

L'INITIALISATION DU MAXIMUM DE VRAISEMBLANCE EN TRAITEMENT D'ANTENNE

Jean-Jacques FUCHS & Hervé CHUBERRE

IRISA / Université de Rennes 1

Campus de Beaulieu, 35042 Rennes Cedex, France

RÉSUMÉ

L'estimation au sens du maximum de vraisemblance des gisements et puissances de P sources éclairant une antenne linéaire à capteurs équirépartis en présence de bruit blanc est un problème mal conditionné pour lequel les algorithmes d'optimisation convergent vers l'optimum local le plus proche du point d'initialisation utilisé. Nous proposons un ensemble de paramétrisations différentes du problème initial. Elles sont naturellement associées à une procédure d'initialisation dont nous étudions les propriétés.

1 Introduction

Nous considérons le problème de la localisation de sources à l'aide des signaux reçus sur une antenne passive à capteurs équidistants et de façon plus précise l'estimation des gisements et puissances de ces sources supposées être placées dans le champ lointain. Nous nous limitons au cas – appelé – à bande étroite où un seul bin de fréquence des transformées de Fourier des signaux issus des capteurs est pris en compte. On peut alors penser aux signaux émis par les sources comme étant des sinusoides de même fréquence.

Ce problème est d'une importance considérable et de nombreuses solutions ont été proposées. L'approche "maximum de vraisemblance" (MV) a été l'une des premières à être considérée [1]. Elle n'est cependant pas devenue très populaire notamment à cause du mauvais conditionnement de la fonctionnelle (nombreux extréma locaux) et du temps de calcul prohibitif des procédures d'optimisation non-linéaire requises. D'autres méthodes sous-optimales – appelé à haute résolution [2] – nécessitant moins de calculs et ne présentant pas de difficultés d'initialisation, ont fini par monopoliser toute l'attention. Leurs performances sont cependant inférieures à celles des approches MV surtout dans des configurations difficiles : faible rapport signal à bruit, sources proches, petit nombre d'observations... Nous présentons une façon originale de paramétrer et d'initialiser un algorithme du type MV.

2 Modélisation

Nous considérons une antenne linéaire à N capteurs équidistants éclairée par P sources ($P < N$) non corrélées en présence de bruit blanc. Les sorties des capteurs sont filtrées, échantillonnées et on prend leurs transformées de Fourier discrète sur des tranches temporelles disjointes. Sous les hypothèses habituelles, le vecteur complexe $X_t(f)$ de dimension N , associé à la $t^{\text{ème}}$ tranche et construit à l'aide des transformées de Fourier pris à un bin de fréquence temporelle f fixée, peut être modélisé comme un tirage d'un processus gaussien complexe, centré, circulaire dont la matrice de covariance R est une matrice de Toeplitz, hermitienne, définie positive dont la plus petite valeur propre est de multiplicité $N - P$.

On peut également voir cette matrice de covariance comme étant la matrice de densité spectrale, évaluée à la fréquence

ABSTRACT

The maximum likelihood estimation of the bearings and powers of P plane waves impinging on a linear array of equispaced sensors in spatially white noise is known to be a highly non-linear ill-conditioned optimization problem. The initialization procedure is thus a major issue since it dictates most of the performances of the overall algorithm. We propose a different set of parametrizations of the problem and investigate the properties of the associated initialization procedure.

f (omise dans la suite) du processus vectoriel, temporel observé par les N capteurs. Cette seconde interprétation permet d'obtenir un modèle paramétrique de R . Avec les notations habituelles, on a :

$$\begin{aligned} R &= \sum_{p=1}^P a_p d\theta_p d\theta_p^* + \sigma^2 I \\ d\theta_p &= [1 e^{i\theta_p} e^{2i\theta_p} \dots e^{i(N-1)\theta_p}]^T \\ \theta_p &= \pi \sin \varphi_p \end{aligned} \quad (1)$$

où $d\theta_p$ est le vecteur-direction, de déphasage θ_p , associé à la source p dont le gisement mesuré par rapport au travers de l'antenne est φ_p . La matrice R apparaît alors comme dépendant de $(2P + 1)$ paramètres que nous notons Y_0

$$Y_0 = \{ \sigma^2 ; (a_p, \theta_p), p = 1 \text{ à } P \} \quad (2)$$

Si on dispose alors de T vecteurs-observations X_t , l'estimation au sens du MV de Y_0 consiste à chercher le maximum en Y_0 de :

$$f(\{X_t\}, Y_0) = \pi^{-TN} (\det R)^{-T} \exp \left\{ - \sum_{t=1}^T X_t^* R^{-1} X_t \right\}$$

où les paramètres Y_0 interviennent dans R de la façon décrite dans (1). En posant :

$$\hat{R} = \frac{1}{T} \sum X_t X_t^* \quad (3)$$

l'estimée de la matrice de covariance et en prenant le logarithme de $f(\cdot)$ on voit qu'il est équivalent de minimiser en Y_0 la fonction.

$$g(Y_0) = \log \det R + \text{tr}(\hat{R} R^{-1}) \quad (4)$$

C'est cette forme qui montre que \hat{R} est une statistique suffisante, que nous conservons dans la suite.

3 La méthode du maximum de vraisemblance

Etant donnée l'observation \hat{R} (3), la méthode du MV consiste donc à chercher la matrice R , classiquement paramétrée à l'aide



des inconnues Y_0 (2), qui minimise le critère (4). Cette opération est difficile notamment à cause de la présence de l'inverse de R . Devant cette difficulté, on peut changer la fonctionnelle ou relâcher la contrainte imposée par la paramétrisation et se contenter d'un modèle de R moins structuré. Les deux options ont été largement explorées. On peut considérer les méthodes du type "sous-espace signal" comme dérivant de la seconde et bon nombre de méthodes du MV approché comme issues de la première ou d'un mélange des deux [3].

La vraie difficulté dans la méthode du MV réside en fait dans l'initialisation : un algorithme d'optimisation itératif va en général converger vers l'optimum local associé au domaine d'attraction dans lequel il a été initialisé [4]. L'important est donc d'initialiser l'algorithme avec un point situé dans le domaine d'attraction de l'optimum global et l'étude de l'algorithme proprement dit est d'un intérêt limité. Dans le cas d'une source unique ($P = 1$) l'optimum de $g(\cdot)$ par rapport aux trois inconnues (σ^2, a_1, θ_1) s'obtient facilement. Le problème est séparable et l'estimée $\hat{\theta}_1$ du déphasage est donnée par :

$$\hat{\theta}_1 = \arg \max_{\theta} d\theta^* \hat{R} d\theta$$

Il s'agit bien entendu du maximum de la formation de voies classique. Il faut alors remarquer que cette fonction a de nombreux maxima locaux et que seule une recherche exhaustive permet de localiser le maximum global. On peut par ailleurs observer que ces extrema locaux n'apparaissent que par rapport à la variable de déphasage θ et que, dans le cas particulier $P = 1$, ils sont distants de $\Delta f_s = 1/N$ (N , le nombre de capteurs ; $f_s = \sin \theta/2$, la fréquence spatiale). Cette observation reste valide dans le cas général et met l'accent sur l'importance de l'initialisation en θ , les autres paramètres apparaissant beaucoup moins sensibles.

Dans la suite, nous proposons une procédure d'initialisation. Elle est basée sur un changement de paramétrisation qui a pour effet de dilater l'échelle en θ et donc d'augmenter la taille du domaine d'attraction et faciliter l'initialisation.

4 Un changement de paramétrisation

La façon habituelle de paramétrer R consiste à prendre les $(2P + 1)$ inconnues réelles

$$Y_0 = \{ \sigma^2 ; (a_p, \theta_p), p = 1 \text{ à } P \},$$

à leurs associer la suite de covariance :

$$r_k = \sum_p a_p e^{ik\theta_p} + \sigma^2 \delta_k \quad (5)$$

et à en déduire la matrice de covariance R d'ordre N , Toeplitz, hermitienne :

$$R = \begin{bmatrix} r_0 & \bar{r}_1 & \bar{r}_2 & \cdots & \bar{r}_{N-1} \\ r_1 & r_0 & & & \\ r_2 & r_1 & & & \\ \cdots & & & & \\ r_{N-1} & & & & r_0 \end{bmatrix} \quad (6)$$

Cette paramétrisation découle naturellement de la modélisation physique (1) et est celle qui intéresse l'utilisateur. Nous proposons de paramétrer différemment la matrice R et introduisons l'ensemble suivant de jeux de paramètres :

$$Y_k = \{ r_0, r_k, r_{2k}, \dots, r_{Pk} \} \quad k = 1, 2, \dots \quad (7)$$

Pour chaque valeur de l'indice k , on a bien un ensemble de $(2P + 1)$ inconnues réelles. De la même façon que dans Y_0 on demande aux puissances σ^2 et a_p d'être positives, des contraintes sont à imposer à la suite Y_k pour correspondre à une paramétrisation valide. Comme ce sont les paramètres Y_0 qui nous intéressent, même si nous minimisons $g(\cdot)$ par rapport à Y_k , nous calculerons toujours l'estimée Y_0 correspondante et serons surtout concernés par la matrice de covariance de cette estimée. Cette matrice ne dépend que du critère $g(\cdot)$ et de la configuration et non du jeu de paramètres utilisé pour trouver l'optimum. Tout ce que l'on peut escompter d'un changement de variables, c'est un domaine d'attraction plus étendu et mieux conditionné et éventuellement une procédure d'initialisation plus aisée.

Une difficulté liée à l'utilisation de cette nouvelle paramétrisation réside dans l'évaluation du critère. En un point Y'_k donné, il faut pour calculer $g(Y'_k)$ commencer par évaluer les valeurs $r_l(Y'_k)$ pour $l = 0$ à $N - 1$, de la suite de covariance. Une manière de procéder consiste alors à chercher la valeur Y'_0 du point associé à Y'_k , dans la paramétrisation standard et à en déduire $r_l(Y'_k)$ à l'aide des relations (5).

Pour le jeu de paramètres (7) obtenu pour $k = 1$, la méthode de Pisarenko [5] réalise précisément le changement des variables Y_1 en Y_0 . On note au passage que pour qu'un point Y'_1 soit valide il faut et il suffit que la matrice de covariance (6) d'ordre $P + 1$ associée soit définie positive et à valeur propre minimale simple. Dans le cas des paramétrisations (7) associées à des indices $k > 1$, les composantes de Y_k peuvent être vues comme les covariances issues d'une antenne présentant un sous-échantillonnage spatial de rapport k . Ce sous-échantillonnage se traduit par un phénomène de repliement du spectre spatial et la méthode de Pisarenko appliquée à Y_k va fournir des valeurs des fréquences spatiales présentes dans le spectre replié. A une valeur $F' \in] -1/2, 1/2[$ donnée par la méthode de Pisarenko, il faut alors associer les k fréquences spatiales (dépliées) :

$$f'_i = \frac{1}{k} (F' + l) \quad (8)$$

obtenues pour les valeurs de l'entier l donnant une fréquence $f'_i \in] -1/2, 1/2[$. Dans le cas de P sources, il y a ainsi k^P jeux de fréquences possibles. A chaque jeu, on associe alors un point-candidat dans l'espace des Y_0 et il est naturel de retenir le candidat (la détermination) dont le coût $g(\cdot)$ est le plus faible. Ceci est la façon habituelle de procéder en présence d'une transformation dont l'inverse n'est pas unique et le point correspondant est parfois appelé l'estimée induite [6]. Cette procédure est extrêmement coûteuse en temps de calcul mais peut être simplifiée.

Un avantage de la paramétrisation Y_k par rapport à la paramétrisation standard Y_0 , réside dans son association naturelle avec le point initial :

$$\hat{Y}_k(0) = \{ \hat{r}_0, \hat{r}_k, \hat{r}_{2k}, \dots, \hat{r}_{Pk} \} \quad (9)$$

où les composantes \hat{r}_{lk} sont des estimées obtenues à partir de la matrice \hat{R} d'une façon qui reste à définir. Ceci nous amène à ne considérer que des paramétrisations Y_k avec k inférieur à :

$$k_{\max} = \text{partie entière } \{ (N - 1)/P \} \quad (10)$$

L'exploitation du point initial (9) passe par la procédure décrite plus haut qui lui associe un point unique dans l'espace Y_0 . Ce point ne présente un intérêt que s'il est situé dans le bassin d'attraction de l'optimum global. La taille et la forme de ce bassin sont liées au conditionnement du critère à l'optimum.



L'inverse de la matrice de covariance asymptotique du minimum global de $g(\cdot)$ qui caractérise localement le critère s'obtient facilement pour toutes les paramétrisations Y_k à partir de la matrice correspondante pour Y_0 (la matrice d'information de Fisher). Son conditionnement varie de façon importante avec k . La meilleure paramétrisation k , en ce sens, dépend de la configuration et est donc inconnue a priori. Il est cependant facile d'imaginer que ce sont les indices k proches de k_{\max} qui correspondent aux "meilleures" paramétrisations notamment vis-à-vis de difficultés liées à des sources de gisements voisins. En effet, pour deux sources séparées de Δf en fréquence spatiale dans la paramétrisation Y_0 ou Y_1 , la séparation devient $k\Delta f$ dans la paramétrisation Y_k , une configuration généralement bien plus favorable.

Une autre façon de se faire une idée des sensibilités relatives des paramétrisations consiste à s'intéresser au déterminant du Jacobien $\nabla_{Y_0} Y_k$ reliant les variations élémentaires :

$$\Delta Y_k = (\nabla_{Y_0} Y_k) \Delta Y_0$$

Quand ce déterminant est grand, il indique qu'une petite variation en Y_0 induit une grande variation en Y_k qui est alors une paramétrisation bien moins sensible aux erreurs d'estimation.

Dans la pratique, pour déterminer la "meilleure" paramétrisation, en ce sens, nous utilisons la procédure décrite plus haut et retenons parmi les points initiaux obtenus pour $k = 1$ à k_{\max} celui dont le coût est le plus faible. On observe sur les simulations que la paramétrisation Y_k ainsi choisie dépend, pour une même configuration de sources, de la réalisation mais que la "meilleure" paramétrisation au sens du conditionnement de la matrice de covariance évaluée aux valeurs exactes des paramètres est, en général, retenue pour plus de la moitié des réalisations.

5 L'algorithme proposé

Résumons la démarche proposée jusqu'à présent. Au lieu de paramétrer le critère (4) à l'aide des inconnues standards Y_0 (2), nous avons introduit un ensemble de k_{\max} (10) nouveaux jeux de paramètres Y_k (7). A chacun d'eux, on associe de façon naturelle un point initial $\hat{Y}_k(0)$ (9) auquel on applique la méthode de Pisarenko. Cette méthode réalise un changement de variables faisant passer de l'espace des Y_k à celui des Y_0 . Cette transformation n'est pas univoque et il existe k^P (k , l'indice de la paramétrisation ; P , le nombre de sources) points Y_0 qui correspondent à l'unique point initial $\hat{Y}_k(0)$. Pour sélectionner la bonne détermination, on décide de retenir le point dans l'espace Y_0 qui optimise le critère (4). Cette opération est réalisée pour les k_{\max} paramétrisations proposées et le point "initial" finalement conservé est à nouveau celui dont le coût est le plus faible parmi les k_{\max} candidats.

Avant de nous intéresser aux propriétés de cette procédure, il reste à préciser la façon dont est construit le point initial $\hat{Y}_k(0)$ à partir de \hat{R} :

$$\hat{Y}_k(0) = \{\hat{r}_0, \hat{r}_k, \hat{r}_{2k}, \dots, \hat{r}_{Pk}\}$$

Cette phase n'est pas sans importance puisqu'elle conditionne largement la qualité du point initial. L'idée la plus simple consiste à prendre ces estimées dans la première colonne de la matrice \hat{R} (3). Cependant, comme la matrice R exacte (6) est une matrice de Toeplitz, il est certainement intéressant d'utiliser la redondance pour améliorer les estimées des covariances. On peut alors avoir recours à diverses manières de toeplerizer (rectifier) la matrice \hat{R} . L'une d'entre elles consiste à minimiser le critère $g(\cdot)$ du MV (4). On peut obtenir la solution à l'aide d'un algorithme itératif [7]. En analyse factorielle

[8], [9], des critères plus faciles à minimiser ont souvent été utilisés. Parmi les plus courants, on trouve celui des moindres carrés non pondérés :

$$U(\cdot) = \frac{1}{2} \text{tr}(\hat{R} - R)^2 \quad (11)$$

et celui des moindres carrés généralisés :

$$W(\cdot) = \frac{1}{2} \text{tr}(I - \hat{R}^{-1}R)^2 \quad (12)$$

Ce sont deux autres manières de définir une distance entre l'observation \hat{R} et le modèle R . Dans le cas où R est seulement contrainte à être Toeplitz, hermitienne, les solutions sont faciles à obtenir et ont déjà été considérées dans le domaine du traitement d'antenne. La solution du critère (11) s'obtient en prenant les moyennes le long des diagonales, celle de (12) s'obtient comme le minimum d'une forme quadratique [3] convexe. Dans la suite, nous retiendrons la solution correspondant à minimiser la norme de Frobenius (11), l'obtention du point initial $\hat{Y}_k(0)$ est alors trivial.

Nous allons démontrer que dans le cas le plus simple (et sans intérêt pratique) d'une source unique, le point correspondant dans l'espace Y_0 atteint quasiment la borne de Cramer-Rao. Il est alors inutile de faire appel à un algorithme d'optimisation itératif initialisé en ce point ! Cette propriété reste vraie dans le cas de deux sources de gisements voisins, comme le confirment les simulations. La justification statistique est dans ce cas plus difficile. Il faut cependant se rendre compte que l'on choisit le meilleur point, au sens du MV, parmi k_{\max} points eux-mêmes déjà de bonnes qualités. Comme par ailleurs, la paramétrisation k^* retenue dépend de la réalisation, la variance de l'estimée qui en résulte est certainement inférieure à celle que donnerait, isolément, chacune des paramétrisations. De toute façon, la variance est limitée inférieurement par celle associée à l'optimum de la fonctionnelle utilisée pour faire la sélection. Mais dans notre cas, il s'agit de la fonctionnelle du MV et la limite est donc la borne de Cramer-Rao.

Il faut également remarquer que la simple sélection de la meilleure estimée (au sens du MV) parmi les k_{\max} disponibles mène à un résultat quasi-optimal. Cela signifie qu'un algorithme de fusion des estimées ne peut améliorer le résultat. Cette constatation est un peu surprenante car les estimées sont loin d'être totalement corrélées.

6 Analyse des propriétés

Nous allons montrer que dans le cas $P = 1$ d'une source unique, les performances de la procédure d'initialisation proposée sont proches de l'optimum. Dans le cas $P = 1$, la méthode de Pisarenko fait passer de $Y_k = \{\hat{r}_0, \hat{r}_k\}$ à un triplet, que nous notons $Y'_0 = \{\hat{\sigma}^2, \hat{a}, \hat{\theta}'\}$, qui vérifie :

$$\begin{aligned} \hat{r}_0 &= \hat{a} + \hat{\sigma}^2 \\ \hat{r}_k &= \hat{a} e^{i\hat{\theta}'} \end{aligned} \quad (13)$$

Nous nous intéressons à la variance de $\hat{\theta}'$ uniquement. De (13), on déduit qu'elle ne dépend que de celle de l'estimée \hat{r}_k . Nous considérons le cas le plus simple (11) où \hat{r}_k est obtenu par moyennage :

$$\hat{r}_k = \frac{1}{N-k} \sum_{j=1}^{N-k} \hat{r}_{j,j+k} \quad (14)$$

où $\hat{r}_{m,n}$ est l'élément (m, n) de \hat{R} . En notant :

$$\Delta r. = \hat{r}. - E(\hat{r}.)$$

on a pour les composantes de \hat{R} , une matrice de Wishart complexe à T degrés de liberté [9] :



$$E(\Delta r_{ij} \Delta r_{mn}) = \frac{1}{T} r_{in} r_{mj}$$

On trouve alors pour l'estimée (14) et en prenant en compte le fait que l'indice k peut être du même ordre de grandeur que N :

$$E(\Delta r_k \bar{\Delta} r_k) = \frac{1}{T(N-k)^2} \{(N-k)r_0^2 + 2 \sum_{j=1}^{N-k-1} (N-k-j) r_j \bar{r}_j\}$$

qui devient pour $P = 1$:

$$E(\Delta r_k \bar{\Delta} r_k) = \frac{a^2}{T} + \frac{2a\sigma^2 + \sigma^4}{T(N-k)} \quad (15)$$

Par un calcul analogue, on obtient :

$$E(\Delta r_k^2) = \frac{a^2}{T} e^{2ik\theta} \quad (16)$$

où θ est le vrai déphasage de l'unique source. Les relations (15) et (16) décrivent entièrement la loi de la variable complexe \hat{r}_k (14). On en déduit la variance de $\hat{\theta}'$ son argument :

$$E(\Delta \theta'^2) = \frac{1}{2\rho^2} \frac{2\rho + 1}{T(N-k)} \quad (17)$$

avec $\rho = \frac{a}{\sigma^2}$ le rapport signal à bruit. Pour passer de l'estimée de θ' à celle de θ , il faut "déplier". Cette opération se traduit par une division par k et une translation (8). En supposant que la bonne translation/détermination est retenue, on a :

$$E(\Delta \theta^2) = \frac{1}{k^2} E(\Delta \theta'^2)$$

où il reste à remplacer $E(\Delta \theta'^2)$ par sa valeur (17). Le minimum en k de cette expression est atteint pour $k^* = 2N/3$ et elle devient :

$$E(\Delta \theta^2) = \frac{27}{8\rho^2} \frac{2\rho + 1}{TN^3} \quad (18)$$

Cette quantité est à comparer à la borne de Cramer-Rao correspondante [2] :

$$E(\Delta \theta^2) \geq \frac{6(N\rho + 1)}{T\rho^2 N(N^2 - 1)}$$

La borne est approchée de très près. Il faut maintenant s'assurer que cette efficacité est conservée dans des cas plus réalistes. Le calcul ne présente pas de difficultés mais ne permet pas de se faire une idée. Pour des sources isolées (au sens de la formation de voies), le calcul précédent reste valide, mais il faut observer que quand P augmente, k_{\max} (10) diminue et il n'est alors plus possible de retenir la valeur k^* qui minimise la variance.

Nous présentons plus loin des résultats de simulation obtenus dans le cas de deux sources très proches en gisement. On observe que les performances restent excellentes et voisines des bornes de Cramer-Rao.

7 Simulations et conclusions

Nous ne présentons de résultats que pour une configuration présentant une difficulté en résolution. Les caractéristiques sont les suivantes : $T = 100$ "snapshot", $N = 11$ capteurs, $P = 2$ sources, le rapport signal à bruit vaut 0dB pour les deux sources qui sont placées à 10 et 12 degrés par rapport au tra-

vers de l'antenne. Dans ces conditions, le borne de CR pour les estimées de gisements donne un écart-type de $0,95$ degré pour chaque source. Comme la vraie séparation est de 2 degrés, on est dans une configuration limite car on ne peut espérer raisonnablement séparer des sources plus proches. L'algorithme proposé les sépare et l'écart-type sur les gisements évalué sur 50 réalisations vaut $1,6$. Il est donc voisin de la borne de CR et il en est de même pour les estimées des amplitudes. On peut également noter que $k_{\max} = 5$ et que les diverses paramétrisations ont respectivement été retenues : $0, 3, 9, 24$ et 14 fois. La paramétrisation Y_4 est donc sélectionnée 24 fois sur 50 .

Dans le cas considéré, la méthode d'initialisation proposée ne nécessite aucune itération ultérieure. Il n'en serait sans doute pas de même pour des configurations à nombre de sources plus importants. Il faut alors éviter qu'une mauvaise détermination soit retenue au départ. Un premier traitement consistant en une formation de voies classique s'impose. Elle permettra de localiser grossièrement les sources. On peut alors retirer les sources isolées et donc faciles à estimer avec précision et poursuivre avec la procédure proposée. Les vertus des paramétrisations Y_k proposées permettent alors d'affiner l'initialisation.

Références

- [1] F.C. SCHWEPPE : "Sensor array processing for multiple sensor sources". *IEEE-T-IT*, vol.14, p. 294-305, 1968.
- [2] G. BIENVENU, L. KOPP : "Optimality of high resolution array processing using the eigensystem approach". *IEEE-T-ASSP*, vol.31, p. 1235-1248, 1983.
- [3] P. FORSTER, G. VEZZOSI : "Optimal toeplitzification with a given rank constraint". *Proc. ICASSP*, p. 2783-2786, 1989.
- [4] M.I. MILLER, D.R. FUHRMANN : "Maximum-likelihood narrow-band direction finding and the EM algorithm". *IEEE-T-ASSP*, vol.38, p. 1560-1577, 1990.
- [5] V.F. PISARENKO : "The retrieval of harmonics from a covariance function". *Geophys. J. roy. Astron. Soc.*, p. 347-366, 1973.
- [6] I.N. ZEHNA : "Invariance of maximum likelihood estimators". *Ann. of Math. Stat.*, vol.37, p. 744, 1966.
- [7] J.P. LE CADRE : *Contribution à l'utilisation des méthodes paramétriques en traitement d'antenne*. Thèse d'état, Grenoble, octobre 1987.
- [8] K.G. JORESKOG : "Factor analysis by least-squares and maximum likelihood methods". In *Statistical Methods for digital computers*, vol.3, p. 125-153, John Wiley, 1978.
- [9] T.W. ANDERSON : "An introduction to multivariable statistical analysis". 2nd Edition, John Wiley, 1984.