

# DIXIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 20 au 24 MAI 1985

---

Estimation d'une matrice de structure imposée.

Application au traitement d'antennes

J.P. LE CADRE

GERDSM LE BRUSC  
83140 SIX FOURS LES PLAGES

CEPHAG - B.P. 46  
38402 ST MARTIN D'HERES

---

## RESUME

On présente diverses utilisations de structure imposées pour l'estimation de matrices interspectrales.

On décrit ici un cadre général pour les méthodes envisagées, les algorithmes explicites (itératifs) et les conditions de convergence ; et enfin les conclusions que l'on a pu tirer de l'utilisation de ces algorithmes.

L'objet de cette étude est de montrer l'intérêt d'inclure une information a priori par l'intermédiaire d'une structure imposée, puis de résoudre les problèmes d'optimisation correspondants.

Ces informations a priori correspondent à l'hypothèse ondes planes, modèle AR, ou structure de bruit à décorrélation rapide.

## SUMMARY

One presents some use of constrained structures for the estimation of cross-spectral matrices.

One describes a general frame for the relevant methods, the explicit algorithms (iterative) and the conditions of convergence, and last the conclusions obtained by use of these algorithms.

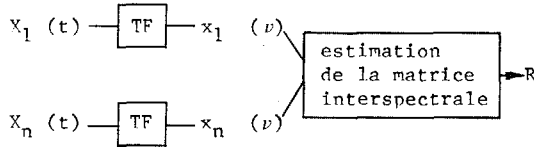
The aim of this article is to prove the interest of inclusion of a priori informations by the way of fixed structures and then to solve the subsequent problems of optimization. These a priori informations correspond to the plane waves hypothesis, or ambient sea noise with fast spatial decorrelation.



1 - INTRODUCTION

Le but général de ces méthodes est d'inclure une information a priori pour l'estimation de matrices interspectrales de structure imposée.

On s'intéresse donc au problème suivant :



R est donc la matrice interspectrale des observations (signaux recus sur les capteurs). R sera considérée comme résumant l'information disponible (ce choix est évidemment contestable). Le problème général sera d'optimiser un critère (maximum de vraisemblance ou d'entropie)  $\mathcal{C}(R, M)$ , M ayant une structure imposée. On considère principalement les problèmes posés par l'estimation de R dans  $\mathcal{C}_n$  ( $\mathcal{C}_n$  s.e.v des matrices de Toeplitz), l'estimation de R comme matrice de covariance d'un processus AR au sens du maximum de vraisemblance et enfin l'estimation des paramètres définissant la covariance spatiale du bruit ambiant.

2 - ESTIMATION DE R DANS  $\mathcal{C}_n$  : [1], [5]

2.1 - On veut déterminer M qui soit la plus proche de R au sens d'un certain critère de contraste (Inclusion de l'hypothèse "ondes planes").

Ainsi si on considère la minimisation de la distance associée à la norme hilbertienne i.e :

$$(1) d(R, M) = \text{tr} [ (R-M) \cdot (R-M)^* ]$$

On obtient que M est la matrice obtenue à partir de R par moyennage des diagonales.

Mais on peut aussi considérer la maximisation de la Log-vraisemblance qui peut s'écrire dans le cas d'observations indépendantes [1], [6] :

$$(2) +f(R, M) = \text{Log det } M + \text{tr} (M^{-1} \cdot R)$$

Ce problème de maximisation n'a pas de solution explicite aussi on doit utiliser une méthode itérative en l'occurrence une méthode de gradient.

On peut prouver que f(R, M) est concave sur l'ensemble des matrices M telles que (2R-M) soient positives, mais la convergence de méthodes de gradient à pas optimal peut être prouvée sur un ensemble bien plus grand [6]. En fait, on n'a jamais observé de divergence de l'algorithme de gradient à pas optimal.

2.2 - Description de l'algorithme de gradient (Minimisation de  $-f(R, M)$ )

$\{U_i\}_{i=1}^n$  étant une base de  $\mathcal{C}_n$ ,

on écrit :

$$M = \sum_{i=1}^n \alpha_i \cdot U_i$$

Le vecteur gradient s'écrit à l'itération k :

$$(3) \bar{\nabla} f(R, M_k) = \begin{bmatrix} + \text{tr} [ M_k^{-1} \cdot (M_k - R) \cdot M_k^{-1} \cdot U_1 ] = d_k^1 \\ \vdots \\ + \text{tr} [ M_k^{-1} \cdot (M_k - R) \cdot M_k^{-1} \cdot U_n ] = d_k^n \end{bmatrix}$$

$$\text{où : } M_k = \sum x_i^k \cdot U_i$$

Un algorithme de gradient consiste donc à écrire :

$$(4) \bar{X}_{k+1} = \bar{X}_k - \rho_k \bar{\nabla} f(R, M_k)$$

$\bar{\nabla} f(R, M_k)$  ayant été calculé, on cherche à déterminer  $\rho$  qui minimise :

$$f(R, M_{k+1}) = f(R, M_k - \rho \Omega_k) \text{ où } \Omega_k = \sum_{i=1}^n d_k^i \cdot U_i$$

Pour cela, on utilise la décomposition en facteur de Choleski de la matrice  $M_k$  (supposée positive), ce qui permet d'écrire :

$$(5) \det(M_{k+1}) = \det(M_k - \rho \Omega_k) = \det M_k \cdot \det(\text{Id} - \rho T_k^{-1*} \cdot \Omega_k \cdot T_k^{-1})$$

et

$$(6) \text{tr} [ M_{k+1}^{-1} \cdot R ] = \text{tr} [ T_k^{-1*} \cdot (\text{Id} - \rho T_k^{-1*} \cdot \Omega_k \cdot T_k^{-1})^{-1} \cdot T_k^{-1} \cdot R ]$$

Introduisant la décomposition en éléments propres de la matrice  $T_k^{-1*} \cdot \Omega_k \cdot T_k^{-1}$  on

obtient f(R,  $M_{k+1}$ ) comme une fonction explicite de  $\rho$  soit :

$$(7) T_k^{-1*} \cdot \Omega_k \cdot T_k^{-1} = \sum_{i=1}^n \lambda_i^k \bar{v}_i \cdot \bar{v}_i^* ; \bar{v}_i \perp \bar{v}_j \quad i \neq j$$

et

$$(8) f(R, M_{k+1}) = \text{Log det } M_k + \sum_{i=1}^n \text{Log} (1 - \rho \lambda_i^k) + \sum_{i=1}^n (1 - \rho \lambda_i^k)^{-1} \bar{v}_i^* \cdot T_k^{-1} R \cdot T_k^{-1*} \cdot \bar{v}_i$$

Cette fonction de  $\rho$  pouvant être minimisée (rapidement) par un algorithme de Newton unidimensionnel. De plus, la contrainte  $M_k$  positive se traduit par la contrainte unidimensionnelle :

$$0 < \rho_k < \frac{1}{\max(\lambda_i^k)}$$

2.3 - Utilisation de l'algorithme : fig 1, 2

Cet algorithme a été utilisé sur des signaux simulés et réels et a donné des résultats intéressants pour des produits RT assez faibles (mais la limitation du produit RT est due à la non stationnarité des signaux). L'utilisation de la vraisemblance peut donner des résultats différents de ceux obtenus par projection orthogonale sur  $\mathcal{C}_n$  et parfois sensiblement meilleurs, mais au prix de calculs bien plus lourds. C'est pourquoi il est .../...



Estimation d'une matrice de structure imposée  
Application au traitement d'antennes

intéressant de développer pour les méthodes AR une méthode récurrente (surl'ordre du modèle) de la matrice M correspondant à un modèle AR.

3 - ESTIMATION AU SENS DU MAXIMUM DE VRAISEMBLANCE  
D'UN MODELE AR

3.1 - Le but est maintenant de déterminer un modèle AR décrivant le vecteur  $\bar{X}$ , i.e :

$$(9) \quad X_n = \sum_{i=1}^{n-1} a_i \cdot X_{n-i} + e$$

Alors la vraisemblance conditionnelle à  $(\bar{A}, \bar{e})$  s'écrit :

$$\begin{aligned} \text{Log } p(\bar{X}_1, \dots, \bar{X}_n / \bar{A}, \bar{e}) \\ = - \text{Log det } M(\bar{A}, \bar{e}) - \text{tr}(M^{-1}(\bar{A}, \bar{e}) \cdot R) \end{aligned}$$

L'équation de Yule-Walker permet d'écrire :

$$(10) \quad \Delta_n = T_n \cdot M_n \cdot T_n^*$$

$\Delta_n$  étant la matrice diagonale des

erreurs de prédiction et  $T_n$  la matrice triangulaire des vecteurs de prédiction aux différents ordres.

On en déduit que :

$$(11) \quad \begin{aligned} \text{Log det } M_n &= n \text{ Log } \sigma_{n-1}^2 - (n-1) \\ &\text{Log}(1 - |K_{n-1}|^2) \dots - \text{Log}(1 - |K_1|^2) \end{aligned}$$

$K_i$  étant le coefficient de réflexion à l'ordre  $i$ .

On se ramène alors à considérer la fonction de  $K_{n-1}$ .

$$(12) \quad \begin{aligned} L_n(\bar{X}_1, \dots, \bar{X}_n / \bar{A}_n) \\ = -m \cdot n \text{ Log}(\bar{A}_n^* R \bar{A}_n) + m(n-1) \text{ Log}(1 - |K_{n-1}|^2) + c \end{aligned}$$

où :

$$(13) \quad \bar{A}_n = \begin{pmatrix} \bar{A}_{n-1} \\ \dots \\ \bar{A}_{n-1} \end{pmatrix} + K_{n-1} \begin{pmatrix} \bar{A}_{n-1} \\ \dots \\ \bar{A}_{n-1} \end{pmatrix}$$

Il s'agit de déterminer  $K_{n-1}$  maximisant  $L_n$  [2], [4], [5], ce qui s'obtient par découpage par blocs de  $R$ . ( $K_{n-1} \in D(0,1)$ ).  $K_{n-1}$  étant obtenu par la résolution d'une équation du 3<sup>ème</sup> degré à coefficients réels.

3.2 - Utilisation de l'algorithme fig. 3, 4

L'algorithme a été testé sur des données simulées et réelles, on observe une amélioration très sensible de la stabilité de l'estimation des puissances relatives des sources. De plus, l'analyse des pôles de la fraction rationnelle associée au modèle prouve que les pôles dus au bruit sont ramenés vers l'intérieur du disque, cf fig. , et ceci est en grande partie du à l'effet stabilisateur du déterminant comme il a été établi par [3].

En outre, cet algorithme est (relativement) peu coûteux en temps de calcul.

.../...

4 - ESTIMATION DES PARAMETRES DE CORRELATION DU  
BRUIT AMBIANT

4.1 - On suppose que M peut s'écrire sous la forme :

$$(14) \quad M = \sum_{i=1}^p \sigma_i \bar{n}_{\theta_i} \cdot \bar{n}_{\theta_i}^* + R \quad (p \text{ ondes planes})$$

R étant la matrice du bruit ambiant, on peut la supposer de Toeplitz. De plus, le bruit étant spatialement assez "vite" décorréolé, il est raisonnable de supposer que R n'est déterminée que par quelques éléments de sa diagonale, i.e :

$$R = \sum_{i=1}^{p_0} \beta_i (U_i)$$

Le but est ici d'estimer les paramètres  $\beta_i$  et pour cela, on considère deux types d'approches dont on expose seulement les grandes lignes.

Méthode 1 : Méthodes des directions fixes de  
scrutation

On écrit :

$$M = \sum_{i=1}^{p'} \tau_i \bar{n}_{\theta_i} \cdot \bar{n}_{\theta_i}^* + R(\beta_1, \dots, \beta_{p_0})$$

$\{\theta_i\}_{i=1}^{p'}$  correspond aux directions fixes de scrutation, on choisit donc  $p'$  assez grand afin que l'estimation des  $\tau_i$  ne soit pas trop erronée du fait du pas de discrétisation.

Puis on détermine  $\{\tau_i\}_{i=1}^{p'}$ ,  $\{\beta_i\}_{i=1}^{p_0}$  par un algorithme de gradient à pas optimal analogue à celui développé en § 2.2, les conditions de convergence étant identiques.

Méthode 2 : Méthode de l'entropie relative

On suppose que l'antenne est constituée de  $n$  capteurs,  $n$  étant "largement" supérieur à  $p_0$ . On considère alors le vecteur :

$$\bar{X}^t = \{(\bar{S} + \bar{R})_1, \bar{R}_1, \dots, \bar{R}_L\}$$

$(\bar{S} + \bar{R})_1$  étant le signal reçu sur le premier groupe de  $p$  capteurs,  $\bar{R}_i$  étant le bruit ambiant reçu le ième groupe de capteurs.

De plus, on suppose (hypothèse fondamentale) les  $\bar{R}_i$  décorréolés, cette hypothèse étant justifiée si les groupes successifs de capteurs sont séparés par  $p_0$  capteurs.

Alors l'entropie du vecteur  $\bar{X}$  (supposé gaussien) s'écrit :

$$(14) \quad H = \text{Log det} \begin{bmatrix} R_p & R & 0 & \dots & 0 \\ R & R & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & 0 \end{bmatrix} \quad \begin{array}{l} R \text{ incon-} \\ \text{nue mais} \\ \text{de struc-} \\ \text{ture connue} \end{array}$$

Utilisant la formule usuelle :

$$\text{Det} \begin{bmatrix} A & B \\ C & n \end{bmatrix} = \text{det } A \text{ det}(n - C \cdot A^{-1} \cdot B), \text{ on}$$

peut établir par récurrence que :



(15)  $H = \text{Log det } (\hat{R}_p - B) + L \cdot \text{Log det } B$   
 $\hat{R}_p$  étant la sous-matrice extraite de  $\hat{R}$  en  
 prenant les  $p$  premiers termes de corrélation  
 ( $\hat{R} \in \mathcal{C}_n$ ).

$L$  qui est ici de l'ordre de  $n/p_0$  détermine l'importance de l'information a priori (ici c'est la longueur de corrélation du bruit ambiant).

La fonction  $H$  est une fonction des paramètres  $\beta_i$  caractérisant la corrélation spatiale du bruit. Il faut donc maximiser  $H$  relativement aux paramètres  $\{\beta_i\}$ . Ceci ne présente pas de problème, en effet :

$$\bar{V}^2 H(V, W) = -\text{tr} [(R - B)^{-1} \cdot V (R - B)^{-1} \cdot W] \\ - \text{tr} [B^{-1} \cdot V \cdot B^{-1} \cdot W]$$

d'où :

$$(16) \quad \bar{\beta}^* \cdot \bar{V}^2 H \cdot \bar{\beta} = -\text{tr} [(R - B)^{-1} \cdot \\ (\sum \beta_i U_i) \cdot (R - B)^{-1} \cdot (\sum \beta_i U_i)] \\ - \text{tr} [B^{-1} \cdot (\sum \beta_i U_i) \cdot B^{-1} \cdot (\sum \beta_i U_i)]$$

La fonction  $H$  est donc concave lorsqu'on fait l'hypothèse que  $B$  et  $R - B$  sont positives, l'algorithme de gradient à pas optimal décrit en § 2.2 s'applique donc.

Examinons le résultat obtenu lorsqu'on suppose que  $B = \lambda \text{Id}$ , on montre aisément le résultat suivant :

$$\left( \frac{L}{L+1} \right) \alpha_{\min} \leq \hat{\lambda} \leq \alpha_{\min} \quad (\alpha_{\min} \text{ plus petite}$$

valeur propre de  $R$ ).

Si  $L$  devient grand l'approximation de la plus petite valeur propre devient donc d'autant meilleure.

La méthode 2 présente l'avantage important de ne faire intervenir que les paramètres  $\beta_i$ .

#### 4.2 - Utilisation

La méthode 1 a été utilisée sur des signaux simulés, la difficulté principale est due au choix du nombre de directions de scrutation relativement au nombre de capteurs. Si le choix est bien adapté, les résultats pour l'estimation des paramètres  $\beta_i$  sont intéressants quoique la con-

vergence soit lente. A cet égard, la méthode 2 apparaît plus convaincante.

L'utilisation sur des signaux réels est assez difficile, les  $\beta_i$  ayant tendance à évoluer

au cours du temps, un algorithme évolutif apparaît nécessaire.

#### 5 - CONCLUSION

On a tenté de mettre en évidence l'intérêt d'inclure une information a priori en imposant une

structure aux matrices interspectrales. Les paramètres correspondants sont alors estimés au sens de vraisemblance ou d'entropie.

Les améliorations ainsi apportées dépendent beaucoup des paramètres définissant la simulation ( $RT$ ,  $S/R$ , caractéristiques du bruit).

#### BIBLIOGRAPHIE

- [1] J.P. Burg, D.G. Luenberger, D.L. Venner, Estimation of structured covariance matrix, Proc. IEEE, 70, n° 9, Sept. 82, pp 963-974.
- [2] H. Clergeot, Estimation spectrale par le critère du maximum de vraisemblance pour un modèle AR, Actes du 9<sup>ème</sup> colloque du GRETSI, Nice, Mai 83.
- [3] U. Grenander and G. Szegö, Toeplitz forms and their applications, University of California Press, San Francisco, 1958.
- [4] S.M. Kay, Recursive maximum likelihood estimation of autoregressive processes, IEEE Trans on ASSP, Vol. ASSP 31, n° 1, Feb. 83, pp 56-65.
- [5] J.P. Le Cadre, Estimation au sens du maximum de vraisemblance d'un modèle autorégressif, Application au traitement d'antennes, soumis à Traitement du Signal.
- [6] J.P. Le Cadre et P. Lopez, Estimation d'une matrice interspectrale de structure imposée, Traitement du Signal, Vol. TS-1, n° 1, Dec. 84, p 3.17.

Estimation d'une matrice de structure imposée  
 Application au traitement d'antennes

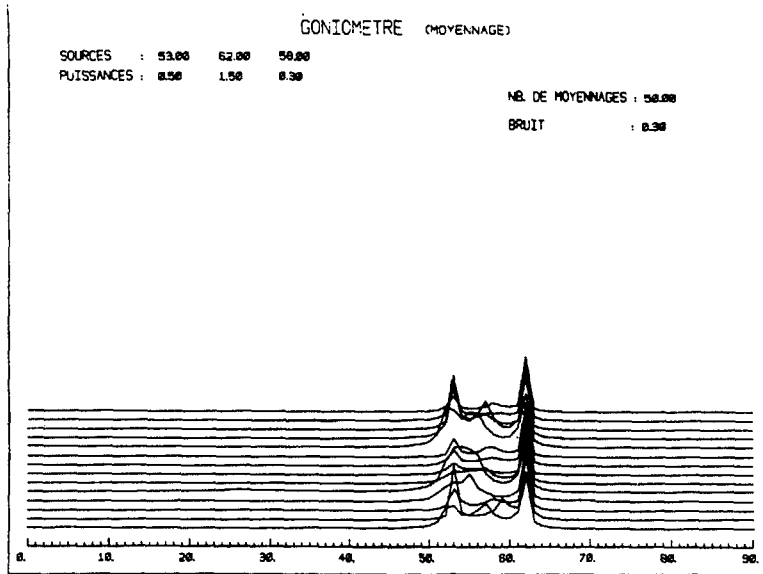


fig 1

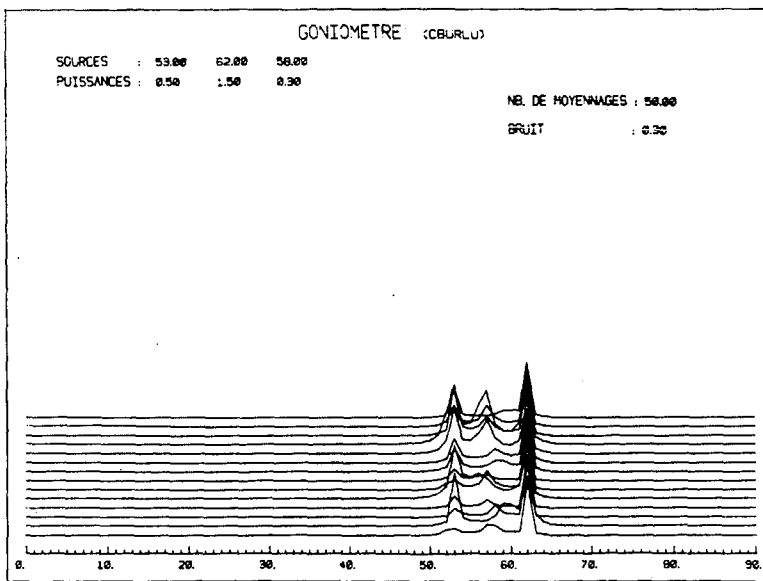


fig 2



Estimation d'une matrice de structure imposée  
 Application au traitement d'antennes

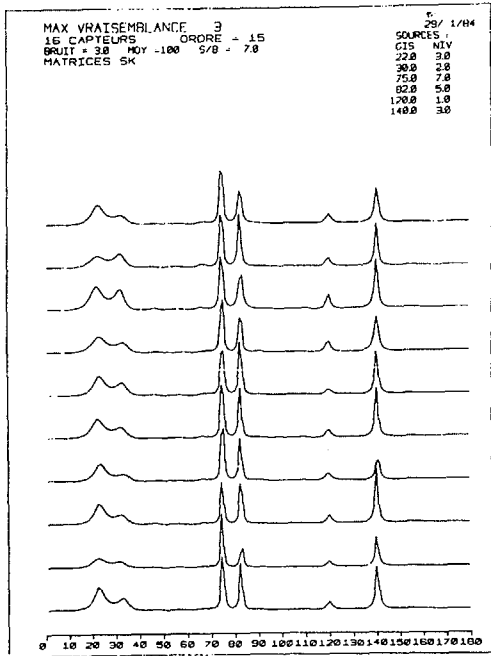


fig 3

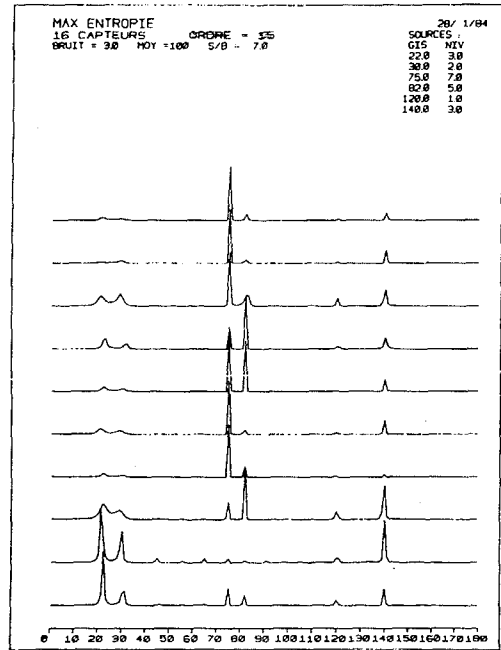


fig 4

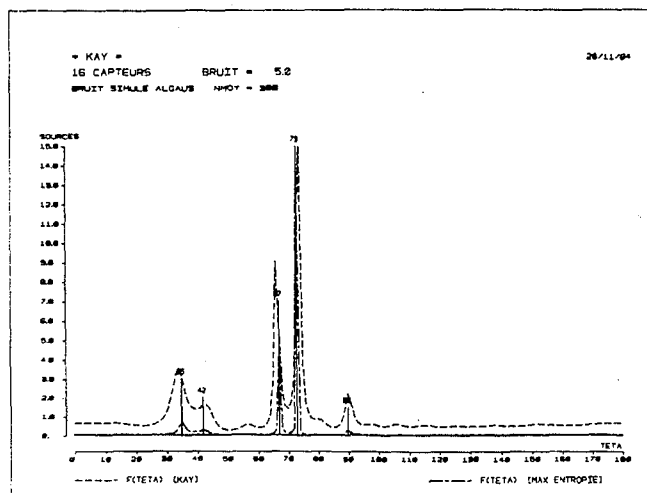


fig 5