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 3 conflict.

Rubric for 3a, 3b, 3c was meant to be:
- 4 pts for 3a only
- 4 pts for 3b only
- 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 0 on 3c,
maybe the grader overlooked your proof
⇒ request a regrade.
Recall: random linear transformation $\mathbb{R}^n \rightarrow \mathbb{R}^0(k \cdot n)$ 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 **sparse vectors**.

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

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

$x$ is $s$-sparse if it has sparsity pattern $J$ for some $|J| \leq s$. ($x$ 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.
If $R$ is a (Gaussian) random matrix representing a linear transformation $\mathbb{R}^d \rightarrow \mathbb{R}^k$ then given $Rx$ where $x$ is sparse we can efficiently "invert" $R$ and recover $x$.

**Lemma.** If $R : \mathbb{R}^d \rightarrow \mathbb{R}^k$ is a linear transform with independent Gaussian $N(0, 1/k)$ entries and $k > \frac{2 \epsilon^2}{\delta^2} \left( \ln \left( \frac{1}{\delta} \right) + 7 \delta \right)$ then with probability at least $1 - \delta$,

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

**Proof.** Observe $\|Rx\|_2^2 = \langle x, R^TRx \rangle$.

Let's calculate $\mathbb{E} [R^TR]$.  

$$R = \begin{bmatrix} -\frac{1}{\sqrt{k}} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ -\frac{1}{\sqrt{k}} & \ldots & 0 \end{bmatrix} \quad R^TR = \sum_{i=1}^{k} r_i r_i^T$$  

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

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

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

$$(1 - \epsilon) \langle x, \mathbb{E}(R^TR)x \rangle \leq \langle x, R^TRx \rangle \leq (1 + \epsilon) \langle x, \mathbb{E}(R^TR)x \rangle$$
Prop 4.3 from lecture notes says for Gaussian $\mathbb{R}^n$,

$$
\Pr \left( \sup_{x \in \mathbb{R}^n} \left| \frac{\langle x, R^T R x \rangle}{\langle x, E(R^T R) x \rangle} - 1 \right| \geq \epsilon \right) \leq \frac{2}{\sqrt{2 \pi}} \epsilon^2 \frac{7 \epsilon^2}{k \epsilon^2}.
$$

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

$$(1 - \epsilon) \langle x, E(R^T R) x \rangle \leq \langle x, R^T R x \rangle \leq (1 + \epsilon) \langle x, E(R^T R) x \rangle.$$

Prop: If $s \in \mathbb{N}$, $s \geq 3$, $n \in \mathbb{N}$, $n \geq 5$, $0 < \delta < 1$ and $k > \frac{72 \epsilon^2}{\sqrt{\ln \left( \frac{2}{\delta} \right) + 7 \epsilon} + 5 \ln(n)}$ then with probability $\geq 1 - \delta$

$$(1 - \epsilon) \| x \|_2^2 \leq \| R x \|_2^2 \leq (1 + \epsilon) \| x \|_2^2,$$

where $\ln \left( \frac{2}{\delta} \right) = \ln \left( \frac{2}{\delta} \right) + 5 \ln(n)$.

Proof: If $J$ is any sparsity pattern, and $x \in \mathbb{R}^n$ has sparsity pattern $J$, $k \begin{bmatrix} \mathbb{R}^n \end{bmatrix} = \mathbb{R}_J \times_J \mathbb{R}_S^k$.
where \( R_J \) selects the columns of \( R \) indexed by \( J \), \( x_J \) selects the entries of \( x \) indexed by \( J \).

\[
S = \mathbb{P}( \forall x \in \mathbb{R}^s \quad (1 - \varepsilon) \|x_J\|_2^2 \leq \|R_j x_J\|_2^2 \leq (1 + \varepsilon) \|x_J\|_2^2 ) > 1 - \delta' \]

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

\[
\frac{\varepsilon}{\delta'} = \frac{\varepsilon}{\delta} \cdot n^s \quad \delta' = \delta \cdot n^s .
\]

Union Bound over all possible \( J \) with \( |J| = s \).

Number of these is \( \binom{n}{s} \leq n^s \).

So, probab that \( \exists J, |J| = s \), and \( \exists x \) with sparsity problem \( J \) s.t.:

\[
\|Rx\|_2^2 - \|x\|_2^2 > \varepsilon \|x\|_2^2
\]

is less than \( 8' \cdot \binom{n}{s} < \delta \). \( \text{Q.E.D.} \)

**Def:** A matrix \( R \) satisfies the \( s \)-restricted isometry property (RIP) with constant \( \varepsilon \) if:

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