MIP and Set Covering approaches for Sparse Approximation

Diego Delle Donne[1], Matthieu Kowalski[2], Leo Liberti[1]

[1] LIX CNRS, Ecole Polytechnique, Institut Polytechnique de Paris, 91128 Palaiseau, France.
{delledonne,liberti}@lix.polytechnique.fr

[2] Laboratoire des Signaux et Systèmes, UMR 8506 Univ Paris-Sud – CNRS – Centralesupelec, 91192 Gif-sur-Yvette Cedex, France
matthieu.kowalski@u-psud.fr

Keywords : *sparse approximation, set covering, mixed integer programming.*

1 Introduction

The *sparse representation* of a vector $y \in \mathbb{R}^n$ in a *dictionary* $H \in \mathbb{R}^{n \cdot m}$ aims to find a solution $x \in \mathbb{R}^m$ to the system $Hx = y$, having the minimum number of non-zero components, i.e., minimizing the so-called ℓ_0 pseudo norm of x, defined by $||x||_0 := \#\{j \mid x_j \neq 0\}$. The *sparse approximation* problem takes into account *noise* and *model errors*, and it relaxes the equality constraint aiming to minimize the misfit data measure $||y - Hx||$, for a given norm $|| \cdot ||$. In this context, several optimization problems my be stated such as:

1. minimize $||x||_0$ subject to a given *threshold* $||y - Hx|| \leq \alpha$,
2. minimize the data misfit $||y - Hx||$ subject to a given bound $||x||_0 \leq k$,
3. minimize a weighted sum $\lambda_1||y - Hx|| + \lambda_2||x||_0$ for some $\lambda_1, \lambda_2 \in \mathbb{R}$.

In this work, we focus on the problem stated in Item 1 when the norm used for the data misfit measure is the Euclidean p-norm, for $p \in \{1, \infty\}$. Following the notation from [1], we define these problems as

$$\mathcal{P}_{0/p} : \min_x ||x||_0 \text{ s. t. } ||y - Hx||_p \leq \alpha_p.$$

Some natural *mixed-integer programming* (MIP) formulations for $\mathcal{P}_{0/1}$ and $\mathcal{P}_{0/\infty}$ are introduced in [1]. These models use decision variables $x_j \in \mathbb{R}$, for each $j \in [m]$ to determine the solution and binary *support* variables b_j to state whether x_j has a non-zero value or not. They require to (artificially) bound $|x_j|$ with a value M in order to properly state the models. Then, $\sum_{j\in[m]} b_j$ is minimized subject to appropriate constraints. We call these formulations $MIP_{0/1}$ and $MIP_{0/\infty}$, respectively, and we omit to state them here due to space limitations. In [1], these formulations are solved directly by CPLEX. As far as we know, no polyhedral studies have been done for these formulations, with the goal of developing more powerful resolution algorithms (e.g., cutting planes based ones).

In this context, we propose some families of valid inequalities for $MIP_{0/1}$ and $MIP_{0/\infty}$ and we test these families in a *branch & cut* scheme, in order to assess their practical contribution to solve the problems. We also introduce a novel IP approach for $\mathcal{P}_{0/p}$ which reduces the latter to a *minimum set covering* problem (with exponentially many covering constraints). We propose two resolution approaches: a 2-stages algorithm to tackle the IP resolution and a combinatorial algorithm to solve the associated covering problem. In the remaining, for any natural number t, we may use $[t]$ as a shortcut for the set $\{1, \dots, t\}$.

2 New valid inequalities and a novel IP approach for $\mathcal{P}_{0/p}$

We say that a set of columns $J \subseteq [m]$ is a *forbidden support* for $\mathcal{P}_{0/p}$ if there exist no solutions with J as support, i.e., if $\min_x\{||y - H^J x^J||_p\} > \alpha_p$, where H^J (resp. x^J) is the submatrix of H (resp. subvector of x) involving only those columns indexed by J.

Proposition 2.1. *If $J \subseteq [m]$ is a forbidden support for $\mathcal{P}_{0/p}$, then the* forbidden support inequality

$$\sum_{j \in [m] \setminus J} b_j \geq 1 \tag{1}$$

is valid for $MIP_{0/p}$.

From Proposition 2.1, we derive some subfamilies of valid inequalities for which we developed separation procedures (both exact and heuristics) and implemented a *branch & cut* algorithm using them as cutting planes. We omit here all these elements due to space limitations. We state next an interesting theoretical result about the *forbidden support inequalities* (1).

Proposition 2.2. *The projection on the variables b_j of all feasible solutions of formulation $MIP_{0/p}$ can be described by the forbidden support inequalities* (1) *as*

$$\mathcal{P}_{fs} = \{b \in \{0,1\}^m \mid b \textit{ satisfies } (1) \textit{ for each forbidden support } J \subseteq [m]\}.$$

Proposition 2.2 lets us obtain a minimum support $\hat{b}$ of a solution to $\mathcal{P}_{0/p}$ by solving the following *integer programming* (IP) formulation:

$$[IP_{0/p}^{cov}] \qquad \hat{b} = \arg\min_{b \in \{0,1\}^m} \{ \sum_{j \in [m]} b_j \mid \sum_{j \in [m] \setminus J} b_j \geq 1, \quad \forall \text{ forbidden support } J \subseteq [m]\}$$

We should remark that $\hat{b}$ is not a solution for $\mathcal{P}_{0/p}$ but just an optimal support. However, a solution for this support can be easily obtained afterwards by solving a linear program. Moreover, as the support is already fixed for this last step, there is no need to use the artificial big-M bounds for x, fact that gives an important advantage against formulation $MIP_{0/p}$. Another interesting characteristic about $IP_{0/p}^{cov}$ is that it represents a *minimum set covering* problem and this kind of problems have been widely studied in the literature.

A strong drawback of $IP_{0/p}^{cov}$ is that the formulation may have exponentially many constraints. Furthermore, a linear program may be eventually need to be solved (in general) for each subset $J \subseteq [m]$ to check if J is a forbidden support or not. To address these obstacles we propose and implement two approaches: a 2-stages algorithm to solve $IP_{0/p}^{cov}$ and a combinatorial algorithm tailored for our covering problem (in which the "elements to cover" are exponentially many).

3 Final remarks

In this work we present some families of valid inequalities for known formulations for Sparse Approximation problems and we test their potential in a branch & cut scheme. We also prove that one of these families is sufficient to describe the support of all feasible solutions and from this fact we re-state the problem as a *set covering* problem, thus opening many possible lines of work to solve Sparse Approximation problems; we inspect some of them proposing two exact algorithms. As an ongoing work, we are trying to complement our branch & cut algorithm for $MIP_{0/p}$ by using some of the many valid inequalities known for set covering polytopes.

References

[1] 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.