

PROBLEMAS NO LINEALES

F. VADILLO

RESUMEN. En estas notas se explican algunos métodos numéricamente para resolver ecuaciones y sistemas no lineales de la forma $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ donde \mathbf{f} es una función vectorial real de variables reales. Dichas soluciones son los ceros de \mathbf{f} .

El orden de exposición es el siguiente: se comienza por el método más "ingenuo" que es el método de bisección para después comentar el método de la secante y las iteraciones de punto fijo con especial atención al método de Newton y su extensión al problema de los sistemas de ecuaciones no lineales.

ÍNDICE

1. Introducción a las ecuaciones no lineales	1
2. El método de bisección	2
3. El método de la secante	3
4. El algoritmo zeroin	4
5. Iteraciones de punto fijo	5
6. El método de Newton	8
7. Introducción a los sistemas no lineales	10
8. El método de Newton para sistemas no lineales	11
9. Modificaciones al método de Newton	12
10. Algunos comentarios para terminar	12
Referencias	13

1. INTRODUCCIÓN A LAS ECUACIONES NO LINEALES

En este capítulo se consideran ecuaciones no lineales de la forma

$$(1.1) \quad f(x) = 0,$$

donde f es una función real de variable real, un caso habitual es cuando la función f es un polinomio. Los polinomios de primero y segundo grado ya fueron tratados con éxito en tiempos muy remotos, los babilonios eran capaces de resolver ecuaciones cuadráticas. La historia del descubrimiento de la solución algebraica de la cúbica enfrentó a dos grandes rivales italianos: Cardano y Tartaglia hacia 1540, y Ferrari, alumno y secretario de Cardano resolvió en 1545 la ecuación de cuarto grado. Posteriormente fueron muchos los matemáticos eminentes que trataron de resolver las ecuaciones de grado superior a cuatro, una búsqueda sin resultados porque el matemático noruego Abel en 1893 probó que es imposible resolver por radicales

la ecuación general de grado mayor que cuatro. En consecuencia, para calcular las raíces de polinomios de grado mayor que cuatro es imprescindible usar técnicas de aproximación numéricas.

2. EL MÉTODO DE BISECCIÓN

El método de bisección es el más sencillo posible. Dada una función f real y continua en el intervalo $I_0 = [a, b]$ con valores de signos contrario en a y b , esto es $f(a)f(b) < 0$, el **teorema de Bolzano** asegura que f se anula en algún punto interior de I_0 .

El método de bisección calcula el punto medio del intervalo $c = \frac{a+b}{2}$ y $f(c)$. Si $f(c) = 0$, c es la raíz el problema está resuelto. En el caso contrario, si $f(c) \neq 0$ tiene el mismo signo que $f(a)$ o $f(b)$, es decir, en uno de los dos intervalos $[a, c]$ o $[c, b]$ la función toma valores de signo contrario en los extremos y por lo tanto contiene una raíz. Llamando ahora I_1 a dicho intervalo estamos en la situación anterior pero con I_1 en lugar de I_0 , la ventaja es que la longitud del intervalo I_1 es la mitad de la I_0 . En el siguiente paso se repite el procedimiento, se evalúa f en el punto medio de I_1 y se toma la mitad I_2 donde haya un cambio de signo en los valores de los extremos. Continuando este proceso se construye una sucesión de intervalos encajados I_n de longitud $\frac{b-a}{2^n}$ que contienen una raíz de f .

Ejemplo 2.1. Para conocer en qué punto de la trayectoria se encuentra un satélite que gira en una órbita elíptica en torno a la Tierra en un instante t , se debe resolver la ecuación de Kepler

$$x - e \sin x = z(t),$$

donde e es la excentricidad y $z(t)$ es un función conocido.

Definiendo la función $f(x) = x - e \sin x - z$, para el caso $e = 0.5$ y $z = 0.7$, $f(0) = -0.7 < 0$ y $f(2) = 0.8 > 0$ por lo que el proceso iterativo se puede comenzar en el intervalo $[0, 2]$. Las cinco primeras bisecciones están recogidas en la siguiente tabla

n	Intervalo	Punto medio	Valor de f en el punto medio
0	[0,2]	1	-0.121
1	[1,2]	1.5	0.301
2	[1, 1.5]	1.25	0.075
3	[1,1.25]	1.125	-0.026
4	[1.125, 1.25]	1.1875	0.023
5	[1.125, 1.1875]	1.1563	-0.001

El método de bisección es **muy robusto** porque se puede aplicar a cualquier función continua y no presenta posibilidades de error, sin embargo, es un método **lento**: en cada paso sólo se divide el intervalo por la mitad. Además **no aprovecha otras características de la función**, por ejemplo no valora el tamaño en los extremos de los intervalos aunque parezca más plausible que la raíz esté más cerca del extremo donde el módulo es menor.

3. EL MÉTODO DE LA SECANTE

En el método de la secante en lugar de obtener una sucesión de intervalos que contienen la raíz, se calcula una sucesión de números que aproximan el cero de la función.

Comenzando con dos aproximaciones x_0, x_1 a la raíz que se busca α , primero se construye la rectas secante que pasa por los puntos $(x_0, f(x_0)), (x_1, f(x_1))$ y se calcula su corte con el eje OX x_2 y después se calcula $f(x_2)$. El proceso se repite con x_1 y x_2 para calcular x_3 y sucesivamente x_4, x_5, \dots . La figura 1 ilustra el primer paso del método de la secante para el ejemplo 2.1.

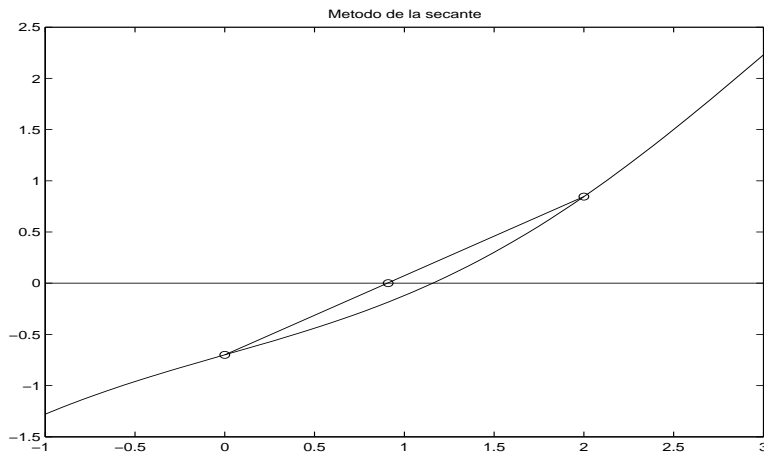


FIGURA 1. Método de la secante para el ejemplo 2.1.

La iteración del método de la secante es la siguiente

$$(3.1) \quad x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad n = 1, 2, \dots$$

Las iteración del método de la secante (3.1) no son infalibles, puede ocurrir que algún de los x_n no se encuentre en el dominio de f o que el denominador sea casi cero, en cuyo caso la secante es horizontal y salvo que x_n fuese ya un cero de f no corta al eje de abscisas.

La siguiente tabla recoge los resultados del método de la secante aplicado al ejemplo 2.1

n	x_n	$f(x_n)$
0	0	-7.0×10^{-1}
1	2	8.45×10^{-1}
2	0.9059	-1.88×10^{-1}
3	1.1046	-4.2×10^{-2}
4	1.1619	3.2×10^{-3}
5	1.1579	-4.91×10^{-5}

Comparado este método con el de bisección se ve claramente que ahora la convergencia es más rápida, en cinco iteraciones se tienen residuos del orden de 10^{-5} mientras que en la bisección era del orden de 10^{-3} .

Los posibles **criterios de parada** del algoritmo son tres:

$$|x_{n+1} - x_n| < Tol_1,$$

donde Tol_1 es una tolerancia elegida por el usuario, o bien

$$|f(x_n)| < Tol_2,$$

o finalmente se supera un número máximo de iteraciones Max también elegida por el usuario

$$n > Max.$$

En la práctica se utilizan una combinación de estos tres criterios de parada.

Una extensión obvia del método de la secante es el **método de Muller**. Comenzando con tres valores $x_{-2}, x_{-1}, x_0 \in [a, b]$ se puede usar interpolación cuadrática, es decir, construiremos el polinomio de segundo grado que interpola por los puntos $(x_{-2}, f(x_{-2})), (x_{-1}, f(x_{-1}))$ y $(x_0, f(x_0))$ y hallar el punto de intersección con el eje de abscisas que sería x_1 . Repitiendo este proceso se obtiene una sucesión que en determinadas condiciones aproximan la raíz de la función (ver pg. 294 de [3]). El principal inconveniente de este método es que las raíces pueden salir complejas y por tanto estaríamos obligados a usar aritmética compleja que de momento es preferible evitar.

4. EL ALGORITMO ZEROIN

La primera versión del algoritmo zeroin fue desarrollado por T.J.Dekker y sus colaboradores en 1969, aunque versiones posteriores lo han mejorado considerablemente. La idea de este algoritmo es la de combinar la versatilidad del método de bisección con la velocidad del método de la secante y el método IQI (inverse quadratic interpolation).

El **método IQI**, en lugar de interpolar en x , interpola en la variable y para determinar el polinomio cuadrático $P(y)$ que verifica las condiciones

$$(4.1) \quad x_n = P(f(x_n)), \quad x_{n-1} = P(f(x_{n-1})), \quad x_{n-2} = P(f(x_{n-2})),$$

que siempre corta al eje de abscisas en $y = 0$. La siguiente iteración es

$$(4.2) \quad x_{n+1} = P(0).$$

El método IQI puede tener serios problemas, por ejemplo necesita que $f(x_n)$, $f(x_{n-1})$ y $f(x_{n-2})$ sean diferentes lo que no está garantizado, pero tiene la gran ventaja de que es muy rápido cerca de la raíz.

A grandes rasgos el algoritmo es como sigue

Algoritmo 4.1.

- Se comienza con a y b tales que $f(a)$ y $f(b)$ tienen signos opuestos.
- Se usa el método de la secante para calcular c .
- Se repiten los siguientes pasos hasta que $|b - a| < \epsilon|b|$ o $f(b) = 0$.
- Se ordenan a, b, c tales que
 - $f(a)$ y $f(b)$ tienen signos opuestos,
 - $|f(b)| \leq |f(a)|$,
 - c es el anterior a b .

- Si $c \neq a$ se da un paso con el algoritmo IQI.
- Si $c = a$ se da un paso con el método de la secante.
- Si el paso dado está en $[a, b]$ se toma.
- Si el paso no está en $[a, b]$ se usa bisección.

La implementación en MATLAB del algoritmo zeroin es el comando `fzero` que permite con gran eficacia aproximar raíces de ecuaciones no lineales. Para el ejemplo 2.1 el resultado de la instrucción `fzero('x-.5*sin(x)-.7',1)` es 1.1580 y $f(1.1580) = -1.1102 \times 10^{-16}$.

Moler en [8] da otra versión llamada `fzerotx` con dos argumentos, el primero especifica la función cuyo cero se quiere calcular y el segundo el intervalo $[a, b]$ de búsqueda, el resulta de `fzerotx('x-.5*sin(x)-.7',[1,1.2])` es otra vez 1.1580.

Finalmente el fichero `fzerogui.m` muestra gráficamente el comportamiento del algoritmo zeroin, si teclea la instrucción `fzerogui('x-.5*sin(x)-.7',[1,1.2])` se abrirá una ventana gráfica donde puede optar por correr el algoritmo eligiendo el botón `auto`, o bien puede ir eligiendo la opción del método: bisección está marcado en rojo, el punto secante en verde y el punto del algoritmo IQI en azul.

5. ITERACIONES DE PUNTO FIJO

La iteración de punto fijo es la técnica más importante para calcular los ceros de funciones no lineales. Se comienza escribiendo la ecuación (1.1) de la forma

$$(5.1) \quad x = g(x),$$

tal que los ceros de la función f son exactamente los puntos fijos de la función g . Geométricamente resolver (5.1) significa hallar la intersección de la gráfica de g con la recta $y = x$.

Existen varias formas de reescribir (1.1) en la forma (5.1), por ejemplo se puede tomar $g(x) = x - f(x)$. La forma de construir la función g diferenciará a unos métodos iterativos de otros.

Después de elegir la llamada **función de iteración** g , el método iterativo construye una sucesión $\{x_n\}$ usando la recurrencia

$$(5.2) \quad x_{n+1} = g(x_n), \quad n = 0, 1, 2, \dots,$$

comenzando con un dato inicial x_0 . La figura 2 ilustra la geometría del método iterativo en el ejemplo 2.1 con $g(x) = 0.7 + 0.5 \sin x$ comenzando en $x_0 = 0$.

La cuestión que ahora se plantea es cuando y en que condiciones la sucesión $\{x_n\}$ converge a la raíz α de la función f que es el punto fijo de g .

Existen varios tipos de teoremas de punto fijo, el siguiente da una condición suficiente de convergencia.

Teorema 5.1. Teorema de convergencia semilocal Sea g una función Lipschitzian, es decir,

$$(5.3) \quad |g(x) - g(x')| \leq \lambda |x - x'|, \quad \forall x, x',$$

con $0 \leq \lambda < 1$ (g **contractiva**) cuyo dominio incluye un intervalo $I = [x_0 - \rho, x_0 + \rho]$. Si además $|x_0 - g(x_0)| \leq (1 - \lambda)\rho$, entonces

- La sucesión está bien definida, es decir, $x_n \in I \quad \forall n$.
- Los iterantes x_n convergen a un punto fijo α de g tal que $|x_n - \alpha| \leq \lambda^n \rho$.
- α es el único punto fijo de g en I .

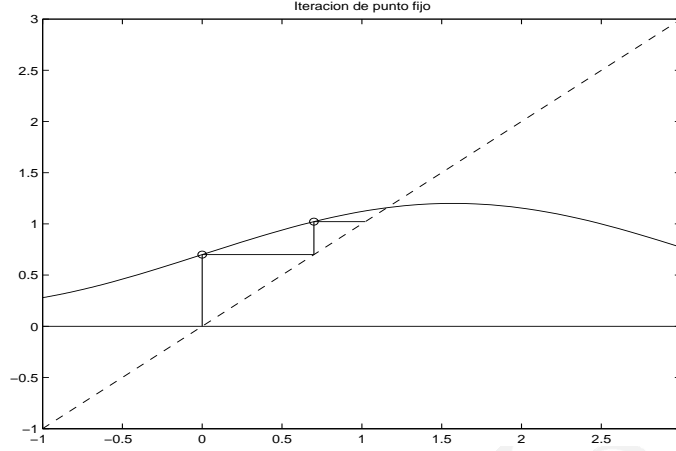


FIGURA 2. Iteración de punto fijo para el ejemplo 2.1.

Demostración. Ver página 251 de [3] ó página 86 de [6]. \square

El hecho de que $|g(x) - g(x_0)| \leq \lambda|x - x_0|$ indica que en la proximidades de x_0 la gráfica de g está incluida en la región del plano $|y - g(x_0)| \leq \lambda|x - x_0|$ cuyas fronteras son las rectas $y - g(x_0) = \lambda(x - x_0)$ e $y - g(x_0) = -\lambda(x - x_0)$. Entonces está claro que si $\lambda > 1$ no se puede garantizar la existencia del punto fijo.

Es de notar que si g es derivable en I con $|g'(x)| \leq \lambda < 1$, entonces g es Lipschitziana por el teorema del valor medio y además

$$\frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = \frac{|g(x_n) - g(\alpha)|}{|x_n - \alpha|} = |g'(\xi_n)|,$$

con ξ_n entre x_n y α . Luego si además g' es continua

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = |g'(\alpha)|.$$

El teorema demostrado se dice que es de tipo semilocal porque:

- (a) No se supone existencia del punto fijo.
- (b) El punto inicial x_0 debe satisfacer alguna condición.
- (c) Se obtienen cotas para el error: $|x_{n+1} - \alpha| \leq \lambda^n |x_0 - \alpha|$.

Los llamados teoremas de punto fijo globales son los resultados con convergencia para cualquier valor inicial

Teorema 5.2. *Teorema de convergencia global* Suponiendo que $g : \mathbb{R} \rightarrow \mathbb{R}$ verifica que $|g(x) - g(x')| \leq \lambda|x - x'|$, $\forall x, x' \in \mathbb{R}$ con $|\lambda| < 1$. Entonces g tiene un único punto fijo al que converge la iteración $x_{n+1} = g(x_n)$ para todo valor inicial x_0 .

Demostración. Similar a la anterior. \square

En los teoremas de convergencia local, se supone de antemano la existencia del punto fijo y el resultado de convergencia se demuestra para cuando el valor inicial x_0 se encuentra en un entorno de punto fijo buscado.

Teorema 5.3. Teorema de convergencia local Sea g una función con derivada continua en un intervalo de la forma $(\alpha - r_0, \alpha + r_0)$ tal que $r_0 > 0$ y $g(\alpha) = \alpha$. Entonces si

$$(5.4) \quad |g'(\alpha)| < 1,$$

existe $0 < r < r_0$ tal que para $x_0 \in [\alpha - r, \alpha + r]$ la sucesión $x_{n+1} = g(x_n)$, $n = 0, 1, \dots$ está en $[\alpha - r, \alpha + r]$ y converge hacia α . Además, si los errores $e_n = x_n - \alpha \neq 0$ sus cocientes $\frac{e_{n+1}}{e_n} \rightarrow g'(\alpha)$ cuando $n \rightarrow \infty$.

Demostración. Página 126 de [12] y página 110 de [6]. □

Es muy importante destacar que este resultado es un teorema de convergencia local porque la convergencia sólo se garantiza si el valor inicial x_0 está próximo a la raíz α que se busca. Además, si el error $e_n = 0$, entonces $x_n = \alpha$ y todos los términos de la iteración también coinciden con α .

Volviendo al ejemplo 2.1 con $g(x) = 0.7 + 0.5 \sin x$, el módulo de su derivada $|g'(x)| = |0.5 \cos x| < 1$ de donde se deduce la convergencia de la figura 2.

Aunque la condición de convergencia (5.4) es sólo suficiente es casi necesaria por el siguiente teorema

Teorema 5.4. Puntos fijos repulsivos Suponiendo que g es una función con derivada continua en $(\alpha - r_0, \alpha + r_0)$ con $r_0 > 0$, $g(\alpha) = \alpha$ y

$$|g'(\alpha)| > 1.$$

Las únicas sucesiones de la forma $x_{n+1} = g(x_n)$, $n = 0, 1, \dots$ que convergen a α son aquellas cuyos términos son, de uno en adelante, iguales a α .

Demostración. Página 132 de [12] □

Además de la convergencia de la sucesión a la raíz α es también importante la velocidad con que lo hace porque ello condiciona el número de pasos necesarios y por tanto el tiempo de computación. Un método iterativo se dice que tiene un orden de convergencia $p \geq 1$ si

$$(5.5) \quad |\alpha - x_{n+1}| \leq \lambda |\alpha - x_n|^p, \quad n \geq 0,$$

para algún $\lambda > 0$. Si $p = 1$ se dice que la convergencia es lineal, si $p = 2$ la convergencia es cuadrática, cúbica cuando $p = 3 \dots$

Ejemplo 5.5. El método de la secante tiene una convergencia superlineal ($p > 1$), exactamente el orden es la famosa razón áurea $\frac{1+\sqrt{5}}{2} \approx 1.62$ (Teorema 2.1 de la página 49 de [2]).

Para la convergencia lineal con $\lambda = |g'(\alpha)| < 1$, para mejorar r decimales en la solución el número m de iteraciones necesarias será

$$\lambda^m < 10^{-r} \quad \Leftrightarrow \quad \frac{r}{-\log_{10} \lambda} < m,$$

donde el factor $-\log_{10} \lambda$ se denomina razón de convergencia del método. De esta forma

1. Si $0 \ll \lambda < 1$ (λ próximo a uno), $\log_{10} \lambda \approx 0$ por lo que m es muy grande y la convergencia muy lenta.

2. Por el contrario si $0 < \lambda \ll 1$ (λ cercano a cero), $-\log_{10} \lambda$ es un valor grande y m pequeño, es decir, con pocas iteraciones se mejora mucho la aproximación.

Teorema 5.6. Convergencia cuadrática Suponiendo que g verifica las condiciones del teorema 5.3, $g'(\alpha) = 0$ y existe $g''(\alpha)$, entonces existe $0 < r < r_0$ tal que para $x_0 \in [\alpha - r, \alpha + r]$ la sucesión $x_{n+1} = g(x_n)$, $n = 0, 1, \dots$ está en $[\alpha - r, \alpha + r]$ y converge hacia α .

Además, si los errores son no nulos $\frac{e_{n+1}}{e_n^2} \rightarrow \frac{1}{2}g''(\alpha)$ cuando $n \rightarrow \infty$.

Demostración. Página 129 de [12] □

6. EL MÉTODO DE NEWTON

Suponga que se quiere sustituir la función f no lineal por una recta que pasa por el punto $(x_n, f(x_n))$ de ecuación

$$y - f(x_n) = m(x - x_n).$$

Esta recta corta al eje de abscisas en el valor de $x = x_n - \frac{1}{m}f(x_n)$ y el método iterativo que sugiere es

$$x_{n+1} = x_n - \frac{1}{m}f(x_n).$$

Introduciendo los errores $e_n = x_n - \alpha$ se tiene que

$$\alpha + e_{n+1} = \alpha + e_n - \frac{1}{m}f(\alpha + e_n),$$

y suponiendo ahora regularidad de f en α

$$e_{n+1} = e_n - \frac{1}{m} \left(e_n f'(\alpha) + \frac{1}{2} e_n^2 f''(\alpha) + \dots \right),$$

o también

$$e_{n+1} = \left(1 - \frac{1}{m} f'(\alpha) \right) e_n - \frac{1}{2m} e_n^2 f''(\alpha) - \dots$$

La mejor elección de la pendiente será entonces $m = f'(\alpha)$, sin embargo como α no es un valor conocido se toma $m = f'(x_n)$ para llegar al conocido como **método de Newton-Raphson**

$$(6.1) \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

suponiendo siempre que la derivada f' no se anula en los puntos x_n . Este método simplemente sustituye la función por la recta tangente a la curva en el punto de abscisa x_n .

La iteración de Newton-Raphson es una iteración de punto fijo para la función de iteración

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Notese que efectivamente si $f(\alpha) = 0$ entonces $g(\alpha) = \alpha$. La velocidad de convergencia depende de $|g'(\alpha)|$, derivando se tiene

$$(6.2) \quad g'(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2}.$$

Como $f(\alpha) = 0$, si $f'(\alpha) \neq 0$, es decir α es raíz simple de f , $g'(\alpha) = 0$ y la convergencia es cuadrática. Más concretamente se tiene el siguiente teorema

Teorema 6.1. *Suponiendo que f tiene dos derivadas continuas en un entorno de α con $f(\alpha) = 0$ y $f'(\alpha) \neq 0$. Entonces si x_0 está suficientemente próximo a α , la sucesión $x_n \rightarrow \alpha$ y además*

$$(6.3) \quad \lim_{n \rightarrow \infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^2} = -\frac{f''(\alpha)}{2f'(\alpha)}.$$

Demostración. Página 54 de [2]. □

La principal ventaja del método de Newton es esta convergencia cuadrática cuando $f'(\alpha) \neq 0$, pero tiene evidentes dificultades cuando la raíz α es múltiple porque entonces $f'(\alpha) = 0$. Si α tiene multiplicidad $p > 1$ entonces $f(x) = (x - \alpha)^p h(x)$ con $h(\alpha) \neq 0$, el método de Newton puede todavía converger pero en general pierde la convergencia cuadrática porque la función de iteración es

$$g(x) = x - \frac{(x - \alpha)h(x)}{ph(x) + (x - \alpha)h'(x)},$$

y por lo tanto $g'(\alpha) = 1 - \frac{1}{p} \neq 0$.

Ejemplo 6.2. Para el polinomio $p(x) = x^m$ con $m > 1$, $\frac{f(x)}{f'(x)} = \frac{x}{m}$ y el método de Newton es

$$x_{n+1} = x_n - \frac{x_n}{m} = \frac{m-1}{m}x_n,$$

que evidentemente tiene convergencia lineal porque

$$|x_{n+1} - \alpha| = \frac{m-1}{m}|x_n - \alpha|, \quad (\alpha = 0).$$

Incluso puede darse el caso perverso de que el método de Newton se eternice oscilando entorno a una raíz.

Ejemplo 6.3. Si se considera la iteración

$$(6.4) \quad x_{n+1} - a = -(x_n - a) \quad \forall n,$$

entonces $f(x)$ debe satisfacer la ecuación diferencial de variables separadas

$$(6.5) \quad x - \frac{f(x)}{f'(x)} - a = -(x - a),$$

cuya solución exacta es

$$(6.6) \quad f(x) = \text{sign}(x - a)\sqrt{|x - a|}.$$

Evidentemente a es una raíz de f pero el método de Newton nunca llegará a alcanzarla porque calculado x_n

$$x_{n+1} = 2a - x_n \quad \Rightarrow \quad x_{n+2} = 2a - x_{n+1} = 2a - (2a - x_n) = x_n.$$

En las implementaciones del método de Newton es importante el criterio de parada de la iteración que son de tres tipos:

1. En el test de tamaño se itera hasta que $|f(x_n)| < \varepsilon_1$.
2. En el segundo test natural se pide que $|x_{n+1} - x_n| \leq \varepsilon_2$.
3. Se fija un número máximo de iteraciones.

Existen ejemplos (ver por ejemplo [2] p. 70) donde fallan uno u otro test por lo que en general es aconsejable siempre usar más de uno.

7. INTRODUCCIÓN A LOS SISTEMAS NO LINEALES

Se supone que f_1, \dots, f_n son funciones de un subconjunto de \mathbb{R}^n en \mathbb{R} y se quieren encontrar las soluciones del sistema de ecuaciones no lineales

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0, \\ \vdots & \\ f_n(x_1, \dots, x_n) &= 0, \end{aligned}$$

que vectorialmente se escribe de la forma

$$(7.1) \quad \mathbf{f}(\mathbf{x}) = \mathbf{0}.$$

con los vectores $\mathbf{x} = (x_1, \dots, x_n)^T$ y $\mathbf{f} = (f_1, \dots, f_n)^T$.

Los métodos numéricos para resolver estos problemas son dos tipos

- (a) **Métodos iterativos** que se estudian en esta lección.
- (b) **Métodos para minimizar** la función escalar

$$F(\mathbf{x}) = \mathbf{f}^T \mathbf{f} = \|\mathbf{f}\|^2 = \sum_{k=1}^n f_k^2(\mathbf{x}).$$

Métodos que no se estudian en esta lección. El capítulo 7 de [11] o el capítulo 9 de [1] son dos buenas introducciones a estos métodos.

Los métodos iterativos de punto fijo, como en el caso escalar, escriben (7.1) de la forma conocida $\mathbf{x} = \mathbf{g}(\mathbf{x})$ y construyen la iteración de punto fijo $\mathbf{x}_{n+1} = \mathbf{g}(\mathbf{x}_n)$ que en condiciones adecuadas debe converger al cero de \mathbf{f} .

Todos los resultados escalar se generalizan a la situación actual, e incluso a la más general de espacios de Banach.

Teorema 7.1. Teorema de convergencia semilocal Sea $\|\cdot\|$ una norma de \mathbb{R}^n , si

$$\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \leq \lambda \|\mathbf{x} - \mathbf{y}\|, \quad 0 \leq \lambda < 1,$$

para \mathbf{x}, \mathbf{y} en una bola cerrada $\bar{B}(\mathbf{x}_0, \rho)$, donde el primer iterante es tal que verifica que $\|\mathbf{g}(\mathbf{x}_0) - \mathbf{x}_0\| \leq (1 - \lambda)\rho$. Entonces

- (i) Cada iteración está en $\bar{B}(\mathbf{x}_0, \rho)$.
- (ii) La iteración converge a un punto fijo α de la función \mathbf{g} .
- (iii) α es el único punto fijo de \mathbf{g} en la bola $\bar{B}(\mathbf{x}_0, \rho)$.

Demostración. Basta en la demostración del caso escalar cambiar los $|\cdot|$ por $\|\cdot\|$. \square

Se trata por tanto de conocer cuando la función \mathbf{g} es Lipschitziana con una constante $\lambda \leq 1$. Suponiendo que \mathbf{g} es de clase \mathcal{C}^1 en una bola

$$\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \leq \sup_{\xi \in B} \|\mathcal{J}(\xi)\| \cdot \|\mathbf{x} - \mathbf{y}\|,$$

donde \mathcal{J} es el jacobiano con la norma matricial deducida de la norma vectorial utilizada. Por tanto la contractividad depende de la matriz jacobina y el valor de su norma.

Por otra parte, si $\mathcal{J}(\alpha) \neq 0$ la iteración es de primer orden porque

$$\lim \frac{\|\mathbf{x}_{n+1} - \alpha\|}{\|\mathbf{x}_n - \alpha\|} \leq \|\mathcal{J}(\alpha)\|.$$

Si $\mathcal{J}(\alpha) \equiv 0$ utilizan Taylor

$$g_i(\mathbf{x}) - g_i(\alpha) = \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n \frac{\partial^2 g_i(\xi_i)}{\partial x_j \partial x_k} (x_j - \alpha_j)(x_k - \alpha_k),$$

si ahora se supone que en una bola de centro α

$$\max_{i,j,k} \left| \frac{\partial^2 g_i}{\partial x_j \partial x_k} \right| \leq \frac{2M}{n^2},$$

entonces

$$\|\mathbf{g}(\mathbf{x}) - \alpha\|_\infty \leq M \|\mathbf{x} - \alpha\|_\infty^2,$$

y por tanto

$$\|\mathbf{x}_{n+1} - \alpha\|_\infty \leq M \|\mathbf{x}_n - \alpha\|_\infty^2,$$

que concluye la convergencia de orden dos en $\|\cdot\|_\infty$ y en cualquier otra norma.

Ejemplo 7.2. Resuelva el sistema no lineal.

$$\begin{aligned} x + 2y &= 3, \\ 2x^2 + y^2 &= 5. \end{aligned}$$

primero de forma gráfica y después con un método iterativo convergente.

8. EL MÉTODO DE NEWTON PARA SISTEMAS NO LINEALES

El método de Newton para sistemas utiliza la misma idea que en el problema escalar. Suponiendo que $\mathbf{f} \in \mathcal{C}^1$, $\mathbf{x}^{(n+1)}$ es el cero de la ecuación linealizada, es decir

$$(8.1) \quad \mathbf{f}(\mathbf{x}_n) + \mathcal{J}(\mathbf{x}_n)(\mathbf{x}_{n+1} - \mathbf{x}_n) = \mathbf{0}.$$

Si la matriz Jacobiana $\mathcal{J}(\mathbf{x}_n)$ es invertible se puede escribir de la forma

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathcal{J}^{-1}(\mathbf{x}_n)\mathbf{f}(\mathbf{x}_n),$$

de modo que la iteración de Newton está asociada a la función

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - \mathcal{J}^{-1}(\mathbf{x})\mathbf{f}(\mathbf{x}).$$

Entonces, cuando $\mathbf{f} \in \mathcal{C}^2$ y el jacobiano no se anula en un entorno de la raíz α , el jacobiano de \mathbf{g} es

$$\mathbf{I} - \mathcal{J}^{-1}(\alpha)\mathcal{J}(\alpha) - (\mathcal{D}\mathcal{J}^{-1})(\alpha)\mathbf{f}(\alpha) \equiv \mathbf{0},$$

y por tanto se concluye que la convergencia es de orden dos. El resultado de [convergencia local](#), tomaría \mathbf{x}_0 próximo a la raíz α y entonces demostraría que todas las iteraciones del método de Newton están definidas y convergen a dicha raíz α con orden dos.

También existen algunos resultados de [convergencia semi-global](#), el más conocido es [Teorema de Newton-Kantorovich de 1949](#), cuyo enunciado se puede encontrar en la página 256 de [3] aunque para la demostración hay que remitirse al clásico [9]. Este y otros resultados parecidos, en general, tienen poco valor práctico porque las hipótesis son demasiados fuertes o complicadas de comprobar.

El método de Newton tienen una ventaja fundamental: su **convergencia cuadrática**, pero también tienen serios inconvenientes

- (a) **Convergencia local**: la sucesión de iteraciones está definida sólo es convergente cuando el valor inicial \mathbf{x}_0 está próximo a la raíz.
- (b) **Cuando la dimensión del problema N es grande el esfuerzo de evaluar N^2 derivadas en cada iteración es prohibitivo.**
- (c) Hay que resolver un sistema lineal en cada iteración.

Ejemplo 8.1. Aplique el método de Newton para la resolver el sistema no lineal:

$$\begin{aligned}x^2 + 4y^2 &= 5, \\2x^2 - 2x - 3y &= 2.5.\end{aligned}$$

9. MODIFICACIONES AL MÉTODO DE NEWTON

En orden de paliar los inconvenientes arriba apuntados, se han introducido modificaciones al método de Newton. La dos más sencillas son las siguientes:

1. Para rebajar el coste de resolver el sistema lineal en cada iteración, se puede fija el jacobiano \mathcal{J} un número de iteraciones, dos, tres,...iteraciones. Como el mayor coste operativo estriba en la factorización LU, se ahorra dicha factorización y de esta manera la iteración es más barato. Desgraciadamente esto bajo el orden de convergencia a uno.
2. Otra modificación muy utilizada en la práctica es la de sustituir las derivadas por cocientes en diferencias

$$\frac{\partial f_i}{\partial x_j} \approx \frac{f_i(x_1, \dots, x_j + h_j, \dots, x_N) - f_i(x_1, \dots, x_j, \dots, x_N)}{h_j}$$

de forma que (8.1) ahora es

$$\mathbf{f}(\mathbf{x}_n) + \Delta \mathbf{f}(\mathbf{x}_n) (\mathbf{x}_{n+1} - \mathbf{x}_n) = \mathbf{0}.$$

En esta nueva versión, en lugar de evaluar N^2 derivadas, sólo tenemos que realizar $N+1$ evaluaciones de la función, aunque se tiene un nuevo problema porque si se toma el paso h_j demasiado grande la aproximación a la derivada es mala y por el contrario si es muy pequeño podemos llegar a una matriz singular.

10. ALGUNOS COMENTARIOS PARA TERMINAR

Para finalizar la lección algunos temas que por razones de tiempo no se han estudiado:

1. Los métodos para acelerar la convergencia, fundamentalmente el método Δ^2 de Aitken y el método de Steffensen que se pueden consultar en la sección 2.5 de [3].
2. Todos los métodos especiales para calcular raíces de polinomios: sucesiones de Sturm, el algoritmo QD de Rutishauser o también el método de Bairstow para determinar factores cuadráticos que se pueden ver la sección 3.5 de [7] y también en la sección 5.8 de [3].

Para seguir aprendiendo y conocer más detalles de este importante tópico en Análisis Numérico se recomiendan las siguientes lecturas: [1, Capítulos 2 y 9], [5, Capítulo 2], [4, Capítulo 4], [13, Capítulo 4] y [10, Capítulo 2].

REFERENCIAS

1. U.M. Ascher and C. Greif, *A Firts Course in Numerical Methods*, SIAM, 2011.
2. K.E. Atkinson, *An Introduction to Numerical Analysis*, Wiley, 1978.
3. R. Bulirsch and J. Stoer, *Introduction to Numerical Analysis*, Springer, 1980.
4. W. Gautschi, *Numerical Analysis. Second Edition*, Birkauer, 2012.
5. M.H. Holmes, *Introduction to Scientific Computing and Data Analysis*, Springer, 2016.
6. E. Isaacson and H. Keller, *Analysis of Numerical Methods*, Wiley, 1966.
7. D. Kincaid and W. Cheney, *Análisis numérico. Las matemáticas del cálculo científico*, Addison-Wesley Iberoamericana, 1994.
8. C.B. Moler, *Numerical Computing with MATLAB*, SIAM, 2004.
9. J.M. Ortega and W.C. Rheinboldt, *Nonlinear Equation in Several Variables*, Academic Press, 1969.
10. A. Quarteroni and F. Saleri, *Scientific Computing with MATLAB and Octave. Second Edition*, Springer, 2006.
11. R.Sacco, F.Saleri, and A.Quarteroni, *Numerical Mathematics*, Springer, 2000.
12. J.M. Sanz-Serna, *Diez lecciones de Cálculo Numérico*, Universidad de Valladolid, 2010.
13. W.Y. Yang, W. Cao, T. Chung, and J. Morris, *Applied Numerical Methods using MATLAB*.

DEPARTAMENTO DE MATEMÁTICAS DE LA UNIVERSIDAD DEL PAIS VASCO
Email address: fernando.vadillo@ehu.es