Remarques sur la diagonalisation tensorielle par la méthode de Jacobi

Pierre Comon
Thomson-Sintra, B.P.157, F-06903 Sophia-Antipolis Cedex
Pierre.Comon@sp1.y-net.fr

Résumé

Le problème de "séparation de sources" est soluble en recourrant uniquement aux statistiques d'ordre 2, comme nous le soulignons. Pourtant, les conditions d'identifiabilité imposent souvent l'usage complémentaire de statistiques d'ordre supérieur. Le problème revient alors à diagonaliser un tenseur cumulant, ce qui ne peut se faire en général que de façon approximative. Un algorithme de type Jacobi semble bien approprié à ce genre de tâche, mais il est difficile de prouver sa convergence vers le maximum absolu du critère. Deux démonstrations restent inachevées.

Abstract

The "sources separation" problem can be solved by resorting only to 2nd order statistics, as is emphasized. Yet, identifiability conditions often impose the complementary use of higher order statistics. The problem then comes to the diagonalization of a cumulant tensor, which can only be performed approximately. A Jacobi like algorithm seems well suited to this task, but it is difficult to prove its convergence to the absolute maximum of the criterion. Two proofs remain uncompleted.

1 Introduction

L'Analyse en Composantes Indépendante (ACI) est un des problèmes regroupés sous la dénomination de *séparation de sources* (cf. section 4). Les approches permettant de mettre en œuvre l'ACI d'une matrice de données sont à présent plusieurs. A l'origine, Hérault et Jutten ont proposé un algorithme adaptatif simple, mais dont la convergence pouvait être facilement mise en défaut. Lacoume a proposé ensuite de minimiser une fonction de plusieurs variables, mais sans proposer d'algorithme pratique si ce n'est une procédure de recherche exhaustive [14]; cette contribution avait le mérite d'être justifiable sur le plan théorique. Cardoso [1] et Comon [4] ont proposé indépendamment deux algorithmes permettant de calculer l'ACI en dimension arbitraire et en un temps polynomial. Ces auteurs ont d'ailleurs proposé depuis d'autres algorithmes plus efficaces [3] [5] [12] [6]. Par ailleurs un certain nombre de méthodes intéressantes ont été proposées, notamment pour l'égalisation adaptative, mais exploitent le fait que les cumulants des sources sont tous de même signe, et que le bruit est gaussien, hypothèses qui ne sont pas faites ici.

Au cours de la dernière conférence *GRETSI*, Souloumiac a exposé une nouvelle procédure de résolution, permettant de faire le lien entre les approches décrites dans [2] et [5]. Les performances de cette nouvelle technique sont analysées en détail dans [15]. Les techniques de Comon et Souloumiac sont particulièrement performantes en dimension 2, mais leur extension aux dimensions supérieures est basée sur la *méthode de Jacobi*. Or pour l'instant, il n'a pas été possible de prouver leur convergence vers l'extrêmum absolu recherché en dimension supérieure 2. Nous tentons ici d'expliquer pourquoi cette preuve est difficile, alors qu'elle est si simple lorsque la méthode de Jacobi est appliquée à des matrices. Nous nous limiterons dans la suite à des variables réelles.

La section 2 rappelle quelques propriétés élémentaires satisfaites par les cumulants de variables aléatoires de dimension 1 et supérieure. La section 3 résume l'essentiel des méthodes que nous qualifions d' *algorithmes de type Jacobi*. La section 4 propose une technique d'identification aveugle utilisant uniquement des statistiques d'ordre 2, et souligne ses conditions d'identifiabilité. Deux approches sont proposées dans la section 5 pour analyser l'existence de maxima parasites dans les algorithmes de type Jacobi d'ordre 2, 3, ou 4. Comme nous ne parvenons à aucune démonstration complète pour les ordres supérieurs à 2, la section 2 a finalement pour fonction essentielle de définir les notations.

2 Cumulants

Soit X une variable aléatoire d'ordre 4 à valeurs dans $\mathbb{R}$. Notons $\mu_{(r)} = E\{(X - E\{X\})^r\}$ ses moments centrés successifs d'ordre $r, r > 1$, et $\kappa_{(r)}$ ses cumulants d'ordre r. Pour $r = 3$ et $r = 4$, les cumulants peuvent s'écrire très simplement en fonction des moments [11]:

$$\kappa_{(3)} = \mu_{(3)} \tag{1}$$

$$\kappa_{(4)} = \mu_{(4)} - 3\mu_{(2)}^2. \tag{2}$$

Rappelons que les cumulants d'ordre supérieur à 2 s'annulent lorsque la variable considérée est gaussienne. On constate que les cumulants dépendent de la variance de la variable, de sorte qu'il est utile sur le plan pratique d'introduire les cumulants de la variable réduite, appelés *cumulants standardisés*:

$$\gamma_{(r)} = \kappa_{(r)}\,\kappa_{(2)}^{-r/2}. \tag{3}$$

Le cumulant $\gamma_{(3)}$ est souvent appelé *asymétrie* (*skewness* en anglais), et $\gamma_{(4)}$ *aplatissement* (*kurtosis* en anglais). Les cumulants standardisés vérifient des inégalités remarquables qu'il est important de souligner. Par exemple, en utilisant le fait que $var\{aX^2 + bX\}$ est toujours positive, quels que soient les coefficients a, b, on montre que:

$$\gamma_{(3)}^2 \leq \gamma_{(4)} + 2. \tag{4}$$

Autrement dit, l'asymétrie est bornée par l'aplatissement.

La standardisation est définie de façon similaire dans le cas multivariable. Soit $\mathbf{X}$ un vecteur aléatoire d'ordre 2, et C sa matrice de covariance. Notons sa décomposition spectrale par $C = U\Lambda U^T$, où Λ est diagonale de rang plein (donc éventuellement de taille inférieure à C). Alors la variable standardisée peut être définie par $\tilde{\mathbf{X}} = \Lambda^{-1/2}U^T\mathbf{X}$. Elle a donc une covariance unité. Les cumulants standardisés seront notés $\gamma_{ijkl} = cum(\tilde{X}_i, \tilde{X}_j, \tilde{X}_k, \tilde{X}_l)$; l'aplatissement de $\tilde{X}_i$ pourra donc être noté γ_{iiii}. Cette nouvelle notation à plusieurs indices devient nécessaire dès que plusieurs variables sont en jeu; il est courant de faire cohabiter les deux notations [13]. Si on note les moments de la même façon, e.g., $\mu_{ijkl} = E\{X_i\, X_j\, X_k\, X_l\}$, on peut remarquer que pour un vecteur standardisé, $\gamma_{1111} = \mu_{1111} - 3$, $\gamma_{1122} = \mu_{1122} - 1$, et $\gamma_{1112} = \mu_{1112}$. Ces relations sont utilisées dans le paragraphe qui suit.

En reprenant le raisonnement précédent avec deux variables aléatoires X_1 et X_2, l'inégalité $var\{aX_1^2 + bX_2^2 + X_1X_2\} \geq 0$ conduit à la relation suivante, pour tout couple (a, b) de $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$:

$$\begin{aligned} a^2(\gamma_{1111}+2)+b^2(\gamma_{2222}+2) & \\ +(1+2ab)\gamma_{1122}+2a\gamma_{1112}+2b\gamma_{1222}+1 & \geq 0. \end{aligned} \tag{5}$$

On notera en particulier que $\gamma_{iiii} \geq -2$, que $\gamma_{iijj} \geq -1$, et que $\gamma_{iiii} + 2\gamma_{iiij} \geq -2$. La relation (4) s'étend quant à elle facilement au cas bivariable en utilisant $var\{aX_1^2 + bX_2\} \geq 0$, et $\gamma_{1111} \geq -2$; on obtient alors $\gamma_{112}^2 \leq \gamma_{1111} + 2$. En outre, la relation simple $\gamma_{112}^2 + \gamma_{122}^2 \leq \gamma_{1122} + 1$ se déduit de $var\{X_1X_2 + aX_1 + bX_2\} \geq 0$... Il est inutile de poursuivre pour comprendre qu'un certain nombre d'inégalités peuvent être obtenues à partir de ce principe.

3 Algorithmes de type Jacobi

L'algorithme de Jacobi a pour but initialement de diagonaliser les matrices symétriques réelles par changement de base orthonormée. Pour diagonaliser une matrice symétrique M de taille $N \times N$, on cherche la matrice de passage orthonormée Q, telle que la fonctionnelle:

$$\Psi_2(Q; M) = \sum_p G_{pp}^2, \tag{6}$$

avec
$$G_{ij} = \sum_{m,n} Q_{im} Q_{jn} M_{mn}, \tag{7}$$

soit maximale. Le cas le plus simple est celui où $N = 2$. Dans ce cas, la matrice Q peut être paramétrée par un angle unique (ou plutôt par sa tangente θ). La maximisation de (6) revient dans ce cas à la résolution d'un polynôme de degré 2 en θ [10]. Lorsque $N > 2$, il est possible de paramétrer la matrice Q par un produit de $N(N-1)/2$ rotations planes (maintenant baptisées *rotations de Givens*), chacune caractérisée par un paramètre unique. On résoud ainsi une séquence de problèmes monovariables au lieu d'un problème d'optimisation multivariable [10]. La recherche d'une rotation plane constitue une *itération*, et la séquence de $N(N-1)/2$ rotations planes dans un ordre préétabli constitue un *balayage*. Il est alors nécessaire d'exécuter plusieurs balayages, de sorte que la méthode devient itérative, et s'apparente un peu à une méthode de relaxation, à la différence près qu'on se déplace dans un groupe (le groupe *Spécial Orthogonal*) et non dans un espace vectoriel.

La convergence de l'algorithme de Jacobi peut se prouver [10] en remarquant que quelle que soit la matrice de départ M, le critère $\Psi_2(Q; M)$ décroit strictement à chaque balayage, d'une part, et que le maximum de $\Psi_2(Q; M)$ est unique et vérifie $G_{ij} = 0$ pour $i \neq j$. Les résultats obtenus dans la section 5 permettront d'ailleurs de le montrer d'une façon originale.

Dans le cas de tenseurs d'ordre supérieur à 2, et plus particulièrement de tenseurs cumulants, cette dernière propriété est fausse: il n'est pas toujours possible de diagonaliser un tenseur complètement symétrique par transformation linéaire. Une condition nécessaire a d'ailleurs déjà été mise en évidence dans [5]. Si un changement de base défini par une matrice orthonormée Q est effectué, les tenseurs cumulants d'ordre 3 et 4 de la variable transformée Qx s'expriment en fonction des tenseurs cumulants de x comme suit:

$$G_{ijk} = \sum_{mno} Q_{im} Q_{jn} Q_{ko} g_{mno}, \tag{8}$$

$$G_{hijk} = \sum_{lmno} Q_{hl} Q_{im} Q_{jn} Q_{ko} g_{lmno}. \tag{9}$$

Les critères à maximiser sont alors respectivement:

$$\Psi_3(Q; g) = \sum_p G_{ppp}^2, \text{et } \Psi_4(Q; g) = \sum_p G_{pppp}^2, \tag{10}$$

et ne peuvent pas forcément atteindre la borne supérieure qui est la somme des carrés de tous les cumulants (e.g. à cause du bruit).

4 Identification de mélanges

Le problème souvent appelé *séparation de sources* consiste à identifier la matrice carrée inversible M dans un modèle d'observation de la forme: $y(t) = Mx(t)$, où $x(t)$ est un vecteur dont les composantes sont statistiquement indépendantes. Il existe une façon d'identifier le mélange M en utilisant seulement les moments d'ordre 2 des observations; l'algorithme de Jacobi pourra y être utilisé. Le principe de base a été proposé à l'origine par Fety [8], mais nous renvoyons à [7] pour une démonstration plus rigoureuse. Notons $\Gamma_y(\tau)$ et $\Gamma_x(\tau)$ les fonctions de corrélation de $x(t)$ et $y(t)$, respectivement. Alors nous avons:

$$\Gamma_y(\tau) = M\, \Gamma_x(\tau)\, M^\dagger, \tag{11}$$

pout tout retard τ. En général, la matrice M peut être identifiée uniquement avec le concours de cette équation, contrairement à certaines croyances. Voici comment procéder. On construit deux matrices

$$\Gamma_1 = \sum_\tau \alpha_\tau\, \Gamma_y(\tau), \text{ et } \Gamma_2 = \sum_\tau \beta_\tau\, \Gamma_y(\tau), \tag{12}$$

où α_τ et β_τ sont des coefficients scalaires réels. On calcule une factorisation (n'importe laquelle) de la forme $\Gamma_1 = H\, H^\dagger$, puis on calcule la décomposition spectrale: $H^{-1}\, \Gamma_2\, H^{-\dagger} = U\, \Lambda^2\, U^\dagger$. Alors il est aisé de voir que la transformation $F = U^\dagger H^{-1}$ diagonalise simultanément les deux matrices Γ_i. Après analyse, on vérifie que $H\, U = M$ à une permutaton près si et seulement si les valeurs propres Λ_{ii} sont toutes distinctes [7]. Il faudra donc choisir les coefficients α_τ et β_τ en conséquence.

Cette dernière condition n'est pas toujours facile à satisfaire. Si $\Gamma_y(\tau)$ et $\Gamma_x(\tau)$ sont proportionnelles, par exemple, il est impossible de la satisfaire (e.g. bruits blancs au second ordre). En particilier, si $x(t)$ et $y(t)$ n'ont pas de cohérence temporelle, mais sont simplement des réalisations indépendantes de variables aléatoires X et Y, respectivement, le modèle d'observation que l'on doit adopter est plutôt de la forme:

$$Y = M\, X. \tag{13}$$

L'identification aveugle de la matrice M est alors plus difficile. Il s'agit d'un autre problème de "séparation de sources", que nous appellerons *Analyse en Composantes Indépendantes* (ACI). Pour de plus amples détails concernant l'ACI, y compris sa dénomination, on se réferrera à [5] [12] [6]. Cette identification, contrairement à celle des mélanges de *signaux*, nécessite le recours aux statistiques d'ordre supérieur à 2.

Au lieu de diagonaliser conjointement 2 matrices par transformation linéaire régulière, on diagonalisera *approximativement* un tenseur d'ordre 3 ou 4 par transformation *orthogonale*. Nous nous interesserons dans la suite uniquement aux algorithmes de type "Jacobi".

5 Conditions d'existence d'un maximum unique

Bien qu'aucune des deux approches décrites ci-après n'ait permis de conclure, la première reste sans doute préférable car à la fois plus élégante et plus générale.

5.1 Première approche

Essayons de caractériser les transformations Q rendant les fonctionnelles Ψ_i stationnaires. Si Q est orthogonale, alors elle vérifie $Q^T Q = I$, et donc en conséquence $dQ^T Q = -Q^T dQ$, ce qui montre que la matrice $A = dQ\,Q^T$ est antisymétrique. On peut donc poser $dQ = dA\,Q$, comme cela a été proposé dans [6]. Contrairement aux matrices orthogonales, l'ensemble des matrices antisymétriques est un espace vectoriel et admet admet pour base les matrices $A(p,q)$ définies par $A(q,r)_{pt} = \delta_{pq}\delta_{tr} - \delta_{pr}\delta_{tq}$. La condition de stationnarité des fonctionnelles Ψ_i:

$$\begin{aligned} \tfrac{1}{2}d\Psi_2 &= 2\sum_{p,t} G_{pp}\,dA_{pt}\,G_{tp} = 0, && (14)\\ \tfrac{1}{2}d\Psi_3 &= 3\sum_{p,t} G_{ppp}\,dA_{pt}\,G_{tpp} = 0, && (15)\\ \tfrac{1}{2}d\Psi_4 &= 4\sum_{p,t} G_{pppp}\,dA_{pt}\,G_{tppp} = 0, && (16) \end{aligned}$$

peut alors s'écrire sur la base des A(q,r):

$$\begin{aligned} G_{qq}G_{qr} - G_{rr}G_{qr} &= 0, && (17)\\ G_{qqq}G_{qqr} - G_{rrr}G_{qrr} &= 0, && (18)\\ G_{qqqq}G_{qqqr} - G_{rrrr}G_{qrrr} &= 0, && (19) \end{aligned}$$

pour tout couple (q,r), $q \neq r$. Penchons-nous à présent sur la différentielle seconde en ces points stationnaires. Après calcul, on obtient sur la base des matrices antisymétriques:

$$\begin{aligned} &\tfrac{1}{4}d^2\Psi_2 = 4\,G_{qr}^2 - (G_{qq} - G_{rr})^2, && (20)\\ &\tfrac{1}{6}d^2\Psi_3 = 4\,G_{qqr}^2 + 4\,G_{qrr}^2\\ &\quad -(G_{qqq} - G_{qrr})^2 - (G_{rrr} - G_{qqr})^2, && (21)\\ &\tfrac{1}{8}d^2\Psi_4 = \tfrac{9}{2}\,G_{qqrr}^2 + 4\,G_{qqqr}^2 + 4\,G_{qrrr}^2\\ &\quad -(G_{qqqq} - \tfrac{3}{2}\,G_{qqrr})^2 - (G_{rrrr} - \tfrac{3}{2}\,G_{qqrr})^2. && (22) \end{aligned}$$

On vérifie d'après (17) que $d\Psi_2$ s'annule dans deux situations: soit $G_{qr} = 0$, soit $G_{qq} = G_{rr}$. L'inspection de (20) permet alors de montrer que si $G_{qr} = 0$ pour tous les couples (q,r), alors on atteint un maximum; si $G_{qq} = G_{rr}$ pour tous les couples (q,r), alors il s'agit d'un minimum; tandis que tous les autres points, où l'annulation de $d\Psi_2$ fait intervenir les deux types d'égalité, sont des points-selle.

Pour Ψ_3, nous avons moins de chance car la relation (18) ne se factorise pas comme (17), de sorte qu'il y a différentes façons d'annuler $d\Psi_3$. La première est que G_{qqr} et G_{qrr} soient nuls, alors (21) est négative et nous avons bien un maximum. Si en revanche l'un des deux est non nul, disons G_{qqr} pour fixer les idées, on peut exprimer G_{qqq} en fonction de G_{rrr}. Alors la négativité de (21) est équivalente à

$$(G_{qqr} + G_{rrr})(3\,G_{qqr} - G_{rrr}) \leq 0, \qquad (23)$$

ce qu'il semble possible de satisfaire sans contredire les inégalités citées dans la section 2. Autrement dit, rien n'interdit apparamment l'existence de maxima parasites. Il est clair que cela ne s'arrange pas à l'ordre 4 avec Ψ_4. Cependant, l'auteur n'a pas de contre-exemple à proposer.

5.2 Seconde approche

Une autre approche consiste à prouver l'existence d'un seul maximum dans le domaine de définition de $\Psi_i(Q,g)$, pour tout type d'entrées et tout système orthonormé de coordonnées. On se limite ici à la dimension 2, et on pose

$$Q = \frac{1}{\sqrt{1+\theta^2}}\begin{pmatrix} 1 & \theta \\ -\theta & 1 \end{pmatrix}. \qquad (24)$$

Avec cette notation, on peut écrire les $\Psi_i(Q;g)$ sous la forme de fonctions de θ, ou de $\xi = \theta - 1/\theta$:

$$\begin{aligned} \psi_3(\theta;g) &= (\theta + \frac{1}{\theta})^{-3}\sum_{i=1}^{3} a_i\,(\theta^i - (-\theta)^{-i}) && (25)\\ \psi_4(\xi;g) &= (\xi^2+4)^{-2}\sum_{i=0}^{4} b_i\,\xi^i. && (26) \end{aligned}$$

On peut alors montrer que leurs points stationnaires vérifient les identités $\omega_i(\xi;g) = 0$ avec:

$$\begin{aligned} \omega_3(\xi;g) &= d_2\,\xi^2 + d_1\,\xi - 4\,d_2,\ \text{si } d_2 \neq 0, && (27)\\ \omega_4(\xi;g) &= \sum_{i=0}^{4} c_i\,\xi^i. && (28) \end{aligned}$$

En outre, si $d_2 = 0$, $\theta = 0$ est aussi racine de ω_3. Les coefficients de Ψ_3 et de ω_3 valent:

$$\begin{aligned} a_3 &= g_{111}^2 + g_{222}^2,\\ a_2 &= 6\,(g_{122}\,g_{222} - g_{111}\,g_{112}),\\ a_1 &= 9\,(g_{122}^2 + g_{112}^2) + 6\,(g_{112}\,g_{222} + g_{111}\,g_{122});\\ d_2 &= a_2/6 \ = \ g_{122}\,g_{222} - g_{111}\,g_{112},\\ d_1 &= a_1/3 - a_3. \end{aligned}$$

Pour Ψ_4 et ω_4, il est utile d'introduire des variables auxiliaires pour alléger les écritures:

$$\begin{aligned} t &= 16\,(g_{1112}^2 + g_{1222}^2),\\ u &= g_{1111} + g_{2222} - 6\,g_{1122},\\ v &= 4\,(g_{1222} - g_{1112}),\\ w &= 6\,g_{1122}\,(g_{1111} + g_{2222}); \end{aligned}$$

et nous avons [5]:

$$\begin{aligned} b_4 &= g_{1111}^2 + g_{2222}^2,\\ b_3 &= -8\,(g_{1111}\,g_{1112} - g_{1222}\,g_{2222}),\\ b_2 &= 4\,b_4 + t + 2\,w,\\ b_1 &= 4\,b_3 - 2\,u\,v,\\ b_0 &= 2\,(b_4 + t + 2\,w + 36\,g_{1122}^2 + 2\,g_{1111}\,g_{2222}\\ &\quad +32\,g_{1112}\,g_{1222});\\ c_4 &= -b_3/8 \ = \ g_{1111}\,g_{1112} - g_{2222}\,g_{1222},\\ c_3 &= 2\,b_4 - b_2/4 \ = \ b_4 - (t + 2\,w)/4,\\ c_2 &= 3\,b_3/2 - 3\,b_1/8 \ = \ 3\,u\,v/4,\\ c_1 &= b_2 - b_0/2,\\ c_0 &= b_1/2 \ = \ 2\,b_3 - u\,v; \end{aligned}$$

On peut toujours choisir la tangente θ dans l'intervalle $]-1,1]$ sans restreindre la généralité, de sorte qu'il existe une correspondance bijective entre les variables ξ et θ. Les fonctions Ψ_i étant bornées et continues, les nombres de maxima et de minima ne peuvent différer de plus de 1. Autrement dit, une condition suffisante pour que Ψ_i admette un maximum unique est que le polynôme ω_i admette au plus deux racines réelles. Ceci est assuré pour Ψ_3, mais cela demande une démonstration pour Ψ_4.

Le problème consiste donc à tester si le polynôme de degré 4, $\omega_4(\xi)$, a ou non exactement 2 racines réelles, ce qui peut se faire en construisant une suite de Sturm ([9], tome 2). Le test se ramène à tester si un "discriminant" est négatif, ce qui est malheureusement de nouveau difficilement décidable, comme nous allons le voir maintenant.

5.3 Suite de Sturm

Pour tester le nombre de racines réelles d'un polynôme $f_1(\xi)$, on construit la suite de Sturm suivante, par divisions euclidiennes successives [9], et on obtient:

$$\begin{array}{rcl} f_1(\xi) & = & a\xi^4 + b\xi^3 + c\xi^2 + d\xi + e; \\ f_2(\xi) & = & f_1'(\xi) \;=\; 4a\xi^3 + 3b\xi^2 + 2c\xi + d; \\ f_3(\xi) & = & \{\Delta_1\xi^2 + 2D_2\xi + D_3\} \,/\, 16a; \\ f_4(\xi) & = & 4a\{\Delta_6\xi + D_7\} \,/\, \Delta_1^2; \\ f_5(\xi) & = & \Delta_8 \,/\, 16a\Delta_6^2; \end{array}$$

$$\begin{array}{rcl} \text{avec : } \Delta_1 & = & 3b^2 - 8ac; \\ D_2 & = & bc - 6ad; \\ D_3 & = & bd - 16ae; \\ D_4 & = & cd - 6be; \\ D_5 & = & 4c^2 - 9bd; \\ \Delta_6 & = & \Delta_1 D_3 - 4D_2^2 + \Delta_1 D_5; \\ D_7 & = & 2(\Delta_1 D_4 - D_2 D_3); \\ \Delta_8 & = & 2D_2\Delta_6 D_7 - \Delta_1 D_7^2 - D_3\Delta_6^2. \end{array}$$

On note $V(\xi)$ le nombre de changements de signe dans la séquence $\{f_1(\xi), .., f_5(\xi)\}$. Le nombre de racines réelles distinctes de $\omega_4(\xi)$ est égal à la différence $V(-\infty) - V(+\infty)$. Si c_4 est positif, ce que l'on peut toujours admettre, alors:

$$\begin{array}{rcl} sign\{f_2(\epsilon\infty)\} & = & \epsilon\, sign\{a\}; \\ sign\{f_3(\epsilon\infty)\} & = & sign\{a\,\Delta_1\}; \\ sign\{f_4(\epsilon\infty)\} & = & \epsilon\, sign\{a\,\Delta_6\}; \\ sign\{f_5(\epsilon\infty)\} & = & sign\{a\,\Delta_8\}. \end{array}$$

Après analyse des différentes combinaisons de signes de $\{\Delta_1, \Delta_6, \Delta_8\}$, on s'aperçoit que le polynôme $f_1(\xi)$ a deux racines distinctes *si et seulement si* $a\Delta_8 < 0$, indépendamment du signe des deux autres Δ_i. Après de longs tâtonnements, il n'a pas été possible de prouver que $\omega_4(\xi)$ vérifiait cette inégalité, en utilisant les propriétés de la section 2. En revanche, si le modèle n'est pas bruité, il a été établi dans [5] que Ψ_4 admet un maximum unique, obtenu pour $g_{1122}\xi = g_{1112} - g_{1222}$.

6 Conclusion

Les contributions des sections 2, 3, et 4, bien qu'ayant un intérêt en elles-mêmes, n'ont pas permis de conclure par une démonstration de convergence globale dans la section 5. Le seul résultat définitif valable en présence de bruit non gaussien que nous ayons obtenu est que le contraste Ψ_3 admet un maximum unique en dimension 2. Nous espérons néanmoins que certains lecteurs seront plus doués et pourront terminer une des deux démonstrations proposées. Il a été en effet constaté expérimentalement par plusieurs chercheurs de différents laboratoires que l'algorithme de type Jacobi maximisant Ψ_4 ne restait jamais bloqué dans un maximum local. Nous sommes obligés d'admettre pour l'instant sans démonstration les conjectures suivantes:

Conjecture 1: Dans un grand nombre de cas (restant à préciser) en dimension 2, la fonction de contraste Ψ_4 admet un maximum unique dans son domaine minimal de définition ($\theta \in]-1, 1]$).

Conjecture 2: Si à chaque itération le maximum absolu du contraste Ψ_r, $r \in \{3, 4\}$, est obtenu par rapport à l'ensemble des rotations planes, l'algorithme de type Jacobi, lorsqu'il devient stationnaire, atteint le maximum global de Ψ_r par rapport à l'ensemble du groupe des rotations.

Références

[1] J.F. CARDOSO. Source separation using higher order moments. In *Proc. ICASSP* Glasgow, pages 2109–2112, 1989.

[2] J.F. CARDOSO. Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem. In *Proc. ICASSP* Albuquerque, pages 2655–2658, 1990.

[3] J.F. CARDOSO. Fourth-order cumulant structure forcing, application to blind array processing. In *Proc. IEEE SP Workshop on SSAP*, October 1992.

[4] P. COMON. Separation of stochastic processes. In *Proc. Workshop on Higher-Order Spectral Analysis*, Vail, Colorado, pages 174–179. IEEE-ONR-NSF, June 1989.

[5] P. COMON. Analyse en composantes indépendantes et identification aveugle. *Traitement du Signal*, 7(5):435–450, Dec. 1990.

[6] P. COMON. Independent Component Analysis, a new concept ? to appear in *Signal Processing*, 1993. submitted on March 12, 1992.

[7] P. COMON and J. L. LACOUME. Statistiques d'ordres superieurs pour le traitement du signal. Ecole Predoctorale de Physique, Les Houches, 30 aout – 10 septembre 1993. P. Flandrin et J. L. Lacoume ed.

[8] L. FETY. *Methodes de Traitement d'Antenne Adaptees aux Radiocommunications*. Doctorat, ENST, 1988.

[9] F. R. GANTMACHER. *Theorie des matrices*. Dunod, 1966.

[10] G. H. GOLUB and C. F. VAN LOAN. *Matrix computations*. The John Hopkins University Press, 1989.

[11] M. KENDALL and A. STUART. *The Advanced Theory of Statistics*. C. Griffin, 1977.

[12] J. L. LACOUME, editor. *Higher-Order Statistics*. Elsevier, 1992.

[13] P. McCULLAGH. *Tensor Methods in Statistics*. Monographs on Statistics and Applied Probability. Chapman and Hall, 1987.

[14] P. RUIZ and J.L. LACOUME. Extraction of independent sources from correlated inputs. In *Proc. Workshop on Higher-Order Spectral Analysis*, Vail, Colorado, pages 146–151, June 1989.

[15] A. SOULOUMIAC. *Utilisation des statistiques d'ordre supérieur pour la séparation et le filtrage*. Doctorat, ENST, Février 1993.