EE5585 Data Compression	April 11, 2013

Lecture 21

Instructor: Arya Mazumdar — *Scribe: Zhongyu He*

Compressed Sensing

From last lecture, compressed sensing can be boiled down to a simple mathematical problem as the following equation:

$$\Phi x = y$$

where Φ is an m $\times$ N matrix, x is an $N \times 1$ vector, y is an $m \times 1$ vector. And the linear system is undetermined.

Given Φ and y, we need to solve for x ,which is approximately sparse, i.e., x has at most k prominent coordinates. Therefore, we want to design such a matrix Φ that can give us stable solutions.

Stable recovery

Suppose $\hat{x}$ is our estimate. x_k is a vector with k largest coefficient of x.

Suppose $x = \begin{bmatrix} 123 \\ 6.4 \\ -2.5 \\ 193 \\ 4.5 \end{bmatrix}$, $x_2 = \begin{bmatrix} 123 \\ 0 \\ 0 \\ 193 \\ 0 \end{bmatrix}$, then $x - x_k = \begin{bmatrix} 0 \\ 0.4 \\ -2.5 \\ 0 \\ 4.5 \end{bmatrix}$

The stable solution must satisfy the following:

$$\|\hat{x} - x\|_{\ell_2} \leq C\|x - x_k\|_{\ell_1}$$

where C is some constant.

If x has exactly k non-zero elements, then the equation above will give us the exact solution. If not, the error will be bounded by some small value. That's what is called stable recovery.

Basis Pursuit Algorithm(Φ,y)

Now we have an algorithm called Basis Pursuit, with Φ and y as input. It states as following:

$$\min \quad \|z\|_{\ell_1}$$

$$Subject\ to : \Phi z = y$$

This is an optimization problem and can be solved by a linear program.

k-RIP

For any k-sparse vector z, if

$$(1 - \delta_k)\|z\|_{\ell_2}^2 \leq \|\Phi z\|_{\ell_2}^2 \leq (1 + \delta_k)\|z\|_{\ell_2}^2$$

is satisfied with the smallest δ_k. Then Φ will be called (k,δ_k)-RIP. RIP is short for Restricted Isometry Property.

Theorem 1(by Candes, Tao 2005/2006)

If Φ is (2k,δ_{2k})-RIP with

$$\delta_{2k} < \sqrt{2} - 1 \approx 0.414$$

that is

$$(1-0.414)\|z\|_{\ell_2}^2 \leq \|\Phi z\|_{\ell_2}^2 \leq (1+0.414)\|z\|_{\ell_2}^2 \quad \forall\, 2k\texttt{-sparse}\ z$$

Then, basis pursuit solution $\hat{x}$ will satisfy

$$\|\hat{x} - x\|_{\ell_2} \leq (c/\sqrt{k})\|x - x_k\|_{\ell_1}$$

This property guarantees that stable recovery would happen.

Theorem 2

If $m = c_1 k \log N$ then there exists $m \times N$ matrix Φ that has $\delta_{2k} < \sqrt{2} - 1$.

If we choose an m×N independent zero-mean random Gaussian matrix satisfying the theorem above, then that matrix with a very high probability will have the RIP. This tells us how to construct Φ, but this is not our concern here.

Proof of **Theorem 1**

Assume $\hat{x} = x + h \Rightarrow h = \hat{x} - x$, where h is the error vector.

We will bound from above $\|h\|_{\ell_2}$

First of all,

$$\|x\|_{\ell_1} \geq \|\hat{x}\|_{\ell_1} = \|x + h\|_{\ell_1}$$

Now assume a vector v and a subset T $\subseteq$ {1,2,$\cdots$ N}. v_T is the projection of v on T.

For example:

$$v = \begin{bmatrix} 5 \\ 3 \\ 4 \\ -1 \\ 2 \end{bmatrix}, \text{T} = \{1,4,5\}, \text{ then } v_t = \begin{bmatrix} 5 \\ 6 \\ 0 \\ -1 \\ 2 \end{bmatrix}$$

$$\begin{aligned}
\|x\|_{\ell_1} &\geq \|x_{T_0} + h_{T_0} + x_{T_0^c} + h_{T_0^c}\|_{\ell_1} && (1)\\
&= \|x_{T_0} + h_{T_0}\|_{\ell_1} + \|x_{T_0^c} + h_{T_0^c}\|_{\ell_1} && (2)\\
&\geq \|x_{T_0}\|_{\ell_1} - \|h_{T_0}\|_{\ell_1} - \|x_{T_0^c}\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} && (3)\\
\Rightarrow \|x_{T_0}\|_{\ell_1} + \|x_{T_0^c}\|_{\ell_1} &\geq \|x_{T_0}\|_{\ell_1} - \|h_{T_0}\|_{\ell_1} - \|x_{T_0^c}\|_{\ell_1} + \|h_{T_0^c}\|_{\ell_1} && (4)\\
\Rightarrow \|h_{T_0^c}\|_{\ell_1} &\leq \|h_{T_0}\|_{\ell_1} + 2\|x_{T_0^c}\|_{\ell_1} && (5)
\end{aligned}$$

where T_0 is the k-largest coordinate of x.

*Reasonnings for (2)(4)

For example,

$$\hat{x}= \begin{bmatrix} 9 \\ 4 \\ 2 \\ 3 \\ 1 \\ -1 \\ -9 \\ 6 \end{bmatrix}, \hat{x}_{T_0}= \begin{bmatrix} 9 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -9 \\ 6 \end{bmatrix}, \hat{x}_{T_0^c}= \begin{bmatrix} 0 \\ 4 \\ 2 \\ 3 \\ 1 \\ -1 \\ 0 \\ 0 \end{bmatrix},$$

$$\begin{aligned} \|\hat{x}\|_{\ell_1} &= |9| + |4| + |2| + |3| + |1| + |-1| + |-9| + |6| \\ &= \|\hat{x}_{T_0}\|_{\ell_1} + \|\hat{x}_{T_0^c}\|_{\ell_1} \end{aligned}$$

****Resonings for (3)**

By Triangle Inequality

Lemma

If z, z' are two vectors that are k_1-sparse and k_2-sparse respectively; moreover, the coordinates where z, z' are non-zero do not overlap. Then

$$| < \Phi z, \Phi z' > | \leq \delta_{k_1+k_2} \|z\|_{\ell_2} \|z'\|_{\ell_2}$$

Proof

$$\begin{aligned} | < \Phi z, \Phi z' > | &= \frac{1}{4} \left[\|\Phi z + \Phi z'\|_{\ell_2}^2 - \|\Phi z - \Phi z'\|_{\ell_2}^2 \right] \\ &= \frac{1}{4} \left[\|\Phi(z + z')\|_{\ell_2}^2 - \|\Phi(z - z')\|_{\ell_2}^2 \right] \end{aligned}$$

$$(1 - \delta_{k_1+k_2}) \|z \pm z'\|_{\ell_2}^2 \leq \|\Phi(z \pm z')\|_{\ell_2}^2 \leq (1 + \delta_{k_1+k_2}) \|z \pm z'\|_{\ell_2}^2$$

$$\begin{aligned} \Rightarrow | < \Phi z, \Phi z' > | &\leq \frac{1}{4} \left[(1 + \delta_{k_1+k_2}) \|z + z'\|_{\ell_2}^2 - (1 - \delta_{k_1+k_2}) \|z + z'\|_{\ell_2}^2 \right] \\ &= \frac{1}{2} \delta_{k_1+k_2} \|z + z'\|_{\ell_2}^2 \end{aligned}$$

Now

$$\begin{aligned} \left| \frac{< \Phi z, \Phi z' >}{\|z\|_{\ell_2} \|z'\|_{\ell_2}} \right| &= | < \Phi \frac{z}{\|z\|_{\ell_2}}, \Phi \frac{z'}{\|z'\|_{\ell_2}} > | \\ &\leq \frac{1}{2} \delta_{k_1+k_2} \| \frac{z}{\|z\|_{\ell_2}} + \frac{z'}{\|z'\|_{\ell_2}} \|_{\ell_2}^2 \\ &= \delta_{k_1+k_2} \\ \Rightarrow | < \Phi z, \Phi z' > | &\leq \delta_{k_1+k_2} \|z\|_{\ell_2} \|z'\|_{\ell_2} \end{aligned}$$

Lemma proved. ■

Let's come back to the proof of **Theorem 1**.
T_0 is the k-largest coefficients of x.
T_1 is the k-largest in absolute value coefficients of $h_{T_0^c}$.
T_2 is the next k-largest in absolute value coefficients of $h_{T_0^c}$.

$$\vdots$$

Next, we will show that both $\|h_{T_0 \cup T_1}\|$ and $\|h_{(T_0 \cup T_1)^c}\|$ are bounded.
Note that for any $j \geq 2$,

$$\begin{aligned} \|h_{T_j}\|_{\ell_2} &\leq \sqrt{k} \texttt{ max value in } h_{T_j} \texttt{ (Definition of l2 norm)} \\ &\leq \sqrt{k} \texttt{ average absolute value in } h_{T_{j-1}} \\ &\leq \frac{\sqrt{k}}{k} \|h_{T_{j-1}}\|_{\ell_1} \end{aligned}$$

Now

$$\begin{aligned} \|h_{(T_0 \cup T_1)^c}\|_{\ell_2} &= \|\sum_{j \geq 2} h_{T_j}\|_{\ell_2} \\ &\leq \sum_{j \geq 2} \|h_{T_j}\|_{\ell_2} \\ &= \frac{1}{\sqrt{k}} \sum_{j \geq 1} \|h_{T_j}\|_{\ell_1} \\ &= \frac{1}{\sqrt{k}} \|h_{T_0^c}\|_{\ell_1} \\ &\leq \frac{1}{\sqrt{k}} (\|h_{T_0}\|_{\ell_1} + 2\|x_{T_0^c}\|_{\ell_1}) \end{aligned} \tag{6}$$

Because

$$\begin{aligned} \|z\|_{\ell_1} &\geq \|z\|_{\ell_2} \\ |z_1| + |z_2| &\geq \sqrt{|z_1|^2 + |z_2|^2} \\ \|z\|_{\ell_2} &\geq \frac{\|z\|_{\ell_1}}{\sqrt{\texttt{number of elements}}} \\ \Rightarrow \sqrt{z_1^2 + z_2^2 + \cdots + z_k^2} &\geq \frac{|z_1| + |z_2| + \cdots + |z_k|}{\sqrt{k}} \end{aligned}$$

So from(6),

$$\begin{aligned} \|h_{(T_0 \cup T_1)^c}\|_{\ell_2} &\leq \|h_{T_0}\|_{\ell_2} + \frac{2}{\sqrt{k}} \|x_{T_0^c}\|_{\ell_1} \\ \Rightarrow \|h_{(T_0 \cup T_1)^c}\|_{\ell_2} &\leq \|h_{(T_0 \cup T_1)}\|_{\ell_2} + \frac{2}{\sqrt{k}} \|x_{T_0^c}\|_{\ell_1} \end{aligned} \tag{7}$$

Now we have

$$\|h_{(T_0 \cup T_1)}\|_{\ell_2}^2 \leq \frac{1}{1 - \delta_{2k}} \|\Phi h_{(T_0 \cup T_1)}\|_{\ell_2}^2 \tag{8}$$

$$\begin{aligned} \|\Phi h_{(T_0 \cup T_1)}\|_{\ell_2}^2 &= < \Phi h_{(T_0 \cup T_1)}, \Phi h_{(T_0 \cup T_1)} > \\ &= < \Phi h_{(T_0 \cup T_1)}, \Phi(h - h_{(T_0 \cup T_1)^c}) > \end{aligned}$$

Since $\Phi h = \Phi(\hat{x} - x) = \Phi\hat{x} - \Phi x = y - y = 0$

$$\begin{aligned}
\Rightarrow \|\Phi h_{(T_0 \cup T_1)}\|_{\ell_2}^2 &= <\Phi h_{(T_0\cup T_1)}, -\Phi h_{(T_0\cup T_1)^c}> \\
&= <\Phi h_{(T_0\cup T_1)}, -\sum_{j\geq 2}\Phi h_{T_j}> \\
&\leq |<\Phi h_{T_0}, -\sum_{j\geq 2}\Phi h_{T_j}>| + |<\Phi h_{T_1}, -\sum_{j\geq 2}\Phi h_{T_j}>| \\
&\leq \sum_{j\geq 2}\left[|<\Phi h_{T_0}, \Phi h_{T_j}>| + |<\Phi h_{T_1}, \Phi h_{T_j}>|\right] \\
&\leq \sum_{j\geq 2}(\delta_{2k}\|h_{T_0}\|_{\ell_2}\|h_{T_1}\|_{\ell_2} + \delta_{2k}\|h_{T_1}\|_{\ell_2}\|h_{T_j}\|_{\ell_2})
\end{aligned}$$

Since $\sqrt{a} + \sqrt{b} \leq \sqrt{2(a+b)}$

$$\Rightarrow \|\Phi h_{(T_0\cup T_1)}\|_{\ell_2}^2 \leq \delta_{2k} 2\|h_{(T_0\cup T_1)}\|_{\ell_2}(\|h_{(T_0\cup T_1)}\|_{\ell_2} + \frac{2}{\sqrt{k}}\|x_{T_0^c}\|_{\ell_1})$$

$$\texttt{from (8)} \Rightarrow \|h_{(T_0\cup T_1)}\|_{\ell_2}^2 \leq \frac{1}{1-\delta_{2k}}\left[\delta_{2k} 2\|h_{(T_0\cup T_1)}\|_{\ell_2}(\|h_{(T_0\cup T_1)}\|_{\ell_2} + \frac{2}{\sqrt{k}}\|x_{T_0^c}\|_{\ell_1})\right]$$

$$\|h_{(T_0\cup T_1)}\|_{\ell_2}(1 - \frac{\sqrt{2}\delta_{2k}}{1-\delta_{2k}}) \leq \frac{2}{\sqrt{k}}\|x_{T_0^c}\|_{\ell_1})$$

$$\|h_{(T_0\cup T_1)}\|_{\ell_2} \leq \frac{1}{(1 - \frac{\sqrt{2}\delta_{2k}}{1-\delta_{2k}})}\frac{2}{\sqrt{k}}\|x_{T_0^c}\|_{\ell_1}) \tag{9}$$

$$\begin{aligned}
\|h\|_{\ell_2} &\leq \|h_{(T_0\cup T_1)}\|_{\ell_2} + \|h_{(T_0\cup T_1)^c}\|_{\ell_2} \\
&= 2\|h_{(T_0\cup T_1)^c}\|_{\ell_2} + \frac{2}{\sqrt{k}}\|x_{T_0^c}\|_{\ell_1} \\
\texttt{from (9)} &\Rightarrow \leq (\frac{2}{(1 - \frac{\sqrt{2}\delta_{2k}}{1-\delta_{2k}})} + 1)\frac{2}{\sqrt{k}}\|x_{T_0^c}\|_{\ell_1} \\
\Rightarrow \|\hat{x} - x\|_{\ell_2} &\leq (\frac{2}{(1 - \frac{\sqrt{2}\delta_{2k}}{1-\delta_{2k}})} + 1)\frac{2}{\sqrt{k}}\|x - x_2\|_{\ell_1}
\end{aligned}$$

where $\left(\frac{2}{(1 - \frac{\sqrt{2}\delta_{2k}}{1-\delta_{2k}})} + 1\right)$ is positive for any $\delta_{2k} > \sqrt{2} - 1$

QED (End of the Proof for **Theorem 1**)

■