ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
DEPARTAMENTO DE MATEMÁTICAS

TESIS DE GRADO

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:
"MAGÍSTER EN INVESTIGACIÓN MATEMÁTICA"

TEMA:
"ANÁLISIS E IMPLEMENTACIÓN DE DOS MÉTODOS NO TRADICIONALES PARA RESOLVER SISTEMAS NO LINEALES"

AUTOR:
CARLOS MANUEL MARTÍN BARREIRO

GUAYAQUIL – ECUADOR
2013

DEDICATORIA

A mi esposa Soraya y a mi hija Alicia a quienes amo con todas mis fuerzas. A mi tía Maruja Martín González y a mi abuela Alicia Camacho de Martín, dos personas que me entregaron su vida. A mi querida Mayeve.

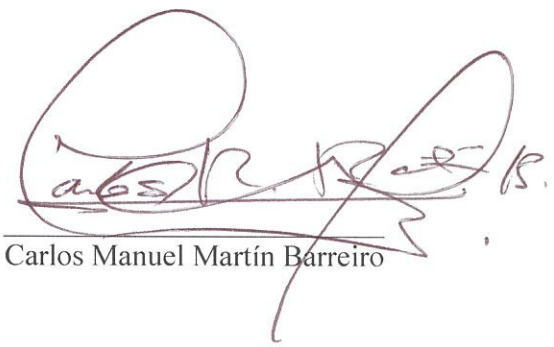
AGRADECIMIENTOS

A DIOS y a mis padres Carlos Julio y Eva Rosalía. A mis hermanos José Antonio y Julio Ricardo. A mis tíos Ricardo, Joel y Lino. A mis tías Rita, Alexandra y Mercedes. A mis primos Nicolás, Adrián y Jael. A mis profesores por todas sus enseñanzas, en especial a los Doctores Borys Álvarez Samaniego, Rafael Martí, Joseph Páez, Polo Vaca, Julio Medina y Francisco Vera. A mis compañeros José Xavier Cabezas, Eduardo Rivadeneira, Erwin Delgado, Roberto Cascante, John Ramírez, Elkin Angulo, Pablo Álvarez y Margarita Martínez por su amistad, consejos y sugerencias.

DECLARACIÓN EXPRESA

La responsabilidad por los hechos y doctrinas expuestas en esta TESIS DE MAESTRÍA, me corresponde exclusivamente; el patrimonio intelectual del mismo, corresponde exclusivamente a la **Facultad de Ciencias Naturales y Matemáticas (FCNM), Departamento de Matemáticas** de la Escuela Superior Politécnica del Litoral (ESPOL).

(Reglamento de Graduación de la ESPOL)



Carlos Manuel Martín Barreiro

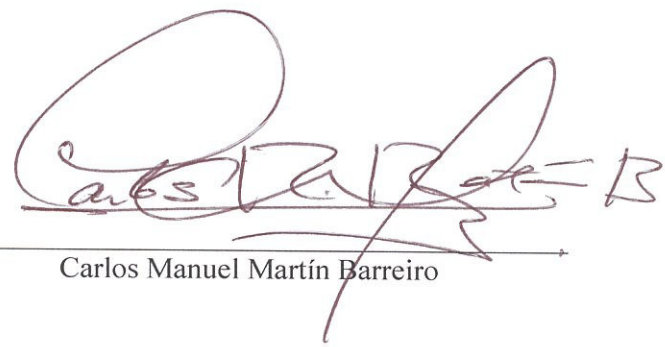
TRIBUNAL DE GRADUACIÓN

M. Sc. Efrén Jaramillo
PRESIDENTE DEL TRIBUNAL

Doctor Francisco Vera
DIRECTOR DE TESIS

M. Sc. Luis Rodríguez
VOCAL DEL TRIBUNAL

AUTOR DE LA TESIS



Carlos Manuel Martín Barreiro

RESUMEN

En este documento de tesis se presentan dos métodos numéricos no clásicos, y por ende poco conocidos, para resolver sistemas de ecuaciones no lineales. Cabe indicar que son de interés únicamente las soluciones reales de dichos sistemas. Por un lado este par de métodos son independientes en el sentido que no dependen de otros métodos para encontrar las soluciones de un sistema, pero por otro lado pueden también trabajar de forma colaborativa con los métodos tradicionales como los métodos de punto fijo en general y los métodos de Newton y sus variantes.

Estos métodos que se presentarán son conocidos como métodos de convergencia global, mientras que los métodos iterativos usuales son métodos de convergencia local. Entonces, para cada método de convergencia global que mostraremos, se va a plantear un problema equivalente. Para el primero el problema equivalente es un problema de optimización, mientras que para el segundo el problema equivalente es un sistema de ecuaciones diferenciales ordinarias de primer orden junto con un dato inicial. Se usarán algoritmos heurísticos y metaheurísticos para el problema de optimización y el método numérico de Runge-Kutta clásico de cuarto orden para el problema de valor inicial.

Se va a demostrar formalmente que resolver un sistema no lineal de ecuaciones es equivalente a resolver un problema de optimización, particularmente un problema de minimización. Se utiliza la metaheurística “scatter search” y las heurísticas del “gradiente” y de “Nelder-Mead” para el problema de optimización. Se demuestra también que resolver un sistema no lineal de ecuaciones es equivalente a resolver un problema de valor inicial de primer orden. Para el segundo método que mostraremos los conceptos de homotopía y continuación numérica son muy importantes. Finalmente, se realizaron experimentos computacionales y se incluyen resultados sobre la ejecución de estos dos métodos numéricos sobre algunos sistemas no lineales de ecuaciones.

INTRODUCCIÓN

Los sistemas no lineales de ecuaciones se presentan con mucha frecuencia en diversas áreas de las ciencias básicas y aplicadas. Resolver un sistema de ecuaciones no lineales no es una tarea trivial, es decir, no es una tarea en general sencilla. Se trata de un problema que ha cautivado a muchos matemáticos durante siglos [6]. Aún en la actualidad se discuten nuevos métodos para resolver sistemas no lineales [1] [2] [3] [9]. Normalmente no es posible, a través de algún método analítico, conocer la solución exacta de un sistema no lineal. Necesitamos entonces métodos numéricos que nos permitan aproximar, tanto como se desee, las soluciones reales de sistemas no lineales.

Existen muchos métodos numéricos iterativos convencionales para resolver sistemas de ecuaciones no lineales, entre los que destacan los métodos de punto fijo en general y los métodos de Newton y sus variantes [6] [10] [11] [12] [13]. Usualmente los métodos iterativos clásicos utilizan fórmulas recursivas que producen una sucesión vectorial que, dadas ciertas condiciones, converge a una solución del sistema no lineal. Para garantizar la convergencia de la sucesión construida, normalmente se necesita, entre otras cosas, que el dato inicial (es decir el primer elemento de la sucesión vectorial) se encuentre "cerca" de una solución del sistema no lineal, además de que las funciones de varias variables que componen el sistema cumplan ciertas condiciones de continuidad y diferenciabilidad. Por esta razón es que estos métodos clásicos se conocen como métodos de convergencia local. Si el dato inicial no se encuentra dentro de la región de convergencia (usualmente una bola abierta centrada en una solución del sistema no lineal), entonces la sucesión será divergente y por lo tanto no va a converger a una solución del sistema no lineal. Dependiendo del problema a resolver, puede ocurrir que el radio de la región de convergencia tenga una medida muy pequeña. Es muy importante tener presente entonces que los métodos clásicos pueden fallar principalmente cuando se tiene poca o ninguna información sobre en qué parte de la región de búsqueda el sistema tiene sus soluciones.

Otra técnica consiste en trabajar sobre un problema equivalente de optimización. Sobre este nuevo problema se pueden aplicar métodos de optimización numérica conocidos, como el método del gradiente o el método BFGS, sin embargo, desde hace menos de una década, se están usando algoritmos metaheurísticos que hacen uso de la computación evolutiva, como los "algoritmos genéticos" y la "búsqueda dispersa", para resolver el problema equivalente de optimización, con bastante éxito en cuanto al consumo de recursos computacionales y a los resultados obtenidos [4] [5].

En esta tesis se muestran dos métodos numéricos que permiten encontrar soluciones reales a sistemas no lineales de ecuaciones. El primer método consiste en plantear como problema equivalente al sistema no lineal un problema de minimización. Se enuncia y se demuestra formalmente una proposición que nos dice que efectivamente todo sistema no lineal tiene un problema equivalente de optimización. Pero para resolver el problema de minimización no se utilizarán métodos de optimización numérica, por el contrario se usará la heurística. Como métodos de búsqueda local se emplearán los algoritmos heurísticos del "gradiente" y de "Nelder-Mead". Detrás de ellos y orquestando todo el funcionamiento de búsqueda de soluciones de un sistema no lineal se tendrá la metaheurística conocida como "búsqueda dispersa". Es decir, el primer método que se presenta consiste básicamente en un solo algoritmo general evolutivo para resolver sistemas no lineales. Al final se muestran los experimentos computacionales realizados.

Existen muchas metaheurísticas, entre ellas: el algoritmo genético, la búsqueda tabú, el recocido simulado, GRASP, que pueden utilizarse, junto con heurísticas, para resolver problemas de optimización [4] [5]. "Búsqueda dispersa" es también un método metaheurístico que se puede emplear para resolver problemas de optimización y forma parte de la familia de los algoritmos evolutivos [4] [5]. Tiene sus orígenes en los años setenta, pero es en la última década cuando ha sido rediseñado y probado en muchos problemas con un alto grado de dificultad y con excelentes resultados [4] [5]. Esta es la principal razón por la cual se escogió el método de búsqueda dispersa para resolver el problema de minimización del primer método que se propone para resolver sistemas no lineales de ecuaciones.

El segundo método que se presenta trata sobre plantear otro tipo de problema equivalente al sistema no lineal: un problema de valor inicial de primer orden. Esta idea la propuso D. F. Davidenko alrededor de 1960 y actualmente tiene muchas aplicaciones [6] [7]. Para esta forma diferente de resolver un sistema no lineal hacemos uso del concepto de homotopía y de la continuación numérica [6] [7]. Se demostrará de manera rigurosa más adelante que resolver el sistema de ecuaciones diferenciales de primer orden con el dato inicial es equivalente a resolver el sistema no lineal de ecuaciones. Se hará uso del método de cuarto orden de Runge-Kutta clásico para resolver numéricamente el problema de valor inicial. Por lo tanto, se tomará como una solución del sistema no lineal la solución obtenida en el problema de valor inicial. Finalmente se comparten algunos experimentos computacionales.

Cabe indicar que lo que se persigue con estos dos métodos numéricos que se presentan es intentar trabajar con métodos que proporcionen un radio de convergencia de mayor medida que lo que ocurre en los métodos iterativos convencionales donde la medida del radio de convergencia puede ser muy pequeña [6] [7]. Debido a lo mencionado anteriormente es que a estos métodos que se muestran en esta tesis se los conoce como métodos de convergencia global. Con estos métodos es más probable tener convergencia a una solución del sistema no lineal cuando se ingresa un dato inicial. Por un lado estos métodos pueden trabajar de forma cooperativa con los métodos clásicos. Por ejemplo, si no se tiene ninguna sospecha de dónde podría encontrarse una solución del sistema se puede utilizar un método de convergencia global para acercarnos lo más posible a una solución y, luego, podríamos tomar esa salida proporcionada por el método de convergencia global como el dato inicial que necesita el método de convergencia local para iterar el número de veces que sea necesario y obtener así una aproximación que se encuentre tan cerca como se desee de una solución del sistema no lineal. Por otro lado, se pueden usar independientemente para que por sí mismos estos métodos de convergencia global carguen con todo el peso de encontrar las soluciones.

Como autor de este documento de tesis, deseo terminar esta introducción indicando que mi contribución se encuentra principalmente en el “método heurístico” para resolver sistemas no lineales de ecuaciones. La metodología de “búsqueda dispersa” proporciona sugerencias y lineamientos generales cuando se pretende resolver un problema de optimización, mi aporte estuvo en el diseño y en el detalle de la implementación para

resolver puntualmente el problema de minimización de una función no lineal de varias variables sobre un conjunto compacto (una región hiperrectangular cerrada en R^n). La búsqueda local, el método de combinación, entre otros aspectos internos de diseño e implementación que utiliza el "método heurístico" de este documento de tesis, son de mi autoría. Quiero también indicar que tuve la fortuna de ser alumno del Doctor Rafa Martí, profesor de la Universidad de Valencia en España, él es uno de los creadores de la metaheurística de "búsqueda dispersa". Sus enseñanzas, sus consejos, fueron una guía fundamental para que yo pudiera terminar con éxito la implementación total del "método heurístico" que resuelve sistemas no lineales de ecuaciones. En la maestría de investigación matemática el Doctor Martí me dictó las materias "Optimización" y "Algoritmos Heurísticos".

El Doctor Joseph Páez, profesor de ESPOL, me sugirió que realizara la investigación del "método homotópico", el segundo método de esta tesis para resolver sistemas de ecuaciones no lineales. En este segundo método mi aporte es mínimo, es más un trabajo de mucha investigación en lo teórico. Mi contribución en este segundo método apunta a la implementación y ciertas cuestiones de la convergencia que se discuten en la sección 3.4 de este documento. También fui alumno del Doctor Páez, con él tuve la oportunidad de estudiar el concepto de homotopía y algunos métodos de continuación numérica. En la maestría de investigación matemática el Doctor Páez me dictó las materias "Análisis Numérico" y "Sistemas Dinámicos". Reitero mi agradecimiento inmenso a los Doctores Martí y Páez.

ÍNDICE DE CONTENIDO

CONTENIDO	PÁGINA
Resumen	
Introducción	
1. Conceptos y Fundamentos	**1**
1.1 Sistema de Ecuaciones NO Lineales	**1**
1.2 Función Objetivo de Minimización	**2**
1.3 Funciones Homotópicas y Homotopía	**3**
1.4 Función Definida Implícitamente	**5**
1.5 Homeomorfismo	**5**
1.6 Teorema de la Función Implícita	**6**
1.7 Teorema de la Función Inversa	**7**
1.8 Heurísticas y Metaheurísticas	**7**
2. El Método Numérico Heurístico	**10**
2.1 El Problema de Optimización	**10**
2.2 El Algoritmo del Método Heurístico	**12**
2.2.1 Generación del Conjunto *P*	**12**
2.2.2 Búsqueda Local	**14**
2.2.3 Construcción del Conjunto *R*	**15**
2.2.4 Actualización y Mantenimiento del Conjunto *R*	**15**
2.2.5 Método de Combinación	**17**
2.2.6 Método Principal del Algoritmo	**17**
2.3 Análisis Comparativo con Métodos Clásicos	**19**
2.4 Análisis de Convergencia	**20**
3. El Método Numérico Homotópico	**23**
3.1 El Problema de Valor Inicial	**23**
3.2 El Algoritmo del Método Homotópico	**27**
3.3 Análisis Comparativo con Métodos Clásicos	**29**
3.4 Análisis de Convergencia	**31**
4. Experimentos Computacionales	**33**
4.1 Una Ecuación NO Lineal	**33**
4.2 Sistemas NO Lineales de Segundo Orden	**37**
4.3 Sistemas NO Lineales de Orden Superior	**41**
Conclusiones	**45**
Recomendaciones	**48**
Referencias Bibliográficas	**50**

ÍNDICE DE FIGURAS

FIGURA	**PÁGINA**
2.1 Etapas Búsqueda Dispersa	**17**
4.1 Gráfico de $f(x)=x^3-e^x\ sen(x)$	**29**
4.2 Gráfico de $g(x)=\left\| x^3-e^x\ sen(x)\right\|$	**29**
4.3 Curvas de Nivel de *H* para $f(x)=x^3-e^x\ sen(x)$	**31**
4.4 Gráfico de $g(x,y)=\left\| x\ sen(y)-1\right\|+\left\| x^2+cos(2y)\right\|$	**32**
4.5 Otra Vista de $g(x,y)=\left\| x\ sen(y)-1\right\|+\left\| x^2+cos(2y)\right\|$	**33**
4.6 Trayectoria de las Soluciones Segundo Ejemplo	**34**
4.7 Gráfica de $g(x,y)=\left\| 5x^2-y^2\ \right\|+\left\| 4y-sen(x)-cos(y)\right\|$	**35**
4.8 Trayectoria de las Soluciones Tercer Ejemplo	**36**

ÍNDICE DE TABLAS

TABLA	PÁGINA
2.1 Análisis Comparativo Métodos Heurísticos con Métodos Tradicionales Iterativos	**19**
3.1 Análisis Comparativo Métodos Homotópicos con Métodos Tradicionales Iterativos	**27**

CAPÍTULO 1

CONCEPTOS Y FUNDAMENTOS

Para poder explicar cómo funcionan los métodos numéricos de convergencia global que se presentan en esta tesis es necesario enunciar los conceptos de: sistema de ecuaciones no lineales, función objetivo de minimización, homotopía y funciones homotópicas, función definida implícitamente, homeomorfismo, heurísticas y metaheurísticas. Además, es muy importante también enunciar dos teoremas muy populares entre los matemáticos que son: el teorema de la función implícita y el teorema de la función inversa. Esta teoría es fundamental para comprender el método numérico de la heurística y el método numérico homotópico que son los métodos de convergencia global que se analizan en este documento de tesis. A medida que se enuncian los conceptos y los teoremas se da una breve explicación de cuál es el papel que juegan dentro de los métodos numéricos.

1.1 SISTEMA DE ECUACIONES NO LINEALES

Un "sistema no lineal de n ecuaciones con n incógnitas" es de la forma:

$$\begin{cases} f_1(x_1, x_2, x_3, \text{K}, x_n) = 0 \\ f_2(x_1, x_2, x_3, \text{K}, x_n) = 0 \\ f_3(x_1, x_2, x_3, \text{K}, x_n) = 0 \\ \text{M} \\ f_n(x_1, x_2, x_3, \text{K}, x_n) = 0 \end{cases}$$

O en presentación vectorial mediante la ecuación $F(X) = \vec{0}$, donde:

$$X = (x_1, x_2, x_3, \text{K}, x_n),\ F(X) = \begin{pmatrix} f_1(x_1, x_2, x_3, \text{K}, x_n) \\ f_2(x_1, x_2, x_3, \text{K}, x_n) \\ f_3(x_1, x_2, x_3, \text{K}, x_n) \\ \text{M} \\ f_n(x_1, x_2, x_3, \text{K}, x_n) \end{pmatrix} \text{ y } \vec{0} \text{ es el vector nulo de } R^n$$

Las n funciones $f_1, f_2, f_3, \text{K}, f_n$ son, en general, funciones no lineales de n variables reales independientes y que tienen un dominio en común el cual se va a denotar como $\Omega \subseteq R^n$. Es decir, para cada $i = 1,2,3,\text{K},n$:

$$\begin{array}{lccc} f_i: & \Omega & \alpha & R \\ & X = (x_1, x_2, x_3, \text{K}, x_n) & \rightarrow & f_i(X) \end{array}$$

A estas n funciones las llamaremos "funciones componentes" del sistema no lineal [3]. Es importante indicar que se pueden tener sistemas no lineales en los cuales el número de ecuaciones puede ser diferente al número de incógnitas. En otros términos, se pueden tener sistemas no lineales con m funciones componentes cada una con n variables reales independientes donde $m \neq n$. Si tal es el caso, la implementación que proponemos para el método heurístico se puede usar sin inconvenientes, simplemente se tendría que definir adecuadamente la función objetivo de minimización empleando las m funciones componentes del sistema no lineal. Pero para el caso del método homotópico no serviría la implementación que se presenta en esta tesis ya que dicha implementación trabaja con matrices cuadradas y con inversas de matrices cuadradas. Si $m \neq n$ sería necesaria una generalización del concepto de matriz inversa, como por ejemplo, la matriz inversa de Moore-Penrose [6], lo cual se escapa del alcance de este documento, sin embargo, es una recomendación que se puede tomar en cuenta para un trabajo a futuro.

1.2 FUNCIÓN OBJETIVO DE MINIMIZACIÓN

Haciendo uso de las n funciones componentes $f_1, f_2, f_3, \text{K}, f_n$ de un sistema no lineal, se procede ahora a definir una función g a la que llamaremos "función objetivo de minimización" la cual tiene también obviamente su dominio en el conjunto Ω, tal como ocurre con las funciones componentes. Se define entonces g mediante la expresión:

$$g(X) = \sum_{i=1}^{n} |f_i(X)| = |f_1(X)| + |f_2(X)| + |f_3(X)| + \text{K} + |f_n(X)|$$

Note que la variable i del operador sigma va desde 1 hasta n. Si se tiene un sistema no lineal con n incógnitas pero con m ecuaciones, entonces el operador sigma que se

utiliza para definir a la función g debería considerar eso y terminar en m. En el método numérico de la heurística se resuelve un problema de minimización que, como ya se dijo, es equivalente al sistema no lineal. La función que se debe minimizar es precisamente la función g que se acaba de definir y la búsqueda del óptimo ocurre en el conjunto Ω.

1.3 FUNCIONES HOMOTÓPICAS Y HOMOTOPÍA

Suponga que las funciones $F, G : \Omega \subseteq R^n \to R^n$ son continuas. Se dice que F y G son homotópicas, si existe una función continua $H : \Omega \times [0,1] \to R^n$ tal que para toda $X \in \Omega$ ocurre que $H(X,0) = G(X)$ y $H(X,1) = F(X)$. Si tal es el caso, llamaremos a la función H una homotopía de G hacia F.

La idea del concepto de homotopía es que una función continua G se pueda "deformar continuamente" hasta convertirse en la función continua F. La función que permite esta transformación es la función continua H mediante el parámetro $\lambda \in [0,1]$. Se podría definir H de muchas maneras, pero es muy común trabajar con homotopías convexas y definir H mediante la expresión:

$$H(X,\lambda) = \lambda F(X) + (1-\lambda) G(X)$$

De esta manera, conforme λ varía continuamente de 0 hasta 1, la función H permite que G se deforme de manera continua hasta convertirse en F [6] [7].

¿Cómo este concepto nos puede ayudar para resolver el sistema no lineal $F(X) = \vec{0}$? La idea es la siguiente: debemos partir de un sistema no lineal $H(X,0) = G(X) = \vec{0}$ del cual se conozca una solución $\tilde{X}$. A medida que "movemos continuamente" λ desde el valor 0 hasta el valor 1, y tomando como base de partida la solución $\tilde{X}$, vamos resolviendo toda la familia de sistemas no lineales $H(X,\lambda) = \vec{0}$ hasta encontrar finalmente la solución deseada X^* del sistema no lineal $H(X,1) = F(X) = \vec{0}$ [6] [7].

¿Qué sistema $G(X)=\vec{0}$ utilizar? Tenemos también muchas respuestas a esta pregunta, pero es típico definir para algún $\tilde{X} \in \Omega$:

$$G(X)=F(X)-F\left(\tilde{X}\right)$$

Obviamente, debido a la manera en que se construyó la función G, es de observar que $\tilde{X}$ es una solución del sistema $G(X)=\vec{0}$. Note que el sistema $H(X,\lambda)=\vec{0}$ define, bajo ciertas condiciones, una curva continua y diferenciable $X=\psi(\lambda)$ en R^n con punto inicial en $\tilde{X}$ cuando $\lambda=0$ y punto final en X^* cuando $\lambda=1$. Es decir, se tiene la función:

$$\begin{aligned}\psi:[0,1] &\to R^n \\ \lambda &\to X=\psi(\lambda)\end{aligned}$$

La continuidad y la diferenciabilidad son condiciones deseadas sobre la función $X=\psi(\lambda)$ en todo su dominio. Consideremos las siguientes preguntas: ¿Cómo se puede garantizar que la curva $X=\psi(\lambda)$ existe, es única y es de clase C^1? ¿Cómo obtener la curva $X=\psi(\lambda)$ numéricamente? ¿Qué garantía existe para que podamos afirmar que $X^*=\psi(1)$ es una solución del sistema no lineal $H(X,1)=F(X)=\vec{0}$? Para responder a estas preguntas necesitamos algunas definiciones y teoremas que se presentan a continuación [6] [7].

1.4 FUNCIÓN DEFINIDA IMPLÍCITAMENTE

Sea $H : U \subseteq R^n \times R \to R^n$ una función con U conjunto abierto. Suponga que existe un punto $(X_0, \lambda_0) \in U$ tal que $H(X_0, \lambda_0) = \vec{0}$. Se dice que $H(X, \lambda) = \vec{0}$ define a X como una función implícita de λ en un entorno de (X_0, λ_0) si existe un conjunto abierto $V \subset R^n$ tal que $X_0 \in V$, un intervalo abierto $(\lambda_0 - \rho, \lambda_0 + \rho)$ con $\rho > 0$ como radio tal que $V \times (\lambda_0 - \rho, \lambda_0 + \rho) \subset U$, y una única función $\psi : (\lambda_0 - \rho, \lambda_0 + \rho) \to V$ tal que para todo $\lambda \in (\lambda_0 - \rho, \lambda_0 + \rho)$ la función $X = \psi(\lambda)$ es solución del problema $H(X, \lambda) = \vec{0}$ (Es decir que $H(\psi(\lambda), \lambda) = \vec{0}$).

Según el concepto anterior, si se cumplen ciertas condiciones, se puede garantizar la existencia de la función $X = \psi(\lambda)$ que define una curva en R^n con extremos en $\tilde{X}$ con $\lambda = 0$ y X^* con $\lambda = 1$. Se enuncia más adelante el teorema de la función implícita pero se lo hace de manera muy particular aplicado a nuestra función H [6] [7].

1.5 HOMEOMORFISMO

Se dice que la función $F : R^n \to R^n$ es un homeomorfismo si se cumple que F es una función biyectiva continua con función inversa F^{-1} también continua.

Recordar que un sistema no lineal de ecuaciones en forma vectorial se lo representa mediante la ecuación $F(X) = \vec{0}$. En la sección 3.1 se enuncia y se demuestra una "proposición 2", en la cual una de las hipótesis es que la función F es un homeomorfismo. Este supuesto es muy importante ya que nos permite probar que existe un problema de valor inicial que es equivalente al problema de resolver un sistema de ecuaciones no lineales. Además, permitirá probar que existe solución y que dicha solución es única para el problema de valor inicial sin importar el dato inicial que se tome $\tilde{X} \in R^n$. Esto garantiza la convergencia del método homotópico [6] [7].

1.6 TEOREMA DE LA FUNCIÓN IMPLÍCITA

Sea $H : U \subset R^n \times R \to R^n$ una función de clase $C^1(U)$ con U conjunto abierto. Si se tiene el punto $(X_0, \lambda_0) \in U$ tal que $H(X_0, \lambda_0) = \vec{0}$, y además $H_X(X_0, \lambda_0)$ es una matriz inversible, entonces la ecuación $H(X, \lambda) = \vec{0}$ define implícitamente la función $X = \psi(\lambda)$ en un entorno abierto del punto (X_0, λ_0). Además, ψ es también una función de clase C^1 en algún intervalo alrededor de λ_0 y es única.

En el enunciado anterior, H_X representa la derivada parcial de la función H respecto de la primera variable, por lo que se trata de una matriz Jacobiana.
Este teorema es muy importante porque proporciona condiciones suficientes para garantizar la existencia y unicidad de la curva $X = \psi(\lambda)$ en el intervalo $\lambda \in [0,1]$. No sólo eso, también garantiza que $X = \psi(\lambda)$ es de clase $C^1[0,1]$. Dependiendo del problema a resolver, puede ocurrir que $U = R^n \times R$. Lo importante es que el dominio de la función H de donde se considera la existencia, unicidad, continuidad y diferenciabilidad continua de $X = \psi(\lambda)$ contenga al conjunto $\Omega \times [0,1]$ de nuestro interés. Dependiendo del problema, podría no ocurrir esto. Es importante entonces estudiar cómo se comporta la curva para el problema en particular que se desea resolver. La curva podría no llegar a alcanzar el punto para el cual $\lambda = 1$ o podría recorrer una distancia "infinita" desde $\lambda = 0$ hasta $\lambda = 1$ debido a su longitud no finita [6] [7].

Procedemos ahora a enunciar el popular teorema de la función inversa para funciones de un subconjunto de R^n en R^n. Este teorema es de gran utilidad para discutir una proposición que se enunciará y demostrará más adelante.

1.7 TEOREMA DE LA FUNCIÓN INVERSA

Sea $F:U\subseteq R^n\to R^n$ una función de clase C^1. Sea $X_0\in U$ tal que la matriz Jacobiana $F'(X_0)$ es inversible y $F(X_0)=Y_0$. Entonces, existen conjuntos abiertos V y W tales que $X_0\in V$, $Y_0\in W$ y $F:V\to W$ es una función biyectiva. Por lo tanto la función inversa $F^{-1}:W\to V$ existe y es también de clase C^1. Además, ocurre que $(F^{-1})'(Y_0)=[F'(X_0)]^{-1}$ [6].

1.8 HEURÍSTICAS Y METAHEURÍSTICAS

La palabra heurística proviene de un término griego cuyo significado es "hallar", "descubrir". Se puede entonces definir una heurística como una técnica para resolver problemas donde el objetivo principal es buscar. Dicha búsqueda está basada en un conjunto de reglas, es decir, se trata de un procedimiento sistemático y organizado. Los pasos a seguir en una heurística nos dirán cómo proceder y qué inconvenientes evitar durante una búsqueda en particular [4] [5]. Por ejemplo, en los problemas de optimización se usan las heurísticas. Un modelo matemático típico de optimización es de la forma:

$$MAX\ \ o\ \ MIN\ \ f(X)\in R$$
$$SUJETO\ \ A\ \ X\in\Omega\subseteq R^n$$

El objetivo en el problema de optimización es buscar aquel elemento $X^*\in\Omega$ que hace que la función f tome el valor más alto posible (MAX), o que permite que f tome el valor más bajo posible (MIN). La función f puede ser lineal o no lineal y se la conoce como la función objetivo. Se dice que el problema es sin restricciones si $\Omega=R^n$ y con restricciones si estrictamente $\Omega\subset R^n$. Al conjunto Ω se lo conoce como el conjunto de soluciones factibles o espacio de búsqueda [4] [5]. Existen métodos exactos para resolver problemas de optimización, y como el nombre mismo lo dice, son métodos que nos permiten conocer la solución óptima, es decir, la solución exacta del problema de optimización. La función objetivo y sobre todo el tamaño del espacio de búsqueda pueden complicar el encontrar el óptimo de un problema de optimización. Al ejecutar

un algoritmo que está intentando encontrar la solución de un problema de optimización, los recursos computacionales actuales, dígase principalmente memoria principal y el procesador, son también un factor importante que puede complicar muchísimo el encontrar el óptimo al menos en un tiempo prudencial [4] [5].

Las heurísticas son métodos inexactos para problemas de optimización, es decir, se trata de métodos que no ofrecen ningún tipo de garantía para encontrar la solución exacta, pero en la práctica son capaces de ofrecer muy buenas soluciones, es decir, soluciones viables en el mundo real que están "cerca" de la solución óptima del problema, y que esa solución cercana y buena la pueden encontrar en un tiempo computacional aceptable. Lo dicho anteriormente es la principal razón por la cual las heurísticas son muy populares y son la técnica preferida en muchos problemas de optimización que son muy complicados [4] [5]. En resumen, la heurística es un método que desea encontrar buenas soluciones (preferiblemente las óptimas) en buenos tiempos de ejecución.

El término heurística se usa también para referirse a algoritmos que resuelven problemas de propósito particular. Si se tiene un tipo de problema de optimización con ciertas características que lo hacen especial, distinto, diferente, de otros tipos de problemas de optimización, se puede diseñar una heurística, es decir, un algoritmo que resuelve ese tipo de problema puntual de optimización. Por ejemplo, es muy conocido el algoritmo de Dijkstra para encontrar el camino más corto entre dos vértices de un grafo.

Por el contrario, una metaheurística es un algoritmo que resuelve problemas generales de optimización, es decir, se le puede aplicar una misma metaheurística a muchos tipos de problemas de optimización sin importar lo diferente que sean. Una metaheurística se la puede ver también como una heurística para heurísticas, es decir, cuando se emplea una metaheurística para resolver un problema de optimización en particular, la metaheurística como parte de las tareas que debe realizar, se apoyará en heurísticas de propósito particular, para resolver el problema de optimización [4] [5].

Existen muchas metaheurísticas muy conocidas, entre ellas: el algoritmo genético, la búsqueda tabú, el recocido simulado, GRASP, que pueden utilizarse, junto con heurísticas, para resolver problemas de optimización. Scatter search es también un

método metaheurístico que se puede emplear para resolver problemas de optimización y forma parte de la familia de los algoritmos evolutivos. Tiene sus orígenes en los años setenta, pero es en la última década cuando ha sido probado en muchos problemas con un alto grado de dificultad y con excelentes resultados [4] [5].

Un objetivo secundario de este documento de tesis es el de explicar cada una de las fases de "scatter search" o en español "búsqueda dispersa", es decir, indicar en detalle cómo trabaja internamente el algoritmo. Para tal efecto, se ha construido un software que optimiza funciones continuas de varias variables reales sobre un conjunto compacto Ω utilizando la metaheurística de "scatter search". A medida que se explique cómo internamente hace su trabajo el algoritmo de búsqueda dispersa, se indicará también en qué momento utiliza heurísticas de propósito particular, y cuál es el trabajo que estas heurísticas realizan.

CAPÍTULO 2

EL MÉTODO NUMÉRICO HEURÍSTICO

Este capítulo trata exclusivamente el primer método de convergencia global de este documento de tesis que se lo ha llamado el método numérico de la heurística. En la primera parte se probará que se puede expresar un sistema no lineal de ecuaciones en términos de un problema de optimización [1] [2] [3]. Luego, para resolver el problema de optimización se utiliza la metaheurística de búsqueda dispersa. Se explica en detalle todo lo que tiene que ver con el diseño del algoritmo y su funcionamiento interno. Se finaliza el capítulo haciendo un análisis comparativo de esta nueva técnica para resolver sistemas no lineales con los métodos iterativos convencionales.

2.1 EL PROBLEMA DE OPTIMIZACIÓN

A continuación se va a enunciar y a demostrar matemáticamente una proposición que justifica el plantear un sistema de ecuaciones no lineales en términos de un problema de minimización.

Proposición.- El vector $X^* = \left(x_1^*, x_2^*, x_3^*, \text{K}\, , x_n^*\right) \in \Omega$ es una solución de $F(X) = \vec{0}$ (es decir: $F\left(X^*\right) = \vec{0}$), si y sólo si, la función g (definida anteriormente) toma como valor mínimo en X^* el valor 0

Demo

Primera Parte: Supongamos que $F\left(X^*\right) = \vec{0}$. Entonces para cada $i = 1,2,3,\text{K}\, ,n$ ocurre que $f_i\left(X^*\right) = 0$. Luego: $g\left(X^*\right) = \sum_{i=1}^{n} \left| f_i\left(X^*\right) \right| = 0$

Además para toda $i = 1,2,3,\text{K}\, ,n$ y para toda $X \in \Omega$ se tiene que $\left| f_i(X) \right| \geq 0$, por lo que $g(X) \geq 0$ para toda $X \in \Omega$.

Entonces la función g toma el valor mínimo cero en $X^* \in \Omega$

Segunda Parte: Ahora supongamos que la función g toma el valor mínimo cero en $X^* \in \Omega$. Se debe entonces demostrar que $F(X^*) = \vec{0}$.

Por hipótesis $g(X^*) = 0$. Entonces $g(X^*) = \sum_{i=1}^{n} |f_i(X^*)| = 0$

Es decir $|f_1(X^*)| + |f_2(X^*)| + |f_3(X^*)| + \mathrm{K} + |f_n(X^*)| = 0$

Se debe probar que para toda $i = 1,2,3,\mathrm{K}\ ,n$ ocurre que $f_i(X^*) = 0$. Hagamos la prueba por contradicción. Supongamos que lo anterior es falso, es decir, que existe una $1 \le i \le n$ tal que $f_i(X^*) \neq 0$. Si esto es así, sucede que $g(X^*) = \sum_{i=1}^{n} |f_i(X^*)| = |f_1(X^*)| + |f_2(X^*)| + |f_3(X^*)| + \mathrm{K} + |f_n(X^*)| > 0$ lo que obviamente contradice la hipótesis que dice que $g(X^*) = 0$, lo cual no puede ser. Por lo tanto, para toda $1 \le i \le n$ se da que $f_i(X^*) = 0$ y $F(X^*) = \vec{0}$. Es decir, X^* es una solución de $F(X) = \vec{0}$.

De aquí en más, para resolver un sistema no lineal en particular, lo primero que haremos es construir la función g utilizando las funciones $f_1, f_2, f_3, \mathrm{K}\ , f_n$ que componen el sistema. Luego, el conjunto Ω, para la ejecución del algoritmo, lo vamos a suponer de la forma:

$$\Omega = \left\{ (x_1, x_2, x_3, \mathrm{K}\ , x_n) \in R^n \,/\, \alpha_i \le x_i \le \beta_i \ \ para\ cada\ i = 1,2,3,\mathrm{K}\ ,n \right\}$$

Es decir, para resolver el sistema de ecuaciones no lineales, se debe minimizar la función no lineal g y se deben fijar cotas inferiores y superiores para cada una de las variables del sistema. Se espera que exista una solución para el sistema no lineal en el conjunto Ω, para que la solución encontrada en el problema de optimización la podamos tomar también como solución del sistema no lineal. Debido a que el sistema no lineal es el modelo matemático de alguna aplicación o problema del mundo real, se puede construir un conjunto Ω donde se sospecha o se considera se encuentra la solución del sistema no lineal. La experiencia del modelador, o de las personas

involucradas en el problema que es representado por el sistema no lineal, es vital y sumamente importante para definir dicho conjunto Ω.

2.2 EL ALGORITMO DEL MÉTODO HEURÍSTICO

El problema de optimización que se pretende resolver es un problema de minimización en el cual la función objetivo $g(x_1, x_2, x_3, \text{K}, x_n)$ es una función no lineal y, cada una de las variables independientes de g, está acotada tanto inferior como superiormente. Es decir, para cada $i = 1,2,3,\text{K}, n$ ocurre que $x_i \in [\alpha_i, \beta_i]$. En otras palabras:

$$\Omega = \prod_{i=1}^{n} [\alpha_i, \beta_i] = [\alpha_1, \beta_1] \text{X} [\alpha_2, \beta_2] \text{X} [\alpha_3, \beta_3] \text{XK X} [\alpha_n, \beta_n] \subseteq R^n$$

Para ejecutar el algoritmo necesitamos ingresar entonces la función g y, los n intervalos para cada una de las n variables donde se cree que se encuentra la solución del sistema no lineal.

Algo importante de mencionar es que al usar cualquier algoritmo metaheurístico estamos concientes que no existe garantía de encontrar el óptimo del problema, pero podemos encontrar una "buena solución" en un tiempo prudencial y con un costo computacional aceptable, en comparación con cualquier método exacto que se pueda usar para el problema de optimización [1] [2] [3]. Ahora, ¿por qué el algoritmo "scatter search"? Porque ha sido muy ampliamente usado y con mucho éxito para la optimización de funciones lineales y no lineales de varias variables sobre un espacio de soluciones que puede ser continuo y también discreto [4] [5]. A continuación se explica cada etapa del algoritmo.

2.2.1 GENERACIÓN DEL CONJUNTO P

Lo primero que debe hacerse al resolver un problema de optimización con "scatter search" es construir un conjunto al que llamaremos P. Una entrada del algoritmo entonces es el tamaño del conjunto P, definida con la variable $PSize$. Los elementos del conjunto P serán vectores de R^n que pertenecen al conjunto Ω. Es decir, cada elemento del conjunto P es una solución factible del problema de minimización. Estos elementos se van a generar aleatoriamente pero de forma tal que podamos cubrir o abarcar todo el conjunto Ω, para así tener diversidad en esta primera generación de soluciones factibles. Sin embargo, el tema de la calidad es también importante en esta

primera etapa del algoritmo. A continuación se explica cómo se van a generar los vectores.

Cada variable independiente x_i de la función g "vive" en un intervalo $[\alpha_i, \beta_i]$. Se va a dividir o a particionar cada intervalo $[\alpha_i, \beta_i]$ en m subintervalos. El número de subintervalos de igual tamaño en el que se va a dividir el intervalo $[\alpha_i, \beta_i]$ se va a representar entonces con la variable m, el cual es también un parámetro de entrada para el algoritmo. De esta manera, la longitud de cada subintervalo (todos los subintervalos para la variable x_i tienen la misma longitud) es: $\Delta x_i = (\beta_i - \alpha_i)/m$

Ahora se van a generar *PSize* elementos con una mezcla de diversidad y calidad. Primero hacemos uso de la diversidad, ¿cómo trabaja? Así: Para cada variable x_i se selecciona aleatoriamente un subintervalo (se genera un número entero aleatorio entre 1 y m), luego se genera un número aleatorio de punto flotante dentro del subintervalo seleccionado anteriormente, lo que va a representar el valor de la *i*-ésima componente del vector. Esto se debe hacer entonces por cada una de las n componentes del vector de R^n. Finalmente, se aplica un método de mejora al vector que se acaba de construir, en otras palabras, a través de una "local search" se busca si existe "cerca" de la solución generada otro vector que mejora el comportamiento de la función objetivo (es decir que minimiza la función g). Posteriormente se explicará la forma de trabajar de la búsqueda local. Pero, para asegurar la diversificación de los elementos en el conjunto P, la probabilidad de elegir un subintervalo, para una variable cualquiera x_i, es inversamente proporcional a la cantidad de números de punto flotante generados en dicho subintervalo de esa variable x_i. De esta manera se trata de "balancear" las soluciones a lo largo de todo el conjunto Ω. Como podemos apreciar, se tiene ahora un conjunto P con una mezcla muy buena de calidad y diversidad.

2.2.2 BÚSQUEDA LOCAL

La "local search" o búsqueda local implementada recibe como entrada la solución a mejorar y un parámetro adicional que le dice al método el número de intentos que debe hacer para buscar un "mejor" vector en una vecindad del vector inicial, este número de intentos está en función del número de veces que se va a evaluar la función g. ¿Cómo trabaja? A la i-ésima componente del vector inicial se le genera un número aleatorio que se le suma o se le resta a esa i-ésima componente (lo cual también es aleatorio) de manera tal que se obtiene un nuevo valor para esa i-ésima componente y, de esta forma, al efectuar esto n veces, se obtiene un nuevo vector. Aquí se debe tener todo el cuidado de que el valor de la i-ésima componente del nuevo vector respete los límites definidos del subintervalo al cual pertenece el valor de la i-ésima componente del vector original (esto es lo que hace que la búsqueda sea local, el nuevo valor de cada variable no se debe salir del subintervalo en el que se encuentra el valor original de esa variable). Cada vez que se construye un vector en la vecindad del vector original se ha realizado "un intento". Como resultado del intento se tiene un nuevo vector que es descartado si no mejora la función g o tomado como el nuevo vector original si la función g mejora con este nuevo vector. Después de una cantidad de intentos (lo que es un argumento de este procedimiento) se detiene la búsqueda. Si se ha llegado al número de intentos tope de búsqueda definido a través del argumento "número de evaluaciones" y no se ha encontrado una mejor solución, el método devuelve la primera solución que se pasó como argumento, es decir, se indica así que "nos quedamos igualitos", no se encontró nada mejor. Cabe indicar que otro método que se implementó para la búsqueda local es el conocido algoritmo heurístico "Nelder-Mead" propuesto por Jhon Nelder y Roger Mead en 1965 para la optimización continua de una función objetivo sin restricciones. Este algoritmo sólo utiliza la regla de correspondencia de la función a optimizar y no utiliza las derivadas parciales de ningún orden para dicha función, motivo por el cual fue escogido también. Además se trata de un algoritmo que ha tenido mucho éxito en diversos problemas presentando buenos resultados, pero hay que tener presente que se trata de un algoritmo local y no global. De esta manera, el algoritmo scatter search que se ha construido tiene la capacidad de ejecutar con cualquiera de los dos algoritmos de búsqueda local antes indicados.

2.2.3 CONSTRUCCIÓN DEL CONJUNTO R

En el método de "scatter search" existe un segundo conjunto importante, el conjunto R, el cual es construido a partir del conjunto P. La cardinalidad de este conjunto R es también un parámetro del algoritmo. Para construir el conjunto R nos preocupamos que la variable $RSize$, que define el tamaño de R, sea par, ya que la mitad de los elementos de R serán escogidos por calidad mientras que la otra mitad por diversificación, como ya se dijo, haciendo uso del conjunto P. Aquí es importante indicar que se pudo haber parametrizado esto, es decir, que b_1 elementos sean escogidos por calidad y b_2 por diversidad teniendo presente que $b_1 + b_2 = RSize$. Retomando, se toman entonces los $RSize/2$ "primeros en calidad" elementos del conjunto P (es decir, los mejores elementos de dicho conjunto) y se incorporan al conjunto R. ¿Cómo escoger los elementos restantes?

De los elementos que quedan de P se deben tomar $RSize/2$ elementos más para completar el conjunto R, ¿cómo se lo hace para garantizar diversidad? Se calcula la distancia de cada elemento de P (ya no se consideran aquellos que se llevan a R) al conjunto R con la fórmula: Sea $v \in P$, luego $d(v,R) = \underset{w \in R}{MIN} \| v - w \|$ y el elemento que se almacena en R es aquel elemento de P que dista más de R, es decir, el más lejano al conjunto R. La norma utilizada para calcular la distancia mencionada es la norma euclidiana.

En este momento se tiene ya construido totalmente el conjunto R con una mezcla muy buena y adecuada de calidad y diversidad.

2.2.4 ACTUALIZACIÓN Y MANTENIMIENTO DEL CONJUNTO R

Durante la búsqueda del óptimo el conjunto R debe actualizarse muchas veces y eso es lo que se va a explicar. Cabe indicar aquí que el número de actualizaciones del conjunto R es también un parámetro del algoritmo. Una vez construido el conjunto R se procede a crear subconjuntos de R. En esta implementación el tamaño de cada subconjunto es de dos elementos de R (pero en otras implementaciones con scatter search la cardinalidad de estos subconjuntos puede ser de más elementos), es decir, usando R se generan todas las parejas posibles de elementos sin repetir, porque se les va a aplicar un método de combinación. ¿Cómo trabaja el método de combinación? Eso

se explica unas líneas más adelante. Al aplicar un método de combinación a cada subconjunto, se generan cuatro elementos y, a cada uno de estos nuevos elementos (llamados vectores prueba) se les aplica un método de mejora, es decir, se emplea la operación de búsqueda local explicada anteriormente para estudiar la vecindad de estas "soluciones prueba" en busca de unas mejores soluciones. Entonces, por cada subconjunto se tienen cuatro "soluciones mejoradas" que se añaden a un conjunto denominado $Pool$. Es decir que finalmente en $Pool$ se tiene el resultado de aplicar métodos de "combinación" y "mejora" a cada uno de los subconjuntos construidos a partir del gran conjunto R.

Una vez que el conjunto $Pool$ está lleno se construye un conjunto de "vida corta" $Temp$ de forma tal que ahí se almacenen todos los elementos de R y de $Pool$, es decir $Temp = R \cup Pool$. Este conjunto temporal, tiene una cardinalidad igual a la cardinalidad de R más la cardinalidad de $Pool$. Los $RSize$ mejores elementos son los que pasan finalmente al conjunto R. ¿Cuántas veces ocurre esto? Hay un parámetro llamado "Número de Iteraciones" que se debe ingresar antes de correr el algoritmo de scatter search, pues este parámetro es el que sirve para configurar el número de veces que el conjunto R se va a actualizar.

La opción de reconstruir el conjunto R, aclaramos, no se ha incorporado en la implementación. Es decir, si el conjunto R permanece sin cambios después de una actualización (debido a que el "combinar y mejorar" no produce nada nuevo) se recomienda una reconstrucción donde participa nuevamente el conjunto P y el conjunto R se actualiza nuevamente por calidad y diversidad. No se implementó esta etapa debido a que no se lo consideró necesario ya que en las pruebas que se realizaron el algoritmo siempre encontraba soluciones muy cercanas al óptimo, es decir, el software encontraba correctamente valores muy próximos al óptimo del problema de minimización en todos los ejemplos de sistemas no lineales en los que se lo puso a prueba. Cabe indicar que se probó el algoritmo en sistemas no lineales de hasta cinco ecuaciones y cinco incógnitas, y con conjuntos Ω "pequeños" donde se sospechaba se encontraba la o las soluciones. Pero sería importante considerar la reconstrucción del conjunto R si se tienen sistemas no lineales de mayor tamaño.

2.2.5 MÉTODO DE COMBINACIÓN

El método de combinación implementado es muy sencillo y opera así: se generan cuatro números aleatorios de punto flotante en el intervalo $(0,1)$ con la idea de aplicar combinaciones lineales "convexas". Es decir, el método de combinación lo único que recibe como parámetro es el subconjunto de R que tiene dos soluciones $SolA$ y $SolB$, y devuelve como salida cuatro combinaciones lineales convexas distintas. Es importante decir que esta salida no pasa directamente a R ya que antes se aplica el método de mejora a cada una de las cuatro soluciones obtenidas por combinación, y ellas sí pasan ya al conjunto R.

2.2.6 MÉTODO PRINCIPAL DEL ALGORITMO

A continuación se va a mostrar el método principal llamado "EjecutarSS" que gobierna el comportamiento y la operación general del algoritmo de scatter search que se implementó. Todas las etapas o fases anteriores coordinan y se enlazan en el método principal. A continuación la figura 2.1 nos muestra las etapas explicadas anteriormente y luego se presenta el método principal del algoritmo:

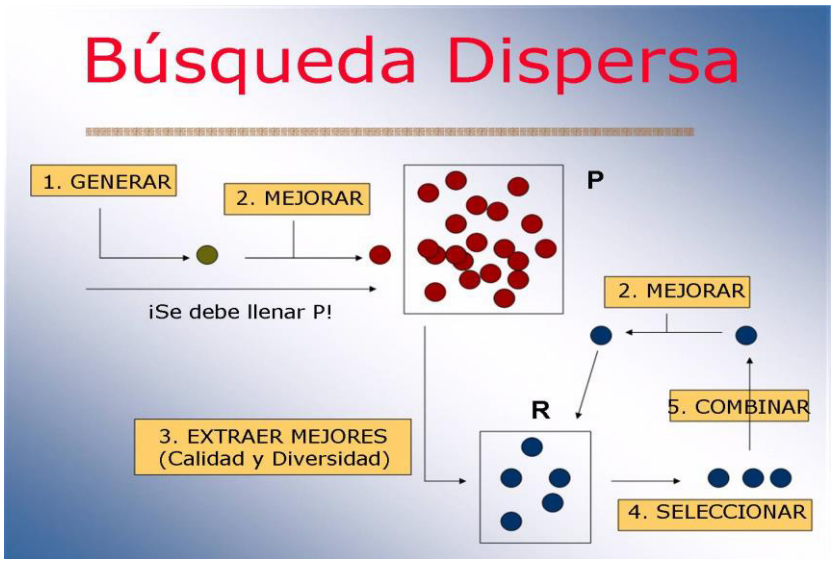


Figura 2.1: Etapas Búsqueda Dispersa

```
public double[] EjecutarSS(out double MejorValor)
{
    ArrayList Pool = new ArrayList();
    ArrayList EvalPool = new ArrayList();
    Stopwatch Tiempo;
    double[,] MejoresValores =
    new double[this.NumIteraciones, 2];

    Tiempo = Stopwatch.StartNew();
    this.GenerarP();
    this.ConstruirR(this.P, this.EvalP, this.R, this.EvalR);

    ArrayList SubSets = new ArrayList();
    int Iteraciones = 0;
    while (true)
    {
        this.GenerarSubSets(SubSets, R);
        while (SubSets.Count != 0)
        {
            double[][] SubSet = (double[][])SubSets[0];
            double[][] TrialSolutions = this.Combinar(SubSet);
            double Valor1 = this.LS(TrialSolutions[0], 100);
            Pool.Add(TrialSolutions[0]);
            EvalPool.Add(ValorUno);
            double Valor2 = this.LS(TrialSolutions[1], 100);
            Pool.Add(TrialSolutions[1]);
            EvalPool.Add(ValorDos);
            double Valor3 = this.LS(TrialSolutions[2], 100);
            Pool.Add(TrialSolutions[2]);
            EvalPool.Add(ValorTres);
            double Valor4 = this.LS(TrialSolutions[3], 100);
            Pool.Add(TrialSolutions[3]);
            EvalPool.Add(ValorCuatro);
            SubSets.RemoveAt(0);
        }
        this.ActualizarR(R, EvalR, Pool, EvalPool);
        MejoresValores[Iteraciones, 0] = EvalR[0];
        MejoresValores[Iteraciones, 1] =
        Tiempo.ElapsedMilliseconds;
        Iteraciones++;
        if (Iteraciones == this.NumIteraciones) break;
    }
    MejorValor = EvalR[0];
    return R[0];
}
```

En el capítulo 4 se muestran resultados de los experimentos computacionales efectuados con el algoritmo del método de la heurística. Se efectuaron pruebas con sistemas no lineales de hasta 5 ecuaciones con 5 incógnitas, en los cuales se obtuvieron muy buenas aproximaciones. Cabe indicar que la precisión depende del conjunto de búsqueda Ω.

2.3 ANÁLISIS COMPARATIVO MÉTODOS CLÁSICOS

A continuación en la tabla 2.1 se muestra un análisis que compara los métodos heurísticos con los métodos iterativos convencionales para la solución de sistemas no lineales. Se presentan ventajas y desventajas de cada uno, pero de manera comparativa. Es importante indicar que básicamente lo único que tienen en común ambos métodos es que resuelven sistemas no lineales, pero la forma en la que encaran el problema es radicalmente distinta.

Métodos Heurísticos	
Ventajas	**Desventajas**
Dominio de convergencia ampliado debido a las estrategias de "diversificación"	La implementación del algoritmo es más complicada, existen muchas fases o etapas involucradas
Se los puede utilizar para determinar el "dato inicial" que necesitan los algoritmos iterativos convencionales o para futuras corridas del mismo algoritmo en un espacio de búsqueda más reducido	No existe garantía de encontrar el óptimo del problema, son algoritmos que utilizan en su trabajo la aleatoriedad. Además, pueden quedarse atrapados en óptimos locales
Las funciones de varias variables que componen el sistema no necesitan cumplir con condiciones de continuidad y diferenciabilidad de ningún orden	Espacio de búsqueda considerable, en general convergencia lenta. La velocidad de convergencia depende del tamaño del espacio de búsqueda
No se necesitan cálculos matriciales tediosos ni calcular derivadas parciales de ningún orden	No se puede estimar el número de iteraciones necesarias para estar tan cerca de la solución exacta como se desee.
En una misma ejecución se pueden encontrar muchas soluciones, si es que existen	La técnica principal en que basan su trabajo es una computación evolutiva. Existe muy poca teoría matemática detrás del método que estudie o analice convergencia y error

Métodos Iterativos Clásicos	
Ventajas	**Desventajas**
La implementación del algoritmo es más sencilla, usualmente mediante fórmulas recursivas	Normalmente un reducido dominio de convergencia
Si el "dato inicial" se encuentra dentro de la región de convergencia se garantiza que el método converge a la solución exacta	Pueden tener dependencia de otros métodos que proporcionen el "dato inicial" para poder comenzar con las iteraciones
En general se tiene convergencia al menos lineal, es decir, son métodos rápidamente convergentes	Dependiendo del método que se utilice, las funciones de varias variables empleadas deben ser de clase C^1 o de clase C^2 o incluso superior, o deben satisfacer o cumplir la condición de Lipschitz
Se puede tener control sobre el error. Se puede estimar el número de iteraciones para que la distancia entre la solución exacta y el valor estimado no supere la tolerancia deseada	Por cada iteración muchos cálculos matriciales (en algunos casos el cálculo de matrices inversas y producto de matrices) y cálculos de derivadas parciales
Muchos teoremas y proposiciones que estudian convergencia y error de los métodos	Si el sistema no lineal tiene muchas soluciones, sólo una se puede encontrar por cada corrida

Tabla 2.1: Análisis Comparativo Métodos Heurísticos con Métodos Tradicionales Iterativos

2.4 ANÁLISIS DE CONVERGENCIA

El método heurístico que se propone en esta tesis es numérico en el sentido que busca aproximar o estimar, con la mayor precisión posible, soluciones para un sistema no lineal de ecuaciones [10] [11] [13], pero al mismo tiempo es un método no determinista debido a los componentes aleatorios que posee [4] [5]. Cabe indicar que no todo se deja al azar, la aleatoriedad se usa de manera criteriosa. Existen reglas que guían la búsqueda de forma que se pueda explorar el espacio de soluciones eficientemente incorporando además mecanismos que intentan evitar caer en óptimos locales. Es decir, se trata de un procedimiento de búsqueda sistemático con moderados tintes aleatorios [4] [5].

El objetivo en esta sección es analizar, dentro de lo posible, el tema de la convergencia del método heurístico, si vale la utilización del término. Los métodos numéricos convencionales, como el conocido método de Newton, son métodos deterministas y se puede efectuar sin muchas complicaciones un análisis de convergencia ya que existe una expresión claramente definida que permite encontrar, conforme se itera, el siguiente elemento de la sucesión si conocemos el elemento anterior, es decir, se tiene una regla de correspondencia para la sucesión, que obviamente depende del dato inicial [11] [13]. En otras palabras, si al ejecutar el método de Newton una y otra vez, usamos el mismo sistema no lineal y partimos del mismo dato inicial, se construirá exactamente la misma sucesión en todas las corridas. Suponga que dado un dato inicial X_0 se ha iterado n veces con el método de Newton y se tiene el elemento X_n de la sucesión vectorial que construye dicho método. Bajo estas condiciones cuando se desea resolver el sistema $F(X)=\vec{0}$ se puede conocer con certeza el elemento $n+1$ de la sucesión, el cual es:

$$X_{n+1} = X_n - \left[F'(X_n)\right]^{-1} F(X_n)$$

El análisis de convergencia para el método heurístico es muchísimo más complicado debido a lo no determinista de su comportamiento. Si ejecutamos el método heurístico una y otra vez, con el mismo sistema no lineal y con los mismos parámetros de entrada, se van a generar distintas sucesiones, es decir, se van a obtener resultados diferentes. En el método de Newton, bajo ciertas condiciones, se puede garantizar la convergencia y los tiempos de ejecución son pequeños en comparación a los tiempos de ejecución del método heurístico que son mucho más considerables y donde además no existe garantía de alcanzar el óptimo global [4] [5] [9]. El caer en un óptimo local es algo que puede ocurrir en general con cualquier metaheurística, no sólo con búsqueda dispersa [4] [5].

Lo que procedo a efectuar ahora es, basado en la operación interna del método heurístico, una discusión sobre la convergencia del método intentando ser lo más formal posible desde una perspectiva matemática.
Recordar que el método heurístico intenta minimizar la función g no lineal definida en la sección 1.2 sobre un conjunto compacto Ω, el cual es una región hiperrectangular en el espacio vectorial R^n. En términos matemáticos, el método heurístico construye dos sucesiones importantes, no sólo una sucesión como lo hacen los métodos de convergencia local, particularmente el método de Newton. La primera sucesión $\{X_n\}$ es vectorial y la segunda sucesión $\{y_n\}$ es escalar.

Una vez que entra en ejecución el método heurístico el primer paso es generar aleatoriamente la cantidad de vectores que permitan "llenar" el conocido conjunto P, el cual tiene un tamaño definido $PSize$. En esta generación es importante intentar distribuir los vectores uniformemente a lo largo de la región que define Ω. Pero estos vectores no ingresan directamente a P, antes se busca localmente si "cerca" de cada vector generado existe otro "mejor" y, si tal es el caso, se almacena este último en el conjunto P y no el vector construido inicialmente. Sea entonces v un vector resultante (generación + mejora). Si $v \notin P$ entonces $P = P \cup \{v\}$, caso contrario se descarta el vector v. Este paso se repite hasta llenar todo el conjunto P. Observe un primer componente aleatorio en esta primera etapa del método heurístico [4] [5].

El segundo paso consiste en llenar un conjunto R con cardinalidad $RSize$. Tomamos los mejores $RSize/2$ elementos del conjunto P para ubicarlos en R. Luego tomamos los $RSize/2$ elementos del conjunto P que se encuentren más lejos del conjunto R. Note como los aspectos de calidad y diversificación son considerados también en la construcción del conjunto R. Aún no se genera el primer elemento de las sucesiones de interés $\{X_n\}$ y $\{y_n\}$.

El tercer paso consiste en formar parejas de manera aleatoria entre los elementos del conjunto R. Con estas parejas se realizan cuatro combinaciones lineales convexas. Es decir, por cada pareja se generan cuatro elementos más, a los cuales se les aplica una búsqueda local. Al final se tienen, además de los elementos del conjunto R, una cantidad adicional de elementos generados que fueron mejorados. De entre todos estos elementos, los $RSize$ mejores son los que formarían ahora parte del conjunto R. En este momento el vector X_1 es el mejor elemento del conjunto R mientras que $y_1 = g(X_1)$. Ya tenemos entonces el primer elemento de cada sucesión.

Cada iteración en el método heurístico consiste en una actualización del conjunto R y, por cada actualización de R, se obtiene como resultado un nuevo elemento para cada sucesión [4] [5]. El tercer paso explicado anteriormente es lo que ocurre en cada iteración, es decir, nuevamente se forman parejas entre todos los elementos del actual conjunto R para hacer combinaciones lineales convexas que serán posteriormente mejoradas mediante una búsqueda local y, de entre todos los elementos que se tienen, incluidos los elementos de R, se toman los $RSize$ mejores para que sean los elementos del "nuevo" conjunto R. En cada iteración entonces se obtiene el nuevo elemento de cada sucesión.

Por lo explicado anteriormente es que se puede afirmar que una iteración en el método heurístico consume más recursos computacionales (memoria + tiempo de procesador) que una iteración en cualquier método clásico de convergencia local. Es muy notorio que en una iteración del método heurístico existen más pasos, más tareas y por ende más operaciones aritméticas que en una iteración de Newton por ejemplo.

Supongamos que nos encontramos en la iteración n del método heurístico. El vector X_n es el "mejor" elemento del conjunto R de esa iteración n por lo que ocurre también que $y_n = g(X_n)$.

Supongamos también que el sistema no lineal tiene al vector X^* como una solución en la región Ω escogida para la ejecución del método heurístico. Es decir que existe $X^* \in \Omega$ tal que $y^* = g(X^*) = 0$ y se puede afirmar que la función g toma el valor mínimo cero en el punto $X^* \in \Omega$, tal como se explicó en la primera parte de la demostración del teorema de la sección 2.1 de este documento de tesis.

El método heurístico tiene memoria en el sentido que el "mejor" elemento del conjunto R que resulta de cada iteración, es decir, cada elemento de la sucesión $\{X_n\}$, es almacenado con el fin de no desmejorar. Quiere decir que en el peor de los casos, si después de la formación de parejas, las combinaciones lineales convexas y las búsquedas locales, no se obtiene un elemento mejor, tomamos entonces como mejor elemento de esa iteración, el mejor elemento del conjunto R de la iteración anterior [4] [5]. Se trata entonces de un proceso evolutivo, conforme itera el método heurístico el objetivo siempre es mejorar. En términos matemáticos eso quiere decir que:

$$\forall n \in N \quad y_{n+1} \leq y_n$$

En palabras quiere decir que la sucesión $\{y_n\}$ es decreciente. En cálculo es conocido el "teorema de las sucesiones convergentes" que nos dice en su enunciado que si una sucesión cualquiera es decreciente y además está acotada inferiormente, entonces esa sucesión es convergente. Es más, nos dice también que el valor al cual converge la sucesión es la mayor de todas las cotas inferiores [6] [13]. El cero es una cota inferior para la sucesión $\{y_n\}$, por lo que se puede afirmar que la sucesión escalar $\{y_n\}$ es convergente. Lo que ocurre aquí es que no podemos garantizar que converge a cero. En general converge a un número que es mayor o igual a cero. Es decir:

$$\exists \alpha \geq 0 \quad \lim_{n \to +\infty} y_n = \alpha$$

Si $\alpha = 0$ entonces la convergencia se ha dado hacia un óptimo global. Si por el contrario $\alpha > 0$ la convergencia ha sido hacia un óptimo local. Podemos apreciar entonces que no podemos garantizar la convergencia a un óptimo global. Lo mejor que podemos afirmar es que la convergencia ocurre hacia un óptimo local, pero, si esto ocurre, en términos del sistema no lineal de ecuaciones quiere decir que no se pudo encontrar una solución. Podemos entonces decir que el método heurístico encuentra una solución del sistema no lineal únicamente cuando $\lim_{n \to +\infty} y_n = 0$, es decir, cuando la sucesión escalar $\{y_n\}$ converge a cero y se alcanza un óptimo global. Si se alcanzó el óptimo global eso quiere decir que la sucesión vectorial $\{X_n\}$ converge a un elemento $X^* \in \Omega$ que es solución del sistema de ecuaciones no lineales.

Si el método heurístico no garantiza convergencia a un óptimo global, ¿por qué usarlo? Entre las razones más importantes está que no requiere fuertes hipótesis de continuidad y diferenciabilidad sobre las funciones componentes del sistema no lineal, permite ampliar la región de búsqueda con posibilidad de éxito de encontrar soluciones, y es mejor que los métodos clásicos en problemas complicados, es decir, cuando se tienen sistemas no lineales grandes de muchas ecuaciones e incógnitas [1] [2] [3].

CAPÍTULO 3

EL MÉTODO NUMÉRICO HOMOTÓPICO

En este capítulo se va a probar que se puede plantear un problema de valor inicial como un problema equivalente a un sistema no lineal de ecuaciones [6]. Esto es muy importante ya que podemos luego resolver el problema de valor inicial con cualquier método numérico y tomar la solución encontrada como una solución del sistema no lineal [6]. Luego se enuncia y se demuestra una proposición que nos proporciona una condición suficiente para la existencia y unicidad de la solución del problema de valor inicial. Esta solución es la conocida curva homotópica que cortará al plano de ecuación $\lambda = 1$ [7] [9]. Se explica luego cómo para esta tesis se resolvió numéricamente el problema de valor inicial y se termina el capítulo haciendo una comparación entre el método homotópico y los métodos numéricos iterativos convencionales para resolver sistemas de ecuaciones no lineales. Se analizan ventajas y desventajas.

3.1 EL PROBLEMA DE VALOR INICIAL

A continuación se va a enunciar y a demostrar formalmente un teorema que justifica el plantear, como problema equivalente de un sistema no lineal de ecuaciones, un problema de valor inicial. Este teorema nos permite tomar la solución del problema de valor inicial como una solución del sistema de ecuaciones no lineales.

Proposición 1.- Sea $H(X,\lambda)$ una función de clase C^1 y la matriz Jacobiana $H_X(X,\lambda)$ inversible. Suponga también que $X = \psi(\lambda)$ es una función de clase $C^1[0,1]$. La función $X = \psi(\lambda)$ es una solución de la ecuación $H(X,\lambda) = \vec{0}$ para todo $0 \le \lambda \le 1$, si y sólo si, $X = \psi(\lambda)$ es una solución del problema de valor inicial $X'(\lambda) = -[H_X(X,\lambda)]^{-1} H_\lambda(X,\lambda)$ con $H(X(0),0) = \vec{0}$ y además $\lambda \in [0,1]$

Demo

Primera Parte: Supongamos que la función $X=\psi(\lambda)$ satisface la ecuación $H(X,\lambda)=\vec{0}$ para todo $\lambda\in[0,1]$ y defínase la función $\phi(\lambda)=H(X=\psi(\lambda),\lambda)$. Podemos derivar entonces aplicando la regla de la cadena, respecto de la variable λ, ambos lados de la ecuación: $\phi(\lambda)=H(X=\psi(\lambda),\lambda)=\vec{0}$

Entonces: $\phi'(\lambda)=H_X(X=\psi(\lambda),\lambda)\,\psi'(\lambda)+H_\lambda(X=\psi(\lambda),\lambda)=\vec{0}$

Al despejar $\psi'(\lambda)$ de la ecuación anterior, suponiendo que $H_X(X,\lambda)$ tiene inversa, tenemos: $\psi'(\lambda)=-[H_X(X=\psi(\lambda),\lambda)]^{-1}H_\lambda(X=\psi(\lambda),\lambda)$

El X que satisface la condición inicial $H(X(0),0)=\vec{0}$ es la solución del sistema $H(X(0),0)=G(X)=\vec{0}$, es decir, $\tilde{X}=\psi(0)$

Segunda Parte: Lo que debemos probar ahora es que $X=\psi(\lambda)$ satisface la ecuación $H(X,\lambda)=\vec{0}$ para todo $0\le\lambda\le1$, suponiendo que se cumple $\psi'(\lambda)=-[H_X(\psi(\lambda),\lambda)]^{-1}H_\lambda(\psi(\lambda),\lambda)$ con $\psi(0)=\tilde{X}$ y $\lambda\in[0,1]$.

Entonces: $[H_X(\psi(\lambda),\lambda)]\,\psi'(\lambda)=-H_\lambda(\psi(\lambda),\lambda)$

Luego: $[H_X(\psi(\lambda),\lambda)]\,\psi'(\lambda)+H_\lambda(\psi(\lambda),\lambda)=\vec{0}$

Después: $\frac{d}{d\lambda}[H(X=\psi(\lambda),\lambda)]=\vec{0}$

Se concluye entonces que $H(X=\psi(\lambda),\lambda)=\vec{C}$ donde $\vec{C}$ es un vector constante de R^n para todo $0\le\lambda\le1$. Pero por la condición inicial $H(X(0),0)=\vec{0}$, el vector constante $\vec{C}$ es el vector nulo $\vec{0}$. Por lo tanto $H(X=\psi(\lambda),\lambda)=\vec{0}$ para todo $\lambda\in[0,1]$.

Otra demostración para esta segunda parte de la prueba es la siguiente: Apliquemos el teorema del valor medio a la función vectorial $\phi(\lambda)=H(X=\psi(\lambda),\lambda)$ [6]. Debido a que ϕ es una función vectorial de clase C^1 en el intervalo $[0,1]$, se procede a fijar un λ en el intervalo $(0,1]$ y aplicamos el teorema del valor medio para funciones vectoriales

en el intervalo $[0,\lambda]$ donde obviamente la función vectorial ϕ es también de clase C^1. Entonces, tendríamos $\|\phi(\lambda)-\phi(0)\| \le (\lambda-0)\|\phi'(t)\|$ para algún t entre 0 y λ. Pero por la hipótesis que $X=\psi(\lambda)$ satisface la ecuación diferencial para todo λ entre 0 y 1 se tiene que $\phi'(t)=H_X(X=\psi(t),\,t)\,\psi'(t)+H_\lambda(X=\psi(t),\,t)=\vec{0}$ para cualquier t entre 0 y λ. Además, de la condición inicial tenemos $H(X=\psi(0),0)=\vec{0}=\phi(0)$. Entonces: $\|\phi(\lambda)\| \le (\lambda-0)\|\phi'(t)\| = 0$

Se puede concluir de la última expresión que $\phi(\lambda)=\vec{0}$, pero $\phi(\lambda)=H(X=\psi(\lambda),\lambda)$. Por lo tanto $H(X=\psi(\lambda),\lambda)=\vec{0}$ para $\lambda\in[0,1]$.

Con la definición de $G(X)=F(X)-F\left(\tilde{X}\right)$, la ecuación para la función $H(X,\lambda)$ es la siguiente:

$$H(X,\lambda)=\lambda F(X)+(1-\lambda)G(X)=\vec{0}$$

$$H(X,\lambda)=\lambda F(X)+(1-\lambda)\left[F(X)-F\left(\tilde{X}\right)\right]=\vec{0}$$

$$H(X,\lambda)=F(X)+(\lambda-1)F\left(\tilde{X}\right)=\vec{0}$$

Entonces, la ecuación que define implícitamente a la curva $X=\psi(\lambda)$ es:

$$F(X)=(1-\lambda)F\left(\tilde{X}\right)$$

Al derivar ambos lados de la ecuación anterior respecto de λ, aplicando regla de la cadena y suponiendo que el Jacobiano de $F(X)$ es no singular, para obtener el problema de valor inicial equivalente al sistema no lineal, tenemos $F'(X)X'(\lambda)=-F\left(\tilde{X}\right)$, de donde la ecuación diferencial es $X'(\lambda)=-[F'(X)]^{-1}F\left(\tilde{X}\right)$ con dato inicial $X(0)=\tilde{X}$.

Proposición 2.- Sea $F:R^n \to R^n$ una función de clase C^1. Suponga que el Jacobiano F' es no singular para todo $X \in R^n$ y que F es un homeomorfismo. Entonces, para cualquier $\tilde{X} \in R^n$ existe una única función $X(\lambda):[0,1] \to R^n$ de clase $C^1[0,1]$ tal que $F(X) = (1-\lambda)F\left(\tilde{X}\right)$. Además, ocurre que $X'(\lambda) = -\left[F'(X)\right]^{-1} F\left(\tilde{X}\right)$ para todo $0 \le \lambda \le 1$ y $X(0) = \tilde{X}$ [6].

Demo

Recordar que de la ecuación $H(X,\lambda) = \lambda F(X) + (1-\lambda)\left[F(X) - F\left(\tilde{X}\right)\right] = \vec{0}$ se obtiene que $F(X) = (1-\lambda)F\left(\tilde{X}\right)$. Pero por hipótesis F es un homeomorfismo y por ende F^{-1} existe y es continua. Además, por el teorema de la función inversa F^{-1} es también de clase C^1 ya que F es de clase C^1 y F' es no singular para todo $X \in R^n$. Entonces, para cada $\lambda \in [0,1]$ se tiene la solución única: $X(\lambda) = F^{-1}\left[(1-\lambda)F\left(\tilde{X}\right)\right]$

Note que $X(0) = F^{-1}\left[F\left(\tilde{X}\right)\right] = \tilde{X}$

Finalmente, derivando ambos lados de la ecuación anterior respecto de λ, aplicando la regla de la cadena y haciendo uso también del teorema de la función inversa, tenemos:

$$X'(\lambda) = D_\lambda F^{-1}\left[(1-\lambda)F\left(\tilde{X}\right)\right] = \left(F^{-1}\right)'\left[(1-\lambda)F\left(\tilde{X}\right)\right] D_\lambda \left[(1-\lambda)F\left(\tilde{X}\right)\right]$$

$$X'(\lambda) = \left[F'(X)\right]^{-1}\left[-F\left(\tilde{X}\right)\right]$$

$$X'(\lambda) = -\left[F'(X)\right]^{-1} F\left(\tilde{X}\right)$$

Por tanto se tiene la ecuación diferencial $X'(\lambda) = -\left[F'(X)\right]^{-1} F\left(\tilde{X}\right)$ para todo $0 \le \lambda \le 1$ con dato inicial $X(0) = \tilde{X}$ con solución única de clase $C^1[0,1]$.

3.2 EL ALGORITMO DEL MÉTODO HOMOTÓPICO

Existen muchos métodos numéricos para resolver problemas de valor inicial, pero los métodos de Runge-Kutta son muy populares por su facilidad de implementación y por el orden de error local y global que presentan [8] [10] [11] [12] [13]. Se usará entonces el método clásico de cuarto orden de Runge-Kutta para resolver el problema de valor inicial antes mencionado.

Sin embargo, suponga por un momento que se desea implementar el método de Euler que presenta un error local $O\left(h^2\right)$ y un error global $O\left(h\right)$ [12] [13]. La ecuación recursiva que resuelve el problema de valor inicial es:

$$X_{i+1} = X_i - h\left[F'\left(X_i\right)\right]^{-1} F\left(X_0\right);\ i = 0,1,2,\text{K}\ ,n-1$$

Recuerde que el método numérico discretiza el problema y la variable $h = 1/n$ representa el tamaño de paso. El número de subintervalos en el cual es particionado el intervalo $[0,1]$ es representado por la variable n. Además $X_0 = X(0) = \tilde{X}$ y el valor de X_n es la estimación numérica de $X(1) = X^*$.

Si ahora introducimos una nueva variable independiente t y la relacionamos con la variable independiente λ de la ecuación diferencial $X'(\lambda) = -\left[F'(X)\right]^{-1} F\left(\tilde{X}\right)$ mediante la ecuación $t = -ln\left(1-\lambda\right)$, al hacer el cambio de variable tendríamos:

$$X'(\lambda) = X'(t)\cdot\frac{dt}{d\lambda} = X'(t)\cdot\frac{1}{1-\lambda} = -\left[F'(X)\right]^{-1} F\left(\tilde{X}\right)$$

Finalmente:

$$X'(t) = -(1-\lambda)\left[F'(X)\right]^{-1} F\left(\tilde{X}\right)$$

Pero antes vimos que: $F(X) = (1-\lambda)F\left(\tilde{X}\right)$

Por lo tanto, la nueva ecuación diferencial ahora con t como variable independiente es:

$$X'(t) = -\left[F'(X)\right]^{-1} F(X)$$

Observe que si $\lambda = 0$, entonces $t = 0$. Además, si $\lambda \to 1^-$, entonces $t \to +\infty$. Por lo tanto, el nuevo problema de valor inicial equivalente a nuestro problema de valor inicial original es $X'(t) = -\left[F'(X)\right]^{-1} F(X)$ para $t \in [0,+\infty]$ con $X(0) = \tilde{X}$. Si ahora aplicamos el método de Euler con tamaño de paso $h = 1$ tenemos la ecuación recursiva:

$$X_{i+1} = X_i - \left[F'(X_i)\right]^{-1} F(X_i);\ i = 0,1,2,\text{K}$$

Note que lo que se ha obtenido por tanto es el método iterativo convencional de Newton para resolver sistemas de ecuaciones no lineales que, bajo ciertas condiciones, converge a la solución deseada X^*. Es de observar que si vemos al método de Newton desde esta última perspectiva, se ha tomado un tamaño de paso que no es bueno, junto con el método de Euler cuya aproximación es pobre en comparación con la diversidad de métodos numéricos existentes para resolver ecuaciones diferenciales que presentan mejores estimaciones de los valores exactos [6].

El método de Runge-Kutta que usaremos presenta un error local $O(h^5)$ y un error global $O(h^4)$ [12] [13]. Al discretizar el problema con un tamaño de paso $h = 1/n$, donde n es el número de subintervalos, tenemos la ecuación recursiva:

$$X_{i+1} = X_i + \frac{h}{6}(K_1 + 2K_2 + 2K_3 + K_4);\ i = 0,1,2,\text{K}\ , n-1$$

Las funciones K_i son:

$$K_1 = -\left[F'(X_i)\right]^{-1} F\left(\tilde{X}\right)$$

$$K_2 = -\left[F'\left(X_i + \frac{h}{2}K_1\right)\right]^{-1} F\left(\tilde{X}\right)$$

$$K_3 = -\left[F'\left(X_i + \frac{h}{2}K_2\right)\right]^{-1} F\left(\tilde{X}\right)$$

$$K_4 = -\left[F'\left(X_i + h\,K_3\right)\right]^{-1} F\left(\tilde{X}\right)$$

La estimación de $X(1) = X^*$ es el valor de X_n. Se hicieron pruebas con sistemas no lineales de hasta 5 ecuaciones con 5 incógnitas con excelentes resultados que se muestran en el capítulo 4. Cabe indicar que la precisión obtenida, además del tamaño de paso h, depende del valor inicial $X(0) = X_0 = \tilde{X}$.

3.3 ANÁLISIS COMPARATIVO MÉTODOS CLÁSICOS

A continuación en la tabla 3.1 se muestra un análisis que compara los métodos homotópicos con los métodos iterativos convencionales para la solución de sistemas no lineales. Se presentan ventajas y desventajas de cada uno, pero de forma comparativa. En el capítulo anterior se dijo que no tienen nada en común, en la manera en que resuelven un sistema no lineal, los métodos heurísticos con los métodos clásicos. Esto no ocurre entre los métodos homotópicos y los métodos convencionales en el sentido de que sí existen elementos en común. Por ejemplo, ambos métodos iteran de manera recursiva con el objetivo de buscar una solución, el grado de dificultad de la implementación de los algoritmos es muy similar, por cada corrida de los algoritmos se puede encontrar una solución y sólo una, se tiene mucha teoría matemática detrás respaldando a los métodos, se necesitan en los dos métodos por cada iteración cálculos matriciales tediosos y de derivadas parciales, en ambos métodos no es suficiente que las funciones componentes se encuentren definidas en Ω ya que son necesarias condiciones de continuidad y diferenciabilidad de cierto orden [7] [9]. No olvidar sin embargo que ambos tipos de métodos pueden trabajar colaborativamente.

Métodos Homotópicos	
Ventajas	**Desventajas**
Dominio de convergencia ampliado. El "dato inicial" del PVI puede estar fuera de la región de convergencia de un método clásico	Análisis de convergencia y error es más complicado, se necesita teoría de la medida, de ecuaciones diferenciales y de sistemas dinámicos
Se los puede utilizar para determinar el "dato inicial" que necesitan los algoritmos iterativos convencionales o para futuras corridas del mismo algoritmo con un nuevo dato inicial más cercano a la solución	No convergen con cualquier dato inicial. Se tiene un radio de convergencia de mayor medida pero se puede tener divergencia si el dato inicial se encuentra considerablemente lejos de una solución
Para resolver el PVI se tiene diversidad de métodos de un paso o multipaso con distinto orden de error según lo que se desee: métodos de Taylor, Runge-Kutta, Predictor-Corrector	La rapidez de convergencia no es tan lenta como puede ocurrir con los métodos heurísticos pero tampoco tan rápida como en los métodos clásicos
	Es más costoso estimar el número de iteraciones necesarias para estar tan cerca de la solución exacta como se desee.

Métodos Iterativos Clásicos	
Ventajas	**Desventajas**
Muchos teoremas y proposiciones que estudian convergencia y error de los métodos con un grado de dificultad menor	Normalmente un "pequeño" dominio de convergencia
Si el "dato inicial" se encuentra dentro de la región de convergencia se garantiza que el método converge a la solución exacta	Pueden tener dependencia de otros métodos que proporcionen el "dato inicial" para poder comenzar con las iteraciones
En general se tiene convergencia al menos lineal, es decir, son métodos rápidamente convergentes	La ecuación iterativa que resuelve el sistema no lineal tiene una sola presentación
Se puede estimar el número de iteraciones para que la distancia entre la solución exacta y el valor estimado no supere la tolerancia deseada	

Tabla 3.1: Análisis Comparativo Métodos Homotópicos con Métodos Tradicionales Iterativos

3.4 ANÁLISIS DE CONVERGENCIA

En [9] se afirma que el método homotópico mejora el inconveniente de los métodos clásicos de convergencia local, en los cuales el dato inicial debe estar "cerca" de la solución para poder tener convergencia, por eso la utilización del término "local". Esto nos dice que con el método homotópico podemos tener una región ampliada de convergencia en relación, por ejemplo, al método de Newton en cualquiera de sus variantes o a un método de punto fijo.

También en [9] se afirma que el tiempo de ejecución del método homotópico es mayor al de los métodos de convergencia local. Algún precio se debe pagar por intentar aumentar la región de convergencia al buscar soluciones de sistemas no lineales. Observe las semejanzas entre el método heurístico y el método homotópico en relación a la convergencia y a los tiempos de ejecución. Esta es la razón por la cual los dos métodos que se explican en este documento de tesis son llamados métodos de convergencia global.

En el método heurístico no se define un punto inicial con el cual comenzar a iterar [1] [2] [3]. El método homotópico, por el contrario, si necesita de un vector inicial X_0 con el cual comenzar las iteraciones. ¿Cómo escoger X_0 en el método homotópico para que haya convergencia a un vector X^* que es solución del sistema no lineal? En [9] se concluye que la región de convergencia en el método homotópico es igual a la región de estabilidad del correspondiente sistema dinámico. En el método homotópico se resuelve, como un problema equivalente al sistema no lineal $F(X)=\vec{0}$, el sistema no lineal autónomo de ecuaciones diferenciales:

$$X'(\lambda) = -\left[F'(X)\right]^{-1} F(X_0)$$

El dato inicial es $X(0)=X_0$. Este problema de valor inicial representa un sistema dinámico de dimensión infinita. Lo que nos dice [9] es que si tomamos como vector X_0 un vector de R^n que pertenezca a la región de estabilidad de dicho sistema dinámico, entonces con ese dato inicial X_0 el método homotópico será convergente al vector X^* que es solución del sistema no lineal correspondiente. En conclusión, podemos entonces aplicar la teoría de estabilidad de sistemas no lineales autónomos para saber si un vector inicial X_0 nos permite o no la convergencia con el método homotópico.

Es muy importante, además de identificar con cuál vector X_0 comenzar las iteraciones para garantizar la convergencia, tener también una medida para el orden p de convergencia del método homotópico. Experimentalmente se puede estimar el valor de p usando regresión lineal simple [12] [13]. El método homotópico, contrario al método heurístico, sí es un método determinista. Es decir que si ejecutamos una y otra vez el método homotópico con el mismo sistema no lineal y el mismo dato inicial, la sucesión que se construye siempre será la misma [6].

Sea E_n el error del método homotópico en la iteración n. Entonces: $E_n = \|X_n - X^*\|$ donde X_n es la aproximación en la iteración n, X^* es la solución exacta y $\| \ \|$ es una norma en R^n. Se dice que un método tiene orden de convergencia p si ocurre que:

$$E_{n+1} = C\, E_n^{\,p}$$

En general $p > 1$ y $C > 0$ o $p = 1$ y $C \in (0,1)$. Si aplicamos logaritmo natural de ambos lados obtenemos:

$$ln(E_{n+1}) = ln(C) + p\, ln(E_n)$$

Podemos efectuar experimentos computacionales de forma tal que analicemos $ln(E_n)$ versus $ln(E_{n+1})$. Es decir, podemos tomar todos los valores que se obtienen en una corrida en particular y ubicar en el eje horizontal los valores de $ln(E_n)$, mientras que en el eje vertical del mismo plano $ln(E_{n+1})$. Luego usando regresión lineal simple podemos determinar la mejor recta que se ajuste a todos esos puntos.
El valor de p representaría la pendiente de esa recta y el valor de $ln(C)$ el corte de la recta con el eje vertical. Es decir:

$$p = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$ln(C) = \frac{\sum_{i=1}^{n} y_i - p\sum_{i=1}^{n} x_i}{n}$$

En estas expresiones los x_i representan los valores de las abscisas de los puntos mientras que los y_i representan los valores de las ordenadas. Todos los sistemas no lineales que se muestran en el capítulo 4 fueron expuestos a este experimento computacional que consiste en determinar el orden de convergencia p del método homotópico. Es interesante observar que en todos los casos ocurrió que $p \in (1,1.5)$. Recordemos que en lo métodos de punto fijo el orden de convergencia es al menos lineal $p = 1$, mientras que para el método de Newton se tiene al menos convergencia cuadrática $p = 2$. Finalmente se puede concluir que en el peor de los casos el método homotópico tiene convergencia casi lineal y en el mejor de los casos presenta un orden de convergencia similar al orden de convergencia del método de la secante.

CAPÍTULO 4

EXPERIMENTOS COMPUTACIONALES

En este capítulo se muestran los resultados computacionales al trabajar con distintos sistemas no lineales. Comenzamos resolviendo una ecuación no lineal con una incógnita para poder discutir ambos métodos desde una perspectiva geométrica. Se resuelve la misma ecuación con el método heurístico y con el método homotópico. Luego se presentan dos sistemas no lineales de dos ecuaciones con dos incógnitas, donde aún se puede incluir en el análisis una interpretación geométrica. Nuevamente los mismos sistemas son resueltos primeramente con el método heurístico y luego con el método homotópico. Finalmente se muestran experimentos computacionales con ambos métodos en un sistema de tres ecuaciones con tres incógnitas y en un sistema con cuatro ecuaciones con cuatro incógnitas, donde esta vez es imposible elaborar interpretaciones geométricas. Se termina este capítulo con un sistema no lineal especial de tercer orden que no se puede resolver con los métodos clásicos, ni con el método homotópico, pero sí con el método heurístico, y se proporcionan las razones del por qué.

4.1 UNA ECUACIÓN NO LINEAL

El primer experimento computacional que se presenta es sobre una ecuación no lineal. Se usará inicialmente el método heurístico. Para empezar, y para poder dar una interpretación geométrica de los resultados, vamos a mostrar lo que ocurre con la ecuación $f(x) = x^3 - e^x \, sen(x)$. Haciendo uso del conocido teorema de Bolzano se puede probar que f tiene un CERO en el intervalo $\Omega = [1,2]$. Por la proposición enunciada y demostrada anteriormente en la sección 2.1 se puede afirmar que x^* satisface $f(x) = x^3 - e^x \, sen(x) = 0$, si y sólo si, x^* minimiza la función $g(x) = \left| x^3 - e^x \, sen(x) \right|$ en el intervalo $[1,2]$ con el valor cero. La figura 4.1 muestra la gráfica de la función f en el intervalo $[1,2]$, mientras que la figura 4.2 nos muestra la gráfica de la función a minimizar g en el mismo intervalo $[1,2]$.

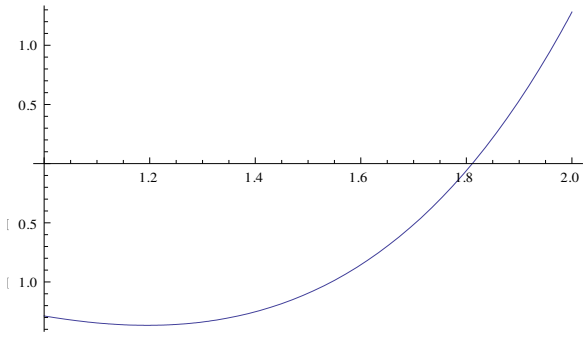


Figura 4.1: Gráfico de $f(x) = x^3 - e^x \, sen(x)$

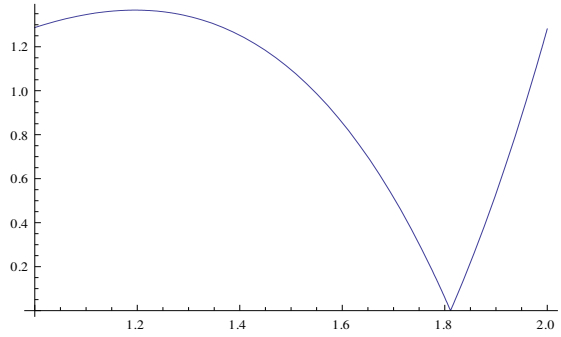


Figura 4.2: Gráfico de $g(x) = \left| x^3 - e^x \, sen(x) \right|$

Se puede observar que mayor a 1.8 es el valor de x^*. En una de las corridas del algoritmo metaheurístico de "scatter search" encontramos que $x^* = 1.8112693646$ y que $g(x^*) = 0.0000655681$. Para esta ejecución se dividió el intervalo $\Omega = [1,2]$ sólo en 10 subintervalos y se solicitaron 5 actualizaciones del conjunto R.

Ahora se usará el método homotópico. Tomemos la misma función de variable real $f(x) = x^3 - e^x\ sen(x)$ cuya derivada es $f'(x) = 3x^2 - e^x\left[\ sen(x) + cos(x)\right]$. Luego tenemos que $H(x,\lambda) = x^3 - e^x\ sen(x) + (\lambda - 1)\left({x_0}^3 - e^{x0}\ sen(x_0)\right)$ con $x_0 \in R$. Finalmente, el problema de valor inicial correspondiente es: $x'(\lambda) = \dfrac{e^{x0}\ sen(x_0) - {x_0}^3}{3x^2 - e^x\left[\ sen(x) + cos(x)\right]}$ para todo $0 \le \lambda \le 1$ y $x(0) = x_0$.

A continuación, en la figura 4.3, se muestra la curva de nivel $H(x,\lambda) = 0$ en R^2 con distintos valores para x_0. Se incluyen los resultados obtenidos al ejecutar el método numérico de Runge-Kutta de cuarto orden en cada caso. También se incluye la gráfica de la función $f(x) = x^3 - e^x\ sen(x)$ en distintos intervalos. El tamaño de paso $h = 0.001$ se utilizó en todas las ejecuciones. En el eje horizontal hemos ubicado los valores que toma la variable x, mientras que los valores de la variable $\lambda \in [0,1]$ se encuentran en el eje vertical. Note que f tiene muchos "ceros".

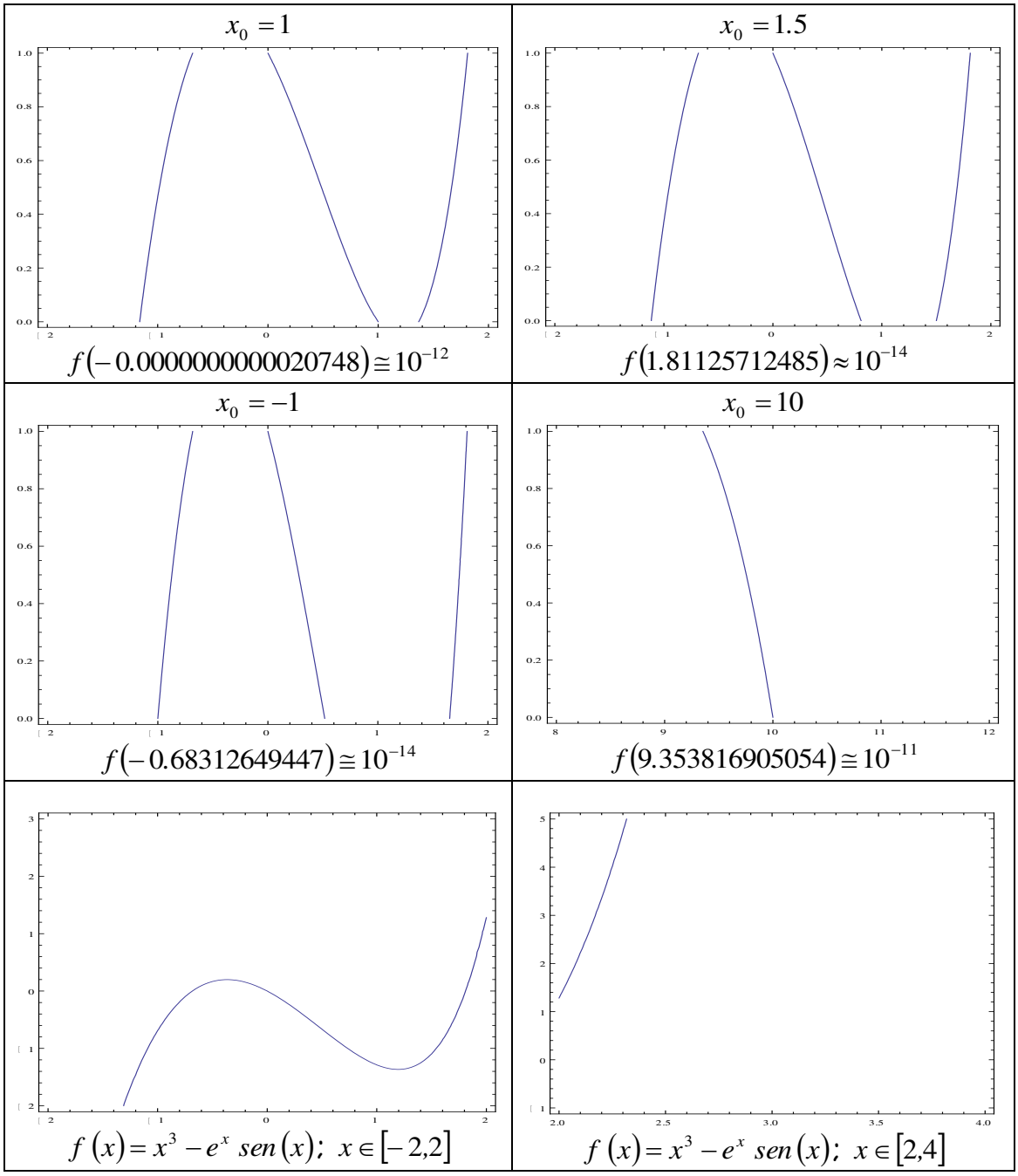


Figura 4.3: Curvas de nivel de H para $f(x) = x^3 - e^x \, sen(x)$

4.2 SISTEMAS NO LINEALES DE SEGUNDO ORDEN

Un segundo ejemplo es el sistema:

$$\begin{cases} x\,sen(y)-1=0 \\ x^2+cos(2y)=0 \end{cases}$$

Se puede apreciar que el par ordenado $(1,\pi/2)$ es una solución del sistema en la región $\Omega=\{(x,y)\in R^2 / 0\le x\le 2\ ,\ 0\le y\le 2\}$.

En la figura 4.4 y en la figura 4.5 se muestra la gráfica de la función $g(x,y)=|\,x\,sen(y)-1\,|+|\,x^2+cos(2y)\,|$ sobre la región Ω del plano xy.

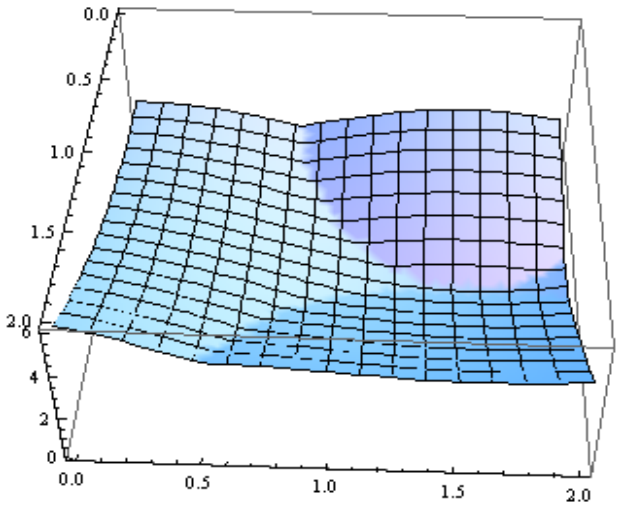


Figura 4.4: Gráfico de $g(x,y)=|\,x\,sen(y)-1\,|+|\,x^2+cos(2y)\,|$

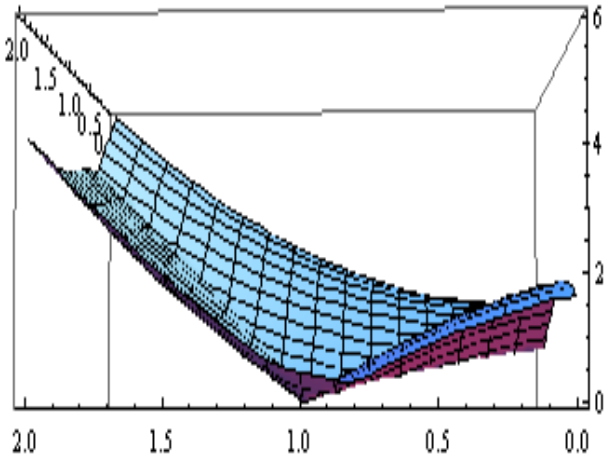


Figura 4.5: Otra Vista de $g(x,y)=\left|\, x\,sen(y)-1\,\right|+\left|\, x^2+cos(2y)\,\right|$

En una de las ejecuciones del algoritmo "scatter search", para este segundo ejemplo, se obtuvo que $(x^*,y^*)=(1.0,1.5707963290621922)$ y que $g(x^*,y^*)$ es cero. Para esta corrida se solicitaron 5 actualizaciones del conjunto R y, para cada variable, se ingresó el intervalo $[-10,10]$, el cual se particionó en 10 subintervalos.

Ahora se usará el método homotópico. Recuerde que el valor de X_n que entrega el método de Runge-Kutta es la estimación del valor exacto X^*. De aquí en adelante usaremos la siguiente expresión para calcular el error: $E(X_n)=\sum_{i=1}^{n}\left|\, f_i(X_n)\,\right|$ donde las f_i son las funciones componentes del sistema no lineal. Los valores de la variable x se encuentran en el eje horizontal, los valores de la variable y en el eje vertical. Se utilizó $h=0.001$ como tamaño de paso. En la figura 4.6 se muestra, para cada corrida, la trayectoria homotópica de las soluciones.

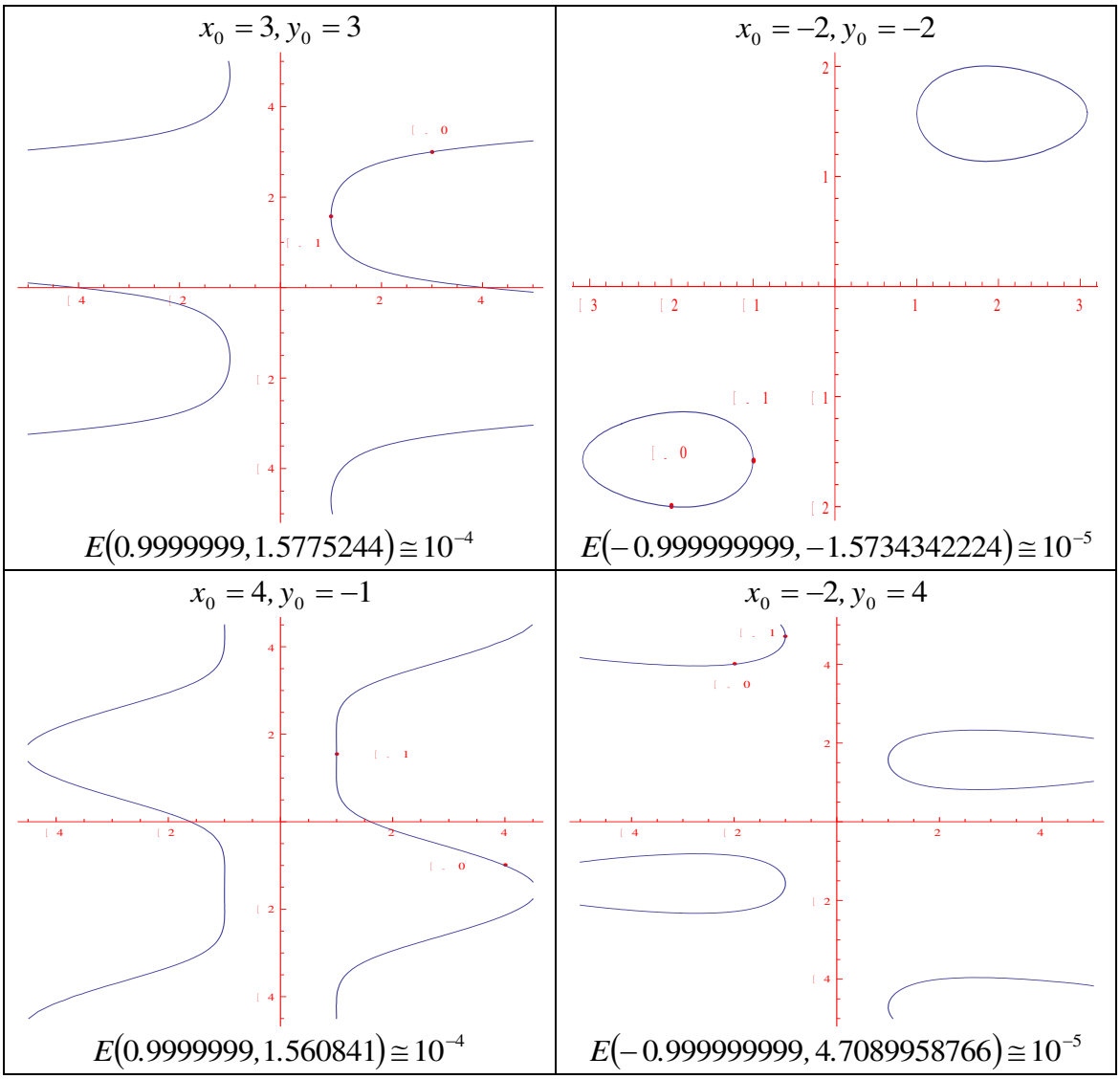


Figura 4.6: Trayectoria de las Soluciones Segundo Ejemplo

Otro ejemplo con interpretación gráfica es el siguiente:

$$\begin{cases} 5x^2 - y^2 = 0 \\ 4y - sen(x) - cos(y) = 0 \end{cases}$$

Este sistema no lineal tiene solución en $\Omega = [0,1] \mathrm{X} [0,1]$. En la figura 4.7 se muestra la gráfica de la función $g(x,y) = \left| 5x^2 - y^2 \right| + \left| 4y - sen(x) - cos(y) \right|$.

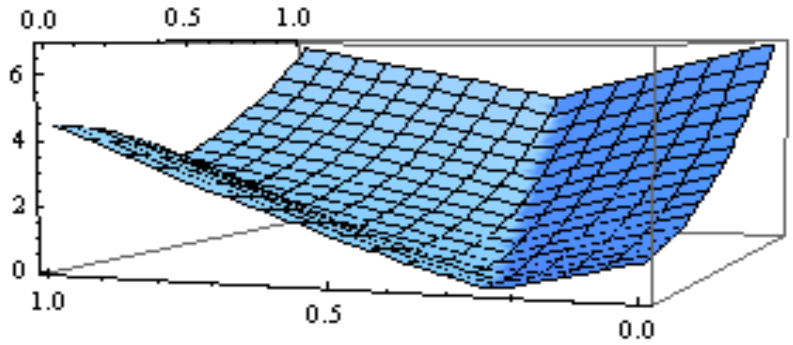


Figura 4.7: Gráfica de $g(x,y)=\left|5x^2-y^2\right|+\left|4y-sen(x)-cos(y)\right|$

Con los mismos parámetros del método heurístico de la ejecución del segundo ejemplo se obtuvo que $(x^*,y^*)=(0.12124191148054805, 0.27110515579243255)$ y que $g(x^*,y^*)$ es del orden de 10^{-14}. Es importante indicar que se pueden solicitar más actualizaciones del conjunto R, aumentar el valor del parámetro m, aumentar la cardinalidad de los conjuntos P y R, todo esto con el objetivo de mejorar la precisión. Sin embargo, los costos computacionales serán mayores.

Con $h=0.001$ como tamaño de paso, se ejecutó el método numérico homotópico con distintos valores iniciales (x_0, y_0). En la figura 4.8 se muestra, para cada ejecución, la trayectoria homotópica de las soluciones de cada uno de los sistemas no lineales que se van resolviendo desde $\lambda=0$ hasta $\lambda=1$.

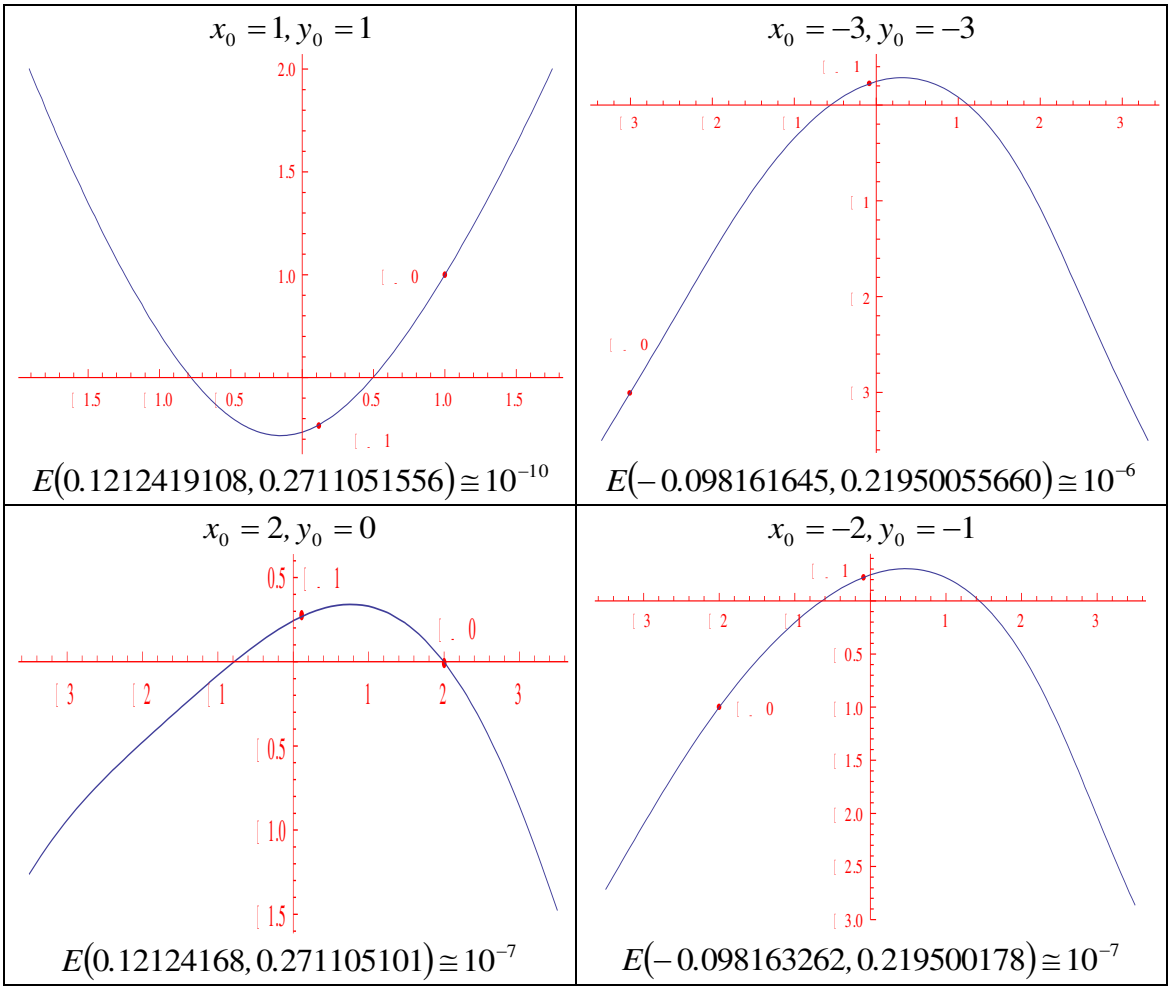


Figura 4.8: Trayectoria de las Soluciones Tercer Ejemplo

4.3 SISTEMAS NO LINEALES DE ORDEN SUPERIOR

En otro ejemplo, considere el sistema no lineal que se muestra a continuación:

$$\begin{cases} (x-2)^2 + (y-1)^2 + x\,y - 3 = 0 \\ x\,e^{x+y} + y\,z - 3 = 0 \\ sen\,(x+z) + x - y - z + 1 = 0 \end{cases}$$

Se ejecutó el algoritmo "scatter search" y se obtuvo como resultado la solución $(x^*, y^*, z^*) = (3.03924738054514, -1.6104828215273, 6.01336437928736)$ y además que $g(x^*, y^*, z^*) = 0.0000000345641$ en la región sólida $\Omega = [-10,10] \text{X} [-10,10] \text{X} [-10,10]$. Se solicitaron nuevamente 5 actualizaciones del conjunto R y además $m = 10$.

Con el mismo tamaño de paso $h = 0.001$ usado en los ejemplos anteriores del método homotópico, se muestran los resultados de dos ejecuciones para este cuarto ejemplo:

$x_0 = 1, y_0 = 1, z_0 = 1$	$x_0 = 4, y_0 = -2, z_0 = 7$
$x_n = 0.3885222610700$	$x_n = 3.039247388237$
$y_n = 1.0345501950139$	$y_n = -1.610482827797$
$z_n = 1.341346675089$	$z_n = 6.013364389595$
$E(x_n, y_n, z_n) \cong 10^{-13}$	$E(x_n, y_n, z_n) \cong 10^{-13}$

Finalmente, como un último ejemplo, se tiene el sistema no lineal:

$$\begin{cases} x^2 + 2y^2 + cos(z) - w^2 = 0 \\ 3x^2 + y^2 + sen^2(z) - w^2 = 0 \\ -2x^2 - y^2 - cos(z) + w^2 = 0 \\ -x^2 - y^2 - 2cos^2(z) + w^2 = 0 \end{cases}$$

Este sistema no lineal presenta muchas soluciones. Dos de ellas son: $(1, -1, 0, 2)$ y $(-1, 1, 0, -2)$. Se ingresaron como parámetros del método heurístico $m = 10$, un total de 5 actualizaciones para el conjunto R y el espacio de búsqueda $\Omega = [-10,10] \text{X} [-10,10] \text{X} [-10,10] \text{X} [-10,10]$. El valor de la función g en el punto encontrado $(x^*, y^*, z^*, w^*) = (0.99896239, -0.99896202, 0.03718569, 1.99827054)$ es igual a 0.000002593939, es decir, del orden de 10^{-6}.

Se muestran también un par de ejecuciones con el método homotópico para este sistema con $h = 0.00001$ como tamaño de paso:

$x_0 = 3, y_0 = 3, z_0 = 3, w_0 = 3$
$x_n = 0.998525959$
$y_n = 0.998525959$
$z_n = 0.0456785223$
$w_n = 1.9975435860$
$E(x_n, y_n, z_n) \cong 10^{-4}$

$x_0 = -2, y_0 = -2, z_0 = -2, w_0 = -2$
$x_n = -0.9989511610$
$y_n = -0.9989511610$
$z_n = -0.0385347375$
$w_n = -1.9982520966$
$E(x_n, y_n, z_n) \cong 10^{-4}$

Cabe mencionar que todos los ejemplos que se mencionan en esta tesis fueron resueltos tanto con el método heurístico como con el método homotópico, con distintos valores de entrada en los parámetros de ejecución. Además se realizaron pruebas con sistemas de hasta cinco ecuaciones con cinco incógnitas y con otros sistemas no lineales de menor tamaño, que no se muestran en este documento, donde se obtuvieron excelentes resultados.

Se desea terminar esta sección con un sistema no lineal especial:

$$\begin{cases} x\mu(x+2) + [\lfloor y-2 \rfloor][\lfloor y+z \rfloor] = 5.15 \\ |x\ y\ z| + ln\,|x-y| = 28.70742 \\ z\ Sgn\left(\dfrac{x+y}{z}\right) = 4.62 \end{cases}$$

En la sección 2.3 se dijo que una ventaja del método heurístico sobre los métodos clásicos es que las funciones componentes del sistema de ecuaciones no lineales, no necesitan cumplir condiciones de continuidad y derivabilidad de orden alguno, lo único que necesitamos con el algoritmo de búsqueda dispersa que se presenta en este documento es que todas las funciones componentes tengan su dominio en el conjunto Ω, es decir, que al menos estén definidas en Ω. El sistema que se muestra es especial en el sentido que las funciones componentes están expresadas en términos de las funciones "valor absoluto", "escalón", "signo" y "entero mayor", las cuales tienen problemas de continuidad y derivabilidad en algunos puntos. Es decir que este sistema no se puede resolver con los métodos clásicos pero sí con el método numérico heurístico que se propone. En la sección 2.4 se dijo que el método homotópico sí necesita, al igual que los métodos convencionales, que las funciones componentes cumplan criterios de continuidad y derivabilidad. Es decir que este sistema tampoco lo

podríamos resolver con el método homotópico. Al ejecutarlo con el método heurístico se obtuvo que:

$$x^* = -1.84991182 \qquad y^* = 3.1701208 \qquad z^* = 4.619976576$$

Además:

$$g(x^*, y^*, z^*) = 0.000501278$$

Es decir, se aproximó la solución con un error alrededor de 10^{-4} y se ingresaron para la corrida como parámetros 5 actualizaciones en el conjunto *R*, y el intervalo $[-10,10]$ para las tres variables con una partición de 10 subintervalos. La terna $(x = -1.85, y = 3.17, z = 4.62)$ es la solución exacta del sistema.

Un ejemplo sencillo y con ilustración gráfica en una variable es la ecuación $x e^{-x} = 0$. Se observa que $x = 0$ es una solución trivial. En el intervalo $[-2,3]$ el método heurístico funciona, es decir, converge a cero. Con el dato inicial "-2" convergen tanto el método homotópico como el método clásico de Newton. Si el dato inicial es ahora "2", los métodos homotópico y de Newton divergen y no son capaces de encontrar una buena aproximación a la solución de la ecuación.

Otro ejemplo con solución obvia $x = 0$ es la ecuación $arctan(x) = 0$. En el mismo intervalo $[-2,3]$ el método heurístico converge a la solución. Si tomamos como dato inicial "1", tanto el método de Newton como el método homotópico construyen sucesiones que convergen a la solución. Sin embargo, si tomamos "2" como dato inicial, el método de Newton diverge pero el método homotópico sí converge a la solución del problema. Principalmente este último ejemplo nos muestra que los métodos "heurístico" y "homotópico" son una alternativa atractiva cuando los métodos convencionales fallan.

CONCLUSIONES

- Para resolver sistemas no lineales de ecuaciones se puede plantear un problema de optimización y un problema de valor inicial, como problemas equivalentes. Para resolver el problema de optimización se tienen métodos heurísticos como los algoritmos genéticos y la búsqueda dispersa, y métodos tradicionales de optimización numérica como el método BFGS; para el problema de valor inicial se tienen diversos métodos numéricos de un paso o métodos multipaso, entre ellos los métodos de Taylor, Runge-Kutta, Adams-Bashforth, Adams-Moulton, Predictor-Corrector, entre otros.

- En el método heurístico no se necesita que las funciones componentes sean de clase C^k para algún número natural k, ni satisfacer condiciones de Lipschitz. Es más, podrían no ser continuas las funciones componentes, sólo se necesita que se encuentren definidas sobre un conjunto compacto Ω.

- El método heurístico y el método homotópico permiten trabajar sobre regiones de convergencia que tienen radios de mayor longitud, y es por eso que son conocidos como métodos de convergencia global ya que amplían las regiones de convergencia, en comparación con los métodos iterativos convencionales los cuales tienen sólo convergencia local.

- Los métodos de convergencia global y los métodos de convergencia local no son mutuamente excluyentes, es decir, pueden trabajar de forma colaborativa. Se puede utilizar un método de convergencia global, como el método heurístico o el método homotópico, para que proporcione el dato inicial que puede utilizar el método de convergencia local para obtener una solución del sistema no lineal de ecuaciones con la precisión requerida.

- La implementación de un algoritmo de convergencia global es más costosa que la implementación de un algoritmo de convergencia local. Hay mucho más trabajo y esfuerzo de programación en el método heurístico o en el método homotópico, que en cualquier método de convergencia local como el método de Newton o una de sus variantes.

- El método heurístico tiene la capacidad de encontrar más de una solución, si es que las hay para un sistema no lineal de ecuaciones en particular, dentro de una misma corrida. Los métodos de convergencia local encuentran una solución y sólo una por cada ejecución, lo que ocurre también con el método homotópico.

- Los métodos de convergencia global tiene una convergencia más lenta que los métodos de convergencia local, es decir, necesitan mayor cómputo y por ende mayor consumo de recursos de memoria y de uso del procesador. Por el contrario, los métodos de convergencia local pueden llegar a una solución con la aproximación deseada en un menor número de iteraciones.

- No existe garantía de convergencia con el método heurístico. Es un método que hace uso de la aleatoriedad y la computación evolutiva, por lo que puede quedarse atrapado en óptimos locales. Sin embargo, en la práctica cuando se lo ha utilizado ha presentado excelentes resultados.

- No existe garantía de convergencia con el método homotópico, la curva homotópica podría no alcanzar el plano de ecuación $\lambda = 1$, pero sí se tiene una "mayor" región de convergencia en comparación con los métodos de convergencia local. Si la curva homotópica no es de longitud finita, se puede hacer uso de otro dato inicial para estudiar el comportamiento de la nueva curva homotópica.

- El método homotópico tiene mucha teoría matemática que respalda su uso, que permite hacer análisis de convergencia y tratar con el error. Igual ocurre con los métodos iterativos tradicionales como son los métodos de convergencia local. Pero en el caso del método heurístico, no se puede tener control sobre el error ni hacer ningún tipo de análisis riguroso y formal de convergencia. Para el método heurístico es únicamente una cuestión de fe. Los métodos heurísticos se están usando ya para resolver numéricamente problemas en áreas como los sistemas dinámicos y las ecuaciones diferenciales.

RECOMENDACIONES

- Construir un "algoritmo híbrido", es decir, un algoritmo donde algunas partes provengan de métodos iterativos clásicos, y por otro lado, también tenga componentes que utilicen inteligentemente la aleatoriedad como lo hacen los algoritmos metaheurísticos. Por ejemplo, en el componente de búsqueda local del algoritmo de búsqueda dispersa que se propone en esta tesis, se puede hacer uso del método de máximo descenso o del método BFGS, obviamente estos métodos emplean derivadas y se necesitan condiciones de continuidad. Finalmente realizar experimentos computacionales con el algoritmo híbrido.

- Rediseñar el algoritmo heurístico de esta tesis para que pueda encontrar no solamente soluciones reales de sistemas de ecuaciones no lineales, sino también soluciones complejas. Sería importante que el rediseño permita determinar en una sola corrida cuántas soluciones tiene el sistema no lineal y cuáles son, sin importar si son reales o complejas.

- Investigar sobre la existencia de un tercer tipo de problema equivalente para resolver sistemas de ecuaciones no lineales. Realizar implementaciones sobre métodos numéricos que permitan resolver ese nuevo problema equivalente, hacer análisis de convergencia y error. No se puede negar la existencia de un tercer tipo de problema equivalente, se debería emprender tal búsqueda.

- En el método homotópico se puede hacer un análisis de cómo escoger el dato inicial analizando el comportamiento de la curva homotópica, es decir, bajo qué condiciones precisas el método homotópico converge. Así mismo, por qué el método podría ser divergente. Se puede, para un sistema en particular, ir variando el dato inicial para luego estudiar el comportamiento de las distintas curvas homotópicas.

- El método numérico usado para resolver el problema de valor inicial es de un paso, podrían implementarse métodos mutipaso y hacer un estudio comparativo de los resultados. De la misma manera, podría utilizarse "recocido simulado", "búsqueda tabú" o un "algoritmo genético" para la implementación del método heurístico y analizar los resultados de ejecución para observar cuál método se comporta mejor en eficiencia y eficacia.

- Investigar el empleo de heurísticas y metaheurísticas para resolver otros tipos de problemas matemáticas. En áreas de la matemática como las ecuaciones diferenciales y los sistemas dinámicos existe la oportunidad de aplicar técnicas de computación evolutiva

- Usar siempre los métodos de convergencia local. Aún cuando los métodos de convergencia global pueden proporcionar una muy buena solución, siempre será importante que el "trabajo final" de encontrar una solución que se encuentre tan cerca como se desee (tolerancia) de la solución exacta, lo realice un método de convergencia local. De esta manera se obtendrá siempre la precisión requerida en la solución aproximada.

REFERENCIAS BIBLIOGRÁFICAS

[1] Grosan C., Abraham A. (2008). "A New Approach for Solving Nonlinear Equations Systems", IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, Vol. 38, No. 3

[2] Sacco W. F., Henderson N. (2011). "Finding All Solutions of Nonlinear Systems Using a Hybrid Metaheuristic with Fuzzy Clustering Means", Applied Soft Computing, Elsevier

[3] Amaya I., Cruz J., Correa R. (2011) "Real Roots of Nonlinear Systems of Equations through a Metaheuristic Algorithm", Dyna, Vol. 78, No. 170

[4] Laguna M., Martí R. (2003). "Scatter Search: Methodology and Implementations in C", Kluwer Academic Publishers, Norwell Massachusetts

[5] Martí R., Laguna M. (2003). "Scatter Search: Diseño Básico y Estrategias Avanzadas", Universidad de Valencia, España

[6] Ortega J. M., Rheinboldt W. C. (2001). "Iterative Solution of Nonlinear Equations in Several Variables", SIAM's Classics in Applied Mathematics, Chapters 5, 7, 10

[7] Allgower E. L., Georg K. (1990). "Introduction to Numerical Continuation Methods", Colorado State University

[8] Hairer E., Norsett S. P., Wanner G. (2008). "Solving Ordinary Differential Equations I", Springer Series in Computational Mathematics Chapters 1, 2

[9] Lee J., Chiang H. (2001). "Convergent Regions of the Newton Homotopy Method for Nonlinear Systems: Theory and Computational Applications", IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications, Vol. 48, No. 1

[10] Isaacson E., Keller H. B. (1994). "Analysis of Numerical Methods", Dover Publications, INC., New York

[11] Stoer J., Bulirsch R. (2010). "Introduction to Numerical Analysis", Third Edition, Springer-Verlag

[12] Rodríguez L. (2012) "Análisis Numérico Básico, Un Enfoque Algorítmico con el Soporte de MATLAB", FCNM-ESPOL

[13] Burden R., Faires D. (2002) "Análisis Numérico", Séptima Edición, International Thompson Editores