Lecture 2: Sparse signal recovery: Analysis of $\ell_1$ minimization via RIP

Yuejie Chi
The Ohio State University
Outline

• Definition of sparse and compressible signals
  
  **Reference:** S. Foucart and H. Rauhut. A Mathematical Introduction to Compressive Sensing, Chapter 1.

• Uniqueness and identifiability using spark and coherence
  
  **Reference:** Donoho, D. L., & Elad, M. Optimally sparse representation in general (nonorthogonal) dictionaries via $\ell_1$ minimization. 2003.

• $\ell_1$ minimization, and sufficient condition for recovery using RIP
  
Signals that are exactly sparse

Consider a signal $x \in \mathbb{R}^n$.

**Definition 1. [Support]** The support of a vector $x \in \mathbb{R}^n$ is the index set of its nonzero entries, i.e.

$$\text{supp}(x) := \{ j \in [n] : x_j \neq 0 \}$$

where $[n] = \{1, \ldots, n\}$.

**Definition 2. [k-sparse signal]** The signal $x$ is called $k$-sparse, if

$$\|x\|_0 := |\text{supp}(x)| \leq k.$$

Note: $\|x\|_0$ is called the sparsity level of $x$. 
Sparse signals belong to union-of-subspace models

There’re $\binom{n}{k}$ subspaces of dimension $k$. 
**Compressible signals**

We’re also interested in signals that are *approximately* sparse. This is measured by how well they can be approximated by sparse signals.

**Definition 3. [Best $k$-term approximation]** Denote the index set of the $k$-largest entries of $|x|$ as $S_k$. The best $k$-term approximation $x_k$ of $x$ is defined as

$$x_k(i) = \begin{cases} 
  x_i, & i \in S_k \\
  0, & i \notin S_k 
\end{cases}$$

The $k$-term approximation error in $\ell_p$ norm is then given as

$$\|x - x_k\|_p = \left( \sum_{i \notin S_k} |x_i|^p \right)^{1/p}.$$ 

**Compressibility:** A signal is called *compressible* if $\|x - x_k\|_p$ decays fast in $k$. 
Example of compressible signals

Proposition 1. [Compressibility] For any $q > p > 0$ and $x \in \mathbb{R}^n$,

$$
\|x - x_k\|_q \leq \frac{1}{k^{1/p-1/q}} \|x\|_p.
$$

Example: set $q = 2$ and $0 < p < 1$, we have

$$
\|x - x_k\|_2 \leq \frac{1}{k^{1/p-1/2}} \|x\|_p.
$$

Consider a signal $x \in B_p^n := \{z \in \mathbb{R}^n : \|z\|_p \leq 1\}$. Then $x$ is compressible when $0 < p < 1$. [Geometrically, the $\ell_p$-ball is pointy when $0 < p < 1$ in high dimension. ]
Proof of Proposition [1]: Without loss of generality we assume the coefficients of \( \mathbf{x} \) is ordered in descending order of magnitudes. We then have

\[
\| \mathbf{x} - \mathbf{x}_k \|_q^q = \sum_{j=k+1}^{n} |x_j|^q \quad \text{(by definition)}
\]

\[
= |x_k|^{q-p} \sum_{j=k+1}^{n} |x_j|^p (|x_j|/|x_k|)^{q-p}
\]

\[
\leq |x_k|^{q-p} \sum_{j=k+1}^{n} |x_j|^p \quad (|x_j|/|x_k| \leq 1)
\]

\[
\leq \left( \frac{1}{k} \sum_{j=1}^{k} |x_j|^p \right)^{\frac{q-p}{p}} \left( \sum_{j=k+1}^{n} |x_j|^p \right)^{\frac{q-p}{p}}
\]

\[
\leq \left( \frac{1}{k} \| \mathbf{x} \|_p^p \right)^{\frac{q-p}{p}} \| \mathbf{x} \|_p^p = \frac{1}{k^{q/p-1}} \| \mathbf{x} \|_p^q.
\]
Compressive acquisition of sparse signals

• Let $A \in \mathbb{R}^{m \times n}$ be the measurement/sensing matrix. Consider, for start, noise-free measurements:

$$y = Ax \in \mathbb{R}^m,$$

where $m \ll n$. We are interested in reconstructing $x$ from $y$.

• Since we want to motivate sparse solutions, we could seek the sparsest signal satisfying the observation:

$$(P0:) \quad \hat{x} = \arg\min_x \|x\|_0 \quad \text{subject to} \quad y = Ax.$$

where $\| \cdot \|_0$ counts the number of nonzero entries.

• Although this algorithm is NP-hard, we can still analyze when it is expected to work.
Spark and uniqueness

Question: What properties do we seek in $A$ regardless of complexity of reconstruction algorithms?

Definition 4. [Spark] Let $\text{Spark}(A)$ be the size of the smallest linearly dependent subset of columns of $A$.

Basic Fact: $2 \leq \text{Spark}(A) \leq m + 1$.

Theorem 1. [Uniqueness, Donoho and Elad 2002] A representation $y = Ax$ is necessarily the sparsest possible if $\|x\|_0 < \text{Spark}(A)/2$.

Proof: If $x$ and $x'$ satisfy $Ax = Ax'$, with $\|x'\|_0 \leq \|x\|_0$, then

$$A(x - x') = 0$$

for $\|x - x'\|_0 < \text{Spark}(A)$, which contradicts with definition of Spark. Therefore, $x = x'$ and $x$ is the sparsest solution of $y = Ax$. 

Page 9


Mutual coherence

Definition 5. [Mutual Coherence]  Let

\[ \mu = \mu(A) := \max_{i \neq j} |\langle a_i, a_j \rangle| \]

where \( a_i \) and \( a_j \) are normalized columns of \( A \).

- \( \mu(A) \leq 1 \) if the columns of \( A \) are pairwise independent.

- \( \text{Spark}(A) > 1/\mu(A) \) [can be shown by the Gershgorin circle's theorem].

- Welch bound asserts

\[ \mu^2 \geq \frac{m-n}{n(m-1)}, \]

which roughly gives \( \mu = O(1/\sqrt{m}) \) for a “well-behaved” \( A \).
Lemma 2. [Gershgorin circle’s theorem] The eigenvalues of an $n \times n$ matrix $M$ with entries $m_{ij}$, $1 \leq i, j \leq n$, lie in the union of $n$ discs $d_i = d_i(c_i, r_i)$, $1 \leq i \leq n$, centered at $c_i = m_{ii}$ and with radius $r_i = \sum_{j \neq i} |m_{ij}|$.

Example: take $M = \begin{bmatrix} 4 & 2 & 3 \\ -2 & -5 & 8 \\ 1 & 0 & 3 \end{bmatrix}$
Sufficient condition using mutual coherence

Theorem 3. [Equivalence, Donoho and Elad 2002] The sparsest solution to $y = Ax$ is unique if $\|x\|_0 < \frac{1}{2} + \frac{1}{2\mu(A)}$.

- The largest recoverable sparsity of $x$ is $k \sim O(1/\mu) = O(\sqrt{m})$, which is square-root in the number of measurements.

- This result is deterministic.

- Requires the signal to be exactly sparse, which is not always practical.
Sparse Recovery via $\ell_1$ Minimization

Since the above $\ell_0$ minimization is NP-hard. We would like to take its convex relaxation, which leads to the $\ell_1$ minimization, or basis pursuit:

$$\text{(BP:)} \quad \hat{x} = \arg\min_x \|x\|_1 \text{ subject to } y = Ax.$$

- The BP algorithm does not assume knowledge of the sparsity level to perform.

- Compare this with the usual wisdom of $\ell_2$ minimization:

$$\hat{x}_{\ell_2} = \arg\min_x \|x\|_2 \text{ subject to } y = Ax.$$

which has a closed form solution

$$\hat{x}_{\ell_2} = A^\dagger y,$$

where $^\dagger$ denotes pseudo-inverse.
A numerical example

Let's run an example using CVX (http://cvxr.com/cvx/).
Geometry of basis pursuit

\[ p = 1 \]

\[ p = 2 \]
Definition 6. [Restricted Isometry Property (RIP)] If $A$ satisfies the restricted isometry property (RIP) with $\delta_{2k}$, then for any two $k$-sparse vectors $x_1$ and $x_2$:

$$1 - \delta_{2k} \leq \frac{\|A(x_1 - x_2)\|_2^2}{\|x_1 - x_2\|_2^2} \leq 1 + \delta_{2k}.$$ 

If $\delta_{2k} < 1$, this implies the $\ell_0$ problem has a unique $k$-sparse solution.
RIP matrices preserve orthogonality between sparse vectors

Proposition 2.

\[ |\langle Ax_1, Ax_2 \rangle| \leq \delta_{s_1+s_2} \|x_1\|_2 \|x_2\|_2 \]

for all \( x_1, x_2 \) that are supported on disjoint subsets \( T_1, T_2 \subset [n] \) with \( |T_1| \leq s_1 \) and \( |T_2| \leq s_2 \).

Proof: Without loss of generality assume \( \|x_1\|_2 = \|x_2\|_2 = 1 \). Applying the parallelogram identity, which says

\[
|\langle Ax_1, Ax_2 \rangle| = \frac{1}{4} \|Ax_1 + Ax_2\|_2^2 - \|Ax_1 + Ax_2\|_2^2 \leq \frac{1}{4} |2(1 + \delta_{s_1+s_2}) - 2(1 - \delta_{s_1+s_2})| \leq \delta_{s_1+s_2}.
\]
Theorem 4. [Performance of BP via RIP, Candès, Tao, Romberg, 2006]

If \( \delta_{2^k} < \sqrt{2} - 1 \), then for any vector \( x \), the solution to basis pursuit satisfies

\[
\| \hat{x} - x \|_2 \leq C_0 k^{-1/2} \| x - x_k \|_1.
\]

where \( x_k \) is the best \( k \)-term approximation of \( x \) for some constant \( C_0 \).

- **exact recovery** if \( x \) is exactly \( k \)-sparse.

- Many random ensembles (e.g. Gaussian, sub-Gaussian, partial DFT) satisfies the RIP as soon as (we'll return to this point)

\[
m \sim \Theta(k \log(n/k))
\]

- The proof of theorem is particularly elegant.
Proof of Theorem 4

Proof of Theorem 4: Set \( \hat{x} = x + h \). We already show \( A h = 0 \). The goal is to establish that \( h = 0 \) when \( A \) satisfies the desired RIP.

The first step is to decompose \( h \) into a sum of vectors \( h_{T_0}, h_{T_1}, h_{T_2}, \ldots \), each of sparsity at most \( k \). Here, \( T_0 \) corresponds to the locations of the \( k \) largest coefficients of \( x \); \( T_1 \) to the locations of the \( k \) largest coefficients of \( h_{T_0^c} \), \( T_2 \) to the locations of the next \( k \) largest coefficients of \( h_{T_0^c} \), and so on.

The proof proceeds in two steps:

1. the first step shows that the size of \( h \) outside of \( T_0 \cup T_1 \) is essentially bounded by that of \( h \) on \( T_0 \cup T_1 \).

2. the second step shows that \( \|h_{T_0 \cup T_1}\|_2 \) is appropriately small.
Step 1: Note that for each $j \geq 2$, 

$$\| h_{T_j} \|_2 \leq \sqrt{k} \| h_{T_j} \|_\infty \leq \frac{1}{\sqrt{k}} \| h_{T_{j-1}} \|_1$$

therefore 

$$\sum_{j \geq 2} \| h_{T_j} \|_2 \leq \frac{1}{\sqrt{k}} \sum_{j \geq 1} \| h_{T_j} \|_1 = \frac{1}{\sqrt{k}} \| h_{T^c_0} \|_1.$$

This allows us to bound 

$$\| h_{(T_0 \cup T_1)^c} \|_2 \leq \| \sum_{j \geq 2} h_{T_j} \|_2 \leq \sum_{j \geq 2} \| h_{T_j} \|_2 \leq \frac{1}{\sqrt{k}} \| h_{T^c_0} \|_1.$$

Given $\hat{x} = x + h$ is the optimal solution, we have 

$$\| x \|_1 \geq \| x + h \|_1 = \sum_{i \in T_0} | x_i + h_i | + \sum_{i \in T^c_0} | x_i + h_i |$$

$$\geq \| x_{T_0} \|_1 - \| h_{T_0} \|_1 + \| h_{T^c_0} \|_1 - \| x_{T^c_0} \|_1,$$  \(\ast\)
which gives

\[ \| h_{T_0^c} \|_1 \leq \| h_{T_0} \|_1 + \| x \|_1 - \| x_{T_0^c} \|_1 + \| x_{T_0^c} \|_1 \]
\[ \leq \| h_{T_0} \|_1 + 2 \| x_{T_0^c} \|_1 := \| h_{T_0} \|_1 + 2 \| x - x_k \|_1. \]

Combining with (*), we have

\[ \| h_{(T_0 \cup T_1)^c} \|_2 \leq \frac{1}{\sqrt{k}} \| h_{T_0^c} \|_1 \leq \frac{1}{\sqrt{k}} \| h_{T_0} \|_1 + \frac{2}{\sqrt{k}} \| x - x_k \|_1. \]

Step 2: We next bound \( \| h_{T_0 \cup T_1} \|_2 \). Note that

\[ 0 = Ah = Ah_{T_0 \cup T_1} + \sum_{j \geq 2} Ah_{T_j}, \]

we have by RIP

\[ (1 - \delta_{2k}) \| h_{T_0 \cup T_1} \|_2^2 \leq \| Ah_{T_0 \cup T_1} \|_2^2 = \| \langle Ah_{T_0 \cup T_1}, \sum_{j \geq 2} Ah_{T_j} \rangle \|. \]
Using Proposition 2, we have for $j \geq 2$

$$|\langle Ah_{T_0 \cup T_1}, Ah_j \rangle| \leq |\langle Ah_{T_0}, Ah_j \rangle| + |\langle Ah_{T_1}, Ah_j \rangle|$$

$$\leq \delta_{2k}(\|h_{T_0}\|_2 + \|h_{T_1}\|_2)\|h_j\|_2$$

$$\leq \delta_{2k}\sqrt{2}\|h_{T_0 \cup T_1}\|_2\|h_j\|_2,$$

which gives

$$\sum_{j \geq 2} (1 - \delta_{2k})\|h_{T_0 \cup T_1}\|_2^2 \leq \sum_{j \geq 2} |\langle Ah_{T_0 \cup T_1}, Ah_j \rangle|$$

$$\leq \sqrt{2}\delta_{2k}\|h_{T_0 \cup T_1}\|_2 \sum_{j \geq 2} \|h_j\|_2$$

$$\leq \sqrt{2}\delta_{2k}\|h_{T_0 \cup T_1}\|_2 \frac{1}{\sqrt{k}}\|h_{T_0^c}\|_1,$$

therefore

$$\|h_{T_0 \cup T_1}\|_2 \leq \frac{\sqrt{2}\delta_{2k}}{(1 - \delta_{2k})\sqrt{k}}\|h_{T_0^c}\|_1 \leq \frac{1}{\sqrt{k}}\rho\frac{1}{\sqrt{k}}(\|h_{T_0}\|_1 + 2\|x - x_k\|_1)$$
where $\rho := \frac{\sqrt{2}\delta_{2k}}{(1-\delta_{2k})}$. Since $\|h_{T_0}\|_1 \leq \sqrt{k}\|h_{T_0}\|_2 \leq \sqrt{k}\|h_{T_0 \cup T_1}\|_2$, we can bound

$$\|h_{T_0 \cup T_1}\|_2 \leq \frac{2\rho}{1-\rho} \frac{\|x - x_k\|_1}{\sqrt{k}}.$$

Finally,

$$\|\hat{x} - x\|_2 = \|h\|_2 \leq \|h_{T_0\cup T_1}\|_2 + \|h_{(T_0 \cup T_1)^c}\|_2$$

$$\leq \|h_{T_0 \cup T_1}\|_2 + \frac{1}{\sqrt{k}}\|h_{T_0}\|_1 + \frac{2}{\sqrt{k}}\|x - x_k\|_1$$

$$\leq 2\|h_{T_0 \cup T_1}\|_2 + \frac{2}{\sqrt{k}}\|x - x_k\|_1$$

$$\leq \frac{2(1+\rho)}{1-\rho} \frac{\|x - x_k\|_1}{\sqrt{k}}.$$

Therefore, $C_0 := \frac{2(1+\rho)}{1-\rho}$. The requirement on $\delta_{2k}$ comes from the fact that we need $1-\rho > 0$ to avoid the bound to blow up.
In the presence of additive measurement noise,

\[ y = Ax + w, \]

where \( \|w\|_2 \leq \epsilon \) is assumed to be bounded.

We can modify the BP algorithm in the following manner:

\[ (\text{BP-noisy:}) \quad \hat{x} = \arg \min_x \|x\|_1 \quad \text{subject to} \quad \|y - Ax\|_2 \leq \epsilon. \]

**Theorem 5. [Performance of BP via RIP, noisy case]**  
\[ \text{If } \delta_{2k} < \sqrt{2} - 1, \text{ then for any vector } x, \text{ the solution to basis pursuit (noisy case) satisfies} \]

\[ \|\hat{x} - x\|_2 \leq C_0 k^{-1/2} \|x - x_k\|_1 + C_1 \epsilon. \]

where \( x_k \) is the best \( k \)-term approximation of \( x \) for some constants \( C_0 \) and \( C_1 \).
Proof of Theorem 5

Again let’s start by assuming $\hat{x} = x + h$. The key difference from the noiseless case is that in Step 2, we now have

$$\|Ah\|_2 = \|A(\hat{x} - x)\|_2 = \|(y - A\hat{x}) - (y - Ax)\|_2$$
$$\leq \|y - A\hat{x}\|_2 + \|y - Ax\|_2 \leq 2\epsilon.$$

Therefore, we need to bound

$$\|Ah_{T_0 \cup T_1}\|_2^2 = \langle Ah - \sum_{j \geq 2} Ah_{T_j}, Ah_{T_0 \cup T_1} \rangle$$
$$\leq \langle Ah, Ah_{T_0 \cup T_1} \rangle - \sum_{j \geq 2} \langle Ah_{T_j}, Ah_{T_0 \cup T_1} \rangle$$
$$\leq 2\epsilon \delta_{2k} \|h_{T_0 \cup T_1}\|_2$$
$$\text{bounded as before}$$
By plugging in this modification, we show

$$\|\hat{x} - x\|_2 = \|h\|_2 \leq \frac{2(1 + \rho)\|x - x_k\|_1}{1 - \rho} \sqrt{k} + \frac{2\alpha}{1 - \rho} \epsilon,$$

where

$$\alpha = \frac{2\sqrt{1 + \delta_{2k}}}{1 - \delta_{2k}}.$$
Remarks

• The theorems are quite strong, in the sense it holds for all signals once $A$ satisfies RIP.

• The reconstruction quality relies on two quantities: the best $k$-term approximation error and the noise level.

• Our generalization of the performance guarantee from the noise-free case to the noisy case is essentially effortless. However, we do need an upper bound of the noise level in order to perform the algorithm.

• A related algorithm is called LASSO, which has the form of

$$\hat{x}_{lasso} = \arg\min_x \frac{1}{2}\|y - Ax\|_2^2 + \lambda\|x\|_1,$$

where $\lambda > 0$ is called a regularization parameter. Another related algorithm is called Dantizg selector. Both can be analyzed in a similar manner as the BP using RIP.
Which matrices satisfy RIP?

- Random matrices with i.i.d. Gaussian entries satisfy RIP with high probability, as long as

  \[ m \gtrsim k \log(n/k). \]

- Random Partial DFT matrices, \( A = I_\Omega F \), where \( I_\Omega \) is an partial identity matrix with rows indexed by the random subset \( \Omega \), and \( F \) is the DFT matrix, satisfy RIP with high probability, as long as

  \[ m = |\Omega| \gtrsim k \log^4 n. \]

- Similar results hold for random Partial Circulant/Toeplitz matrices, random matrices with i.i.d. sub-Gaussian entries, etc...

- All these are probabilistic, in the sense if we draw a random matrix following the stated distribution, it will satisfy the RIP with high probability (i.e. \( 1 - \exp(-cm) \)).
Deterministic matrices satisfying RIP

Constructing deterministic matrices that satisfy RIP is difficult.

There’re many benefits of having deterministic constructions: fast computation, less storage, etc..