CS 15-859: Algorithms for Big Data Fall 2019

Lecture lecture 11 — Nov 21

Prof. David Woodruff *Scribe: Yifan Sun*

1 Compressed Sensing

In compressed sensing problem, we try to recover a vector $x \in \mathbb{R}^d$ through random linear measurements of x. We choose a random skeching matrix $S \in \mathbb{R}^{r \times n}$ and observe $S \cdot x$, the algorithm output a vector $x' \in \mathbb{R}^d$ such that

$$\|x - x'\|_p = D \cdot \min_{k-sparse\ z} \|x - z\|_q$$

where D is the distortion and the above problem is known as ℓ_p/ℓ_q-guarantee.

There are two main schemes for compressed sensing, the **Randomized** ('for-each') and **Deterministic** ('for-all').

Let x_k be the best $k - sparse$ approximation to x, i.e. the vector with largest k entries in magnitude. CountSketch with a randomized scheme achieves ℓ_2/ℓ_1 guarantee with high probability:

$$\|x - x'\|_2 = O(1) \cdot \|x - x_k\|_2 \tag{1}$$

1.1 CountSketch for Compressed Sensing

Multiplying by a CountSketch matrix with $O(k \log n)$ rows can be thought of as $O(\log n)$ repetitions of hashing into $O(k)$ buckets. S is a random linear map, it estimates every entry of x vector up to an additive error $\|x - x_k\|_2/\sqrt{k}$.

If the algorithm outputs $2k-$sparse vector $x' \in \mathbb{R}^d$ that keeps the top $2k$ entries in magnitude, then we can prove equation (1) holds with high probability. We list two definitions before going into the proof.

Definition: A coordinate i is **heavy** if:

$$|x_i| \geq \frac{\|x - x_k\|_2}{\sqrt{k}}$$

A direct fact is there are at most $2k$ heavy coordinates.

Definition: A coordinate i is **super-heavy** if:

$$|x_i| \geq 3 \cdot \frac{\|x - x_k\|_2}{\sqrt{k}}$$

We can observe that the set T of super-heavy coordinates is in the support of x'. Hence:

$$
\begin{aligned}
\|x - x'\| &\le \|(x-x')_T\|_2 + \|(x-x')_{[n]\setminus T}\|_2 \\
&\le \sqrt{2}k \cdot \frac{\|x - x_k\|_2}{\sqrt{k}} + \|(x-x')_{[n]\setminus T}\|_2 \\
&\le \sqrt{2}\|x - x_k\|_2 + \|(x-x_k)_{[n]\setminus T}\|_2 + \|(x_k - x')_{[n]\setminus T}\|_2 \\
&= O(\|x - x_k\|_2) \qquad (2)
\end{aligned}
$$

1.2 No Deterministic Scheme for ℓ_2/ℓ_2 Distortion

Consider the case for $k = 1$ and suppose that S is a deterministic sketching matrix with $r = o(n)$ rows. Then it sufficies to show that there exists a vector $x \in Ker(S)$ and for any constant C we have:

$$
\|x\|_\infty \ge C\|x - x_1\|_2 \qquad (3)
$$

Without loss of generality, assume that S has orthonormal rows, therefore $\sum_i \|Se_i\|_2^2 = r$ and there exists a coordinate i with $\|Se_i\|_2^2 \le \frac{r}{n}$. Let x be the vector:

$$
x = e_i - S^T S e_i. \qquad (4)
$$

It is clear that $x \in Ker(S)$, and:

$$
\begin{aligned}
\|x\|_\infty^2 &\ge |x_i|^2 \\
&= (e_i^T e_i - e_i^T S^T S e_i)^2 \\
&\ge (1 - \frac{r}{n})^2. \qquad (5)
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\|x - x_1\|_2 &\le \|x - e_i\|_2 \\
&= \|S^T S e_i\|_2 \\
&= \|Se_i\|_2 \\
&\le \sqrt{\frac{r}{n}} \\
&= o(1) \qquad (6)
\end{aligned}
$$

1.3 Deterministic Scheme for ℓ_2/ℓ_1 Distortion

Definition: Matrix S has $(\epsilon, k)-$restricted isometry property (RIP) if for all $k-$sparse vector $x \in \mathbb{R}^d$, the following equation holds:

$$
(1-\epsilon)\|x\|_2^2 \le \|Sx\|_2^2 \le (1+\epsilon)\|x_2\|_2^2
$$

A major open question is if there exists an explicit matrix with $(\epsilon, k)-$RIP with $o(k^2)$ rows. In [1], authors showed it is possible to get a matrix with $k^{2-\gamma}$ rows for a constant $\gamma > 0$ and $k \approx n^{1/2}$.

It can be shown that if S has the $(\epsilon, k)-$RIP property, then one can get an $x' \in \mathbb{R}^d$ for which:

$$\|x - x'\|_2 = O(\frac{1}{\sqrt{k}})\|x - x_k\|_1 \tag{7}$$

from solving the following Linear Program:

$$\min_{z \in \mathbb{R}^d} |z|_1 \tag{8}$$

$$s.t. \ Sz = Sx \tag{9}$$

If x' is the solution, then

$$\|x - x'\|_2 \leq O(1/\sqrt{k})\|x - x_k\|_1$$

.

References

[1] Jean Bourgain, S.J. Dilworth, Kevin Ford, Sergei Konyagin and Denka Kutzarova. *Explicit constructions of RIP matrices and related problems* , Duke Matehmatical Journal, Vol. 159 (2011), No.1, pages 145-185.

[2] Kenneth L. Clarkson and David P. Woodruff. *Input Sparsity and Hardness for Robust Subspace Approximation.* IEEE Symposium on Foundations of Computer Science (FOCS), 2015.