

Stable Signal Recovery

Jacob Denson

September 25, 2020

Let us begin by introducing some notation:

- We let $[N] = \{1, \dots, N\}$.
- We view \mathbf{R}^N as the space of functions $f : [N] \rightarrow \mathbf{R}$.
- We write $A(n) \lesssim B(n)$ if there exists a universal constant $C > 0$ such that $A(n) \leq C \cdot B(n)$ for all $n \geq 1$. We write $A(n) \lesssim_K B(n)$ if there exists a universal constant $C > 0$ and $K > 0$ such that $A(n) \leq C \log(1 + n)^K B(n)$.
- For $0 < p \leq \infty$ we let

$$\|f\|_{L^p} = \left(\sum_{s \in S} |f(s)|^p \right)^{1/p}.$$

and set $\|f\|_{L^0} = \#(\text{supp}(f))$, where

$$\text{supp}(f) = \{s \in [N] : f(s) \neq 0\}.$$

More generally, given $E \subset [N]$, we let $\|f\|_{L^p(E)} = \|f \mathbf{I}_E\|_{L^p}$.

In signals processing, a common problem is to reconstruct a function f given limited data. We are interested in the following toy problem we call the *signal recovery problem*: Given a class of functions $\mathcal{F} \subset \mathbf{R}^N$, and a matrix $A : \mathbf{R}^N \rightarrow \mathbf{R}^M$, when is it possible to recover $f^0 \in \mathcal{F}$ from Af^0 , where $M \ll N$?

If $\mathcal{F} = \mathbf{R}^N$, the problem is underdetermined unless $M = N$. More generally, if \mathcal{F} is a K dimensional subspace of \mathbf{R}^N , then we must have $M \geq K$. More generally, the signal recovery problem is impossible if there exists distinct $f_1, f_2 \in \mathcal{F}$ such that $Af_1 = Af_2$. Of course, this condition is actually necessary to determine if the problem is unsolvable. Of course, even if this problem is ‘solvable’ in a mathematical sense, we still require our solutions to be practically solvable, i.e. computationally efficient, which adds additional constraints.

An often useful assumption is that f^0 is *sparse*. Given $0 \leq S \leq N$ we let

$$\mathcal{F}_S = \{f \in \mathbf{R}^N : \|f\|_{L^0} \leq S\}.$$

We say f is S -sparse if $f \in \mathcal{F}_S$. If $M < 2S$, we can find two functions $f_1, f_2 \in \mathcal{F}_S$ such that $Af_1 = Af_2$, so we must assume $M \geq 2S$. But we might hope that

$M \gtrsim 2S$ is sufficient to recover sparse data. Indeed, a necessary condition that such a problem is solvable is that there exists a nonzero $f \in \mathcal{F}_{2S}$ such that $Af = 0$. One assumption guaranteeing this to be true is that A is a *restricted approximate isometry*. Recall that an approximate isometry $A : \mathbf{R}^N \rightarrow \mathbf{R}^M$ is a map such that for all $x \in \mathbf{R}^N$,

$$(1 - \delta)\|x\|_{L^2} \leq \|Ax\|_{L^2} \leq (1 + \delta)\|x\|_{L^2}. \quad (1)$$

for some $\delta > 0$. A *restricted approximate isometry* is a map $A : \mathbf{R}^N \rightarrow \mathbf{R}^M$ satisfying 1 for all $x \in \mathcal{F}_S$. We let δ_S be the optimal constant in 1 for A . Clearly, if $\delta_{2S} < 1$, the signal recovery problem is mathematically solvable, since if $Af = 0$ and $f \in \mathcal{F}_{2S}$, $f = 0$.

Thus assuming $\delta_{2S} < 1$, given $g = Af^0$, we have

$$f^0 = \operatorname{argmin} \{\|f\|_{L^0} : Af = g\}.$$

But computing the minimum on the right is NP hard. We note that for any $f \in \mathbf{R}^N$,

$$\lim_{p \rightarrow 0} \|f\|_p^p = \|f\|_0,$$

so one might expect the values $\|f\|_{L^p}$ for small p to model the case $\|\cdot\|_{L^0}$. For any $p \geq 1$, the function $\|\cdot\|_p$ is convex, so in particular, we can use convex optimization to compute

$$f^* = \operatorname{argmin} \{\|f\|_{L^1} : Af = g\}$$

in a computationally efficient manner. In particular, a reduction to linear programming can be used to solve this problem in $O(N^{2+\varepsilon_0})$ time, where $\varepsilon_0 = 1/18$.

Another reason to expect f^* to be close to sparse is because for each $R > 0$, the set

$$B_{L^1}(R) = \{f \in \mathbf{R}^N : \|f\|_{L^1} \leq R\}$$

consists of many sharp corners lying on sparse coordinate points. An optimum solution to the optimization problem is therefore likely to be close to one of these corners.

In fact, under the slightly stronger condition that $\delta_S + \delta_{2S} + \delta_{3S} < 1$, we find $f^* = f^0$ (proved in Candes and Tao, *Decoding by Linear Programming*).

The restricted isometry property might seem quite restrictive but, in fact, by choosing *random matrices*, it is often quite easy to find matrices satisfying the restricted isometry property:

- s

We now ask how robust this result is. In particular:

- Suppose only that f^0 is only *approximately sparse*. For each $S \geq 0$, let f_S^0 be the projection in L^1 norm onto \mathcal{F}_S , i.e. obtained by keeping the S values s where $|f(s)|$ is largest. Instead of suppose f^0 is S -sparse, we only assume $\|f^0 - f_S^0\|_{L^1}$ is small.

- Suppose Af is corrupted by noise, i.e. we are given g such that $\|Af_0 - g\|_{L^2} \leq \varepsilon$ rather than Af_0 itself.

Despite how ill-posed this problem is, we are still able to recover f in a stable manner under noise. In particular, if $\delta_{3S} + 3\delta_{4S} < 2$, and we define

$$f^* = \operatorname{argmin}\{\|f\|_{L^1} : \|Af - g\|_{L^2} \leq \varepsilon\},$$

then $\|f - f^*\| \lesssim_{\delta_{4S}} \varepsilon$. In particular, if $\varepsilon = 0$, we can recover the signal exactly.