Plan

Logistics

Review

Fast JL

Game Night (6pm Wednesday)
↳ drop lowest HW problem
(come talk to me if you can't attend)

Problems

Project
↳ Presentation
↳ Work wednesday

Course Response Forms
Tomorrow (please bring computer)

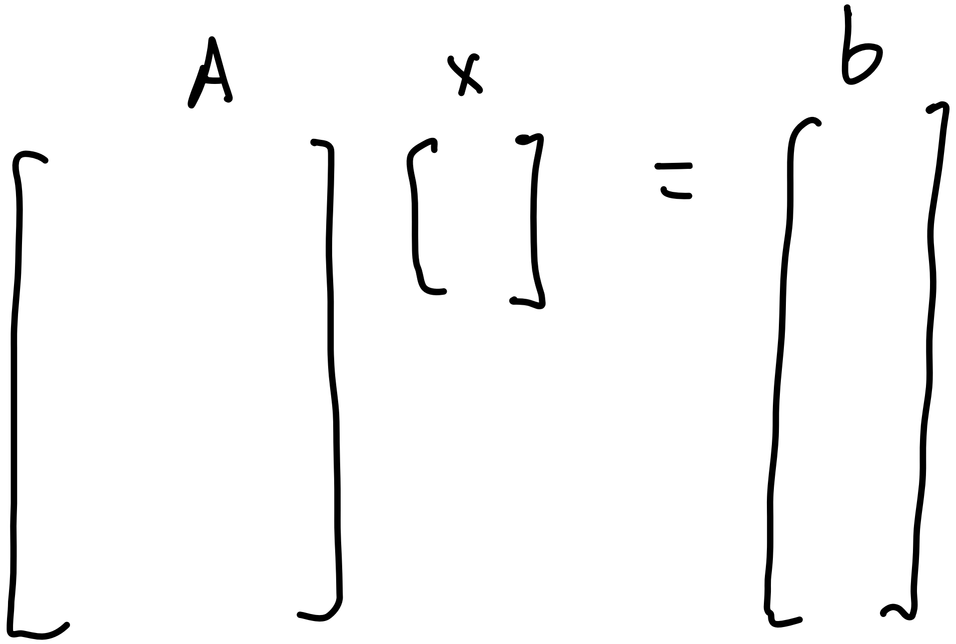
Regression

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

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

$$x^* = (A^T A)^{-1} A^T b$$

compute in $O(nd^2)$



Sketched Regression

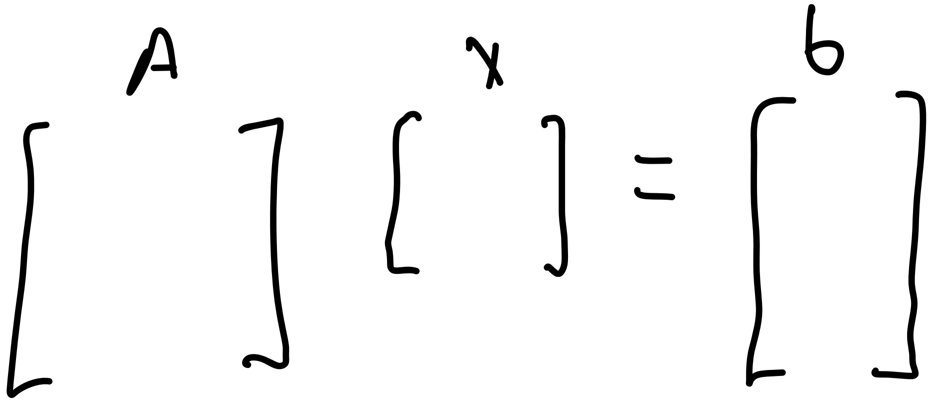
$\Pi \in \mathbb{R}^{m \times n}$

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

$$\tilde{x} = (A^T \Pi^T \Pi A)^{-1} A^T \Pi^T b$$

Compute in $O(md^2)$

~~if we have~~ ~~$\underset{m \times n}{\Pi}\underset{n \times d}{A}$~~



Fast JL

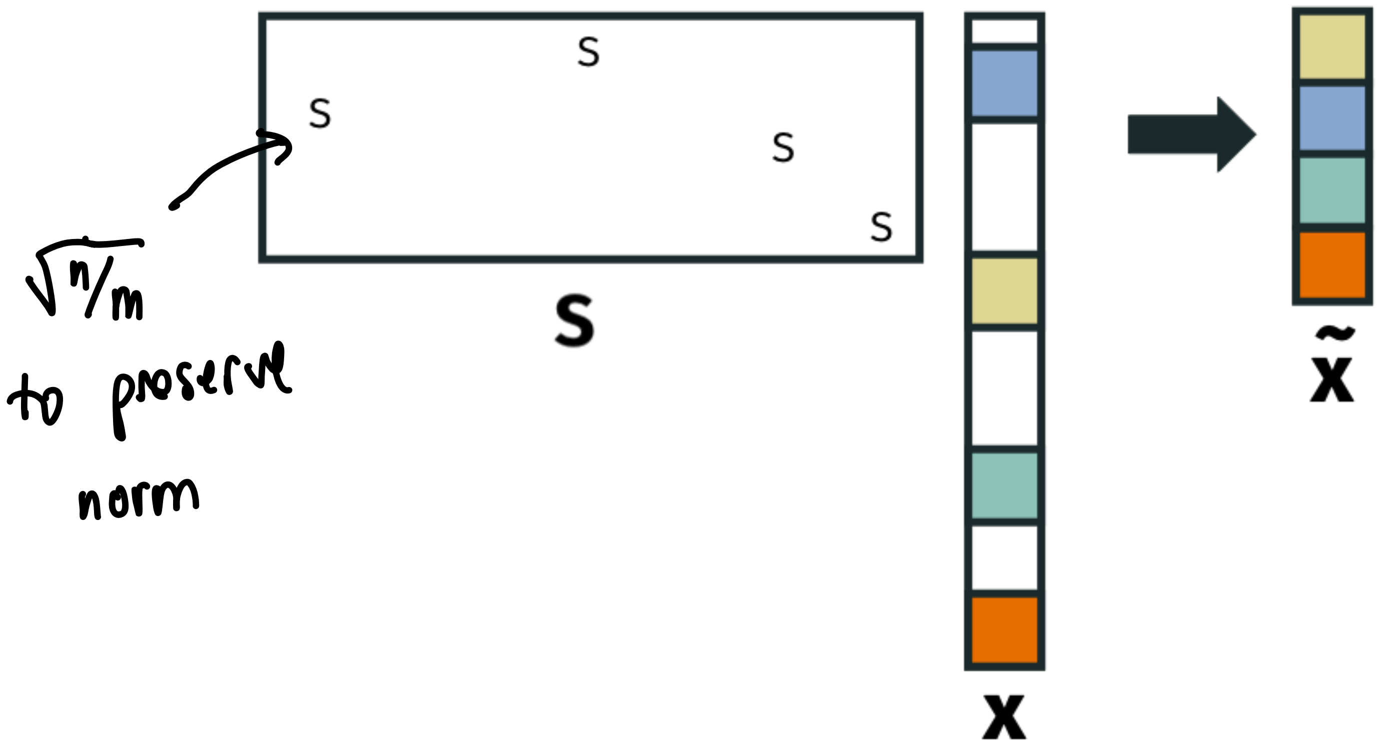
close to reading A

↓

Goal: Compute $\underset{m\times n}{\Pi}\ \underset{n\times d}{A}$ in $O(n\log n \cdot d)$ time

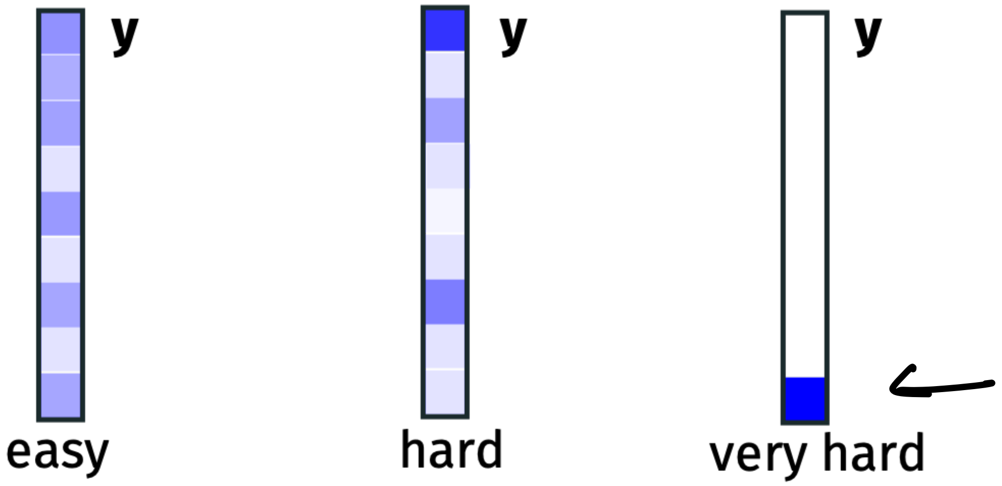
Approach: Compute Πx in $O(n\log n)$ time for each column

$S \in \mathbb{R}^{m\times n}$ is sampling matrix



What could go wrong?

Approach only works if x is "flat"



← unlikely to get this dimension

Claim (from Hoeffding): If $x_i^2 \le \frac{c}{n} \|x\|_2^2$ for all i then $m = O\left(\frac{c \cdot \log(1/\delta)}{\epsilon^2}\right)$ samples suffice to preserve ℓ_2 norm within $1 \pm G$ wp $1-\delta$

How do we make x flatter?

Use a mixing matrix M

$$\Pi x = \underset{\uparrow \atop \text{sampling}}{S} \; \underset{\uparrow \atop \text{mixing}}{M} \; x$$

- $\|Mx\|_2^2 = \|x\|_2^2$ exactly
- $[Mx]_i^2 \leq \frac{c}{n} \|x\|_2^2$ whp
- Compute Mx in $O(n \log n)$ time

Does M have to be random?

We will make M pseudorandom: $M = HD$

- $D \in \mathbb{R}^{n \times n}$ diagonal with entries $D_{i,i} = \pm 1$ uniform
- $H \in \mathbb{R}^{n \times n}$ is Hadamard matrix

Assume n is power of 2

$$H_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{bmatrix} \qquad H_k = \frac{1}{\sqrt{n}} \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix}$$

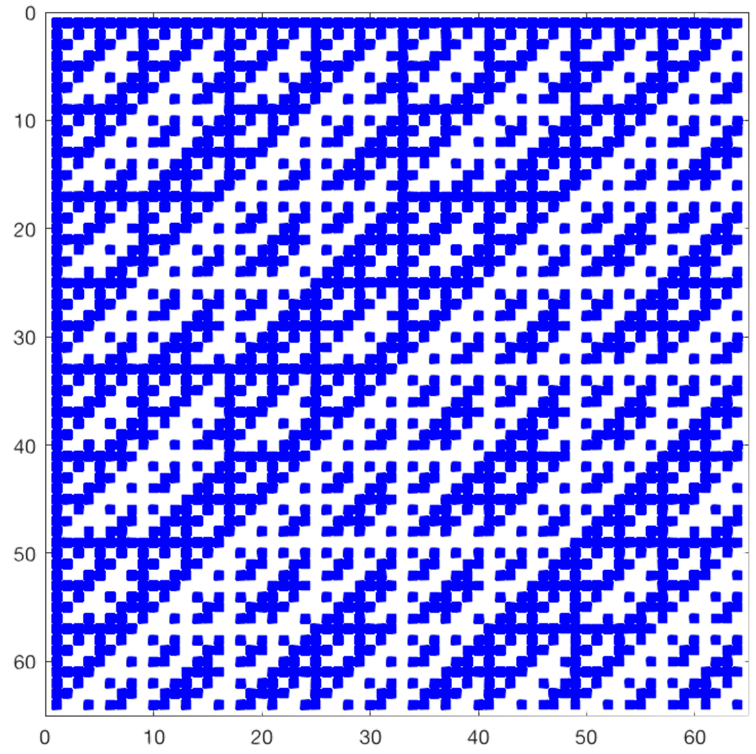
Property 1: $\|H_k x\|_2^2 \underset{\substack{\uparrow \\ \text{Show}}}{=} \|x\|_2^2$

Property 2: Compute $\Pi_x = SHD_x$ in $O(n \log n)$ time

↑

Show

Property 3: The randomized Hadamard HD is good for mixing



Lemma: $z = HDx$ $\qquad z_i^2 \le \frac{c \log(n/\delta)}{n} \|z\|_2^2$ wp $1-\delta$

$h_i^T =$ ith row of H

$$\frac{1}{\sqrt{n}} \; h_i^T \rightarrow \underset{H}{\begin{bmatrix} & & & & \\ -1 & +1 & +1 & -1 & -1 \\ & & & & \end{bmatrix}} \underset{D}{\begin{bmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & +1 \end{bmatrix}} \underset{x}{\begin{bmatrix} \\ \\ \\ \end{bmatrix}}$$

$$h_i^T D = \frac{1}{\sqrt{n}} [-1 \; +1 \; +1 \; -1 \; -1] \begin{bmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & +1 \end{bmatrix} = \frac{1}{\sqrt{n}} [R_1 \quad R_2 \quad \dots \quad R_n]$$

← Rademacher ±1

$$z_i = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_i \cdot x_i \qquad \mathbb{E}[z_i] = 0 \qquad \mathrm{Var}(z_i) = \frac{1}{n} \|x\|_2^2$$

$z_i = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_i \cdot x_i$ $\qquad$ $\mathbb{E}[z_i] = 0$ $\qquad$ $\mathrm{Var}(z_i) = \frac{1}{n} \|x\|_2^2$

Concentration (Rademacher)

$$\Pr\left(\sum_{i=1}^{n} R_i a_i \geq t \|a\|_2 \right) \leq e^{-t^2/2}$$

$$\Pr\Bigg(\underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} R_i x_i}_{z_i} \geq \underbrace{t \cdot \frac{1}{\sqrt{n}}}_{\left(\frac{c \log(n/\delta)}{n}\right)^{1/2}} \|x\|_2 \Bigg) \leq e^{-t^2/2}$$

$$\Rightarrow t = \sqrt{\log n/\delta}$$

$$\Updownarrow$$

$$\Pr\left(z_i^2 \geq \frac{c \log(n/\delta)}{n} \right) \leq e^{-\log n/\delta} = \frac{\delta}{n}$$