Resolución numérica de ecuaciones no lineales

Javier Segura

Cálculo Numérico I. Tema 2.

Contenidos:

1. **Métodos básicos**
   - Bisección
   - Método de la secante
   - El método de Newton
2. **Métodos de punto fijo**
3. **Raices de polinomios**
   - Algunos resultados de acotación de ceros de polinomios
   - El método de Newton para polinomios
   - Ceros complejos y el método de Müller
   - Sobre la estabilidad de los ceros de un polinomio

Estructura de la presentación:

1. **Métodos básicos**
   - Bisección
   - Método de la secante
   - El método de Newton
2. **Métodos de punto fijo**
3. **Raices de polinomios**
   - Algunos resultados de acotación de ceros de polinomios
   - El método de Newton para polinomios
   - Ceros complejos y el método de Müller
   - Sobre la estabilidad de los ceros de un polinomio

Bisección

Teorema (Bolzano)

.
Si $f(x)$ es continua en el intervalo $[a, b]$ con $f(a)f(b) \leq 0$, entonces existe al menos un $\alpha \in [a, b]$ tal que $f(\alpha) = 0$.

Algoritmo (Bisección)

Algoritmo: Bisección en un intervalo [a,b], tal que $f(a)f(b) < 0$

(1) *Sea* $c = (b + a)/2$

(2) *Si* $b - c \leq \epsilon$, *aceptar c; parar*

(3) *Si* $f(b)f(c) \leq 0$, *tomar* $a = c$; *y si no:* $b = c$.

(4) *Volver a* **(1)**

1. Bisección es lo óptimo para resolver una ecuación $f(x) = 0$ cuando no sabemos nada de f salvo calcular su signo.
2. El cero está siempre dentro de un intervalo de acotación y converge siempre que f cambie de signo, aunque también puede converger a una discontinuidad.
3. Lento aunque seguro: a lo sumo, el orden de convergencia es 1
4. Inespecífico: el número de iteraciones sólo depende del intervalo $[a, b]$ y del cero $\alpha \in [a, b]$.

Orden de convergencia

Sea $\{x_n\}$ una sucesión que converge a α y definamos $\epsilon_n = x_n - \alpha$. Si exite un número p, una constante $C \neq 0$ y un $n_0 \in \mathbb{N}$ tales que para todo $n \geq n_0$

$$|\epsilon_{n+1}| \leq C|\epsilon_n|^p$$

y p es el máximo valor para el que se verifica tal relación, entonces diremos que p es el orden de convergencia.
En particular, si

$$\lim_{n\to\infty} \frac{|\epsilon_{n+1}|}{|\epsilon_n|^p} = C$$

el orden de convergencia es p, siendo C la constante de error asintótica.

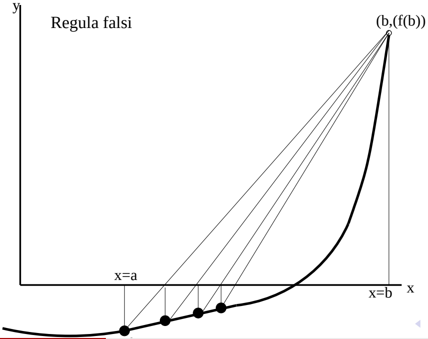
Regula falsi

Regula falsi es un intento de mejorar bisección estimando c así:

$$c = b - \frac{b-a}{f(b)-f(a)} f(b) = \frac{af(b)-bf(a)}{f(b)-f(a)}$$

Como en bisección, nos quedamos en cada iteración con el intervalo donde cambia de signo f.

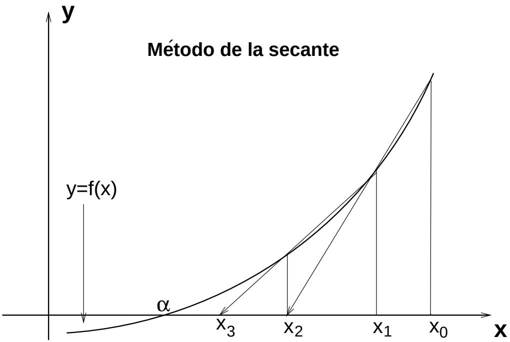
El intervalo de acotación puede ser peor que para bisección:



Método de la secante

Igual que regula falsi, pero sin tener en cuenta los signos (se rechaza x_{n-1} una vez calculado x_n):

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n) \quad (1)$$



Orden de convergencia:
Con $\epsilon_k = x_k - \alpha$, desarrollando $f(x_k) = f(\alpha + \epsilon_k)$ y llevándolo a (1), tenemos que

$$\epsilon_{n+1} = \frac{f''(\alpha)}{2f'(\alpha)}\epsilon_n \epsilon_{n-1}$$

Con esto, se puede comprobar que el orden es $p = (1 + \sqrt{5})/2$
(Vamos mejorando, si es que converge!)

Criterio de parada: como el orden es $p > 1$:

$$\frac{|x_{n+1} - x_n|}{|x_n - \alpha|} \leq 1 + \frac{|\epsilon_{n+1}|}{|\epsilon_n|} \leq 1 + C|\epsilon_n|^{p-1} \to 1$$

cuando $n \to +\infty$. Se puede entonces utilizar $|x_{n+1} - x_n|$ para estimar el error.

Recordemos que con el método de la secante no tenemos intervalos de acotación de la raíz.

Cómo garantizar que el método converge?

Cómo escoger los valores iniciales?

Existen algoritmos que combinan bisección y secante que:

1. Garantizan la convergencia y dan un intervalo de acotación
2. Convergen más rápidamente que bisección

Los métodos de Dekker (bisección + secante) y Brent (bisección + interpolación cuadrática inversa) son dos ejemplos populares (pero quedan fuera de los objetivos de este curso).

El método de Newton

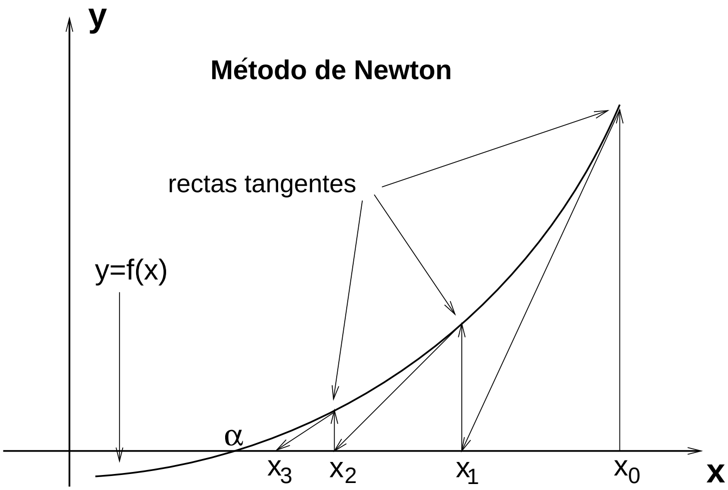
El método de Newton consiste en resolver la ecuación $f(x) = 0$ invirtiendo la aproximación lineal de Taylor alrededor de $x = \alpha$ ($f(\alpha) = 0$)

$$0 = f(x_n - \epsilon_n) = f(x_n) - f'(x_n)\epsilon_n + \frac{f''(\zeta_n)}{2}\epsilon_n^2$$

donde $\epsilon_n = x_n - \alpha$
Despreciando el error

$$\alpha \simeq x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Orden de convergencia = 2
Parece buena idea caso de que se disponga de la derivada $f'(x)$


y
Método de Newton
rectas tangentes
y=f(x)
α
x3
x2
x1
x0
x

Convergencia local

Tanto el método de Newton como el de la secante son convergentes localmente.

Teorema (Convergencia local del método de Newton)

Si g', con $g(x) = x - f(x)/f'(x)$, es continua en un entorno de α ($f(\alpha) = 0$) entonces existe un entorno de α tal que el método de Newton converge para todo x_0 en este entorno.

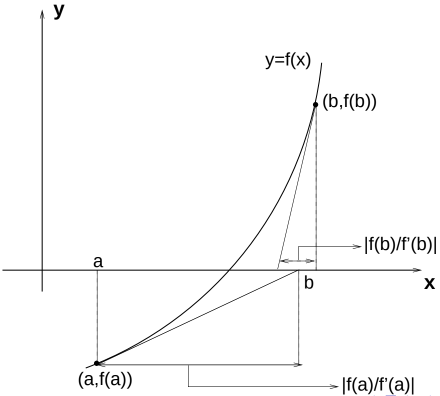
Teorema (Convergencia local del método de la secante)

Si f'' es continua en un entorno de α siendo $f(\alpha) = 0$ y $f'(\alpha) \neq 0$ entonces existe un entorno de α para el que el método de la secante converge para cualesquiera valores iniciales x_0, x_1 en ese entorno.

Un resultado de convergencia en un intervalo:
Si $f(x)$ es tal que $f(a)f(b) < 0$ y los dos primeras derivadas son continuas y no cambian de signo en $[a, b]$ y además

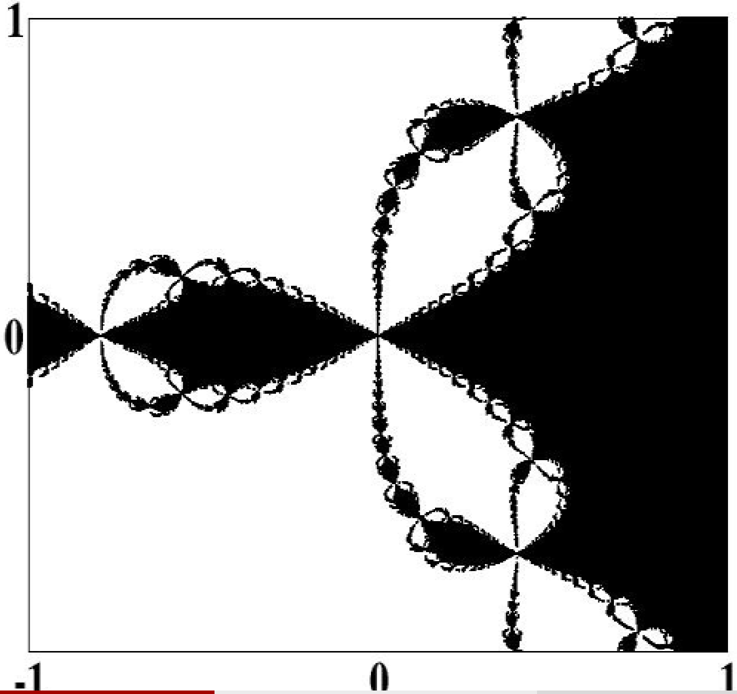
$$\left|\frac{f(a)}{f'(a)}\right| < b - a, \; \left|\frac{f(b)}{f'(b)}\right| < b - a$$

entonces el método de Newton converge $\forall x_0 \in [a, b]$.



En el plano complejo los problemas aumentan

El método de Newton también se puede utilizar para calcular ceros en el plano complejo.
Supongamos que queremos resolver $f(z) = 0$ donde $f(z) = z^3 - 1$.
Tenemos $z_{n+1} = z_n - \dfrac{z_n^3 - 1}{3z_n^2}$. El conjunto de los valores iniciales z_0 que dan convergencia a la raíz $z = 1$ es...


1
0
-1
0
1

Estructura de la presentación:

Métodos de punto fijo

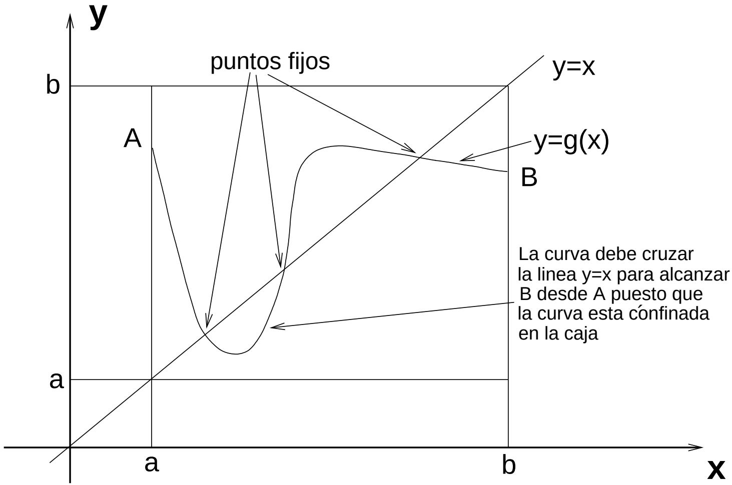
Los métodos de punto fijo responden a esquemas

$$x_{n+1} = g(x_n)$$

Ejemplo: el método de Newton, g(x)=x-f(x)/f'(x)
Si g is continua y, dado cierto x_0, existe $\lim_{n\to\infty} x_n = \alpha$ entonces $\alpha = g(\alpha)$ (α punto fijo de T).

Teorema

Si $g(x)$ es continua en $I = [a, b]$ y $g(I) \subseteq I$ entonces existe al menos un punto fijo de $g(x)$, $\alpha \in [a, b]$.


y
puntos fijos
y=x
b
A
y=g(x)
B
La curva debe cruzar
la linea y=x para alcanzar
B desde A puesto que
la curva esta confinada
en la caja
a
a
b
x

Teorema (Continuación)

Si, además $g(x)$ es diferenciable en (a, b) y

$$|g'(x)| \leq M < 1 \,, \quad \forall x \in (a, b) \,,$$

entonces

1. **$g(x)$ tiene un único punto fijo α en $[a, b]$**
2. **Para cada $x_0 \in [a, b]$, la sucesión definida como**

$$x_{n+1} = g(x_n)$$

   converge al único punto fijo α.
3. **La estimación del error de la enésima iteración está dada por**

$$|x_n - \alpha| \leq \frac{M^n}{1 - M} |x_1 - x_0|$$

4. **Si $g'(\alpha) \neq 0$, $a_n \neq \alpha \; \forall n$ y g' es continua en α entonces**

$$\lim_{n \to \infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = g'(\alpha).$$

   Es decir, que si $g'(\alpha) \neq 0$ el orden de convergencia es 1.

Observación: si $g'(\alpha) = 0$ el orden puede ser mayor. En efecto:

Dado un método de punto fijo $x_{n+1} = g(x_n)$ y $g(\alpha) = \alpha$, si la primera derivada no nula en α de $g(x)$ es la m-ésima (y es continua), entonces el método tiene orden de convergencia m con constante de error asintótica $C = g^{(m)}(\alpha)/m!$.
En efecto:

$$x_{n+1} = g(\alpha + \epsilon_n) = g(\alpha) + \frac{g^{(m)}(\zeta_n)}{m!}\epsilon_n^m$$

Luego

$$\epsilon_{n+1} = \frac{g^{(m)}(\zeta_n)}{m!}\epsilon_n^m$$

con ζ_n entre α y x_n, luego

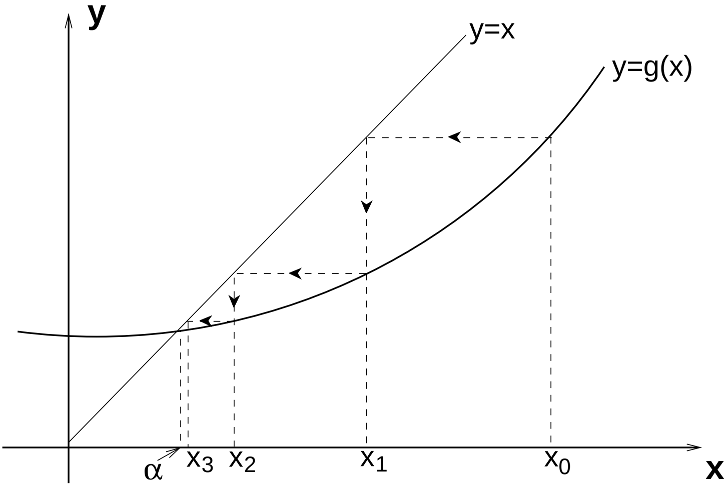
$$\lim_{n\to\infty} \frac{\epsilon_{n+1}}{\epsilon_n^m} = g^{(m)}(\alpha)/m!$$

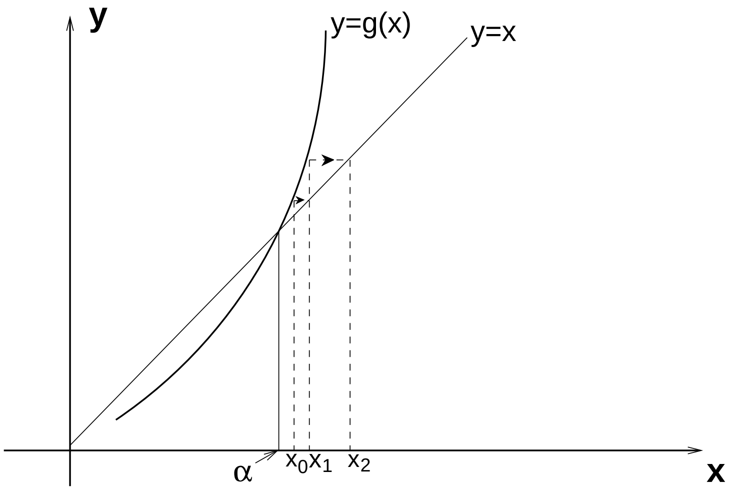
A continuación damos un resultado en el que no es necesario encontrar un cerrado I tal que $g(I) \subseteq I$.

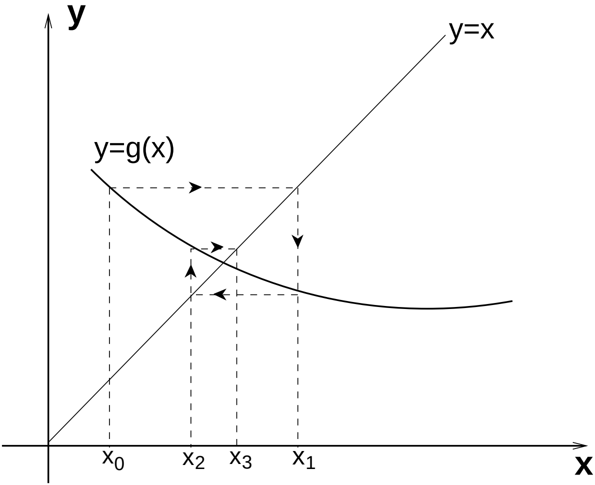
Teorema (Convergencia monótona)

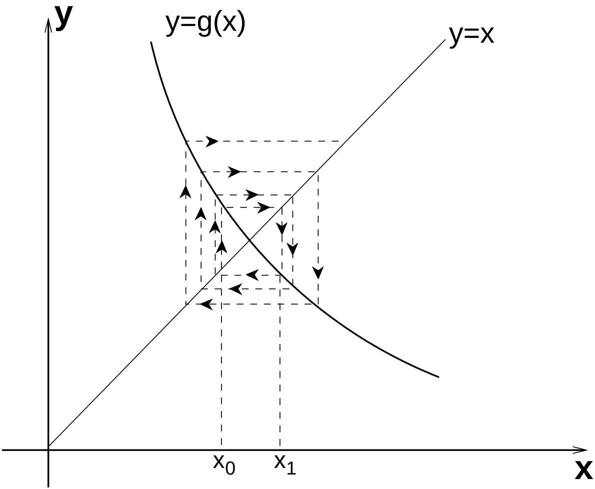
Si $g(x)$ tiene derivada primera en un intervalo I que contiene un punto fijo de $g(x)$ y $0 \leq g'(x) < 1$ en I, entonces hay sólo un punto fijo de $g(x)$ en I, la iteración de punto fijo $x_{n+1} = g(x_n)$ es convergente $\forall x_0 \in I$ y la convergencia es de tipo monótono, es decir:
Si $x_0 > \alpha$, entonces $x_0 > x_1 > \ldots > x_n > \ldots > \alpha$
Si $x_0 < \alpha$, entonces $x_0 < x_1 < \ldots < x_n < \ldots < \alpha$

La demostración a partir del teorema del valor medio. Interpretación gráfica en la próxima página. Las siguientes tres describen otras situaciones (no todas convergentes).


y
y=x
y=g(x)
x
α
x3
x2
x1
x0


y
y=g(x)
y=x
α
x0
x1
x2
x


y
y=x
y=g(x)
x
x_0
x_2
x_3
x_1


y
y=g(x)
y=x
x0
x1
x

Teorema (Convergencia local)

Sea $g(x)$ derivable en un entorno U de α siendo $g(\alpha) = \alpha$ y cumpliéndose que $|g'(x)| \leq M < 1$ en U. Sea $I = [\alpha - \epsilon, \alpha + \epsilon]$ tal que $I \subset U$. Entonces el anterior teorema se verifica en I y la iteración $x_{n+1} = g(x_n)$ converge a α para todo $x_0 \in I$.

Como Corolario, considerando el método de Newton ($g(x) = x - f(x)/f'(x)$), si α es una raíz simple de $f(x)$ ($f(\alpha) = 0$, $f'(\alpha) \neq 0$), y $f''(x)$ es continua en α, entonces existe un entorno de α tal que el método converge para todo x_0 en este entorno.

Estructura de la presentación:

Raices de polinomios: algunas razones para estudiar este caso en particular

1. Es un caso importante.
2. Es importante explicar como calcular valores de los polinomios y derviadas de forma eficiente.
3. Existen una serie de resultados que sirven para acotar la posición de los ceros.
4. Existen métodos específicos.

Teorema (Teorema fundamental del álgebra)

Un polinomio de grado n, $P_n(x) = a_n x^n + ... + a_1 x + a_0$, $a_n \neq 0$, tiene exactamente n raíces (sean puramente reales o complejas) contando con su multiplicidad.

Teorema

Si los coeficientes del polinomio a_i, $1 \leq i \leq n$, son todos reales, el número de raíces complejas es par y aparecen en pares complejo-conjugados.

Teorema (Regla de los signos de Descartes)

Dado un polinomio $P(x) = a_n x^n + ... + a_1 x + a_0$ con coeficientes reales, el número de raíces reales positivas de la ecuación $P(x) = 0$, que denotamos como N_p, está acotado por el número de cambios de signo en los coeficientes (ordenados) del polinomio (V), de tal forma que $V - N_p$ puede ser 0 o un número positivo y par. Los coeficientes nulos no cuentan.

Ejemplos:

1. $P(x) = x^4 + 2x^2 - x - 1$. Como sólo hay un cambio de signo hay a lo sumo un cero positivo, pero como $V - N_p = 0, 2, ...$, necesariamente $N_p = 1$.
   Por otra parte $Q(x) = P(-x) = x^4 + 2x^2 + x - 1$ también tiene sólo un cero positivo.
   Por lo tanto, $P(x)$ tiene una raíz positiva y otra negativa.
2. $P(x) = x^4 - x^3 - x^2 + x - 1$.
   Hay tres cambios de signo, luego puede haber una o tres raíces positivas. Por otra parte $P(-x) = x^4 + x^3 - x^2 - x - 1$, que sólo tiene un cambio de signo, luego $P(x)$ tiene una única raíz negativa.
   Dos posibilidades: que hayan dos raíces reales (una positiva y otra negativa) y dos complejo-conjugadas o bien que hayan cuatro raíces reales, una negativa y tres positivas. La situación real es la primera.

Teorema

Dado $P(x) = a_n x^n + \dots + a_1 x + a_0$, $a_n \neq 0$, entonces $P(x)$ tiene al menos un cero (real o complejo) verificando $|x| \leq \min\{\rho_1, \rho_n\}$, siendo
$\rho_1 = n\dfrac{|a_0|}{|a_1|}$, $\rho_n = \sqrt[n]{\left|\dfrac{a_0}{a_n}\right|}$

Teorema

Sea R el cero positivo de $P(x) = |a_n|x^n - |a_{n-1}|x^{n-1} - \dots - |a_0|$ y r el cero positivo de $Q(x) = |a_n|x^n + |a_{n-1}| + \dots + |a_1|x - |a_0|$, entonces todos los ceros de $a_n x^n + \dots + a_1 x + a_0 = 0$ verifican que $r \leq |x| \leq R$

Teorema

Dado $P(x) = a_n x^n + \dots + a_0$, $a_n \neq 0$, todos los ceros de $P(x)$ satisfacen $|x| \leq r$, donde $r = 1 + \max_{0 \leq k \leq n-1} \left|\dfrac{a_k}{a_n}\right|$.

Por otra parte, es importante tener en cuenta que sólo necesitaremos calcular ceros en el intervalo $[0, 1]$. Si $P(x)$ es de grado n, los ceros mayores que 1 de $P(x)$ son los ceros en $(0, 1)$ de $x^n P(1/x)$ (polinomio recíproco).

El método de Newton para polinomios

Vamos a ver como evaluar eficientemente polinomios y su derivada, y lo aplicaremos al método de Newton.
Para evaluar un polinomio lo más eficiente es el método de Horner (es decir, Ruffini)

$$\begin{array}{c|cccc} & a_n & a_{n-1} & \dots & a_0 \\ z & & & & \\ \hline & b_n & b_{n-1} & \dots & b_0 \end{array}$$

así que tenemos que

$$P(x) = (x - z)Q(x) + b_0, b_0 = P(z) \qquad (2)$$

donde

$$Q(x) = b_n x^{n-1} + \dots + b_2 x + b_1$$

Además $P'(z) = Q(z)$, con lo que a la vez que calculamos $P(z)$ podemos calcular $P'(z)$. De esta forma, podemos escribir el siguiente algoritmo

Algoritmo de Newton para polinomios

Input: x_0 (valor inicial), ϵ tolerancia de error relativo, a_n, ..., a_0
Output: x_i, iteraciones del método de Newton

er $= \epsilon + 1$; $i = 0$; $z = x_0$
Repetir mientras er $> \epsilon$
    $b_n = a_n$; $c_n = b_n$
    Desde $k = n - 1$ hasta 1; paso -1
        $b_k = a_k + z b_{k+1}$ (Horner para $P(z)$)
        $c_k = b_k + z c_{k+1}$ (Horner para $P'(z)$)
    Fin bucle
    $b_0 = a_0 + z b_1$
    $i = i + 1$; $x_i = z - b_0/c_1$; er $= \left|1 - \frac{z}{x_i}\right|$; $z = x_i$
Fin bucle

Si calculamos una raíz α de $P(x)$ el algoritmo proporciona una aproximación del polinomio cociente $Q(x) = P(x)/(x-\alpha)$.
En efecto, si el algoritmo se detiene en la n-ésima iteración tras obtenerse α con la precisión relativa requerida, tenemos que

$$P(x) = b_0 + (x - x_n)(b_1 + ... + b_n x^{n-1})$$

y como $b_0 = P(x_n) \approx P(\alpha) = 0$ pues $x_n \approx \alpha$, entonces

$$Q(x) \approx \frac{P(x)}{x - x_n} \approx \frac{P(x)}{x - \alpha}$$

Entonces, si α es un cero simple de $P(x)$, $Q(x)$ será un polinomio en el que α ya no es una raíz.
Este proceso se conoce como **deflación**.

Método de Müller: iterativo, convergencia cercana a cuadrática, no precisa del cálculo de derivadas y sirve para raices complejas.
Dada la función $f(x)$ se traza la parábola que pasa por los tres puntos $(x_{n-2}, f(x_{n-2}))$, $(x_{n-1}, f(x_{n-1}))$ y $(x_n, f(x_n))$. x_{n+1} es el punto de corte de esta parábola con el eje X que esté más cercano a la anterior iteración. Si escribimos la parábola como
$P(x) = a(x - x_n)^2 + b(x - x_n) + c$ los coeficientes son:

$$c = f(x_n)$$

$$b = \frac{(x_n - x_{n-2})^2(f(x_n) - f(x_{n-1})) - (x_n - x_{n-1})^2(f(x_n) - f(x_{n-2}))}{(x_n - x_{n-1})(x_n - x_{n-2})(x_{n-1} - x_{n-2})}$$

$$a = \frac{(x_n - x_{n-2})(f(x_n) - f(x_{n-1})) - (x_n - x_{n-1})(f(x_n) - f(x_{n-2}))}{(x_n - x_{n-1})(x_n - x_{n-2})(x_{n-1} - x_{n-2})}$$

y

$$x_{n+1} = x_n + \frac{2c}{-b - \mathrm{signo}(b)\sqrt{b^2 - 4ac}}$$

Los ceros de un polinomio pueden depender críticamente de los valores de los coeficientes.

Consideremos que en la evaluación de los ceros de $P(x) = (x-1)^{10}$ los coeficientes de $P(x)$ se han visto perturbados numéricamente de la siguiente forma

$$P_\epsilon(x) = P(x) - \epsilon$$

$\epsilon = 10^{-10}$. Veremos que esta pequeña perturbación produce soluciones considerablemente perturbadas.
$P_\epsilon(x) = 0$ tiene soluciones:

$$z_k = 1 + e^{ik\pi/5}\sqrt[10]{\epsilon},\ k = 0, ..., 9$$

y la variación en los valores de los ceros es, en todos los casos:

$$|z_i - x_i| = 0.1$$

lo que es un error importante. Además, sólo una raíz de $P_\epsilon(x)$ sigue siendo real.