Note Of Lecture 8 Probabilistic Dimension Reduction

Fang Pin

October 26, 2017

1 Why Dimension Reduction In Distance Preserving

- n-dimension data ==> O(n)
- log n -dimension data ==> O(log n)

2 Johnson and Lindenstrauss Theorem

`https://en.wikipedia.org/wiki/Johnson%E2%80%93Lindenstrauss_lemma`, according to the theory, we propose the probabilistic dimension reduction idea.

- project a d-dimensional random variable to a k-dimensional random subspace with a low error rate for its square length.
- the project has the same probability distribution as projection to a fixed k-dimension space.

3 A Gaussian Example.

- Let X_1, X_2, ... , X_d be i.i.d N(0,1) random variables.
- Let Y = (X_1, X_2, ... , X_d) / ||X||.
- Let Z be a projection of Y into its first k-coordinates.

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

3.1 prove

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

$$Pr[d(X_1^2 + \cdots + X_k^2) \leq k\beta = k\beta(X_1^2 + \cdots + X_d^2)] =$$
$$= Pr[d(X_1^2 + \cdots + X_k^2) - k\beta(X_1^2 + \cdots + X_k^2) \leq 0] \leq$$

markov inequality; d

$$\leq E[exp\{t(k\beta(X_1^2 + \cdots + X_d^2) - d(X_1^2 + \cdots X_k^2))\}] \leq$$

t is any fixed constant, choose t to minimize estimation

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

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

3.2 some details

3.2.1 markov inequality step

https://en.wikipedia.org/wiki/Markov%27s_inequality

$$P(X \geq a) \leq \frac{E(X)}{a}$$

so that $Pr[d(x_1^2 + ... + x_k^2) - k\beta(x_1^2 + ... + x_d^2) \leq 0] = Pr[exp\{k\beta(x_1^2 + ... + x_d^2) - d(x_1^2 + ... + x_k^2)\} \geq 1] \leq E[exp\{k\beta(x_1^2 + ... + x_d^2) - d(x_1^2 + ... + x_k^2)\}]$

3.2.2 the last inequality.

- Let $f(x) = xln(1 + \frac{1}{x}) - 1(x > 0)$
- so that $f'(x) = \ln(1 + \frac{1}{x}) - \frac{1}{x+1}(x > 0)$

- and $f''(x) = -\frac{1}{x(x+1)^2}(x > 0)$

so we can get $f''(x) < 0$
so that $f'(x) < lim_{x->0} f'(0) = 0$
so that $f(x) < lim_{x->0} f(0) = 0$

- so we know $1 > x \ln(1 + \frac{1}{x})(x > 0)$. i.e $(1 + \frac{1}{x})^x <= e$

- $\beta^{\frac{k}{2}}[1+\frac{(1-\beta)k}{(d-k)}]^{\frac{(d-k)}{2}} = exp\{\ln(LHS)\} = exp\{\frac{k}{2}\ln\beta + \ln(1+\frac{1}{\frac{d-k}{(1-\beta)k}})^{\frac{d-k}{(1-\beta)k}\frac{(1-\beta)k}{2}}\} \leq exp\{\frac{k}{2}(1-\beta+\ln\beta)\}$