

Raíces de ecuaciones no lineales

Escuela de Ingeniería Informática de Oviedo

Contenidos

- 1 **Introducción**
- 2 **Método de bisección**
- 3 **Iteración de punto fijo**
- 4 **El método de Newton-Raphson**
- 5 **Método de la secante**

Problema: resolver ecuaciones

Una ecuación escalar es una expresión de la forma

$$f(x) = 0, \tag{1}$$

donde $f : \Omega \subset \mathbb{R} \rightarrow \mathbb{R}$ es una función.

Cualquier número α tal que $f(\alpha) = 0$ se dice que es una **solución de (1)** o una **raíz de f** .

Ejemplos:

- Ecuación: $x^2 - 2x + 1 = 0$ ($f(x) = x^2 - 2x + 1$).
Raíces: $x = 1$.
- Ecuación: $\sin(\log(x)) = 0$ ($f(x) = \sin(\log(x))$).
Raíces: $x = 1, x = e^\pi, \dots$ (infinitas raíces).

Multiplicidad de una raíz: Si

$$f(\alpha) = f'(\alpha) = \dots = f^{(m-1)}(\alpha) = 0, \quad \text{and } f^{(m)}(\alpha) \neq 0,$$

entonces decimos que α es una raíz de f con **multiplicidad** m . En el caso de que $m = 1$ decimos que la raíz es **simple**.

Decimos que la raíz α de $f(x)$ está **separada en un subconjunto** $\omega \subset \Omega$ si α es la única raíz de $f(x)$ en ω .

Ejemplos:

- Ecuación: $x^2 - 2x + 1 = 0$ ($f(x) = x^2 - 2x + 1$).
Raíces: $x = 1$ de multiplicidad $m = 2$.
- Ecuación: $\sin(\log(x)) = 0$ ($f(x) = \sin(\log(x))$).
Raíces: $x = 1, x = e^\pi, \dots$ (infinitas raíces simples).

Separación de raíces:

- *Métodos gráficos.* El gráfico de una función permite determinar aproximadamente (con las debidas precauciones) qué intervalo contiene una única raíz.
- *Métodos analíticos.* Basados en las propiedades de las funciones continuas (Bolzano, Rolle)

Teorema (Bolzano)

Sea $f : [a, b] \rightarrow \mathbb{R}$ una función continua con $f(a)f(b) < 0$. Entonces existe al menos una raíz de f en $[a, b]$.

Además, si f es monótona estricta en $[a, b]$ entonces la raíz es única y está separada.

Teorema (Rolle)

Sea $f : [a, b] \rightarrow \mathbb{R}$ continua en $[a, b]$ y derivable en (a, b) con $f(a) = f(b)$. Entonces existe al menos una raíz de f' en $[a, b]$.

Consecuencias:

- Entre dos raíces de f existe, al menos, una raíz de f' .
- Entre dos raíces consecutivas de f' existe, como máximo, una raíz de f .
- El número de raíces de f en (a, b) es
 - par, si $f(a)f(b) < 0$,
 - impar o cero, si $f(a)f(b) > 0$.

Orden de convergencia

Los métodos numéricos de cálculo de raíces de una función son *métodos iterativos*, es decir, construimos una sucesión

$$x_0, x_1, \dots, x_k, \dots$$

tal que

$$\lim_{k \rightarrow \infty} f(x_k) = 0.$$

El orden de convergencia de un métodos está relacionado con la *velocidad de convergencia* de la sucesión con respecto a k .

Este concepto es útil para comparar métodos.

Orden de convergencia

Supongamos que la sucesión $(x_k)_k$ converge a $\alpha \in \mathbb{R}$. Decimos que $(x_k)_k$ converge a α con orden de convergencia p si

$$\lim_{k \rightarrow \infty} \frac{|x_k - \alpha|}{|x_{k-1} - \alpha|^p} = \lambda \neq 0.$$

En los casos particulares

- $p = 1$, decimos que la convergencia es lineal.
- $p = 2$, decimos que la convergencia es cuadrática.

Un método numérico se dice de orden p si genera una sucesión que converge a la solución con un orden de convergencia p .

Criterios de parada

Como la sucesión $(x_k)_k$ tal que $x_k \rightarrow \alpha$, con $f(\alpha) = 0$ es, en general, infinita, vamos a introducir un criterio para decidir cuando paramos.

Los criterios de parada se basan en

- El error absoluto,

$$|x_k - x_{k-1}| < \varepsilon.$$

- El error relativo,

$$\frac{|x_k - x_{k-1}|}{|x_k|} < \varepsilon.$$

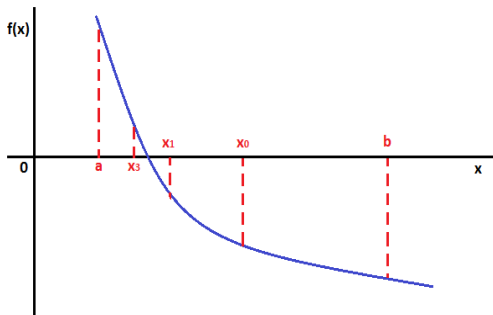
- El residual,

$$|f(x_k)| < \varepsilon.$$

Método de bisección

Sea $F : [a, b] \rightarrow \mathbb{R}$ tal que $f(a)f(b) < 0$. El teorema de Bolzano garantiza la existencia de una raíz de f en (a, b) .

La idea del método de bisección es:



Método de bisección

Algoritmo:

- Sea $a_1 = a$, $b_1 = b$.
- Para $k = 1, 2, \dots, \text{MaxIteraciones}$
 - Calcular el punto medio $x_k = \frac{a_k + b_k}{2}$ y evaluar $f(x_k)$.
 - Si x_k satisface el criterio de parada, parar.
 - En el caso contrario,
 - si $f(a_k)f(x_k) < 0$ entonces

$$a_{k+1} = a_k, \quad b_{k+1} = x_k,$$

- en el caso contrario

$$a_{k+1} = x_k, \quad b_{k+1} = b_k.$$

Método de bisección

Propiedad: Para cada k definimos el intervalo $[a_{k+1}, b_{k+1}]$ que satisface

- $[a_{k+1}, b_{k+1}] \subset [a_k, b_k]$,
- $f(a_{k+1})f(b_{k+1}) < 0$,
- $b_{k+1} - a_{k+1} = \frac{b_k - a_k}{2} = \dots = \frac{b - a}{2^k}$.

Por lo tanto, siempre tenemos $x_k \rightarrow \alpha$, una raíz de f , con

$$|x_k - \alpha| \leq \frac{b - a}{2^k}.$$

Si el criterio de parada se basa en el error absoluto, $|x_k - \alpha| < \varepsilon$ entonces necesitamos $\lceil \log_2 \frac{b - a}{\varepsilon} \rceil + 1$ iteraciones para conseguirlo.

Método de bisección

Ventajas:

- Fácil de programar.
- si aplicamos el teorema de Bolzano:
 - El método es convergente,
 - el error absoluto es fácil de estimar,
 - podemos saber a priori cuantas iteraciones tenemos que realizar.

Desventajas:

- La velocidad de convergencia es lenta,
- Podemos estar cerca de la raíz y en la siguiente iteración alejarnos.

En la práctica, el método de bisección se utiliza para inicializar métodos más rápidos.

Iteración de punto fijo

Decimos que x es un punto fijo de g si

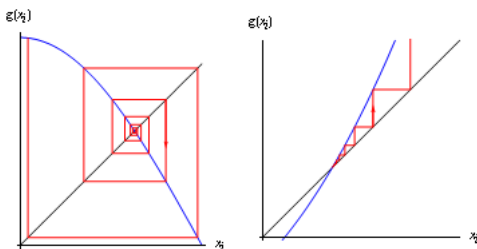
$$g(x) = x.$$

Como cualquier ecuación $f(x) = 0$ puede ser reescrita como $g(x) = x$, con $g(x) = f(x) + x$, resolver la primera ecuación o encontrar un punto fijo de la segunda función es equivalente.

Ejemplo: En la ecuación $x^3 - 4x^2 + 3 = 0$, podemos escribir $f(x) = 0$, con $f(x) = x^3 - 4x^2 + 3$, o $g(x) = x$, con $g(x) = x^3 - 4x^2 + x + 3$.

Iteración de punto fijo

El algoritmo de iteración de punto fijo es sencillo (pero no siempre converge):



- Sea x_0 un punto suficientemente próximo a la raíz de f
- Para $k = 1, 2, \dots, \text{MaxIter}$:
 - Calcular $x_k = g(x_{k-1})$
 - Probar si se cumple el criterio de parada. Por ejemplo, si

$$|x_k - x_{k-1}| < \varepsilon,$$

entonces aceptamos x_k como una aproximación del punto fijo de g (raíz de f).

Contractividad

¿Cuándo converge el método de punto fijo?

Decimos que g es **contractiva** en $[a, b]$ si existe una constante (**constante de contractividad**) $\gamma \in (0, 1)$ tal que

$$|g(x) - g(y)| \leq \gamma |x - y|$$

para todo $x, y \in [a, b]$.

Señalar que:

- Si g es contractiva entonces es continua en $[a, b]$, pero no necesariamente derivable.
- Si g está definida en el intervalo $[a, b]$, podría suceder que $x_k = g(x_{k-1}) \notin [a, b]$, y entonces el algoritmo debería parar (sin encontrar la solución).
- Se llama *contractiva* porque la función g comprime o contrae cualquier subintervalo de $[a, b]$.

Comprobando la contractividad

A partir de la definición tenemos que

$$\frac{|g(x) - g(y)|}{|x - y|} \leq \gamma.$$

Si tomamos el límite $y \rightarrow x$, si f es derivable,

$$|g'(x)| \leq \gamma.$$

Por lo tanto, si

$$|g'(x)| \leq \gamma < 1,$$

entonces g es contractiva.

Resultados matemáticos

Teorema (Existencia y unicidad del punto fijo)

Sea $g : [a, b] \rightarrow \mathbb{R}$ una función que cumple:

- $g(x) \in [a, b]$ para cualquier $x \in [a, b]$,
- g es contractiva en $[a, b]$, con constante de contractividad γ .

Entonces g tiene un único punto fijo $\alpha \in [a, b]$ y la sucesión

$$x_0, \quad x_k = g(x_{k-1}) \quad \text{for } k \geq 1, \dots$$

converge a α para cualquier valor inicial $x_0 \in [a, b]$.

Señalar que: El método de punto fijo tiene, al menos, orden de convergencia lineal. El error absoluto se puede estimar con

$$|x_k - \alpha| \leq \frac{\gamma^k}{1 - \gamma} |x_1 - x_0|.$$

Resultados matemáticos

Teorema (Existencia y unicidad del punto fijo)

Sea $g : [a, b] \rightarrow \mathbb{R}$ una función que cumple:

- $g(x) \in [a, b]$ para cualquier $x \in [a, b]$,
- g es contractiva en $[a, b]$, con constante de contractividad γ .

Entonces g tiene un único punto fijo $\alpha \in [a, b]$ y la sucesión

$$x_0, \quad x_k = g(x_{k-1}) \quad \text{for } k \geq 1, \dots$$

converge a α para cualquier valor inicial $x_0 \in [a, b]$.

Señalar que: El método de punto fijo tiene, al menos, orden de convergencia lineal. El error absoluto se puede estimar con

$$|x_k - \alpha| \leq \frac{\gamma^k}{1 - \gamma} |x_1 - x_0|.$$

El método de Newton-Raphson

Es el método más usado cuando f es derivable.

- Sea x_0 un punto inicial.
- Para $k = 1, 2, \dots, \text{MaxIter}$:
 - Calcular $x_k = x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}$
 - Comprobar el criterio de parada. Por ejemplo, si

$$|x_k - x_{k-1}| < \varepsilon,$$

aceptamos x_k como una aproximación de la raíz de f .

Señalar que: Podríamos escribirlo como un método de punto fijo utilizando la función:

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

¿De dónde sale la fórmula de Newton? La tangente a f en $\hat{x}$ es la línea recta

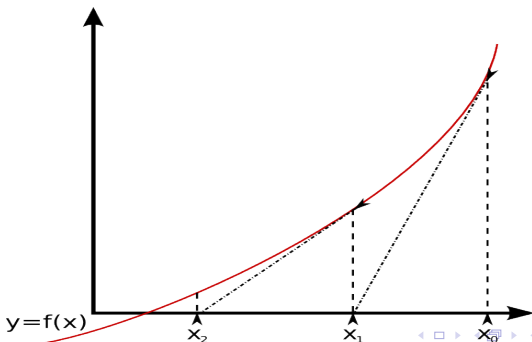
$$T_f(x) = f(\hat{x}) + f'(\hat{x})(x - \hat{x}).$$

La tangente cruza el eje horizontal cuando $T_f(x) = 0$, es decir, en

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

Ahora escribimos las k para hacer iteraciones,

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



Señalar que:

- En general, el método sólo converge si el punto inicial, x_0 está *suficientemente próximo* a la raíz, α . En la práctica, se usa el método de bisección unas cuantas veces, o se selecciona sobre una gráfica de f para inicializar el algoritmo.
- Si $f''(x)$ es continua y $f'(\alpha) \neq 0$ (la raíz es simple) entonces el método de Newton-Raphson tiene convergencia cuadrática, es decir

$$|x_k - \alpha| \leq \lambda |x_{k-1} - \alpha|^2,$$

para alguna constante $\lambda > 0$.

- Existen **muchas variantes** de este método.

Método de la secante

Es la variante más importante del método de Newton-Raphson.

Problema: Para el método de N-R necesitamos tener $f'(x)$.

En la secante aproximamos $f'(x)$ con *diferencias finitas*:

$$f'(x_{k-1}) = \frac{f(x_{k-1}) - f(x_{k-2})}{x_{k-1} - x_{k-2}}.$$

Inconveniente: necesitamos **dos puntos iniciales**, x_0 y x_1 para la primera iteración:

$$x_k = x_{k-1} - f(x_{k-1}) \frac{x_{k-1} - x_{k-2}}{f(x_{k-1}) - f(x_{k-2})}.$$