

PRUEBAS NUMÉRICAS DEL MÉTODO DL
PARA RESOLVER
SISTEMAS DE ECUACIONES NO LINEALES

WILLIAM FAJARDO ESCAMILLA

Abril 2010

PRUEBAS NUMÉRICAS DEL MÉTODO DL
PARA RESOLVER
SISTEMAS DE ECUACIONES NO LINEALES

WILLIAM FAJARDO ESCAMILLA

Trabajo de grado presentado como requisito
para optar al título de
Matemático

Asesor

RUBEN DARIO ORTIZ ORTIZ

Universidad de Cartagena
Facultad de Ciencias Exactas y Naturales
Programa de Matemáticas
Abril 2010

*Dedico este trabajo a mis padres William y Ladys
agradeciéndoles por todo su apoyo.*

*Menciono también a Nubia y Andrés, mis amigos bogotanos,
quienes siempre me alentaron a lograr mis metas*

AGRADECIMIENTOS

A todos aquellos quienes de una u otra forma y en poca o gran medida me ayudaron de manera desinteresada o, más bien, con el único interés de verme salir adelante.

Al Profesor Ruben Darío Ortiz por su colaboración y acompañamiento durante el proceso de graduación. Lo mismo para el funcionario Germán Cardona y la secretaria de la Facultad de Ciencias, Ana Teresa.

Contenido

1	INTRODUCCIÓN	8
2	MARCO TEÓRICO	11
2.1	Aproximaciones y Errores	11
2.1.1	Manejo de números en la computadora	11
2.1.2	Errores	12
2.1.3	Causas de errores graves en computación	13
2.1.4	Propagación de errores	14
2.2	Fórmula de Taylor de Primer Orden	15
2.2.1	Fórmula de Taylor de primer orden en dimensión 1	15
2.2.2	Fórmula de Taylor de primer orden en dimensión $m \times n$	16
2.3	Ecuaciones No Lineales con una Incógnita	16

2.3.1	Método de la Bisección	17
2.3.2	Método del Punto Fijo	17
2.3.3	Método de Newton-Raphson	18
2.4	Sistemas de Ecuaciones No Lineales	19
3	RESUMEN DEL ARTÍCULO	21
3.1	Introducción	21
3.2	El método DL y el cálculo de órbitas periódicas	23
3.3	Análisis de convergencia	25
3.4	Pruebas numéricas: una propuesta para la elección de β	27
3.5	Conclusiones	31
4	PRUEBAS NUMÉRICAS	32
4.1	Diseño de las pruebas	32
4.1.1	Problema fijo	33
4.1.2	Matriz de intercambio fija	34
4.2	Resultados	34

4.2.1	Problema Fijo	34
4.2.2	Matriz de intercambio fija	34
5	CONCLUSIONES	37
	APÉNDICE A	42
	APÉNDICE B	43
	APÉNDICE C	44

Capítulo 1

INTRODUCCIÓN

Gran parte del funcionamiento del mundo tecnológico moderno se relaciona con la cada vez más extendida aplicación de los métodos numéricos en la resolución de problemas que surgen en la vida real. Claro está, el constante avance de las ciencias informáticas tanto en hardware (procesadores más rápidos) como en software (programas inteligentes) permite una implementación fácil y eficiente de dichos métodos.

El estado actual de este arte permite, por ejemplo, mantener y dirigir un avión en vuelo mediante un sistema electrónico llamado “piloto automático”, el cual debe resolver, en tiempo real, problemas numéricos cuyas soluciones determinan la maniobra a ejecutar en cada instante. Otra aplicación importante en la actualidad es el control del sistema de generación y distribución de energía eléctrica de un país; las frecuentes variaciones que se presentan debido a diversos factores como el clima o el propio funcionamiento de los dispositivos eléctricos, requieren respuestas rápidas en el mencionado sistema como, por ejemplo, una reducción de voltaje, un cambio en la red de interconexión, etc.

Otra aplicación de los métodos numéricos muy importante para las ciencias puras y aplicadas tiene que ver con la Simulación de Sistemas, tanto a nivel numérico como gráfico. Fenómenos naturales como reacciones químicas, dinámicas poblacionales, la tectónica de placas y la formación de los planetas a partir de asteroides, así como el funcionamiento de un helicóptero y la estabilidad de un edificio, pueden estudiarse mediante técnicas avanzadas de Simulación, lo cual requiere, no sólo una rápida resolución de problemas matemáticos de gran complejidad sino, la posibilidad de mostrar en pantalla lo que sucede en cada caso.

Si bien los métodos numéricos han sido ideados y estudiados desde el mismo nacimiento del Cálculo, su aplicación a problemas reales que requieren no sólo gran precisión sino rapidez, se viene logrando desde hace más o menos 50 años gracias al desarrollo tecnológico requerido. Una vez que los matemáticos y científicos en general asumieron el uso de los métodos numéricos en sus investigaciones, se presentaron situaciones en las que era necesario mejorar el desempeño de los mismos o extender su aplicabilidad. De ahí que últimamente han acaparado gran interés, no solamente la búsqueda de nuevos métodos, sino la modificación de métodos existentes tratando de hacerlos más eficientes.

Particularmente, en diversos problemas de ciencias e ingeniería, surge la necesidad de resolver sistemas de ecuaciones no lineales, es decir, se hace indispensable resolver el problema de encontrar, si existe, un vector $x \in \mathbb{R}^n$ que satisfaga la ecuación

$$F(x) = 0$$

donde $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función no lineal y continuamente diferenciable.

En el artículo tomado como base para elaborar el presente trabajo se describe, se analiza y se prueba un método para resolver sistemas de ecuaciones no lineales bautizado como el Método DL. Si bien las pruebas realizadas y expuestas en el artículo [14, sec. 4] demuestran las bondades de dicho método, los autores del

mismo proponen continuar con las pruebas en el espacio bidimensional * variando, no sólo las funciones que especifican el sistema sino, algunos parámetros propios del método en aras de establecer algún tipo de relación entre ellos teniendo en cuenta los resultados [14, pag. 34].

*Se modificó la propuesta original teniendo en cuenta el nivel académico de quien realiza el presente trabajo.

Capítulo 2

MARCO TEÓRICO

2.1 Aproximaciones y Errores

En esta sección se comentan algunos conceptos concernientes a la aplicación de la Informática en la Matemática.

2.1.1 Manejo de números en la computadora

Por razones prácticas, sólo puede manejarse una cantidad finita de bits para cada número en una computadora, y esta cantidad o longitud varía de una máquina a otra [18, pag. 8]. Por ejemplo, cuando se realizan cálculos en ingeniería y ciencias, es mejor trabajar con una longitud grande; por otro lado, una longitud pequeña es más económica y útil para cálculos y procedimientos administrativos.

2.1.2 Errores

De lo anterior se deduce que toda cantidad representada en un computador es en realidad sólo una aproximación [18, pag. 11] y, por lo tanto, implica la existencia de un error. Algunas definiciones relevantes al error:

DEFINICIÓN 1.1 Si p' es una aproximación a p , el **error** E (el cual puede provenir de una medición) se define como

$$E \equiv p' - p$$

DEFINICIÓN 1.2 Si p' es una aproximación a p , el **error absoluto** EA es

$$EA \equiv |E| = |p' - p|$$

DEFINICIÓN 1.3 Si p' es una aproximación a p , y si $p \neq 0$, el **error relativo** ER está dado por

$$ER \equiv \frac{EA}{p} = \left| \frac{p' - p}{p} \right|$$

DEFINICIÓN 1.4 Se dice que p' aproxima a p **con t dígitos significativos**, si t es el entero positivo más grande para el cual $ER < 5 \times 10^{-t}$.

Por ejemplo, supóngase que $p = 10$; para que p' aproxime a p con 2 cifras significativas, p' debe cumplir:

$$\frac{|p' - 10|}{10} < 5 \times 10^{-2} \quad \Leftrightarrow \quad -5 \times 10^{-2} < \frac{p' - 10}{10} < 5 \times 10^{-2} \quad \Leftrightarrow$$

$$10 - 5 \times 10^{-1} < p' < 10 + 5 \times 10^{-1} \quad \Leftrightarrow \quad 9.5 < p' < 10.5$$

En general, para t dígitos significativos

$$\frac{|p' - p|}{p} < 5 \times 10^{-t} \quad \Leftrightarrow \quad p - 5p \times 10^{-t} < p' < p + 5p \times 10^{-t}$$

2.1.3 Causas de errores graves en computación

Algunas de las causas de los errores más comunes de computación se describen a continuación [18, pag. 13]. Se supondrá que la computadora maneja los números en notación científica de la forma 0.1234×10^n , es decir, con 4 cifras significativas a manera de mantisa y un exponente, para el 10, de un solo dígito n , $-9 < n < 9$; se indicará esto mostrando los números en **negrita**.

Suma de números de muy distinta magnitud: El número menor se pierde al quedar por fuera de la precisión con la cual se manejan los números.

$$\text{Ejemplo: } 0.0003 + 25.67 = \mathbf{0.3000} \times 10^{-4} + \mathbf{0.2567} \times 10^2 = 0.256703 \times 10^2 = \mathbf{0.2567} \times 10^2 = 25.67 .$$

Resta de números casi iguales: Se pierden cifras significativas, lo cual impide confiar demasiado en la respuesta.

$$\text{Ejemplo: } 2.145 - 2.144 = \mathbf{0.2145} \times 10^1 - \mathbf{0.2144} \times 10^1 = 0.0001 \times 10^1 = \mathbf{0.1000} \times 10^{-3} .$$

Overflow y Underflow (desbordamiento y colapso de memoria): Después de una operación aritmética con números de tamaño válido, es posible que resulte una cantidad demasiado grande o demasiado pequeña como para que la computadora sea capaz de almacenarla.

Ejemplo: $(\mathbf{0.2000} \times 10^9)(\mathbf{0.5000} \times 10^8) = \mathbf{0.1000} \times 10^{17}$ OVERFLOW: la respuesta tiene un exponente de dos dígitos y no puede almacenarse. Generalmente, un overflow es un error grave y la computadora podría emitir un mensaje e incluso bloquearse.

Ejemplo: $(\mathbf{0.3000} \times 10^{-5})(\mathbf{0.2000} \times 10^{-4}) = 0.0600 \times 10^{-9} = \mathbf{0.6000} \times 10^{-10}$ UNDERFLOW: la respuesta tiene un exponente de dos dígitos y no puede almacenarse. Contrariamente al overflow, el underflow no se considera un error demasiado grave ya que la cantidad simplemente se almacena como 0. Obviamente, el programador desearía que tal cosa no ocurriera; por ejemplo, si $A = \mathbf{0.3000} \times 10^{-5}$, $B = \mathbf{0.2000} \times 10^{-4}$, $C = \mathbf{0.4000} \times 10^7$ y se va a calcular el producto de los tres números, en el orden $(AB)C$ el resultado es 0, pero en el orden $(AC)B$ se obtiene $\mathbf{0.1200} \times 10^2$.

2.1.4 Propagación de errores

Una vez que se comete un error al realizar una operación aritmética, éste se propaga a través de los cálculos subsiguientes hasta el final [18, pag. 17]. A continuación se describe cómo las distintas operaciones básicas (suma, resta, multiplicación y división) transmiten los errores de sus argumentos, los cuales se escriben como $p' = p + \varepsilon_p$, donde p es la cantidad exacta, p' es una aproximación a p y ε_p es el error correspondiente.

Suma: Se desea hallar $c = a + b$; la computadora hace $c' = a' + b' = (a + \varepsilon_a) + (b + \varepsilon_b) = (a + b) + (\varepsilon_a + \varepsilon_b) = c + \varepsilon_c$, luego $\varepsilon_c = \varepsilon_a + \varepsilon_b$ y así $|\varepsilon_c| \leq |\varepsilon_a| + |\varepsilon_b|$.

Resta: Se desea hallar $c = a - b$, lo cual es lo mismo que $a + (-b)$ y así, $|\varepsilon_c| \leq |\varepsilon_a| + |\varepsilon_b|$.

Multiplicación: Se desea hallar $c = a \cdot b$; la computadora hace $c' = a' \cdot b' = (a + \varepsilon_a) \cdot (b + \varepsilon_b) = ab + (b\varepsilon_a + a\varepsilon_b + \varepsilon_a\varepsilon_b) = c + \varepsilon_c$, luego $\varepsilon_c = b\varepsilon_a + a\varepsilon_b + \varepsilon_a\varepsilon_b$; si ε_a y ε_b son suficientemente pequeños, se puede despreciar $\varepsilon_a\varepsilon_b$ y así $\varepsilon_c = b\varepsilon_a + a\varepsilon_b$. Si $ab \neq 0$, el error relativo es $ER_c = \left| \frac{\varepsilon_c}{c} \right| = \left| \frac{b\varepsilon_a + a\varepsilon_b}{ab} \right| = \left| \frac{\varepsilon_a}{a} + \frac{\varepsilon_b}{b} \right| \leq \left| \frac{\varepsilon_a}{a} \right| + \left| \frac{\varepsilon_b}{b} \right|$, luego $ER_c \leq ER_a + ER_b$.

División: Se desea hallar $c = \frac{a}{b}$, si $b \neq 0$; la computadora hace $c' = \frac{a'}{b'} = \frac{a + \varepsilon_a}{b + \varepsilon_b} = \frac{a + \varepsilon_a}{b + \varepsilon_b} \cdot \frac{b - \varepsilon_b}{b - \varepsilon_b} = \frac{ab + b\varepsilon_a - a\varepsilon_b - \varepsilon_a\varepsilon_b}{b^2 - \varepsilon_b^2}$; si ε_a y ε_b son suficientemente pequeños, se pueden despreciar $\varepsilon_a\varepsilon_b$ y ε_b^2 ; así $c' = \frac{ab + b\varepsilon_a - a\varepsilon_b}{b^2} = \frac{a}{b} - \frac{a\varepsilon_b}{b^2} + \frac{\varepsilon_a}{b} = c + \varepsilon_c$, luego $\varepsilon_c = -\frac{a\varepsilon_b}{b^2} + \frac{\varepsilon_a}{b}$. Si $a \neq 0$, el error relativo es $ER_c = \left| \frac{\varepsilon_c}{c} \right| = \left| \frac{-\frac{a\varepsilon_b}{b^2} + \frac{\varepsilon_a}{b}}{\frac{a}{b}} \right| = \left| \frac{\varepsilon_a}{a} - \frac{\varepsilon_b}{b} \right| \leq \left| \frac{\varepsilon_a}{a} \right| + \left| \frac{\varepsilon_b}{b} \right|$. Por tanto $ER_c \leq ER_a + ER_b$.

2.2 Fórmula de Taylor de Primer Orden

Gran parte del Análisis Numérico descansa sobre la famosa **Fórmula de Taylor**, la cual, cuando se restringe al primer orden, equivale a la condición de diferenciabilidad. Aquí se enunciará esta versión corta en dos teoremas, primero en dimensión 1 y después en dimensión $m \times n$.

2.2.1 Fórmula de Taylor de primer orden en dimensión 1

TEOREMA 1.1 Ver [16, pag. 341]. Si una función $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ es diferenciable en $a \in A$, entonces para $x \in A$ se tiene que

$$f(x) = f(a) + f'(a) \cdot (x - a) + \varepsilon_a(x)$$

en donde $\lim_{x \rightarrow a} \frac{\varepsilon_a(x)}{x - a} = 0$.

2.2.2 Fórmula de Taylor de primer orden en dimensión $m \times n$

TEOREMA 1.2 Ver [17, pag. 328]. Si $F : S \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es un campo vectorial diferenciable en $p \in S$, entonces para $v \in S$ se tiene que

$$F(v) = F(p) + DF(p)(v - p) + \varepsilon_p(v)$$

en donde $\lim_{v \rightarrow p} \frac{\varepsilon_p(v)}{\|v - p\|} = 0$.

Si $F = (f_1, f_2, \dots, f_m)$, donde $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ para cada $i = 1, 2, \dots, m$, entonces $DF(p)$, la **diferencial** de F en p , es la matriz

$$DF(p) = \begin{bmatrix} D_1 f_1(p) & D_2 f_1(p) & \dots & D_n f_1(p) \\ D_1 f_2(p) & D_2 f_2(p) & \dots & D_n f_2(p) \\ \vdots & \vdots & \vdots & \vdots \\ D_1 f_m(p) & D_2 f_m(p) & \dots & D_n f_m(p) \end{bmatrix}$$

2.3 Ecuaciones No Lineales con una Incógnita

DEFINICIÓN 1.5 Una **ecuación no lineal con una incógnita real** es de la forma $f(x) = 0$, donde $f : \mathbb{R} \longrightarrow \mathbb{R}$ es una función no lineal. Resolverla, consiste en hallar todos los valores reales de x que la satisfacen; dichos valores se denominan **raíces** de la ecuación y de f .

Los métodos tradicionales, más sencillos y, por lo tanto, con mayor grado de generalidad, de los cuales se derivan aquellos para dimensiones superiores, son los siguientes [15, Cap. 10]. Antes, una definición concerniente a la informática:

DEFINICIÓN 1.6 Un **algoritmo** para resolver numéricamente una ecuación $f(x) = 0$, consiste en una secuencia finita de pasos que, a partir de un valor inicial $x_0 \in \mathbb{R}$, genera una sucesión de números $x_k \in \mathbb{R}$, los cuales se pretende que converjan a una raíz de f .

2.3.1 Método de la Bisección

Este método que tiene como ventaja el hecho de que siempre converge, pero la desventaja de hacerlo muy lentamente, es atribuido a Bernardo Bolzano y se basa en el conocido Teorema del Valor Intermedio (TVI):

TEOREMA 1.3 Ver [15, pag. 90]. Si $f : [a, b] \subseteq \mathbb{R} \rightarrow \mathbb{R}$ es una función continua y $W \in \mathbb{R}$ es tal que $f(a) < W < f(b)$ ó $f(b) < W < f(a)$, existe entonces $c \in (a, b)$ para el cual $f(c) = W$.

El Método de la Bisección consiste entonces en encerrar sucesivamente una raíz de f en un intervalo cada vez más pequeño; más precisamente, en cada paso la raíz queda en el semi-intervalo cuyos extremos satisfacen las condiciones del TVI tomando $W = 0$.

2.3.2 Método del Punto Fijo

Este método de rápida convergencia (obviamente, cuando es aplicable) se basa en el famoso Teorema del Punto Fijo (TPF):

DEFINICIÓN 1.7 Sea A un conjunto no vacío y $g : A \rightarrow A$ una función; se dice entonces que g es un **mapeo** sobre A . Un elemento $r \in A$ es un **punto fijo** de g , si y sólo si, $g(r) = r$.

TEOREMA 1.4 Ver [15, pag. 506]. Si $g : [a, b] \rightarrow [a, b]$ es una función continua, entonces tiene al menos un punto fijo r en $[a, b]$. Además, si g es diferenciable y si existe $M \in \mathbb{R}$ para el cual $|g'(x)| < M < 1$ para cada $x \in [a, b]$, entonces el punto fijo r es único y la sucesión $x_{k+1} = g(x_k)$, con $x_0 \in [a, b]$, converge a r .

El Método del Punto Fijo se aplica entonces escribiendo la ecuación $f(x) = 0$ en una forma equivalente $x = g(x)$, donde g satisface las condiciones del TPF en un intervalo $[a, b]$ del que se sabe, por análisis previos, contiene una raíz r de f . Observando que $f(r) = 0$ equivale a $g(r) = r$, se concluye que la sucesión $x_{k+1} = g(x_k)$, con $x_0 \in [a, b]$, converge a la raíz r de f en $[a, b]$.

2.3.3 Método de Newton-Raphson

Tiene como argumento principal el reemplazar a la función $f(x)$ por su aproximación afín en el punto x_k

$$f(x) \approx L_k(x) \equiv f(x_k) + f'(x_k)(x - x_k)$$

en cada iteración del algoritmo. Una vez escogido un valor inicial x_0 , se determinan los demás términos de la sucesión mediante la ley de recurrencia

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

la cual resulta de tomar x_{k+1} como la solución de la ecuación lineal $L_k(x) = 0$. Gráficamente, $y = L_k(x)$ es la recta ℓ tangente a la curva $y = f(x)$ en el punto $(x_k, f(x_k))$ y $(x_{k+1}, 0)$ es el punto donde ℓ corta al eje x .

Al igual que el anterior y como la mayoría de los métodos distintos al de la Bisección, el Método de Newton-Raphson es muy sensible a las propiedades de la función f y por ello no garantizan la convergencia; cuando ésta se da, lo hace de manera rápida.

Algunas variantes de este método, como el **Método de la Secante** [18, pag. 44], son útiles cuando no se conoce una fórmula para $f(x)$ sino puntos aislados como mediciones experimentales o registros históricos. En estos métodos, la derivada de $f(x)$ se aproxima como la pendiente de la recta que une dos de los puntos aislados.

2.4 Sistemas de Ecuaciones No Lineales

DEFINICIÓN 1.8 *Un Sistema de Ecuaciones No Lineales de dimensión n se conforma de n ecuaciones numéricas que involucran n incógnitas y, resolverlo, consiste en hallar, si existen, todos los valores que éstas pueden tomar de modo que se satisfagan simultáneamente las n ecuaciones [18, pag. 255].*

En símbolos, se trata de resolver

$$F(x) = 0$$

donde $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ y $x \in \mathbb{R}^n$.

Escribiendo $F = (f_1, f_2, \dots, f_n)$, donde $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ para cada $i = 1, 2, \dots, n$, el sistema se expresa más ampliamente como:

$$f_1(x_1, x_2, \dots, x_n) = 0$$

$$f_2(x_1, x_2, \dots, x_n) = 0$$

$$\vdots$$

$$f_n(x_1, x_2, \dots, x_n) = 0$$

DEFINICIÓN 1.9 *Un vector $x_* \in \mathbb{R}^n$ tal que $F(x) = 0$ se denomina una **raíz** de F y de la ecuación misma.*

DEFINICIÓN 1.10 Un **algoritmo** para resolver numéricamente un sistema de ecuaciones $F(x) = 0$, consiste en una secuencia finita de pasos que, a partir de un vector inicial $x_0 \in \mathbb{R}^n$, genera una sucesión de vectores $x_k \in \mathbb{R}^n$, los cuales se pretende que converjan a una raíz de F .

Los métodos más comúnmente utilizados para resolver numéricamente un sistema de ecuaciones no lineales son simples extensiones de aquellos correspondientes al caso unidimensional; uno de ellos es el **Método de Newton multivariable**, el cual se describe en el artículo base de este trabajo.

Capítulo 3

RESUMEN DEL ARTÍCULO

En el presente capítulo se resume el artículo [14] base de este trabajo, teniendo en cuenta el objetivo general del mismo.

3.1 Introducción

Se trata de resolver la ecuación

$$F(x) = 0 \tag{1}$$

para $x \in \mathbb{R}^n$, donde $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ es una función no lineal y continuamente diferenciable [4] [8] [12].

Moré [10], reportó una colección de ejemplos prácticos los cuales incluyen, entre otros, problemas de estabilidad de aeronaves, ecuaciones de transferencia radiactiva, problemas elípticos de valor en la frontera, el problema inverso de varas elásticas, problemas de potencia de flujos, discretización de problemas de evolución y problemas de equilibrio de plantas químicas. El alcance de las aplicaciones es aún mayor

si se incluye, además, la familia de problemas de programación no lineal, ya que las condiciones de optimalidad de tales problemas conllevan a la resolución de sistemas de ecuaciones no lineales.

Uno de los métodos más frecuentemente utilizados, debido a sus buenas propiedades de convergencia, es el tradicional *Método de Newton* [4] [8]. Dado $x_0 \in \mathbb{R}^n$ como una estimación inicial de la solución de (1), se define x_{k+1} en cada iteración del método como la solución del sistema lineal

$$L_k(x) = 0$$

donde

$$F(x) \approx L_k(x) \equiv F(x_k) + J(x_k)(x - x_k) \quad (2)$$

siendo $J(x_k)$ la matriz jacobiana de F evaluada en x_k . Cabe aclarar que esta solución existe y es única si y sólo si $J(x_k)$ es no singular.

De manera más breve, el algoritmo es el siguiente:

- Resolver el sistema lineal $J(x_k)s_k = -F(x_k)$
- Hacer $x_{k+1} = x_k + s_k$

Como alternativa al alto costo computacional que implica calcular la matriz jacobiana de F en cada iteración, han aparecido los denominados *Métodos cuasi-Newton*. Estos métodos fueron concebidos para ser tan eficientes como el de Newton pero más baratos computacionalmente. Una iteración cuasi-Newton sigue el esquema:

- Resolver el sistema lineal $B_k s_k = -F(x_k)$
- Hacer $x_{k+1} = x_k + s_k$

donde B_k es una aproximación a $J(x_k)$ [4] [8] [3].

Por otro lado, el método que se propone y estudia teórica y numéricamente en este trabajo, el cual fue ideado por Ruslan L. Davidchack y Ying-Cheng Lai [2] para encontrar *puntos de periodo p de sistemas dinámicos*, representa una alternativa interesante para considerar en la resolución de sistemas de ecuaciones no lineales en general, porque permite encontrar, en los casos donde es posible, puntos solución distintos simplemente variando la dirección de búsqueda mediante una técnica conocida, por la manera en que se define, como *Matriz de Intercambio*.

3.2 El método DL y el cálculo de órbitas periódicas

En el contexto de la Teoría del Caos surge la necesidad de resolver sistemas de ecuaciones no lineales. Como se ha dicho, el método más frecuentemente empleado es el de Newton. Debido a una dificultad relevante al tema de las dinámicas caóticas, Schmelcher-Diakonos proponen *el Método SD*, una variante del Método de Newton, cuya iteración está dada por:

$$x_{k+1} = x_k + \lambda CF(x_k)$$

en donde λ es un número positivo “pequeño” y C es una matriz $n \times n$ llamada *matriz de intercambio* o *matriz bandera*, cuyas componentes pertenecen al conjunto $\{-1, 0, 1\}$, con la imposición de que cada fila o columna contenga sólo un elemento no nulo [1] [5] *.

Si en la iteración SD se define $s_k = \lambda CF(x_k)$, lo cual equivale a:

$$[-\lambda C]^{-1} s_k = -F(x_k)$$

se tiene que $x_{k+1} = x_k + s_k$.

Se observa entonces que $[-\lambda C]^{-1}$ es una aproximación a $J(x_k)$ y así el método de Schmelcher-Diakonos es un método cuasi-Newton [3] [7] [8].

*Similares a las matrices elementales de la teoría de los sistemas de ecuaciones lineales y matrices.

Volviendo al contexto de las dinámicas caóticas, el método de Schmelcher-Diakonos posee la desventaja, respecto al de Newton, en una lenta velocidad de convergencia pero, a la vez, la ventaja no sólo de conducir a distintas soluciones de (1) sino de ampliar el conjunto de puntos iniciales para los cuales el método converge hacia una de ellas en un problema específico, el cual se denomina *el “basin” de la solución*.

Con el fin de conciliar las ventajas de los dos métodos descritos anteriormente, Davidchack y Lai proponen el **Método DL**, el cual se obtiene a partir del siguiente esquema similar al SD:

$$x_{k+1} = x_k + \lambda CF(x_{k+1})$$

Sustituyendo $F(x)$ por su aproximación afín, dada en (2), se obtiene:

$$x_{k+1} = x_k + \lambda CF(x_k) + \lambda CJ(x_k)(x_{k+1} - x_k)$$

Haciendo $\lambda = \frac{1}{\beta}$ y despejando x_{k+1} , se tiene que una iteración del Método DL está dada por:

$$x_{k+1} = x_k + [\beta I - CJ(x_k)]^{-1} CF(x_k) \quad (4)$$

Además, Davidchack y Lai, modifican (4), usando el esquema iterativo

$$x_{k+1} = x_k + [\beta \|F(x_k)\| I - CJ(x_k)]^{-1} CF(x_k) \quad (5)$$

Geoméricamente, la matriz C establece direcciones que permiten encontrar diferentes soluciones del sistema no lineal. Algebraicamente, permite estabilizar la matriz jacobiana $J(x_k)$ en el sentido de que sus valores propios tengan parte real negativa [1] [13]. El factor $\|F(x_k)\|$ acelera la convergencia del método, garantizando que el esquema (5) tenga convergencia cuadrática, en contraste con la convergencia de (4) que es sólo lineal, como se demuestra en la siguiente sección.

3.3 Análisis de convergencia

En esta parte se presenta el análisis de convergencia local del Método DL en la solución de ecuaciones no lineales (TEOREMA 2.4). Este análisis se realiza, tomando las hipótesis usuales de los métodos tipo Newton [3] [4] [8]. Además, se presenta un resultado análogo para la convergencia del esquema iterativo (4) (TEOREMA 2.5), en el cual se nota el papel que juega el factor $\|F(x_k)\|$ en la rapidez de convergencia de la iteración (5). Se inicia este estudio mostrando tres resultados que se usarán en las demostraciones de los teoremas mencionados.

En lo que sigue, el símbolo $\|\cdot\|$ denotará una norma matricial inducida por alguna norma en \mathbb{R}^n , es decir,

$$\|A\| = \max_{x \in \mathbb{R}^n, \|x\|=1} \|Ax\|, \quad \|AB\| \leq \|A\| \cdot \|B\|, \quad \|I\| = 1$$

donde, A , B e I son matrices de orden n con I la matriz identidad.

LEMA 2.1 ([4], pag. 45) Sean A , B , C matrices reales de orden n .

1. Si $\|A\| < 1$, entonces $(I - A)^{-1}$ existe y se tiene

$$\|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|} \quad (6)$$

2. Si B es no singular y $\|B^{-1}(C - B)\| < 1$, entonces C es no singular y se tiene

$$\|C^{-1}\| \leq \frac{\|B^{-1}\|}{1 - \|B^{-1}(C - B)\|} \quad (7)$$

LEMA 2.2 ([4], pág 75) Si la matriz jacobiana $J(x)$ existe para todo x en un subconjunto abierto y convexo D de \mathbb{R}^n y si existe una constante $\gamma > 0$ tal que para todo $x, y \in D$, $\|J(x) - J(y)\| \leq \gamma\|x - y\|$, entonces para todo $x, y \in D$ se cumple que

$$\|F(x) - F(y) - J(y)(x - y)\| \leq \frac{\gamma}{2}\|x - y\|^2$$

LEMA 2.3 ([5], pág. 1309) Existen $2^n n!$ matrices de intercambio de orden n . Además, si C es una matriz bandera de orden n , entonces la matriz C es invertible, $C^{-1} = C^T$, $\|C\| = 1$ y $\|Cx\| = \|x\|$ para todo $x \in \mathbb{R}^n$.

TEOREMA 2.4 Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función continuamente diferenciable en un subconjunto abierto y convexo D de \mathbb{R}^n . Para $\beta > 0$ y C una matriz de intercambio de orden n , se define la matriz

$$M(x) = \beta s(x)I - CJ(x)$$

donde $s(x) = \|F(x)\|$, I es la matriz identidad de orden n y $J(x)$ es la matriz jacobiana de F en x .

Supóngase que $x_* \in D$ y r, α, γ son constantes positivas tales que

- i) $F(x_*) = 0$
- ii) $N(x_*, r) = \{x \in \mathbb{R}^n \mid \|x - x_*\| \leq r\} \subseteq D$
- iii) $M(x_*)^{-1}$ existe y $\|M(x_*)^{-1}\| \leq \alpha$
- iv) $J(x) \in Lip_\gamma(N(x_*, r))$

existe entonces $\varepsilon > 0$ tal que para todo $x_0 \in N(x_*, \varepsilon) = \{x \in \mathbb{R}^n \mid \|x - x_*\| \leq \varepsilon\}$ la sucesión generada por

$$x_{k+1} = x_k + M(x_k)^{-1}CF(x_k) \tag{8}$$

está bien definida, converge a x_* y, además, existe una constante $\mu > 0$ tal que

$$\|x_{k+1} - x_*\| \leq \mu \|x_k - x_*\|^2 \quad (9)$$

Demostración. Ver [14, pag. 28].

TEOREMA 2.5 Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función continuamente diferenciable en un subconjunto abierto y convexo D de \mathbb{R}^n . Para $\beta > 0$ y C una matriz de intercambio de orden n , se define la matriz

$$M(x) = \beta I - CJ(x)$$

donde I es la matriz identidad de orden n y $J(x)$ es la matriz jacobiana de F en x .

Si $x_* \in D$ y r, α, γ son constantes positivas que cumplen las hipótesis (I)-(IV) del TEOREMA 2.4, entonces, para $\beta < \frac{1}{4\alpha}$, existe $\varepsilon > 0$ tal que para todo $x_0 \in N(x_*, \varepsilon) = \{x \in \mathbb{R}^n \mid \|x - x_*\| \leq \varepsilon\}$ la sucesión generada por

$$x_{k+1} = x_k + M(x_k)^{-1}CF(x_k) \quad (13)$$

está bien definida, converge a x_* y satisface

$$\|x_{k+1} - x_*\| \leq \mu \|x_k - x_*\| + \rho \|x_k - x_*\|^2 \quad (14)$$

donde μ y ρ son constantes positivas. Demostración. Ver [14, pag. 30].

3.4 Pruebas numéricas: una propuesta para la elección de β

A partir de (4), se tiene que una iteración del Método DL es de la forma

- Resolver el sistema lineal $[CJ(x_k) - \beta I]s_k = -CF(x_k)$

- Hacer $x_{k+1} = x_k + s_k$

así, en cada iteración, se debe resolver un sistema de ecuaciones lineales cuya matriz de coeficientes sugiere una forma de elegir, en cada iteración, el parámetro β : escogerlo ligeramente mayor que el radio espectral [†] de la matriz $CJ(x_k)$.

Esto garantiza no sólo que el algoritmo esté bien definido pues en cada iteración la matriz $[CJ(x_k) - \beta I]$ es no singular, sino que da una primera idea hacia la globalización del método [4]. Efectivamente, es ésta la propuesta de elección de β .

En la implementación numérica del método, inicialmente se escogen puntos iniciales arbitrarios y posteriormente los puntos iniciales son generados aleatoriamente. El valor de β considerado es el radio espectral de la matriz $CJ(x_k)$ incrementado en 0.1. Los códigos de los algoritmos y funciones fueron escritos en SCILAB 3.0, y hechos en un computador con procesador AMD Athlon(tm) de 1.10 GHz.

Se utilizó como criterio de convergencia $\|F(x_k)\| < 10^{-6}$. Se suspendió la búsqueda en el algoritmo cuando el número de iteraciones excedió 500 iteraciones o cuando $\|F(x_k)\| > 10^2$. En el último caso diremos que el método diverge.

Con el fin de ilustrar numéricamente la propuesta de elección de β , se considera el siguiente sistema de ecuaciones no lineales,

$$\begin{aligned} 2 + \text{sen}(x) - y &= 0 \\ 2(x + 2)e^{-(x-1)^2} - y &= 0 \end{aligned}$$

y las ocho matrices intercambio correspondientes a $n = 2$,

$$\begin{aligned} C_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & C_3 &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} & C_5 &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} & C_7 &= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \\ C_2 &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & C_4 &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} & C_6 &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & C_8 &= \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \end{aligned}$$

[†]El máximo del conjunto formado por las normas de los valores propios.

Inicialmente, se resolvió el sistema con la matriz de intercambio C_1 y con el vector inicial $x_0 = (1, 2.6)$. En este caso, el Método DL converge en 31 iteraciones a $x_* = (0.167229, 2.166459)$. Además, se implementó el Método DL generado a partir de (5), al cual se hará referencia en lo que sigue como *Método DL modificado*. En este caso, se obtiene convergencia a x_* en 17 iteraciones, lo cual, es de esperarse a la luz de la teoría de convergencia desarrollada en la sección anterior.

Posteriormente, y para ilustrar una de las bondades del Método DL, se cambió la matriz de intercambio por C_3 y se consideró el mismo punto inicial. El método converge en 26 iteraciones a $x_* = (2.007147, 2.906299)$, que corresponde a la otra solución del sistema no lineal. Con las otras matrices de intercambio los resultados obtenidos se presentan en la Tabla 1.

matriz	x_*	iter.
C_1	(0.1672295, 2.1664508)	31
C_2	no converge	500
C_3	(2.0071477, 2.9062993)	26
C_4	(0.1672295, 2.1664508)	44
C_5	no converge	500
C_6	no converge	500
C_7	(0.1672295, 2.1664508)	31
C_8	no converge	500

Tabla 3.1: Convergencia del método DL variando la matriz intercambio.

Para un tercer grupo de pruebas numéricas, los puntos iniciales fueron generados aleatoriamente en el intervalo $I = [-1, 3] \times [1, 3]$. Para cada uno de ellos, se resolvió el sistema no lineal usando los Métodos DL y DL modificado. La Tabla 2 muestra los resultados obtenidos para el valor inicial $x_0 = (1.6615244, 2.2567836)$ y para cada una de las matrices de intercambio correspondientes a $n = 2$. Esta tabla permite comparar claramente la convergencia lineal del Método DL frente a la convergencia cuadrática del Método DL Modificado.

	DL		DL modificado	
matriz	x_*	iter.	x_*	iter.
C_1	no converge	500	(0.167230 , 2.166452)	19
C_2	(2.007147 , 2.906299)	42	(2.007147 , 2.906299)	8
C_3	(2.007147 , 2.906299)	24	(2.007147 , 2.906299)	8
C_4	(0.167230 , 2.166452)	51	(0.167230 , 2.166452)	7
C_5	(2.007147 , 2.906299)	31	(2.007147 , 2.906299)	8
C_6	no converge	500	no converge	500
C_7	(0.167230 , 2.166452)	33	no converge	500
C_8	no converge	500	no converge	500

Tabla 3.2: Comparación, en iteraciones, de los métodos DL y DL modificado.

A partir de la teoría del Método DL se deduce que el parámetro β debe ser pequeño. Se realizó una exploración numérica en esta dirección utilizando un grupo de cuatro funciones no lineales que aparecen en los problemas de prueba propuestos por Luksan y Vlcek [6]. Los problemas estudiados son los siguientes: sistema trigonométrico exponencial (P1), problema singular de Broyden (P2), problema jacobiano estructurado (P3), problema de valor de frontera discreta (P4).

Para cada uno de estos problemas, se fijó $n = 5$ con lo cual, existen 3840 matrices de intercambio. Para cada una de estas matrices, se consideraron los valores de β que se consignan en la siguiente tabla.

β	P1	P2	P3	P4
0.2	751	3834	3840	3840
2	91	811	459	3840
5	47	14	98	2671
15	8	1	46	630
30	1	1	27	150

Tabla 3.3: Relación de β con la convergencia del Método DL.

En la tabla se relaciona β con el número total de matrices de intercambio con las

que se obtuvo convergencia. En la Tabla 3, se observa que a medida que β crece, el número de matrices de intercambio con las que se consigue convergencia disminuye. No obstante, con este pequeño grupo de funciones se puede observar la variedad de direcciones que el método proporciona para localizar la solución.

3.5 Conclusiones

Se propone el Método DL para resolver sistemas de ecuaciones no lineales en general y se hace su análisis de convergencia local a la luz de la teoría de los métodos tipo Newton. Se propone una elección del parámetro del método, β , aplicándola a un ejemplo. Además, se presentan algunos resultados del desempeño del método considerando valores crecientes de β y cuatro funciones de prueba propuestas en [6].

Cabe destacar que aún quedan tareas por realizar tanto a nivel teórico como numérico:

- Continuar con el estudio del desempeño del método aumentando la cantidad de funciones de prueba utilizadas y analizar la relación entre el parámetro β , y el número de iteraciones que requiere el método para la convergencia. En esta parte se debe tener en cuenta, el número de soluciones distintas que se consiguen con un único parámetro β , considerando todas las matrices de intercambio.
- Realizar el trabajo propuesto en el párrafo anterior, pero fijando ahora una matriz de intercambio y considerando varios valores del parámetro β .
- Implementar el Método DL para valores grandes de n . El inconveniente en esta parte radica en el gran costo computacional que implica generar las $2^n n!$ matrices de intercambio.
- Introducir en el algoritmo una estrategia de globalización y realizar el análisis teórico y numérico correspondiente.

Capítulo 4

PRUEBAS NUMÉRICAS

4.1 Diseño de las pruebas

Las pruebas se clasificaron teniendo en cuenta el problema, el punto inicial, la matriz de intercambio y el parámetro β . Para enfrentar lo expuesto en la sección 1.1 acerca de los errores y su propagación, se incluyó siempre en los cálculos el valor de $\|F(x_k)\|$ para verificar que, en los casos de convergencia, esta cantidad realmente tiende a 0.

Algo importante es que, en lugar de hallar el radio espectral de la matriz $J(x_k)$ en cada iteración (ver sección 2.4), lo cual hubiese aumentado inmensamente la complejidad del software, se tomó un valor fijo para β en cada ejecución del programa.

Los problemas estudiados fueron los siguientes:

- P1:

$$2 + \text{sen}(x) - y = 0$$

$$2(x + 2)e^{-(x-1)^2} - y = 0$$

- P2:

$$4x^2 + 2\sqrt{3}xy + 2y^2 + 10\sqrt{3}x + 10y - 5 = 0$$

$$x^2 - 4 - y = 0$$

- P3:

$$(y - 1)(y^2 - 2y + 2) - 3x(x - 2)(x - 5) = 0$$

$$5(x + y - 3) - 8(x - y + 1)^2 + (x - y + 1)^4 = 0$$

Obsérvese que el problema P1 es el mismo que aparece en el artículo [14, pag. 32]

Es claro que una buena escogencia del punto inicial depende de un análisis previo del problema; en el caso que se estudia en este trabajo, dimensión 2, las gráficas (ubicadas en los anexos) “revelan”, sin mucha precisión obviamente, la(s) solución(es) del problema; sin embargo, esto último no significa que se hayan escogido valores iniciales demasiado cercanos a las soluciones.

Las pruebas se llevaron a cabo con el programa `mDLd2.exe`, el cual fue desarrollado e implementado en lenguaje C/C++, versión Borland C++ 5.02 de Borland International, Inc. El código fuente, `mDLd2.cpp`, y los archivos resultantes de todas las pruebas mencionadas se encuentran en el CD anexo a este trabajo.

4.1.1 Problema fijo

Se resolvió el problema P2 con todas las matrices de intercambio, desde el mismo punto inicial $x_0 = (0, 0)$, tomando el parámetro β del conjunto $\{0, 0.2, 1, 2, 5, 15, 30\}$ el cual contiene los valores usados en el artículo. Antes de realizarlas, se esperaba que estas pruebas hablaran de qué tan buena es la idea de haber introducido las matrices de intercambio para hallar distintas soluciones partiendo siempre del mismo punto inicial.

4.1.2 Matriz de intercambio fija

Con la matriz C_5 , se aplicó el Método DL a los tres problemas, variando el punto inicial y tomando el parámetro β del conjunto $\{0, 0.2, 1, 2, 5, 15, 30\}$. Como control, se hizo lo mismo con la matriz identidad C_1 escogiendo un único punto inicial para cada problema, que en el caso del problema P1 se tomó como $(1, 2.6)$ ya que éste es idéntico al del artículo. Al planearla, se esperaba que esta prueba dijera si es conveniente o no tratar de resolver un problema con una única matriz de intercambio observando si ésta siempre sigue alguna dirección particular desde el punto inicial.

4.2 Resultados

4.2.1 Problema Fijo

La Tabla 4 muestra la(s) solución(es) del problema P2 a las cuales conduce el Método DL, comenzando en el mismo punto inicial $x_0 = (0, 0)$, con cada una de las ocho matrices de intercambio; asimismo los valores menor y mayor del parámetro β para los cuales se obtiene cada solución y el número de iteraciones con las cuales se logra la precisión deseada, la cual fue escogida como 1×10^{-6} , lo cual garantiza 5 decimales de exactitud en cada solución.

4.2.2 Matriz de intercambio fija

La Tabla 5 ilustra lo que sucedió al resolver los tres problemas con el Método DL usando la misma matriz de intercambio, C_5 , pero variando el punto inicial. Como

Matriz de Intercambio	x_*	menor β	iter.	mayor β	iter.
C_1	(1.254799 , -2.425479)	0.0	5	5.0	56
C_2	(1.254799 , -2.425479)	0.0	5	2.0	31
C_2	(-2.682370 , 3.195109)	5.0	31	5.0	31
C_3	(1.254799 , -2.425479)	0.0	5	30.0	279
C_4	(1.254799 , -2.425479)	0.0	5	2.0	17
C_5	(1.254799 , -2.425479)	0.0	5	0.2	6
C_5	(-2.682370 , 3.195109)	2.0	14	5.0	25
C_6	(1.254799 , -2.425479)	0.0	5	1.0	37
C_6	(-2.682370 , 3.195109)	2.0	17	30.0	124
C_7	(1.254799 , -2.425479)	0.0	5	0.2	9
C_7	(-2.682370 , 3.195109)	2.0	13	30.0	144
C_8	(1.254799 , -2.425479)	0.0	5	30.0	132

Tabla 4.1: Problema P2 resuelto con las ocho matrices de intercambio desde el mismo punto inicial $x_0 = (0, 0)$.

en la tabla anterior, también se muestran los valores menor y mayor del parámetro β para los cuales se obtiene cada solución y el número de iteraciones con las cuales se logra la precisión deseada.

La última tabla es similar a la anterior pero usando la matriz de intercambio, C_1 , y tomando sólo un punto inicial. Como se indicó en la sección 3.1.2, esto se hizo como un mecanismo de control ya que esta prueba (problema P1 con la matriz de intercambio fija C_1) y sus resultados se mencionan en el artículo. En este caso el resultado es favorable y permite así confiar en el programa `mDLd2.exe` desarrollado para este trabajo.

Problema	x_0	x_*	menor β	iter.	mayor β	iter.
P1	(-1.0 , 1.0)	(0.167230 , 2.166452)	0.0	5	1.0	21
	(1.0 , 2.6)	(0.167230 , 2.166452)	0.0	6	0.0	6
	(1.0 , 2.6)	(2.007147 , 2.906299)	0.2	9	1.0	19
	(3.0 , 3.0)	(2.007147 , 2.906299)	1.0	26	2.0	50
P2	(-4.0 , 4.0)	(-2.682370 , 3.195109)	0.0	4	5.0	35
	(0.0 , 0.0)	(1.254799 , -2.425479)	0.0	5	0.2	6
	(0.0 , 0.0)	(-2.682370 , 3.195109)	2.0	14	5.0	25
	(1.0 , -2.0)	(1.254799 , -2.425479)	0.0	3	0.2	7
	(1.0 , -2.0)	(-2.682370 , 3.195109)	2.0	17	5.0	30
P3	(-2.0 , -3.0)	(1.156153 , 3.092198)	0.0	14	0.2	10
	(-2.0 , -3.0)	(1.869183 , 2.070362)	1.0	9	1.0	9
	(-2.0 , -3.0)	(1.156153 , 3.092198)	2.0	10	5.0	19
	(1.0 , 1.0)	(2.076843 , 0.170736)	0.0	8	2.0	17
	(1.0 , 1.0)	(1.869183 , 2.070362)	5.0	50	5.0	50
	(3.0 , 2.0)	(1.156153 , 3.092198)	0.0	20	5.0	18

Tabla 4.2: Los tres problemas resueltos con la matriz C_5 desde varios puntos iniciales.

Problema	x_0	x_*	menor β	iter.	mayor β	iter.
P1	(1.0 , 2.6)	(0.167230 , 2.166452)	1.0	20	15.0	501
P2	(0.0 , 0.0)	(1.254799 , -2.425479)	0.0	5	0.2	6
P3	(-2.0 , -3.0)	(1.156153 , 3.092198)	0.0	14	0.2	15
	(-2.0 , -3.0)	(1.869183 , 2.070362)	1.0	35	1.0	35
	(-2.0 , -3.0)	(2.076843 , 0.170736)	2.0	20	2.0	20
	(-2.0 , -3.0)	(-2.000383 , -4.458077)	5.0	35	30.0	31

Tabla 4.3: Los tres problemas resueltos con la matriz C_1 desde un punto inicial.

Capítulo 5

CONCLUSIONES

En lo que sigue, se considera que una ejecución particular del método converge, si antes de 501 iteraciones la norma de $F(x_k)$ es menor que 1×10^{-6} ; en cualquier otro caso, se dirá que el método diverge.

- i) Aunque en la teoría expuesta en el artículo se exige que $\beta > 0$, es posible ensayar el método DL tomando $\beta = 0$ y en caso de haber convergencia, el número de iteraciones es el mínimo.
- ii) Una matriz de intercambio puede conducir, desde un punto inicial fijo, a una o a varias soluciones dependiendo del valor de β . Al parecer, en algunos casos, existen intervalos en los que β conduce a soluciones diferentes, claro está, manteniendo fijos los otros parámetros.
- iii) En general, si para valores pequeños de β el método diverge, es probable que para valores más grandes (pero “no tanto”) el método converja; y viceversa, si para un cierto valor de β se da la convergencia, lo más probable es que, a medida que aumenta β , el método diverja; en éste último caso, también puede

sucedir que, sin otros cambios, se obtenga una solución diferente.

iv) Como en el artículo:

- (a) Cuando hay convergencia, mayores valores de β , implican mayor número de iteraciones.
- (b) En general, a mayor β , menor número de matrices conducen a la convergencia.
- (c) Si una matriz de intercambio particular conlleva a divergencia, es posible que otra conduzca a la convergencia, sin cambiar los otros parámetros.

Bibliografía

- [1] DAVIDCHACK, R. L. y LAI, Y. C.: Efficient algorithm for detecting unstable periodic orbits in chaotic systems. *Physics Review Letters E*. 60 (1999) 61-72.
- [2] DAVIDCHACK, R. L., LAI, Y. C., KLEBANOFF, A. and BOLT, E.: Towards complete detection of unstable periodic orbits in chaotic systems. *Physics Letters A*. 287 (2001) 99-104.
- [3] DENNIS, J. E. y MORE, J. J.: Quasi-Newton methods, motivation and theory. *SIAM Review*. 19 (1977) 46-89.
- [4] DENNIS, J. E. y SCHNABEL, R. B.: Numerical methods for unconstrained optimization and nonlinear equations. Prentice-Hall, New Jersey, 1983.
- [5] KLEBANOFF, A. y BOLT, E.: Convergence Analysis of Davidchack and Lai's algorithm for finding periodic orbits. *Chaos Solitons and Fractals*. (2001) 1305-1322.
- [6] LUKSAN, L. y VLCEC, J.: Sparse and partially separable test problems for unconstrained and equality constrained optimization. Institute of Computer

Science Academy of Sciences of the Czech Republic. Technical report. 767, ICS AS CR, 1998.

- [7] MARTINEZ, J. M.: Algorithms for solving nonlinear systems of equations. Continuous optimizations: The state of art. Spedicato Kluwer, 1994.

- [8] MARTINEZ, J. M.: Practical quasi-Newton methods for solving nonlinear systems. Journal of Computational and Applied Mathematics. 124 (2000) 97-121.

- [9] MILLER, J. R. y YORKE, J. A.: Finding all periodic orbits of maps using Newton methods: sizes of basins. Physica D. 135 (2000) 195-211.

- [10] MORE, J. J.: A collection of nonlinear model problems. Lectures in Applied Mathematics. 26 (1990) 723-762.

- [11] PLAZA, A. M. y PEREZ, R.: Un estudio numérico de métodos tipo Newton para encontrar puntos de periodo p de sistemas dinámicos. Unicauca Ciencia. 7 (2004) 116-130.

- [12] ORTEGA, J. M. y RHEINBOLT, W. G.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York, 1970.

- [13] SCHMELCHER, P. y DIAKONOS, F. K.: A general approach to the localization of unstable periodic orbits in chaotic dynamical systems. Physics Review. 57 (1998) 2739-2746.

- [14] ACEVEDO, R. , ARENAS, F. , PEREZ R. El Método DL Para Resolver Sistemas de Ecuaciones No Lineales. Enseñanza Universitaria, Vol. XVII, 2008. Pag. 23-36.
- [15] VARBERG, Dale y PURCELL, Edwin. Calculus. 7th Ed . Prentice Hall, 1997.
- [16] APOSTOL, Tom. Calculus.2nd Ed. Ed. Reverté, 1988. Vol.1.
- [17] APOSTOL, Tom. Calculus.2nd Ed. Ed. Reverté, 1988. Vol.2.
- [18] NIEVES, Antonio y DOMÍNGUEZ, Federico. Métodos Numéricos Aplicados a la Ingeniería. 2nd Ed. CECSA, 2002.
- [19] DE LA FUENTE O'CONNOR, José Luis. Métodos Matemáticos de Especialidad Ingeniería Eléctrica. Sistemas de Ecuaciones No Lineales. Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid.
- [20] DEITEL, H. M. y DEITEL, P. J. Cómo Programar en C/C++. 2nd Ed. Prentice Hall, 1994.
- [21] <http://www.angelfire.com/scifi/jzavalar/apuntes/pseudo.html>
Cómo escribir en pseudocódigo.

APÉNDICE A

Gráfica del problema P1

APÉNDICE B

Gráfica del problema P2

APÉNDICE C

Gráfica del problema P3