

1.- Práctica 10: algoritmo genético

Los algoritmos genéticos trabajan con una población de individuos, cada uno de ellos representa una solución para un problema determinado. En esta práctica se realizó un algoritmo genético (mutaciones y recombinaciones) obteniendo la mejor combinación aleatoria para tener una mejor especie, se comparó la mejor solución alcanzada con la solución óptima obtenida con un método exacto.

2.- Descripción de la tarea

- Paraleliza el algoritmo genético y estudia los efectos en su tiempo de ejecución con pruebas estadísticas y visualizaciones.
- El primer reto es cambiar la selección de padres para reproducción a que use *selección de ruleta*: cada solución se selecciona como padre con una probabilidad que es proporcional a su valor de función objetivo y a su factibilidad, combinando los dos a alguna función que parezca conveniente (recomiendo aprovechar el parámetro `prob en sample`). Estudia si este cambio produce una mejora estadísticamente significativa en la calidad de la solución, incluyendo visualizaciones informativas

3.- Tarea

Utilizando el programa proporcionado por la práctica, se realizó una paralelización para el algoritmo genético de la práctica, además se comparó los tiempos de ejecución obtenidos con esta versión paralela y la versión secuencial. Los cambios que se realizaron para poder paralelizar, fueron los apartados donde se encontraban las funciones: `fun.mutar`, `fun.reproducción`, `val.objetivo` y `fun.factibilidad`, las cuales fueron paralelizadas. Se ejecutaron las versiones secuencial y paralela para comparar los tiempos de ejecución para veinte replicas con cincuenta objetos, treinta generaciones y población inicial de 50, 100, 200 y 400 individuos.

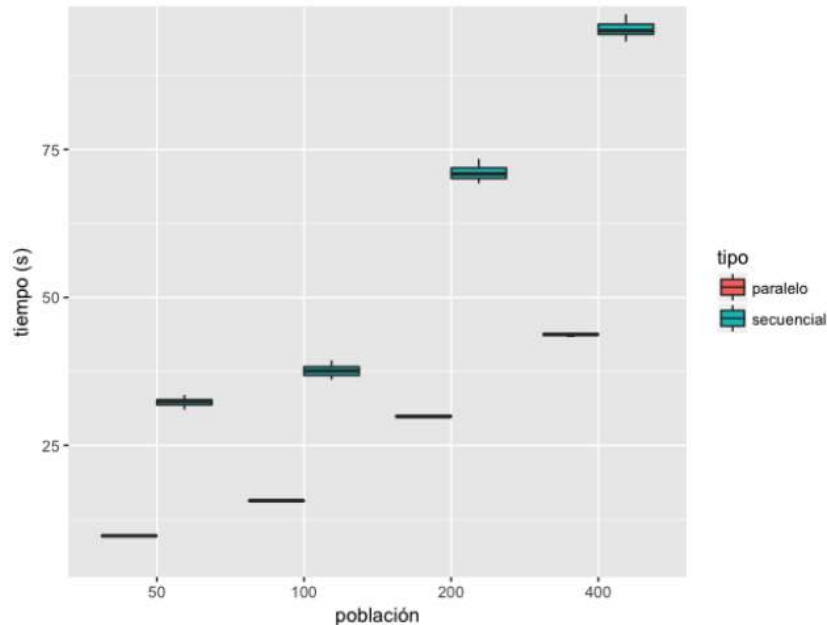


Figura 1: Tiempo de ejecución de veinte replicas con población inicial de 50, 100, 200 y 400 individuos.

En la figura 1 se observa en la programación con paralelo que los tiempos son menores a comparación con los de ejecución normal del programa sin paralelizar, dando como resultado que los tiempos se reducen en un cantidad de tres veces del valor que muestran las secuencias sin paralelizar, arrojando una mayor eficiencia en tiempo.

4.- Reto 1

Tomando como punto de partida para el reto uno se usara el programa paralelizado con modificaciones de la manera en seleccionar los individuos que serán los padres en la reproducción, esto mediante la adición de una probabilidad que será directamente proporcional a su valor genético. En el código original se utiliza la función `sample` para elegir a los padres al azar, pero para cumplir este reto se utilizó una herramienta `prob` de la función. Se determinó la probabilidad en base a sus valores, primero se creó un vector de manera paralela, que contuviera los valores para cada individuo (padres valor), posteriormente se guardó la división de cada valor entre la suma de todos los valores en el vector `pr`. Se agregó una función `cruza-ruleta` para implementar la probabilidad `pr` en la selección de los padres para la cruce. Se graficó mediante la librería `ggplot2`, para esto se guardaron los datos en un `data.frame` para poder llamarlos al hacer las gráficas.

En la figura 2, se puede observar una comparación entre el método de ruleta y el método tradicional al azar de la práctica y lo primero que salta a la vista es el método ruleta se mantiene encima del método al azar en casi todo el tiempo. Aunque la diferencia no es tan pronunciada se puede observar que al aumentar el número de generación se llega en ambos métodos en una estabilidad pero en todo el método ruleta es el más acercado al optimo esperado.

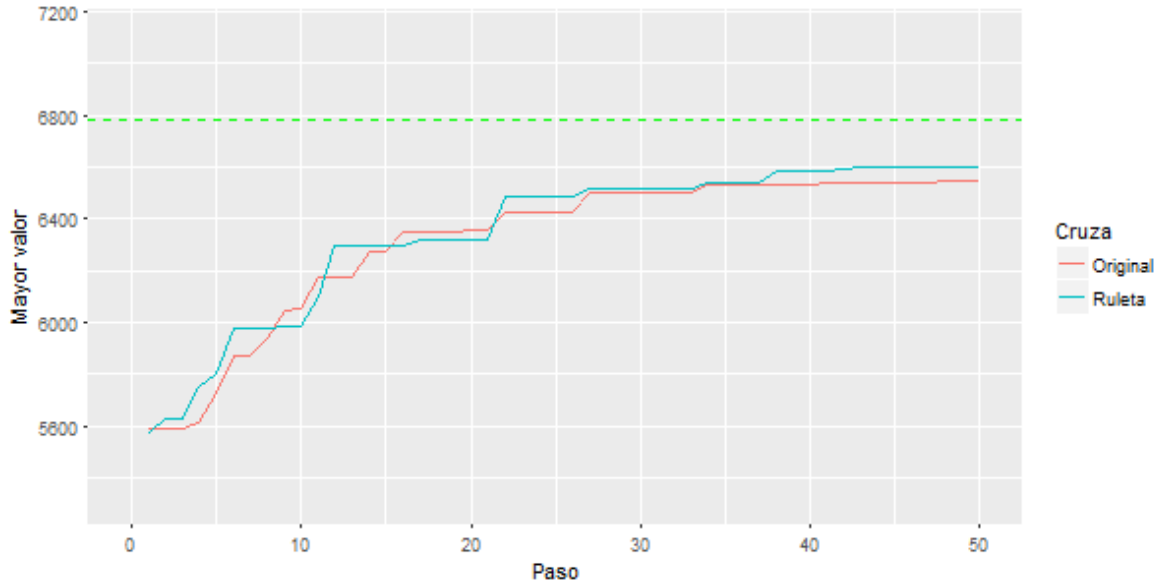


Figura 2. Gráfica de líneas de ambos métodos de selección de padres para la cruce (ruleta y original).

5.- Conclusiones

Se demostró que si es posible paralelizar el algoritmo genético, dando como resultado un menor tiempo de ejecución, por consecuencia de la paralización de las subrutinas muta, cruce y fc. Se pudo observar que los tiempos de ejecución están relacionados con la cantidad de población inicial, es decir, entre más población mayor tiempo. Se modificó el método de selección de los padres mediante la adición de la probabilidad pr en la función cruce-ruleta con la herramienta prob de sample. Los resultados del programa ruleta demostraron valores más acercados al óptimo de generaciones. Otro dato a considerar fue que al momento de tener un número mayor de generaciones la diferencia es muy poco dando como resultado valores muy cercanos entre sí.

6.- Especificaciones de equipo

Modelo del sistema Inspiron 5420, fabricada por Dell Inc, procesador Intel(R) Core(TM) i5-3210M CPU @2.50Hz 2.50 Hz, memoria instalada (RAM) 8 GB (7.86 GB utilizable), tipo de sistema operativo de 64 bits procesador x64, edición de Windows 10 Pro.