Branch-and-Bound Algorithm for Sparse Approximation in Signal Processing and Statistics

Ramzi Ben Mhenni, Sel̀ bastien Bourguignon and Jordan Ninin

February 5, 2020

Algorithme *Branch-and-Bound* pour l'approximation parcimonieuse en traitement du signal et en statistiques

Ramzi Ben Mhenni[1], Sébastien Bourguignon[1], Jordan Ninin[2]

[1] École Centrale de Nantes, LS2N , 1 rue de la Noë, F-44321 Nantes, France
{ramzi.ben-mhenni,sebastien.bourguignon}@ec-nantes.fr
[2] ENSTA-Bretagne, Lab-STICC , 2 rue François Verny, F-29806 Brest cedex 9, France
jordan.ninin@ensta-bretagne.fr

Mots-clés : *Optimisation en norme ℓ_0, Cardinalité, Programmation en Nombres Mixtes, Algorithme Branch-and-Bound.*

1 Introduction

L'approximation parcimonieuse de données $\boldsymbol{y} \in \mathbb{R}^m$ consiste à estimer, sous un modèle $\boldsymbol{y} \simeq \boldsymbol{A}\boldsymbol{x}$, un vecteur $\boldsymbol{x} \in \mathbb{R}^n$ avec peu de composantes non nulles, *i.e.*, de faible "norme" ℓ_0 : $\|\boldsymbol{x}\|_0 := \text{Card}\,\{i|\; x_i \neq 0\}$. En introduisant des variables de décision binaires ($b_i = 1 \Leftrightarrow x_i \neq 0$), la norme ℓ_0 peut être reformulée exactement et le problème s'écrit alors, selon la formulation choisie, sous la forme d'un MIP (*Mixed-integer Program*) [1, 2, 3] :

$$\begin{aligned}
\mathcal{P}_{2/0} : & \min_{\boldsymbol{b}\in\{0,1\}^n,\ \boldsymbol{x}\in\mathbb{R}^n} \tfrac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2 \ \text{ s.c. } \sum_{i=1}^n b_i \leq K \ \text{ et } \ |\boldsymbol{x}| \leq M\boldsymbol{b};\\
\mathcal{P}_{0/2} : & \min_{\boldsymbol{b}\in\{0,1\}^n,\ \boldsymbol{x}\in\mathbb{R}^n} \sum_{i=1}^n b_i \ \text{ s.c. } \tfrac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2 \leq \epsilon \ \text{ et } \ |\boldsymbol{x}| \leq M\boldsymbol{b};\\
\mathcal{P}_{2+0} : & \min_{\boldsymbol{b}\in\{0,1\}^n,\ \boldsymbol{x}\in\mathbb{R}^n} \tfrac{1}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu\sum_{i=1}^n b_i \ \text{ s.c. } \ |\boldsymbol{x}| \leq M\boldsymbol{b};
\end{aligned}$$

où M est une constante suffisamment grande (*bigM*). Il a ainsi été montré dans [3] que la résolution exacte de ces problèmes était envisageable par un solveur MIP générique (CPLEX) pour des problèmes de petite taille mais difficiles, où les méthodes d'approximation sous-optimales standard échouent à trouver l'optimum global. Ces MIPs sont cependant très spécifiques et nous proposons donc ici un algorithme Branch-and-Bound dédié [1, 2]. Notre contribution majeure est basée sur le fait que les relaxations continues à chaque nœud peuvent être réduites à des problèmes d'optimisation sans variables binaires faisant intervenir la norme ℓ_1, pour lesquels nous construisons un algorithme particulièrement efficace.

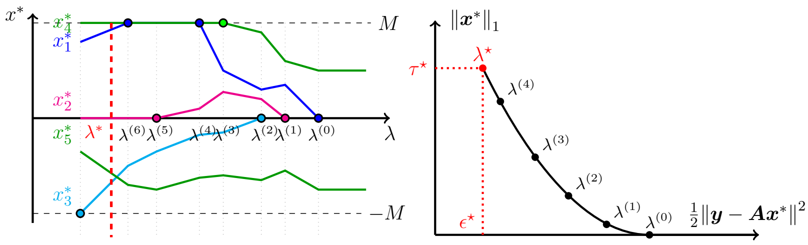
2 Algorithme Branch-and-Bound spécifique

Notre algorithme fait une recherche en profondeur d'abord. Puisque les variables de décision sont binaires, la séparation correspond à une décision : $b_i = 1$ ou $b_i = 0$. Nous privilégions l'exploration du côté $b_i = 1$ d'abord pour limiter la profondeur de l'arbre de recherche, contrainte par la parcimonie (pour $\mathcal{P}_{2/0}$, le nombre de branches $b_i = 1$ est limité à K). Chaque nœud est évalué *via* la relaxation continue des variables binaires et la variable de branchement correspond à la coordonnée d'amplitude maximale dans la solution relâchée.

Avec des notations évidentes, on a, à un nœud donné, $\boldsymbol{b}_{\mathrm{S}^1} = 1$, $\boldsymbol{b}_{\mathrm{S}^0} = 0$ et les variables $\boldsymbol{b}_{\bar{\mathrm{S}}}$ sont indéterminées. Nous avons montré que la relaxation continue ($b_i \in [0,1], i \in \bar{\mathrm{S}}$) peut être calculée en résolvant un problème équivalent sans variables binaires, impliquant un terme en norme ℓ_1. Par exemple, pour $\mathcal{P}_{2+0}$, celui-ci s'écrit :

$$\mathcal{R}_{2+1} : \min_{\boldsymbol{x}_{\mathrm{S}^1\cup\bar{\mathrm{S}}}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{A}_{\mathrm{S}^1}\boldsymbol{x}_{\mathrm{S}^1} - \boldsymbol{A}_{\bar{\mathrm{S}}}\boldsymbol{x}_{\bar{\mathrm{S}}}\|_2^2 + \lambda\|\boldsymbol{x}_{\bar{\mathrm{S}}}\|_1 \quad \text{s.c.} \quad \|\boldsymbol{x}_{\mathrm{S}^1\cup\bar{\mathrm{S}}}\|_\infty \leq M, \quad \text{avec } \lambda := \mu/M.$$

Des résultats similaires sont obtenus pour $\mathcal{P}_{2/0}$ et $\mathcal{P}_{0/2}$. Nous avons alors construit un algorithme homotopique [4] pour ces problèmes. Ce dernier considère la forme pénalisée $\mathcal{R}_{2+1}$ et exploite le fait que le chemin de solution est linéaire par morceaux en fonction de λ (Figure 1 à gauche). L'algorithme calcule itérativement toutes les solutions en diminuant de façon



Chemin de solutions pour $\mathcal{R}_{2+1}$ — Front de Pareto correspondant

FIG. 1 – Méthode homotopique : chemin de solution $\boldsymbol{x}^* = \arg\min \mathcal{R}_{2+1}$ en fonction de λ.

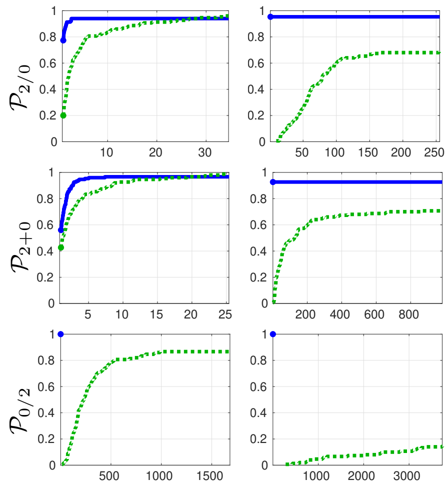
continue le paramètre λ jusqu'à atteindre le valeur cible. Par conséquent, il peut également résoudre les problèmes $\mathcal{R}_{2/1}$ ou $\mathcal{R}_{1/2}$, en s'arrêtant lorsque la contrainte correspondante est atteinte (Figure 1 à droite).

3 Résultats expérimentaux

Nous comparons notre algorithme, noté B&B$_{\text{Hom}}$, au solveur MIP de CPLEX (v12.8), MIP$_{\text{Cplex}}$, sur :

- des problèmes de déconvolution parcimonieuse [3] avec $n = 100$ inconnues et un degré de parcimonie de $K = 5$ à $K = 9$. Malgré leur petite taille, ce sont des problèmes inverses mal posés et très compliqués ;
- des problèmes de sélection de variables avec $\boldsymbol{A}$ et $\boldsymbol{x}$ aléatoires, $n = 1000$ inconnues et un degré de parcimonie de $K = 5$ à $K = 15$.

Les profils de performance représentés en figure 2 montrent que notre algorithme est bien meilleur que MIP$_{\text{Cplex}}$. Sur les problèmes de déconvolution, même si MIP$_{\text{Cplex}}$ est capable de résoudre plus de problèmes de type $\mathcal{P}_{2/0}$ et $\mathcal{P}_{2+0}$, B&B$_{\text{Hom}}$ reste plus rapide dans 80% des cas pour $\mathcal{P}_{2/0}$ et 58% pour $\mathcal{P}_{2+0}$. Surtout, pour $\mathcal{P}_{0/2}$, notre algorithme reste de performances comparables aux autres cas alors que celles de MIP$_{\text{Cplex}}$ s'écroulent. Sur les problèmes de sélection de variables, B&B$_{\text{Hom}}$ surpasse de loin MIP$_{\text{Cplex}}$, reflétant la difficulté du solveur générique à monter en dimension.



déconv. — sél. variables

FIG. 2 – Profils de performance obtenus sur 150 problèmes, pour B&B$_{\text{Hom}}$ $(-)$, MIP$_{\text{Cplex}}$ $(\cdots)$.

Références

[1] D. Bertsimas and R. Shioda. Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1) :1–22, 2009.

[2] D. Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74(2) :121–140, 1996.

[3] S. Bourguignon, J. Ninin, H. Carfantan, and M. Mongeau. Exact sparse approximation problems via mixed-integer programming : Formulations and computational performance. *IEEE Transactions on Signal Processing*, 64(6) :1405–1419, March 2016.

[4] M. R. Osborne, B. Presnell B, and B. A. D. Turlach BAD. A new approach to variable selection in least squares problems. *IMA Journal of Numerical Analysis*, 2000.