

Algoritmo de Evolución Diferencial con restricciones

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

Resumen. En este artículo se presenta un estudio acerca de un algoritmo que emplea la estrategia de búsqueda de Evolución Diferencial (ED). Los problemas a resolver son de optimización y cuentan con un conjunto de restricciones de igualdades o desigualdades. ED es una estrategia de búsqueda poblacional propuesta por Storn & Price [1995], entre sus diferencias con respecto a los algoritmos genéticos están los operadores de cruce y mutación, la ED ha sido aplicada ampliamente para resolver problemas artificiales o del mundo real, obteniendo resultados que muestran un desempeño superior a los mostrados por otros AEs [1].

1 Introducción

En 1995, Price y Storn propusieron un algoritmo llamado Evolución Diferencial (ED), este algoritmo forma parte de los algoritmos evolutivos, por lo que utiliza parámetros de población, generaciones y operadores de cruce y mutación.

La Evolución Diferencial es una estrategia de búsqueda poblacional que ha demostrado tener un desempeño muy competitivo en problemas de optimización numérica. ED ha sido aplicada ampliamente para resolver problemas artificiales y del mundo real.

Características de la evolución diferencial:

- El operador de mutación consiste en una simple diferencia aritmética entre pares de vectores.
- Cada vector target (padre) generará solo un vector trial (hijo).
- El vector trial es mejor que el target, lo reemplazará para la siguiente generación (determinista).

Para el manejo de restricciones se emplearon las siguientes reglas de factibilidad [3]:

1. Entre 2 soluciones factibles, la solución con un mejor valor de la función objetivo gana.
2. Si una solución es factible y la otra es no factible, la solución factible gana.
3. Si ambas soluciones son no factibles, aquella con el valor más bajo en la suma de la violación de restricción gana.

En este trabajo se propone realizar un estudio empírico del algoritmo de Evolución Diferencial aplicándolo a resolver problemas con restricciones. El documento se encuentra organizado de la siguiente manera, en la sección 2 se menciona todo lo que ha motivado la realización de este trabajo, incluyendo trabajos relacionados. En la sección 3 se describe a detalle la propuesta con los problemas que se pretenden resolver con Evolución Diferencial y restricciones. En la sección 4 se discuten los resultados obtenidos por el algoritmo ED en cada problema. En el Anexo I se incluyen ejecuciones con otras instancias.

2 Trabajos relacionados

Debido a que en el mundo real los problemas de optimización tienen distintos tipos de restricciones en donde su espacio de búsqueda es modificado, resulta conveniente poder evaluar el desempeño del algoritmo de Evolución Diferencial en situaciones de este tipo. Otro punto importante es que surge la necesidad de medir el desempeño del algoritmo ED con respecto funciones que tienen restricciones de igualdad y que son presentadas en el reporte técnico "Problem Definitions and Evaluation Criteria for the CEC 2010 Competition on Constrained Real-Parameter Optimization" [2], actualmente no existe información acerca de los resultados óptimos para las función presentadas en este reporte por lo que en las pruebas que se realizaron sólo están tomando en consideración el ingreso a las zonas factibles.

En este trabajo se realizaron evaluaciones de desempeño utilizando un conjunto de ejecuciones por cada problema, asimismo se analizó el comportamiento del algoritmo en cada uno de los problemas que se presentan en la sección 3.

3 Propuesta

A continuación se pretende resolver un conjunto de problemas de optimización con restricciones, estos problemas son de minimización. En la tabla 1 se muestran las funciones a resolver:

Problema	E	I	Region de factibilidad	
			p	
			10 variables	30 variables
C01 [0,10]	0	2	0.997689	1.000000
C02 [-5,12, 5.12]	1	2	0.0000000	0.0000000
C05 [-600,600]	2	0	0.0000000	0.0000000

Tabla 1 - Resumen de las características de los 3 problemas, donde p es el radio estimado de la región factible con respecto al espacio de búsqueda. E son las restricciones de desigualdades, I son las restricciones de igualdad.

C01:

$$\text{Min } f(x) = - \frac{\left| \sum_{i=1}^D \cos^4(z_i) - 2 \prod_{i=1}^D \cos^2(z_i) \right|}{\sqrt{\sum_{i=1}^D i z_i^2}} \quad z = x - o$$

$$g_1(x) = 0.75 - \prod_{i=1}^D z_i \leq 0$$

$$g_2(x) = \sum_{i=1}^D z_i - 7.5D \leq 0$$

$$x \in [0, 10]^D$$

C02:

$$\text{Min } f(x) = \max(z) \quad z = x - o, y = z - 0.5$$

$$g_1(x) = 10 - \frac{1}{D} \sum_{i=1}^D [z_i^2 - 10 \cos(2\pi z_i) + 10] \leq 0$$

$$g_2(x) = \frac{1}{D} \sum_{i=1}^D [z_i^2 - 10 \cos(2\pi z_i) + 10] - 15 \leq 0$$

$$h(x) = \frac{1}{D} \sum_{i=1}^D [y_i^2 - 10 \cos(2\pi y_i) + 10] - 20 = 0$$

$$x \in [-5.12, 5.12]^D$$

C05

$$\text{Min } f(x) = \max(z) \quad z = x - o$$

$$h_1(x) = \frac{1}{D} \sum_{i=1}^D (-z_i \sin(\sqrt{|z_i|})) = 0$$

$$h_2(x) = \frac{1}{D} \sum_{i=1}^D (-z_i \cos(0.5\sqrt{|z_i|})) = 0$$

$$x \in [-600, 600]^D$$

Las restricciones de igualdad se transformaron en desigualdades de la siguiente manera [2]:

$$|h_j(X)| - \varepsilon \leq 0, \text{ for } j = p+1, \dots, m$$

Donde una solución factible de X es sí:

$$g_i(X) \leq 0, \text{ for } i = 1, \dots, p \text{ and } |h_j(X)| - \varepsilon \leq 0$$

En el algoritmo implementado $\varepsilon = 0.0001$.

Criterios de evaluación:

- Número de problemas: 3.
- Número de corridas por cada función: 30.
- Máximo de evaluaciones con 10 variables: 200,000.
- Máximo de evaluaciones con 30 variables: 600,000.

4 Resultados y discusión

La implementación del algoritmo con Evolución Diferencial está en el lenguaje de programación Java 1.6 y con el entorno de desarrollo NetBeans 6.9.

La selección de parámetros se hizo con una serie de pruebas que precedieron a las corridas mostradas a continuación, las pruebas permitieron elegir los parámetros donde el algoritmo obtenía mejores resultados, en la tabla 1 se muestran los parámetros iniciales para cada instancia.

Función	Población	Max. Gen (10 variables).	Max. Gen. (30 variables).	CR (Prob. Cruza)	F (Factor escalar)
C01	20	10000	30000	0.5	0.51
C02	20	10000	30000	0.5	0.51
C05	20	10000	30000	0.872	0.3

Tabla 1 – Parámetros de calibración para cada función.

De las medidas de desempeño consideradas para medir el algoritmo está la desviación estándar, media, mediana, la mejor y peor solución encontrada, además de la probabilidad de factibilidad (FP). FP se obtiene con la siguiente fórmula:

$$FP = f/t, \text{ donde } f \text{ número corridas factibles y } t \text{ es el total de corridas}$$

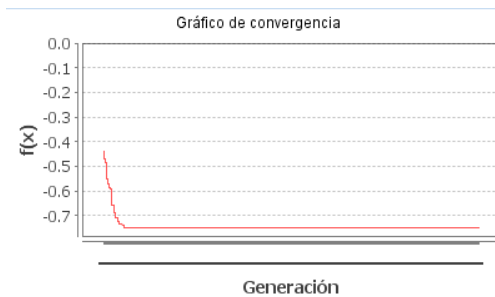
Función	NV	Desv. Est.	Mejor	Peor	Mediana	Media	Probabilidad de factibilidad
C01	10	0.02998272	-0.7473103	-0.6457599	-0.73026156	-0.74731016	1
C01	30	0.043571502	-0.82188314	-0.6056949	-0.78392947	-0.7686661	1
C02	10	1.0573767	-0.15187693	4.559508	2.9955602	2.90618	0.93
C02	30	0.6435007	2.170576	5.099135	3.673405	3.62685	0.96
C05	10	106.358864	161.32149	550.99567	483.90485	441.76514	0.86
C05	30	73.478165	294.34326	579.6369	483.90485	465.58215	0.86

Tabla 2 - Resultados del algoritmo ED (NV = Número de variables).

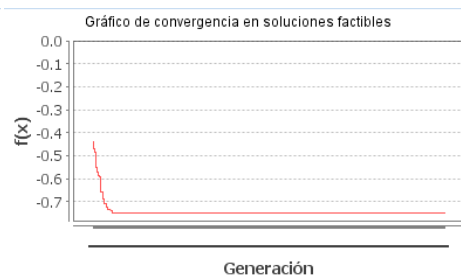
Las medidas de desempeño y estadísticas se presentan en la tabla 2, las filas remarcadas representan los mejores resultados, como puede verse en la medida de desempeño de factibilidad de probabilidad, las seis instancias lograron obtener resultados aceptables, pues el valor más bajo que se obtuvo es de .86 e incluso en algunas instancias se logro alcanzar el óptimo que es 1 de probabilidad. Las desviaciones estándar son bajas para las primeras cuatro instancias de la tabla 2, sin embargo, en las últimas dos resultó ser mucho más alta, por lo que los datos que se obtuvieron estaban más dispersos que en los otros casos, algo no deseado en estos problemas. Y al igual que la desviación estándar, la diferencia que hay de la mediana con la media es baja en las primeras 4 instancias, mientras que en las últimas dos estos valores no se encuentran tan cercanos.

A continuación se presentan las graficas de convergencia con respecto a la mediana de cada instancia ejecutada. A la izquierda se ve la grafica representativa de toda la corrida y a la derecha la grafica a partir de cuando el algoritmo entra a la región factible.

C01 – 10 variables

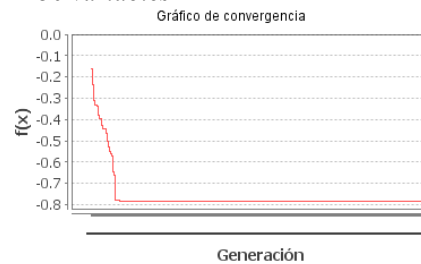


Gráfica de convergencia de toda la corrida.

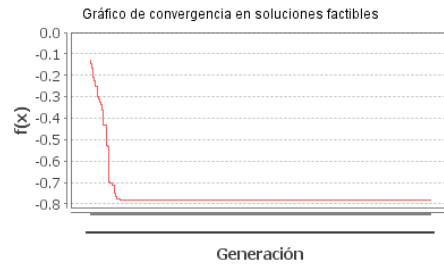


Gráfica de convergencia a partir de la región factible.

C01 - 30 variables

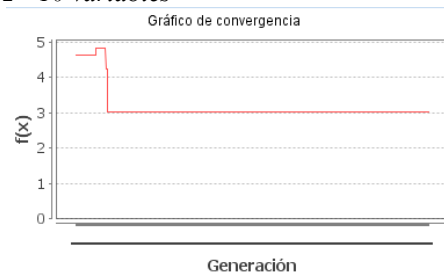


Gráfica de convergencia de toda la corrida.

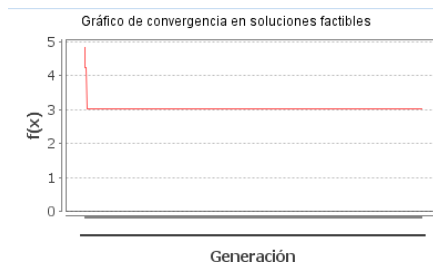


Gráfica de convergencia a partir de la región factible.

C02 - 10 variables

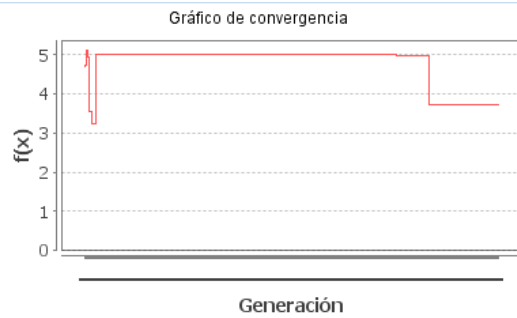


Gráfica de convergencia de toda la corrida.

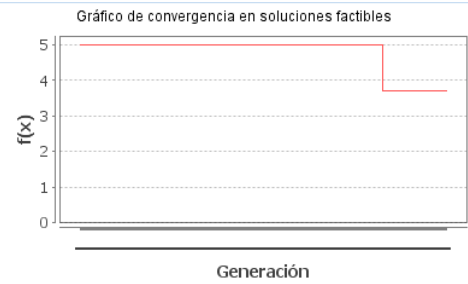


Gráfica de convergencia a partir de la región factible.

C02 - 30 variables

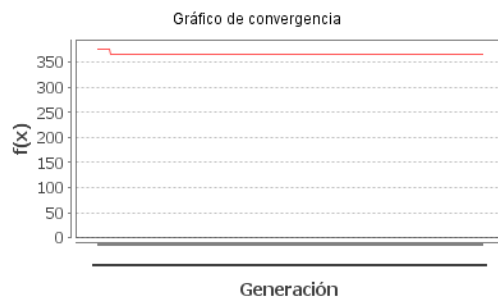


Gráfica de convergencia de toda la corrida.

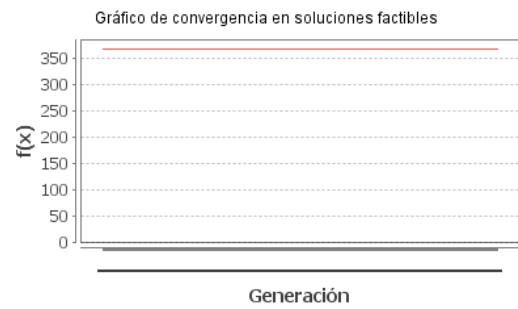


Gráfica de convergencia a partir de la región factible.

C05 - 10 variables

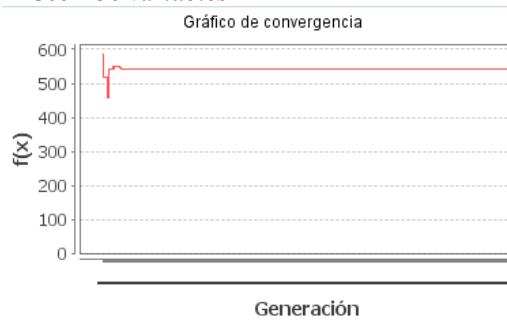


Gráfica de convergencia de toda la corrida.

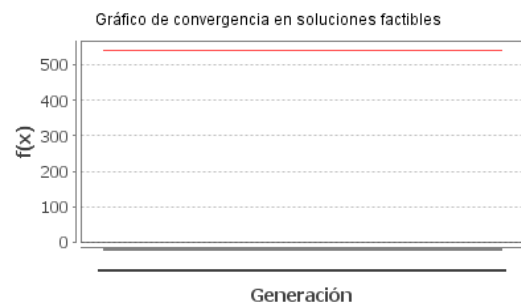


Gráfica de convergencia a partir de la región factible.

C05 - 30 variables



Gráfica de convergencia de toda la corrida.



Gráfica de convergencia a partir de la región factible.

En las gráficas anteriores se pudo apreciar el comportamiento de la mediana con respecto a la zona factible, en C05 con 10 variables y C05 con 30 variables al entrar a la zona factible se quedó constante hasta terminar con las evaluaciones correspondientes, caso contrario en las demás gráficas pues se observa que al entrar a la zona factible sigue bajando sustancialmente al pasar las generaciones.

5 Conclusiones y trabajos futuros

En general los algoritmos evolutivos se han convertido en una alternativa que resulta viable para resolver problemas complejos de optimización de búsqueda.

Como se mencionó anteriormente la Evolución Diferencial es un algoritmo evolutivo reciente y existen versiones que han sido aplicadas con éxito en distintos problemas de optimización con restricciones reales o artificiales.

Entre los trabajos futuros está el aplicar el algoritmo con otras funciones de prueba que tienen restricciones H y que se encuentran en el reporte "Problem Definitions and Evaluation Criteria for the CEC 2010 Competition on Constrained Real-Parameter Optimization" [2], además de realizar la calibración del algoritmo utilizando herramientas como Spot.

Otro aspecto importante que resulta factible de realizar en el futuro son pruebas con problemas del mundo real y adaptar el algoritmo para que pueda ser ejecutado en un Grid (p.e. Java Parallel Processing Framework (JPPF)).

Anexo I

A continuación se muestran otras funciones a minimizar: g01, g02, g03, g04 g05, g06, g07, g08, g09, g10. Estas funciones fueron probadas con el algoritmo ED y en la tabla I.2 es donde se muestran los resultados y las medidas de desempeño obtenidas.

g01:

$$f(\vec{x}) = 5 \sum_{i=1}^4 x_i - 5 \sum_{i=1}^4 x_i^2 - \sum_{i=5}^{13} x_i$$

$$g_1(\vec{x}) = 2x_1 + 2x_2 + x_{10} + x_{11} - 10 \leq 0$$

$$g_2(\vec{x}) = 2x_1 + 2x_3 + x_{10} + x_{12} - 10 \leq 0$$

$$g_3(\vec{x}) = 2x_2 + 2x_3 + x_{11} + x_{12} - 10 \leq 0$$

$$g_4(\vec{x}) = -8x_1 + x_{10} \leq 0$$

$$g_5(\vec{x}) = -8x_2 + x_{11} \leq 0$$

$$g_6(\vec{x}) = -8x_3 + x_{12} \leq 0$$

$$g_7(\vec{x}) = -2x_4 - x_5 + x_{10} \leq 0$$

$$g_8(\vec{x}) = -2x_6 - x_7 + x_{11} \leq 0$$

$$g_9(\vec{x}) = -2x_8 - x_9 + x_{12} \leq 0$$

$$0 \leq x_i \leq 1 \ (i = 1, \dots, 9), \ 0 \leq x_i \leq 100 \ (i = 10, 11, 12) \text{ and } 0 \leq x_{13} \leq 1$$

g02:

$$f(\vec{x}) = - \left| \frac{\sum_{i=1}^n \cos^4(x_i) - 2 \prod_{i=1}^n \cos^2(x_i)}{\sqrt{\sum_{i=1}^n i x_i^2}} \right|$$

$$g_1(\vec{x}) = 0.75 - \prod_{i=1}^n x_i \leq 0$$

$$g_2(\vec{x}) = \sum_{i=1}^n x_i - 7.5n \leq 0$$

$$n = 20 \text{ and } 0 < x_i \leq 10 \ (i = 1, \dots, n)$$

g03:

$$f(\vec{x}) = -(\sqrt{n})^n \prod_{i=1}^n x_i$$

$$h_1(\vec{x}) = \sum_{i=1}^n x_i^2 - 1 = 0$$

$$n = 10 \text{ and } 0 \leq x_i \leq 1 \ (i = 1, \dots, n)$$

g04:

$$f(\vec{x}) = 5.3578547x_3^2 + 0.8356891x_1x_5 + 37.293239x_1 - 40792.141$$

$$\begin{aligned} g_1(\vec{x}) &= 85.334407 + 0.0056858x_2x_5 + 0.0006262x_1x_4 - 0.0022053x_3x_5 - 92 \leq 0 \\ g_2(\vec{x}) &= -85.334407 - 0.0056858x_2x_5 - 0.0006262x_1x_4 + 0.0022053x_3x_5 \leq 0 \\ g_3(\vec{x}) &= 80.51249 + 0.0071317x_2x_5 + 0.0029955x_1x_2 + 0.0021813x_3^2 - 110 \leq 0 \\ g_4(\vec{x}) &= -80.51249 - 0.0071317x_2x_5 - 0.0029955x_1x_2 - 0.0021813x_3^2 + 90 \leq 0 \\ g_5(\vec{x}) &= 9.300961 + 0.0047026x_3x_5 + 0.0012547x_1x_3 + 0.0019085x_3x_4 - 25 \leq 0 \\ g_6(\vec{x}) &= -9.300961 - 0.0047026x_3x_5 - 0.0012547x_1x_3 - 0.0019085x_3x_4 + 20 \leq 0 \end{aligned}$$

$$78 \leq x_1 \leq 102, 33 \leq x_2 \leq 45 \text{ and } 27 \leq x_i \leq 45 \ (i = 3, 4, 5)$$

g05:

$$f(\vec{x}) = 3x_1 + 0.000001x_1^3 + 2x_2 + (0.000002/3)x_2^3$$

$$\begin{aligned} g_1(\vec{x}) &= -x_4 + x_3 - 0.55 \leq 0 \\ g_2(\vec{x}) &= -x_3 + x_4 - 0.55 \leq 0 \\ h_3(\vec{x}) &= 1000 \sin(-x_3 - 0.25) + 1000 \sin(-x_4 - 0.25) + 894.8 - x_1 = 0 \\ h_4(\vec{x}) &= 1000 \sin(x_3 - 0.25) + 1000 \sin(x_3 - x_4 - 0.25) + 894.8 - x_2 = 0 \\ h_5(\vec{x}) &= 1000 \sin(x_4 - 0.25) + 1000 \sin(x_4 - x_3 - 0.25) + 1294.8 = 0 \end{aligned}$$

$$0 \leq x_1 \leq 1200, 0 \leq x_2 \leq 1200, -0.55 \leq x_3 \leq 0.55 \text{ and } -0.55 \leq x_4 \leq 0.55$$

g06

$$f(\vec{x}) = (x_1 - 10)^3 + (x_2 - 20)^3$$

$$\begin{aligned} g_1(\vec{x}) &= -(x_1 - 5)^2 - (x_2 - 5)^2 + 100 \leq 0 \\ g_2(\vec{x}) &= (x_1 - 6)^2 + (x_2 - 5)^2 - 82.81 \leq 0 \end{aligned}$$

$$13 \leq x_1 \leq 100 \text{ and } 0 \leq x_2 \leq 100$$

g07:

$$\begin{aligned} f(\vec{x}) &= x_1^2 + x_2^2 + x_1x_2 - 14x_1 - 16x_2 + (x_3 - 10)^2 + 4(x_4 - 5)^2 + (x_5 - 3)^2 \\ &\quad + 2(x_6 - 1)^2 + 5x_7^2 + 7(x_8 - 11)^2 + 2(x_9 - 10)^2 + (x_{10} - 7)^2 + 45 \end{aligned}$$

$$\begin{aligned} g_1(\vec{x}) &= -105 + 4x_1 + 5x_2 - 3x_7 + 9x_8 \leq 0 \\ g_2(\vec{x}) &= 10x_1 - 8x_2 - 17x_7 + 2x_8 \leq 0 \\ g_3(\vec{x}) &= -8x_1 + 2x_2 + 5x_9 - 2x_{10} - 12 \leq 0 \\ g_4(\vec{x}) &= 3(x_1 - 2)^2 + 4(x_2 - 3)^2 + 2x_3^2 - 7x_4 - 120 \leq 0 \\ g_5(\vec{x}) &= 5x_1^2 + 8x_2 + (x_3 - 6)^2 - 2x_4 - 40 \leq 0 \\ g_6(\vec{x}) &= x_1^2 + 2(x_2 - 2)^2 - 2x_1x_2 + 14x_5 - 6x_6 \leq 0 \\ g_7(\vec{x}) &= 0.5(x_1 - 8)^2 + 2(x_2 - 4)^2 + 3x_5^2 - x_6 - 30 \leq 0 \\ g_8(\vec{x}) &= -3x_1 + 6x_2 + 12(x_9 - 8)^2 - 7x_{10} \leq 0 \end{aligned}$$

$$-10 \leq x_i \leq 10 \ (i = 1, \dots, 10)$$

g08:

$$f(\vec{x}) = -\frac{\sin^3(2\pi x_1) \sin(2\pi x_2)}{x_1^3(x_1 + x_2)}$$

$$g_1(\vec{x}) = x_1^2 - x_2 + 1 \leq 0$$

$$g_2(\vec{x}) = 1 - x_1 + (x_2 - 4)^2 \leq 0$$

$$0 \leq x_1 \leq 10 \text{ and } 0 \leq x_2 \leq 10$$

g09:

$$f(\vec{x}) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 \\ + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7$$

$$g_1(\vec{x}) = -127 + 2x_1^2 + 3x_2^4 + x_3 + 4x_4^2 + 5x_5 \leq 0$$

$$g_2(\vec{x}) = -282 + 7x_1 + 3x_2 + 10x_3^2 + x_4 - x_5 \leq 0$$

$$g_3(\vec{x}) = -196 + 23x_1 + x_2^2 + 6x_6^2 - 8x_7 \leq 0$$

$$g_4(\vec{x}) = 4x_1^2 + x_2^2 - 3x_1x_2 + 2x_3^2 + 5x_6 - 11x_7 \leq 0$$

$$-10 \leq x_i \leq 10 \text{ for } (i = 1, \dots, 7)$$

g10:

$$f(\vec{x}) = x_1 + x_2 + x_3$$

$$g_1(\vec{x}) = -1 + 0.0025(x_4 + x_6) \leq 0$$

$$g_2(\vec{x}) = -1 + 0.0025(x_5 + x_7 - x_4) \leq 0$$

$$g_3(\vec{x}) = -1 + 0.01(x_8 - x_5) \leq 0$$

$$g_4(\vec{x}) = -x_1x_6 + 833.33252x_4 + 100x_1 - 83333.333 \leq 0$$

$$g_5(\vec{x}) = -x_2x_7 + 1250x_5 + x_2x_4 - 1250x_4 \leq 0$$

$$g_6(\vec{x}) = -x_3x_8 + 1250000 + x_3x_5 - 2500x_5 \leq 0$$

$$100 \leq x_1 \leq 10000, 1000 \leq x_i \leq 10000 \text{ } (i = 2, 3) \text{ and } 10 \leq x_i \leq 1000 \text{ } (i = 4, \dots, 8).$$

P	n	F	p	DL	DN	IL	IN	a
g01	13	Cuadrática	0.0111%	9	0	0	0	6
g02	20	No lineal	99.9971%	0	2	0	0	1
g03	10	Polinomial	0.0000%	0	0	0	1	1
g04	5	Cuadrática	52.1230%	0	0	0	0	2
g05	4	Cúbica	0.0000%	2	0	0	3	3
g06	2	Cúbica	0.0066%	0	2	0	0	2
g07	10	Cuadrática	0.0003%	3	5	0	0	6
g08	2	No lineal	0.8560%	0	2	0	0	0
g09	7	Polinomial	0.5121%	0	4	0	0	2
g10	8	Lineal	0.0010%	3	3	0	0	6

Tabla I.1 – Resumen de las características de los 10 problemas, donde p es el radio estimado de la región factible con respecto al espacio de búsqueda. DL son las desigualdades lineales, DN desigualdades no lineales, IL igualdad lineal, IN igualdad no lineal, a las restricciones activas.

Función	NV	Desv.Est.	Mejor	Peor	Mediana	Media	FP
G01	13	0.36514837	-15	-13	-15	-14.933333	1
G02	20	0.032286994	-0.8036165	-0.6779491	-0.7889347	-0.7762515	1
G03	10	0.18065295	-0.9422657	-0.06577042	-0.20437562	-0.24188143	1
G04	5	0.0062609348	-30665.54	-30665.537	-30665.54	-30665.545	1
G05	4	214.55083	5126.545	5877.986	5172.7075	5277.239	.9
G06	2	0.0011646972	-6961.8135	-6961.811	-6961.8135	-6961.813	1
G07	10	0.024852501	24.308716	24.442837	24.313028	24.319887	1
G08	2	0.0000000151559	-0.09582504	-0.09582504	-0.09582504	-0.095825024	1
G09	7	0.0010059155	680.6303	680.6346	680.6315	680.63184	1
G10	8	15.875416	7049.3877	7108.391	7054.45	7063.983	1

Tabla I.2 – Resultados de desempeño (NV = Número de variables).

Como se puede ver en la tabla I.2 la factibilidad de probabilidad resultó bastante buena, el valor más bajo que se obtuvo es de .9. Además de que los mejores valores remarcados llegan a los valores óptimos conocidos si se utiliza una tolerancia de error.

En la figura I.1 se puede ver la interfaz de usuario que permitió ejecutar las corridas de cada instancia, en esta aplicación existe una serie de pestañas como la de configuración, la cual hace posible elegir el número de corridas a realizar junto con un conjunto de semillas, además de otras dos pestañas que permiten elegir y configurar los parámetros a la función que se pretende ejecutar. Las gráficas que se muestran son la del comportamiento de toda la generación en la corrida y la del comportamiento en la región factible por corrida, al finalizar las ejecuciones existen otras dos pestañas que muestran resultados y datos estadísticos.

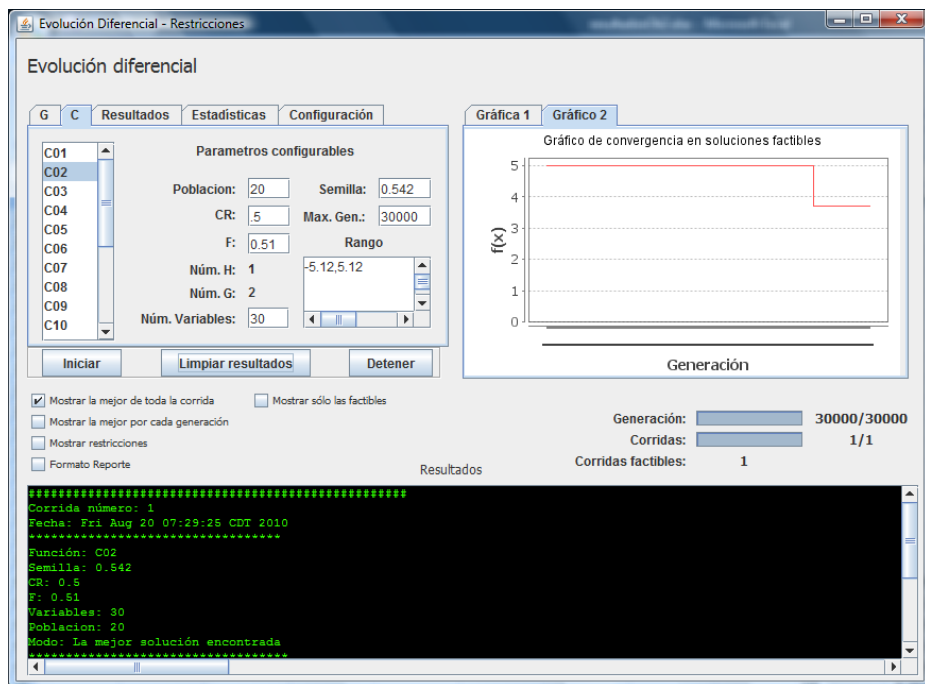


Figura I.1 – Pantalla de la herramienta que permite la ejecución del ED.

Referencias

- [1] Notas de clase 10, Curso de Cómputo Inteligente, Maestría en Computación Aplicada, Laboratorio Nacional de Informática Avanzada, 2010.
- [2] Problem Definitions and Evaluation Criteria for the CEC 2010 Competition on Constrained Real-Parameter Optimization", *Technical Report*, Nanyang Technological University, Singapore, 2010.
- [3] Global and Local Selection in Differential Evolution for Constrained Numerical Optimization, Efrén Mezura-Montes and Carlos A. Monterrosa-López, LANIA, 2009.