COMS E6998-9: Algorithms for Massive Data (Fall'25) Sep 22, 2025

Lecture 6: Compressed Sensing

Instructor: *Alex Andoni* Scribe: *Dev Goyal*

1 Review

Definition 1. *A mapping $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is called a (k,δ) Oblivious Subspace Embedding (OSE) if for any k-dimensional subspace $U \subseteq \mathbb{R}^n$:*

$$Pr_{\varphi}\left[\forall z \in U : \frac{\|\varphi(z)\|^2}{\|z\|^2} \in (1 \pm \epsilon)\right] \geq 1 - \delta$$

Definition 2. *A matrix $A \in \mathbb{R}^{m \times n}$ is said to satisfy the (k,ϵ)Restricted Isometry Property (RIP) if for all k-sparse vectors $z \in \mathbb{R}^n$:*

$$\|Az\|_2 \in (1 \pm \epsilon)\|z\|_2$$

2 Guarantees of OSE and RIP

We will start by showing that OSE implies RIP. So we want to prove the following theorem:

Theorem 3. *if A is an (k,$\frac{0.1}{\binom{n}{k}}$)-OSE, then A satisfies (k,ϵ)-RIP.*

Proof. We want to prove:

$$Pr_A\left[\forall z \in S_k : \frac{\|Az\|}{\|z\|} \in (1 \pm \epsilon)\right] \geq 0.9$$

where S_k is the set of all k-sparse vectors in $\mathbb{R}^n$.

We will write S_k as a union of subspaces:

$$S_k = \bigcup_{T \subseteq [n], |T| = k} U_T$$

where $U_T = \{z \in \mathbb{R}^n : z_i = 0 \text{ for } i \notin T\}$
Note that U_T is a k-dimensional subspace of $\mathbb{R}^n$ and the number of such subspaces is $\binom{n}{k}$.

By the definition of OSE and using union bound, we have:

$$
\begin{aligned}
&Pr_A\left[\exists z \in S_k : \frac{\|Az\|}{\|z\|} \notin (1 \pm \epsilon)\right] \\
&\leq Pr_A\left[\exists z \in U_{T_1} : \frac{\|Az\|}{\|z\|} \notin (1 \pm \epsilon)\right] + Pr_A\left[\exists z \in U_{T_2} : \frac{\|Az\|}{\|z\|} \notin (1 \pm \epsilon)\right] + \ldots \\
&\leq \sum_{T \subseteq [n], |T| = k} Pr_A\left[\exists z \in U_T : \frac{\|Az\|}{\|z\|} \notin (1 \pm \epsilon)\right] \\
&\leq \binom{n}{k} \cdot \frac{0.1}{\binom{n}{k}} = 0.1
\end{aligned}
$$

□

Corollary 4. *For fixed $\epsilon > 0$, if A is a scaled Gaussian $m \times n$ matrix with $m = O(k + \log \frac{1}{\delta})$ then there exists A that satisfies (k,ϵ)-RIP $m = O(\log \binom{n}{k}) \leq O(\log (\frac{n}{k})^k) = O(k \cdot \log \frac{n}{k})$*

3 proving the L1 theorem

Recall the Problem setup from last lecture:

We are given an input x and measurements $y = Ax$ and we want to recover x from y using L1 minimization.

$$
L_1(y) = \underset{Ax=y}{\operatorname{argmin}} \|x\|_1 \triangleq x^*
$$

Recall that we define the distance from k sparsity as:

$$
\operatorname{Err}_1^k(x) = \min_{x' \text{ is k-sparse}} \|x - x'\|_1
$$

We want to prove the following theorem:

Theorem 5. *if A is $(4k, \epsilon)$-RIP for a small enough ϵ, then $\|x^* - x\| \leq O(1) \cdot Err_1^k(x)$*

The proof will rely on some intermediate definitions and lemmas.

Definition 6. *We say that A satisfies (k,ϵ) - nullspace property if $\forall T \subseteq [n], |T| \leq k, \forall \eta \in \mathbb{R}^n$ s.t. $A\eta = 0$*

$$
\|\eta_T\|_1 \leq \epsilon \|\eta_{-T}\|_1 \iff \|\eta\|_1 \leq (1 + \epsilon)\|\eta_{-T}\|_1
$$

We will also use the following two lemmas to prove the above theorem.

Lemma 7. *Fix $\forall r \in N$. If A is $((2 + r)k, \epsilon)$-RIP for a small enough ϵ, then A satisfies $(2k, \sqrt{\frac{2}{r}} \cdot \frac{1+\epsilon}{1-\epsilon})$ - nullspace property.*

Lemma 8. *if A satisfies (2k,ϵ)-nullspace property for a small enough ϵ, then for any x, the output of L1 minimization x^* satisfies:*

$$\|x^* - x\|_1 \leq 2 \cdot \frac{1+\epsilon}{1-\epsilon} \cdot Err_1^k(x)$$

We will first prove Lemma 8.

In the proof of the lemma, k-RIPness will be actually good enough. But it's easier to see why $2k$ is useful: consider k-sparse x and a k-sparse potential solution x^* such that $Ax = Ax^* = y$. Then, then the nullspace property of $\eta = x - x^*$ (a $2k$-sparse vectors) immediately implies that $\eta = 0$, and hence $x = x^*$.

Proof of Lemma 8:

let $\eta \triangleq x^* - x \implies A\eta = Ax^* - Ax = y - y = 0$ Let T be the indices of the k largest entries of x in absolute value.

By the definition of L1 minimization, we have:

$$\|x^*\|_1 \leq \|x\|_1 \implies \|x_T^*\|_1 + \|x_{-T}^*\|_1 \leq \|x_T\|_1 + \|x_{-T}\|_1$$

by triangle inequality, we have:

$$\|x_T^*\|_1 \geq \|x_T\|_1 - \|x_T^* - x_T\| = \|x_T\|_1 - \|\eta_T\|_1$$

$$\|x_{-T}^*\|_1 \geq \|\eta_{-T}\|_1 - \|x_{-T}\|_1$$

So we have:

$$\|x_T\|_1 - \|\eta_T\|_1 + \|\eta_{-T}\|_1 - \|x_{-T}\|_1 \leq \|x_T\|_1 + \|x_{-T}\|_1$$

$$\implies \|\eta_{-T}\|_1 \leq \|\eta_T\|_1 + 2\|x_{-T}\|_1$$

By the nullspace property, we have:

$$\|\eta_{-T}\|_1(1-\epsilon) \leq 2\|x_{-T}\|_1$$

$$\implies \|\eta\|_1 \leq \frac{2 \cdot (1+\epsilon)}{1-\epsilon}\|x_{-T}\|_1$$

$$\implies \|x^* - x\|_1 \leq 2 \cdot \frac{1+\epsilon}{1-\epsilon} \cdot \mathrm{Err}_1^k(x)$$

Where the last step follows from the definition of $\mathrm{Err}_1^k(x)$.

This completes the proof of Lemma 8.

We will now prove Lemma 7.

Proof of Lemma 7:

Let $M \triangleq rk$, define $T_0 = T$, T_1 = indices of the largest M entries of η_{-T}, T_2 = indices of the next largest M entries of η_{-T} and so on up to T_s.

define $\eta_0 = \eta_{T_0}, \eta_1 = \eta_{T_1}, \ldots, \eta_s = \eta_{T_s}$.

We can write $\eta = \eta_0 + \eta_1 + \ldots + \eta_s$.

Since $A\eta = 0$, we have:

$$A(\eta_0 + \eta_1) = -A(\eta_2 + \ldots + \eta_s)$$

Taking norms on both sides, we have:

$$\|A\eta_0 + \eta_1\|_2 = \|A(\eta_2 + \ldots + \eta_s)\|_2$$

By triangle inequality (and since η_0 and η_1 are non-zero on disjoint coordinates), we have:

$$\|A\eta_0\|_2 \leq \|A\eta_0 + A\eta_1\|_2 \leq \|A\eta_2\|_2 + \ldots + \|A\eta_s\|_2$$

By the RIP property, we have:

$$(1-\epsilon)\|\eta_0\|_2 \leq \|A\eta_0\|_2 \leq (1+\epsilon)\|\eta_0\|_2$$

and hence

$$\begin{aligned}\|\eta_T\|_2 &\leq \frac{1}{1-\epsilon}(\|A\eta_2\|_2 + \ldots + \|A\eta_s\|_2) \\ &\leq \frac{1+\epsilon}{1-\epsilon}(\|\eta_2\|_2 + \ldots + \|\eta_s\|_2)\end{aligned}$$

Let $\eta^{(i)} = i^{th}$ coord of η. for all $j \geq 2$ and for all $i \in T_j, |\eta^{(i)}| \leq \frac{\|\eta_{T_{j-1}}\|_1}{M}$
This implies that:

$$\|\eta_j\|_2^2 \leq \sum_{i \in T_j} \left(\frac{\|\eta_{T_{j-1}}\|_1}{M}\right)^2 = M \cdot \left(\frac{\|\eta_{T_{j-1}}\|_1}{M}\right)^2 = \frac{\|\eta_{T_{j-1}}\|_1^2}{M}$$

$$\implies \|\eta_j\|_2 \leq \frac{\|\eta_{T_{j-1}}\|_1}{\sqrt{M}}$$

So we have:

$$\|\eta_T\|_2 \leq \frac{1+\epsilon}{1-\epsilon} \cdot \sum_{j=2}^{s} \frac{\|\eta_{T_{j-1}}\|_1}{\sqrt{M}} \leq \frac{1+\epsilon}{1-\epsilon} \cdot \frac{\|\eta_{-T}\|_1}{\sqrt{M}}$$

Now, by the standard inequality between ℓ_1 and ℓ_2 norms (for vectors of dimension $2k$ or sparsity $2k$), we have:

$$\|\eta_T\|_1 \leq \sqrt{2k}\|\eta_T\|_2 \leq \sqrt{2k} \cdot \frac{1+\epsilon}{1-\epsilon} \cdot \frac{\|\eta_{-T}\|_1}{\sqrt{M}}$$

substituting $M = rk$, we have:

$$\|\eta_T\|_1 \leq \sqrt{\frac{2}{r}} \cdot \frac{1+\epsilon}{1-\epsilon} \cdot \|\eta_{-T}\|_1.$$