Ecuaciones no lineales

Damián Ginestar Peiró

Departamento de Matemática Aplicada

Universidad Politécnica de Valencia

Curso 2011-2012

Índice

Ecuaciones no lineales

- Los métodos para la búsqueda de raíces de una función generalmente se clasifican en
  - Métodos cerrados. Se parte de un intervalo en el que se sabe que hay al menos una raíz y convergen siempre. Método de la Bisección.
  - Métodos abiertos. Se parte de una aproximación inicial y tienen un cierto radio de convergencia. Método del punto fijo, Método de Newton, Método de la secante.

Método de la bisección

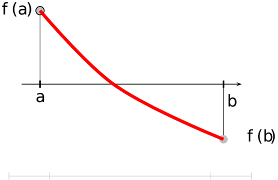
Supongamos, por ejemplo, que se quiere calcular una solución de la ecuación

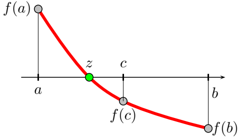
$$x^2 = 2 \ ,$$

sabiendo que la solución está en un intervalo $[a, b] = [1, 2]$.
Probamos con un punto c que sea el punto medio del intervalo

$$c = \frac{a+b}{2} \ ,$$

calculamos $c^2 = (1{,}5)^2 = 2{,}25$. Como $c^2 > 2$, entonces tendremos que la raíz estará en un nuevo intervalo $[a', b'] = [a, c]$.
Repitiendo esta estrategia se van obteniendo intervalos cada vez más pequeños que contienen la raíz buscada.





Método de la bisección

```
M=2;
a=1;
b=2;
k=0;
tol=1.0e-5;
while (b-a) >tol
  x=(a+b)/2;
  if x^2>M
     b=x;
  else
     a=x;
  end
k=k+1;
end
sol=(a+b)/2
```

Método de la bisección

Bisección

Se parte de una intervalo $[a_n, b_n]$ tal que $f(a_n) f(b_n) < 0$. Entonces $f(x)$ tiene una raíz en el intervalo. Se construye un intervalo $[a_{n+1}, b_{n+1}]$ que también la contiene del siguiente modo. Se calcula

$$p_n = a_n + \frac{b_n - a_n}{2} \; ,$$

y luego

$$a_{n+1} = a_n \;\; \text{y} \;\; b_{n+1} = p_n \;\; \text{si} \;\; f(a_n) f(p_n) < 0 \; ,$$

o bien

$$a_{n+1} = p_n \;\; \text{y} \;\; b_{n+1} = b_n \;\; \text{si} \;\; f(a_n) f(p_n) > 0 \; ,$$

Método del punto fijo

Punto fijo

Diremos que un punto p es un punto fijo de una función $\varphi(x)$ si se satisface $\varphi(p) = p$.

Supongamos que se busca una raíz de una ecuación

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

el método del punto fijo consiste en reescribir esta ecuación de la forma

$$x = \varphi(x) \ ,$$

y construir una sucesión de la forma

$$\begin{aligned} x_1 &= \varphi(x_0) \\ x_2 &= \varphi(x_1) \\ &\vdots \\ x_{n+1} &= \varphi(x_n) \end{aligned}$$

Método del punto fijo

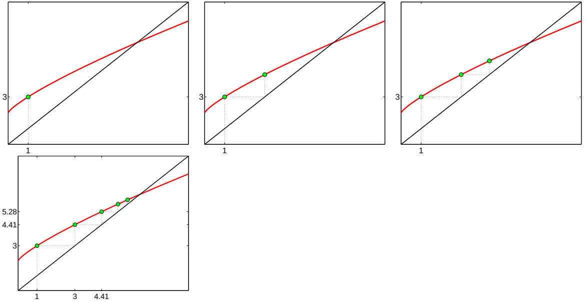
Dada la ecuación

$$f(x) = x - x^{4/5} - 2 = 0$$

si tomamos la función de iteración

$$g(x) = x^{4/5} + 2$$

Las iteraciones



Método del punto fijo

Si existe un m tal que $x_{m+1} \approx x_m$, se cumplirá que $x_m \approx \varphi(x_m)$ y, por tanto, se podrá tomar como valor aproximado de la raíz x_m.

Ejemplo

Suponemos que se quiere buscar una raíz de la función

$$f(x) = x^3 + 4x^2 - 10$$

en $[1, 2]$. *Se pueden hacer diferentes elecciones de la función* $\varphi(x)$, *por ejemplo,*

a) $\varphi_1(x) = x - x^3 - 4x^2 + 10;$

b) $\varphi_2(x) = \left(\frac{10}{x} - 4x\right)^{1/2};$

c) $\varphi_3(x) = \frac{1}{2}\left(10 - x^3\right)^{1/2};$

d) $\varphi_4(x) = (10/(4 + x))^{1/2};$

e) $\varphi_5(x) = x - \left(x^3 + 4x^2 - 10\right) / \left(8x + 3x^2\right)$

Método del punto fijo

Resultado para la iteración del punto fijo $x = \varphi(x)$.

n	$a)$	$b)$	$c)$	$d)$	$e)$
0	1.5	1.5	1.5	1.5	1.5
1	-0.875	0.8165	1.286953768	1.348399725	1.37333333
2	6.732	2.9969	1.402540804	1.367376372	1.365262015
3	-469.7	–	1.345458374	1.364957015	1.365230014
4	1.03×10^8		1.375170253	1.365264748	1.365230013
5			1.360094193	1.365225594	
6			1.367846968	1.365230576	
7			1.363887004	1.365229942	
8			1.365916734	1.365230022	
9			1.364878217	1.365230012	
10			1.365410062	1.365230014	

Método del punto fijo

Teorema (Teorema del punto fijo)

Sea $\varphi : [a, b] \to [a, b]$ *una función continua en* $[a, b]$ *y derivable en* $]a, b[$*, y además que cumple que* $|\, \varphi'(x) \,| \leq k < 1$, $\forall x \in]a, b[$*. Entonces existe un único* $c \in]a, b[$ *tal que* $\varphi(c) = c$*. Además para todo* $x_0 \in]a, b[$*, la sucesión* $x_0, \ \{x_n\}_{n=0}^{\infty}$ *obtenida de la forma* $x_n = \varphi(x_{n-1})$ *converge a* c*.*

Este teorema ayuda a elegir una función $\varphi(x)$ adecuada,

$$\begin{aligned}
\varphi_4(x) &= (10/(4+x))^{1/2}\,; \\
\varphi_4^{'}(x) &= \sqrt{10}\,(-1/2)\,(4+x)^{-3/2}; \\
\left|\varphi_4^{'}(x)\right| &= \sqrt{10}/2\left|(4+x)^{-3/2}\right| < \sqrt{10}/(2 \cdot 5^{3/2}) = 0{,}141 < 1\ .
\end{aligned}$$

La sucesión converge en $[1, 2]$.

Método de Newton

Este método se basa en utilizar el desarrollo de Taylor. Escribimos $x_1 = x_0 + \Delta x$,

$$f\left(x_1\right) = f\left(x_0\right) + f'\left(x_0\right)\Delta x + o\left(\Delta x^2\right) \ ,$$

y suponiendo que $f\left(x_1\right) = 0$, queda

$$x_1 = x_0 - \frac{f\left(x_0\right)}{f'\left(x_0\right)} \ .$$

El método de Newton-Raphson se basa en esta ecuación y consiste en calcular los valores de una sucesión de la forma

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

Método de Newton

Otro modo de obtener este método consiste en suponer que $f : [a, b] \to \mathbb{R}$ es continua en $[a, b]$ y tal que $f''(x)$ no cambia de signo en $[a, b]$ con $f(a)f(b) < 0$. El proceso para encontrar un x tal que $f(x) = 0$ consiste en lo siguiente:

1) Fijamos $c = a$ ó b, tal que $f(c)f''(x) > 0$, $\forall x \in (a, b)$.
2) $x_0 = c$.
3) Hallamos la ecuación de la tangente que pasa por $(x_0, f(x_0))$ y el punto de corte de dicha tangente con el eje X. El proceso se repite hasta conseguir una sucesión de aproximaciones que converge a la raíz de $f(x) = 0$.

Método de Newton

La ecuación de la recta tangente a la curva $y = f(x)$ en $(x_{n-1}, f(x_{n-1}))$ viene dada por

$$y = f(x_{n-1}) + f'(x_{n-1})(x - x_{n-1}) \quad .$$

La abscisa del punto de intersección de la recta tangente con el eje X,

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

y mediante esta relación obtenemos una sucesión, $\{x_n\}_{n=1}^{\infty}$ de aproximaciones al valor de la raíz buscada.
El error que se comete en la iteración n-ésima será

$$\mid r - x_n \mid < \frac{M}{2m} \mid x_n - x_{n-1} \mid ,$$

donde $0 < m \leq \mid f(x) \mid$ y $\mid f''(x) \mid \leq M$, $\forall x \in (a, b)$.

Método de la secante

- Una de las desventajas del método de Newton para obtener las raíces de una ecuación de la forma
$$f(x) = 0 \ ,$$
es que es necesario conocer la derivada $f'(x)$.
- En ocasiones esta derivada es difícil de calcular o no se dispone de la misma, y se utiliza una aproximación de la forma
$$s_n = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \ ,$$
obteniendo el método de la secante, que se basa en iteraciones de la forma
$$x_{n+1} = x_n - \frac{f(x_n)}{s_n} \ .$$
El método de la secante hace uso de dos aproximaciones iniciales para la raíz.

Sistemas de ecuaciones

Partimos, por ejemplo, de un sistema de dos ecuaciones de la forma

$$\begin{aligned} f_1(x,y) &= 0 \; , \\ f_2(x,y) &= 0 \; , \end{aligned}$$

del que se pretende obtener la solución. Para utilizar el método del punto fijo, se reescribe el sistema de la forma

$$\begin{aligned} x &= g_1(x,y) \; , \\ y &= g_2(x,y) \; . \end{aligned}$$

Se construye la sucesión

$$(x_{n+1}, y_{n+1}) = (g_1(x_n, y_n), g_2(x_n, y_n))$$

que como ocurre en el caso escalar la eficiencia del método dependerá de la elección de las funciones g_1 y g_2.
Se puede hacer una implementación similar al método de Jacobi y otra como la de Gauss-Seidel.

Método iterativo del punto fijo

Ejemplo

Suponemos que se quiere buscar una solución del sistema

$$\begin{aligned} x^2 - 10x + y^2 + 8 &= 0 \ , \\ xy^2 + x - 10y + 8 &= 0 \ . \end{aligned}$$

Se eligen

$$\begin{aligned} x &= \frac{1}{10}\left(x^2 + y^2 + 8\right) \ , \\ y &= \frac{1}{10}\left(xy^2 + x + 8\right) \ . \end{aligned}$$

Se parte de $(x, y) = (0, 0)$ *y, utilizando el método del punto fijo,*

Método iterativo del punto fijo

Resultado de la iteración del punto fijo para el sistema.

Jacobi		Gauss-Seidel	
x_n	y_n	x_n	y_n
0	0	0	0
0.80000	0.80000	0.8	0.88
0.92800	0.93120	0.94144	0.96705
0.97283	0.97327	0.98215	0.99006
0.98937	0.98944	0.99448	0.99693
0.99578	0.99579	0.99829	0.99905
0.99832	0.99832	0.99947	0.9997
0.99933	0.99933	0.99983	0.99991
0.99973	0.99973	0.99995	0.99997
0.99989	0.99989	0.99998	0.99999
0.99996	0.99996	0.99999	1
0.99998	0.99998	1	1
0.99999	0.99999		
1	1		

Método de Newton-Raphson

Este método se basa en utilizar el desarrollo de Taylor. Partimos de un sistema de la forma

$$\begin{aligned} f_1(x, y) &= 0\,, \\ f_2(x, y) &= 0\,, \end{aligned}$$

del que se pretende obtener la solución. Se supone que

$$x = x_0 + \Delta x \;\text{ e }\; y = y_0 + \Delta y$$

Utilizando el desarrollo de Taylor alrededor de (x_0, y_0) y quedándonos con el primer orden, tenemos

$$\begin{aligned} f_1(x_0, y_0) + \frac{\partial f_1(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f_1(x_0, y_0)}{\partial y}\Delta y &\approx 0\,, \\ f_2(x_0, y_0) + \frac{\partial f_2(x_0, y_0)}{\partial x}\Delta x + \frac{\partial f_2(x_0, y_0)}{\partial y}\Delta y &\approx 0, \end{aligned}$$

Método de Newton-Raphson

Si introducimos la notación $\vec{x} = (x, y)$, $\vec{F} = (f_1, f_2)$ y

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}$$

queda

$$\vec{F}(\vec{x_0}) + J(\vec{x_0})(\vec{x} - \vec{x_0}) \approx 0$$

entonces

$$\vec{x} = \vec{x_0} - J^{-1}(\vec{x_0})\,\vec{F}(\vec{x_0})$$

y el método de Newton consiste en calcular la sucesión

$$\vec{x}_{n+1} = \vec{x}_n - J^{-1}(\vec{x}_n)\,\vec{F}(\vec{x}_n).$$

Método de Newton-Raphson

A la hora de implementar el método se hace en dos pasos:

1. Se resuelve el sistema

$$J\left(\vec{x}_n\right)\Delta\vec{x}_{n+1} = -\vec{F}\left(\vec{x}_n\right) \ ,$$

2. Se calcula

$$\vec{x}_{n+1} = \vec{x}_n + \Delta\vec{x}_{n+1} \ .$$

Método de Broyden

- El método de Broyden generaliza el método de la secante para resolver ecuaciones no lineales par el caso de un sistema de ecuaciones.
- Para el caso de un sistema, se ha de obtener una aproximación de la matriz derivada, S_n, que cumpla

$$S_n\left(x_n - x_{n-1}\right) = f\left(x_n\right) - f\left(x_{n-1}\right) \ .$$

Método de Broyden

Introducimos la siguiente notación

$$
\begin{aligned}
b_n &= x_{n+1} - x_n \ , \\
\Delta f_n &= f\left(x_{n+1}\right) - f\left(x_n\right) \ , \\
S_{n+1} &= S_n + C_n \ .
\end{aligned}
$$

Con lo que tenemos que

$$S_{n+1}\left(x_{n+1} - x_n\right) = f\left(x_{n+1}\right) - f\left(x_n\right) \ ,$$

o sea,

$$\left(S_n + C_n\right) b_n = \Delta f_n \ ,$$

y, por tanto,

$$C_n b_n = \Delta f_n - S_n b_n \ .$$

Método de Broyden

De este modo, dado un vetor w_n^T tal que $w_n^T b_n \neq 0$, podemos elegir

$$C_n = \frac{1}{w_n^T b_n} \left(\Delta f_n - S_n b_n \right) w_n^T \ ,$$

ya que

$$C_n b_n = \Delta f_n - S_n b_n \ .$$

Si elegimos $w_n^T = b_n$, se obtiene el primer método de Broyden, y si elegimos $w_n = S_n^T b_n$ se obtiene el segundo método de Broyden.

Con este esquema el método de Broyden no es muy eficiente, ya que lo que se gana al sustituir la matriz jacobiana por la S_n, se pierde al pasar de tener una convergencia cuadrática a una convergencia superlineal.

Fórmula de Sherman-Morrison

Si A es una matriz no singular y z e y son dos vectores, entonces la matrix $A + zy^T$ es no singular siempre que $y^T A^{-1} z \neq -1$. Además,

$$(A + zy^T)^{-1} = A^{-1} - \frac{A^{-1} z y^T A^{-1}}{1 + y^T A^{-1} z} \quad .$$

Método de Broyden

Tomando

$$A = S_{i-1}\ , \qquad z = \frac{(f(x_i) - f\left(x_{i-1}\right) - A_{i-1}\left(x_i - x_{i-1}\right))}{\parallel x_i - x_{i-1} \parallel^2}\ , \qquad y = x_i - x_{i-1}$$

y utilizando la Fórmula de inversión

$$A_i^{-1} = \left(A_{i-1} + \frac{(f\left(x_i\right) - f\left(x_{i-1}\right) - A_{i-1}\left(x_i - x_{i-1}\right))\left(x_i - x_{i-1}\right)^T}{\parallel x_i - x_{i-1} \parallel^2} \right)^{-1}$$

como

$$A_i^{-1} = A_{i-1}^{-1} + \frac{(x_i - x_{i-1} - A_{i-1}^{-1}(f(x_i) - f(x_{i-1})))(x_i - x_{i-1})^T A_{i-1}^{-1}}{(x_i - x_{i-1})A_{i-1}^{-1}(f(x_i) - f(x_{i-1}))}\quad .$$

Método de Broyden

- Esta expresión nos evita construir la matriz A_i y resolver el sistema de ecuaciones asociado, en cada iteración.
- Hay que tener en cuenta que la matriz A_i no es más que una aproximación de la matriz jacobiana y, por tanto, para mejorar la convergencia del método es conveniente alternar un cierto número de iteraciones del método de Broyden con una iteración del método de Newton, en donde se vuelve a calcular la matriz jacobiana.

Métodos de continuación

Es conocido que, en general, el radio de convergencia de los métodos de resolución de sistemas de ecuaciones no lineales es muy pequeño. Por ello, es conveniente utilizar estos métodos tomando como punto de partida un punto cercano a la solución real del problema.
Una estrategia posible para hacer esto consiste en el siguiente método.

Se parte de un sistema de ecuaciones de la forma

$$f(x) = 0$$

y de un punto inicial x_0. Se construye la sucesión de problemas no lineales

$$G(x, k) = 0; \qquad k = N, \ldots, 0 \quad ,$$

siendo

$$G(x, k) = f(x) - \frac{k}{N} f(x_{i-1}) \quad .$$

Métodos de continuación

Etapa primera.- Se toma $k = N$, con lo que $x = x_0$ es una solución trivial del sistema.

Etapa Segunda.- Se toma $k = N - 1$ y se utiliza la solución de la etapa anterior, x_0, como punto inicial en la resolución del problema

$$G(x, N-1) = f(x) - \frac{N-1}{N} f(x_0) = 0 \quad ,$$

La solución obtenida para este problema la denotamos por x_1.

Etapa i-ésima.- Se toma $k = N - i$ y se utiliza la solución de la etapa anterior, x_{i-1}, como punto inicial en la resolución del problema

$$G(x, N-i) = f(x) - \frac{N-i}{N} f(x_{i-1}) = 0 \quad .$$

La solución obtenida para este problema la denotamos por x_i.

Métodos de continuación

Por otra parte, supongamos que $x(\lambda)$ es una única solución de

$$G(\lambda, x) = 0\ , \quad \lambda \in [0,1]\ .$$

El conjunto $\{x(\lambda) \,/\, 0 \leq \lambda \leq 1\}$ puede verse como una curva parametrizada por λ que va desde $x(0) = x^0$ hasta $x(1) = x$. Si las funciones $G(\lambda, x)$ y $x(\lambda)$ son diferenciables, se tiene que

$$\frac{\partial G(\lambda, x(\lambda))}{\partial \lambda} + \sum_{i=1}^{n} \frac{\partial G(\lambda, x(\lambda))}{\partial x_i}\, x_i'(\lambda) = 0\ .$$

Métodos de continuación

Como

$$G\left(\lambda, x(\lambda)\right) = f\left(x(\lambda)\right) + (\lambda - 1)\, f\left(x(0)\right) \ ,$$

tenemos que,

$$\frac{\partial G}{\partial x_i} = \frac{\partial f}{\partial x_i} = DF\left(x\left(\lambda\right)\right) \ , \quad \frac{\partial G}{\partial \lambda} = f\left(x(0)\right) \ ,$$

Con lo que se tiene el sistema

$$DF\left(x(\lambda)\right)x'\left(\lambda\right) = -f\left(x(0)\right)\ ,$$

que es un ecuación diferencial ordinaria para $x\left(\lambda\right)$, con la condición inicial $x(0) = x^0$. Distintos métodos de continuación se obtienen aplicando distintos métodos numéricos a la resolución de este problema de valores iniciales.

Métodos con dirección de búsqueda

- Se puede modificar el método de Newton para ampliar el radio de convergencia del método a costa de reducir la velocidad de convergencia.
- Supongamos que se quiere encontrar la raíz $x^* = 0$ de la función $f(x) = \arctan(x)$ y se parte de $x^0 = 10$. Es fácil ver que el método de Newton diverge. La razón de esto es $f' = (1 + x^2)^{-1}$ es pequeña cuando x es grande, mientras que $f(x)$ se mantiene en valor cercano a $\pm\pi/2$. Esto hace que las iteraciones de Newton diverjan.
- Para solucionar este problema, se interpreta $-f(x)/f'(x)$ com una dirección de búsqueda y se introduce un factor que va reduciendo la magnitud de las iteraciones hasta conseguir que el método converja.

Algoritmo

1. seleccionar: `tol`, x (inicial).
2. mientras $|f(x)| > \texttt{tol}$, hacer:
   1. si $f'(x) = 0$ parar con un error.
   2. $s = -f(x)/f'(x)$ (dirección de búsqueda)
   3. $x_p = x + s$
   4. Si $|f(x_p)| < |f(x)|$ entonces $x = x_p$, (se acepta el punto).
      Si no
      $s = s/2$ e ir al paso 2(2).

Métodos con dirección de búsqueda

Esta estrategia se puede generalizar de distintos modos. En general, se calcula una dirección de búsqueda, como la de Newton

$$d = -\frac{f(x_c)}{f'(x_c)} \ ,$$

y las distintas iteraciones se hacen mediante pasos de la forma $s = \lambda d$ con $\lambda = 2^{-j}$ y $j \geq 0$, hasta que $f(x_c + s)$ satiface

$$\left| f\left(x_c + \lambda d\right)\right| < (1 - \alpha\lambda) \left| f\left(x_c\right)\right| \ ,$$

donde α es un número pequeño que, generalmente, se toma com $\alpha = 10^{-4}$. Este tipo de estrategia de conoce como la Regla de Armijo.

- Supongamos que se pretende usar el método de Newton para resolver un problema no lineal de gran dimensión de la forma

$$\vec{F}\left(\vec{x}\right) = \vec{0} \ .$$

- Ya hemos visto que cundo se tratan problemas de gran dimensión pueden aparecer problemas de almacenamiento de las matrices y que el coste computacional crece mucho cuando se utilizan métodos directos para la resolución de sistemas de ecuaciones.

Métodos de Newton inexactos

- El método de Newton se basa en la solución de problemas lineales de la forma
$$DF\left(\vec{x}_i\right)\Delta\vec{x} = -\vec{F}\left(\vec{x}_i\right) \ . \tag{1}$$
Supongamos que para resolver el sistem (1) se utiliza un método iterativo, entonces el método resultante se conoce como un método de Newton inexacto.
- Entre los métodos iterativos posibles es ineteresante considerar los métodos iterativos basados en los subespacios de Krylov. Estos métodos se basan en productos matriz-vector.

- Para el problema (1), dado un vector $\vec{v}$, interesa ver cómo se podría calcular el producto $DF\left(\vec{x}_i\right)\vec{v}$. Si tenemos en cuenta que

$$\vec{F}\left(\vec{x}_i + \varepsilon\vec{v}\right) \approx \vec{F}\left(\vec{x}_i\right) + \varepsilon DF\left(\vec{x}_i\right)\vec{v} + o\left(\varepsilon^2\right) \ ,$$

tenemos que

$$DF\left(\vec{x}_i\right)\vec{v} \approx \frac{\vec{F}\left(\vec{x}_i + \varepsilon\vec{v}\right) - \vec{F}\left(\vec{x}_i\right)}{\varepsilon} \ . \tag{2}$$

Con esta aproximación podemos calcular el producto matriz-vector mediante una evaluación de la función no lineal $\vec{F}$ y sin necesidad de almacenar su matriz jacobiana. Utilizando esta estrategia obtenemos un método de Newton-Krylov sin matriz jacobiana.
Para que la aproximación (2) funcione bien, hace falta tomar un valor adecuado de ε. Algunos autores proponen elegir ε como un número algo mayor que la raíz cuadrada del épsilon de la máquina.