Escuela Técnica Superior de Ingenieros Aeronáuticos

Universidad Politécnica de Madrid

TESIS DOCTORAL

**UTILIZACIÓN DE LOS ALGORITMOS GENÉTICOS
PARA LA RESOLUCIÓN NUMÉRICA DE LOS
PROBLEMAS DE OPTIMIZACIÓN EN
MECÁNICA DEL VUELO**

Miguel Ángel Gómez Tierno

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS AERONÁUTICOS
UNIVERSIDAD POLITÉCNICA DE MADRID

TESIS DOCTORAL

"UTILIZACIÓN DE LOS ALGORITMOS GENÉTICOS PARA
LA RESOLUCIÓN NUMÉRICA DE LOS PROBLEMAS DE
OPTIMIZACIÓN EN MECÁNICA DEL VUELO"

por: Miguel Ángel Gómez Tierno

Director de la Tesis: Prof. Dr. Juan J. Martínez García

Madrid, 11 de Septiembre de 1984

RESUMEN

Esta es una tesis interdisciplinaria.

Se pretende resolver numéricamente y mediante evaluaciones de la función objetivo, un problema de optimización de Mecánica del Vuelo matemático, sin ligaduras de igualdad, paramétrico, continuo y determinístico.

Para ello se emplean los Algoritmos Genéticos, que son métodos probabilísticos adaptativos, basados en la imitación de aquellas estrategias naturales que han permitido y permiten a los seres vivos adaptarse casi totalmente a su entorno.

Por tratarse de métodos probabilísticos, los Algoritmos Genéticos presentan ventajas respecto a los métodos determinísticos cuando la dimensión del espacio en estudio es alta y cuando la función objetivo es muy complicada.

Por tanto, para resolver un problema típico de una Ciencia Aplicada (la Mecánica del Vuelo), se utilizan técnicas de una Ciencia Pura (la Matemática) o, más concretamente, de uno de sus campos (la Teoría de la Optimización), basadas en la imitación de hechos cuyo estudio corresponde a una Ciencia Natural (la Biología) y, en particular, a una de sus ramas (la Genética).

De esta forma se subsanan en parte los problemas de la incomunicación entre las ciencias y de la excesiva especialización del científico.

A mis padres y a mi hermano.

Y a Valdeavellano de Tera.

Agradecimientos

La presente Tesis ha sido desarrollada durante los dos últimos años merced a un contrato sobre el tema, firmado por la Cátedra de Mecánica del Vuelo, con el Centro de Operaciones de la Agencia Espacial Europea (E.S.O.C.)

Vaya mi agradecimiento, en general, a todos los que de una u otra forma, han hecho posible esta Tesis.

En particular, y en especial, agradezco al Dr. G. Janin de E.S.O.C. y al Dr. J.J. Martínez-García, Catedrático de Mecánica del Vuelo, sus valiosísimas directrices y sugerencias y su inestimable estímulo y apoyo.

También deseo expresar mi agradecimiento a todas las personas relacionadas con la Cátedra de Mecánica del Vuelo. Concretamente, al profesor C. Bachmaier, por sus útiles indicaciones y comentarios; al profesor T. Prieto, por sus desvelos y ayuda con los programas de ordenador; al colaborador M. Belló por la misma razón anterior; al delineante F. Huertas, por sus impecables figuras y dibujos y al mecanógrafo D. Franco por el excelente acabado final de esta Tesis.

Madrid, Septiembre de 1984

ÍNDICE

1. INTRODUCCIÓN	1
2. OPTIMIZACIÓN. CLASIFICACIÓN DE LOS PROBLEMAS DE OPTIMI- ZACIÓN	3
3. EL PROBLEMA DE OPTIMIZACIÓN ESTUDIADO	7
4. MÉTODOS DE RESOLUCIÓN DEL PROBLEMA DE OPTIMIZACIÓN ESTUDIADO	9
4.1 Introducción	9
4.2 Métodos determinísticos	11
4.2.1 Generalidades	11
4.2.2 Métodos determinísticos globales	12
4.2.2.1 Método de Hooke y Jeeves	14
4.2.3 Métodos determinísticos globales	18
4.3 Métodos probabilísticos	20
4.3.1 Generalidades	20
4.3.2 Categoría 1 ó Métodos de Montecarlo	24
4.3.2.1 Búsqueda Aleatoria Pura	25
4.3.2.2 Aproximación Chichinadze	28
4.3.3 Categoría 2 ó Métodos de Rastreo Aleatorio	32
4.3.3.1 Métodos de Rastreo Aleatorio Discretos	33
4.3.3.2 Método de Bremmerman	37
4.3.3.3 Método Adaptativo de Matyas	39
4.3.3.4 Método de Beltrami e Indusi	42

4.3.4	Categoría 3 ó Métodos Mixtos	46
4.3.4.1	Método Multipartida	47
4.3.4.2	Método de Hartman S2	48
4.3.4.3	Método de las Constelaciones de Törn ..	50
4.3.4.4	Método Estocástico-Determinístico de Fagioli	56
4.3.4.5	Método de Gaviano	57
5.	LOS ALGORITMOS GENÉTICOS	59
5.1	Introducción	59
5.2	El Modelo Simplificado de la Evolución	62
5.3	Los Algoritmos (μ, λ) y $(\mu + \lambda)$	65
5.3.1	Generalidades	65
5.3.2	Desarrollo de los Algoritmos (μ, λ) y $(\mu + \lambda)$	66
5.3.3	El Valor Esperado del Tamaño de Paso	69
5.3.4	Los Modelos de Función Objetivo Utilizados	72
5.4	Los Algoritmos $(1, \lambda)$ y $(1 + \lambda)$	76
5.4.1	Generalidades	76
5.4.2	Características Teóricas Estudiadas	77
5.4.3	Probabilidad de Éxito y Velocidad de Convergen- cia para el Algoritmo $(1 + 1)$	80
5.4.3.1	Modelo de Corredor	81
5.4.3.2	Modelo Esférico	83
5.4.4	Probabilidad de Éxito y Velocidad de Convergen- cia para los Algoritmos $(1, \lambda)$ y $(1 + \lambda)$	85
5.4.4.1	Modelo de Corredor	86
5.4.4.2	Modelo Esférico	92
5.4.5	Tamaño de Paso Óptimo y Control del Tamaño de Paso	97

5.4.5.1	Tamaño de Paso Óptimo y Control del Tamaño de Paso para el Algoritmo $(1+1)$..	98
5.4.5.2	Tamaño de Paso Óptimo y Control del Tamaño de Paso para los Algoritmos $(1,\lambda)$ y $(1+\lambda)$	101
5.4.6	Conclusiones	108
5.5.	Mejora de los Algoritmos Genéticos	109
6.	DISEÑO DE PROGRAMAS PARA LOS ALGORITMOS GENÉTICOS $(1,\lambda)$ Y $(1+\lambda)$	113
7.	PRUEBA Y AJUSTE DE LOS PROGRAMAS DISEÑADOS	116
7.1	Introducción	116
7.2	Funciones objetivo de prueba	117
7.2.1	Generalidades	117
7.2.2	Funciones Objetivo de Prueba para Problemas de Optimización sin Ligaduras	118
7.2.3	Funciones Objetivo de Prueba para Problemas de Optimización con Ligaduras	120
7.2.4	Funciones Objetivo de Prueba para Problemas Globales de Optimización	122
7.2.5	Funciones Objetivo de Prueba que Aparecen en Problemas Técnicos Típicos	124
7.2.6	Funciones Objetivo de Prueba con Analogía Biológica	126
7.3	Resultados obtenidos	132
8.	APLICACIÓN DE LOS ALGORITMOS GENÉTICOS PARA LA RESOLU- CIÓN DE UN PROBLEMA DE OPTIMIZACIÓN EN MECÁNICA DEL VUELO	135

8.1	Introducción	135
8.2	El problema estudiado. Resolución teórica	136
8.3	Resolución de un caso concreto mediante los Algor- ritmos Genéticos	141
9.	CONCLUSIONES	143
APÉNDICE I: LA REVOLUCIÓN DARWINISTA. EL EVOLUCIONISMO		145
APÉNDICE II: VOCABULARIO DE GENÉTICA BÁSICA		150
REFERENCIAS Y BIBLIOGRAFÍA		155

*"Una mirada a los libros,
dos a la vida"*

J.W. Goethe

1. INTRODUCCIÓN

Esta Tesis es interdisciplinaria.

Se pretende resolver numéricamente un problema típico de optimización en Mecánica del Vuelo, como es la maximización del alcance de un misil cuyas masas y velocidades iniciales y finales están fijadas. La formulación del problema y su resolución teórica que, afortunadamente, existe en el caso estudiado, se recogen en el Capítulo 8.

Para resolver numéricamente el problema, se recuerda en primer lugar qué se entiende por optimización y cuál es la clasificación de los problemas de optimización (Capítulo 2).

En segundo lugar, se sitúa el problema de optimización de Mecánica del Vuelo, cuya resolución se pretende, dentro de la clasificación mencionada (Capítulo 3). Además se establece la restricción adicional de que sólo se utilizarán evaluaciones de la función objetivo.

Una vez centrado el problema, se pasa revista a los métodos que permiten resolverlo (Capítulo 4). Existen dos grandes grupos: los métodos determinísticos y los métodos probabilísticos. De los métodos determinísticos se hace un somero resumen, ya que se dispone de muy buenos tratados en el área. Se considerarán únicamente aquellas técnicas que no utilicen derivadas. En cuanto a los métodos probabilísticos, el estudio será más detallado. Se propondrá una clasificación de estos métodos totalmente original y se hará un breve resumen de las técnicas más importantes dentro de cada una de las Categorías definidas.

El Capítulo 5 está ya completamente centrado en los Algoritmos Genéticos. La definición que se propone en esta Tesis para

estos algoritmos es la siguiente: son métodos probabilísticos adaptativos basados en la imitación de las estrategias utilizadas en la evolución natural de los seres vivos. En los distintos apartados de este Capítulo se han obtenido conclusiones originales muy interesantes. Entre ellas destacan el método de cálculo y control del tamaño de paso óptimo para estos algoritmos y la idea de que existe un algoritmo mejor que todos los demás y que además es el más sencillo.

Una vez analizados los aspectos más importantes de los Algoritmos Genéticos, se pasa al diseño de programas de ordenador que se basen en las estrategias estudiadas (Capítulo 6) y posteriormente a la prueba y ajuste de los programas diseñados, mediante una serie de funciones objetivo de prueba (Capítulo 7).

Por último, y como ya se ha mencionado, el Capítulo 8 estará dedicado a formular el problema de optimización de Mecánica del Vuelo en estudio, a resolverlo teóricamente y a resolverlo prácticamente en un caso concreto, mediante la aplicación de los programas de ordenador desarrollados.

Se han incluido en esta Tesis dos Apéndices. En el primero de ellos se pasa revista, de forma simplificada, a lo que fué la Revolución Darwinista y a lo que es el Evolucionismo. Y en el segundo, se ofrece un Vocabulario de Genética Básica, asimismo muy simplificado. Con estos Apéndices se pretende resumir los conocimientos biológicos y genéticos necesarios para el perfecto entendimiento de esta Tesis.

2. OPTIMIZACIÓN. CLASIFICACIÓN DE LOS PROBLEMAS DE OPTIMIZACIÓN

La optimización de un objeto o de un sistema cualquiera es el proceso que permite obtener la manera más eficiente de realizar sus objetivos, o, lo que es lo mismo, que permite obtener la mejor opción entre todas las existentes.

Según esto, el proceso completo de optimización consiste en:

- La proposición de todos los conceptos alternativos razonables que conduzcan a los objetivos propuestos.
- La optimización de cada concepto.
- La elección del mejor de los conceptos optimizados.

En esta Tesis se trata exclusivamente de la optimización de un concepto o diseño dado de un sistema, es decir, del punto segundo de la relación anterior.

Ante todo, es necesario disponer de una medida cuantitativa de la efectividad o de la bondad del concepto del sistema estudiado, en función de una serie de variables de éste.

Las características independientes que distinguen una solución de otra se denominan variables o parámetros del sistema u objeto bajo consideración.

La variable dependiente que proporciona la medida de la bondad del sistema se denomina función objetivo. Obsérvese que este nombre de función objetivo se utiliza indistintamente en los problemas de maximización y en los de minimización. Si se trata de obtener un máximo, pueden emplearse otras dos denominaciones: función utilidad y función calidad. Si, por el contra-

rio, el extremo a obtener es un mínimo, el nombre alternativo es función de costo.

A continuación se realiza una clasificación de los distintos problemas de optimización que habitualmente aparecen en la literatura especializada.

Se denomina problema experimental de optimización a aquel problema en el que la relación funcional entre la función objetivo y los parámetros es desconocida. Entonces será necesario realizar una serie de experimentos con el objeto físico propiamente dicho o con un modelo físico adecuado que tenga su mismo comportamiento.

Muchas veces es imposible una investigación exhaustiva de todas las posibles soluciones para un problema de este tipo y, al mismo tiempo, una selección arbitraria de casos puede conducir a resultados erróneos. Por ello, se deben utilizar las llamadas estrategias, que son esquemas de experimentación que tienen en cuenta la información obtenida en experimentos previos con el fin de seleccionar el próximo caso a investigar.

En caso de que sea posible la construcción de un modelo matemático del objeto o sistema considerado, es decir, cuando sea conocida la relación funcional entre la función objetivo y las variables o parámetros, se hablará de un problema matemático de optimización.

En muchos problemas de optimización los valores de los parámetros solución del problema han de satisfacer, además, una serie de restricciones llamadas ligaduras. Estas ligaduras pueden ser de igualdad o de desigualdad, según se utilicen los signos $(=)$ ó $(\geq)$ respectivamente. Los problemas en que aparezcan restricciones se denominan problemas de optimización con liga-

duras, y aquellos en los que no aparezcan problemas de optimización sin ligaduras.

Si la función objetivo y las ligaduras están sujetas a perturbaciones estocásticas, se estará en presencia de un problema estocástico de optimización y en caso contrario de un problema determinístico de optimización.

Se habla de problema estático de optimización o problema paramétrico de optimización cuando la solución del problema no cambia con el tiempo. Si la función objetivo y las ligaduras son funciones lineales de las variables se tiene un problema de programación lineal y en caso contrario un problema de programación no lineal.

Cuando la solución del problema es función del tiempo, la denominación utilizada es problema dinámico de optimización. Este problema se describe normalmente mediante un conjunto de ecuaciones diferenciales y una función objetivo, que puede ser una integral. Matemáticamente estos problemas han sido tratados de forma adecuada desde distintos puntos de vista, como son el cálculo de variaciones y la programación dinámica.

En algunos casos las variables del sistema sólo pueden tomar valores discretos, en cuyo caso se trata de un problema discreto de optimización.

Por último, aquel problema en el que se pretenda determinar un óptimo local se denomina problema local de optimización mientras que si lo que se busca es el óptimo global se hablará de un problema global de optimización. Un óptimo global en un punto, implica que el valor de la función objetivo en ese punto es mejor que en cualquier otro de su dominio de definición. Por

otra parte, un óptimo local en un punto significa que el valor de la función objetivo en ese punto es mejor que en cualquier otro situado en sus cercanías. El óptimo global de una función objetivo puede obtenerse eligiendo el mejor de todos sus óptimos locales, ya que un óptimo global también lo es local.

3. EL PROBLEMA DE OPTIMIZACIÓN ESTUDIADO

Entre todos los posibles problemas de optimización resumidos en el capítulo anterior, esta Tesis únicamente se centrará en aquellos problemas de optimización

matemáticos, sin ligaduras de igualdad, paramétricos, continuos y determinísticos

Se considerarán problemas matemáticos sin ligaduras de igualdad, es decir, se pretende optimizar (maximizar o minimizar) la función objetivo

$$y = f(x) \quad y \in \mathbb{R} \quad . \quad 3/1$$

$$x \in \mathbb{R}^n ; x^T \equiv (x_1, x_2, \dots, x_n)$$

donde

$y \equiv$ valor de la función objetivo

$x \equiv$ vector de variables o parámetros

$n \equiv$ dimensión del espacio

sin ligaduras o, como máximo, con las ligaduras de desigualdad

$$g_j(x) \geq 0 \quad j = 1(1)m \quad 3/2$$

Para adoptar un punto de vista lo mas general posible, se supondrá que la única forma de obtener información del problema es a través de evaluaciones de la función objetivo.

Esta falta de información a priori o situación de "black box" aparece por ejemplo, cuando las funciones tienen una forma analítica conocida pero muy complicada, por lo que es imposible calcular sus derivadas, o cuando las funciones tienen cambios muy bruscos, tanto si son continuas como si no lo son.

Por último, es conveniente aclarar lo siguiente: si una función objetivo $f(x)$ tiene un máximo (o mínimo) en el punto x^1 , la función $-f(x)$ tendrá en ese mismo punto un mínimo (o máximo). Por ello, a lo largo de esta Tesis se hablará de minimización (de igual forma podría hablarse de maximización) y de función objetivo o función de costo.

4. MÉTODOS DE RESOLUCIÓN DEL PROBLEMA DE OPTIMIZACIÓN ESTUDIADO

4.1 INTRODUCCIÓN

En este capítulo se resumirán aquellos métodos de resolución numérica de los problemas de optimización que constituyen la base fundamental de los distintos algoritmos de cálculo existentes.

Ante todo, hay que señalar que un problema de optimización, en general no tiene por qué ser resoluble. Es más, normalmente sólo podrá obtenerse una aproximación del óptimo e incluso a veces ni siquiera eso. Por ejemplo, parece difícil encontrar un método que obtenga el mínimo de la siguiente función:

$$f(x) = \begin{cases} 0, & \forall x \neq x^1 \\ -1, & x = x^1 \end{cases} \quad 4.1/1$$

Los métodos existentes para resolver la optimización de problemas matemáticos sin ligaduras de igualdad, paramétricos, continuos y determinísticos pueden clasificarse en dos grandes categorías:

- Métodos determinísticos
- Métodos probabilísticos

Los métodos determinísticos calculan el mínimo de la función objetivo mediante la aplicación de reglas determinísticas, mientras que los métodos probabilísticos utilizan para ello, al menos parcialmente, reglas probabilísticas.

Es importante resaltar la diferencia existente entre clasificación de problemas de optimización y clasificación de mé-

todos de resolución. Por ejemplo, un problema determinístico de optimización puede resolverse mediante un método determinístico y también mediante uno probabilístico.

A continuación, en el apartado 4.2 se resumen sucintamente los métodos determinísticos y en el apartado 4.3 se estudian, con mayor amplitud, los métodos probabilísticos.

4.2 MÉTODOS DETERMINÍSTICOS

4.2.1 Generalidades

Los métodos determinísticos permiten calcular el mínimo de la función objetivo mediante la aplicación de reglas determinísticas.

Estos métodos pueden clasificarse en

- Métodos locales, y
- Métodos globales

dependiendo de su capacidad para resolver problemas locales de optimización, en el primer caso, y problemas globales, en el segundo.

En este apartado se presenta a su vez una sucinta clasificación de los métodos locales y globales utilizados en la resolución de problemas matemáticos, sin ligaduras de igualdad, paramétricos, continuos, determinísticos y locales o globales respectivamente.

De nuevo se hace notar que sólo se considerarán aquí aquellos métodos que utilicen evaluaciones de la función objetivo como único instrumento para calcular el mínimo.

4.2.2 Métodos determinísticos locales

Los métodos determinísticos locales permiten la resolución de problemas locales de optimización.

Estos métodos pueden dividirse en métodos directos e indirectos. Los métodos directos calculan el mínimo mediante comparación directa de los valores de la función objetivo en dos o más puntos y los indirectos lo hacen a través de una condición necesaria para el mínimo (Ref. 1).

De alguna forma los métodos indirectos utilizan el concepto de derivada de la función objetivo, por lo que no se tendrán en cuenta en esta Tesis.

La tabla 4.2.2/1 muestra la clasificación de los métodos determinísticos locales directos que no utilizan derivadas, empleados en la optimización de problemas matemáticos sin ligaduras de igualdad, paramétricos, continuos, determinísticos y locales.

En el párrafo 4.2.2.1 se esboza uno de estos métodos, conocido como Método de Hooke y Jeeves, el cual ha servido de base a varios métodos probabilísticos.

TABLA 4.2.2/1

MÉTODOS DETERMINÍSTICOS LOCALES DIRECTOS, QUE NO UTILIZAN DERIVADAS, EMPLEADOS EN LA OPTIMIZACIÓN DE PROBLEMAS MATEMÁTICOS, SIN LIGADURAS DE IGUALDAD, PARAMÉTRICOS, CONTINUOS, DETERMINÍSTICOS Y LOCALES

Espacio Unidimensional

- Métodos de Eliminación
 - Búsqueda Lineal Simultánea
 - Búsqueda uniforme
 - Búsqueda con Distribución Asimétrica
 - Búsqueda con Distribución Uniforme de Pares de Puntos
 - Búsqueda Lineal Secuencial
 - Búsqueda Dicotómica
 - Búsqueda Fibonacci
 - Método de la Sección Áurea

Espacio Multidimensional

SIN LIGADURAS

- Métodos de Eliminación
 - Método Simple de la Malla
 - Método Secuencial de las Tangentes
 - Búsqueda Dicotómica Multivariable
- Métodos de Descenso
 - Método Cíclico en las Coordenadas
 - Método de Hooke y Jeeves
 - Método de Rosenbrock
 - Método Simplex de Spendley, Hext y Himsworth

CON LIGADURAS DE DESIGUALDAD

- Método de las Funciones de Penalización y Barrera

4.2.2.1 Método de Hooke y Jeeves

El método de Hooke y Jeeves se basa en la idea de que es conveniente utilizar de nuevo una dirección de búsqueda que haya tenido éxito en etapas anteriores.

Se realizan dos tipos de búsqueda (ver fig. 4.2.2.1/1):
búsqueda exploratoria y búsqueda guiada.

La minimización a lo largo de las distintas direcciones se puede efectuar de forma continua o mediante etapas discretas.

A continuación se presenta un algoritmo de minimización para la versión de etapas discretas (Ref. 2).

Etapa 0. Elegir, como criterio de parada, un escalar ε mayor que cero y suficientemente pequeño. Elegir el tamaño de paso inicial, $\Delta \geq \varepsilon$ y un factor de aceleración $\alpha > 0$.

Elegir también un punto de partida, x^1 . Hacer $y = x^1$ y $k = j = 1$.

Etapa 1. Si $f(y^j + \Delta d^j) - f(y^j) < 0$, donde d^j es la dirección coordenada j -ésima, el intento ha tenido éxito. Hacer $y^{j+1} = y^j + \Delta d^j$ e ir a 2.

De otra forma, el intento ha sido un fracaso.

Si $f(y^j - \Delta d^j) - f(y^j) < 0$, hacer $y^{j+1} = y^j - \Delta d^j$ e ir a 2.

De otra forma, hacer $y^{j+1} = y^j$ e ir a 2.

Etapa 2. Si $j < n$, hacer $j = j+1$ e ir a 1.

De otra forma, si $f(y^{n+1}) - f(x^k) < 0$ ir a 3.

De otra forma, ir a 4.

Etapa 3. Hacer $x^{k+1} = y^{n+1}$ e $y^1 = x^{k+1} + \alpha(x^{k+1} - x^k)$. Hacer

$k = k+1$, $j = 1$ e ir a 1.

Etapa 4. Si $\Delta \leq \epsilon$, parar; x^k es la solución. De otra forma, hacer $\Delta = \Delta/2$, $y^1 = x^k$, $x^{k+1} = x^k$, $k = k+1$, $j = 1$, e ir a 1.

La fig. 4.2.2.1/2 ilustra los puntos generados mediante este método para la función objetivo $f(x) = (x_1-2)^4 + (x_1-2x_2)^2$ y para el punto inicial $(0.0, 3.0)$.

FIGURA 4.2.2.1/1

MÉTODO DE HOOKE Y JEEVES. BÚSQUEDAS EXPLORATORIA Y GUIADA.

FIGURA 4.2.2.1/2

MÉTODO DE HOOKE Y JEEVES CON ETAPAS DISCRETAS. (Los números indican el orden en que se generan los puntos).

4.2.3 Métodos determinísticos globales

Los métodos determinísticos globales pretenden resolver los problemas globales de optimización.

En general, estos métodos encuentran muchas dificultades para obtener el óptimo global, y a veces sólo se puede asegurar la convergencia hacia un óptimo local.

No obstante, han sido desarrollados distintos métodos que no utilizan derivadas y que proporcionan buenos resultados cuando se efectúan algunas simplificaciones en la forma de la función objetivo. Entre los que resuelven problemas sin ligaduras están (Ref. 3, Art. 2) :

- Técnicas de Cobertura Espacial
- Técnicas de las Regiones de Atracción

Las primeras cubren el espacio a estudiar con un conjunto de puntos seleccionados de forma no aleatoria y posteriormente, mediante la comparación de los valores de la función objetivo en esos puntos, obtienen el óptimo global. Entre las Técnicas de Cobertura Espacial destacan tres métodos:

- Método Simple de la Malla
- Método de Evtushenko
- Método de Shubert

Las Técnicas de las Regiones de Atracción se basan en la identificación de un conjunto de condiciones iniciales desde las que es posible calcular óptimos locales mediante algún algoritmo

de optimización local. Una vez que han sido obtenidos todos los óptimos locales, la comparación de los valores correspondientes de la función objetivo proporciona el óptimo global.

4.3 MÉTODOS PROBABILÍSTICOS

4.3.1 Generalidades

Los métodos probabilísticos pueden definirse como aquellos métodos capaces de determinar el mínimo de una función objetivo utilizando para ello, al menos parcialmente, números aleatorios y probabilidades.

La aplicación de los métodos determinísticos a problemas donde la función objetivo es muy complicada o bien depende de muchos parámetros (la dimensión del espacio es alta), presenta grandes dificultades. Asimismo, los métodos determinísticos globales, para resolver problemas globales de optimización, solamente pueden asegurar la convergencia global si la función objetivo satisface exigencias bastante estrictas.

Para resolver este tipo de problemas fueron y están siendo desarrollados los métodos probabilísticos. Estos métodos presentan una serie de ventajas y de desventajas.

Entre las ventajas pueden señalarse:

- Resuelven problemas muy generales con poca información.
- Normalmente se requiere poco esfuerzo para obtener una primera aproximación del óptimo, lo cual es suficiente para muchos problemas prácticos. Sin embargo, puede necesitarse un gran esfuerzo para mejorar esa primera aproximación.

Entre las desventajas más importantes están:

- Base teórica muy limitada. La mayor parte de los métodos han sido desarrollados como algoritmos particulares.
- Difícil tratamiento de las ligaduras. La única forma existente de tenerlas en cuenta consiste en determinar para cada valor de la variable dependiente si se cumplen o no las ligaduras, aceptando o rechazando respectivamente el resultado obtenido.

Los aspectos positivos pueden ser fundamentales para resolver los problemas que aparecen en algunas misiones espaciales típicas, mientras que las desventajas tienen menos importancia y pueden ser subsanadas desarrollando algoritmos especiales que resuelvan el problema específico en consideración.

Realizar un estudio sistemático de los métodos probabilísticos es una tarea de gran dificultad, debido a la diversidad de métodos particulares existentes, a la falta total de sistematización en el área y a las distintas y a veces contradictorias nomenclaturas utilizadas.

No obstante, en esta Tesis se realiza un primer intento de clasificación.

Los métodos probabilísticos hacen uso de dos clases de técnicas distintas:

- Técnicas de Cobertura Espacial
- Técnicas de Rastreo Aleatorio

Una Técnica de Cobertura Espacial es un procedimiento que cubre el espacio a estudiar con puntos seleccionados de forma aleatoria y que después procede a calcular el óptimo mediante

investigación directa de los valores de la función objetivo en todos los puntos seleccionados.

Una Técnica de Rastreo Aleatorio es un procedimiento que a partir de un punto $\mathbf{x}^k \in S^n \subset \mathbb{R}^n$ (o un conjunto de puntos, $\mathbf{x}^k, \mathbf{x}^{k-1}, \dots, \mathbf{x}^1$) define, utilizando números aleatorios y probabilidades, un nuevo punto $\mathbf{x}^{k+1} \in S^n \subset \mathbb{R}^n$ para el que el valor de la función objetivo satisface una condición de desigualdad determinada (por ejemplo, $f(\mathbf{x}^{k+1}) < f(\mathbf{x}^1)$, $\forall 1 \leq k$).

De acuerdo con la definición dada, un método se dice que es probabilístico si incluye elementos estocásticos. Por tanto, este tipo de métodos pueden utilizar técnicas determinísticas en alguna de sus partes.

Tomando en consideración todo lo anteriormente dicho, en esta Tesis se propone la siguiente clasificación para los métodos probabilísticos:

Categoría 1 ó Métodos de Montecarlo: métodos que utilizan únicamente Técnicas de Cobertura Espacial.

Categoría 2 ó métodos de rastreo aleatorio: métodos que utilizan únicamente Técnicas de Rastreo Aleatorio.

Categoría 3 ó métodos mixtos: métodos que utilizan simultáneamente técnicas determinísticas y probabilísticas o diferentes tipos de técnicas probabilísticas.

Cada una de las categorías anteriores pueden subdividirse a su vez en dos grupos:

- Métodos Adaptativos
- Métodos no Adaptativos

Un método se llamará adaptativo cuando en cierta etapa se utiliza la información proporcionada por etapas anteriores con el fin de seleccionar algunos parámetros internos del método y de esta forma guiar la búsqueda. El método será no adaptativo cuando esta guía no se realiza.

Los métodos probabilísticos de optimización, englobados en las Categorías 1, 2 y 3, se estudian a continuación en los subapartados 4.3.2, 4.3.3 y 4.3.4, respectivamente.

4.3.2 Categoría 1 ó métodos de Montecarlo

Los métodos de Montecarlo utilizan únicamente Técnicas de Cobertura Espacial.

La aproximación probabilística a la solución de los problemas de optimización mediante las Técnicas de Cobertura Espacial se basa en el siguiente resultado (Ref. 4, Art. 1):

Se supone que la región de interés de la función objetivo a optimizar es el conjunto finito S^n contenido en $\mathbb{R}^n$, $S^n \subset \mathbb{R}^n$

Si A^n es un subconjunto de S^n con medida m , tal que

$$\frac{m(A^n)}{m(S^n)} \geq \alpha > 0 \quad 4.3.2/1$$

y si $p(A^n, N)$ es la probabilidad de que al menos un punto de una secuencia de N puntos distribuidos uniformemente de forma aleatoria sobre S^n , caiga en A^n , entonces

$$\lim_{N \rightarrow \infty} p(A^n, N) = 1 \quad 4.3.2/2$$

Estas técnicas son especialmente útiles para resolver problemas globales de optimización.

Existen diversos métodos pertenecientes a este grupo. Algunos de ellos han sido seleccionados en esta Tesis y se presentan a continuación.

4.3.2.1 Búsqueda Aleatoria Pura

El método más simple dentro de la Categoría 1 es la búsqueda aleatoria pura. Consta de las siguientes etapas (Ref. 4, Art. 1 y Ref. 5):

Etapla 1. Evaluar $f(x)$ en N puntos elegidos aleatoriamente según una distribución uniforme sobre S^n .

Etapla 2. Mediante la comparación directa de los resultados de la etapa 1, determinar el mínimo global.

En este algoritmo puede introducirse un test de hipótesis. El mínimo global estimado se acepta únicamente si tests adicionales confirman su validez con un cierto nivel de confianza.

La búsqueda aleatoria pura está basada en la siguiente propiedad: puede demostrarse que si $A^n \subset S^n \subset \mathbb{R}^n$ (donde A^n es un subconjunto del dominio de definición de $f(x)$, S^n) de forma que

$$\frac{m(A^n)}{m(S^n)} = \alpha \quad 4.3.2.1/1$$

donde $m(A^n)$ es la medida de A^n , entonces

$$p(A^n, N) = 1 - (1 - \alpha)^N, \text{ y} \quad 4.3.2.1/2$$

$$N = \frac{\ln(1-p)}{\ln(1-\alpha)} \quad 4.3.2.1/3$$

donde $p(A^n, N)$ es la probabilidad de que al menos un punto de los N caiga en A^n .

El número de puntos N requeridos para tener una probabilidad alta de caer en una pequeña región que contenga al óptimo global es muy alto. Por ejemplo, para $\alpha = 0.001$ y $p = 0.9$, se obtiene $N = 2300$.

Esta es la razón por la que se han propuesto distintas variantes para mejorar este método. Entre ellas, se pueden mencionar:

- Realizar optimizaciones locales desde puntos iniciales distintos distribuidos uniformemente. Este tipo de métodos pertenecen a la Categoría 3.
- Dividir el dominio S^n en subdominios en cada uno de los cuales sólo se permite un punto. Con este procedimiento es posible evitar puntos que estén demasiado juntos entre sí.

La figura 4.3.2.1/1 ilustra este método para una función objetivo arbitraria.

FIGURA 4.3.2.1/1

BÚSQUEDA ALEATORIA PURA

4.3.2.2 Aproximación Chichinadze

Este método se basa en la utilización de una función de probabilidad $P(V)$, donde $P(V)$ es la probabilidad de que se satisfaga la condición $f(x) > V$ siempre que x se elija de forma aleatoria y con distribución uniforme sobre el espacio S^n (Ref. 3, Art. 2).

Obsérvese que el mínimo global de la función objetivo (ver fig. 4.3.2.2/1) se produce cuando

$$P(V) = 1 \quad 4.3.2.2/1$$

y el máximo global cuando

$$P(V) = 0 \quad 4.3.2.2/2$$

Obviamente, no se conoce la expresión analítica de la función $P(V)$. Sin embargo, si se calcula la función objetivo $f(x)$ en N puntos aleatoria y uniformemente distribuidos y si se representa el número de puntos en los que $f(x) > V_i$, donde V_i indica el nivel considerado, por $m_i(V_i, N)$, entonces la relación m_i/N es una aproximación de $P(V_i)$ que mejora cuando $N \rightarrow \infty$. Chichinadze demostró que

$$\left[\frac{m_i}{N} - P(V_i) \right] \sim O\left(\frac{1}{N}\right) \quad 4.3.2.2/3$$

Para obtener la solución $P(V) = 1$, él sugiere ajustar las relaciones m_i/N mediante una combinación lineal de polinomios

ortogonales $P_j(V)$. Los coeficientes λ_j se determinan minimizando alguna medida de la distancia existente entre $\sum_j \lambda_j P_j(V)$ y los valores m_i/N en los distintos V_i . La raíz de la ecuación

$$\sum_j \lambda_j P_j(V) = 1 \quad 4.3.2.2/4$$

representada por V^* en la figura 4.3.2.2/2, proporcionará una estimación del mínimo global de $f(x)$.

La aproximación Chichinadze consta, por tanto, de las siguientes etapas:

Etapla 1. Seleccionar N puntos sobre S^n , de forma aleatoria, mediante una distribución uniforme.

Etapla 2. Evaluar $f(x)$ en esos N puntos.

Etapla 3. Fijar los diferentes niveles V_i .

Etapla 4. Calcular el número de puntos, m_i , en los que $f(x) > V_i$ así como las relaciones m_i/N .

Etapla 5. Ajustar las relaciones m_i/N mediante una combinación lineal de polinomios ortogonales, $P_j(V)$, obteniendo los coeficientes λ_j .

Etapla 6. Obtener V^* , raíz de la ecuación

$$\sum_j \lambda_j P_j(V) = 1$$

V^* es una estimación del valor de $f(x)$ en el mínimo global.

La base teórica de esta aproximación se describe de forma más rigurosa en la Ref. 3, Art. 9.

FIGURA 4.3.2.2/1

APROXIMACIÓN CHICHINADZE. NIVELES CONSIDERADOS DE $f(x)$

FIGURA 4.3.2.2/2

APROXIMACIÓN CHICHINADZE. PROBABILIDAD DE SATISFACER LA CONDICIÓN $f(x) > v$ Y AJUSTE MEDIANTE COMBINACIÓN LINEAL DE POLINOMIOS ORTOGONALES.

4.3.3 Categoría 2 ó métodos de rastreo aleatorio

Los métodos de rastreo aleatorio utilizan únicamente Técnicas de Rastreo Aleatorio para resolver el problema de optimización considerado.

En estos métodos los requerimientos para lograr una convergencia local rápida son diametralmente opuestos a los requerimientos para conseguir la convergencia global. La condición más importante para asegurar la convergencia local es la selección adecuada del tamaño de paso (Ref. 6).

Hay muchos métodos dentro de esta Categoría. La diferencia entre ellos se basa, fundamentalmente, en la forma de obtener nuevos puntos a partir de los anteriores (selección del vector aleatorio Δx^i tal que $x^{i+1} = x^i + \Delta x^i$).

Los métodos de rastreo aleatorio más significativos se han extraído de la literatura especializada y se describen a continuación.

4.3.3.1 Métodos de Rastreo Aleatorio Discretos

Los métodos de rastreo aleatorio discretos son algoritmos de rastreo aleatorio con un número finito de direcciones de búsqueda (Ref. 3, Art. 18).

A partir de un punto, $x^i \in S^n$, donde $S^n \subset \mathbb{R}^n$, se obtiene el punto $x^{i+1} \in S^n$ mediante

$$x^{i+1} = x^i + \Delta x^i \quad 4.3.3.1/1$$

donde los incrementos Δx^i tienen longitud unidad y son de la forma:

$$\Delta x^i = \left(\pm \frac{1}{\sqrt{n}}, \pm \frac{1}{\sqrt{n}}, \dots, \pm \frac{1}{\sqrt{n}} \right)^T. \quad 4.3.3.1/2$$

Se consideran, por tanto, 2^n direcciones de búsqueda. Para cada paso, la dirección de búsqueda concreta se elige aleatoriamente con distribución uniforme, entre las 2^n posibles.

Existen dos algoritmos básicos de este tipo:

- Rastreo aleatorio con retorno

Este algoritmo consta de las etapas siguientes:

Etapas 0. Seleccionar $x^0 \in S^n$. Calcular $f(x^0)$. Hacer $i = 0$.

Etapas 1. Elegir aleatoriamente con distribución uniforme, un valor Δx^i entre los 2^n posibles.

Etapas 2. Calcular $x^{i+1} = x^i + \Delta x^i$ y $f(x^{i+1})$.

Etapla 3. Si $x^{i+1} \in S^n$ ir a 4. De otra forma, ir a 1.

Etapla 4. Si $f(x^{i+1}) - f(x^i) < 0$, hacer $i = i+1$ e ir a 1.

De otra forma, hacer $x^{i+1} = x^i$, $i = i+1$ e ir a 1.

Esto significa que si en una etapa se ha tenido éxito, se elige aleatoriamente nueva dirección de búsqueda. Si no, se vuelve al último punto válido.

La figura 4.3.3.1/1 ilustra este método para la función objetivo $f(x) = (x_1-2)^4 + (x_1-2x_2)^2$ y para el punto inicial $(0.0, 2.0)$.

- Descenso aleatorio

Este algoritmo consta de las siguientes etapas:

Etapla 0. Seleccionar $x^0 \in S^n$. Calcular $f(x^0)$. Hacer $i = 0$.

Etapla 1. Elegir aleatoriamente con distribución uniforme, un valor de Δx^i entre los 2^n posibles.

Etapla 2. Calcular $x^{i+1} = x^i + \Delta x^i$ y $f(x^{i+1})$.

Etapla 3. Si $x^{i+1} \in S^n$ ir a 4. De otra forma, ir a 1.

Etapla 4. Si $f(x^{i+1}) - f(x^i) < 0$, hacer $\Delta x^{i+1} = \Delta x^i$, $i = i+1$ e ir a 2. De otra forma, hacer $x^{i+1} = x^i$, $i = i+1$ e ir a 1.

Esto significa que cada vez que la búsqueda tiene éxito se continua investigando en esa misma dirección. En caso de fallo se elige aleatoriamente una nueva dirección de búsqueda.

La figura 4.3.3.1/2 ilustra este método para la función objetivo $f(x) = (x_1-2)^4 + (x_1-2x_2)^2$ y para el punto inicial $(0.0, 2.0)$

FIGURA 4.3.3.1/1

RASTREO ALEATORIO CON RETORNO

FIGURA 4.3.3.1/2

DESCENSO ALEATORIO

4.3.3.2 Método de Bremmerman

El método de Bremmerman consta de las siguientes etapas (Ref. 4, Art. 2 y Ref. 5):

- Etapa 0. Seleccionar $x^0 \in S^n$. Calcular $f(x^0)$. Hacer $i = 0$.
- Etapa 1. Elegir un vector aleatorio, p^i , con una función de densidad de probabilidad normal.
- Etapa 2. Calcular los cuatro valores $x^{i+1} = x^i + tp^i$ y los cuatro $f(x^{i+1})$ para $t = \pm h, \pm 2h$, donde h es un parámetro escalar fijado.
- Etapa 3. Ajustar los cuatro valores de $f(x^{i+1})$ y el valor de $f(x^i)$ mediante un polinomio de cuarto grado, $P(t)$.
- Etapa 4. Determinar el valor de t^{*i} correspondiente al mínimo de $P(t)$, resolviendo la cúbica $P'(t) = 0$.
- Etapa 5. Si $f(x^i + t^{*i}p^i) - f(x^i) < 0$, hacer $x^{i+1} = x^i + t^{*i}p^i$. De otra forma, hacer $x^{i+1} = x^i$.
- Etapa 6. Si se satisface algún criterio de parada, parar. De otra forma, hacer $i = i+1$ e ir a 1.

Este algoritmo es muy similar al método de Gaviano (ver sección 4.3.4.5) pero tiene una eficiencia peor. Esto es debido a que, en lugar de realizar una minimización global a lo largo de una línea, sólo se pseudominimiza basándose en el cálculo del mínimo de un polinomio de cuarto grado. Por esta razón la eficiencia del método depende mucho de la elección del parámetro h .

La figura 4.3.3.2/1 ilustra el método para la función objetivo $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ y para el punto inicial $(0.0, 2.0)$

FIGURA 4.3.3.2/1

MÉTODO DE BREMERMAN

4.3.3.3 Método Adaptativo de Matyas

El método de Matyas consta de las siguientes etapas (Ref. 7, Art. 9):

Etapa 0. Seleccionar $\mathbf{x}^0 \in S^n$ y los escalares $\alpha_s, \alpha_f, r_s, r_f, c_s, c_f$. Calcular $f(\mathbf{x}^0)$. Hacer $d^0 = 0, t^0 = 1, i = 0$ y $f_{\min} = f(\mathbf{x}^0)$.

Etapa 1. Elegir un vector aleatorio, ξ^i , mediante una distribución uniforme y generado con el algoritmo de Brown-Muller.

Etapa 2. Calcular $\Delta \mathbf{x}^i = d^i + t^i \xi^i, \mathbf{x}^{i+1} = \mathbf{x}^i + A \mathbf{x}^i$ y $f(\mathbf{x}^{i+1})$.

Etapa 3. Si $f(\mathbf{x}^{i+1}) - f_{\min} < 0$, ir a 4. De otra forma, ir a 5.

Etapa 4. Hacer $f_{\min} = f(\mathbf{x}^{i+1})$. Calcular $d^{i+1} = r_s d^i + c_s \Delta \mathbf{x}^i$ y $t^{i+1} = \alpha_s t^i$. Hacer $i = i+1$ e ir a 6.

Etapa 5. Hacer $\mathbf{x}^{i+1} = \mathbf{x}^i$. Calcular $d^{i+1} = r_f d^i + c_f \Delta \mathbf{x}^i$ y $t^{i+1} = \alpha_f t^i$. Hacer $i = i+1$ e ir a 1.

Etapa 6. Si se satisface algún criterio de parada, parar. De otra forma, ir a 1.

Cada $A \mathbf{x}^i$ se determina a partir de un elemento aleatorio, $t^i \xi^i$ y de un sesgo, d^i . Los coeficientes r_s, r_f controlan la contribución del sesgo de la etapa anterior al sesgo de la presente. Se les llama coeficientes de rechazo. Los coeficientes c_s y c_f controlan la contribución de la etapa anterior al sesgo de la presente etapa. Son denominados coeficientes de detección. Los coeficientes α_s y α_f incrementan y disminuyen respectivamente el tamaño de paso, t .

De acuerdo con Matyas, estos coeficientes deben cumplir las siguientes desigualdades:

$$r_s \geq 0, r_f \leq 1, c_s \geq 0, c_f \leq 0, r_s + c_s > 1, |r_f + c_f| < 1$$

$$\alpha_s > 1, \alpha_f < 1$$

La figura 4.3.3.3/1 ilustra el método para la función objetivo $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ y para el punto inicial $(0.0, 2.0)$.

FIGURA 4.3.3.3/1

MÉTODO ADAPTATIVO DE MATYAS

4.3.3.4 Método de Beltrami e Indusi

El método de Beltrami e Indusi es un método adaptativo muy similar al método determinístico de Hookes y Jeeves con etapas discretas y factor de aceleración $\alpha = 1$ (Ref. 8 y Ref. 5).

El algoritmo es como sigue:

Etapa 0. Seleccionar $x^1 \in S^n$ y un valor inicial para R. Hacer

$$y^1 = x^1 \text{ e } i = j = 1.$$

Etapa 1. Elegir un vector p^j cuyo módulo es R y cuya dirección se toma aleatoriamente con función de densidad uniforme.

Etapa 2. Si $f(y^j + p^j) - f(y^j) < 0$, el intento ha tenido éxito. Hacer $y^{j+1} = y^j + p^j$ e ir a 3.

De otra forma, el intento ha sido un fracaso:

$$\text{Si } f(y^j - p^j) - f(y^j) < 0, \text{ hacer } y^{j+1} = y^j - p^j \text{ e ir a 3.}$$

$$\text{De otra forma, hacer } y^{j+1} = y^j \text{ e ir a 3.}$$

Etapa 3. Si $j < n$, hacer $j = j + 1$ e ir a 1.

De otra forma, ir a 4.

Etapa 4. Hacer $x^{i+1} = y^{n+1}$ e $y^1 = 2x^{i+1} - x^i$.

Si $i < 2$, hacer $i = i + 1$, $j = 1$ e ir a 1.

De otra forma: si $f(x^3) < f(x) < f(x^2)$ 4.3.3.4/1

donde $x = 2x^2 - x^1$, se dice

del movimiento guiado $2x^2 - x^1$ que ha

sido una "búsqueda con éxito". Si la

búsqueda no puede encontrar un punto x^3

tal que se satisfaga la fórmula 4.3.3.4/1

entonces se vuelve a x^2 y se hace desde

allí una búsqueda exploratoria con

$R = R\Delta$, donde

$$\Delta = 1 - \frac{1}{0.27} \left(1 - \sqrt{1 - \frac{0.406}{n}} \right) \quad 4.3.3.4/2$$

En cualquier caso, se obtiene un punto, x^3 .

Etapla 5. Hacer $y^1 = 2x^3 - x^2$, $i = j = 1$, e ir a 1.

Si se hacen cinco movimiento guiados sucesivos que sean "búsqueda con éxito", entonces se aumenta el radio R .

Si el próximo movimiento guiado tiene éxito se retiene el radio incrementado y de otra forma el radio antiguo es tomado de nuevo.

Cuando el radio R se hace menor que un valor predeterminado suficientemente pequeño, se habrá conseguido la convergencia del algoritmo.

Se hace notar que las etapas 2 y 3 son una búsqueda exploratoria y la etapa 4 una búsqueda guiada.

Este método exhibe una rápida convergencia hasta que se llega a las cercanías de un mínimo. Entonces la convergencia es lenta.

Entre las características más interesantes del método se pueden citar las siguientes:

- El radio adaptativo de búsqueda, R .
- El coste de generación de un punto $y^j + p^j$ no se pierde en el caso de que éste sea desfavorable. Antes de rechazarlo se investiga en el sentido contrario, $y^j - p^j$.

La figura 4.3.3.4/1 ilustra el método para la función objetivo $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ y para el punto inicial $(0.0, 2.0)$.

4.3.4 Categoría 3 o métodos mixtos

Los métodos mixtos utilizan simultáneamente técnicas determinísticas y probabilísticas o diferentes tipos de técnicas probabilísticas.

Los métodos mixtos más significativos se han extraído de la literatura especializada y se describen a continuación.

4.3.4.1 Método Multipartida

El método multipartida, que es una extensión natural de la búsqueda aleatoria pura (ver 4.3.2.1), puede que sea el método más utilizado para intentar encontrar el mínimo global de una función objetivo cualquiera (Ref. 4, Art. 1).

El algoritmo consta de las siguientes etapas:

Etapa 0. Seleccionar aleatoriamente, de acuerdo con una distribución dada, un punto $x^0 \in S^n$.

Etapa 1. Utilizar un algoritmo de minimización local, con x^0 como punto inicial, hasta que se cumpla un criterio de parada predeterminado.

Etapa 2. Si puede considerarse que el punto encontrado es el mínimo global, parar.

De otra forma, ir a 0.

Como puede verse, el usuario de este método utiliza su técnica de optimización favorita desde distintos puntos iniciales x^0 . El conjunto de todos los puntos finales puede que incluya el mínimo global, x^* .

El método multipartida tiene la ventaja de que las características de la función objetivo que vayan siendo conocidas al ir aplicando el método pueden utilizarse para guiar la elección siguiente del punto inicial x^0 (en este caso se diría que es un método adaptativo). Sin embargo, cuando se conocen pocas características de la función objetivo el método necesita demasiadas evaluaciones de $f(x)$.

4.3.4.2 Método de Hartman S2

Para evitar el alto número de evaluaciones de $f(x)$ que utiliza el método multipartida cuando se conocen pocas características de la función objetivo se han desarrollado diversos métodos. Uno de ellos es el método de Hartman S2, que consta de las etapas siguientes (Ref. 4, Art. 1):

Etapla 0. Hacer $V^+ = \infty$

Etapla 1. Seleccionar aleatoriamente, de acuerdo con una distribución dada, un punto $x^0 \in S^n$.

Etapla 2. Si $f(x^0) > V^+$, ir a 1.

De otra forma, ir a 3.

Etapla 3. Realizar una minimización local desde x^0 hasta M^i (punto para el que se cumple un criterio de parada predeterminado). Hacer $V^+ = f(M^i)$.

Etapla 4. Ir a 1.

Con este método sólo se efectúa una búsqueda local adicional si el nuevo punto inicial es mejor que el mínimo global actual.

Normalmente este procedimiento acelera la obtención del mínimo global. Sin embargo, en algunos casos se rechazan puntos que conducirían a él. Esta situación especial se ilustra en la figura 4.3.4.2/1, donde será rechazado el punto A.

FIGURA 4.3.4.2/1

MÉTODO DE HARTMAN S2. SITUACIÓN ESPECIAL

4.3.4.3 Método de las Constelaciones de Törn

Para evitar el alto coste del método multipartida, se ha introducido en la minimización global el concepto de "constelación". La idea básica de este concepto es que los puntos con valores más bajos de la función objetivo estarán agrupados en las cercanías de los mínimos locales, por lo que una búsqueda posterior desde esos puntos conduciría al mismo mínimo local. De esta forma, se podría obtener un ahorro de tiempo considerable si la búsqueda se continuara desde un número de puntos reducido.

Törn ha efectuado varias implementaciones de su método, el cual está descrito en la Ref. 4, Art. 4, analizado en la Ref. 4, Art. 5 y resumido en la Ref. 5.

El método de Törn consta de las siguientes etapas:

- Etapla 1. Tomar una muestra aleatoria de puntos, mediante una distribución uniforme, sobre el espacio de interés, S^n .
- Etapla 2. Con la ayuda de una subrutina de minimización local, llevar cada punto hacia un mínimo local.
- Etapla 3. Utilizar un método conocido en el reconocimiento de patrones como "agrupamiento" para agrupar los puntos en constelaciones definidas como conjuntos de puntos cuya distancia entre sí es menor que la distancia media desde otros.
- Etapla 4. Descartar una fracción de puntos determinada en cada grupo (por ejemplo, retener uno de cada dos o tres puntos).

Etapa 5. Ir a 2 hasta que se satisfaga un criterio de parada predeterminado.

Este método ha demostrado ser un método fiable para obtener el mínimo global.

Las figuras 4.3.4.3/1, 2, 3 y 4 ilustran las distintas etapas del método para una función objetivo arbitraria.

FIGURA 4.3.4.3/1

MÉTODO DE LAS CONSTELACIONES DE TÖRN. ETAPA 1

FIGURA 4.3.4.3/2

MÉTODO DE LAS CONSTELACIONES DE TÖRN. ETAPA 2

FIGURA 4.3.4.3/3

MÉTODO DE LAS CONSTELACIONES DE TÖRN. ETAPA 3

FIGURA 4.3.4.3/4

MÉTODO DE LAS CONSTELACIONES DE TÖRN. ETAPA 4

4.3.4.4 Método Estocástico-Determinístico de Fagioli

Todos los métodos de la Categoría 1, que han sido descritos anteriormente, utilizan un conjunto de puntos distribuidos aleatoriamente en el espacio S^n . Sin embargo, Fagioli y otros (Ref. 4, Art. 8) definen una técnica de cobertura espacial donde el espacio a investigar se divide en celdas. A continuación se seleccionan de forma aleatoria una celda y un punto dentro de ella. Y, por último, desde ese punto, se emplea un método de minimización determinístico local.

El algoritmo consta de las etapas siguientes:

Etapla 0. Dividir el espacio S^n en N celdas, C_j , $j = 1(1)N$.

Tomar un valor V^+ suficientemente grande.

Etapla 1. Elegir una celda, C_k , de acuerdo con una distribución de probabilidad dada.

Etapla 2. Dentro de la celda, elegir un punto, x^k , de acuerdo con una distribución de probabilidad uniforme.

Etapla 3. Calcular $f(x^k)$. Si $f(x^k) - V^+ > 0$, ir a 1.

Si $f(x^k) - V^+ \leq 0$, emplear el algoritmo determinístico de minimización local dentro de la celda. Sea M^k el resultado del proceso de minimización. Hacer $V^+ = f(M^k)$.

Etapla 4. Ir a 1 hasta que se satisfaga un criterio de parada predeterminado.

4.3.4.5 Método de Gaviano

El método propuesto por Gaviano consta de las etapas siguientes (Ref. 3, Art. 8):

Etapa 0. Seleccionar $x^0 \in S^n$. Hacer $i = 0$.

Etapa 1. Elegir un vector, p^i cuyo módulo sea la unidad y cuya dirección se tome aleatoriamente con función de densidad uniforme.

Etapa 2. Calcular el escalar λ_i tal que

$$f(x^i + \lambda_i p^i) = \min\{f(x^i + \lambda p^i) / \lambda \geq 0\}$$

Etapa 3. Calcular $x^{i+1} = x^i + \lambda_i p^i$. Hacer $i = i+1$ e ir a 1.

Según esto, se realiza una minimización global según direcciones elegidas arbitrariamente. Gaviano ha demostrado que el algoritmo converge hacia el mínimo global.

La figura 4.3.4.5/1 ilustra el método para la función objetivo $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ y para el punto inicial $(0.0, 2.0)$.

FIGURA 4.3.4.5/1

MÉTODO DE GAVIANO

5. LOS ALGORITMOS GENÉTICOS

5.1 INTRODUCCIÓN

De acuerdo con la Teoría de la Evolución o Evolucionismo (ver Apéndice I), todos los seres vivos son el resultado de un proceso muy largo de desarrollo durante el cual han evolucionado, han mejorado o, lo que es lo mismo, han luchado contra el incremento de entropía de su entorno.

La observación de la Naturaleza muestra que las estructuras, procesos y sistemas biológicos están muy bien adaptados para realizar sus funciones. Incluso ha sido posible demostrar la optimización total de ciertos subsistemas biológicos mediante la utilización de métodos matemáticos. Algunos ejemplos típicos de tales soluciones óptimas son:

- a) La reducción del diámetro de las arterias en sus ramificaciones.
- b) La relación volumétrica de partículas sólidas en la sangre (valor hemocrítico).
- c) La posición de los puntos límites de un sistema circulatorio plano.

Para un caso general, no puede demostrarse esta optimización total. Sin embargo, sí parece claro, a la luz del presente conocimiento biológico, que los seres vivos están, al menos parcialmente, optimizados. Esto es el motivo por el que la evolución biológica representa una estrategia pseudoóptima para la adaptación de los seres vivos a su entorno.

Según esto, la imitación de las estrategias utilizadas por los organismos para resolver sus problemas de supervivencia y evolución podría conducir a una resolución muy eficiente de los problemas técnicos de optimización.

La pregunta que surge ahora es la siguiente: ¿Cómo imitar las estrategias de los seres vivos, si éstas no se conocen completamente? La respuesta es obvia: sólo se imitarán aquellos aspectos del Evolucionismo más importantes y perfectamente conocidos, dejando de lado las cuestiones secundarias.

La definición de Algoritmos Genéticos que se propone en esta Tesis es la siguiente: Los Algoritmos Genéticos son métodos probabilísticos adaptativos basados en la imitación de las estrategias utilizadas en la evolución natural de los seres vivos.

Son métodos probabilísticos, pues hacen uso de números aleatorios y de probabilidades y son adaptativos pues utilizan la información recogida en etapas anteriores para guiar la búsqueda presente.

Algunos Algoritmos Genéticos son métodos de rastreo aleatorio (Categoría 2) y otros métodos mixtos (Categoría 3). Asimismo, puede haber métodos locales y métodos globales.

El presente capítulo se divide en cinco apartados.

El apartado 5.2 presenta el modelo muy simplificado de la Evolución que se estudiará en esta Tesis.

El apartado 5.3 establece la formulación matemática de los algoritmos (μ, λ) y $(\mu + \lambda)$, basada en el modelo de Evolución estudiado en el apartado anterior.

El apartado 5.4 particulariza los resultados del apartado anterior para el caso $\mu = 1$. Son de una importancia extraordinaria

ria los subapartados 5.4.5, en el que se propone un método para calcular el tamaño de paso óptimo de estos algoritmos y además para controlarlo, y el 5.4.6, en el que se recogen las conclusiones obtenidas.

Por último, el apartado 5.5 investiga las posibles mejoras de los algoritmos genéticos en estudio.

5.2 EL MODELO SIMPLIFICADO DE LA EVOLUCIÓN

La base fundamental del Evolucionismo de Darwin está constituida por los dos principios siguientes (ver Apéndice I):

- Mutación
- Selección

Un modelo altamente simplificado de la Evolución, que incluye estos dos principios, puede desarrollarse como sigue (Ref. 6):

0. INICIALIZACIÓN: Se considera una población formada por μ individuos denominados padres. Estos μ individuos constituyen una generación de padres. Cada individuo se caracteriza por su genotipo.
1. MUTACIÓN: A partir de los μ padres se obtiene una generación de λ nuevos individuos, llamados hijos, donde $\lambda \geq \mu$. Los cambios en los genes (mutaciones) son pequeños, estocásticos e independientes entre sí. Por tanto, el genotipo de cada descendiente diferirá muy ligeramente del de su progenitor.
2. SELECCIÓN: Los $\mu + \lambda$ individuos de los que ahora se dispone (padres e hijos) tienen distinta vitalidad por tener distintos genotipos. Sólo sobrevivirán y evolucionarán los μ mejores individuos y ellos constituirán la siguiente generación de padres, con lo que el proceso se repite.

Para seleccionar esos μ mejores individuos se pueden utilizar dos estrategias:

- a) Estrategia (μ, λ) : Los μ padres serán los mejores individuos entre los λ (hijos) posibles. Según esta estrategia los padres, tras procrear, mueren.
- b) Estrategia $(\mu + \lambda)$: Los μ padres serán los mejores individuos entre los $\mu + \lambda$ (padres e hijos) posibles. Según esta estrategia, los padres pueden vivir y generar descendientes indefinidamente.

Este modelo asume lo siguiente:

- a) El número total de individuos en la población permanece constante.
- b) Sólo hay mutaciones puntuales de genes.
- c) El entorno, y por lo tanto el criterio de vitalidad, se mantiene constante.

Para aplicar a la resolución de los problemas matemáticos de optimización el modelo simplificado de la Evolución, esbozado anteriormente, se establecen las siguientes analogías:

- Cromosoma: Equivalente a vector de n componentes en el espacio $\mathbb{R}^n$ (espacio de variables), x .
- Individuo: Equivalente a punto en el espacio de variables, x .
- Población: Equivalente a conjunto de puntos en el espacio de variables.

- Generación de hijos: Equivalente a conjunto de λ puntos en el espacio de variables, H_g .
- Generación de padres: Equivalente a conjunto de μ puntos en el espacio de variables, G_g .
- Vitalidad de un individuo; Equivalente a valor de la función objetivo en un punto, $-f(x)$.

Obsérvese que según este modelo, los conceptos de cromosoma e individuo son idénticos, cosa que en la realidad, obviamente, no ocurre.

5.3 LOS ALGORITMOS (μ, λ) Y $(\mu + \lambda)$

5.3.1 Generalidades

En este apartado se considera el caso más general posible, en el cual μ y λ pueden ser cualesquiera, con una única restricción: cada padre genera λ/μ hijos, y por ello λ ha de ser múltiplo de μ .

El subapartado 5.3.2 presenta los algoritmos (μ, λ) y $(\mu + \lambda)$ los cuales han sido obtenidos traduciendo al lenguaje matemático el modelo simplificado de la Evolución del apartado 5.2.

En el subapartado 5.3.3 se hacen una serie de consideraciones respecto al valor esperado del tamaño de paso.

Por último, el subapartado 5.3.4 recoge las funciones objetivo a las que serán aplicados los algoritmos descritos, ya que no es posible realizar un análisis general válido para una función objetivo arbitraria.

5.3.2 Desarrollo de los algoritmos (μ, λ) y $(\mu + \lambda)$

- El algoritmo (μ, λ) es el resultado obtenido al traducir al lenguaje matemático la estrategia (μ, λ) del modelo simplificado de la Evolución. Consta de las etapas siguientes:

Etapla 0. Seleccionar, mediante un criterio adecuado, una serie de puntos iniciales, x^{ok} , $k = 1(1)\mu$ (Población de μ padres de la 0-ésima generación).

Hacer $g = 0$ (contador de generaciones).

Etapla 1. Calcular $y^{g1} = x^{gk} + z^{g1}$, $l = 1(1)\lambda$ (Población de λ hijos de la g -ésima generación), donde

$$k = 1 + \text{Int}[(l-1)/\frac{\lambda}{\mu}]$$

$$y \text{ } z^{g1} \sim N(0, \sigma^g).$$

Etapla 2. Clasificar los y^{g1} , $l=1(1)\lambda$ de tal forma que $f(y^{g11}) \leq f(y^{g12})$, para $l_1 = 1(1)\mu$ y $l_2 = \mu + 1(1)\lambda$.

Hacer $x^{g+1,k} = y^{g11}$, $k = l_1 = 1(1)\mu$.

Hacer $g = g+1$ e ir a 1, hasta que se satisfaga un criterio de parada predeterminado.

- El algoritmo $(\mu + \lambda)$ es el resultado obtenido al traducir al lenguaje matemático la estrategia $(\mu + \lambda)$ del modelo simplificado de la Evolución. Consta de las etapas siguientes:

Etapla 0. Seleccionar, mediante un criterio adecuado, una serie de puntos iniciales, x^{ok} , $k = 1(1)\mu$ (Población de μ padres de la 0-ésima generación).

Hacer $g = 0$ (contador de generaciones)

Etapla 1. Calcular $y^{g1} = x^{gk} + z^{g1}$, $l = 1(1)\lambda$ (Población de λ hijos de la g -ésima generación), donde

$$k = 1 + \text{Int}\left[(l-1)/\frac{\lambda}{\mu}\right]$$

$$y \text{ } z^{g1} \sim N(0, \sigma^g)$$

Hacer $x^{g, \mu+1} = y^{g1}$, $l = 1(1)\lambda$.

Etapla 2. Clasificar los x^{g1} , $l = 1(1) \mu + \lambda$ de tal forma que

$$f(x^{g11}) \leq f(x^{g12}), \text{ para } l_1 = 1(1)\mu \text{ y } l_2 = \mu + 1(1) \mu + \lambda.$$

Hacer $x^{g+1, k} = x^{g11}$, $k = l_1 = 1(1)\mu$.

Hacer $g = g + 1$ e ir a 1, hasta que se satisfaga un criterio de parada predeterminado.

Es necesario realizar una aclaración. El vector aleatorio z^{g1} representa la mutación en el l -ésimo hijo de la g -ésima generación. De acuerdo con el presente conocimiento biológico, las mutaciones se consideran como sucesos estocásticos producidos por multitud de causas. Entonces, de acuerdo con el teorema central del limite, debe admitirse que el vector aleatorio está distribuido normalmente. Además, la analogía establecida con la evolución natural indica que:

$$\begin{aligned} E\{z^{g1}\} &= 0 & 5.3.2/1 \\ E\{z_i^{g1} z_j^{g1}\} &= \begin{cases} (\sigma_i^g)^2, & i = j, i = 1(1)n, j = 1(1)n \\ 0, & i \neq j, i = 1(1)n, j = 1(1)n \end{cases} & 5.3.2/2 \end{aligned}$$

La figura 5.3.2/1 ilustra una iteración del algoritmo (2,4) para la función objetivo $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$.

FIGURA 5.3.2/1

ALGORITMO (2,4)

5.3.3 El valor esperado del tamaño de paso

El tamaño de paso desde el punto x^{gk} hasta el punto y^{g1} viene dado por la longitud, s , del vector z^{g1} . Esta longitud no puede ser predicha antes de realizar la prueba.

Por ello, debe utilizarse el concepto de valor esperado del tamaño de paso, definido como $E\{s\}$, en lugar del concepto de tamaño de paso empleado en los métodos determinísticos.

Se demostrará a continuación que este valor esperado es directamente proporcional a la desviación típica del vector aleatorio z . Esta es la razón por la que en la literatura especializada sobre métodos probabilísticos se denominan tamaños de paso a las desviaciones típicas.

La función de densidad gaussiana de un vector aleatorio, z , de n componentes, viene dada por

$$p(z) = \frac{1}{(2\pi)^{n/2} |P|^{1/2}} \exp\left[-\frac{1}{2}(z-\bar{z})^T P^{-1} (z-\bar{z})\right] \quad 5.3.3/1$$

donde

$$\bar{z} = E\{z\} \quad 5.3.3/2$$

$$P = E\{(z-\bar{z})(z-\bar{z})^T\} \quad 5.3.3/3$$

En el caso estudiado se ha supuesto que:

$$\bar{z} = 0 \quad 5.3.3/4$$

$$P = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix} \quad 5.3.3/5$$

Según esto, la ecuación 5.3.3/1 se reduce a

$$p(z) = \frac{1}{(2\pi)^{n/2} \prod_{i=1}^n \sigma_i} \exp\left[-\frac{1}{2} \sum_{i=1}^n \left(\frac{z_i}{\sigma_i}\right)^2\right] \quad 5.3.3/6$$

Para el caso particular muy comúnmente estudiado de $\sigma_i = \sigma$, $i = 1(1)n$, la ecuación 5.3.3/6 tiene la forma:

$$p(z) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n z_i^2\right] \quad 5.3.3/7$$

Notar que el cuadrado de la longitud, s , del vector aleatorio z es la suma de los cuadrados de n variables normalmente distribuidas, con desviaciones típicas σ . Por ello,

$$\frac{s^2}{\sigma^2} \sim \chi_n^2 \quad 5.3.3/8$$

donde χ_n^2 es la distribución de Pearson. Dos de sus propiedades son:

$$E\{\chi_n^2\} = n \quad 5.3.3/9$$

$$E\{(\chi_n^2 - n)^2\} = 2n \quad 5.3.3/10$$

Cuando el número de variables n es grande, la distribución de Pearson tiende a la distribución normal. En este caso, se obtienen las expresiones siguientes:

$$\bar{s} = E\{s\} = \sigma\sqrt{n} \quad 5.3.3/11$$

$$\sigma_s^2 = E\{(\mathbf{s} - \tilde{\mathbf{s}})^2\} = \sigma^2 \sqrt{2n}$$

5.3.3/12

La ecuación 5.3.3/11 muestra que para una dimensión del espacio de variables dada, el valor esperado del tamaño de paso es proporcional a la desviación típica.

5.3.4 Los modelos de función objetivo utilizados

El análisis teórico de las estrategias (μ, λ) y $(\mu + \lambda)$ no ha podido ser realizado, hasta el momento, para funciones objetivo cualesquiera (Ref. 6 y Ref. 9).

Por ello, es preciso simplificar el problema general y estudiar las propiedades teóricas que resultan para funciones objetivo simples. Se analizarán dos funciones objetivo: el modelo de corredor y el modelo esférico.

El modelo de corredor representa una variación de $f(x)$ a lo largo de una dirección. Mediante un cambio de variables, esta dirección puede hacerse coincidir siempre con el eje x_1 . La función se define como sigue:

$$f(x) = \begin{cases} f(x_1) & , -b < x_1 < b, \quad i = 2(1)n \\ \infty & , \text{ en el resto de } \mathbb{R}^n \end{cases} \quad 5.3.4/1$$

donde $f(x_1)$ es una función monótona decreciente.

La figura 5.3.4/1 muestra un modelo de corredor en un espacio bidimensional.

El modelo esférico representa una reducción monótona radial hacia el mínimo de la función objetivo. Esta función se define como sigue:

$$f(x) = \sum_{i=1}^n x_i^2 \quad 5.3.4/2$$

La figura 5.3.4/2 muestra un modelo esférico en un espacio bidimensional.

El interés fundamental de estos dos modelos, aparte de su sencillez, es que mientras el modelo esférico puede considerarse como una buena aproximación de cualquier función objetivo en las cercanías de un mínimo, el modelo de corredor representa adecuadamente el comportamiento de la función suficientemente lejos del mínimo.

FIGURA 5.3.4/1

MODELO DE CORREDOR

FIGURA 5.3.4/2

MODELO ESFÉRICO

5.4 LOS ALGORITMOS $(1,\lambda)$ Y $(1+\lambda)$

5.4.1 Generalidades

Como casos particulares de los algoritmos (y,λ) y $(y+\lambda)$, se estudian en este apartado los algoritmos $(1,\lambda)$ y $(1+\lambda)$ obtenidos a partir de los anteriores haciendo $y = 1$.

La razón para estudiar estos casos particulares es doble. Por una parte, constituyen un primer paso necesario para posteriores generalizaciones. Y, por otra parte, como se demostrará, su sencillez les confiere una extraordinaria potencia.

En el subapartado 5.4.2 se definen las características teóricas que serán analizadas cuando estos algoritmos son aplicados al modelo de corredor y al modelo esférico.

El subapartado 5.4.3 presenta la probabilidad de éxito y la velocidad de convergencia obtenidas para el algoritmo $(1+1)$.

El subapartado 5.4.4 analiza, análogamente, estas características para los algoritmos $(1,\lambda)$ y $(1+\lambda)$.

En el subapartado 5.4.5 se calculan los tamaños de paso óptimos y se propone un método para controlar el tamaño de paso basado en la probabilidad de éxito cuando este tamaño de paso es óptimo.

Por último, el subapartado 5.4.6 presenta las conclusiones obtenidas en este estudio.

5.4.2 Características teóricas estudiadas

Se aplicarán los algoritmos $(1,\lambda)$ y $(1+\lambda)$ a dos funciones objetivo que sigan el modelo de corredor y el modelo esférico y se estudiarán tres características teóricas:

- La probabilidad de éxito, P_λ
- La velocidad de convergencia, Φ
- La relación Φ/λ

La probabilidad de éxito, P_λ , se define como la probabilidad de que alguno de los λ hijos sea mejor (tenga un valor de la función objetivo mejor) que el padre. Obsérvese que la probabilidad de éxito de la estrategia $(1,\lambda)$ coincide con la de la estrategia $(1+\lambda)$.

La velocidad de convergencia, Φ , se define como el valor esperado de la reducción de la distancia, en cada iteración, entre el óptimo y el padre de la generación. Si se denota mediante x^* el óptimo y mediante x^{g1} el padre de la generación g , la distancia entre x^* y x^{g1} viene dada por:

$$d_g = \sqrt{\sum_{i=1}^n (x_i^* - x_i^{g1})^2} \quad 5.4.2/1$$

y, por tanto, la velocidad de convergencia es:

$$\phi = E\{d_g - d_{g+1}\} \quad 5.4.2/2$$

Por último, el estudio de la relación Φ/λ tiene un gran interés por lo siguiente. Se consideran los padres de las generacio-

nes, g y $g+1$, situados a distancias del óptimo d_g y d_{g+1} (ver figura 5.4.2/1). El tiempo empleado en calcular la generación de hijos H_g es proporcional al número de éstos. Es decir:

$$t = k \lambda \quad 5.4.2/3$$

El número de generaciones necesarias para llegar al óptimo será:

$$N = \frac{d_g}{\phi} \quad 5.4.2/4$$

y el tiempo total empleado en llegar al óptimo

$$T = Nt = kd_g \frac{\lambda}{\phi} \quad 5.4.2/5$$

Por tanto, la relación Φ/λ es inversamente proporcional al tiempo total necesario para alcanzar el óptimo. El mejor de todos los algoritmos posibles será aquel para el que Φ/λ sea máximo.

• Generación de hijos, H_g

FIGURA 5.4.2/1

VELOCIDAD DE CONVERGENCIA, Φ

5.4.3 Probabilidad de éxito y velocidad de convergencia para el algoritmo (1+1)

En este subapartado se recogen, sin demostración, los resultados obtenidos por Rechenberg (Ref. 9) con el algoritmo (1+1) para el modelo de corredor y el modelo esférico. Para ambos modelos se calcularon la probabilidad de éxito y la velocidad de convergencia.

5.4.3.1 Modelo de corredor

a) Probabilidad de éxito

La probabilidad de éxito es independiente de la coordenada x_1 del punto considerado, pero es dependiente del resto de las coordenadas. Su valor es:

$$P_1(x) = \frac{1}{2} \left\{ \frac{1}{2} \left[\Phi\left(\frac{b-x_2}{\sqrt{2}\sigma}\right) + \Phi\left(\frac{b+x_2}{\sqrt{2}\sigma}\right) \right] \cdot \dots \cdot \right. \\ \left. \cdot \frac{1}{2} \left[\Phi\left(\frac{b-x_n}{\sqrt{2}\sigma}\right) + \Phi\left(\frac{b+x_n}{\sqrt{2}\sigma}\right) \right] \right\} \quad 5.4.3.1/1$$

donde

$\Phi \equiv$ función integral de error de Gauss:

$$\Phi(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-z^2} dz \quad 5.4.3.1/2$$

$b \equiv$ semianchura del corredor

$\sigma \equiv$ desviación típica del vector aleatorio z , normalmente distribuido, igual para todas las coordenadas ($\sigma_i = \sigma$, $i = 1(1)n$).

A pesar de que la expresión 5.4.3.1/1 proporciona el valor de la verdadera probabilidad de éxito, Rechenberg utiliza en vez de ésta su valor medio en la sección x_1 . Es decir, se supone que el padre de la generación está situado a cierta distancia del óptimo, sin importar su localización exacta en la sección correspondiente. La expresión que se obtiene es:

$$P_1 = \frac{1}{2} \left[\Phi\left(\sqrt{2} \frac{b}{\sigma}\right) - \frac{1}{\sqrt{2\pi}} \frac{\sigma}{b} \left(1 - e^{-2\frac{b^2}{\sigma^2}}\right) \right]^{n-1} \quad 5.4.3.1/3$$

Si además se cumple $\sigma/b \ll 1$, la anterior expresión conduce

a:

$$P_1 = \frac{1}{2} \left(1 - \frac{1}{\sqrt{2\pi}} \frac{\sigma}{b}\right)^{n-1} \quad \left(\frac{\sigma}{b} \ll 1\right) \quad 5.4.3.1/4$$

b) Velocidad de convergencia

Para la velocidad de convergencia se obtiene:

$$\begin{aligned} \phi(x) = \frac{\sigma}{\sqrt{2\pi}} \left\{ \frac{1}{2} \left[\Phi\left(\frac{b-x_2}{\sqrt{2}\sigma}\right) + \Phi\left(\frac{b+x_2}{\sqrt{2}\sigma}\right) \right] \cdot \dots \cdot \right. \\ \left. \cdot \frac{1}{2} \left[\Phi\left(\frac{b-x_n}{\sqrt{2}\sigma}\right) + \Phi\left(\frac{b+x_n}{\sqrt{2}\sigma}\right) \right] \right\} \end{aligned} \quad 5.4.3.1/5$$

donde Φ ya ha sido definida mediante la expresión 5.4.3.1/2.

Haciendo las mismas consideraciones que en el caso de la probabilidad de éxito, la velocidad de convergencia media en la sección x_1 es:

$$\phi = \frac{\sigma}{\sqrt{2\pi}} \left[\Phi\left(\sqrt{2} \frac{b}{\sigma}\right) - \frac{1}{\sqrt{2\pi}} \frac{\sigma}{b} \left(1 - e^{-2\frac{b^2}{\sigma^2}}\right) \right]^{n-1} \quad 5.4.3.1/6$$

En el caso particular de $\sigma/b \ll 1$, resulta

$$\phi = \frac{\sigma}{\sqrt{2\pi}} \left(1 - \frac{1}{\sqrt{2\pi}} \frac{\sigma}{b}\right)^{n-1} \quad \left(\frac{\sigma}{b} \ll 1\right) \quad 5.4.3.1/7$$

5.4.3.2 Modelo esférico

a) Probabilidad de éxito

La probabilidad de éxito, debido a la simetría de la función objetivo, sólo depende de la distancia desde el padre hasta el mínimo.

La fórmula general que permite obtener esta probabilidad de éxito es:

$$P_1(r_o) = \frac{ae^{-a}}{2v} [I_{v-1}(a) + I_v(a)] - \frac{1}{2v} \int_0^a e^{-t} I_v(t) dt \quad 5.4.3.2/1$$

donde

$I_p(z)$ $\equiv$ solución de la ecuación de Bessel modificada

$$I_p(z) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(p+k+1)} \left(\frac{z}{2}\right)^{p+2k} \quad 5.4.3.2/2$$

Γ $\equiv$ función Gamma

r_o $\equiv$ distancia desde el padre hasta el mínimo

σ $\equiv$ desviación típica del vector aleatorio z , normalmente distribuido, igual para todas las coordenadas ($\sigma_i = \sigma$, $i = 1(1)n$)

$$a = r_o^2 / \sigma^2 \quad 5.4.3.2/3$$

$$v = n/2 \quad 5.4.3.2/4$$

La fórmula general, 5.4.3.2/1 puede simplificarse si se consideran las relaciones adicionales $n \gg 1$ y $n\sigma^2/r_o^2 \ll 1$. Con éstas se obtendría:

$$P_1(r_o) = \frac{1}{2} \left[1 - \Phi\left(\frac{n\sigma}{\sqrt{8}r_o}\right) \right] \quad (n \gg 1, \frac{n\sigma^2}{r_o^2} \ll 1) \quad 5.4.3.2/5$$

donde Φ viene dada por la expresión 5.4.3.1/2.

b) Velocidad de convergencia

Para la velocidad de convergencia se obtiene la expresión general siguiente:

$$\frac{\phi(r_o)}{r_o} = e^{-a} I_\nu(a) - a e^{-a/2} \int_0^1 e^{-\frac{a}{2}u^2} u^{\nu+1} [I_{\nu-1}(au) - I_\nu(au)] du \quad 5.4.3.2/6$$

donde I , a , ν vienen dadas por las expresiones 5.4.3.2/2, 3 y 4 respectivamente.

En el caso particular en que $n \gg 1$ y $n\sigma^2/r_o^2 \ll 1$, resulta:

$$\phi = \frac{\sigma}{\sqrt{2\pi}} \left\{ e^{-\left(\frac{n\sigma}{\sqrt{8}r_o}\right)^2} - \sqrt{\pi} \frac{n\sigma}{\sqrt{8}r_o} \left[1 - \Phi\left(\frac{n\sigma}{\sqrt{8}r_o}\right) \right] \right\} \quad (n \gg 1, \frac{n\sigma^2}{r_o^2} \ll 1)$$

5.4.3.2/7

donde Φ ha sido definida mediante la expresión 5.4.3.1/2.

5.4.4 Probabilidad de éxito y velocidad de convergencia para los algoritmos $(1,\lambda)$ y $(1+\lambda)$

En este subapartado se estudian, para el modelo de corredor y el modelo esférico, la probabilidad de éxito y la velocidad de convergencia obtenidas mediante la aplicación de los algoritmos $(1,\lambda)$ y $(1+\lambda)$.

Como ya se ha señalado, la probabilidad de éxito es la misma para los algoritmos $(1,\lambda)$ y $(1+\lambda)$. Sin embargo, la velocidad de convergencia responderá a distintas expresiones para los dos algoritmos estudiados.

5.4.4.1 Modelo de corredor

a) Probabilidad de éxito

La probabilidad de éxito para los algoritmos $(1, \lambda)$ y $(1 + \lambda)$ se obtiene fácilmente a partir de la expresión análoga para el algoritmo $(1 + 1)$, mediante la fórmula:

$$P_{\lambda} = 1 - [1 - P_1]^{\lambda} \quad 5.4.4.1/1$$

donde P_1 es la probabilidad de éxito para el algoritmo $(1 + 1)$.

Si se quiere obtener una fórmula general que proporcione la probabilidad de éxito, P_1 vendrá dada por la expresión 5.4.3.1/3, mientras que si se considera suficiente la aproximación obtenida con $\sigma/b \ll 1$, P_1 deberá sustituirse por la expresión 5.4.3.1/4.

La figura 5.4.4.1/1 representa la probabilidad de éxito en función de la desviación típica adimensionalizada,

$$\sigma^* = \frac{\sigma n}{b} , \quad 5.4.4.1/2$$

para distintos valores de λ . La curva de trazo continuo corresponde a los resultados teóricos obtenidos a partir de la expresión 5.4.4.1/1, con $n = 10$, $b = 5$. Para estos mismos datos se ha realizado una simulación con los programas de que se dispone, representándose los resultados obtenidos mediante puntos.

Obsérvese que la variación de la probabilidad de éxito con la desviación típica viene dada por:

$$\frac{dP_\lambda}{d\sigma} = \lambda (1-P_1)^{\lambda-1} \frac{dP_1}{d\sigma} \quad 5.4.4.1/3$$

Según esto, los signos de $dP_\lambda/d\sigma$ y $dP_1/d\sigma$ son los mismos.

Como $dP_1/d\sigma$ siempre es negativo, se deduce que

$$\frac{dP_\lambda}{d\sigma} < 0 \quad 5.4.4.1/4$$

b) Velocidad de convergencia

De acuerdo con la Ref. 6, la velocidad de convergencia para el algoritmo (1, λ) viene dada por la expresión

$$\lambda = 1 + \sqrt{\pi} \frac{\phi^{*'}}{\sqrt{2}\sigma^*} \exp\left[\left(\frac{\phi^{*'}}{\sqrt{2}\sigma^*}\right)^2\right] \left[\Phi\left(\frac{\phi^{*'}}{\sqrt{2}\sigma^*}\right) + 2\exp\left(\frac{\sigma^{*2}}{\sqrt{2}\pi}\right) - 1\right] \quad 5.4.4.1/5$$

donde:

$$\Phi(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-z^2} dz \quad 5.4.4.1/6$$

$$\phi^{*'} = \frac{\phi^*}{1 - \left[1 - \exp\left(\frac{-\sigma^{*2}}{\sqrt{2}\pi}\right)\right]^\lambda} \quad 5.4.4.1/7$$

$$\phi^* = \frac{\phi n}{b} \quad 5.4.4.1/8$$

$$\sigma^* = \frac{\sigma n}{b} \quad 5.4.4.1/9$$

La ecuación 5.4.4.1/5 no puede resolverse explícitamente de forma que proporcione $\Phi^* = f(\sigma^*)$ para distintos λ . La figura 5.4.4.1/2 muestra los resultados teóricos obtenidos mediante iteración para el algoritmo (1, λ) así como los resultados de las

simulaciones realizadas.

No se dispone de resultados teóricos para el algoritmo $(1+\lambda)$. Solamente se han efectuado simulaciones que se recogen en la misma figura 5.4.4.1/2.

Por último, la figura 5.4.4.1/3 proporciona la velocidad de convergencia adimensionalizada máxima, $\Phi_{\max}^*$, en función del número de hijos, λ , para el algoritmo $(1,\lambda)$.

FIGURA 5.4.4.1/1

PROBABILIDAD DE EXITO. MODELO DE CORREDOR. ($n = 10$, $b = 5$)

FIGURA 5.4.4.1/2

VELOCIDAD DE CONVERGENCIA ADIMENSIONALIZADA PARA EL MODELO DE CORREDOR.

FIGURA 5.4.4.1/3

VELOCIDAD DE CONVERGENCIA ADIMENSIONALIZADA MÁXIMA PARA EL MODELO DE CORREDOR Y PARA EL ALGORITMO $(1, \lambda)$.

5.4.4.2 Modelo esférico

a) Probabilidad de éxito

El desarrollo que se hace aquí es idéntico al realizado en el párrafo 5.4.4.1.

La probabilidad de éxito para los algoritmos $(1, \lambda)$ y $(1 + \lambda)$ se obtiene mediante la fórmula:

$$P_{\lambda} = 1 - [1 - P_1]^{\lambda} \quad 5.4.4.2/1$$

donde P_1 es la probabilidad de éxito para el algoritmo $(1+1)$, que viene dada por la expresión 5.4.3.2/1 en el caso general y por la expresión 5.4.3.2/5 en el caso particular de $n \gg 1$ y $n\sigma^2/r_o^2 \ll 1$.

La figura 5.4.4.2/1 representa la probabilidad de éxito en función de la desviación típica adimensionalizada,

$$\sigma^* = \frac{\sigma n}{r_o} \quad 5.4.4.2/2$$

para distintos valores de λ y para $n = 10$ y $r_o = 5$. Los resultados obtenidos son unos teóricos y otros simulados.

Al igual que en el modelo de corredor, se hace notar que el signo de $dP_{\lambda}/d\sigma$ es siempre negativo.

b) Velocidad de convergencia

De acuerdo con la Ref. 6, la velocidad de convergencia para el algoritmo $(1, \lambda)$ viene dada por la expresión

$$\lambda = 1 + \sqrt{\pi} \left(\frac{\phi^*}{\sqrt{2}\sigma^*} + \frac{\sigma^*}{\sqrt{8}} \right) \exp \left[\left(\frac{\phi^*}{\sqrt{2}\sigma^*} + \frac{\sigma^*}{\sqrt{8}} \right)^2 \right] \left[1 + \phi \left(\frac{\sigma^*}{\sqrt{8}} \right) \right] \quad 5.4.4.2/3$$

donde:

$$\phi(u) = \frac{2}{\sqrt{\pi}} \int_0^u e^{-z^2} dz \quad 5.4.4.2/4$$

$$\phi^* = \frac{\phi n}{r_o} \quad 5.4.4.2/5$$

$$\sigma^* = \frac{\sigma n}{r_o} \quad 5.4.4.2/6$$

La figura 5.4.4.2/2 muestra los resultados teóricos y simulados obtenidos para el algoritmo $(1, \lambda)$ y los resultados simulados para el algoritmo $(1 + \lambda)$.

La figura 5.4.4.2/3 proporciona la velocidad de convergencia adimensionalizada máxima, $\Phi_{\max}^*$ en función del número de hijos, λ , para el algoritmo $(1, \lambda)$.

FIGURA 5.4.4.2/1

PROBABILIDAD DE EXITO. MODELO ESFERICO. ($n = 10$, $r_0 = 5$)

FIGURA 5.4.4.2/2

VELOCIDAD DE CONVERGENCIA ADIMENSIONALIZADA PARA EL MODELO ESFÉRICO

FIGURA 5.4.4.2/3

VELOCIDAD DE CONVERGENCIA ADIMENSIONALIZADA MÁXIMA PARA EL MODELO ESFÉRICO Y PARA EL ALGORITMO $(1, \lambda)$.

5.4.5 Tamaño de paso óptimo y control de tamaño de paso

El tamaño de paso óptimo será aquel tamaño de paso para el que la velocidad de convergencia sea máxima. Obsérvese que el concepto de tamaño de paso es equivalente al de desviación típica (ver 5.3.3).

Si las desviaciones típicas son muy pequeñas, la probabilidad de éxito será muy grande, pero cada hijo que tenga éxito avanzará poco hacia el mínimo. Sin embargo, para desviaciones típicas grandes la probabilidad de éxito empeorará y el avance de los hijos hacia el mínimo mejorará. Esta es la razón por la que existirá una desviación típica óptima ni muy grande ni muy pequeña.

Cualquier algoritmo $(1, \lambda)$ ó $(1 + \lambda)$ que se desarrolle deberá contar con un método que controle el tamaño de paso y lo ajuste siempre a su valor óptimo.

A continuación, en el párrafo 5.4.5.1, se estudia el tamaño de paso óptimo y la forma de controlarlo para el algoritmo $(1 + 1)$. Y después, en el párrafo 5.4.5.2 se generalizan estos resultados para los algoritmos $(1, \lambda)$ y $(1 + \lambda)$.

5.4.5.1 Tamaño de paso óptimo y control del tamaño de paso para el algoritmo (1+1)

Según la Ref. 9, para obtener el tamaño de paso óptimo es necesario derivar con respecto a σ las expresiones 5.4.3.1/7 y 5.4.3.2/7 que proporcionan la velocidad de convergencia, Φ , para los modelos de corredor y esférico, respectivamente. Es decir:

$$\frac{d\Phi}{d\sigma} = 0 \quad 5.4.5.1/1$$

a) Modelo de corredor

Derivando la expresión 5.4.3.1/7 se obtiene:

$$\sigma_{\text{opt}} = \frac{b}{n} \sqrt{2\pi} \quad 5.4.5.1/2$$

$$\Phi_{\text{max}} = \frac{b}{n} \left(1 - \frac{1}{n}\right)^{n-1} \quad 5.4.5.1/3$$

Estas dos fórmulas son difícilmente aplicables en la práctica, pues contienen la semianchura de corredor, b . Sin embargo, puede sustituirse el valor de σ_{opt} en la expresión 5.4.3.1/4 y así se obtiene la probabilidad de éxito para velocidad de convergencia máxima. Esto es:

$$P_{1\text{opt}} = \frac{1}{2} \left(1 - \frac{1}{n}\right)^{n-1} \quad 5.4.5.1/4$$

Si $n \gg 1$:

$$P_{1\text{opt}} = \frac{1}{2e} = 0.184 \quad (n \gg 1, \sigma/b \ll 1) \quad 5.4.5.1/5$$

b) Modelo esférico

Derivando la expresión 5.4.3.2/7 se obtiene:

$$\sigma_{\text{opt}} = \frac{r_o}{n} 1.224 \quad 5.4.5.1/6$$

$$\phi_{\text{max}} = \frac{r_o}{n} 0.202 \quad 5.4.5.1/7$$

Al igual que las fórmulas deducidas para el modelo de corredor, estas expresiones tienen difícil aplicación práctica, pues la distancia desde el padre hasta el mínimo está contenida en ellas. Sustituyendo el valor de σ_{opt} en la expresión 5.4.3.2/5 se obtiene:

$$P_{1\text{opt}} = 0.270 \quad (n \gg 1, \quad n\sigma^2/r_o^2 \ll 1) \quad 5.4.5.1/8$$

Por tanto, los resultados obtenidos han sido:

a) Modelo de corredor: $P_{1\text{opt}} = 0.184$

b) Modelo esférico: $P_{1\text{opt}} = 0.270$

Es decir, la regla práctica para adaptar el tamaño de paso a su valor óptimo puede ser la siguiente:

Se sabe que si se selecciona σ_{opt} , se obtendrá 1 éxito de cada 5.4 intentos cuando la función objetivo estudiada se aproxime al modelo de corredor y 1 éxito de cada 3.7 intentos cuando la función objetivo se ajuste al modelo esférico.

Por ello, para cualquier función objetivo puede tomarse

1/5 como la frecuencia relativa de éxito óptima, Fr_{opt} . Si durante la optimización $Fr < 0.20$, entonces σ debe ser reducida y si $Fr > 0.20$, σ debe ser incrementada (recuérdese que $dP_1/d\sigma < 0$).

5.4.5.2 Tamaño de paso óptimo y control del tamaño de paso para los algoritmos $(1,\lambda)$ y $(1+\lambda)$

Se determinan en este párrafo los valores de Φ^{max} , $\Phi^{\text{max}}/\lambda$, σ_{opt}^* y $P_{\lambda\text{opt}}$ en función de λ para los modelos de corredor y esférico y para los algoritmos $(1,\lambda)$ y $(1+\lambda)$.

a) Algoritmo $(1,\lambda)$

Se opera de la forma siguiente:

- 1) Fijar λ .
- 2) Con λ , obtener Φ^{max} mediante las figuras 5.4.4.1/3 y 5.4.4.2/3 para los modelos de corredor y esférico respectivamente.
- 3) Con λ y Φ^{max} obtener $(\Phi^{\text{max}}/\lambda)$.
- 4) Con λ y Φ^{max} obtener σ_{opt}^* mediante las figuras 5.4.4.1/2 y 5.4.4.2/2 para los modelos de corredor y esférico respectivamente.
- 5) Con λ , σ_{opt}^* y n , obtener $P_{1\text{opt}}$ mediante la expresión 5.4.3.1/4 y $P_{\lambda\text{opt}}$ mediante la expresión 5.4.4.1/1, para el modelo de corredor.

Con λ y σ_{opt}^* obtener $P_{1\text{opt}}$ mediante la expresión 5.4.3.2/5 y $P_{\lambda\text{opt}}$ mediante la expresión 5.4.4.2/1 para el modelo esférico.

Los resultados obtenidos mediante el proceso descrito anteriormente se han recogido en la tabla 5.4.5.2/1 y en las figuras 5.4.5.2/1 y 2.

b) Algoritmo $(1+\lambda)$

Se opera de la forma siguiente:

- 1) Fijar λ .
- 2) Con λ , obtener $\Phi_{\max}^*$ y σ_{opt}^* mediante los resultados de simulación de las figuras 5.4.4.1/2 y 5.4.4.2/2 para los modelos de corredor y esférico respectivamente.
- 3) Con λ y $\Phi_{\max}^*$ obtener $(\Phi_{\max}^*/\lambda)$.
- 4) Con λ , σ_{opt}^* y n , obtener $P_{1\text{opt}}$ mediante la expresión 5.4.3.1/4 y $P_{\lambda\text{opt}}$ mediante la expresión 5.4.4.1/1, para el modelo de corredor.
Con λ y σ_{opt}^* , obtener $P_{1\text{opt}}$ mediante la expresión 5.4.3.2/5 y $P_{\lambda\text{opt}}$ mediante la expresión 5.4.4.2/1 para el modelo esférico.

Los resultados obtenidos se han recogido en la tabla 5.4.5.2/2 y en las figuras 5.4.5.2/1 y 2.

Para controlar el tamaño de paso en los algoritmos $(1,\lambda)$ y $(1+\lambda)$ y adaptarlo a su valor óptimo, se utiliza el mismo método ya descrito para el algoritmo $(1+1)$. La tabla 5.4.5.2/3 muestra las frecuencias relativas de éxito óptimas, Fr_{opt} , las cuales han sido obtenidas a partir de las probabilidades de éxito óptimas para el modelo de corredor y el modelo esférico.

La forma de controlar el tamaño de paso es la siguiente: si durante el proceso de optimización de cualquier función objetivo, $Fr < Fr_{\text{opt}}$, entonces σ debe ser reducida, y si $Fr > Fr_{\text{opt}}$ σ debe ser incrementada.

TABLA 5.4.5.2/1

ALGORITMO (1,λ) . RESULTADOS OBTENIDOS

λ	Modelo de Corredor					Modelo Esférico			
	ϕ* _{max}	$\frac{\phi^*_{max}}{\lambda}$	σ* _{opt}	P _{λopt}		ϕ* _{max}	$\frac{\phi^*_{max}}{\lambda}$	σ* _{opt}	P _{λopt}
				n→∞	n=10				
1	0	0	0	0.50	0.50	0	0	0	0.50
2	0.31	0.16	1.44	0.49	0.50	0.17	0.09	0.49	0.65
3	0.62	0.21	1.77	0.57	0.59	0.35	0.12	0.73	0.73
4	0.92	0.23	2.07	0.63	0.65	0.52	0.13	0.91	0.79
5	1.20	0.24	2.28	0.68	0.70	0.65	0.13	1.02	0.84
6	1.44	0.24	2.49	0.71	0.73	0.76	0.13	1.10	0.87
7	1.66	0.24	2.66	0.74	0.76	0.87	0.12	1.18	0.90
8	1.88	0.24	2.82	0.76	0.78	0.96	0.12	1.24	0.92
9	2.10	0.23	2.99	0.77	0.79	1.05	0.12	1.31	0.93
10	2.32	0.23	3.15	0.78	0.80	1.13	0.11	1.37	0.94
	$\phi^* = \frac{\phi n}{b} \quad , \quad \sigma^* = \frac{\sigma n}{b}$					$\phi^* = \frac{\phi n}{r_o} \quad , \quad \sigma^* = \frac{\sigma n}{r_o}$			

TABLA 5.4.5.2/2

ALGORITMO (1+λ) . RESULTADOS OBTENIDOS

λ	Modelo de Corredor					Modelo Esférico			
	ϕ [*] _{max}	$\frac{\phi^*_{\max}}{\lambda}$	σ [*] _{opt}	P _{λopt}		ϕ [*] _{max}	$\frac{\phi^*_{\max}}{\lambda}$	σ [*] _{opt}	P _{λopt}
				n→∞	n=10				
1	0.37	0.37	2.51	0.18	0.19	0.20	0.20	1.22	0.27
2	0.69	0.35	2.67	0.32	0.33	0.37	0.19	1.29	0.45
3	0.99	0.33	2.82	0.41	0.43	0.52	0.17	1.36	0.58
4	1.27	0.32	2.97	0.49	0.50	0.65	0.16	1.42	0.66
5	1.52	0.30	3.10	0.54	0.56	0.77	0.15	1.48	0.73
6	1.76	0.29	3.21	0.59	0.63	0.88	0.15	1.52	0.78
7	1.98	0.28	3.30	0.64	0.65	0.98	0.14	1.55	0.82
8	2.18	0.27	3.36	0.67	0.69	1.07	0.13	1.57	0.86
9	2.37	0.26	3.41	0.71	0.73	1.15	0.13	1.59	0.88
10	2.55	0.26	3.46	0.74	0.76	1.22	0.12	1.60	0.91
	$\phi^* = \frac{\phi n}{b} \quad , \quad \sigma^* = \frac{\sigma n}{b}$					$\phi^* = \frac{\phi n}{r_o} \quad , \quad \sigma^* = \frac{\sigma n}{r_o}$			

TABLA 5.4.5.2/3

FRECUENCIA RELATIVA DE ÉXITO ÓPTIMA PARA EL CONTROL DEL TAMAÑO DE PASO DE LOS ALGORITMOS $(1,\lambda)$ Y $(1+\lambda)$

λ	F_{ropt}	
	ALGORITMO $(1,\lambda)$	ALGORITMO $(1+\lambda)$
1	—	0.20
2	0.60	0.40
3	0.65	0.50
4	0.70	0.60
5	0.75	0.65
6	0.80	0.70
7	0.80	0.75
8	0.85	0.80
9	0.85	0.80
10	0.85	0.85

FIGURA 5.4.5.2/1

PROBABILIDAD DE EXITO OPTIMA. MODELO DE CORREDOR

FIGURA 5.4.5.2/2
PROBABILIDAD DE EXITO OPTIMA. MODELO ESFERICO

5.4.6 Conclusiones

Las conclusiones originales que se han obtenido en esta parte de la Tesis son:

- 1) Para los algoritmos $(1,\lambda)$ y $(1+\lambda)$, el tamaño de paso puede controlarse y adaptarse a su valor óptimo mediante las frecuencias relativas de éxito óptimas que se recogen en la tabla 5.4.3.2/3.
- 2) Entre los posibles algoritmos $(1,\lambda)$, los mejores son los correspondientes a $\lambda=5$ y $\lambda=6$.
- 3) Entre los posibles algoritmos $(1+\lambda)$ el mejor de ellos es el correspondiente a $\lambda = 1$.
- 4) Entre todos los posibles algoritmos $(1,\lambda)$ y $(1+\lambda)$, el mejor de ellos es el algoritmo $(1+1)$. Parece ser que la sencillez de los programas de ordenador relacionados con estrategias naturales les proporciona gran potencia y robustez, ya que si el algoritmo $(1+1)$ ha demostrado ser el mejor de todos los estudiados para resolver problemas de optimización, también el algoritmo TIT FOR TAT (Ref. 10), cuya simplicidad es casi insultante, proporciona resultados extraordinarios en los problemas de evolución de la conducta cooperativa.

5.5 MEJORA DE LOS ALGORITMOS GENÉTICOS

En esta Tesis se han analizado sucíntamente los algoritmos (μ, λ) y $(\mu + \lambda)$ y, exhaustivamente, sus casos particulares para $\mu = 1$.

Estos algoritmos están basados en el modelo altamente simplificado de la Evolución, recogido en el apartado 5.2, que sólo incluye el principio de mutación (de genes)-selección.

Sin embargo, según la teoría moderna de la Evolución, el proceso evolutivo se basa fundamentalmente en los siguientes mecanismos:

1. Mutación de genes
2. Mutación de cromosomas
3. Cruce o recombinación
4. Selección
5. Aislamiento

Los factores 1 a 3 son responsables de la generación de nuevos puntos en el espacio de nucleótidos. La selección elimina aquellos puntos que corresponden a regiones con vitalidad asociada baja. El mecanismo de aislamiento mantiene separados los puntos del espacio de nucleótidos que forman "canales de éxito" distintos.

Dentro del factor 2 (mutación de cromosomas) se pueden distinguir los casos siguientes:

- Duplicación
- Delección
- Translocación
- Inversión

De estos casos, sólo se considera importante para un primer intento de mejora de los algoritmos genéticos el mecanismo de inversión.

El cruce o la recombinación consiste en un intercambio de secciones entre dos cromosomas.

La introducción del cruce en el modelo de Evolución estudiado debe proporcionar buenos resultados para la evolución, como se desprende del hecho de que únicamente los organismos muy simples renuncian a él. Además, existen dos importantes mecanismos de herencia, los cuales utilizan el cruce: la herencia diploide y la herencia haploide.

Los seres vivientes más avanzados tienen un mecanismo de herencia diploide, por lo que la implementación de este mecanismo en un algoritmo genético parece más prometedora que la de la herencia haploide.

Se pueden utilizar las analogías del apartado 5.2 para implementar el cruce y la inversión en los algoritmos genéticos analizados anteriormente.

a) Cruce

Se define el Operador Cruce Elemental, c_i , como:

$$c_i: \{S^n \times S^n\} \rightarrow \{S^n \times S^n\} ,$$

$$\begin{aligned} c_i\{x^1, x^2\} &= c_i\{(x^1_1, \dots, x^1_n)^T, (x^2_1, \dots, x^2_n)^T\} = \\ &= (x^1_1, \dots, x^1_{i-1}, x^2_i, x^1_{i+1}, \dots, x^1_n)^T, (x^2_1, \dots, x^2_{i-1}, x^1_i, x^2_{i+1}, \dots, x^2_n)^T \\ &= y^1, y^2 \end{aligned} \quad 5.5/1$$

El algoritmo de generación de nuevos puntos mediante este Operador Cruce Elemental, puede ser como sigue:

Etapas 1: Elegir un punto x^1 dentro de la población considerada, de acuerdo con determinada probabilidad, función de su vitalidad (valor asociado de la función objetivo).

Etapas 2: Elegir un punto x^2 de la misma forma que se ha elegido x^1 .

Etapas 3: Elegir una componente i , entre 1 y n , con función de densidad uniforme.

Etapas 4: Mediante la expresión 5.5/1, obtener dos nuevos puntos, y^1, y^2 .

Etapas 5: Eliminar de la población el punto o los puntos con peor valor de la función objetivo.

b) Inversión

Se define la estructura de inversión como una permutación de la secuencia $1, 2, \dots, n$ representada por $i_1, i_2, \dots, i_n$, o más abreviadamente por i , de tal forma que:

$$i: \{S^n\} \rightarrow \{S^n\} ,$$

$$i\{x^1\} = y^1, \text{ donde } y_k^1 = x_k^1 i_k \quad k = 1(1)n \quad 5.5/2$$

El algoritmo de generación de un nuevo punto a partir de otro, puede ser como sigue:

Etapa 1: Elegir un punto x^1 dentro de la población considerada, de acuerdo con determinada probabilidad, función de su vitalidad (valor asociado de la función objetivo).

Etapa 2: Elegir con función de densidad uniforme, una estructura de inversión, i , entre las $n!$ posibles.

Etapa 3: Mediante la expresión 5.5/2, obtener un nuevo punto y^1 .

Etapa 4: Eliminar de la población el punto con peor valor de la función objetivo.

Tanto el cruce como la inversión pueden contribuir a mejorar aquellos algoritmos genéticos que utilicen solamente el mecanismo de mutación-selección. Sin embargo, la mutación será siempre la regla genética más importante, pues produce puntos totalmente nuevos en la población, mientras que el cruce y la inversión generan únicamente un número finito de variantes.

Para terminar, hay que decir que es necesaria una investigación teórica más profunda en el área.

6. DISEÑO DE PROGRAMAS PARA LOS ALGORITMOS GENÉTICOS $(1,\lambda)$ y $(1+\lambda)$

Se han desarrollado dos subrutinas FORTRAN, una para los algoritmos $(1,\lambda)$ y otra para los $(1+\lambda)$, las cuales han sido implementadas en el ordenador PDP 11/60 de la Cátedra de Mecánica del Vuelo. La base de las subrutinas mencionadas se encuentra en el programa EVOL de Schwefel (Ref. 6).

Las características más sobresalientes de estas subrutinas se indican a continuación.

Objeto:

- Búsqueda del mínimo de $f(x)$, $x^* \in \mathbb{R}^n$

sin utilizar $\frac{\partial f(x)}{\partial x_i}$ $i = 1(1)n$ 6/1

- Pueden tenerse en cuenta ligaduras de desigualdad.

$g_j(x) \geq 0$ $j = 1(1)m$. 6/2

Entradas:

λ : número de hijos

n : número de variables o dimensión del espacio

m : número de restricciones de desigualdad

T : tiempo de cálculo máximo permitido

PC : parámetro de control del tamaño de paso. La frecuencia relativa de éxito óptima está directamente relacionada con PC , mediante:

$$F_{ropt} = \frac{1}{5 \cdot PC} \quad 6/3$$

CA: parámetro auxiliar para comprobar la convergencia del algoritmo. Los criterios de convergencia se comprueban en las $10 \cdot n \cdot PC \cdot CA$ últimas búsquedas.

ε_a : límite inferior absoluto para el tamaño de paso.

ε_b : límite inferior, relativo al valor de las variables, para el tamaño de paso.

ε_c : criterio de convergencia absoluto.

$$f_2 - f_1 \leq \varepsilon_c \quad 6/4$$

$$f_1 = \min_{10n \cdot PC \cdot CA \text{ últimas búsquedas}} f(x) \quad 6/5$$

$$f_2 = \max_{10n \cdot PC \cdot CA \text{ últimas búsquedas}} f(x) \quad 6/6$$

ε_d : criterio de convergencia relativo

$$f_2 - f_1 \leq \frac{\varepsilon_d}{10n \cdot PC \cdot CA} \left| \sum_{10n \cdot PC \cdot CA \text{ últimas búsquedas}} f(x) \right| \quad 6/7$$

donde f_1 y f_2 vienen dadas por las expresiones 6/5 y 6/6.

SC: magnitud auxiliar para reducción o aumento del tamaño de paso.

x^o : punto inicial.

σ^o : desviación típica inicial (las componentes de este vector pueden ser distintas).

Salidas:

$f(x^*)$: valor mínimo de la función objetivo.

x^* : punto para el que la función objetivo es mínima.

Observaciones:

- La determinación de mínimos en bordes y esquinas es lenta y, a veces, poco precisa.
- No se garantiza la convergencia global.

7. PRUEBA Y AJUSTE DE LOS PROGRAMAS DISEÑADOS

7.1 INTRODUCCIÓN

En este capítulo se comentan las pruebas y posteriores ajustes de los programas de ordenador diseñados para los algoritmos $(1, \lambda)$ y $(1 + \lambda)$.

En el apartado 7.2 se recopilan una serie de funciones objetivo de prueba que se han encontrado al realizar una revisión exhaustiva de la literatura especializada en el tema, algunas de las cuales pueden utilizarse para probar y ajustar los programas.

Y el apartado 7.3 presenta algunos resultados interesantes que se han obtenido al efectuar estas pruebas y ajustes.

7.2 FUNCIONES OBJETIVO DE PRUEBA

7.2.1 Generalidades

Este apartado presenta algunas funciones objetivo de prueba, frecuentemente utilizadas en la literatura revisada sobre la solución de los problemas de optimización.

Las funciones objetivo de prueba han sido clasificadas en cinco grupos:

- a) Funciones Objetivo de Prueba para Problemas de Optimización sin Ligaduras.
- b) Funciones Objetivo de Prueba para Problemas de Optimización con Ligaduras.
- c) Funciones Objetivo de Prueba para Problemas Globales de Optimización.
- d) Funciones Objetivo de Prueba que aparecen en Problemas Técnicos Típicos.
- e) Funciones Objetivo de Prueba con Analogía Biológica.

Estas funciones objetivo de prueba se han obtenido de las Refs. 3 , 4 , 6 , 8 , 11 , 12 y 13.

7.2.2 Funciones Objetivo de Prueba para Problemas de Optimización sin Ligaduras

1) Función Esférica

$$f(x) = \sum_{i=1}^n x_i^2 \quad 7.2.2/1$$

$$\text{Mínimo: } x^* = 0, \quad f(x^*) = 0$$

2) Función de Índices

$$f(x) = \sum_{i=1}^n i x_i^2 \quad 7.2.2/2$$

$$\text{Mínimo: } x^* = 0, \quad f(x^*) = 0$$

3) Función de Índices Cuadrados

$$f(x) = \sum_{i=1}^n i^2 x_i^2 \quad 7.2.2/3$$

$$\text{Mínimo: } x^* = 0, \quad f(x^*) = 0$$

4) Función de Expansión de Coordenadas (fig. 7.2.2/1)

$$f(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2 \quad 7.2.2/4$$

$$\text{Mínimo: } x^* = 0, \quad f(x^*) = 0$$

5) Función de Wood (Generalización de la Función de Rosenbrock, fig. 7.2.2/2)

$$f(\mathbf{x}) = 100(\mathbf{x}_2 - \mathbf{x}_1^2)^2 + (1 - \mathbf{x}_1)^2 + 90(\mathbf{x}_4 - \mathbf{x}_3^2)^2 + (1 - \mathbf{x}_3)^2 + \\ + 10.1[(\mathbf{x}_2 - 1)^2 + (\mathbf{x}_4 - 1)^2] + 19.8(\mathbf{x}_2 - 1)(\mathbf{x}_4 - 1) \quad 7.2.2/5$$

$$\text{Mínimo: } \mathbf{x}^* = (1, 1, 1, 1)^T, \quad f(\mathbf{x}^*) = 0$$

7.2.3 Funciones Objetivo de Prueba para Problemas de Optimización con Ligaduras

1) Función de Chichinadze

$$f(x) = x_1^2 - 12x_1 + 11 + 10\cos\frac{\pi}{2}x_1 + 8\sin\pi x_1 - \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x_2 - 0.5}{2}\right] \quad 7.2.3/1$$

$$\begin{aligned} 0 &\leq x_1 \leq 7 \\ 0 &\leq x_2 \leq 5 \end{aligned} \quad 7.2.3/2$$

Mínimo: $x^* = (5.62, 0.50)^T$, $f(x^*) = -40.964$.

2) Función de Bracken y McCormick (fig 7.2.3/1)

$$f(x) = -x_1^2 - x_2^2 \quad 7.2.3/3$$

$$x_i \geq 0 \quad i = 1, 2$$

$$-x_1 - 4x_2 + 5 \geq 0 \quad 7.2.3/4$$

$$-x_1 + 1 \geq 0$$

Mínimo: $x^* = (1, 1)^T$, $f(x^*) = -2$

3) Función de Johnson

$$f(x) = x_1 x_2 x_3 x_4 \quad 7.2.3/5$$

$$x_1^3 + x_2 - 1 = 0$$

$$x_1^2 x_4 - x_3 = 0 \quad 7.2.3/6$$

$$x_4^2 - x_2 = 0$$

Mínimo: $f(x^*) = -0.25$

4) Función de Danskin

$$\min_y \max_x f$$

7.2.3/7

$$f(x,y) = \sum_{i=1}^2 (1 - e^{-10x_i} e^{-y_i})$$

7.2.3/8

$$x_i, y_i \geq 0 \quad i = 1, 2$$

$$x_1 + x_2 = 1$$

7.2.3/9

$$y_1 + y_2 = 1$$

Mínimo: $f^* = 1.88$

7.2.4 Funciones Objetivo de Prueba para Problemas Globales de Optimización

1) Función de Shekel

$$f(x) = - \sum_{i=1}^l \frac{1}{(x-a_i)^T (x-a_i) + c_i} \quad 7.2.4/1$$

$$x = (x_1, \dots, x_n)^T, \quad a_i = (a_{i1}, \dots, a_{in})^T, \quad c_i > 0 \quad i = 1(1)l$$

$$0 \leq x_j \leq 10 \quad j = 1(1)n \quad 7.2.4/2$$

Esta función objetivo de prueba tiene l mínimos cuya situación es a_i y cuyos valores dependen de c_i .

2) Función de Hartman

$$f(x) = - \sum_{i=1}^l c_i \exp\left(- \sum_{j=1}^n \alpha_{ij} (x_j - p_{ij})^2\right) \quad 7.2.4/3$$

$$x = (x_1, \dots, x_n)^T, \quad p_i = (p_{i1}, \dots, p_{in})^T; \quad \alpha_i = (\alpha_{i1}, \dots, \alpha_{in})^T$$

$$c_i > 0, \quad i = 1(1)l$$

$$0 \leq x_j \leq 1 \quad j = 1(1)n \quad 7.2.4/4$$

Esta función objetivo de prueba tiene l mínimos cuya situación es aproximadamente p_i y cuyos valores dependen de c_i .

3) Función de Branin

$$f(x) = a(x_2 - bx_1^2 + cx_1 - d)^2 + e(1-f)\cos x_1 + e \quad 7.2.4/5$$

$$a = 1, \quad b = 5.1 / (4\pi)^2, \quad c = 5/\pi, \quad d = 6, \quad e = 10, \quad f = 1 / (8\pi)$$

$$-5 \leq x_1 \leq 10$$

7.2.4/6

$$0 \leq x_2 \leq 15$$

En esta región hay tres mínimos, todos ellos globales.

4) Función de Goldstein y Price

$$f(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \cdot \\ \cdot [30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 - 48x_2 - 36x_1x_2 + 27x_2^2)]$$

7.2.4/7

$$-2 \leq x_i \leq 2 \quad i = 1, 2$$

7.2.4/8

Hay cuatro mínimos locales. El mínimo global se da para

$$x^* = (0, -1)^T, \quad f(x^*) = 3.$$

5) Función de Hosaki

$$f(x) = (1 - 8x_1 + \frac{7}{3}x_1^2 - \frac{7}{3}x_1^3 + \frac{1}{4}x_1^4)x_2^2e^{-x_2} \quad 7.2.4/9$$

$$0 \leq x_1 \leq 5$$

7.2.4/10

$$0 \leq x_2 \leq 6$$

$$\text{Hay dos mínimos: } x = (1, 2)^T, \quad f(x) = -1.127 \quad (\text{local})$$

$$x = (4, 2)^T, \quad f(x) = -2.345 \quad (\text{global})$$

7.2.5 Funciones Objetivo de Prueba que aparecen en Problemas Técnicos Típicos

1) Problema de Diseño de un Transformador

$$f(x) = 0.0204x_1x_4(x_1+x_2+x_3) + 0.0187x_2x_3(x_1+1.57x_2+x_4) + \\ + 0.0607x_1x_4x_5^2(x_1+x_2+x_3) + 0.0437x_2x_3x_6^2(x_1+1.57x_2+x_4)$$

7.2.5/1

$$x_i \geq 0 \quad i = 1(1)6$$

$$x_1x_2x_3x_4x_5x_6 \geq 2.07 \times 10^3 \quad 7.2.5/2$$

$$1 - 0.00062x_1x_4x_5^2(x_1+x_2+x_3) - 0.00058x_2x_3x_6^2(x_1+1.57x_2+x_4) \geq 0$$

2) Problema de Planificación de Potencia Estática

$$f(x) = 3000x_1 + 1000x_1^3 + 2000x_2 + 666.667x_2^3 \quad 7.2.5/3$$

$$0.4 - x_1 + 2Cx_5^2 + x_5x_6[Dsen(-x_8) - Ccos(-x_8)] + \\ + x_5x_7[Dsen(-x_9) - Ccos(-x_9)] = 0$$

$$0.4 - x_2 + 2Cx_6^2 + x_6x_5[Dsen(x_8) - Ccos(x_8)] + \\ + x_6x_7[Dsen(x_8 - x_9) - Ccos(x_8 - x_9)] = 0$$

$$0.8 + 2Cx_7^2 + x_7x_5[Dsen(x_9) - Ccos(x_9)] + \\ + x_7x_6[Dsen(x_9 - x_8) - Ccos(x_9 - x_8)] = 0$$

$$0.2 - x_3 + 2Dx_5^2 + x_5x_6[csen(-x_8) + Dcos(-x_8)] - \\ - x_5x_7[Csin(-x_9) + Dcos(-x_9)] = 0 \quad 7.2.5/4$$

$$0.2 - x_4 + 2Dx_6^2 - x_6x_5 [C\text{sen}(x_8) + D\cos(x_8)] - \\ - x_6x_7 [C\text{sen}(x_8 - x_9) + D\cos(x_8 - x_9)] = 0$$

$$-0.337 + 2Dx_1^2 - x_7x_5 [C\text{sen}(x_9) + D\cos(x_9)] - \\ - x_7x_6 [C\text{sen}(x_9 - x_8) + D\cos(x_9 - x_8)] = 0$$

$$1.0909 \geq x_i \geq 0.90909 \quad i = 5, 6, 7$$

donde:

$$C = \text{sen}(0.25)48.4/50.176$$

$$D = \cos(0.25)48.4/50.176$$

3) Problema de Modelado de un Transistor

$$f(x) = \gamma^2 + \sum_{k=1}^4 (\alpha_k^2 + \beta_k^2) \quad 7.2.5/5$$

donde

$$\alpha_k = (1 - x_1x_2)x_3 \{ \exp [x_5 (g_{1k} - g_{3k}x_7 \times 10^{-3} g_{5k}x_8 \times 10^{-3})] - 1 \} - \\ - g_{5k} + g_{4k}x_2$$

$$\beta_k = (1 - x_1x_2)x_4 \{ \exp [x_6 (g_{1k} - g_{2k} - g_{3k}x_7 \times 10^{-3} + g_{4k}x_9 \times 10^{-3})] - 1 \} \\ - g_{5k}x_1 + g_{4k}$$

y las constantes numéricas g_{ik} vienen dadas por la matriz:

$$\begin{bmatrix} 0.485 & 0.752 & 0.869 & 0.982 \\ 0.369 & 1.254 & 0.703 & 1.455 \\ 5.2095 & 10.0677 & 22.9274 & 20.2153 \\ 23.3037 & 101.779 & 111.461 & 191.267 \\ 28.5132 & 111.8467 & 134.3884 & 211.4823 \end{bmatrix}$$

$$x_i \geq 0 \quad i = 1, \dots, 9$$

7.2.5/6

7.2.6 Funciones Objetivo de Prueba con Analogía Biológica

1) Evolución Mimética

$$f(x) = ||x^* - x||_H \quad 7.2.6/1$$

$$x = (x_1, \dots, x_n)^T, \quad x_i \in \{0, 1\} \quad i = 1(1)n$$

$||\cdot||_H$: Distancia de Hamming, definida como el número de elementos x_i que son diferentes entre los puntos x^* y x .

Mínimo: $x = x^*, f(x^*) = 0$

Ejemplo: el problema anteriormente expuesto puede ser aplicado a la Evolución Mimética de la forma siguiente:

$n = 81$

Punto inicial: $x^0 = 0$

Este punto representa una mariposa con alas blancas (fig 7.2.6/1a). Esta mariposa agrada a los pájaros.

Mínimo: $x^* = (110000001110001001110000001110011001110011001110000001110011001110000001111111111)^T$

El ala de la mariposa se divide en 81 cuadrados y cada cuadrado se dibuja de negro si $x_i=1$ y se deja blanco si $x_i=0$ (fig. 7.2.6/1b) Esta mariposa no agrada a los pájaros.

El número de etapas necesarias para ir desde x^0 hasta x^* , cuando se aplica un método de optimización basado en los algoritmos genéticos, da una idea acerca del tiempo requerido por la evolución natural para obtener la segunda mariposa a partir de la primera.

FIGURA 7.2.2/1

FUNCIÓN DE EXPANSIÓN DE COORDENADAS: $f(x) = x_1^2 + (x_1+x_2)^2$.

FIGURA 7.2.2/2

FUNCIÓN DE ROSENBROCK: $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$.

FIGURA 7.2.3/1

FUNCIÓN DE BRACKEN Y MACCORMICK: $f(x) = -x_1^2 - x_2^2$

$$x_i \geq 0, i = 1, 2$$

$$-x_1 - 4x_2 + 5 \geq 0$$

$$-x_1 + 1 \geq 0$$

DOS MARIPOSAS

(A)

LES GUSTA A LOS PAJAROS

$X^0 = 0$

(B)

NO LES GUSTA A LOS PAJAROS

$X^1 = (110000001110001001110$
 $000001110011001110011001110$
 00000111001100111000000
 $111111111)$

FIGURA 7.2.6/1

EVOLUCION MIMETICA

7.3 RESULTADOS OBTENIDOS

Los algoritmos diseñados han sido probados y ajustados con varias funciones objetivo, demostrando su buen comportamiento.

Una de las funciones más utilizadas ha sido la función de Shekel, la cual presenta características tan interesantes como la posibilidad de que el usuario la diseñe incluso desde el punto de vista de multimodalidad.

En todos los casos analizados se corroboran, en general, las conclusiones obtenidas en el subapartado 5.4.6. Por ejemplo, para el caso en que la función objetivo sea la función de Shekel, con

$$\begin{array}{l|l} n = 10 & a_{1j} = 5.0, \quad j = 1(1)10 \\ l = 1 & C_1 = 0.025 \end{array}$$

y sin ligaduras de desigualdad, se obtienen los resultados que presentan las tablas 7.3/1 y 7.3/2.

TABLA 7.3/1

RESULTADOS OBTENIDOS CON EL PROGRAMA $(1+\lambda)$ Entradas:

$n = 10$, $m = 0$, $T = 900.0s$, $CA = 3$, $\varepsilon_a = \varepsilon_b = \varepsilon_c = \varepsilon_d = 10^{-5}$,
 $SC = 0.85$, $x_i^\circ = 0.0$ ($i = 1(1)9$), $x_{10}^\circ = 10.0$, $\sigma_i^\circ = 0.5$ ($i = 1(1)10$)

Salidas:

$f(x^*) = -39.99999$, $x_i^* = 5.0000$ ($i = 1(1)10$)

λ	PA	TIEMPO EN ALCANZAR EL MÍNIMO (s)	NUMERO DE EVALUACIONES DE $f(x)$
1	1.000	12.48	1201
2	0.500	12.76	1201
3	0.400	14.82	1441
4	0.333	19.40	1921
5	0.308	24.22	2401
6	0.286	21.74	2161
7	0.267	25.32	2521
8	0.250	28.70	2881
9	0.250	32.30	3241
10	0.235	34.52	3601

TABLA 7.3/2

RESULTADOS OBTENIDOS CON EL PROGRAMA (1,λ)

Entradas:

$n = 10, m = 0, T = 900.0s, CA = 3, \varepsilon_a = \varepsilon_b = \varepsilon_c = \varepsilon_d = 10^{-5},$
 $SC = 0.85, x_i^o = 0.0 \text{ (i=1(1)9)}, x_{10}^o = 10.0, \sigma_i^o = 0.5 \text{ (i=1(1)10)}$

Salidas:

$f(x^*) = -39.9999, x_i^* = 5.0000 \text{ (i = 1(1)10)}$

λ	PA	TIEMPO EN ALCANZAR EL MÍNIMO (s)	NUMERO DE EVALUACIONES DE f(x)
2	0.333	*	*
3	0.308	*	*
4	0.286	18.68	1801
5	0.267	22.94	2251
6	0.250	21.84	2161
7	0.250	31.66	3151
8	0.235	28.78	2881
9	0.235	32.40	3241
10	0.235	34.60	3601

* El mínimo no se ha alcanzado. El algoritmo se ha detenido en un punto no válido.

8. APLICACIÓN DE LOS ALGORITMOS GENÉTICOS PARA LA RESOLUCIÓN DE UN PROBLEMA DE OPTIMIZACION EN MECÁNICA DEL VUELO

8.1 INTRODUCCIÓN

Una vez diseñados, probados y ajustados los programas de ordenador cuya base fundamental son los Algoritmos Genéticos, se pasa a la resolución, mediante ellos, de un problema típico de Mecánica del Vuelo.

El apartado 8.2 propone el problema a estudiar y ofrece su solución teórica que, por fortuna, existe en este caso.

Y el apartado 8.3 muestra la solución obtenida mediante los programas de ordenador disponibles para unos valores concretos de los parámetros del problema teórico formulado en el apartado anterior.

8.2 EL PROBLEMA ESTUDIADO. RESOLUCIÓN TEÓRICA

Enunciado:

Se considera un misil no sustentador, cuya resistencia es proporcional al cuadrado de la velocidad y cuyo empuje es proporcional al gasto másico de combustible del motor. Este gasto másico puede variar entre cero y un valor máximo. Se supone, además, que el ángulo de asiento de la velocidad es muy pequeño y que la componente del peso del misil según la tangente a la trayectoria es despreciable. Se pretende determinar la ley de control que maximice el alcance del misil, para unos valores prefijados de las masas y de las velocidades iniciales y finales.

Formulación

- Hipótesis previas

$$L = 0 \quad 8.2/1$$

$$D = kV^2 \quad 8.2/2$$

$$T = I_{sp} \beta \quad 8.2/3$$

$$0 \leq \beta \leq \beta_{\max} \quad 8.2/4$$

$$\gamma \ll 1 \quad 8.2/5$$

$$W\gamma \ll T \quad 8.2/6$$

- Ecuaciones del movimiento en un plano vertical

$$\dot{x} - V = 0 \quad 8.2/7$$

$$\dot{h} - V\gamma = 0 \quad 8.2/8$$

$$\dot{V} + \frac{kV^2 - I_{sp}\beta}{m} = 0 \quad 8.2/9$$

$$\dot{\gamma} + \frac{g}{V} = 0 \quad 8.2/10$$

$$\dot{m} + \beta = 0 \quad 8.2/11$$

Si se utiliza la masa como variable independiente

$$\frac{dx}{dm} = - \frac{V}{\beta} \quad 8.2/12$$

$$\frac{dh}{dm} = - \frac{V\gamma}{\beta} \quad 8.2/13$$

$$\frac{dv}{dm} = \frac{kv^2}{m\beta} - \frac{I_{sp}}{m} \quad 8.2/14$$

$$\frac{d\gamma}{dm} = \frac{g}{V\beta} \quad 8.2/15$$

$$\frac{dt}{dm} = - \frac{1}{\beta} \quad 8.2/16$$

Variables dependientes $\equiv x, h, V, \gamma, t, \beta$

Variable independiente $\equiv m$

Ecuaciones $\equiv 5$

Grados de libertad $\equiv 1$ ($0 \leq \beta \leq \beta_{max}$)

- Región de desplazamientos admisibles en el diagrama V,m

Si se parte de un punto inicial, I, y se quiere llegar a un punto final, F, los desplazamientos admisibles se muestran en la figura 8.2/1.

Las líneas ID y CF vienen dadas por la ecuación:

$$m = m_0 \quad 8.2/17$$

Las líneas IC y DF responden a la ecuación:

$$\left(\frac{m}{m_0}\right) \sqrt{\frac{4kI_{sp}}{\beta_{max}}} = \left(\frac{v - \sqrt{\frac{I_{sp}\beta_{max}}{k}}}{v + \sqrt{\frac{I_{sp}\beta_{max}}{k}}}\right) \cdot \left(\frac{v_0 + \sqrt{\frac{I_{sp}\beta_{max}}{k}}}{v_0 - \sqrt{\frac{I_{sp}\beta_{max}}{k}}}\right) \quad 8.2/18$$

En las expresiones 8.2/17 y 8.2/18, m_0 y V_0 representan la masa y velocidad en el punto de partida.

- Maximización del alcance

De 8.2/12 y de 8.2/14 se obtiene:

$$dx = - \frac{v}{\beta} dm \quad 8.2/19$$

$$\frac{v}{\beta} = \frac{m}{kV} \left(\frac{dv}{dm} + \frac{I_{sp}}{m} \right) \quad 8.2/20$$

Sustituyendo 8.2/20 en 8.2/19 e integrando:

$$x = \int_I^F \left(- \frac{m}{kV} dv - \frac{I_{sp}}{kV} dm \right) \quad 8.2/21$$

La cuestión a resolver es cual de todos los posibles caminos que van desde el punto I al punto F de la figura 8.2/1, maximiza la ecuación 8.2/21.

Resolución teórica

Según la Ref. 14 para resolver este problema sólo es necesario estudiar el comportamiento del signo de la llamada función fundamental, la cual, para una integral lineal de la forma

$$H = \int_I^F [\Psi(x,y) + \psi(x,y) y'] dx, \quad 8.2/22$$

se define como

$$\omega(x,y) = \frac{\partial \psi}{\partial x} - \frac{\partial \Psi}{\partial y} \quad 8.2/23$$

En el caso estudiado:

$$\Psi(V,m) = - \frac{m}{kV} \quad 8.2/24$$

$$\psi(V,m) = - \frac{I_{sp}}{kV} \quad 8.2/25$$

$$\omega(V,m) = \frac{I_{sp} + V}{kV^2} > 0 \quad 8.2/26$$

Según la citada Ref. 14, si $\omega > 0$ la integral lineal 8.2/21 se maximizará por el camino IDF y se minimizará por el ICF.

FIGURA 8.2/1

REGIÓN DE DESPLAZAMIENTOS ADMISIBLES

8.3 RESOLUCIÓN DE UN CASO CONCRETO MEDIANTE LOS ALGORITMOS GENÉTICOS

El caso concreto que será resuelto mediante los programas desarrollados es el siguiente:

$$\begin{array}{ll}
 I_{sp} = 2000 \text{ Ns/kg} & m_I = 200 \text{ kg} \\
 k = 0.02695 \text{ kg/m} & m_F = 170 \text{ kg} \\
 \beta_{max} = 1.5 \text{ kg/s} & V_I = V_F = 300 \text{ m/s}
 \end{array}$$

Con los valores anteriores se obtiene la región de desplazamientos admisibles que recoge la Fig 8.3/1.

Los resultados obtenidos se han representado mediante puntos en la figura mencionada. El alcance máximo conseguido es 10243 m.

Es de destacar el excelente comportamiento de los programas desarrollados.

FIGURA 8.3/1

REGION DE DESPLAZAMIENTOS ADMISIBLES Y RESULTADOS OBTENIDOS, PARA EL CASO CONCRETO ESTUDIADO.

9. CONCLUSIONES

En la presente Tesis se han obtenido los siguientes resultados originales:

- 1) Sistematización en el área de los Algoritmos Genéticos.
- 2) Clasificación de los métodos probabilísticos y resumen de las técnicas más importantes encuadradas en cada Categoría propuesta.
- 3) Obtención de las probabilidades de éxito, cuando el tamaño de paso es óptimo, para los modelos de corredor y esférico y para los algoritmos $(1,\lambda)$ y $(1+\lambda)$.
- 4) Obtención de un método que permite controlar el tamaño de paso y adaptarlo a su valor óptimo, para una función objetivo cualquiera.
- 5) Conclusión de que entre los algoritmos $(1,\lambda)$, el mejor de ellos es el correspondiente a $\lambda = 5$ ó $\lambda = 6$.
- 6) Conclusión de que entre los algoritmos $(1+\lambda)$, el mejor de ellos es el correspondiente a $\lambda = 1$.
- 7) Conclusión de que entre todos los algoritmos $(1,\lambda)$ y $(1+\lambda)$, el mejor de ellos es el $(1+1)$.
- 8) Diseño y puesta a punto de programas de ordenador basados en los algoritmos $(1,\lambda)$ y $(1+\lambda)$.

- 9) Conclusiones idénticas a las mencionadas en los puntos 5), 6) y 7), obtenidas al aplicar los programas diseñados a funciones objetivo de prueba.
- 10) Resolución de un problema típico de Mecánica del Vuelo (maximización del alcance de un misil no sustentador) mediante la utilización de los Algoritmos Genéticos.

APÉNDICE I

LA REVOLUCIÓN DARWINISTA. EL EVOLUCIONISMO

La idea de la Evolución, en contra de lo que pudiera pensarse, proviene de muy antiguo. En Grecia y Roma ya existían opiniones que basaban el origen de los organismos en leyes naturales.

Sin embargo, cualquier especulación en torno al proceso evolutivo quedó truncada por dos hechos históricos trascendentales: la aparición de los sistemas filosóficos de Platón y Aristóteles, y la aparición y posterior expansión del cristianismo.

La Teoría de las Ideas de Platón postula la existencia transcendente del mundo que contemplamos; todas las cosas de este mundo tienen un modelo transcendente, por lo cual es absurda cualquier separación, por pequeña que sea, de ese modelo. La Metafísica de Aristóteles también impide el paso al Evolucionismo, aunque este autor no postule la realidad de una Idea independiente de la cosa representada.

Estas filosofías, unidas al relato que realiza el Génesis de la Creación, condujeron a una concepción estática del mundo, la cual perduró hasta el final de la Edad Media.

Si dos hechos fueron los causantes principales de la mencionada imagen estática del mundo, otros dos acontecimientos debilitaron esa concepción. El primero fue el nacimiento de una Física nueva, que culmina con los "Principia" de Newton, la cual puso en tela de juicio la aceptación textual de los relatos bíblicos.

Y el segundo acontecimiento fue el adelanto realizado en ciencias geológicas y biológicas como, por ejemplo, toda una serie de descubrimientos de distintas formas fósiles. Este hecho concre-

to hizo que se cuestionara el que la edad de la Tierra fuese tan corta como parecía desprenderse de la Biblia.

La autoridad de la Biblia, por tanto, disminuyó algo a finales del siglo XVII aunque, concretamente en Gran Bretaña, se reforzó de nuevo en el siglo XVIII como consecuencia del evangelismo predicado por John Wesley. Y esta autoridad continuó siendo un importante factor en el siglo XIX.

Independientemente de esto, hacia finales del siglo XVIII y principios del siglo XIX las especulaciones acerca de la evolución de los seres vivos ya no era novedosas, aunque tampoco estaban muy extendidas ni eran aceptadas en su totalidad.

El primer estudio sistemático a favor de un punto de vista evolucionista que explicara el origen de los seres vivos fue realizado por el biólogo francés Jean Baptiste de Lamarck. La obra de Lamarck, a pesar de estar escrita en un lenguaje muy poco preciso y a pesar de contener varios errores importantes, constituyó una aceptable base para la posterior obra de Darwin.

Lamarck creía que todo el reino animal se podía ordenar en una escala descendente (la llamada "Cadena del Ser" ó "Escala de la Naturaleza") en la que el hombre ocupaba la parte más alta y los infusorios la parte más baja. Se sospecha que Lamarck veía a los seres vivos transformarse siguiendo una dirección predeterminada cuyo extremo y fin, en el reino animal, era el hombre. Esta es la razón por la que, de alguna forma, Lamarck era un teólogo ya que explicaba los hechos en estudio mediante causas finales.

El rasgo más conocido de la teoría de Lamarck es que cualquier novedad adquirida por un individuo antes de la procreación gracias a un comportamiento determinado y a un entorno específico, puede ser transmitida a la progenie, si no por completo, si al

menos parcialmente. Esta idea, denominada "lamarckismo" es insostenible, pues el fundamento sobre el que descansa (las propiedades adquiridas son heredables) es falso, según se demostró posteriormente.

Y este era el estado de la cuestión cuando surgió la figura de Charles Robert Darwin (1809-1882).

La Revolución Darwinista no se puede contemplar desde un punto de vista único; tuvo aspectos, causas y efectos muy diferentes. Con frecuencia se la presenta como el triunfo de la ciencia sobre la religión, pero a pesar de que esta idea sea fundamental, la evaluación así hecha es inadecuada. Un análisis exhaustivo del fenómeno darwinista, el cual no será realizado aquí por razones obvias, debería incluir los aspectos científico, filosófico, teológico, social y político, todos ellos interrelacionados entre sí.

Vistos todos estos aspectos de forma global, se puede decir que la Revolución Darwinista completa la Revolución Copernicana que comenzara a mediados del siglo XVI y que prosiguieran después Galileo, Kepler y Newton.

La importancia fundamental de la Revolución Copernicana fue llegar a la conclusión de que el Universo obedece a leyes inmanentes, las cuales determinan los fenómenos naturales. De esta forma se sustituye una concepción animista del Universo por una concepción causal.

Darwin extiende al mundo orgánico el concepto de Naturaleza derivado de la Física y la Geología, es decir, el que no es necesario el postulado de la existencia de agentes sobrenaturales para explicar los fenómenos orgánicos.

En este momento el hombre occidental alcanza su madurez in-

telectual, ya se enfrenta a la Naturaleza con la confianza de que ésta es comprensible y, en algunos aspectos, modificable.

A continuación, y una vez resumidos los aspectos históricos más sobresalientes de la Revolución Darwinista, se exponen, también abreviadamente, las ideas fundamentales del Evolucionismo.

En 1859 Darwin publicó su célebre obra: "On the Origin of Species by Means of Natural Selection" ó "The Preservation of Favoured Races in Struggle for Life". En ella estableció las bases del Evolucionismo, las cuales, como casi todas las ideas revolucionarias que han cambiado el mundo, son de extrema sencillez y, al mismo tiempo, de gran potencia.

El Evolucionismo explica la adaptación de los organismos a su entorno por medio del mecanismo fundamental de mutación-selección.

En una población determinada, algunos individuos sufren cambios espontáneos heredables, o lo que es lo mismo, mutaciones. Las causas por las que aparecen estas mutaciones no son completamente conocidas en la actualidad. Parece ser que influyen en ellas muchos factores (uno de los más conocidos es la radiación de ondas electromagnéticas) y que el que se produzcan o no es un proceso aleatorio.

A continuación, de esa población en estudio sobreviven los individuos mejor adaptados a su medio ambiente. Este es el mecanismo de la selección natural. Obsérvese que ésta favorece el carácter continuamente adaptativo de los organismos, pues de todos los individuos sobreviven aquellos que han sido favorecidos por mutaciones aleatorias que aumenten la probabilidad de supervivencia, las cuales son transmitidas a la descendencia, con lo que la po-

blación mejora globalmente.

Si no existiesen mutaciones (el azar), no podría darse el proceso evolutivo pues no habría variantes alternativas sobre las que pudiese actuar la selección y si no existiese la selección (el determinismo) se obtendría la misma conclusión, ya que las mutaciones, por ser aleatorias, son desfavorables en su mayor parte y de esta forma se llegaría a la desorganización y extinción de la población.

El problema que surge ahora es el siguiente: ¿Están los seres vivos completamente adaptados o, lo que es lo mismo, completamente optimizados con respecto a su entorno? A pesar de que los seguidores de Leibniz y los biólogos neodarwinistas, defensores de las teorías panglossianas, opinen que la optimización es total y de que se haya comprobado que existen en la Naturaleza soluciones óptimas para algunos problemas, la respuesta no parece ser esa en la actualidad.

Sin embargo, sí parece claro a la luz del presente conocimiento biológico que los seres vivos están, al menos, parcialmente optimizados en relación con su medio ambiente.

Para terminar esta breve síntesis del Evolucionismo, conviene citar textualmente a F.J. Ayala (Ref. 15): "La teoría de la evolución nos muestra el azar y la necesidad intrincados en el meollo de la vida, la casualidad y el determinismo entrelazados en un proceso natural que ha producido las realidades más complejas del Universo, los seres vivientes; y entre ellos, el hombre, capaz de pensar y de amar, capaz del libre albedrío y de analizar el proceso mismo de la evolución que le ha traído a la existencia".

APÉNDICE II

VOCABULARIO DE GENÉTICA BÁSICA

AISLAMIENTO: El mecanismo de aislamiento, mediante el desarrollo de barreras de reproducción, mantiene separados los puntos del espacio de nucleótidos que forman los distintos "canales de éxito". Así, pueden seguir evolucionando separadamente.

ALELO: Cada una de las diversas formas alternativas en que se puede presentar un gen. Un alelo se distingue de los demás por pequeños cambios de su estructura bioquímica.

CÉLULA: Unidad morfológica y fisiológica que compone el cuerpo de las plantas y animales. Tiene vida propia y está constituida por protoplasma, en general, y por citoplasma y núcleo, en particular.

CENTROMETRO: Punto central de un cromosoma donde se unen las dos cromátidas.

CITOPLASMA: Parte del protoplasma que constituye la sustancia viva de la célula, con exclusión del núcleo.

CROMATIDA: Cada uno de los dos brazos paralelos de un cromosoma. Las dos cromátidas están unidas, en sus puntos centrales, por el centrómetro.

CROMOSOMA: Orgánulo lineal, fuertemente tingible por cualquier colorante (propiedad de la que obtiene su nombre), que contiene los genes. Los cromosomas constituyen el núcleo de las células de los seres vivos eucariotas y se presentan en número constante en cada especie.

CRUCE: También denominado recombinación o incluso crossover. Consiste en un cambio de secciones de cromosomas entre pares de estos. El proceso tiene cuatro etapas: en la primera los cromosomas se agrupan en pares; en la segunda, aparecen puntos de fusión en los cromosomas homólogos; en la tercera etapa, se intercambian secciones, llamadas quiasmas, entre los dos cromosomas homólogos; y por último, en la cuarta etapa, se separan los cromosomas homólogos.

DELECIÓN: Anomalía cromosómica estructural por la que una sección de un cromosoma se rompe y se pierde. El número de genes del cromosoma disminuye.

DIPLOIDE: Una célula se dice que es diploide si tiene n parejas de cromosomas, es decir, si el número total de cromosomas es $2n$. Se dice que un mecanismo de herencia es diploide si el cruce entre los cromosomas de los padres tiene lugar después de la generación del nuevo ser.

DUPLICACIÓN: Anomalía cromosómica estructural por la que una sección de un cromosoma se duplica. El número de genes del cromosoma aumenta.

EUCARIOTA: Dícese de los seres vivos cuyas células tienen núcleo. Constituyen la casi totalidad de los organismos.

FENOTIPO: Se denomina así al conjunto de rasgos somáticos, físicos, bioquímicos, psíquicos, etc. que se presentan en un ser vivo. Es la manifestación externa del genotipo.

GEN: Es, por definición, la unidad de factor genético; aquella capaz y suficiente para dar lugar a la producción de un cierto carácter somático o bioquímico.

GENERACIÓN: En determinada especie, conjunto de todos los seres vivos coetáneos. Dícese también del hecho de crear o producir algo (descendientes, en el caso particular de la Biología).

GENOTIPO: Conjunto de todos los genes de un ser vivo. Es la característica o carga genética total del individuo.

HAPLOIDE: Una célula se dice que es haploide si tiene solamente n cromosomas (sin formar parejas). Se dice que un mecanismo de herencia es haploide si el cruce entre los cromosomas de los padres tiene lugar antes de la generación del nuevo ser.

HETEROCIGOTICO: Un individuo será heterocigótico si los dos alelos que ocupan un determinado locus son distintos.

HOMOCIGOTICO: Un individuo será homocigótico si los dos alelos que ocupan un determinado locus son idénticos.

INDIVIDUO: Sinónimo de ser vivo u organismo diferenciado.

INVERSIÓN: Anomalía cromosómica estructural por la que una sección de un cromosoma se rompe, gira 180° y se fija de nuevo en el mismo cromosoma. El número de genes del cromosoma no cambia.

LOCUS: Lugar geométrico dentro de un cromosoma, ocupado por un determinado gen.

MEIOSIS: División doble peculiar, que sólo tiene lugar en las gónadas (testículos y ovarios) y cuya finalidad es la de producir gametos (espermatozoides y óvulos) haploides.

MITOSIS: División regular, por la cual una célula diploide origina dos células hijas asimismo diploides.

MUTACIÓN: Cambio aleatorio de un gen producido por multitud de causas, algunas de las cuales no son totalmente conocidas, y

que se transmite hereditariamente a los descendientes.

NÚCLEO: Corpúsculo especializado del protoplasma de todas las células vivas. Está constituido, esencialmente, por los cromosomas y está rodeado por el citoplasma. Actúa como órgano rector de las funciones de nutrición y reproducción de la célula.

NUCLEOTIDO: Los ácidos nucleicos (ácido desoxiribonucleico, DNA y ácido ribonucleico, RNA) son polímeros formados por unidades menores llamadas nucleótidos. Cada nucleótido consta, a su vez, de un fosfato, una pentosa y una base orgánica.

POBLACIÓN: Dícese de un conjunto de individuos que componen una determinada comunidad.

PROCARIOTA: Organismos celulares primitivos, como las bacterias y algas verdeazuladas, desprovistas de núcleo. Sin embargo, sí disponen de un área nuclear.

PROTOPLASMA: Sustancia constitutiva de las células, de consistencia más o menos líquida, estructura coloidal y composición química muy compleja. Contiene gran cantidad de agua en la que están disueltos numerosos cuerpos orgánicos y sales inorgánicas.

RECOMBINACIÓN: Sinónimo de cruce (ver cruce).

SELECCIÓN: Mecanismo evolutivo mediante el cual, de una determinada población sólo sobreviven aquellos individuos mejor adaptados a su medio ambiente.

TRANSLOCACION: Anomalía cromosómica estructural por la que una sección de un cromosoma se rompe y se fija en otro cromosoma. El número de genes del primer cromosoma disminuye y el del segundo aumenta.

VITALIDAD: Capacidad de un individuo para sobrevivir en un determinado medio ambiente.

REFERENCIAS Y BIBLIOGRAFIA

1. Leitman, G.: "Optimization Techniques with Applications to Aerospace Systems", Academic Press, New Yor, 1962.
2. Bazaraa, S.B. & Shetty, CM.: "Nonlinear Programming. Theory and Algorithms", John Wiley and Sons, New York, 1979.
3. Dixon, L.C.W. & Szego, G.P. (eds.): "Towards Global Optimization", North-Holland American Elsevier, 1975.
4. Dixon, L.C.W. & Szego, G.P. (eds.): "Towards Global Optimization, 2", North-Holland American Elsevier, 1978.
5. Janin, G.: "The Genetic Operator for the Solution of Optimization Problems", MAO TR-3, ESOC, 1983.
6. Schwefel, H.P.: "Numerische Optimierung von Computer-Modellen- Mittels der Evolutionsstrategie", Birkhauser Verlag, Basel, 1977.
7. Lootsma, F.A. (ed.): "Numerical Methods for Non-Linear Optimization", Academic Press, London, 1972.
8. Beltrami, E.J. & Indusi, J.P.: "An Adaptative Random Search Algorithm for Constrained Minimization", IEEE, Transactions on Computers, September, 1972.
9. Rechenberg, I.: "Evolutionsstrategie. Optimierung Technischer Systeme nach Prinzipier der biologischen Evolution". Friedrich Fromman Verlag, Stuttgart, 1973.
10. Hofstadter, D.R.: "Torneos Computerizados del Dilema del Preso que sugieren cómo evoluciona la conducta cooperativa", Temas Metamágicos, Investigación y Ciencia, Julio 1983.
11. Price, W.L.: "Global Optimization by Controlled Random Search", J.O.T.A., Vol. 40, No. 3, July 1983.

12. Bekey, G.A. & Ung, M.T.: "A Comparative Evaluation of Two Global Search Algorithms", IEEE Transactions on Computers, January 1974.
13. Borsworth, J., Foo, N. and Zeigler, B.P.: "Comparison of Genetic Algorithms with Conjugate Gradient Methods", NASA CR-2093, August 1972.
14. Martínez García, J.J.: "Problemas Especiales de Mecánica del Vuelo", IUPFAN-ETSIN, Sept. 1980.
15. Ayala, F.J.: "Darwin contra las causas sobrenaturales", Tribuna Libre, "El País", Abril 1982.
16. Wilde, D. & Beighter, C.S.: "Formulations of Optimizations", Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1967.
17. Wismer, D.A. & Chattergy, R.: "Introduction to Nonlinear Optimization. A Problem Solving Approach", North-Holland, New York, 1973.
18. Kushner, H.J. & Clark, D.S.: "Stochastic Approximation Methods for Constrained and Unconstrained Systems", Springer Verlag, New York, 1978.
19. Collatz, L. & Wetterling, W.: "Optimization Problems" (Applied Mathematical Sciences, 17), Springer-Verlag, New York, 1975.
20. Bellman, R.E. & Dreyfus, S.E.: "La Programmation Dynamique et ses Applications", Dunod, Paris, 1965.
21. Stoecker, W.F.: "Design of Thermal Systems", McGraw-Hill, 1980.
22. Sóbol, I.M.: "Método de Montecarlo", Mir, Moscú, 1976.
23. Masry, S.F., Bekey, G.A. & Safford, F.B.: "A Global Optimization Algorithm Using Adaptive Random Search", Applied Mathematics and Computations 7, 1980.

24. Gran, R.: "On the Convergence of Random Search Algorithms in Continuous Time with Applications to Adaptative Control", IEEE, Vol. SMC-3, No. 1, Jan., 1973.
25. Schumer, M.A. & Steiglitz, K.: "Adaptative Step Size Random Search", IEEE, Vol. AC-13, No. 3, Junio, 1968.
26. Foo, N.Y. & Bosworth, J.L.: "Algebraic, Geometric and Stochastic Aspects of Genetic Operators", NASA CR-2099, Agosto 1972.
27. Holland, J.H.: "Genetic Algorithms and the Optimal Allocation of Trials", SIAM J. Comp., Vol. 2, No. 2, Junio 1973.
28. Gaidosch, H.: "Optimization of Missile Trajectories by Means of Evolution Strategy", AIAA-80-0323, 1980.
29. Bethkle, A.D., Zeigler, B.P. & Strauss, D.M.: "Convergence Properties of Simple Genetic Algorithms", NASA TR-159, 1974.
30. Ruse, M.: "La Revolución Darwinista", Alianza Universidad, Madrid, 1983.
31. Ford, E.B.: "Mendelismo y Evolución", Nueva Colección Labor, Barcelona, 1973.
32. Sánchez Cascos, A.: "Manual de Genética Médica", Editorial Científico-Médica.
33. Cerdá Olmedo, E.: "Nuestros Genes", Aula Abierta Salvat, Barcelona, 1981.
34. Schrödinger, E.: "Mente y Materia", Cuadernos Infimos 110, Tusquets Editores, Barcelona, 1984.
35. Landsberg, P.R.: "The Enigma of Time", Adam Hilger, Ltd., Bristol, 1982.
36. Gómez-Tierno, M.A.: "Solución Numérica de los Problemas de Optimización", GMV 207/83, ETSI Aeronáuticos, UPM, 1983.

37. Gómez-Tierno, M.A.: "Métodos Numéricos Generales por Búsqueda no Aleatoria Utilizados en la Resolución de los Problemas de Optimización Paramétrica", GMV 209/83, ETSI Aeronáuticos, UPM, 1983.