SOME USEFUL APPLICATIONS

LARGE-SCALE MATRIX COMPUTATIONS

- from Hutchinson we know (Lemma 12.2)

$$tr(A) = \mathbb{E}[\vec{m}^T A \vec{m}] \quad : A: SPD$$

$m \in \mathbb{R}^n$ $\quad A \in \mathbb{R}^{n \times n}$

$\mathbb{E}[m] = 0$

$\mathbb{E}[m m^T] = I$

Monte Carlo

$\vec{m}_i$ are i.i.d.

$$\frac{1}{N} \sum_{i=1}^{N} \underbrace{\vec{m}_i^T A \vec{m}_i}_{z_i} =: S_N$$

z_i are i.i.d.

- It can be shown (by Gauss-quadrature) Golub:

$$L_i \leq \vec{m}_i^T A m_i \leq U_i$$

where L_i and U_i are computable from Gauss quadrature.

- Now from Hutchison:

$$\mathbb{E}[z_i] = E[\vec{m}_i^T A m_i] = tr(A)$$

- From Hoeffding inequality

- from Hoeffding inequality

$$\mathbb{P}\left[\,|S_N - tr(A)| > t\,\right] \leq 2e^{-2N^2 \frac{t^2}{\Sigma (U_i - L_i)^2}}$$

$$\Downarrow$$

$$-t \leq S_N - tr(A) \leq t$$

with probability:

$$1 - 2e^{-2N^2 \frac{t^2}{\Sigma (U_i - L_i)^2}}$$

- If we pick an error t and a successful probability β, we can find N by

$$1 - 2e^{-2N^2 t^2 / \Sigma (U_i - L_i)^2} = \beta$$

$$\Updownarrow$$

$$N = \sqrt{\frac{\Sigma (U_i - L_i)^2}{2t^2} \ln\left(\frac{2}{1-\beta}\right)}$$

DIMENSION REDUCTION

- Let $\vec{x} \in \mathbb{R}^N$, $N \gg 1$, consider a random matrix $A \in \mathbb{R}^{n \times N}$ whose entries are i.i.d. RVs with zero mean and unity variance. Define a random "projection" $\mathcal{P}$:

$$\vec{z} := \mathcal{P}\vec{x} := \frac{1}{\sqrt{n}} A x$$

- we can show that $z_i = A(i,:)\vec{x}$

$$\mathbb{E}[z_i] = 0$$

$$\mathbb{V}[z_i] = \mathbb{E}[z_i^2] = \frac{\|\vec{x}\|^2}{n}$$

$$\mathbb{E}[\|\vec{z}\|^2] = \|x\|^2$$

Remark: this is true on average, but what we are interested in is the behavior of $\vec{z}$ for a particular realization of A.

By chernoff

$$\mathbb{P}[\|z\|^2 - \|x^2\| \geq \varepsilon\|x\|^2]$$

$$\|$$

$$\mathbb{P}[n\|z\|^2 \geq n(1+\varepsilon)\|x\|^2]$$

// chernoff

$$\min \; e^{-n\lambda(1+\varepsilon)\|x\|^2} \prod^{n} \mathbb{E}[e^{\lambda n z_i^2}]$$

$$\min_{\lambda>0} \; e^{-n\lambda(1+\varepsilon)\|x\|^2} \prod_{i=1}^{n} \mathbb{E}[e^{\lambda n z_i^2}]$$

Thus what we need is to bound the MGF

$$\mathbb{E}\left[e^{\lambda n z_i^2}\right]$$

A_{ij} are Gaussians $N(0,1)$

Then from Ex 15.3 we have

$$\mathbb{E}[e^{\lambda n z_i^2}] = \frac{1}{\sqrt{1-2\lambda\|x^2\|}} \qquad \lambda \le \frac{1}{2\|x\|^2}$$

and

$$\mathbb{P}[\|z\|^2 - \|x\|^2 \geq \varepsilon\|x\|^2] \le \min_{\lambda>0} e^{\frac{n}{2} f(\lambda)}$$

where

$$f(\lambda) := -2\lambda(1+\varepsilon)\|x\|^2 - \ln(1-2\lambda\|x\|^2)$$

- It can be shown (elementary) that the minimizer is

$$\lambda^* = \frac{\varepsilon}{2(1+\varepsilon)\|x\|^2} \qquad 0 \le \lambda \le \frac{1}{2\|x\|^2}$$

and thus

$$\mathbb{P}[\|z\|^2 - \|x\|^2 \ldots] \le e^{\frac{n}{2}(\ln(1+\varepsilon)-\varepsilon)}$$

$$\mathbb{P}\left[\|z\|^2 - \|x\|^2 \geq \varepsilon \|x\|^2\right] \leq e^{\frac{n}{2}\left(\ln(1+\varepsilon) - \varepsilon\right)}$$

$\varepsilon \in [0,1]$

$$\ln(1+\varepsilon) - \varepsilon \leq -\frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3}$$

$$e^{\frac{n}{2}\left(-\frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3}\right)}$$

<u>Conclusion</u>:

$$\boxed{\mathbb{P}\left[\|z\|^2 - \|x\|^2 \geq \varepsilon \|x\|^2\right] \leq e^{n/2\left(-\frac{\varepsilon}{2} + \frac{\varepsilon^3}{3}\right)}}$$

$$\Downarrow$$

$$-\varepsilon \|x\|^2 \leq \|z\|^2 - \|x\|^2 \leq \varepsilon \|x\|^2$$

with probability

$$1 - 2e^{\frac{n}{2}\left(-\frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3}\right)}$$

- suppose that we have m vectors $\vec{x}_i \in \mathbb{R}^N$
$N \gg 1$. then

$$\vec{y}_i := P \vec{x}_i$$

and we have show $\mathbb{E}\left[\|\vec{y}_i\|^2\right] = \|\vec{x}_i\|^2$

$$\vec{y}_i - \vec{y}_j = P(\vec{x}_i - \vec{x}_j)$$

and we have shown that f preserves the distance with high probability.

$$\mathbb{P}[\text{ Some pair has } \varepsilon\text{-distortion }]$$

$$\|$$

$$\mathbb{P}\left[\bigcup_{i=1}^{m(m-1)/2} \text{pair } i \text{ has } \varepsilon\text{-distortion}\right]$$

$$\wedge\!\!/\!\!/$$

$$\sum_{i=1}^{m(m-1)/2} \mathbb{P}[\text{ Pair } i \text{ has } \varepsilon\text{-distortion }]$$

$$\frac{m(m-1)}{2} e^{\frac{n}{2}\left(-\frac{\varepsilon^2}{2} + \frac{\varepsilon^3}{3}\right)}$$

Thus if we want

$$\mathbb{P}[\text{ at least one pair has } \varepsilon\text{-distortion }] \leq \frac{1}{m^\beta}$$

then this can be attained if

$$m\frac{(m-1)}{2} e^{\frac{n}{2}\left(-\frac{\varepsilon^2}{2} + \varepsilon^3/3\right)} \leq \frac{1}{m^\beta}$$

$$\Updownarrow$$

$$\geq \quad 2\beta \ln(m) + 2\ln(m(m-1))$$

$$(*)\quad \boxed{n \geq \frac{2\beta \ln(m) + 2\ln(m(m-1))}{\frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3}}}$$

Lem 15.1 (Johnson-Lindenstrauss lemma)

Consider m vectors $\vec{x}_i \in \mathbb{R}^N$, $i=1,\dots,m$, $(N \gg 1)$. Define a random matrix $P := \frac{1}{\sqrt{n}} A$ where $A \in \mathbb{R}^{n\times N}$, A_{ij} are i.i.d. with zero mean and unity variance. $\forall$ any $\beta > 0$ if we choose n as in $(*)$, then with probability at leas $1 - m^{-\beta}$ we have

$$(1-\varepsilon)\|\vec{x}_j - \vec{x}_i\|^2 \leq \|P(\vec{x}_i - \vec{x}_j)\|^2 \leq (1+\varepsilon)\|\vec{x}_i - \vec{x}_j\|^2$$

(Restricted Isometry Property RIP).

SUB-GAUSSIAN RVs:

A_{ij} are i.i.d. α-sub-gaussian RV with $\mathbb{E}[A_{ij}] = 0$; $\mathbb{V}[A_{ij}] = 1$

observations:

1) $z_i := A(i,:)\vec{x}$ is a sub-gaussian RV

with proxy $\alpha^2 \frac{\|x\|^2}{n}$

2) $n z_i^2 - \|x\|^2$ is a zero mean RVs.

3)

$$\mathbb{P}\left[\|z\|^2 \geq (1+\varepsilon)\|x\|^2 \right]$$

$$\wedge\!\!/\!/$$

$$\min_{\lambda > 0} e^{-n\lambda(1+\varepsilon)\|x\|^2} \prod_{i=1}^{n} \mathbb{E}\left[e^{\lambda n z_i^2} \right]$$

How to bound this?

Lem 15.2: Let m be a zero-mean RV with the tail bound

$$(*) \quad \boxed{\mathbb{P}\left[|m| \geq t \right] \leq 2 e^{-\frac{2t}{\beta}} \quad \forall t \geq 0}$$

for some $\beta > 0$. Then the MGF of m satisfies

$$\mathbb{E}\left[e^{sm} \right] \leq e^{2 s^2 \beta^2}, \quad \forall\ |s| \leq \frac{1}{2\beta}$$

Proof: the proof is similar to the equivalence of sub-gaussian properties. In particular $(*)$ implies

$$\mathbb{E}[|m|^p] \overset{!}{=} 2(\beta/2)^p \, p\, \Gamma(p) \overset{!}{\leq} \beta^p p! \quad (**)$$

which, together with $\mathbb{E}[m] = 0$, gives

$$\mathbb{E}[e^{sm}] = 1 + \sum_{p=2}^{\infty} \frac{s^p \mathbb{E}[m^p]}{p!}$$

$$\bigwedge (**) \qquad |s\beta| < 1$$

$$1 + \sum_{p=2}^{\infty} (s\beta)^p = 1 + \frac{s^2\beta^2}{1 - s\beta}$$

$$1 + 2x^2 \leq e^{x^2} \qquad \bigwedge \quad |s\beta| \leq 1/2$$
$$|x| \leq 1/2$$

$$1 + 2s^2\beta^2 \leq e^{s^2\beta^2}$$

$\Rightarrow$ conclude:

$$\mathbb{E}[e^{sm}] \leq e^{s^2\beta^2} \quad \forall \; |s| \leq \frac{1}{2\beta}$$

Back to bounding $\mathbb{E}[e^{\lambda n z_i^2}]$

- First we have

$$\mathbb{P}\left[n z_i^2 - \|x\|^2 \geq \varepsilon \|x\|^2 \right]$$

$$\Updownarrow$$

$$\mathbb{P}\left[|z_i| \geq \sqrt{1+\varepsilon}\, \|x\| \right]$$

$$\mathbb{P}\left[|z_i| \geq \sqrt{\frac{1+\varepsilon}{n}} \|x\| \right]$$

$$\leq \qquad z_i \text{ is } \alpha^2 \frac{\|x\|^2}{n}\text{-sub-gaussian}$$

$$2 e^{-\frac{(1+\varepsilon)\|x\|^2}{2\alpha^2 \|x\|^2}}$$

Thus

$$\mathbb{P}\left[n z_i^2 - \|x\|^2 \geq \varepsilon \|x\|^2 \right] \leq 2 e^{-\frac{(1+\varepsilon)\|x\|^2}{2\alpha^2\|x\|^2}}$$

If we define $t := \varepsilon \|x\|^2$

$p := 4\alpha^2 \|x\|^2$

$m = n z_i^2 - \|x\|^2$

Now we can apply Lem 15.2 :

$$(**) \quad \boxed{\mathbb{E}\left[e^{\lambda(n z_i^2 - \|x\|^2)} \right] \leq e^{32\alpha^4 \|x\|^4 \lambda^2} \Leftarrow \lambda \leq \frac{1}{8\alpha^2 \|x\|^2}}$$

Thus:

$$\mathbb{P}\left[\|z\|^2 - \|x\|^2 \geq \varepsilon \|x\|^2 \right]$$

$$\leq$$

$$e^{-n\varepsilon\lambda\|x\|^2} \prod^{n} \mathbb{E}\left[e^{\lambda(n z_i^2 - \|x\|^2)} \right]$$

$$e^{-n\varepsilon\lambda\|x\|^2}\prod_{i=1}^{n}\mathbb{E}\left[e^{\lambda(n z_i^2-\|x\|^2)}\right]$$

$$\overset{(***)}{\leq}$$

$$\min_{\lambda} e^{32\alpha^4\|x\|^4\lambda^2-\varepsilon\lambda\|x\|^2}$$

$$\|$$

$$e^{-n\frac{\varepsilon^2}{128\alpha^4}}$$

By the union bound we have

$$(1-\varepsilon)\|x\|^2\leq\|z\|^2\leq(1+\varepsilon)\|x\|^2$$

with probability $1-2e^{-n\,\varepsilon^2/128\alpha^4}$

which is sub-gaussian version of JL lemma.

Observations:

1) If m is a sub-gaussian then m^2 is a sub-exponential

$$\mathbb{P}[|m|\geq t]\leq 2e^{-\frac{2t}{\beta}}$$

2) Thm 15.1. (Bernstein's inequality).

- Let m_i, $i=1,\dots,N$ be zero mean

sub-exponential RVs., i.e. m_i satisfies

$$\mathbb{P}[\,|m_i| \geq t\,] \leq 2e^{-2t/\beta_i}$$

Thm. for any $t \geq 0$, we have

$$\mathbb{P}\left[\left|\sum_{i=1}^N m_i\right| \geq t\right] \leq 2\min\left\{e^{-\frac{t^2}{8\sum \beta_i^2}}, ce^{-\frac{t}{2\max\beta_i}}\right\}$$

Proof: we have

$$\mathbb{P}\left[\sum_{i=1}^N m_i \geq t\right]$$

$$\leq$$

$$\min_{\lambda>0} \frac{\prod_{i=1}^N \mathbb{E}[e^{\lambda m_i}]}{e^{\lambda t}}$$

$$\leq \quad \text{Lem 15.2}$$

$$\min_{\lambda \leq \frac{1}{2\max\beta_i}} e^{2\lambda^2\sum_{i=1}^N \beta_i^2 - \lambda t}$$

$$\Big\| \quad \text{minimizer } \lambda^* = \min\left\{\frac{t}{4\sum \beta_i^2}, \frac{1}{2\max\beta_i}\right\}$$

$$\min\left\{e^{-\frac{t^2}{8\sum\beta_i^2}}, ce^{-\frac{t}{2\max\beta_i}}\right\}$$

$$\min \left(e^{-8\sum \beta_i^2} ,\ c\, e^{2\max \beta_i} \right)$$