Compressed Sensing

Lecture 25
Dec 1, 2022
Sparse recovery

Recall:

- Vector $x \in \mathbb{R}^n$ and integer $k$
- $x$ updated in streaming setting one coordinate at a time (can be positive or negative changes)
- Want to find best $k$-sparse vector $\tilde{x}$ that approximates $x$.
  \[ \min_{y, \|y\|_0 \leq k} \| y - x \|_2. \]  
  Optimum solution is clear: take $y$ to be the largest $k$ coordinates of $x$ in absolute value.
- Using Count-Sketch: $O\left(\frac{k}{\epsilon^2 \text{polylog}(n)}\right)$ space one can find $k$-sparse $z$ such that $\| z - x \|_2 \leq (1 + \epsilon)\| y^* - x \|_2$ with high probability.
- Count-Sketch can be seen as $\Pi x$ for some $\Pi \in \mathbb{R}^{m \times n}$ where $m = O\left(\frac{k}{\epsilon^2 \text{polylog}(n)}\right)$. 
**Compressed Sensing**

**Compressed sensing:** we want to create projection matrix $\Pi$ such that for *any* $x$ we can create from $\Pi x$ a good $k$-sparse approximation to $x$

Doable! With $\Pi$ that has $O(k \log(n/k))$ rows. Creating $\Pi$ requires randomization but once found it can be used. Called RIP matrices. First due to Candes, Romberg, Tao and Donoho. Lot of work in signal processing and algorithms.
Theorem (Candes-Romberg-Tao, Donoho)

For every $n, k$ there is a matrix $\Pi \in \mathbb{R}^{m \times n}$ with $m = O(k \log(n/k))$ and a polytime algorithm such that for any $x \in \mathbb{R}^n$, the algorithm given $\Pi x$ outputs a $k$-sparse vector $\tilde{x}$ such that $\|\tilde{x} - x\|_2 \leq O\left(\frac{1}{\sqrt{k}}\right)\|x_{\text{tail}(k)}\|_1$. In particular it recovers $x$ exactly if it is $k$-sparse.

Matrix that satisfies above property are called RIP matrices (restricted isometry property)

Closely connected to JL matrices
Understanding RIP matrices

Suppose \( x, x' \) are two distinct \( k \)-sparse vectors in \( \mathbb{R}^n \)

Basic requirement: \( \Pi x \neq \Pi x' \) otherwise cannot recover exactly

Let \( S, S' \subset [n] \) be the indices in the support of \( x, x' \) respectively. \( \Pi x \) is in the span of columns of \( \Pi_S \) and \( \Pi x' \) is in the span of columns of \( \Pi_{S'} \)

Thus we need columns of \( \Pi_{S \cup S'} \) to be linearly independent for any \( S, S' \) with \( S \neq S' \) and \( |S| \leq k \) and \( |S'| \leq k \). Any \( 2k \) columns of \( \Pi \) should be linearly independent.
Understanding RIP matrices

Suppose $x, x'$ are two distinct $k$-sparse vectors in $\mathbb{R}^n$

Basic requirement: $\Pi x \neq \Pi x'$ otherwise cannot recover exactly

Let $S, S' \subset [n]$ be the indices in the support of $x, x'$ respectively. $\Pi x$ is in the span of columns of $\Pi_S$ and $\Pi x'$ is in the span of columns of $\Pi_{S'}$.

Thus we need columns of $\Pi_{S \cup S'}$ to be linearly independent for any $S, S'$ with $S \neq S'$ and $|S| \leq k$ and $|S'| \leq k$. Any $2k$ columns of $\Pi$ should be linearly independent.

Sufficient information theoretically. Computationally?
Recovery

Suppose we have $\Pi$ such that any $2k$ columns are linearly independent.

Suppose $x$ is $k$-sparse and we have $\Pi x$. How do we recover $x$?

Solve the following:

$$\min \|z\|_0 \quad \text{such that} \quad \Pi z = \Pi x$$

Guaranteed to recover $x$ by uniqueness but NP-Hard!
Suppose we have $\Pi$ such that any $2k$ columns are linearly independent.

Suppose $x$ is $k$-sparse and we have $\Pi x$. How do we recover $x$?

Solve the following:

$$\min \|z\|_0 \text{ such that } \Pi z = \Pi x$$

Guaranteed to recover $x$ by uniqueness but NP-Hard!
Recovery

Instead of solving

\[
\min \| z \|_0 \quad \text{such that} \quad \Pi z = \Pi x
\]

solve

\[
\min \| z \|_1 \quad \text{such that} \quad \Pi z = \Pi x
\]

which is a linear/convex programming problem and hence can be solved in polynomial-time.

If \( \Pi \) satisfies additional properties then one can show that above recovers \( x \).
RIP Property

**Definition**

A \( m \times n \) matrix \( \Pi \) has the \((\epsilon, k)\)-RIP property if for every \( k \)-sparse \( x \in \mathbb{R}^n \),

\[
(1 - \epsilon) \|x\|_2^2 \leq \|\Pi x\|_2^2 \leq (1 + \epsilon) \|x\|_2^2
\]

Equivalent, whenever \(|S| \leq k\) we have

\[
\|\Pi^T_S \Pi_S - I_k\|_2 \leq \epsilon
\]

which is equivalent to saying that if \( \sigma_1 \) and \( \sigma_k \) are the largest and smallest singular value of \( \Pi_S \) then

\[
\frac{\sigma_1^2}{\sigma_k^2} \leq (1 + \epsilon)
\]

Every \( k \) columns of \( \Pi \) are approximately orthonormal.
Recovery theorem

Suppose $\Pi$ is $(\epsilon, 2k)$-RIP with $\epsilon < \sqrt{2} - 1$ and let $\tilde{x}$ be optimum solution to the following LP

$$\min ||z||_1 \text{ such that } \Pi z = \Pi x$$

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

Called $\ell_2/\ell_1$ guarantee. Proof is somewhat similar to the one for sparse recovery with Count-Sketch.

More efficient “combinatorial” algorithms that avoid solving LP.
RIP matrices and subspace embeddings

**Definition**

A $m \times n$ matrix $\Pi$ has the $(\epsilon, k)$-RIP property if for every $k$-sparse $x \in \mathbb{R}^n$,

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

Fix $S \subset [n]$ with $|S| = k$. $S$ defines a subspace of $k$-sparse vectors.

Total of $\binom{n}{k}$ different subspaces. Want to preserve the length of vectors in all of these subspaces.
Fix $S \subset [n]$ with $|S| = k$. $S$ defines a subspace of $k$-sparse vectors. Total of $\binom{n}{k}$ different subspaces. Want to preserve the length of vectors in all of these subspaces.

Given a subspace $\mathcal{W}$ of dimension $d$ we saw that if $\Pi$ is JL matrix with $m = O(d/\epsilon^2)$ rows we have the property that for every $x \in \mathcal{W}$: $\|\Pi x\|_2^2 \simeq (1 \pm \epsilon)\|x\|_2^2$. Via a net argument where net size is $e^{O(k)}$.

If we want to preserve $\binom{n}{k}$ different subspaces need to preserve nets of all subspaces.

Hence via union bound we get $m = O\left(\frac{1}{\epsilon^2} \log\left(e^{O(k)}\binom{n}{k}\right)\right)$ which is $O\left(\frac{k}{\epsilon^2} \log n\right)$.
Fix $S \subset [n]$ with $|S| = k$. $S$ defines a subspace of $k$-sparse vectors. Total of $\binom{n}{k}$ different subspaces. Want to preserve the length of vectors in all of these subspaces.

Given a subspace $W$ of dimension $d$ we saw that if $\Pi$ is JL matrix with $m = O(d/\epsilon^2)$ rows we have the property that for every $x \in W$: $\|\Pi x\|_2^2 \simeq (1 \pm \epsilon)\|x\|_2^2$. Via a net argument where net size is $e^{O(k)}$.

If we want to preserve $\binom{n}{k}$ different subspaces need to preserve nets of all subspaces

Hence via union bound we get $m = O\left(\frac{1}{\epsilon^2} \log(e^{O(k)}\binom{n}{k})\right)$ which is $O\left(\frac{k}{\epsilon^2} \log n\right)$.

Other techniques give $m = O(k^2/\epsilon^2)$. 