Mathematics of Compressive Sensing: Sparse recovery using $\ell_1$ minimization

Justin Romberg

Georgia Tech, School of ECE

Tsinghua University

October 17, 2013
Beijing, China
We say that a $M \times N$ matrix $\Phi$ is a restricted isometry for sets of size $2S$ if there exists a $0 \leq \delta_{2S} < 1$ such that

$$(1 - \delta_{2S})\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_{2S})\|x\|_2^2$$

for all $2S$-sparse $x$.

It is clear that if $\Phi$ is $2S$-RIP and $y = \Phi x_0$, where $x_0$ is an $S$-sparse vector, then we can recover $x_0$ by solving

$$\min_x \|x\|_0 \quad \text{subject to} \quad \Phi x = y.$$ 

Now we will see that this condition also means we can recover by minimizing the $\ell_1$ norm...
Our goal

**Theorem:** Let $\Phi$ be an $M \times N$ matrix that is an approximate isometry for $3S$-sparse vectors. Let $x_0$ be an $S$-sparse vector, and suppose we observe $y = \Phi x_0$. Given $y$, the solution to

$$\min_x \|x\|_1 \quad \text{subject to} \quad \Phi x = y$$

is exactly $x_0$. 
Moving to the solution

\[
\min_x \| x \|_1 \quad \text{such that} \quad \Phi x = y
\]

Call the solution to this \( x^\# \). Set

\[
h = x^\# - x_0.
\]
Moving to the solution

\[
\min_x \|x\|_1 \quad \text{such that} \quad \Phi x = y
\]

Call the solution to this \(x^\#\). Set

\[
h = x^\# - x_0.
\]

Two things must be true:

- \(\Phi h = 0\)
  Simply because both \(x^\#\) and \(x_0\) are feasible: \(\Phi x^\# = y = \Phi x_0\)

- \(\|x_0 + h\|_1 \leq \|x_0\|_1\)
  Simply because \(x_0 + h = x^\#\), and \(\|x^\#\|_1 \leq \|x_0\|_1\)
Moving to the solution

\[
\min_x \|x\|_1 \text{ such that } \Phi x = y
\]

Call the solution to this \( x^\# \). Set

\[
h = x^\# - x_0.
\]

Two things must be true:

- \( \Phi h = 0 \)
  
  Simply because both \( x^\# \) and \( x_0 \) are feasible: \( \Phi x^\# = y = \Phi x_0 \)

- \( \|x_0 + h\|_1 \leq \|x_0\|_1 \)
  
  Simply because \( x_0 + h = x^\# \), and \( \|x^\#\|_1 \leq \|x_0\|_1 \)

We’ll show that if \( \Phi \) is 3S-RIP, then these conditions are \textit{incompatible} unless \( h = 0 \)
Two things must be true:

- $\Phi h = 0$
- $\|x_0 + h\|_1 \leq \|x_0\|_1$

\(H = \{x : \Phi x = y\}\)
Cone condition

For $\Gamma \subset \{1, \ldots, N\}$, define $h_\Gamma \in \mathbb{R}^N$ as

$$h_\Gamma(\gamma) = \begin{cases} h(\gamma) & \gamma \in \Gamma \\ 0 & \gamma \notin \Gamma \end{cases}$$

Let $\Gamma_0$ be the support of $x_0$. For any "descent vector" $h$, we have

$$\|h_{\Gamma_0^c}\|_1 \leq \|h_{\Gamma_0}\|_1$$
Cone condition

For $\Gamma \subset \{1, \ldots, N\}$, define $h_\Gamma \in \mathbb{R}^N$ as

$$h_\Gamma(\gamma) = \begin{cases} h(\gamma) & \gamma \in \Gamma \\ 0 & \gamma \notin \Gamma \end{cases}$$

Let $\Gamma_0$ be the support of $x_0$. For any “descent vector” $h$, we have

$$\|h_{\Gamma_0^c}\|_1 \leq \|h_{\Gamma_0}\|_1$$

Why? The triangle inequality..

$$\|x_0\|_1 \geq \|x_0 + h\|_1 = \|x_0 + h_{\Gamma_0} + h_{\Gamma_0^c}\|_1$$

$$\geq \|x_0 + h_{\Gamma_0^c}\|_1 - \|h_{\Gamma_0}\|_1$$

$$= \|x_0\|_1 + \|h_{\Gamma_0^c}\|_1 - \|h_{\Gamma_0}\|_1$$
Cone condition

For $\Gamma \subset \{1, \ldots, N\}$, define $h_\Gamma \in \mathbb{R}^N$ as

$$h_\Gamma(\gamma) = \begin{cases} h(\gamma) & \gamma \in \Gamma \\ 0 & \gamma \notin \Gamma \end{cases}$$

Let $\Gamma_0$ be the support of $x_0$. For any “descent vector” $h$, we have

$$\|h_{\Gamma_0^c}\|_1 \leq \|h_{\Gamma_0}\|_1$$

We will show that if $\Phi$ is $3S$-RIP, then

$$\Phi h = 0 \quad \Rightarrow \quad \|h_{\Gamma_0}\|_1 \leq \rho \|h_{\Gamma_0^c}\|_1$$

for some $\rho < 1$, and so $h = 0$. 
Some basic facts about $\ell_p$ norms

- $\|h_\Gamma\|_{\infty} \leq \|h_\Gamma\|_{2} \leq \|h_\Gamma\|_{1}$

- $\|h_\Gamma\|_{1} \leq \sqrt{|\Gamma|} \cdot \|h_\Gamma\|_{2}$

- $\|h_\Gamma\|_{2} \leq \sqrt{|\Gamma|} \cdot \|h_\Gamma\|_{\infty}$
Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

\[ \Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \]
\[ \Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \]
\[ \vdots \]
Dividing up $h_{\Gamma_0^c}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

$$\Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c},$$
$$\Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c},$$
$$\vdots$$

Then

$$0 = \|\Phi h\|_2 = \|\Phi \left( \sum_{j \geq 0} h_{\Gamma_j} \right)\|_2 \geq \|\Phi (h_{\Gamma_0} + h_{\Gamma_1})\|_2 - \| \sum_{j \geq 2} \Phi h_{\Gamma_j} \|_2$$
Dividing up $h_{\Gamma_0^c}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

$$\Gamma_1 = \text{locations of 2}\mathcal{S} \text{ largest terms in } h_{\Gamma_0^c},$$

$$\Gamma_2 = \text{locations next 2}\mathcal{S} \text{ largest terms in } h_{\Gamma_0^c},$$

$$\vdots$$

Then

$$0 = \| \Phi h \|_2 = \| \Phi \left( \sum_{j \geq 0} h_{\Gamma_j} \right) \|_2 \geq \| \Phi (h_{\Gamma_0} + h_{\Gamma_1}) \|_2 - \| \sum_{j \geq 2} \Phi h_{\Gamma_j} \|_2$$

$$\geq \| \Phi (h_{\Gamma_0} + h_{\Gamma_1}) \|_2 - \sum_{j \geq 2} \| \Phi h_{\Gamma_j} \|_2$$
Dividing up \( h_{\Gamma_0^c} \)

Recall that \( \Gamma_0 \) is the support of \( x_0 \)

Fix \( h \in \text{Null}(\Phi) \). Let

\[
\Gamma_1 = \text{locations of 2S largest terms in } h_{\Gamma_0^c}, \\
\Gamma_2 = \text{locations next 2S largest terms in } h_{\Gamma_0^c}, \\
\vdots
\]

Then

\[
\| \Phi(h_{\Gamma_0} + h_{\Gamma_1}) \|_2 \leq \sum_{j \geq 2} \| \Phi h_{\Gamma_j} \|_2
\]
Dividing up $h_{\Gamma_0^c}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

\[ \Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c}, \]
\[ \Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c}, \]
\[ \vdots \]

Applying the $3S$-RIP gives

\[
\sqrt{1 - \delta_{3S}} \| h_{\Gamma_0} + h_{\Gamma_1} \|_2 \leq \| \Phi(h_{\Gamma_0} + h_{\Gamma_1}) \|_2 \\
\leq \sum_{j \geq 2} \| \Phi h_{\Gamma_j} \|_2 \leq \sum_{j \geq 2} \sqrt{1 + \delta_{2S}} \| h_{\Gamma_j} \|_2
\]
Dividing up $h_{\Gamma^c_0}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

$$
\Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma^c_0},
\Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma^c_0},
\vdots
$$

Applying the $3S$-RIP gives

$$
\|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \leq \sqrt{\frac{1 + \delta_{2S}}{1 - \delta_{3S}}} \sum_{j \geq 2} \|h_{\Gamma_j}\|_2
$$
Dividing up $h_{\Gamma_0^c}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

\[\Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c},\]
\[\Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c},\]
\[\vdots\]

Then

\[\|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \leq \sqrt{\frac{1 + \delta_{2S}}{1 - \delta_{3S}}} \sum_{j \geq 2} \sqrt{2S}\|h_{\Gamma_j}\|_\infty\]

since \[\|h_{\Gamma_j}\|_2 \leq \sqrt{2S}\|h_{\Gamma_j}\|_\infty\]
Dividing up $h_{\Gamma_0}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

\[ \Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0}^c, \]
\[ \Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0}^c, \]
\[ \vdots \]

Then

\[ \|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \leq \sqrt{\frac{1 + \delta_{2S}}{1 - \delta_{3S}}} \sum_{j \geq 1} \frac{1}{\sqrt{2S}} \|h_{\Gamma_j}\|_1 \]

since $\|h_{\Gamma_j}\|_{\infty} \leq \frac{1}{2S} \|h_{\Gamma_{j-1}}\|_1$
Dividing up $h_{\Gamma_0^c}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

$$\Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c},$$
$$\Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c},$$
$$\vdots$$

Which means

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \leq \sqrt{\frac{1 + \delta_{2S}}{1 - \delta_{3S}}} \frac{\|h_{\Gamma_0^c}\|_1}{\sqrt{2S}}$$
Dividing up $h_{\Gamma_0}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

$$
\Gamma_1 = \text{locations of 2S largest terms in } h_{\Gamma_0},

\Gamma_2 = \text{locations next 2S largest terms in } h_{\Gamma_0},

\vdots
$$

Working to the left

$$
\|h_{\Gamma_0}\|_2 \leq \|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \leq \sqrt{\frac{1 + \delta_{2S}}{1 - \delta_{3S}}} \frac{\|h_{\Gamma_0}\|_1}{\sqrt{2S}}
$$
Dividing up $h_{\Gamma_0^c}$

Recall that $\Gamma_0$ is the support of $x_0$

Fix $h \in \text{Null}(\Phi)$. Let

$$\Gamma_1 = \text{locations of } 2S \text{ largest terms in } h_{\Gamma_0^c},$$
$$\Gamma_2 = \text{locations next } 2S \text{ largest terms in } h_{\Gamma_0^c},$$
$$\vdots$$

Working to the left

$$\frac{\|h_{\Gamma_0}\|_1}{\sqrt{S}} \leq \|h_{\Gamma_0}\|_2 \leq \|h_{\Gamma_0} + h_{\Gamma_1}\|_2 \leq \sqrt{1 + \frac{\delta_{2S}}{1 - \delta_{3S}}} \frac{\|h_{\Gamma_0^c}\|_1}{\sqrt{2S}}$$
Wrapping it up

We have shown

\[ \| h_{\Gamma_0} \|_1 \leq \sqrt{\frac{1 + \delta_{2S}}{1 - \delta_{3S}}} \sqrt{\frac{S}{2S}} \| h_{\Gamma_0^c} \|_1 \]

\[ = \rho \| h_{\Gamma_0^c} \|_1 \]

for

\[ \rho = \sqrt{\frac{1 + \delta_{2S}}{2(1 - \delta_{3S})}} \]
Wrapping it up

We have shown

\[ \|h_{\Gamma_0}\|_1 \leq \sqrt{\frac{1 + \delta_{2S}}{1 - \delta_{3S}}} \sqrt{\frac{S}{2S}} \|h_{\Gamma_0^c}\|_1 \]

\[ = \rho \|h_{\Gamma_0^c}\|_1 \]

for

\[ \rho = \sqrt{\frac{1 + \delta_{2S}}{2(1 - \delta_{3S})}} \]

Taking \( \delta_{2S} \leq \delta_{3S} < 1/3 \) \( \Rightarrow \) \( \rho < 1 \).
Theorem: Let $\Phi$ be an $M \times N$ matrix that is an approximate isometry for $3S$-sparse vectors. Let $x_0$ be an $S$-sparse vector, and suppose we observe $y = \Phi x_0$. Given $y$, the solution to

$$\min_x \|x\|_1 \quad \text{subject to} \quad \Phi x = y$$

is exactly $x_0$.  

SUCCESS!!