

Estimation d'un signal complexe à partir d'un modèle parcimonieux perturbé.

Anisia FLORESCU¹, Emilie CHOUZENOUX², Jean-Christophe PESQUET² et Silviu CIOCHINA¹

¹Politehnica University of Bucharest, Telecommunications Dept., Romania

²Université Paris-Est, LIGM, UMR CNRS 8049, 77454 Marne-la-Vallée

anisia.florescu@ugal.ro, emilie.chouzenoux@univ-mlv.fr, pesquet@univ-mlv.fr, silviu@comm.pub.ro

Thème – Représentations et modèles

Problème traité – Nous abordons le problème de l'estimation, en présence de bruit, d'un signal à valeurs complexes admettant une représentation parcimonieuse dans une famille paramétrique de vecteurs dont les paramètres sont connus de manière imprécise.

Originalité – Nous formulons ce problème d'estimation comme celui de la minimisation d'une fonction de coût non convexe sous contraintes non convexes. Un algorithme de type explicite-implicite (*forward-backward*) permet d'apporter une solution numérique à ce problème d'optimisation. La mise en œuvre de cet algorithme nécessite le calcul d'un opérateur proximal dont nous donnons la forme exacte. Nous montrons que la méthode proposée peut être vue comme une généralisation d'une technique de seuillage dur itératif.

Résultats – En utilisant des résultats récents d'analyse non lisse, nous étudions la convergence de l'algorithme d'optimisation employé. Par ailleurs, nous illustrons le bon comportement numérique de l'approche proposée sur des exemples d'analyse spectrale parcimonieuse.

1 Introduction

Dans ce travail, nous nous intéressons au problème classique de l'estimation d'un signal $\bar{x} \in \mathbb{C}^Q$ à partir d'observations $y = \bar{x} + w$, où $w \in \mathbb{C}^Q$ est une réalisation d'un vecteur aléatoire. On suppose que ce signal admet une représentation parcimonieuse sur une famille $\mathcal{E} = \{e_\nu \mid \nu \in \mathbb{R}\}$ de vecteurs de \mathbb{C}^Q paramétrés par une variable scalaire $\nu \in \mathbb{R}$. Plus précisément, le signal $\bar{x} \in \mathbb{C}^Q$ peut s'écrire sous la forme

$$\bar{x} = \sum_{n=1}^N \bar{c}_n e_{\bar{\nu}_n} \quad (1)$$

où $\{e_{\bar{\nu}_n} \mid 1 \leq n \leq N\}$ est un sous-ensemble fini de \mathcal{E} , un grand nombre de composantes du vecteur $\bar{c} = (\bar{c}_n)_{1 \leq n \leq N} \in \mathbb{C}^N$ étant supposées nulles.

L'originalité de ce travail est de se placer dans le cas où les paramètres $(\bar{\nu}_n)_{1 \leq n \leq N}$ sont connus de manière imprécise, *i.e.* ils sont tels que, pour tout $n \in \{1, \dots, N\}$, $\bar{\nu}_n = \theta_n + \bar{\delta}_n$ où $\theta_n \in \mathbb{R}$ est une valeur fixée et $\bar{\delta}_n \in \mathbb{R}$ est une erreur inconnue qui doit être estimée. Sous les hypothèses que les perturbations $(\bar{\delta}_n)_{1 \leq n \leq N}$ sont petites et que la fonction $\nu \mapsto e_\nu$ est différentiable, un développement de Taylor au premier ordre conduit à

$$(\forall n \in \{1, \dots, N\}) \quad e_{\bar{\nu}_n} \simeq e_{\theta_n} + \bar{\delta}_n e'_{\theta_n} \quad (2)$$

où e'_{θ_n} est le gradient de $\nu \mapsto e_\nu$ en θ_n . Après cette approximation, le modèle (1) prend une forme bilinéaire :

$$\bar{x} = \sum_{n=1}^N (\bar{c}_n e_{\theta_n} + \bar{c}_n \bar{\delta}_n e'_{\theta_n}). \quad (3)$$

Notons que l'étude d'un modèle parcimonieux perturbé a été récemment réalisée dans [1] en optimisant un critère de type moindres carrés totaux (*total least squares*) pénalisé par un terme ℓ_1 . Notre approche diffère par les points

suivants : (i) elle est basée sur l'emploi d'une pénalisation ℓ_0 , (ii) les erreurs $(\bar{\delta}_n)_{1 \leq n \leq N}$ sont supposées telles que

$$(\forall n \in \{1, \dots, N\}) \quad |\bar{\delta}_n| \leq \Delta_n, \quad (4)$$

les bornes $(\Delta_n)_{1 \leq n \leq N} \in [0, +\infty[^N$ étant fixées au préalable, (iii) un terme de fidélité quelconque peut être employé pourvu qu'il soit de gradient Lipschitzien.

2 Méthode proposée

Formulation variationnelle. Afin d'identifier le modèle parcimonieux perturbé, nous proposons de résoudre le problème d'optimisation suivant :

$$\underset{\substack{c=(c_n)_{1 \leq n \leq N} \in \mathbb{C}^N \\ \delta=(\delta_n)_{1 \leq n \leq N} \in B}}{\text{minimiser}} \quad \Phi \left(\sum_{n=1}^N (c_n e_{\theta_n} + c_n \delta_n e'_{\theta_n}) - y \right) + \lambda \ell_0(c) + \frac{\epsilon}{2} \|c\|^2 \quad (5)$$

où B est le pavé de \mathbb{R}^N défini par $[-\Delta_1, \Delta_1] \times \dots \times [-\Delta_N, \Delta_N]$, $\Phi: \mathbb{C}^Q \rightarrow \mathbb{R}$ est le terme de fidélité aux données souvent lié à l'anti-log-vraisemblance du bruit entâchant les observations, $\lambda \in]0, +\infty[$ est un paramètre de régularisation servant à imposer la "sparsité" de la représentation et $\epsilon \in [0, +\infty[$. Le dernier terme joue un rôle similaire à celui d'une pénalisation élastique [2].

Dans la suite, on introduit les matrices $E = [e_{\theta_1} \dots e_{\theta_N}] \in \mathbb{C}^{Q \times N}$, $E' = [e'_{\theta_1} \dots e'_{\theta_N}] \in \mathbb{C}^{Q \times N}$ et l'on définit, pour tout $n \in \{1, \dots, N\}$, la fonction

$$(\forall (c_n, d_n) \in \mathbb{C}^2) \quad \psi_n(c_n, d_n) = \lambda \chi_{\{0\}}(c_n) + \iota_{S_n}(c_n, d_n) + \frac{\epsilon}{2} |c_n|^2, \quad (6)$$

où ι_{S_n} est la fonction indicatrice du cône fermé $S_n = \{(c_n, d_n) \in \mathbb{C}^2 \mid \exists \delta_n \in [-\Delta_n, \Delta_n], d_n = \delta_n c_n\}$. On peut alors reformuler le problème (5) comme celui de la minimisation de la fonction

$$(c, d) = ((c_n)_{1 \leq n \leq N}, (d_n)_{1 \leq n \leq N}) \mapsto \Phi \left([E \ E'] \begin{bmatrix} c \\ d \end{bmatrix} - y \right) + \sum_{n=1}^N \psi_n(c_n, d_n). \quad (7)$$

Algorithme. Si l'on fait l'hypothèse que la fonction Φ est différentiable, on peut apporter une solution numérique au problème (5) en employant l'algorithme explicite-implicite suivant :

$$\begin{aligned} & (c^{(0)}, d^{(0)}) \in (\mathbb{C}^N)^2 \\ & (\underline{\gamma}, \bar{\gamma}) \in]0, +\infty[^2 \\ & \text{Pour } k = 0, 1, \dots \\ & \left[\begin{array}{l} \gamma^{(k)} \in (\underline{\gamma}, \bar{\gamma}) \\ D^{(k)} = \nabla \Phi \left([E \ E'] \begin{bmatrix} c^{(k)} \\ d^{(k)} \end{bmatrix} - y \right) \\ (\tilde{c}_n^{(k)})_{1 \leq n \leq N} = c^{(k)} - \gamma^{(k)} E^H D^{(k)} \\ (\tilde{d}_n^{(k)})_{1 \leq n \leq N} = d^{(k)} - \gamma^{(k)} (E')^H D^{(k)} \\ (c_n^{(k+1)}, d_n^{(k+1)})_{1 \leq n \leq N} = (\text{prox}_{\gamma^{(k)} \psi_n}(\tilde{c}_n^{(k)}, \tilde{d}_n^{(k)}))_{1 \leq n \leq N}. \end{array} \right. \quad (8) \end{aligned}$$

Remarquons que, pour mettre en œuvre cet algorithme, il est nécessaire de calculer les opérateurs proximaux [3] des fonctions $(\psi_n)_{1 \leq n \leq N}$. Ces fonctions sont non convexes, mais on peut montrer que leurs opérateurs proximaux prennent une forme relativement simple. Il est intéressant de souligner que, dans le cas particulier où Φ est un critère de moindres carrés, $\Delta_n \equiv 0$ et $\epsilon = 0$, l'algorithme se réduit à un seuillage dur itératif.

Théorème de convergence. Le problème d'optimisation considéré étant non convexe, la convergence de l'algorithme explicite-implicite n'est pas évidente. En utilisant des résultats récents d'analyse non lisse pour les fonctions à valeurs réelles [4], nous avons cependant été en mesure de montrer la propriété suivante :

Proposition 2.1. *Si Φ est une fonction semi-algébrique de gradient L -Lipschitzien avec $L \in]0, +\infty[$ et si les bornes $\underline{\gamma}$ et $\bar{\gamma}$ sur le pas de l'algorithme sont choisies telles que*

$$0 < \underline{\gamma} \leq \bar{\gamma} < L^{-1}(\|E\|^2 + \|E'\|^2)^{-1}, \quad (9)$$

alors toute suite $(c^{(k)}, d^{(k)})_{k \in \mathbb{N}}$ générée par l'algorithme (8) converge vers un point critique de la fonction (7).

Par ailleurs, nous avons pu établir que, si Φ est convexe, tout point critique de la fonction (7) est, sous une condition peu restrictive, minimiseur local de cette fonction.

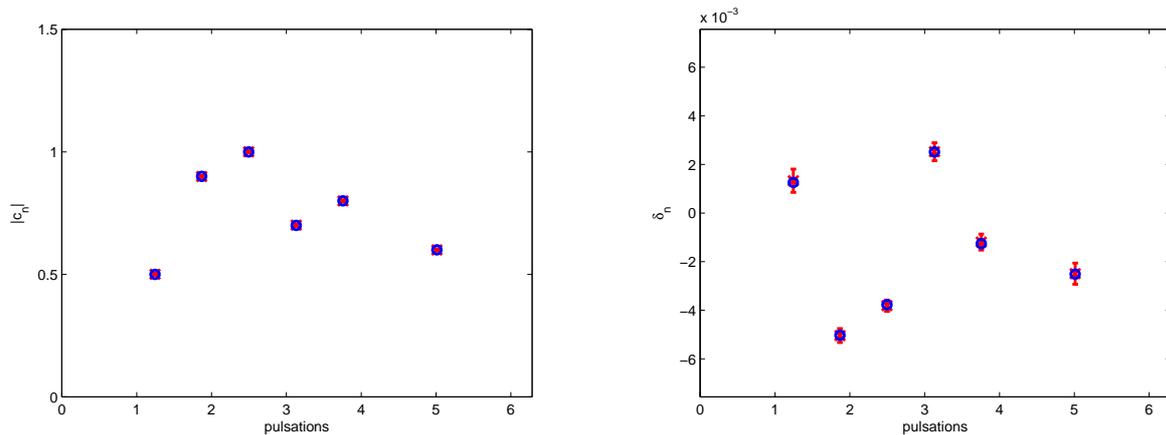


FIGURE 1 – Estimation de 6 cisoïdes (RSB = 23.27 dB) : à gauche, amplitudes $(|c_n|)_{1 \leq n \leq N}$; à droite, perturbations $(\delta_n)_{1 \leq n \leq N}$ en fonction des pulsations $(\theta_n)_{1 \leq n \leq N}$ sur la grille. Les valeurs exactes sont représentées par des cercles bleus et les estimées avec leur intervalle de confiance en rouge (la moyenne est indiquée par une croix).

3 Application à l’analyse spectrale

Afin d’illustrer le bon comportement de la méthode proposée, nous considérons un problème d’estimation spectrale parcimonieuse [5]. Nous proposons ici une alternative aux méthodes de raffinement de grilles de recherche [6, 7]. Plus précisément, nous disposons de $Q = 50$ observations d’un signal à valeurs complexes correspondant à la somme de 6 cisoïdes bruitées, échantillonnées irrégulièrement à des instants $(\tau_q)_{1 \leq q \leq Q}$ choisis de manière aléatoire sur $[0, Q]$. Les observations discrètes ont été dégradées par l’addition d’un bruit blanc gaussien complexe circulaire, centré et de variance σ^2 . Le dictionnaire employé comporte $N = 500$ cisoïdes dont les pulsations sont régulièrement échantillonnées sur $[0, 2\pi]$. On a ainsi : $(\forall \nu \in \mathbb{R}) e_\nu = (\exp(i\nu\tau_q))_{1 \leq q \leq Q}$ et $(\forall n \in \{1, \dots, N\}) (\theta_n, \Delta_n) = (2\pi(n-1)/N, \pi/N)$.

Dans les cas traités, les pulsations $(\bar{\nu}_n)_{1 \leq n \leq N}$ des composantes parcimonieuses n’appartiennent pas à la grille de recherche. La figure 1, résultat d’une étude de Monte Carlo menée sur 100 réalisations, présente un exemple typique des estimations obtenues. Notre approche conduit à des gains substantiels par rapport à l’utilisation d’une stratégie basée sur la simple minimisation d’un coût ℓ_1 ainsi que sur l’emploi direct d’un algorithme de seuillage dur itératif. Notons que les résultats de la minimisation ℓ_1 servent d’initialisation à notre algorithme. Par ailleurs, le paramètre de régularisation λ est choisi de manière automatique.

Références

- [1] H. Zhu, G. Leus et G. B. Giannakis, « Sparsity-cognizant total least-squares for perturbed compressive sampling », *IEEE Trans. Signal Process.*, vol. 59, n°5, pp. 2002–2016, Mai 2011.
- [2] H. Zou et T. Hastie, « Regularization and variable selection via the elastic net », *J. R. Statist. Soc. A*, vol. 67, n°2, pp. 301–320, 2005.
- [3] P. L. Combettes et J.-C. Pesquet, « Proximal splitting methods in signal processing », in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke et H. Wolkowicz, Eds., pp. 185–212. Springer-Verlag, New York, 2011.
- [4] H. Attouch, J. Bolte et B. F. Svaiter, « Convergence of descent methods for semi-algebraic and tame problems : proximal algorithms, forward backward splitting, and regularized Gauss-Seidel methods », *Math. Prog.*, vol. 137, n°1-2, pp. 1–39, Fév. 2013.
- [5] S. Bourguignon, H. Carfantan et J. Idier, « A sparsity-based method for the estimation of spectral lines from irregularly sampled data », *IEEE J. Sel. Topics Signal Process.*, vol. 1, n°4, pp. 575–585, Déc. 2007.
- [6] S. Sahnoun, E. H. Djermoune, C. Soussen et D. Brie, « Sparse multidimensional modal analysis using a multigrad dictionary refinement », *EURASIP J. Adv. Sig. Proc.*, vol. 60, Mar. 2012.
- [7] P. Stoica et P. Babu, « Sparse estimation of spectral lines : Grid selection problems and their solutions », *IEEE Trans. Signal Process.*, vol. 60, n°2, pp. 962–967, Fév. 2012.