Lecture 15 — October 24, 2013

Prof. Jelani Nelson *Scribe: Mark Bun*

1 Overview

In this lecture we started the fourth module of the course, on compressed sensing. We gave an overview of the general idea and goals of compressed sensing, and then proved an "ℓ_2/ℓ_1 guarantee" that is essentially the best possible.

2 Compressed Sensing

In compressed sensing, we are given a "compressible" signal $x \in \mathbb{R}^n$, and our goal is use few linear measurements of x to approximately recover x. Here, a linear measurement of x is its dot product with another vector in $\mathbb{R}^n$. We can arrange m such linear measurements to form the rows of a matrix $\Pi \in \mathbb{R}^{m\times n}$, so the goal now becomes to approximately recover x from Πx using $m \ll n$.

Note that if $m < n$, then any Π has a non-trivial kernel, so we have no hope of exactly recovering every $x \in \mathbb{R}^n$. This motivates our relaxed objective of only recovering *compressible* signals.

So what exactly do we mean by compressible? A compressible signal is one which is (approximately) sparse in some basis – but not necessarily the standard basis. Here an approximately sparse signal is a sum of a sparse vector with a low-weight vector.

Example: Consider an image consisting of $2^N \times 2^N$ pixels, each with a numeric color value. A real-world image is unlikely to be sparse in the standard basis over $\mathbb{R}^n$ with $n = 2^{2N}$. However, an image is likely to be approximately sparse over the *Haar wavelet basis*. Weights in this alternative basis capture changes between adjacent pixels in the image, which are unlikely to be extreme in most places.

The Haar wavelet transform is easier to describe than the orthonormal basis itself. An image is broken up into $2^{N-1} \times 2^{N-1}$ blocks of 2×2 pixels. Let's focus on a single block at index (i, j), and label its pixels p_1, p_2, p_3, p_4. Now consider another $2^N \times 2^N$ grid divided into four quadrants. The (i, j) entry of the upper-left quadrant will take the value $\frac{1}{4}(p_1 + p_2 + p_3 + p_4)$, and the (i, j)-entries of the other quadrants will take the values

$$\frac{1}{4}(p_1 + p_3 - p_2 - p_4), \quad \frac{1}{4}(p_1 + p_2 - p_3 - p_4), \quad \frac{1}{4}(p_1 + p_4 - p_2 - p_3),$$

which are all likely to be close to zero. However, the upper-left quadrant is still dense, so we recurse the transformation on this quadrant, eventually yielding a 2×2 block in the upper-left with large entries and relatively low weights elsewhere.

The point of this example is to illustrate that a signal can be compressible even if it's dense in the standard basis. In fact, the wavelet transform is one of the (many) tricks that goes into practical image compression.

2.1 Algorithmic Goals

The compressed sensing algorithms we discuss will achieve the following. If x is actually sparse, we will recover x exactly in polynomial time. And if x is only approximately sparse, then we will recover x approximately, again in poly-time.

More formally, we seek to meet "ℓ_p/ℓ_q guarantees": Given Πx, we will recover $\tilde{x}$ such that

$$\|x - \tilde{x}\|_p \leq C_{k,p,q} \min_{\|y\|_0 \leq k} \|y - x\|q,$$

where the ℓ_0-norm of a vector is the count of its non-zero coordinates. Observe that the minimizer y in the statement above picks out the largest (in absolute value) k coordinates of x and zeroes out the rest of them. Also, note that the right-hand side is zero when x is actually k-sparse.

3 Main Result

The goal of this lecture is to prove the following ℓ_2/ℓ_1 guarantee. It is due (independently) to Candes, Romberg, and Tao [CRT04] and to Donoho [Don04].

Theorem 1. *There is a polynomial-time algorithm which, given Πx for $\Pi \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$, can recover $\tilde{x}$ such that*

$$\|x - \tilde{x}\|_2 \leq O\left(\frac{1}{\sqrt{k}}\right) \|x_{tail(k)}\|_1$$

where $x_{tail(k)}$ is x with its top k coordinates zeroed out.

3.1 Exact recovery in the sparse case

As a first step toward proving the theorem, let's examine what we need to recover x exactly when we actually have $\|x\|_0 \leq k$. Information-theoretically, it's necessary and sufficient to have $\Pi x \neq \Pi x'$ whenever $x \neq x'$ are both k-sparse. This is equivalent to requiring any $2k$-sparse vector to lie outside $\ker \Pi$, i.e., requiring each restriction Π_S of Π to the columns in a set S to have full column rank for every $S \subseteq [n]$ with $|S| \leq 2k$.

How can we use this characterization to recover x given $y = \Pi x$ when Π has this property? One way is to find the minimizer z in

$$\begin{aligned} \min_{z \in \mathbb{R}^n} \quad & \|z\|_0 \\ \text{s.t.} \quad & \Pi z = y. \end{aligned}$$

Unfortunately, this optimization problem is NP-hard to solve in general [GJ79, Problem MP5]. In what follows, we will show that with an additional constraint on Π, we can approximately solve this optimization problem using linear programming.

3.2 RIP matrices

Observe that Π_S has full column rank if and only if $\Pi_S^T \Pi_S$ has no zero eigenvalues. RIP matrices are matrices Π for which this property is robust: i.e., all eigenvalues of $\Pi_S^T \Pi_S$ are approximately equal, and hence $\Pi_S^T \Pi_S$ is far from being non-invertible.

Definition 2 (RIP Matrix)**.** *A matrix $\Pi \in \mathbb{R}^{m \times n}$ satisfies the (ε, k)-restricted isometry property (RIP) if for all k-sparse vectors x,*

$$(1-\varepsilon)\|x\|_2^2 \leq \|\Pi x\|_2^2 \leq (1+\varepsilon)\|x\|_2^2.$$

Equivalently, whenever $|S| \leq k$, we have

$$\|\Pi_S^T \Pi_S - I_k\|_2 \leq \varepsilon.$$

Where do we get RIP matrices? There are a few methods.

1. Use Johnson-Lindenstrauss to preserve each k-dimensional subspace consisting of vectors supported on k coordinates. Applying JL to the requisite $\binom{n}{k} e^{O(k)}$ vectors yields, by Stirling's approximation,
$$m \lesssim \frac{1}{\varepsilon^2} k \log\left(\frac{n}{k}\right).$$
(Note that for our application, we will want to think of ε as a constant.)
2. Use incoherent matrices. The upshot over JL is that these constructions are explicit from codes, rather than being samples from a distribution. The tradeoff is that we get worse parameters: the best constructions require $m \geq k^2/\varepsilon^2$.
3. From first principles. A matrix Π might not satisfy JL, but might still preserve the norms of k-sparse vectors. For example, we can take Π to sample m rows from a Fourier matrix. Recall that for the FJLT, we needed to subsequently multiply by a diagonal sign matrix, but there is no need to do so in the particular case of sparse vectors.

From the section of the course on dimensionality reduction, we already have a good handle on method 1, so we'll instead focus on method 2.

Incoherent matrices are RIP. Recall that a matrix Π is α-incoherent if, letting Π^i denote the ith column of Π,

- $\|\Pi^i\|_2 = 1$ for all i, and
- $|\langle \Pi^i, \Pi^j \rangle| \leq \alpha$ for all $i \neq j$.

Our analysis of incoherent matrices is based on the following lemma.

Lemma 3 (Gershgorin Circle Theorem)**.** *Every eigenvalue of a matrix $A = (a_{ij})$ lies in a disk about some a_{ii} of radius $\sum_{j \neq k} |a_{ij}|$ in the complex plane.*

Proof. Let x be an eigenvector of A with corresponding eigenvalue λ. Let i be an index such that $|x_i| = \|x\|_\infty$. Then $(Ax)_i = \lambda x_i$ so

$$\lambda x_i = \sum_{j=1}^{n} a_{ij} x_j \Rightarrow |(\lambda - a_{ii}) x_i| = \left| \sum_{j \neq i} a_{ij} x_j \right|$$
$$\Rightarrow |\lambda - a_{ii}| \leq \sum_{j \neq i} \left| \frac{a_{ij} x_j}{x_i} \right| \leq \sum_{j \neq i} |a_{ij}|$$

by our choice of i. □

Now suppose Π is α-incoherent. We'll apply the lemma to $A = \Pi_S^T \Pi_S$ for some $|S| \leq k$. Note that A is symmetric, so it has real eigenvalues by the Spectral Theorem. Moreover, $a_i i = 1$ for each i and $|a_{ij}| \leq \alpha$ by the definition of incoherence. Therefore, the eigenvalues of A lie in intervals about 1 of radius

$$\sum_{j \neq i} |\langle \Pi_S^i, \Pi_S^j \rangle| \leq \alpha(k-1).$$

Hence Π is (ε, k)-RIP as long as $\alpha \leq \varepsilon/k$.

Remark: Finding explicit RIP matrices with $m \ll k^2$ is a challenging open problem. Some headway appears in the work of Bourgain et al. [BDFKK11], who obtained $m \ll k^{2-\gamma}$ for some small constant γ and for some $k \approx \sqrt{n}$. Their technique combines tools from analytic number theory and additive combinatorics.

3.3 From RIP matrices to recovery

Theorem 4. *If Π is $(\varepsilon_{2k}, 2k)$-RIP with $\varepsilon_{2k} \leq \sqrt{2} - 1$, and $\tilde{x} = x + h$ is the solution to the "basis pursuit" linear program*

$$\begin{aligned} \min_{z \in \mathbb{R}^n} \quad & \|z\|_1 \\ \text{s.t.} \quad & \Pi z = \Pi x, \end{aligned}$$

then

$$\|h\|_2 \leq O\left(\frac{1}{\sqrt{k}}\right) \|x_{tail(k)}\|_1.$$

Remark: A *linear program* (LP) is an optimization problem in which one seeks to optimize a linear objective function subject to linear constraints. The above problem is indeed a linear program with polynomially many variables and constraints, since it is equivalent to

$$\begin{aligned} \min_{y \in \mathbb{R}^n} \quad & \sum_i y_i \\ \text{s.t.} \quad & \Pi z = \Pi x, \\ & z_i \leq y_i \quad \forall i, \\ & -z_i \leq y_i \quad \forall i. \end{aligned}$$

It is known (e.g. via Khachiyan's analysis of the ellipsoid method) that LPs can be solved in polynomial time.

We will now present a proof along the lines of [Candes08].

Proof of Theorem 4. First, we define some notation. For a vector $x \in \mathbb{R}^n$ and a set $S \subseteq [n]$, let x_S be the vector with all of its coordinates outside of S zeroed out.

- Let $T_0 \subseteq [n]$ be the indices of the largest (in absolute value) k coordinates of x.
- Let T_1 be the indices of the largest k coordinates of $h_{T_0^c} = h_{tail(k)}$.
- Let T_2 be the indices of the *second* largest k coordinates of $h_{T_0^c}$.
- ...and so forth, for $T_3, \ldots$

By the triangle inequality, we can write

$$\begin{aligned}\|h\|_2 &= \|h_{T_0 \cup T_1} + h_{(T_0 \cup T_1)^c}\|_2 \\ &\leq \|h_{T_0 \cup T_1}\|_2 + \|h_{(T_0 \cup T_1)^c}\|_2.\end{aligned}$$

Our strategy for bounding h will be to show

1. $$\|h_{(T_0 \cup T_1)^c}\|_2 \leq \|h_{T_0 \cup T_1}\|_2 + O\left(\frac{1}{\sqrt{k}}\right) \|x_{tail(k)}\|_1.$$

2. $$\|h_{T_0 \cup T_1}\|_2 \leq O\left(\frac{1}{\sqrt{k}}\right) \|x_{tail(k)}\|_1)$$

Both parts rely on the following lemma.

Claim 5.

$$\sum_{j \geq 2} \|h_{T_j}\|_2 \leq \frac{2}{\sqrt{k}} \|x_{T_0^c}\|_1 + \|h_{T_0 \cup T_1}\|_2.$$

Proof of Claim 5. We first get an upper bound on the left-hand side by applying a technique known as the "shelling trick."

$$\begin{aligned}\sum_{j \geq 2} \|h_{T_j}\|_2 &\leq \sqrt{k} \sum_{j \geq 2} \|h_{T_j}\|_\infty \\ &\leq \frac{1}{\sqrt{k}} \sum_{j \geq 2} \|h_{T_{j-1}}\|_1 \\ &\leq \frac{1}{\sqrt{k}} \|h_{T_0^c}\|_1. \qquad (1)\end{aligned}$$

The first inequality holds because each h_{T_j} is k-sparse, and the second holds because the size of every term in h_{T_j} is bounded from above by the size of every term in $h_{T_{j-1}}$.

Now since $\tilde{x} = x + h$ is the minimizer of the LP, we must have

$$\begin{aligned}
\|x\|_1 &\geq \|x+h\|_1 \\
&= \|(x+h)_{T_0}\|_1 + \|(x+h)_{T_0^c}\|_1 \\
&\geq \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1
\end{aligned}$$

by two applications of the reverse triangle inequality. Rearranging, we obtain

$$\begin{aligned}
\|h_{T_0^c}\|_1 &\leq \|x\|_1 - \|x_{T_0}\|_1 + \|h_{T_0}\|_1 + \|x_{T_0^c}\|_1 \\
&= 2\|x_{T_0^c}\|_1 + \|h_{T_0}\|_1 \\
&\leq 2\|x_{T_0^c}\|_1 + \sqrt{k}\|h_{T_0}\|_2 && \text{by Cauchy-Schwarz} \\
&\leq 2\|x_{T_0^c}\|_1 + \sqrt{k}\|h_{T_0 \cup T_1}\|_2
\end{aligned}$$

Combining this upper bound with Inequality (1) yields the claim. □

Returning to the proof of Theorem 4, let us first upper bound the size of $h_{(T_0 \cup T_1)^c}$. We get

$$\begin{aligned}
\|h_{(T_0 \cup T_1)^c}\|_2 &= \left\| \sum_{j \geq 2} h_{T_j} \right\|_2 \\
&\leq \sum_{j \geq 2} \|h_{T_j}\|_2 \\
&\leq \|h_{T_0 \cup T_1}\|_2 + \frac{2}{\sqrt{k}}\|x_{T_0^c}\|_1 && \text{by the claim} \\
&= \|h_{T_0 \cup T_1}\|_2 + \frac{2}{\sqrt{k}}\|x_{tail(k)}\|_1.
\end{aligned}$$

Now to bound the size of $h_{T_0 \cup T_1}$, we need another lemma.

Lemma 6. *If x, x' are supported on disjoint sets T, T' respectively, where $|T| = k$ and $|T'| = k'$, then*

$$|\langle \Pi x, \Pi x' \rangle| \leq \varepsilon_{k+k'} \|x\|_2 \|x'\|_2,$$

where Π is $(\varepsilon_{k+k'}, k+k')$-RIP.

Proof. We can assume WLOG that x, x' are unit vectors. Write

$$\|\Pi x + \Pi x'\|_2^2 = \|\Pi x\|_2^2 + \|\Pi x'\|_2^2 + 2\langle \Pi x, \Pi x' \rangle, \quad \text{and}$$

$$\|\Pi x - \Pi x'\|_2^2 = \|\Pi x\|_2^2 + \|\Pi x'\|_2^2 - 2\langle \Pi x, \Pi x' \rangle.$$

Taking the difference gives

$$
\begin{aligned}
|\langle \Pi x, \Pi x' \rangle| &= \frac{1}{4} \left| \|\Pi(x+x')\|_2^2 - \|\Pi(x-x')\|_2^2 \right| \\
&\le \frac{1}{4}((1+\varepsilon_{k+k'})\|x+x'\|_2^2 - (1-\varepsilon_{k+k'})\|x-x'\|_2^2) \\
&= \frac{1}{4}((1+\varepsilon_{k+k'})2 - (1-\varepsilon_{k+k'})2) \\
&= \varepsilon_{k+k'}
\end{aligned}
$$

since $x \pm x'$ are $(k+k')$-sparse, and x, x' are disjointly supported. This proves the lemma. □

To bound the size of $h_{T_0 \cup T_1}$, first observe that

$$
\Pi h_{T_0 \cup T_1} = \Pi h - \sum_{j \ge 2} \Pi h_{T_j} = -\sum_{j \ge 2} \Pi h_{T_j}
$$

since $h \in \ker \Pi$. Therefore,

$$
\|\Pi h_{T_0 \cup T_1}\|_2^2 = -\sum_{j \ge 2} \langle \Pi h_{T_0 \cup T_1}, \Pi h_{T_j} \rangle \le \sum_{j \ge 2} (|\langle \Pi h_{T_0}, \Pi h_{T_j} \rangle| + |\langle \Pi h_{T_1}, \Pi h_{T_j} \rangle|).
$$

By Lemma 6, each summand is at most

$$
\varepsilon_{2k}(\|h_{T_0}\|_2 + \|h_{T_1}\|_2)\|h_{T_j}\|_2 \le \varepsilon_{2k}\sqrt{2}\|h_{T_0 \cup T_1}\|_2 \|h_{T_j}\|_2.
$$

Thus

$$
\begin{aligned}
(1-\varepsilon_{2k})\|h_{T_0 \cup T_1}\|_2^2 &\le \|\Pi h_{T_0 \cup T_1}\|_2^2 \\
&\le \varepsilon_{2k}\sqrt{2}\|h_{T_0 \cup T_1}\|_2 \sum_{j \ge 2} \|h_{T_j}\|_2 \\
&\le \varepsilon_{2k}\sqrt{2}\|h_{T_0 \cup T_1}\|_2 \left(\frac{2}{\sqrt{k}} \|x_{T_0^c}\|_1 + \|h_{T_0 \cup T_1}\|_2 \right)
\end{aligned}
$$

by Claim 5. Cancelling a factor of $\|h_{T_0 \cup T_1}\|_2$ from both sides and rearranging gives

$$
\|h_{T_0 \cup T_1}\|_2 \le \frac{\varepsilon_{2k} 2\sqrt{2}}{(1-\varepsilon_{2k}-\varepsilon_{2k}\sqrt{2})\sqrt{k}} \|x_{T_0^c}\|_1 = O\left(\frac{1}{\sqrt{k}}\right) \|x_{tail(k)}\|_1.
$$

Putting everything together:

$$
\begin{aligned}
\|h\|_2 &\le \|h_{(T_0 \cup T_1)^c}\|_2 + \|h_{T_0 \cup T_1}\|_2 \\
&\le 2\|h_{T_0 \cup T_1}\|_2 + \frac{2}{\sqrt{k}} \|x_{tail(k)}\|_1 \\
&\le O\left(\frac{1}{\sqrt{k}}\right) \|x_{tail(k)}\|_1.
\end{aligned}
$$

□

References

[BDFKK11] Jean Bourgain, Stephen Dilworth, Kevin Ford, Sergei Konyagin, and Denka Kutzarova. Breaking the k^2 Barrier for Explicit RIP Matrices. In *STOC*, pages 637–644, 2011.

[Candes08] Emmanuel Candès. The restricted isometry property and its implications for compressed sensing. *C*. R. Acad. Sci. Paris, Ser. I 346:589–592, 2008.

[CRT04] Emmanuel Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, 2006.

[Don04] David Donoho. Compressed Sensing. *IEEE Trans. Inf. Theory*, 52(4):1289–1306, 2006.

[GJ79] Michael R. Garey, David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* San Francisco, CA: Freeman, 1979.