Solver for Modified-CS

João F. C. Mota

October 29, 2014

We solve

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \|x_{T^c}\|_1 \\ \text{subject to} & Ax = b\,, \end{array} \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $T \subseteq \{1,\ldots,n\}$ are given. The set T^c is the complement of T in $\{1,\ldots,n\}$, and x_{T^c} denotes the subvector of $x \in \mathbb{R}^n$ containing the components of x that are indexed by T^c. Therefore, $\|x_{T^c}\|_1 = \sum_{i\notin T}|x_i|$. Problem (1) was proposed in [1] and is known as *Modified-CS*.

This document describes a solver for (1) that uses the Alternating Direction Method of Multipliers (ADMM) [2, 3, 4]. First, we recast the problem as

$$\begin{array}{ll} \underset{x,y}{\text{minimize}} & f(x) + g(y) \\ \text{subject to} & x = y\,, \end{array} \tag{2}$$

where $f(x) = \|x_{T^c}\|_1$ and $g(y) = \mathrm{i}_{\{x\,:\,Ax=b\}}(y)$, where $\mathrm{i}_S(x)$ is the indicator function of the set S, i.e.,

$$\mathrm{i}_S(x) = \begin{cases} 0 & ,\ \text{if } x \in S \\ +\infty & ,\ \text{if } x \notin S. \end{cases}$$

The augmented Lagrangian of (2) is

$$L_\rho(x,y;\lambda) = f(x) + g(y) + \lambda^\top(x-y) + \frac{\rho}{2}\|x-y\|^2\,,$$

and ADMM becomes

$$x^{k+1} = \arg\min_x\ f(x) + {\lambda^k}^\top x + \frac{\rho}{2}\|x - y^k\|^2 \tag{3}$$

$$y^{k+1} = \arg\min_y\ g(y) - {\lambda^k}^\top y + \frac{\rho}{2}\|x^{k+1} - y\|^2 \tag{4}$$

$$\lambda^{k+1} = \lambda^k + \rho(x^{k+1} - y^{k+1})\,.$$

It turns out that both (3) and (4) have closed-form solutions.

Problem in x. Developing the square, problem (3) is equivalent to

$$x^{k+1} = \arg\min_x\ f(x) + (\lambda^k - \rho y^k)^\top x + \frac{\rho}{2}\|x\|^2\,. \tag{5}$$

Let $v := \lambda^k - \rho y^k$. Replacing the expression for function f in (5),

$$x^{k+1} = \arg\min_x\ \sum_{i\notin T}|x_i| + v^\top x + \frac{\rho}{2}\|x\|^2\,, \tag{6}$$

whose solution can be computed in closed-form. First note that (6) decomposes into n independent problems. Then if $i \in T$, we have

$$x_i^{k+1} = \arg\min_x \; v_i x_i + \frac{\rho}{2} x_i^2 \; = -\frac{1}{\rho} v_i \, .$$

If $i \notin T$, then

$$x_i^{k+1} = \arg\min_x \; |x_i| + v_i x_i + \frac{\rho}{2} x_i^2 \; = \begin{cases} -\frac{1}{\rho}(v_i + 1) & , \; v_i < -1 \\ 0 & , \; -1 \leq v_i \leq 1 \\ -\frac{1}{\rho}(v_i - 1) & , \; v_i > 1 \, . \end{cases}$$

Problem in $\boldsymbol{y}$. Problem (4) is equivalent to

$$\begin{array}{ll} \arg\min_y & \frac{\rho}{2}\|y\|^2 - \rho\, {x^{k+1}}^\top y - {\lambda^k}^\top y \\ \text{s.t.} & Ay = b \end{array} \quad \Longleftrightarrow \quad \begin{array}{ll} \arg\min_y & \|y\|^2 - \frac{2}{\rho}(\lambda^k + \rho\, x^{k+1})^\top y \\ \text{s.t.} & Ay = b \end{array}$$

$$\Longleftrightarrow \quad \begin{array}{ll} \arg\min_y & \|y - \frac{1}{\rho}(\lambda^k + \rho\, x^{k+1})\|^2 \\ \text{s.t.} & Ay = b \end{array}$$

Defining $z := (1/\rho)(\lambda^k + \rho\, x^{k+1})$, this is equivalent to projecting a point onto $\{y \,:\, Ay = b\}$:

$$\begin{array}{ll} \underset{y}{\text{minimize}} & \frac{1}{2}\|y - z\|^2 \\ \text{subject to} & Ay = b \, , \end{array} \tag{7}$$

which has the closed-form solution

$$y^\star = z - A^\top (AA^\top)^{-1}(Az - b) \, . \tag{8}$$

For large-scale problems, or if we only have access to the operations Ax and $A^\top y$ but not to the full matrix A, (8) can be computed via the conjugate gradient method.

References

[1] N. Vaswani, W. Lu, "Modified-CS: Modifying Compressive Sensing for Problems With Partially Known Support," *IEEE Transaction on Signal Processing*, Vol. 58, No. 9, 2010

[2] R. Glowinski and A. Marrocco, "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de dirichelet non linéaires," *Revue Française d'Automatique, Informatique, et Recherche Opérationelle*, vol. 9, no. 2, pp. 41-76, 1975.

[3] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximations," *Computers and Mathematics with Applications*, vol. 2, no. 1, pp. 17-40, 1976

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1-122, 2011