Tuesday, April 14

If second midterm grade is higher, I will replace first midterm grade.

- Midterm next Thursday (4/23)
- No class next Tuesday (4/21)
- Office hours:
  - ↳ Canceled next Thursday (4/23)
  - ↳ Virtual Monday (4/20)
  - ↳ Extra tomorrow (4/15)

This week: Continue doing cool things with matrices, fast

Randomized Numerical Linear Algebra
- related to my research
- beautiful, heavy proof

Spectral Graph Theory Correction

$L = D - A$

Incorrect : If 0 is smallest eigenvalue of L,
then 1 is top eigenvalue of A

Let $\bar{L} = D^{-1/2} L D^{-1/2}$, $\bar{A} = D^{-1/2} A D^{-1/2}$, $\bar{D} = D^{-1/2} D D^{-1/2} = I$

$\bar{L} = I - \bar{A}$ Consider $\bar{A}v = \lambda v$.

$\bar{L}v = (I - \bar{A})v = v - \lambda v = (1-\lambda)v$

Since $\bar{L}$ has eigenvalues ≥ 0, $(1-\lambda) \geq 0 \iff \lambda \leq 1$ for all λ

Since 0 is smallest eigenvalue of $\bar{L}$ with eigenvector $D^{1/2}\mathbb{1}$,
$\therefore$ 1 is top eigenvalue of $\bar{A}$ with same eigenvector

Sketched Regression

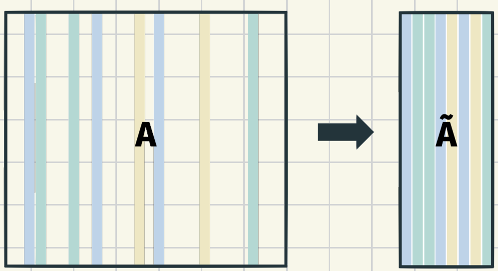
Goal: Solve least squares regression problem faster

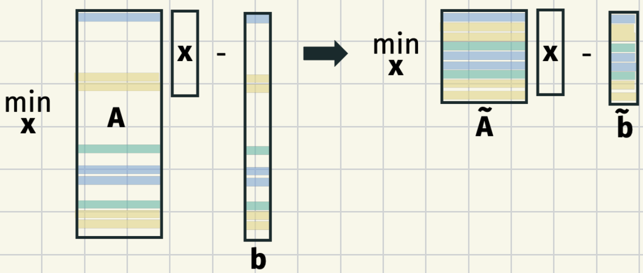
$A \in \mathbb{R}^{n \times d}$ $\quad b \in \mathbb{R}^n$

$$x^* = \underset{x \in \mathbb{R}^d}{\arg\min} \|Ax - b\|_2^2 = (A^T A)^{-1} A^T b$$

Time complexity:

Q: What if n (# cols) or d (# rows) is huge?





Idea: Find optimal solution to approximate problem

$$x^* = \underset{x \in \mathbb{R}^d}{\arg\min} \|Ax - b\|_2^2 \quad \text{vs.} \quad \tilde{x} = \underset{x \in \mathbb{R}^d}{\arg\min} \|\Pi A x - \Pi b\|_2^2$$

$\Pi \in \mathbb{R}^{m \times n}$ is JL projection matrix for $m << n$

Theorem: when $m = O(d/\epsilon^2)$ then, wp 9/10,

$$\|A\tilde{x} - b\|_2^2 \leq (1+\epsilon)\|Ax^* - b\|_2^2$$

That is, $\tilde{x}$ gives <u>almost</u> as good of a solution as x^*

Claim: Under the same conditions, for all $x \in \mathbb{R}^d$,

$$(1-\epsilon)\|Ax-b\|_2^2 \le \|\Pi Ax - \Pi b\|_2^2 \le (1+\epsilon)\|Ax-b\|_2^2.$$

Proof of Theorem using claim:

Plan: Let's prove the claim :)

Distributional JL: If $m = O\left(\frac{\log(1/\delta)}{\epsilon^2}\right)$ then, for a fixed y, wp $1-\delta$,

$$(1-\epsilon)\,\|y\|_2^2 \le \|\Pi y\|_2^2 \le (1+\epsilon)\,\|y\|_2^2 .$$

Challenge: JL holds for one y whereas claim holds for infinite $Ax-b$

Idea: Yes, $Ax-b \in \mathbb{R}^n$ but lies in d-dimensional subspace (i.e., linear combo of $d+1$ vecs).

Subspace Embedding: Let $\mathcal{U} \subset \mathbb{R}^n$ be a d-dim subspace.

If $m = O\left(\frac{d \log(1/\epsilon) + \log(1/\delta)}{\epsilon^2}\right)$ then, wp $1-\delta$, for all $y \in \mathcal{U}$,

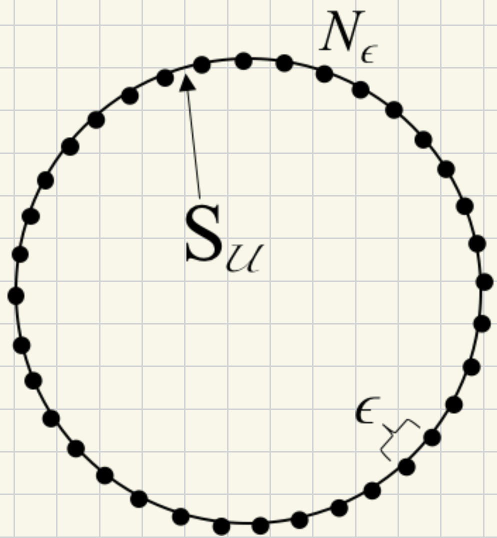
$$(1-\epsilon)\|y\|_2^2 \le \|\Pi y\|_2^2 \le (1+\epsilon)\|y\|_2^2.$$

Notice the claim follows since $\mathcal{U} = \{Ax - b : \forall x \in \mathbb{R}^d\}$ is $(d+1)$-dim.

Notice it suffices to prove the embedding result for w with $\|w\|_2^2 = 1$ since the statement is scale-invariant.

Let $S_{\mathcal{U}} = \{w \in \mathcal{U} : \|w\|_2 = 1\}$

Epsilon-Net : There is a "small" set of points that cover unit sphere



For $\epsilon < 1$, there exists $N_\epsilon \subset S_u$ with

① $|N_\epsilon| \leq \left(\frac{3}{\epsilon}\right)^d$

② For all $v \in S_u$, $\min_{w \in N_\epsilon} \|v - w\|_2 \leq \epsilon$

Thursday, April 16

Midterm Thursday, April 23

↳ virtual OH Monday

↳ no class Tuesday

↳ no OH Thursday

Project

Proposal due April 27

↳ explore topic/algorithm of your choice from class

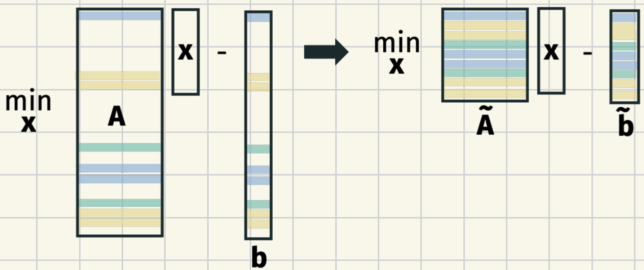
Today

Continue sketched regression!

Sketched Regression

$A \in \mathbb{R}^{n \times d}$, $b \in \mathbb{R}^n$

Idea: Project Columns from $\mathbb{R}^n$ to $\mathbb{R}^d$



$$x^* = \underset{x}{\operatorname{argmin}} \|Ax - b\|_2^2$$

$\Pi \in \mathbb{R}^{m \times n}$ JL matrix

$$\tilde{x} = \underset{x}{\operatorname{argmin}} \|\Pi A x - \Pi b\|_2^2$$

Theorem: When $m = O(d/\epsilon^2)$, wp $9/10$,

$$\|A\tilde{x} - b\|_2^2 \le (1+\epsilon)\|Ax^* - b\|_2^2.$$

Claim: when $m = O(d/\epsilon^2)$, wp $9/10$, for all $x \in \mathbb{R}^d$,

$$(1-\epsilon)\|Ax-b\|_2^2 \le \|\Pi Ax - \Pi b\|_2^2 \le (1+\epsilon)\|Ax-b\|_2^2.$$

$$\mathcal{U} = \{Ax - b : x \in \mathbb{R}^d\} \subset \mathbb{R}^n$$

Embedding Theorem: Let $\mathcal{U}$ be d-dim.

When $m = O\left(\frac{d \log(1/\epsilon) + \log(1/\delta)}{\epsilon^2}\right)$, wp $1-\delta$,

$$(1-\epsilon)\|y\|_2^2 \le \|\Pi y\|_2^2 \le (1+\epsilon)\|y\|_2^2 \quad \forall y \in \mathcal{U}$$

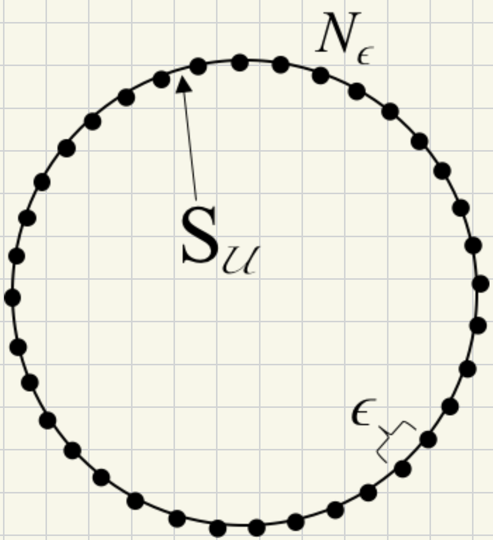
By scale-invariance, enough to show

for $S_{\mathcal{U}} = \{w \in \mathcal{U} : \|w\|_2 = 1\}$.

Distributional JL: If $m = O\left(\frac{\log(1/\delta)}{\epsilon^2}\right)$ for one $y \in \mathbb{R}^n$, wp $1-\delta$,

$$(1-\epsilon)\|y\|_2^2 \le \|\Pi y\|_2^2 \le (1+\epsilon)\|y\|_2^2$$

Epsilon-Net: There is a "small" set of points that cover unit sphere



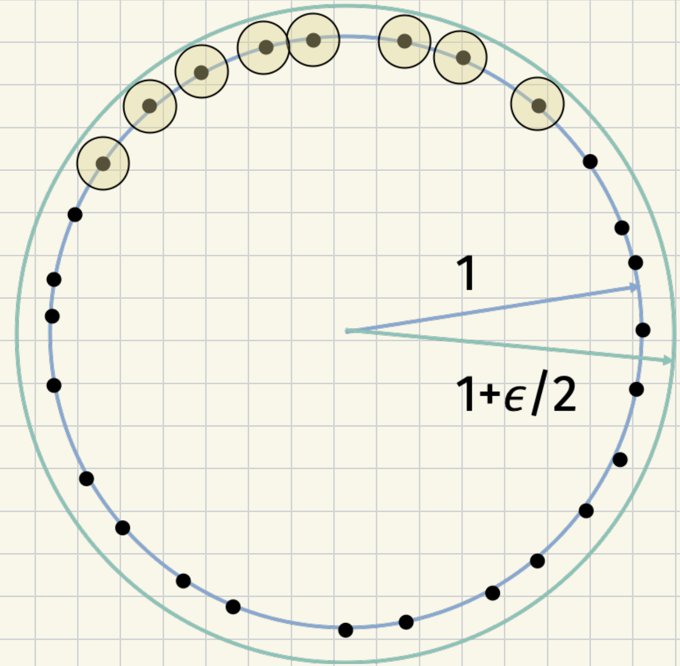
For $\epsilon < 1$, there exists $N_\epsilon \subset S_u$ with

① $|N_\epsilon| \le \left(\frac{3}{\epsilon}\right)^d$

② For all $v \in S_u$, $\min_{w \in N_\epsilon} \|v-w\|_2 \le \epsilon$

(Hypothetical) Epsilon-Net Construction:

Greedily add points that are at least ϵ-away from other points while they exist



② clearly holds, let's prove ① ...

Recall $\text{vol}(d, r) = c_d r^d$.

$$\text{vol}\left(d, \tfrac{\epsilon}{2}\right) |\mathcal{N}_\epsilon| \leq \text{vol}\left(d, 1 + \tfrac{\epsilon}{2}\right)$$

$$\Rightarrow |\mathcal{N}_\epsilon| \leq \frac{\left(1+\frac{\epsilon}{2}\right)^d}{(\epsilon/2)^2} = \left(\frac{2+\epsilon}{2} \cdot \frac{2}{\epsilon}\right)^d \leq \left(\frac{3}{\epsilon}\right)^d$$

Set $\delta = \frac{1}{10|\mathcal{N}_\epsilon|}$. By JL, wp $1-1/10$,

$$(1-\epsilon)\|w\|_2^{\cancel{2}} \le \|\Pi w\|_2^{\cancel{2}} \le (1+\epsilon)\|w\|_2^{\cancel{2}}$$

when $m = O\left(\frac{\log(1/\delta)}{\epsilon^2}\right) =$

Q: How about points $v \in S_u$ but not $\mathcal{N}_\epsilon$?

Write $v = w_0 + c_1 w_1 + c_2 w_2 + \dots$ where

1. $w_1, w_2, \dots \in \mathcal{N}_\epsilon$
2. $|c_\ell| \le \epsilon^\ell \quad \forall$

$w_0 = \operatorname{argmin}_{w \in \mathcal{N}_\epsilon} \|v - w\|_2$ $\quad r_1 = v - w_0$ $\quad c_1 = \|r_1\|_2$

$w_1 = \operatorname{argmin}_{w \in \mathcal{N}_\epsilon} \|\frac{r_1}{c_1} - w\|_2$ $\quad r_2 = v_1 - w_0 - c_1 w_1$ $\quad c_2 = \|r_2\|_2$

$w = \operatorname{argmin}_{w \in \mathcal{N}_\epsilon} \|\frac{r_2}{c_2} - w\|_2$ $\quad r_3 = v_1 - w_0 - c_1 w_1 - c_2 w_2$ $\quad c_3 = \|r_3\|_2$

Induction: $|c_\ell| = \|r_\ell\|_2 \leq \epsilon^\ell$

Base case:

Inductive step:

$a, b \in \mathbb{R}^n$

$$\|a+b\|_2^2 = \|a\|_2^2 + 2\langle a,b\rangle + \|b\|_2^2$$

$$\le \|a\|_2^2 + 2\|a\|_2\|b\|_2 + \|b\|_2^2 \quad \text{by} \quad 1 \ge \cos\theta = \frac{\langle a,b\rangle}{\|a\|_2\|b\|_2}$$

$$\le (\|a\|_2 + \|b\|_2)^2$$

For $v \in S_u$,

$$\|\Pi v\|_2 = \|\Pi w_0 + c_1 \Pi w_1 + c_2 \Pi w_2 + \dots\|_2$$

$$\le \|\Pi w_0\|_2 + c_1\|\Pi w_1\|_2 + c_2\|\Pi w_2\|_2 + \dots$$

$$\le (1+\epsilon)\|w_0\|_2 + \epsilon(1+\epsilon)\|w_1\|_2 + \epsilon^2(1+\epsilon)\|w\|_2 + \dots$$

$$\le (1+\epsilon)(1+\epsilon+\epsilon^2+\dots) = \frac{1+\epsilon}{1-\epsilon} = 1 + O(\epsilon)$$

$$\|a\|_2 = \|a+b-b\|_2^2 \leq \|a+b\|_2 + \|b\|_2$$

$$\Rightarrow \|a+b\|_2 \geq \|a\|_2 - \|b\|_2$$