An alternating $l_1$ approach to the compressed sensing problem

Stéphane Chrétien$^{1,*}$

$^1$ Department of Mathematics, UMR CNRS 6623 and University of Franche Comté, 16 route de Gray, 25030 Besançon Cedex, France.

The purpose of this note is to present a simple algorithm motivated by Lagrangian duality and generalizing the $l_1$-relaxation approach to the recent Compressed Sensing (CS) problem introduced by Candes, Romberg and Tao.

1 Introduction

From the algorithmic perspective, the Compressed Sensing problem can be stated as follows. Let $x$ be a $k$-sparse vector in $\mathbb{R}^n$, i.e. a vector with no more than $k$ nonzero components. The observations are simply given by

$$y = Ax$$

where $A \in \mathbb{R}^{m \times n}$ and $m$ small compared to $n$, and the goal is to recover $x$ exactly from these observations. Following [1], recovery can be obtained by simply finding the sparsest solution to (1). If for any $x$ in $\mathbb{R}^n$ we denote by $\|x\|_0$ the $l_0$-norm of $x$, i.e. the cardinal of the set of indices of nonzero components of $x$, the compressed sensing problem is equivalent to

$$\min_{x \in \mathbb{R}^n} \|x\|_0 \quad \text{s.t.} \quad Ax = y. \quad (2)$$

Using the notations of [2], we denote by $\Delta_0(y)$ the solution of problem (2) and $\Delta_0(y)$ is called a decoder. Thus, the CS problem is a combinatorial optimization problem. The main problem in this approach is that obtaining $\Delta_0(y)$ is generally computationally intractable. In a recent series of wonderful papers (see precise references in [1]) Candes, Romberg and Tao proposed a convex $l_1$-relaxation instead and proved that this relaxation works with high probability for certain ensembles of random observation matrices $A$ using the new notion of Restricted Isometry Property (RIP).

One among many challenging problems motivated by these important discoveries is to find new methods to allow exact recovery for less sparse signals at a reasonable computational cost compared to the current $l_1$ relaxation. The goal of our paper is to present such a new method for solving the CS problem. Our approach is based on ideas drawn from Lagrangian duality theory.

2 The Alternating $l_1$ method

One very exciting questions motivated by these works is how much can one increase the value of $k$ for which every $k$-signal can be reconstructed exactly for a given pair $(n, m)$ using a polynomial-time algorithm? The purpose of this section is to describe a algorithm based on Lagrangian duality with improved practical efficiency compared to the current $l_1$ relaxation and even to the recent interesting Reweighted $l_1$ of Boyd, Candes and Wakin.

2.1 The Lagrangian dual

Let us start by writing down problem (2), to which $\Delta_0(y)$ is the solution map, as the following equivalent problem

$$\sup_{z \in \{0, 1\}^n, x \in \mathbb{R}^n} e^t z \quad \text{s.t.} \quad z_i x_i = 0, \quad i = 1, \ldots, n, \quad Ax = y \quad (3)$$

where $e$ denotes the vector of all ones. Here since the sum of the $z_i$’s is maximized, the variable $z$ plays the role of an indicator function for the event that $x_i = 0$. This problem is clearly nonconvex due to the quadratic equality constraints $z_i x_i = 0, \quad i = 1, \ldots, n$. However, these constraints can be merged into the unique constraint $\|D(z)x\|_1 = 0$, leading to the following equivalent problem

$$\sup_{z \in \{0, 1\}^n, x \in \mathbb{R}^n} e^t z \quad \text{s.t.} \quad \|D(z)x\|_1 = 0, \quad Ax = y. \quad (4)$$

* Corresponding author E-mail: stephane.chretien@univ-fcomte.fr, Phone: +33 3 81 66 64 65, Fax: +33 3 81 66 66 23
Turning back to problem (4) and choosing to keep the constraints \(Ax = y\) and \(z \in \{0, 1\}^n\) implicit, the Lagrangian function is given by

\[
L(x, z, u) = e^Tz + u\|D(z)x\|_1
\]

where \(D(z)\) is the diagonal matrix with diagonal vector equal to \(z\). The dual function (with values in \(\mathbb{R} \cup \{+\infty\}\)) is defined by

\[
\theta(u) = \sup_{z \in \{0, 1\}^n, x \in \mathbb{R}^n, Ax = y} L(x, z, u)
\]

and the dual problem is

\[
\inf_{u \in \mathbb{R}} \theta(u).
\]

Notice that if \(u > 0\) then \(\theta(u) = +\infty\).

### 2.2 A practical alternative

The main problem with the dual problem (6) is that \(\theta\) maybe as difficult to compute for each \(u\) as the solution of the original problem (4) due to the nonconvexity of the Lagrangian function \(L\). However, we may try and solve (6) in a suboptimal way using an alternating minimization approach which we stop after an fixed arbitrary number \(L\) of steps so as to control the complexity of the method.

**Algorithm 1 Alternating \(l_1\) algorithm (Alt-\(l_1\))**

**Require:** \(u < 0\) and \(L \in \mathbb{N}_*\)

\[
\begin{align*}
    z_u^{(0)} &= e \\
x_u^{(0)} &= \max_{x \in \mathbb{R}^n, Ax = y} L(x, z_u^{(0)}, u) \\
l &= 1
\end{align*}
\]

while \(l \leq L\) do

\[
\begin{align*}
    z_u^{(l)} &= \text{argmax}_{z \in \{0, 1\}^n, Ax = y} L(x_u^{(l)}, z, u) \\
x_u^{(l)} &= \text{argmax}_{x \in \mathbb{R}^n, Ax = y} L(x, z_u^{(l)}, u) \\
l &\leftarrow l + 1
\end{align*}
\]

end while

Output \(z_u^{(L)}\) and \(x_u^{(L)}\).

### 2.3 Monte Carlo experiments

Comparison between the success rate of \(l_1\) and Alternating \(l_1\) is shown in Figure 1. Optimization of the Lagrange multiplier \(u\) was performed using coarse dichotomic search and we finally used \(u = 3\) and \(L = 4\) iterations in the Alternating \(l_1\). We also incorporated the results obtained using Boyd, Candès and Wakin’s recent proposal [3] called the Reweighted \(l_1\) relaxation. Our proposal outperformed both the plain \(l_1\) and the Reweighted \(l_1\) relaxations for the given data sizes.

![Fig. 1 Rate of success over 1000 Monte Carlo experiments in recovering the support of the signal vs. signal sparsity \(k\) for \(n = 128\), \(m = 50\), \(L = 4\), \(u = -3\). A and nonnull components of \(x\) were drawn from the Gaussian \(N(0,1)\) distribution. The black line is for the \(l_1\) relaxation, the blue line for Boyd, Candès and Wakin’s new Reweighted \(l_1\) relaxation with \(\epsilon = .1\), the best parameter value which was found in [3], and the green line is for our Alternating \(l_1\) relaxation.](image)

**References**

[1] E. Candès, Compressive Sampling, International Congress of Mathematics, Madrid, Spain 3, 1433–1452 (2006).
[2] A. Cohen, W. Dahmen and R. DeVore, Compressed Sensing and Best \(k\)-term Approximation, preprint (2007).
[3] E. Candés, M. Wakin and S. Boyd, Enhancing Sparsity by Reweighted \(l_1\) Minimization, preprint (2007).