**CS 395T: Sublinear Algorithms** Fall 2014

Lecture: 12 – More compressed sensing, Oct 7, 2014

*Prof. Eric Price* *Scribe: Tianyang Li*

# 1 Overview

In the last lecture: *regular compressed sensing.*

In this lecture: *model-based compressed sensing.*

# 2 Compressed sensing

- $x$ is $k$-sparse
- observe $y = Ax + e$
- recover $\hat{x} \approx x$ where $\|\hat{x} - x\|_2 \lesssim \|e\|_2$

(or $x$ is "approximately" $k$-sparse and we recover $\hat{x}$ where $\|\hat{x} - x\|_2 \le \|e\|_2 + C \min_{k\text{-sparse } x'} \underbrace{\|x - x'\|}_{\text{various norm}}$ )

**Some notes about $A$**

- If $A \in \mathbb{R}^{m \times n}$ satisfies RIP, then recovery is possible.
- When each entry in $A$ is sampled from a Gaussian with mean 0 and variance 1, then $m = O(n \log \frac{n}{k})$ suffices.

**How good is this?**

- to store the positions of the entries: $\log \binom{n}{k} \approx k \log \frac{n}{k}$
- to store the values of the entries: $k$ words

Define "sparsity ratio" $R = \frac{n}{k}$.

Compressed sensing saves $\frac{R}{\log R}$ factor relative to naive sampling.

Storage saves approximately $R$ factor.

Can't use $O(k)$ measurements <u>*in general*</u>.

<u>*But*</u> can for more structured signals, e.g. *block-sparse* signals:

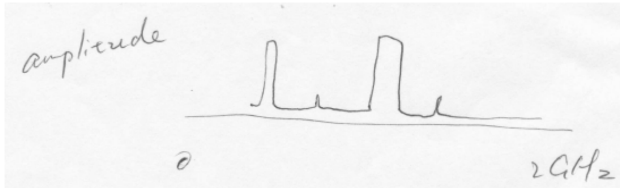


Figure 1: $\frac{k}{B}$ "blocks" of length $B$ where each block is all on/all off

For block-sparse signals, the number of support is $\binom{\frac{n}{B}}{\frac{k}{B}} = 2^{O(\frac{k}{B}\log\frac{n}{k})}$. When $B \geq \log\frac{n}{k}$ this is $2^{O(k)}$.

# 3 Tree sparsity

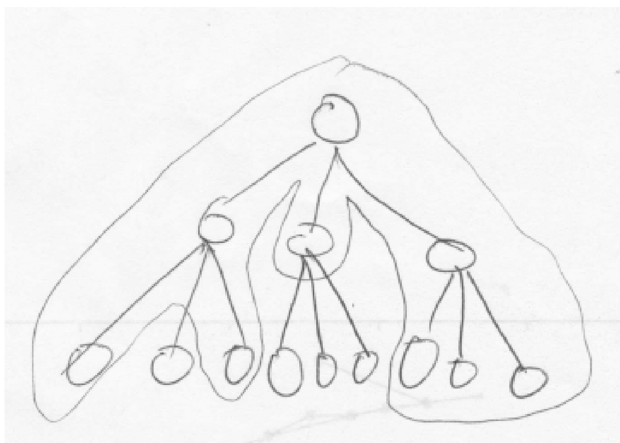

Figure 2: Sparsity pattern is contiguous rooted subtree

**Number of trees with $k$ terms in size $n$ binary tree**

- path that visits all vertices $\leq 2k$ edges
- at each vertex there are 3 possible directions to go

number of trees $\leq 3^{2k} = 2^{O(k)}$

# 4 Model sparsity

$\mathcal{F}$ is a family of supports, each $S$ in $\mathcal{F}$ satisfies $S \subseteq [n]$, $|S| \leq k$.

**Theorem 1.** *$m = O(k + \log|\mathcal{F}|)$ Gaussian measurements suffice.*

**Model based compressed sensing** Given $y = Ax + e$, $\text{supp}(x) \in \mathcal{F}$, recover $\hat{x}$ such that $\|\hat{x} - x\|_2 \lesssim \|e\|_2$

**Model RIP** $\forall x$ with $\text{supp}(x) \in \mathcal{F} \oplus \mathcal{F} = \{S \cup T | S, T \in \mathcal{F}\}$

$$\|Ax\|_2 = (1 \pm \epsilon)\|x\|_2 \tag{1}$$

**Model IHT**

$$x^{i+1} = H_{\mathcal{F}}(x^i + A^T(y - Ax^i)) \tag{2}$$

where

$$H_{\mathcal{F}} = \underset{T \in \mathcal{F}}{\arg\min} \|z_T\|_2 \tag{3}$$

(Before in *regular* compressed sensing if $A$ satisfies $2k$-RIP then IHT works.)

**First iteration analysis**

$$z = A^T y = A^T Ax + A^T e \tag{4}$$

$\forall T$ and $S = \text{supp}(x)$, if $A$ is model RIP on $\mathcal{F} \oplus \mathcal{F}$,

$$\begin{aligned} \|z - x_{S \cup T}\|_2 &\le \|A^T A - I\|_2 \|x\|_2 + \|A^T_{S \cup T \times [n]}\|_2 \|e\|_2 && (5)\\ &\le \epsilon \|x\|_2 + (1 + \epsilon)\|e\|_2 && (6) \end{aligned}$$

$\forall z$ and $\underbrace{T}_{\text{top } k \text{ of } z} \in \mathcal{F}$ we want

$$\|x - z_T\|_2 \lesssim \|(x - z)_{S \cup T}\|_2 \tag{7}$$

To prove (7):

$$\begin{aligned} \|x_{S \setminus T}\|_2 &\le \|(x - z)_{S \setminus T}\|_2 + \|z_{S \setminus T}\|_2 && (8)\\ &\le \|(x - z)_{S \setminus T}\|_2 + \|z_{T \setminus S}\|_2 && (9)\\ \Rightarrow \|x_{S \setminus T}\|_2^2 &\le 2\|(x - z)_{S \setminus T}\|_2^2 + 2\|z_{T \setminus S}\|_2^2 && (10) \end{aligned}$$

$$\begin{aligned} \|x - z_T\|_2^2 &= \|x_{s \setminus T}\|_2^2 + \|z_{T \setminus S}\|_2^2 + \|(x - z)_{T \cap S}\|_2^2 && (11)\\ &\le 2\|(x - z)_{S \setminus T}\|_2^2 + 3\|z^{T \setminus S}\|_2^2 + \|(x - z)^{T \cap S}\|_2^2 && (12)\\ &\le 3\|(x - z)_{S \cup T}\|_2^2 && (13) \end{aligned}$$

**Running time**

- regular IHT: $\log \frac{\|x\|_2}{\|e\|_2}$ (matrix vector multiplication time for $A$)
- model IHT: $\log \frac{\|x\|_2}{\|e\|_2}$ (matrix vector multiplication time for $A + H_{\mathcal{F}}$)

**Computing $H_{\mathcal{F}}$ for trees**

- exact: $O(nk^2)$, $O(nk)$
- approximate (find $T'$ such that $\|z_{T'}\|_2 \lesssim \min_T \|z_T\|_2$): $\tilde{O}(n)$

# 5 Compressed sensing using $L^1$ minimization

For

$$y = Ax + e \tag{14}$$

$$\min \|x\|_1 \tag{15}$$

given

$$\|A\hat{x} - y\|_2 \leq \epsilon \tag{16}$$

**Theorem 2.** *If $\epsilon \geq \|e\|_2$ and $A$ satisfies RIP or RE then $\|\hat{x} - x\|_2 \lesssim \epsilon$.*

## 5.1 Restricted Eigenvalue (RE)

**IHT fails for $A = 2I$**

$$\begin{aligned} z &= A^T Ax + A^T e && (17) \\ &= 4x + 2e && (18) \end{aligned}$$

**Definition 3.** *Restricted Eigenvalue (RE)*

$$\frac{\|Az\|_2}{\|z\|_2} \geq \epsilon \tag{19}$$

*whenever*

$$\begin{aligned} |S| &= k && (20) \\ \|z_S\|_1 &\geq \alpha \|z_{\bar{S}}\|_1 && (21) \end{aligned}$$

For example, $\epsilon = \frac{1}{10}$ and $\alpha = 1$.

*Proof.* (Theorem 2) Set $\epsilon = \|e\|_2$.

Let $z = \hat{x} - x$.

$$\begin{aligned} \|Az - e\|_2^2 &\leq \|e\|_2^2 \\ \|Az\|_2^2 - 2e^T Az + \|e\|_2^2 &\leq \|e\|_2^2 \\ &\Rightarrow \|Az\|_2 \leq 2\|e\|_2 \end{aligned}$$

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

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

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

RE $\Rightarrow \|z\|_2 \lesssim \|Az\|_2 \leq 2\|e\|_2$. □

### 5.2 RIP ⇒ RE

**"Shelling argument"** Suppose $A$ satisfies the RIP of order $2k$. We would like to show for any $z$ and $S \subset [n]$ of size $k$ with $\|z_S\|_1 \geq \|z_{\overline{S}}\|_1$ that $\|Az\|_2 \gtrsim \|z\|_2$.

Split $z$ into blocks $z^1, z^2, \ldots$ of decreasing magnitude, so $z^1$ has the largest $k$ coordinates, and each next $z^i$ has the next largest $2k$ coordinates. Then for $i \geq 3$ we have that

$$\frac{\|z^i\|_2}{\sqrt{2k}} \leq \|z^i\|_2 \leq \frac{\|z^{i-1}\|_2}{\sqrt{2k}} \tag{22}$$

By assumption, $\|z^1\|_1 \geq \|\sum_{i=2}^{\infty} z^i\|_1$. Then

$$\begin{aligned}
\|Az\| &= \|A(z^1 + z^2 + \ldots)\| \\
&\geq \|A(z^1+z^2)\| - \|Az^3\| - \ldots \\
&\geq (1-\epsilon)\|z^1+z^2\|_2 - (1+\epsilon)(\sum_{i=3}^{\infty} \|z^i\|_2) \\
&\geq (1-\epsilon)\|z^1\|_2 - (1+\epsilon)(\sum_{i=2}^{\infty} \|z^i\|_1)/\sqrt{2k} \\
&= (1-\epsilon)\|z^1\|_2 - \frac{(1+\epsilon)}{\sqrt{2k}}\|\sum_{i=2}^{\infty} z^i\|_1 \\
&= (1-\epsilon)\|z^1\|_2 - \frac{(1+\epsilon)}{\sqrt{2k}}\|\sum_{i=2}^{\infty} z^i\|_1 \\
&\geq (1-\epsilon)\|z^1\|_2 - \frac{(1+\epsilon)}{\sqrt{2k}}\|z^1\|_1 \\
&\geq (1-\epsilon)\|z^1\|_2 - \frac{(1+\epsilon)}{\sqrt{2}}\|z^1\|_2 \\
&\geq \frac{1}{10}\|z^1\|_2
\end{aligned}$$

for $\epsilon < 1/10$.

## References

[CRT06] Candes, E. J., Romberg, J. K., Tao, T. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8), 1207-1223.

[BCDH10] R. Baraniuk, V. Cevher, M. F. Duarte, C. Hegde. Model-based compressive sensing. *IEEE Transactions on Information Theory*, 59(8), vol. 56, num. 4, p. 1982-2001, 2010.