15-859E: Advanced Algorithms **CMU, Spring 2015**
Lecture #13: Compressed Sensing January 26, 2016
Lecturer: Anupam Gupta Scribe: Balaji N and Abhishek Sharma

1 Classical Wine Bottle problem:

Given: 1000 wine bottles out of which one is poisoned.

Poisoned bottle can only be determined, if someone drinks in which results his/her death. You can mix two wine from two bottles and if any of the bottle is poisoned, the mixture becomes poisonous.

Goal:

To figureout the minimum number of people required to determine the poisoned bottle.

Trivial solution: 1000

Decision Tree Based Strategy:

- Divide 1000 in two groups of 500 each and give the first part to person-1.
- Divide both groups of 500 in two groups of 250 each, give first part of each group to person-2.
- Continue the same till the group had one bottle remaining.

In general, if there are N bottles of wine, and one of them is poisoned, $\lfloor logN \rfloor +$ bottles are necessary (by counting argument) and $\lceil logN \rceil$ are sufficient.It is easier to see this in matrix form.

y = observed boolean vectors of illness of m people

$$A = \begin{pmatrix} \cdots A_1 \cdots \\ \cdots A_2 \cdots \\ \vdots \\ \cdots A_m \cdots \end{pmatrix}$$

Input matrix for m people.

$y = Ax$, where x is unknown vector with exactly one, rest are zero. So we observed that i^{th} column of A if $X_i = 1$ $X_{[n]-i}$ =0.

Now, if k bottles were poisoned, the problem would be called Group Testing Problem.

- Main difficulty is that, we see only boolean outputs zero or non-zero; there is nothing to say how sick a person had fallen.
- For a moment, let us assume we can observe $y = Ax$. Not just the boolean form but the exact value.
- We claim that recovery is possible as long as
  $\sum_{i \in S} A_i \neq \sum_{i \in S'} A_i$
  for all S, S' such that $S \neq S'$ and $|S|= |S'|= k$

Proof : Fix $S \neq S'$

$\sum_{i \in S} A_i = \sum_{i \in S'} A_i \Leftrightarrow \sum_{i \in S-S'} A_i = \sum_{i \in S'-S} A_i$

Therefore, $\mathbf{Pr}[\sum_{i \in S-S'} A_i = \sum_{i \in S'-S} A_i] \leq (\frac{1}{2})^m$

$\mathbf{Pr}[\sum_{i \in S-S'} A_i - \sum_{i \in S'-S} = A_{i_0}] \leq (\frac{1}{2})^m$

- Random A is a goal matrix, where $x \neq x' \leftarrow Ax = x'$, for sparse x, x', given A, $y = Ax$, find x.

2 Overview: Compressed sensing

Observations are expensive but computation is cheaper, inputs provided are sparse.

Question : We have here is, can we observe the input in compressed form and uniquely recover all sparse inputs?

Secondly, we also need efficient recovery algorithms.

Idea : Use LP Approach.

The goal is to recover x from $y = Ax$. Recovery is not possible in general as x is not unique if $m < n$. But, from the input assumption, we want sparse set x.

Thus, minimize $||X||_0$ given $Ax = y$, where x is a variable and y is a given observation. But this minimization or optimization problem is $NP-hard$, because of its non-convex behavior. So, we relaxed to L_1.

Minimize $||x||_1, Ax = y$

- L_1 is solvable because it is aan LP.
- L_1 is better than L_2, because a little deviation in the slope in L_1 would change location drastically whereas not so in L_2.

3 Algorithm:Candy-Tao:

- Choose ϕ to be (δ, k) be RIP matrix. $\Rightarrow [\forall ||x||_2 = 1$,x$isksparse 1 - \delta \leq ||\phi x||_2^2 \leq 1 + \delta]$
- $y = \phi x$ is observed.
- $x^* \equiv \min_{\tilde{x}} ||x||_1, \tilde{x} = y$
- Output x^*

Proof : Consider the error vector $h = x^* - x$,

we will show that $||h||_2 = 0$

Breakup the coordinates into disjoint sets $T_0, T_1, T_2 ....$

$T_0 = k$ non-zero coordinates in the unknown x.

$T_1 = k$ largest coordinates of $h_{[n]-T_0}$

$T_2 = k$ largest coordinates of $h_{[n]-T_0-T_1}$

So, $h = \sum_{j \geq 0} h_{Tj}$

$Part-1$: bound $||h_{T_2\cup T_3\cup ....}||_2$

$||h_{T_2}||_2\leq \sqrt{k}||h_{T2}||_\infty$

$||h_{T_2}||_2\leq \frac{1}{\sqrt{k}}||h_{T2}||_1$

$||h_{T_3}||_2\leq \frac{1}{\sqrt{k}}||h_{T2}||_1 \rightarrow A$

$||h_{T_0\cup T_1}||_2\leq \sum_{j\geq 2}||h_{Tj}||_2$

$||h_{T_0\cup T_1}||_2\leq \frac{1}{\sqrt{k}}|h_{T_1\cup T_2\cup ....}||_1$

$||h_{T_0\cup T_1}||_2\leq \frac{1}{\sqrt{k}}|h_{T_0^c}||_1$

we optimize the $||x^*||_1$

$||x^*||\leq ||x||_1$

$\Rightarrow ||x-h||_1\leq ||x||_1$

$||(x-h)_{T_0}||_1+||(x-h)_{T_o^c}||_1\leq ||x_{T_0}||_1$

$||x_{T_0}||_1-||h_{T_0}||_1+||h_{T_0^c}||_1\leq ||x_{T0}||_1$

$\Rightarrow ||h_{T0^c}||_1\leq ||h_{T0}||_1$

$||x_{T_0^c}||_1\leq ||(x-h)_{T0}||_1+||h_{T0}||_1 \rightarrow B$

From A

$||h_{T0\cup T1}||_2\leq \frac{1}{\sqrt{k}}||h_{T_0^c}||$

$Part-2$:

Bound $||h_{T_0\cup T_1}||_2$

Use the RIP property and use the property that ϕ is $(2k-\delta)RIP$

$(1-\delta)||h_{T_0\cup T_1}||_2^2\leq ||\phi h_{T_o\cup T_1}||_2^2$

$(1-\delta)||h_{T_0\cup T_1}||_2^2= ||\phi h_{T_o\cup T_1}, \phi h - \sum_{j\geq 2}\phi h_{T_j}||$

$(1-\delta)||h_{T_0\cup T_1}||_2^2= (\phi h_{T_0\cup T_1}, -\sum_{j\geq 2}\phi h_{Tj})$ we need to bound this, notice that,

$4\langle \phi h_{T_0}, \phi h_{T_j}\rangle = ||\phi(h_{T_0}+h_{T_j})||_2^2 - ||\phi(h_{T_0}-h_{T_j})||_2^2$

$4\langle \phi h_{T_0}, \phi h_{T_j}\rangle = 4\delta \; ||(h_{T_0})||_2||h_{T_j}||_2$

Overall we conclude that $|\langle \phi h_{T_0}, _- \sum_{j\geq 2}\phi h_{Tj}\rangle|\leq \delta ||h_{T_0}||_2 \sum_{j\geq 2}||h_{T_j}||$

$|\langle \phi h_{T_0}, _- \sum_{j\geq 2}\phi h_{Tj}\rangle|\leq \delta ||h-T_0\cup T_1||_2$

Similarly for T_1

$\langle h_{T_1}, \sum_{j\geq 2}\phi h_{T_j}\rangle$

$|\langle h_{T_0\cup T_1}, \sum_{j\geq 2}\phi h_{T_j}\rangle|\leq 2\delta ||h_{T_0\cup T_1}||_2 \sum_{j\geq 2}||h_{T_j}||_2$

$\Rightarrow (1-\delta)||h_{T_0\cup T_1}||_2^2\leq 2\delta ||h_{T0\cup T_1}||_2 \sum_{j\geq 2}||h_{T_j}||_2$

$\Rightarrow (1-\delta)||h_{T_0\cup T_1}||_2^2\leq 2\delta ||h_{T_0\cup T_1}||_2 \frac{1}{\sqrt{k}}||h_{T_0}||_1$

$\Rightarrow ||h_{T_0\cup T_1}||_2^2\leq \frac{2\delta}{1-\delta}\frac{1}{\sqrt{k}}||h_{T_0}||_1$

$\Rightarrow ||h_{T_0\cup T_1}||_2^2\leq \frac{2\delta}{1-\delta}\frac{1}{\sqrt{k}}\sqrt{k}||h_{T_0}||_2$

$\Rightarrow (1-\delta)||h_{T_0\cup T_1}||_2\leq 2\delta ||h_{T_0\cup T_1}||_2$

$\Rightarrow (1-\delta)||h_{T_0 \cup T_1}||_2 \leq 0$

$\Rightarrow ||h_{T_0 \cup T_1}||_2 = 0 (\delta < \frac{1}{3})$

$\Rightarrow h = 0$

Hence the proof

4 How to desigh good RIP matrices ?

$\phi \in R^{mn}$ is (δ, k) RIP if $(1-\delta) \leq ||\phi x||_2^2$

$(1-\delta) \leq (1+\delta)$

5 Reconstruction of Johnson-Linderstrauss lemma:

like JL but for all K sparse vectors.

$$\phi = \begin{pmatrix} \cdots g_1 \cdots \\ \cdots g_2 \cdots \\ \cdots g_3 \cdots \end{pmatrix}$$

$\phi = [g_{ij}]$, where $[g_{ij}] = N(0,1)$ be a random varible.

- Fix $x \in R^n$, where x is a k sparse matrix, $||x||_2 = 1$
- Show that with high probability $||\phi x||_2^2 \approx 1 \pm \delta$
- Union bound overall x, $||x_0||_0 \leq K, ||x_1||_2 = 1$

$$y = \phi x = \begin{pmatrix} \cdots g_1 \cdots \\ \cdots g_2 \cdots \\ \vdots \\ \cdots g_m \cdots \end{pmatrix}$$

$y = \sum_{j=1}^{k} g_{ij} x_j$

Each g_{ij} is $N(0,1)$

if Z is $N(0,1)$ then αZ is $N(0, \alpha^2)$

$$y = \begin{pmatrix} \cdots N(0,1) \cdots \\ \cdots N(0,1) \cdots \\ \vdots \\ \cdots N(0,1) \cdots \end{pmatrix}$$

$E[||y||^2] = E[\sum y_{i^2}]$

$E[||y||^2] = \sum E[y_{i^2}]$

$E[||y||^2] = m$

$R = ||y^2||$ is a random variable.

$E[R] = m$

$var[R] = E[R^2] - E[R]^2 \leq \mathcal{O}(m)$

6 Bernstein's Inequality:

$\mathbf{Pr}[|R - M| > t\sqrt{m}] \leq e^{-t^2}$

Set $t = \epsilon\sqrt{m}$

$\mathbf{Pr}[|R - M| > \epsilon m] \leq e^{-\epsilon^2 m}$

for fixed u, $wp \geq 1 - \epsilon^2 m, ||\phi u||_2^2 \in (1 \pm \epsilon)^m$

$\forall R$ sparse vectors, the length to be preserved.

Morally, number of vectors to union bound over

$$\leq \binom{n}{k} o(1)^k = n^{o(k)}$$

if $e^{-\epsilon^2 m}.n^{o(k)} < \frac{1}{2}$ this is the overall probability.

We have to show that $m = \theta(k \log n)$ suffices.

$\rightarrow$ How to bound over infinitely many vectors?

Fix h coordinates as 1,2,3,.... k WLOG there are infinitely many unit vectors $||x|| = 1$

$X = x : \sum_{i=1}^{k} x_i^2 = 1$ are unit vectors.

Find a small set N such that $\forall x \in X, \exists n \in N$ such that $d(x, N) > \epsilon$

7 How small N can be?

Greedy ϵ - net, start with an arbitrary x_0, add to N, as long as $\exists x \in X$ such that $d(x, N) > \epsilon$, add x to N.

$\forall n_1, n_2 \in N$, we have $||n_1 - n_2||_2 > \epsilon$

$\Rightarrow B[n_i, \frac{\epsilon}{2}]n[n_j, \frac{\epsilon}{2}] = \phi, \forall i, j \in N$

Moreover, $N \subset X$, all points are in unit ball $B(0, 1)$ Number of net points,

$$|N| \leq \frac{volume(B(0,1))}{volume(B(\frac{\epsilon}{2}))}$$

is directly proportional to $(\frac{2}{\epsilon})^2$. Volume is directly proportional to r^k in k-d space. Now we show that all points in N are good for ϕ that means $\forall n \in N$.

$1 - \frac{\delta}{2} \leq ||\phi n||_2^2 \leq 1 + \frac{\delta}{2}$

Union bound as long as $e^{-\epsilon^2 m}$. $\binom{n}{k}$ $(\frac{2}{\epsilon})^k < \frac{1}{2}$

if ϕ is $1 \pm \frac{\delta}{2}$ is good for N, will show ϕ is good for X also

$ExtremalityArgument$ let $1 + A$ be the largest distortion for ϕ , meaning $\exists \tilde{x}$ such that $||\phi\tilde{x}||_2 \leq \frac{\delta}{4}$.

We want to bound A

$\exists \tilde{n} \in N$ such that $||(\tilde{x} - \tilde{n})||_2 \leq \frac{\delta}{4}$

$$||\phi\tilde{x}||_2 = 1 + A = ||\phi n^2 + \phi(\tilde{x} - \tilde{n})||_2 \leq ||\phi n_2||_2 + ||\phi(\tilde{x} - \tilde{n}||_2$$

$(1 + A) \leq 1 + \frac{\delta}{2} + \frac{\delta}{4} + \frac{\delta}{4}(1 + A)$

$A \leq \frac{\delta}{2} + \frac{\delta}{4} + \frac{A\delta}{4}$

$A(1 - \frac{\delta}{4} \leq \frac{3\delta}{4}$

$A \leq \delta$

$\therefore \phi$ is δ good for all points.