Tuesday, April 28

Exam:

50 79/79.5 100

$\xleftrightarrow{\quad 15 \quad}$

Grades: All B- or above

Hypothetical

(details on Discord and Canvas)

Plan

Proposals due

↳ I'll message w/ feedback

Fast JL Transform today

Explainable AI next

Talks

Ahmet 4-5pm Wednesday @ Davidson

Feyza 4:15-5:15pm Friday @ Estella 1021

Regression Review

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

$x^* = \underset{x \in \mathbb{R}^d}{\operatorname{argmin}} \|Ax - b\|_2^2$ takes $O(nd^2)$ time to solve

Let $m \ll n$ where $m \approx O(d)$, $\Pi \in \mathbb{R}^{m \times n}$ JL transform

$\tilde{x} = \underset{x \in \mathbb{R}^d}{\operatorname{argmin}} \|\Pi A x - \Pi b\|_2^2$ takes $O(md^2)$ to solve

If we already have ΠA

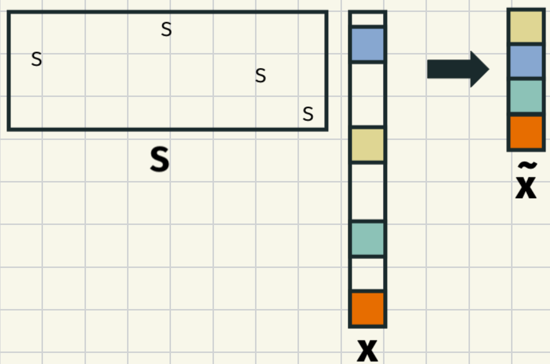
Time to compute: $O(mnd) \approx O(nd^2)$

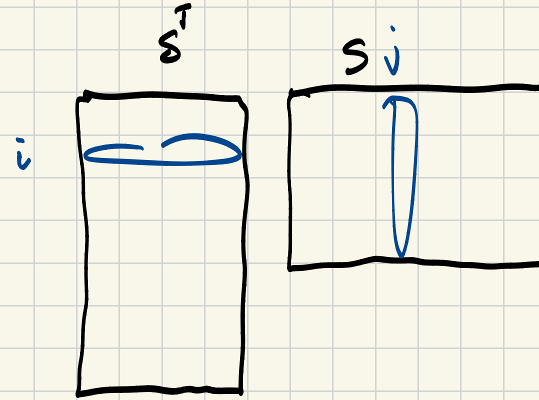
Today: Build Π so that

1. $(1-\epsilon)\|x\|_2 \le \|\Pi x\|_2 \le (1+\epsilon)\|x\|_2$ whp
2. Compute Πx in $O(n \log n)$ time

Sampling Matrix $S \in \mathbb{R}^{m \times n}$

$\boxed{\text{Time}: O(m)}$



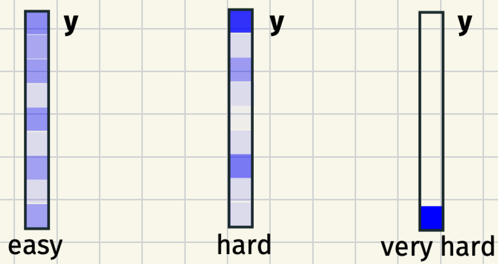


Every row contains a single value $s = \sqrt{\frac{n}{m}}$ in a random column

$$\mathbb{E}\, \|Sx\|_2^2 = \mathbb{E}\, x^T S^T S x$$
$$= x^T \, \mathbb{E}[S^T S] \, x$$
$$= x^T x$$
$$= \|x\|_2^2$$

$$\mathbb{E}[(S^T S)_{i,j}] = \mathbb{E} \sum_{k=1}^{m} [S^T]_{i,k} [S]_{k,j}$$
$$= \mathbb{E} \sum_{k=1}^{m} \mathbb{1}[i \text{ in row } k] \, \mathbb{1}[j \text{ in row } k] \frac{n}{m}$$
$$= \sum_{k=1}^{m} \Pr(i \text{ in row } k, j \text{ in row } k) \frac{n}{m}$$
$$= \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

Concentration



Claim: If $x_i^2 \le \frac{c}{n}\|x\|_2^2$ for all i,

$m = O\left(c \log(1/\delta)/\epsilon^2\right)$ preserve ℓ_2-norm whp.

Challenge: Guarantee $x_i^2 \le \frac{c}{n}\|x\|_2^2$ whp.

Build M so that

1. $\|Mx\|_2^2 = \|x\|_2^2$ $\iff$ $M^T M = I$ $\iff$ $M^{-1} = M^T$

so $\exists\, y$ s.t. $My = e_1 \iff y = M^{-1} e_1$

2. $[Mx]_i^2 \le \frac{c}{n}\|x\|_2^2$

very hard input so make M random

3. Compute Mx in $O(n \log n$ time$)$

$M = HD$ where $D \in \mathbb{R}^{n \times n}$ is diagonal with ± 1

$H \in \mathbb{R}^{n \times n}$ is Hadamard

Assume n is a power of 2.

$H_0 = [1]$

$H_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

$H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} H_1 & H_1 \\ H_1 & -H_1 \end{bmatrix}$

$\vdots$

$H_k = \frac{1}{\sqrt{2}} \begin{bmatrix} H_k & H_k \\ H_k & -H_k \end{bmatrix}$

Each entry is $\pm \frac{1}{\sqrt{2}^k} = \pm \frac{1}{\sqrt{n}}$

Property ①: $H_k^T H_k = I \iff \|H_k x\|_2 = \|x\|_2$

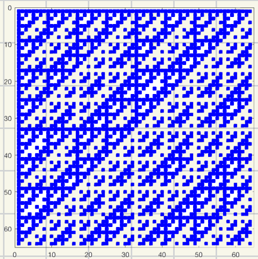
Hint: Use induction

Property ②: Computing $H_k x$ is $O(n \log n)$

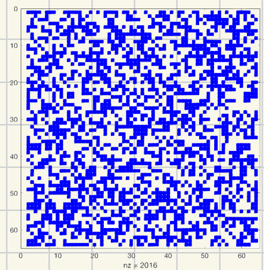
Hint: Write $x = \begin{bmatrix} x_a \\ x_b \end{bmatrix}$ and use master method

Property ③ : HD is a good mixing matrix

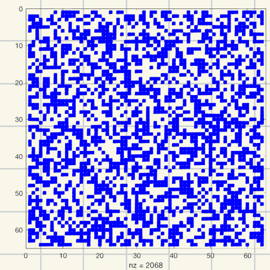
H



HD



Uniform ± 1



SHRT Mixing Lemma: Let $z = HDx$. wp $1-\delta$

uses rademacher concentration

$$z_i^2 \le \frac{c \log \frac{n}{\delta}}{n} \|z\|_2^2 .$$

$$\Pi x = \underbrace{S \underbrace{H \underbrace{D\, x}_{O(n)}}_{O(n \log n)}}_{O(m)}$$

$$(1-\epsilon)\|x\|_2 \le \|\Pi x\|_2 \le (1+\epsilon)\|x\|_2 \quad \text{whp.}$$

Thursday, April 30

- Final time slot Fri. May 15 2-5pm
- Project Proposals
- Thank you for coming to Ahmet's talk
- Feyza's talk tomorrow at 4:15pm in Estella 1021

Explainable AI

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix} \rightarrow \blacksquare \rightarrow f(x)$$

Q: How does changing i^{th} feat. change f?

Linear Regression

$$f(x) = \sum_{i=1}^{d} w_i x_i$$

A: When we change i^{th} feature from b_i to x_i, f changes by

$$\phi_i = w_i(x_i - b_i)$$

Application:

- ↪ In mortgage model, check ϕ_{race} is 0
- ↪ In hiring model, check ϕ_{gender} is 0

Big Question: What about non-linear models like neural networks?

Idea: Change b_i to x_i in context of all possible subsets

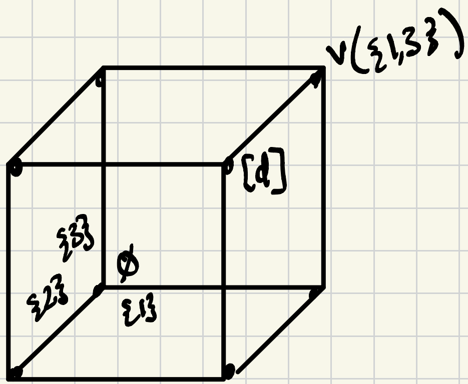
Baseline: $b \in \mathbb{R}^d$

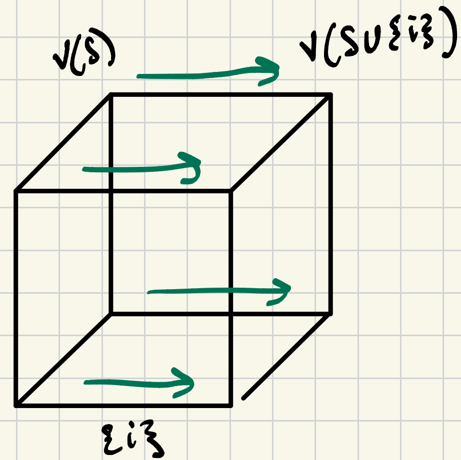
Explicand: $x \in \mathbb{R}^d$

Given $S \subseteq \{1, \dots, d\}$,

$$x^S_i = \begin{cases} x_i & \text{if } i \in S \\ b_i & \text{if } i \notin S \end{cases}$$

Let $v : 2^{[d]} \to \mathbb{R}$, $v(S) = f(x^S)$.



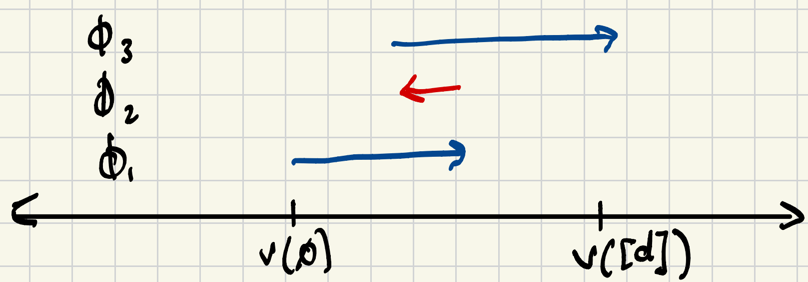


Shapley Value

$$\phi_i = \sum_{S \subseteq [d] \setminus \{i\}} [v(S \cup i) - v(S)]\, \mu_S \qquad \text{where} \quad \mu_S = \frac{1}{d\binom{d-1}{|S|}}$$

$$= \underbrace{\frac{1}{d} \sum_{\ell=0}^{d-1} \underbrace{\frac{1}{\binom{d-1}{\ell}} \sum_{\substack{S \subseteq [d]\setminus\{i\} \\ |S| = \ell}} v(S \cup i) - v(S)}_{\text{average for size } \ell}}_{\text{average over size}}$$

Efficiency: $\displaystyle\sum_{i=1}^{d} \phi_i = v([d]) - v(\emptyset)$



Monte Carlo Estimator

Sample m sets S wp p_S

choose $\mu_S = p_S$

$$\tilde{\phi}_i^{MC} = \frac{1}{m} \sum_{j=1}^{m} \left[v(S_j \cup i) - v(S_j) \right] \frac{\mu_S}{p_S} = \frac{1}{m} \sum_{j=1}^{m} v(S_j \cup i) - v(S_j)$$

$$\mathbb{E}[\tilde{\phi}_i^{MC}] = \mathbb{E} \frac{1}{m} \sum_{j=1}^{m} \sum_{S \subseteq [d] \setminus i} \mathbb{1}[S = S_j] \left[v(S \cup i) - v(S) \right]$$

$$= \frac{1}{m} \cdot m \sum_{S \subseteq [d] \setminus i} \Pr(S = S_j) \left[v(S \cup i) - v(S) \right]$$

$$= \sum_{S} \left[v(S \cup i) - v(S) \right] \mu_S = \phi_i$$

$$\mathrm{Var}(\tilde{\phi}_i^{MC}) =$$

$$\leq \frac{1}{m} \sum_{S} [v(S \cup i) - v(S)]^2 \mu_S$$

Disadvantage:

Maximum Sample Reuse

$$\phi_i = \sum_{S \subseteq [d] \setminus i} [v(S \cup i) - v(S)]\, \mu_s$$

$$= \sum_{S : i \in S} v(S)\, \mu_{s-1} - \sum_{S : i \notin S} v(S)\, \mu_s$$

$$= \sum_{S} v(S) \left(\mathbb{1}[i \in S]\, \mu_{s-1} - \mathbb{1}[i \notin S]\, \mu_s \right)$$

$$\tilde{\phi}_i^{MSR} = \frac{1}{m} \sum_{j=1}^{m} v(S_j) \left(\mathbb{1}[i \in S]\, \mu_{s-1} - \mathbb{1}[i \notin S]\, \mu_s \right) \frac{1}{p_s}$$

$\mathrm{Var}(\tilde{\phi}_i^{MSR})$ depends on ...

Linear Regression Perspective

$\boxed{v(\emptyset)=0 \quad \text{WLOG}}$

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_d \end{bmatrix} = \underset{\beta \in \mathbb{R}^d : \langle \beta, \mathbb{1} \rangle = v([d])}{\arg\min} \sum_{S : 0 < |S| < d} w_S \left(v(S) - \sum_{i \in S} \beta_i \right)^2$$

$$\text{where} \quad w_S = \frac{1}{\binom{d}{|S|} |S| (d - |S|)}$$

$$A' \in \mathbb{R}^{2^d - 2 \times d}, \quad b' \in \mathbb{R}^{2^d - 2}$$

For $\emptyset \subset S \subset [d]$, $\quad [A']_{S,i} = \mathbb{1}[i \in S] \sqrt{w_S}, \quad [b']_S = v(S) \sqrt{w_S}$

$$\phi = \underset{\beta \in \mathbb{R}^d : \langle \beta, \mathbb{1} \rangle = v([d])}{\arg\min} \| A'\beta - b' \|_2^2$$

weird form...

$$\phi = \underset{\beta \in \mathbb{R}^d : \langle \beta, 1 \rangle = v([d])}{\operatorname{argmin}} \| A'\beta - b' \|_2^2 \qquad \left\lfloor \langle 1, 1 \tfrac{v(d)}{d} \rangle = v(d) \right.$$

$$= \underset{\beta_\perp \in \mathbb{R}^d : \langle \beta_\perp, 1 \rangle = 0}{\arg\min} \| A'(\beta_\perp + 1 \tfrac{v([d])}{d}) - b' \|_2^2 + \frac{1 v([d])}{d}$$

$$= \underset{\beta \in \mathbb{R}^d}{\operatorname{argmin}} \| \underbrace{A' P}_{A} \beta + \underbrace{A' 1 \tfrac{v(d)}{d} - b'}_{-b} \|_2^2 + 1 \frac{v(d)}{d}$$

where $P = I - \frac{1}{d} 1 1^T$ so $\langle 1, P\beta \rangle = 1^T P \beta$

$$= 1^T (\beta - \tfrac{1}{d} 1 1^T \beta)$$

$$= 1^T \beta - \tfrac{d}{d} 1^T \beta = 0$$

$$= \underset{\beta \in \mathbb{R}^d}{\operatorname{argmin}} \| A\beta - b \|_2^2 + 1 \frac{v(d)}{d}$$