



noyau de R_s . Si on convient maintenant d'associer à b_i , la $i^{\text{ème}}$ colonne de B , un polynôme $B_i(q^{-1})$ de degré $P-1$, de la manière dont on a associé $A(q^{-1})$ à $a(2)$, alors à $v_2(i)$, la $i^{\text{ème}}$ colonne de V_2 , sera associée le polynôme $A(q^{-1}) B_i(q^{-1})$. Cette observation est la base de la généralisation de la méthode de Pisarenko.

ESTIMATION DE LA MATRICE DE COVARIANCE

Pour la matrice \hat{R}_y estimée à l'aide de T observations $\{y_1, y_2, \dots, y_T\}$ les propriétés précédentes ne sont vraies qu'asymptotiquement en T . Nous proposons d'estimer R_y de la manière suivante, parfois appelé méthode de covariance, qui fournit une estimée non biaisée et consistante.

$$\hat{R}_y = \frac{1}{T-N+1} \sum_{n=N}^T \phi_y(n) \phi_y(n)^T \tag{7}$$

$$\phi_y^T(n-1) = [y_{n-1} \ y_{n-2} \ \dots \ y_{n-N}]$$

Dans la pratique, on "centro-symétrisera" cette estimée. Si dans la théorie qui suit ce type de modification est sans réelle importance, il n'en est pas de même en pratique où le nombre T d'observations est souvent limité. Le choix de N , l'ordre de la matrice, que nous n'abordons pas, est alors important. Notons également dans ce contexte que "toeplitziser le long des diagonales" n'a pas vraiment de sens pour un signal temporel scalaire mais que l'on pourrait songer à un estimateur de la forme :

$$(1/\Sigma \alpha_i) \cdot \Sigma \alpha_i \phi_y(i) \phi_y(i)^T$$

(une fenêtre d'apodisation sur les dyades) pour se rapprocher d'une structure de Toeplitz sans perdre la propriété de rang.

En définissant maintenant $\phi_s(\cdot)$ et $\phi_e(\cdot)$ de manière analogue à $\phi_y(\cdot)$ (7) on a :

$$\phi_y(\cdot) = \phi_s(\cdot) + \phi_e(\cdot)$$

$$\hat{R}_y = \hat{R}_s + \sigma_e^2 I + \Delta$$

avec
$$\hat{R}_s = \frac{1}{T-N+1} \sum \phi_s(n) \phi_s(n)^T \tag{8}$$

$$\Delta = \Delta_1 + \Delta_2$$

$$\Delta_2 = \frac{1}{T-N+1} \sum \phi_e(n) \phi_e(n)^T - \sigma_e^2 I$$

et avec Δ_1 les termes croisés. On vérifie alors que \hat{R}_s est une matrice de rang $2L$ admettant les colonnes de $A(6)$ comme base de son noyau, que Δ_1 est symétrique de rang $4L$ et que Δ_2 est génériquement de rang plein. Il apparaît alors que \hat{R}_y peut être vu comme la somme d'une matrice exacte $\hat{R}_s + \sigma_e^2 I$ et d'une matrice de perturbation Δ .

RESULTATS DE PERTURBATION DE MATRICE

Afin de pouvoir préciser les perturbations subies par les éléments propres de la matrice "exacte" $R_s + \sigma_e^2 I$ nous définissons sa décomposition :

$$\begin{aligned} \hat{R}_s + \sigma_e^2 I &= U D U^T \\ &= U_1 D_1 U_1^T + \sigma_e^2 U_2 U_2^T \end{aligned} \tag{9}$$

$$D_2 = \sigma_e^2 I_P \ ; \ D_1 = \text{diag}((\lambda_i))$$

avec : $\lambda_1 \gg \lambda_2 \gg \dots \gg \lambda_{2L} > \sigma_e^2$
 et
$$U_2 = AB \tag{10}$$

Bien que la matrice U_2 ne soit définie qu'à une transformation orthogonale près, la matrice A (6) dans la décomposition (10) de U_2 est invariante. On vérifie que les composantes des matrices Δ_1 et Δ_2 (8) sont des estimées de covariances de e_n ou de covariances croisées entre e_n et s_n , elles sont donc asymptotiquement en T , des échantillons de variables gaussiennes centrées dont la variance est d'ordre T^{-1} (notée $O(T^{-1})$). On peut donc écrire $\Delta_1 = O(T^{-2})$. Finalement en désignant par $\hat{U}_1, \hat{U}_2 \dots$ les quantités perturbées/estimées associées à $U, U_1 \dots$ on a la proposition suivante.

Proposition : Pour T suffisamment grand, étant donné une matrice perturbée \hat{R}_y (7,8) il existe une base, notée \hat{U}_2 , du noyau de \hat{R}_s (9) pour laquelle les relations suivantes sont vérifiées :

$$\begin{aligned} \hat{U}_1 &= U_1 + O(T^{-\frac{1}{2}}) \\ \hat{U}_2 &= \check{U}_2 + O(T^{-\frac{1}{2}}) \\ \hat{D} &= D + O(T^{-\frac{1}{2}}) \end{aligned} \tag{11}$$

et plus précisément :

$$\begin{aligned} \hat{U}_2 - \check{U}_2 &= -U_1(D_1 - \sigma_e^2 I)^{-1} U_1^T \Delta \check{U}_2 + O(T^{-1}) \\ U_2^T(\hat{U}_1 - U_1) &= (D_1 - \sigma_e^2 I)^{-1} U_2^T \Delta U_1 + O(T^{-1}) \quad \blacktriangle \end{aligned} \tag{12}$$

En utilisant ces résultats déterministes et la définition (8) des perturbations Δ_1 et Δ_2 qui peuvent être vues comme des matrices aléatoires, il est possible de calculer la loi des différents vecteurs propres estimés $\hat{u}_2(i)$, $i=1$ à P de \hat{U}_2 . Ces vecteurs sont, asymptotiquement en T , gaussiens de moyenne $\check{u}_2(i)$ et de variance Q_{ij} de l'ordre de T^{-1} . L'expression de

$$\begin{aligned} Q_{ij} &= E(\tilde{u}_2(i) \tilde{u}_2(j)^T) \\ \tilde{u}_2(i) &= \hat{u}_2(i) - \check{u}_2(i) \end{aligned} \tag{13}$$

est donnée en annexe A. Notons au passage que l'effet de Δ_1 est négligeable dans les termes du 1^{er} ordre que nous considérons. Des résultats analogues sur les perturbées des valeurs propres [4] mènent à des tests statistiques permettant d'estimer le nombre L de sinusoïdes présentes.

METHODE DE PISARENKO GENERALISEE

En utilisant les résultats (12,13) qui précèdent nous indiquons comment utiliser de manière optimale l'information contenue dans \hat{U}_2 pour obtenir des estimées des pulsations ω_k (2) de variance minimale. Les détails sont donnés en Annexe B. Dans le texte nous nous contentons d'exposer l'idée générale. Nous présentons également de façon encore plus succincte la manière de retrouver ces mêmes estimées à partir de l'espace signal \hat{U}_1 . Dans les 2 cas l'approche est paramétrique puisqu'elle passe par l'estimation des coefficients \underline{a} du polynôme $A(q^{-1})$ (2).

La relation (10) est à la base de la méthode. Elle se réécrit, en utilisant les notations introduites à la suite de (6), sous forme matricielle et polynomiale :

$$\begin{aligned} u_2(i) &= A b_i \quad i = 1 \text{ à } P \tag{14} \\ u_{2i}(q^{-1}) &= A(q^{-1}) B_i(q^{-1}) \end{aligned}$$

De même que l'on a associé au vecteur (2) \underline{a} de dimension $(2L+1,1)$ la matrice (6) $A(N,P)$ avec $P=N-(2L+1)+1$, nous allons associer au $i^{\text{ème}}$ vecteur colonne b_i de B de dimension $(P,1)$ une matrice notée $B_i(N,2L+1)$ et vérifier que

$$u_2(i) = A b_i = B_i \underline{a} \tag{15}$$

relation qui traduit la commutativité du produit des polynômes $A(q^{-1})$ et $B_i(q^{-1})$. Mais (14) signifie en fait que tous les polynômes $u_{2i}(\cdot)$, $i = 1$ à P pouvaient

se factoriser en produit de deux polynômes dont l'un $B_i(q^{-1})$ dépend de i et dont l'autre $A(q^{-1})$ n'en dépend pas et est le polynôme (2) qui nous intéresse pour estimer les pulsations ω_k .

L'idée de la méthode de Pisarenko généralisée est la suivante. On associe le polynôme $\hat{u}_{2i}(q^{-1})$ à $\hat{u}_2(i)$ de \hat{U}_2 et on le factorise en $\hat{A}_i(q^{-1}) \cdot \hat{B}_i(q^{-1})$ avec $\hat{A}_i(q^{-1})$ une estimée de $A(q^{-1})$. A l'aide des résultats (12,13) on sait alors, pour chacune des P estimées $\hat{A}_i(q^{-1})$, évaluer la loi de \hat{a}_i , le vecteur des coefficients associé, ainsi que les lois jointes de \hat{a}_i et \hat{a}_j . De l'ensemble de ces estimées et de leurs lois, on déduit alors une estimée unique de variance minimale. Dans la pratique pour effectuer la factorisation de $\hat{u}_{2i}(q^{-1})$ en produit de $\hat{A}_i(q^{-1})$ et $\hat{B}_i(q^{-1})$, on commence par chercher les racines des P polynômes $\hat{u}_{2i}(q^{-1})$ et on reconnaît celles qui sont à retenir pour former les estimées $\hat{A}_i(q^{-1})$ par le fait qu'elles sont complexes conjuguées, proches ou sur le cercle unité et surtout "communes" à tous les polynômes.

Nous indiquons en Annexe B les différentes étapes ultérieures du calcul et notamment la manière de déduire de la variance (13) du vecteur propre $\hat{u}_2(i)$, la variance du vecteur \hat{a}_i des coefficients du polynôme $\hat{A}_i(q^{-1})$.

On peut remarquer que cette variance va dépendre du polynôme $\hat{B}_i(q^{-1})$ qui dépend fortement des données (de la réalisation du bruit) observées. En effet si on revient à la relation (10), la matrice carrée B est définie, comme U_2 la base du sous-espace propre associé à la valeur propre multiple σ_e^2 , à une transformation orthogonale près. Mais si la variance des différentes estimées \hat{a}_i dépend ainsi de la réalisation du bruit, on vérifie qu'il n'en est pas de même pour l'estimée unique, finale et optimale dont la variance est indépendante de la matrice B (ou plutôt \hat{B} puisque la base estimée \hat{U}_2 est proche de la base exacte $U_2 = AB$). On peut alors se demander si parmi toutes les bases possibles U_2 il n'en existe pas une pour laquelle les différentes estimées \hat{a}_i sont indépendantes, ou encore mieux une pour laquelle toutes les estimées sont identiques, totalement corrélées et par conséquent de variance minimale ?

Avant de présenter quelques résultats numériques dans le cas d'une sinusoïde unique, nous allons nous intéresser à la manière (duale) d'obtenir cette même estimée de variance minimale en partant de la base U_1 de l'espace signal. La relation de départ n'est plus alors l'équation (10) mais

$$A^T U_1 = 0 \quad (16)$$

qui exprime l'orthogonalité entre U_2 et U_1 . (On peut noter que l'ambiguïté due à la non-unicité de U_2 a disparu, puisque la base U_1 est définie de manière unique). Pour le i ème vecteur, $i = 1$ à $2L$, $u_1(i)$ de U_1 on a alors

$$A^T u_1(i) = 0$$

Cette relation, vu la structure de la matrice A (6), peut se mettre sous la forme d'un système linéaire de $P = N - 2L$ équations non homogènes à $2L$ inconnues (les $2L$ premières composantes de \underline{a} (2) par exemple). Si P est alors supérieur à $2L$, cela semble indiquer que chaque vecteur propre signal contient suffisamment d'information pour estimer \underline{a} et donc les pulsations. Mais rien ne garantit en fait que ce système soit de rang plein et permettre de déterminer l'ensemble du vecteur \underline{a} . (Pour des vecteurs sources orthogonaux, il est bien sûr de rang déficient). Il faut également observer que $u_1(i)$ comme vecteur propre d'une matrice de Toeplitz symétrique associé à une valeur propre simple est "symétrique" ou "antisymétrique" et que le vecteur \underline{a} est lui-même symétrique. Le système précédent est donc largement redondant et se simplifie en fait en un système de N_e équations non homogène à L inconnues (les éléments a_1 à a_L de \underline{a} (2)) avec :

$$N_e = \text{partie entière} \left\{ \frac{N+1-2L}{2} \right\}$$

En regroupant alors toutes les équations obtenues pour

les $2L$ vecteurs de U_1 on aboutit à un système largement surdéterminé de $2LN_e$ équations à L inconnues. En utilisant les résultats (12) on sait alors calculer la loi de la solution de ce système au sens des moindres carrés non dans la norme euclidienne associée à la matrice identité d'ordre $2LN_e$ mais plus généralement dans une norme associée à une matrice définie positive Q . La loi, la variance de la solution (de l'estimée) est alors une fonction de Q dont on sait trouver le minimum (en Q). En cherchant alors la solution dans cette norme Q , elle a pour variance la même valeur minimale que celle obtenue plus haut en exploitant de manière optimale U_2 . Il s'agit évidemment d'appliquer ce raisonnement à \hat{U}_1 et la matrice Q dont on obtient une estimée traduit alors le fait que les $2LN_e$ différentes équations du système ne sont pas connues avec la même précision et sont corrélées entre elles.

SIMULATION ET CONCLUSIONS

On prend T échantillons du processus

$$y_n = \frac{1}{\sqrt{2}} \cos(n) + e_n$$

avec e_n un bruit blanc gaussien de variance unité et on estime la pulsation $\omega = 1$ de l'unique sinusoïde ($L = 1$) avec la méthode "optimale" précédente d'une part et avec le "goniomètre" d'autre part pour des matrices \hat{R}_N d'ordre N allant de 3 à 16. La variance des estimées (pour T grand) est de la forme V/T et dans le tableau nous donnons les valeurs théoriques de V en fonction de N .

N	Pisarenko généralisée	Goniomètre
3	0,2618	0,2618
4	0,2607	0,2621
5	0,0799	0,1166
6	0,0275	0,0399
7	0,0262	0,0309
8	0,0168	0,0224
9	0,0081	0,0128
10	0,0074	0,0098
12	0,0034	0,0055
16	0,0015	0,0023

On peut noter que pour le goniomètre la variance ne diminue pas systématiquement (entre $N = 3$ et 4) pour N croissant, ce qui est caractéristique d'une méthode non optimale. Le gain de performance est net mais difficile à apprécier dans la mesure où pour un signal temporel N , l'ordre de la matrice, n'est pas fixé et qu'on ne peut l'augmenter indéfiniment pour T donné. En traitement d'antenne, la situation est différente puisque N est fixé par la taille de l'antenne. Il serait alors intéressant d'étudier le pouvoir discriminant de cette approche.

ANNEXE A

$$Q_{ij} = R_s^+ S_{ij} R_s^+$$

$$R_s^+ = U_1^T (D_1 - \sigma_e^2 I)^{-1} U_1$$

pour simplifier l'écriture nous remplaçons $\check{u}_2(k)$ par \check{u}_k dans l'expression qui suit :

$$S_{ij} = \frac{\sigma_e^4}{T} \left\{ \check{u}_j^T \check{u}_i^T + \check{u}_i^T \check{u}_j^T I_{ij} + \sum_{k=1}^{N-1} (S^k \check{u}_j^T \check{u}_i^T S^k + S^T k \check{u}_j^T \check{u}_i^T S^T k + (\check{u}_i^T S^k \check{u}_j^T)^k + (\check{u}_i^T S^T k \check{u}_j^T)^k) \right\}$$

où S est la matrice carrée d'ordre N dont l'élément i, j est égal à $\delta_{i, j-1}$ (matrice de décalage). Dans la pratique on obtient une estimée de C_{ij} en remplaçant les quantités exactes inconnues par leurs estimées.



ANNEXE B

RÉFÉRENCES

Avec les notations introduites dans le texte (14,15) on a successivement

$$\hat{u}_2(i) = \hat{A}_i \hat{b}_i = \hat{E}_i \hat{a}_i$$

où dans \hat{a}_i nous convenons de normaliser la dernière composante à 1. Ceci entraîne que dans $\hat{a}_i = \hat{a}_i - \frac{a}{a_i}$ la dernière composante est nulle et nous obtenons \hat{a}_i le vecteur de dimension réduite obtenue en enlevant cette dernière composante

$$\begin{aligned} \check{u}_2(i) &= \check{u}_2(i) - \hat{u}_2(i) \\ &= A(\check{b}_i - \hat{b}_i) + (A - \hat{A}_i)\hat{b}_i \\ &= A \check{b}_i + \hat{A}_i \hat{b}_i \\ &\cong A \check{b}_i + \check{B}_i \check{a}_i \\ &= A \check{b}_i + \check{B}_i \check{a}_i \end{aligned}$$

avec $\check{B}_i(N, 2L)$ les 2L premières colonnes de \check{B}_i . Cette dernière relation s'écrit sous forme partitionnée :

$$\begin{bmatrix} \check{u}_{21}(i) \\ \check{u}_{22}(i) \end{bmatrix} = \begin{bmatrix} \check{B}_{i1} & A_1 \\ \check{B}_{i2} & A_2 \end{bmatrix} \begin{bmatrix} \check{a}_i \\ \check{b}_i \end{bmatrix}$$

d'où l'on déduit

$$\begin{aligned} \check{a}_i &= (\check{B}_{i1} - A_1 A_2^{-1} \check{B}_{i2})^{-1} \check{u}_{21}(i) \\ & \quad [I_{2L} \quad - A_1 A_2^{-1}] \check{u}_2(i) \end{aligned} \quad (B1)$$

Le vecteur $\check{u}_2(i)$ étant gaussien, centré de variance Q_{ij} (annexe A et (13)), le vecteur \check{a}_i est également gaussien, centré de variance aisément déduite de (B1) et Q_{ij} . Nous noterons $\Sigma_{ij} = E(\check{a}_i \check{a}_j^T)$.

La matrice Σ_{ij} est en fait singulière. Les 2L composantes de \check{a}_i sont redondantes (\check{a} (2) est un vecteur symétrique), elles ne dépendent que de L paramètres comme par exemple les L pulsations ω_k que nous voulons estimer. En notant alors $\underline{\omega}$ et $\hat{\underline{\omega}}_i$ les vecteurs de dimension (L,1) des pulsations exactes et des estimées obtenues à l'aide de $\hat{u}_2(i)$, on a $\hat{\underline{\omega}}_i = G \check{a}_i$ avec G une matrice (L,2L) exprimant le gradient de $\underline{\omega}$ par rapport à \check{a} . Il en découle

$$\Sigma_{ij}^r = E(\hat{\underline{\omega}}_i \hat{\underline{\omega}}_j^T) = G \Sigma_{ij} G^T$$

avec Σ_{ij}^r une matrice de dimension réduite (L,L) et Σ_{ii}^r une matrice inversible.

On introduit maintenant les notations suivantes :

$$\hat{\underline{\Omega}}^T = [\hat{\underline{\omega}}_1^T : \hat{\underline{\omega}}_2^T : \dots : \hat{\underline{\omega}}_P^T] \quad (1, PL)$$

$$\mathbb{1}^T = [I_L : I_L : \dots : I_L] \quad (L, PL)$$

$$\Sigma = ((\Sigma_{ij}^r)) \quad (PL, PL)$$

où dans le vecteur colonne $\hat{\underline{\Omega}}$ on a réuni les P estimées de $\underline{\omega}$. On sait que ce vecteur est asymptotiquement gaussien de moyenne $\mathbb{1} \underline{\omega}$ et de variance Σ , l'estimée au maximum de vraisemblance de $\underline{\omega}$ utilisant cette observation $\hat{\underline{\Omega}}$ est alors

$$\hat{\underline{\omega}}^* = (\mathbb{1}^T \Sigma^{-1} \mathbb{1})^{-1} \mathbb{1}^T \Sigma^{-1} \hat{\underline{\Omega}}$$

et sa variance est égale à $(\mathbb{1}^T \Sigma^{-1} \mathbb{1})^{-1}$.

[1] H. MERMOZ, "Imageries, corrélations et modèles", Annales des Télécommunications, vol. 31, n° 1-2, pp. 17-36, Fév. 1976.

[2] G. BIENVENU, L. KOPP, "Optimality of high resolution array processing using the eigensystem approach", IEEE-ASSP-31, pp. 1235-1248, 1983.

[3] V.F. PISARENKO, "The retrieval of harmonics from covariance functions", Geophys. J. Royal Astronomical Soc., pp. 347-366, 1973.

[4] J.J. FUCHS, "Estimation du nombre de sinusoides dans un bruit coloré", 11e Colloque GRETSI, Nice, Juin 1987.