HUITIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 1er au 5 JUIN 1981

Mélange de lois statistiques sphériquement invariantes.

D. de BRUCQ*

Laboratoire de Traitement de l'Information U.E.R. des Sciences et des Techniques
BP 67 76130 Mont-Saint-Aignan

RESUME

Les lois statistiques de processus centrés non gaussiens mais sphériquement invariants se généralisent pour des processus sommes de processus sphériquement invariants. Ces nouvelles lois modélisent la réception sur n capteurs, de l'émission des s sources indépendantes gaussiennes centrées ayant traversé un milieu qui évolue aléatoirement d'une expérience à l'autre. Chaque expérience fournit n mesures $Z_1, Z_2, \ldots, Z_n$, une par capteurs. De l'étude de la structure statistique, on déduit que la matrice de covariance des espérances conditionnelles $E(Z_i/Z_1)$ des Z_i par rapport à Z_1 est au plus d'ordre s ! un test statistique simple résulte de cette propriété pour déduire le nombre s de sources dès que le nombre n de capteurs vérifie $n \geq s+1$.

SUMMARY

Statistical distributions of centered spherically invariant - and so non gaussian processes are generalized : Sums of such processes are considered. These new distributions describe the reception, furmished by n pressure catchment devices, of s gaussian independent sources, emission which passes previously through a random medium. This one changes from an experiment to another. Each measurement furnisches n numbers $Z_1, Z_2, \ldots, Z_n$ one by pressure catchment device.

We deduce from the property of the statistical structure that the covariance matrix order of the conditional means $E(Z_i/Z_1)$ of Z_i with respect to Z_1 is at most s ! When the number n of pressure catchement devices is greater than s + 1, a simple statistical test results to obtain the number s of sources.

*Ce travail a été effectué dans le cadre d'un contrat Cethedec-Dret.

INTRODUCTION : Cet article introduit une nouvelle structure statistique pouvant décrire les bruits, notamment acoustiques marins. Cette structure généralise la notion de processus sphériquement invariants et conduit à un algorithme simple de détection du nombre de sources émettrices.

L'origine du modèle est expérimentale. On cherche à expliquer à partir de r sources physiques $S^1, S^2, \ldots, S^s$ les propriétés de corrélation des bruits observés sur n capteurs $C_1, C_2, \ldots, C_n$. Chaque source S^j émet une onde X^j aléatoire gaussienne centrée se propageant dans un milieu aléatoire. Ces processus X^j $j = 1, 2, \ldots, s$ sont supposés (statistiquement) indépendants deux à deux. Chaque capteur C_i $i = 1, 2, \ldots, n$ fournit une observation Z_i à partir des champs $Y^{(j)}$ provenant des X^j émis par la source C^j puis transformés par le milieu aléatoire.

S^1	X^1		Y_1^1	Y_1^s	C_1	Z_1
--	--	Milieu Aléatoire	--	--	--	--
S^j	X^j		Y_i^1	Y_i^s	C_i	Z_i
--	--		--	--	--	--
S^s	X^s		Y_n^1	Y_n^s	C_n	Z_n

Ainsi $(Z_1, Z_2, \ldots, Z_n)$ constitue une variable aléatoire à valeurs dans $(\mathbb{R}^n, \mathbb{R}^{n\otimes})$ et de loi de probabilité notée P. Seule cette loi de probabilité P peut être estimée en effectuant de nombreuses expériences. On cherche à décrire la structure statistique de $\mathcal{P}$ ensemble des probabilités P de $Z_1, Z_2, \ldots, Z_n$.

Un certain nombre d'hypothèses d'origine physique vont restreindre cette ensemble $\mathcal{P}$.

Pour tout i, on associe une fonction φ_i donnant la réponse Z_i du capteur excité par un champ aléatoire Pr (de Pression acoustique) soit

$$Z_i(w) = \int \varphi_i(x, y, z, t)\, P_r(x, y, z, t, w)\, dx\, dy\, dz\, dt.$$

Le capteur C, a clairement une étendue spatiale ; il peut se déplacer lors de la mesure d'où la présence des 4 variables x, y, z, t. Si le capteur physique fournit une réponse en temps continue, on échantillonne le temps et on obtient une succession de mesures du type précédent soit $Z_{1\Delta t}, Z_{2\Delta t}, \ldots, Z_{n\Delta t}$.

Le champ P provient de l'émission X des sources et on admet l'existence d'un noyau de Green G aléatoire pour décrire l'effet du milieu fournissant P_r soit

$$P_r(x, y, z, t, \omega) = \int G\, x, y, z, t, \xi, \eta, \zeta, \tau, \omega\,)X(\xi, \eta, \zeta, \tau, \omega)\, d\xi\, d\eta\, d\zeta\, d\tau.$$

L'émission X se décomposent en s émissions $X^{(j)}$ $j = 1, \ldots, s$ supposées gaussiennes centrées et deux à deux statistiquement indépendantes. Ainsi avec des notations évidentes

$$P_r(x, y, z, t, \omega) = \sum_{j=1}^{s} \int G^j\, x, y, z, t, \xi, \eta, \zeta, \tau, \omega\,)\, X^j(\xi, \eta, \zeta, \tau, \omega)\, d\xi\, d\eta\, d\zeta\, d\tau.$$

Chacun des noyaux G est aléatoire. Finalement on admet que cet effet est multiplicatif du type suivant $G^j(x, y, z, t, \omega) = A^j(\omega)\, g^j(x, y, z, t, \xi, \eta, \zeta, \tau)$.

L'effet de la source S^j sur l'ensemble des capteurs $C_1, \ldots, C_n$ varie d'une expérience à l'autre par un coefficient multiplicatif uniquement.

Posons

$$Y_i^{(j)}(\omega) \triangleq \int \varphi_i(x, y, z\, t)\, dx\, dy\, dz\, dt\; g^{(j)}(x, y, z, t, \xi, \eta, \zeta, \tau) X^{(j)}(\xi, \eta, \zeta, \tau)\, d\xi\, d\eta\, d\zeta\, d\tau$$

d'où la structure statisitique qui aurait pu être introduite axiomatiquement suivant l'usage actuel des mathématiciens.

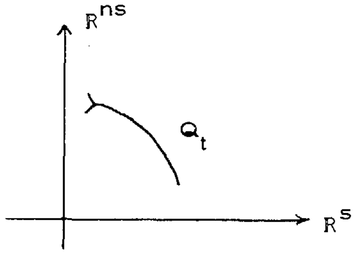
II Modèle Statistique

Les variables $A_1, A_2, \ldots, A_s$ sont à valeurs dans R^s et les vecteurs $Y^1, Y^2, \ldots, Y^s$ sont à valeurs dans $(R^n)^s$.

On note $\underline{\underline{R}}$, la tribu des boréliens de R.

On considère l'espace mesurable $(R^s \times R^{ns}, \underline{\underline{R}}^{s\otimes} \otimes \underline{\underline{R}}^{ns\otimes})$

Lorsque $A_1 = a_1, A_2 = a_2, \ldots, A_s = a_s$ sont fixés, on désire que les vecteurs $Y^1, Y^2, \ldots, Y^s$ soient gaussiens centrés et deux à deux indépendants :



On introduit une probabilité μ quelconque sur $(R^s, \underline{\underline{R}}^{\otimes s})$ et une probabilité de transition $Q_t = (Q_t(a, B)\ ;\ a \in R^s,\ B \in \underline{\underline{R}}^{ns\otimes})$ vérifiant

a) $\forall a \in R^s$ $Q_t(a, .)$ est une probabilité sur $\underline{\underline{R}}^{ns\otimes}$ telle que $Y^1, Y^2, \ldots, Y^s$ soient gaussiens centrés et deux à deux indépendants.

b) $\forall B \in \underline{\underline{R}}^{ns\otimes}$ $a \mapsto Q_t(a, B)$ est $\underline{\underline{R}}^{s\otimes}$ mesurable.

<u>Proposition II-1</u> : soit $Z_1 = A_1 Y_1^1 + \ldots + A_s Y_1^s$

$$Z_n = A_1 Y_n^1 + \ldots + A_1 Y_n^s$$

alors conditionnellement à $A_1 = a_1, A_2 = a_2, \ldots, A_n = a_n$, les variables aléatoires $Z_1, \ldots, Z_n$ sont gaussiennes centrées de covariance

$$\Gamma = E\left[\begin{pmatrix} Z_1 \\ - \\ Z_n \end{pmatrix} (Z_1, \ldots, Z_n)\right] = \sum_{j=1}^{s} a_j^2 \Gamma^j \text{ où}$$

$\Gamma^j = E(Y^j Y^{j*})$

est la covariance du vecteur Y^j pour $j = 1, \ldots, s$.

Dém. : On calcule la fonction caractéristique de $Z_1, \ldots, Z_n$ conditionnellement à $A_1, A_2, \ldots, A_n$:

Soit $\forall u \in R^n$

$E(e^{i<u, Z>} / A_1 = a_1, \ldots, A_n = a_n) =$

$$\int e^{i<u, z>} Q_t(a, dy) = \int e^{i<u, \sum_{j=1}^{s} a_j y^j>} Q_t(a, dy) =$$

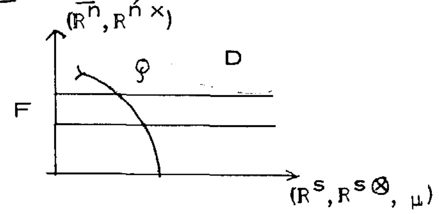
$$\int e^{i \sum_{j=1}^{s} a_j <u, y^j>} Q_t(a, dy) = e^{\frac{1}{2} \sum_{j=1}^{s} a_j^2 u^* \Gamma^j u}$$

On a noté u^* la matrice transposée de u vecteur colonne ■

<u>Corollaire II - 2</u> : Il existe une probabilité de transition $Q = (Q(a, C)\ ;\ a \in R^s, C \in \underline{\underline{R}}^{n\otimes})$ vérifiant

$$\forall a \in R^n \quad \forall \mu \in R^n \int e^{i<u, z>} Q(a, dz) = e^{-\frac{1}{2} \sum_{j=1}^{s} a_j^2 u^* \Gamma^j u}$$

La loi P de probabilité du vecteur Z est déterminée à l'aide de la loi μ marginale de $A_1, A_2, \ldots, A_s$ et de la transition Q : Pour tout ensemble D de $\underline{\underline{R}}^{s\otimes} \otimes \underline{\underline{R}}^{n\otimes}$ on définit



$$\mu . Q(D) \triangleq \int \mu(da) Q(a, dz) 1_D(a, z)$$

et pour tout F de $\underline{\underline{R}}^{n\otimes}$:

$$P(F) \triangleq \mu . Q(F \times R^s) = \int_{R^s} \mu(da) Q(a, F)$$

Voici la structure statistique $\mathcal{P}$ retenue :

Les lois de probabilités P sur $(R^n, \underline{\underline{R}}^{n\otimes})$ provenant du n upple de variables aléatoires observées $Z_1, Z_2, \ldots, Z_n$ constituent un ensemble dépendant d'une <u>probabilité μ arbitraire sur</u> $(R^s, \underline{\underline{R}}^{s\otimes})$, la transition Q est connue.

On note $M \triangleq \{\mu . Q\ ;\ \mu$ probabilité $\}$ l'énsemble des probabilités sur l'espace produit $R^s \times R^n$. Les observations $Z_1, Z_2, \ldots, Z_n$ ne permettent d'atteindre que la loi P marginale sur R^n de $\mu . Q$!

<u>Remarque</u> : La structure statistique introduite est une généralisation de la notion de processus gaussien centré qui correspond au cas $r = 1$, A certain - une source, milieu certain. Il généralise la notion de processus sphériquement invariants (cf [1] Picinbono [2] Besson) $r = 1$, A aléatoire.

Chacun des vecteurs $A_j Y^j$ est sphériquement invariant d'où le nom Z mélange de processus sphériquement invariants.

<u>Remarque</u> : Ce type de structure : $M = \{\ \mu . Q\ ;\ \mu$ probabilité $\}$ avec Q transition quelconque est un cas particulier de la notion de simplexe fort étudiée en statistique (cf [3] Dynkin). Les points extrémaux du simplexe M sont associés à $\mu = \delta_a$ mesure de dirac concentrée en a. Dans l'exemple les probabilités extrémales sont gaussiennes et la structure est constituée de toutes les moyennes définies par μ quelconque de ces gaussiennes.

III - Propriétés de la structure $\mathcal{P}$

Proposition III-1 : La variable aléatoire vectorielle Z a pour fonction caractéristique $\varphi(u)$

$$\varphi(u) = \int_{R^s} e^{-1/2 \sum_{i=1}^{s} a_j^2 u^* \Gamma^j u} \mu(da)$$

Dim : $\varphi(u) = E(e^{i<u, z>}) = E(E(e^{i<u, z>}/A_1, \dots, A_s))$ ■

Est-il possible de trouver une équation aux dérivées partielles linéaires du premier ordre, vérifiée par φ pour tout P de $\mathcal{P}$?

Observons que $\text{grad}\,\varphi(u) = -\int_{R^s} \sum_j a_j^2 \Gamma^j u \; e^{-\frac{1}{2} \sum a_j^2 u^* \Gamma^j u} \mu(da)$

Une telle équation aux dérivées partielles s'écrirait à l'aide de $L \in R^n$ vecteur des coefficients de l'équation :

$<L, \text{grad}\,\varphi> = 0$ donc $\forall a \in R^s$, il faut $\sum_{j=1}^{s} a_j^2 <L, \Gamma^j u> = 0$

soit encore $\forall j = 1, \dots, s$ $<L, \Gamma^j u> = 0$

Le vecteur L devrait être dans l'orthogonal de l'espace vectoriel V engendré par $\Gamma^1 u, \Gamma^2 u, \dots, \Gamma^s u$ avec u éventuellement restreint à un sous espace W de R^n c'est à dire $V \triangleq V\{\Gamma^1 u, \dots, \Gamma^s u ; u \in W\}$. Si W est de dimension k alors V est de dimension au plus égale à s k. Dès que $sk < n$ l'orthogonal $V^\perp$ de V dans R^n est de dimension $n-sk > 0$ au moins et tout vecteur L de $V^\perp$ convient.

Proposition III-2 : Si $n > s$, il existe un vecteur $L \in R^n$ telle que $\forall u_1 \in R$

$$L_1 \frac{\partial \varphi}{\partial u_1}(u_1, 0, \dots, 0) + L_2 \frac{\partial \varphi}{\partial u_2}(u_1, 0, \dots, 0) + \dots + L_n \frac{\partial \varphi}{\partial u_n}(u_1, 0, \dots, 0) = 0$$

Dém. : Soit $W \triangleq v(1, 0, \dots, 0)^*$ l'espace vectoriel engendré par le premier vecteur de base. Par suite $\Gamma^1 u, \Gamma^2 u, \dots, \Gamma^s u$, avec $u \in W$ est au plus de dimension s dans R^n et il existe L orthogonal à ces vecteurs, $<L, \text{grad}\,\varphi(u)> =$

$$-\int \sum_j a_j^2 <L, \Gamma^j u> e^{-\frac{1}{2} \sum_j a_j^2 u^* \Gamma^j u} \mu(da)$$

est nul pour $u \in W$ quelque soit μ. ■

Théorème III-3 : Si $Z_1, Z_2, \dots, Z_n$ sont intégrables alors $Z_1, E(Z_2/Z_1), \dots, E(Z_n/Z_1)$ est au plus de dimension s.

Dém. : Il suffit de vérifier que $E(Z_{s+1}/Z_1)$ est combinaison linéaire de $Z_1, E(Z_2/Z_1) \dots E(Z_s/Z_1)$. Or $\exists L \in R^{s+1}$ tel que $<L, \text{grad}\,\varphi(u_1, 0, \dots, 0)> = 0$ $\forall u_1 \in R$ d'où

$$\sum_{i=1}^{s+1} L_i \frac{\partial}{\partial u_i} E(e^{i<u, z>}) \Big|_{u_1, 0, \dots, 0} = 0 \Leftrightarrow$$

$$E(E(\sum_{i=1}^{s+1} L_i Z_i/Z_1) e^{i u_1 Z_1}) = 0 \Leftrightarrow$$

$$L_1 Z_1 + L_2 E(Z_2/Z_1) + \dots + L_{s+1} E(Z_{s+1}/L_1) = 0$$ ■

IV - Algorithme d'identification du nombre de sources.

Soit s le nombre de sources S_i à déterminer à l'aide des observations $Z_1, Z_2, \dots, Z_n$ fournies par les capteurs C, le milieu est aléatoire : d'une expérience à l'autre le milieu évolue et modifie l'effet de chaque source sur les capteurs.

Rappelons que pour le vecteur aléatoire

$Z_1, E(Z_2/Z_1), \dots, E(Z_n/Z_1)$, la covariance de ce vecteur est la matrice de Gram :

$$= E[\begin{pmatrix} Z_1 \\ E(Z_2/Z_1) \\ ---- \\ E(Z_n/Z_1) \end{pmatrix} (Z_1, E(Z_2/Z_1), \dots, E(Z_n/Z_1))]$$

Cette matrice Γ est différente de la covariance de $Z_1, Z_2, \dots, Z_n$ qui vaut

$$E[\begin{pmatrix} Z_1 \\ Z_2 \\ --- \\ Z_n \end{pmatrix} (Z_1, Z_2, \dots, Z_n)]$$ et qui est utilisée usuellement pour déterminer le nombre de sources.

Le résultat suivant est connu :

Proposition IV-1 : L ordre d'une matrice de Gram est égale à la dimension de l'espace vectoriel engendré par les variables aléatoires dont elle est la covariance.

Ainsi Γ a pour ordre, le nombre de sources.
Calculer l'ordre de Γ c'est obtenir le nombre de sources.
Au lieu de la référence Z_1 associé au premier capteur il vaut mieux prendre

$$Z_1' \triangleq \frac{Z_1 + Z_2 + \dots + Z_n}{n}$$

pour lequel les erreurs de mesure sont diminuées.
Ensuite on peut normaliser Z_1' en divisant par l'écart type provenant des mesures. Les espérances conditionnelles $E(Z_i/Z_1)$ $i = 1, 2, \dots, n$ sont estimées à partir des tableaux statistiques des variables bidimensionnelles :

Z_1 \ Z_i	$E(Z_i/Z_1)$
classe 1	$E(Z_i/Z_1 \in 1)$
classe 2	$E(Z_i/Z_1 \in 2)$
---	---
classe k	$E(Z_i/Z_1 \in k)$
---	---
classe l	$E(Z_i/Z_1 \in l)$

Les diverses classes $1, 2, \dots, l$ doivent être sensiblement équiprobables et contenir chacune plus de 30 valeurs de $Z_1(\omega)$.

L'estimation des moyennes conditionnelles ne nécessite pas de regroupement par classe pour Z_i, les valeurs expérimentales sont utilisées directement.
Finalement la méthode d'inversion d'une matrice quelconque par la méthode du pivot fournit un procédé simple pour obtenir l'ordre de la matrice Γ :
Si les coefficients de la matrice Γ sont connus avec une précision de 1%, à l'étape m de l'inversion par la méthode du pivot, celui-ci inférieur à 1% signifie que la colonne m est dans l'espace vectoriel des m-1 colonnes antérieures et l'ordre de la matrice reste celui calculé pour les m-1 premières colonnes.

<u>BIBLIOGRAPHIE :</u>

[1] Picinbono, Spherically Invariant and Compound gaussian stochastic processes. IEEE Transactions on Information Theory Vol 1 T-6 1970 p 77-79.

[2] T.L. Besson, Continuité des fonctions aléatoires sphériquement invariantes. Publications du Département de Mathématiques 1974 T II-3 p59-68.

[3] E. B. Dynkin, Sufficient statistics and extreme points.The annals of Probability N° 5 v 6 1978 p 705-730.