Lecture 8: Probabilistic Dimension Reduction

Xiaotie Deng

AIMS Lab
Department of Computer Science
Shanghai Jiaotong University

October 25, 2017

Johnson and Lindenstrauss Theorem

Johnson and Lindenstrauss Theorem

- $\forall \epsilon \in (0,1)$ and $\forall n$ *integer*, set $k \geq 4(\epsilon^2/2 - \epsilon^3/3)^{-1} \ln n$.
- $\forall V \subseteq \mathcal{R}^d$, $|V| = n$, there is a map $f : \mathcal{R}^d \to \mathcal{R}^k$ such that $\forall v \in V$:
  - $(1-\epsilon)||u-v||^2 \leq ||f(u)-f(v)||^2 \leq (1+\epsilon)||u-v||^2$.
- Further, f can be found in randomized polynomial time.

Reference: Dasgupta, Sanjoy; Gupta, Anupam (2003), "An elementary proof of a theorem of Johnson and Lindenstrauss", Random Structures & Algorithms, 22 (1): 60-65.

Main Indeas

- square length of a random vector is (sharply) concentrated around its mean if projected into a k-dimensional random subspace.
  - With probability $O(1/n^2)$, its (scaled) length is not distorted by more than $(1 \pm \epsilon)$.
  - It holds for all n vectors with probability $\frac{1}{n}$.
- The project has the same probability distribution as projection to a fixed k-dimension space.

A Gaussian Model

- Let $X_1, X_2, \cdots, X_d$ be iid Gaussian $N(0,1)$ random variables.
- Let $Y = \frac{1}{||X||}(X_1, X_2, \cdots, X_d)$, where $X = (X_1, X_2, \cdots, X_d)$.
- Y equals to uniformly random variable on unit sphere S^{d-1}.
- Let Z be a projection of Y into its first k-coordinates, $L = Z^2$.
- $E[Z^2] = E[L] = k/d$ by symmetry.

Concentration Lemma

- For $k < d$ we have:
  (a) If $\beta < 1$, then
  $$Pr[L \leq \frac{\beta k}{d}] \leq \beta^{k/2}(1 + \frac{(1-\beta)k}{(d-k)})^{(d-k)/2} \leq exp(\frac{k}{2}(1 - \beta + \ln \beta)).$$
  (b) If $\beta > 1$, then
  $$Pr[L \leq \frac{\beta k}{d}] \leq \beta^{k/2}(1 + \frac{(1-\beta)k}{(d-k)})^{(d-k)/2} \leq exp(\frac{k}{2}(1 - \beta + \ln \beta)).$$

Proof: Concentration Lemma (a)

- $X \sim N(0,1)$ & $s < \frac{1}{2}$,
  $\Longrightarrow E[e^{sX^2}] = \frac{1}{\sqrt{2\pi}} \int e^{(sx^2 - \frac{x^2}{2})} dx = 1/\sqrt{1-2s}.$
- $Pr[d(X_1^2 + X_2^2 + \cdots + X_k^2) \leq k\beta(X_1^2 + X_2^2 + \cdots + X_d^2)] \leq \beta^{k/2}(1 + \frac{k(1-\beta)}{d-k})^{(d-k)/2}$
- Proof:

$$Pr[d(X_1^2 + \cdots + X_k^2) \leq k\beta = k\beta(X_1^2 + \cdots + X_d^2)] =$$
$$= Pr[d(X_1^2 + \cdots + X_k^2) - k\beta(X_1^2 + \cdots + X_k^2) \leq 0] \leq$$
$$\leq E[exp\{t(k\beta(X_1^2 + \cdots + X_d^2) - d(X_1^2 + \cdots X_k^2))\}] \leq$$
$$\leq E[exp\{t(k\beta)X_1^2\}]^{(d-k)} * E[exp\{t(k\beta - d)X_1^2\}]^k =$$
$$= (1 - 2tk\beta)^{-(d-k)/2}(1 - 2t(k\beta - d))^{-k/2}$$

$RHS \leq (\frac{d-k\beta}{d-k})^{(d-k)/2}\beta^{k/2}$

Proof: Concentration Lemma (b)

- Proof is Similar.

Proof of Main Theorem

- $\forall V \subseteq \mathcal{R}^d$, $|V| = n$, $\exists$ map $f : \mathcal{R}^d \to \mathcal{R}^k$ such that $\forall v \in V$:
  - $(1-\epsilon)||u-v||^2 \leq ||f(u)-f(v)||^2 \leq (1+\epsilon)||u-v||^2$.
- Trivial if $d \leq k$.
- $k < d$: Consider subspace S, and map $v_i \in V$ to $v_i' \in S$. Let $L = ||v_i' - v_j'||^2$ and $\mu = (k/d)||v_i - v_j||^2$.
- By Main Lemma (a) (and $\ln(1-\epsilon) \leq -x - \frac{x^2}{2}$)

$$Pr[L \leq (1-\epsilon)\mu] \leq$$

$$\leq exp(\frac{k}{2}(1-(1-\epsilon)+\ln(1-\epsilon))) \leq$$

$$\leq exp(\frac{k}{2}(\epsilon-(\epsilon+\frac{\epsilon^2}{2}))) = exp(-\frac{-k\epsilon^2}{4}) \leq$$

$$\leq exp\{-2\ln n\} = \frac{1}{n^2}$$

Proof of Main Theorem

- By Main Lemma (b) (and $\ln(1+\epsilon) \leq x - \frac{x^2}{2} + \frac{x^3}{3}$)

$$Pr[L \leq (1+\epsilon)\mu] \leq$$

$$\leq exp(\frac{k}{2}(1-(1+\epsilon)+\ln(1-\epsilon))) \leq$$

$$\leq exp(\frac{k}{2}(-\epsilon+(\epsilon-\frac{\epsilon^2}{2}+\frac{\epsilon^3}{3})) = exp(-\frac{k(\frac{\epsilon^2}{2}-\frac{\epsilon^3}{3})}{2}) \leq$$

$$\leq exp\{-2\ln n\} = \frac{1}{n^2}$$

Proof of Main Theorem

- Setting $f(v) = \sqrt{d/k}v'$.
- Repeat the random choice n times to boost up the probability.