

Algoritmos Genéticos Simples

Luis Gerardo Montané Jiménez
Laboratorio Nacional de Informática Avanzada
lmontane@lania.edu.mx

Resumen. Los algoritmos genéticos (GA) están inspirados en la evolución natural de las especies con el fin de resolver problemas complejos de optimización, entre sus elementos se encuentran la representación de los individuos, mutación, recombinación, selección de padres y reemplazo generacional. El primer paso para la utilización de un GA es elegir el tipo y nivel de la representación de soluciones, en el presente trabajo se realiza un análisis de dos algoritmos genéticos, el primero con representación binaria y el segundo con números reales[1].

1 Introducción

A continuación se hace una comparativa entre dos algoritmos genéticos simples, el *primer algoritmo* es con representación binaria para números reales y el *segundo algoritmo* con representación directa con números reales, en ambos casos se incluye elitismo [2]. La función a minimizar es:

$$f(x) = \sum_{i=1}^{10} x^2 i .$$

En la siguiente sección se discuten los resultados que se obtuvieron con ambos algoritmos incluyendo ventajas o desventajas que se presentaron en el transcurso del análisis e implementación.

2 Discusión de Resultados

El desarrollo de los algoritmos fue empleando el lenguaje de programación Java con el entorno de desarrollo NetBeans 6.8. En la implementación, las 10 variables (x) requeridas para el problema fueron tomadas con una precisión de dos decimales en un rango de -10 a 10. Para el algoritmo 1 se utilizó cruce de dos puntos, mutación simple y selección por ruleta. En el algoritmo 2 se empleó cruce aritmética completa, mutación uniforme y selección por torneo binario probabilístico. En la tabla 1 se muestran los parámetros utilizados en los dos algoritmos para realizar las 30 ejecuciones que son mostradas en la tabla 2. La selección de parámetros fue realizada con pruebas que precedieron las 30 corridas, para esto se calibraron los parámetros que permitieron llegar a tener un mejor comportamiento en ambos algoritmos.

Parámetros	Algoritmo 1	Algoritmo 2
Población	200	10
Generaciones	1250	25000
Porcentaje de cruce	0.6	0.45
Porcentaje de mutación	0.3	0.96
Porcentaje de torneo	NO APLICA	0.3
Número de evaluaciones realizadas	250,000	250,000

Tabla 1 - Parámetros para los dos algoritmos genéticos.

Ejecución	Semilla	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	f(x)
1	0.12	0	0	0	0	0	0	0	0	0	-0.01	0.0001
2	0.15	0	0	0	0	0	0	0	0	0	0	0
3	0.22	0	0	0	0	0.01	0	0	0	0	0	0.0001
4	0.23	0	0	0	0	0	0	0	0	0	0	0
5	0.25	0	0	0	0	0	0	0	0	0	0	0
6	0.33	0	0	0	0	0	0	0	0	0	0	0
7	0.36	0	0	0	0	0	0	0	0	0	0	0
8	0.414	0	0	-0.02	0	0	0	0	0	0	0	0.0004
9	0.42	0	0	0	0	0	0	0	0	0	0	0
10	0.43	0	0	0	0	0	0	0	0.01	0	0	0.0001
11	0.44	0	0	0	0	0	0	0	0	0	0	0
12	0.443	0	0	0	0	0	0	0	0	0	0	0
13	0.45	0	0	0	0	0	0	0	0	0	0	0
14	0.46	0	0	0	0	0	0	0	0	0	0	0
15	0.47	0	0	0	0	0	0	0	0	0	0	0
16	0.59	0	0	0	0	0	0	0	0	0	0	0
17	0.65	0	0	0	0	0	0	0	0	0	0	0
18	0.67	0	0	0	0	0	0	0	0	0	0	0
19	0.73	0	0	0.02	0	0	0	0	0	0	0	0.0004
20	0.77	0	0	0	0	0	0	-0.01	0	0	0	0.0001
21	0.78	0	0	0	0	0	0	0	0	0	0	0
22	0.8	0	0	0	0	0	0.02	0	0	0	0	0.0004
23	0.812	0	0	0	0	0	0	0	0	0	0	0
24	0.94	0	0	0	0	0	0	0	0	0	0	0
25	0.9	0	0	0	0	0	0	0	0	0	0	0
26	0.91	0	0	0	0	0	0	0	0	0	0	0
27	0.9143	0	0	0	0	0	0	0	0	0	0	0
28	0.92	0	0	0	0	0	0	0.01	0	0	0	0.0001
29	0.94	0	0	0	0	0	0	0	0	0	0	0
30	0.98	0	0	0	0	0	0	0	0	0	0	0

Tabla 2 – 30 ejecuciones del algoritmo 1.

En la tabla 2 y 3 se observa que el algoritmo 1 con representación binaria obtuvo un mayor número de soluciones en las que las 10 variables del problema llegan a cero, sin embargo, en ambos algoritmos se obtuvieron buenas soluciones, pues aunque en el algoritmo 2 se tenían resultados como 0.0015 estás son consideradas como soluciones válidas pues como ya se mencionó anteriormente la precisión es de dos dígitos. Con respecto a la implementación se concluye que el algoritmo 1 tuvo una complejidad mayor, pues fue necesario realizar más procedimientos para el manejo de los individuos, además de que se utilizaron estructuras de datos más grandes que en el segundo algoritmo; la representación binaria para cada una de las variable del individuo era de 11 bits.

Las estadísticas de la tabla 4 muestran como la desviación estándar del algoritmo 1 es menor que la del segundo, lo que permite ver que los datos se encuentran menos dispersos, no obstante, en ambos casos las diferencias entre estos valores son mínimas, al igual que los valores de la mediana con la media.

Problema y Mejor solución conocida	Técnicas comparadas		
	Estadísticas	Algoritmo 1	Algoritmo 2
Función $f(\vec{x}) = \sum_{i=1}^{10} x_i^2$	Mejor	0	0
	Media	0.000056	0.001156667
	Mediana	0	0.0009
	Peor	0.0004	0.0045
	Desv. Estándar	0.00012229	0.001022061

Tabla 4 –Tabla comparativa de resultados en ambos algoritmos.

Ejecución	Semilla	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	f(x)
1	0.12	0	0	0	0	0	0.02	-0.02	0	0	0	0.0008
2	0.15	-0.01	-0.01	0	-0.03	-0.01	-0.01	-0.01	0	0	0	0.0014
3	0.22	0	0	0	0	0	0.01	0.01	-0.01	0	0	0.0003
4	0.23	0	0	0	0.01	0	0	0	0.02	0	-0.02	0.0009
5	0.25	0	0	-0.01	0	-0.01	0	0.01	0	-0.01	-0.01	0.0005
6	0.33	0.01	0	0	0	0.02	-0.01	0	0	0	0	0.0006
7	0.36	0	-0.02	-0.04	0	0	-0.01	0	0	0	-0.01	0.0022
8	0.414	0	0	0	-0.02	-0.01	-0.02	0	0	0.01	0.01	0.0011
9	0.42	0	0	0	0	0	0	0	0	0	0	0
10	0.43	0	-0.02	0	0.03	0	0	0	0	0	0	0.0013
11	0.44	0	0	0	0	0	0	0	0	0	0	0
12	0.443	0	0.01	0	0	0	0	0	0	0	0.01	0.0002
13	0.45	0.04	0.01	0	0.05	0.01	0	-0.01	0	-0.01	0	0.0045
14	0.46	0	0	0.01	0	0	0	0	0	0.02	0.03	0.0014
15	0.47	0	-0.01	0	0	0.01	0	0	0	0.01	0	0.0003
16	0.59	0	-0.01	-0.02	-0.01	0	0	0	0	0.01	0.01	0.0008
17	0.65	0	0.02	0	0.01	0	0.02	0	0	0	0	0.0009
18	0.67	0	-0.01	0	0	0	0	0.02	0	0.03	0	0.0014
19	0.73	0	0	0	0.01	0.01	0.01	0	0	0.02	0	0.0007
20	0.77	0.01	0	0	-0.02	0	0.03	-0.02	0	-0.02	0	0.0022
21	0.78	0.02	0	0.01	-0.01	-0.01	0	0	0	0	0	0.0007
22	0.8	-0.01	0.01	0	0	0	-0.02	0	0	0.01	0	0.0007
23	0.812	0.01	0	0	0	0.01	-0.01	-0.03	0	0	-0.01	0.0013
24	0.94	0	0.02	0	0	0.02	0	-0.04	-0.01	-0.02	-0.01	0.003
25	0.9	0	0	0	0	0	0	0	0	0	0	0
26	0.91	0	0	0	0	0	0	0	0	0	0	0
27	0.9143	0.02	0	0	0	0.02	0	0	0	0.02	0.01	0.0013
28	0.92	0	0	0	0	0.01	0	0	-0.04	0	0	0.0017
29	0.94	0	0.02	0	0	0.02	0	-0.04	-0.01	-0.02	-0.01	0.003
30	0.98	-0.01	-0.02	0	0	-0.03	0.01	0	0	0	0	0.0015

Tabla 3 – 30 ejecuciones del algoritmo 2.

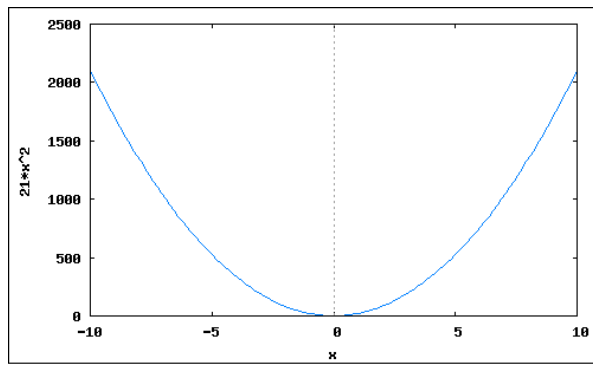


Figura 1 - Función a optimizar $f(\vec{x}) = \sum_{i=1}^{10} x_i^2$

Las gráficas de convergencia de la figura 2 permiten ver que los dos algoritmos no presentan convergencias prematuras, pues tanto el algoritmo 1 y 2 llegan a encontrar una solución en un número distinto de generaciones, pero ambos compiten con el mismo número de evaluaciones, que en este caso es de 250,000.

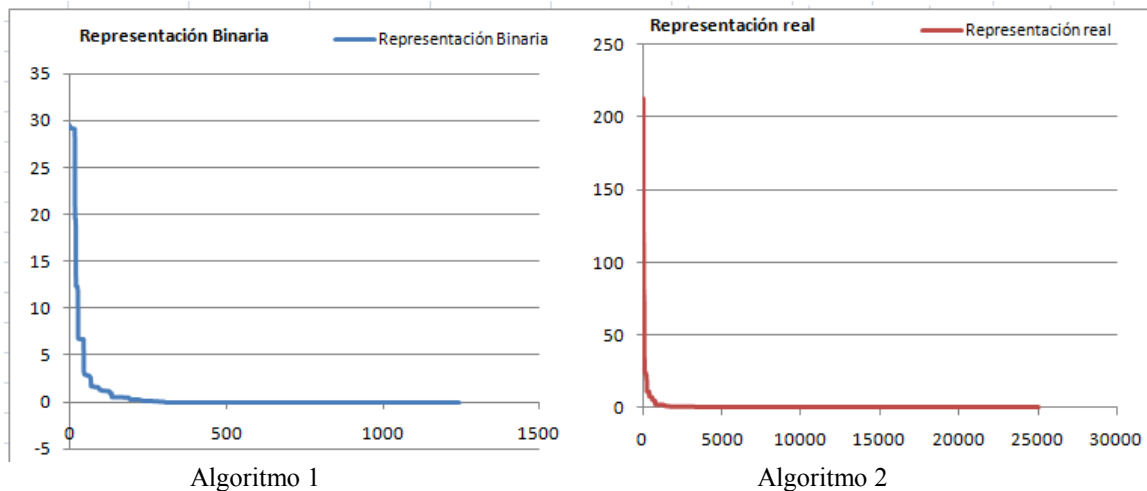


Figura 2 - Gráficas de convergencia del valor $f(x)$ con las generaciones.

3 Conclusiones

En la sección de resultados pudo observarse que los dos algoritmos logran alcanzar una solución válida, el primer algoritmo fue el que presentó más soluciones óptimas; pues en varias corridas las 10 variables de x tenían el valor de 0, pero un punto en contra es que este algoritmo de representación binaria ocupaba mayor poder de procesamiento y recursos del equipo de cómputo, pues cabe mencionar que existe codificación y decodificación de las cadenas de bits, lo cual se conoce como mapeo de genotipo a fenotipo.

Otro punto importante del algoritmo 1 es que para la representación de un individuo eran necesarias vectores de bits de longitud de 110, donde cada 11 bits representaba una única variable. Mientras tanto que en el algoritmo 2 el mapeo de genotipo a fenotipo no existía, y aunque se obtuvieron menos soluciones óptimas donde las 10 variables de x fueran 0, las otras soluciones que presentaba también eran válidas pues se obtenían resultados como $f(x) = 0.0045$ y teniendo en cuenta que la precisión es de dos dígitos esto equivale a 0.

Referencias

- [1] Notas de clase 6 y 7, Curso de Cómputo Inteligente, Maestría en Computación Aplicada, Laboratorio Nacional de Informática Avanzada, 2010.
- [2] Especificación de tarea 4, Curso de Cómputo Inteligente, Maestría en Computación Aplicada, Laboratorio Nacional de Informática Avanzada, 2010.