Mathematical Programming manuscript No.
(will be inserted by the editor)

A Note on the Complexity of L_p Minimization

Dongdong Ge · Xiaoye Jiang · Yinyu Ye

Abstract We show that the L_p ($0 \le p < 1$) minimization problem arising from sparse solution construction and compressed sensing is both hard and easy. More precisely, for any fixed $0 < p < 1$, we prove that finding the global minimal value of the problem is strongly NP-Hard; but approximating a local minimizer of the problem is polynomial-time doable. We also develop an interior-point algorithm with a provable complexity bound and demonstrate preliminary computational results of effectiveness of the algorithm.

Keywords nonconvex programming · global optimization · interior-point method · sparse solution reconstruction
MSC2010 Classification Code: 90C26, 90C51

1 Introduction

In this note, we consider the following optimization problem:

$$\begin{array}{lll} \text{(P)} & \text{Minimize} & p(x) := \sum_{1 \le j \le n} x_j^p \\ & \text{Subject to} & x \in \mathcal{F} := \{x : Ax = b, x \ge 0\}, \end{array} \tag{1}$$

and

$$\begin{array}{ll} \text{Minimize} & \sum_{1 \le j \le n} |x_j|^p \\ \text{Subject to} & x \in \mathcal{F}' := \{x : Ax = b\}, \end{array} \tag{2}$$

where data $A \in R^{m \times n}, b \in R^m$, and $0 < p < 1$.

Sparse signal or solution reconstruction by solving optimization problem (1) or (2), especially for the cases of $0 \le p \le 1$, recently received great attentions, e.g., see [4, 7–9, 11, 13]. In signal reconstruction, one typically has linear measurements $b = Ax$ where x is a sparse signal, and the sparse signal would be recovered by solving inverse problem (1) or (2) with $p = 0$, that is, to find the sparsest or smallest support cardinality solution of a linear system (here $|x|^0 = 1$ if $x \neq 0$ and 0 otherwise). From the computational complexity point of view, when $p = 0$, problem (1) or (2) is shown to be NP-hard [15] to solve; when $p = 1$, both problems are linear programs, hence they are polynomial-time solvable. For more general convex programs involving L_p norm($p > 1$) in the objective(or constraints), readers are referred to [16].

In [5, 10], it was shown that if certain *restricted isometry property* holds for A, then the solutions of (2) for $p = 0$ and $p = 1$ are identical. Hence, problem (2) with $p = 0$ can be relaxed to problem (2) with $p = 1$. However, *restricted isometry property* may be too strong for practical basis design matrices A to

Dongdong Ge
Antai School of Economics and Management, Shanghai Jiao Tong University, Shanghai, China 200052.
E-mail: ddge@sjtu.edu.cn

Xiaoye Jiang
Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94305
E-mail: xiaoye@stanford.edu

Yinyu Ye
Department of Management Science and Engineering, Stanford University, Stanford, CA 94305
E-mail: yinyu-ye@stanford.edu

hold. Instead, one may consider sparse recovery by solving relaxation problem (1) or (2) for a fixed p, $0 < p < 1$. Recently, this approach has attracted a lot of research efforts in variable selection and sparse reconstruction, e.g., [11]. It exhibits desired threshold bounds on any non-zero entry of a computed solution [9], and computational experiences show that by replacing $p = 1$ with $p < 1$, reconstruction can be done equally fast with many fewer measurements while being more robust to noise and signal non-sparsity [8].

In this note we present some interesting properties of the L_p $(0 \le p < 1)$ minimization problem. The L_p $(0 \le p < 1)$ minimization problem (1) or (2) is strongly NP-Hard. However, any basic feasible solution of (1) or (2) is a local minimizer. Moreover, a feasible point satisfying the first order and second order necessary conditions is always a local minimizer. This motivates us to design interior-point algorithms to approximate a local minimum point satisfying the Karush-Kuhn-Tucker(KKT) conditions.

1.1 Notations and Preliminaries

For the simplicity of our analysis, throughout this paper we assume the feasible set $\mathcal{F}$ is bounded and A is full-ranked.

A feasible point x is called a *local minimum point* or *local minimizer* of problem (1) if there exists $\epsilon > 0$, such that $p(x') \ge p(x)$ for any $x' \in B(x, \epsilon) \bigcap \mathcal{F}$ where $B(x, \epsilon) = \{x' : \| x' - x \|_2 \le \epsilon\}$.

Let $X := \text{Diag}(x_i)_{1 \le i \le n}$ and $X^p := \text{Diag}(x_i^p)_{1 \le i \le n}$. At a local minimum point x^* of problem (1), the *first order necessary condition* or *KKT condition* [2] holds: there exists unique Lagrange multiplier vector $y^* \in R^m$, such that

$$p(X^*)^p - X^* A^T y = 0. \tag{3}$$

If $p(x)$ is differentiable at x^*, i.e., $x^* > 0$, we have the *second order necessary condition*

$$\lambda^T \nabla^2_{xx} p(x^*) \lambda \ge 0, \tag{4}$$

for all $\lambda \in R^n$ such that $A\lambda = 0$.

In the general case a local minimizer x^* may lie on the boundary so that $p(x)$ is not differentiable at x^*. Let S be the support set of x^*, so $|S| = k \le n$. Let $z^* \in R^k$ be a vector of the nonzero coordinates of x^* and $A_S \in R^{m \times k}$ be the submatrix of A exactly consisting of the columns in A corresponding to $|S|$. It is not difficult to verify that z^* is a local minimizer of problem $\min p_S(z) := \sum_{i \in S} z_i^p,$ s.t. $A_S z = b,\ z \ge 0$. The second order necessary condition holds:

$$\lambda^T \nabla^2_{zz} p_S(z^*) \lambda \ge 0, \tag{5}$$

for all $\lambda \in R^k$ such that $A_S \lambda = 0$. We also call (5) the *second order necessary condition* for general x^*.

There is also a dual problem associated with (P):

$$\begin{array}{lll} \text{(D)} & \text{Maximize} & d(x, y) := (1 - p)p(x) + b^T y \\ & \text{Subject to} & (x, y, s) \in \mathcal{F}_d := \{(x, y, s) : pX^p - XA^T y - Xs = 0, x, s \ge 0\}, \end{array} \tag{6}$$

The fact that $\mathcal{F}$ is bounded and nonempty implies $p(x)$ has a minimum value(Denote by $\underline{z}$) and a maximum value(Denote by $\bar{z}$). An ϵ-minimal solution or ϵ-minimizer is defined as a feasible solution x such that

$$\frac{p(x) - \underline{z}}{\bar{z} - \underline{z}} \le \epsilon. \tag{7}$$

Vavasis [20] demonstrated the importance to have the term $\bar{z} - \underline{z}$ in the criterion for continuous optimization. Similarly Ye [18] defined an ϵ-KKT (or ϵ-stationary) point as an (x, y, s) such that $x \in \mathcal{F}, (x, y, s) \in \mathcal{F}_d$, and

$$\frac{x^T s}{\bar{z} - \underline{z}} \le \epsilon. \tag{8}$$

Ye also developed an interior point algorithm converging to an ϵ-KKT point for the nonconvex quadratic progrm. In this note similarly we present an interior point algorithm to find an ϵ-KKT point for the L_p minimization problem. This potential reduction algorithm is shown to be a fully polynomial-time approximation scheme(FPTAS) and it converges to an ϵ-KKT point in no more than $O(\frac{n}{\epsilon} \log \frac{1}{\epsilon})$ iterations.

The note is organized as follows. In section 2 We show that the L_p $(0 \le p < 1)$ minimization problem (1) or (2) is strongly NP-Hard. In section 3 we prove that the set of all basic feasible solutions of (1) or (2) is identical with the set of all local minimizers. In section 4 we present our FPTAS algorithm to find an ϵ-KKT point of problem (1). Numerical experiments are conducted to test its efficiency in section 5.

2 The Hardness

Theorem 1 *The L_p $(0 \le p < 1)$ minimization problem (1) (or (2)) is strongly NP-hard.*

Proof We present a polynomial time reduction from the well known strongly NP-complete 3-partition problem [12]. The partition problem can be described as follows: given a multiset S of $n = 3m$ integers $\{a_1, a_2, \dots, a_n\}$. The sum of S is equal to mB and each integer in S is strictly between $B/4$ and $B/2$. Can S be partitioned into m subsets, such that the sum of the numbers in each subset is equal to B, which implies each subset has exactly three elements?

We describe a reduction from an instance of the 3-partition problem to an instance of the L_p $(0 \le p < 1)$ minimization problem (1) that has the optimal value n if and only if the former has an equitable 3-partition. Given an instance of the partition problem, let vector $a = (a_1, a_2, \dots, a_n) \in R^n$ and $n = 3m$. Let the sum of a be mB and each $a_i \in (B/4, B/2)$. Consider the following minimization problem in form (1):

$$\begin{array}{ll} \text{Minimize} & P(x) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}^p \\ \text{Subject to} & \sum_{j=1}^{m} x_{ij} = 1, \quad\quad i = 1, 2, \cdots, n, \\ & \sum_{i=1}^{n} a_i x_{ij} = B, \quad j = 1, 2, \cdots, m, \\ & x_{ij} \ge 0, \quad\quad\quad\quad i = 1, 2, \cdots, n; j = 1, 2, \cdots, m, \end{array} \tag{9}$$

From the strict concavity of the objective function, and $x_{ij} \in [0, 1]$,

$$\sum_{j=1}^{m} x_{ij}^p \ge \sum_{j=1}^{m} x_{ij}(= 1),\ i = 1, 2, \cdots, n.$$

The equality holds if and only if $x_{ij_0} = 1$ for some j_0 and other $x_{ij} = 0$, $\forall j \ne j_0$. Thus, $P(x, y, z) \ge n$ for any feasible solution of (9).

If there is a feasible (optimal) solution x such that $P(x) = n$, it must be true that $\sum_{j=1}^{m} x_{ij}^p = 1 = \sum_{j=1}^{m} x_{ij}$ for all i so that for any i, $x_{ij_0} = 1$ for some j_0 and other $x_{ij} = 0$, $\forall j \ne j_0$. This generates an equitable 3-partition of the entries of a. On the other hand, if the entries of a have an equitable 3-partition, then (9) must have a binary solution x such that $P(x) = n$. Thus we prove the strong NP-hardness of problem (1).

For the same instance of the 3-partition problem, we consider the following minimization problem in form (2):

$$\begin{array}{ll} \text{Minimize} & \sum_{i=1}^{n}\sum_{j=1}^{m} |x_{ij}|^p \\ \text{Subject to} & \sum_{j=1}^{m} x_{ij} = 1, \quad i = 1, 2, \cdots, n, \\ & \sum_{i=1}^{n} a_i x_{ij} = B,\ j = 1, 2, \cdots, m, \end{array} \tag{10}$$

Note that this problem has no non-negativity constraints on variables x. However, for any feasible solution x of the problem, we still have

$$\sum_{j=1}^{m} |x_{ij}|^p \ge \sum_{j=1}^{m} x_{ij}(= 1),\ i = 1, 2, \cdots, n.$$

This is because the minimal value of $\sum_{j=1}^{m} |x_{ij}|^p$ is 1 if $\sum_{j=1}^{m} x_{ij} = 1$, and the quality holds if and only if for any i, $x_{ij_0} = 1$ for some j_0 and other $x_{ij} = 0$, $\forall j \ne j_0$.

Therefore, similarly we can prove that this instance of the partition problem has an equitable 3-partition if and only if the objective value of (10) is n. This leads to the strong NP-hardness of problem (2).

3 The Easiness

The above discussion reveals that finding a global minimizer for the L_p norm minimization problem is NP-hard as long as $p < 1$. Thus, relaxing $p = 0$ to some $p < 1$ gains no advantage in terms of the (worst-case) computational complexity. We now turn our attention to local minimizers. Note that, for many optimization problems, finding a local minimizer, or checking if a solution is a local minimizer, remains NP-hard. What about local minimizers of problems (1) and (2)? The answer is that they are easy to specify.

Theorem 2 *The set of all basic feasible solutions of (1) is exactly the set of all local minimizers.*

Proof If x is a basic feasible solution(or extreme point), without loss of generality, assume $x = (x_1, \cdots, x_m, 0, \cdots, 0)$ where $x_i > 0, i = 1, 2, \cdots, m$. Consider an arbitrary nonzero feasible direction $d = (d_1, d_2, \cdots, d_n)$ at x. There exists an appropriate $\epsilon_0 > 0$, such that $x + \epsilon d$ is feasible for any $0 < \epsilon < \epsilon_0$. By the property of a basic feasible solution, there must exist some i, $m + 1 \leq i \leq n$, such that $d_i > 0$. Assume $i = m + 1$.

Thus

$$p(x + \epsilon d) - p(x) \geq (\sum_{i=1}^{m} (x_i + \epsilon d_i)^p - x_i^p) + (\epsilon d_{m+1})^p.$$

Define the index set $I^- = \{i : d_i < 0\}$. Without loss of generality, assume I^- is nonempty. We can also choose sufficiently small ϵ such that $x_i + \epsilon d_i \geq \frac{x_i}{2}$ for any $i \in I^-$.

Then

$$p(x + \epsilon d) - p(x) \geq (\sum_{i \in I^-} ((x_i + \epsilon d_i))^p - x_i^p) + (\epsilon d_{m+1})^p \geq \sum_{i \in I^-} (\epsilon d_i) p (\frac{x_i}{2})^{p-1} + (\epsilon d_{m+1})^p \geq 0.$$

The last inequality holds if

$$\epsilon \leq d_{m+1}^{\frac{p}{1-p}} (p \sum_{i \in I^-} (-d_i) (\frac{x_i}{2})^{p-1})^{\frac{1}{p-1}}. \tag{11}$$

This shows that any feasible direction at x is an increasing direction. Let V be the vertex set of $\mathcal{F}$ and choose d_i in the same direction as the line connecting x to vertex v_i. Since d_i is feasible, we can choose the corresponding ϵ_i as in (11) such that $p(x_i + \epsilon d_i) \geq p(x)$ for any $0 \leq \epsilon \leq \epsilon_i$. Denote by $\text{Conv}\{x + \epsilon_i d_i : i = 1, 2, \cdots, |V|\}$ the convex hull spanned by $x + \epsilon_i d_i, i = 1, 2, \cdots, |V|$.

For any $x' \in \mathcal{F} \bigcap \text{Conv}\{x + \epsilon_i d_i : i = 1, 2, \cdots, |V|\}$, we have $p(x') \geq p(x)$ by the strict concavity of $p(x)$. Thus we can always choose a sufficiently small $\epsilon > 0$, such that $B(x, \epsilon) \bigcap \mathcal{F} \subset \mathcal{F} \bigcap \text{Conv}\{x + \epsilon_i d_i : i = 1, 2, \cdots, |V|\}$. By the definition x is a local minimizer.

On the other hand, if x is a local minimizer but not a basic feasible solution, there must exist a feasible direction $d \neq 0$ such that both $x + \epsilon d$ and $x - \epsilon d$ are feasible for sufficiently small $\epsilon > 0$. The strict concavity of $p(x)$ implies that either d or $-d$ will be a descent direction. Thus, x cannot be a local minimizer.

Similarly, we can prove

Theorem 3 *The set of all basic solutions of (2) is exactly the set of all of its local minimizers.*

Furthermore, one can observe the following property of a local minimizer.

Theorem 4 *If the first order necessary condition (3) and the second order necessary condition (5) hold at x and $x \in \mathcal{F}$, then x is a local minimizer of problem (1).*

Proof If x is in the interior of the feasible region $\mathcal{F}$ and satisfies the first order condition (3), The strict concavity of $p(x)$ implies x must be the only maximum point.

If x lies on the boundary, define its support set S and positive vector z as in the second order necessary condition (5). We check its Hessian. $\nabla^2_{zz}(z)$ cannot be positive semidefinite unless $\{\lambda : A_S \lambda = 0\} = \{0\}$. Thus the second order necessary condition (5) holds only when x is a basic feasible solution. By Theorem 2, x is a local minimizer.

4 Interior-Point Algorithm

These local minimizer results show that we can approximate a local minimizer problem (1) by searching for an ϵ-KKT point of it. Considering the differentiability of $p(x)$ in the interior feasible region, one would naturally consider an interior point algorithm approach. It starts from an interior point, follows a interior feasible path and finally converges to either an ϵ-global point or an ϵ-KKT point. At each step, it chooses a new interior point which produces the maximal potential reduction to a potential function by an affine-scaling operation. See [14] or [19] for an intensive introduction on interior-point algorithms.

Our algorithm starts from an interior-point feasible solution such as the analytic center x^0 of the feasible polytope. Similar to potential reduction algorithms for linear programming, one could consider the potential function

$$\phi(x) = \rho \log(\sum_{j=1}^{n} x_j^p - \underline{z}) - \sum_{j=1}^{n} \log x_j = \rho \log(p(x) - \underline{z}) - \sum_{j=1}^{n} \log x_j, \tag{12}$$

where $\underline{z}$ is a lower bound on the global minimal objective value of (1) and parameter $\rho > n$. For simplicity, we set $\underline{z} = 0$ throughout the discussion. Note that

$$\frac{\sum_{j=1}^{n} x_j^p}{n} \geq \left(\prod_{j=1}^{n} x_j^p\right)^{1/n}.$$

So

$$\frac{n}{p}\log(p(x)) - \sum_{j=1}^{n} \log x_j \geq \frac{n}{p}\log n.$$

Thus, if $\phi(x) \leq (\rho - n/p)\log(\epsilon) + (n/p)\log n$, we must have $p(x) \leq \epsilon$, which implies that x must be an ϵ-global minimizer.

Starting from the analytic center x^0, an interior algorithm looks for the best possible potential reduction at each iteration. In a manner similar to the algorithm discussed in [18] for non-convex quadratic minimization, one can consider the potential reduction $\phi(x^+) - \phi(x)$ by one-iteration update from x to x^+.

Note that

$$\phi(x^+) - \phi(x) = \rho(\log(p(x^+)) - \log(p(x))) + (-\sum_{j=1}^{n}\log(x_j^+) + \sum_{j=1}^{n}\log(x_j)).$$

Let d_x, $Ad_x = 0$, be a vector such that $x^+ = x + d_x > 0$. Then, from the concavity of $\log(p(x))$, we have

$$\log(p(x^+)) - \log(p(x)) \leq \frac{1}{p(x)}\nabla p(x)^T d_x.$$

On the other hand, by restricting $\|X^{-1}d_x\| \leq \beta < 1$, where $X = \mathrm{Diag}(x)$, we have(See [3] for details)

$$-\sum_{j=1}^{n}\log(x_j^+) + \sum_{j=1}^{n}\log(x_j) \leq -e^T X^{-1} d_x + \frac{\beta^2}{2(1-\beta)}.$$

From the analysis above, if $\|X^{-1}d_x\| \leq \beta < 1$, then $x^+ = x + d_x > 0$, and

$$\phi(x^+) - \phi(x) \leq (\frac{\rho}{p(x)}\nabla p(x)^T X - e^T)X^{-1}d_x + \frac{\beta^2}{2(1-\beta)}.$$

Let $d' = X^{-1}d_x$. Then, to achieve a potential reduction, one can minimize an affine-scaled linear function subject to a ball constraint as it is done for linear programming(See Chapter 1 and 4 in [19] for more details):

$$\begin{array}{rl} Z(d') := \text{Minimize} & \left(\frac{\rho}{p(x)}\nabla p(x)^T X - e^T\right) d' \\ \text{Subject to} & AXd' = 0 \\ & \|d'\|^2 \leq \beta^2. \end{array} \tag{13}$$

This is simply a linear projection problem. The minimal value is $Z(d') = -\beta \cdot \|g(x)\|$ and the optimal direction is given by $d' = \frac{\beta}{\|g(x)\|}g(x)$. Here

$$\begin{aligned} g(x) &= -(I - XA^T(AX^2A^T)^{-1}AX)(\frac{\rho}{p(x)}X\nabla p(x) - e) \\ &= e - \frac{\rho}{p(x)}X(\nabla p(x) - A^T y), \end{aligned}$$

where $y = (AX^2A^T)^{-1}AX(X^p - \frac{p(x)}{\rho}e)$.

If $\|g(x)\| \geq 1$, then the minimal objective value of the subproblem is less than $-\beta$ so that

$$\phi(x^+) - \phi(x) < -\beta + \frac{\beta^2}{2(1-\beta)}.$$

Thus, the potential value is reduced by a constant if setting $\beta = 1/2$. If this case would hold for $O((\rho - n/p)\log\frac{1}{\epsilon})$ iterations, we would have produced an ϵ-global minimizer of (1).

On the other hand, if $\|g(x)\| \leq 1$, from $g(x) = e - \frac{\rho}{p(x)}X(\nabla p(x) - A^T y)$, we must have

$$\frac{\rho}{p(x)}X(\nabla p(x) - A^T y) \geq 0, \quad \frac{\rho}{p(x)}X(\nabla p(x) - A^T y) \leq 2e, \forall j.$$

In other words,

$$\left(\nabla p(x) - A^T y\right)_j \geq 0, \quad \frac{x_j}{p(x)} \left(\nabla p(x) - A^T y\right)_j \leq \frac{2}{\rho}, \ \forall j.$$

The first condition indicates that the Lagrange multiplier y is feasible. For the second inequality, by choosing $\rho \geq \frac{2n}{\epsilon}$ we have $\frac{1}{p(x)} x^T (\nabla p(x) - A^T y) \leq \epsilon$. Therefore,

$$\frac{x^T s}{\bar{z} - \underline{z}} = \frac{x^T (\nabla p(x) - A^T y)}{\bar{z} - \underline{z}} \leq \frac{x^T (\nabla p(x) - A^T y)}{p(x)} \leq \epsilon,$$

which implies that x is an ϵ-stationary solution.

Concluding the analysis above, we have the following lemma.

Lemma 1 *The interior point algorithm returns an ϵ-stationary or ϵ-global point of (1) in no more than $O(\frac{n}{\epsilon} \log \frac{1}{\epsilon})$ iterations.*

A more careful computation will make the complementarity point satisfy the second order optimality condition; see [18]. By combining these observations with Theorem 4, we immediately conclude the convergence of interior-point algorithms.

Corollary 1 *Interior-point algorithms, starting from the analytic center of the polytope, generate a sequence of interior points converging to a local minimizer and compute an approximate local minimizer in FPTAS time for the L_p minimization problem.*

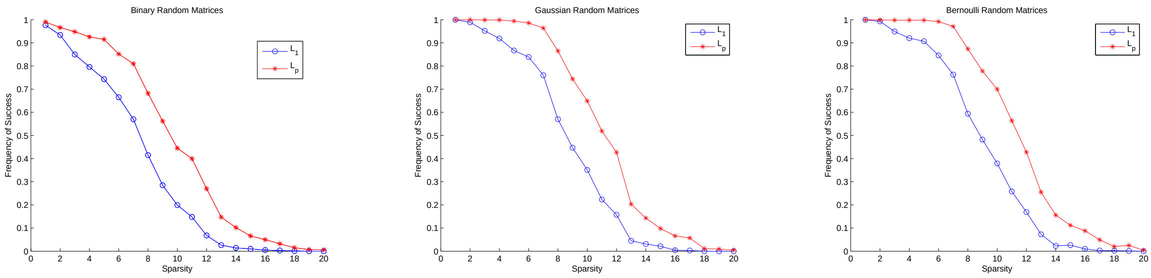
Thus, interior-point algorithms, including the simple affine-scaling algorithm that always goes along a descent direction, can be effective (with a FPTAS time) in tackling the L_p minimization problem as well.

5 Computational Experiment of the Interior-Point Algorithm

In this section, we compute a solution of (1) using the interior-point algorithm above and compare it with the solution of L_1 problem, i.e., the solution of (1) with $p = 1$, which is also computed by an interior-point algorithm for linear programming [17]. Our preliminary results reinforce our reasoning that interior-point algorithms likely avoid some local minimizers on the boundary and recover the sparse solution.

We construct 1000 random pairs (A, x) with matrices A size of 30 by 120 and vectors $x \in R^{120}$ for sparsity $\|x\|_0 = s$ with $s = 1, 2, \cdots, 20$. With basis matrix given, vector $b = Ax$ is known. We use several basis design matrices A to test our algorithms. In particular, let $A = [M, -M]$ and M is one of the following matrices: (1) Sparse binary random matrices where there are only a small number of ones in each column; (2) Gaussian random matrices whose entries are i.i.d. Gaussian random variables; (3) Bernoulli random matrices whose entries are i.i.d. Bernouli random variables. We note that Gaussian or Bernoulli random matrices satisfy the *restricted isometry property* [5,6], i.e., columns of the basis matrix are nearly orthogonal, while sparse binary random matrix only satisfies a weaker form of restricted isometry property [1]. Because two copies of the same random matrices are concatenated in A, column orthogonality will be hard to hold, which will lead to easy failure of sparse recovery by L_1 problem. We solve (1) with $p = 1/2$ by the interior-point algorithm developed above and compare the successful sparse recovery rate with the solutions of the L_1 problem. A solution is considered successfully recovering x if the l_2 distance between the solution and x is less than 10^{-3}.

The phase transitions of successful sparse recovery rates for L_p and L_1 problem for three cases of basis matrices are plotted in figure 5. We observe that the $L_{0.5}$ interior-point algorithm performs better in successfully identifying the sparse solution x than the L_1 algorithm does. Moreover, we note that when basis matrix are binary, we have comparatively lower rates of successful sparse recovery (when sparsity $s = 4$, for example, successful recovery rate for L_p solutions is about 95% for binary random matrices compared with almost 100% for Gaussian or Bernoulli random matrices; for L_1 solutions it is about 80% for binary random matrices compared with about 90% for Gaussian or Bernoulli random matrices). This may be supported by the fact that sparse binary random matrix has even worse column orthogonality than Gaussian or Bernoulli cases. Our simulation results also show that the interior-point algorithm for the L_p problem with $0 < p < 1$ runs as fast as the interior-point method for L_1 problem [17], which makes the interior-point method competitive for large scale sparse recovery problems.



(a) (b) (c)

Fig. 1 Successful sparse recovery rates of L_p and L_1 solutions, with matrix A constructed from (a) binary matrix; (b) Gaussian random matrix; (c) Bernoulli random matrix.

References

1. R. Berinde, A. C. Gilbert, P. Indyk, H. J. Karloff and M. J. Strauss, "Combining geometry and combinatorics: A unified approach to sparse signal recovery", preprint, 2008.
2. D. P. Bertsekas, "Nonlinear programming", Athena Scientific, (1999).
3. D. Bertsimas and J. Tsitsiklis, "Introduction to linear Optimization", Athena Scientific, (1997), 414.
4. J. M. Borwein and D. R. Luke, "Entropic Regularization of the l_0 function", in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering in Springer Optimization and Its Applications*, (2010).
5. E. J. Candès and T. Tao, "Decoding by linear programming", *IEEE Transaction of Information Theory*, 51(2005), 4203-4215.
6. E. J. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Transaction of Information Theory*, 52(2006), 5406-5425.
7. R. Chartrand "Exact Reconstruction of sparse signals via nonconvex minimization", *IEEE Signal Processing Letters*, 14-10(2009).
8. R. Chartrand "Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data", in *IEEE International Symposium on Biomedical Imaging (ISBI)*, (2009).
9. X. Chen, F. Xu and Y. Ye, "Lower Bound Theory of Nonzero Entries in Solutions of ℓ_2-ℓ_p Minimization", Technical Report, The Hong Kong Polytechnic University, (2009).
10. D. Donoho, "For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest Solution", Technical Report, Stanford University, (2004).
11. J. Fan and R. Li, "Variable selection via nonconcave penalized likelihood and its oracle properties", *Journal of American Statistical Society*, 96(2001), 1348-1360.
12. M. R. Garey and D. S. Johnson, "Computers and Intractability; A Guide to the Theory of NP-Completeness", W. H. Freeman & Company, (1979), 96-105.
13. N. Mourad and P. Reilly, "Minimizing nonconvex functions for sparse vector reconstruction", *IEEE Transactions on Signal Processing*, 58(2010), 3485–3496.
14. Y. Nesterov and A. Nemirovski, Interior Point Polynomial Methods in Convex Programming, SIAM, Philadelphia, PA, (1994).
15. B. K. Natarajan, "Sparse approximate solutions to linear systems", *SIAM Journal on Computing*, 24(1995), 227-234.
16. T. Terlaky, "On ℓ^p programming", *European Journal of Operational Research*, 22(1985), 70- 100.
17. Y. Ye, M. J. Todd and S. Mizuno, "An $O(\sqrt{n}L)$-Iteration Homogeneous and Self-dual Linear Programming Algorithm", Mathematics of Operations Research 19(1994) 53-67.
18. Y. Ye, "On the complexity of approximating a KKT point of quadratic programming", *Mathematical Programming*, 80(1998), 195-211.
19. Y. Ye, "Interior point algorithms: theory and analysis", John Wiley & Sons, Inc., New York, (1997).
20. S. A. Vavasis, "Polynomial time weak approximation algorithms for quadratic programming", in C.A. Floudas and P.M. Pardalos: *Complexity in Numerical Optimization*, World Scientific, New Jersey, (1993).