Fast Sketching with Fast Walsh-Hadamard Transform

Beatrice Aragones[1], Henry Boateng[1], Xiaoye S. Li[2]

[1]San Francisco State University, [2]Lawrence Berkeley National Laboratory

Background Information

Matrix sketching preserves some properties of a very large matrix into a much smaller matrix with reduced dimension.

$$\mathbf{A} \in \mathbb{R}^n \rightarrow \mathbb{R}^d, d < n$$

By having a low rank approximation of a big matrix, it effectively reduces the size of the problem and makes it significantly easier and faster to compute.

We ensure a high probability sketch through **random sampling**. Randomization reduces the complexity of algorithms and increases its speed.

Fast Walsh-Hadamard Transform

When performing matrix sketching by SRHT, we use the **Fast-Walsh Hadamard transform (FWHT)** over the regular Hadamard transform.

Both transformations make use of the Hadamard matrix, which is a square matrix that is defined recursively and has dimensions of a power of 2.

$$H_n = H_2 \otimes H_{\frac{n}{2}}$$
$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

where ⊗ is the Kronecker product

Rather than multiplying with the Hadamard matrix, FWHT is computed algorithmically without needing to create the full Hadamard matrix.

$$x = [y_1 \quad y_2] \rightarrow xH_4 = [y_1 \quad y_2] \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix}$$
$$xH_4 = [y_1 H_2 + y_2 H_2 \quad y_1 H_2 - y_2 H_2]$$

Sketching Methods

The current sketching method consists of **random sampling**.

$$Y = \mathbf{A}\Omega$$

A $\in \mathbb{R}^{m,n}$ **:** an original $m \times n$ matrix
Ω $\in \mathbb{R}^{n,d}$ **:** a Gaussian random matrix
Cost: $O(mnd)$

To optimize, we produce and modify the Gaussian random matrix using the **Subsampled randomized Hadamard transform (SRHT)**.

$$\Omega = DHS$$

D $\in \mathbb{R}^{n,n}$ **:** a diagonal matrix whose diagonal entries are ±1
H $\in \mathbb{R}^{n,n}$ **:** Walsh-Hadamard transform matrix
S $\in \mathbb{R}^{n,d}$**:** a random subset of d columns from an $n \times n$ identity matrix
Cost: $O(mn \log(n))$

Bit-Reversal

Bit-reversal is a permutation where the indices of elements within a sequence are switched according to their binary representation. It simplifies the process of accessing elements in an array, thus making the algorithm run faster.

Bit-reversal can be implemented in FWHT for potential improved fast memory access. When performing the calculations, bit-reversal allows immediate addition and subtraction on the same numbers.

Omitting the permutation results in calculating all additions first on different values, followed by all subtractions in each iteration.

$$\begin{aligned} x_0 + x_1 &= \hat{x}_0 \rightarrow \hat{x}_0 + \hat{x}_2 \\ x_0 - x_1 &= \hat{x}_1 \rightarrow \hat{x}_1 + \hat{x}_3 \\ x_2 + x_3 &= \hat{x}_2 \rightarrow \hat{x}_0 - \hat{x}_2 \\ x_2 - x_3 &= \hat{x}_3 \rightarrow \hat{x}_1 - \hat{x}_3 \end{aligned}$$

FWHT without bit-reversal

$$\begin{aligned} x_0 + x_2 &= \hat{x}_0 \rightarrow \hat{x}_0 + \hat{x}_2 \\ x_0 - x_2 &= \hat{x}_1 \rightarrow \hat{x}_0 - \hat{x}_2 \\ x_1 + x_3 &= \hat{x}_2 \rightarrow \hat{x}_1 + \hat{x}_3 \\ x_1 - x_3 &= \hat{x}_3 \rightarrow \hat{x}_1 - \hat{x}_3 \end{aligned}$$

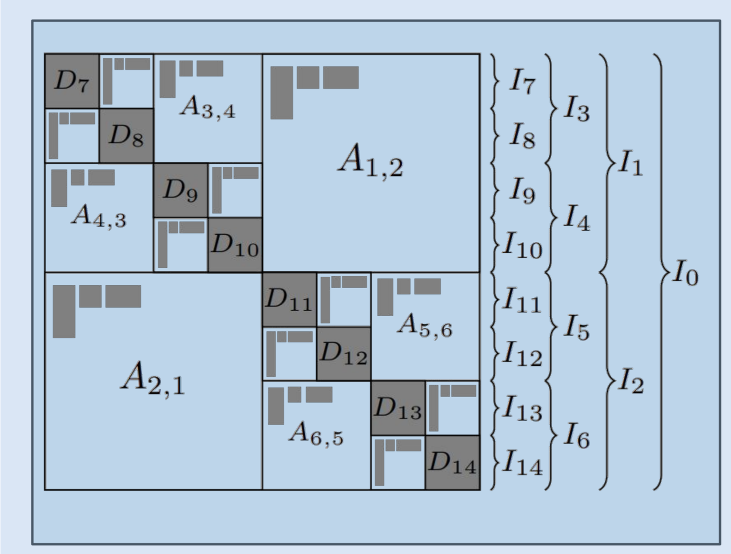
FWHT with bit-reversal

Application

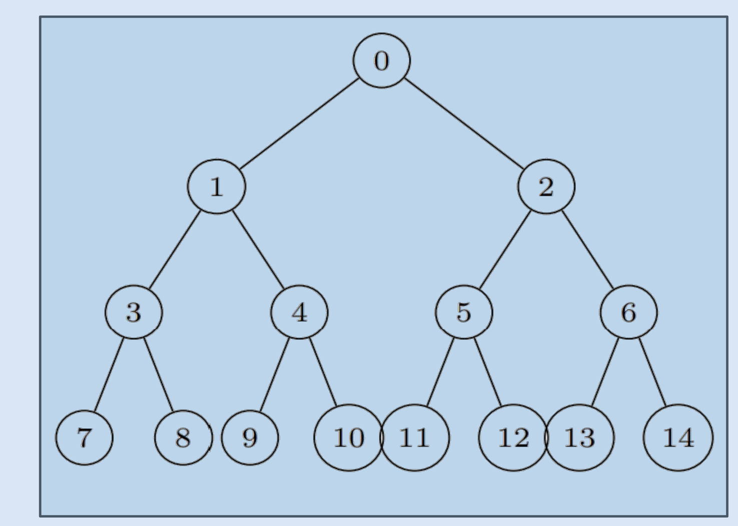
SRHT is one of the many sketching algorithms that can be used in a **Hierarchically Semi-Separable Matrix (HSS Matrix)**.

Other sketching algorithms that can be used are:
- Sparse Johnson-Lindenstrauss transform
- Subsampled randomized fast Fourier transform
- Subsampled randomized discrete cosine transform

The HSS matrix is a dense matrix where the off-diagonal blocks are recursively partitioned into smaller rank-deficient blocks.



An HSS matrix with 4 levels. Grey blocks denote the basis matrices, U, B, V, and D



The HSS matrix represented as a top-down level-by-level tree.

By compressing a matrix into its HSS form, we are able to perform computations significantly faster with a complexity of $O(rn)$ as opposed to $O(n^2)$, where r is the maximum rank and n is the size of the dense matrix.

References

- Fino and Algazi, "Unified Matrix Treatment of the Fast Walsh-Hadamard Transform," in IEEE Transactions on Computers, vol. C-25, no. 11, pp. 1142-1146, Nov. 1976, doi: 10.1109/TC.1976.1674569.
- Gorman, Christopher et al. "Matrix-free construction of HSS representation using adaptive randomized sampling." ArXiv abs/1810.04125 (2018): n. pag.
- Wang, Shen, Xia, Jianlin, Situ, Yingchong, and Hoop, Maarten V. de. Efficient scalable algorithms for hierarchically semiseparable matrices. United States: N. p., 2011. Web.