Stable Recovery

Not only does ℓ_1 minimization recover signals which are exactly sparse perfectly, it also recovers good approximations to signals which are approximately sparse.

Suppose that x_0 is an arbitrary signal.
Let

$$x_{0,S} = \text{best } S\text{-term approximation to } x_0$$

($x_{0,S}$ contains the S largest magnitude terms from x_0 and is zero elsewhere.)

Let x^* be the solution to

$$\min_x \|x\|_{\ell_1} \quad \text{s.t.} \quad \Phi x = y \qquad (P1)$$

Since x_0 is not exactly sparse, $x^* \neq x_0$.
But we will show that

$$\|x^* - x_0\|_{\ell_1} \leq C_\delta \cdot \|x_0 - x_{0,S}\|_{\ell_1}$$

constant depends on RIP constants

approximation error

Given what we have done before, this is actually pretty straightforward. The key is modifying the (cone) condition for the case where x_0 is not exactly sparse.

Take

$$T_0 = \text{locations of } S \text{ largest terms in } x_0$$

$$h = x^* - x_0 \quad \text{(as before)}.$$

Then

$$\|x_0 + h\|_{\ell_1} = \sum_{i \in T_0} |x_0[i] + h[i]| + \sum_{i \in T_0^c} |x_0[i] + h[i]|$$

$$\geq \sum_{i \in T_0} |x_0[i]| - |h[i]| + \sum_{i \in T_0^c} |h[i]| - |x_0[i]|$$

$$= \|x_0\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} - \|h_{T_0}\|_{\ell_1} - 2 \underbrace{\|x_{0,T_0^c}\|_{\ell_1}}_{= \|x_0 - x_{0,S}\|_{\ell_1}}$$

Since $\|x_0 + h\|_{\ell_1} - \|x_0\|_{\ell_1} \leq 0$

$$\Rightarrow \quad \|h_{T_0^c}\|_{\ell_1} \leq \|h_{T_0}\|_{\ell_1} + 2\|x_0 - x_{0,S}\|_{\ell_1}$$

(cone 2)

We have seen before that if Φ obeys a 2S-RIP, then $\exists\ \rho < 1$ (that depends only on the isometry constants) such that

$$\|h_{T_0}\|_{\ell_1} \leq \rho\, \|h_{T_0^c}\|_{\ell_1}$$

Thus

$$\boxed{\|h_{T_0^c}\|_{\ell_1} \leq \frac{2}{1-\rho} \cdot \|x_0 - x_{0,S}\|_{\ell_1}}$$

and

$$\begin{aligned}
\|x^* - x_0\|_{\ell_1} = \|h\|_{\ell_1} &= \|h_{T_0}\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} \\
&\leq (1+\rho) \cdot \|h_{T_0^c}\|_{\ell_1} \\
&\leq \frac{2(1+\rho)}{1-\rho} \cdot \|x_0 - x_{0,S}\|_{\ell_1}
\end{aligned}$$

So the <u>recovery error</u> (measured in the ℓ_1 norm) is the same, to within a constant, as the <u>approximation error</u> (measured in the ℓ_1 norm).

Stable Recovery – ℓ_2 norm

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

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

We divide the (squared) error into two parts:

$$\|h\|_2^2 = \|h_{T_0 \cup T_1}\|_2^2 + \|h_{(T_0 \cup T_1)^c}\|_2^2$$

Our work from before (e.g. box on p. 58 w/ $S = 2S = |T_j|$, $j \geq 1$) tells us that

$$\|h_{T_0 \cup T_1}\|_2^2 \leq \frac{\rho^2}{2S} \|h_{T_0^c}\|_{\ell_1}^2$$

For the second term, we use the fact that the k^{th} largest value in $h_{T_0^c}$ obeys

$$\underbrace{|h_{T_0^c}|_{(k)}}_{k^{th}\text{ largest value}} \leq \underbrace{\frac{1}{k}\sum_{j=1}^{k} |h_{T_0^c}|_{(j)}}_{\text{average of } k \text{ largest values}} \leq \underbrace{\frac{\|h_{T_0^c}\|_{\ell_1}}{k}}_{\text{since } \|h_{T_0^c}\|_{\ell_1} = \sum_{j \geq 1} |h_{T_0^c}|_{(j)}}$$

As a result

$$\|h_{(T_0 \cup T_1)^c}\|_2^2 = \sum_{k=2S+1}^{n} |h_{T_0^c}|_{(k)}^2$$

$$\leq \sum_{k=2S+1}^{n} \|h_{T_0^c}\|_{\ell_1}^2 / k^2$$

$$\leq \frac{\|h_{T_0^c}\|_{\ell_1}^2}{2S} \qquad \text{since } \sum_{k=2S+1}^{n} \frac{1}{k^2} \leq \frac{1}{2S}$$

Thus

$$\|h\|_2^2 \leq \frac{\rho^2}{2S} \|h_{T_0^c}\|_{\ell_1}^2 + \frac{1}{2S} \|h_{T_0^c}\|_{\ell_1}^2$$

$$= \frac{1+\rho^2}{2S} \|h_{T_0^c}\|_{\ell_1}^2$$

$$\leq \frac{2(1+\rho^2)}{(1-\rho)^2} \frac{\|x_0 - x_{0,S}\|_{\ell_1}^2}{S} \qquad \text{using box on page 62}$$

so

$$\boxed{\|h\|_2 \leq C \cdot \frac{\|x_0 - x_{0,S}\|_{\ell_1}}{\sqrt{S}}}$$

this assumed Φ obeyed a 3S-RIP. But again, the argument can be tweaked for a 2S-RIP

$$C = \sqrt{\frac{2(1+\rho^2)}{(1-\rho)^2}}$$

Robust recovery in the presence of noise

We now consider the case where our observations are corrupted by noise.

$$y = \Phi x_0 + e$$

e is an arbitrary noise vector with $\|e\|_2 \le \epsilon$.

To recover x_0, we solve

$$\min_x \|x\|_{\ell_1} \quad \text{s.t.} \quad \|\Phi x - y\|_2 \le \epsilon$$

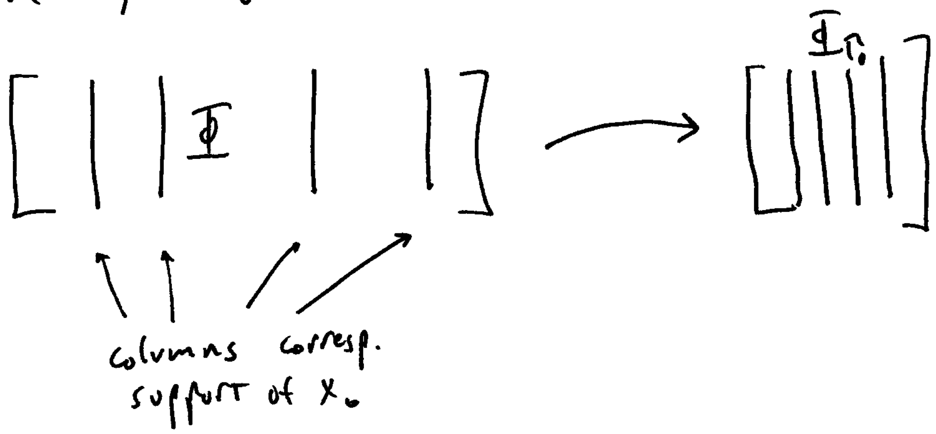
rather than $\Phi x = y$, we as that Φx be within ϵ of y

For now, we will just assume x_0 is perfectly sparse and supported on T_0 — we can extend what's below in the same manner as before for non-sparse x_0.

What is the best we can hope for in this situation?

Say that an oracle makes our life easy by telling us the support T_0.

Let Φ_{T_0} = $m \times S$ matrix of columns from Φ indexed by T_0:



Also, let x_{T_0} be the S-vector consisting of the "active" components of x_0. Note that

$$\Phi x_0 = \Phi_{T_0} x_{T_0}$$

So now

$$y = \underbrace{\Phi_{T_0}}_{m \times S \text{ system}} x_{T_0} + e$$

Since $m > S$, this is an overdetermined system of linear equations.

Not knowing anything else, a good way to recover x_{T_0} from y is using <u>least-squares</u>.

The LS solution is

$$x^*_{orc} = (\Phi^T_{\Gamma_0}\Phi_{\Gamma_0})^{-1}\Phi^T_{\Gamma_0} y$$

and the recovery error is

$$\|x^*_{orc} - x_0\|_2 = \|(\Phi^T_{\Gamma_0}\Phi_{\Gamma_0})^{-1}\Phi^T_{\Gamma_0} e\|_2$$

If the error e is chosen adversarialy (w.r.t. Φ), we can make

$$\|x^*_{orc} - x_0\|_2 = \frac{1}{\lambda_{min}(\Phi^T_{\Gamma_0}\Phi_{\Gamma_0})}\|e\|_2$$

smallest eigenvalue of $\Phi^T_{\Gamma_0}\Phi_{\Gamma_0}$

If Φ obeys an S-RIP, then

$$\|x^*_{orc} - x_0\|_2 \leq \frac{1}{1-\delta_S}\|e\|_2$$

$$\leq \text{Const} \cdot \epsilon$$

So if we are interested in protecting against arbitrary errors, a recovery error of $\sim C \cdot \epsilon$ is the best we can hope for.

Now, what can we say about ℓ_1 recovery?

Call x^* the solution to

$$\min \|x\|_{\ell_1} \quad \text{s.t.} \quad \|\Phi x - y\|_2 \le \varepsilon$$

We can bound the recovery error $\|x^* - x_0\|_2 = \|h\|_2$ using two facts:

① $\|x^*\|_{\ell_1} \le \|x_0\|_{\ell_1}$

This essentially means (and we will see this below) that

$$\|\Phi h\|_2 \ge C'_\delta \cdot \|h\|_2$$

constant depending only on RIP constants

i.e. not only is h not in the null space of Φ, but there is a uniform lower bound on $\frac{\|\Phi h\|_2}{\|h\|_2}$.

(2) x_0 is feasible, since $\|y - \Phi x_0\|_2 = \|e\|_2 \leq \varepsilon$

Then the triangle inequality tells us

$$
\begin{aligned}
\|\Phi h\|_2 &= \|\Phi(x^* - x_0)\|_2 \\
&= \|\Phi x^* - y - \Phi x_0 + y\|_2 \\
&\leq \|\Phi x^* - y\|_2 + \|\Phi x_0 - y\|_2 \\
&\leq 2\varepsilon
\end{aligned}
$$

Combining these facts

$$
C'_\delta \cdot \|h\|_2 \leq \|\Phi h\|_2 \leq 2\varepsilon
$$

$$
\begin{aligned}
\Rightarrow \|h\|_2 &\leq \frac{2}{C'_\delta} \cdot \varepsilon \\
&= \text{Const} \cdot \varepsilon \qquad !
\end{aligned}
$$

It remains to verify that

$$
\|\Phi h\|_2 \geq C'_\delta \cdot \|h\|_2
$$

We start by writing

$$\|\Phi h\|_2 \geq \|\Phi h_{T_0 \cup T_1}\|_2 - \|\Phi h_{(T_0 \cup T_1)^c}\|_2$$

$$\geq \|\Phi h_{T_0 \cup T_1}\|_2 - \sum_{j \geq 2} \|\Phi h_{T_j}\|_2$$

$$\geq \sqrt{1-\delta_{s+s'}} \cdot \|h_{T_0 \cup T_1}\|_2 - \sqrt{1+\delta_{s'}} \cdot \sum_{j \geq 2} \|h_{T_j}\|_2$$

on page 58, we showed

since (cone) still holds

$$\sum_{j \geq 2} \|h_{T_j}\|_2 \leq \frac{\|h_{T_0^c}\|_{\ell_1}}{\sqrt{s'}} \leq \frac{\|h_{T_0}\|_{\ell_1}}{\sqrt{s'}}$$

$$\leq \|h_{T_0}\|_2 \cdot \frac{1}{\sqrt{2}}$$

if we take $s' = 2s$

$$\leq \|h_{T_0 \cup T_1}\|_2 \cdot \frac{1}{\sqrt{2}}$$

Thus

$$\|\Phi h\|_2 \geq \underbrace{\left(\sqrt{1-\delta_{3s}} - \frac{\sqrt{1+\delta_{2s}}}{\sqrt{2}}\right)}_{C_1} \cdot \|h_{T_0 \cup T_1}\|_2$$

Also on page 58, we showed

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq \frac{\rho}{\sqrt{2s}} \cdot \|h_{T_0^c}\|_{\ell_1} \leq \frac{\rho}{\sqrt{2}} \cdot \|h_{T_0 \cup T_1}\|_2$$

And so

$$\|h\|_2^2 = \|h_{T_0 \cup T_1}\|_2^2 + \|h_{(T_0 \cup T_1)^c}\|_2^2$$

$$\leq \left(1 + \frac{\rho^2}{2}\right) \cdot \|h_{T_0 \cup T_1}\|_2^2$$

$$\Rightarrow \|h\|_2 \leq \underbrace{\sqrt{1 + \frac{\rho^2}{2}}}_{C_2} \cdot \|h_{T_0 \cup T_1}\|_2$$

and we have

$$\|\Phi h\|_2 \geq C_1 \|h_{T_0 \cup T_1}\|_2 \geq \underbrace{\frac{C_1}{C_2}}_{C_8'} \cdot \|h\|_2$$

Thus, if Φ obeys a 3S-RIP and x_0 is S-sparse, the recovery error is $\sim \epsilon$

$$\|x^* - x_0\|_2 \leq \text{Const} \cdot \epsilon$$

Again, this can be weakened to a 2S-RIP with a slightly different argument.

If x_0 is arbitrary, then the recovery error obeys

$$\|x^* - x_0\|_2 \leq \text{Const} \cdot \left(\epsilon + \frac{\|x_0 - x_{0,S}\|_{\ell_1}}{\sqrt{S}} \right)$$

Proving this is a straightforward extension of the above.