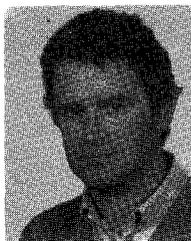


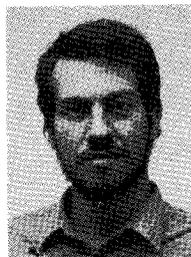
Estimation d'une matrice interspectrale de structure imposée. Applications



Jean-Pierre LE CADRE

(CEPHAG, BP n° 46, 38402 SAINT-MARTIN-D'HÈRES
GERDSM, Le Brusac, 83140 SIX-FOURS-LES-PLAGES

Jean-Pierre LE CADRE a soutenu une thèse de doctorat de 3^e cycle en 1982 à l'INPG en traitement du signal. Depuis 1980 son activité au Groupe d'Études et de Recherche de Détection Sous-Marine porte essentiellement sur les traitements d'antennes appliqués à l'acoustique sous-marine.



Philippe LOPEZ

GERDSM, Le Brusac, 83140 SIX-FOURS-LES-PLAGES

Philippe LOPEZ est diplômé de l'École Française de Radioélectricité, d'Électronique et d'Informatique de Paris. En tant que scientifique du contingent au Groupe d'Études et de Recherches de Détection Sous-Marine au Brusac (GERDSM), il a été chargé de l'implantation sur calculateurs d'algorithmes de traitement d'antennes, faisant appel à l'application de méthodes numériques et algébriques pour la résolution de problèmes de détection et d'estimation.

RÉSUMÉ

Cet article a pour but d'examiner les algorithmes d'estimation d'une matrice ayant une structure imposée. Les conditions de convergence des méthodes de gradient sont examinées ainsi que la détermination explicite d'algorithmes à pas optimaux; puis on examine diverses extensions. Enfin un exemple d'application au traitement spatial d'antennes (sur simulations et signaux enregistrés à la mer) est proposé.

MOTS CLÉS

Matrices, maximum de vraisemblance, moindres carrés, Toeplitz, gradient, convergence, convexité, algorithme.

SUMMARY

The aim of this paper is to consider estimation algorithms of a matrix with a fixed structure. The convergence conditions of the steepest descent methods as well as the explicit determination of optimal steps algorithms are studied, along with some possible extensions. Finally, application examples to the area of spatial signal processing are considered both on simulated data and signals recorded at sea.

KEY WORDS

Matrices, maximum Likelihood, least Square, Toeplitz, gradient, convergence, algorithm.

TABLE DES MATIÈRES

Principales notations utilisées

1. Introduction
 2. Modèles et présentation du problème
 3. Convergence de méthodes itératives
 - 3.1. Étude de la convexité et de l'ellipticité dans le cas général
 - 3.2. Étude de la convexité et de l'ellipticité dans le cas paramétrique
 - 3.3. Extension des conditions de convergence
 - 3.4. Conséquences d'une inégalité fondamentale
 4. Description d'un algorithme de gradient à pas optimal
 - 4.1. Détermination du pas optimal
 - 4.2. Utilisation pratique de l'algorithme
 - 4.3. Opérateur de projection et convergence
 5. Mise en œuvre de l'algorithme
 6. Extensions de ce type de méthodes
 7. Présentation de résultats de simulations
 8. Présentation de résultats expérimentaux
- Conclusion
Annexe 1
Annexe 2
Bibliographie

Principales notations utilisées

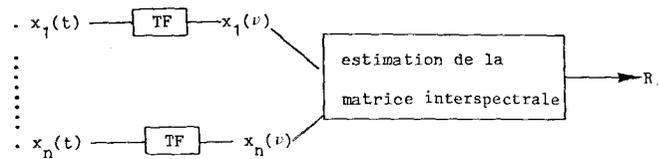
- tr, trace d'une matrice;
 $*$, A^* est la matrice transposée conjuguée de A ;
 det, déterminant d'une matrice;
 \mathcal{T}_n , sous-espace vectoriel des matrices de Toeplitz réelles de dimension n ;
 ∇f_M , gradient de f en M ;
 $\nabla^2 f_M$, différentielle seconde de f en M ;
 $\|M\|_{\mathcal{H}}$, norme hilbertienne de M
 i. e. $\text{tr}(M^* \cdot M) = \|M\|_{\mathcal{H}}^2$;
 $\{U_i\}_{i=1}^n$, base de \mathcal{T}_n définie par :

$$U_i(k, l) = \begin{cases} 1 & \text{si } [k-l]=i, \\ 0 & \text{sinon.} \end{cases}$$
 Cette base est orthogonale pour la norme \mathcal{H} .
 \mathcal{M}_n , espace vectoriel des matrices carrées de dimension n ;
 \mathcal{U}_n , groupe unitaire de dimension n ;
 Id, matrice identité de l'élément neutre de la multiplication.

1. Introduction

Le but poursuivi ici est l'estimation d'une matrice interspectrale (ou de covariance) dans un sous-espace (ou un sous-ensemble) matriciel donné. Plus précisément, on

s'intéresse au problème suivant rencontré très fréquemment en acoustique sous-marine :



La figure ci-dessus décrit la situation où les $\{X_i(t)\}$ sont les sorties de n capteurs d'une antenne linéaire à capteurs équirépartis. R est la matrice interspectrale estimée par exemple par la méthode du périodogramme. Cette matrice sert à décrire les champs rayonnés par des sources éloignées [7].

L'hypothèse « ondes planes » est généralement admise en acoustique sous-marine (à grande distance de la source), elle est à la base de tous les traitements d'antennes. Cette hypothèse implique que la matrice interspectrale asymptotique des sources est de Toeplitz pour une antenne linéaire à capteurs équirépartis. Réciproquement, l'hypothèse « matrice asymptotique de Toeplitz » implique par le théorème de Carathéodory [5] que les signaux reçus sur les capteurs peuvent s'interpréter comme la somme de p ondes planes (l'équivalent des fréquences pures dans le domaine temporel) et de bruit ambiant. L'estimation de la matrice interspectrale dans \mathcal{T}_n est donc essentielle pour l'application des méthodes à « haute résolution ».

Si on fait l'hypothèse ondes planes R est asymptotiquement une matrice de Toeplitz. On se pose donc le problème suivant : comment déterminer la matrice M (modèle) appartenant au sous-espace vectoriel des matrices de Toeplitz de dimension n (noté \mathcal{T}_n) qui soit la plus « proche » de R au sens d'un certain critère de contraste ? L'intérêt de cette estimation est d'inclure l'hypothèse « ondes planes » dans l'estimation de la matrice interspectrale, on peut donc espérer améliorer ainsi le pouvoir de séparation et de détection des méthodes à haute résolution (Méthodes de Pisarenko et du maximum d'entropie). Cet aspect a fait l'objet de nombreuses simulations et expérimentations.

L'estimation de M dans \mathcal{T}_n peut être obtenue par les estimateurs habituels. On peut ainsi considérer l'estimation au sens des moindres carrés ou du maximum de vraisemblance. Si le problème associé au premier estimateur a une solution explicite (en fonction des éléments de R), il n'en est pas de même du second qui nécessite l'utilisation de méthodes itératives. Il est évident que si l'on dispose de signaux stationnaires durant un temps d'analyse « grand » (c'est-à-dire si la matrice interspectrale des observations peut être estimée avec une erreur faible), les deux méthodes donneront des résultats très voisins. Dans la réalité, ce n'est pas le cas; les caractéristiques des bruiteurs, le canal sous-marin évoluent au cours du temps et le temps de stationnarité est donc limité.

Au-delà de cette limite les performances des traitements se dégradent. C'est un aspect opérationnel important qu'il est nécessaire de considérer pour chaque type d'applications. C'est pourquoi, la comparaison entre ces deux méthodes d'estimation de M est particulièrement intéressante dans ce contexte.

Ce type d'idées a été développé précédemment dans [3]. On s'est attaché ici à examiner la convergence de ce

type de méthodes, à déterminer le pas optimal pour la maximisation de la fonctionnelle considérée et à étudier l'application des algorithmes ainsi déterminés au traitement d'antennes, et plus particulièrement à la méthode de Pisarenko.

Par ailleurs, de nombreuses extensions de ce type d'algorithmes sont possibles. On peut ainsi imposer que M soit une matrice définie positive, ce qui est d'un intérêt certain pour l'utilisation de tests d'ordre (du type Akaike) ou de sphéricité pour la détermination du nombre de sources, ou pour l'estimation de la densité spectrale de puissance du bruit ambiant. Cette méthode peut également s'appliquer à l'estimation des gains d'une antenne afin d'apporter les corrections nécessaires.

Dans les simulations, on s'attachera essentiellement à comparer l'influence de l'estimation de la matrice interspectrale pour la méthode de Pisarenko appliquée au traitement spatial. Rappelons que ce type de traitement d'antenne a été étudié par Mermoz et Bienvenu [7; 2] et qu'il conduit à une description globale du champ des bruiteurs en considérant que tous les bruiteurs jouent un rôle symétrique. On est alors amené à déterminer le sous-espace bruit et le sous-espace des sources, les positions des sources s'obtenant par projection des vecteurs directions sur le sous-espace bruit; les maximums de l'inverse de cette projection correspondent aux positions estimées des sources.

2. Modèles et présentation du problème

Supposons que R soit la matrice interspectrale estimée, on veut déterminer la matrice $M(\varepsilon, \mathcal{F})$ la plus proche de R au sens d'un certain critère de contraste.

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

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

On aboutit au résultat remarquable (par sa simplicité) : M est la matrice de Toeplitz obtenue par moyenne des diagonales de R , soit :

$$(2) \quad M(1, k) = \frac{1}{n+1-k} \left[\sum_{i=1}^{n+1-k} R(i, i+k-1) \right].$$

Ce résultat s'obtenant très simplement par orthonormalisation de la base orthogonale $\{U_i\}_{i=1}^n$ de \mathcal{F} (on notera cette matrice M_p). (Le résultat est identique dans le cas complexe et est obtenu par le même raisonnement.) M_p est l'estimée au sens des moindres carrés.

Supposons que l'observation soit un vecteur gaussien X de covariance (matrice interspectrale) M , alors sa densité de probabilité s'écrit :

$$(3) \quad P_M(X) = \pi^{-n} (\det M)^{-1} \exp(-X^* \cdot M^{-1} \cdot X).$$

Si, de plus, on dispose de m vecteurs X_1, X_2, \dots, X_m supposés indépendants, alors :

$$(4) \quad P_M(X_1, X_2, \dots, X_m) = \pi^{-nm} (\det M)^{-m} \times \exp\left(-\sum_{i=1}^m X_i^* \cdot M^{-1} \cdot X_i\right).$$

Cette hypothèse est assez réaliste si la durée d'estimation (TF) des vecteurs $X(v)$ est assez longue, elle a été vérifiée en simulation et sur données réelles.

L'estimateur du maximum de vraisemblance de M connaissant $\{X_1, \dots, X_m\}$ maximise :

$$(5) \quad \begin{aligned} \text{Log } P_M(X_1, \dots, X_m) &= -m \text{Log}(\det M) - \sum_{i=1}^m X_i \cdot M^{-1} \cdot X_i \\ &= m \left(-\text{Log} \det M - \text{tr} \left(M^{-1} \cdot \frac{1}{m} \sum_{i=1}^m X_i \cdot X_i^* \right) \right) \end{aligned}$$

((puisque $X^* \cdot A \cdot X = \text{tr}(A \cdot X \cdot X^*)$))

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

avec :

$$(6) \quad R = \frac{1}{m} \cdot \sum_{i=1}^m X_i \cdot X_i^*.$$

Le problème sera donc la minimisation de la fonctionnelle (de vraisemblance) :

$$(7) \quad f(R, M) = \text{Log} \det M + \text{tr}(M^{-1} \cdot R).$$

M peut s'écrire :

$$M = \sum_{i=1}^n \alpha_i \cdot U_i \quad (\text{dans le cas réel})$$

et

$$f(R, M) = f(\alpha_1, \dots, \alpha_n, R).$$

Nous allons étudier la minimisation de la fonctionnelle $f(R, M)$ dans les paragraphes suivants. Les buts principaux étant de démontrer la possibilité d'utiliser des méthodes itératives pour la minimisation de $f(R, M)$ (étude de la convergence des algorithmes), de déterminer les algorithmes à pas optimal assurant la convergence, et enfin de comparer (à m fixé) les résultats obtenus par l'estimateur du maximum de vraisemblance avec ceux obtenus par l'estimateur des moindres carrés.

3. Convergence de méthodes itératives

Le but est ici d'étudier la convergence des méthodes itératives pour la minimisation de $f(R, M)$. On déterminera une condition de convergence (prop. 1, 2, 3) étendue par les annexes 1 et 2.

3.1. ÉTUDE DE LA CONVEXITÉ ET DE L'ELLIPTICITÉ DANS LE CAS GÉNÉRAL

On se place (pour l'instant) dans \mathcal{M}_n , M n'a donc pas de structure imposée, alors :

$$(8) \quad \nabla f_M(U) = \text{tr}(M^{-1} \cdot U) - \text{tr}(M^{-1} \cdot U \cdot M^{-1} \cdot R).$$

En effet :

$$\bar{V} \det M(U) = \det M \cdot \text{tr}(M^{-1} \cdot U)$$

et

$$\nabla^2 f_M(U, U) = -\text{tr}(M^{-1} \cdot U \cdot M^{-1} \cdot U) + 2 \text{tr}(M^{-1} \cdot U \cdot M^{-1} \cdot U \cdot M^{-1} \cdot R).$$

soit encore, en regroupant les termes :

$$(9) \quad \nabla^2 f_M(U, U) = \text{tr}[M^{-1} \cdot (2R - M) \cdot M^{-1} \cdot (U \cdot M^{-1} \cdot U)],$$

étudions le signe de $\nabla^2 f_M(U, U)$:

Supposons $(2R - M)$ positive, M étant positive, on constate que $M^{-1} \cdot (2R - M) \cdot M^{-1}$ est encore positive. Utilisons les décompositions de Choleski de $M^{-1} \cdot (2R - M) \cdot M^{-1}$ et de M^{-1} , soit :

$$(10) \quad \begin{cases} M^{-1} \cdot (2R - M) \cdot M^{-1} = T \cdot T^*, \\ T \text{ et } Q \text{ triangulaires,} \\ M^{-1} = Q \cdot Q^* \end{cases}$$

$$\Rightarrow \nabla^2 f_M(U, U) = \text{tr}(T \cdot T^* \cdot U \cdot Q \cdot Q^* \cdot U)$$

$$= \text{tr}(T(T^* \cdot U \cdot Q) \cdot Q^* \cdot U)$$

$$= \text{tr}((T^* \cdot U \cdot Q)(Q^* \cdot U T))$$

$$= \|T^* \cdot U \cdot Q\|_H^2 > 0,$$

donc, si on suppose $2R - M$ positive, $\nabla^2 f_M(U, U)$ est positif quel que soit U , on peut donc conclure :

Propriété 1 : La fonction de vraisemblance $f(R, M)$ est convexe sur l'ensemble (convexe) des matrices M hermitiennes telles que $(2R - M)$ soit positive.

L'ellipticité de $f(M, R)$ s'étudie de la même façon.

Posons :

$$A = M^{-1} \cdot (2R - M) \cdot M^{-1}.$$

Supposons $(2R - M)$ positive, alors A est positive, et soit λ_{\min} sa plus petite valeur propre ($\lambda_{\min} > 0$).

Soit : $0 < \lambda < \lambda_{\min}$.

Alors : $A - \lambda \text{Id}$ est positive ($\lambda > 0$).

De même, il existe μ (inférieur à la plus petite valeur propre de M^{-1}), tel que : $M^{-1} - \mu \text{Id}$ soit positive ($\mu > 0$).

Alors :

$$(11) \quad \nabla^2 f_M(U, U) = \text{tr}(A \cdot U \cdot M^{-1} \cdot U)$$

$$= \text{tr}[(A - \lambda \text{Id} + \lambda \text{Id}) \cdot U (M^{-1} - \mu \text{Id} + \mu \text{Id}) \cdot U]$$

$$= \text{tr}[(A - \lambda \text{Id}) \cdot U \cdot (M^{-1} - \mu \text{Id}) \cdot U]$$

$$+ \lambda \text{tr}[U (A - \lambda \text{Id}) \cdot U]$$

$$+ \mu \text{tr}[U (M^{-1} - \mu \text{Id}) \cdot U] + \lambda \mu \text{tr}[U \cdot U].$$

Or $\text{tr}[(A - \lambda \text{Id}) \cdot U \cdot (M^{-1} - \mu \text{Id}) \cdot U]$ est positif, car $(A - \lambda \text{Id})$ et $(M^{-1} - \mu \text{Id})$ sont positives (utiliser leurs décompositions de Choleski).

De même, on montre que :

$$\text{tr}[U \cdot (A - \lambda \text{Id}) \cdot U] \quad \text{et} \quad \text{tr}[U \cdot (M^{-1} - \mu \text{Id}) \cdot U]$$

sont positifs, donc :

$$\nabla^2 f_M(U, U) \geq \lambda \mu \|U\|^2,$$

d'où :

Propriété 2 : La fonction de vraisemblance $f(R, M)$ est elliptique sur l'ensemble des matrices M hermitiennes telles que $(2R - M)$ soit positive.

Donc sur cet ensemble convexe, les méthodes de gradient à pas optimal convergent [4].

3.2. ÉTUDE DE LA CONVEXITÉ ET DE L'ELLIPTICITÉ DANS LE CAS PARAMÉTRIQUE

Maintenant, on considère que M a une structure imposée, par exemple :

$$(12) \quad M = \sum_{i=1}^n \alpha_i \cdot U_i.$$

Alors :

$$\frac{\partial^2 f}{\partial \alpha_i \partial \alpha_j}(M, R) = \text{tr}(M^{-1} \cdot (2R - M) \cdot M^{-1} \cdot U_j \cdot M^{-1} \cdot U_i).$$

Considérons la matrice (hessienne) de f de terme général :

$$(13) \quad H(i, j) = \frac{\partial^2 f}{\partial \alpha_i \partial \alpha_j}(M, R).$$

Soit X un vecteur quelconque :

$$X^t = (x_1, \dots, x_n),$$

$$(14) \quad X^* \cdot H \cdot X = \sum_{i,j=1}^n X_i^* \cdot H(i, j) \cdot X_j$$

$$= \text{tr}[M^{-1} \cdot (2R - M) \cdot M^{-1} \cdot (\sum_i x_i^* U_i^*) \cdot M^{-1} \cdot (\sum_j x_j U_j)].$$

Par le même raisonnement qu'au paragraphe 3.1, on voit que :

Propriété 3 : $f(R, M)$ est convexe et elliptique sur l'ensemble des matrices de Toeplitz, telles que $2R - M$ soit positive.

Sur cet ensemble (convexe) de \mathcal{S} , les méthodes de gradient à pas optimal convergent. Toutefois, ce sous-ensemble peut être restrictif, aussi peut-on essayer de déterminer des conditions de convergences plus générales.

3.3. EXTENSION DES CONDITIONS DE CONVERGENCE

On trouvera en annexes 1 et 2 un exposé de conditions plus générales permettant de garantir la convergence de la méthode de gradient à pas optimal. Dans la pratique, on a pu vérifier que l'algorithme de gradient converge, bien que la condition $2R - M$ positive ne soit pas vérifiée. Cette condition étant suffisante mais non nécessaire. Sur les données simulées aussi bien que sur les données réelles l'algorithme de gradient à pas optimal a toujours convergé.

3. 4. CONSÉQUENCES D'UNE INÉGALITÉ FONDAMENTALE

L'inégalité des moyennes arithmético-géométrique établit que :

$$(15) \quad \frac{1}{n} \cdot \left(\sum_{i=1}^n \lambda_i \right) \geq \left(\prod_{i=1}^n \lambda_i \right)^{1/n} \quad (\lambda_i \geq 0, \forall i),$$

ce qui est conséquence de la convexité de l'exponentielle, l'égalité n'ayant lieu que :

$$\lambda_1 = \dots = \lambda_n = \alpha \quad (\alpha \geq 0).$$

On en déduit la propriété 4.

Propriété 4 : *M* et *R* étant deux matrices hermitiennes positives, on a l'inégalité :

$$\text{tr} [M^{-1} \cdot R] \geq n [\det (M^{-1} R)]^{1/n},$$

l'égalité n'ayant lieu que si :

$$M = \alpha \cdot R \quad (\alpha > 0).$$

En effet, *R* étant positive, on peut la décomposer en facteurs triangulaires, soit :

$$R = T \cdot T^* \quad \Rightarrow \quad \begin{cases} \text{tr} (M^{-1} \cdot R) = \text{tr} (T^* \cdot M^{-1} \cdot T), \\ \det (M^{-1} \cdot R) = \det (T^* \cdot M^{-1} \cdot T), \end{cases}$$

de plus $T^* \cdot M^{-1} \cdot T$ est hermitienne et positive (puisque *M* est positive), l'inégalité résulte donc de l'application de l'inégalité des moyennes arithmético-géométrique.

Supposons maintenant que l'égalité ait lieu et donc que $M^{-1} \cdot R$ ait toutes ses valeurs propres confondues.

On peut remarquer que les valeurs propres de $M^{-1} \cdot R$ sont identiques à celles de $T^* \cdot M^{-1} \cdot T$ puisque :

$$\begin{aligned} \det (M^{-1} \cdot R - \lambda \text{Id}) &= \det (M^{-1} \cdot T \cdot T^* - \lambda T^* \cdot T) \\ &= \det ((M^{-1} T - \lambda T^* \cdot T) T^*) \\ &= \det (T^* (M^{-1} \cdot T - \lambda T^* \cdot T)) \\ &= \det (T^* \cdot M^{-1} \cdot T - \lambda \text{Id}). \end{aligned}$$

Alors $T^* \cdot M^{-1} \cdot T$ a également toutes ses valeurs propres confondues.

Mais $T^* \cdot M^{-1} \cdot T$ étant hermitienne, elle est diagonalisable, donc il existe $U (\in \mathcal{U}_n)$ telle que :

$$U^* (T^* \cdot M^{-1} \cdot T) \cdot U = \alpha \text{Id} \quad (\alpha > 0),$$

d'où :

$$T^* \cdot M^{-1} \cdot T = \alpha U \cdot U^* = \alpha \text{Id}$$

et

$$T^* \cdot M^{-1} \cdot T \cdot T^* = \alpha T^* \quad \Rightarrow \quad M^{-1} \cdot R = \alpha \text{Id}$$

et donc :

$$(16) \quad M = \alpha \cdot R.$$

Application : Supposons que l'on veuille minimiser Log det M sous la contrainte $\text{tr}(M^{-1} \cdot R) = n$.

λ étant un multiplicateur de Lagrange, on obtient la condition nécessaire :

$$\langle \nabla (\text{Log det M} + \lambda \text{tr} (M^{-1} R)), U \rangle = 0, \quad \forall U \in \mathcal{M}_n,$$

soit :

$$\text{tr} [(M^{-1} - \lambda M^{-1} \cdot R \cdot M^{-1}) U] = 0, \quad \forall U \in \mathcal{M}_n,$$

en particulier si : $U = M$ on obtient :

$$(17) \quad \lambda = \frac{n}{\text{tr}(M^{-1} \cdot R)} = 1.$$

Donc minimiser Log det M sous la contrainte $\text{tr}(M^{-1} \cdot R) = n$ revient à dire que *M* est un extrémum de $f(R, M)$, et donc dans le cas général (sans structure imposée), le minimum de $f(R, M)$ est atteint en $M = R$. Dans le cas où *M* a une forme paramétrique, l'optimalité est seulement asymptotique (i. e. si $M \rightarrow R$).

Une fois étudiée la convergence des méthodes itératives, on va donc définir un algorithme de gradient à pas optimal dont les conditions de convergence seront donc déterminées.

4. Description d'un algorithme de gradient à pas optimal

Le calcul du vecteur gradient ne pose pas de difficultés, on a :

$$(18) \quad \begin{aligned} \nabla f(R, M_k) &= \text{tr} [M_k^{-1} \cdot (M_k - R) \cdot M_k^{-1} \cdot U_i] = d_k^i \\ \text{tr} [M_k^{-1} \cdot (M_k - R) \cdot M_k^{-1} \cdot U_n] &= d_k^n, \end{aligned}$$

où :

$$M_k = \sum x_i^k \cdot U_i.$$

Un algorithme de gradient consiste donc à écrire :

$$(19) \quad X_{k+1} = X_k - \rho_k \nabla f(R, M_k).$$

5. 1. DÉTERMINATION DU PAS OPTIMAL

$\nabla f(R, M_k)$ ayant été calculé, on cherche à déterminer ρ qui minimise :

$$f(R, M_{k+1}) = f(R, M_k - \rho D_k),$$

où :

$$D_k = \sum_{i=1}^n d_i^k \cdot U_i.$$

Pour cela, on va obtenir une expression explicite $f(R, M_k - \rho D_k)$ (en ρ).

(a) Calcul du déterminant :

$$\det(M_{k+1}) = \det(M_k - \rho D_k) = \det[M_k (\text{Id} - \rho M_k^{-1} \cdot D_k)],$$

or :

$$M_k = T_k \cdot T_k^* \quad (T_k \text{ triangulaire})$$

M_k étant positive par hypothèse

$$\Rightarrow M_k^{-1} = S_k^* \cdot S_k$$

(avec $S_k = T_k^{-1}$, S_k est donc triangulaire)

$$\begin{aligned} &\Rightarrow \det(M_{k+1}) \\ &= \det M_k \cdot \det(\text{Id} - \rho S_k^* S_k \cdot D_k) \\ &= \det M_k \cdot \det(\text{Id} - \rho S_k D_k \cdot S_k^*), \end{aligned}$$

or : $S_k \cdot D_k \cdot S_k^*$ est hermitienne, donc diagonalisable,

$$\left\{ \begin{aligned} &\Rightarrow S_k \cdot D_k \cdot S_k^* = U \cdot \Delta_k \cdot U^*, \quad U \in U_n, \\ &\Rightarrow \det(M_{k+1}) = \det M_k \prod_{i=1}^n (1 - \rho \lambda_i^k), \end{aligned} \right.$$

où les $\{\lambda_i^k\}$ sont les valeurs propres de $S_k \cdot D_k \cdot S_k^*$ (ou celles de $M_k^{-1} \cdot D_k$).

(b) Calcul de la trace :

$$\begin{aligned} \text{tr}[M_{k+1}^{-1} \cdot R] &= \text{tr}[(T_k \cdot T_k^* - \rho D_k)^{-1} \cdot R] \\ &= \text{tr}[T_k^{*-1} \cdot (\text{Id} - \rho S_k \cdot D_k \cdot S_k^*)^{-1} \cdot T_k^{-1} \cdot R], \end{aligned}$$

or :

$$\begin{aligned} S_k \cdot D_k \cdot S_k^* &= \sum_{i=1}^n \lambda_i^k V_i \cdot V_i^* \\ &\quad (V_i \text{ vecteurs propres}) \\ &\quad (V_i \perp V_j, i \neq j \text{ et } \|V_i\| = 1) \\ &\Rightarrow \text{tr}[M_{k+1}^{-1} \cdot R] = \sum_{i=1}^n (1 - \rho \lambda_i^k)^{-1} \\ &\quad \times V_i^* T_k^{-1} \cdot R \cdot T_k^{-1*} \cdot V_i, \end{aligned}$$

d'où finalement :

$$(21) \quad \boxed{\begin{aligned} f(R, M_{k+1}) &= \text{Log} \det M_k + \sum_{i=1}^n \text{Log}(1 - \rho \lambda_i^k) \\ &\quad + \sum_{i=1}^n (1 - \rho \lambda_i^k)^{-1} \times V_i^* \cdot T_k^{-1} \cdot R \cdot T_k^{-1*} \cdot V_i. \end{aligned}}$$

$f(R, M_{k+1})$ est ainsi exprimé comme une fonction explicite du paramètre ρ .

Pour minimiser cette fonction, on utilise la méthode de Newton initialisée par une approximation quadratique de f , soit :

$$(22) \quad \rho_0 = \frac{\text{tr}(M_k^{-1} \cdot (M_k - R) M_k^{-1} \cdot D_k)}{2 \text{tr}[M_k^{-1} \cdot (2R + M_k) \cdot M_k^{-1} \cdot D_k M_k^{-1} D_k]}.$$

4. 2. UTILISATION PRATIQUE DE L'ALGORITHME

Dans la pratique, on utilise l'algorithme de gradient à pas optimal durant les premières itérations (typiquement, une dizaine), puis un algorithme de gradients conjugués par la suite, le pas optimal étant calculé par la méthode exposée en paragraphe 4. 1. (La détermination du pas optimal garantissant la convergence des algorithmes de gradient sous les conditions définies au paragraphe 3.)

4. 3. OPÉRATEUR DE PROJECTION ET CONVERGENCE

On doit (pour que la matrice M_k soit positive) considérer le problème suivant :

minimisation de $f(R, M)$ sous la contrainte M positive. On définira alors un algorithme de gradient projeté par :

$$(23) \quad M_{k+1} = P_H(M_k - \rho_k D_k).$$

H étant une partie convexe et fermée de l'espace vectoriel topologique des matrices de forme paramétrique donnée. Examinons les implications pratiques de cet algorithme.

On fait l'hypothèse de récurrence que M_k est positive : or : $M_k - \rho D_k$ a mêmes valeurs propres que :

$$T_k^* (\text{Id} - \rho M_k^{-1} D_k) \cdot T_k,$$

donc :

$(M_k - \rho D_k)$ positive

$$\Leftrightarrow \text{Id} - \rho T_k^{-1} \cdot D_k \cdot T_k^{*-1} \text{ positive,}$$

$$(M_k - \rho D_k) \text{ positive} \Leftrightarrow 0 < \rho < \frac{1}{\lambda_i^k},$$

$$i = 1, 2, \dots, n.$$

P_H est donc particulièrement simple, puisque la condition nécessaire est suffisante pour que M_{k+1} soit positive, est que :

$$(24) \quad 0 < \rho_k < \frac{1}{\max(\lambda_i^k)}.$$

5. Mise en œuvre de l'algorithme

Après avoir estimé une matrice interspectrale R (par la méthode du périodogramme), on veut déterminer la matrice M de Toeplitz et positive maximisant $f(M, R)$.

M est alors initialisée à Id ($M_0 = \text{Id}$) ou par M_p , l'algorithme utilisé est identique à celui décrit au paragraphe 4. R et M étant complexes, on écrit $M = M_1 + i \cdot M_2$, M_1 et M_2 réelles.

La donnée de M équivaut à celle de la matrice $(2n, 2n)$

réelle et symétrique $\begin{pmatrix} M_1 & -M_2 \\ M_2 & M_1 \end{pmatrix}$. Il en est de même

pour les bases considérées.

Ceci est pénalisant pour le temps de calcul, par contre le passage au cas complexe en dimension

$$n \left(M = \sum_{i=1}^n \alpha_i U_i + \sum_{i=1}^n \bar{\alpha}_i U_i^* \quad U_i \text{ triangulaire supérieure} \right)$$

n'a pas donné de bons résultats (bien que les deux approches paraissent identiques!).

La convergence est assez rapide (test d'arrêt :

$$\frac{f(M_k) - f(M_{k+1})}{f(M_k)} < \varepsilon,$$

typiquement de l'ordre de quelques dizaines d'itérations, bien qu'elle dépende grandement des données. En initialisant l'algorithme avec M_p , la convergence est plus rapide qu'en initialisant avec la matrice identité.

La matrice finale (M_f) est alors utilisée dans une application de traitement d'antennes afin de caractériser les paramètres des bruiteurs.

Le calcul est sensiblement accéléré par l'emploi de méthodes de calculs spécifiques aux matrices de Toeplitz. Ainsi les déterminants des matrices de Toeplitz peuvent être calculées par un algorithme déduit de l'algorithme de Durbin-Levinson [1].

6. Extensions de ce type de méthodes

On a étudié le cas où les signaux reçus sur l'antenne sont affectés de gains (constants), mais inconnus. Soit Δ (diagonale) la matrice des gains des capteurs. On a :

$$(25) \quad R = E(\Delta \cdot XX^* \Delta^*) = \Delta \cdot R_0 \cdot \Delta^*,$$

on modélisera donc M par :

$$(26) \quad M = \Delta \cdot M_0 \cdot \Delta^*.$$

M_0 étant une matrice de Toeplitz et Δ étant diagonale inconnue.

Le but de l'algorithme est alors d'estimer les coefficients α_i de M_0 ($M_0 = \sum_{i=1}^n \alpha_i \cdot U_i$) et ceux de $\Delta(\delta_1, \dots, \delta_n)$.

L'algorithme précédent peut être transposé à ce cas, toutefois il n'existe pas de condition simple de convergence pour l'estimation simultanée des $\{\alpha_i\}$ et des $\{\delta_i\}$.

Une procédure globale de minimisation par rapport aux paramètres $\{\alpha_i, \delta_i\}$ ne donne pas de bons résultats sur des simulations (l'algorithme ne pouvant calculer le pas optimal). Par contre, la procédure (sous-optimale!) consistant à minimiser par rapport aux $\{\delta_i\}$, puis par rapport aux $\{\alpha_i\}$ a donné de meilleurs résultats.

7. Présentation de résultats de simulations

On a étudié l'influence de cet algorithme sur les résultats obtenus par une méthode de traitement d'antennes, en l'occurrence la méthode de Pisarenko [7].

La méthode de Pisarenko est appliquée concurremment à M_p et à M_f , R était une matrice simulée par périodogramme (le nombre de moyennages, les positions des

sources, leurs puissances et celle du bruit étant indiquée sur les figures correspondantes) :

$$(26) \quad R = \frac{1}{N} \cdot \sum_{i=1}^N X_i \cdot X_i^*.$$

X étant un vecteur gaussien de matrice de covariance exacte :

$$(27) \quad T = \sum_{i=1}^n \sigma_i^2 \mathbf{D} \theta_i \cdot \mathbf{D} \theta_i^* + b^2 \text{Id},$$

$$(28) \quad \mathbf{X} = \sum_{j=1}^n \alpha_j \cdot \mathbf{D} \theta_j + \mathbf{B}.$$

Les $\{\alpha_j\}$ étant des v. a. gaussiennes centrées indépendantes et \mathbf{B} un vecteur gaussien; et $\mathbf{D} \theta_j$ étant le vecteur d'exponentielles complexes associé à la source j .

$\{\sigma_i^2, b^2\}$ étant les puissances des sources et du bruit, $\{\theta_i\}$ étant les gisements des sources. Ces paramètres sont indiqués sur chaque figure. La courbe en traits pointillés correspond à l'application de la méthode de Pisarenko à M_f , alors que la courbe en traits pleins correspond à l'application de cette méthode à M_p . Dans tous les cas, on considère qu'il y a 16 capteurs équirépartis (à une demi-longueur d'onde) sur une antenne linéaire.

Toutes les simulations correspondent à des sources indépendantes, parfaitement cohérentes. Le nombre de moyennage détermine ici la variance des éléments de la matrice R . L'axe horizontal représente les gisements (en degrés) et l'axe vertical la réponse de la méthode de Pisarenko (dans une échelle linéaire).

Sur la figure 1, on peut constater l'amélioration apportée par l'utilisation de l'algorithme pour la détection de la source faible (dans le gisement 74) et pour la séparation des deux sources fortes (dans les gisements 57 et 60).

Cette conclusion est également illustrée par les figures 2 et 3.

On présente sur les figures 4 et 5 les résultats obtenus sur plusieurs matrices R (15) aléatoires. On peut constater que les sources de gisements 53 et 62 sont mieux séparées et que la source faible apparaît plus fortement. L'algorithme de maximisation de la vraisemblance étant initialisé, on a pu constater que dans tous ces cas la fonction de Log-vraisemblance augmentait sensiblement durant les premières itérations et convergait au bout d'une trentaine d'itérations. C'est de là que résulte l'amélioration apportée à la méthode de Pisarenko.

(La mention « moyennage » portée sur les figures correspond à l'estimation de M par moyennage des diagonales de R , la mention « CBURLU » correspond à l'estimation de M par la méthode itérative du maximum de vraisemblance.)

Sur les figures 6 et 7 sont présentés les résultats obtenus par le même type de simulation que précédemment. Les deux sources fortes apparaissent mieux séparées et la source faible apparaît plus nettement; dans ce cas, l'amélioration tient beaucoup à la diminution de la variance d'estimation des gisements des bruiteurs. Ceci est illustré par les figures 6 bis et 7 bis (le biais et la variance d'estimation des gisements sont exprimés en degrés).

ESTIMATIONS D'UNE MATRICE INTERSPECTRALE DE STRUCTURE IMPOSÉE

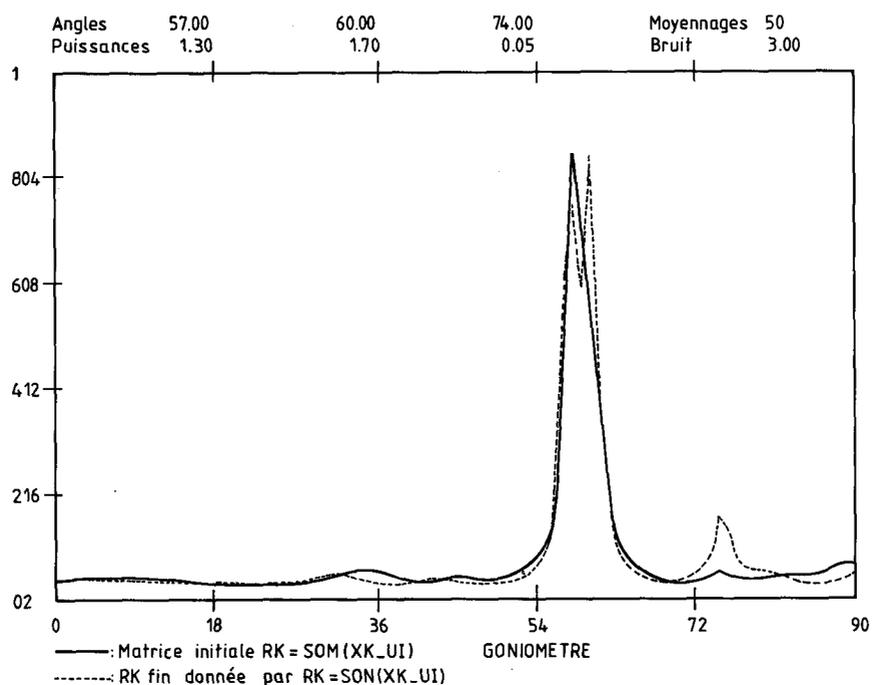


Fig. 1. — Comparaison de deux méthodes d'estimation.

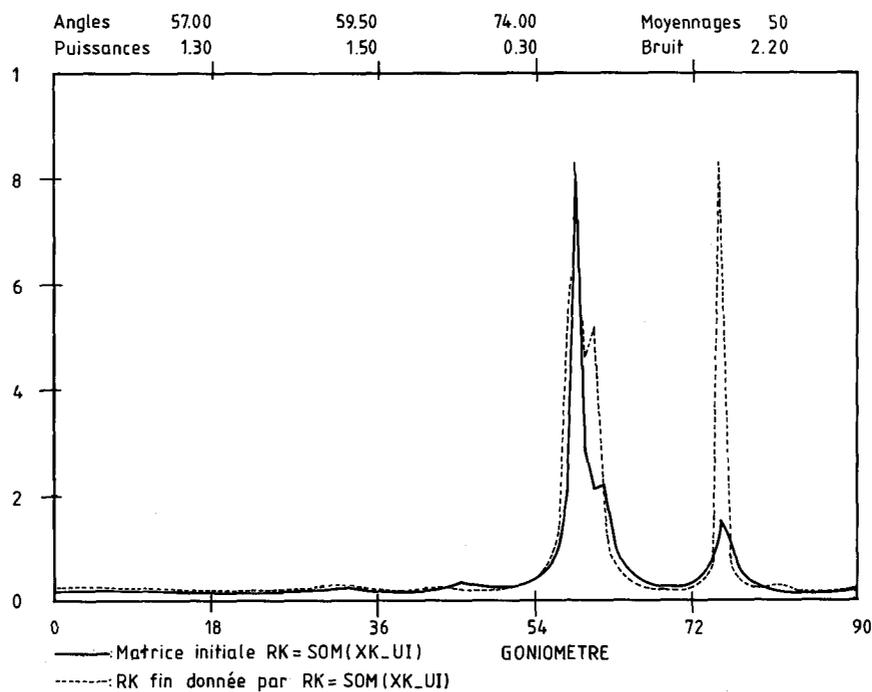


Fig. 2. — Comparaison deux deux méthodes d'estimation.

RECHERCHES

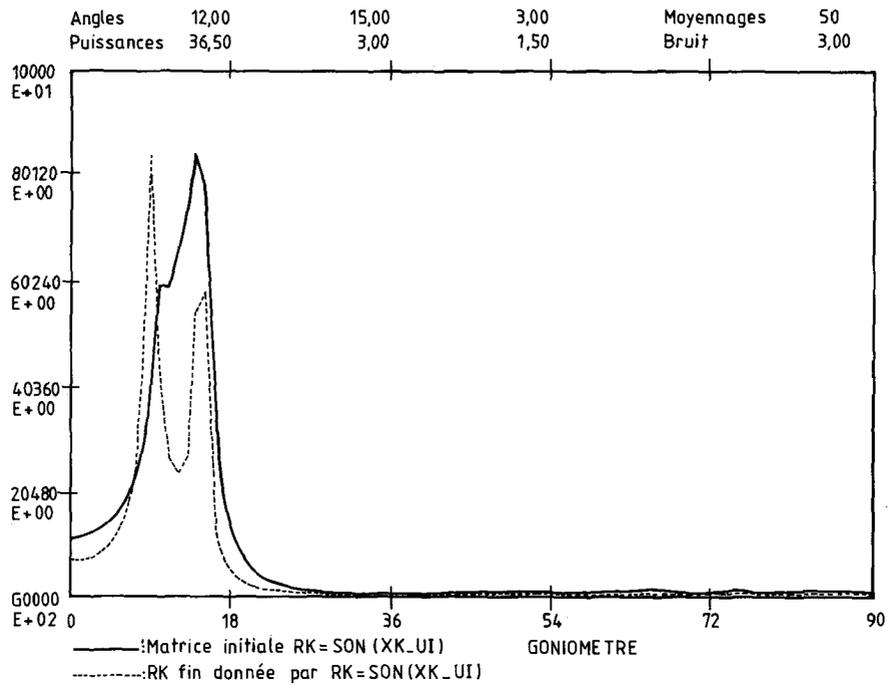


Fig. 3. — Comparaison de deux méthodes d'estimation.

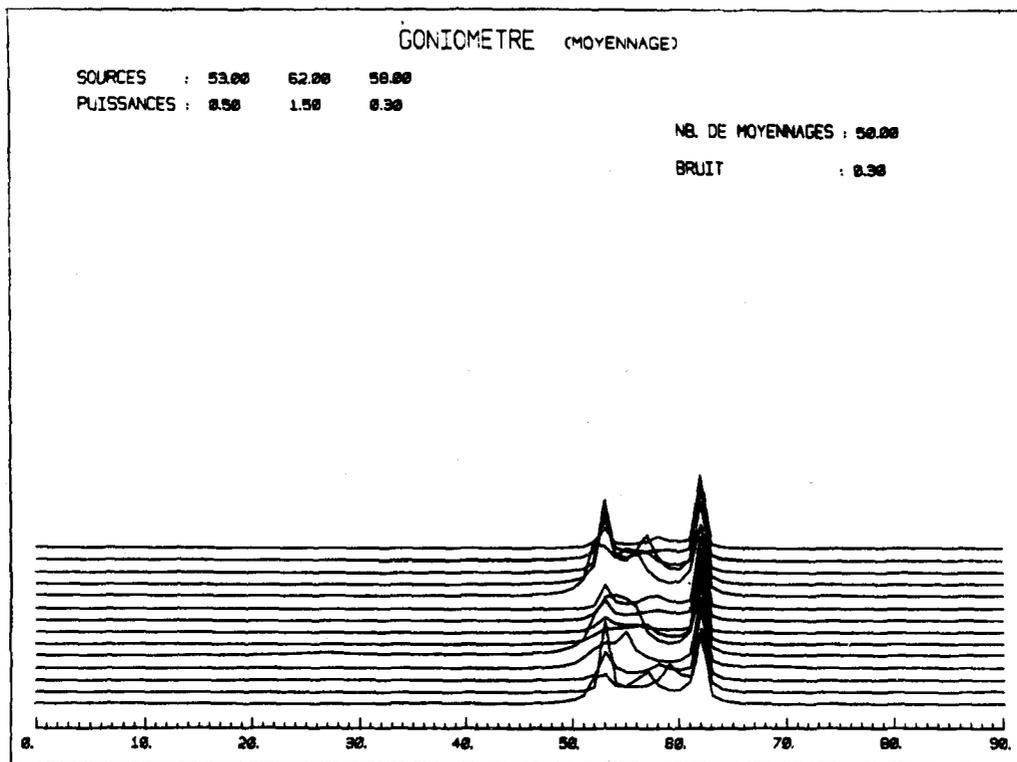


Fig. 4. — Comparaison obtenue sur plusieurs réalisations.

ESTIMATIONS D'UNE MATRICE INTERSPECTRALE DE STRUCTURE IMPOSÉE

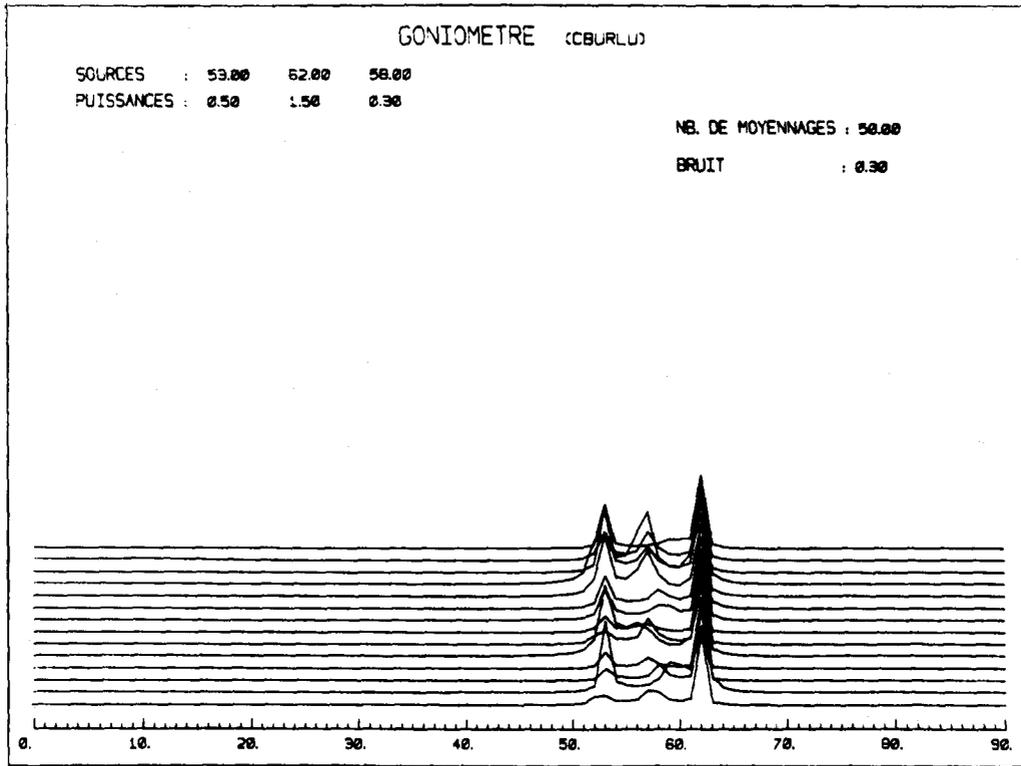


Fig. 5. - Comparaison obtenue sur plusieurs réalisations.

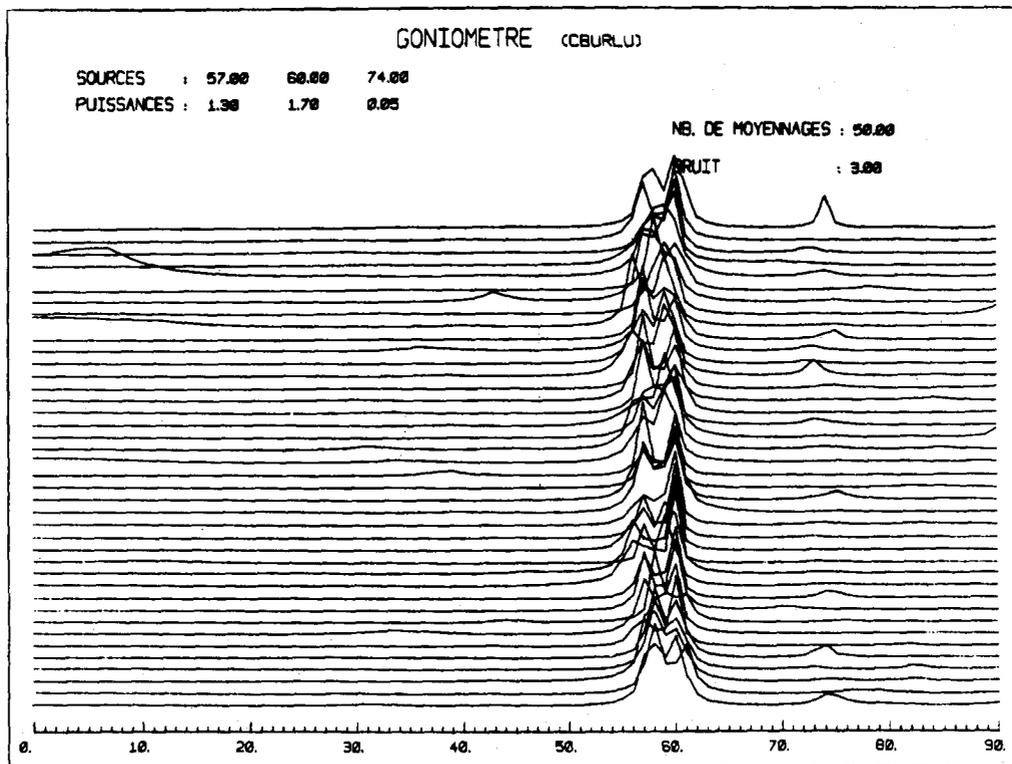


Fig. 6. - Comparaison obtenue sur plusieurs réalisations.

RECHERCHES

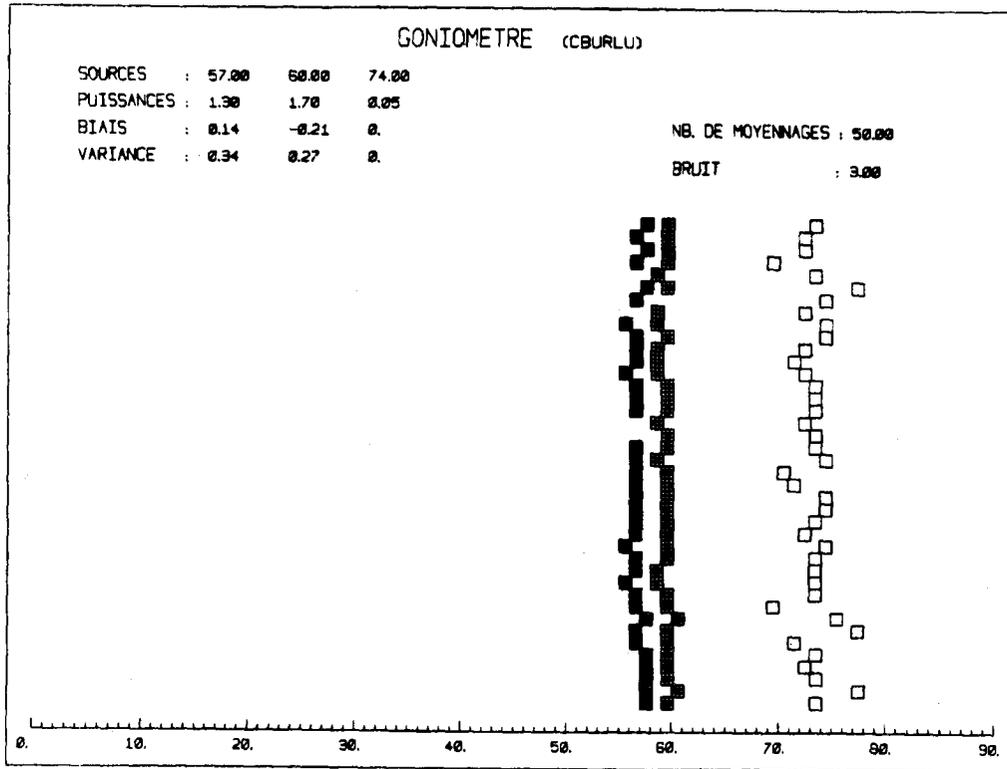


Fig. 6 bis. — Étude du biais et de la variance en gisement, relative aux deux estimateurs.

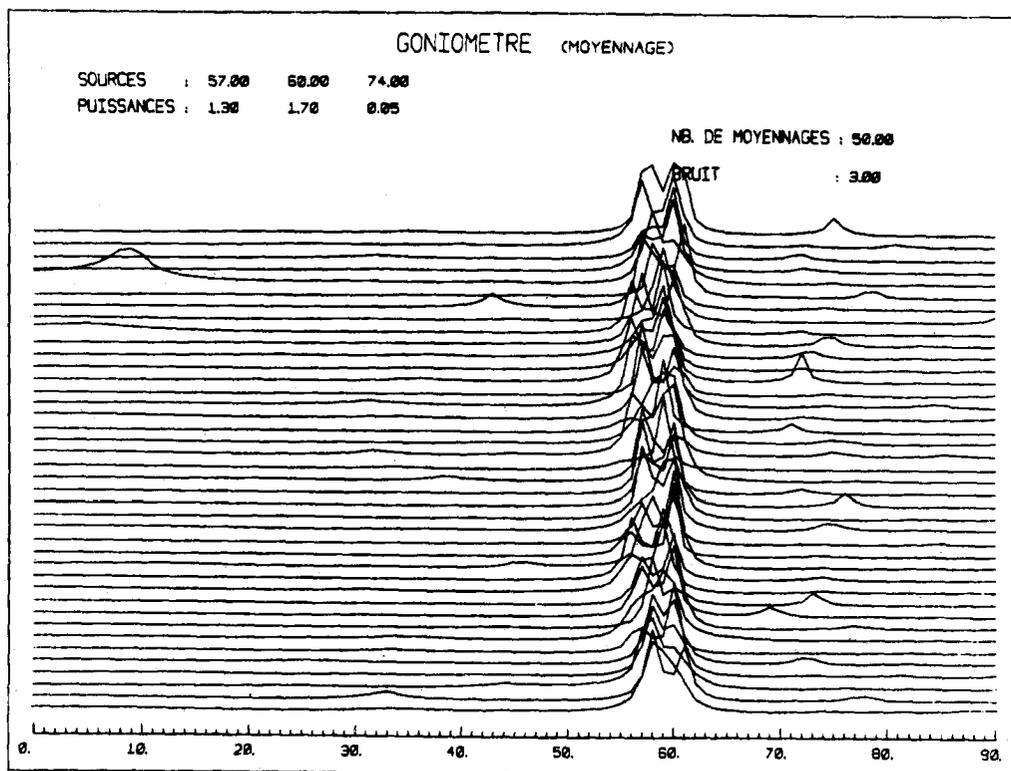


Fig. 7. — Comparaison obtenue sur plusieurs réalisations.

ESTIMATIONS D'UNE MATRICE INTERSPECTRALE DE STRUCTURE IMPOSÉE

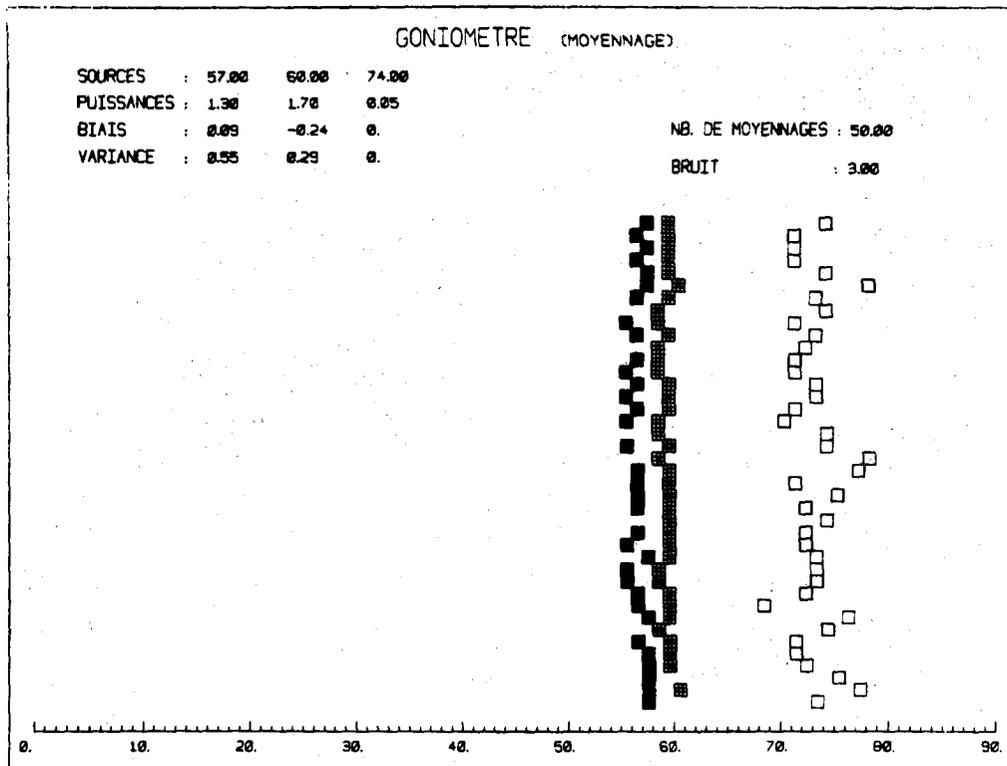


Fig. 7 bis. — Étude du biais et de la variance en gisement, relative aux deux estimateurs.

Il est pratiquement impossible de présenter un résultat quantitatif pour les probabilités de détection et de fausse alarme relatives aux deux estimations, puisque la détection d'une source s'effectue sur les figures présentées par des critères assez intuitifs (coïncidence des pics au cours du temps, importances relatives de ces pics). Pour les données expérimentales, cette tâche s'avère actuellement impossible.

8. Présentation de résultats expérimentaux

Pour cette expérimentation, on disposait d'une antenne linéaire à 16 capteurs équirépartis. On utilisait deux bruiteurs de forte puissance émettant à 800 Hz (à fréquence pure); auquel s'ajoutait le trafic habituel (bâtiments de commerce) dont une situation radar est indiquée en figure 8.

On présente sur les figures 9 et 10 une comparaison des résultats obtenus par la méthode de Pisarenko appliquée respectivement à la matrice M_p et à la matrice M_r . On peut constater une amélioration sensible de la détection des sources faibles (relativement aux deux bruiteurs à 800 Hz). Sur l'axe horizontal se trouve le gisement et sur l'axe vertical le temps, $D = 60$ dB signifie que les résultats sont présentés en échelle logarithmique avec une dynamique de 60 dB.

Ainsi, une source faible (bâtiment de commerce) défilant lentement du gisement 80 au gisement 60 est détectée sur la figure 10 pendant la première partie de l'expérimentation alors qu'elle ne l'est pas sur la figure 9. Il en est de même de plusieurs autres sources « faibles » qui apparaissent sur la figure 10 alors qu'elles n'étaient pas

visibles sur la figure 9. Il faut toutefois constater que l'aspect général de la figure 10 est plus confus que celui de la figure 9; ceci est dû en grande partie à la présentation logarithmique dont l'usage nous a cependant semblé indispensable pour comparer valablement les deux méthodes et aussi à la présence de nombreuses sources faibles.

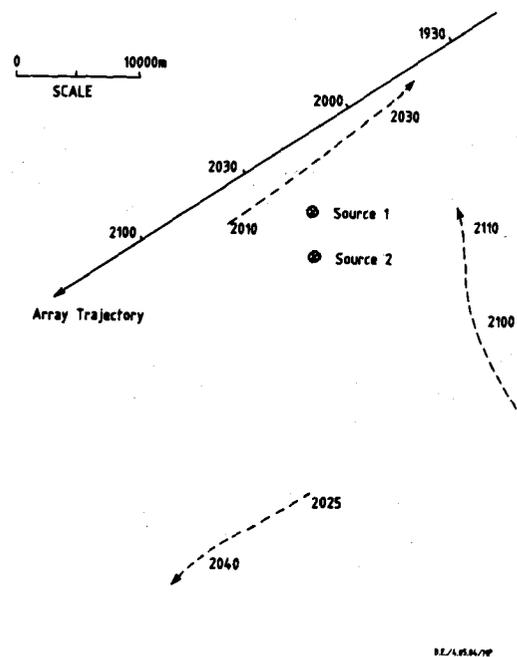


Fig. 8. — Situation radar lors de l'expérimentation en mer.

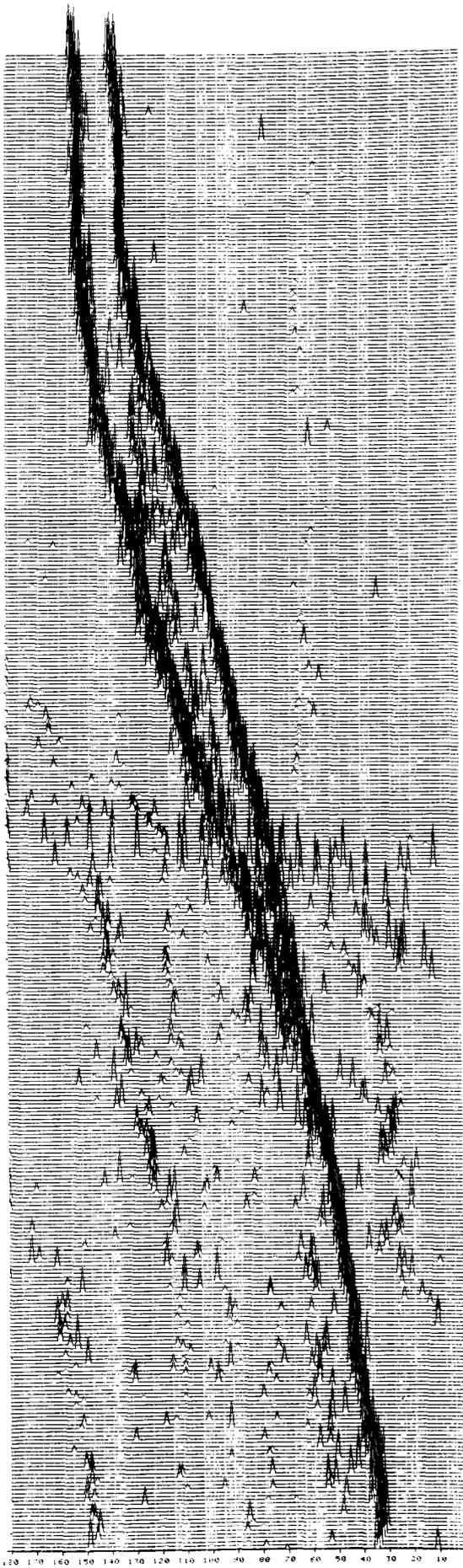


Fig. 9. — Méthode de Pisarenko appliquée à la matrice M_w .

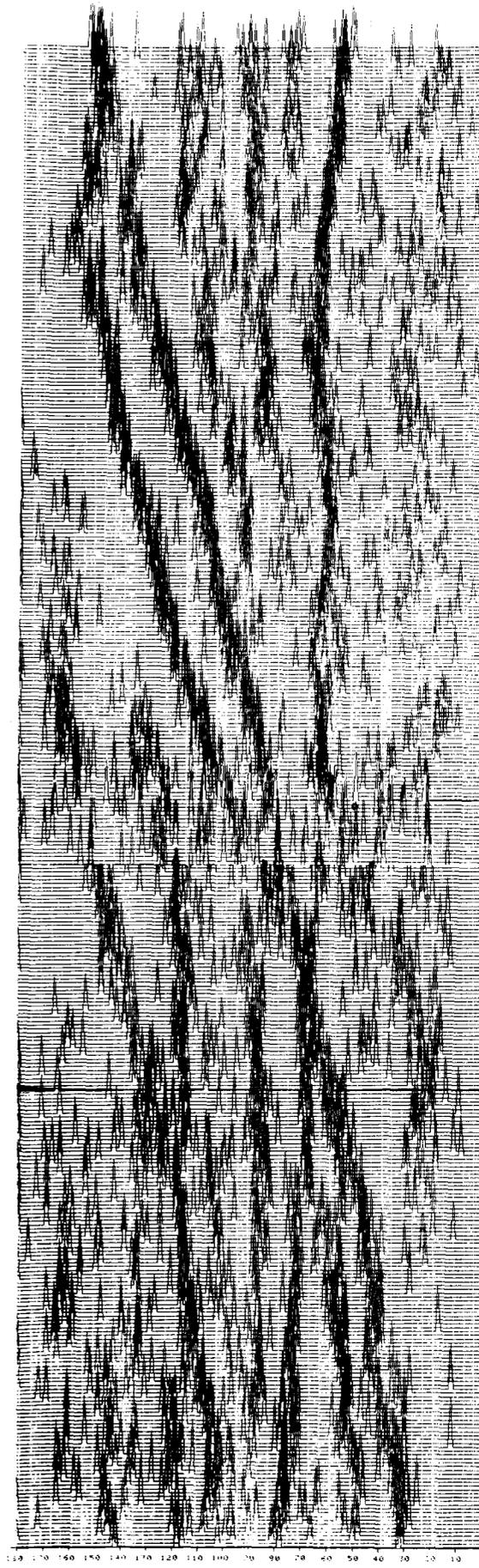


Fig. 10. — Méthode de Pisarenko appliquée à la matrice M_f .

Conclusion

L'estimation d'une matrice de covariance (ou interspectrale) par des méthodes itératives de maximisation de la vraisemblance est intéressante à plusieurs titres. D'une part elle améliore le pouvoir séparateur de la méthode de Pisarenko, d'autre part elle améliore aussi son pouvoir de détection. Cet avantage est d'autant plus marqué qu'il y a des bruiteurs de puissances très différentes et que le temps d'estimation de la matrice est faible (et la variance d'estimation grande).

A faible rapport signal sur bruit, l'avantage se porte surtout sur l'estimation du niveau du bruit ambiant, ce qui peut améliorer sensiblement l'estimation du nombre de sources.

Par contre, l'estimateur du maximum d'entropie (en traitement spatial) ne semble guère amélioré par l'usage de cet algorithme (relativement à l'estimateur des moindres carrés).

Enfin, on s'est attaché à comparer les résultats obtenus par maximisation de la vraisemblance à ceux obtenus par minimisation d'un critère quadratique, car il faut bien considérer les temps de calcul nécessaires pour chaque méthode. Si le gain obtenu par maximisation de la vraisemblance (par rapport au critère quadratique) est parfois déterminant, il est toutefois nécessaire d'étudier des méthodes rapides d'exécution de cet algorithme. Il reste aussi à étudier théoriquement l'influence de tels prétraitements sur les méthodes de traitement spatial.

Annexe 1

Le but de cette annexe est d'établir des conditions moins restrictives de convergence des algorithmes de gradient à pas optimal que la condition obtenue au paragraphe 3 (i. e. $2R - M$ positive). Pour cela, en utilisant une inégalité démontrée en annexe 2, on montre que la suite $f(M_k, R)$ est une suite décroissante minorée :

$$(28) \quad \nabla f_M(U_i) = -\text{tr} [M^{-1} \cdot (R - M) \cdot M^{-1} \cdot U_i],$$

or $(R - M)$ est diagonalisable (hermitienne), soit :

$$R - M = \sum_{j=1}^n \beta_j W_j \cdot W_j^*,$$

$$(29) \quad \nabla f_{M,R} = 0 \Leftrightarrow f_{M,R}(U_i) = 0, \quad \forall i \in \{1, \dots, n\},$$

$$\Leftrightarrow \begin{cases} \beta_1 W_1^* \cdot M^{-1} \cdot U_1 \cdot M^{-1} \cdot W_1 + \dots \\ \vdots \\ \beta_1 W_1^* \cdot M^{-1} \cdot U_n \cdot M^{-1} \cdot W_1 + \dots \\ \vdots \\ \beta_n W_n^* \cdot M^{-1} \cdot U_n \cdot M^{-1} \cdot W_n = 0. \end{cases}$$

Si on fait l'hypothèse que la matrice de terme général : $W_i^* M^{-1} \cdot U_j \cdot M^{-1} \cdot W_i$ est inversible, on déduit que :

$$\nabla f_{M,R} = 0 \Leftrightarrow \beta_1 = \beta_2 = \dots = \beta_n = 0 \Leftrightarrow R = M.$$

On peut considérer cette hypothèse vérifiée dans le cas général, il reste à établir la convergence de l'algorithme de gradient à pas optimal. On montre aisément que :

$$(30) \quad \|\nabla f_{M,R} - \nabla f_{M',R}\| \\ = \|\mathbf{M}^{-1} - \mathbf{M}'^{-1} - (\mathbf{M}^{-1} \cdot \mathbf{R} \cdot \mathbf{M}^{-1} - \mathbf{M}'^{-1} \cdot \mathbf{R} \cdot \mathbf{M}'^{-1})\| \\ \leq C \cdot \|\mathbf{M} - \mathbf{M}'\|$$

(C constante positive) (cf. Annexe 2).

Alors :

$$f(M, R) - f(M', R) \\ = \int_0^1 \langle \nabla f(M' + \tau(M - M')), M - M' \rangle d\tau \\ = \langle \nabla f(M), M - M' \rangle \\ + \int_0^1 \langle \nabla f(M' + \tau(M - M')) \\ - \nabla f(M), M - M' \rangle d\tau \geq \langle \nabla f(M), M - M' \rangle \\ - \int_0^1 \|\nabla f(M' + \tau(M - M')) - \nabla f(M)\| \\ \times \|M - M'\| d\tau \\ \geq \nabla f_M(M - M') - \frac{C}{2} \|M - M'\|^2.$$

Dans l'algorithme de gradient à pas optimal :

$$(31) \quad (\rho_k) \leq \frac{1}{C} \text{ et :} \\ f(M_k, R) - f(M_{k+1}, R) \geq \langle \nabla f(M_k), \rho_k \nabla f(M_k) \rangle \\ - (C/2) \cdot \rho_k \|\nabla f(M_k)\|^2 \\ \Rightarrow f(M_k, R) - f(M_{k+1}, R) \\ \geq \left(1 - \frac{C \cdot \rho_k}{2}\right) \cdot \rho_k \cdot \|\nabla f(M_k)\|^2.$$

La suite $\{f(M_k, R)\}$ est donc une suite décroissante minorée, il existe donc une sous-suite extraite de la suite $\{M_k\}$ qui converge vers R puisque d'après l'inégalité ci-dessus ∇f_R est nul au point de convergence de la suite M_k (noté M^*).

On voit donc qu'une algorithme de gradient à pas optimal converge sous des conditions bien moins restrictives que celles énoncées au paragraphe 3. (On peut supposer $\rho_k \leq 1/C$, puisque l'algorithme considéré est à pas optimal.)

De plus, dans un ensemble borné de diamètre n , le théorème des estimations conduit à l'inégalité :

$$f(M_k, R) - f(M^*, R) \\ \leq \mu \left[1 + \mu \sum_{i=0}^{k-1} \frac{f(M_i, R) - f(M_{i+1}, R)}{\|\nabla f(M_i, R)\|^2} \right]^{-1}.$$

Soit d'après la minoration précédente ($\rho_i \leq 1/C$) :

$$(32) \quad f(M_k, R) - f(M^*, R) \leq \mu \cdot \left[1 + \frac{\mu}{n^2} \cdot \sum_{i=0}^{k-1} \rho_i \right]^{-1},$$

où :

$$\mu = f(M_0, R) - f(M^*, R).$$

Annexe 2

Le but de cette annexe est la démonstration de l'inégalité (30).

Démonstration de l'inégalité :

$$(30) \quad \|\nabla f_{M,R} - \nabla f_{M',R}\| \leq C \|M - M'\|,$$

C constante positive,

$$\|\nabla f_{M,R} - \nabla f_{M',R}\| \leq \|M^{-1} - M'^{-1}\| + \|M^{-1} \cdot R \cdot M^{-1} - M'^{-1} \cdot R \cdot M'^{-1}\|.$$

Examinons d'abord le terme :

$$\|M^{-1} - M'^{-1}\|.$$

Posons :

$$\begin{aligned} M^{-1} \cdot M' &= \text{Id} - V \\ \Rightarrow V &= \text{Id} - M^{-1} \cdot M' = M^{-1} \cdot (M - M') \\ \Rightarrow \|V\| &\leq \|M^{-1}\| \cdot \|M - M'\|, \end{aligned}$$

donc si :

$$\|M - M'\| \leq \frac{k}{\|M^{-1}\|} \quad (k < 1)$$

alors : $\|V\| \leq 1$ (on convient que $\|\text{Id}\| = 1$) d'autre part :

$$\begin{aligned} M'^{-1} &= [M(\text{Id} - V)]^{-1} = (\text{Id} - V)^{-1} \cdot M^{-1} \\ \Rightarrow M'^{-1} - M^{-1} &= \{(\text{Id} - V)^{-1} - \text{Id}\} \cdot M^{-1}, \end{aligned}$$

or :

$$\begin{aligned} (\text{Id} - V)^{-1} &= \sum_{n \geq 1} V^n \\ \Rightarrow M'^{-1} - M^{-1} &= \left(\sum_{n \geq 1} V^n \right) \cdot M^{-1} \\ \Rightarrow \|M'^{-1} - M^{-1}\| &\leq \frac{\|V\|}{1 - \|V\|} \cdot \|M^{-1}\|, \end{aligned}$$

or :

$$\|V\| \leq \|M^{-1}\| \cdot \|M - M'\|.$$

On a donc finalement obtenu que :

$$\|M^{-1} - M'^{-1}\| \leq \left(\frac{1}{1-k} \right) \cdot \|M^{-1}\|^2 \cdot \|M - M'\|,$$

d'autre part :

$$\begin{aligned} &\|M^{-1} \cdot R \cdot M^{-1} - M'^{-1} \cdot R \cdot M'^{-1}\| \\ &= \|(MR^{-1} \cdot M)^{-1} - (M'R^{-1} \cdot M')^{-1}\| \\ &\leq k' \cdot \|M^{-1} \cdot R \cdot M^{-1}\|^2 \cdot \|M \cdot R^{-1} \cdot M - M' \cdot R^{-1} \cdot M'\|, \end{aligned}$$

d'après le calcul précédent or :

$$\begin{aligned} &M' \cdot R^{-1} \cdot M' - M \cdot R^{-1} \cdot M \\ &= (M' - M) \cdot R^{-1} \cdot M - M' \cdot R^{-1} \cdot (M - M') \\ &\Rightarrow \|M \cdot R^{-1} \cdot M' - M \cdot R^{-1} \cdot M\| \\ &\leq \|M - M'\| \cdot \|R^{-1}\| \cdot (\|M\| + \|M'\|), \end{aligned}$$

finalement :

$$\begin{aligned} &\|M^{-1} \cdot R \cdot M^{-1} - M'^{-1} \cdot R \cdot M'^{-1}\| \\ &\leq k' \cdot \|M^{-1}\|^4 \cdot \|R\|^2 \cdot \|R^{-1}\| \cdot (\|M - M'\|). \end{aligned}$$

On suppose que dans le compact étudié :

$$\|M\| < m \quad \text{et} \quad \|M^{-1}\| > m',$$

alors :

$$\|\nabla f_{M,R} - \nabla f_{M',R}\| \leq \left(\frac{m'^2}{1-k} + 2k'm \cdot m'^4 \|R\|^2 \cdot \|R^{-1}\| \right) \cdot \|M - M'\|.$$

Posons :

$$C = m'^2 + 2m \cdot m'^4 \cdot \|R\|^2 \cdot \|R^{-1}\|,$$

on a donc : sous les conditions précédentes :

$$\|\nabla f_{M,R} - \nabla f_{M',R}\| \leq C \cdot \|M - M'\|.$$

Manuscrit reçu le 25 juin 1984.

BIBLIOGRAPHIE

- [1] A. ARCESE, On the method of Maximum Entropy Spectrum Estimation, *IEEE Trans. Inform. Theory*, IT-29, n° 1, janvier 1983, p. 161-164.
- [2] G. BIENVENU et L. KOPP, Principe de la Goniométrie passive adaptative, *Acte du septième colloque sur le traitement du signal et ses applications*, Nice, juin 1979.
- [3] J. P. BURG, D. G. LUENBERGER et D. L. WENGER, Estimation of structured covariance matrix, *Proc. IEEE*, 70, n° 9, septembre 1982, p. 963-974.
- [4] P. G. CIARLET, *Introduction à l'analyse numérique matricielle*, Masson, Paris, 1982.
- [5] U. GRENANDER et G. SZEGO, *Toeplitz Forms and their application*, University of California Press, San Francisco, 1958.
- [6] V. F. PISARENKO, The retrieval of harmonics from a covariance function, *Geophys. J.R. Ast. Soc.*, 33, 1973, p. 347-366.
- [7] H. MERMOZ, Imagerie, Corrélation et Modèles, *Annales des Télécommunications*, 31, n° 1-2, 1976, p. 17-36.