13 Apr 2022 Sparse recovery

Announcement:

Quiz 2 is now graded.

Approx mapping of Q2 grade to letter grade

18-24	A
10-18	B
< 10	C

Quiz 3 will be a week from today.

Email me if $\exists$ conflict.

Rubric for 3a, 3b, 3c was meant to be:

- 4 pts for 3a algo
- 4 pts for 3b algo
- 4 pts for proving (at least) one of those two algorithms is correct.

If you had a correctness proof (resembling the solution set) and you got $\emptyset$ on 3c, maybe the grader overlooked your proof

$\Longrightarrow$ request a regrade.

Recall: random linear transformation
$\mathbb{R}^n \longrightarrow \mathbb{R}^{O(\log n)}$ approximately
preserves distances between all pairs
in a set of n points.

Today we'll see:
(a) It also preserves distances between
all pairs of <u>sparse</u> <u>vectors</u>.

<u>Def.</u> A vector $x \in \mathbb{R}^n$ has <u>sparsity pattern</u>
$J \subset [n]$ if $x_i = 0 \ \forall i \notin J$.

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$ has sparsity pattern $\{3,5\}$.

x is s-sparse if it has
sparsity pattern J for some $|J| \le s$.
(x has at most s non-zero coordinates)

E.g. partial derivatives of natural images
are nearly sparse.
"bag of words" representation of a
document is usually very sparse.

(b) If R is a (Gaussian) random matrix representing a lin. transformation $\mathbb{R}^n \to \mathbb{R}^{O(\log n)}$ then given Rx where x is sparse we can efficiently "invert" R and recover x.

<u>Lem.</u> If $R: \mathbb{R}^s \longrightarrow \mathbb{R}^k$ is a lin transf. with indep Gaussian $N(0, \frac{1}{k})$ entries and $k > \frac{72}{\varepsilon^2}\left(\ln\left(\frac{2}{\delta}\right) + 7s\right)$ then with probability at least $1-\delta$

$$\forall x \in \mathbb{R}^s \quad (1-\varepsilon)\|x\|_2^2 \leq \|Rx\|_2^2 \leq (1+\varepsilon)\|x\|_2^2.$$

<u>Proof.</u> Obs. $\|Rx\|_2^2 = \langle x, R^T R x\rangle$.

Let's calculate $\mathbb{E}[R^T R]$.

$$R = \begin{bmatrix} - r_1^T - \\ - r_2^T - \\ \vdots \\ - r_k^T - \end{bmatrix} \qquad R^T R = \sum_{i=1}^{k} r_i r_i^T$$

Each r_i is $N(0, \frac{1}{k}\mathbb{1})$ so $\mathbb{E}[r_i r_i^T] = \frac{1}{k}\mathbb{1}$.

$$\mathbb{E}[R^T R] = \sum_{i=1}^{k} \mathbb{E}[r_i r_i^T] = k \cdot \frac{1}{k} \cdot \mathbb{1} = \mathbb{1}.$$

We're trying to prove with prob. $\geq 1-\delta$

$$(1-\varepsilon)\langle x, \mathbb{E}(R^T R)x\rangle \leq \langle x, R^T R x\rangle \leq (1+\varepsilon)\langle x, \mathbb{E}(R^T R)x\rangle$$

Prop 4.3 from lecture notes says for Gaussian R,

$$\Pr\left(\sup_{x\in\mathbb{R}^s}\left|\frac{\langle x, R^TRx\rangle}{\langle x, \mathbb{E}(R^TR)x\rangle}-1\right|\ge\varepsilon\right)\le 2e^{7s-\frac{1}{72}\varepsilon^2 k}$$

The assumption $k>72\varepsilon^{-2}\left(\ln\left(\frac{2}{\delta}\right)+7s\right)$ was designed so RHS would be $\le\delta$.
When event on LHS doesn't happen then $\forall x$

$$(1-\varepsilon)\langle x,\mathbb{E}(R^TR)x\rangle \le \langle x, R^TRx\rangle \le (1+\varepsilon)\langle x,\mathbb{E}(R^TR)x\rangle.$$

Prop. If $s\in\mathbb{N}$, $s\ge 3$, $n\in\mathbb{N}$, $n\ge s$, $0<\varepsilon,\delta<1$ and $k>\frac{72}{\varepsilon^2}\left(\ln\left(\frac{2}{\delta}\right)+7s+s\ln(n)\right)$ then with probability $\ge 1-\delta$

$\ln\left(\frac{2}{\delta'}\right)+7s$ where $\ln\left(\frac{2}{\delta'}\right)=\ln\left(\frac{2}{\delta}\right)+s\ln n$

$\forall$ s-sparse $x\in\mathbb{R}^n$

$$(1-\varepsilon)\|x\|_2^2\le\|Rx\|_2^2\le(1+\varepsilon)\|x\|_2^2.$$

Proof. If J is any sparsity pattern, and $x\in\mathbb{R}^n$ has sparsity pattern J,

$$k\underbrace{\begin{bmatrix} & R & \end{bmatrix}}_{n}\overbrace{\begin{bmatrix}0\\0\\*\\0\\*\\0\\*\end{bmatrix}}^{x\in\mathbb{R}^n}=\underbrace{R_J}_{(k\times s)}\underbrace{x_J}_{\in\mathbb{R}^s}$$

where R_J selects the columns of R indexed by J, x_J selects the entries of x indexed by J.

So $\Pr\left(\forall x_J \in \mathbb{R}^s \quad (1-\varepsilon)\|x_J\|_2^2 \leq \|R_J x_J\|_2^2 \leq (1+\varepsilon)\|x_J\|_2^2\right)$
$\geq 1-\delta'$

and $\ln\left(\frac{2}{\delta'}\right) = \ln\left(\frac{2}{\delta}\right) + s\ln(n)$

$$\frac{2}{\delta'} = \frac{2}{\delta}\cdot n^s$$

$$\delta' = \delta \cdot n^{-s}.$$

Union Bound over all possible J with $|J|=s$.
Number of these is $\binom{n}{s} \leq n^s$.
So probab that $\exists J$, $|J| \leq s$, and $\exists x$ with sparsity ~~pattern~~ J s.t.

$$\|Rx\|_2^2 - \|x\|_2^2 > \varepsilon\|x\|_2^2$$

is less than $\delta' \cdot \binom{n}{s} < \delta$. QED.

<u>Def:</u> A matrix R satisfies the s-restricted isometry property (RIP) with constant ε if

$\forall$ s-sparse $x \in \mathbb{R}^n$ $\quad (1-\varepsilon)\|x\|_2^2 \leq \|Rx\|_2^2 \leq (1+\varepsilon)\|x\|_2^2.$