**CS 395T: Sublinear Algorithms** Fall 2016

Lecture: 16 – $L^1$ minimization, Oct 20, 2016

*Prof. Eric Price* *Scribe: David Liau and Ger Yang*

# 1 Overview

In the last lecture: *Model-based Compressed Sensing* .

In this lecture: $L^1$ *minimization*.

# 2 $L^1$ minimization

Given

$$A \in R^{m \times n} \quad with \quad \big(O(k), 1\big) RIP$$

For a k-sparse vector $x$, define $y = Ax + e$. The problem is to find $\hat{x}$ with $\|\hat{x} - x\|_2 \lesssim \|e\|_2$. Now if we pick $\hat{x}$ as the solution to the following optimization problem,

$$\begin{aligned} \underset{\hat{x}}{\text{minimize}} \quad & \|\hat{x}\|_1 \\ \text{subject to} \quad & \|A\hat{x} - y\|_2 \leq \epsilon. \end{aligned}$$

We have the theorem which,

**Theorem 1.** *If* $\epsilon \geq \|e\|_2$ *then* $\|\hat{x} - x\|_2 \lesssim \epsilon$.

*Proof.* Lets set $z = \hat{x} - x$, by definition we have,

$$\begin{aligned} \|Az - e\|_2^2 = \|A\hat{x} - y\|_2^2 &\leq \epsilon^2 \\ \|Az\|_2^2 - 2(Az)^T e + \|e\|_2^2 &\leq \epsilon^2 \\ \|Az\|_2^2 = 2(Az)^T e - \|e\|_2^2 &+ \epsilon^2 \\ &\leq 2\|e\|_2\|Az\|_2 + \epsilon^2 - \|e\|_2^2 \\ \big(\|Az\|_2 - \|e\|_2\big)^2 &\leq \epsilon^2 \\ \|Az\|_2\|e\|_2 + \epsilon &\leq 2\epsilon \end{aligned}$$

Now we just need $\|z\|_2 \leq \|Az\|_2$ to complete the proof. Fortunately, the property is known to be true given by the Restricted Eigenvalue property discussed in the following section. Thus we have proved that by utilizing the $L^1$ minimization vector, it is possible to construct the sparse vector we desired. $\square$

### 2.1 Restricted Eigenvalue (RE)

For $S = \mathrm{supp}(x)$,

$$\begin{aligned}\|x_S\|_1 = \|x\|_1 &\geq \|\hat{x}\|_1 \\ &= \|x+z\|_1 \\ &\geq \|(x+z)_S\|_1 + \|z_{\bar{S}}\|_1 \\ &\geq \|x_S\|_1 + \|z_{\bar{S}}\|_1 - \|z_S\|_1\end{aligned}$$

Thus we have that

$$\|z_S\|_1 \geq \|z_{\bar{S}}\|_1.$$

The last equation states the following important property we would like to introduce in this section.

**Definition 1.** *Uniform Restricted Eigenvalue Condition (RE)*
*We said that the matrix $A$ satisfies the uniform restricted eigenvalue condition if for all $S$ and $z$ which $\|z_S\|_1 \geq \|z_{\bar{S}}\|_1$ holds, $\|Az\|_2 \gtrsim \|z\|_2$.*

### 2.2 RIP $\Rightarrow$ RE via Shelling Argument

In this section, we are going to show that RIP implies RE via a "shelling argument". More precisely, suppose $A$ satisfies the RIP of order $k$. We would like to show that for any $z$ and $S \subset [n]$ with $|S| \leq k$ such that $\|z_S\|_1 \geq \|z_{\overline{S}}\|_1$, we have $\|Az\|_2 \gtrsim \|z\|_2$.

The *shelling argument* works as follows. First, we split $z$ into blocks $z^1, z^2, \ldots$ of size $k$ in the order of decreasing magnitude, i.e. $z = z^1 + z^2 + \cdots + z^{n/k}$, and $\|z^1\|_2 \geq \|z^2\|_2 \geq \cdots \geq \|z^{n/k}\|_2$. Then, for $i \geq 2$ we have that

$$\frac{\|z^i\|_1}{\sqrt{k}} \leq \|z^i\|_2 \leq \frac{\|z^{i-1}\|_1}{\sqrt{k}} \tag{1}$$

Also, for $i \geq 3$ we have

$$\|z^i + z^{i+1}\|_2 \leq \frac{\|z^{i-1} + z^{i-2}\|_1}{\sqrt{2k}} \tag{2}$$

Next, we look at $\|Az\|_2$:

$$\begin{aligned}\|Az\|_2 &= \|A(z^1 + z^2 + \cdots + z^{n/k})\|_2 \\ &\geq \|A(z^1+z^2+z^3)\|_2 - \|A(z^4+z^5)\|_2 - \|A(z^6+z^7)\|_2 - \ldots \\ &\geq (1-\epsilon)\|z^1+z^2+z^3\|_2 - (1+\epsilon)\left(\sum_{i=2}^{\frac{n/k-1}{2}} \|z^{2i}+z^{2i+1}\|_2\right)\end{aligned} \tag{3}$$

where the first inequality is a result of triangle inequality, and the second one uses the RIP property.

Then, we can further bound the second term in (**??**) by

$$
\begin{aligned}
(1+\epsilon)\left(\sum_{i=2}^{\frac{n/k-1}{2}} \|z^{2i}+z^{2i+1}\|_2\right) &\le \frac{(1+\epsilon)}{\sqrt{2k}}\left(\sum_{i=2}^{\frac{n/k-1}{2}} \|z^{2i-1}+z^{2i-2}\|_1\right) \\
&= \frac{(1+\epsilon)}{\sqrt{2k}}\sum_{i=2}^{\frac{n/k-1}{2}} \left(\|z^{2i-1}\|_1+\|z^{2i-2}\|_1\right) \\
&\le \frac{(1+\epsilon)}{\sqrt{2k}}\left\|\sum_{i=2}^{n/k} z^i\right\|_1 \le \frac{(1+\epsilon)}{\sqrt{2k}}\left\|z^1\right\|_1
\end{aligned}
$$

where the first inequality uses (**??**). The equality in the second line and the first inequality in the third line follows the fact that each $z^i$ has distinct support. The last inequality is a result of the assumption that for any $|S| \le k$ such that $\|z_S\|_1 \ge \|z_{\overline{S}}\|_1$, we have $\|Az\|_2 \gtrsim \|z\|_2$. Then, we are able to get a bound of (**??**).

$$
\begin{aligned}
\|Az\|_2 &\ge (1-\epsilon)\|z^1+z^2+z^3\|_2 - (1+\epsilon)\left(\sum_{i=2}^{\frac{n/k-1}{2}} \|z^{2i}+z^{2i+1}\|_2\right) \\
&\ge (1-\epsilon)\|z^1\|_2 - \frac{(1+\epsilon)}{\sqrt{2k}}\|z^1\|_1 \ge (1-\epsilon)\|z^1\|_2 - \frac{(1+\epsilon)}{\sqrt{2}}\|z^1\|_2 \\
&\ge 0.1\|z^1\|_2
\end{aligned}
$$

for small enough $\epsilon$. Next, note that

$$
\|z^1\|_2 \ge \frac{\|z^1\|_1}{\sqrt{k}} \ge \frac{\|z^2+\cdots+z^{n/k}\|_1}{\sqrt{k}} = \frac{\sum_{i=2}^{n/k}\|z^i\|_1}{\sqrt{k}} \ge \sum_{i=3}^{n/k}\|z^i\|_2 \ge \left\|\sum_{i=3}^{n/k} z^i\right\|_2
$$

Hence we have $3\|z^1\|_2 \ge \|z\|_2$, and finally we have $\|Az\|_2 \gtrsim \|z\|_2$.

### 2.3 Recover a non-Sparse Vector

In Theorem **??**, we have shown how to recover a sparse $x$. Now, what if $x$ is not sparse? It turns out that we may use the similar technique. Denote $x_a$ as the vector that zero out all but $a$ largest items from $x$. We leave it as an exercise that showing the following property that suggests, essentially, we can solve the problem assuming that $x$ is a sparse vector:

**Exercise 1.** *If $A$ is a RIP matrix, then*

$$
\|A(x-x_{2k})\|_2 \le \frac{\|x-x_k\|_1}{\sqrt{k}}(1+\epsilon)
$$

This property suggests that we may use the technique is the previous lectures to get $\hat{x}$ from $y = Ax$ with $l2/l1$ guarantee

$$
\|\hat{x}-x\|_2 \lesssim \frac{1}{\sqrt{k}}\|x-x_k\|_1
$$

And after further sparsifying, we can get a solution $\hat{x}^*$ that satisfies the $l1/l1$ guarantee:

$$
\|\hat{x}^*-x\|_1 \lesssim \|x-x_k\|_1
$$

## 3 Comparison

In this section we provide a quick reference to the comparison between different recovery methods and their properties within the following table. The Count-min and Count-sketch results are based on the previous lectures.

| lable | $L^1$ minimization | Count-min | Count Sketch |
|---|---|---|---|
| Recovery time | $n^3$ | $n \log n$ | |
| Space | $k \log \frac{n}{k}$ | $k \log n$ | |
| Guarantee | $l_2/l_1$ | $l_\infty/l_1 \implies l_1/l_1$ | $l_\infty/l_2 \implies l_2/l_2$ |
| Notes | If $A$ satisfy RIP, works for all $x$ | For each $x$, $A$ works with high probability | |

Table 1: table of comparison

In class we also discussed a related theorem without a detailed proof. The theorem is stated as below and the proof will be given in the next lecture if possible.

**Theorem 2.** *We cannot get an algorithm with deterministic $l_2/l_2$ guarantee with $m = o(n)$.*