'a rendu dégénéré.

Remarque 1 :

Si une famille de variables présente des covariances "voisines" de celles d'un processus stationnaire sur \mathbb{Z}_p , alors son analyse factorielle ressemblera à une décomposition spectrale, ce qui peut aider l'interprétation.

Remarque 2 :

Comme tout groupe fini et un produit [11] de groupes isomorphes à des \mathbb{Z}_p on sait faire l'analyse factorielle ou spectrale sur tout groupe fini.

Remarque 3 :

Les figures qu'on obtient sur des graphes-plans avec ce type de décomposition sont très simples. On pourrait songer à les utiliser, dans une A.C.P. ordinaire, pour distinguer des familles de variables ayant des structures particulières représentables comme des processus stationnaires sur des groupes finis et obéissant donc à des règles de transformation simples.

5 - PROCESSUS SUR LE CERCLE TRIGONOMETRIQUE \mathbb{T}

Commençons par la représentation spectrale. Comme $\hat{\mathbb{T}} = \mathbb{Z}$, celle-ci prend l'aspect d'une série convergente uniformément en moyenne quadratique :

$$X_t = \sum_{n=-\infty}^{+\infty} e^{int} z_n \quad (5-1)$$

Posons $\phi(n) = \text{Var } z_n$ (densité spectrale) on a aussi

$$X_t = \sum_{-\infty}^{+\infty} \phi^{1/2}(n) e^{int} \xi_n \quad (5-2)$$

avec ξ_n système orthonormé base hilbertienne de $H(X)$. Comme les exponentielles sont un système orthonormé de $L^2(\mathbb{T}, dt)$ cette expression ressemble bigrement à celle de l'analyse factorielle des processus.

En fait c'est bien cela que l'on retrouve. On peut, pour s'en convaincre, procéder à peu près comme au numéro précédent. Pour tout h de \mathbb{T} et toute fonction sur \mathbb{T} , f , posons $S_h f(t) = f(t+h)$. Les S_h forment un groupe d'opérateurs unitaires sur $L^2(\mathbb{T}, dt)$ - (rappelons que dt est la mesure de Lebesgue normalisée), et on a pour l'opérateur de covariance (défini sans aucun problème

puisque Π est compact et X_t continu en moyenne quadratique) :

$$\begin{aligned} Cf(t) &= \int C(t,s) f(s) ds = \int \rho(t-s) f(s) ds = \int \rho(-h) f(t+h) dh \\ &= \int \overline{\rho(h)} S_h f(t) dh \end{aligned}$$

et donc :

$$C = \int \overline{\rho(h)} S_h dh \quad (5-3)$$

Le même genre de calcul montre que l'opérateur d'Escoufier s'écrit :

$$V = \int \overline{\rho(h)} U_h dh \quad (5-4)$$

Introduisons sur $L^2(\Pi)$ l'opérateur autoadjoint non borné :

$$A = -i D \quad \text{où } D \text{ est la dérivation.}$$

$$\text{On a :} \quad S_h = \exp(-ihA) \quad (5-5)$$

Ce résultat classique exprime le développement en série de Taylor valable au sens le plus étroit sur l'espace (partout dense) des fonctions analytique sur Π . Le spectre de A est \mathbb{Z} tout entier, les fonctions propres étant les exponentielles $t \mapsto e^{int}$ ($Ae^{int} = ne^{int}$). De ce fait les S_h admettent les mêmes fonctions propres associées aux valeurs propres e^{-inh} . Si h n'est pas commensurable avec 2π cette suite est dense dans Π qui est donc le spectre de S_h ; dans le cas contraire S_h n'a qu'un nombre fini de valeurs spectrales, chacune d'elle étant associée à un sous espace propre de dimension infinie.

L'opérateur compact C admet toujours les mêmes fonctions propres associées cette fois-ci d'après (5-3) et (5-5) aux valeurs propres :

$$\lambda_n = \int \rho(h) e^{inh} dh = \phi(n) \quad (5-6)$$

Les composantes s'obtiennent alors par :

$$\xi_n = \lambda_n^{-\frac{1}{2}} \int_{\Pi} X_t e^{-int} dt \quad (5-7)$$

L'utilisation de la représentation spectrale (5-1) montre que les ξ_n intervenant maintenant sont bien les mêmes que ceux qui interviennent dans la représentation spectrale. Du coup les opérateurs U_h et V admettent les ξ_n comme famille propre orthonormée, les valeurs propres associées étant respectivement e^{-inh} et $\lambda_n = \phi(n)$.

On retrouve cependant la même difficulté qu'au numéro précédent : la fonction ϕ est symétrique dès que X_t est réel, de sorte qu'à l'exception de $\phi(0)$ toutes les valeurs propres de C et V sont doubles. La formation de C (resp. V) comme intégrale sur S_h (resp. U_h) a dégénéré le spectre.

Si on veut représenter les choses en termes réels, on prendra plutôt la bse formée par les $\sqrt{2} \sin(nt)$ et $\sqrt{2} \cos(nt)$, $n=0,1,\dots$. Les composantes ξ_n deviendront alors réelles alors que dans l'écriture (5-2). On a :

$$\xi_{-n} = \bar{\xi}_n \text{ (et aussi } E \xi_n^2 = 0 \text{ !)}.$$

Remarque 1 :

Ce qui vient d'être dit reste vrai pour tout groupe compact à quelques détails près. On sait, par ailleurs [11], que ce type de groupe est généré par des produits de groupes isomorphes à \mathbb{R} ou \mathbb{Z}_p .

Remarque 2 :

On a rarement, dans la pratique statistique, formalisé des données à l'aide d'une structure de groupe compact. Cela devrait pourtant se faire naturellement pour des données du style tour d'horizon d'une sonde spatiale, direction des vents, saisonnalité de certains phénomènes.

6 - PROCESSUS EN TEMPS DISCRET : HYPOTHESE ET GENERALITES

Jusqu'à la fin nous admettrons l'hypothèse suivante relativement à la fonction de covariance :

Hypothèse : Il existe une fonction positive $r(h)$ sur \mathbb{Z} , vérifiant $\sum_h r(h) = a < \infty$ telle que la covariance du processus $(X_t)_{t \in \mathbb{Z}}$ vérifie :

$$\forall t, s \in \mathbb{Z} \quad |C(t,s)| \leq r(t-s) \quad (6-1)$$

Cette hypothèse n'a aucun caractère minimal. Elle permet d'avoir les résultats auxiliaires qui vont suivre et de ne travailler qu'avec des opérateurs bornés de l^2 (ensemble des suites de carré sommable) ou de $H(X)$. Sur le plan probabiliste, cette condition exprime une sorte de non-corrélation asymptotique, de mélange relativement fort, excluant des trajectoires trop "constantes". Elle implique aussi que $\text{Var}(X_t)$ reste borné : on exclut tout processus explosif. Ce genre d'hypothèse est souvent postulée dans les questions relatives à l'estimation du spectre d'un processus stationnaire [12].

Conséquence 1 :

Si $x = (x_t) \in \ell^2$ alors $\sum x_t X_t$ converge en moyenne quadratique et est dans $H(X)$. L'opérateur $A : x \mapsto \xi = \sum x_t X_t$ est borné de norme inférieure à $a^{1/2}$.

Soit ℓ_0 la partie partout dense de ℓ^2 constituée par les suites nulles en dehors d'une partie finie de \mathbb{Z} . Alors $Ax = \sum_t x_t X_t$ est définie de ℓ_0 dans H . Pour tout x de ℓ_0 on a :

$$||Ax||^2 = \text{Var} (\sum x_t X_t) = \sum_{t,s} x_t \bar{x}_s C(t,s) \leq \sum_{t,s} |x_t| |x_s| r(t-s)$$

D'où :

$$||Ax||^2 \leq \sum_h r(h) \sum_t |x_t| |x_{t-h}| \leq a \sum_t |x_t|^2$$

L'opérateur A se prolonge donc par continuité à ℓ^2 tout entier ce qui établit le résultat.

Conséquence 2 :

Sur ℓ^2 et L^2 respectivement, les formules suivantes définissent des opérateurs hermitiens positifs de norme inférieure à a :

$$\cdot (Cx)_t = \sum_s C(t,s) x_s \tag{6-2}$$

$$\cdot V\xi = \sum_s E(\xi \bar{X}_s) X_s \tag{6-3}$$

L'opérateur A a pour adjoint $J : L^2 \rightarrow \ell^2$ défini par $(J\xi)_t = E\xi \bar{X}_t$. Dès lors on a (cf numéro 1 p.4) $C = A^*A$ et $V = AA^*$ d'où le résultat.

Introduisons l'espace (de Schwartz) $s \subset \ell^2$ des suites à décroissance rapide. Plus précisément :

$$s = \{(x_t)_{t \in \mathbb{Z}} ; \forall p \text{ entier } \lim_{t \rightarrow +\infty} x_t |t|^p = 0\}$$

Conséquence 3 :

L'opérateur C est stable sur s dès que $(r(h))_{h \in \mathbb{Z}} \in s$.

Soit $y_t = \sum_s C(t,s) x_s$. On a toujours $|y_t| \leq \sum_h r(h) |x_{t-h}|$. La suite $|y_t|$ est le produit de convolutions de deux suites de s et est donc dans s .

Remarque :

L'hypothèse qui vient d'être introduite sur les $r(h)$ est assez restrictive. Cependant tous les processus ARMA la satisfont de même que certains schémas autorégressifs variables avec t . C'est le cas par exemple pour un processus vérifiant $x_{t+1} = a_t x_t + \varepsilon_t$, où ε_t est un bruit blanc et où on a $|a_t| \leq h < 1$ pour tout t .

Elle a été formulée pour pouvoir appliquer la théorie de Gelfand-Chilov-Vilenkin [13][14]. On peut s'en affranchir avec un appareillage mathématique un peu plus abstrait dont on trouvera l'ébauche au n° 11.

7 - PROCESSUS STATIONNAIRES EN TEMPS DISCRET

En procédant comme au 4, avec S opérateur de translation à indice de ℓ^2 ($(Sx)_t = x_{t+1}$), on obtient immédiatement :

$$. C = \sum_{h \in \mathbb{Z}} \overline{\rho(h)} S^h \quad (7-1)$$

$$. V = \sum_{h \in \mathbb{Z}} \overline{\rho(h)} V^h \quad (7-2)$$

La convergence de ces séries est assurée par l'hypothèse de sommabilité de $\rho(h)$. La série $\sum_h \rho(h) e^{-i\omega h}$ est uniformément convergente et définit une fonction continue $\phi(\omega)$ qui est la densité spectrale du processus. Si $\{\rho(h)\} \in s$, cette densité est même indéfiniment dérivable.

Le spectre de S est Π , à chaque valeur propre $e^{i\omega}$ étant associé le "pseudo-vecteur propre" $u_\omega = \{e^{i\omega t}\}_{t \in \mathbb{Z}}$. De même u_ω est "pseudo vecteur propre" de C en ce sens que $(Cu_\omega)_t = \phi(\omega) e^{i\omega t}$. Mais les u_ω ne sont pas dans ℓ^2 . Le spectre de C (ou de V) est, d'après (7-1) l'ensemble des valeurs prises par ϕ quand ω varie. Dès que X_t est réel chaque valeur spectrale de C (sauf peut-être un nombre fini correspondant aux extrêmes de ϕ) est au moins double.

Ceci posé, on a du mal à définir des pseudo-vecteurs propres de V . Il faudrait s'intéresser à un être "indéfinissable" comme $\sum e^{-i\omega t} X_t = z_\omega$. En fait la représentation spectrale $X_t = \int_{[-\Pi, +\Pi[} e^{i\omega t} dZ(\omega)$ donne une piste.

La mesure stochastique $Z(d\omega)$ peut s'écrire $\sqrt{\phi(\omega)} B(d\omega)$, où B est une mesure blanche. Si on se permet d'écrire celle-ci $\xi_\omega d\omega$, on a

$$X_t = \int_{[-\pi, \pi[} \sqrt{\phi(\omega)} u_\omega(t) \xi_\omega d\omega, \text{ ce qui ressemble beaucoup à l'équation de l'analyse factorielle des processus. Reste à définir une telle écriture.}$$

8 - ANALYSE HARMONIQUE DE PROCESSUS EN TEMPS DISCRET

L'espace s du numéro 6 est classiquement muni de la topologie suivante, définie la famille de normes :

$$p \text{ entier} : N_p(x) = \sum_t (1+|t|)^p |x_t|^2 \tag{8-1}$$

Le dual s' de s s'identifie à l'ensemble des suites vérifiant :

$$s' = \{x_t ; \exists p \text{ entier} > 0, K > 0 \text{ et } M : |t| > \tilde{M} \quad |x_t| \leq K|t|^p\}$$

Il est clair que :

$$s \subset \ell^2 \subset s' \tag{8-2}$$

avec s dense dans ℓ^2 , ℓ^2 dense dans s' . On obtient ainsi un "sandwich au Hilbert" ⁽¹⁾ ; si la majorante $r(h)$ est dans s , C laisse s stable de sorte qu'on peut appliquer la théorie de Gelfand-Chilov-Vilenkin ([13] et [14]).

Sur $K = Sp(C) \subset \mathbb{R}^+$, on peut donc trouver une mesure finie σ et des familles $u_\lambda^n \in s'$ ($\lambda \in K, n = 1, 2, \dots$) telles que l'on ait :

$$C = \sum_n \int_K \lambda u_\lambda^n \otimes u_\lambda^n d\sigma(\lambda) \tag{8-3}$$

en ce sens que :

$$\forall v, w \in s : (v|Cw)_{\ell^2} = \sum_n \int_K \lambda \langle u_\lambda^n, v \rangle \overline{\langle u_\lambda^n, w \rangle} d\sigma^2(\lambda) \tag{8-4}$$

(Les crochets $\langle \rangle$ notent la dualité entre s' et s).

Les u_λ^n sont des vecteurs propres généralisés de C en ce sens que :

$$\forall v \in s \quad \langle u_\lambda^n, Cv \rangle = \lambda \langle u_\lambda^n, v \rangle = \lambda \sum_t u_\lambda^n(t) v_t \tag{8-5}$$

(1) Un sondage non représentatif indique que cette traduction serait plus satisfaisante, pour le terme "Rigged Hilbert Space" que la locution "triade hilbertienne" qui figure dans la traduction officielle.

Elles forment un système complet dans le sandwich en ce sens que :

$$\forall v, w \in s \quad (v|w)_{\ell^2} = \sum_n \int_K \langle u_\lambda^n, v \rangle \overline{\langle u_\lambda^n, w \rangle} d\sigma(\lambda) \quad (8-6)$$

Appelons "petite variable" toute $\xi = \sum_t v_t X_t$ avec $v = \{v_t\} \in s$.

L'ensemble H_0 des petites variables est un espace partout dense dans H . Soit z_λ^n l'être définit formellement par $\sum_t u_\lambda^n(t) X_t$, et rigoureusement comme forme linéaire sur H_0 , continue si on t munie H_0 d'une topologie déduite de celle de s qui le rend complet et même nucléaire :

$$\forall \xi \in H_0 \quad \langle z_\lambda^n, \xi \rangle_H = E(\sum_t u_\lambda^n(t) X_t) (\sum_s v_s X_s) = \langle u_\lambda^n, Cv \rangle_{\ell^2} = \lambda \langle u_\lambda^n, v \rangle \quad (8-7)$$

Posons, pour $\lambda \neq 0$, $\xi_\lambda^n = \lambda^{-1/2} z_\lambda^n$.

Nous pouvons maintenant montrer que :

V laisse stable l'espace des petites variables. Les ξ_λ^n forment un système complet dans H de pseudo vecteurs propres de V qui admet donc la représentation :

$$V = \sum_n \int_K \lambda \xi_\lambda^n \otimes \xi_\lambda^n d\sigma(\lambda) \quad (8-8)$$

Si $v \in s$ et $\xi = \sum_t v_t X_t$, on a $V\xi = \sum_u E(\xi \bar{X}_u) X_u = \sum_u X_u (\sum_t C(u,t) v_t)$

Comme C laisse s stable, V laisse H_0 stable. De plus si $w = Cv$, on a :

$$\langle \xi_\lambda^n, V\xi \rangle_H = \lambda^{-1/2} \langle u_\lambda^n, Cw \rangle_{\ell^2} = \lambda \lambda^{-1/2} \langle u_\lambda^n, Cv \rangle_{\ell^2} = \lambda \langle \xi_\lambda^n, v \rangle_H$$

Enfin, pour tout couple $\xi = \sum_t v_t X_t$ et $\eta = \sum_s w_s X_s$ de H_0 on a :

$$\begin{aligned} E\xi\bar{\eta} &= \sum_{t,s} v_t \bar{w}_s C(t,s) = \sum_n \int_K \lambda \langle u_\lambda^n, v \rangle \overline{\langle u_\lambda^n, w \rangle} d\sigma(\lambda) \\ &= \sum_n \int_K \lambda^{1/2} \langle \xi_\lambda^n, \xi \rangle_H \overline{\langle u_\lambda^n, w \rangle} d\sigma(\lambda) \\ &= \sum_n \int_K \langle \xi_\lambda^n, \xi \rangle \overline{\langle \xi_\lambda^n, \eta \rangle} d\sigma(\lambda) \end{aligned} \quad (8-9)$$

La représentation de V résulte d'un calcul immédiat analogue de $(\xi|V\eta)$.

Comme X_t est une petite variable, associée à $w_t = 1$ $w_s = 0$, on obtient en remplaçant η par X_t dans la formule du milieu : (8-9) $s \neq t$

$$\forall \xi \in H_0 \quad E \xi \bar{X}_t = \sum_n \int_K \lambda^{1/2} \overline{u_\lambda^n(t) \langle \xi_\lambda^n, \xi \rangle} d\sigma(\lambda) \quad (8-10)$$

Ce qu'on peut aussi écrire :

$$X_t = \sum_n \int_K \lambda^{1/2} u_\lambda^n(t) \xi_\lambda^n d\sigma(\lambda) \quad (8-11)$$

ce qui est l'exacte généralisation de l'analyse harmonique d'un processus à espace des temps compacts.

L'intérêt statistique de ces représentations provient de deux faits :

1 - L'espace H_0 contient toutes les combinaisons linéaires finies formées sur les X_t , c'est-à-dire la seule catégorie "accessible" de variable.

2 - Les u_λ^n sont des suites (à "croissance modérée") qu'on pourrait, en principe, avoir l'ambition de calculer, au moins de façon approximative. On peut montrer, en particulier et de plus, que pour tout t $u_\lambda^n(t)$ peut être une fonction continue de λ sur $Sp C = Sp V$.

9 - RETOUR AUX PROCESSUS STATIONNAIRES

Les pseudo-vecteurs propres de C et des S^h , $u_\omega = \{e^{i\omega t}\}$ sont dans s' et forment un système complet (cf. théorie des séries de Fourier). Comme dans ce qui précède, on définit des formes linéaires (continues) sur l'espace des petites variables répondant formellement à la définition :

$$\langle \xi_\omega, \xi \rangle_H = \phi(\omega)^{-1/2} \overline{E \xi (\sum_t e^{i\omega t} X_t)}$$

La légitimité des représentations suivantes ne demande qu'une vérification triviale :

$$S = \int_{]-\pi, +\pi]} e^{i\omega} u_\omega \otimes u_\omega d\omega \quad (9-1)$$

$$U = \int_{]-\pi, +\pi]} e^{i\omega} \xi_\omega \otimes \xi_\omega d\omega \quad (9-2)$$

De même que :

$$C = \int_{]-\pi, +\pi]} \phi(\omega) u_\omega \otimes u_\omega d\omega \quad (9-3)$$

$$V = \int_{]-\pi, +\pi]} \phi(\omega) \xi_\omega \otimes \xi_\omega d\omega \quad (9-4)$$

Le fait que les ξ_ω forment aussi un système complet dans H permet, comme au paragraphe précédent, d'arriver à la représentation suivante du processus, toute proche de la représentation spectrale et réalisant le but de la fin du numéro 7 :

$$X_t = \int_{]-\pi, +\pi]} \sqrt{\phi(\omega)} e^{i\omega t} \xi_\omega d\omega \quad (9-5)$$

en ce sens que : $\forall \xi$ petite variable :

$$E(\xi \bar{X}_t) = \int_{]-\pi, +\pi]} \sqrt{\phi(\omega)} e^{-i\omega t} \langle \xi_\omega, \xi \rangle_H d\omega \quad (9-6)$$

La mesure spectrale $P(d\omega)$ associée à V peut s'écrire $P(d\omega) = \xi_\omega \otimes \xi_\omega d\omega$ en ce sens que si f est essentiellement P- bornée et que $F = \int f(\omega) dP(\omega)$, on aura pour tout couple ξ, η de petites variables :

$$(\xi | F\eta) = \int_{]-\pi, +\pi]} \langle \xi_\omega, \xi \rangle_H f(\omega) \langle \xi_\omega, \eta \rangle_H d\omega \quad (9-7)$$

10 - COMPARAISONS ENTRE LES REPRESENTATIONS DES DEUX NUMEROS PRECEDENTS

Obtenues par des moyens analogues, ces deux représentations se relient par la notion de transport d'une mesure spectrale.

Si $P(d\omega)$ est une mesure sur \hat{T} dans un espace de Hilbert H, que ϕ est une application mesurable dans un espace (B, \mathcal{B}) on vérifie immédiatement qu'on définit sur B une mesure spectrale transportée par ϕ en posant :

$$\forall D \in \mathcal{B} \quad Q(D) = P(\phi^{-1}(D)) \quad (10-1)$$

Si ϕ est à valeurs complexes et est P_essentielle-ment bornée, on a :

$$C = \int_{\hat{T}} \phi(\omega) P(d\omega) = \int_{SpC} \lambda Q(d\lambda) \quad (10-2)$$

On peut le vérifier de la façon suivante : soit m_x la mesure sur \hat{T} définie par $m_x(D) = ||P(D)x||^2$ et m_x^ϕ sa transportée par ϕ . On a :

pour tout $x \in H$

$$(x|Cx) = \int_{\hat{T}} \phi(\omega) m_x(d\omega) = \int_{SpC} \lambda m_x^\phi(d\lambda) \text{ par définition du transport d'une mesure.}$$

$$\text{Or } m_x^\phi(D) = ||P(\phi^{-1}(D))x||^2 = ||Q(D)x||^2.$$

Ainsi, dans le cas d'un processus stationnaire, l'opérateur C (resp V) a une représentation spectrale transportée de celle de S (resp \hat{V}) par la densité spectrale.

On dit que P est de dimension 1 s'il existe un vecteur x tel que H soit engendré par les $P(A)x$ $A \subset \hat{T}$ mesurable ; $Z_x(A) = P(A)x$ est alors une mesure stochastique orthogonale dont la mesure structurelle est m_x .

Toute mesure spectrale dans un Hilbert séparable peut se décomposer en une somme au plus dénombrable de mesures spectrales de dimension 1. Le procédé de construction est le suivant : si les $P(A)x_1$ n'engendrent qu'un sous espace strict H_1 de H, on prend $x_2 \in H_1^\perp$ et les $P(A)x_2$ engendrent un sous espace orthogonal à H_1 , et on continue l'opération jusqu'à reconstituer H. C'est cette construction qui, fondamentalement, permet d'obtenir le résultat de [14] utilisé au numéro 8. Si \hat{T} était discret, elle reviendrait à choisir, dans chaque sous espace $P(\omega)$, une base orthonormée dont les éléments seraient numérotés de 1 à $\dim P(\omega)$. La représentation factorielle des processus, obtenue en 8 présente donc un caractère arbitraire qui est exactement de cette nature.

Si maintenant ϕ est injective et P de dimension 1, la transportée de P par ϕ l'est également. Or la mesure spectrale de 0 est évidemment de dimension 1 (par construction même) ; supposons que la densité spectrale continue (et même éventuellement C^∞) ϕ soit suffisamment régulière pour vérifier la propriété suivante :

$$\begin{aligned} \text{il existe } a_0 = -\Pi, a_1, \dots, a_n = \Pi \text{ tels que } \phi \text{ soit monotone} \\ \text{sur } D_i =]a_{i-1}, a_i] , i = 1, \dots, n. \end{aligned}$$

Chacune des mesures spectrales $P_i(d\omega) = 1_{D_i}(\omega) P(d\omega)$ se transporte par ϕ en une mesure Q_i à support contenu dans le spectre de C et de dimension 1. On a alors :

$$C = \sum_{j=1}^n \int_{\text{Sp}C} \lambda Q_j(d\lambda) \quad (10-3)$$

La représentation de V se transporte de la même manière ainsi que celle du processus qui s'explicité en :

$$X_t = \sum_{j=1}^n \int_{D_j} \sqrt{\phi(\omega)} e^{i\omega t} \xi_j d\omega \quad (10-4)$$

équivalent à :

$$X_t = \sum_{j=1}^n \int_{\phi(D_j)} \lambda^{1/2} e^{i\phi_j^{-1}(\lambda)t} \xi_{\phi_j^{-1}(\lambda)} d\sigma_j(\lambda) \quad (10-5)$$

où ϕ_j^{-1} est la fonction inverse de ϕ restreinte à D_j et σ_j la mesure $d\omega$ restreinte à D_j transportée par ϕ . Comme les σ_j sont absolument continue par rapport à une même mesure (par exemple $\Sigma \sigma_j$, mais si ϕ est C^∞ par rapport aussi à la mesure de Lebesgue sur $\text{Sp}C$) on retrouve exactement la formulation établie au numéro 8. Les vecteurs générateurs des mesures spectrales Q_j sont, naturellement, les $P(D_j) X_0$.

L'analyse harmonique "généralisée" du numéro 8 est, visiblement, une extension de l'analyse harmonique "compacte" habituelle. Elle redonne (ou peut redonner) la représentation spectrale des processus stationnaires moyennant une "bonne" décomposition de la mesure spectrale de l'opérateur de covariance en mesures spectrales de dimension 1. Elle peut donc apparaître comme une synthèse des deux techniques.

11 - EN ROUTE VERS D'AUTRES AVENTURES

Première aventure : La théorie qui vient d'être exposée cherche à utiliser la théorie des sandwich hilbertiens de Gelfand pour aboutir à une représentation explicite des processus sous forme d'intégrale paramétrée par le temps. C'est ce qui a amené une limitation aux processus sur \mathbb{Z} à fonction de covariance rapidement décroissante quand $|t-s|$ augmente.

On aurait pu pourtant, être beaucoup plus général ; voici un aperçu rapide des choses. On suppose toujours que la formule $\int E(\xi X_t) X_t dt$ définit un opérateur V sur H éventuellement non borné d'ailleurs. Notons K le spectre de V . Dans la construction de Gelfand Vilenkin [14 p. 29] de V comme une intégrale hilbertienne intervient une représentation de H comme somme d'espace $L^2(K, \sigma_n)$ où σ_n est une mesure finie sur K . Tout $\xi \in H$ se représente donc comme une famille ϕ_n de fonctions sur K avec :

$$||\xi||^2 = \sum_n \int_K |\phi_n|^2 d\sigma_n$$

Considérons alors $L \subset H$ le sous espace partout dense représenté par des (ϕ_n) continues sur chaque copie de K et munissons le de la topologie de la convergence uniforme sur chaque copie de K , par exemple, de façon à ce que les $\xi_\lambda^p(\xi) = \xi_\lambda^p(\{\phi_n\}) = \phi_p(\lambda)$ soient des formes linéaires continues sur L . On a alors une formule de représentation de V :

$$V = \sum_n \int_K \lambda \xi_\lambda^n \otimes \xi_\lambda^n d\sigma_n(\lambda)$$

et les ξ_λ^n forment un système complet de H . On montre aussi, d'ailleurs que sur chaque (K, σ_n) on peut construire une mesure stochastique orthogonale $Z_n(\lambda)$ tel que l'isomorphisme entre H et $\sum_n (K, \sigma_n)$ prennent la forme :

$$\xi = \sum_n \int_K \phi_n(\lambda) dZ_n(\lambda)$$

On peut maintenant écrire $dZ_n(\lambda) = \xi_\lambda^n d\sigma_n(\lambda)$ dans la mesure où, dès que $\eta \in L$, on a :

$$E(\xi \bar{\eta}) = \sum_n \int_K \phi_n(\lambda) \overline{\langle \xi_\lambda^n, \eta \rangle} d\sigma_n(\lambda)$$

Les variables X_t elles-mêmes admettent une représentation comme intégrale stochastique :

$$X_t = \sum_n \int_K \phi_n^t(\lambda) dZ_n(\lambda)$$

Après avoir vu qu'on peut prendre les applications $(t, \lambda) \rightarrow \phi_n^t(\lambda)$ mesurables sur $T \times K$, on montre que les $\lambda^{-1/2} \phi_n^t(\lambda) = f_\lambda^n(t)$ forment un système complet de $L^2(T)$ et que le processus admet la représentation :

$$X_t = \sum_n \int_K \lambda^{1/2} f_\lambda^n(t) \xi_\lambda^n d\sigma_n(\lambda) \text{ en ce sens que}$$

pour presque tout T et tout $\xi \in L$ on a :

$$E(X_t \bar{\xi}) = \sum_n \int_K \lambda^{1/2} f_\lambda^n(t) \overline{\langle \xi_\lambda^n, \xi \rangle} d\sigma_n(\lambda)$$

On généralise ainsi l'analyse harmonique "généralisée" à une classe très large de processus.

Deuxième aventure : Les "généralisations" de 8 ou du paragraphe précédent permettent de retrouver l'analyse spectrale sur \mathbb{Z} ou \mathbb{R} . La clé est toujours le transport de la décomposition spectrale par la fonction ϕ , densité spectrale qui envoie une variété de dimension 1 dans \mathbb{R}^+ . Dès que le processus devient un champ aléatoire stationnaire, par exemple sur \mathbb{Z}^2 ou \mathbb{R}^n , ça ne va plus marcher : dans le premier cas ϕ , sauf pathologie, est inversible par morceaux. Ce ne sera jamais le cas (sauf pathologie) si ϕ envoie, par exemple, \mathbb{R}^2 sur \mathbb{R}^+ ; les deux représentations, spectrales et factorielles "surgénéralisées", ne pourront jamais coïncider. C'est d'ailleurs déjà ce qui se produit s'il se trouve que la densité spectrale est constante sur un ensemble de mesure positive de $]-\pi, +\pi]$.

Troisième aventure : Que peut faire, après tout ce déballage, un statisticien qui, pour des individus $i=1, \dots, N$, a recueilli des séries x_t^i ($t = T_0^i, T_0^i + 1, \dots, T_0^i + L(i)$), considérées comme extraites d'un processus justiciable de cette Théorie ?

Il aimerait trouver une représentation approchée analogue à la représentation théorique :

$$x_t^i \approx \sum_{k=1}^p u_k(t) z_k(i)$$

avec p petit, z_k variables deux à deux orthogonales de variances λ_k décroissantes avec k et les $u_k(t)$ "proches" des $u_{\lambda_k}^n(t)$ de la théorie.

Dans le cas où l'on assume la stationnarité on dispose d'une certaine technologie. On sait [15], par exemple, que la matrice des corrélations empiriques admet des valeurs propres qui sont réparties comme des $\phi(\omega_i)$, les ω_i étant tirés au hasard sur $]\pi, +\pi]$ et les vecteurs propres voisins des $\cos \omega_i t$ ou $\sin \omega_i t$. L'exemple suivant, élémentaire et classique [16] ou même [15] donne pourtant à réfléchir : X_t est stationnaire de fonction de covariance $\rho(0) = 1$, $\rho(\pm 1) = r$ ($|r| \leq \frac{1}{2}$), $\rho(h) = 0$ pour $|h| \geq 2$. C'est une moyenne mobile sur deux périodes

d'un bruit blanc, et sa densité spectrale est $\phi(\omega) = 1 + 2r \cos \omega$. Elle prend deux fois exactement toutes les valeurs comprises entre $1-2r$ et $1+2r$ et à chacune de ces valeurs correspond deux pseudo-vecteurs propres : $(e^{i\omega t})$ et $(e^{-i\omega t})$ $t \in \mathbb{Z}$.

Passons en temps fini $t = 1, 2, \dots, T$. La matrice de covariance :

$$\begin{bmatrix} 1 & r & 0 & \dots & 0 \\ r & 1 & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & r \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & r & 1 \end{bmatrix}$$

admet les valeurs propres :

$$\lambda_k = 1 + 2r \cos \frac{\pi}{T+1} k \quad k = 1, \dots, T$$

Les vecteurs propres étant :

$$V_k = \left\{ \sin \frac{k\pi}{T+1} t \right\}_{t=1, \dots, T} \quad k = 1, \dots, T$$

L'échantillonnage des valeurs propres suit bien un schéma correct mais celles-ci sont devenues toutes simples et associées à un vecteur propre qui est la trace sur $[1, T]$ de $(e^{i\omega t} - e^{-i\omega t})$.

Cet exemple pose donc la question de savoir ce qui se passe quand on travaille sur un intervalle fini et ce qui amène cette réduction des sous espaces propres, (et naturellement en quoi cet exemple a-t'il un caractère général).

12 - EN GUISE DE CONCLUSION

Dans le cas des processus non stationnaires la situation est encore plus angoissante.

Ou beaucoup moins si on se dit qu'il est de toute façon assez illusoire de vouloir inférer à partir d'observations finies sur toute la durée d'un processus. Il faut nécessairement soit s'en tenir compte à une description finie avec des méthodes "compactes", soit se servir de "modèles" de processus, dont une étude finie ou compacte permettra de déceler, d'estimer des paramètres "éternels".

Cette démarche devient nécessaire dès qu'il s'agit de prédire, c'est-à-dire de décrire le futur du processus à l'aide d'observations portant sur le passé et sous une hypothèse de régularité. Nous restons toujours, en somme, dans une problématique d'analyse des données, c'est-à-dire de statistique.

REMERCIEMENTS : L'auteur tient à remercier particulièrement Jacques DAUXOIS et Antoine de FALGUEROLLES pour l'aide qu'ils lui ont apporté dans la préparation de cet article.

Références

- [1] J.C. DEVILLE Méthodes statistiques et numériques de l'analyse harmonique. Annales de l'INSEE n° 15 (1974)
- [2] L.H. LOOMIS An introduction to abstract harmonic analysis. VAN NOSTRAND (1952)
- [3] N. BOURBAKI Théories spectrales. HERMANN (1967)
- [4] M. NAIMARK Normed Rings. WOLTERS-NOORHOFF (1972)
- [5] A. KIRILOV Eléments de la théorie des représentations. MIR (1979)
- [6] M. NAIMARK et A. STERN Théorie des représentations des groupes. MIR (1980)
- [7] ASH-GARDNER Topics in stochastic processes. ACADEMIC PRESS (1975)
- [8] G. MELARD Processus purement indéterminables à paramètre discret. Thèse université libre de Bruxelles (1975)
- ..
- [9] D. TJOSTHEIM- B. THOMAS Some properties and examples of random processes that are almost wide sense stationary. IEEE transactions on information theory - May (1975)
- ..
- [9bis] D. TJOSTHEIM Measuring deviations from stationarity. Stochastic processes and their applications Vol. 10 (1980)
- [10] R. TRUONG VAN MANH Contributions à l'étude des processus V bornés multidimensionnels. Thèse Université Paul Sabatier TOULOUSE (1979)
- [10bis] DANG DUC HAI Contributions à l'étude des processus non stationnaires. Applications à la prévision. Thèse Université Paul Sabatier TOULOUSE (1978)
- [11] BIRKHOFF-Mac LANE Algèbre Tome II - GAUTHIER VILLARS
- [12] D. BRILLINGER Time series : Data Analysis and Theory. HOLT, RINEHART, WINSTON (1975)
- [13] I.M. GELFAND-CHILOV Generalized Functions : Volume 3 : Theory of differential Equations. ACADEMIC PRESS (1966)
- [14] I.M. GELFAND-VILENKIN Generalized Functions : Volume 4 : Applications of harmonic Analysis. ACADEMIC PRESS (1964)
- [15] U. GRENNANDER-M. ROSENBLATT Statistical Analysis on Stationary Time Series ALMQUIST et WIKSELL (1956)
- [16] T.W. ANDERSON The Statistical Analysis of time Series. WILEY (1971)