Stable recovery

So far, our analysis of ℓ_1 minimization has been based on recovery signal which are **exactly sparse** and measurements that are **free of noise**.

We will now look at how ℓ_1 recovery performs in more realistic situations.

We will not be able to recover x_0 exactly; we bound the **recovery error**

$$h = x^{\#} - x_0$$

We will have three essential results:

1. when x_0 is not sparse, we bound $\|h\|_1$
2. when x_0 is not sparse, we bound $\|h\|_2$
3. when x_0 is sparse, but there is error in the measurements

It also easy to combine (3) with (1) and (2).

Not necessarily sparse signals, ℓ_1 recovery error

Here is the result we will establish:

Let x_0 be an arbitrary vector, and let

$$x_{0,S} \;=\; \text{the best } S\text{-term approximation of } x_0$$

$x_{0,S}$ simply contains the S largest entries in x_0

Let $y = \Phi x_0$ be the observations, and let $x^\sharp$ be the solution to

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

We will show that there exists a constant C_δ such that

$$\|x^\sharp - x_0\|_1 \;\leq\; C_\delta \|x_0 - x_{0,S}\|_1$$

where C_δ depends on the RIP constants δ for Φ

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\Gamma_0 = \text{locations of } S \text{ largest terms in } x_0$$
$$h = x^\sharp - x_0$$

Then

$$\|x_0 + h\|_1 = \sum_{\gamma \in \Gamma_0} |x_0(\gamma) + h(\gamma)| + \sum_{\gamma' \in \Gamma_0^c} |x_0(\gamma') + h(\gamma')|$$

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\begin{aligned}\Gamma_0 &= \text{locations of } S \text{ largest terms in } x_0 \\ h &= x^\sharp - x_0\end{aligned}$$

Then

$$\begin{aligned}\|x_0 + h\|_1 &= \sum_{\gamma \in \Gamma_0} |x_0(\gamma) + h(\gamma)| + \sum_{\gamma' \in \Gamma_0^c} |x_0(\gamma') + h(\gamma')| \\ &\geq \sum_{\gamma \in \Gamma_0} |x_0(\gamma)| - |h(\gamma)| + \sum_{\gamma' \in \Gamma_0^c} |h(\gamma')| - |x_0(\gamma')|\end{aligned}$$

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\begin{aligned} \Gamma_0 &= \text{locations of } S \text{ largest terms in } x_0 \\ h &= x^\sharp - x_0 \end{aligned}$$

Then

$$\begin{aligned} \|x_0 + h\|_1 &= \sum_{\gamma \in \Gamma_0} |x_0(\gamma) + h(\gamma)| + \sum_{\gamma' \in \Gamma_0^c} |x_0(\gamma') + h(\gamma')| \\ &\geq \sum_{\gamma \in \Gamma_0} |x_0(\gamma)| - |h(\gamma)| + \sum_{\gamma' \in \Gamma_0^c} |h(\gamma')| - |x_0(\gamma')| \\ &= \|x_0\|_1 + \|h_{\Gamma_0^c}\|_1 - \|h_{\Gamma_0}\|_1 - 2 \sum_{\gamma' \in \Gamma_0^c} |x_0(\gamma')| \end{aligned}$$

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\begin{aligned} \Gamma_0 &= \text{locations of } S \text{ largest terms in } x_0 \\ h &= x^\sharp - x_0 \end{aligned}$$

Then

$$\begin{aligned} \|x_0 + h\|_1 &= \sum_{\gamma \in \Gamma_0} |x_0(\gamma) + h(\gamma)| + \sum_{\gamma' \in \Gamma_0^c} |x_0(\gamma') + h(\gamma')| \\ &\geq \sum_{\gamma \in \Gamma_0} |x_0(\gamma)| - |h(\gamma)| + \sum_{\gamma' \in \Gamma_0^c} |h(\gamma')| - |x_0(\gamma')| \\ &= \|x_0\|_1 + \|h_{\Gamma_0^c}\|_1 - \|h_{\Gamma_0}\|_1 - 2\|x - x_{0,s}\|_1 \end{aligned}$$

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\begin{aligned}\Gamma_0 &= \text{locations of } S \text{ largest terms in } x_0 \\ h &= x^\sharp - x_0\end{aligned}$$

Then

$$\|x_0 + h\|_1 \geq \|x_0\|_1 + \|h_{\Gamma_0^c}\|_1 - \|h_{\Gamma_0}\|_1 - 2\|x - x_{0,s}\|_1$$

Since $\|x_0 + h\|_1 - \|x_0\|_1 \leq 0$,

$$\|h_{\Gamma_0^c}\|_1 \ \leq \ \|h_{\Gamma_0}\|_1 + 2\|x_0 - x_{0,S}\|_1$$

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\Gamma_0 = \text{locations of } S \text{ largest terms in } x_0$$
$$h = x^\sharp - x_0$$

$$\|h_{\Gamma_0^c}\|_1 \ \leq \ \|h_{\Gamma_0}\|_1 + 2\|x_0 - x_{0,S}\|_1$$

We have seen before that if Φ obeys a $2s$-RIP, then $\exists\ \rho < 1$ such that

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

Combining this with the above ...

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\Gamma_0 = \text{locations of } S \text{ largest terms in } x_0$$
$$h = x^\sharp - x_0$$

Two main facts about the descent vector:

$$\|h_{\Gamma_0}\|_1 \leq \rho \|h_{\Gamma_0^c}\|_1$$
$$\|h_{\Gamma_0^c}\|_1 \leq \frac{2}{1-\rho} \|x_0 - x_{0,S}\|_1$$

(Modified) cone constraint

We follow the same line of argumentation as in the perfect recovery case. The main difference is that we need to modify the *cone condition*.

Take

$$\Gamma_0 = \text{locations of } S \text{ largest terms in } x_0$$
$$h = x^\sharp - x_0$$

Two main facts about the descent vector:

$$\|h_{\Gamma_0}\|_1 \le \rho \|h_{\Gamma_0^c}\|_1$$
$$\|h_{\Gamma_0^c}\|_1 \le \frac{2}{1-\rho}\|x_0 - x_{0,S}\|_1$$

and so

$$\begin{aligned}\|x^\sharp - x_0\|_1 = \|h\|_1 &= \|h_{\Gamma_0}\|_1 + \|h_{\Gamma_0^c}\|_1 \\ &\le (1+\rho)\|h_{\Gamma_0^c}\|_1 \\ &\le \frac{2(1+\rho)}{1-\rho}\|x_0 - x_{0,S}\|_1\end{aligned}$$

Not necessarily sparse signals, ℓ_1 recovery error

Here is the result we will establish:

Let x_0 be an arbitrary vector,
let $y = \Phi x_0$ be the observations,
let $x^\sharp$ be the solution to

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

Then

$$\|x^\sharp - x_0\|_1 \;\leq\; \frac{2(1+\rho)}{1-\rho}\|x_0 - x_{0,S}\|_1, \qquad \rho = \sqrt{\frac{1+\delta_{2S}}{2(1-\delta_{3S})}}$$

Summary: The *recovery error* (measured in the ℓ_1 norm) is the same (to within a constant) as the *approximation error* (measured in the ℓ_1 norm)

Not necessarily sparse signals, ℓ_2 recovery error

We can also bound the ℓ_2 norm of the recovery error:

$$\|h\|_2 = \|x^\sharp - x_0\|_2$$

Start by dividing the (squared error) into two parts:

$$\|h\|_2^2 = \|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 + \|h_{(\Gamma_0 \cup \Gamma_1)^c}\|_2^2$$

We have see before that since $h \in \mathrm{Null}(\Phi)$,

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 \ \leq \ \frac{\rho^2}{2S}\|h_{\Gamma_0^c}\|_1^2$$

Not necessarily sparse signals, ℓ_2 recovery error

The recovery error has two parts:

$$\|x^\sharp - x_0\|^2 = \|h\|_2^2 = \|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 + \|h_{(\Gamma_0 \cup \Gamma_1)^c}\|_2^2$$

For the first part:

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 \;\leq\; \frac{\rho^2}{2S}\|h_{\Gamma_0^c}\|_1^2,$$

for the second part, note that the kth largest term in $h_{\Gamma_0^c}$ obeys

$$|h_{\Gamma_0^c}|_{(k)} \quad \leq \quad \frac{1}{k}\sum_{j=1}^{k}|h_{\Gamma_0^c}|_{(j)} \quad \leq \frac{1}{k}\|h_{\Gamma_0^c}\|_1$$

kth largest term $\leq$ average of k largest terms $\leq \frac{1}{k}$sum of all terms

Not necessarily sparse signals, ℓ_2 recovery error

The recovery error has two parts:

$$\|x^\sharp - x_0\|^2 = \|h\|_2^2 = \|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 + \|h_{(\Gamma_0 \cup \Gamma_1)^c}\|_2^2$$

For the first part:

$$\|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 \;\leq\; \frac{\rho^2}{2S}\|h_{\Gamma_0^c}\|_1^2,$$

for the second part, note that

$$|h_{\Gamma_0^c}|_{(k)} \leq \frac{1}{k}\|h_{\Gamma_0^c}\|_1$$

and so

$$\begin{aligned}\|h_{(\Gamma_0 \cup \Gamma_1)^c}\|_2^2 &= \sum_{k=2S+1}^{N} |h_{\Gamma_0^c}|_{(k)}^2 \\ &\leq \sum_{k=2S+1}^{N} \|h_{\Gamma_0^c}\|_1^2 / k^2 \leq \frac{\|h_{\Gamma_0^c}\|_1^2}{2S}\end{aligned}$$

since $\sum_{k \geq 2S+1} 1/k^2 \leq \frac{1}{2S}$

Not necessarily sparse signals, ℓ_2 recovery error

The recovery error has two parts:

$$\|x^\sharp - x_0\|^2 = \|h\|_2^2 = \|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 + \|h_{(\Gamma_0 \cup \Gamma_1)^c}\|_2^2$$

$$\leq \frac{1+\rho^2}{2S} \|h_{\Gamma_0^c}\|_1^2$$

Not necessarily sparse signals, ℓ_2 recovery error

The recovery error has two parts:

$$\|x^\sharp - x_0\|^2 = \|h\|_2^2 = \|h_{\Gamma_0} + h_{\Gamma_1}\|_2^2 + \|h_{(\Gamma_0 \cup \Gamma_1)^c}\|_2^2$$

$$\begin{aligned} &\le \frac{1+\rho^2}{2S}\|h_{\Gamma_0^c}\|_1^2 \\ &\le \frac{2(1+\rho^2)}{(1-\rho)^2}\frac{\|x_0 - x_{0,S}\|_1^2}{S} \end{aligned}$$

Not necessarily sparse signals, ℓ_2 recovery error

Here is the result we have established:

Let x_0 be an arbitrary vector,
let $y = \Phi x_0$ be the observations,
let $x^\sharp$ be the solution to

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

Then

$$\|x^\sharp - x_0\|_2 \;\leq\; C_\delta \frac{\|x_0 - x_{0,S}\|_1}{\sqrt{S}}$$

where

$$C_\delta = \frac{2(1+\rho^2)}{(1-\rho)^2} \qquad \rho = \sqrt{\frac{1+\delta_{2S}}{2(1-\delta_{3S})}}$$

Summary: The *recovery error* (measured in the ℓ_2 norm) scales like the *approximation error* (measured in the ℓ_1 norm) scaled by $S^{-1/2}$

Stability in the presence of noise

Suppose now we observe corrupted measurements of an S-sparse signal

$$y = \Phi x_0 + e$$

where e is an arbitrary perturbation with $\|e\|_2 \leq \epsilon$.

Given y, we solve

$$\min_x \ \|x\|_1 \quad \text{subject to} \quad \|\Phi x - y\|_2 \leq \epsilon$$

and call the solution $x^\sharp$. Then there exists a small constant C_δ such that

$$\|x^\sharp - x_0\|_2 \ \leq \ C_\delta \cdot \epsilon.$$

(See the supplemental notes for a proof.)