

# FASTER LEAST SQUARE APPROXIMATION

Let  $[n] = \{1, 2, 3, \dots, n\}$

Recall Hadamard transform:  $H_n$  defined recursively:

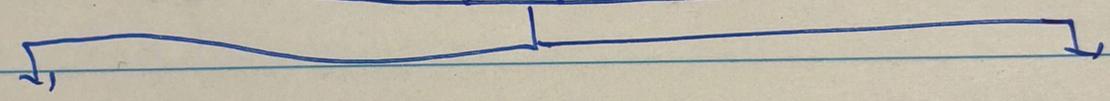
$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad H_n = \begin{bmatrix} H_{n/2} & H_{n/2} \\ H_{n/2} & -H_{n/2} \end{bmatrix} \rightarrow (1)$$

$H = \frac{1}{\sqrt{n}} H_n$  - normalized version:  $n = 2^k$

Let  $D = \text{diag}(d_{11}, d_{12}, \dots, d_{nn}) \in \mathbb{R}^{n \times n}$  }  $\rightarrow (2)$   
 $d_{ii} = \pm 1$  with probability  $1/2$

$DH$  is known as the randomized Hadamard transform.

## Two properties of $(DH)$



When applied to a vector  $x$ , it spreads its energy in the sense that  $\|x\|_\infty$  is bounded after application of  $(DH)$

The matrix vector product  $(DH)x$  can be computed in  $O(n \log n)$  time as in standard Fast Fourier transform

In computing  $(DH)x$ , there is no multiplication involved.

Linear least squares problem:-

Let  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$   $m > n$ . A-Full rank

Our goal is to find  $x$  such that

$$z = \min_{x \in \mathbb{R}^n} \|Ax - b\|_2^2 \longrightarrow (3)$$

$$\Rightarrow x_{LS} = (A^T A)^{-1} A^T b = A^+ b, \quad A^+ = (A^T A)^{-1} A^T \longrightarrow (4)$$

Let  $B$  be preconditioning matrix.  $B \in \mathbb{R}^{n \times n}$   
( $n \leq m$ )

Define  ~~$\tilde{z}$~~  such that

$$\tilde{z} = \min_{x \in \mathbb{R}^n} \|B(Ax - b)\|_2^2 \longrightarrow (5)$$

$$\Rightarrow \tilde{x}_{LS} = (BA)^+ (Bb), \quad (BA)^+ = [(BA)^T (BA)]^{-1} (BA)^T \longrightarrow (6)$$

We can find  $\tilde{x}_{LS}$  by using conjugate gradient algorithm.

$B$  also viewed as downscaling operator

Two choices for B

Random Sampling

Random Projection

$B = S^T H D \rightarrow (7)$

$B = T H D \rightarrow (8)$

• HD Randomized Hadamard Transform

• HD-Randomized Hadamard Transform

• S - Sampling matrix

• T is a random projection matrix

Choice of S:  $S \in \mathbb{R}^{m \times n}$  - Initially empty

1) For  $i = 1, 2, \dots, n$ , select uniformly at random an integer from  $[n]$  - iid trials with replacement.

• If  $i$  is that integer, add  $\sqrt{\frac{m}{n}} e_i$  to S  
 $e_i$  = Standard unit vector in  $\mathbb{R}^m$

End For

Choice of T:  $T = [T_{ij}] \in \mathbb{R}^{n \times m}$

1)  $T_{ij} \sim \text{i.i.d. } \mathcal{N}(0, 1)$

2)  $T_{ij} \sim \{-1, 1\}$  with probability  $1/2, 1/2$

3)  $T_{ij} \sim \{\pm\sqrt{3}, 0\}$   $T_{ij} = 0$  with prob =  $2/3$   
 $\quad \quad \quad \quad \quad \quad \quad = \pm\sqrt{3}$  " " =  $1/6$

(4)

Claim: Let  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $\epsilon \in (0, 1)$ . Then

there exists a randomized algorithm that returns a vector  $\tilde{x}_{\text{opt}} \in \mathbb{R}^n$  such that

$$(a) \quad \tilde{x}_{\text{opt}} : \|A \tilde{x}_{\text{opt}} - b\|_2 \leq (1 + \epsilon) \zeta \rightarrow (9)$$

where  $\zeta$  is defined in (3), with a large probability.

(b) If  $\kappa(A)$  is the Condition number of  $A$ ,  $\gamma \in (0, 1)$ ,

the fraction of the norm of  $b$  that lies in the column space of  $A$ , is

$$\gamma = \frac{\|U_A U_A^T b\|_2}{\|b\|_2} \rightarrow (10)$$

where  $U_A$  is the orthonormal basis for the column space of  $A$ , then

$$\|x_{\text{opt}} - \tilde{x}_{\text{opt}}\|_2 \leq \epsilon \left[ \kappa(A) \sqrt{\frac{1}{\gamma} - 1} \right] \|x_{\text{opt}}\|_2$$

When structural conditions are true:  $\rightarrow (11)$

$$\sigma_{\min}^2(B U_A) \gg \frac{1}{\sqrt{2}} \rightarrow (12)$$

$$\|U_A^T B^T b\|_2 \leq \frac{\epsilon}{2} \zeta^2 \rightarrow (13)$$

where  $\zeta$  is in (3).

Recall:  $A = U_A \Sigma_A^{1/2} V_A^T$  - SVD of  $A$ .