An Unconstrained ℓ_q Minimization with $0 < q \leq 1$ for Sparse Solution of Under-determined Linear Systems

Ming-Jun Lai *and Jingyue Wang
Department of Mathematics
The University of Georgia
Athens, GA 30602.

October 28, 2009

Abstract

We study an unconstrained version of the ℓ_q minimization for the sparse solution of under-determined linear systems for $0 < q \leq 1$. Although the minimization is non-convex, we introduce a regularization and develop an iterative algorithm. We show that the iterative solutions converge to the sparse solution. Numerical experiments will be demonstrated to show that our approach works very well.

Key Words and Phrases: Compressed Sensing, Sparse Solution, ℓ_q Minimization
Mathematics Subject Classification 2000: 15A09, 41A15, 65F10, 90C26
Short Title: An Unconstrained ℓ_q Minimization with $0 < q \leq 1$

1 Introduction

We are interested in computing the sparse solution of under-determined linear systems in the following sense: letting A be a matrix of size $m \times N$ with $m << N$ and $\mathbf{b}$ be a vector which is compressible, i.e., there exists a vector $\mathbf{x}^*$ with $\|\mathbf{x}^*\|_0 < m$ such that $\mathbf{b} = A\mathbf{x}^*$, we would like to find the solution of the following minimization

$$\min_{\mathbf{x}\in\mathbf{R}^N} \{\|\mathbf{x}\|_0, \quad A\mathbf{x} = \mathbf{b}\}, \tag{1}$$

where $\|\mathbf{x}\|_0$ denotes the number of nonzero components of $\mathbf{x}$. The solution is called the sparse solution of $A\mathbf{x} = \mathbf{b}$. This is one of critical problems in compressed sensing research.

*mjlai@math.uga.edu. This author is partly supported by the National Science Foundation under grant DMS-0713807

This problem is motivated by data compression, error correction decodes, n-term approximation, and etc.. (See, e.g. [23]). It is known that the problem (1) needs non-polynomial time to solve (cf. [25]). It is crucial to recognize that one natural approach to tackle (1) is to solve the following convex minimization problem:

$$\min_{\mathbf{x}\in\mathbf{R}^N}\{\|\mathbf{x}\|_1, \quad A\mathbf{x}=\mathbf{b}\}, \tag{2}$$

where $\|\mathbf{x}\|_1 = \sum_{j=1}^N |x_j|$ is the standard ℓ_1 norm. Note that it is equivalent to linear programming problem which can be solved by the simplex method or interior point method and their variations.

The study of this problem (2) was pioneered by Donoho, Candés and their collaborators. Many other researchers have made a lot of contributions. A lot of results related to the existence, uniqueness, and other properties of the sparse solution as well as computational algorithms and their convergence analysis to tackle Problem (1) are available in the literature. See survey papers in [1], [3], and [2].

To motivate our study, let us outline some research results related to numerical algorithms for the computation of sparse solutions of (1). First of all, the ℓ_1 minimization (2) by Candés and his collaborators (cf. [5]) is a successful approach to find sparse solutions (1) if the sparsity $s = \|\mathbf{x}\|_0$ is not very large. A matlab program based on a linear programming method for the sparse solution is available on-line at the Candés webpage. The performance of the ℓ_1 method is further improved based on the ideas of repeating reweighted iteration (cf. [8]). Another approach is based on orthogonal greedy algorithm (OGA). See [28] and [29] for some theoretic study and [27] for an efficient numerical algorithm. The performance of the OGA in [27] is much improved based on the greedy ℓ_1 algorithm proposed recently in [22]. Another approach for the computation of the sparse solutions is based on ℓ_q minimization with $0 < q < 1$. That is, we consider the following

$$\min_{\mathbf{x}\in\mathbf{R}^N}\{\|\mathbf{x}\|_q^q, \quad A\mathbf{x}=\mathbf{b}\}, \tag{3}$$

where $\|\mathbf{x}\|_q^q = \sum_{j=1}^N |x_j|^q$ for $0 < q \le 1$. This minimization is motivated by the following fact:

$$\lim_{q\to 0_+} \|\mathbf{x}\|_q^q = \|\mathbf{x}\|_0.$$

This approach was initiated by [21] and many researchers have worked on this direction. There are at least three advantages of using this approach to the best of the authors knowledge. One is the result in [10]: for a Gaussian random matrix A, the restricted q-isometry property of order s holds if s is almost proportional to m when $q \to 0_+$. Another advantage demonstrated in [18] is that the sufficient conditions for the solution of the ℓ_q minimization to be a sparse solution is weaker that the Candés result in [4] for the ℓ_1 minimization. The third advantage is that the ℓ_q minimization can be applied to a wider class of random matrices A, e.g., when A is a random matrix whose entries are iid copies of a pre-Gaussian random variable. See [19]. In addition, there are many other approaches, e.g., optimal basis pursuit(OMB) method (for problem (2)), soft-thresholding iterations, standard and damped Landweber iterations ([11]) for problem (2), iterative reweighted least squares (IRLS) method (cf. [13]) (for problems (2) and (3)) and etc..

In this paper we shall consider another version of ℓ_q minimization:

$$\min_{\mathbf{x}\in\mathbf{R}^N} \|\mathbf{x}\|_q^q + \frac{1}{2\lambda}\|A\mathbf{x}-\mathbf{b}\|_2^2, \tag{4}$$

which $\|\mathbf{x}\|_2^2 = \sum_{j=1}^N x_j^2$ and $\lambda > 0$ is a parameter which is sufficiently small, e.g., $\lambda = 10^{-8}$. Clearly, this is a standard unconstrained version of the original ℓ_q minimization (3). Due to the singularity of the gradient of the associated functional above because of the sparsity of the solution $\mathbf{x}$, we introduce the following regularized version of the unconstrained ℓ_q minimization:

$$\min_{\mathbf{x}\in\mathbf{R}^N} \|\mathbf{x}\|_{q,\epsilon} + \frac{1}{2\lambda}\|A\mathbf{x}-\mathbf{b}\|_2^2, \tag{5}$$

where

$$\|\mathbf{x}\|_{q,\epsilon} = \sum_{j=1}^N (\epsilon + x_j^2)^{q/2}$$

and $\epsilon > 0$ is another parameter which will go to zero in order to approximate $\|\mathbf{x}\|_q^q$. This is the main minimization problem we study in this paper. Although many unconstrained versions of problem (2) have been studied in the literature (cf. [30] and references therein), the unconstrained ℓ_q minimization (4) has not been analyzed yet in the literature. We shall show that the above problem (5) has a solution for any $q \in (0, 1]$ and $\epsilon > 0$. We also derive an iterative algorithm to compute the solution $\mathbf{x}^{\epsilon,q}$ of (5). We prove that the iterative algorithm is convergent for any starting point. We shall show that $\mathbf{x}^{\epsilon,q}$ has a convergent subsequence which converges to a minimizer $\mathbf{y}^q$ of the unconstrained ℓ_q minimization problem (4). Finally we show that $\mathbf{y}^q$ converge to the sparse solution of our original under-determined linear system (1) by using the Γ-convergence notation. See our main results in Theorem 2.3. According to our theory, we can completely determine a sparse solution without any conditions, e.g., the restricted isometry property (RIP) on matrix A. This is an another advantage of the ℓ_q minimization for $0 < q < 1$ over the classic ℓ_1 minimization approach.

However, it is easy to see that there is a lot of computations for various ϵ, q and λ to perform. Namely we have to compute $\mathbf{x}^{\epsilon,q}$ for many different small values ϵ for a fixed q to get a limit $\mathbf{y}^q$ of a convergent subsequence. Each $\mathbf{x}^{\epsilon,q}$ is computed by using an iterative algorithm which is proved to be convergent. Then we look for a convergent subsequence of $\{\mathbf{y}^q, q > 0\}$. Thus, the computation is expensive. However, if we know the information about the sparsity, say s of the sparse solution of (1), we can determine immediately if $\mathbf{x}^{\epsilon,q}$ is already a solution or not by checking if $\|\mathbf{x}^{\epsilon,q}\|_0 = s$ or not and if $A\mathbf{x}^{\epsilon,q} - \mathbf{b} = 0$ or not. The iteration can be quickly stopped. Suppose we do not know the sparsity. Letting $\mathbf{y}^*$

be the limit of a subsequence of $\{\mathbf{y}^q, q > 0\}$, if the sparsity of $\mathbf{y}^*$ is $\leq m/2$, the chance of $\mathbf{y}^*$ to be the sparse solution is big. Reduce ϵ and use $\mathbf{y}^*$ as an initial solution to compute the minimization problem (5) again. Otherwise, if the sparsity of $\mathbf{y}^*$ is m or closed to m, then the chance of $\mathbf{y}^*$ to be a solution of the following minimization (6) is slim. It is better to start with a completely new initial guess and construct another subsequence $\mathbf{y}^q$. This gives reasons of the better performance of our algorithm than the other schemes in our numerical experiments. Note that when $q \to 0_+$, the unconstrained ℓ_q minimization converges to

$$\min_{\mathbf{x}\in\mathbf{R}^N} \|\mathbf{x}\|_0 + \frac{1}{2\lambda}\|A\mathbf{x} - \mathbf{b}\|_2^2, \tag{6}$$

which is the unconstrained version of our main problem (1). With λ small enough, the solution can be viewed as a good numerical approximation of the sparse solution. Extensive numerical experiments have been performed to compare with many other methods as explained above. Our unconstrained ℓ_q minimization does indeed perform much better. In particular, our approach performs the best for under-determined linear systems $A\mathbf{x} = \mathbf{b}$ with uniform random matrices A, i.e., the entries of A are random variables of uniform distribution.

The paper consists three sections, in addition to this introductory section: our analysis of the unconstrained ℓ_q minimization when $q < 1$ in §2, some additional properties of the unconstrained ℓ_q minimization when $q = 1$ in §3, and finally in §4 numerical results to demonstrate how well our unconstrained ℓ_q minimization can find the sparse solutions.

2 Analysis of Unconstrained ℓ_q Minimization

We begin with some elementary properties of the minimization problem (5). Let $L_q(\epsilon, \mathbf{x})$ be the function associated with (5). It is easy to see that the problem has a solution. We use $\mathbf{x}^{\epsilon,q}$ to denote a minimizer of (5).

Consider the following one variable function of α

$$L_q(\epsilon, \mathbf{x}^{\epsilon,q} + \alpha\mathbf{y}) = \|\mathbf{x}^{\epsilon,q} + \alpha\mathbf{y}\|_{q,\epsilon} + \frac{1}{2\lambda}\|A\mathbf{x}^{\epsilon,q} + \alpha A\mathbf{y} - \mathbf{b}\|^2.$$

A minimizer $\mathbf{x}^{\epsilon,q}$ satisfies the following gradient equations

$$\left[\frac{q x_j^{\epsilon,q}}{(\epsilon + (x_j^{\epsilon,q})^2)^{1-q/2}}\right]_{1\leq j\leq N} + \frac{1}{\lambda}A^T(A\mathbf{x}^{\epsilon,q} - \mathbf{b}) = 0. \tag{7}$$

In fact, this is only a necessary condition. We call any solution satisfying the above equations a critical point.

We now derive an iterative method to solve the above equations due to their nonlinearity. Starting with any initial $\mathbf{x}^{(1)}$, we solve the following system of linear equations for

$\mathbf{x}^{k+1}$:

$$\left[\frac{qx_j^{(k+1)}}{(\epsilon+(x_j^{(k)})^2)^{1-q/2}}\right]_{1\le j\le N}+\frac{1}{\lambda}A^T(A\mathbf{x}^{(k+1)}-\mathbf{b})=0. \tag{8}$$

or

$$\left(A^TA+\ \mathrm{diag}\ \left[\frac{q\lambda}{(\epsilon+|x_j^{(k)}|^2)^{1-q/2}}, j=1,\cdots,N\right]\right)\mathbf{x}^{k+1}=A^T\mathbf{b} \tag{9}$$

for $k=1,2,3,\cdots,$. It is easy to see that the above linear system is invertible for any $\mathbf{x}^{(k)}$ as long as $\epsilon>0$. Thus, the iterative method is well defined. We now show that these $\mathbf{x}^{(k)}$ converge to a critical point of the minimization problem (5). We begin with

Lemma 2.1 *Fix any $\epsilon>0$. Let $\mathbf{x}^{k+1}$ be the solution of (9) for $k=1,2,3,\cdots$. Then*

$$L_q(\epsilon,\mathbf{x}^{(k+1)})\le L_q(\epsilon,\mathbf{x}^{(k)}). \tag{10}$$

Furthermore,

$$\|A\mathbf{x}^{(k)}-A\mathbf{x}^{(k+1)}\|^2\le 2\lambda(L_q(\epsilon,\mathbf{x}^{(k)})-L_q(\epsilon,\mathbf{x}^{(k+1)})). \tag{11}$$

Proof. Mainly we need the following inequality

$$(\epsilon+|x|^2)^{q/2}-(\epsilon+|y|^2)^{q/2}-\frac{qy(x-y)}{(\epsilon+|x|^2)^{1-q/2}}\ge 0. \tag{12}$$

This inequality can be verified by a direct computation. We now compute as follows:

$$\begin{aligned}L_q(\epsilon,\mathbf{x}^{(k)})-L_q(\epsilon,\mathbf{x}^{(k+1)}) &= \sum_{j=1}^N(\epsilon+|x_j^{(k)}|^2)^{q/2}-\sum_{j=1}^N(\epsilon+|x_j^{(k+1)}|^2)^{q/2}\\ &\quad+\frac{1}{2\lambda}\left(\|A\mathbf{x}^{(k)}-\mathbf{b}\|^2-\|A\mathbf{x}^{(k)}-\mathbf{b}\|^2\right)\\ &= \sum_{j=1}^N(\epsilon+|x_j^{(k)}|^2)^{q/2}-(\epsilon+|x_j^{(k+1)}|^2)^{q/2}+\frac{1}{2\lambda}\|A\mathbf{x}^{(k)}-A\mathbf{x}^{(k+1)}\|^2\\ &\quad+\frac{1}{\lambda}(A\mathbf{x}^{(k+1)}-\mathbf{b})^T(A\mathbf{x}^{(k)}-A\mathbf{x}^{(k+1)}).\end{aligned}$$

The last term can be simplified to be

$$-\sum_{j=1}^N\frac{qx_j^{(k+1)}(x_j^{(k)}-x_j^{(k+1)})}{(\epsilon+|x_j^{(k)}|^2)^{1-q/2}}$$

by using (8) based on a dot-product with $\mathbf{x}^{(k)}-\mathbf{x}^{(k+1)}$. With this we have

$$L_q(\epsilon,\mathbf{x}^{(k)})-L_q(\epsilon,\mathbf{x}^{(k+1)})$$

$$
\begin{aligned}
&= \sum_{j=1}^{N}\left((\epsilon+|\mathbf{x}_j^{(k)}|^2)^{q/2}-(\epsilon+|\mathbf{x}_j^{(k+1)}|^2)^{q/2}-\frac{x_j^{(k+1)}(x_j^{(k)}-x_j^{(k+1)})}{(\epsilon+|x_j^{(k)}|^2)^{1-q/2}}\right)\\
&\quad +\frac{1}{2\lambda}\|A\mathbf{x}^{(k)}-A\mathbf{x}^{(k+1)}\|^2\\
&\geq \frac{1}{2\lambda}\|A\mathbf{x}^{(k)}-A\mathbf{x}^{(k+1)}\|^2 \geq 0
\end{aligned}
$$

by using (12). The two results (10) and (11) follow immediately. ■

The proof of the Lemma above can be extended to show

Proposition 2.1 *Let $\mathbf{x}^{\epsilon,q}$ be any critical point for problem (5). Then for any $\mathbf{y}$*

$$\|A\mathbf{y}-A\mathbf{x}^{\epsilon,q}\|^2 \leq 2\lambda(L_q(\epsilon,\mathbf{y})-L_q(\epsilon,\mathbf{x}^{\epsilon,q})). \tag{13}$$

Next we introduce the concept of completely full rank. We call a matrix A of size $m\times n$ with $m<n$ *completely full rank* if any submatrices of A of size $m\times m$ are full rank. For example, $A=[(x_j)^{i-1}]_{1\leq i\leq m, 1\leq j\leq n}$ with x_j distinct is completely full rank.

Lemma 2.2 *Suppose that A is completely full rank. Let*

$$\mathcal{A}=\begin{bmatrix} A & 0_m \\ I_n & R_m \end{bmatrix}.$$

where 0_m is a zero block matrix of size $m\times m$, I_n is the identity matrix of size $n\times n$, and R_m is a zero matrix except for $R(r(i),i)=1$ for $i=1,\cdots,m$ with $r(1),\cdots,r(m)$ being the first m entries of a random permutation of integers $1,2,\cdots,n$. Then $\mathcal{A}$ is invertible and $\|\mathcal{A}^{-1}\|_2$ is bounded above by a constant C which is dependent on A for any random permutation.

Proof. Without loss of generality, we may assume that

$$R_m=\begin{bmatrix} I_m \\ 0_{n-m,m} \end{bmatrix}$$

with $0_{n-m,m}$ being a zero block matrix of size $(n-m)\times m$. Since A is completely full rank, we use the rows from $m+1$ to $2m$ of $\mathcal{A}$ to make $\mathcal{A}(1:m,1:m)$ to be zero. Then we use the rows $2m+1$ to $m+n$ to make $\mathcal{A}(1:m,m+1:n)$ to zero. Note that $\mathcal{A}(1:m,n+1:n+m)$ is $-A(1:m,1:m)$. Clearly, the resulting matrix of $\mathcal{A}$ is invertible and the norm of the inverse of the resulting matrix is dependent on the norm of the inverse of $A(1:m,1:m)$. Similar for other random matrix R_m. ■

With this auxiliary result, we show

Proposition 2.2 *Suppose that A is completely full rank. Let $\mathbf{x}^{\epsilon,q}$ be a sparse solution for problem (5) with sparsity $\|\mathbf{x}^{\epsilon,q}\|_0\leq m/2$. Then for any $\mathbf{y}$ with sparsity $\|\mathbf{y}\|_0\leq m/2$,*

$$\|\mathbf{y}-\mathbf{x}^{\epsilon,q}\|^2 \leq C^2\lambda(L_q(\epsilon,\mathbf{y})-L_q(\epsilon,\mathbf{x}^{\epsilon,q})). \tag{14}$$

Proof. Consider $\mathbf{x}^{\epsilon,q} - \mathbf{y}$ whose sparsity is at most m. Without loss of generality, we may assume that the first m entries of $\mathbf{x}^{\epsilon,q} - \mathbf{y}$ are nonzero. Let $\mathbf{z}$ be a vector of size $(n+m) \times 1$ which consists of all $\mathbf{x}^{\epsilon,q} - \mathbf{y}$ and then the negative of the first m components of $\mathbf{x}^{\epsilon,q} - \mathbf{y}$. It is easy to see that

$$\mathcal{A}\mathbf{z} = \begin{bmatrix} A\mathbf{x}^{\epsilon,q} - A\mathbf{y} \\ 0_{m,1} \end{bmatrix}.$$

$$\|\mathbf{z}\|_2^2 = 2\|\mathbf{x}^{\epsilon,q} - \mathbf{y}\|_2^2 \le \|\mathcal{A}^{-1}\|_2^2 \|A\mathbf{x}^{\epsilon,q} - A\mathbf{y}\|_2^2 \le 2C^2\lambda(L_q(\epsilon,\mathbf{y}) - L_q(\epsilon,\mathbf{x}^{\epsilon,q})).$$

where we have used Lemma 2.2 with C being a constant dependent on A. This completes the proof. ∎

Similarly, we can prove the following

Proposition 2.3 *Suppose that A is completely full rank. Let $\mathbf{x}^{\epsilon,q}$ be a solution for problem (5) with sparsity $\|\mathbf{x}^{\epsilon,q}\|_0 \le m/2$. Let $\mathbf{x}^{(k)}, k = 1,2,3,\cdots$ be a sequence from the iterative solution above starting from the zero initial guess $\mathbf{x}^{(0)} = 0$. Suppose that all $\mathbf{x}^{(k)}$ has the sparsity $\|\mathbf{x}^{(k)}\|_0 \le m/2$. Then*

$$\sum_{k=1}^{\infty} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|_2^2 \le C2\lambda L(\epsilon, \mathbf{x}^{(0)}).$$

We are now ready to prove the convergence of our iterative algorithm.

Proposition 2.4 *Fix each $\epsilon > 0$. There exists $\mathbf{x}^* \in \mathbf{R}^N$ such that*

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*$$

and $\mathbf{x}^*$ is a critical point of the problem (5).

Proof. By Lemma 2.1, $L_q(\epsilon, \mathbf{x}^{(k)})$ is decreasing. Let $\lim_{k\to\infty} L_q(\epsilon,\mathbf{x}^{(k)}) = M$. It is clear that $\|\mathbf{x}^{(k)}\|_q$ is bounded and hence, there exist a vector $\mathbf{x}$ and a convergent subsequence $\mathbf{x}^{k_j}$ such that $\mathbf{x}^{k_j} \to \mathbf{x}$. Note that $\mathbf{x}^{(k_j+1)}$ solves (8). By (9), $\mathbf{x}^{(k_j+1)}$ is also a convergent subsequence. Let us say $\mathbf{x}^{(k_j+1)} \to \mathbf{y}$. Then by (11),

$$\|A\mathbf{x} - A\mathbf{y}\|^2 \le 2\lambda(L_q(\epsilon,\mathbf{x}) - L_q(\epsilon,\mathbf{y})).$$

Due to the monotonicity, i.e., (10), we conclude that $L_q(\epsilon,\mathbf{x}) = L_q(\epsilon,\mathbf{y})$. Thus we have $A\mathbf{y} = A\mathbf{x}$. Then it follows that $\|\mathbf{x}\|_{q,\epsilon} = \|\mathbf{y}\|_{q,\epsilon}$.

We now claim that $\mathbf{x} = \mathbf{y}$. Indeed, using dot product in (8) with $\mathbf{x} - \mathbf{y}$ for $k = k_j$ and letting $j \to \infty$, we have

$$\sum_{j=1}^{N} \frac{q y_j (x_j - y_j)}{(\epsilon + |x_j|^2)^{1-q/2}} + \frac{1}{\lambda}(A\mathbf{x} - A\mathbf{y})^T (A\mathbf{y} - \mathbf{b}) = 0$$

As we have proved that $A\mathbf{x} = A\mathbf{y}$ above, we have

$$\sum_{j=1}^{N} \frac{y_j(x_j - y_j)}{(\epsilon + |x_j|^2)^{1-q/2}} = 0.$$

Combining the above equation with the fact $\|\mathbf{x}\|_{q,\epsilon} = \|\mathbf{y}\|_{q,\epsilon}$ we just proved above, we end with

$$\|\mathbf{x}\|_{q,\epsilon} - \|\mathbf{y}\|_{q,\epsilon} - \sum_{j=1}^{N} \frac{qy_j(x_j - y_j)}{(\epsilon + |x_j|^2)^{1-q/2}} = 0.$$

In other words,

$$\sum_{j=1}^{N} \left((\epsilon + |x_j|^2)^{q/2} - (\epsilon + |y_j|^2)^{q/2} - \frac{qy_j(x_j - y_j)}{(\epsilon + |x_j|^2)^{1-q/2}} \right) = 0$$

By inequality (12), each summation term is nonnegative and hence has to be zero term by term. Furthermore, each term can be rewritten as

$$\begin{aligned} 0 &= (\epsilon + |x_j|^2)^{q/2} - (\epsilon + |y_j|^2)^{q/2} - \frac{y_j(x_j - y_j)}{(\epsilon + |x_j|^2)^{1-q/2}} \\ &= \frac{q|x_j|^2 - 2qx_jy_j + q|y_j|^2}{2(\epsilon + |x_j|^2)^{1-q/2}} + \frac{2\epsilon + (2-q)|x_j|^2 + q|y_j|^2 - 2(\epsilon + |x_j|^2)^{1-q/2}(\epsilon + |y_j|^2)^{q/2}}{2(\epsilon + |x_j|^2)^{1-q/2}}. \end{aligned}$$

Since both of the above two terms are nonnegative, it follows from the first term above that $x_j = y_j$ for all j and hence, $\mathbf{x} = \mathbf{y}$. Let us denote $\mathbf{x}$ by $\mathbf{x}^*$. Therefore, from (8) for $k = k_j$ with $j \to \infty$, we know $\mathbf{x}^*$ satisfies (7) with $\mathbf{x}^*$ in place of $\mathbf{x}^{\epsilon,q}$. Thus $\mathbf{x}^*$ is a critical point. ■

When $q = 1$, we can see that the functional $L_1(\epsilon, \mathbf{x})$ is strictly convex. The uniqueness of the minimizer which satisfies the gradient equation implies that our iterative limit $\mathbf{x}^* = \mathbf{x}^{\epsilon,1}$, the minimizer of (5) with $q = 1$. Although we only proved that $\mathbf{x}^*$ is a critical point when $q < 1$, numerical experiments show that the $\mathbf{x}^*$ achieves the minimum for some ϵ small enough and hence, these $\mathbf{x}^*$ for all $\epsilon \to 0_+$ lead to the sparse solution of the under-determined linear system (1). Indeed, we manage to prove the following two results:

Theorem 2.1 *Fix $q > 0$. Let $\mathbf{x}^k$ be a critical point of (5) associated with ϵ_k for $k = 1, 2, 3, \cdots$ with ϵ_k decreasing to 0. There exists a subsequence from $\mathbf{x}^k$ which converges to $\mathbf{y}^q$ and $\mathbf{y}^q$ is a minimizer of (4).*

Proof. By using Lemma 2.1, it is easy to see that

$$L_q(\epsilon_{k+1}, \mathbf{x}^{(k+1)}) \le L_q(\epsilon_{k+1}, \mathbf{x}^{(k)}) \le L_q(\epsilon_k, \mathbf{x}^{(k)}) \le L_q(\epsilon_k, \mathbf{x}^*)$$

for any minimizer $\mathbf{x}^*$ of (4). Now it is clear that all these $\mathbf{x}^{(k)}$ are bounded and hence, there exists a convergent subsequence. For simplicity, let us say $\mathbf{x}^{(k)}$ converges to $\mathbf{y}^q$. Thus we have

$$L_q(0, \mathbf{y}^q) = \lim_{k\to\infty} L_q(0, \mathbf{x}^{(k+1)}) \le \lim_{k\to\infty} L_q(\epsilon_k, \mathbf{x}^{(k+1)}) \le \lim_{k\to\infty} L_q(\epsilon_k, \mathbf{x}^*) = L_q(0, \mathbf{x}^*).$$

Since the functional $L_q(0, \mathbf{x})$ is the functional in (4), the definition of the minimizer of (4) implies $L_q(0, \mathbf{x}^*) \le L_q(0, \mathbf{y}^q)$. It follows that

$$L_q(0, \mathbf{x}^*) = L_q(0, \mathbf{y}^q)$$

That is, $\mathbf{y}^q$ is also a minimizer of (4). ■

We give a rough error bound for $\|\mathbf{y} - \mathbf{x}^{\epsilon,q}\|$ by the result of Theorem 2.1.

Proposition 2.5 *Suppose that A is completely full rank. Let $\mathbf{x}^*$ be the sparse solution of (1) and $\mathbf{y}^q$ be the minimizer of the unconstrained ℓ_q minimization problem (4), $\|\mathbf{x}^*\|_0 \le m/2$, $\|\mathbf{y}^q\|_0 \le m/2$. Then*

$$\|\mathbf{y}^q - \mathbf{x}^*\|^2 \le C^2\lambda(\|\mathbf{x}^*\|_q^q - \|\mathbf{y}^q\|_q^q) \tag{15}$$

Proof. Let $\mathbf{y} = \mathbf{x}^*$ and sequence $\{\mathbf{x}^{\epsilon_k,q}\}$ to be the subsequence in Theorem 2.1. We apply Proposition 2.2 and let ϵ_k tend to zero. The result follows. ■

Next we show that $\mathbf{y}^q$ converges to the sparse solution of the original problem (1) as $q \to 0_+$. We shall use the concept of Γ-convergence which was introduced by E. De Giorgi and T. Franzoni in 1975 (cf. [20]). We first give the definition for the Γ-convergence.

Definition 2.1 *Let (X, d) be a metric space with metric d. We say that a sequence of functionals $E_k : X \to [-\infty, \infty]$ is Γ-convergent to a functional $E : X \to [-\infty, \infty]$ as $k \to \infty$ if for all $u \in X$ we have*

(i) for every sequence $\{u_k \in X\}$ converging to u

$$E(u) \le \liminf_k E_k(u_k)$$

(ii) there exists a sequence $\{u_k \in X\}$ converging to u such that

$$E(u) \ge \limsup_k E_k(u_k),$$

or equivalently

$$E(u) = \lim_k E_k(u_k).$$

Next we prove that if the minimizers of E_k have a cluster point, it is a minimizer of E under the assumption of the γ-convergence of E_k to E. We start with the following

Lemma 2.3 *If a sequence of functionals E_k is $\Gamma-$convergent to a functional E on X as $k \to \infty$, for any subsequence $\{E_{k_j}\}$ of $\{E_k\}$,*

$$\limsup_{k_j \to \infty} \inf_{u \in X} E_{k_j}(u) \leq \inf_{v \in X}(v)$$

Proof. For any vector $v \in X$, by the definition of $\Gamma-$convergence, there exists $\{u_k\}$ converging to v such that,

$$\limsup_{k \to \infty} E_k(u_k) \leq E(v).$$

Note that $\inf_{u \in X} E_{k_j}(u) \leq E_{k_j}(u_{k_j})$,

$$\begin{aligned} \limsup_{k_j \to \infty} \inf_{u \in X} E_{k_j}(u) &\leq \limsup_{k_j \to \infty} E_{k_j}(u_{k_j}) \\ &\leq \limsup_{k \to \infty} E_k(u_k) \\ &\leq E(v). \end{aligned}$$

Since v is arbitrarily chose, we have

$$\limsup_{k_j \to \infty} \inf_{u \in X} E_{k_j}(u) \leq \inf_{v \in X} E(v).$$

■

One important consequence of a Γ-convergent sequence of functionals is the following standard result (cf. [24])

Theorem 2.2 *Suppose that a sequence of functionals E_k is $\Gamma-$convergent to a functional E on X as $k \to \infty$. Letting E_{k_j} be a subsequence and u_{k_j} be the minimizer of E_{k_j}, if u_{k_j} converges to u in X, then u is a minimizer of E.*

Proof. By the definition of $\Gamma-$convergence,

$$\begin{aligned} E(u) &\leq \liminf_{k_j \to \infty} E_{k_j}(u_{k_j}) \\ &\leq \limsup_{k_j \to \infty} E_{k_j}(u_{k_j}) \\ &= \limsup_{k_j \to \infty} \inf_{v \in X} E_{k_j}(v) \\ &\leq \inf_{v \in X} E(v). \end{aligned}$$

The first line follows from the definition of $\Gamma-$convergence and the last line is the result of Lemma 2.3. ■

More details on Γ convergence can be found in [24]. The preparatory results above are enough for our current purpose.

Consider $L_q(0,\mathbf{x}), q \in (0,1)$ to be a sequence of functionals. Let $L_0(0,\mathbf{x})$ be another functional associated with the minimization in (6). We claim that $L_q(0,\mathbf{x}), q \in (0,1)$ are Γ-convergent to $L_0(0,\mathbf{x})$. Indeed, for any sequence $\mathbf{x}^q \in \mathbf{R}^N$ which converge to $\mathbf{x}$ as $q \to 0_+$, we can see $\|\mathbf{b} - A\mathbf{x}^q\|_2^2$ converge to $\|\mathbf{b} - A\mathbf{x}\|_2^2$ easily. Writing $\mathbf{x} = (x_1, \cdots, x_N)^T$, let $\delta = \min\{|x_i| > 0\}$. We have

$$\begin{aligned} L_q(0,\mathbf{x}^q) &\geq \sum_{|x_i|>0} |x_{q,i}|^q + \frac{1}{2\lambda}\|\mathbf{b} - A\mathbf{x}^q\|_2^2 \\ &\geq \sum_{|x_i|>0} |\frac{1}{2}\delta|^q + \frac{1}{2\lambda}\|\mathbf{b} - A\mathbf{x}^q\|_2^2 \end{aligned}$$

for q sufficiently small. It follows that

$$\liminf_{q\to 0_+} L_q(0,\mathbf{x}^q) \geq \sum_{|x_i|>0} 1 + \frac{1}{2\lambda}\|\mathbf{b} - A\mathbf{x}\|_2^2 = L_0(0,\mathbf{x}).$$

On the other hand, for any $\mathbf{x}$, we choose a particular sequence $\mathbf{x}^q = \mathbf{x}$. Then we have

$$\limsup_{q\to 0} L_q(0,\mathbf{x}^q) = L_0(0,\mathbf{x}).$$

These show that $L_q(0,\mathbf{x}), q \in (0,1]$ are Γ-convergent to $L_0(0,\mathbf{x})$.

Assuming $\mathbf{x}^*$ be the sparse solution of our original problem (1) with sparsity $s = \|\mathbf{x}^*\|_0$, let

$$D = \min_{\|\mathbf{x}\|_0 \leq s-1} \|A\mathbf{x} - \mathbf{b}\|_2^2. \tag{16}$$

It is easy to see that $D > 0$. Then by Theorem 2.2, we conclude the following

Theorem 2.3 *Let $\mathbf{y}^q$ be a minimizer of the unconstrained ℓ_q minimization problem (4) with λ satisfying $\frac{D}{\lambda} > 2s$. Then $\mathbf{y}^q$ contains at least one convergent subsequence and the limit of any subsequence of $\mathbf{y}^q, q > 0$ is a sparse solution of (1).*

Proof. Since $\mathbf{y}^q$ are bounded, $\mathbf{y}^q$ contains a convergent subsequence. We use Theorem 2.2 to conclude that the limit, say $\mathbf{x}^0$ of the convergent subsequence is a minimizer of $L_0(0,\mathbf{x})$. Since the under-determined linear system has a sparse solution $\mathbf{x}^*$, $L_0(0,\mathbf{x}^*) = \|\mathbf{x}^*\|_0$ which is the minimal value for any λ.

Since $\mathbf{x}^0$ is a minimizer of $L_0(0,\mathbf{x})$, we have $L_0(0,\mathbf{x}^0) = \|\mathbf{x}^*\|_0$ by Lemma 2.2. That is, $\mathbf{x}^0$ has to be a vector such that $A\mathbf{x}^0 = \mathbf{b}$ and $\|\mathbf{x}^0\|_0 = \|\mathbf{x}^*\|_0$. Otherwise, if $A\mathbf{x}^0 \neq \mathbf{b}$, then $\|\mathbf{x}^0\|_0 \leq s-1$ and

$$L_0(0,\mathbf{x}^0) = \|\mathbf{x}^0\|_0 + \frac{1}{2\lambda}\|A\mathbf{x}^0 - \mathbf{b}\|_2^2 \geq \|\mathbf{x}^0\|_0 + \frac{1}{2\lambda}D > 1 + s$$

which contradicts to the fact $L_0(0,\mathbf{x}^0) = \|\mathbf{x}^*\|_0 = s$. This completes the proof. ■

As the sparse solution $\mathbf{x}^*$ may not be unique, the sequence $\mathbf{y}^q, q \in (0,1)$ does not converge in general. The result in Theorem 2.3 above shows that the limit of any subsequence is a sparse solution of (1).

3 Some Additional Properties of Unconstrained ℓ_1 Minimization

In this section we exhibit more properties of the unconstrained ℓ_q minimization when $q = 1$. We have the following stability property of the unconstrained ℓ_1 minimization.

Proposition 3.1 (Stability) *Suppose that $q = 1$. Let $\mathbf{x_b}$ be a minimizer for input data $\mathbf{b}$ in problem (5). Similarly, for an input data $\mathbf{c}$, let $\mathbf{x_c}$ be a minimizer of (5) with $\mathbf{b}$ replaced by $\mathbf{c}$. Then*

$$\|A\mathbf{x_b} - A\mathbf{x_c}\|_2 \leq \|\mathbf{b} - \mathbf{c}\|_2 .$$

In particular, the above property holds for a minimizer $\mathbf{x_b}$ of (2) and a minimizer $\mathbf{x_c}$ of the minimization problem (2) with $\mathbf{b}$ replaced by $\mathbf{c}$. In addition, if A is completely full ranked and $\|\mathbf{x_b}\|_0 \leq m/2$ as well $\|\mathbf{x_c}\|_0 \leq m/2$, then there exists a positive constant C d

$$\|\mathbf{x_c} - \mathbf{x_b}\| \leq C\|\mathbf{b} - \mathbf{c}\|_2.$$

Proof. We mainly use the following inequality:

$$\left(\frac{x}{\sqrt{\epsilon + x^2}} - \frac{y}{\sqrt{\epsilon + y^2}}\right)(x - y) \geq 0. \tag{17}$$

which can be verified easily. Fix an $\epsilon > 0$. Let $\mathbf{x_b}$ be the minimizer satisfying the equation (7) with $q = 1$ associated with $\mathbf{b}$. Similarly, let $\mathbf{x_c}$ satisfy (7) associated with $\mathbf{c}$ replacing $\mathbf{b}$. For convenience, let us write $\mathbf{x} := \mathbf{x_b}$ and $\mathbf{y} := \mathbf{x_c}$. Multiplying $\mathbf{x} - \mathbf{y}$ to both sides of (7), we have

$$\sum_{j=1}^{N} \frac{x_j(x_j - y_j)}{\sqrt{\epsilon + (x_j)^2}} + \frac{1}{\lambda}(A(\mathbf{x} - \mathbf{y}))^T(A\mathbf{x} - \mathbf{b}) = 0. \tag{18}$$

Similarly, we have

$$\sum_{j=1}^{N} \frac{y_j(x_j - y_j)}{\sqrt{\epsilon + (x_j)^2}} + \frac{1}{\lambda}(A(\mathbf{x} - \mathbf{y}))^T(A\mathbf{y} - \mathbf{c}) = 0. \tag{19}$$

The subtractions of the above equations yields

$$\frac{1}{\lambda}(A(\mathbf{x} - \mathbf{y}))^T(A\mathbf{x} - A\mathbf{y} - \mathbf{b} + \mathbf{c}) = -\sum_{j=1}^{N}\left(\frac{x_j}{\sqrt{\epsilon + (x_j)^2}} - \frac{y_j}{\sqrt{\epsilon + (y_j)^2}}\right)(x_j - y_j)$$

which is less than or equal to zero by (17). It follows that

$$(A\mathbf{x} - A\mathbf{y}))^T(A\mathbf{x} - A\mathbf{y} - \mathbf{b} + \mathbf{c}) \leq 0$$

or

$$\|A\mathbf{x}-A\mathbf{y}\|_2^2 \le (A\mathbf{x}-A\mathbf{c})^T(\mathbf{b}-\mathbf{c}).$$

An application of Cauchy-Schwarz inequality and let $\epsilon \to 0_+$ together with Theorem 2.1 yield the proof of this Lemma. ■

A simple corollary is the following

Corollary 3.1 *Let $N(A)$ be the null space of A. For any two minimizers $\mathbf{x}^*$ and $\mathbf{x}$ of the unconstrained ℓ_1 minimization (2), $\mathbf{x}-\mathbf{x}^* \in Null(A)$. Similar for any two minimizers of (3).*

Next we prove the following

Proposition 3.2 (Extremal Value) *Suppose that $q=1$. Let $\mathbf{x}_{\mathbf{b},\epsilon}$ be the minimizers for input data $\mathbf{b}$ in problem (5) with $q=1$. Then*

$$\min_{\mathbf{x}\in\mathbf{R}^N} L_1(\epsilon,\mathbf{x}) = \frac{1}{2\lambda}\left(\|\mathbf{b}\|_2^2 - \|A\mathbf{x}_{\mathbf{b},\epsilon}\|_2^2\right) + \sum_{j=1}^N \frac{\epsilon}{\sqrt{\epsilon+|(\mathbf{x}_{\mathbf{b},\epsilon})_j|_2^2}}. \tag{20}$$

Consequently, for a minimizer $\mathbf{x}_\mathbf{b}$ of Problem (4), we have

$$\min_{\mathbf{x}\in\mathbf{R}^N}\{\|\mathbf{x}\|_1 + \frac{1}{2\lambda}\|A\mathbf{x}-\mathbf{b}\|_2^2\} = \frac{1}{2\lambda}\left(\|\mathbf{b}\|_2^2 - \|A\mathbf{x}_\mathbf{b}\|_2^2\right). \tag{21}$$

Proof. For convenience, let us write $\mathbf{x} := \mathbf{x}_{\mathbf{b},\epsilon}$. Multiplying $\mathbf{x}$ to the both sides of equation (7), we have

$$\sum_{j=1}^N \frac{x_j^2}{\sqrt{\epsilon+|x_j|^2}} + \frac{1}{\lambda}(A\mathbf{x})^T(A\mathbf{x}-\mathbf{b}) = 0.$$

The first term can be rewritten to

$$\sum_{j=1}^N \frac{x_j^2}{\sqrt{\epsilon+|x_j|^2}} = \|\mathbf{x}\|_\epsilon - \sum_{j=1}^N \frac{\epsilon}{\sqrt{\epsilon+|x_j|^2}}.$$

The second term can be rewritten to

$$\frac{1}{\lambda}(A\mathbf{x})^T(A\mathbf{x}-\mathbf{b}) = \frac{1}{2\lambda}\left(\|A\mathbf{x}-\mathbf{b}\|_2^2 + (A\mathbf{x}+\mathbf{b})^T(A\mathbf{x}-\mathbf{b})\right).$$

Combining these two equations together, we have

$$L_1(\epsilon,\mathbf{x}) = \sum_{j=1}^N \frac{\epsilon}{\sqrt{\epsilon+|x_j|^2}} - \frac{1}{2\lambda}\left((A\mathbf{x}+\mathbf{b})^T(A\mathbf{x}-\mathbf{b})\right).$$

which yields (20) in this proposition. By letting ϵ go to zero in (20), we get (21). ■

It follows from (21) that for a minimizer $\mathbf{x_b}$ of Problem (4), we have the following Pythagorean inequality:

$$\|A\mathbf{x_b} - \mathbf{b}\|_2^2 + \|\mathbf{Ax_b}\|_2^2 \leq \|\mathbf{b}\|_2^2.$$

Next we generalize the proof of Lemma 2.1 to have

Proposition 3.3 *Suppose that $q = 1$. Let $\mathbf{x}$ be the minimizer for problem (4) with $q = 1$. Then for any $\mathbf{y}$*

$$\|A\mathbf{y} - A\mathbf{x}\|^2 \leq 2\lambda(L_1(\mathbf{y}) - L_1(\mathbf{x})). \tag{22}$$

Proof. We follow the idea of proving Lemma 2.1. First we consider the problem (5) with $q = 1$. It is easy to verify

$$\sqrt{\epsilon + y^2} - \sqrt{\epsilon + x^2} \geq \frac{(y - x)x}{\sqrt{\epsilon + x^2}} \tag{23}$$

Then

$$\begin{aligned} L_1(\epsilon, \mathbf{y}) - L_1(\epsilon, \mathbf{x}) &= \sum_{j=1}^{N} \sqrt{\epsilon + |y_j|^2} - \sum_{j=1}^{N} \sqrt{\epsilon + |x_j|^2} \\ &\quad + \frac{1}{2\lambda}\left(\|A\mathbf{y} - \mathbf{b}\|^2 - \|A\mathbf{x} - \mathbf{b}\|^2\right) \\ &= \sum_{j=1}^{N} \sqrt{\epsilon + |y_j|^2} - \sqrt{\epsilon + |x_j|^2} + \frac{1}{2\lambda}\|A\mathbf{y} - A\mathbf{x}\|^2 \\ &\quad + \frac{1}{\lambda}(A\mathbf{x} - \mathbf{b})^T(A\mathbf{y} - A\mathbf{x}) \end{aligned}$$

The last term can be reduced to

$$-\sum_{j=1}^{N} \frac{(y_i - x_i)x_i}{\sqrt{\epsilon + x^2}}$$

Then

$$\begin{aligned} & L_1(\epsilon, \mathbf{y}) - L_1(\epsilon, \mathbf{x}) \\ = & \sum_{j=1}^{N} \left(\sqrt{\epsilon + |y_j|^2} - \sqrt{\epsilon + |x_j|^2} - \frac{x_j(y_j - x_j)}{\sqrt{\epsilon + |x_j|^2}} \right) + \frac{1}{2\lambda}\|A\mathbf{y} - A\mathbf{x}\|^2 \\ \geq & \frac{1}{2\lambda}\|A\mathbf{y} - A\mathbf{x}\|^2 \end{aligned}$$

by using (23). Next we let $\epsilon \to 0$ to conclude the result in this proposition. ■

4 Computation of Sparse Solutions of Under-determined Linear Systems

In this section we compare with our unconstrained ℓ_q minimization described in Section 2 with six other existing algorithms, namely the orthogonal greedy algorithm (OGA)(cf. [27]), the ℓ_1 greedy algorithm (cf. [22]), the standard ℓ_1 (cf. [5]) and the reweighted ℓ_1 algorithms (cf. [8]) which can be obtained on-line from the Candés webpage, the regularized orthogonal matching pursuit (ROMP) (cf. [26]), in addition to the ℓ_q algorithm developed in [18]. In our minimization, we choose $\lambda = 10^{-6}$ and run our iterative algorithm explained in Section 2 for many $\epsilon > 0$ and $q > 0$. We have to admit that it takes very long time to get the sparse solution. However, if the sparsity is known, our algorithm finds the sparse solution quickly since an intermediate iterative solution $\mathbf{x}^{\epsilon,q}$ may have already been a solution. We use this additional assumption for our algorithm to perform a numerical comparison. In this comparison, we used 500 random pairs $(A, \mathbf{x})$ with matrices A of size $m \times N$ with $N = 250$ and $m = 50$ as in [13] and vectors $\mathbf{x} \in \mathbf{R}^N$ for sparsity $\|\mathbf{x}\|_0 = s$ for $s = 1, 2, \cdots, 30$, where matrices A are Gaussian random matrices whose entries are iid of $N(0, \sigma)$ with $\sigma = 1/50$. Thus $\mathbf{b} = A\mathbf{x}$ are known given vectors. We use all 7 methods to solve $A\mathbf{x} = \mathbf{b}$. For each s and each pair, we run each of the 7 algorithms to obtain a vector $\widetilde{\mathbf{x}}$, and we considered the recovery a success if $\|\mathbf{x} - \widetilde{\mathbf{x}}\|_\infty < 10^{-5}$. We plot the percentages of successfully finding the sparse solutions for each method in Fig. 1. Here nL_q denotes our unconstrained ℓ_q minimization method. From the graph, we can see that our method is very close to the best performer: the ℓ_1 greedy algorithm. The interested reader may use our graph to compare with other methods which are not listed in this paper, e.g., the one in [13].

We have to point out that our algorithm is slower than the ℓ_1 greedy algorithm since we have to do many iterative solutions for various ϵ and q. Nevertheless, our algorithm does offer some advantage when the matrices A are an uniform random matrix. That is, we also use these above 7 algorithms to test the recovery of the sparse solutions of under-determined linear systems for uniform random matrices A and vector $\mathbf{x}$. All the procedures are exactly the same as above except for replacing Gaussian random matrices by uniform random matrices. Our method is clearly better as shown in Fig. 2.

References

[1] R. Baraniuk, Compressive sensing, IEEE Signal Processing Magazine, (2007), vol. 24, 118–121.

[2] A. M. Bruckstein, D. L. Donoho, M. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images, SIAM Review, (2009) vol. 51, No . 1, pp. 34–81.

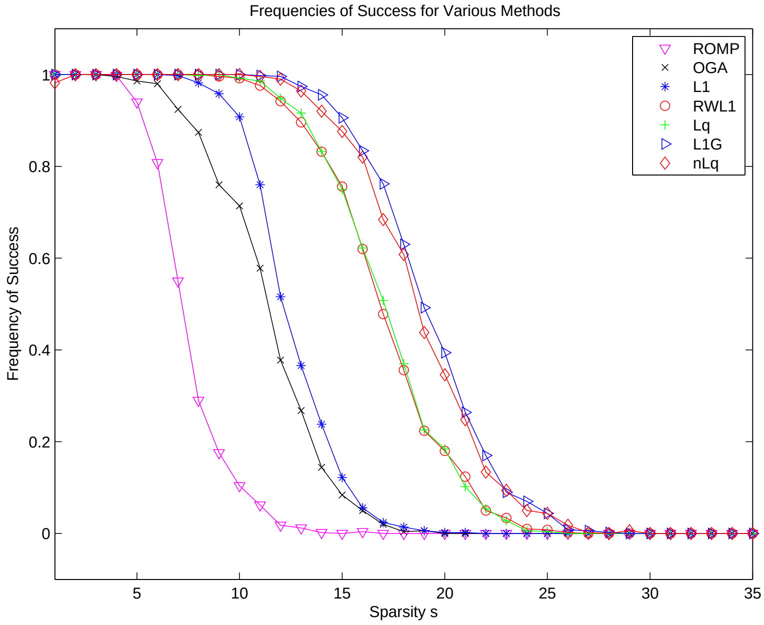

Figure 1: Comparison of the 7 algorithms for sparse solution of underdetermined linear system associated with Gaussian random matrices

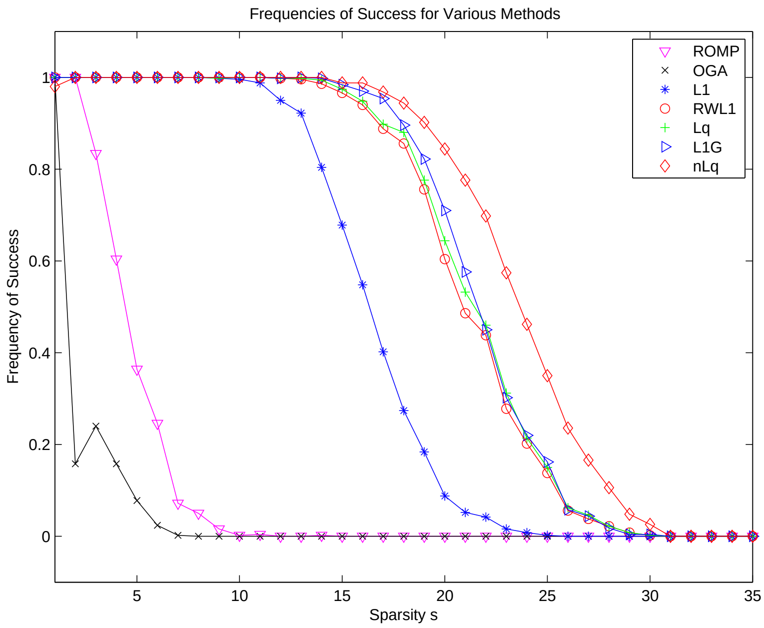

Figure 2: Comparison of the 7 algorithms for sparse solution of under-determined linear system associated with uniform random matrices

[3] E. J. Candés, Compressive sampling, International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 1433–1452.

[4] Candés, E. J., The restricted isometry property and its implications for compressed sensing, Comptes Rendus de l'Académie des Sciences, Série I, 346 (2008), 589–592.

[5] Candés, E. J., J. K. Romberg, and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Comm. Pure Appl. Math. 59 (2006), 1207–1223.

[6] Candés, E. J. and T. Tao, Decoding by linear programming, IEEE Trans. Inform. Theory 51 (2005), no. 12, 4203–4215.

[7] Candés, E. J. and T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies, IEEE Trans. Inform. Theory 52 (2006), no. 12, 5406–5425.

[8] Candés, E. J., M. Watkin, and S. Boyd, Enhancing Sparsity by Reweighted l_1 Minimization, To appear in J. Fourier Anal. Appl., 2009.

[9] Chartrand, R., Exact reconstruction of sparse signals via nonconvex minimization, IEEE Signal Process. Lett., 14 (2007), 707–710.

[10] R. Chartrand and V. Staneva, Restricted isometry properties and nonconvex compressive sensing, Inverse Problem, 24(2008), 1–14.

[11] I. Daubechies, M. Defreise, and C. De Mol, An iterative thresholding algorithm for linear inverse problem with a sparsity constraint, Comm. Pure and Applied Math., (2004) vol LVII, 1413–1457.

[12] I. Daubechies, M. Fornasier, I. Loris, Accelerated projected gradient method for linear inverse problems with sparsity constraints, J. Fourier Anal. Appl. 14 (2008), no. 5-6, 764–792.

[13] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntuk, Iteratively re-weighted least squares minimization for sparse recovery, manuscript, 2008.

[14] D. L. Donoho, Compressed sensing, IEEE Trans. Inf. Theory 52 (2006), no. 4, 1289–1306.

[15] Donoho, D. L. and M. Elad, Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization, Proc. Natl. Acad. Sci. USA 100 (2003), no. 5, 2197–2202.

[16] Donoho, D. L., M. Elad, and V. N. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise. IEEE Trans. Inform. Theory, 52 (2006), 6–18.

[17] D. L. Donoho and B. F. Logan, Signal recovery and the large sieve, SIAM J. Appl. Math. 52 (1992), no. 2, 557–591.

[18] S. Foucart and M. J. Lai, Sparsest Solutions of Under-determined Linear Systems via ℓ_q minimization for $0 < q \leq 1$, Applied Comput. Harmonic Analysis, 26(2009) 395–407.

[19] S. Foucart and M. J. Lai, Sparse Recovery with Pre-Gaussian Random Matrices, submitted, 2009.

[20] E. de Giorgi and T. Franzoni, Su un tipo di convergenza variazionale; Atti Accad. Naz. Lincei Rend. Cl. Sci. Mat. (8) 58 (1975), 842–850.

[21] R. Gribnoval and M. Nielsen, Sparse decompositions in unions of bases, IEEE Trans. Info. Theory, 49(2003), 3320–3325.

[22] I. Kozlov and A. Petukhov, Sparse Solutions of Underdetermined Linear Systems, Chapter in Handbook of Geomathematics, edited by W. Freeden, Springer, 2010.

[23] M. J. Lai, On sparse solution of underdetermined linear systems, accepted by Journal of Concrete and Applicable Mathematics, 2009.

[24] G. dal Maso, An introduction to Γ-convergence, Birkhauser, Basel 1993.

[25] Natarajan, B. K., Sparse approximate solutions to linear systems, SIAM J. Comput., vol. 24, pp. 227–234, 1995.

[26] D. Needell, R. Vershynin, Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit, Found. Comput. Math., in Press, 2009.

[27] Petukhov, A., Fast implementation of orthogonal greedy algorithm for tight wavelet frames, Signal Processing, 86 (2006), 471–479.

[28] Temlyakov, V. N., Nonlinear methods of approximation, Foundations of Comp. Math., 3 (2003), 33–107.

[29] Tropp, J. A., Greed is good: algorithmic results for sparse approximation, IEEE Trans. Inf. Theory, 50 (2004), 2231–2242.

[30] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, Bregman iterative algorithms for ℓ_1-minimization with applications to compressed sensing, SIAM J. Imaging Sciences, 1(2008), 143–168.