CS 395T: Sublinear Algorithms Fall 2014

Lecture 14 — Oct 14, 2014

Prof. Eric Price *Scribe: Xiangru Huang*

1 Overview

In the last lecture : lower bounds

In this lecture: problems for final project, RIP-1, expanders, SMP, SSMP

2 Problems for final project

- **Adaptivity in Sparse Recovery**
  So far, we can choose matrix A independent of x and estimate x from Ax using $O(klog\frac{n}{k})$ space. What if we choose $< v_1, x >, \ < v_2, x >, \ \cdots, \ < v_m, x >$ where v_i depend on $v_1, \ v_2, \ \cdots, \ v_{i-1}$?

- **k-sparse**
  Given $A \sim \mathcal{N}(0, I_{M\times N})$, we know that if $m = O(klog\frac{n}{k})$ (hence RIP), then L1 minimization and IHT works.
  But for $m << klog\frac{n}{k}$, doing ℓ_1/ℓ_1 recovery is impossible. What if x is exactly k-sparse?

- **Compressed Sensing with priors**
  e.g. If x follows some distribution, what could happen? This is a general question.

- **Count-sketch**
  Using top $2k$ coordinates, we can do $(1+\epsilon)$ ℓ_2/ℓ_2 approximate recovery where $m = O(\frac{k}{\epsilon}logn)$
  If we use top k coordinates, then $m = O(\frac{k}{\epsilon^2}logn)$.
  Better analysis used in [MP14] might give $O(\frac{k}{\epsilon^2} + \frac{k}{\epsilon}logn)$.

- **LASSO vs sqrt LASSO**
  LASSO finds $\underset{x}{argmin}\|y - Ax\|_2^2 + \lambda\|x\|_1$. And sqrt LASSO finds $\underset{x}{argmin}\|y - Ax\|_2 + \lambda\|x\|_1$.
  Compare these two algorithms. Which one is better under certain situations?

- **Random order streams**
  Given $x_1, \ x_2, \ \cdots, \ x_m \sim D$ over F_2^n. e.g. $x = Ay$ (mod 2) where A is a n by $\frac{n}{2}$ matrix. One of the two cases is true:

  - D is $\frac{n}{2}$-dimension subspace
  - D is uniform

  How many samples to distinguish these two cases?

3 Sparse Matrix with RIP-1

Consider 0-1 matrix $A \in \mathbb{R}^{m\times n}$ with $d = O(logn)$ ones per column. For A, we can achieve fast multiplication and $O(logn)$ update time. The problem is A can't satisfy RIP unless $m = O(k^2)$. But A can satisfy the following RIP-1 property [BGIKS08].

Definition 1. *A has RIP-1 of (k, ϵ) if $\forall k$-sparse x, $\|Ax\|_1 = (1 \pm \epsilon)\|x\|_1$.*

Definition 2. *$G = (U, V, E)$ is a bipartite graph with left-degree d. $n = |U|$, $m = |V|$. $N(S)$ denotes the neighbors of S. G is a unbalanced bipartite expander of (k, ϵ) if*

$$\forall S \subset U, |S| \leq k \Rightarrow N(S) \geq (1-\epsilon)d|S|$$

Claim 3. *random graph with $d = \frac{1}{\epsilon}log\frac{n}{k}, m = \frac{1}{\epsilon^2}klog\frac{n}{k}$ is expander w.h.p.*

Claim 4. *There exist explicit expander constructions for $\forall\alpha > 0, d = O(logn\frac{logk}{\epsilon})^{1+\frac{1}{\alpha}}, m = k^{1+\alpha}d^2$*

The adjacency matrices of expander graphs, scaled by a factor of $\frac{1}{d}$, satisfy RIP-1.

Theorem 5. *(k, ϵ) expander $\Rightarrow$ $(k, 2\epsilon)$ RIP-1*

Proof. $\forall S$ of size k, $d = \frac{1}{\epsilon}log\frac{n}{k}$, $m = \frac{2}{\epsilon}kd$. Consider all kd edges, define $V_1, V_2, \cdots, V_{kd} \in [m] i.i.d.$ Let C_j denotes the event that V_j collide with $V_1, V_2, \cdots, V_{j-1}$. We have

$$Pr[C_j] \leq \frac{j-1}{m} \leq \frac{kd}{m} \leq \frac{\epsilon}{2}$$

$$Pr[\sum_{i=1}^{kd} C_i > \epsilon kd] \leq 2^{-\frac{\epsilon kd}{2}} = 2^{-\Omega(\frac{\epsilon kd}{2})}$$

By union bound, the probability that $N(S) \geq (1-\epsilon)d|S|$ holds for all $S \subset U, |S| = k$ is at least $1 - 2^{-\Omega(klog\frac{n}{k})}\binom{n}{k} = 1 - 2^{-\Omega(klog\frac{n}{k})}$. Then sum over all sizes we can show $\sum_{i=0}^{n} 2^{-ilog\frac{n}{k}} << 1$, which means the theorem is true w.h.p. □

4 From RIP-1 to sparse recovery

There is a viable natual algorithm from count-sketch. You can check out the explanation in the book [FSR].

Algorithm: Natural Alg
$x^{(0)} = 0$
For $r = 0, 1, \cdots, T\{$
  $u_i \leftarrow \underset{j\in N(i)}{median}(y - Ax^{(r)})_j$
  $x^{(r+1)} \leftarrow H_k(x(r) + u)$
$\}$
output $x^{(T)}$

Here we introduce SMP [BIR08] and SSMP [BI09]. The following algorithm is SMP.

Algorithm: SMP
$x^{(0)} = 0$
Repeat T times

$$u_i \leftarrow \underset{j\in N(i)}{median}(y - Ax^{(r)})_j$$
$$x^{(r+1)} \leftarrow H_k(x(r) + H_{2k}(u))$$

output $x^{(T)}$

SSMP is similar to SMP except the updates is done sequentially.

Algorithm: SSMP
1) Let $x^{(0)} = 0$
2) For $r = 1, 2, \cdots, T = O(log(\|x\|_1/\|e\|_1))$

   a) For $t = 1, 2, \cdots, 10k$

   - $u_i \leftarrow \underset{j\in N(i)}{median}(y - Ax^{(r)})_j$
   - Let i be the largest term of u
   - Let $x^{(r)} = x^{(r)} + u_i e_i$

   b) Let $x^{(r)} = H_k(x^{(r)})$

3) Report $x' = x^{(T)}$

Here we prove SSMP. A more detailed proof can be found at [BI09]

Proof. $y = Ax = (\sum_{i\in S} a_i x_i) + e$
If $\sum \|a_i x_i\|_1 \leq (1+\epsilon)\|\sum a_i x_i\|_1$, [1] we can show $\exists a_i, i, s.t.$

$$\|y - a_i x_i\|_1 \leq (1 - \frac{1}{10k})\|y\|_1$$

Therefore each step $x^{(r)} = x^{(r)} + u_i e_i$ decreases $\|y - Ax^{(r)}\|$ by $1 - \frac{\Omega(1)}{k}$, after $O(k)$ steps, we have

$$\|y - Ax^r\| \leq \frac{1}{10}\|y - Ax^{(r-1)}\|$$
$$\Rightarrow \|x^{(r)} - x\|_1 \leq \frac{1}{5}\|x^{(r)} - x\|_1 + O(\|e\|_1)$$
$$\Rightarrow \textit{It takes } O(log(\|x\|_1/\|e\|_1) \textit{ iterations to get error to } O(\|e\|_1)$$

[1] you will prove this in your homework

□

References

[AMS99] Noga Alon, Yossi Matias, Mario Szegedy. The Space Complexity of Approximating the Frequency Moments. *J. Comput. Syst. Sci.*

[MP14] Gregory T. Minton and Eric Price. Improved Concentration Bounds for Count-Sketch SODA(best student paper) *2014.*, 58(1):137–147, 1999.

[BGIKS08] Radu Berinde and Anna C. Gilbert and Piotr Indyk and Howard J. Karloff and Martin J. Strauss. Combining geometry and combinatorics: A unified approach to sparse signal recovery. *CoRR.*

[BI09] Berinde, Radu, and Piotr Indyk. Sequential sparse matching pursuit. *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on. IEEE*, 2009.

[BIR08] Berinde, Radu, Piotr Indyk, and Milan Ruzic. Practical near-optimal sparse recovery in the l1 norm. *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on. IEEE*, 2008.

[FSR] Foucart, Simon, and Holger Rauhut. A mathematical introduction to compressive sensing. *Berlin: Springer, 2013.*