

Détection et estimation de sinusôides dans du bruit coloré de matrice de covariance inconnue

Nicolas MOAL, Jean-Jacques FUCHS

IRISA/Université de Rennes I
Campus de Beaulieu - 35042 Rennes Cedex - France
nmoal@irisa.fr, fuchs@irisa.fr

Résumé — Nous considérons le problème de la détection et de l'estimation des paramètres de sinusôides réelles noyées dans du bruit coloré de matrice de covariance inconnue. Nous supposons que le bruit peut être modélisé par un processus ARMA. Nous étendons à la situation présente une approche basée sur de la programmation quadratique qui a déjà été appliquée avec succès au cas du bruit blanc. Nous construisons des bases redondantes spécifiques pour les sinusôides et pour le bruit. Nous reconstruisons les covariances estimées des observations comme une combinaison linéaire parcimonieuse des éléments de ces bases.

Abstract — We address the issue of the detection and estimation of the parameters of real sinusoids embedded in additive Gaussian noise with unknown covariance matrix. We assume that the noise can be modeled by an autoregressive moving average process. We extend the quadratic programming approach that we have already successfully applied to the white noise case to the present situation. We design specific redundant bases for the sinusoids and for the noise and we reconstruct the estimated covariances of the observations as a parsimonious linear combination of elements of these bases.

1 Introduction

La détection et l'estimation des paramètres de sinusôides noyées dans du bruit gaussien constituent des problèmes importants qui apparaissent dans de nombreux domaines allant du radar au sonar, des radio communications à la sismologie. Beaucoup de travaux ont été publiés sur le développement et l'étude des performances d'algorithmes pour de tels signaux. La grande majorité d'entre eux considère le cas où le bruit additif est blanc gaussien. Le cas où le bruit n'est pas blanc mais de matrice de covariance connue peut être ramené au cas précédent par une opération de *blanchiment*.

Lorsque le bruit additif a une *fonction de covariance inconnue*, le problème devient plus difficile. Il faut en général estimer conjointement les paramètres des sinusôides et du bruit pour obtenir une estimée non biaisée du signal. Récemment, on a proposé d'utiliser les méthodes haute-résolution pour résoudre ce problème. La plupart de ces méthodes cependant ont été initialement conçues pour du bruit blanc additif et conduisent par conséquent à des estimées médiocres dans le cas du bruit coloré. Dans [4] nous avons présenté un algorithme basé sur de la programmation quadratique pour la détection et l'estimation de sinusôides dans bruit blanc. Ici nous proposons une extension de cette méthode au cas d'un bruit coloré inconnu. Nous illustrons ses bonnes performances par plusieurs simulations numériques.

2 Formulation du problème

Nous supposons que nous observons T échantillons d'un signal bruité, une somme de sinusôides réelles noyées dans

du bruit coloré inconnu :

$$y_t = \sum_{i=1}^P A_i \cos(2\pi f_i t + \phi_i) + n_t, \quad t = 1, \dots, T \quad (1)$$

où le bruit et les sinusôides sont des processus indépendants. Les amplitudes $A_i > 0$ et les fréquences normalisées f_i ($f_i \in]0, 1/2[$) sont supposées constantes inconnues. Les phases ϕ_i sont des variables aléatoires indépendantes uniformément distribuées sur $[0, 2\pi[$. Ce dernier point garantit la stationnarité des données.

Lorsque le bruit n_t a une densité spectrale de puissance inconnue, on suppose en général qu'il peut être approximé par une fonction de transfert rationnelle (des modèles autorégressifs (AR), moyenne mobile (MA) ou autorégressifs à moyenne mobile (ARMA)). Dans notre cas, nous supposons que le processus du bruit peut être modélisé par un modèle ARMA, c'est-à-dire qu'il satisfait la relation :

$$n_t + \sum_{i=1}^p a_i n_{t-i} = e_t + \sum_{j=1}^q c_j e_{t-j} \quad (2)$$

où n_t est un bruit blanc gaussien de variance σ_e^2 .

Etant données T observations, nous voulons estimer le nombre de sinusôides P , et pour chacune d'entre elles ses paramètres utiles, son amplitude A_i et sa fréquence normalisée f_i ($\omega_i = 2\pi f_i$).

Comme nous allons le voir rapidement, les covariances des y_t contiennent toutes ces informations, c'est pourquoi nous proposons une méthode d'estimation reposant sur le vecteur de ces covariances r . Nous allons préciser l'expression générale des covariances de (1), ceci afin de construire une base redondante adaptée à notre problème. Nous chercherons alors à reconstruire r comme une combinaison linéaire parcimonieuse d'éléments de cette base ([1], [4]).

3 Covariances

3.1 Covariances des observations

Commençons par préciser la forme générique des covariances du signal (1). En raison de l'indépendance du bruit additif et des sinusoides, nous avons ainsi :

$$r_k = r_k^s + r_k^n$$

où r_k^s et r_k^n sont respectivement les covariances des sinusoides et du bruit additif ARMA. Plus précisément :

$$r_k^s = \sum_{i=1}^P \frac{A_i^2}{2} \cos k\omega_i$$

tandis que les r_k^n sont de la forme suivante :

$$r_k^n = \sum_{i=1}^p \alpha_i x_i^n \quad (3)$$

où les x_i sont les pôles de la partie AR du bruit :

$$x_i^p + \sum_{l=1}^p a_l x_i^{p-l} = 0$$

Nous trouverons donc deux types de termes dans l'expression générale des covariances d'un processus ARMA : des exponentielles amorties si le pôle est réel et des sinusoides amorties si les pôles sont complexes conjugués. En résumé, pour l'indice k nous avons :

$$r_k^n = \sum_i \alpha_i a_i^k + \sum_j \rho_j^k \{ \beta_j \cos(k\theta_j) + \gamma_j \sin(k\theta_j) \} \quad (4)$$

où les a_i sont les pôles réels et les $\rho_j e^{\pm i\theta_j}$ les pôles complexes conjugués.

3.2 Covariances estimées

Dans la pratique, les covariances r des données doivent être estimées. Par la suite nous considérerons donc le vecteur des estimées biaisées des covariances (1)

$\hat{r} = [\hat{r}_0, \dots, \hat{r}_{K-1}]^T$ où :

$$\hat{r}_k = \frac{1}{T} \sum_{t=1}^{T-k} y(t)y(t+k)$$

Dans le cas du bruit blanc, le comportement asymptotique (quand $T \rightarrow \infty$) de ces estimées est bien connu (cf. [5]). Le théorème qui suit donne une extension de ce résultat dans le cas d'un bruit additif de type ARMA.

3.3 Comportement asymptotique des erreurs d'estimation

Nous avons montré le théorème suivant (dont nous avons omis la démonstration pour des questions de place) :

Théorème

Les erreurs d'estimation normalisées $\sqrt{T}(\hat{r}_k - r_k)$, $k = 0, 1, \dots, K$ des covariances r_k du processus y_t dans (1) ont asymptotiquement une distribution gaussienne centrée de matrice de covariance Σ :

$$\sqrt{T} \begin{bmatrix} (\hat{r}_0 - r_0) \\ \vdots \\ (\hat{r}_K - r_K) \end{bmatrix} \xrightarrow[T \rightarrow \infty]{dist} \mathcal{N}(0, \Sigma) \quad (5)$$

L'élément i, j de la matrice Σ est donné par :

$$V_{i,j} = \sum_{\tau=-\infty}^{+\infty} \{ r_\tau^s r_{\tau+i-j}^n + r_{\tau-j}^s r_{\tau+i}^n + r_{\tau+i}^s r_{\tau-j}^n + r_{\tau+i-j}^s r_\tau^n + r_\tau^n r_{\tau+i-j}^n + r_{\tau-j}^n r_{\tau+i}^n \} \quad 0 \leq i, j \leq K \quad (6)$$

où les r_τ^s sont les covariances des sinusoides et les r_τ^n les covariances du bruit gaussien additif (ARMA).

Dans le cas particulier du bruit blanc additif, nous retrouvons le résultat de [5].

Ce résultat est intéressant d'un point de vue théorique mais malheureusement peu utilisable en pratique. Pour cette raison, nous préférons plutôt utiliser l'expression approchée suivante du terme i, j de la matrice Σ ([5]) :

$$\hat{\Sigma}_{i,j} = \text{cov}(\hat{r}_i, \hat{r}_j) = \frac{1}{T^2} \sum_{t=0}^{T-i-1} \sum_{s=0}^{T-j-1} \{ \hat{r}_{t-s} \hat{r}_{t-s+i-j} + \hat{r}_{t-s+i} \hat{r}_{t-s-j} \} \quad (7)$$

4 Construction d'une base adaptée

Nous avons vu à la section précédente que les covariances du signal (1) pouvaient se décomposer comme une combinaison linéaire de sinusoides, de sinusoides amorties et d'exponentielles amorties (cf. (4)). Fort de ce résultat, nous allons construire une base redondante sur laquelle nous décomposerons ensuite le vecteur des estimées des covariances \hat{r} .

Notons \mathbf{t} le vecteur des instants discrets $\mathbf{t} = [0, 1, \dots, K-1]^T$. Considérons donc la base construite sur les trois familles de fonctions suivantes :

Sinusoides discrètes : Nous discrétisons uniformément le domaine des fréquences $f \in]0, 1/2[$ avec un pas $.5/(n_{sin} + 1)$. En notant $f_k = k * .5/(n_{sin} + 1)$ les fréquences discrètes, les premiers vecteurs de notre base sont :

$$\{ \cos(2\pi f_k \mathbf{t}), k = 1, \dots, n_{sin} \}$$

Exponentielles discrètes amorties : Nous discrétisons les modules des pôles réels sur l'intervalle $] -1, 1[$ (pour garantir la stabilité) avec un pas $1/n_{AR1}$ ($a_k = k/(n_{AR1} + 1)$, $k = 1, \dots, n_{AR1}$). Nous obtenons ainsi les vecteurs :

$$\{ (\pm a_k)^t, k = 1, \dots, n_{AR1} \}$$

Sinusoides discrètes amorties : Pour les pôles complexes conjugués nous avons deux paramètres à discrétiser, le module (sur $]0, 1[$, $\rho_k = k/(n_{mod} + 1)$) et l'argument (sur $[0, 2\pi[$, $\theta_l = 2 * \pi * l/(n_{arg} + 1)$). A chaque paire (ρ_k, θ_l) nous associons alors deux vecteurs : pour $k = 1, \dots, n_{mod}$, $l = 1, \dots, n_{arg}$,

$$\{ [1, \rho_k \cos \theta_l, \dots, \rho_k^{K-1} \cos(K-1)\theta_l]^T \}$$

$$\text{et } \{ [0, \rho_k \sin \theta_l, \dots, \rho_k^{K-1} \sin(K-1)\theta_l]^T \}$$

Nous notons $N = n_{sin} + 2 * n_{AR1} + 2 * n_{mod} * n_{arg}$ la taille de notre base, A la matrice dont les colonnes sont les vecteurs de la base que nous venons de construire et dont la norme euclidienne a été normalisée à un.

5 La méthode proposée

5.1 Réécriture du problème

Nous cherchons maintenant à décomposer le vecteur des covariances estimées \hat{r} dans la base A que nous venons de construire. Nous devons pour cela résoudre le problème linéaire sous déterminé suivant (car A a bien plus de colonnes que de lignes $N \gg K$) : $\hat{r} = r + B = AX$, $X \geq 0$, où B est le vecteur du bruit d'estimation des covariances, de matrice de covariance Σ dont nous avons précisé l'expression (7). Pour nous placer dans le cadre de [1], nous blanchissons le système précédent en prémultipliant r (respectivement B et A) par $\hat{\Sigma}^{-1/2}$. Nous définissons ainsi $\tilde{r} = \hat{\Sigma}^{-1/2}\hat{r}$ (respectivement \tilde{B} et \tilde{A}).

5.2 Critère quadratique

Nous voulons trouver une *solution parcimonieuse* (c'est-à-dire une solution avec un petit nombre de composantes non nulles) au système :

$$\tilde{r} = \tilde{A}X, \quad X \geq 0 \quad (8)$$

car nous savons que par nature le modèle (1) ne contient qu'un petit nombre P de sinusoïdes. Nous définissons ainsi le critère quadratique ([1]) :

$$\min_X \|\tilde{r} - \tilde{A}X\|_2^2 + \lambda \|X\|_1 \quad (9)$$

où $\|X\|_1$ est la norme ℓ_1 de X , $\|X\|_1 = \sum_{i=1}^n |x_i|$.

λ est un paramètre qui doit être ajusté. Pour une valeur de λ bien choisie (de l'ordre de l'écart type du bruit d'estimation des covariances ([1]), ici $\lambda \simeq 1$ car on a blanchi ce bruit B), les seules composantes de X non nulles devraient être celles associées à des fréquences et phases discrètes encadrant les vrais paramètres des sinusoïdes, ainsi que celles modélisant les covariances du bruit ARMA. La fréquence de chaque sinusoïde est alors simplement obtenue par interpolation linéaire entre les deux fréquences discrètes qui l'encadrent tandis que sa puissance ($A_i^2/2$) est donnée par la somme des pondérations associées.

Le nombre de sinusoïdes est estimé en considérant le nombre de paires de composantes non nulles associées à des sinusoïdes. Dans les cas favorables, si l'on ne regarde que la partie de la base qui modélise les sinusoïdes, seules les composantes encadrant les vraies fréquences devraient être non nulles (ou tout au moins les composantes supplémentaires qui modélisent principalement le bruit d'estimation négligeables). Dans les cas plus difficiles (très peu de données, puissance du bruit ARMA forte à certaines fréquences, faible séparation en fréquence, ...), on utilise un critère comme le MDL pour estimer le nombre de sinusoïdes en considérant les paires de composantes de sinusoïdes par amplitude décroissante.

La situation est en fait un peu plus complexe que dans [1] car dans le cas des sinusoïdes amorties qui modélisent le bruit ARMA, la notion de voisinage est un peu plus délicate à définir. On associe en effet à chaque colonne deux paramètres, la phase et le module du pôle complexe correspondant. Or chaque module devrait génériquement tomber entre deux valeurs discrètes des modules, de la

même façon que chaque phase tombera entre deux phases discrètes. Dans ce cas, les paramètres estimés du pôle complexe seront obtenus comme la moyenne pondérée de ces 3 ou 4 voisins de pondérations non nulles, et l'amplitude correspondante comme la somme des amplitudes associées à ces 3 ou 4 colonnes.

Les différents pas de discrétisation pour les sinusoïdes ($.5/(n_{sin} + 1)$), les exponentielles amorties ($1/(n_{AR1} + 1)$) et les sinusoïdes amorties ($1/n_{mod}$ et $2\pi/(n_{arg} + 1)$) sont choisis de l'ordre de l'écart type du bruit.

6 La procédure complète

Nous résumons à présent la procédure complète.

1. Estimer le vecteur des estimées non biaisées des covariances \hat{r} et la matrice de covariance des erreurs d'estimation des covariances $\hat{\Sigma}$.
2. Calculer la matrice A et blanchir le système $AX = \hat{r}$.
3. Chercher l'optimum du critère :

$$\min_X \|\tilde{A}X - \tilde{r}\|_2^2 + \lambda \|X\|_1 \quad (10)$$

Il faut choisir λ de l'ordre de l'écart-type du bruit d'estimation. Comme nous avons blanchi \hat{r} , nous prenons $\lambda = 1$.

4. Sélectionner un nombre n_{max} de paires de composantes non nulles de X , construire la matrice \tilde{A} , la sous-matrice de A dont les colonnes correspondent à ces composantes non nulles et résoudre $\tilde{A}\tilde{X} = \tilde{r}$ au sens des moindres carrés pour obtenir une estimée non biaisée de X ([1, 4]).
5. Répéter les étapes précédentes pour plusieurs valeurs de λ (2 ou 3) et de n_{max} ($n_{max} = 1, \dots$ jusqu'à un certain nombre maximum de sinusoïdes).
6. Choisir la meilleure estimée X^* en utilisant un critère comme le MDL.
7. Estimer les fréquences des sinusoïdes comme la moyenne pondérée de chaque paire de colonnes de sinusoïdes et leur amplitudes comme la somme des poids.

Une remarque : Il est intéressant de préciser ici que nous avons choisi de traiter de manière légèrement différente les colonnes associées au bruit et celles associées aux sinusoïdes. En effet, c'est un fait bien connu que le problème de l'estimation de sinusoïdes amorties/exponentielles amorties est un problème difficile ([2]), et l'on ne devrait trouver que des estimées assez grossières des covariances du bruit (la partie de la matrice correspondant au bruit est mal conditionnée). Pour cette raison, lorsque nous choisissons les colonnes pour construire la matrice \tilde{A} (qui est utilisée pour obtenir un X non biaisé), nous avons décidé de garder toutes les colonnes qui correspondent à des processus AR/ARMA plutôt qu'en choisir un petit nombre. Nous conservons ainsi le maximum d'information sur le bruit.

7 Simulations numériques

Nous présentons les résultats de plusieurs simulations numériques pour retrouver des sinusoïdes dans différents

bruits additifs. Nous considérons le cas de deux sinusoides d'amplitude unité ($A_1 = A_2 = 1$), de fréquences respectives f_1 et f_2 , noyées dans du bruit additif gaussien ARMA suivant le modèle (1). Les rapports signal-à-bruit sont définis comme $SNR_i = A_i^2/2\sigma_n^2$, où la variance du bruit est simplement $\sigma_n^2 = r_0^n$. Nous effectuons systématiquement 50 réalisations indépendantes, et nous donnons la moyenne et l'écart type des fréquences estimées. Nous ne détaillons pas le problème de la détection. Des renseignements complémentaires peuvent être trouvés dans [1, 4].

Exemple 1 Nous considérons tout d'abord 2 sinusoides dans un bruit AR(1) ([6]) avec un pôle réel en 0.5 et une variance $\sigma_n^2 = 0.5$ ($SNR = 0dB$). Le nombre de données est $T = 300$.

Nous prenons ensuite du bruit de type AR(4), ayant pour paramètres $[1 - 1.352, 1.338, -0.662, 0.240]$ et $\sigma_n^2 = 0.5$ ([6]). La densité spectrale de puissance de ce processus est plus grande dans le voisinage des fréquences du signal. Nous avons toujours $T = 300$. Nous estimons $K = 200$ covariances. La matrice A a 1200 colonnes ($n_{sin} = 200$ sinusoides, $n_{AR1} = 100$, $n_{mod} \times n_{arg} = 20 \times 20$). Dans les deux cas nous avons pris $\lambda = 1$. Le tableau 1 présente les résultats.

TAB. 1: Exemple 1

	Vraie valeur	Moyenne	Ecart-type
AR1			
f_1	0.1397	0.1398	$7.47 \cdot 10^{-5}$
f_2	0.2000	0.2001	$4.20 \cdot 10^{-5}$
AR4			
f_1	0.1397	0.1399	$3.20 \cdot 10^{-3}$
f_2	0.2000	0.2001	$9.45 \cdot 10^{-5}$

Exemple 2 Nous présentons des simulations dans le même contexte que dans [7]. Nous considérons du bruit ARMA avec la partie AR $[1, 1.4563, 0.81]$ et la partie MA $[1, 2, 1]$. Le bruit a un pôle fort en $f = 0.4$, avec un facteur d'amortissement 0.9, et pourrait ainsi être vu comme un sinusoides supplémentaire. Heureusement ce n'est pas le cas lorsque nous appliquons notre méthode. Nous considérons simplement le cas où nous avons chaque fois 16 réalisations indépendantes, chacune consistant en $T = 64$ échantillons. Nous présentons tout d'abord les résultats avec les fréquences comme dans [7] et nous considérons ensuite des cas plus difficiles.

Nous présentons finalement le cas où le bruit additif suit un modèle AR(2) ($[1, -1.058, 0.81]$), tout d'abord dans les mêmes conditions que [7], avec une variance du bruit $\sigma_n^2 = 0.1$ et $\sigma_n^2 = 1$, puis dans des cas plus délicats. Dans ce cas, le bruit a un pôle fort entre les deux sinusoides à la fréquence $f = 0.15$.

8 Conclusion

Dans [4], nous avons proposé une nouvelle méthode de détection et d'estimation de sinusoides dans du bruit blanc. Cette communication a été l'occasion de présenter une extension de ces résultats dans le cas plus général où les sinusoides sont noyées dans du bruit coloré dont

TAB. 2: Exemple 2

	Vraie valeur	Moyenne	Ecart type
ARMA(2,2)	16 réalisations		
f_1	0.1	0.0998	$4.31 \cdot 10^{-4}$
f_2	0.2	0.1997	$2.99 \cdot 10^{-4}$
ARMA(2,2)	16 réalisations		
f_1	0.1	0.0996	$5.38 \cdot 10^{-4}$
f_2	0.105	0.1069	$2.48 \cdot 10^{-4}$
AR(2)	$\sigma_n^2 = 0.1$		
f_1	0.1	0.0999	$4.20 \cdot 10^{-4}$
f_2	0.2	0.1996	$6.40 \cdot 10^{-4}$
AR(2)	$\sigma_n^2 = 1$		
f_1	0.1	0.1001	$8.75 \cdot 10^{-4}$
f_2	0.2	0.1997	$8.90 \cdot 10^{-4}$
AR(2)	$\sigma_n^2 = 1$		
f_1	0.13	0.1299	$1.30 \cdot 10^{-3}$
f_2	0.17	0.1705	$9.60 \cdot 10^{-4}$

la matrice de covariance est inconnue. C'est un problème difficile et peu souvent traité.

Nous avons ici vu comment construire une base redondante discrète sur laquelle décomposer des estimées des covariances des observations bruitées. Nous effectuons alors la recherche d'une solution parcimonieuse en définissant un critère quadratique, qui peut être facilement optimisé au moyen d'algorithmes présents dans la majorité des bibliothèques de programmes scientifiques. Cette approche présente en outre l'avantage de traiter simultanément les problèmes de détection et d'estimation. Nous avons mis en évidence les bons résultats de cette méthode par quelques simulations numériques. Nous présentons enfin une formule théorique plutôt intéressante pour la matrice des covariances des erreurs d'estimation des covariances de sinusoides dans du bruit de type ARMA.

Références

- [1] J.J. Fuchs, Une approche à l'estimation et à l'identification simultanées, Actes du seizième colloque GRETSI'97, Grenoble, pp. 1273-1276, 1997.
- [2] C. Lanczos, *Applied Analysis*, Prentice-Hall Mathematics Series, 1957.
- [3] N. Moal et J.J. Fuchs, Estimation de l'ordre et identification des paramètres d'un processus ARMA Actes du seizième colloque GRETSI'97, Grenoble, pp. 511-514, 1997.
- [4] N. Moal and J.J. Fuchs, Sinusoids in white noise: a quadratic programming approach, Proc. ICASSP 1998, Seattle, pp. 2221-2224.
- [5] B. Porat, *Digital Processing of Random Signals*, Prentice Hall, 1994.
- [6] P.J. Sherman and A.E. Frazho, High resolution spectral estimation of sinusoids in colored noise using a modified Pisarenko decomposition, ICASSP 1986, Tokyo, pp. 181-184
- [7] A. Swami and J.M. Mendel, Cumulant based approach to harmonic retrieval and related problems, *IEEE-ASSP*, Vol.39, No.5, 1991.