Curso de Métodos Numéricos

DEMAT, Universidad de Guanajuato

Clase 4: **Solución de ecuaciones no lineales (Parte 1)**

- Librería GSL
- Método de bisección
- Métodos de punto fijo
- Método de Newton-Raphson

MAT–251

Dr. Joaquín Peña Acevedo
CIMAT A.C.
e-mail: joaquin@cimat.mx

Librería GSL (I)

La librería GSL comprende una amplia gama de métodos numéricos robustos.

- Está escrita en C.
- En la documentación se puede ver la cantidad y variedad de funciones que tiene.
- Hay que consultar la página del curso para las instrucciones de instalación, en el apartado "Extra".
- Para poder usar una función, hay que ver en que archivo .h se encuentra declarada e incluirlo en el código del programa.

Por ejemplo, para usar las funciones matemáticas básicas y algunas constantes, hay que incluir:

```
#include <gsl/gsl_math.h>
```

Algunas constantes que están definidas en ese archivo son

`M_E`	e
`M_PI`	π
`M_PI_2`	$\pi/2$
`M_SQRTPI`	$\sqrt{\pi}$

Librería GSL (II)

Revisar los ejemplos

- `representacionIEEE.c` para ver la representación de punto flotante de un número.
- `serie.c` para modificar la forma en que se hace el redondeo (funciona en linux)
- `pruebaGsl.c` muestra que las funciones de la librería son robustas y eficientes.

En linux, para compilar desde la línea de comandos:

```
gcc -o pruebaGsl pruebaGsl.c -lgsl -lgslcblas -lm
```

En Code::Blocks, ver las instrucciones en la página del curso.

Unidad 2:

Solución de ecuaciones no lineales

Introducción (I)

Tenemos una función $f : D \longrightarrow \mathbb{R}$ continua, $D \subset \mathbb{R}$, y queremos determinar un valor x^* tal que

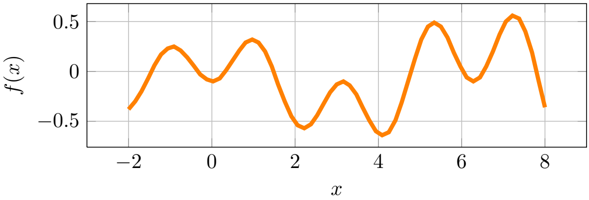
$$f(x^*) = 0.$$

Tal valor se dice que es un *raíz* o un *cero* de la función f. Entonces, dada la ecuación

$$\boxed{f(x) = 0,}$$

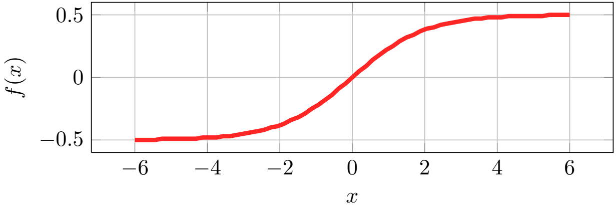
pueden ocurrir los siguientes casos:

Múltiples raíces

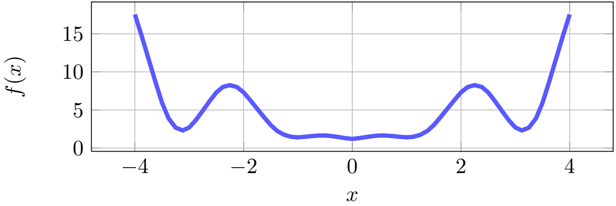


Introducción (II)

Una raíz única



Sin raíces



- En general, no hay fórmulas cerradas para calcular todas las raíces de una ecuación no lineal, por lo que hay que recurrir a algoritmos numéricos iterativos.

Introducción (III)

- Un caso particular son las ecuaciones polinomiales,

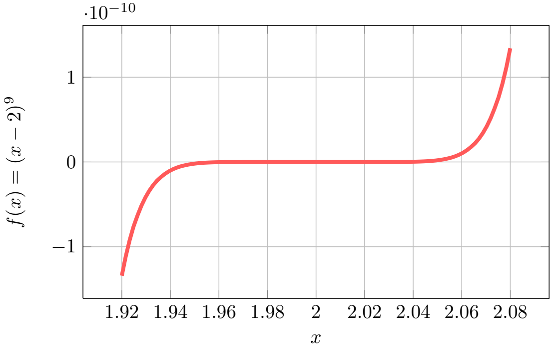
$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = \sum_{i=0}^{n} a_i x^i = 0, \qquad x, a_i \in \mathbb{R},$$

  tiene a lo más n raíces reales.
- Un problema que se tiene es la evaluación de las funciones usando aritmética de punto flotante.
- Tomando en cuenta la propagación de errores por la representación de los números en la computadora, es conveniente cambiar el problema y buscar un punto $\bar{x}$ tal que

$$|f(\bar{x})| < \text{tolerancia}$$

Ejemplo. 1 Sea $f(x) = (x-2)^9$. Evaluando esta función se obtiene la siguiente gráfica.

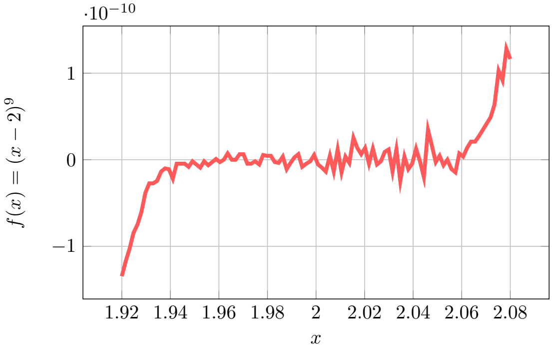
Introducción (IV)



Al generar la gráfica evaluando la expresión al desarrollar el binomio:

$$\begin{aligned} f(x) &= x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 \\ &\quad + 5376x^3 - 4608x^2 + 2304x - 512 \end{aligned}$$

Introducción (V)



Esquema general de los métodos iterativos para encontrar raíces

La función $f(x) : D \longrightarrow \mathbb{R}$ es conocida. Los métodos iterativos trabajan de la siguiente forma:

1. Se da un valor inicial $x_0 \in D$.
2. El algoritmo itera generando una sucesión de valores $x_1, x_2, \ldots, x_k, \ldots$
3. Si el algoritmo tiene éxito, se debe tener que

$$x_k \longrightarrow x^* \qquad \text{con} \quad f(x^*) = 0.$$

4. En la práctica, sólo se aplica un cierto número de iteraciones, de modo que en la iteración final n sólo se tiene

$$f(x_n) \approx 0.$$

5. Si se quiere obtener una raíz diferente, hay que dar otro punto inicial.

Esquema general de los métodos iterativos para encontrar raíces

- En general, los algoritmos sólo obtienen una raíz a la vez. Hay que aplicarlos con diferentes condiciones iniciales para recuperar diferentes raíces.
- Dada una ecuación no lineal, no sabemos cuantas raíces tiene en total.
- Por tanto, no sabemos cuantas veces hay que aplicar el algoritmo para recuperar todas las raíces.

Condiciones de paro del método iterativo

Las condiciones que se utilizan para parar al algoritmo iterativo pueden una combinación de las siguientes restricciones:

$$
\begin{aligned}
|x_n - x_{n-1}| &< \delta_x \\
|x_n - x_{n-1}| &< \delta_r(\epsilon + |x_n|) \\
|f(x_n)| &< \delta_f \\
n &< N_{\max}
\end{aligned}
$$

donde $\delta_x, \delta_r, \delta_f > 0$ son tolerancias dadas y $N_{\max}$ es el número máximo de iteraciones.

Estos valores dependen de la precisión en la computadora y del tipo de variables que se usan en el programa que calcula la solución.

Propiedades deseadas en los métodos iterativos

- *Eficiente* – que no requiera demasiadas evaluaciones de la función
- *Robusto* – que generalmente obtenga un resultado sin importar del punto inicial.
- *Requerimientos mínimos* – que se necesite pocos datos adicionales, como la derivada de la función.
- *Mínimos supuestos sobre f* – para que el algoritmo pueda ser aplicada a una gran variedad de funciones.
- *Generalizable* – para poder aplicarlo a ecuaciones con más variables.

Nota: Ningún algoritmo satisface todas estas condiciones.

Método de bisección (I)

Supongamos que tenemos un intervalo $[a, b]$ en el que se cumple que

$$f(a)f(b) < 0.$$

Si f es continua, debe haber un cero de f en $[a, b]$.

Calculamos el punto medio del intervalo:

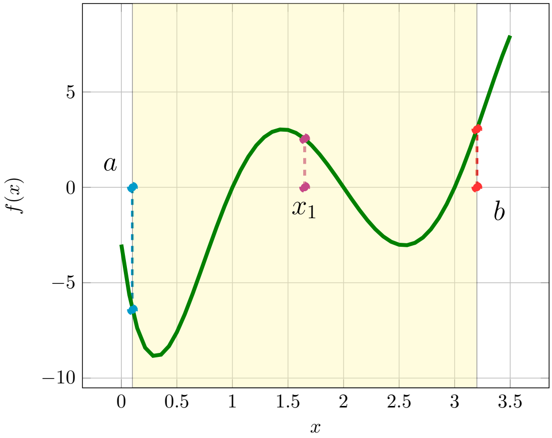
$$c = \frac{1}{2}(a + b)$$

Entonces tenemos los siguiente casos:

- Si $|f(c)| < \tau$, terminamos.
- Si $f(c)f(a) < 0$, entonces debe haber un cero de f en $[a, c]$. Reemplazamos el intervalo $[a, b]$ por $[a, c]$ y repetimos el proceso.
- Si $f(c)f(b) < 0$, entonces debe haber un cero de f en $[c, b]$ Reemplazamos el intervalo $[a, b]$ por $[c, b]$ y repetimos el proceso.

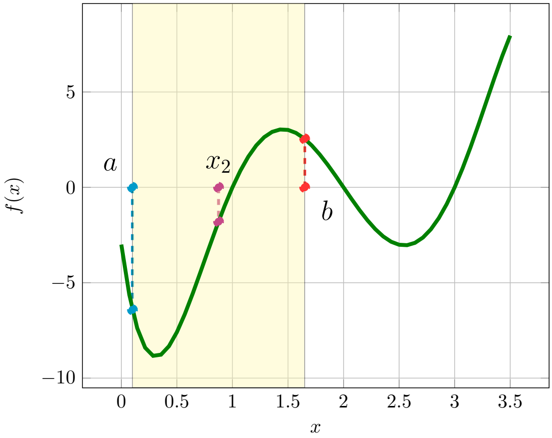
Método de bisección (II)

Ejemplo 2. Iteración 1



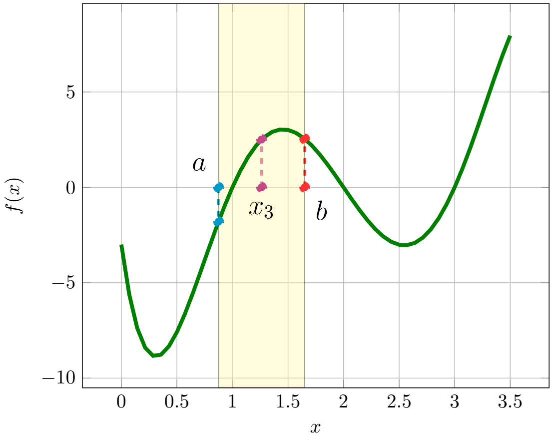
Método de bisección (III)

Ejemplo 2. Iteración 2



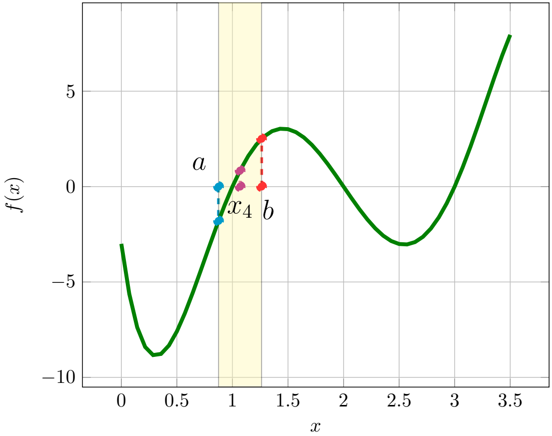
Método de bisección (IV)

Ejemplo 2. Iteración 3



Método de bisección (V)

Ejemplo 2. Iteración 4



Ejemplo del método de bisección (I)

Ejemplo 3. Sea

$$
\begin{aligned}
f(x) &= x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 \\
&\quad + 5376x^3 - 4608x^2 + 2304x - 512
\end{aligned}
$$

Tomando $a = -2$, $b = 8$ y una tolerancia $\tau = 10^{-14}$ se obtiene lo siguiente:

k	$\mathbf{x}_k$	$\mathbf{f}(x_k)$
1	3.000000	1.0000
2	0.500000	-38.4430
3	1.750000	-3.8147×10^{-6}
4	2.375000	1.4665×10^{-4}
5	2.062500	1.4552×10^{-11}
6	1.906250	-5.6025×10^{-10}
7	1.984375	0.0000

Métodos basados en la teoría de punto fijo (I)

El problema

$$f(x) = 0$$

puede ser escrito como

$$x = g(x).$$

En esta formulación, buscamos un *punto fijo*, es decir, un punto x^* tal que

$$x^* = g\left(x^*\right),$$

de modo que también se debería cumplir que

$$f\left(x^*\right) = 0.$$

De esta forma la función $g(x)$ no es arbitraria, y generalmente se define en términos de la función $f(x)$.

Ejemplo 4. Algunas opciones para la función g son las siguientes:

- $g(x) = x - f(x)$.
- $g(x) = x + 2f(x)$.
- $g(x) = x - \frac{f(x)}{f'(x)}$ si $f'(x) \neq 0$.

Algoritmo iterativo de punto fijo

Dada una función de una variable $f(x)$ y una función $g(x)$ tal que si $f(x) = 0$ si y sólo si $x = g(x)$, un punto inicial x_0, una tolencia τ, un número máximo de iteraciones N y $k = 0$. Hacer

1. Terminar si $|x_k - g(x_k)| < \tau$ o $k > N$
2. Si no, calcular $x_{k+1} = g(x_k)$.
3. Hacer $k \leftarrow k + 1$ y volver al paso 1.

Métodos basados en la teoría de punto fijo (III)

Este proceso genera una sucesión $x_1, x_2, ..., x_k, ...$ mediante

$$x_{k+1} = g(x_k), \qquad k = 0, 1, 2, \ldots$$

iniciando en un punto x_0. Si tal sucesión converge, debe hacerlo a un punto fijo de la ecuación.
Las propiedades de la sucesión dependen de la función g. Lo que queremos ver es

- ¿Bajo que condiciones converge la sucesión $\{x_k\}$ a un punto fijo x^*?
- Si converge, ¿el punto x^* es único?
- ¿Qué tan rápido converge?
- Si no converge, ¿esto significa que no existe una raíz del problema?

Teorema de punto fijo (I)

Teorema de punto fijo

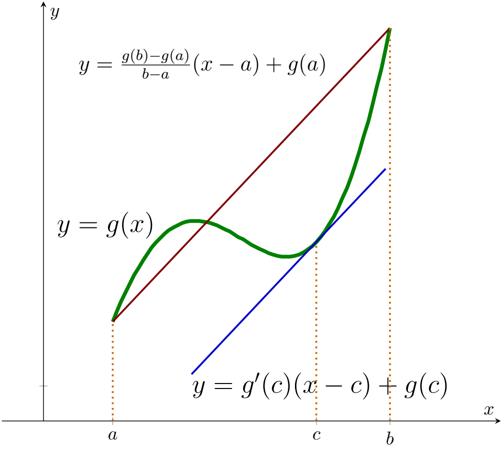
Si $g : [a, b] \to \mathbb{R}$ es una función continua y $a \leq g(x) \leq b$ para todo $x \in [a, b]$, entonces hay un punto fijo $x^* \in [a, b]$.
Además, si la derivada de g existe en el intervalo $[a, b]$ y hay una constante $0 < \rho < 1$ tal que

$$|g'(x)| \leq \rho \qquad \forall x \in (a, b),$$

entonces el punto fijo es único en este intervalo.

Teorema de punto fijo (II)

El teorema establece una condición suficiente. Para demostrarlo, hay que aplicar el teorema del valor intermedio y el teorema del valor medio.



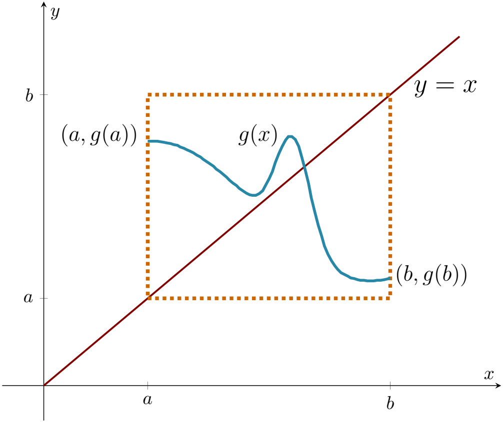
El teorema del valor medio garantiza la existencia de un punto c en el que la derivada es igual a la pendiente de la secante que une a $(a, g(a))$ y $(b, g(b))$:

$$g'(c) = \frac{g(b)-g(a)}{b-a}$$

o

$$g(b)-g(a) = g'(c)\,(b-a)$$

Teorema de punto fijo (III)



Si $g(a) = a$ o $g(b) = b$, el punto fijo existe.

Si $g(a) > a$ y $g(b) < b$, definimos

$$h(x) = g(x) - x.$$

El teorema del valor intermedio garantiza la existencia del punto fijo:

Puesto que h es continua $h(a) > 0$ y $h(b) < 0$, existe $x^* \in (a, b)$ tal que

$$h(x^*) = 0.$$

Teorema de punto fijo (IV)

Para la convergencia de la sucesión $\{x_k\}$ a x^*, hay que notar que

$$|x_{k+1} - x^*| = |g(x_k) - g(x^*)| = |g'(\xi_k)(x_k - x^*)| \leq \rho |x_k - x^*|$$

Así,

$$|x_{k+1} - x^*| \leq \rho |x_k - x^*| \leq \rho^2 |x_{k-1} - x^*| \leq \cdots \leq \rho^{k+1} |x_0 - x^*|$$

Como $\rho^k \to 0$, entonces

$$|x_k - x^*| \to 0 \qquad \text{i.e.} \qquad x_k \to x^*.$$

Ejemplo (I)

Considere la función $f(x) = 2xe^x - 1$.

Si definimos $g_1(x) = x - f(x)$, al aplicar el algoritmo de punto fijo con una tolerancia $\tau = \sqrt{\epsilon_m}$ se obtiene:

k	x_k	$g_1(x_k)$	$\lvert x_k - g_1(x_k)\rvert$
1	0.5000	-0.1487	0.65
2	-0.1487	1.1076	1.26
3	1.1076	-4.5982	5.71
4	-4.5982	-3.5056	1.09
5	-3.5056	-2.2951	1.21
6	-2.2951	-0.8326	1.46
7	-0.8326	0.8916	1.72
8	0.8916	-2.4580	3.35
9	-2.4580	-1.0372	1.42
10	-1.0372	0.6981	1.74
11	0.6981	-1.1080	1.81
12	-1.1080	0.6237	1.73
13	0.6237	-0.7040	1.33
14	-0.7040	0.9924	1.70
15	0.9924	-3.3623	4.35
16	-3.3623	-2.1292	1.23
17	-2.1292	-0.6228	1.51
18	-0.6228	1.0454	1.67
19	1.0454	-3.9020	4.95
20	-3.9020	-2.7444	1.16

Ejemplo (II)

En cambio, si se define $g_2(x) = x - \frac{1}{2}e^{-x}f(x) = \frac{1}{2}e^{-x}$.

k	x_k	$g_2(x_k)$	$\lvert x_k - g_1(x_k)\rvert$
1	0.50000000	0.30326533	1.97e-01
2	0.30326533	0.36920157	6.59e-02
3	0.36920157	0.34564303	2.36e-02
4	0.34564303	0.35388255	8.24e-03
5	0.35388255	0.35097870	2.90e-03
6	0.35097870	0.35199937	1.02e-03
7	0.35199937	0.35164028	3.59e-04
8	0.35164028	0.35176658	1.26e-04
9	0.35176658	0.35172215	4.44e-05
10	0.35172215	0.35173778	1.56e-05
11	0.35173778	0.35173228	5.50e-06
12	0.35173228	0.35173421	1.93e-06
13	0.35173421	0.35173353	6.80e-07
14	0.35173353	0.35173377	2.39e-07
15	0.35173377	0.35173369	8.41e-08
16	0.35173369	0.35173372	2.96e-08
17	0.35173372	0.35173371	1.04e-08

El valor de la función f en el punto fijo es

$$f(x_{17}) \approx 2.96 \times 10^{-8}$$

Método de Newton-Raphson (I)

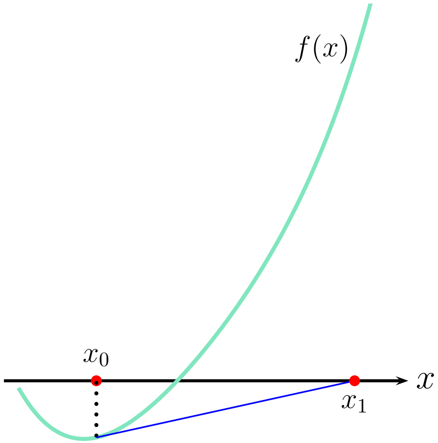
Se da un punto x_0. Si $f(x_0) \neq 0$, consideramos la aproximación de la función f

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

Seleccionamos x_1 de modo que sea un cero de la función $m(x)$. Así, si $m(x_1) = 0$, entonces

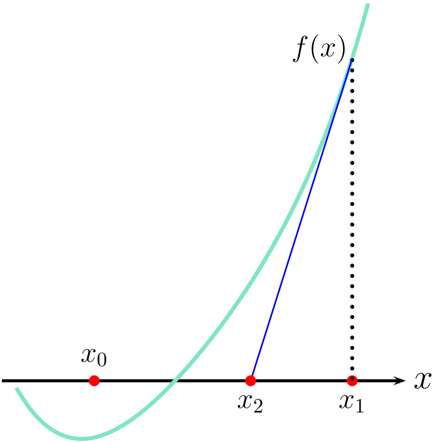
$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Método de Newton-Raphson (II)



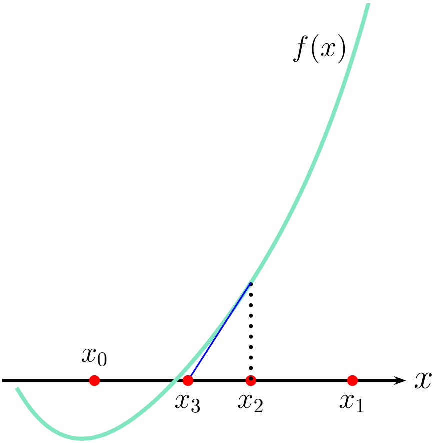
$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Método de Newton-Raphson (III)



$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

Método de Newton-Raphson (IV)



$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$$

Método de Newton-Raphson (V)

Si $|f(x_1)| < \epsilon$, terminamos. Si no, continuamos. En general, definimos la secuencia

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

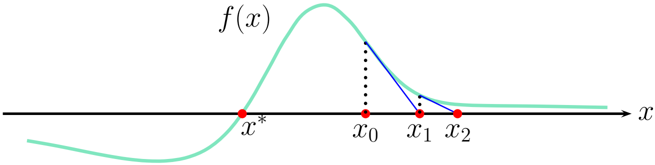
Terminamos cuando $|f(x_{k+1})| < \epsilon$.

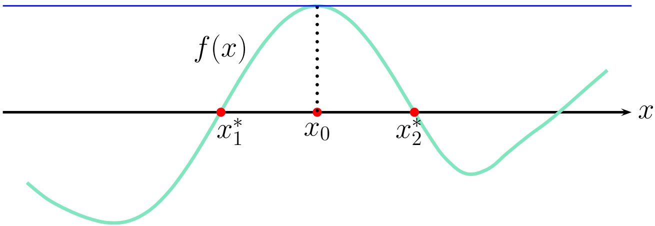
Observaciones:

- El punto al que converge el algoritmo depende del valor inicial x_0.
- El método funciona bien para raíces simples.
- Cerca de la raíz, el método tiene convergencia cuadrática.
- Puede que no converja, dependiendo de la inicialización,

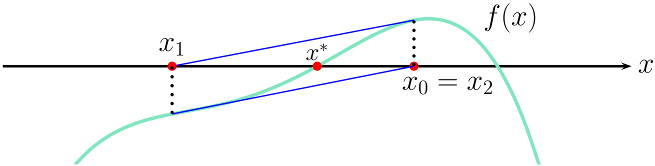
Los siguientes casos ilustran algunos problemas del método de Newton-Raphson.

Método de Newton-Raphson (VI)





Método de Newton-Raphson (VII)



Las aritmética de punto flotante en la computadora puede hacer que este último caso no ocurra tan fácilmente.

Ejemplo 1. Sea $f(x) = 2x^3 - 3x - 6$. Queremos hallar una raíz de $f(x) = 0$. Aplicando Newton-Raphson:

k	$\mathbf{x}_k$	$\mathbf{f}(x_k)$
1	1.000000	-7.00000000
2	3.333333	58.07407407
3	2.421175	15.12276195
4	1.951123	3.00202004
5	1.799821	0.26106567
6	1.783938	0.00271644
7	1.783769	0.00000030
8	1.783769	2.6×10^{-15}

Ejemplo 2 (I)

Aplicamos el método de Newton-Raphson para encontrar la raíz de $f(x) = (x-2)^9 = x^9 - 18\,x^8 + 144\,x^7 - 672\,x^6 + 2016\,x^5 - 4032\,x^4 + 5376\,x^3 - 4608\,x^2 + 2304\,x - 512$.

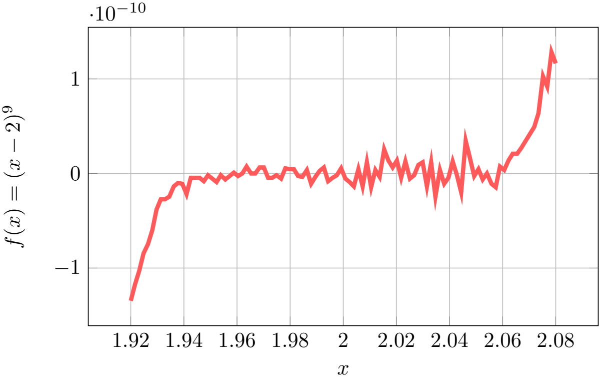
Nota: Sabemos que no se debería aplicar este método a esta función, pero queremos ver como se comporta este método.

Para hacer esto, requerimos evaluar la derivada de la función:

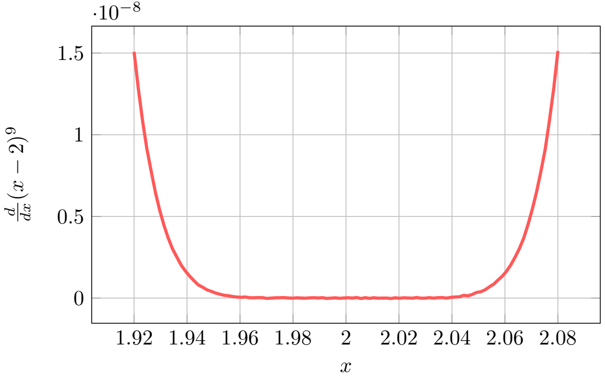
$$
\begin{aligned}
f'(x) &= 9x^8 - 144\,x^7 + 1008\,x^6 - 4032\,x^5 + 10080\,x^4 \\
&\quad -16128\,x^3 + 16128\,x^2 - 9216\,x + 2304
\end{aligned}
$$

Las siguientes figuras muestran las gráficas de la función $f(x)$ y su derivada.

Ejemplo 2 (II)

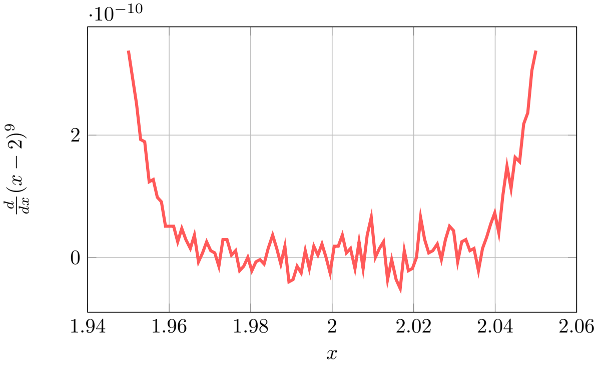


Ejemplo 2 **(III)**



Ejemplo 2 (IV)

Gráfica de la derivada de $f(x)$ más cerca del punto $x = 2$.



Aplicando el método de Newton-Raphson a esta función se obtiene:

Ejemplo 2 (V)

k	x_k	$f(x_k)$	$f'(x_k)$	x_{k+1}
1	10.0000	1.3422e+08	1.5099e+08	9.1111
2	9.1111	4.6498e+07	5.8849e+07	8.3210
3	8.3210	1.6109e+07	2.2936e+07	7.6187
4	7.6187	5.5807e+06	8.9393e+06	6.9944
5	6.9944	1.9334e+06	3.4840e+06	6.4394
6	6.4394	6.6980e+05	1.3579e+06	5.9462
7	5.9462	2.3205e+05	5.2923e+05	5.5077
8	5.5077	8.0390e+04	2.0626e+05	5.1180
9	5.1180	2.7850e+04	8.0390e+04	4.7715
10	4.7715	9.6484e+03	3.1331e+04	4.4636
11	4.4636	3.3426e+03	1.2211e+04	4.1898
12	4.1898	1.1580e+03	4.7593e+03	3.9465
13	3.9465	4.0118e+02	1.8549e+03	3.7302
14	3.7302	1.3898e+02	7.2294e+02	3.5380
15	3.5380	4.8150e+01	2.8176e+02	3.3671
16	3.3671	1.6681e+01	1.0981e+02	3.2152
17	3.2152	5.7789e+00	4.2800e+01	3.0802
18	3.0802	2.0020e+00	1.6681e+01	2.9602

Ejemplo 2 (VI)

k	x_k	$f(x_k)$	$f'(x_k)$	x_{k+1}
19	2.9602	6.9359e-01	6.5013e+00	2.8535
20	2.8535	2.4029e-01	2.5338e+00	2.7586
21	2.7586	8.3245e-02	9.8755e-01	2.6744
22	2.6744	2.8839e-02	3.8489e-01	2.5994
23	2.5994	9.9910e-03	1.5001e-01	2.5328
24	2.5328	3.4613e-03	5.8465e-02	2.4736
25	2.4736	1.1991e-03	2.2787e-02	2.4210
26	2.4210	4.1542e-04	8.8809e-03	2.3742
27	2.3742	1.4392e-04	3.4613e-03	2.3326
28	2.3326	4.9859e-05	1.3490e-03	2.2957
29	2.2957	1.7273e-05	5.2577e-04	2.2628
30	2.2628	5.9842e-06	2.0492e-04	2.2336
31	2.2336	2.0731e-06	7.9865e-05	2.2077
32	2.2077	7.1822e-07	3.1127e-05	2.1846
33	2.1846	2.4882e-07	1.2131e-05	2.1641
34	2.1641	8.6196e-08	4.7280e-06	2.1458
35	2.1458	2.9872e-08	1.8427e-06	2.1296
36	2.1296	1.0351e-08	7.1797e-07	2.1152

Ejemplo 2 (VII)

k	x_k	$f(x_k)$	$f'(x_k)$	x_{k+1}
37	2.1152	3.5761e-09	2.7957e-07	2.1024
38	2.1024	1.2442e-09	1.0905e-07	2.0910
39	2.0910	4.3292e-10	4.2419e-08	2.0808
40	2.0808	1.6098e-10	1.6353e-08	2.0710
41	2.0710	5.5479e-11	5.7880e-09	2.0614
42	2.0614	-1.8190e-12	1.8263e-09	2.0624

Se puede ver que se genera una secuencia decreciente de puntos que tiende hacia la raíz y se detiene cuando el valor de $|f(x_k)|$ es menor que la tolerancia.