COMS E6998-9: Algorithms for Massive Data (Fall '23) September 28, 2023

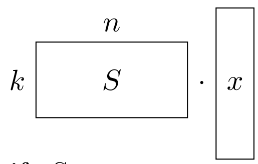
Lecture 7: Fast Dimensional Reduction

Instructor: *Alex Andoni* Scribes: *Justin Hudis*

Theorem: $\exists$ a distribution of $\varphi : \mathbb{R}^n \to \mathbb{R}^k, \varphi(x) = Sx$ s.t.

1. φ satisfies DJL: $\forall x \ \Pr_\varphi \left[\frac{||\varphi(x)||_2^2}{||x||_2^2} = 1 \pm \epsilon\right] \geq 1 - \delta$
2. φ can be computed in $O(n \lg n + k)$ time
3. $k = O\left(\frac{\lg \frac{1}{\delta}}{\epsilon^2} \cdot \lg \frac{n}{\delta}\right)$

Proof

DJL: $\varphi(x) =$ 

Idea 1: Sparsify S

$$S = \begin{bmatrix} S_1 \\ \vdots \\ S_k \end{bmatrix}$$

Original analysis: $S_1 x \sim ||x||_2 g$
Suppose S_1 has s non-zeros
S_{1i} is $N(0,1)$ with probability $\frac{s}{n}$, 0 otherwise

$$\begin{aligned} \frac{n}{s} \cdot \mathbb{E}\left[(S_1 \cdot x)^2\right] &= \frac{n}{s} \cdot \mathbb{E}\left[\sum_{i,j=1}^{n} S_{1i} S_{1j} x_i x_j\right] \\ &= \frac{n}{s} \cdot \sum_{i,j=1}^{n} \mathbb{E}\left[S_{1i} S_{1j}\right] x_i x_j \\ &= \frac{n}{s} \cdot \sum_{i=1}^{n} x_i^2 \cdot \frac{s}{n} \cdot \mathbb{E}_g[g^2] \\ &= ||x||_2^2 \end{aligned}$$

$$
\begin{aligned}
\mathrm{Var}\left[\frac{n}{s}(S_1x)^2\right] &= \frac{n^2}{s^2}\mathbb{E}_{S_1}\left[(S_1x)^4\right] - ||x||_2^4 \\
&\geq \frac{n^2}{s^2}\sum_{i=1}^{n}\mathbb{E}\left[S_{1i}^4\right]\cdot x_i^4 \\
&= \frac{n^2}{s^2}\sum_{i=1}^{n} x_i^4 \cdot \frac{s}{n} \\
&= ||x||_4^4 \cdot \frac{n}{s}
\end{aligned}
$$

For this to work, the following must be true:

$$
\begin{aligned}
\mathbb{E} &\geq \alpha\sqrt{\mathrm{Var}} \\
||x||_2^2 &\geq \alpha \cdot \sqrt{\frac{n}{s}} \cdot ||x||_4^2 \\
s &\geq \alpha^2 \cdot n \cdot \left(\frac{||x||_4}{||x||_2}\right)^2
\end{aligned}
$$

Case 1: $x = (1, 1, 1, \ldots, 1)$

$$||x||_2 = \sqrt{n}$$

$$||x||_4 = n^{\frac{1}{4}}$$

$$s \approx \alpha^2 n \cdot \left(\frac{n^{\frac{1}{4}}}{\sqrt{n}}\right) \approx \alpha^n$$

Case 2: $x = (1, 0, \ldots, 1)$

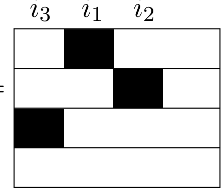
$$||x||_2 = 1$$

$$||x||_4 = 1$$

$$s \approx \alpha^2 n \cdot 1 = \alpha^2 n$$

Conclusion: This only works for x's s.t. $||x||_4 << ||x||_2$
Lemma: Define P of size $k \times n$ s.t.

$$Px = (x_{i_1}, x_{i_2}, \ldots, x_{i_k}), \quad i_1 \ldots i_k \in_r [n]$$

$P =$ 

$$\forall x \in \mathbb{R}^n : \Pr_p\left[\frac{n}{k} \cdot \frac{||Px||^2}{||x||^2} = 1 \pm \epsilon\right] \geq 1 - \delta$$

$$k = O\left(\frac{\lg\frac{1}{\delta}}{\epsilon^2}\cdot n\cdot\frac{||x||_\infty^2}{||x||_2^2}\right)$$

Proof: Use Chernoff bound:

$$||Px||^2 = \sum_{j=1}^{k} x_{ij}^2 \triangleq \sum_{j=1}^{k} z_j^2$$

$$O \le z_j \le ||x||_\infty^2$$

$$0 \le y_j = \frac{z_j}{||x||_\infty^2} \le 1$$

$$\mu = \mathbb{E}\left[\sum_{i=1}^{k} y_j\right] = \frac{1}{||x||_\infty^2}\sum_{j=1}^{k}\mathbb{E}_{i_j}\left[x_{i_j}^2\right] = \frac{k}{||x||_\infty^2}\cdot\frac{\sum_i x_i^2}{n} = \frac{k}{n}\cdot\frac{||x||_2^2}{||x||_\infty^2}$$

By Chernoff bound,

$$Pr_p\underbrace{\left[\left|\sum_{j=1}^{k} y_j - \mu\right| \le \epsilon\mu\right]}_{\sum y_j = (1\pm\epsilon)\mu} \ge 1 - 2e^{-\frac{\mu\epsilon^2}{3}}$$

$$\frac{n}{k}\cdot||x||_\infty^2\cdot\sum y_j = (1\pm\epsilon)\cdot||x||_2^2$$
$$\frac{n}{k}\cdot||Px||_2^2 =$$

$$2e^{-\frac{\mu\epsilon^2}{3}} = 2\exp\left[-\frac{\frac{||x||_2^2}{||x||_\infty^2}\cdot\frac{1}{n}\cdot\lg\frac{1}{\delta}\cdot\frac{1}{\epsilon^2}\cdot n\cdot\frac{||x||_\infty^2}{||x||_2^2}\cdot\epsilon^2}{3}\right]$$
$$\le \delta$$

Idea 1: Densify S
Use Hadamed transform

$$n = 2^\ell$$

$$H_0 = 1$$

$$H_1 = \begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\cdot\frac{1}{\sqrt{2}}$$

$$H_2 = \left[\begin{array}{c|c} H_{i-1} & H_{i-1} \\ \hline H_{i-1} & -H_{i-1}\end{array}\right]\cdot\frac{1}{\sqrt{2}}$$

$$H_\ell = \begin{pmatrix}\pm1 & \pm1 & \\ & & \dots\end{pmatrix}\cdot\frac{1}{\sqrt{2}}$$

Properties of H:

1. H is a rotation matrix in $\mathbb{R}^n$

$$||Hx||_2 = ||x||_2$$

2. Can compute Hx in $O(n \lg n)$ time (FFT)

3. Uncertainty principle: x is sparse $\Rightarrow$ Hx is dense

Goal: Use $P \cdot \underbrace{Hx}_{\textbf{dense}}$

Issue: If $x = H^{-1} \cdot$ (sparse)

Idea 3: Premultiply x by $D = \begin{pmatrix} \pm 1 & & & 0 \\ & \pm 1 & & \\ & & \ddots & \\ 0 & & & \pm 1 \end{pmatrix}$

Overall: $\varphi(x) \triangleq P \cdot H \cdot D \cdot x$

Lemma: Fix $||x||_2 = 1$. Then for $y \triangleq H \cdot D \cdot x, \Pr\left[||y||_\infty \leq O\left(\sqrt{\frac{\lg \frac{n}{\delta}}{n}}\right)\right] \geq 1 - \delta$

Observation: $||y||_2 = ||x||_2$

Proof: Analyze each y_i

$$\begin{aligned} y_i &= \sum_{j=1}^{n} (H_\ell)_{ij} \cdot D_{jj} \cdot x_j \\ &\sim \frac{1}{\sqrt{n}} \sum_{i=1}^{n} r_j x_j, r_j \in \{\pm 1\} \\ &\leq ||x||_2 \end{aligned}$$

If $r_j \sim N(0,1)$, then $\sum r_j x_j \sim ||x||_2 \cdot g$

Theorem: $\forall a_1 \dots a_n \in \mathbb{R},\ r_1 \dots r_n \in \pm 1,\ \Pr\left[|\sum a_i r_i| \geq t \cdot ||a||_2\right] \leq e^{-\frac{t^2}{2}}$

$$\Pr\left[|y_i| \geq \frac{1}{\sqrt{n}} \cdot t \cdot ||a||_2\right] \leq e^{-\frac{t^2}{2}}$$

Set $t = \theta\left(\sqrt{\lg \frac{n}{\delta}}\right)$

Then $\Pr\left[|y_i| \geq O\left(\sqrt{\lg \frac{n}{\delta}}\right) ||x||\right] \leq \frac{\delta}{n}$

By union bound, $\Pr\left[||y||_\infty \geq O\left(\sqrt{\lg \frac{n}{\delta}}\right) ||x||\right] \leq \delta$

This finishes the proof of FDR/FJL.

Using the lemma on P,

$$k = O\left(\frac{\lg \frac{1}{\delta}}{\epsilon^2} \cdot \cancel{n} \cdot \frac{\frac{\lg \frac{n}{\delta}}{\cancel{n}} \cdot \cancel{||x||^2}}{\cancel{||x||^2}} \right)$$
$$= O\left(\frac{\lg \frac{1}{\delta} \cdot \lg \frac{n}{\delta}}{\epsilon^2} \right)$$