

Reporte:Singular Value Decomposition on GPU using CUDA

Oliab Herrera Coria 107863

Contents

El artículo *Singular Value Decomposition on GPU using CUDA* nos habla sobre la adopción de GPU's en el desarrollo y adaptación de métodos para el procesamiento y manejo de datos, en particular habla sobre una adaptación del método de SVD implementada en GPU's a través de CUDA.

Si bien la mayoría de los programas para hacer estadística, como R, cuentan con una función para calcular esta descomposición el problema ocurre cuando tratamos con muchos datos. Las matrices de grandes datos tienden a ser complicadas desde el punto de vista numérico, en particular el tiempo que tardan en procesar y el error que arrojan puede ser un problema a la hora de implementar nuestras soluciones, incluso a nivel no tan bajo ya que se menciona la adaptación de MATLAB a manera que este pueda tomar ventajas de la arquitectura de los GPU's para optimizar procesos.

Pero es gracias a los GPU's que este problema puede ser resuelto, aquí es cuando, a propósito del curso, este artículo ofrece un algoritmo para la implementación de la descomposición en valores singulares utilizando CUDA.

Este algoritmo utiliza el método Golub-Reinsch el cuál consta de dos etapas:

En la primer etapa se hace la Bidiagonalización de la matriz A:

$$A = QBP^T$$

En la segunda etapa se hace la diagonalización de B de modo que :

$$\Sigma = X^T B Y$$

$$U = QX$$

$$V^T = (PY)^T$$

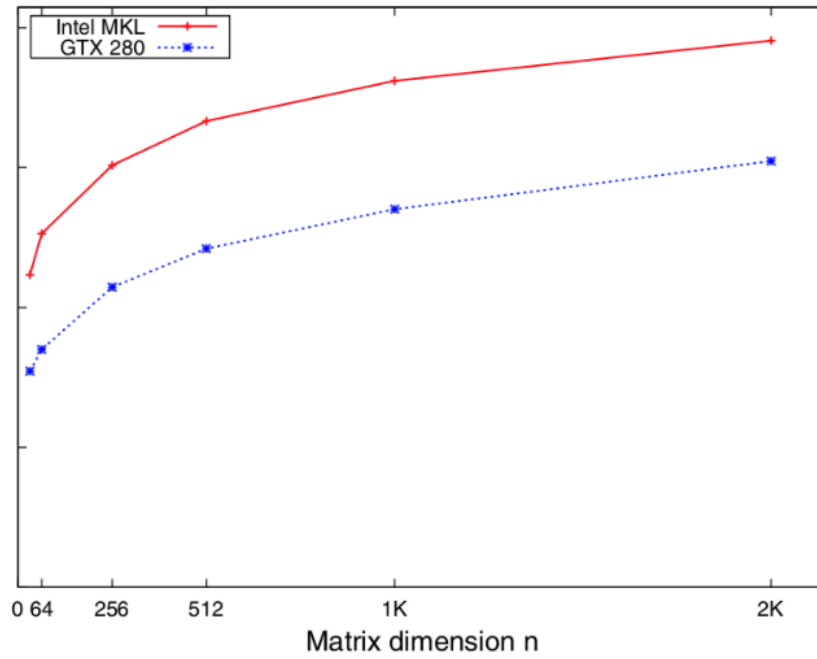
Una matriz bidiagonal es del tipo:

$$Y_k = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Y en este caso se obtiene haciendo transformaciones de householder.

Este proceso se puede paralelizar de manera eficiente en CUDA, dividiendo cada renglón de la matriz en bloques y haciendo que cada thread se ocupe de un elemento.

El autor menciona el trabajo de otros investigadores y los enfoques que utilizaron pero compara el rendimiento de su método contra las implementaciones de Matlab e Intel, podemos observar en la siguiente gráfica que la descomposición que proponen tiene un mejor desempeño que la solución de Intel para matrices no cuadradas:



Para evaluar sus resultados, se escogen distintas muestras de matrices y tamaños para mostrar que, efectivamente, su implementación fue mejor que las ofrecidas comercialmente.