

MONOGRAFÍA: COMPUTACIÓN EVOLUTIVA

Algoritmos genéticos con reducción de varianza

Pedro Cuesta, Blas Galván, David Greiner y Gabriel Winter

Centro de Aplicaciones Numéricas en Ingeniería (CEANI)
 Universidad de Las Palmas de Gran Canaria (ULPGC)
 Edificio de Ingenierías, Tafira Baja S/N
 Las Palmas de Gran Canaria, España

pedro@titan.ulpgc.es
bgalvan@dma.ulpgc.es
davidg@hptitan.ulpgc.es
gwinter@cicei.ulpgc.es

Resumen

El objetivo del trabajo es incrementar la eficiencia computacional por reducción del tratamiento de la función objetivo introduciendo un patrón en el esquema que se acerque a la solución deseada. Con ello se reduce el cálculo de la función objetivo que en casi todos los casos representa aproximadamente el 80% de computación. El cromosoma patrón es calculado basándose en las experiencias ocurridas en las poblaciones iniciales.

Como caso test de implementación del algoritmo se presentan varios casos académicos

Palabras clave: Algoritmos Genéticos, Selección, Cruce, Patrón

1. Introducción

Los Algoritmos Genéticos simples [1], [2] operan sobre una población finita de tamaño "n" y elementos binarios de longitud fija "l" para cada cromosoma, usando operadores de selección, cruce y mutación.

La población inicial debe incluir individuos con muy alta diversidad cuyos "alelos" son generados aleatoriamente. Un operador selección identifica a los individuos con mejor valor de la función objetivo de la población actual, para que sean padres de la siguiente generación. La función objetivo provee de la retroalimentación necesaria para el proceso de selección.

La eficiencia de los métodos de optimización que utilizan la estrategia anteriormente descrita [5] depende principalmente de:

- 1) La población inicial
- 2) Dominio de las variables

- 3) Operador de selección
- 4) Operador de cruce
- 5) Operador de mutación
- 6) Coste de la función objetivo

Determinados procesos de optimización de tipo real, sobre todo los concernientes a problemas propuestos por la industria [3][4][6], son demasiado lentos y de escasa precisión, fundamentalmente debido a:

- 1) Excesivo tiempo de cálculo de la función objetivo.
- 2) Excesiva varianza en los individuos de las poblaciones

Esto conduce a realizar las siguientes preguntas:

- 1) ¿ Es posible reducir el tiempo de cálculo de la función objetivo sin pérdida de precisión en los cálculos ?

- 2) ¿ Es posible reducir la varianza de los individuos de las poblaciones sin pérdida de la robustez de los GAs ?

En este trabajo se pretende dar respuesta a estas interrogantes mediante el uso de un cromosoma 'patrón' que contemple el conocimiento que genera el propio método en su dinámica de búsqueda del óptimo.

2. Estrategia Elista de Selección

Una de las estrategias más conocida en el proceso de selección es la *elista* basada en copiar los mejores individuos de cada generación a la siguiente, donde algunos padres son admitidos para la creación de nuevos individuos de la generación, con lo que se disminuye la posibilidad de que los mejores individuos de una generación puedan fallar al producir descendencia en la nueva generación.

Una selección lineal se basa en la clasificación ordenada. Así, después de un reordenamiento de la población según los valores de la función objetivo y siendo s^* los mejores individuos, la probabilidad de selección es un valor constante que se obtiene de la fórmula:

$$P_i(s^*) = \frac{1}{\lambda} \left(\eta_{\max} - (\eta_{\max} - \eta_{\min}) \frac{i-1}{\lambda-1} \right)$$

Donde λ es el número de individuos que representan a una generación, en algoritmos genéticos estándar $\lambda=n$ y

$$\eta_{\min} = 2 - \eta_{\max}, \quad (1 \leq \eta_{\max} \leq 2)$$

El operador cruce selecciona aleatoriamente una pareja de individuos entre aquellos previamente seleccionados para recombinar su material genético y producir su descendencia.

Es posible mejorar la estrategia de cruce de individuos, eliminando la posibilidad de destrucción de buenos esquemas que debieran ser transmitidos a la siguiente generación, como veremos en el siguiente apartado.

3. Utilización de un Patrón Fijo

Es un hecho conocido que las sucesivas poblaciones de un proceso de optimización mediante algoritmo genético almacenan el conocimiento generado por el proceso sobre la función que optimiza. Es posible extraer parte de ese

conocimiento mediante la creación de un cromosoma patrón utilizando estimadores estadísticos, y usar dicho cromosoma para mejorar la estrategia de cruce, reduciendo la varianza de los individuos y disminuyendo el número de evaluaciones de la función objetivo.

Sea una generación 't' obtenida mediante un algoritmo genético simple, y definamos la siguiente distribución de probabilidades:

$$p_i^t = P(\text{Valor en posición } i \text{ del patrón} = 1) = \frac{\text{Numero de '1' en la posición } i \text{ de cada individuo}}{\text{Numero de individuos}}$$

$$p_i^t = P(\text{Valor en la posición } i \text{ del patrón} = 0) = 1 - p_i^t$$

Lo que corresponde con una variable aleatoria Benouilli para cada posición, para la cual el éxito consistirá en obtener un '1' y el fracaso obtener un '0' en la posición respectiva.

La probabilidad de cruce uniforme más fácilmente definible es:

$p_c^t = 0$ si la posición es fija en el patrón

$p_c^t = 1$ si la posición no es fija en el patrón

$p_c^t = 1 - p_i^t$ si $p_i^t > p$

$p_c^t = p_i^t$ si $1 - p_i^t > p$

$p_c^t = p_c$ en otros casos

4. Formación del Patrón Fijo

En el siguiente esquema se representa la obtención del patrón fijo.

Tabla 1. Cromosomas de una generación utilizados para la obtención de un patrón. Probabilidades de los bits. Patrón Fijo

1	0	0	1	1	0
1	0	1	0	0	1
1	1	0	1	1	1
1	0	1	0	1	0
0	0	0	1	0	0
Probabilidad de la existencia de un '1'					
4/5	1/5	2/5	3/5	3/5	2/5

Patrón $p = 4/5$ (# = '0' o '1')

1	0	#	#	#	#
---	---	---	---	---	---

5. Cruce Utilizando el Patrón Fijo

Son varias las estrategias que se pueden seguir utilizando el patrón fijo:

- Elegir el punto para cruzar de manera aleatoria que no coincida con una posición de un bit fijo del patrón.
- Una vez cruzados dos individuos de forma estándar hacer iguales los bits, al menos en uno de los descendientes, con los bits fijos del patrón.
- Los puntos correspondientes a posiciones de bits fijos del patrón tendrán una baja probabilidad de ser seleccionados como punto de cruce de los progenitores.

Se muestra un ejemplo de cruce utilizando la segunda estrategia en la tabla siguiente:

Tabla 2. Cruce de un punto utilizando el patrón fijo

Patrón $p = 4/5$ (# = '0' o '1')

1	0	#	#	#	#
---	---	---	---	---	---

Padres seleccionados

0	0	0	1	0	1
1	1	1	1	0	0

Descendientes antes de imponer patrón. Cruce $n = 4$

0	0	0	1	0	0
1	1	1	1	0	1

Imponer el patrón al segundo descendiente

0	0	0	1	0	0
1	0	1	1	0	1

La utilización de este patrón debe hacerse, según se ha mencionado, de forma tal que no perjudique la necesaria diversidad del proceso. Para ello se deben alternar según una secuencia determinada ejecuciones usando algoritmos genéticos con y sin cruce mejorado con el patrón. Distinguiremos dos fases: una inicial en la que se ejecuta un algoritmo genético normal durante 'n' generaciones, al final de esta se seleccionan los mejores individuos y se confecciona el patrón a partir de l cálculo de las probabilidades de cada bit del cromosoma como ya se ha visto. A partir de entonces se desencadena la segunda fase en la cual de cada 'M' ejecuciones se aplicará el cruce

mejorado con el patrón en 'k' de ellas (ejemplo: una de cada tres veces). Esta estrategia de aplicación se puede condensar indicándola como:

[n, k/M]

Dependiendo de la magnitud del problema 'n' deberá ser de distinto valor. En las aplicaciones comentadas al final como casos test, ha sido suficiente con la estrategia

[1, 2/3]

para mejorar considerablemente la eficiencia del proceso.



Fig. 1. Diagrama de la estrategia [n,2/3]

6. Utilización de un Patrón Móvil

El patrón fijo desarrollado en el apartado anterior contiene información estadísticamente 'aceptable' como cierta en cualquier posición (locus) de los individuos. Sin embargo dada la forma de codificar en binario los individuos, es claro que al principio de la optimización es más relevante la información 'aprendida' en las primeras posiciones del patrón (ajuste de las cifras más significativas) para situarse en un entorno cercano al óptimo buscado, mientras que al final es más relevante la información 'aprendida' en las últimas posiciones del patrón (ajuste de las cifras menos significativas) para localizar el valor óptimo con precisión suficiente. En consecuencia, al principio de la optimización la información 'aprendida' en las últimas posiciones del patrón es irrelevante (poco influyente para buscar el óptimo) y viceversa.

Por otra parte el valor 'aprendido' por el patrón para una posición determinada es mejor si se cumplen (en el óptimo buscado) los valores anteriores 'aprendidos' en el mismo. Esto es, en términos probabilísticos, las probabilidades que se obtengan para una posición del patrón están condicionadas a las probabilidades de las posiciones anteriores. Esto será rigurosamente cierto si se conocen todos los valores de los elementos situados a la izquierda de uno dado, y perderá certeza a medida que se desconozca un mayor número de estos.

Los dos razonamientos anteriores permiten definir un patrón que denominaremos 'MOVIL', cuyas características fundamentales serán:

- Utiliza un grupo de posiciones (esquema) de tamaño menor que el total de posiciones de los individuos, el cual se va desplazando desde las primeras posiciones a las últimas conforme va avanzando el proceso de optimización.
- Las probabilidades de las posiciones que utilice se calcularán condicionadas a las probabilidades de las posiciones anteriores.

Las características de este patrón móvil, empleando estrategias para su aplicación como las mencionadas en el apartado anterior, para no perder la diversidad necesaria en los GAs, permite abordar problemas de alta complejidad con gran ganancia de eficiencia en el proceso.

Consideremos el mismo ejemplo del apartado anterior:

Tabla 3. Cromosomas elegidos para la obtención del patrón móvil

1	0	0	1	1	0
1	0	1	0	0	1
1	1	0	1	1	1
1	0	1	0	1	0
0	0	0	1	0	0

Las probabilidades condicionadas para elegir el patrón serían:

Tabla 4. Árbol de probabilidades

		P(1) = 0
	P(1) = 1/4	P(0) = 1
P(1) = 4/5		P(1) = 2/3
	P(0) = 3/4	P(0) = 1/3

El patrón elegido según el árbol de probabilidades anterior es:

Tabla 5. Patrón móvil obtenido a partir del árbol de probabilidades

Patrón $p = 4/5$ (# = '0' o '1')					
i	#	#	#	#	#

7. Sustitución de la Función Objetivo

Si la información contenida en el patrón después de un determinado número de generaciones es cierta, con un determinado nivel de confianza, esto significa que la solución buscada contendrá dicha información. Por tanto para un nivel de confianza alto y después de ejecutarse un número estadísticamente representativo de generaciones, es posible sustituir durante un número determinado de generaciones, la función objetivo por el patrón, disminuyendo considerablemente el coste computacional, sobre todo en problemas reales y complejos propuestos por la industria.

El proceso continúa su búsqueda del óptimo alternando la función objetivo real con el patrón hasta que los valores de la función objetivo coinciden en un porcentaje alto con el patrón.

8. Casos Test

En las figuras 2 y 3 se puede apreciar los resultados obtenidos mediante el uso de patrones comparados con los de estrategias elitistas conocidas, para funciones ampliamente usadas en la validación de códigos de algoritmos genéticos en casos en que la búsqueda del óptimo es muy compleja.

El tamaño de población en los dos casos representados en las figuras adjuntas es de 90 individuos por generación, cada individuo de 30 bits por variable. El eje de abscisas representa la función objetivo y el de ordenadas el número de generaciones.

Es de destacar la gran ganancia de eficiencia obtenida con el empleo del patrón. En las figuras se aprecia que aproximadamente esta es superior al 50%, llegándose a un óptimo con la mitad de generaciones e incluso menos.

Con el empleo de patrón móvil la ganancia es reflejada en tiempo de computación, dependiendo esta de la complejidad del problema tratado.

9. Conclusiones

En el presente trabajo se han analizado las causas fundamentales de pérdida de eficiencia en los métodos de optimización mediante Algoritmos Genéticos en el caso de funciones complejas, y se ha estudiado el incremento de eficiencia que produce el uso de individuos 'patrón' que reflejan lo 'aprendido' por la propia dinámica del proceso de optimización generación tras generación. La forma de caracterizar el conocimiento ha sido mediante la definición de funciones de probabilidad que confirman valores para ciertas posiciones (locus) con cierto nivel de confianza. Este tipo de aprendizaje resulta ser independiente del problema que se este resolviendo y puede aplicarse sin afectar negativamente la necesaria diversidad de los procesos, y por tanto usarse libremente ya que, en el peor de los casos, la eficiencia del proceso original no se verá grandemente afectada. En los casos estudiados se ha logrado una reducción considerable del número de generaciones necesarios para obtener una solución dada.

Estos procedimientos se están aplicando a la optimización de la demanda y distribución de energía eléctrica con minimización del coste que supone la introducción de una nueva subestación de Medio Voltaje en la red teniendo en cuenta la pérdida de carga y el coste de la propia subestación, en colaboración con la compañía estatal UNELCC.

Referencias

- [1] Addison-Wesley, (1989) Goldberg, D., Genetic Algorithm in Search, Optimization, and Machine Learning, Addison-Wesley, Reading, Mass. (1989).
- [2] Winter G, Periaux J, Galán M, & Cuesta P. Editors, Genetic Algorithms in Engineering and Computer Science, Proc. EUROGEN95, John, Wiley & Sons, (1995).
- [3] Winter G, Galan, M. Cuesta P., Greiner D., "Genetic Algorithms: A Stochastic in provement technique. Tools, Skills, Pitfalls and examples", presented in EUROGEN95, pp. 217-245 (Las Palmas de Gran Canaria, España, Dec. 1995)
- [4] Quagliarella D., Periaux J, Poloni C. & Winter G. Editors, Genetic Algorithms and Evolution Strategies in Engineering and Computer Science, Proc. EUROGEN97, John, Wiley & Sons, (1997).
- [5] Winter G, Cuesta P., Galvan B., Greiner D. and Sanchez L, "Some proposals to improve GAs efficiency introducing Variance Reduction", presented in EUROGEN97 (Trieste, Italy, Dec. 1997), submitted for publication to John Wiley & sons in second edition.
- [6] Kargupta, H. "Blacbox and non-blacbox optimization: A common perspective. Algorithms for large scale systems. State of the art and applications in science and engineering. Submitted for publication (1998).

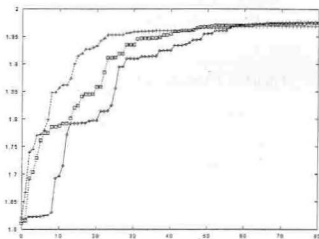


Fig. 2. Función $\frac{\text{sen } x}{x} + e^{-x}$ $[0, 100]$

cruce uniforme entre código binario y código gray
 cruce con patrón y código binario (+)
 cruce con patrón y código gray ■

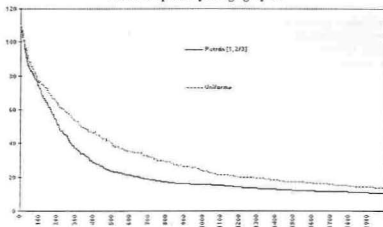


Fig. 3. $f_9 = \sum (x^i - 10 \cos(2\pi x) + 10)$