Clase 6 - Solución de ecuaciones no lineales (3)

Método de la secante

Continuamos con el mismo problema de resolver una ecuación no lineal. Hasta ahora vimos el método de bisección y el método de Newton. Recordemos que la iteración del método de Newton está dada por:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \text{para } n \geq 0. \tag{1}$$

Si bien este método tiene convergencia cuadrática local, tiene como desventaja que requiere la evaluación de la derivada de la función f en cada iteración. Uno de los métodos más conocidos que evita esto es el **método de la secante.**

La idea del método de la secante consiste en reemplazar $f'(x_n)$ en la iteración de Newton (1) por una aproximación dada por el cociente incremental, dado por la pendiente de la recta secante que pasa por los puntos $(x_n, f(x_n))$ y $(x_n + h, f(x_n + h))$:

$$f'(x_n) \approx a_n = \frac{f(x_n + h) - f(x_n)}{h},$$

para algún h suficientemente pequeño. Ver Figura 1.

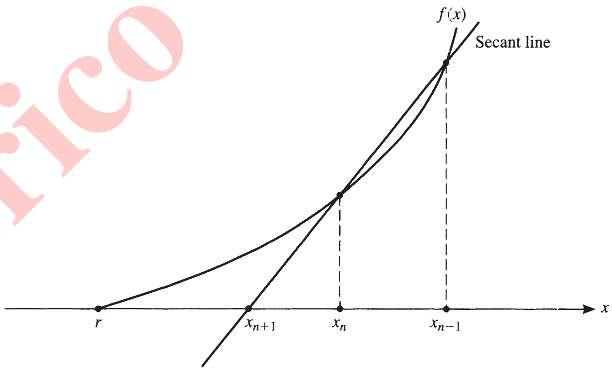

Figura 1: Interpretación gráfica del método secante.

Para evitar evaluar la función f en un punto adicional $(x_n + h)$ se elige $h = x_{n-1} - x_n$, entonces:

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

Así, la iteración del método secante consiste en:

$$x_{n+1} = x_n - \frac{f(x_n)}{\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}} \qquad \text{para } n \geq 1,$$

es decir,

$$x_{n+1} = x_n - f(x_n)\left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\right] \qquad \text{para } n \geq 1. \tag{2}$$

Algoritmo del método de la secante

Dados los siguientes datos de entrada y parámetros algorítmicos: a la penúltima aproximación de r, b la última aproximación de r, el número máximo de iteraciones permitido M, δ la tolerancia para el error e (en la variable x) y ε la tolerancia para los valores funcionales.

input $a, b, M, \delta, \varepsilon$

$fa \leftarrow f(a)$

$fb \leftarrow f(b)$

output $0, a, fa$

output $1, b, fb$

for $k = 2, 3, \ldots, M$ **do**

    if $|fa| > |fb|$ **do**

        $a \leftrightarrow b; fa \leftrightarrow fb$

    endif

    $s \leftarrow (b-a)/(fb-fa)$

    $b \leftarrow a$

    $fb \leftarrow fa$

    $a \leftarrow a - fa * s$

    $fa \leftarrow f(a)$

    output k, a, fa

    if $|b-a| < \delta$ **or** $|fa| < \varepsilon$ **then** **STOP**

end do

Observaciones:

1. En el algoritmo los puntos a y b pueden intercambiarse para lograr que $|f(a)| \leq |f(b)|$. Así, para el par $\{x_n, x_{n-1}\}$ se satiface que $|f(x_n)| \leq |f(x_{n-1})|$, y para el par siguiente $\{x_{n+1}, x_n\}$ se tiene que $|f(x_{n+1})| \leq |f(x_n)|$. Esto garantiza que la sucesión $\{|f(x_n)\}$ es no creciente.

2. El algoritmo se detiene por el número máximo de iteraciones permitidas, por satisfacer la tolarancia para los valores funcionales, o por la tolerancia en la diferencia de dos aproximaciones sucesivas.

3. En cuanto al análisis de errores, es posible probar que:

$$e_{n+1} \approx c e_n^{\alpha},$$

donde $\alpha = (1+\sqrt{5})/2 = 1.618334\ldots$ Como $1 < \alpha < 2$ se dice que el método de la secante tiene convergencia superlineal. Además, por recurrencia

$$e_{n+1} \approx c e_n^{\alpha} \approx c^{1+\alpha} e_{n-1}^{\alpha^2}$$

donde $\alpha^2 = (1+\sqrt{5})^2/4 = (3+\sqrt{5})/2 = 2.618334\ldots$, esto dice que dos iteraciones de método de la secante es mejor que una iteración del método de la secante.

Iteración de punto fijo

En esta sección veremos que es un punto fijo de una función dada, como encontrarlos numéricamente y cual es la conexión con el problema de determinar raíces de funciones.

Definición 1. *un punto fijo de una función g es un número p, en el dominio de g, tal que $g(p) = p$.*

Por un lado, si p es una raíz de una función f, esto es, $f(p) = 0$, entonces es posible definir diferentes funciones g con un punto fijo en p, por ejemplo: $g(x) = x - f(x)$, o $g(x) = x + 3f(x)$.

Por otro lado, si g tiene un punto fijo en p, esto es, $g(p) = p$, entonces la función $f(x) = x - g(x)$ tiene una raíz en p.

Aunque estamos interesados en el problema de determinar soluciones de una ecuación no lineal, o equivalentemente, encontrar raíces de funciones no lineales veremos que la forma de punto es muy fácil de estudiar y analizar. Además algunas opciones de punto fijo dan origen a técnicas matemáticas y computacionales muy poderosas para determinar raíces.

Ejemplo 1: se busca determinar los posibles puntos fijos de la función $g(x) = x^2 - 2$ en el intervalo $[-2, 3]$.

Para esto se plantea la ecuación $g(x) = x$, es decir, se debe resolver la ecuación cuadrática $x^2 - 2 = x$, o sea, $x^2 - x - 2 = 0$. Luego es fácil verificar que $x = -1$ y $x = 2$ son puntos fijos $(g(p) = p)$ pues

$$g(-1) = (-1)^2 - 2 = -1 \qquad \text{y} \qquad g(2) = (2)^2 - 2 = 2.$$

Ver Figura 2.

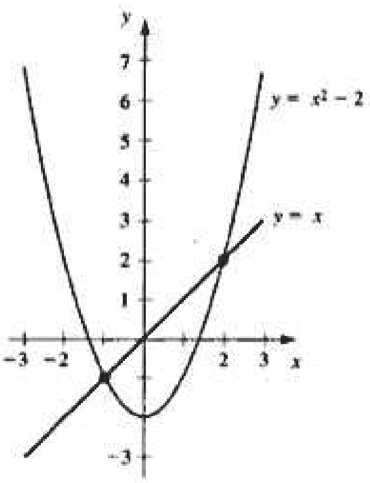

Figura 2: Puntos fijos de g.

A continuación veremos un resultado que da condiciones suficientes para la existencia y unicidad del punto fijo.

Teorema 1.

1. *Si $g \in C[a,b]$ (es decir, g es una función continua en el intervalo $[a,b]$) y $g(x) \in [a,b]$ para todo $x \in [a,b]$ entonces existe $p \in [a,b]$ tal que $g(p) = p$.* **(EXISTENCIA)**

2. *Si además existe $g'(x)$ para todo $x \in (a,b)$ y existe una constante positiva $k < 1$ tal que $|g'(x)| \leq k$ para todo $x \in (a,b)$, entonces el punto fijo en (a,b) es único. (Ver Figura 3).* **(UNICIDAD)**

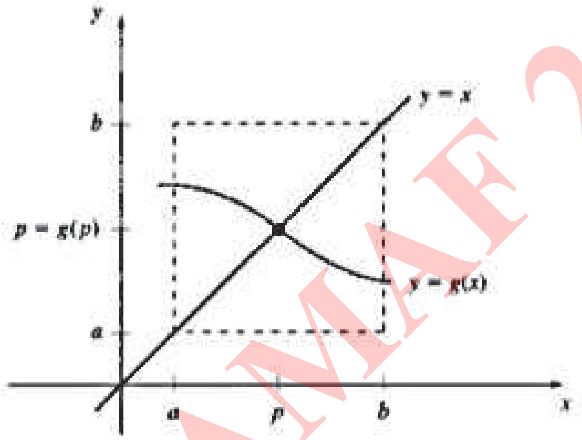

Figura 3: Unicidad del punto fijo.

Demostración.

1. Si $g(a) = a$ o $g(b) = b$, el punto fijo está en uno de los extremos del intervalo y ya estaría probado.

   Supongamos que esto no es cierto, entonces $g(a) > a$ y $g(b) < b$. Sea $h(x) = g(x) - x$ una función definida en $[a,b]$, que además es continua en $[a,b]$ pues g y x lo son y resta de funciones continuas es una función continua. Además, por lo anterior, tenemos que

   $$h(a) = g(a) - a > 0 \qquad \text{y} \qquad h(b) = g(b) - b < 0.$$

   Entonces, por el Teorema del Valor Intermedio, existe $p \in (a,b)$ tal que $h(p) = 0$, esto es, $g(p) = p$ y por lo tanto p es un punto fijo de g.

2. Ahora supongamos que existen dos puntos fijos distintos p y q en $[a,b]$, es decir, $p, q \in [a,b]$, $p \neq q$ tal que $g(p) = p$ y $g(q) = q$. Por el Teorema del Valor Medio existe ξ entre p y q, y por lo tanto en $[a,b]$ tal que

   $$g(p) - g(q) = g'(\xi)(p - q).$$

   Luego usando la hipótesis que $|g'(x)| \leq k < 1$ tenemos que

   $$|p - q| = |g(p) - g(q)| = |g'(\xi)||p - q| \leq k|p - q| < |p - q|,$$

   lo que es una contradicción que provino de suponer que habrían dos puntos fijos distintos en $[a,b]$, y por lo tanto el punto fijo es único.

□

Analicemos la existencia y unicidad de punto fijo en los siguientes ejemplos.

Ejemplo 1: considerar $g(x) = \dfrac{(x^2-1)}{3} = \dfrac{x^2}{3} - \dfrac{1}{3}$ en el intervalo $[-1,1]$.

Es fácil ver g tiene un mínimo absoluto en $x = 0$ y $g(0) = -1/3$. Además tiene máximos absolutos en $x = -1$ y $x = 1$ donde $g(-1) = 0$ y $g(1) = 0$. Además, claramente g es continua en $[-1,1]$ y $|g'(x)| = \left|\dfrac{2}{3}x\right| \leq \dfrac{2}{3}$ para todo $x \in [-1,1]$. Por lo tanto existe un único punto fijo p de g en el intervalo $[-1,1]$. Para determinar p, planteamos

$$g(p) = p, \quad \Rightarrow \quad \frac{p^2-1}{3} = p, \quad \Rightarrow \quad p^2 - 3p - 1 = 0,$$

cuyas raíces son $p_1 = \dfrac{(3-\sqrt{13})}{2} = -0{,}302776...$ y $p_2 = \dfrac{(3\sqrt{13})}{2} = 3{,}302776...$ El único punto fijo es claramente p_1 pues este punto está en $[-1,1]$ en cambio p_2 no. Ver Figura 4.

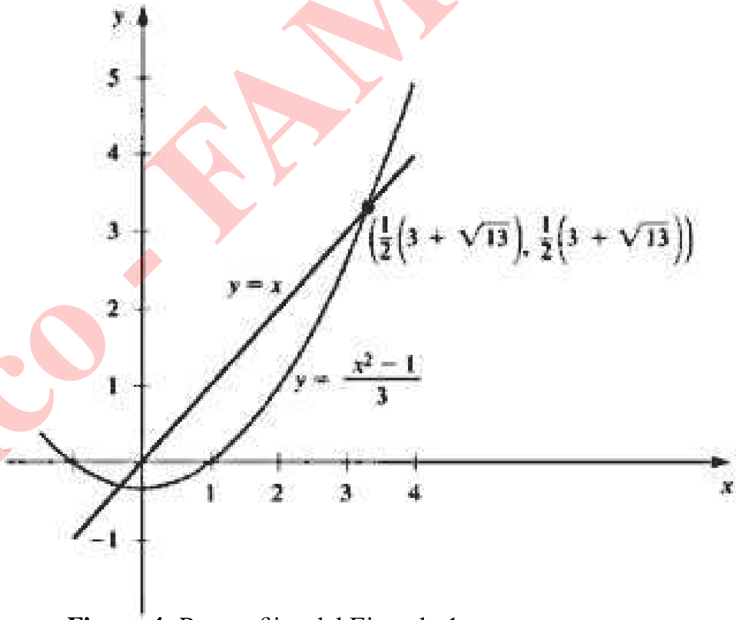

Figura 4: Puntos fijos del Ejemplo 1.

Notar en el gráfico que p_2 también es un punto fijo de g en el intervalo $[3,4]$. Sin embargo, $g(4) = 5$ y $g'(4) = \dfrac{8}{3} > 1$, por lo que no se satisfacen las hipótesis del teorema anterior en el intervalo $[3,4]$. Esto dice que tales hipótesis son condiciones suficientes para garantizar la existencia y unicidad de un punto fijo, pero no son necesarias.

Ejemplo 2: considerar $g(x) = 3^{-x} = e^{-(\ln 3)x}$ en el intervalo $[0,1]$.

Su derivada es $g'(x) = -(\ln 3)e^{-(\ln 3)x} = -(\ln 3)3^{-x} < 0$, por lo tanto g es decreciente en el intervalo $[0,1]$. Entonces,

$$g(1) = \frac{1}{3} \leq g(x) \leq 1 = g(0),$$

y por el teorema anterior, existe un punto fijo en $[0,1]$. Por otro lado, $g'(0) = -\ln 3 = -1.0986...$ y por lo tanto no se puede usar el teorema anterior para garantizar unicidad pues $|g'(x)| \not< 1$ en el intervalo $(0,1)$. Ver Figura 5.

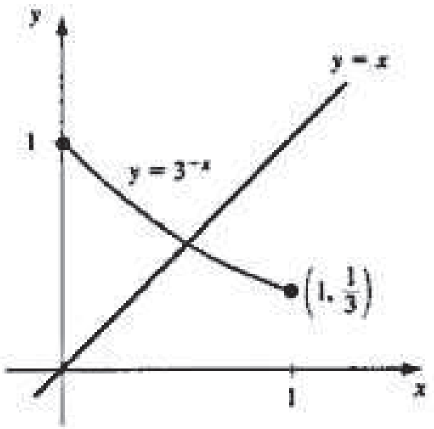

Figura 5: Puntos fijos del Ejemplo 1.

Idea del algoritmo de punto fijo

Para calcular aproximadamente el punto fijo de una función g primero se inicia con una aproximación inicial p_0 y calulando $p_n = g(p_{n-1})$ para $n \geq 1$ se obtiene una sucesión de aproximaciones $\{p_n\}$. Si la función g es continua y la sucesión converge entonces lo hace a un punto fijo p de g pues

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} g(p_{n-1}) = g(\lim_{n\to\infty} p_{n-1}) = g(p).$$

Algoritmo del método de punto fijo

Dados los siguientes datos de entrada y parámetros algorítmicos: p_0 una aproximación inicial, el número máximo de iteraciones permitido M y δ la tolerancia para el error e (en la variable x)

input p_0, M, δ

output $0, p_0$

$i \leftarrow 1$

while $i \leq M$ **then do**

$p \leftarrow g(p_0)$

output i, p

if $|p - p_0| < \delta$ **then** **STOP**

$i \leftarrow i + 1$

$p_0 \leftarrow p$

end do

Observaciones:

1. El algoritmo es muy fácil de implementar.

2. los criterios de parada utilizados son la distancia entre dos iteraciones sucesivas y el número máximo de iteraciones.

Teorema 2.

Sea $g \in C[a,b]$ tal que $g(x) \in [a,b]$ para todo $x \in [a,b]$. Supongamos que existe $g'(x)$ para todo $x \in (a,b)$ y existe una constante positiva $0 < k < 1$ tal que $|g'(x)| \leq k$ para todo $x \in (a,b)$, entonces para cualquier $p_0 \in [a,b]$ la sucesión definida por $p_n = g(p_{n-1})$, para $n \geq 1$, converge al único punto fijo p en (a,b).

Demostración. Por el teorema anterior, se sabe que bajo estas hipótesis existe un único punto fijo $p \in [a,b]$. Como $g(x) \in [a,b]$ para todo $x \in [a,b]$, la sucesión de aproximaciones $\{p_n\}$ está bien definida para todo n, es decir, $p_n \in [a,b]$ para todo n.

Para probar la convergencia se usará el Teorema del Valor Medio en lo siguiente

$$|p_n - p| = |g(p_{n-1}) - g(p)| = |g'(\xi_n)||p_{n-1} - p| \leq k|p_{n-1} - p|,$$

luego, por recurrencia, se tiene que

$$|p_n - p| \leq k|p_{n-1} - p| \leq k^2|p_{n-2} - p| \leq \cdots \leq k^n|p_0 - p|.$$

Como $0 < k < 1$ entonces $\lim\limits_{n\to\infty} k^n = 0$, luego

$$\lim_{n\to\infty} |p_n - p| \leq \lim_{n\to\infty} k^n|p_0 - p| = |p_0 - p| \lim_{n\to\infty} k^n = 0,$$

y por lo tanto, la sucesión $\{p_n\}$ converge al punto fijo p. □

Corolario 1. *Si g es una función que satisface las hipótesis del teorema anterior, se tienen las siguientes cotas de error*

$$\begin{aligned} |p_n - p| &\leq k^n \max\{p_0 - a, b - p_0\} \\ |p_n - p| &\leq \tfrac{k^n}{1-k}|p_1 - p_0| \qquad \textit{para todo} \quad n \geq 1 \end{aligned}$$

Demostración. La demostración puede consultarse en el libro de Burden-Faires. □

Análisis de error en métodos de punto fijo

Supongamos que p es un punto fijo de una función g y que $\{p_n\}$ es la sucesión, que converge a p, definida por $p_{n+1} = g(p_n)$.

Sea p_n una aproximación del punto fijo p, es decir $p_n = p + h$, y consideremos el desarrollo de Taylor de g centrado en p:

$$\begin{aligned} p_{n+1} = g(p_n) = g(p+h) &= g(p) + g'(p)(p_n - p) + \tfrac{g''(p)}{2}(p_n - p)^2 + \ldots \\ &\quad + \tfrac{g^{(r-1)}(p)}{(r-1)!}(p_n - p)^{r-1} + \tfrac{g^{(r)}(\xi)}{r!}(p_n - p)^r, \end{aligned} \tag{3}$$

para algún ξ entre p_n y p.

Supongamos ahora que $g'(p) = g''(p) = \cdots = g^{(r-1)}(p) = 0$ pero $g^{(r)}(p) \neq 0$, entonces

$$e_{n+1} = p_{n+1} - p = g(p_n) - g(p) = \frac{g^{(r)}(\xi)}{r!}(p_n - p)^r = \frac{g^{(r)}(\xi)}{r!} e_n^r,$$

esto es,

$$e_{n+1} = \frac{g^{(r)}(\xi)}{r!} e_n^r,$$

y tomando límite se obtiene

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^r} = \frac{g^{(r)}(p)}{r!} = C,$$

por lo tanto el método tiene orden de convergencia (al menos) r.

Conclusión: si las derivadas de la función de iteración de punto fijo se anulan en el punto fijo p hasta el orden $(r-1)$ entonces el método tiene orden de convergencia (al menos) r.

Usando este resultado se obtienen tres corolarios interesantes que relacionan el método de Newton con el método de punto fijo para una función f que satisface las hipótesis del teorema de convergencia de Newton.

Corolario 2. *Si f es una función que tiene una raíz simple p, entonces el método de Newton es un método de punto fijo y tiene orden de convergencia (al menos 2).*

Demostración. Sea $g(x) = x - \dfrac{f(x)}{f'(x)}$, la función de iteración del método de Newton. Es claro que si p es una solución de $f(x) = 0$, entonces p es un punto fijo de g pues

$$g(p) = p - \frac{f(p)}{f'(p)} = p,$$

ya que $f(p) = 0$.

Ahora calculemos la derivada de g en p:

$$g'(x) = 1 - \frac{(f'(x))^2 - f''(x)f(x)}{(f'(x))^2} = 1 - 1 + \frac{f''(x)f(x)}{(f'(x))^2} = \frac{f''(x)f(x)}{(f'(x))^2},$$

entonces

$$g'(p) = \frac{f''(p)f(p)}{(f'(p))^2} = 0,$$

y por lo tanto el método tiene orden de convergencia (al menos) 2. □

Corolario 3. *Si p es una raíz de multiplicidad $r \geq 2$ de f, entonces el método de Newton tiene orden 1.*

Demostración. Ya vimos que si $g(x) = x - \frac{f(x)}{f'(x)}$ entonces $g'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$.

Ahora, supongamos que p es una raíz de multiplicidad r de f, esto es, $f(x) = (x-p)^r h(x)$, con h una función tal que $h(p) \neq 0$ y $r \geq 2$.

La derivada primera de f es

$$f'(x) = r(x-p)^{r-1}h(x) + (x-p)^r h'(x) = (x-p)^{r-1}\left[rh(x) + (x-p)h'(x)\right],$$

y la derivada segunda de f es

$$\begin{aligned} f''(x) &= r(r-1)(x-p)^{r-2}h(x) + 2r(x-p)^{r-1}h'(x) + (x-p)^r h''(x), \\ &= (x-p)^{r-2}\left[r(r-1)h(x) + 2r(x-p)h'(x) + (x-p)^2 h''(x)\right]. \end{aligned}$$

Luego

$$g'(x) = \frac{(x-p)^r h(x)(x-p)^{r-2}\left[r(r-1)h(x) + 2r(x-p)h'(x) + (x-p)^2 h''(x)\right]}{(x-p)^{2r-2}\left[rh(x) + (x-p)h'(x)\right]^2},$$

entonces

$$g'(p) = \frac{h(p)r(r-1)h(p)}{r^2(h(p))^2} = \frac{r-1}{r} \neq 0,$$

pues $r \geq 2$. $\square$

Por último, modificando levemente la función de iteración del método de Newton se puede recuperar la convergencia cuadrática aún en casos de raíces múltiples.

Corolario 4. *Si p es una raíz de multiplicidad $r \geq 2$ de f, entonces la siguiente modificación del método de Newton recupera la convergencia cuadrática:*

$$x_{n+1} = x_n - r\frac{f(x_n)}{f'(x_n)}, \qquad \textit{esto es,} \qquad g(x) = x - r\frac{f(x)}{f'(x)}.$$

Demostración. Usando las expresiones de f, f' y f'' obtenidas en el corolario anterior, se tiene que

$$\begin{aligned} g'(x) &= 1 - r\tfrac{(f'(x))^2 - f''(x)f(x)}{(f'(x))^2} = 1 - r + r\tfrac{f(x)f''(x)}{(f'(x))^2} \\ &= 1 - r + r\tfrac{(x-p)^r h(x)(x-p)^{r-2}\left[r(r-1)h(x) + 2r(x-p)h'(x) + (x-p)^2 h''(x)\right]}{(x-p)^{2r-2}\left[rh(x) + (x-p)h'(x)\right]^2} \\ &= 1 - r + r\tfrac{h(x)\left[r(r-1)h(x) + 2r(x-p)h'(x) + (x-p)^2 h''(x)\right]}{\left[rh(x) + (x-p)h'(x)\right]^2}. \end{aligned}$$

Evaluando en el punto fijo p se obtiene

$$g'(p) = 1 - r + r\frac{h(p)r(r-1)h(p)}{r^2(h(p))^2} = 1 - r + (r-1) = 0,$$

y por lo tanto el método de Newton modificado tiene convergencia (al menos) cuadrática. $\square$