Project Ideas in Distributed Environment

A few ideas that were shared in the class for project topics:

- Consider 2 binary matrices A and B. Is the sum of these two matrices invertible. Known lower bound $\omega(n)$ and upper bound is $O(n^2)$ in bits

- Consider a set of constraints for a linear programming problem, which are exclusively split between multiple processes. Is it possible to solve the linear programming problem efficiently.

- Consider two users with partial information about a graph. Is there a perfect matching that based on the union of the information both the users have?

$\ell_1$ Regression

Until now we have looked at $\ell_2$ regression. However in $\ell_2$ regression, the cost is more sensitive to outliers in the data set than some other $\ell_p$ regressions (like $\ell_1$). We now want to use $\ell_1$ regression in problems that are prone to the issue of outliers.

The $\ell_1$ regression problem is defined using $\ell_1$ norm. For any vector $x \in \mathbb{R}^d$, the $\ell_1$ norm is defined as

$$\|x\|_1 = \sum_{i=1}^{d} |x_i|$$

and for any matrix $A \in \mathbb{R}^{n \times d}$, and vector $b \in \mathbb{R}^n$, the $\ell_1$ regression is defined as

$$\min_x \|Ax - b\|_1$$

The $\ell_1$ regression problem can be solved optimally in $O(nd)$ time using linear programming. However, with datasets (large $n$), this can become computationally challenging to solve perfectly. Therefore to achieve better time complexity for the problem, we want to use sketching to obtain an approximate solution that is close to the ideal solution with high probability. We want to use our trusty sketching matrices to reduce the problem complexity.

Well-Conditioned Bases

$\ell_1$ norm does not have all of the nice properties of the $\ell_2$ norm. For example, while orthonormal bases do not change the $\ell_2$ norm, they can change the $\ell_1$ norm of a vector.
Therefore we need to derive some properties of the $\ell_1$ norm that can help us, when we try to sketch the matrices.

Some simple properties for $\ell_1$ norms:

- If $x \in \mathbb{R}^d$, then $\|x\|_1 \leq \sqrt{d}\|x\|_2$: Using Cauchy-Schwarz inequality
  \[
  \sum_i |x_i \cdot 1| \leq \sqrt{\sum_i (x_i)^2 \sum_i (1)^2}
  \Rightarrow \sum_i |x_i \cdot 1| \leq \sqrt{d \cdot \sum_i (x_i)^2}
  \Rightarrow \|x\|_1 \leq \sqrt{d} \cdot \|x\|_2
  \]

- $\|x\|_2 \leq \|x\|_1$
  \[
  \sum_i x_i^2 = \sum_i \|x_i\|^2 \leq (\sum_i \|x_i\|)^2
  \Rightarrow \|x\|_2 \leq \|x\|_1
  \]

We now see an analogue of orthonormal basis for the case of $\ell_1$ norm. For this let $A = QW$, where $Q$ is a $n \times d$ matrix with full column rank. We define the $Q,1$ norm of a vector as follows:

$$\|z\|_{Q,1} = \|Qz\|_1$$

Aside: Q,1 norm is a norm

To show that a function is a norm, we need to show 3 properties:

- $\|x\|_{Q,1} = 0$ iff $x = 0$
  \[
  \|0\|_{Q,1} = \|Q \cdot 0\|_1 = \|0\|_1 = 0
  \]
  and if $\|z\|_{Q,1} = 0$ then $\|Qz\|_1 = 0$ which implies $Qz = 0$. As $Q$ has full column rank $z = 0$.

- $\|cx\|_{Q,1} = |c|\|x\|_{Q,1}$
  \[
  \|cx\|_{Q,1} = \|Q(cx)\|_1 = |c|\|Qx\|_1 = |c|\|x\|_{Q,1}
  \]

- $\|x + y\|_{Q,1} \leq \|x\|_{Q,1} + \|y\|_{Q,1}$
  \[
  \|x + y\|_{Q,1} = \|Q(x + y)\|_1
  = \|Qx + Qy\|_1
  \leq \|Qx\|_1 + \|Qy\|_1
  \leq \|x\|_{Q,1} + \|y\|_{Q,1}
  \]
Consider the unit ball formed with the $Q,1$ norm. Let $C = \{z \in \mathbb{R}^d : \|z\|_{Q,1} \leq 1\}$. C has two properties:

- C is symmetric: For each vector $z$ in C, $-z$ also belongs in C
- C is convex: $\|\lambda x + (1 - \lambda)y\|_{Q,1} \in C, \forall \lambda \in [0,1], x, y \in C$

\[
\|\lambda x + (1 - \lambda)y\|_{Q,1} \leq \|\lambda x\|_{Q,1} + (1 - \lambda)\|y\|_{Q,1} \\
\leq \lambda\|x\|_{Q,1} + (1 - \lambda)\|y\|_{Q,1} \\
\leq \lambda + (1 - \lambda) = 1
\]

**Theorem 1** (Lowner-John Ellipsoid Theorem). Let $K$ be a convex body in $\mathbb{R}^d$ which is symmetric about origin. Then there is a ellipsoid $E$, also symmetric about origin so that

$$E \subseteq K \subseteq \sqrt{d}E$$

where $E = \{z \in \mathbb{R}^d : z^TFz \leq 1\}$ and $F = G^TG$ for some $G$. Here $F$ defines an ellipsoid.

For our case letting $K = \{z : \|z\|_{Q,1} \leq 1\}$, it can be shown that for the $Q,1$ norm, the matrix $F$ given by the above theorem satisfies

$$\sqrt{z^TFz} \leq \|z\|_{Q,1} \leq \sqrt{d}\sqrt{z^TFz}$$

(1)

Let’s define $U = QG^{-1}$ and let $z = G^{-1}x$, then using the previous inequality we get

$$\|Ux\|_1 = \|QG^{-1}x\|_1 = \|Qz\|_1 = \|z\|_{Q,1}$$

and

$$z^TFz = x^T(G^{-1})^TG^TG^{-1}x = \|x\|_2^2.$$

Now, using (1),

$$\|x\|_2 \leq \|Ux\|_1 \leq \sqrt{d}\|x\|_2$$

Using the upper and lower bounds on $\|x\|_1$ in terms of $\|x\|_2$, we obtain that

$$\|x\|_1/\sqrt{d} \leq \|x\|_2 \leq \|Ux\|_1 \leq \sqrt{d}\|x\|_2 \leq \sqrt{d}\|x\|_1$$

and therefore $\|x\|_1 \leq \sqrt{d}\|Ux\|_1 \leq d\|x\|_1$. Defining $U' := \sqrt{d}U$, we get $\|x\|_1 \leq \|U'x\|_1 \leq d\|x\|_1$. $U'$ here is called a well-conditioned basis as it preserves the $\ell_1$ norm of a vector up to small factors that depend only on the dimension of the column space. Importantly, note that the column space of $U'$ is the same as that of the matrix $A$ that we started with.

**Net construction**

We construct $\ell_1$ subspace embeddings using a net argument similar to that of the $\ell_2$ subspace embedding using Gaussian random variables. Here we construct a net for the the set

$$\{U'x : \|x\|_1 = 1\}$$
where $U'$ is as defined previously.

If a subspace embedding $S$ satisfies $\|Sy\|_1 = (1 \pm \varepsilon)\|y\|_1$ for all $y$ in the above set, then it is clear that $\|SAx\|_1 = (1 \pm \varepsilon)\|Ax\|_1$ for all $x$. Let $B := \{x \in \mathbb{R}^d \mid \|x\|_1 = 1\}$ be the $\ell_1$ unit ball. Similar to the $\ell_2$ net, we can show that there is a $\gamma$-net $N$ of size $\leq (1 + 2/\gamma)^d$ for the set $B$ i.e., for every $x \in \mathbb{R}^d$ with $\|x\|_1 = 1$, there is a $y \in N$ such that

$$\|x - y\|_1 \leq \gamma.$$ 

Now let $N'$ be a $\gamma/d$-net for $B$. From above, we have $|N'| \leq (1 + 2d/\gamma)^d$. Now consider the set

$$M = \{U'y \mid y \in N'\}.$$ 

Now, for any $x$ with $\|x\|_1 = 1$, we have that there is a vector $y \in N'$ with $\|y\|_1 = 1$ and $\|x - y\|_1 \leq \gamma/d$ which implies that $\|U'(x - y)\|_1 \leq d\|x - y\|_1 \leq d(\gamma/d) = \gamma$. Additionally, $U'y \in M$. Thus, $M$ is a $\gamma$ net for the set $\{U'x \mid \|x\|_1 = 1\}$. 

4