

4. Ecuaciones No Lineales

4.1 Introducción

En general no es posible obtener las raíces de una ecuación no lineal $f(x)=0$ en forma explícita, debiéndose utilizar métodos iterativos. Partiendo de una raíz aproximada, x_0 , se obtiene una secuencia $x_1, x_2, x_3 \dots$ que converge a la raíz deseada. Para algunos métodos es suficiente conocer el intervalo $[a, b]$ en que se halla la raíz; otros procedimientos, de convergencia más rápida, requieren una aproximación inicial cercana a la raíz. Puede ser conveniente empezar los cálculos con un método del primer tipo y cambiar a un método de convergencia más rápida en la etapa final.

La primera parte de este capítulo considera el caso en que la raíz, \bar{x} , es una raíz simple, es decir, $f'(\bar{x}) \neq 0$. Las dificultades que se presentan en el caso de raíces múltiples se discuten en la sección 4.6. La sección 4.7 revisa métodos específicos para extraer "ceros" (raíces) de polinomios.

La parte final da algunas ideas para la solución de sistemas de ecuaciones no lineales, un problema que ciertamente puede demandar mucho esfuerzo de cómputo.

4.2 Aproximaciones Iniciales

Pueden obtenerse aproximaciones iniciales a las raíces de $f(x)=0$ graficando o tabulando la función.

Si $f(a_0) \cdot f(b_0) < 0$, hay por lo menos una raíz, \bar{x} , en el intervalo (a_0, b_0) . En el método de Bisección se definen una serie de intervalos $(a_0, b_0) \supset (a_1, b_1) \supset (a_2, b_2) \dots$

El punto medio de un intervalo (a_i, b_i) es $x_i = (a_i + b_i) / 2$. Suponiendo que $f(x) \neq 0$ (si este no es el caso, se ha hallado la raíz) se define el sub-intervalo (a_{i+1}, b_{i+1}) mediante:

$$(a_{i+1}, b_{i+1}) = \begin{cases} (x_i, b_i) & \text{si } f(x_i) \cdot f(b_i) < 0 \\ (a_i, x_i) & f(a_i) \cdot f(x_i) < 0 \end{cases}$$

Introduciendo la notación $\varepsilon_n = x_n - \bar{x}$, se tiene que $\varepsilon_{n+1} = O(\frac{1}{2}\varepsilon_n)$. Como $10^{-1} \approx 2^{-3.3}$, se requieren 3 ó 4 pasos para mejorar un dígito decimal en la aproximación. La convergencia es prácticamente independiente de $f(x)$.

Por ejemplo, para la función:

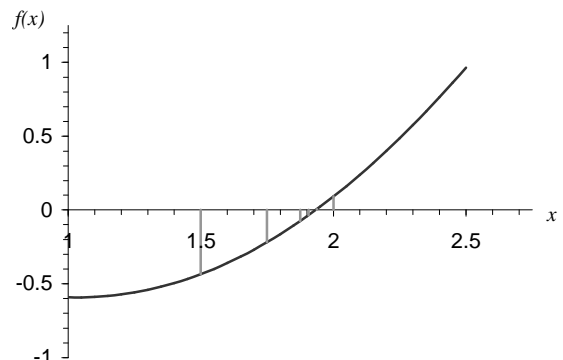
$$f(x) = \frac{x^2}{4} - \text{sen}(x) = 0$$

puede iniciarse la iteración con

$$a_0 = 1.5$$

$$b_0 = 2$$

obteniéndose:



i	a_i	b_i	c_i	$f(c_i)$
1	<u>1.5</u>	2.	<u>1.75</u>	<0
2	<u>1.75</u>	2.	<u>1.875</u>	<0
3	<u>1.875</u>	2.	<u>1.9375</u>	>0
4	<u>1.875</u>	<u>1.9375</u>	<u>1.90625</u>	<0
5	<u>1.90625</u>	<u>1.9375</u>	...	

Se han subrayado las cifras correctas. Puede observarse que la convergencia es lenta.

4.3 Método de Newton – Raphson

Si x_0 es suficientemente cercano a la raíz \bar{x} :

$$f(\bar{x}) = f(x_0) + (\bar{x} - x_0) f'(x_0) + \frac{1}{2} (\bar{x} - x_0)^2 f''(x_0) + \dots = 0$$

Y despreciando términos de orden superior:

$$\bar{x} \approx x_0 - \frac{f(x_0)}{f'(x_0)}$$

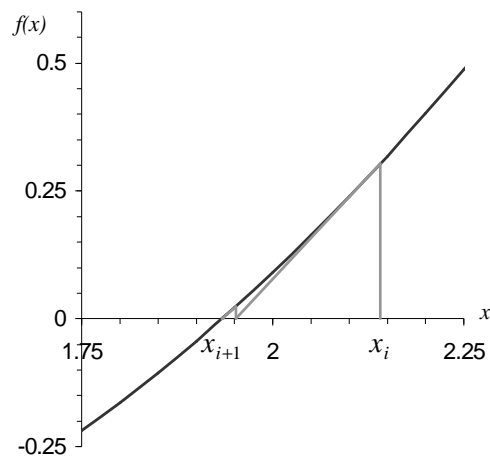
Puede entonces pensarse en iterar con:

$$x_{i+1} = x_i + h_i$$

donde:

$$h_i = - \frac{f(x_i)}{f'(x_i)}$$

expresión que define el método de *Newton – Raphson* (o simplemente de Newton). La figura muestra una interpretación geométrica:



Por ejemplo, puede emplearse el método de Newton para extraer la raíz p de un número c , lo que equivale a resolver:

$$f(x) = x^p - c = 0$$

$$f'(x) = p x^{p-1}$$

Con el método de Newton:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

y por lo tanto:

$$x_{i+1} = x_i - \frac{x_i^p - c}{p x_i^{p-1}} = \frac{1}{p} \left[(p-1) x_i + \frac{c}{x_i^{p-1}} \right]$$

Para obtener la raíz cúbica de 2 se tendría: $x_{i+1} = \frac{1}{3} \left(2 x_i + \frac{2}{x_i^2} \right)$

e iniciando los cálculos con $x_0 = 1$:

i	x_i	$\varepsilon_i = x_i - \bar{x}$
0	1	-0.259921
1	1.333	0.073412
2	<u>1.2638889</u>	0.003968
3	<u>1.2599334934</u>	$1.244 \cdot 10^{-5}$
4	<u>1.25992105001778</u>	$1.229 \cdot 10^{-10}$

Puede observarse que la convergencia es muy rápida. Sin embargo este no siempre es el caso. Considérese, por ejemplo, la ecuación:

$$f(x) = 2 - x - \operatorname{ctg}(x) = 0$$

para la que $f'(x) = -1 + \operatorname{cosec}^2(x)$

Los resultados obtenidos con dos distintas aproximaciones iniciales, $x_0 = 0.5$ y $x_0 = 2$, son:

i	x_i para $x_0 = 0.5$	x_i para $x_0 = 2$
0	0.5	2.0
1	0.5986	-0.18504
2	<u>0.628703</u>	-0.44878
3	<u>0.6308034</u>	-1.49817
4	<u>0.630812760</u>	-676.133
5		-1140.538
6		-1163.343 ¡Diverge!

Las condiciones para la convergencia del método de Newton se revisan a continuación.

De la expansión de f en series de Taylor:

$$0 = f(\bar{x}) = f(x_n) + (\bar{x} - x_n) \cdot f'(x_n) + \frac{1}{2} (\bar{x} - x_n)^2 \cdot f''(\eta) \quad \eta \text{ en intervalo } (x, \bar{x})$$

Dividiendo entre $f'(x_n)$ (que se supone distinto de cero):

$$\frac{f(x_n)}{f'(x_n)} + (\bar{x} - x_n) = \bar{x} - x_{n+1} = -\frac{1}{2} (\bar{x} - x_n)^2 \frac{f''(\eta)}{f'(x_n)}$$

$$\therefore \varepsilon_{n+1} = \frac{1}{2} \frac{f''(\eta)}{f'(x_n)} \varepsilon_n^2$$

y cuando $x_n \rightarrow \bar{x}$ se tiene que $\varepsilon_{n+1} \rightarrow \frac{1}{2} \frac{f''(\bar{x})}{f'(\bar{x})} \varepsilon_n^2$.

Esto es válido sólo si $f'(x) \neq 0$, es decir si la raíz es simple. El caso de raíces múltiples se revisa más adelante. El proceso converge cuando $\varepsilon_{n+1} < \varepsilon_n$, es decir si: $\varepsilon_0 f''/f' < 1$. Si la raíz es simple y la aproximación inicial es adecuada, la convergencia del método de Newton es cuadrática.

Con el método de Newton pueden también obtenerse raíces complejas. Esto requiere que la aproximación inicial sea un número complejo. Por ejemplo, si:

$$f(x) = x^2 + 1 = 0$$

se tendría:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} = \frac{1}{2} \left(x_i - \frac{1}{x_i} \right)$$

e iniciando los cálculos con $x_0 = 1 + i$:

i	x_i
0	$1 + i$
1	$0.25 + 0.75 i$
2	$-0.075 + 0.975 i$
3	$0.0017 + 0.9968 i$
4	$-0.000005 + 1.000004 i$

4.4 Método de la Secante y Otros Procesos del Mismo Tipo

Si en $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ la derivada $f'(x_n)$ se aproxima por $\frac{(f_n - f_{n-1})}{(x_n - x_{n-1})}$, donde f_n

denota $f(x_n)$, se tiene el método de la Secante:

$$x_{n+1} = x_n + h_n$$

$$h_n = -f_n \frac{(x_n - x_{n-1})}{(f_n - f_{n-1})} \quad \text{Se supone que } f_n \neq f_{n-1}$$

Para cada punto, n , solo debe evaluarse una función, f_n , mientras que el método de Newton requiere también la evaluación de f'_n .

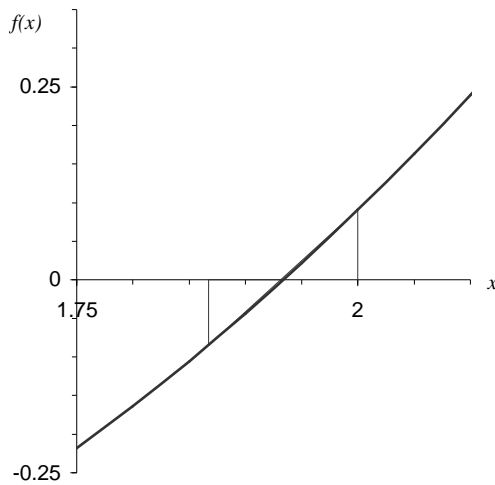
Es importante hacer notar que NO es conveniente reescribir estas expresiones en la forma:

~~$$x_{n+1} = \frac{(x_{n-1} \cdot f_n - x_n \cdot f_{n-1})}{(f_n - f_{n-1})}$$~~

ya que con esta última expresión pueden introducirse fuertes errores numéricos cuando $x_n \approx x_{n-1}$ y $f_n \cdot f_{n-1} > 0$.

En la tabla siguiente se resuelve $f(x) = \frac{1}{4}x^2 - \text{sen}(x)$ para $\bar{x} \approx 2$ por el método de la secante:

n	x_n	$f(x_n)$	$\varepsilon_n = x_n - \bar{x}$
0	1.	-0.59147	-0.93
1	2.	0.090703	0.066
2	1.86704	-0.084980	-0.067
3	1.93135	-0.003177	-0.0024
4	1.93384	-0.000114	+0.00009
5	1.93375	-0.000005	$< \frac{1}{2} \cdot 10^{-5}$



(en la gráfica adjunta casi no se aprecia la diferencia entre la función y la secante).

La convergencia de método de la secante es en general más lenta que la del método de Newton. Sin embargo, se justifica usar este método si la evaluación de $f'(x)$ requiere más del 44% del trabajo que se emplea en evaluar $f(x)$, ya que el mayor número de iteraciones queda más que compensado por el menor número de operaciones realizadas en cada caso. A esta conclusión se llega comparando la convergencia de ambos métodos.

Para el método de la secante:

$$x_{n+1} - x_n = \varepsilon_{n+1} - \varepsilon_n = -f_n \left(\frac{\varepsilon_n - \varepsilon_{n-1}}{f_n - f_{n-1}} \right)$$

y siendo:

$$f_n = f(x_n) = f(\bar{x}) + \varepsilon_n f'(\bar{x}) + \frac{1}{2} \varepsilon_n^2 f''(\bar{x}) + \dots$$

se obtiene (despreciando términos de orden superior):

$$\begin{aligned} \varepsilon_{n+1} &= \varepsilon_n + \frac{(\varepsilon_n f'(\bar{x}) + \frac{1}{2} \varepsilon_n^2 f''(\bar{x}) + \dots)(\varepsilon_n - \varepsilon_{n-1})}{(\varepsilon_n - \varepsilon_{n-1})f'(\bar{x}) + \frac{1}{2}(\varepsilon_n - \varepsilon_{n-1})^2 f''(\bar{x}) + \dots} = \\ &= \varepsilon_n + (\varepsilon_n f'(\bar{x}) + \frac{1}{2} \varepsilon_n^2 f''(\bar{x}) + \dots)(f'(\bar{x}) + \frac{1}{2}(\varepsilon_n + \varepsilon_{n-1})f''(\bar{x}) + \dots)^{\frac{1}{2}} \\ \therefore \varepsilon_{n+1} &\approx \frac{1}{2} \frac{f''(\bar{x})}{f'(\bar{x})} \varepsilon_n \varepsilon_{n-1} \end{aligned}$$

Suponiendo entonces: $\varepsilon_{n+1} = C \varepsilon_n^\beta$ y por lo tanto $\varepsilon_{n+1} = C^{\beta+1} \varepsilon_{n-1}^{\beta^2} = \frac{1}{2} \frac{f''}{f'} (C \varepsilon_{n-1}^\beta) \varepsilon_{n-1}$, de

donde se obtiene: $\beta^2 = \beta + 1 \rightarrow \beta = \frac{1}{2}(1 + \sqrt{5}) = 1.618\dots$ Es decir, la convergencia del método de la secante es entre lineal y cuadrática.

En el método de Falsa Posición la secante se toma entre los puntos (x_n, f_n) y (x_r, f_r) , donde $r < n$ es el mayor índice para el cual $f_n f_r < 0$. Si $f(x)$ es continuo, este método es siempre convergente. Sin embargo, la convergencia es de primer orden.

Otra alternativa es el método de Steffensen: $x_{n+1} = x_n - \frac{f(x_n)}{g(x_n)}$

donde: $g(x_n) = \frac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)}$

cuya convergencia es de segundo orden: $\epsilon_{n+1} \approx \frac{1}{2} \frac{f''(\bar{x})}{f'(\bar{x})} (1 + f'(\bar{x})) \epsilon_n^2$

4.5. Otros Métodos Iterativos

Una ecuación de la forma $f(x) = 0$ puede describirse como $x = g(x)$. Dada entonces una aproximación x_0 a una raíz \bar{x} de $f(x) = 0$, la secuencia x_1, x_2, x_3, \dots definida por: $x_{n+1} = g(x_n)$ converge a \bar{x} siempre que $|g'(x)| < 1$ en la región de interés. El procedimiento puede ser más simple que otros y en algunos casos puede incluso converger más rápidamente.

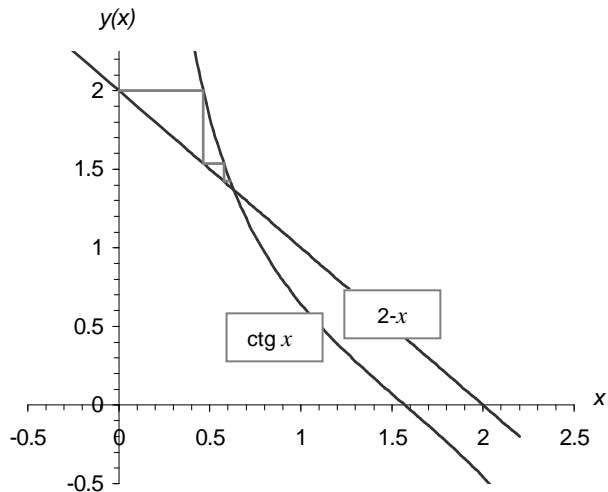
Por ejemplo, la ecuación:

$$f(x) = (2 - x) \operatorname{ctg} x - 1 = 0$$

es equivalente a $\operatorname{ctg} x = (2 - x)$. En este caso podría iterarse con $\operatorname{ctg} x_{n+1} = 2 - x_n$ o, lo que es lo mismo, $x_{n+1} = \operatorname{arc} \operatorname{ctg}(2 - x_n)$. Sin embargo, la iteración escrita al revés, es decir $x_{n+1} = 2 - \operatorname{ctg} x_n$, no funciona.

En efecto, tomando $x_0 = 0$ e iterando con $\operatorname{ctg} x_{n+1} = 2 - x_n$ se obtienen:

n	x_n
0	0.
1	0.464
2	0.577
3	<u>0.6125</u>
4	<u>0.6245</u>
5	<u>0.6286</u>
6	<u>0.6301</u>
...	
10	<u>0.6308017</u>
...	
20	0.630812760



Una interpretación gráfica del proceso se muestra en la figura.

En lo que sigue se analiza cómo se reducen los errores en este caso.

Si se considera $x_n = \bar{x} + \epsilon_n$
 $x_{n+1} = \bar{x} + \epsilon_{n+1}$

se tiene que $\operatorname{ctg}(\bar{x} + \epsilon_{n+1}) = 2 - (\bar{x} + \epsilon_n)$

Suponiendo que $\varepsilon_{n+1} \ll \bar{x}$ puede hacerse una expansión de la cotangente en series:

$$\operatorname{ctg} \bar{x} - \frac{\varepsilon_{n+1}}{\operatorname{sen}^2 \bar{x}} + O(\varepsilon_{n+1}^2) = (2 - \bar{x}) - \varepsilon_n$$

y siendo \bar{x} la solución exacta, ésta satisface idénticamente la ecuación: $\operatorname{ctg} \bar{x} = (2 - \bar{x})$; de donde se concluye que:

$$\varepsilon_{n+1} \approx (\operatorname{sen}^2 \bar{x}) \varepsilon_n \approx \frac{1}{3} \varepsilon_n$$

es decir $\varepsilon_{n+1} < \varepsilon_n$ y por lo tanto el proceso es convergente.

Si en cambio se considera la iteración $x_{n+1} = 2 - \operatorname{ctg} x_n$ se tiene que $\varepsilon_{n+1} \approx 3\varepsilon_n$ y el proceso no converge, aún cuando x_0 sea muy cercano a la raíz \bar{x} :

n	x_n
0	0.6
1	0.538
2	0.325
3	0.965
4	2.69
5	4.01

4.6. Condicionamiento de las Raíces: Raíces Múltiples

Si x_n es una aproximación a una raíz \bar{x} de $f(x) = 0$, se tiene:

$$f(x_n) = f(\bar{x}) + (x_n - \bar{x})f'(\bar{x}) + \frac{1}{2}(x_n - \bar{x})^2 f''(\bar{x}) + \dots$$

Para $x_n \approx \bar{x}$ puede escribirse:

$$|x_n - \bar{x}| \approx \frac{f(x_n)}{f'(\bar{x})}$$

Si el error en la evaluación de $f(x_n)$ es de $O(\delta)$ (δ depende de la precisión de la computadora y es independiente de x) el error en la aproximación de la raíz $|x_n - \bar{x}|$ es de $O\left(\frac{\delta}{f'(\bar{x})}\right)$. Entonces, si $f'(\bar{x})$ es muy pequeño se tiene que $\frac{\delta}{f'(\bar{x})} \gg \delta$ y se dice que la raíz \bar{x} está mal condicionada.

El mismo argumento puede repetirse cuando se trata de una raíz de multiplicidad m , caso en el que::

$$f'(\bar{x}) = f''(\bar{x}) = \dots = f^{(m-1)}(\bar{x}) = 0.$$

En tal caso:

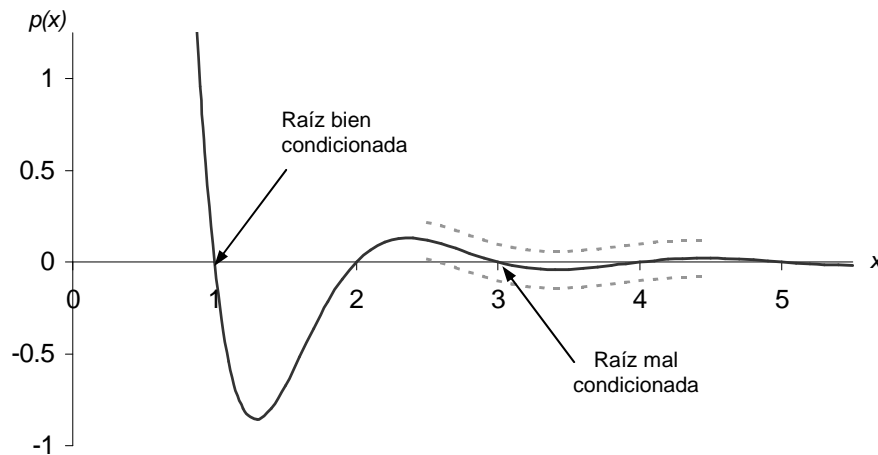
$$|x_n - \bar{x}| = O\left(\frac{\delta \cdot m!}{f^{(m)}(\bar{x})}\right)^{\frac{1}{m}}$$

El exponente $1/m$ implica que las raíces múltiples son en general mal condicionadas.

Considérese por ejemplo: $f(x) = x^2 - 2x + 1 = 0$, que tiene una raíz doble $\bar{x} = 1$. Supóngase que se trabaja con aritmética de punto flotante y 8 decimales en la mantisa. Entonces $\delta = \frac{1}{2} \cdot 10^{-8}$, y siendo $f''(\bar{x}) = 2$ se obtiene:

$$|x_n - \bar{x}| = O\left(\frac{1}{2} \cdot 10^{-8} \cdot \frac{2}{2}\right)^{\frac{1}{2}} = \frac{\sqrt{2}}{2} \cdot 10^{-4} = 0.7071\dots 10^{-4}$$

es decir, cualquier valor en el rango: $0,99992929 \leq x \leq 1,00007071$ sería aceptado como exacto, pudiéndose tener un error relativo de orden 10^{-5} .



Siendo las raíces múltiples mal condicionadas, es de esperarse que la separación relativa de las raíces afecte el condicionamiento. Particularmente crítico es el caso en que las raíces están más o menos uniformemente espaciadas, por ejemplo en:

$$p(x) = (x-1)(x-2)(x-3)\dots(x-20) = x^{20} - 210x^{19} + 20615x^{18} - \dots + 20! = 0$$

Si en lugar del coeficiente 210 se tuviera un coeficiente $(210 + \epsilon)$, se obtendrían las raíces:

$\epsilon = 0$	$\epsilon = 10^{-9}$
1 a 10	1 a 10
11	10.85
12	$12.38 \pm 0.11i$
13	
14	$14.37 \pm 0.72i$
15	
16	$16.57 \pm 0.88i$
17	
18	$18.67 \pm 0.35i$
19	
20	20

Este ejemplo es particularmente mal condicionado, pero son frecuentes las dificultades análogas al evaluar los ceros (es decir, las raíces) de polinomios. La evaluación de las raíces de un polinomio es un problema numérico que debe evitarse, a menos que los datos iniciales hayan sido los coeficientes del polinomio en forma explícita.

Aparte de un mal condicionamiento, las raíces múltiples reducen el orden de la convergencia. Así, por ejemplo, el método de Newton, cuya convergencia es cuadrática cuando x_n es aproximadamente una raíz simple:

$$\varepsilon_{n+1} = \left(\frac{1}{2} \frac{f''(\bar{x})}{f'(\bar{x})} \varepsilon_n^2 \right)$$

tiene convergencia lineal cuando la raíz es múltiple. Para una raíz de multiplicidad m :

$$\varepsilon_{n+1} = \left(1 - \frac{1}{m} \right) \varepsilon_n + O(\varepsilon_n^2)$$

El método de Newton modificado:

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

tiene todavía convergencia cuadrática, pero en general no es posible conocer m a priori. Alternativamente, puede pensarse en procesos del tipo:

$$u(x_n) = \frac{f(x_n)}{f'(x_n)}$$

$$u'(x_n) = 1 - \frac{f''(x_n)}{f'(x_n)} u(x_n)$$

$$x_{n+1} = x_n - \frac{u(x_n)}{u'(x_n)}$$

Pero nótese que esto es poco efectivo cuando la raíz es simple, puesto que se requiere evaluar una función adicional, $f''(x_n)$.

4.7. Métodos para Calcular Raíces de Polinomios

En esta sección se revisan algunos de los muchos métodos específicos para evaluar las raíces (ceros) de polinomios: $p(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_{n-1} x + a_n = 0$
 $(x - \alpha_1)(x - \alpha_2)(x - \alpha_3) \dots (x - \alpha_n) = 0$

Es frecuente subestimar las dificultades que se presentan en problemas de este tipo.

Los métodos mencionados en los acápites anteriores (Newton, secante y otros) son también aquí aplicables, aunque pueden ser poco eficientes.

4.7.1 Método de Bernoulli

Este es un proceso "clásico", que en general resulta poco apropiado, porque se obtiene primero la raíz de mayor módulo.

Se toman n valores arbitrarios $t_1, t_2, t_3 \dots t_n$ (con $t_n \neq 0$) y se calculan nuevos valores por recursión:

$$t_p = - \sum_{i=1}^n a_i t_{p-i}$$

Esta es una ecuación de diferencias, cuya solución puede escribirse como:

$$t_p = c_1 r_1^p + c_2 r_2^p + \dots + c_n r_n^p$$

Las r_i son raíces (distintas de cero) de:

$$r^n = -\sum_{i=1}^n a_i r^{(n-i)}$$

es decir, las r_i son las raíces del polinomio: $r_i = \alpha_i$.

Por tanto:

$$t_p = c_1 \alpha_1^p + c_2 \alpha_2^p + \dots + c_n \alpha_n^p$$

y si $|\alpha_n| > |\alpha_{n-1}| \geq \dots \geq |\alpha_1|$ se tiene que:

$$\lim_{p \rightarrow \infty} \frac{t_p}{t_{p-1}} = \alpha_n$$

Esto es análogo a una iteración directa para determinar un valor propio de una matriz. También aquí la convergencia es lenta cuando $|\alpha_n| \approx |\alpha_{n-1}|$. El proceso converge a α_n , aún cuando $c_n = 0$ (gracias a los errores de redondeo). El inconveniente principal de este método es que extrae primero la raíz de mayor módulo y la "deflación" con esta raíz (que no es exacta) puede introducir errores importantes en las otras raíces.

Supóngase, por ejemplo, que:

$$f(x) = x^4 - 2x^3 + 1.25x^2 - 0.25x - 0.75$$

Con $t_0 = t_1 = t_2 = 0$ y $t_3 = 1$ se obtienen:

k	t_k	$\frac{t_k}{t_{k-1}} \approx \alpha_n$
4	2.0	2.0
5	2.75	1.375
6	3.25	1.1818
7	4.3125	1.3269
8	6.7500	1.5652
9	10.9844	1.6273
10	17.0469	1.5519
...		
15	125.113	1.5104
16	188.240	1.5046
...		
19	632.809	1.4986
20	949497	1.5004

Nótese que la convergencia es lenta. Esto es frecuente. Después de pocos pasos es mejor cambiar a otros métodos (v.g. Newton - Rapshon).

4.7.2. Método de Graeffe (método de los cuadrados de las raíces)

Este es un proceso que puede ser empleado para mejorar el condicionamiento de las raíces, pero es inadecuado para completar la extracción de las mismas.

Dado:
$$p_1(x) = x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n = 0$$

O bien:
$$p_1(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)\dots(x - \alpha_n) = 0.$$

Puede formarse un nuevo polinomio

$$\phi(x) = (-1)^n p_1(x) p_1(-x) = (x^2 - \alpha_1^2)(x^2 - \alpha_2^2)(x^2 - \alpha_3^2)\dots(x^2 - \alpha_n^2).$$

Como $\phi(x)$ solo contiene potencias pares de x , se puede definir:

$$p_2(x) = \phi(\sqrt{x}) = (x - \alpha_1^2)(x - \alpha_2^2)(x - \alpha_3^2)\dots(x - \alpha_n^2)$$

Es decir, $p_2(x) = 0$ tiene raíces $\alpha_1^2 \dots \alpha_i^2 \dots \alpha_n^2$, que son justamente los cuadrados de las raíces de $p_1(x) = 0$. Del mismo modo pueden obtenerse $p_4(x) p_8(x) \dots p_m(x)$. Las raíces de $p_m(x) = 0$, son las α_i^m donde $m = 2^r$.

Si se tienen coeficientes $a_1^{(r)}, a_2^{(r)}, a_3^{(r)} \dots$ tales que:

$$p_m(x) = x^n + a_1^{(r)}x^{n-1} + a_2^{(r)}x^{n-2} + \dots + a_{n-1}^{(r)}x + a_n^{(r)} = 0$$

los coeficientes $a_1^{(r+1)}, a_2^{(r+1)}, a_3^{(r+1)} \dots$ de $p_{2m}(x)$ resultan:

$$a_i^{(r+1)} = (-1)^{n-i} \left[(a_i^{(r)})^2 + 2 \sum_{j=1}^{\min(n-i,i)} (-1)^j a_{i+j}^{(r)} a_{i-j}^{(r)} \right]$$

Y para m (o r) suficientemente grande, puede escribirse:

$$\alpha_1^m \approx |a_1^{(r)}| \quad \alpha_2^m \approx \left| \frac{a_2^{(r)}}{a_1^{(r)}} \right| \dots \quad \alpha_k^m \approx \left| \frac{a_k^{(r)}}{a_{k-1}^{(r)}} \right|$$

Estas expresiones permiten, en teoría, determinar los valores absolutos de las raíces. Los signos deben obtenerse por sustitución en el polinomio original $p(x) = 0$.

Por ejemplo, considérese el polinomio $p(x) = x^3 - 6x^2 + 11x - 6$ (cuyas raíces 1, 2, 3 están uniformemente espaciadas y por tanto están mal condicionadas):

r	m	a_1	a_2	a_3
0	1	-6.	11.	-6.
1	2	-14.	49.	-36.
2	4	-98.	1393.	-1296.
3	8	-6818.2	1.6864×10^6	-1.6796×10^6
4	16	-4.3112×10^7	2.8212×10^{12}	-2.8212×10^{12}
....				

de donde $\alpha_1^{16} \approx 4.3112 \cdot 10^7$

$$\alpha_2^{16} \approx \frac{2.8212 \cdot 10^{12}}{4.3112 \cdot 10^7} = 6.5438 \cdot 10^4$$

$$\alpha_3^{16} \approx \frac{2.8211 \cdot 10^{12}}{2.8212 \cdot 10^{12}} = 0.99998$$

es decir $\alpha_1 \approx 3.0003$ $\alpha_2 \approx 1.9998$ $\alpha_3 \approx 1.0000$

Las raíces de $p_2(x)$ $p_4(x)$ $p_8(x) \dots$ están mejor condicionadas en cada paso. Sin embargo, se observa un crecimiento muy rápido de los coeficientes (posible "overflow") por lo que no pueden realizarse muchos pasos.

4.7.3. Método de Laguerre

Un muy buen método para extraer raíces de polinomios es el de Laguerre. En este método, las sucesivas aproximaciones a una raíz se calculan mediante:

$$x_{k+1} = x_k - \frac{n p(x_k)}{p'(x_k) \pm \sqrt{H(x_k)}}$$

donde n es el grado del polinomio y:

$$H(x_k) = (n-1) \left[(n-1) (p'(x_k))^2 - n p(x_k) p''(x_k) \right]$$

Como en otros casos, debe considerarse el signo que evita la cancelación, es decir, aquel para el que $|x_{k+1} - x_k|$ resulta lo más pequeño posible.

El método de Laguerre requiere evaluar $p(x_k)$, $p'(x_k)$ y $p''(x_k)$ en cada paso, pero la convergencia (para raíces simples) es cúbica. Si las raíces son reales, este método converge siempre. Suponiendo que la aproximación inicial x_0 esté en el intervalo entre α_r y α_{r+1} , converge a una de esas dos raíces. Si en cambio $x_0 < \alpha_1$ o $x_0 > \alpha_n$ el proceso converge a α_1 o α_n , respectivamente.

Por ejemplo, considérese el polinomio:

$$p(x) = (x-1)(x-2)(x-3)(x-4)(x-5)(x-6)(x-7)(x-8) = x^8 - 36x^7 + 546x^6 - 4536x^5 + 22449x^4 - 67284x^3 + 118124x^2 - 109584x + 40320 = 0$$

Con la aproximación inicial $x_0 = 0$ se obtienen:

x_k	$p(x_k)$	$p'(x_k)$	$p''(x_k)$	$H(x_k)$
0	40320	-109584	236248	5.499492E+10
0.937418	369.916	-6836.714	31394.812	1.639940E+09
0.999940	0.302	-5041.569	26140.729	1.245010E+09
<u>1.000000</u>				

En forma similar, con la aproximación inicial $x_0 = 2.5$:

x_k	$p(x_k)$	$p'(x_k)$	$p''(x_k)$	$H(x_k)$
2.5	121.816	-132.750	-977.625	7.532588E+06
2.838696	42.391	-277.131	60.892	3.618725E+06
2.994301	1.374	-242.118	367.294	2.844179E+06
<u>3.000000</u>				

Aunque este método puede ser usado también para extraer raíces complejas, en ese caso no puede garantizarse la convergencia. Si se tuvieran pares de raíces complejas,

pero los coeficientes del polinomio fueran todos reales, sería aconsejable extraer factores cuadráticos con el procedimiento descrito en la sección 4.7.4

4.7.4. Método de Bairstow (factorización iterativa de polinomios).

Un polinomio $p(x) = x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n$ puede expresarse como:

$$p(x) = (x^2 + rx + s) \cdot q(x, r, s) + xF(r, s) + G(r, s).$$

Se requiere determinar r, s para que $x^2 + rx + s$ sea un factor exacto de $p(x)$, es decir, para que se tenga $F(r, s) = G(r, s) = 0$. Siendo éste el caso, dos de las raíces (o un par de raíces complejas) pueden obtenerse de $x^2 + rx + s = 0$; el resto de las raíces son los ceros de $q(x)$. Dados valores aproximados de r, s tales como r_n y s_n , se obtienen valores mejorados, r_{n+1} y s_{n+1} , considerando:

$$F(r_{n+1}, s_{n+1}) \approx F(r_n, s_n) + (r_{n+1} - r_n) \left. \frac{dF}{dr} \right|_{r_n, s_n} + (s_{n+1} - s_n) \left. \frac{dF}{ds} \right|_{r_n, s_n} \approx 0$$

$$G(r_{n+1}, s_{n+1}) \approx G(r_n, s_n) + (r_{n+1} - r_n) \left. \frac{dG}{dr} \right|_{r_n, s_n} + (s_{n+1} - s_n) \left. \frac{dG}{ds} \right|_{r_n, s_n} \approx 0$$

de donde, con la notación $F_r = \frac{dF}{dr} \dots G_s = \frac{dG}{ds}$, se tiene:

$$r_{n+1} = r_n - \left(\frac{G \cdot F_s - F \cdot G_s}{F_r G_s - F_s G_r} \right)_{r_n, s_n}$$

$$s_{n+1} = s_n - \left(\frac{G \cdot F_r - F \cdot G_r}{F_r G_s - F_s G_r} \right)_{r_n, s_n}$$

Dado que el polinomio original $p(x)$ es independiente de r y s , al derivar el polinomio $p(x) = (x^2 + rx + s) \cdot q + xF + G$ con relación a r y s se obtienen:

$$xq + (x^2 + rx + s) \cdot q_r + (xF_r + G_r) = 0$$

$$q + (x^2 + rx + s) \cdot q_s + (xF_s + G_s) = 0$$

es decir, los F_r y G_r son los coeficientes del residuo al dividir $-xq(x)$ entre $x^2 + rx + s$ mientras los F_s y G_s se obtienen como coeficientes del residuo al dividir $-q(x)$ también entre $x^2 + rx + s$.

La convergencia de este proceso depende de una buena aproximación inicial.

El siguiente ejemplo ilustra un paso del proceso. Supóngase que se tiene el polinomio:

$$p(x) = x^4 - 5x^3 + 14x^2 - 25x + 25 = 0$$

y un factor aproximado $x^2 - 4x + 4$ (es decir $r_0 = -4$, $s_0 = +4$). Dividiendo $p(x)$ entre este factor (utilizando el procedimiento de Ruffini) se obtiene:

$p(x)$	1	-5	14	-25	25
$-r_0 = 4$		4	-4	24	
$-s_0 = -4$			-4	4	-24
$q(x)$	1	-1	6	3	1

Y análogamente:

$$\begin{array}{rcccc}
 -xq(x) & & -1 & 1 & -6 & 0 \\
 4 & & & -4 & -12 & \\
 -4 & & & & 4 & 12 \\
 \hline
 & & -1 & -3 & -14 & 12 \\
 \\
 -q(x) & & & -1 & 1 & -6 \\
 4 & & & & -4 & \\
 -4 & & & & & 4 \\
 \hline
 & & -1 & & -3 & -2
 \end{array}$$

y de estos resultados:

$$\begin{array}{lll}
 F = 3 & F_r = -14 & F_s = -3 \\
 G = 1 & G_r = 12 & G_s = -2
 \end{array}$$

$$w = F_r \cdot G_s - F_s \cdot G_r = 64$$

$$r_1 = -4 - \frac{(1)(-3) - (3)(-2)}{64} = -4.0469$$

$$s_1 = 4 - \frac{(1)(-14) - (3)(12)}{64} = 4.7813$$

es decir: $x^2 + r_1x + s_1 = x^2 - 4.0469x + 4.7813$

En el siguiente paso se obtienen:

$$\begin{array}{ll}
 F = 1.2547 & G = -0.6350 \\
 F_r = -13.1003 & G_r = 14.7920 \\
 F_s = -3.0938 & G_s = -0.5803
 \end{array}$$

de donde $r_2 = -4.0973$, $s_2 = 4.9732$. El factor exacto es $x^2 - 4x + 5$.

El método antes descrito es adecuado solo si la aproximación inicial es cercana al factor exacto (esto es análogo a lo que ocurre con el método de Newton). Pueden también utilizarse las mismas ideas para separar factores de grado mayor que 2.

4.7.5. Método de Jenkins y Traub

Este proceso tiene algunas similitudes con el método de iteración inversa para hallar vectores característicos. Converge a la raíz de menor módulo (lo que es beneficioso para una "deflación" adecuada; véase la sección 4.7.6) y permite efectuar translaciones para acelerar la convergencia. Es probablemente el mejor de los métodos para extraer raíces de polinomios (pero en algunos casos no es el más eficiente). El algoritmo tiene 3 etapas:

a. Dado el polinomio $p(x) = \sum_{i=0}^n a_i x^{n-i}$ con $a_0 = 1, a_n \neq 0$, se determina:

$$h^{(0)}(x) = p'(x) = \sum_{i=0}^n (n-i) a_i x^{n-i-1}$$

y por recursión:

$$h^{(k+1)}(x) = \frac{1}{x} \left[h^{(k)}(x) - \frac{h^{(k)}(0)}{p(0)} p(x) \right] \quad k = 0, 1, 2 \dots$$

A medida que k crece, las razones entre los coeficientes de $h^{(k)}(x)$ y aquellos de $h^{(k+1)}(x)$ tienden a hacerse constantes e iguales a la raíz de menor módulo:

$$\alpha_1 \approx \frac{h^{(k)}(0)}{h^{(k+1)}(0)}$$

b. En una segunda etapa, con una “translación” fija $s_0 \approx r_1$, se determinan:

$$h^{(k+1)}(x) = \frac{1}{(x - s_0)} \left[h^{(k)}(x) - \frac{h^{(k)}(s_0)}{p(s_0)} p(x) \right]$$

c. Y finalmente se refina la aproximación con translaciones variables en cada paso:

$$s_{k+1} = s_k + \frac{h^{(k-1)}(s_k)}{h^{(k)}(s_k)} \approx r_1$$

$$h^{(k+1)}(x) = \frac{1}{(x - s_{k+1})} \left[h^{(k)}(x) - \frac{h^{(k)}(s_{k+1})}{p(s_{k+1})} p(x) \right]$$

Cada paso de este método requiere $O(2n)$ operaciones.

En lo que sigue se analiza cómo el proceso converge. El polinomio cuyas raíces se buscan puede expresarse en la forma:

$$p(x) = (x - \alpha_1)^{m_1} (x - \alpha_2)^{m_2} (x - \alpha_3)^{m_3} \dots = 0$$

donde $m_1, m_2, m_3 \dots$ son las multiplicidades de las raíces.

El método se inicia con $h^{(0)}(x) = p'(x)$, que es un polinomio de grado $n - 1$. Este polinomio puede escribirse en la forma:

$$h^{(0)}(x) = \sum c_i^{(0)} \frac{p(x)}{(x - \alpha_i)}$$

donde $c_i^{(0)} = m_i$. Puede probarse por inducción que los sucesivos polinomios $h^{(1)}(x) h^{(2)}(x) h^{(3)}(x) \dots$ son también polinomios de grado $n - 1$. Supóngase que:

$$h^{(k-1)}(x) = \sum c_i^{(k-1)} \frac{p(x)}{(x - \alpha_i)}$$

$$\text{Entonces: } h^{(k)}(x) = \frac{1}{(x - s_k)} \left[h^{(k-1)}(x) - \frac{h^{(k-1)}(s_k)}{p(s_k)} p(x) \right]$$

$$= \frac{1}{(x - s_k)} \left[\sum c_i^{(k-1)} \frac{p(x)}{(x - \alpha_i)} - \frac{p(x)}{p(s_k)} \sum c_i^{(k-1)} \frac{p(s_k)}{(s_k - \alpha_i)} \right]$$

$$= \sum \frac{c_i^{(k-1)}}{(\alpha_i - s_k)} \frac{p(x)}{(x - \alpha_i)} = \sum c_i^{(k)} \frac{p(x)}{(x - \alpha_i)}$$

es decir, si $h^{(k+1)}(x)$ es un polinomio de grado $n - 1$, también lo es $h^{(k)}(x)$.

Además:
$$c_i^{(k)} = \frac{c_i^{(k-1)}}{(\alpha_i - s_k)} = \frac{m_i}{(\alpha_i - s_1)(\alpha_i - s_2)(\alpha_i - s_3)\dots(\alpha_i - s_k)}$$

y si se consideran translaciones s_j tales que: $|\alpha_1 - s_j| < |\alpha_2 - s_j| < |\alpha_3 - s_j| \dots$
se tiene que (considerando multiplicidades m_i iguales): $c_1^{(k)} > c_2^{(k)} > c_3^{(k)} \dots$ y por lo tanto:

$$(\alpha_i - s_k) = \lim_{k \rightarrow \infty} \frac{h^{(k-1)}(x)}{h^{(k)}(x)}$$

Esta expresión puede también usarse para determinar las translaciones más adecuadas:

$$s_{k+1} = s_k + \frac{h^{(k-1)}(s_k)}{h^{(k)}(s_k)} = s_k - \frac{w^{(k)}(s_k)}{w'^{(k)}(s_k)}$$

donde:

$$w^{(k)}(x) = \frac{p(x)}{h^{(k)}(x)}$$

La última expresión puede probarse considerando que:

$$h^{(k)}(s_k) = \left. \frac{1}{(x - s_k)} \left[h^{(k-1)}(x) - \frac{h^{(k-1)}(s_k)}{p(s_k)} p(x) \right] \right|_{x=s_k}$$

y utilizando la regla de L' Hospital.

Por otro lado, dado que $p(\alpha_i) = 0$, se tiene que $w^{(k)}(\alpha_i) = 0$ y por lo tanto:

$$s_{k+1} = s_k - \frac{w^{(k)}(s_k)}{w'^{(k)}(s_k)}$$

Esta expresión es análoga a la utilizada en el método de Newton. Considerando $c_k = \alpha_1 - s_k$ puede escribirse:

$$\varepsilon_{k+1} \approx \frac{1}{2} \frac{w''^{(k)}(\alpha_1)}{w'^{(k)}(\alpha_1)} \varepsilon_k^2$$

pero en este caso la función $w^{(k)}$ es variable en cada paso:

$$w^{(k)}(x) = \frac{p(x)}{h^{(k)}(x)}$$

$$h^{(k)}(x) = c_i^{(k)} \frac{p(x)}{(x - \alpha_1)} + g(x)p(x)$$

Para $x \approx \alpha_1$ se tiene: $\frac{1}{(x - \alpha_1)} \gg g(x)$ y por lo tanto:

$$w^{(k)}(x) = \frac{p(x)}{h^{(k)}(x)} = \frac{(x - \alpha_1)}{c_1^{(k)} + g(x)(x - \alpha_1)}$$

puede aproximarse por:

$$w^{(k)}(x) = (x - \alpha_1) \left[(c_1^{(k)})^{-1} - g(x)(x - \alpha_1)(c_1^{(k)})^{-2} \dots \right]$$

de donde:

$$\varepsilon_{k+1} \approx \frac{g_1(\alpha_1)}{c_1^{(k)}} \varepsilon_k^2$$

pero
$$c_1^{(k)} = \frac{c_1^{(k-1)}}{(\alpha_1 - s_k)} = \frac{c_1^{(k-1)}}{\varepsilon_k} = \frac{c_1^{(k-2)}}{(\varepsilon_k \varepsilon_{k-1})} \dots$$

y entonces:

$$\varepsilon_{k+1} = (\beta \varepsilon_k \varepsilon_{k-1} \dots) \varepsilon_k^2$$

$$\varepsilon_k = (\beta \varepsilon_{k-1} \dots) \varepsilon_{k-1}^2$$

de donde $\varepsilon_{k+1} = \varepsilon_k^3 \cdot \varepsilon_{k-1}^{-1}$ y considerando $\varepsilon_{k+1} = C \cdot \varepsilon_k^\gamma$ se obtiene finalmente $\gamma = 2.62$. El orden de convergencia es superior al del método de Newton.

Considérese, por ejemplo, el polinomio: $p(x) = x^2 - 2x + 1$ (cuyas raíces son iguales $\alpha_1 = \alpha_2 = 1$). En este caso:

$$h^{(0)}(x) = p'(x) = 2x - 2$$

$$h^{(1)}(x) = \frac{1}{x} \left[h^{(0)}(x) - \frac{h^{(0)}(0)}{p(0)} p(x) \right] = 2x - 2$$

y es evidente que, si se continuara con el proceso, $h^{(n)}(x) = 2x - 2$. Es decir, en un paso se tiene convergencia. Esto se debe a que en realidad se tiene una sola raíz con multiplicidad 2.

$$\alpha_1 = s + \frac{h^{(k-1)}(s)}{h^{(k)}(s)} = 0 + \frac{-2}{-2} = 1$$

Para el siguiente ejemplo: $p(x) = x^2 - 3x + 2 = 0$ (cuyas raíces son $\alpha_1 = 1, \alpha_2 = 2$):

	x^2	x^1	x^0	$h(s_{k-1})$	s_k	$h(s_k)$	$p(s_k)$
1 $p(x)$	1.	-3.	2.				
2 $h^{(0)}(x) = p'(x)$		2.	-3.		0.	-3.	2.
3 $1.5 p(x)$	1.5	-4.5	3.				
4 (2)+(3)	1.5	-2.5	0.				
5 (4)/x = $h^{(1)}(x)$		1.5	-2.5	-2.5	1.2	-0.70	-0.16
6 $-4.375 p(x)$	-4.375	13.125	-8.750				
7 (5)+(6)	-4.375	14.625	-11.250				
8 (7)/(x-1.2)		-4.375	9.375	4.125	1.0303	4.8674	-0.029385
9 $165.645 p(x)$	165.645	-496.934	331.290				
10 (8)+(9)	165.645	-501.309	340.664				
11 (10)/(x-1.00303)		165.645	-330.644	-159.979	.999878	-165.019	0.000122
12 $1352450 p(x)$	1 352 450	-4 057 350	2 704 900				
13 (11)+(12)	1 352 450	-4 057 184	2 704 569				
14 (13)/(x-.999878)		1 352 450	-2 704 899	-1 352 614	1.000000	← La primera raíz	

4.7.6 Deflación

El polinomio $p(x)$, de grado $n \geq 1$, puede escribirse como: $p(x) = (x - a)q(x) + r$, donde $q(x)$ es un polinomio de grado $n - 1$ y $r = p(a)$. Si $a = \alpha_1$ es una raíz exacta de $p(x) = 0$, se tiene $r = 0$. Las raíces restantes son las raíces de $q(x) = 0$. Para el cómputo de estas otras raíces puede entonces trabajarse con el cociente $q(x)$ en lugar de $p(x) = 0$. Esto es una "deflación". La deflación puede repetirse a medida que se obtienen otras raíces. Es evidente que esto ahorra operaciones.

Sin embargo, a medida que se calculan las raíces éstas no pueden obtenerse en forma exacta y la deflación con una raíz que es solo aproximada produce un polinomio $q(x)$ que está afectado por esos errores.

Considérese, por ejemplo, $x^2 - 101x + 100 = 0$, cuyas raíces exactas son $\alpha_1 = 1$, $\alpha_2 = 100$. Si se obtiene primero la raíz α_2 (aquella de mayor módulo) con un 0,1% de error: $\alpha_2 \approx 100.1$, la deflación produce:

$p(x)$		1	-101.	100.
α_2	100.1			
$q(x)$		1	-0.9	9.91

Es decir, se obtiene $q(x) = (x - 0.9)$, y haciendo $q(x) = 0$ resulta $\alpha_1 \approx 0.9$. Nótese que un error de 0,1% en α_2 produce un error de 10% en α_1 . Si en cambio se determina primero la raíz menor, también con un error de 0,1%, $\alpha_1 \approx 1.001$:

$p(x)$		1	-101.	100.
α_1	1.001			
$q(x)$		1	-99.999	0.099

y de $q(x) = (x - 99.999) = 0$: se obtiene $\alpha_2 \approx 99.999$; con sólo un 0,001% de error.

En conclusión, los errores introducidos por la deflación son menores si las raíces se determinan en orden ascendente de valores absolutos.

4.8. SISTEMAS DE ECUACIONES NO LINEALES

En algunos problemas académicos la solución de sistemas de ecuaciones no lineales puede ser hecha por simple eliminación. Por ejemplo, para determinar los puntos en los que se intersectan el círculo $x^2 + y^2 = 3$ y la hipérbola $x^2 - y^2 = 1$, basta restar y sumar ambas expresiones (obteniéndose $x = \pm\sqrt{2}$, $y = \pm 1$).

Sin embargo, en la mayor parte de los casos prácticos será necesario iterar. Muchos de los métodos antes descritos para resolver una ecuación no lineal pueden generalizarse para resolver sistemas de ecuaciones no lineales:

$$f_i(x_1, x_2, x_3, \dots, x_n) = 0 \quad i = 1, 2, 3, \dots, n$$

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

Iteración directa

Un proceso simple puede ser obtenido rescribiendo estas ecuaciones como:

$$x_i = g_i(x_1, x_2, x_3, \dots, x_n) \quad i = 1, 2, 3, \dots, n$$

$$\mathbf{x} = \mathbf{g}(\mathbf{x})$$

lo que permite iterar con

$$x_i^{(k+1)} = g_i(x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, \dots, x_n^{(k)})$$

o bien con

$$x_i^{(k+1)} = g_i(x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}, x_{i+1}^{(k)}, \dots, x_n^{(k)})$$

Estos son los equivalentes no lineales de los métodos de Jacobi y Gauss – Seidel.

También aquí puede establecerse un criterio de convergencia similar a $|g'(x)| < 1$.

Definiendo la matriz \mathbf{D} , con coeficientes:

$$d_{ij}(x) = \frac{\delta g_i}{\delta x_j}$$

se tiene que los valores característicos de la matriz \mathbf{D} deben ser tales que: $|\lambda_i| < 1$. Con frecuencia se presentan sistemas no lineales de la forma: $\mathbf{x} = \mathbf{a} + h\mathbf{g}(\mathbf{x})$ donde las $\delta g_i / \delta x_j$ son finitas. Entonces, para h suficientemente pequeño, el proceso iterativo $\mathbf{x}^{(k+1)} = \mathbf{a} + h\mathbf{g}(\mathbf{x}^{(k)})$ satisface las condiciones de convergencia.

Por ejemplo, sea el sistema de ecuaciones (no lineales):

$$x = \frac{1}{2} \text{sen}(x + y)$$

$$y = \frac{1}{2} \text{cos}(x - y)$$

Puede iterarse con:

$$x^{(k+1)} = \frac{1}{2} \text{sen}(x^{(k)} + y^{(k)})$$

$$y^{(k+1)} = \frac{1}{2} \text{cos}(x^{(k+1)} - y^{(k)})$$

Obteniéndose (con la aproximación inicial $x^{(0)} = y^{(0)} = 0$):

k	$x^{(k)}$	$y^{(k)}$
1	0	0.5
2	0.239 712 729	<u>0.483 158 048</u>
3	<u>0.330 770 125</u>	<u>0.494 205 706</u>
4	<u>0.367 265 691</u>	<u>0.495 976 965</u>
5	<u>0.379 977 102</u>	<u>0.496 639 778</u>
...		
10	<u>0.386 424 387</u>	<u>0.496 949 307</u>
...		
20	<u>0.386 450 795</u>	<u>0.496 950 555</u>

Método de Newton

También el método de Newton - Raphson puede ser generalizado para sistemas de ecuaciones no lineales:

$$\mathbf{f}(\mathbf{x}^{(k+1)}) \approx \mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{D}(\mathbf{x}^{(k)}) (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = 0$$

En la sección 4.7.4 se presentó una aplicación de estas ideas. El proceso converge si $\mathbf{x}^{(k)}$ es suficientemente cercano a la solución exacta $\bar{\mathbf{x}}$, lo que es relativamente fácil

cuando las no linealidades no son muy fuertes y se combina el método de Newton - Raphson con un proceso incremental.

Para el caso particular de dos ecuaciones no lineales:

$$\begin{aligned} f_1(x, y) &= 0 \\ f_2(x, y) &= 0 \end{aligned}$$

puede escribirse:

$$\begin{Bmatrix} f_1 \\ f_2 \end{Bmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} \begin{Bmatrix} x^{(k+1)} - x^{(k)} \\ y^{(k+1)} - y^{(k)} \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix}$$

expresión en la que f_1 , f_2 y sus derivadas se calculan con la aproximación $x^{(k)}$, $y^{(k)}$. Luego se obtienen:

$$\begin{Bmatrix} x^{(k+1)} \\ y^{(k+1)} \end{Bmatrix} = \begin{Bmatrix} x^{(k)} \\ y^{(k)} \end{Bmatrix} - \frac{1}{d} \begin{Bmatrix} f_1 \frac{\partial f_2}{\partial y} - f_2 \frac{\partial f_1}{\partial y} \\ -f_1 \frac{\partial f_2}{\partial x} + f_2 \frac{\partial f_1}{\partial x} \end{Bmatrix}$$

siendo $d = \det(\mathbf{D}) = \frac{\partial f_1}{\partial x} \frac{\partial f_2}{\partial y} - \frac{\partial f_1}{\partial y} \frac{\partial f_2}{\partial x}$.

Considérese, por ejemplo:

$$\begin{aligned} f_1(x, y) &= x^2 + y - 1 = 0 \\ f_2(x, y) &= (x-1)^2 + (y-0.5)^2 - 1 = 0 \end{aligned}$$

Estas ecuaciones tienen dos soluciones:

$$\begin{aligned} r_1 &= (0.125\ 122\ 549\ 762, 0.984\ 344\ 347\ 550) \\ r_2 &= (1.215\ 146\ 790\ 092, -0.476\ 581\ 721\ 472) \end{aligned}$$

Con la aproximación inicial (0,0) se obtienen:

k	$x^{(k)}$	$y^{(k)}$	f_1	f_2
0	0.000000	0.000000	-1.000000	0.250000
1	-0.375000	1.000000	0.140625	1.140625
2	0.125000	1.234375	0.250000	0.304932
3	0.095595	0.991726	0.000865	0.059743
4	0.125088	0.985223	0.000870	0.000912
5	0.125122	0.984344	0.000000	0.000001
6	0.125123	0.984344	0.000000	0.000000

Con otras aproximaciones iniciales, como por ejemplo (1,0) ó (1,-1), se obtiene r_2

Método de Máxima Gradiente

El método de Newton tiene convergencia cuadrática, pero requiere una buena aproximación inicial. Con tal fin, puede emplearse el método de máxima gradiente, cuya convergencia es lenta pero está garantizada (siempre que \mathbf{D} sea no singular).

La solución del sistema de ecuaciones no lineales $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ es aquella que hace mínima la función:

$$F(\mathbf{x}) = [\mathbf{f}(\mathbf{x})]^T \mathbf{f}(\mathbf{x}) = \sum_{i=1}^n f_i(x_1, x_2, \dots, x_n)$$

Partiendo de una aproximación $\mathbf{x}^{(0)}$ se hacen correcciones de la forma:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha_k \nabla F = \mathbf{x}^{(k)} - 2\alpha_k \mathbf{D}^T \mathbf{f}$$

α_k debe ser tal que $F(\mathbf{x}^{(k+1)}) < F(\mathbf{x}^{(k)})$.

Aproximaciones

La evaluación de los n^2 coeficientes de la matriz $\mathbf{D}(\mathbf{x}^{(k)})$, es decir: $d_{ij} = \partial f_i / \partial x_j$ puede requerir demasiadas operaciones, por lo que es común volver a evaluar \mathbf{D} solo cada cierto número, m , de pasos (y no en cada paso). Se tiene así el método de Newton - Raphson modificado:

$$\mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{D}(\mathbf{x}^{(p)}) (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = 0 \quad k = p, p+1, \dots, p+m.$$

Esto también ahorra muchas operaciones al resolver el sistema de ecuaciones lineales para determinar las $\mathbf{x}^{(k+1)}$, puesto que la matriz \mathbf{D} solo debe reducirse cada m pasos.

También pueden utilizarse métodos análogos al método de la secante. Una aproximación frecuente es:

$$d_{ij} = \frac{\delta f_i}{\delta x_j}(\mathbf{x}) \approx \frac{f_i(\mathbf{x} + h_j \mathbf{e}_j) - f_i(\mathbf{x})}{h_j}$$

donde \mathbf{e}_j es la columna j de la matriz identidad de orden n , y los $h_j \neq 0$ son arbitrarios, por ejemplo: $h_j = x_j^{(k-1)} - x_j^{(k)}$ (esto es el método de la secante, que da un orden de convergencia de aproximadamente 1.6). Si en cambio se toma $h_j = f_j(\mathbf{x}^{(k)})$ se obtiene una generalización del método de Steffensen.

Otra posibilidad consiste en derivar $\mathbf{f}(\mathbf{x}) = 0$ con respecto a un parámetro, α :

$$\frac{\delta \mathbf{f}}{\delta \mathbf{x}} \cdot \frac{\delta \mathbf{x}}{\delta \alpha} + \frac{\delta \mathbf{f}}{\delta \alpha} = 0$$

lo que permite utilizar muchos de los procesos para resolver sistemas de ecuaciones diferenciales ordinarias descritos en el capítulo 6. Por ejemplo, pueden considerarse

$$\frac{\delta \mathbf{f}}{\delta \mathbf{x}} = \mathbf{D} \quad \frac{\delta \mathbf{x}}{\delta \alpha} \Delta \alpha = \Delta \mathbf{x} \quad \frac{\delta \mathbf{f}}{\delta \alpha} \Delta \alpha = \Delta \mathbf{f}$$

y entonces:

$$\mathbf{D}(\mathbf{x}^{(k)}) \cdot \Delta \mathbf{x}^{(k)} = \Delta \mathbf{f}^{(k)}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta \mathbf{x}^{(k)}$$

Este es el método de Euler, un proceso simple pero no siempre adecuado.