Optimizing Matrix Multiplication

Vasu Jain Ritvik Vij 29/01/19

1 Subtask 2

This subtask involved improving our Linear Algebra operations such as matrix multiplication using pthreads and commercial libraries such as multi-threading, OpenBlas and Intel MKL.

- Using pthreads if we have a N×N Matrix, suppose we use k threads for the matrix multiplication every k rows is applied to the same pthread, a thread contains a row which can be multiplied with the other matrix. The pthread is a much faster implementation than the normal matrix multiplication. Using the toeplitz matrix, the matrix multiplication can be used to implement the convolution of a matrix with a filter.
- Using OpenBlas OpenBlas is a linear algebra library for fast implementation of linear algebra operations, sgemm function was used to implement matrix multiplication and is considerably fast for larger sizes of matrix upto approx 800×800.
- Using Intel MKL MKL linear algebra library is built upon the OpenBlas library but is considerably faster than OpenBlas and Pthreads, approximately 19 times faster for a matrix of size 200×200, the function used is the same sgemm(for operations on float matrices).

A GNU Plot was made to compare the time taken to perform this operation using pthreads, OpenBlas and Intel MKL.

Further usage of convolution in the LeNet architecture for the digit recognition software has been done using our very own pthread implementation.

1.1 Steps to run

1.1.1 Pthreads

./Week2 convolutionMM padsize InputMatrix.txt numrows1 Kernel.txt numrows2

1.1.2 OpenBlas

- \bullet g++ -std=c++11 -o Week
2 Week 2.cpp -I /"C:/OpenBLAS-0.2.20/OpenBLAS-0.2.20"/OpenBLAS/include/ -L/your_path/OpenBLAS/lib -lopenblas -lm -lpthread
- ./Week2 MMOpenBlas padsize InputMatrix.txt numrows1 Kernel.txt numrows2

1.1.3 Intel MKL

- Give Command For Compilation
- cl Week2_MKL.cpp mkl_intel_lp64_dll.lib mkl_intel_thread_dll.lib
- $\bullet \ \, mkl_core_dll\dot{l}ib \ libiomp5md.lib \\$
- Week2_MKL [commands]