

Índice

3. Resolución de ecuaciones no lineales	3
3.1. Método de la bisección	5
3.2. Método de punto fijo o de iteración funcional	7
3.2.1. Teorema de existencia y unicidad de punto fijo	8
3.2.2. Iteración de punto fijo	9
3.2.3. Convergencia de la iteración de punto fijo	9
3.2.4. Método de Newton-Raphson	10
3.2.5. Convergencia del método de Newton-Raphson	11
3.2.6. Método de la secante	12
3.3. Métodos para encontrar ceros de polinomios	18
3.3.1. Método de Horner	18
3.3.2. Método de deflación	19
3.4. Método de Müller	19

Capítulo 3

Resolución de ecuaciones no lineales

Vamos a considerar ahora el problema de encontrar las *raíces de una ecuación* o, lo que es lo mismo, el *cero de una función*.

Decimos que α_r es un **cero de la función** $\varphi(x)$ si

$$\varphi(\alpha_r) = 0. \quad (3.1)$$

Por otra parte, decimos que α_r es una **raíz de la ecuación**

$$\varphi(x) = 0, \quad (3.2)$$

si se cumple que

$$\varphi(\alpha_r) = 0, \quad (3.3)$$

es decir, si verifica la ecuación.

(Espacio reservado para discutir el caso $\varphi(x) = e^{-x} - x$.)

Una forma de resolver este tipo de problemas es a través de métodos iterativos que permiten obtener una aproximación a la raíz buscada. En la práctica, se trata de generar una secuencia x_0, x_1, x_2, \dots , con x_0 propuesto arbitrariamente, de manera tal de que se cumpla que

$$\lim_{n \rightarrow \infty} x_n = \alpha_r, \quad (3.4)$$

Es decir, se busca generar una secuencia que converja a la raíz buscada (en teoría, lo hace en un número infinito de pasos). Obviamente, las secuencias que generemos deberán ser finitas, lo que nos obliga a establecer ciertos criterios que determinen si uno está “suficientemente cerca” de la raíz, como para frenar el proceso iterativo. Estos criterios pueden ser:

1. la distancia **absoluta** entre números consecutivos

$$|x_n - x_{n-1}| < \Delta \quad (3.5)$$

2. la distancia **relativa** entre números consecutivos

$$\frac{|x_n - x_{n-1}|}{|x_n|} < \Delta \quad (3.6)$$

3. módulo de la función cuyo cero se busca

$$|\varphi(x_n)| < \Delta \quad (3.7)$$

donde Δ es una tolerancia prefijada.

El primer criterio tiene el inconveniente que existen secuencias divergentes para las cuales la distancia entre elementos tiende a cero. Veremos un ejemplo en clase. El tercer criterio puede fallar también, ya que existen funciones cuyo módulo se hace muy pequeño “lejos” del cero. En general, el segundo criterio es el más seguro (siempre que x_n no se anule), o una combinación del segundo y el tercero.

Si el procedimiento iterativo tiene éxito, iremos obteniendo valores que convergerán a la solución del problema. Para determinar a qué ritmo nos acercamos a la solución es que se define lo siguiente.

Velocidad de convergencia

Definición. Sea $\{a_n\}_{n=1}^{\infty}$ una sucesión que converge a un número α_r y sea $\{b_n\}_{n=1}^{\infty}$ una sucesión que converge a cero, entonces, si existe una constante K tal que

$$|a_n - \alpha_r| \leq K|b_n|, \quad (3.8)$$

a partir de cierto n , decimos que la secuencia $\{a_n\}_{n=1}^{\infty}$ converge a α_r con **velocidad de convergencia** $O(b_n)$, y lo denotamos así

$$a_n = \alpha_r + O(b_n). \quad (3.9)$$

Esta definición nos permite tipificar la velocidad con la que la sucesión $\{a_n\}_{n=1}^{\infty}$ se acerca a α_r utilizando otra sucesión (sencilla) que nos sirve para acotar la diferencia. En general, se utilizan sucesiones de comparación de la forma

$$b_n = \frac{1}{n^p}, \quad (3.10)$$

con $p > 0$.

Ejemplo 1

$$a_n = \frac{n+1}{n^2}, \quad n \geq 1. \quad (3.11)$$

Buscamos la sucesión $\{b_n\}_{n=1}^{\infty}$ para este caso.

$$\left| \frac{n+1}{n^2} - 0 \right| \leq \frac{n+n}{n^2} = \frac{2}{n} = 2 \cdot \frac{1}{n}, \quad n \geq 1. \quad (3.12)$$

Por lo tanto, puedo concluir que

$$a_n = 0 + O\left(\frac{1}{n}\right) \quad (3.13)$$

es decir, que $\{a_n\}$ converge a cero al menos con la misma velocidad con que lo hace $1/n$, a partir de $n = 1$.

Ejemplo 2

$$a_n = \frac{n+3}{n^3}, \quad n \geq 1. \tag{3.14}$$

Buscamos la sucesión $\{b_n\}_{n=1}^\infty$

$$\left| \frac{n+3}{n^3} - 0 \right| \leq \frac{2n}{n^3} = \frac{2}{n^2} = 2 \cdot \frac{1}{n^2}, \quad n \geq 3. \tag{3.15}$$

Por lo tanto,

$$a_n = 0 + O\left(\frac{1}{n^2}\right). \tag{3.16}$$

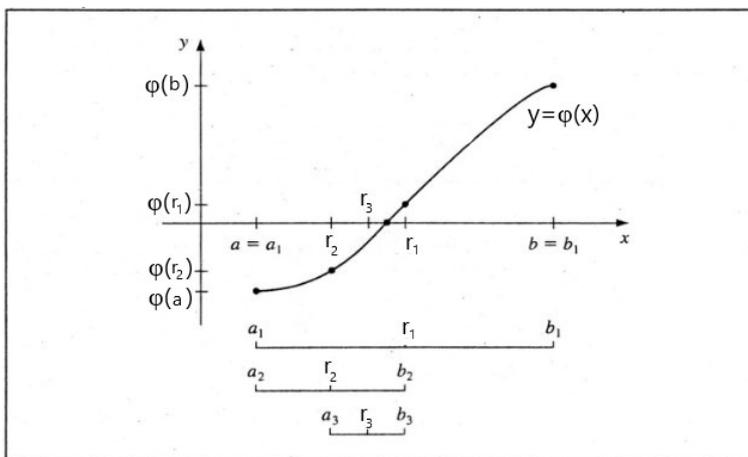
* * *

Volviendo al tema de encontrar los ceros de una función, si bien es posible utilizar métodos para funciones con múltiples ceros, vamos a suponer que hemos aislado el cero que nos interesa, es decir, que hemos determinado un intervalo que lo contiene y donde no hay otro. Comenzaremos explicando un método muy sencillo para hallar dicho cero.

3.1. Método de la bisección

Sea $\varphi(x)$ continua en el intervalo $[a, b]$ y sea $\varphi(a) \cdot \varphi(b) < 0$. Con estas hipótesis podemos asegurar que la función $\varphi(x)$ tiene un cero en $[a, b]$ (por el teorema de Bolzano¹).

El método de la bisección consiste en dividir sucesivamente por la mitad al intervalo dado, tomando como aproximación de la raíz el punto medio del intervalo original.



Sean $a_1 \equiv a$ y $b_1 \equiv b$, definimos la primera aproximación al cero α_r de $\varphi(x)$ como

$$r_1 = \frac{a+b}{2}. \tag{3.17}$$

Evaluamos $\varphi(r_1)$. Si $\varphi(r_1) = 0$, r_1 es el cero buscado. Sin embargo, resulta muy improbable que hallemos así el cero. Lo más normal será que

$$\varphi(r_1) \neq 0, \tag{3.18}$$

¹Teorema de Bolzano. Si una función $f(x)$ está definida y es continua en un intervalo cerrado $[a, b]$, y toma valores de distinto signo en los extremos a y b , entonces existe al menos un punto c del intervalo abierto (a, b) en el que se anula la función.

Consideraciones numéricas

- Cuando se implementa este método en la computadora, para calcular la raíz conviene hacer

$$r_n = a_n + \frac{(b_n - a_n)}{2}, \quad (3.27)$$

en lugar de

$$r_n = \frac{b_n + a_n}{2} \quad (3.28)$$

ya que introduce menor error de redondeo.

- En vez de evaluar $\varphi(a_n) \cdot \varphi(b_n)$ conviene evaluar

$$\text{sgn}(\varphi(a_n)) \cdot \text{sgn}(\varphi(b_n)), \quad (3.29)$$

para evitar un posible *overflow* o *underflow*.

- El método de la bisección es **lento, pero converge siempre** a la solución. En los paquetes disponibles se lo suele utilizar para generar una mejor aproximación inicial para métodos más veloces.

3.2. Método de punto fijo o de iteración funcional

Para enunciar este método, es necesario primero definir lo que es el punto fijo de una función.

Definición. Un número α_r es un *punto fijo de una función* $\psi(x)$ si

$$\psi(\alpha_r) = \alpha_r. \quad (3.30)$$

Los problemas de encontrar puntos fijos y raíces de ecuaciones son equivalentes.

- Dado un problema de encontrar la raíz α_r de la ecuación

$$\varphi(x) = 0, \quad (3.31)$$

podemos definir funciones $\psi(x)$ con un punto fijo en α_r en múltiples formas, por ejemplo, haciendo

$$\psi_1(x) = x - \varphi(x), \quad (3.32)$$

o

$$\psi_2(x) = x + 3\varphi(x). \quad (3.33)$$

- Análogamente, si $\psi(x)$ tiene un punto fijo en α_r , podemos construir una función

$$\varphi(x) = x - \psi(x), \quad (3.34)$$

tal que α_r sea la raíz de la ecuación

$$\varphi(x) = 0. \quad (3.35)$$

Aunque lo que **queremos** es **buscar raíces** de ecuaciones, vamos a transformar el problema en uno de **buscar puntos fijos**, ya que existen métodos muy poderosos para lograr esto último y, como vimos, encontrar el punto fijo de $\psi(x)$ es equivalente a encontrar el cero de $\varphi(x)$.

3.2.1. Teorema de existencia y unicidad de punto fijo

(Existencia) Sea $\psi \in C[a, b]$ y sea $\psi(x) \in [a, b]$, para todo $x \in [a, b]$, entonces $\psi(x)$ tiene uno o más puntos fijos en $[a, b]$.

(Unicidad) Si, además, $\psi'(x)$ existe en (a, b) y existe $0 \leq k < 1$ tal que

$$|\psi'(x)| \leq k, \quad \forall x \in (a, b) \quad (3.36)$$

entonces el punto fijo de $\psi(x)$ en $[a, b]$ es único.

Demostración

Existencia

Si $\psi(a) = a$ o $\psi(b) = b$, entonces $\psi(x)$ tiene un punto fijo en alguno de los extremos del intervalo y la existencia queda asegurada. Pero este sería un caso fortuito. Consideremos el caso más desfavorable en el que $\psi(a) \neq a$ y $\psi(b) \neq b$. Entonces, como $\psi : [a, b] \rightarrow [a, b]$, tendríamos que

$$\psi(a) > a,$$

$$\psi(b) < b.$$

Definiendo una nueva función

$$\eta(x) = \psi(x) - x, \quad (3.37)$$

continua en $[a, b]$, se cumplirá que

$$\eta(a) = \psi(a) - a > 0, \quad (3.38)$$

$$\eta(b) = \psi(b) - b < 0. \quad (3.39)$$

Luego, por el teorema del valor intermedio (teorema de Bolzano), existirá al menos un punto α_r tal que

$$\eta(\alpha_r) = 0. \quad (3.40)$$

En consecuencia, α_r verificará la igualdad

$$\psi(\alpha_r) - \alpha_r = 0 \quad \Rightarrow \quad \boxed{\psi(\alpha_r) = \alpha_r}, \quad (3.41)$$

con lo que queda asegurada la existencia de al menos un punto fijo.

Unicidad

Vamos a demostrar la unicidad proponiendo la existencia de más de un punto fijo y hallando que eso nos conduce a un absurdo.

Sean α_r y α_q puntos fijos de $\psi(x)$ en $[a, b]$, con $\alpha_r < \alpha_q$. Como $\psi(x)$ es derivable en (a, b) , por el teorema del valor medio del cálculo diferencial², existe un valor $\beta \in (\alpha_r, \alpha_q)$ tal que

$$\frac{\psi(\alpha_r) - \psi(\alpha_q)}{\alpha_r - \alpha_q} = \psi'(\beta). \quad (3.43)$$

Por lo tanto,

$$|\psi(\alpha_r) - \psi(\alpha_q)| = |\psi'(\beta)| \cdot |\alpha_r - \alpha_q| \leq k |\alpha_r - \alpha_q| < |\alpha_r - \alpha_q|. \quad (3.44)$$

Utilizando la definición de punto fijo, obtenemos finalmente que

$$|\alpha_r - \alpha_q| < |\alpha_r - \alpha_q|, \quad (3.45)$$

lo cual es un absurdo. Concluimos que el absurdo proviene de haber propuesto la existencia de más de un punto fijo en el intervalo $[a, b]$.

²El teorema dice que dada cualquier función f continua en el intervalo $[a, b]$ y derivable en el intervalo abierto (a, b) , entonces existe al menos algún punto c en el intervalo (a, b) tal que la tangente a la curva en c es paralela a la recta secante que une los puntos $(b, f(b))$ y $(a, f(a))$. Es decir:

$$\frac{f(b) - f(a)}{b - a} = f'(c). \quad (3.42)$$

Este teorema se conoce también como el *teorema del valor medio de Lagrange*.

Interpretación gráfica

(la haremos en clase)

Ejemplo

Determinése si la función

$$\psi(x) = \frac{x^2 - 1}{3} \quad (3.46)$$

tiene un punto fijo en $[-1, 1]$ y, de existir, si es único.

3.2.2. Iteración de punto fijo

Definición. Llamamos **iteración de punto fijo** al método para aproximar al punto fijo de una función $\psi(x)$ que consiste en elegir un valor inicial x_0 y generar una secuencia $\{x_n\}_{n=0}^{\infty}$ mediante la relación de recurrencia

$$x_n = \psi(x_{n-1}), \quad n \geq 1. \quad (3.47)$$

Si la secuencia converge, obtendremos una solución del problema $x = \psi(x)$.

3.2.3. Convergencia de la iteración de punto fijo

Teorema. Sea $\psi(x) \in C[a, b]$, $\psi(x) : [a, b] \rightarrow [a, b]$. Supongamos, además, que $\psi(x)$ es derivable en (a, b) y que existe $k > 0$ tal que

$$|\psi'(x)| \leq k < 1, \quad \forall x \in (a, b). \quad (3.48)$$

Entonces, para cualquier $x_0 \in [a, b]$, la secuencia definida por

$$x_n = \psi(x_{n-1}), \quad n \geq 1 \quad (3.49)$$

converge al único punto fijo de $\psi(x)$ en $[a, b]$.

Demostración

De acuerdo con las hipótesis dadas, sabemos que el punto fijo (α_r) de $\psi(x)$ existe y es único en $[a, b]$. Como $\psi(x): [a, b] \rightarrow [a, b]$, la secuencia $\{x_n\}_{n=0}^{\infty}$ está definida para todo $n \geq 0$ y $x_n \in [a, b]$, para todo n . Podemos escribir entonces la distancia entre el punto fijo y x_n como

$$|x_n - \alpha_r| = |\psi(x_{n-1}) - \psi(\alpha_r)| = |\psi'(\xi)(x_{n-1} - \alpha_r)| = |\psi'(\xi)||x_{n-1} - \alpha_r| \leq k|x_{n-1} - \alpha_r|, \quad (3.50)$$

donde hemos hecho uso del teorema del valor medio del cálculo diferencial. En consecuencia,

$$|x_n - \alpha_r| \leq k|x_{n-1} - \alpha_r| \leq k^2|x_{n-2} - \alpha_r| \leq \dots \leq k^n|x_0 - \alpha_r|. \quad (3.51)$$

Como $k < 1$, la expresión anterior implica que

$$\lim_{n \rightarrow \infty} |x_n - \alpha_r| = 0, \quad (3.52)$$

y queda demostrado que $\{x_n\}_{n=0}^{\infty}$ converge a α_r .

Corolario (sin demostración)

Si $\psi(x)$ satisface las hipótesis del teorema anterior, se puede establecer una cota para el error que se da al aproximar α_r con x_n :

$$|x_n - \alpha_r| \leq \frac{k^n}{1-k} |x_1 - x_0|. \quad (3.53)$$

La velocidad de convergencia depende del valor de la constante k : si $k \rightarrow 1^-$, la convergencia se hace lenta, si $k \rightarrow 0^+$ es más rápida.

3.2.4. Método de Newton-Raphson

Es uno de los métodos más poderosos y conocidos. Presentaremos primero el método y luego hablaremos de sus ventajas y desventajas.

Deducción

Sea $\varphi(x) \in C^2[a, b]$ y sea x_0 una aproximación al cero α_r de $\varphi(x)$. Desarrollando $\varphi(x)$ alrededor de x_0 y evaluando en α_r tenemos

$$\varphi(\alpha_r) = \varphi(x_0) + (\alpha_r - x_0)\varphi'(x_0) + \frac{(\alpha_r - x_0)^2}{2}\varphi''(\xi), \quad (3.54)$$

donde ξ se encuentra entre x_0 y α_r . Si el aporte del término con la derivada segunda es pequeño, cosa que sucederá siempre para $|\alpha_r - x_0|$ suficientemente chico, entonces podemos despreciarlo, y tendremos que

$$\varphi(\alpha_r) \approx \varphi(x_0) + (\alpha_r - x_0)\varphi'(x_0). \quad (3.55)$$

Por hipótesis, α_r es un cero de $\varphi(x)$, por lo tanto la expresión anterior se reduce a

$$0 \approx \varphi(x_0) + (\alpha_r - x_0)\varphi'(x_0), \quad (3.56)$$

donde α_r es aún desconocido. Despejando α_r obtenemos

$$\alpha_r \approx x_0 - \frac{\varphi(x_0)}{\varphi'(x_0)}, \quad (3.57)$$

válida siempre y cuando $\varphi'(x_0) \neq 0$.

Si la aproximación que hicimos es válida, entonces la evaluación del miembro de la derecha nos permitirá obtener una mejor aproximación a α_r . Este proceso se puede repetir con la idea de seguir mejorando el α_r calculado. Así, obtenemos la expresión de recurrencia del método de Newton-Raphson

$$\boxed{x_n \approx x_{n-1} - \frac{\varphi(x_{n-1})}{\varphi'(x_{n-1})}, \quad n = 1, 2, \dots} \quad (3.58)$$

Observaciones

- El método de Newton-Raphson construye las sucesivas aproximaciones mediante tangentes a la curva.

- Un problema es que necesita que $\varphi(x_n) \neq 0$, para todo n ; si la derivada es no nula, pero muy pequeña, el método puede no converger.
- Suele considerarse una desventaja tener que calcular analíticamente la derivada $\varphi'(x)$; también el que en cada paso haya que hacer la evaluación de dos funciones (φ y φ'), en lugar de sólo una. Veremos luego métodos que son una variante de éste y que lo evitan.
- Observen que la deducción del método se apoya en que la aproximación inicial esté lo “suficientemente cerca” del cero de la función. Obviamente esto no se sabe de antemano. Lo que se hace es comenzar con alguna estimación que, si es mala, seguramente provocará que el método no converja. También puede comenzarse con algunas iteraciones del método de la bisección para luego continuar con Newton-Raphson.
- Noten que el método de Newton-Raphson es un método de iteración funcional:

$$x_n = \psi(x_{n-1}) = x_{n-1} - \frac{\varphi(x_{n-1})}{\varphi'(x_{n-1})}. \quad (3.59)$$

3.2.5. Convergencia del método de Newton-Raphson

Teorema. Sea $\varphi(x) \in C^2[a, b]$. Si $\alpha_r \in [a, b]$ es tal que $\varphi(\alpha_r) = 0$ y $\varphi'(\alpha_r) \neq 0$, entonces existe $\delta > 0$ tal que la secuencia $\{x_n\}_{n=1}^{\infty}$ generada con el método de Newton-Raphson converge a α_r para cualquier aproximación inicial $x_0 \in [\alpha_r - \delta, \alpha_r + \delta]$.

Demostración

Sea

$$\psi(x) = x - \frac{\varphi(x)}{\varphi'(x)}, \quad (3.60)$$

y sea k una constante tal que $0 \leq k < 1$. Para probar el teorema vamos a demostrar que la función $\psi(x)$ verifica las hipótesis del teorema de convergencia de la iteración de punto fijo en un intervalo como el señalado. Recuerden que dichas hipótesis implican la continuidad de la función $\psi(x)$, que $\psi : [a, b] \rightarrow [a, b]$ y que $|\psi'(x)| < 1$, para todo $x \in (a, b)$.

Comencemos por demostrar la continuidad de $\psi(x)$. Por hipótesis sabemos que $\varphi(x)$ y $\varphi'(x)$ son funciones continuas. La continuidad de $\psi(x)$ queda asegurada si $\varphi'(x)$ no se anula en el intervalo que consideremos. Dado que $\varphi'(\alpha_r) \neq 0$ (por hipótesis), podemos asegurar que existe un entorno de α_r en donde $\varphi'(x) \neq 0$. Por lo tanto, existe $\delta_1 > 0$ tal que

$$\varphi'(x) \neq 0, \quad \forall x \in [\alpha_r - \delta_1, \alpha_r + \delta_1], \quad (3.61)$$

y, en ese entorno, $\psi(x)$ estará bien definida y será continua.

Pasemos ahora a analizar a la derivada de $\psi(x)$:

$$\psi'(x) = 1 - \frac{[\varphi'(x)]^2 - \varphi(x)\varphi''(x)}{[\varphi'(x)]^2} = \frac{\varphi(x)\varphi''(x)}{[\varphi'(x)]^2} \quad (3.62)$$

Por hipótesis, el numerador es continuo y, por lo visto más arriba, el denominador no se anula si $x \in [\alpha_r - \delta_1, \alpha_r + \delta_1]$. Por lo tanto,

$$\psi(x) \in C^1[\alpha_r - \delta_1, \alpha_r + \delta_1]. \quad (3.63)$$

Evaluando $\psi'(x)$ en α_r y teniendo en cuenta que $\varphi(\alpha_r) = 0$ y que $\varphi'(\alpha_r) \neq 0$, tenemos que

$$\psi'(\alpha_r) = \frac{\varphi(\alpha_r)\varphi''(\alpha_r)}{[\varphi'(\alpha_r)]^2} = 0. \quad (3.64)$$

Por ser $\psi'(x)$ una función continua, podemos afirmar que dada $0 \leq k < 1$, existe $\delta > 0$, $\delta < \delta_1$ tal que

$$|\psi'(x)| \leq k, \quad \forall x \in [\alpha_r - \delta, \alpha_r + \delta]. \quad (3.65)$$

En palabras, estoy diciendo que si una función tiene derivada primera continua y esa derivada se anula en un punto, luego el módulo de la derivada primera está acotado en el entorno de dicho punto.

Sólo nos falta probar que $\psi(x)$ va del intervalo $[\alpha_r - \delta, \alpha_r + \delta]$ en sí mismo. Consideremos $x \in [\alpha_r - \delta, \alpha_r + \delta]$. Usando el teorema del valor medio del cálculo diferencial podemos decir que para η entre x y α_r vale que

$$\psi(x) - \psi(\alpha_r) = \psi'(\eta)(x - \alpha_r). \quad (3.66)$$

Tomando módulo

$$|\psi(x) - \psi(\alpha_r)| = |\psi'(\eta)| |x - \alpha_r|, \quad (3.67)$$

que podemos acotar

$$|\psi(x) - \psi(\alpha_r)| \leq k |x - \alpha_r| < |x - \alpha_r|. \quad (3.68)$$

En consecuencia, como $x \in [\alpha_r - \delta, \alpha_r + \delta]$ tenemos que $|x - \alpha_r| \leq \delta$, lo que implica que

$$\psi(x) \in [\alpha_r - \delta, \alpha_r + \delta], \quad \forall x \in [\alpha_r - \delta, \alpha_r + \delta]. \quad (3.69)$$

De esta manera, hemos demostrado que todas las hipótesis del teorema de convergencia de la iteración de punto fijo se cumplen. Consecuentemente, podemos afirmar que la secuencia $\{x_n\}_{n=1}^{\infty}$ generada por

$$x_n = \psi(x_{n-1}) = x_{n-1} - \frac{\varphi(x_{n-1})}{\varphi'(x_{n-1})}, \quad (3.70)$$

converge a α_r para todo $x_0 \in [\alpha_r - \delta, \alpha_r + \delta]$.

3.2.6. Método de la secante

Se comentó antes que una eventual desventaja del método de Newton-Raphson era el tener que calcular la derivada de la función $\varphi(x)$. Así, cada paso iterativo implica la evaluación numérica tanto de φ como de φ' en el punto en consideración.

El método de la secante es una variante del método de Newton-Raphson que busca evitar el cálculo de la derivada en forma analítica. Del Análisis Matemático sabemos que

$$f'(x_{n-1}) = \lim_{x \rightarrow x_{n-1}} \frac{f(x) - f(x_{n-1})}{x - x_{n-1}}. \quad (3.71)$$

En particular, si x_{n-2} está cerca de x_{n-1} , podemos aproximar

$$f'(x_{n-1}) \approx \frac{f(x_{n-2}) - f(x_{n-1})}{x_{n-2} - x_{n-1}}. \quad (3.72)$$

Reemplazando en el método de Newton-Raphson

$$\boxed{x_n = x_{n-1} - \varphi(x_{n-1}) \frac{x_{n-2} - x_{n-1}}{\varphi(x_{n-2}) - \varphi(x_{n-1})}}, \quad n \geq 2. \quad (3.73)$$

Observen que en este método se necesitan dos valores iniciales, en lugar de uno, y que, además, no se necesita calcular la expresión de la derivada de la función. También es importante señalar que, entre iteraciones, sólo es necesario hacer una única evaluación nueva de la función.

(En clase veremos la interpretación gráfica del método.)

Cuando uno desea comparar los métodos de iteración funcional suele analizar el orden de convergencia, que se define de la siguiente manera.

Definición. Sea $\{x_n\}_{n=0}^{\infty}$ una secuencia que converge a α_r , con $x_n \neq \alpha_r$ para todo n . Si existen constantes positivas γ y λ tales que

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha_r|}{|x_n - \alpha_r|^\gamma} = \lambda, \quad (3.74)$$

decimos que la secuencia $\{x_n\}_{n=0}^{\infty}$ converge a α_r con orden γ y con constante de error asintótico λ .

Observaciones

- Un método $x_n = \psi(x_{n-1})$ se dice de orden γ si la secuencia $\{x_n\}_{n=0}^{\infty}$ converge a la solución $\alpha_r = \psi(\alpha_r)$ de orden γ .
- En general, una secuencia de orden γ grande converge más rápidamente (es decir, necesita menor número de iteraciones) que una secuencia de menor orden. La constante de error asintótico afecta la velocidad de convergencia, pero no en la misma medida que el orden γ .
- Dos casos que se presentan normalmente son:

$$\begin{cases} \gamma = 1, (\lambda < 1) & \text{(convergencia lineal)} \\ \gamma = 2. & \text{(convergencia cuadrática)} \end{cases} \quad (3.75)$$

Ejemplo (convergencias lineal vs. cuadrática)

Sea $\{x_n\}_{n=0}^{\infty}$ una sucesión que converge linealmente a $\alpha_r = 0$ con $\lambda = \frac{1}{2}$

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1}|}{|x_n|} = \frac{1}{2}, \quad (3.76)$$

y sea $\{\hat{x}_n\}_{n=0}^{\infty}$ una que lo hace cuadráticamente con la misma constante de error asintótico

$$\lim_{n \rightarrow \infty} \frac{|\hat{x}_{n+1}|}{|\hat{x}_n|^2} = \frac{1}{2}. \quad (3.77)$$

Supongamos, para hacer la comparación más simple, que se cumple que

$$(a) \frac{|x_{n+1}|}{|x_n|} \approx \frac{1}{2}, \quad (3.78)$$

$$(b) \frac{|\hat{x}_{n+1}|}{|\hat{x}_n|^2} \approx \frac{1}{2}. \quad (3.79)$$

Entonces, para el caso (a) tendremos

$$\begin{aligned} |x_n - 0| &= |x_n| \\ &\approx 0,5 |x_{n-1}| \\ &\approx 0,5^2 |x_{n-2}| \\ &\approx \dots \\ &\approx \underbrace{0,5^n}_{v_1} |x_0| \end{aligned}$$

y para el caso (b)

$$\begin{aligned} |\hat{x}_n - 0| &= |\hat{x}_n| \\ &\approx 0,5 |x_{n-1}|^2 \\ &\approx 0,5 (0,5 |x_{n-2}|^2)^2 = 0,5^3 |x_{n-2}|^4 \\ &\approx 0,5^3 (0,5 |x_{n-3}|^2)^4 = 0,5^7 |x_{n-3}|^8 \\ &\approx \dots \\ &\approx \underbrace{0,5^{2^n - 1}}_{v_2} |x_0|^{2^n} \end{aligned}$$

Si $x_0 = \hat{x}_0 = 1$, podemos comparar las velocidades de convergencia v_1 y v_2

n	v_1	v_2
1	0,5	0,5
3	$1,25 \times 10^{-1}$	$7,8125 \times 10^{-3}$
7	$7,8125 \times 10^{-3}$	$5,8775 \times 10^{-39}$

Se necesitan 126 iteraciones de la convergencia lineal para llegar al coeficiente v_2 obtenido para $n = 7$.

Claramente, la convergencia cuadrática es mucho más rápida que la convergencia lineal.

* * *

Vamos a ver unos teoremas que prueban que la iteración funcional converge sólo linealmente, mientras que, en general, Newton-Raphson lo hace cuadráticamente.

Teorema. Sea $\psi(x) \in C[a, b]$, $\psi(x) : [a, b] \rightarrow [a, b]$. Si además $\psi'(x)$ es continua en (a, b) y existe $0 \leq k < 1$ tal que $|\psi'(x)| \leq k$, $\forall x \in (a, b)$, entonces, si $\psi' \neq 0$, para cualquier $x_0 \in [a, b]$ la secuencia $\{x_n\}_{n=0}^{\infty}$ obtenida de

$$x_n = \psi(x_{n-1}), \quad n \geq 1 \quad (3.80)$$

converge sólo linealmente al único punto fijo α_r de $[a, b]$.

Demostración

Las hipótesis aseguran la existencia y unicidad del punto fijo y la convergencia al mismo. Como antes, aplicamos ahora el teorema del valor medio del cálculo diferencial a $\psi(x)$

$$x_{n+1} - \alpha_r = \psi(x_n) - \psi(\alpha_r) = \psi'(\xi_n)(x_n - \alpha_r), \quad (3.81)$$

donde ξ_n está entre x_n y α_r . Ahora, como $\{x_n\}_{n=0}^{\infty}$ converge a α_r , $\{\xi_n\}$ también lo hará. Como pedimos que $\psi'(x)$ sea continua, tendremos que

$$\lim_{n \rightarrow \infty} \psi'(\xi_n) = \psi'(\alpha_r), \quad (3.82)$$

luego

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha_r|}{|x_n - \alpha_r|} = \lim_{n \rightarrow \infty} |\psi'(\xi_n)| = |\psi'(\alpha_r)|, \quad (3.83)$$

donde $|\psi'(\alpha_r)| < 1$. Por lo tanto, como se verifica que

$$\boxed{\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha_r|}{|x_n - \alpha_r|^1} = |\psi'(\alpha_r)|}, \quad (3.84)$$

podemos concluir que la iteración de punto fijo converge linealmente con constante de error asintótico $|\psi'(\alpha_r)|$.

* * *

El teorema anterior sugiere que sólo se puede lograr un orden de convergencia mayor en los métodos de iteración funcional si se levanta la condición $\psi'(x) \neq 0$. El siguiente teorema describe las condiciones adicionales que aseguran la convergencia cuadrática.

Teorema (sin demostración). Sea α_r un punto fijo de $\psi(x)$ y sea además $\psi'(\alpha_r) = 0$, $\psi''(x)$ continua tal que $|\psi''(x)| < M$ en un intervalo abierto que contenga a α_r . Entonces, existe $\delta > 0$ tal que para $x_0 \in [\alpha_r - \delta, \alpha_r + \delta]$, la secuencia definida por

$$x_n = \psi(x_{n-1}), \quad \forall n \geq 1, \quad (3.85)$$

converge **al menos** cuadráticamente a α_r . Más aún, para n grandes

$$|x_{n+1} - \alpha_r| < |x_n - \alpha_r|^2 \frac{M}{2}. \quad (3.86)$$

Observen que este teorema nos dice que para tener métodos de punto fijo con convergencia cuadrática debemos buscar funciones $\psi(x)$ con derivada primera nula en el punto fijo. Veamos cómo tendríamos que construir $\psi(x)$ para lograr esto.

* * *

Supongamos que queremos hallar el cero α_r de la función $\varphi(x)$. Vamos a transformar este problema en el (equivalente) de encontrar el punto fijo de otra función, $\psi(x)$, asociada a $\varphi(x)$ de la siguiente manera

$$\psi(x) = x - \eta(x) \varphi(x). \quad (3.87)$$

Si $\eta(x)$ es una función bien comportada en el intervalo de interés, vemos que si α_r es un cero de $\varphi(x)$ también será un punto fijo de $\psi(x)$. Según el teorema que acabamos de ver (sin demostración), para que la convergencia sea al menos cuadrática necesitamos exigir que

$$\psi'(\alpha_r) = 0. \quad (3.88)$$

Calculemos $\psi'(x)$:

$$\psi'(x) = 1 - \eta'(x) \varphi(x) - \eta(x) \varphi'(x). \quad (3.89)$$

Evaluando en $x = \alpha_r$

$$\psi'(\alpha_r) = 1 - \eta'(\alpha_r) \underbrace{\varphi(\alpha_r)}_{=0} - \eta(\alpha_r) \varphi'(\alpha_r). \quad (3.90)$$

obtenemos entonces que

$$\psi'(\alpha_r) = 1 - \eta(\alpha_r) \varphi'(\alpha_r). \quad (3.91)$$

Por lo tanto,

$$\psi'(\alpha_r) = 0 \quad \Leftrightarrow \quad \eta(\alpha_r) = \frac{1}{\varphi'(\alpha_r)}. \quad (3.92)$$

En consecuencia, si $\varphi'(x) \neq 0$ en el intervalo, podemos hacer

$$\eta(x) = \frac{1}{\varphi'(x)}, \quad (3.93)$$

y asegurarnos que cuando $\varphi(\alpha_r) = 0$ se verifique que $\psi'(\alpha_r) = 0$. Así, podemos definir nuestro iterador de punto fijo de la siguiente manera

$$x_n = \psi(x_{n-1}) = x_{n-1} - \frac{\varphi(x_{n-1})}{\varphi'(x_{n-1})} \quad (3.94)$$

que converge cuadráticamente. (Redescubrimos el método de Newton-Raphson.)

* * *

Hay casos en los que el método de Newton-Raphson converge, pero sólo linealmente. Veamos cuáles son y cómo podemos recuperar la velocidad de convergencia cuadrática.

Cero de multiplicidad m

Definición. Decimos que α_r es un cero de multiplicidad m de $\varphi(x)$ si para $x \neq \alpha_r$ podemos escribir

$$\varphi(x) = (x - \alpha_r)^m h(x), \quad (3.95)$$

con

$$\lim_{x \rightarrow \alpha_r} h(x) \neq 0. \quad (3.96)$$

Teorema (sin demostración). Una función $\varphi(x) \in C^m[a, b]$ tiene un cero de multiplicidad m en (a, b) sí y sólo sí

$$0 = \varphi(\alpha_r) = \varphi'(\alpha_r) = \varphi''(\alpha_r) = \dots = \varphi^{m-1}(\alpha_r), \quad (3.97)$$

y

$$\varphi^m(\alpha_r) \neq 0. \quad (3.98)$$

Este teorema provee de un método sencillo para encontrar la multiplicidad de un cero.

¿Cómo afecta este hecho al método de Newton-Raphson? Recordemos que cuando se demostró la convergencia del método de N-R calculamos $\psi'(x)$ y la evaluamos en α_r :

$$\psi'(\alpha_r) = \frac{\varphi(\alpha_r)\varphi''(\alpha_r)}{[\varphi'(\alpha_r)]^2} = 0, \quad (3.99)$$

donde se pedía que $\varphi'(\alpha_r) \neq 0$. De acuerdo con el teorema anterior, esto no se va a dar si la multiplicidad del cero es mayor que uno. ¿Qué pasa con el método en este caso? Si $\varphi'(\alpha_r) = 0$, es decir, si la multiplicidad del cero de $\varphi(x)$ es mayor que uno, el método de Newton-Raphson converge, pero lo hace más lentamente.

Ejemplo

Sea

$$\varphi(x) = e^x - x - 1. \quad (3.100)$$

Encuéntrese el cero de $\varphi(x)$ usando el método de Newton-Raphson, partiendo de $x_0 = 1$.

La función tiene un cero de multiplicidad 2 en $x = 0$, como lo demuestra lo siguiente

$$\varphi(0) = 0, \quad \varphi'(0) = 0, \quad \varphi''(0) = 1 \neq 0. \quad (3.101)$$

Aplicando el método de Newton-Raphson obtenemos los siguientes valores

n	x_n
0	1
5	0,0438
10	$1,3881 \times 10^{-3}$
15	$4,2610 \times 10^{-5}$

que claramente muestra que la convergencia no es cuadrática. Para recuperar la convergencia cuadrática, se aplica el siguiente teorema.

Teorema. Si α_r es un cero de multiplicidad m de $\varphi(x)$, con $\varphi^{(m)}(x)$ continua en un intervalo abierto que incluya a α_r , entonces

$$\psi(x) = x - m \frac{\varphi(x)}{\varphi'(x)}, \quad (3.102)$$

es tal que $\psi'(\alpha_r) = 0$.

Demostración

Consideremos la función propuesta

$$\psi(x) = x - m \frac{\varphi(x)}{\varphi'(x)}. \quad (3.103)$$

Si α_r es un cero de multiplicidad m de $\varphi(x)$, por definición podemos escribir

$$\varphi(x) = (x - \alpha_r)^m h(x), \quad (h(\alpha_r) \neq 0). \quad (3.104)$$

Reemplazando en $\psi(x)$

$$\begin{aligned} \psi(x) &= x - m \frac{(x - \alpha_r)^m h(x)}{m(x - \alpha_r)^{m-1} h(x) + (x - \alpha_r)^m h'(x)} \\ &= x - m \frac{(x - \alpha_r) h(x)}{m h(x) + (x - \alpha_r) h'(x)}. \end{aligned} \quad (3.105)$$

Derivando esta última expresión y evaluándola en $x = \alpha_r$ obtenemos

$$\psi'(\alpha_r) = 0. \quad (3.106)$$

* * *

Con este resultado, podemos repetir el cálculo del cero de

$$\varphi(x) = e^x - x - 1, \quad (3.107)$$

pero ahora usando como iterador de punto fijo a la función

$$\psi(x) = x - 2 \frac{\varphi(x)}{\varphi'(x)}. \quad (3.108)$$

Arrancando de $x_0 = 1$ obtenemos la siguiente secuencia:

n	x_n
0	1
1	$1,639 \times 10^{-1}$
3	$3,342 \times 10^{-6}$
4	$-2,416 \times 10^{-11}$

que converge a la solución mucho más rápidamente que antes.

3.3. Métodos para encontrar ceros de polinomios

Los polinomios son un caso especial de funciones que permiten una implementación muy eficiente del método de Newton-Raphson. Recordemos que este método requiere de la evaluación de la función cuyo cero queremos encontrar, y de su derivada. Como la derivada de un polinomio también es un polinomio, conviene evaluarlos de la manera más eficiente posible, como veremos a continuación. La idea puede visualizarse mejor con un ejemplo. Consideremos el siguiente polinomio

$$P(x) = 3x^3 - 2x^2 + 6x + 5, \quad (3.109)$$

cuya evaluación requiere de 6 multiplicaciones y tres operaciones de adición. El polinomio puede ser evaluado también haciendo

$$((3x - 2)x + 6)x + 5, \quad (3.110)$$

que implica 3 multiplicaciones y 3 adiciones. Por lo tanto, escribiendo el polinomio adecuadamente podemos ahorrarnos operaciones y, con eso, disminuir el error de redondeo. Aumentar el grado del polinomio incrementa esta diferencia sustancialmente.

Basado en esta idea, el método de Horner nos permite calcular $P(x)$ y $P'(x)$ en forma eficiente.

3.3.1. Método de Horner

Teorema. Sea

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0. \quad (3.111)$$

Si definimos $b_n = a_n$ y

$$b_k = a_k + b_{k+1} x_0, \quad k = n-1, n-2, \dots, 0, \quad (3.112)$$

entonces $b_0 = P(x_0)$. Además, si

$$Q(x) = b_n x^{n-1} + b_{n-1} x^{n-2} + \dots + b_2 x + b_1, \quad (3.113)$$

luego

$$P(x) = (x - x_0)Q(x) + b_0. \quad (3.114)$$

Demostración

Basta con comprobar que

$$P(x) = (x - x_0)Q(x) + b_0, \quad (3.115)$$

para probar que $P(x_0) = b_0$. Partimos de

$$(x - x_0)Q(x) + b_0 = (x - x_0) [b_n x^{n-1} + b_{n-1} x^{n-2} + \dots + b_2 x + b_1] + b_0, \quad (3.116)$$

donde hemos reemplazado la definición de $Q(x)$. Haciendo la distributiva

$$\begin{aligned} & (x - x_0)Q(x) + b_0 \\ &= (b_n x^n + b_{n-1} x^{n-1} + \dots + b_2 x^2 + b_1 x) - (x_0 b_n x^{n-1} + x_0 b_{n-1} x^{n-2} + \dots + x_0 b_2 x + x_0 b_1) + b_0 \\ &= \underbrace{b_n}_{a_n} x^n + \underbrace{(b_{n-1} - x_0 b_n)}_{a_{n-1}} x^{n-1} + \underbrace{(b_{n-2} - x_0 b_{n-1})}_{a_{n-2}} x^{n-2} + \dots + \underbrace{(b_0 - x_0 b_1)}_{a_0}. \end{aligned} \quad (3.117)$$

Por lo tanto, vemos que

$$P(x) = (x - x_0)Q(x) + b_0, \quad (3.118)$$

y, en consecuencia, que

$$\boxed{P(x_0) = b_0.} \quad (3.119)$$

Observación

Además de ofrecer un mecanismo recursivo muy eficiente para calcular $P(x_0)$, el método de Horner nos permite calcular $P'(x_0)$ con un costo mínimo. Si tenemos en cuenta que

$$P(x) = (x - x_0)Q(x) + b_0, \quad (3.120)$$

entonces

$$P'(x) = Q(x) + (x - x_0)Q'(x). \quad (3.121)$$

Evaluando en $x = x_0$ tenemos que

$$P'(x_0) = Q(x_0). \quad (3.122)$$

Dado que los coeficientes de $Q(x)$ son calculados en el proceso de hallar $P(x_0)$, se puede establecer una doble recursividad para computar simultáneamente $P(x_0)$ y $P'(x_0)$. Lo veremos en clase.

3.3.2. Método de deflación

Hemos visto cómo evaluar polinomios eficientemente a fin de implementar el método de Newton-Raphson en forma eficaz. Supongamos que hemos encontrado una aproximación a un cero x_1 del polinomio $P(x)$ utilizando el método de Newton-Raphson. Entonces, de acuerdo con lo demostrado para el método de Horner, sabemos que podemos escribir $P(x)$ como

$$P(x) = (x - x_1)Q(x) + b_0 = (x - x_1)Q(x) + P(x_1), \quad (3.123)$$

con $Q(x)$ otro polinomio. Si x_1 es una buena aproximación al cero de $P(x)$ tendremos que

$$P(x_1) \approx 0, \quad (3.124)$$

y tenemos entonces que

$$P(x) \approx (x - x_1)Q_1(x), \quad (3.125)$$

donde hemos rebautizado a $Q(x)$ como $Q_1(x)$ para facilitar la notación que sigue. Ahora, para seguir buscando los ceros de $P(x)$ podemos trabajar con el polinomio $Q_1(x)$ y proceder como lo hicimos más arriba. Así, a medida que vayamos hallando ceros podremos escribir

$$P(x) \approx (x - x_1)(x - x_2)\dots(x - x_k)Q_k(x). \quad (3.126)$$

A este procedimiento se denomina de deflación y se puede utilizar para encontrar los ceros de $P(x)$. Algunas dificultades que puede presentar:

- los ceros de $P(x)$ son aproximados con mayor error a medida que aumenta k ; para mejorarlo, puedo usar los x_k como valores iniciales del polinomio $P(x)$ e iterarlos con el método de Newton-Raphson,
- un polinomio de coeficientes reales puede tener raíces complejas, que no vamos a encontrar si la aproximación inicial es un número real; para hallarlas tenemos que comenzar con una aproximación inicial compleja y hacer los cálculos usando aritmética compleja.

3.4. Método de Müller

Este método puede aplicarse a cualquier problema de búsqueda de raíces, pero es muy utilizado para aproximar ceros de polinomios.

Es similar al método de la secante. Recordemos que éste utiliza dos aproximaciones iniciales y construye la siguiente con la intersección de la recta que une esos puntos con el eje x . El método de Müller utiliza 3 puntos iniciales, y construye la siguiente aproximación con la intersección de la parábola determinada por esos 3 puntos con el eje x .

Sea $\varphi(x)$ el polinomio al cual le queremos calcular el cero. Construimos una parábola $P(x)$ que pase por los puntos $(x_0, P(x_0))$, $(x_1, P(x_1))$ y $(x_2, P(x_2))$:

$$P(x) = a(x - x_2)^2 + b(x - x_2) + c, \quad (3.127)$$

donde a , b y c son hallados de las condiciones

$$\begin{aligned} P(x_0) &= \varphi(x_0), \\ P(x_1) &= \varphi(x_1), \\ P(x_2) &= \varphi(x_2). \end{aligned} \tag{3.128}$$

Haciendo las cuentas, obtenemos

$$\begin{aligned} c &= \varphi(x_2), \\ b &= \frac{(x_0 - x_2)^2 (\varphi(x_1) - \varphi(x_0)) - (x_1 - x_2)^2 (\varphi(x_0) - \varphi(x_2))}{(x_0 - x_2)(x_1 - x_2)(x_0 - x_1)}, \\ a &= \frac{(x_1 - x_2) (\varphi(x_0) - \varphi(x_2)) - (x_0 - x_2) (\varphi(x_1) - \varphi(x_2))}{(x_0 - x_2)(x_1 - x_2)(x_0 - x_1)}. \end{aligned} \tag{3.129}$$

Para determinar x_3 , un cero de $P(x)$, usamos la fórmula cuadrática en la forma que minimiza el error de redondeo:

$$P(x_3) = 0, \tag{3.130}$$

$$x_3 - x_2 = r_{1,2} \tag{3.131}$$

$$r_1 = \frac{q}{a}, \quad r_2 = \frac{c}{q}, \tag{3.132}$$

con

$$q = -\frac{1}{2} \left(b + \operatorname{sgn}(b) \sqrt{b^2 - 4ac} \right). \tag{3.133}$$

De las dos raíces posibles, en el método de Müller se elige la de menor valor absoluto, resultando x_3 el cero de $P(x)$ más próximo a x_2 . El proceso se repite, usando x_3 , x_2 y x_1 para construir x_4 , luego x_4 , x_3 y x_2 para hallar x_5 , y así sucesivamente hasta que se satisfaga un criterio de finalización elegido.

* * *

Observaciones

- Como el método involucra el cálculo de $\sqrt{b^2 - 4ac}$ debe considerarse la posibilidad de que esta raíz sea compleja, por lo que usando aritmética compleja, aún para coeficientes y valores iniciales reales, el método converge a la(s) posible(s) raíz(raíces) complejas.
- Es posible probar que el método de Müller es al menos de orden 1.84.
- Las condiciones iniciales deben elegirse de modo tal de garantizar que la parábola que pase por los tres puntos corte al eje x , en caso contrario el método no convergerá.