Resolución de ecuaciones de una variable

Computación 2021

FaMAF

14 abr. 2021

Contenidos

Códigos

Códigos de la clase en:

https://github.com/mlares/compuprof2021

Python

Strings

Resolución de ecuaciones de una variable

Ecuaciones de una variable

Queremos resolver el problema de hallar x tal que

$$f(x) = 0,$$

donde $f : \mathbb{R} \mapsto \mathbb{R}$ es una función con características a determinar. O sea, vamos a encontrar una solución de esta ecuación, que equivale a determinar una **raíz** de la función f. A las raíces de una función también se las suele llamar también los "ceros" de la función.

Notar que no estamos interesados en hallar todas las raíces, en caso de existir más de una, sino en obtener al menos una de ellas.

Para resolver nuestro problema utilizaremos **métodos de aproximación**, esto es, generar una sucesión que converja a la solución del problema.

Raíces de una ecuación

Def./

Una raíz de una función $f(x)$ es un número $\bar{x}$ tan que

$$f(\bar{x}) = 0$$

Ecuaciones de una variable

Dependiendo de ciertas propiedades de f, existen algoritmos capaces de resolver el problema. Por ejemplo, si f es continua, convexa, Lipschitz continua, diferenciable, continuamente diferenciable, etc.

Algunas funciones son simples y se puede encontrar la raíz analíticamente, sin necesidad de métodos numéricos. ▷ e.g.:

$$f(x) = 2x - 6$$

Por otro lado, hay funciones cuya resolución no resulta simple, ▷ e.g.:

$$xe^x = 1$$

Vamos a considerar funciones cuya solución analítica no es fácil o no es posible, pero es fácil evaluar la función en cualquier punto x.

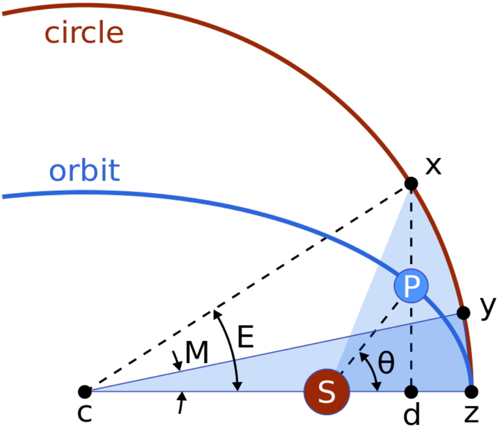
▷ *e.g.: Ecuación de Kepler*

Las leyes de Kepler describen de manera aproximada el movimiento de los planetas alrededor del Sol.
Los planetas describen una órbita elíptica con el Sol en uno de sus focos, y se mueven con velocidad variable en función de la posición en la órbita. La tercera ley de Kepler establece que el radio vector Sol-planeta barre áreas iguales en tiempos iguales.

$$M = E - e\ sin(E),$$

donde M es la anomalía media o ángulo que recorrería un planeta ficticio que se moviese con movimiento uniforme por la circunferencia principal, e es la excentricidad de la elipse ($0 < e < 1$) y E es la anomalía excéntrica, que queremos encontrar.

▷ *e.g.: Ecuación de Kepler*



▷ *e.g.: Modelo de crecimiento exponencial*

En ciertas situaciones, en periodos cortos de tiempo y bajo ciertas condiciones de aislamiento, el crecimiento de una población se puede modelar considerando que la variación en la misma es proporcional al tamaño de la población:

$$\frac{dN(t)}{dt} = \lambda\, N(t)$$

que tiene como solución:

$$N(t) = N_0\, e^{\lambda\, t}$$

donde N_0 es la población inicial. Si además se permite una "inmigración" constante a una tasa ν (individuos por unidad de tiempo), entonces

$$\frac{dN(t)}{dt} = \lambda\, N(t) + \nu; \qquad N(t) = N_0\, e^{\lambda\, t} + \frac{\nu}{\lambda}(e^{\lambda t} - 1)$$

▷ e.g.: Modelo de crecimiento exponencial

$$\frac{dN(t)}{dt} = \lambda\, N(t) + \nu; \qquad N(t) = N_0\, e^{\lambda\, t} + \frac{\nu}{\lambda}(e^{\lambda t} - 1)$$

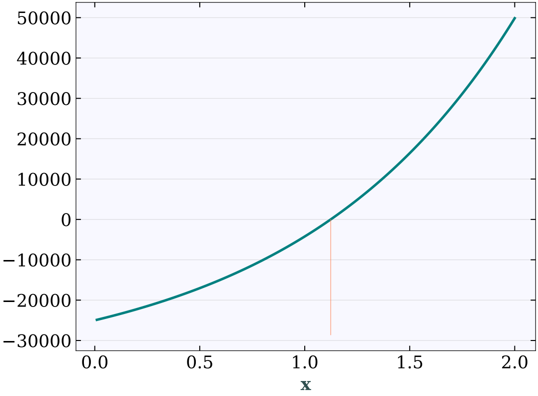
Supongamos ahora que cierta población tiene inicialmente $N(0) = 10000$ individuos, y que cruzan la frontera $\nu = 5000$ individuos por año. Al cabo de un año la población total asciende a $N(1) = 16000$.
Para determinar la tasa de "crecimiento" de la población (por ejemplo, la tasa de nacimientos o la tasa de contagios) se quiere encontrar λ que cumpla:

$$16000 = 10000\, e^{\lambda} + \frac{5000}{\lambda}(e^{\lambda} - 1)$$

es decir,

$$f(\lambda) = 10000\, e^{\lambda} + \frac{5000}{\lambda}(e^{\lambda} - 1) - 16000 = 0$$

▷ e.g.: Modelo de crecimiento exponencial



Método de Bisección

El método de bisección permite hallar x_0 solución de $f(x) = 0$ en un intervalo $[a, b]$ si:

- f es continua en $[a, b]$,
- $f(a)$ y $f(b)$ tienen signos distintos

Teorema

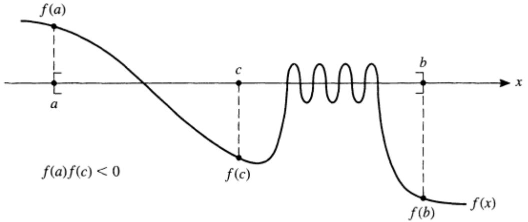
El siguiente teorema asegura la existencia de una raíz real:

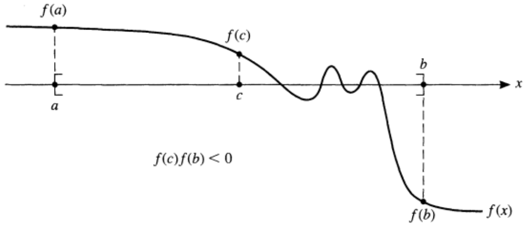
Teorema

Sea f una función con valores reales definida y contínua en un intervalo $[a, b]$, con $f(a)f(b) \leq 0$,

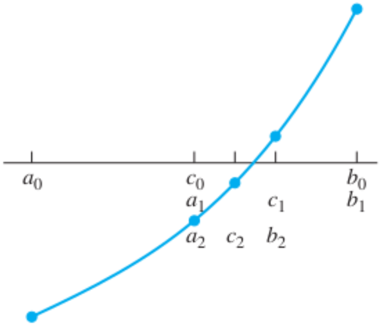
entonces $\exists \ \xi \ \in \ [a, b] \ / \ f(\xi) = 0$

Método de Bisección





Método de Bisección



Construcción del método

Veamos como generar una sucesión $\{x_n\}_{n=0}^{\infty}$ que converja a x_0.
Si $f(a) = 0$ o $f(b) = 0$, ya tenemos resuelto nuestro problema. Por lo tanto, supongamos que $f(a) \neq 0$ y $f(b) \neq 0$. Tomemos $a_0 = a$ y $b_0 = b$.
Como f es continua en el intervalo $[a_0, b_0]$ y tiene signos distintos en sus extremos, entonces podemos garantizar que f se anula en (a_0, b_0). En este caso, definimos x_0 como el punto medio del intervalo, o sea

$$x_0 = \frac{a_0 + b_0}{2} = a_0 + \frac{b_0 - a_0}{2}.$$

Si $f(x_0) = 0$ entonces $\bar{x} = x_0$. Si $f(x_0) \neq 0$ entonces debe tener signo distinto al de $f(a_0)$ o $f(b_0)$.
Si $f(x_0)$ y $f(a_0)$ tienen signos distintos, entonces f se anula en (a_0, x_0). En este caso definimos

$$a_1 = a_0 \quad \text{y} \quad b_1 = x_0.$$

Construcción del método

Por el contrario, si $f(x_0)$ y $f(a_0)$ no tienen signos distintos entonces $f(x_0)$ y $f(b_0)$ deben tener signos distintos, luego f se anula en (x_0, b_0). En este caso definimos

$$a_1 = x_0 \quad \text{y} \quad b_1 = b_0.$$

Ahora podemos aplicar el mismo procedimiento en el intervalo $[a_1, b_1]$. Note que con este procedimiento se van generando intervalos encajados $[a_0, b_0] \supset [a_1, b_1] \supset \ldots [a_n, b_n] \ldots$ donde cada uno contiene al menos una raíz y su longitud es la mitad de la longitud del intervalo precedente.

Pseudocódigo

1. Dar a y b tal que $f(a)f(b) < 0$. Tomar $a_0 = a$, $b_0 = b$ y $k = 0$.
2. Tomar
$$x_k = a_k + \frac{b_k - a_k}{2}.$$
3. Si $f(x_k) = 0$, parar y retornar $\bar{x} = x_k$.
4. Si $f(x_k)f(a_k) < 0$, tomar
$$a_{k+1} = a_k \quad \text{y} \quad b_{k+1} = x_k.$$
Sino, tomar
$$a_{k+1} = x_k \quad \text{y} \quad b_{k+1} = b_k.$$
5. Hacer $k \leftarrow k + 1$ e ir al paso 2.

Implementación del método

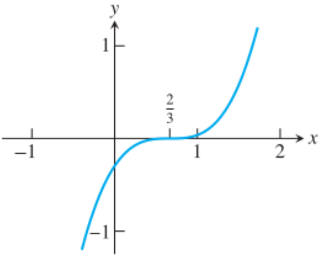
El criterio de parada

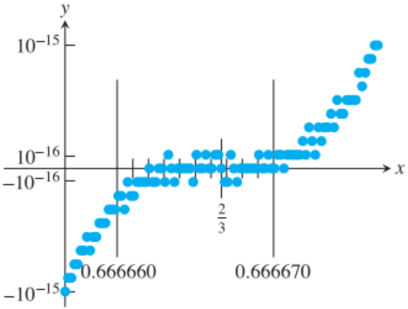
$$f(x_k) = 0$$

no es una condición apropiada para trabajar en aritmética finita.

- Puede que nunca se cumpla debido al error de redondeo del número $f(x_k)$,
- puede que nunca se cumpla debido a que el intervalo $[a_k, b_k]$ es muy pequeño y el error de redondeo hace que x_k sea igual a uno de los extremos del intervalo,
- puede que $\bar{x}$ sea el límite de $\{x_k\}$ y necesitaríamos realizar infinitas iteraciones.

Método de Bisección





Construcción del método

Este tipo de inconvenientes se resuelven fijando tolerancias deseadas. En este caso, considere una tolerancia $\varepsilon > 0$ para el error del valor funcional, $\delta > 0$ para el error a la solución y $N \in \mathbb{N}$ para el número máximo de iteraciones. De esta manera, detendremos la ejecución del algoritmo si:

1. $|f(x_k)| < \varepsilon$
2. $|x_k - x_{k-1}| < \delta$
3. $k > N$.

Construcción del método

En la mayoría de los casos es preferible que las condiciones 1 y 2 se verifiquen simultáneamente.

Notar que en la k-ésima iteración tendremos que $\bar{x} \in (a_k, x_k)$ o $\bar{x} \in (x_k, b_k)$, y que $|x_k - a_k| = |x_k - b_k|$. Por construcción, $a_k = x_{k-1} \vee b_k = x_{k-1}$. Por lo tanto

$$|x_k - \bar{x}| \leq |x_k - x_{k-1}|.$$

Por otro lado, como la longitud de los intervalos se reduce a la mitad, tendremos que

$$b_{k+1} - a_{k+1} = \frac{b_k - a_k}{2} = \frac{b_{k-1} - a_{k-1}}{2^2} = \ldots = \frac{b_0 - a_0}{2^{k+1}}.$$

Construcción del método

Es decir,

$$\log_2(b_{k+1} - a_{k+1}) = \log_2\left(\frac{b_0 - a_0}{2^{k+1}}\right) = \log_2(b_0 - a_0) - (k+1)$$

Si queremos que para un valor de N,

$$(b_{N+1} - a_{N+1}) < \log_2(\delta)$$

entonces

$$\log_2(b_0 - a_0) - (N+1) < \log_2(\delta) \implies N > \log_2\left(\frac{b_0 - a_0}{\delta}\right) - 1$$

Construcción del método

Luego, si tomamos

$$N > \log_2\left(\frac{b_0 - a_0}{\delta}\right) - 1,$$

entonces

$$\frac{b_0 - a_0}{2^{N+1}} < \delta.$$

Obteniendo que

$$|x_N - x_{N-1}| = b_{N+1} - a_{N+1} = \frac{b_0 - a_0}{2^{N+1}} < \delta.$$

Con esto relacionamos los criterios de parada 2 y 3. Obteniendo que el criterio 2 se debe cumplir luego de N iteraciones.

Teorema del método de bisección

Teorema

Si $[a_0, b_0], [a_1, b_1], ..., [a_n, b_n], ...$ *son los intervalos del método de bisección, entonces* $\exists$ r *tal que*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = r$$

y es un cero de f.

Si $c_n = \frac{1}{2}(a_n + b_n)$ *y* $r = \lim_{n\to\infty} c_n$, *entonces*

$$|r - c_n| \leq 2^{-(n+1)}(b_0 - a_0)$$

$\boxed{\text{Ej.:}}$ Se comienza el método de bisección en el intervalo $[12, 19]$. ¿Cuántas iteraciones se requieren para asegurar un error relativo menor que 10^{-3}?

Teorema de convergencia del método de Bisección

Teorema

Sea f una función continua en el intervalo $[a, b]$ tal que $f(a)f(b) < 0$. Suponga que $\{x_k\}$ es una sucesión generada por el método de bisección. Entonces ocurre alguna de las siguientes:

- $f(x_k) = 0$ *para algún* $k \geq 0$,
- $\lim_{k\to\infty} x_k = \bar{x}$ *y* $f(\bar{x}) = 0$.

Demostración del teorema de convergencia del método de Bisección

Si no vale que $f(x_k) = 0$ para algún $k \geq 0$, entonces el método genera una sucesión $\{x_k\}$.

La sucesión $\{x_k\}$ generada por el método tiene algunos de sus miembros en la sucesión $\{a_k\}$ y otros en la sucesión $\{b_k\}$. Por construcción tenemos que $\{a_k\}$ es una sucesión no decreciente acotada superiormente por b_0, y $\{b_k\}$ es una sucesión no creciente acotada inferiormente por a_0. Por lo tanto, ambas sucesiones son convergentes. Además, como

$$0 < b_k - a_k = \frac{b_0 - a_0}{2^k},$$

obtenemos que existe $\bar{x} = \lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k$. En consecuencia, $\lim_{k\to\infty} x_k = \bar{x}$.

Además, como para todo k vale que $f(a_k)f(b_k) < 0$, tomando límite y usando que f es continua obtenemos

$$0 \geq f(\bar{x})f(\bar{x}) = \big(f(\bar{x})\big)^2.$$

Por lo tanto, $f(\bar{x}) = 0$.

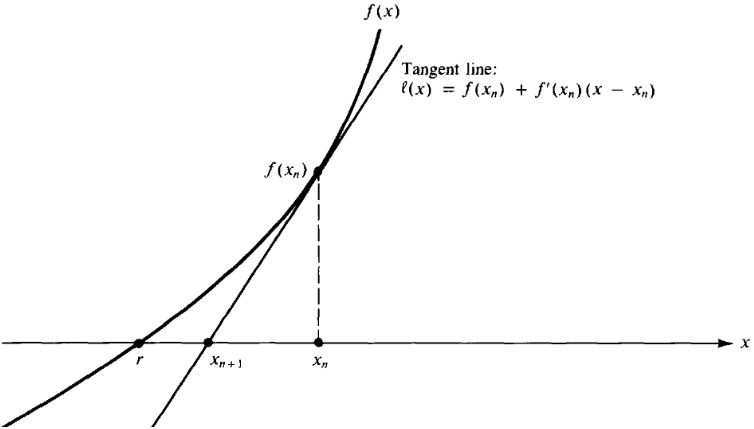
Método de Newton

El método de Newton permite hallar $\bar{x}$ solución de $f(x) = 0$ si:

- f es continuamente diferenciable en un entorno de $\bar{x}$,
- $f'(\bar{x}) \neq 0$,
- x_0 es cercano a $\bar{x}$.

Este método fue propuesto por Isaac Newton (1669) para encontrar raíces de polinomios cúbicos, aunque también lo utilizó para una función no lineal. Luego Joseph Raphson (1690) reescribió el método para encontrar raíces de polinomios, construyéndolo de manera distinta a la propuesta por Newton. Finalmente Thomas Simpson (1740) lo escribió como lo conocemos en la actualidad y lo utilizó para resolver un sistema de dos ecuaciones con dos incógnitas.

Método de Newton



Construcción del método

Veamos cómo el hace el método de Newton para generar una sucesión $\{x_n\}_{n=0}^{\infty}$ que converja a una raíz $\bar{x}$ de f.
El razonamiento usado en este método puede encontrarse en diversas áreas de la matemática y responde al hecho que toda función diferenciable en un punto puede aproximarse localmente por la recta tangente a ese punto. O sea, aproximar f en un punto x_0 por su polinomio de Taylor de grado 1:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

para todo x en un entorno de x_0. Entonces, el problema de hallar x tal que $0 = f(x)$ se cambia por el problema auxiliar de hallar x tal que $0 = f(x_0) + f'(x_0)(x - x_0)$. La solución de este problema auxiliar, llamémosla x_1, servirá como nuevo punto para definir el siguiente problema auxiliar.

Construcción del método

Entonces, dado x_0 el método de Newton genera una sucesión $\{x_k\}$ tal que para cada $k \geq 0$ definimos x_{k+1} como la solución de

$$0 = f(x_k) + f'(x_k)(x - x_k).$$

Resolviendo la ecuación lineal, se obtiene que

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

La condición que f' es continua con $f'(\bar{x}) \neq 0$ (o sea $\bar{x}$ es una raíz simple de f) es de suma importancia para el método. Esto garantiza que $f'(x) \neq 0$ para todo x en un entorno de $\bar{x}$. Luego

- la ecuación lineal tiene solución si x_k está próximo a $\bar{x}$ (pues $f'(x_k) \neq 0$), y
- no existe otra raíz de f cerca de $\bar{x}$.

Construcción del método

Para ver que no existe otra raíz de f cerca de $\bar{x}$, consideremos que si existiera x^* raíz de f en un entorno de $\bar{x}$ obtendríamos que para un ξ entre estos dos puntos vale que:

$$0 = f(x^*) - f(\bar{x}) = f'(\xi)(x^* - \bar{x}),$$

y como $f'(\xi) \neq 0$ entonces $x^* = \bar{x}$.

Pseudocódigo

1. Dar x_0 y $k = 0$.
2. Si $f(x_k) = 0$, parar y retornar x_k.
3. Si $f'(x_k) = 0$, parar porque el método no puede continuar.
4. Definir
$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$
5. Hacer $k \leftarrow k + 1$ e ir al paso 2

Implementación del método

Los criterios de parada $f(x_k) = 0$ y $f'(x_k) = 0$ no son apropiados para trabajar en aritmética finita. Para la implementación, considere una tolerancia $\varepsilon > 0$ para el error del valor funcional, $\delta > 0$ para el error a la solución y $N \in \mathbb{N}$ para el número máximo de iteraciones. De esta manera, detendremos la ejecución del algoritmo si:

1. $|f(x_k)| < \varepsilon$
2. $|x_k - x_{k-1}| < \delta$
3. $k > N$

En general queremos que suceda $(1 \wedge 2) \vee 3$

Para no dividir por un $f'(x_k)$ próximo a cero, podemos detener la ejecución si $|f'(x_k)|$ es menor al épsilon de la máquina, o sea, paramos si no vale que $1 + |f'(x_k)| > 1$.

Criterio de convergencia

1. $C_1 = |f(x_k)| < \varepsilon$
2. $C_2 = |x_k - x_{k-1}| < \delta$
3. $C_3 = k > N$

En general queremos que suceda $C = (C_1 \wedge C_2) \vee C_3$

Notar que si el criterio que debe cumplirse es C, quiere decir que mientras no se cumpla C se debe seguir iterando.
Teniendo en cuenta que para dos proposiciones lógicas P y Q, se cumple que (Leyes de Morgan):

$$\begin{aligned} \neg(P \wedge Q) &\iff \neg P \vee \neg Q \\ \neg(P \vee Q) &\iff \neg P \wedge \neg Q, \end{aligned}$$

la negación de C es:

$$\neg C = \neg(C_1 \wedge C_2) \wedge \neg C_3 = (\neg C_1 \vee \neg C_2) \wedge \neg C_3$$

Criterio de convergencia

```
tolerancia_x = 1.e-6
tolerancia_f = 1.e-6
tolerancia_N = 50

while condicion:

    ...

    condicion1 = delta_x > tolerancia_x
    condicion2 = delta_f > tolerancia_f
    condicion3 = i < tolerancia_N
    condicion = (condicion1 or condicion2) and condicion3
```

Teorema: método de Newton

Teorema

Sea f una función continuamente diferenciable en un entorno de un punto $\bar{x}$ tal que $f(\bar{x}) = 0$ y $f'(\bar{x}) \neq 0$. Entonces existe un intervalo $(\bar{x} - \varepsilon, \bar{x} + \varepsilon)$ tal que si x_0 se toma en ese intervalo, el método de Newton o finaliza con $f(x_k) = 0$ para algún $k \geq 0$, o genera una sucesión infinita $\{x_k\}$ tal que:

- $f'(x_k) \neq 0$ *para todo* k,
- $\lim_{k\to\infty} x_k = \bar{x}$

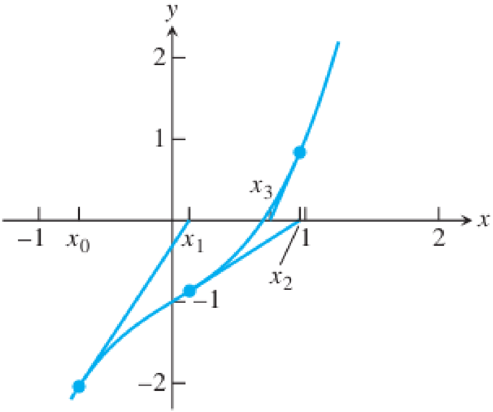
Además, se tiene que

$$\lim_{k\to\infty} \frac{|x_{k+1} - \bar{x}|}{|x_k - \bar{x}|} = 0,$$

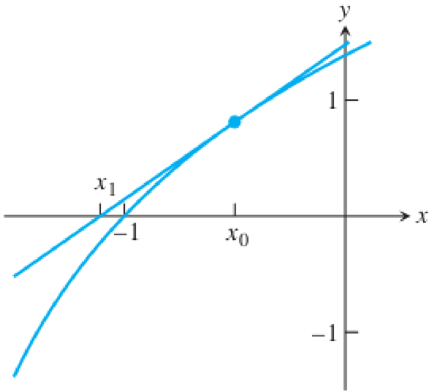
y si existe L tal que $|f'(y) - f'(x)| \leq L|y - x|$ entonces

$$\limsup_{k\to\infty} \frac{|x_{k+1} - \bar{x}|}{|x_k - \bar{x}|^2} < \infty.$$

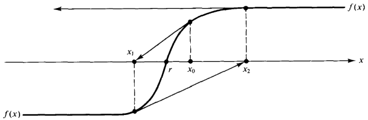
Método de Newton

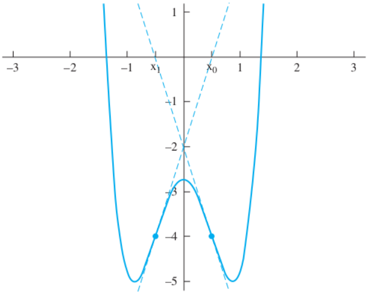


Método de Newton



Método de Newton: casos patológicos





Ref.: Donovan, G. C., Miller, A. R., & Moreland, T. J. (1993). Pathological Functions for Newton's Method. The American Mathematical Monthly, 100(1)

Demostración de convergencia

Si no vale que $f(x_k) = 0$ para algún $k \geq 0$, entonces el método genera una sucesión $\{x_k\}$.
Por continuidad de f', tomando $c = |f'(\bar{x})| > 0$, existe $\varepsilon > 0$ tal que si $x \in (\bar{x} - \varepsilon, \bar{x} + \varepsilon)$ entonces

$$|f'(x)| > \frac{c}{2}, \quad \sup_{t \in [0,1]} |f'(x + t(\bar{x} - x)) - f'(x)| < \frac{c}{4}.$$

Demostración de convergencia

Veamos por inducción que si $x_k \in (\bar{x} - \varepsilon, \bar{x} + \varepsilon)$ entonces $x_{k+1} \in (\bar{x} - \varepsilon, \bar{x} + \varepsilon)$.
De la iteración del método y usando que $f(\bar{x}) = 0$ tenemos

$$\begin{aligned} f(\bar{x}) = 0 &= f(x_k) + f'(x_k)(x_{k+1} - x_k) \\ &= f(x_k) + f'(x_k)(x_{k+1} - \bar{x}) + f'(x_k)(\bar{x} - x_k). \end{aligned}$$

Como $f'(x_k) \neq 0$ pues x_k está en el intervalo, entonces

$$\begin{aligned} x_{k+1} - \bar{x} &= \frac{1}{f'(x_k)} \big(f(\bar{x}) - f(x_k) - f'(x_k)(\bar{x} - x_k) \big) \\ &= \frac{1}{f'(x_k)} \left(\int_0^1 \big(f'(x_k + t(\bar{x} - x_k)) - f'(x_k) \big)(\bar{x} - x_k) dt \right), \end{aligned}$$

Donde usamos que para $\phi(t) = f(x_k + t(\bar{x} - x_k))$ se tiene $\phi'(t) = f'(x_k + t(\bar{x} - x_k))(\bar{x} - x_k)$.

Demostración de convergencia

Tomando valor absoluto obtenemos

$$|x_{k+1} - \bar{x}| \leq \frac{2}{c} \left(\sup_{t \in [0,1]} |f'(x_k + t(\bar{x} - x_k)) - f'(x_k)| \right) |x_k - \bar{x}|.$$

Usando la cota para el supremo y el hecho que $|x_k - \bar{x}| < \varepsilon$ se consigue que

$$|x_{k+1} - \bar{x}| \leq \frac{1}{2} |x_k - \bar{x}| < \frac{\varepsilon}{2} < \varepsilon.$$

Por lo tanto x_{k+1} también está en el intervalo. Concluyendo que si x_0 está en el intervalo toda la sucesión queda en el intervalo. Con esto, garantizamos que

$$|x_k - \bar{x}| \leq \frac{1}{2} |x_{k-1} - \bar{x}| \leq \ldots \leq \frac{1}{2^k} |x_0 - \bar{x}|,$$

y por lo tanto $\lim_{k \to \infty} x_k = \bar{x}$.

Demostración de convergencia

De aquí se obtiene la velocidad de convergencia superlineal

$$\frac{|x_{k+1} - \bar{x}|}{|x_k - \bar{x}|} \leq \frac{2}{c} \left(\sup_{t \in [0,1]} |f'(x_k + t(\bar{x} - x_k)) - f'(x_k)| \right) \to 0.$$

y si $|f'(y) - f'(x)| \leq L|y - x|$ obtenemos la velocidad de convergencia cuadrática

$$\frac{|x_{k+1} - \bar{x}|}{|x_k - \bar{x}|^2} \leq \frac{2L}{c}.$$