

# Accelerating Matrix Multiplication

Yash Malviya 2016CS50403  
Aniket Kumar 2016CS50397

February 17, 2019

## 1 Linear Algebra Libraries

The following was done to speed up convolution process required for solving Digit Recognition Problem using Convolutional Neural Networks.

Convolution is done by converting the image and kernel into corresponding Toeplitz matrices (Wikipedia [2]).

Function from Basic Linear Algebra Subprograms used **cblas\_sgemv**. Which perform following mathematical operation.

$$C \leftarrow \alpha A * B + \beta C$$

where A is processed Toeplitz image matrix, B is processed kernel column vector, C is zero matrix and  $\alpha(= 1.0)$  and  $\beta(= 0.0)$  are constants.

### 1.1 OpenBLAS

*Open source* implementation of the Basic Linear Algebra Subprograms: xianyi/OpenBLAS [3]

### 1.2 Intel MKL

Intel's Math Kernel Library, another implementation of BLAS: Intel [1]

## 2 Performance comparison

We get a increase in performance of calculating matrix convolution using these libraries. Intel MKL library give best performance among all these implementations.

As size of input matrix increases, time taken by PThread tends to time taken by IntelMKL library.

The comparison of performance of the three implementations(OpenBLAS , MKL and PThread) are shown as below :

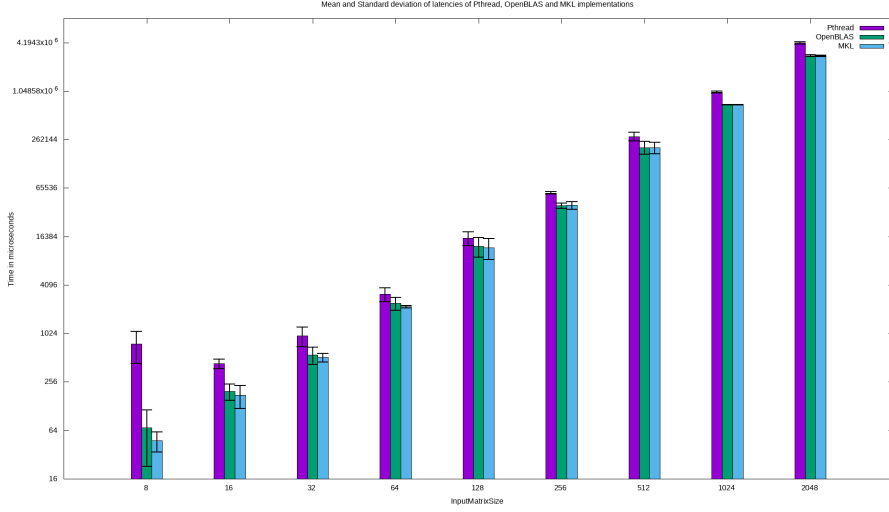


Figure 1: Plot of time V/S Matrix Size in different implementations

### 3 Acceleration with pthreads

#### 3.1 Parallelization Strategy

- Inner product of row of A and the column vector B is done by each thread.
- Each thread gets the row it operates on, in a round robin fashion.  
*For Ex:* say there are k threads, then thread with tid 0 gets 1<sup>st</sup> row, k+1<sup>th</sup> row 2k+1<sup>th</sup> row and so on.

#### 3.2 Difficulties encountered

1. Writing to only different sections of shared memory space.
2. Ensuring all shared memory which was allocated on heap doesn't leak.
3. Passing thread ID to each thread through shared space.

### References

- [1] Intel. *Intel MKL Documentation*. URL: <https://software.intel.com/en-us/mkl/documentation/get-started>.
- [2] Wikipedia. *Toeplitz matrix*. URL: [https://en.wikipedia.org/wiki/Toeplitz\\_matrix](https://en.wikipedia.org/wiki/Toeplitz_matrix).
- [3] xianyi/OpenBLAS. *OpenBLAS Wiki*. URL: <https://github.com/xianyi/OpenBLAS/wiki>.