ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 10, No. 4, 2015, *pp.* 249-254

Linear Variational Inequality Model of L1-norm Minimization Problems with Applications to Compressive Sensing

Min Sun
School of Mathematics and Statistics, Zaozhuang University ,
Zaozhuang, 277160, China.
(*Received March 03, 2015, accepted July 21, 2015)*

Abstract. This paper considers the l_1 norm minimization (L1NM) problem which is a well known problem in compressive sensing. We first transform the L1NM problem into a bounded-constrained quadratic programming (BCQP), and then into an equivalent linear variational inequality (LVI) problem. To solve the resulting LVI problem, a modified extra-gradient method proposed by Han is introduced, whose global convergence can be guaranteed by Han's paper. The method is easily performed, since it only make a projection to the nonnegative orthant and calculate some matrix-vector products to get the next iterate. Numerical simulations are conducted to verify the efficiency of the proposed method.

Keywords: l_1 norm minimization problem; linear variational inequality problem; compressive sensing.

1. Introduction

In this paper we consider the following l_1 norm minimization problem, denoted by L1NM problem,

$$\min_{x\in R^n} \frac{1}{2}\|Ax-b\|_2^2+\mu\|x\|_1 \tag{1}$$

where $A\in R^{m\times n}(m<<n)$ is a linear operator, $b\in R^m$ is an observation, $x\in R^n$ is the vector of unknowns, $\|x\|_1=\sum_{i=1}^{n}|x_i|$ is the l_1 norm of x, and parameter $\mu>0$ is used to trade off both terms for minimization, whose value is quite important, for example, if μ is too large then the solution is the trivial one: $x=0$. Model (1) mainly appeared in statistical and signal processing fields, in which a sparse original signal $\bar{x}\in R^n$ is desirable to be recovered by solving the L1NM problem. In fact, preliminary work in this area showed that if the original signal is sparse or approximately sparse in some orthogonal basis, an exact restoration can be produced via solving (1) so long as certain conditions, such as the Restricted Isometry Property (RIP) hold [1,2].

In recent years, the study in numerical methods for (1) has taken good progress, and many efficient iterative algorithms have been proposed, analyzed, and tested. Among them, the most popular methods are the iterative shrinkage/thresholding (IST) type methods, including IST fixed-point continuation algorithm (FPC) in [3], two-step IST (TwIST) in [4], the fast IST algorithm (FISTA) in [5], and the latter two algorithms have virtually the same complexity as IST, but have better convergence performance. Gradient based algorithm is also quite efficient for solving the L1NM problem due to its simplicity. Gradient projection method for sparse reconstruction (GPSR) proposed by Figueiredo et al.[6] first transform the L1NM problem to a bound-constrained quadratic programming (BCQP) by splitting x and solves BCQP using Barzilai-Borwein gradient method with an efficient nonmonotone line search. Other gradient based methods can be found in [7,8]. In addition, the alternating direction method (ADM) algorithms are also introduced to solve the L1MN problem. For example, Yang and Zhang [9] investigates the L1MN problem from either the primal or the dual forms and solves some l_1 regularized problems related to L1MN problem.

In this paper, we continue to study the L1MN problem based on its BCQP transformation. As is pointed by Xiao et al.[10], the resulting BCQP is equivalent to a linear variational inequality (LVI) problem. To our knowledge, researchers haven't investigate the iterative method for the L1MN problem based on its LVI formulation. Here, the resulting LVI problem is solved by the modified extra-gradient method proposed by Han [11], which is an efficient method with quite low computational load. Therefore, the global convergence is followed directly in this literature. To do so, the rest of the paper is organized as follows. Section 2, we summarize some basic definitions used in the paper, and list the steps of our algorithm. In Section 3, we present and analyze the experimental results, which indicate that the proposed algorithm is quite efficient. Finally, we summarize our paper in Section 4.

2. Preliminaries and the algorithm

In this section, we briefly review some related knowledge, and state our algorithm.

Firstly, we give the definition of projection operator, which is defined as a mapping from R^n to its nonempty closed convex subset Ω:

$$P_\Omega[x] := \arg\min\{\|y-x\| \mid y\in\Omega\}, \forall x\in R^n.$$

In [6], Figueiredo et al. express the L1NM problem as a quadratic programming by splitting the variable x into its positive and negative parts. That is, for any vector $x\in R^n$, it can be formulated for

$$x = u - v, u\ge 0, v\ge 0,$$

where $u\in R^n, v\in R^n$, and $u_i=(x_i)_+, v_i=(-x_i)_+$ for all $i=1,2,\ldots,n$ with $(\cdot)_+=\max\{0,|\cdot|\}$. We thus have $\|x\|_1 = e_n^T u + e_n^T v$, where e_n is an n-dimensional vector with all elements one, so the L1NM problem (1) can be written as the following bound-constrained quadratic programming (BCQP):

$$\min_{u,v} \frac{1}{2}\|y - A(u-v)\|_2^2 + \mu e_n^T u + \mu e_n^T v$$
$$s.t. u\ge 0, v\ge 0.$$

Then, the above problem is further written in more standard BCQP form:

$$\min_{u,v} \frac{1}{2} z^T H z + c^T z$$
$$s.t. z \ge 0,$$

where $z=\begin{bmatrix} u \\ v\end{bmatrix}$, $y=A^T b, c=\mu e_{2n} + \begin{bmatrix} -y \\ y\end{bmatrix}$, and $H=\begin{bmatrix} A^TA & -A^TA \\ -A^TA & A^TA\end{bmatrix}$. Obviously, H is a positive semi-definite matrix, and for a given $z=\begin{bmatrix} u \\ v\end{bmatrix}$, the operations involving H can be performed economically,

$$Hz = H\begin{bmatrix} u \\ v\end{bmatrix} = \begin{bmatrix} A^TA(u-v) \\ -A^TA(u-v)\end{bmatrix}$$

and

$$z^T H z = (u-v)^T A^T A(u-v) = \|A(u-v)\|_2^2.$$

In compressive sensing (CS), matrix A is often formulated as discrete Fourier transform (DFT), discrete cosine transform (DCT), or discrete Walsh-Hadamard transform (DWT), and the multiplication involving A can be computed very fast.

Now, we are ready to transform the BCQP into a linear variational inequality (LVI) problem: Find $z \in R_+^{2n}$ such that

$$\langle Hz + c, z' - z \rangle \geq 0, \forall z' \geq 0. \tag{2}$$

It is well known that the above LVI can be equivalently solved by seeking a zero point of the mapping

$$e(z, \beta) = z - P_{R_+^{2n}}[z - \beta(Hz + c)],$$

where $\beta > 0$ is a constant.

Note that the function $Hz + c$ is monotone on R^{2n} since the matrix H is positive semi-definite.

This property ensures the global convergence of the following modified extra-gradient method for the L1NM problem which is based on the method proposed by Han [11].

Algorithm 2.1. (MEGM)

Step 0. Given $\varepsilon > 0$. Choose the starting point $z_0 = [u_0, v_0] \in R_+^{2n}$. Set the parameters $\gamma \in (0,2)$, $0 < \upsilon < \delta < 1, \beta > 0, \rho \in (0,1)$ and $k = 0$.

Step 1. Stop if $\|e(z, \beta)\| \leq \varepsilon$; else, go to Step 2.

Step 2. Compute the temporal point $\tilde{z}^k$ via

$$\tilde{z}^k = P_{R_+^{2n}}[z^k - \alpha_k (Hz^k + c)],$$

where $\alpha_k = \beta\rho^{m_k}$ with m_k being the smallest nonnegative integer m such that

$$\alpha_k \langle F(z^k) - F(\tilde{z}^k), e(z^k, \beta) - e(\tilde{z}^k, \beta) \rangle \leq \delta \| e(z^k, \beta) - e(\tilde{z}^k, \beta) \|^2. \tag{3}$$

Step 3. Compute the descent direction d_k by

$$d_k = \alpha_k (F(z^k) - F(\tilde{z}^k)).$$

Then, compute the stepsize $\rho(\tilde{z}^k)$ by

$$\rho(\tilde{z}^k) = \frac{\|e(\tilde{z}^k, \alpha_k)\|^2 - \alpha_k \langle F(z^k) - F(\tilde{z}^k), e(\tilde{z}^k, \alpha_k) \rangle}{\|d_k\|^2}$$

and get the next iterate

$$z^{k+1} = P_{R_+^{2n}}[z^k - \gamma\rho(\tilde{z}^k)d_k].$$

Step 4. If

$$\alpha_k \left\langle F(z^k) - F(\tilde{z}^k), e(z^k,\beta) - e(\tilde{z}^k,\beta) \right\rangle \le \upsilon \left\| e(z^k,\beta) - e(\tilde{z}^k,\beta) \right\|^2.$$

then, set $\beta_{k+1} = \beta_k / \eta$. Set $k = k+1$ and go to Step 1.

Remark 2.1. The MEGM method can be extended to solve the box constrained compressing sensing problem, that is $l \le x \le h$. In fact, we only need to modify the projection operator in Step 1. However, the non-smooth equations based method proposed by Xiao et al. [10] cannot deal this issue.

Based on Theorem 3 in [11], the global convergence of MEGM method can be stated as follows.

Theorem 2.1. The iterate sequence $\{z^k\}$ generated by MEGM method either terminates in a finite

number of steps, or converges to a solution point z^*, a solution of the BCQP (set $\varepsilon = 0$).

3. Numerical experiments

In this section, we present some experiments to illustrate the efficiency of the proposed method. All experiments were performed under Windows XP operating system and Matlab 7.0 running on a Lenovo laptop with AMD C-50 Processor at 797 MHz, 1.60GB of memory. We measure the quality of restoration by means of squared error (MSE) to the original signal $\tilde{x}$, that is

$$MSE = \frac{1}{n}\left\| \tilde{x} - x^* \right\|_2^2,$$

where x^* is the restored signal. For the MEGM method, we set $\rho = 0.6, \delta = 0.75, \gamma = 1.95, \beta_0 = 1$, $\upsilon = 0.45$. In our first experiment, we consider a typical CS scenario, where the goal is to reconstruct a length-n sparse signal from m observation, where $m < n$. Here, we set $n = 2^{10}, m = 2^8$, and the original signal contains 128 randomly placed spikes. The $m \times n$ matrix A is obtained by first filling it with independent samples of a standard Gaussian distribution and then ortho-normalizing the rows. The observation b is generated by:

$$b = Ax + \omega,$$

where ω is the Gaussian noise distributed as $N(0, \sigma^2 I)$. Here we choose $\sigma^2 = 10^{-4}$. Parameter

μ is set as

$$\mu = 0.01\left\| A^T b \right\|_\infty.$$

We use $f(x) = \mu\|x\|_1 + \|Ax - b\|^2 / 2$ as the merit function and stop the MEGM method when the relative change of the objective function is below 10^{-5}, i.e.,

$$\frac{\left|f_k - f_{k-1}\right|}{f_{k-1}} \le 10^{-5}$$

The original signal, the measurement and the reconstructed signal by MEGM method are given in Figure 1.

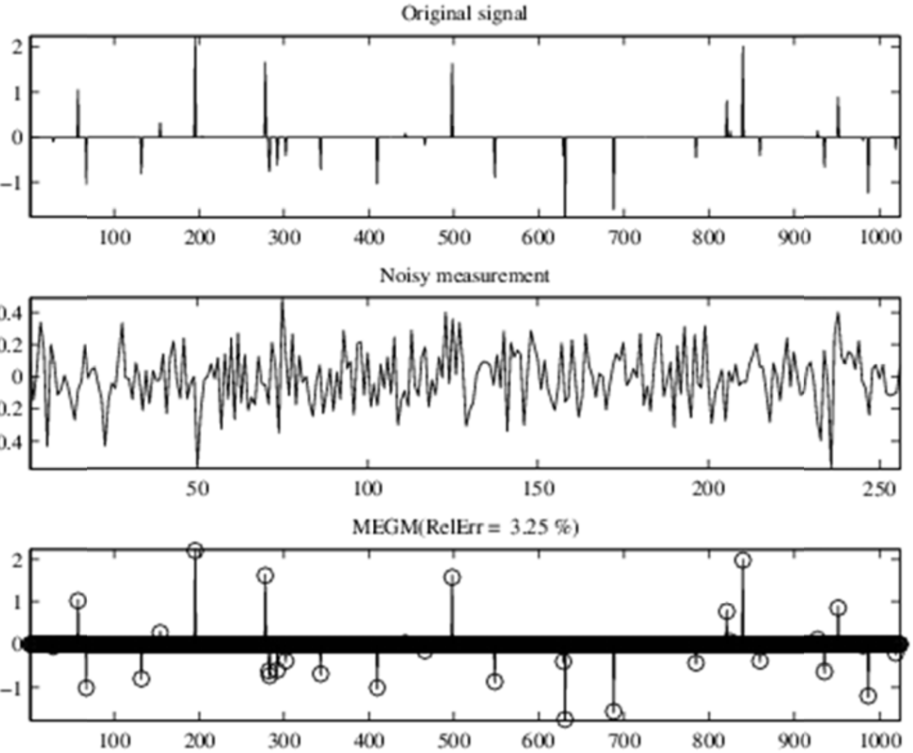

Figure 1: The original signal, noisy measurement and reconstruction results

Compared the first and the last plots in Figure 1, we clearly see that the original signal is recovered almost exactly. In addition, the RelErr=3.25%, the computing time is 11.8594s, the number of iteration is 298. The above results indicate that MEGM is efficient for the given simple problem.

4. Conclusion

In this paper, based on the nonsmooth equations formulation of compressive sensing, we propose a new method for such problem, which is an extension of the modified extra-gradient method for variational inequality problems proposed by Han.The numerical results indicate that the proposed method is efficient. In the future, we shall investigate other useful applications of the nonsmooth equations.

5. References

[1] D. Donoho, *Compressed sensing,* IEEE Trans. Inform. Theory, 52 (2006), pp. 1289 – 1306.

[2] M. Figueiredo and R. Nowak, *An EM algorithm for wavelet-based image restoration*, IEEE Trans. Image Process., 12 (2003), pp. 906 – 916.

[3] E. Hale, W. Yin, and Y. Zhang, *Fixed-point continuation for L1-minimization: Methodology and convergence,* SIAM J. Optim., 19 (2008), pp. 1107 – 1130.

[4] T. Goldstein and S. Osher, *The split Bregman method for L1-regularized problems*, SIAM J. Imaging Sci., 2 (2009), pp. 323 – 343.

[5] A. Beck and M. Teboulle, *A fast iterative shrinkage-thresholding algorithm for linear inverseproblems*, SIAM J.

Imaging Sci., 2 (2009), pp. 183 – 202.

[6] M. Figueiredo, R. Nowak, and S. J. Wright, *Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems*, IEEE J. Sel. Top. Signa., 1 (2007), pp. 586 – 597.

[7] E. Hale, W. Yin, and Y. Zhang, *Fixed-point continuation applied to compressed sensing: Implementation and numerical experiments*, J. Comput. Math., 28 (2010), pp. 170 – 194.

[8] J.-F. Yang, Y. Zhang, and W. Yin, *A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data*, IEEE J. Sel. Top. Signa., 4 (2010), pp. 288 – 297.

[9] J. F. Yang, Yin Zhang, *Alternating direction algorithms for L1-problems in compressive sensing*, SIAM Journal on Scientific Computing, 33(1), 250 – 278, 2011.

[10] Y.H. Xiao, Q.Y. Wang, Q.J. Hu, *Non-smooth equations based method for image-norm problems with applications to compressed sensing,* Nonlinear Analysis: Theory, Methods and Applications, 74(11), 2011, pp.3570-3577

[11] D.R. Han, *A new modified extra-gradient method for variational inequality problems*, Applied Mathematics and Optimization 45 (2002) , pp.63-74.