Reconstruction of Sparse Signals by Minimizing a Re-Weighted Approximate $\ell_0$-Norm in the Null Space of the Measurement Matrix

J. K. Pant, W.-S. Lu, and A. Antoniou

University of Victoria

August 3, 2010
Compresive Sensing
Compressive Sensing

Signal Recovery by $\ell_1$ Minimization
Compressive Sensing

Signal Recovery by $\ell_1$ Minimization

Signal Recovery by $\ell_p$ Minimization with $p < 1$
Outline

- Compressive Sensing
- Signal Recovery by $\ell_1$ Minimization
- Signal Recovery by $\ell_p$ Minimization with $p < 1$
- Performance Evaluation
A signal $x(n)$ of length $N$ is $K$-sparse if it contains $K$ nonzero components with $K \ll N$. 
A signal $x(n)$ of length $N$ is $K$-sparse if it contains $K$ nonzero components with $K \ll N$.

A signal is near $K$-sparse if it contains $K$ significant components.
Sparsity is a generic property of signals: A real-world signal always has a sparse or near-sparse representation with respect to an appropriate basis.
Sparsity is a generic property of signals: A real-world signal always has a sparse or near-sparse representation with respect to an appropriate basis.

An Image

An equivalent 1-D signal

A wavelet representation of the image
Compressive sensing (CS) is a data acquisition process whereby a sparse signal $x(n)$ represented by a vector $x$ of length $N$ is determined using a small number of projections represented by a matrix $\Phi$ of dimension $M \times N$. 

$$y = \Phi \cdot x$$
Compressive sensing (CS) is a data acquisition process whereby a sparse signal $x(n)$ represented by a vector $x$ of length $N$ is determined using a small number of projections represented by a matrix $\Phi$ of dimension $M \times N$.

In such a process, measurement vector $y$ and signal vector $x$ are interrelated by the equation $y = \Phi \cdot x$.

\[
\begin{bmatrix}
\text{measurements} \\
\end{bmatrix}_{M \times 1} = 
\begin{bmatrix}
\text{projection matrix} \\
\end{bmatrix}_{M \times N} 
\begin{bmatrix}
\text{sparse signal of interest} \\
\end{bmatrix}_{N \times 1}
\]
CS theory shows that these random projections contain much, sometimes all, the information content of signal $x$. 

A condition for this to be possible is 

$$M \geq c \cdot K \cdot \log \left( \frac{N}{K} \right)$$ 

where $c$ is a small constant.
Compressive Sensing, cont’d

- CS theory shows that these random projections contain much, sometimes all, the information content of signal \( x \).

- If a sufficient number of such measurements are collected, recovering signal \( x \) from measurements \( y \) is possible.

\[ M \geq c \cdot K \cdot \log \left( \frac{N}{K} \right) \]

where \( c \) is a small constant.
CS theory shows that these random projections contain much, sometimes all, the information content of signal $\mathbf{x}$.

If a sufficient number of such measurements are collected, recovering signal $\mathbf{x}$ from measurements $\mathbf{y}$ is possible.

A condition for this to be possible is

$$M \geq c \cdot K \cdot \log\left(\frac{N}{K}\right)$$

where $c$ is a small constant.
CS theory shows that these random projections contain much, sometimes all, the information content of signal \( \mathbf{x} \).

If a sufficient number of such measurements are collected, recovering signal \( \mathbf{x} \) from measurements \( \mathbf{y} \) is possible.

A condition for this to be possible is

\[
M \geq c \cdot K \cdot \log(N/K)
\]

where \( c \) is a small constant.

Typically,

\[
K < M < N
\]
Recovering signal vector $\mathbf{x}$ from measurement vector $\mathbf{y}$ such that

$$
\Phi_{M \times N} \cdot \mathbf{x}_{N \times 1} = \mathbf{y}_{M \times 1}
$$

is an ill-posed problem.
Signal Recovery by $\ell_1$ Minimization

- Recovering signal vector $x$ from measurement vector $y$ such that

$$\Phi \cdot x = y$$

is an ill-posed problem.

- Given that $x$ is sparse, $x$ can be reconstructed by solving the $\ell_1$-minimization problem

$$\min_{x} \quad \|x\|_1$$

subject to

$$\Phi x = y$$

where $\|x\|_1 = \sum_{i=1}^{N} |x_i|$. 
Why $l_1$-norm minimization?
Why $\ell_1$-norm minimization?

As $c$ increases, the contour of $||x||_1 = c$ grows and touches the hyperplane $\Phi x = y$, yielding a sparse solution

$$x^* = \begin{bmatrix} 0 \\ c \end{bmatrix}$$

Contours for $||x||_1 = c$
Why $\ell_2$-norm minimization fails to work?
Why $\ell_2$-norm minimization fails to work?

As $r$ increases, the contour of $\|x\|_2 = r$ grows and touches the hyperplane $\Phi x = y$.

The solution $x^*$ obtained is not sparse.
Theorem

If $\Phi = \{\phi_{ij}\}$ where $\phi_{ij}$ are independent and identically distributed random variables with zero-mean and variance $1/N$ and $M \geq cK \log(N/K)$, the solution of the $\ell_1$-minimization problem would recover exactly a $K$-sparse signal with high probability.
Theorem

If $\Phi = \{\phi_{ij}\}$ where $\phi_{ij}$ are independent and identically distributed random variables with zero-mean and variance $1/N$ and $M \geq cK \log(N/K)$, the solution of the $\ell_1$-minimization problem would recover exactly a $K$-sparse signal with high probability.

- For real-valued data $\{\Phi, y\}$, the $\ell_1$-minimization problem is a linear programming problem.
Signal Recovery by $\ell_1$ Minimization, cont’d

Example: $N = 512$, $M = 120$, $K = 26$
Example: \( N = 512, M = 120, K = 26 \)
The sparsity of a signal can be measured by using its $\ell_0$ pseudonorm

$$||x||_0 = \sum_{i=1}^{N} |x_i|^0$$
The sparsity of a signal can be measured by using its $\ell_0$ pseudonorm

$$||x||_0 = \sum_{i=1}^{N} |x_i|^0$$

Hence the sparsest solution of $\Phi x = y$ can be obtained by solving the $\ell_0$-norm minimization problem

$$\text{minimize}_{x} ||x||_0 \quad \text{subject to} \quad \Phi x = y$$
The sparsity of a signal can be measured by using its $\ell_0$ pseudonorm

$$\|x\|_0 = \sum_{i=1}^{N} |x_i|^0$$

Hence the sparsest solution of $\Phi x = y$ can be obtained by solving the $\ell_0$-norm minimization problem

$$\min_{x} \|x\|_0$$
subject to $\Phi x = y$

Unfortunately, the $\ell_0$-norm minimization problem is nonconvex with combinatorial complexity.
An effective signal recovery strategy is to solve the $\ell_p$-minimization problem

$$\min_{x} ||x||_p^p \quad \text{with} \quad 0 < p < 1$$

subject to

$$\Phi x = y$$

where $||x||_p^p = \sum_{i=1}^{N} |x_i|^p$. 

The $\ell_p$-norm minimization problem is nonconvex.
An effective signal recovery strategy is to solve the $\ell_p$-minimization problem

$$\min_{x} \|x\|_p^p \quad \text{with} \quad 0 < p < 1$$

subject to

$$\Phi x = y$$

where $\|x\|_p^p = \sum_{i=1}^{N} |x_i|^p$.

The $\ell_p$-norm minimization problem is nonconvex.
Contours of $||x||_p = 1$ with $p < 1$
Why $\ell_p$ minimization with $p < 1$?

As $c$ increases, the contour $||x||_p = c$ grows and touches the hyperplane $\Phi x = y$, yielding a sparse solution $x^* = [0 \ldots c]$. The possibility that the contour will touch the hyperplane at another point is eliminated.
Why $\ell_p$ minimization with $p < 1$?

As $c$ increases, the contour $\|x\|_p^p = c$ grows and touches the hyperplane $\Phi x = y$, yielding a sparse solution $x^* = \begin{bmatrix} 0 \\ c \end{bmatrix}$.

The possibility that the contour will touch the hyperplane at another point is eliminated.

Contours of $\|x\|_p^p = c$ with $p < 1$
We propose to minimize an approximate \( \ell_0 \)-norm

\[
\| x \|_{0, \sigma} = \sum_{i=1}^{N} \left( 1 - e^{-x_i^2/2\sigma^2} \right)
\]

where \( x \) lies in the solution space of \( \Phi x = y \), namely,

\[
x = x_s + V_r \xi
\]

where \( x_s \) is a solution of \( \Phi x = y \) and \( V_r \) is an orthonormal basis of the null space of \( \Phi \).
Why norm $||x||_{0,\sigma}$ works?

With $\sigma$ small,

\[
\left| \left( 1 - e^{-x^2_i/2\sigma^2} \right) x_i = 0 \right| \approx 0
\]
and

\[
\left| \left( 1 - e^{-x^2_i/2\sigma^2} \right) x_i \neq 0 \right| \approx 1
\]

Therefore, for a $K$-sparse signal,

\[
||x||_{0,\sigma} = N \sum_{i=1}^{K} \left| \left( 1 - e^{-x^2_i/2\sigma^2} \right) x_i \right| \approx K = ||x||_0
\]
Why norm $||x||_{0,\sigma}$ works?

With $\sigma$ small,

$$\left(1 - e^{-x_i^2/2\sigma^2}\right) \bigg|_{x_i=0} = 0$$

and

$$\left(1 - e^{-x_i^2/2\sigma^2}\right) \bigg|_{x_i \neq 0} \approx 1$$
Why norm $\|x\|_{0, \sigma}$ works?

With $\sigma$ small,

$$
\left(1 - e^{-x_i^2/2\sigma^2}\right) \bigg|_{x_i=0} = 0
$$

and

$$
\left(1 - e^{-x_i^2/2\sigma^2}\right) \bigg|_{x_i \neq 0} \approx 1
$$

Therefore, for a $K$-sparse signal,

$$
\|x\|_{0, \sigma} = \sum_{i=1}^{N} \left(1 - e^{-x_i^2/2\sigma^2}\right) \approx K = \|x\|_0
$$
Improved recovery rate can be achieved by using a re-weighting technique.

\[
\min_{\mathbf{x}} \sum_{i=1}^{n} w_i \left\{ \frac{1}{2} \left( \frac{x_s(i) + v_i^T \xi}{\sigma^2} \right)^2 \right\}
\]

where \(w_i(k+1) = 1 |x(i)(k)| + \epsilon\).
Improved recovery rate can be achieved by using a re-weighting technique.

This involves solving the optimization problem

\[
\min_{\xi} \sum_{i=1}^{n} w_i \left\{ 1 - e^{-\left[x_s(i) + v_i^T \xi\right]^2/2\sigma^2} \right\}
\]

where

\[
w_i^{(k+1)} = \frac{1}{|x_i^{(k)}| + \epsilon}
\]
Performance Evaluation

Number of perfectly recovered instances versus sparsity $K$ by various algorithms with $N = 256$ and $M = 100$ over 100 runs.

IR: Iterative re-weighting (Chartrand and Yin, 2008)
SL0: Smoothed $\ell_0$-norm minimization (Mohimani et. al., 2009)
NRAL0: Proposed
Average CPU time versus signal length for various algorithms with $M = N/2$ and $K = M/2.5$.

IR: Iterative re-weighting (Chartrand and Yin, 2008)
SL0: Smoothed $\ell_0$-norm minimization (Mohimani et. al., 2009)
NRAL0: Proposed
Performance comparison of $\ell_1$ minimization with approximate $\ell_0$ minimization for $N = 512$, $M = 80$, $K = 30$. 

![Graphs comparing original signal with recovery by $\ell_1$ and $\ell_0$ minimization, showing reconstruction error.](image)
Conclusions

- Compressive sensing is an effective technique for signal sampling.
Conclusions

- Compressive sensing is an effective technique for signal sampling.
- $\ell_1$ minimization works in general for the reconstruction of sparse signals.
Compressive sensing is an effective technique for signal sampling.

- $\ell_1$ minimization works in general for the reconstruction of sparse signals.

- $\ell_p$ minimization with $p < 1$ can improve the recovery performance for signals that are less sparse.
Compressive sensing is an effective technique for signal sampling.

- $\ell_1$ minimization works in general for the reconstruction of sparse signals.

- $\ell_p$ minimization with $p < 1$ can improve the recovery performance for signals that are less sparse.

- Approximate $\ell_0$-norm minimization offers good performance with improved complexity.
Thank you for your attention.