

Proyecto: Multiplicación Matriz-Vector y Matriz-Matriz usando Mapreduce.

W.G & J.S GD ICD

Objetivo:

Ciertas operaciones básicas sobre matrices, como el producto Matriz-vector o Matriz-Matriz, son esenciales en el cálculo de indicadores muy importantes en el análisis de grafos. Con esta asignación se pretende introducir al estudiante en elementos prácticos asociados con el manejo de matrices gigantes mediante el uso del paradigma MapReduce. El cálculo de operaciones Matriz-Vector permite la obtención de vectores como el **PageRank** que permite jerarquizar la importancia de las páginas web y que es la base del análisis de enlaces, la implementación de algunas redes neuronales, *et al.* Operaciones similares permiten hacer cálculos de centralidad importantes en la detección de, por ejemplo, influentes en las redes sociales. El objetivo es, dados dos matrices **A** , **B** y un vector **x**, realizar una función que implemente, con MapReduce, el producto Matriz-Vector cuando el vector **x** por su gran tamaño, no cabe en la memoria RAM. Para esto le haremos llegar una implementación del producto **Ax**, en la que se asume que el vector **x** si cabe en la memoria.

Consideraciones de forma:

- Debe proveer un script que contenga la definición de su función propuesta para la multiplicación Matriz-Vector(**productmv.r**) que tome parámetros:
 - `matriz(string)`: ruta del archivo que contiene la matriz(cuadrada) a multiplicar, por ejemplo: `'.../tblAkv10x10.csv'`.
 - `vector(string)`: ruta del archivo que contiene el vector a multiplicar, por ejemplo: `'.../tblxkv10.csv'`.
 - `N(int)`: dimensión asociada a la multiplicación, por ejemplo: 10.
 - `límite de memoria(int)`: tamaño límite (en bytes) que puede ocupar en la sesión de r donde ejecutará su función, por ejemplo: 480.
- Debe proveer un script que contenga la definición de su función propuesta para la multiplicación Matriz-Matriz(**productmm.r**) que tome parámetros:
 - `matrizA(string)`: ruta del archivo que contiene la matriz(cuadrada) a multiplicar, por ejemplo: `'.../tblAkv10x10.csv'`.
 - `matrizB(string)`: ruta del archivo que contiene la matriz(cuadrada) a multiplicar, por ejemplo: `'.../tblAkv10x10ident.csv'`.
 - `N(int)`: dimension asociada a la multiplicación, por ejemplo: 10.
 - `límite de memoria(int)`: tamaño límite (en bytes) que puede ocupar en la sesión de r donde ejecutará su función, por ejemplo: 480.
- Debe proveer un script que contenga la definición de su función propuesta para la detección de la memoria límite (**memlimit.r**) que tome parámetros:
 - `size(int)`: si es igual a -1 detecte y retorne la memoria límite de la máquina, para cualquier otro valor simplemente lo retorne, por ejemplo: `memlimit(80)` retorna 80.
- Debe proveer un script que contenga la ejemplificación del uso de los scripts previamente mencionados (**test.r**)
- Las funciones deben ser nombradas de igual forma que el archivo que las contiene.

- Los archivos de entrada tendran como estructura: i,j,v donde para cada uno el string que los represente tendrá la misma longitud que el mas largo de ellos, i.e, si $N = 9000$ y el mayor/menor elemento en la matriz ocupa 7 caracteres, la primera entrada sera por ejemplo 0001,0001,0000001 o puede ser: 0001,0001,3.14159 .
- Elabore un breve informe con los resultados de las acciones propuestas, el código fuente y sus comentarios(**informe.pdf**).

Consideraciones de contenido:

- Recuerde que se considerarán las convenciones de estilo que utilice en su código (véase [Google's R Style Guide](#)).

Condiciones de entrega:

- La entrega será realizada mediante el repositorio de la asignación: [Proyecto ICD](#)
- De tener problemas usando con el repositorio de la materia. Cree un repositorio personal con la misma convencionde nombres del repositorio, por favor inmediatamente informe sobre la situación y provea el enlace del repositorio creado(personal) para hacer seguimiento de su trabajo.
- La asignación es grupal de máximo 3 personas. Se promueve la participación y discusión de la misma en un ambiente responsable. Sin embargo, cualquier evidencia de copia será severamente sancionada colocando una nota mínima de cero (0) puntos según lo establecido en la Ley de Universidades. Cualquier proyecto entregado debe ser fruto del trabajo del equipo.
- Fecha de Entrega: **Viernes 9 de abril de 2016** (Hasta este día se aceptarán push's en los repositorios). No se recibirá ninguna tarea por correo electrónico.