Solving Overdetermined Systems in ℓ_p Quasi-Norms*

Miguel Argáez†
Department of Department of Mathematical Sciences,
The University of Texas at El Paso, El Paso TX 79968,
USA
margaez@utep.edu

Abstract

The theory of compressed sensing has shown that sparse signals can be reconstructed exactly from remarkably few measurements by solving a nonconvex underdetermined ℓ_p-regularized quasi-norm problem via an iterative weighted least-squares problem. In this work, we consider the problem of recovering an input signal by solving a nonconvex overdetermined ℓ_p-regularized quasi-norm problem. In order to do this, we carry over a fixed-point algorithm, presented in [17], [10] and [1] from a nonconvex underdetermined to a nonconvex overdetermined ℓ_p quasi-norm problem. Then, we reformulate this procedure by a sequential quadratic program, and use two alternative algorithms for solving its associated linear systems so called augmented system: a direct method and a projected conjugate gradient. The sequential quadratic program takes into account the signals and its associated error. While the direct method scheme works with a sequence of approximations of the signals and its errors simultaneously, the projected conjugate gradient algorithm finds first an approximation error, and later, using this error, an approximate signal is obtained using just a least-squares problem. The numerical advantage of using a direct method for solving the augmented system is that it allows a sparser and cheaper factorization than the Cholesky factorization for solving the weighted normal equation for dense matrices. Besides, the projected conjugate gradient needs only one matrix factorization in all the optimization procedures which is appealing to solve large-scale problems. We implemented these strategies and compare their capabilities to recover signals. Specifically, our interest is to identify at what rate of corruption each formulation fails to recover the signal exactly for different values $0 \leq p < 1$, and compare with the convex problem when $p = 1$.

Keywords: ℓ_p quasi-norms, compressed sensing, fixed-point theory, quadratic programming, conjugate gradient
AMS subject classifications: 65F20, 65F22, 65F50

*Submitted: January 20, 2009; Revised: March 10, 2010; Accepted: April 1, 2010.
†This work was supported in part by Grant from HPTi No W911NF-07-2-0027 administrated by Stanford University and by Grant from NSF No HRD-0734825.

1 Introduction

State of the art research in compressed sensing is aimed at providing efficient algorithms capable to recover an input vector x^* from some corrupted measurements $b = Ax^* + e$, where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ are known data and $e \in \mathbb{R}^m$ is unknown.
In compressed sensing problems the recovering vector x^* is sparse, which means that the number of components different from zero is much smaller than n. In case that $m < n$ the problem is formulated by an underdetermined sparsity norm problem of the form

$$\min_x \|x\|_0 \quad \text{subject to} \quad Ax = b, \tag{1}$$

where $\|x^*\|_0 := \#\{i : x_i^* \neq 0\}$, that is, the number of nonzero elements in x^*.

There are several state of the art algorithms for solving this problem efficiently see [1], [14], [13] and [15]. The purpose of this work is to carry over a fixed-point algorithm presented in [17], [10], and [1], from a nonconvex underdetermined to a nonconvex overdetermined l_p quasi-norm problem.

Toward this end, we consider the problem of recovering an input signal $x^* \in \mathbb{R}^n$ from corrupted measurements $b = Ax^* + e$, where $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$ is a full rank matrix, $m > n$, and $e \in \mathbb{R}^n$ is an unknown error. We consider the nonconvex unconstrained problem

$$\min_x \|Ax - b\|_p^p \quad 0 \le p < 1, \tag{2}$$

for recovering successfully the input signal x^*. Then, we compare the numerical results with the convex problem when $p = 1$. To obtain an optimal solution, which is not necessarily a global solution, we apply a fixed-point procedure to the following unconstrained minimization problem:

$$\min_x \sum_{i=1}^{n} (|b - Ax|_i)^{p-2}((Ax - b)_i)^2, \tag{3}$$

where the first term of the sum is considered as a constant term. This procedure generates a sequence $\{x_k\}$ of approximate signals that, under some suitable assumptions, converges to the true input signal x^* without taking into account its error explicitly. A similar procedure is due for noncovex underdetermined ℓ_ρ quasi-norm problems in [10] and [17] and they reported that by replacing $p = 1$ by $0 < p < 1$, an exact reconstruction is possible with substantially fewer measurements than previously observed. Moreover, further decreasing the value of p yields improvement of the reconstructed signals for a relatively high rate of corruption as compared with $p = 1$.

Then, we present a new procedure for obtaining an optimal solution to the nonconvex problem (1). The approach consists in solving a sequence of convex optimization problems that reconstruct the error e, and recovering the input signal x^*. To do this, we reformulate the iterative weighted least-squares method by a sequential quadratic program. This new procedure takes into account an extra variable associated with the problem, which can be considered as a weighted residual r_w, and it has the property that is contained in the null space of A^T. We propose to solve each quadratic problem using the augmented system as a central framework, and its linear systems using two schemes: a direct method, and a projected conjugate gradient algorithm introduced by Argáez in [2]. The first scheme generates a sequence $\{r_{w_k}, x_k\}$ such that the second component converges to the signal x^* under some suitable assumptions. The second is an iterative method that generates, first, an approximation error e_k, and later uses this approximation and the corrupt measurements b to obtain an approximation x_k to the signal x^* via a least-squares problem. This procedure uses only one matrix factorization in the entire optimization procedure, and has the capability to work efficiently with sparse and dense matrices A.

These two schemes should reduce the computational cost to reconstruct a signal when the measurement matrix A has some dense columns or very large scale. Our interest in this work is to show, from a numerical point of view, that these strategies allow the efficient recovery of signals. We let the implementation of the challenge of recovering signals by using the state of the art technology in numerical linear algebra as a future work. We present a numerical experimentation to show the effectiveness of the new procedures to recover signals for high rate of corruption.

2 Problem Formulation

An optimal solution x^* of Problem (1), which is not necessarily a global solution, satisfies the following nonlinear equation:

$$A^T diag(|r^*|^{p-2})Ax^* = A^T diag(|r^*|^{p-2})b,\ r^* = Ax^* - b. \tag{4}$$

This equation is solved by applying an iterative weighted least-squares problem given by

$$A^T diag(|r_k|^{p-2})Ax_{k+1} = A^T diag(|r_k|^{p-2})b,\ r_k = Ax_k - b. \tag{5}$$

Under some assumptions, the sequence $\{x_k\}$ converges, and converges to an optimal solution

$$x_k \rightarrow x^* = argmin\|Ax - b\|_p^p. \tag{6}$$

Since $p - 2 < 0$, the weighted matrix $diag(|r_k|^{p-2})$ is undefined whenever at least one component of r_k is zero. Therefore, to overcome this situation, the weighted diagonal matrix is regularized by a positive parameter $\mu_k > 0$. That is,

$$A^T diag((|r_k| + \mu_k)^{p-2})Ax_{k+1} = A^T diag((|r_k| + \mu_k)^{p-2})b. \tag{7}$$

In the theory of compressed sensing, we propose the following algorithm to reconstructed signals exactly by solving (1) for values $0 \leq p \leq 1$:

Algorithm 1
Step 0. **Initialization**
a. Input data: A, b, x^*, and $p \in [0, 1]$.
b. Set: μ, ϵ_1, ϵ_μ, and ϵ_2.
c. Set: k_{max} and $k = 0$.
Step 1. Solve: $A^T Ax = A^T b$.
Step 2. **while** $k \leq k_{\max}$ **do**
Step 3. Define the residual vector $r = Ax - b$,
and the matrix $D = diag(|r| + \mu)$.
Step 4. Set: $x_{prev} = x$. Solve for x
$A^T D^{p-2} Ax = A^T D^{p-2} b$.
Step 5. **If** $\frac{\|x - x_{prev}\|}{1+\|x\|} > \epsilon_1$, $k = k + 1$ **go to** Step 2.
Step 6. **If** $\mu > \epsilon_\mu$
updates: μ and ϵ_1. Set $k = 0$; **go to** Step 2
else
display 'x is an optimal solution'
end
Step 7. **If** $\|x - x^*\|_\infty \leq \epsilon_2$
display 'the signal is recovered'
else
display 'fail to recover the signal'
end

This algorithm works as follows:

At Step 0, initialization, is given the decoding matrix A, the corrupt measurements b, the true signal x^*, and the value of the quasi-norm problem p, for which we want to recover the signal. In Step 1, the first approximation x to the true signal x^* is calculated by solving a normal equation, which gives the 2-norm approximation to the true signal. Then, in Step 4 a new approximation to the signal is generated by solving a weighted normal equation. The weight is a positive diagonal matrix that depends on the quasi-norm p, the residual vector r, and a regularization parameter μ as defined in Step 3. Now, for a fixed parameter μ, in Step 5 we determine if a solution x of the weighted normal equation is a good approximation to (6). Then, if the parameter μ is less than a prescribed value ϵ_μ, we claim that the algorithm found an optimal solution in Step 6. Finally, the algorithm recovers the true signal if the ∞-norm of the difference between the optimal solution and the true signal is less than a value ϵ_2 in Step 7. In our implementation, we use a sparse Cholesky factorization to obtain a lower triangular matrix L to the positive definite matrix $A^T D^{p-2} A$, and this matrix is substituted by $L^T L$.

In the next section, we present a quadratic problem for solving the weighted normal equation, and therefore a sequential quadratic program is used to obtain an optimal solution of (1).

3 A Quadratic Problem

We present a quadratic problem as an equivalent approach to solve the weighted normal equation that appears in Step 4 in Algorithm 1. For practical purposes, the subindices are omitted resulting in

$$A^T D^{p-2} Ax = A^T D^{p-2} b, \tag{8}$$

where $D = diag(|r| + \mu)$ with $r = Ax_- - b$, and x_- is the previous value of x.

The quadratic problem associated to (7) is given by:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2} r_w^T D^{2-p} r_w - b^T r_w \\ \text{subject to} & A^T r_w = 0, \end{array} \tag{9}$$

where $r_w \in \mathbb{R}^m$ is considered the weighted residual of the equation (7).

The Lagrange function associated to (8) is

$$L(r_w, x) = \frac{1}{2} r_w^T D^{2-p} r_w - b^T r_w + r_w^T A x, \tag{10}$$

where $x \in \mathbb{R}^n$ is the Lagrange multiplier associated to the equality constraint. Since A has full rank and D is a positive definite matrix, then (9) has a unique global solution which can be determined by solving the following augmented system:

$$\begin{pmatrix} D^{2-p} & A \\ A^T & 0 \end{pmatrix} \begin{pmatrix} r_w \\ x \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}. \tag{11}$$

It is straightforward to show that the Lagrange multiplier x associated to (8) is the solution of the weighted normal equation (7). The advantages of this formulation with respect to (7) is that we can handle simultaneously the corrupt error given by $e = D^{2-p} r_w$ and the signal x since $Ax + e = b$, and the augmented system is reasonably well-conditioned since the exponent of the diagonal matrix D is positive. Another advantage of using (10) is that for dense matrices A, the LDL^T sparse Cholesky factorization of the symmetric and indefinite augmented system could lead to a sparser and cheaper factorization than the sparse Cholesky factorization for the positive definite matrix associated to the weighted normal equation (7).

Therefore, we have a sequential quadratic program for solving nonconvex overdetermined ℓ_ρ-regularized quasi-norm problems using the augmented system as a central framework.

We use a direct method to obtain the solution (r_w, x) of (10). Then x is used in Algorithm 1, at Step 4, as a signal approximation, instead of the value obtained using the sparse Cholesky factorization for solving the weighted normal equation. In the next section, we present a projected conjugate gradient algorithm to obtain optimal inexact directions associated to the augmented system (10) that have the potential to reduce the computational time to recover input signals for large measurement matrices A.

4 A Projected Conjugate Gradient

We propose to solve (10) using the projected conjugate gradient algorithm presented by Argáez in [2]. The idea consists in reducing the problem into the null space of A^T. Solving the first block of equations of (10) for x, and substituting the result in the same block of equations, we obtain the following projected equation:

$$PD^{2-p} r_w = Pb,$$

where $P = I - A(A^T A)^{-1} A^T$ is an orthogonal projector on the null space of A^T.

This equation is consistent since both sides of it are preceded by the projector operator P, and has infinitely many solutions since P is a singular matrix. Moreover, the minimum 2-norm solution is unique and it is in the null space of A^T see ([2]). Therefore, the unique solution r_w is the solution of (10) and the solution x is obtained by solving a least-squares problem.

We propose to use the conjugate gradient algorithm for solving $PD^{2-p}r_w = Pb$ with an initial approximation $r_w = 0$ to obtain the unique solution in the null space of A^T. The algorithm generates a sequence $\{(r_{w_j}, x_j)\}$, with r_{w_j} contained in the null space of A^T, that converges to the solution (r_{w_j}, x_j) of (10) see ([2]). Now, our interest is to verify if the conjugate gradient algorithm, applied to the projected equation, allows the recovery of signals with a high rate of corruption. To study this procedure, we use Algorithm 1 as a central framework. The idea consists of substituting direction x, at Step 4, by an optimal direction x, given by conjugate gradient algorithm applied to the projected equation. The procedure works as follows:

Given a regularized positive value μ and a weighted matrix D^{2-p}, the algorithm first calculates an approximation error $e_j = D^{2-p}r_{w_j}$, where r_{w_j} is obtained by applying the conjugate gradient algorithm to solve $PD^{2-p}r_w = Pb$. Then, solving a least-squares problem $\min_x \frac{1}{2}\|Ax - (b-e)\|_2^2$ the algorithm obtains an approximate signal x_j. The conjugate gradient is repeated until the 2-norm of the residual $r_j = b - (Ax_j + e_j)$ is less than a stopping value, which means that an optimal inexact direction is obtained. The rest of the steps are the same as in algorithm 1. In order to carry out this implementation, it is necessary to determine a stopping value for the conjugate gradient algorithm and a maximum number of iterations for obtaining an optimal direction. To recover signals, we present the following algorithm that uses inexact directions instead of exact directions via a conjugate gradient algorithm:

Algorithm 2. Projected Conjugate Gradient Algorithm.
Step 0. **Initiliazation**
  a. Input data: A, b, x^*, and $p \in [0, 1]$.
  b. Set: μ, ϵ_1, ϵ_μ, ϵ_2, and ϵ_3.
  c. Set: $k_{\max}$ and $k = 0$.
  d. Set: $itercg_{max}$ and $r_0 = 0$.
Step 1. **First approximation to the true signal**
  a. Solve: $x = argmin \frac{1}{2}\|Ax - b\|_2^2$.
  b. Initial residual: $r = b - Ax$.
  c. Set: $d_0 = r$ and $\beta_{n0} = r^T r$.
Step 2. **Outer loop**:
  a. **while** $k \leq k_{\max}$ **do**
  b. Set: $itercg = 1$, $r_w = 0$,
    $d = d_0$, and $\beta_n = \beta_{n0}$.
Step 3. Update: $D = diag(|r| + \mu)$,
    and set $x_{prev} = x$.
Step 4. **Inner loop**:
  a. **while** $itercg \leq itercg_{\max}$ **do**
  b. Set: $\alpha_d = d^T D^{2-p} d$.
  c. Weighted residual: $r_w = r_w + \frac{\beta_n}{\alpha_d} d$.
  d. Error approximation: $e = D^{2-p} r_w$.
  e. Signal approximation:
    $x = argmin \frac{1}{2}\|Ax - (b - e)\|_2^2$.
  f. Residual: $r = b - (Ax + e)$.
  g. Stopping criteria: **If** $\|r\| > \epsilon_3$
  h. Set: $\beta_d = \beta_n$, $\beta_n = r^T r$, and
  $d = r + \frac{\beta_n}{\beta_d} d$.
  i. Set: $itercg = itercg + 1$, **go** to Step 4.
Step 5. **If** $\frac{\|x - x_{prev}\|}{1 + \|x\|} > \epsilon_1$, $k = k + 1$, **go** to Step 2.
Step 6. **If** $\mu > \epsilon_\mu$
    updates: μ and ϵ_1. Set $k = 0$; **go** to Step 2.
    else
     display 'x is an optimal solution'
    end
Step 7. **If** $\|x - x^*\|_\infty \leq \epsilon_2$
     display 'the signal is recovered'
    else
     display 'fail to recover the signal'
    end

This algorithm follows the same structure as algorithm 1. We named Step 2 outer loop in order to differentiate between the number of iterations needed to obtain a good approximation to the signal for a fixed regularization parameter μ, and Step 4 inner loop to highlight the optimal direction obtained from the projected conjugate gradient algorithm. It is important to note that from Substeps 4.a to 4.i the algorithm carried out the projected conjugate gradient to find an optimal direction x. The procedure first obtains an approximation to the weighted residual r_w by using a conjugate gradient strategy, and then finds an approximation to the error e using r_w. Then, by solving a least-squares problem, an approximation to the signal x is found. It is important to realize that the directions, d, the weighted residuals, r_w, and the residuals, r, obtained in Substeps 4.h, 4.c, and 4.f, respectively, are on the null space of A^T. Moreover, the residuals r are orthogonal projections of $b - e$ on the null space of A^T. An important numerical property is that just one matrix factorization is carried out for solving the linear systems that appear in Steps 1.a and 4.e of the algorithm. In our numerical implementation, we use a sparse Cholesky factorization to solve the linear least-squares problems that appears in these steps. Also, we use the augmented system $\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix}$ to obtain the signal approximations x and the residuals r in Steps 1.a and 1.b, and 4.e and 4.f, respectively. In this case the algorithm is carried out without matrix-vector multiplication.

5 Numerical Experimentation

We study the numerical behavior of Algorithms 1 and 2, with the different strategies for solving the linear systems associated to the problem, to recover signals. Our interest is to verify if both algorithms with the different schemes to solve the linear systems associated to the problem allow the recovery of signals at a highly corrupt rate. For now, we are not interested in reducing or in studying the computational cost of recovering signals, but in the ability of the algorithms with the different strategies to recover input signals for values $0 \leq p \leq 1$. We will present a numerical experimentation that shows the ability of our algorithms to reconstruct input signals with substantially fewer measurements for values $0 \leq p \leq 1$.

The numerical experimentation was done on an Intel Xeon 3.06 GHz processor with 2 GB of main memory. The algorithms were written in MATLAB Version 7.1.0. Our experimentation has the objective to investigate the capabilities of recovering signals by the strategies discussed in this paper: Algorithm 1 uses a direct method to solve the weighted least-squares problems and the augmented systems. Algorithm 2, the projected conjugate gradient, uses the normal equations and the augmented systems for obtaining inexact directions. Our special interest is the location of the breakpoint beyond at which ℓ_p fails to recover the signal exactly for values $p \in \Omega = \{1, 0.9, ..., 0.1, 0\}$. To study these issues, we performed a series of experiments as follows:

1. Select an input signal of size $n = 128$. Select randomly a matrix $A_{m \times n}$ with $m = 2n$. Sample A with independent Gaussian with mean zero and standard deviation 1, and select randomly a signal $x^* \in \mathbb{R}^n$.

2. Select a set of 40 percentages of m that goes from 1% to 40%, i.e.,

$$\Theta = \left\{ \frac{m}{100} t, t = 1, 2, \ldots, 40 \right\}.$$

   Then, for each element $\rho \in \Theta$, define a support set $\Gamma = \{e \in \mathbb{R}^m, \|e\|_o = \rho\}$. Then sample an e with independent and identically distributed Gaussian entries, and with standard deviation about the coordinates of Ax^*.

3. Compute the corrupt measurements $b = Ax^* + e$, and find an optimal solution x of problem (1), for each $p \in \Omega$, using the four strategies mentioned in this paper.

4. Compared the solution x with the true signal x^*.

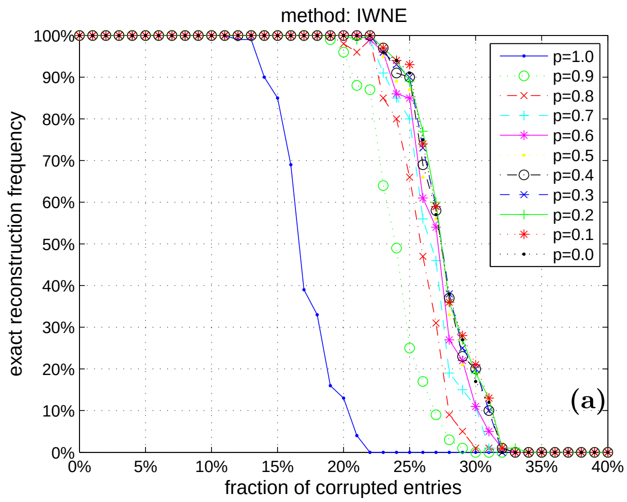
5. Repeat the experiment 100 times for each $p \in \Omega$ and $\rho \in \Theta$.

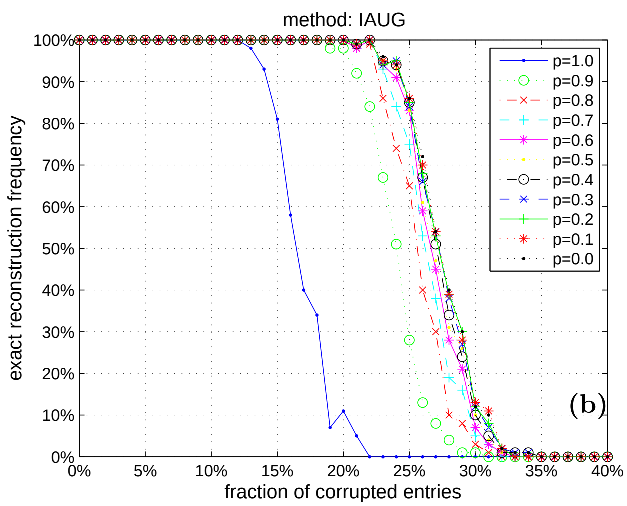
We implemented the algorithms 1 and 2 as follows: The Algorithms take the least-squares solution x as the initial approximation to the signal. The regularization parameter μ is initially set to 1, and then is reduced by a factor of 10^2 at each iteration. For each regularization parameter μ, x is improved if the relative change between two consecutive iterates is below $\epsilon_1 = \frac{\sqrt{\mu}}{100}$. The algorithms claim to find an optimal solution to problem (1) when the regularization parameter is smaller than $\epsilon_\mu = 10^{-10}$. In Algorithm 2 appears a new stopping value ϵ_3 that controls inexact approximations to x and is updated dynamically by $\frac{\sqrt{\mu}}{100}$. We claim that an optimal solution recovers the true signal if $\|x - x^*\|_\infty \leq \epsilon_2$ with $\epsilon_2 = 10^{-6}$.

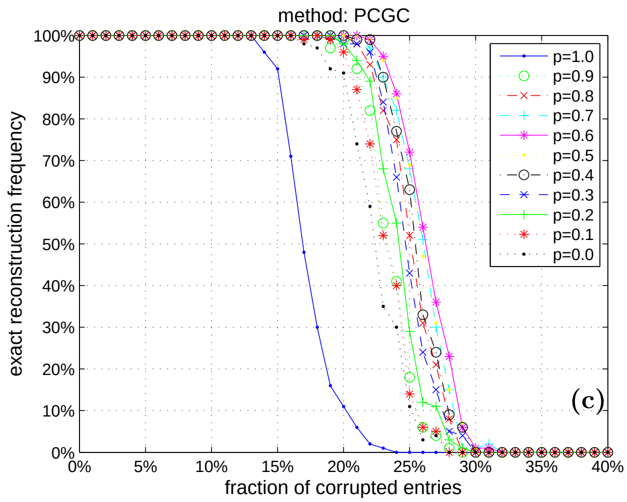
We denote Algorithm 1 that uses the sparse Cholesky factorization to solve the weighted normal equation by **IWNE**, and if uses the LDL^T sparse Cholesky factorization to solve the augmented system by **IAUG**. The Algorithm 2 that uses the sparse Cholesky factorization to solve the normal equation by **PCGC**, and if uses the LDL^T sparse Cholesky factorization to solve the augmented system by **PCGA**. The numerical results are presented in Figures 1.a, 1.b, 1.c and 1.d. From Figures 1.a and 1.b, we observe that the numerical results for **IWNE** and **IAUG** are almost similar. In particular for $p = 1$ exact reconstruction occurred at all 100 times for a corruption rate $\rho \leq 13\%$, 93 times for a corruption rate $\rho \leq 14\%$, and 85 times for a corruption rate $\rho \leq 15\%$. For $p \in \{0.8, 0.7, ..., 0.1\}$, an exact reconstruction occurred all 100 times for a corruption rate $\rho \leq 20\%$, and for $p = 0.9$ exact reconstruction occurred 99 times. It is interesting to observe that after 20% corruption, it is more likely to obtain an exact reconstruction when the value of p is further reduced.

From Figures 1.c and 1.d, **PCGC** and **PCGA** respectively, we have, for $p = 1$, that the reconstruction behavior is similar to the methods **IWNE** and **IAUG**. For **PCGC** and **PCGA** with $p \in \{0.8, 0.7, ..., 0.3\}$ exact reconstruction occurred the 100 times for a corrupted rate $\rho \leq 20\%$, and for $p = \{0.9, 0.2, 0.1\}$ exact reconstruction occurred 100 times just for $\rho \leq 15\%$. It is interesting to observe that for values of p around 0.5 that is more likely to obtain exact reconstructed signals for corruption higher than 20%. For the extreme values of p, that is for $p = 1, 0.9, 0.1, 0$, the projected conjugate gradient algorithm becomes more unstable in recovering the true signals independent of the factorization matrix we use.

In summary, we observe that decreasing p from 1 to $0 \leq p \leq 0.9$ results in a substantial improvement in recovering true signals. By decreasing p even further, the algorithm yields an improvement of recovering signals using direct methods. This is not the situation for the projected conjugate gradient, which yields only improvement for values of p around 0.5.


method: IWNE
exact reconstruction frequency
fraction of corrupted entries
p=1.0
p=0.9
p=0.8
p=0.7
p=0.6
p=0.5
p=0.4
p=0.3
p=0.2
p=0.1
p=0.0
(a)


method: IAUG
exact reconstruction frequency
fraction of corrupted entries
p=1.0
p=0.9
p=0.8
p=0.7
p=0.6
p=0.5
p=0.4
p=0.3
p=0.2
p=0.1
p=0.0
(b)



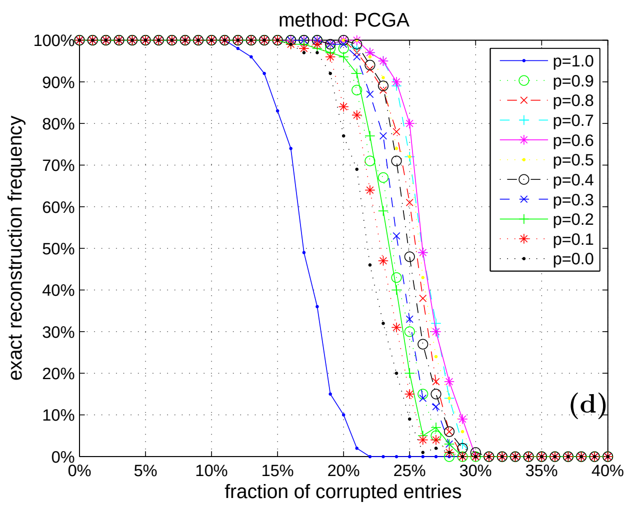


Figure 1: The four figures illustrate the behavior of the frequency of the exact reconstruction signals versus the corrupted entries for the three methods to $n = 128$ and $m = 2n$.

6 Conclusions

We presented fixed-point algorithms for solving a nonconvex overdetermined ℓ_ρ quasi-norm problems using an iteratively weighted normal equation. Then, we posed this procedure by a sequential quadratic program where the central framework is the solution of an augmented system. We propose two schemes for solving the augmented system: a direct method and a projected conjugate gradient method to obtain optimal inexact directions. We apply these methodologies to recover input signals. Our numerical experimentation shows that the four strategies are capable to recover the signals for a relatively high rate of corruption as compared with $p = 1$. We observe that the use of direct methods improves the chance of recovering signals with a higher rate of corruption, while the values p are closer to zero. In particular, we observed for $p = 0.1$ improves the chance of recovering signals independent of what direct method is implemented. Instead, the projected conjugate gradient method, which works with inexact directions, improves the chance of recovering signals for values of p around 0.5.

References

[1] M. Argaez, C. Quintero, C. Ramirez and R. Sanchez, "A path following method for large-scale and dense ℓ_1-underdetermined problems", submmited to *IEEE Transactions on Signal Processing*, Mathematical Sciences Dept., University of Texas at El Paso, 2009.

[2] M. Argaez, *A projected conjugate gradient algorithm for KKT systems*, Tech. Rep. TR01-2010, Mathematical Sciences Dept., University of Texas at El Paso, 2010.

[3] E. Candes, M. Walkin, and S. Boyd, "Enhancing sparsity by reweighted $\ell 1$ minimization", *J. Fourier Anal. Appl.*, vol. 14, 2007.

[4] E. Candes, J. Romberg,, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", *IEEE Trans. Inf. Theory*, vol. 52, 2006.

[5] E. Candes, "Compressive sampling", *Proceeding of the International Congress of Mathematicians*, Madrid, Spain, 2006.

[6] E. Candes, and T. Tao, "Decoding by Linear Programming", *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203–4215, 2005.

[7] E. Candes, and J. Romberg, "Sparsity and Incoherence in Compressive Sampling", *Inverse Problems in IOP* (electronic journal) http://www.iop.org/EJ/IP, vol. 23, pp. 969–985, 2006.

[8] R. Chartrand, "Nonconvex compressed sensing and error correction", *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2007.

[9] R. Chartrand, "Exact Reconstruction of Sparse Signals via Nonconvex Minimization", *IEEE Signal Process Lett.*, vol. 14, pp. 707–710, 2007.

[10] R. Chartrand, and W. Yin, "Iteratively Reweighted Algorithms for Compressive Sensing", *Proceedings of the 33rd International Conference on Acoustics, Speech, and Signal Processing ICASSP*, 2008.

[11] D. Donoho, "Compressed sensing", *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, 2006.

[12] D. Donoho, and Y. Tsaig, "Fast solution of ℓ_1-norm minimization problems when the solution may be sparse", *IEEE Transactions on Information Theory*, vol. 54, pp. 4789–4812, 2008.

[13] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems", *IEEE J. Select. Topics Signal Process.*, 2007.

[14] E. Hale, W. Yin, and Y. Zhang, *A Fixed-point Continuation Method for ℓ_1-regularized Minimization with Applications to Compressed Sensing*, Technical Report TR-07-07, Department of Computational and Applied Mathematics, Rice University, Houston, TX, USA, 2007.

[15] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An Interior-point Method for Large-scale ℓ_1-regularized Least-Squares", *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007. Available at `www.stanford.edu/~boyd/l1_ls.html`

[16] T. Le, R. Chartrand, and T. Asaki, "A Variational Approach to Reconstructing Images Corrupted by Poisson Noise", *Journal of Mathematical Imaging and Vision*, vol. 27, pp. 257–263, 2007.

[17] C. Miosso, R. von-Borries, M. Argáez, L. Velázquez, C. Quintero, and C. Potes, "Compressive Sensing Reconstruction With Prior Information by Iteratively Reweighted Least-Squares", *IEEE Transactions on Signal Processing*, vol. 57, pp. 2424–2431, 2009.

[18] Y. Wang, W. Yin, and Y. Zang, *A Fast Algorithm for Image Deblurring with Total Variation Regularization*, CAAM Technical Report TR07-10, Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005, 2007.