

Mètodes Numèrics — Grau de Matemàtiques

Zeros de funcions

Lluís Alsedà

adaptat dels *Apunts de Mètodes Numèrics* de Josep Maria Mondelo, 2009

Departament de Matemàtiques
Universitat Autònoma de Barcelona

<http://www.mat.uab.cat/~alseda>

març 2016 (versió 1.0.5)

UAB

Universitat Autònoma
de Barcelona

DEPARTAMENT DE MATEMÀTIQUES



Subjecte a una llicència *Creative Commons* de *Reconeixement-NoComercial-CompartirIgual 4.0 Internacional* (<http://creativecommons.org/licenses/by-nc-sa/4.0/>)

Continguts

Introducció	▶ 1
Mètode de la bisecció	▶ 2
Mètode de Regula Falsi	▶ 7
Mètode de la secant	▶ 9
Multiplicitat dels zeros	▶ 11
Mètode de Newton	▶ 12
Iteració directa — Mètodes de punt fix	▶ 23
Teorema de l'Aplicació Contractiva i fites de l'error	▶ 29
Ordre de convergència i control de l'error	▶ 54
Mètode de Txebishev	▶ 80
Eficiència computacional	▶ 85
Mètode d'acceleració Aitken	▶ 89
Mètode d'acceleració Steffensen	▶ 95
Arrels de polinomis de grau baix	▶ 110
Localització de les arrels de polinomis	▶ 119
Separació de zeros de polinomis. Mètode de Sturm	▶ 121
Més sobre localització de zeros de polinomis: Teorema de Budan-Fourier i Regla de Descartes	▶ 146
Localització de valors propis: Teorema de Gershgorin	▶ 160

En aquest tema considerem equacions del tipus $f(x) = 0$ on f és una funció real de variable real. També considerem equacions del tipus $x = g(x)$.

Anomenem *arrel* o *solució* de l'equació $f(x) = 0$ un valor α tal que $f(\alpha) = 0$. Es diu també que α és un *zero* de f .

En el cas d'una equació del tipus $x = g(x)$, anomenem *punt fix* de g un valor α tal que $g(\alpha) = \alpha$.

L'objectiu d'aquest tema és desenvolupar mètodes numèrics que ens permetin obtenir aproximacions numèriques d'arrels o punts fixos.

Suposem que tenim $f: [a, b] \rightarrow \mathbb{R}$ contínua i volem trobar α amb $f(\alpha) = 0$. Si $f(a)f(b) \leq 0$ (i.e., $f(a)$ i $f(b)$ tenen signes diferents), el teorema de Bolzano ens assegura que llavors existeix $\alpha \in [a, b]$ amb $f(\alpha) = 0$. L'aplicació recursiva del teorema de Bolzano a subdivisions de l'interval original dóna lloc al mètode de la bisecció, que podem descriure com segueix:

Algorisme (Mètode de la bisecció)

procedure BISECCIO(f , a , b , Tol x , Tol f)

$fa \leftarrow f(a)$; $fb \leftarrow f(b)$;

▷ Inicialització

if $fabs(fa) < Tolf$ **then return** a ; **end if**

▷ Control: a és la solució

if $fabs(fb) < Tolf$ **then return** b ; **end if**

▷ Control: b és la solució

if $fa * fb > 0$ **then Stop**(Error: No és un interval vàlid); **end if**

while $b - a > Tolx$ **do**

▷ Iteració fins trobar l'arrel amb precisió Tol x

$m \leftarrow \frac{b+a}{2}$; $fm \leftarrow f(m)$;

▷ Punt mig

if $fabs(fm) < Tolf$ **then return** m ; **end if**

▷ Control: m és la solució

if $fa * fm < 0$ **then**

▷ El canvi de signe és a $[a, m]$

$b \leftarrow m$; $fb \leftarrow fm$;

▷ Redefinició de l'interval $[a, b]$

else

▷ Ara el canvi de signe és a $[m, b]$

$a \leftarrow m$; $fa \leftarrow fm$;

▷ Redefinició de l'interval $[a, b]$

end if

end while

return $\frac{b+a}{2}$;

▷ Retornem el punt mig de l'interval final

end procedure

Anem a provar que l'algorisme anterior aproxima arbitràriament una arrel de f .

L'interval $[a, b]$ i el punt m de l'algorisme anterior a la iteració i els denotarem per $I_i = [a_i, b_i]$ i m_i , respectivament. Per fixar la notació prenem $i = 1$ i $I_1 = [a, b]$ a l'inici del programa.

Proposició

Sigui $f : [a, b] \rightarrow \mathbb{R}$ contínua amb $f(a)f(b) \leq 0$ i suposem que apliquem l'algorisme de bisecció. Llavors, $\forall i = 1, 2, 3, \dots$

$$\exists \alpha \in I_i \text{ tal que } f(\alpha) = 0 \text{ i } |m_i - \alpha| \leq \frac{b - a}{2^i}.$$

Demostració

Anem a veure primer per inducció que, $\forall i$,

$$f(a_i)f(b_i) \leq 0 \quad \text{i} \quad b_i - a_i = \frac{b - a}{2^{i-1}}. \quad (1)$$

El cas $i = 1$ se segueix directament de les hipòtesis. Suposem (1) cert per i qualsevol i provem-ho per $i + 1$. Si $f(a_i)f(m_i) \leq 0$, prendrem $a_{i+1} = a_i$, $b_{i+1} = m_i$ i tindrem $f(a_{i+1})f(b_{i+1}) \leq 0$. Si no, s'ha de complir $f(m_i)f(b_i) \leq 0$, donat que, si fos $f(m_i)f(b_i) > 0$, com que estem suposant $f(a_i)f(m_i) > 0$ tindríem $f(a_i)f(b_i) > 0$, contràriament a la hipòtesi d'inducció. Per tant $f(m_i)f(b_i) \leq 0$, i, com que en aquest cas prenem $a_{i+1} = m_i$, $b_{i+1} = b_i$, tenim $f(a_{i+1})f(b_{i+1}) \leq 0$. Respecte de la segona igualtat de (1)

$$b_{i+1} - a_{i+1} = \frac{b_i - a_i}{2} \stackrel{(1)}{=} \frac{1}{2} \frac{b - a}{2^{i-1}} = \frac{b - a}{2^i},$$

i això finalitza la inducció.

Per a acabar, per la primera igualtat de (1), el Teorema de Bolzano ens assegura que existeix $\alpha \in I_i$ amb $f(\alpha) = 0$. A més,

$$|m_i - \alpha| \leq \frac{b_i - a_i}{2} \stackrel{(1)}{=} \frac{1}{2} \frac{b - a}{2^{i-1}} = \frac{b - a}{2^i}$$

tal com volíem veure. □

Observació

La desigualtat $|m_i - \alpha| \leq \frac{b-a}{2^i}$:

- 1 Ens diu que, si suposem m_i aproximació d' α : $|e_a(m_i, \alpha)| \leq \frac{b-a}{2^i}$.
- 2 En cas que α sigui l'única arrel de f a $[a, b]$, demostra la convergència de la successió $\{m_i\}_i$ cap a α .
- 3 Ens permet estimar el nombre d'iterats necessaris a l'algorisme de bisecció per a assolir una certa precisió.

Exemple

Sigui $f(x) = \cos(x) - x$.

Donat que $f(0) = 1 > 0$ i $f(\frac{\pi}{2}) = -\frac{\pi}{2} < 0$, existeix $\alpha \in (0, \frac{\pi}{2})$ amb $f(\alpha) = 0$.

Anem a veure quantes iteracions cal fer per assegurar que l'error sigui més petit que 10^{-6} :

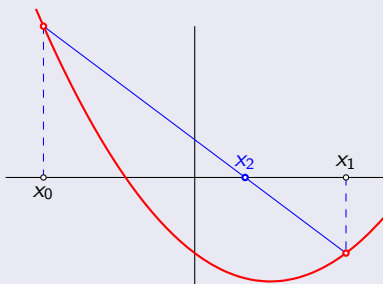
$$|m_n - \alpha| \leq \frac{b-a}{2^n} = \frac{\pi}{2^{n+1}} < 10^{-6}.$$

Per tant,

$$\ln\left(\frac{\pi}{2^{n+1}}\right) = \ln(\pi) - (n+1)\ln(2) < -6\ln(10) \iff 6\ln(10) + \ln(\pi) < (n+1)\ln 2$$

que dona $n > 20.5830$. És a dir, és suficient prendre $n = 21$.

Representació gràfica del mètode de Regula Falsi



El mètode de Regula Falsi funciona com el de bisecció però en comptes de triar el punt mig de l'interval es tria el punt d'intersecció de la recta secant a la gràfica de la funció als punts x_i i x_{i+1} amb l'eix x :

$$y = f(x_i) + (x - x_i) \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i},$$
$$x_{i+2} = x_i - f(x_i) \frac{x_{i+1} - x_i}{f(x_{i+1}) - f(x_i)}.$$

Observació

Si $f(x_{i+1}) = -f(x_i)$ obtenim $x_{i+2} = \frac{x_{i+1} + x_i}{2}$ com a bisecció.

Mètode de Regula Falsi 1/2

Algorisme (Mètode de Regula Falsi)

procedure REGULAFALSI(f , a , b , Tolx, Tolf)

$f_a \leftarrow f(a)$; $f_b \leftarrow f(b)$;

▷ Inicialització

if $fabs(f_a) < Tolf$ **then return** a ; **end if**

▷ Control: a és la solució

if $fabs(f_b) < Tolf$ **then return** b ; **end if**

▷ Control: b és la solució

if $f_a * f_b > 0$ **then Stop**(Error: No és un interval vàlid); **end if**

while $b - a > Tolx$ **do**

▷ Iteració fins trobar l'arrel amb precisió Tolx

$m \leftarrow a - f_a \frac{b - a}{f_b - f_a}$; $f_m \leftarrow f(m)$;

▷ Punt Regula Falsi

if $fabs(f_m) < Tolf$ **then return** m ; **end if**

▷ Control: m és la solució

if $f_a * f_m < 0$ **then**

▷ El canvi de signe és a $[a, m]$

$b \leftarrow m$; $f_b \leftarrow f_m$;

▷ Redefinició de l'interval $[a, b]$

else

▷ Ara el canvi de signe és a $[m, b]$

$a \leftarrow m$; $f_a \leftarrow f_m$;

▷ Redefinició de l'interval $[a, b]$

end if

end while

return $a - f_a \frac{b - a}{f_b - f_a}$;

▷ Retornem el punt mig de l'interval final

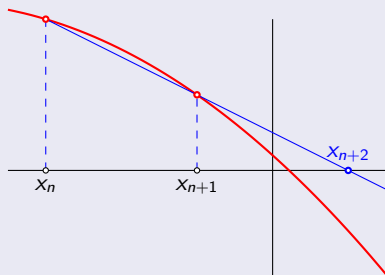
end procedure

Mètode de Regula Falsi 2/2

Mètode de la secant

És com el mètode de Regula Falsi però sense controlar que l'interval contingui un canvi de signe. Usarem la recta secant entre dues iteracions consecutives. És més eficient que els anteriors però no tenim certesa de convergència.

Representació gràfica del mètode de la secant



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

Algorisme (Mètode de la secant)

Escollim

- x, y aproximacions inicials d' α ,
- ε , tolerància, i
- n_{\max} , nombre màxim d'iterats que permetrem

i fem

procedure SECANT($f, x, y, \varepsilon, n_{\max}$)

$fx \leftarrow f(x); fy \leftarrow f(y);$

▷ Inicialització

for $i = 0 \div n_{\max} - 1$ **do**

▷ Bucle d'iteracions permeses

$z \leftarrow x - fx \frac{y - x}{fy - fx}; fz \leftarrow f(z);$

▷ Nova iteració

if $\text{fabs}(fz) \leq \varepsilon$ **then return** z **end if**

▷ Control de convergència

$x \leftarrow y; y \leftarrow z; fx \leftarrow fy; fy \leftarrow fz;$

▷ Actualització

end for

Error(No s'ha trobat la solució amb n_{\max} iteracions);

end procedure

Definició

Donada una funció $f: \mathbb{R} \rightarrow \mathbb{R}$ prou diferenciable, direm que α és una **arrel d'f de multiplicitat m** si

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

Notació: Si $m = 1$ l'arrel s'anomena **simple**, si $m = 2$ l'arrel s'anomena **doble**, etc

Observació (que justifica el terme *multiplicitat*)

Si α té multiplicitat m , el Teorema de Taylor ens diu:

$$\begin{aligned} f(x) &= f(\alpha) + \sum_{i=1}^m \frac{f^i(\alpha)}{i!} (x - \alpha)^i + \frac{f^{m+1}(\xi(x))}{(m+1)!} (x - \alpha)^{m+1} \\ &= (x - \alpha)^m \left[\frac{f^m(\alpha)}{m!} + \frac{f^{m+1}(\xi(x))}{(m+1)!} (x - \alpha) \right] = (x - \alpha)^m g(x) \end{aligned}$$

amb $g(\alpha) = \frac{f^m(\alpha)}{m!} \neq 0$.

El mètode de Newton permet aproximar arrels d'equacions de la forma $f(x) = 0$. Anem a deduir-lo de dues maneres. Suposem que $f: [a, b] \rightarrow \mathbb{R}$ és de classe \mathcal{C}^2 i que existeix $\alpha \in (a, b)$ amb $f(\alpha) = 0$.

Dedució analítica com a mètode de Taylor directe: Suposem que x és una aproximació d' α , i expandim f per Taylor fins a ordre 1 al voltant de x :

$$\begin{aligned} 0 = f(\alpha) &= f(x) + f'(x)(\alpha - x) + \mathcal{O}((\alpha - x)^2) \\ &\approx f(x) + f'(x)(\alpha - x), \end{aligned}$$

d'on
$$\alpha - x \approx -\frac{f(x)}{f'(x)}.$$

Així, $x - f(x)/f'(x) \approx x + (\alpha - x) = \alpha$ i esperem que

$$x_{n+1} = x_n - f(x_n)/f'(x_n)$$

sigui millor aproximació de α que x .

Això suggereix el següent algorisme iteratiu:

Algorisme (Mètode de Newton)

Escollim

- x , aproximació inicial d' α ,
- ε , tolerància (fita de $|e_a(x_n, \alpha)|$), i
- n_{\max} , nombre màxim d'iterats que permetrem

procedure NEWTON($f, x, \varepsilon, n_{\max}$)

for $i = 0 \div n_{\max} - 1$ **do**

$$x_{\text{new}} \leftarrow x - \frac{f(x)}{f'(x)}$$

if $\text{fabs}(x_{\text{new}} - x) \leq \varepsilon$ **then return** x_{new} **end if**

$$x \leftarrow x_{\text{new}}$$

end for

Error(No s'ha trobat la solució amb n_{\max} iteracions);

end procedure

▷ Bucle d'iteracions permeses

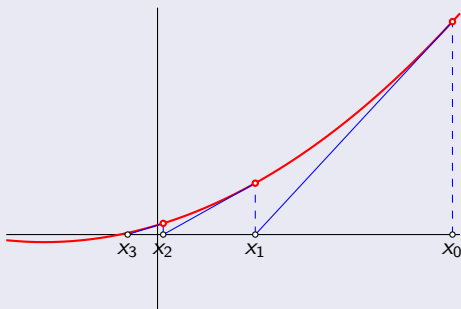
▷ Iteració del mètode de Newton

▷ Control de convergència

▷ Actualització

Mètode de Newton (cont.)

Representació gràfica del mètode de Newton



Donat x_i , aproximació d' α , prenem la recta tangent a f a x_i :

$$y = f(x_i) + f'(x_i)(x - x_i),$$

la tallem amb l'eix x (fem $y = 0$ a dalt): $x = x_i - \frac{f(x_i)}{f'(x_i)}$

i prenem $x_{i+1} := x$ com a aproximació millorada.

Nota:

Òbviament les dues deduccions donen la mateixa fórmula, donat que l'aproximació de Taylor d'ordre 1 de f al voltant de x no és més que la recta tangent a la gràfica de f al punt $(x, f(x))$.

Mètode de Newton 3/4

Observació

El mètode de Newton és el límit del mètode de la secant quan els punts x i y s'acosten:

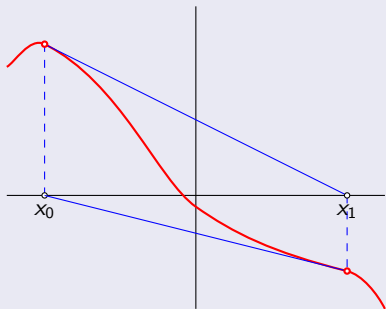
$$\lim_{x_{n-1} \rightarrow x_n} \left(x_{n-1} - f(x_{n-1}) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \right) = x_n - \frac{f(x_n)}{f'(x_n)}.$$

El mètode de Newton és molt més ràpid que el mètode de bisecció si $f'(\alpha) \neq 0$ (i.e., si α és una arrel simple). Veurem més endavant que, quan som a prop de l'arrel, a cada iterat aproximadament es duplica nombre de decimals correctes. En aquestes condicions $|x_{i+1} - x_i| \approx |\alpha - x_i|$, i és per això que a l'algorisme Newton la condició $|x_{i+1} - x_i| < \varepsilon$ és part del criteri d'aturada.

No obstant, quan som lluny de l'arrel el mètode de Newton pot convergir molt lentament (o no convergir), i és per això que a l'algorisme Newton es limita el nombre màxim d'iterats.

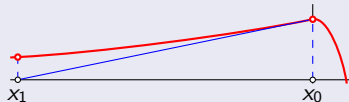
El mètode de Newton no sempre convergeix

Òrbita periòdica (de període 2)



Es perd

Aquest cas s'il·lustra analíticament a l'exemple següent



Exercici (per lliurament suplementari)

Construir exemples analítics en que el mètode de Newton tingui òrbites periòdiques de període el més alt possible (2,3,4,...).

Exemple comparatiu: bisecció versus Newton

Suposem que volem trobar un zero α de $f(x) = e^x - 2$ (que ja sabem que és $\ln 2$). Si prenem $a = 0$ i $b = 1$, com que $f(a)f(b) < 0$, el teorema de Bolzano ens assegura que tenim una arrel $\alpha \in [a, b]$.

Iterats de bisecció

i	m_i	i	m_i
1	0.5	\vdots	\vdots
2	0.75	30	0.6931 4717 9670
3	0.625	31	0.6931 4718 0135
4	0.6875	32	0.6931 4718 0370
\vdots	\vdots	33	0.6931 4718 0485
10	0.6923 8281 2500	\vdots	\vdots
11	0.6928 7019 3750	35	0.6931 4718 0570
\vdots	\vdots	36	0.6931 4718 0555
20	0.6931 4670 5625	37	0.6931 4718 0560
21	0.6931 4718 2460	38	ja no canvia

Iterats de Newton

i	x_i
0	0.5
1	0.7130 6131 9425
2	0.6933 4415 7318
3	0.6931 4719 9958
4	0.6931 4719 0558
5	ja no canvia

Observem que Newton és molt més ràpid.

Bisecció versus Newton 1/3

Exemple comparatiu: bisecció versus Newton (cont.)

Ara prenem $a = -9$ i $b = 1$ a l'exemple anterior ($\Rightarrow f(a)f(b) < 0$):

Iterats de bisecció

i	m_i	i	m_i
1	-4	21	0.6931 5052 0320
2	-1.5	⋮	⋮
3	-0.25	30	0.6931 5717 6870
4	0.375	31	0.6931 5718 1530
5	0.6875	⋮	⋮
⋮	⋮	39	0.6931 5718 0570
10	0.6972 5652 5	40	0.6931 5718 0560
11	0.6923 8281 25	41	0.6931 5718 0555
⋮	⋮	42	0.6931 5718 0560
20	0.6931 5528 8670	43	ja no canvia

Iterats de Newton

i	x_i
0	-4
1	104.1963 0006 6
2	103.1963 0006 6
3	102.1963 0006 6
⋮	⋮
103	2.3213 8700 863
⋮	⋮
107	0.6936 4539 4385
108	0.6931 4730 4647
109	0.6931 4718 0562
110	ja no canvia

Observem que Newton precisa més iterats que bisecció!

Exemple comparatiu: bisecció versus Newton (cont.)

El motiu és que a $x = -4$ la gràfica de $f(x)$ és molt plana ($f'(-4) = e^{-4} \approx 0.0183$), de manera que la tangent per -4 envia el primer iterat al 104.1963.

A partir d'aquí, i fins que s'apropa a l'arrel, baixa d'un en un perquè

$$x_{n+1} = x_n - \frac{e^{x_n} - 2}{e^{x_n}} = x_n - 1 + \frac{2}{e^{x_n}} \approx x_n - 1,$$

donat que x_n és gros. Un cop som a prop de l'arrel, la convergència torna a ser ràpida, com en el cas anterior. Aquest és un comportament típic del mètode de Newton quan l'aproximació inicial és lluny de l'arrel.

Mètode de Newton: Deducció analítica com a mètode de Taylor invers

Suposem que α és una arrel simple de f : $f(\alpha) = 0$. Llavors, existeix $g = f^{-1}$, la inversa local d' f al voltant de 0: $\alpha = g(0)$. Suposem que x_n és una aproximació d' α . Tenim $x_n = g(y_n)$ on $y_n = f(x_n)$. Expandim $g(0)$ per Taylor fins a ordre 1 amb punt base y_n :

$$\alpha = g(0) = g(y_n) - g'(y_n)y_n + \mathcal{O}(y_n^2).$$

Per altra banda $f(g(y_n)) = y_n$ que, per la regla de la cadena, dóna

$$g'(y_n) = \frac{1}{f'(g(y_n))} = \frac{1}{f'(x_n)}.$$

Per tant, com que $g(y_n) = x_n$ i $y_n = f(x_n)$,

$$\alpha \approx x_{n+1} = g(y_n) - g'(y_n)y_n = x_n - \frac{f(x_n)}{f'(x_n)}.$$

x_{n+1} hauria de ser una millor aproximació d' α que x_n .

Exemple: Càlcul de \sqrt{a} a l'estil Babiloni (≈ 3000 anys)

Idea geomètrica

\sqrt{a} és el costat d'un quadrat d'àrea a .

Suposem que x_n és una aproximació de \sqrt{a} .

Això determina un rectangle d'àrea a i costats x_n i $\frac{a}{x_n}$. La solució s'obté quan tenim un quadrat; és a dir, quan $x_n = \frac{a}{x_n}$.

Que podem fer per a millorar l'aproximació de \sqrt{a} ? fer la mitjana dels dos costats

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right).$$

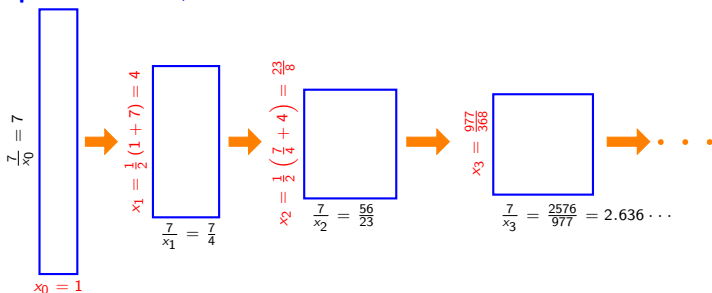
El mètode de Newton per a \sqrt{a} :

Prenem $f(x) = x^2 - a$. Clarament \sqrt{a} és la única arrel positiva d' f (simple).

El mètode de Newton és ($f'(x) = 2x$):

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^2 - a}{2x_n} = \frac{x_n^2 + a}{2x_n} \\ &= \frac{1}{2} \left(x_n + \frac{a}{x_n} \right). \end{aligned}$$

Exemple: càlcul de $\sqrt{7} = 2.64575131\dots$:



Exercici (per lliurament suplementari)

Donat $b \neq 0$ trobeu una funció ad-hoc que tingui $\frac{1}{b}$ com a zero simple i tal que el mètode de Newton aplicat a aquesta funció solament contingui sumes, restes i multiplicacions.

La utilitat d'aquest exercici ve del fet que, en alta precisió (per exemple 1000 xifres decimals), el quocient $\frac{a}{b}$ es calcula a partir del producte $a \cdot \frac{1}{b}$ i $\frac{1}{b}$ s'obté amb el mètode proposat a l'exercici prenent com a llavor el nombre $\text{fl}(1/b)$ en doble precisió.

Un *punt fix* per a una funció g definida a $[a, b]$ és un nombre $\alpha \in [a, b]$ tal que $g(\alpha) = \alpha$. En aquesta secció ens dedicarem a trobar solucions a problemes de punt fix (i.e., equacions de la forma $x = g(x)$) i a estudiar la connexió entre aquests i els problemes de trobar zeros de funcions. Per això passarem de l'equació $f(x) = 0$ a $g(x) = x$, i estudiarem quina és la forma més interessant des del punt de vista numèric.

Per exemple, si hem de resoldre $f(x) = e^x - 3x = 0$ (o $e^x = 3x$) podem passar a resoldre

- $g_1(x) = \frac{e^x}{3} = x$ o bé
- $g_2(x) = \log(3x) = x$ o bé
- $g_3(x) = \frac{e^x - x}{2} = x$ o bé
- $g_4(x) = e^x - 2x = x$ o bé
- $g_5(x) = x - \frac{e^x - 3x}{e^x - 3} = x$.

Iteració directa — Mètodes de punt fix (cont.)

A la dreta es pot veure una comparativa dels mètodes anteriors, en nombre d'iterats, per a calcular la solució de $e^x = 3x$ començant amb $x_0 = 0.6$ i amb precisió 10^{-8} .

Aquesta comparativa mostra que l'elecció del mètode no és arbitrària.

Comparativa dels mètodes

Mètode	# d'iterats
g_1	36
g_2	divergeix
g_3	21
g_4	10
g_5	3

Exercici (per lliurament suplementari)

Considerem la família de mètodes:

$$g_\mu(x) = \frac{e^x + \mu x}{3 + \mu}$$

amb $\mu \in \left\{ \frac{p}{q} : p \in \mathbb{Z}; q \in \{2, 3, 5\} \right\}$. Demostreu que qualsevol mètode d'aquesta família és equivalent a resoldre $e^x = 3x$ i determineu el millor mètode de la família per a resoldre aquest problema.

Iteració directa — Mètodes de punt fix (cont.)

Una equació $x = g(x)$ suggereix de manera natural la iteració:

Algorisme (Iteració directa)

Escollim

- x , aproximació inicial del punt fix α ,
- ε , tolerància (fita de $|e_a(x_n, \alpha)|$), i
- n_{\max} , nombre màxim d'iterats que permetrem

procedure ITERACIODIRECTA($f, x, \varepsilon, n_{\max}$)

for $i = 0 \div n_{\max} - 1$ **do**

$x_{\text{new}} \leftarrow g(x)$

if $\text{fabs}(x_{\text{new}} - x) \leq \varepsilon$ **then return** x_{new} **end if**

$x \leftarrow x_{\text{new}}$

end for

Error(No s'ha trobat la solució amb n_{\max} iteracions);

end procedure

▷ *Bucle d'iteracions permeses*

▷ *Iteració*

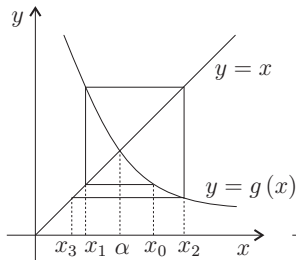
▷ *Control de convergència*

▷ *Actualització*

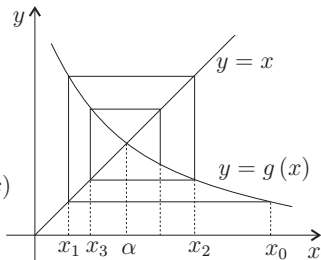
Ens preguntem: la successió d'iterats generada, convergeix? Si sí, a on? Sota quines condicions? A quina velocitat?

Seguidament estudiarem aquestes qüestions gràficament i analíticament.

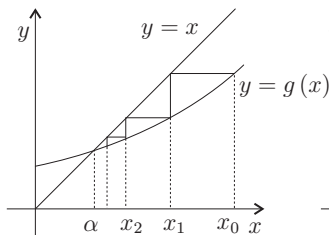
Iterant prop d'un punt fix hiperbòlic ($|g'(\alpha)| \neq 1$)



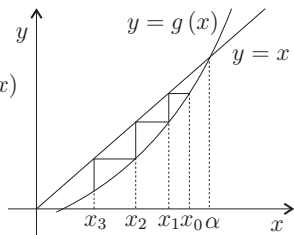
$$g'(\alpha) < -1$$



$$-1 < g'(\alpha) < 0$$



$$0 < g'(\alpha) < 1$$



$$1 < g'(\alpha)$$

Grafs *de teranyina*:

Donat x_0 , obtenim $x_1 = g(x_0)$ "pujant" a la gràfica de g , però el volem representar sobre l'eix de les x per poder tornar a iterar.

Per a això, ens desplacem des de $(x_0, g(x_0))$ fins a $y = x$ a través d'una recta horitzontal i "baixem" el valor $x_1 = g(x_0)$ a l'eix x .

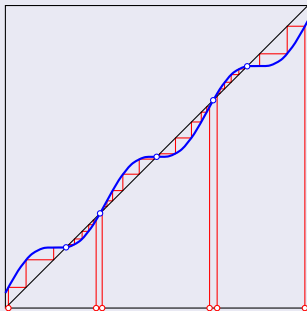
Fem el mateix amb x_1 , i així successivament.

Iterant prop d'un punt fix hiperbòlic 1/2

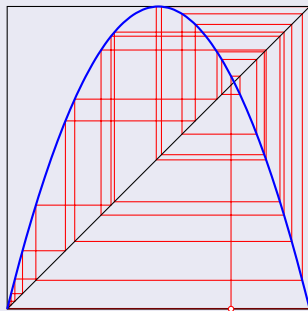
Iterant prop d'un punt fix hiperbòlic ($|g'(\alpha)| \neq 1$) (cont.)

- Les iteracions d'un mètode de punt fix es poden representar en una gràfica (sovint anomenada *de teranyina*) amb les corbes $y = g(x)$ i $y = x$.
- El "cas crític" és $|g'(\alpha)| = 1$: localment tenim convergència per $|g'(\alpha)| < 1$ i divergència per $|g'(\alpha)| > 1$.
- El cas $|g'(\alpha)| \neq 1$ s'anomena *hiperbòlic*.

En general la situació és molt més complicada



Fins i tot caòtica globalment:
 $f(x) = 4x(1 - x)$ a l'interval $[0, 1]$



Iterant prop d'un punt fix hiperbòlic 2/2

Exercici (per lliurament suplementari)

Estudieu els punts periòdics que pot tenir la recurrència

$g_\mu(x) = \mu x(1 - x)$ amb $\mu \in [0, 4]$ a l'interval $[0, 1]$ en funció de μ .

Definició

Donada $g: [a, b] \rightarrow [a, b] \subset \mathbb{R}$ denotarem: $g^0 = Id$, $g^1 = g$ i $g^n = g \circ g^{n-1}$ per $n \geq 2$ (és a dir, $g^n(x) = g(g^{n-1}(x))$).

Un punt $z \in [a, b]$ és *n-periòdic* si $g^n(z) = z$ i $g^j(z) \neq z$ per $j = 1, 2, \dots, n - 1$.

Els punts 1-periòdics són els punts fixos.

Consirem un espai mètric (E, d) on $d: E \times E \rightarrow \mathbb{R}^+$ és una distància:

- $d(x, y) = 0 \Leftrightarrow x = y$,
- $d(x, y) = d(y, x)$,
- $d(x, y) \leq d(x, z) + d(z, y)$

per a tot $x, y, z \in E$.

Exemples

- $E = \mathbb{R}$ i $d(x, y) = |x - y|$.
- $E = \mathbb{R}^n$ i $d(x, y) = \|x - y\|_2 = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$.
- $E = \{f: [a, b] \rightarrow \mathbb{R} \text{ cont.}\}$ i $d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$.
- E conjunt finit i $d(x, y) = \begin{cases} 1 & \text{si } x \neq y \\ 0 & \text{si } x = y \end{cases}$ (distància discreta).

Exercici (per lliurament suplementari)

Demostreu que la funció distància com a funció de l'espai producte amb una mètrica producte $(E \times E, d \times d)$ a \mathbb{R}^+ amb la mètrica Euclídia $(\mathbb{R}^+, |\cdot|)$:

$$\begin{aligned} d : (E \times E, d \times d) &\longrightarrow (\mathbb{R}^+, |\cdot|) \\ (x, y) &\longmapsto d(x, y) \end{aligned}$$

és una funció contínua.

Com a mètrica producte podeu prendre, per exemple, per simplicitat,

$$(d \times d)((x, y), (\tilde{x}, \tilde{y})) := \max\{d(x, \tilde{x}), d(y, \tilde{y})\},$$

encara que el resultat segueix essent cert amb qualsevol altra mètrica producte.

Definició

Direm que un espai mètric (E, d) és **complet** si tota successió de Cauchy és convergent a E .

Exemple

$(\mathbb{Q}, |x - y|)$ no és complet.

Definició

Sigui (E, d) un espai mètric i $g: E \rightarrow E$. Direm que g és una **contracció** o una **aplicació contractiva** si $\exists K \in \mathbb{R}$, $0 \leq K < 1$ tal que

$$d(g(x), g(y)) \leq Kd(x, y)$$

per a cada $x, y \in E$.

Notem que una contracció sempre és Lipschitz i, per tant, contínua.

Teorema (de l'Aplicació Contractiva)

Sigui (E, d) un espai mètric complet i $g: E \rightarrow E$ una aplicació contractiva. Llavors, g té un únic punt fix $\alpha \in E$ i, per a cada $x_0 \in E$,

$$\alpha = \lim_{n \rightarrow \infty} x_n \quad \text{on} \quad x_n = g(x_{n-1}) \text{ per a cada } n \in \mathbb{N}.$$

Exercici (per lliurament suplementari)

Sigui (E, d) un espai mètric complet i $g: E \rightarrow E$ una aplicació (no necessàriament contínua) tal que g^m és contractiva per algun $m \in \mathbb{N}$. Demostreu que g té un únic punt fix $\alpha \in E$ i, per a cada $x_0 \in E$, $\alpha = \lim_{n \rightarrow \infty} x_n$ on $x_n = g(x_{n-1})$ per a cada $n \in \mathbb{N}$. Recordeu que denotem $g^0 = Id$, $g^1 = g$ i $g^n = g \circ g^{n-1}$ per $n \geq 2$ (és a dir, $g^n(x) = g(g^{n-1}(x))$).

Demostració

Demostrem primer que $\{x_n\}_{n \in \mathbb{N}}$ és de Cauchy. Donat que g és contractiva,

$$d(x_n, x_{n+1}) \leq Kd(x_{n-1}, x_n) \leq \dots \leq K^n d(x_0, x_1).$$

Sigui $\varepsilon > 0$ i $q \geq 1$ arbitrari. Com que $K < 1$, existeix $p = p(x_0, x_1, K) \in \mathbb{N}$ (prou gran) tal que:

Prou gran: Per concretar podem prendre

$$p > \frac{\log(\varepsilon(1-K)/d(x_0, x_1))}{\log(K)}$$

$$\begin{aligned} d(x_p, x_{p+q}) &\leq \sum_{n=p}^{p+q-1} d(x_n, x_{n+1}) \leq d(x_0, x_1) \sum_{n=p}^{p+q-1} K^n \\ &= d(x_0, x_1) K^p \sum_{n=0}^{q-1} K^n \leq \frac{K^p}{1-K} d(x_0, x_1) < \varepsilon. \end{aligned}$$

Per tant, $\{x_n\}_{n \in \mathbb{N}}$ és de Cauchy i, com que E és complet, la successió té límit $\alpha = \lim_{n \rightarrow \infty} x_n$.

Demostració (cont.)

Considerem ara la successió $\{g(x_n)\}_{n \in \mathbb{N}}$. La continuïtat de g ens dóna:

$$g(\alpha) = g\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} g(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = \alpha,$$

Per tant, α és un punt fix de g .

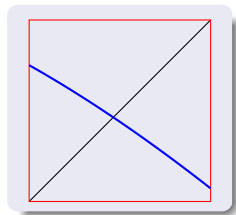
Finalment hem de veure la unicitat del punt fix. Suposem que n'hi ha dos: $g(\alpha) = \alpha$, $g(\beta) = \beta$ i $d(\alpha, \beta) \neq 0$. Com que $K < 1$ tenim:

$$d(\alpha, \beta) = d(g(\alpha), g(\beta)) \leq Kd(\alpha, \beta) < d(\alpha, \beta);$$

contradició. □

Exemple d'utilització de l'aplicació contractiva

Volem trobar la solució de l'equació $x = \cos(x)$



Usarem el Teorema de l'Aplicació Contractiva amb $g(x) = \cos(x)$ a l'interval $[0.6, 0.9]$ (aquest interval es pot trobar buscant un canvi de signe de la funció $\cos(x) - x$ per Bolzano).

En aquest interval la funció $\cos(x)$ és decreixent i $g(0.6) \approx 0.82$ i $g(0.9) \approx 0.62$. Per tant, $g([0.6, 0.9]) \subset [0.6, 0.9]$.

Exemple d'utilització de l'aplicació contractiva (cont.)

Ara anem a veure que g és una contracció a $[0.6, 0.9]$: Pel Teorema del valor mig,

$$|g(x) - g(y)| \leq \max_{\xi \in [0.6, 0.9]} |g'(\xi)| |x - y|$$

i

$$K := \max_{\xi \in [0.6, 0.9]} |g'(\xi)| = \max_{\xi \in [0.6, 0.9]} |\sin(\xi)| \leq 0.79 < 1.$$

Quants iterats necessitarem per a calcular la solució amb precisió 10^{-8} començant a $x_0 = 0.75$? Usarem la fórmula

$d(x_n, \alpha) \leq \frac{K^n}{1-K} |x_1 - x_0| \leq 10^{-8}$ (que justificarem a continuació):

$$x_1 = \cos(x_0) \approx 0.7316888 \implies |x_1 - x_0| \approx 0.01831113.$$

i, per tant,

$$n \approx \frac{\log(10^{-8}(1 - 0.79)) - \log(0.01831113)}{\log(0.79)} \approx 67.79 < 68.$$

Lema (en les hipòtesis del Teorema de l'Aplicació Contractiva)

$$d(x_n, \alpha) \leq \frac{K^n}{1-K} d(x_1, x_0) \quad (\text{error a priori}),$$

$$d(x_n, \alpha) \leq \frac{K}{1-K} d(x_{n-1}, x_n) \quad (\text{error a posteriori}).$$

Demostració

De la demostració del Teorema de l'Aplicació Contractiva tenim:

$d(x_n, x_{n+q}) \leq \frac{K^n}{1-K} d(x_0, x_1)$. Per tant, donat que la funció distància és contínua,

$$d(x_n, \alpha) = \lim_{q \rightarrow \infty} d(x_n, x_{n+q}) \leq \frac{K^n}{1-K} d(x_0, x_1).$$

Demostració (cont.)

Ara demostrem la segona desigualtat:

$$\begin{aligned} d(x_n, x_{n+q}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{n+q-1}, x_{n+q}) \\ &\leq d(x_n, x_{n+1}) + Kd(x_n, x_{n+1}) + \cdots + K^{q-1}d(x_n, x_{n+1}) \\ &= d(x_n, x_{n+1}) \sum_{j=0}^{q-1} K^j \leq \frac{1}{1-K} d(x_n, x_{n+1}). \end{aligned}$$

Com abans, com que la funció distància és contínua,

$$\begin{aligned} d(x_n, \alpha) &= \lim_{q \rightarrow \infty} d(x_n, x_{n+q}) \leq \frac{1}{1-K} d(x_n, x_{n+1}) \\ &\leq \frac{K}{1-K} d(x_{n-1}, x_n). \end{aligned}$$



La primera de les fites anteriors és una *estimació a priori*, donat que es pot avaluar abans de dur a terme el procediment iteratiu. Ens permet predir el nombre d'iterats necessaris per a assolir una tolerància prefixada.

Concretament, si volem trobar l'arrel α amb error absolut més petit que ε , imposarem

$$\frac{K^n}{1-K} |x_1 - x_0| \leq \varepsilon \iff n \geq \frac{\log(\varepsilon(1-K)/|x_1 - x_0|)}{\log(K)}.$$

La segona fita és una *estimació a posteriori*, donat que, per poder avaluar-la, s'han de fer tots els iterats fins a x_n . És, de fet, un criteri de parada.

Notem que, si $K > 0.5$, llavors $|x_n - x_{n-1}| < \varepsilon$ (el criteri que fem servir al mètode de Newton) no és un criteri d'aturada acurat, donat que $\frac{K}{K-1} > 1$ i no podem assegurar que $|x_n - \alpha| < \varepsilon$.

Corol·lari del Teorema de l'Aplicació Contractiva

Corol·lari (del Teorema de l'Aplicació Contractiva)

Sigui $g: \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . Suposem que α és un punt fix de g ($g(\alpha) = \alpha$) i que $|g'(\alpha)| < 1$. Llavors, existeix $\varepsilon > 0$ tal que g és contractiva a l'interval $[\alpha - \varepsilon, \alpha + \varepsilon]$.

Demostració

Com que g és de classe \mathcal{C}^1 , existeixen K i $\varepsilon > 0$ tals que

$$|g'(\alpha)| < K < 1 \quad \text{i} \quad \max_{x \in [\alpha - \varepsilon, \alpha + \varepsilon]} |g'(x)| = K.$$

Aquesta tria d' ε i el Teorema del valor mig fixen $K < 1$ com a constant de contracció.

Per veure que g és contractiva a l'interval $[\alpha - \varepsilon, \alpha + \varepsilon]$ solament ens queda demostrar que l'interval $[\alpha - \varepsilon, \alpha + \varepsilon]$ és invariant per g . Sigui $x \in [\alpha - \varepsilon, \alpha + \varepsilon]$. Tenim,

$$|g(x) - \alpha| = |g(x) - g(\alpha)| = |g'(\xi_x)| |x - \alpha| \leq K |x - \alpha| < |x - \alpha|.$$

Per tant, $g(x) \in [\alpha - \varepsilon, \alpha + \varepsilon]$. □

Observació

Si, al corol·lari anterior, tenim $g'(x) \neq 0, \forall x \in [a, b]$, aleshores qualsevol successió generada per iteració directa dins de l'interval convergirà a α però cap dels seus elements pot “aterrar” a α (llevat que agafem el mateix α com a condició inicial).

Ho provem per reducció a l'absurd. Sigui $\{x_n\}_{n \in \mathbb{Z}^+}$ una successió generada per iteració directa amb $x_0 \neq \alpha$, i sigui k el primer subíndex tal que $x_{k-1} \neq g(x_{k-1}) = x_k = \alpha$. Llavors, $g(x_k) = x_k$ i

$$0 = g(x_k) - g(x_{k-1}) = g'(\xi_k)(x_k - x_{k-1}) \neq 0;$$

contradicció.

Així malgrat que x_n pot ésser tan a prop d' α com vulguem, mai arribarà a ser igual a α .

Notació (sobre la contractivitat local)

- El cas estudiat al corol·lari anterior, és dir $|g'(\alpha)| < 1$, s'anomena **punt fix atractor** per g . Tota successió d'iterats $x_{n+1} = g(x_n)$ a $[\alpha - \varepsilon, \alpha + \varepsilon]$ convergeix a α .
- Si $|g'(\alpha)| > 1$, una aplicació del Teorema del valor mig com la d'abans ens mostra que, localment,

$$|g(x) - \alpha| > |x - \alpha|$$

i diem que α és un **punt fix repulsor** per g . En particular, la successió d'iterats $x_{n+1} = g(x_n)$ **no** convergeix a α , per a tot x_0 dins un entorn arbitràriament petit de α . Tot i així, en aquest cas, $g'(\alpha) \neq 0$ i per tant,

$$|(g^{-1})'(\alpha)| = \frac{1}{|g'(\alpha)|} < 1$$

És a dir, α és un punt fix atractor per g^{-1} . Això ens dóna algorismes efectius de càlcul d' α suposant que la funció g^{-1} és efectivament calculable.

Notació (sobre la contractivitat local) 1/2

Notació (sobre la contractivitat local) (cont.)

• $|g'(\alpha)| = 1$. En aquest cas la successió d'iterats pot ser o no convergent, com mostren els següents exemples:

- ① $g(x) = \sin(x)$, $x \in [0, \frac{\pi}{2}]$. El zero és l'únic punt fix de g a l'interval especificat i es té $g'(0) = 1$. Per $x_0 \in (0, \frac{\pi}{2})$, la successió d'iterats:

$$x_{n+1} = g(x_n) = g(x_n) - g(0) = g'(c_n) \cdot x_n = \cos(c_n) \cdot x_n < x_n$$

és monòtona decreixent i està afitada inferiorment per zero. Per tant x_n és convergent a un punt fix que ha de ser zero (és l'únic punt fix disponible en aquest interval).

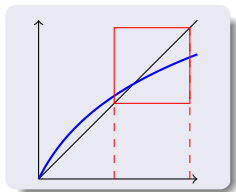
- ② $g(x) = \sinh(x) = \frac{e^x - e^{-x}}{2}$, $x \in [0, +\infty)$. El zero és un punt fix per g amb $g'(0) = 1$. Ara la successió d'iterats satisfà que per a cada n existeix c_n tal que

$$x_{n+1} = g(x_n) = \frac{e^{c_n} + e^{-c_n}}{2} \cdot x_n > x_n$$

i per tant la successió és monòtona creixent i no pot convergir al punt fix $\alpha = 0$.

Exemple sobre la contractivitat local

Suposem que volem resoldre $e^x - 2x - 1 = 0$, que podem reescriure com $x = g(x)$, on $g(x) = \ln(2x + 1)$.



Com que g és creixent, $g(1) = \ln 3 \approx 1.1 > 1$ i $g(2) = \ln 5 \approx 1.6 < 2$. Resumint, tenim $g([1, 2]) \subset [1, 2]$. La derivada, $g'(x) = 2/(2x + 1)$ és decreixent, d'on, per a $x \in [1, 2]$

$$|g'(x)| \leq |g'(1)| = \frac{2}{3} =: K < 1,$$

i per tant, segons el Teorema de l'Aplicació Contractiva, l'esquema iteratiu $x_{k+1} = g(x_k)$ és convergent a l'única arrel α . [Exemple sobre la contractivitat local 1/4](#)

Exemple sobre la contractivitat local (cont.)

Si prenem $x_0 = 1.5$, llavors

$$x_1 = g(x_0) \approx 1.3863 \implies |x_1 - x_0| \approx 0.1137.$$

Si volem trobar α amb error absolut fitat per $\varepsilon := 10^{-5}$ en tindrem prou de fer

$$\frac{\log(10^{-5}(1 - 2/3)) - \log(0.1137)}{\log(2/3)} \approx 25.74 < 26$$

iterats. Iterant amb 12 xifres decimals (només n'escrivim 6), surt

n	x_n	n	x_n	n	x_n	n	x_n
0	1.5	7	1.26052	14	1.25651	21	1.25643
1	1.38629	8	1.25876	15	1.25648	22	1.25643
2	1.32776	9	1.25775	16	1.25646	23	1.25643
3	1.29624	10	1.25718	17	1.25645	24	1.25643
4	1.27884	11	1.25686	18	1.25644	25	1.25643
5	1.26911	12	1.25668	19	1.25644	26	1.25643
6	1.26362	13	1.25657	20	1.25643		

Exemple sobre la contractivitat local 2/4

Exemple sobre la contractivitat local (cont.)

A la taula anterior podem observar que els iterats s'estabilitzen abans del previst. De fet, sembla que l'iterat 20 ja satisfà la tolerància demanada.

Per a explicar això, anem a fer servir l'estimació a posteriori trobada abans amb una petita trampa: substituïrem K per $|g'(x_n)|$. Això està justificat pel fet que, per cada n , podríem aplicar el teorema de l'aplicació contractiva amb l'interval

$$I_n := [\alpha - |x_n - \alpha|, \alpha + |x_n - \alpha|].$$

Anomenem $K_n := \sup_{x \in I_n} |g'(x)|$ el valor de K que hauríem d'emprar al teorema de l'aplicació contractiva. Com que $\lim_{n \rightarrow \infty} K_n = |g'(\alpha)| = \lim_{n \rightarrow \infty} |g'(x_n)|$, podem considerar $K_n \approx |g'(x_n)|$. Així, a cada iterat calculem també la següent aproximació de la fita a posteriori:

$$f_n := \frac{|g'(x_n)|}{1 - |g'(x_n)|} |x_n - x_{n-1}|.$$

Exemple sobre la contractivitat local 3/4

Els resultats són

n	x_n	f_n	n	x_n	f_n	n	x_n	f_n
0	1.5	_____	9	1.25775	1.32×10^{-3}	18	1.25644	8.30×10^{-6}
1	1.38629	1.28×10^{-1}	10	1.25718	7.52×10^{-4}	19	1.25644	4.73×10^{-6}
2	1.32776	7.07×10^{-2}	11	1.25686	4.28×10^{-4}	20	1.25643	2.69×10^{-6}
3	1.29624	3.96×10^{-2}	12	1.25668	2.44×10^{-4}	21	1.25643	1.53×10^{-6}
4	1.27884	2.23×10^{-2}	13	1.25657	1.39×10^{-4}	22	1.25643	8.72×10^{-7}
5	1.26911	1.27×10^{-2}	14	1.25651	7.90×10^{-5}	23	1.25643	4.97×10^{-7}
6	1.26362	7.18×10^{-3}	15	1.25648	4.50×10^{-5}	24	1.25643	2.83×10^{-7}
7	1.26052	4.08×10^{-3}	16	1.25646	2.56×10^{-5}	25	1.25643	1.61×10^{-7}
8	1.25876	2.32×10^{-3}	17	1.25645	1.46×10^{-5}	26	1.25643	9.17×10^{-8}

D'acord amb això, ens podríem aturar a l'iterat 18.

Exemple: Fites d'error i els criteris d'aturada

Suposem que volem trobar un zero de

$$f(x) = x - \sqrt{1 + x^2} + 0.05$$

i considerem la iteració simple $x_{k+1} = g(x_k)$ amb

$$g(x) = \sqrt{1 + x^2} - 0.05.$$

La solució de $x = g(x)$ es pot trobar exactament, i és

$$\alpha = 9.975.$$

Suposem que prèviament hem localitzat l'arrel entre 9 i 10, i prenem com a K una fita de $|g'(x)|$ a $[9, 10]$, i.e.,

$$K := |g'(10)| \approx 0.9950372.$$

Exemple: Fites d'error i els criteris d'aturada (cont.)

Suposem que volem trobar l'arrel amb tolerància $\varepsilon = 10^{-4}$, i iterem a partir de $x_0 = 9$:

n	x_{n-1}	x_n	$ x_n - x_{n-1} $	$\frac{K}{1-K} x_n - x_{n-1} $	$ x_n - \alpha $
1	9.0000 000	9.0053 851	0.0053 851	1.0797 135	0.9699 149
2	9.0053 851	9.0107 374	0.0053 522	1.0731 136	0.9642 626
⋮	⋮	⋮	⋮	⋮	⋮
300	9.7726 149	9.7736 450	0.0010 301	0.2013 550	0.2013 550
⋮	⋮	⋮	⋮	⋮	⋮
500	9.9015 003	9.9018 696	0.0003 693	0.0731 304	0.0731 304
⋮	⋮	⋮	⋮	⋮	⋮
758	9.9548 754	9.9549 760	0.0001 006	0.0201 648	0.0200 240
759	9.9549 760	9.9550 761	0.0001 001	0.0200 639	0.0199 239
760	9.9550 761	9.9551 756	0.0000 996	0.0198 244	0.0198 244
⋮	⋮	⋮	⋮	⋮	⋮
1500	9.9745 084	9.9745 108	0.0000 025	0.0004 916	0.0004 892
⋮	⋮	⋮	⋮	⋮	⋮
1817	9.9748 992	9.9748 997	0.0000 005	0.0001 008	0.0001 003
1818	9.9748 997	9.9749 002	0.0000 005	0.0001 002	0.0000 998
1819	9.9749 002	9.9749 007	0.0000 005	0.0000 997	0.0000 993

Exemple: Fites d'error i els criteris d'aturada 2/3

Exemple: Fites d'error i els criteris d'aturada (cont.)

Si féssim servir $|x_n - x_{n-1}| < \varepsilon$ com a criteri d'aturada, ens aturariem a l'iterat 760, per al qual la diferència entre x_n i l'arrel (darrera columna) és més gran que 0.01. En canvi, si fem servir $\frac{K}{1-K} |x_n - x_{n-1}| < \varepsilon$ com a criteri d'aturada, ens aturem a l'iterat 1819 i tenim la precisió que volíem.

Considerem l'equació $f(x) = x^3 - 4x + 1 = 0$.

És fàcil veure $f(x) = 0$ té tres zeros reals $\alpha_1 < \alpha_2 < \alpha_3$, que estan localitzats a $I_1 = [-2.5, -2]$, $I_2 = [0, 0.5]$ i $I_3 = [1.5, 2]$, respectivament.

L'equació $f(x) = 0$ es pot reescriure $x = \varphi(x)$ amb $\varphi(x) = \frac{1}{4}(x^3 + 1)$. Per tant $\varphi'(x) = \frac{3}{4}x^2$, $\varphi'' = \frac{3}{2}x$ i tenim els següents resultats.

- Sobre $I_1 = [-2.5, -2]$, $\varphi'(x) \geq \varphi'(-2) = 3$
- Sobre $I_2 = [0, 0.5]$, $\varphi'(0) = 0 \leq \varphi'(x) \leq \varphi'(\frac{1}{2}) = 0.1875$
- Sobre $I_3 = [1.5, 2]$, $\varphi'(1.5) = 1.6875 \leq \varphi'(x)$

Solament $\alpha_2 \in I_2$ és un punt fix *atractor* de φ .

Donat que $\varphi'(x) \geq 0$ per a tot $x \in I_2$, obtenim que φ és creixent i

$$\varphi(I_2) = [\varphi(0), \varphi(\frac{1}{2})] = [\frac{1}{4}, \frac{9}{32}] \subset [0, \frac{1}{2}] = I_2.$$

A més, $0 \leq \varphi'(x) \leq 0.1875 = K$.

Per tant, pel Teorema de l'Aplicació Contractiva, la successió $x_{n+1} = \varphi(x_n)$ convergeix al punt fix α_2 per a tot $x_0 \in I_2$. A més,

$$|x_n - \alpha_2| \leq \frac{(0.1875)^n}{0.8125} |x_1 - x_0|.$$

Exemple — diversos zeros (cont.)

Per a aproximar $\alpha_1 \in [-2.5, -2]$, prendrem la funció inversa de φ :
 $\psi(x) = \sqrt[3]{4x - 1}$.

Sigui $\bar{I}_1 = \varphi(I_1) = [-3.65625, -1.75]$, que satisfà $\psi(\bar{I}_1) = I_1 \subset \bar{I}_1$
i per a tot $x \in \bar{I}_1$:

$$0.0163819 \dots \simeq \psi'(-3.65625) \leq \psi'(x) \leq \psi'(-1.75) \simeq 0.0625$$

Per tant, $K = 0.0626$ i tindrem convergència usant el mètode ψ .

Per a trobar el tercer zero, $\alpha_3 \in I_3$, cal usar tècniques similars.

Començarem amb una observació i un parell d'exemples motivadors, en l'esperit del corol·lari anterior.

Observació

Sigui $g: \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 i suposem que α és un punt fix de g amb $|g'(\alpha)| < 1$. Pel Teorema del valor mig tenim,

$$\frac{x_{n+1} - \alpha}{x_n - \alpha} = \frac{g(x_n) - g(\alpha)}{x_n - \alpha} = g'(\xi_n),$$

amb ξ_n entre x_n i α . Per tant,

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{x_n - \alpha} = g'(\alpha) \iff |x_{n+1} - \alpha| \approx |g'(\alpha)| |x_n - \alpha|$$

i, per tant, $|g'(\alpha)|$ és una mesura de la *velocitat de convergència* de la successió $\{x_n\}_{n \in \mathbb{Z}^+}$.

Exemple (re-visitat)

Volem resoldre l'equació $e^x = 3x$ que té per solució:
 $\alpha = 0.6190612867\dots$

La taula següent mostra un estudi comparatiu de diverses funcions d'iteració, les seves derivades al punt fix i el nombre d'iterats requerits per iteració directa començant a $x_0 = 0.6$ amb precisió 10^{-8} . Observeu la correlació entre $|g'(\alpha)|$ i el nombre d'iterats.

Comparativa entre $|g'(\alpha)|$ i nombre d'iterats

Mètode	$ g'(\alpha) $	# d'iterats
$g_1(x) = \frac{e^x}{3}$	α	36
$g_2(x) = \log(3x)$	1.61...	divergeix
$g_3(x) = \frac{e^x - x}{3}$	0.42...	21
$g_4(x) = e^x - 2x$	-0.142...	10
$g_{\text{Newton}}(x) = x - \frac{e^x - 3x}{e^x - 3}$	0	3

Exemple (Newton)

Volem resoldre l'equació $f(x) = 0$ per a calcular un zero simple α d' f ($f'(\alpha) \neq 0$). El Mètode de Newton per aquest problema és:

$$g_{\text{Newton}}(x) = x - \frac{f(x)}{f'(x)}.$$

Com que $f(\alpha) = 0$, tenim

$$g'_{\text{Newton}}(\alpha) = 1 - \frac{(f'(\alpha))^2 - f(\alpha)f''(\alpha)}{(f'(\alpha))^2} = 1 - 1 = 0$$

Això mostra que la darrera línia de la taula anterior no és una casualitat pròpia de la funció f concreta considerada, i que el mètode de Newton *per zeros simples* sempre és molt ràpid.

Definició (Ordre de convergència)

Sigui $\{x_n\}_n \subset \mathbb{R}$ successió amb límit α . Direm que la successió té *ordre de convergència (exactament) p* (pot ser $p \notin \mathbb{N}$, $p > 0$) si $\exists C > 0$ tal que

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = C \quad (\neq 0, \infty).$$

La quantitat C s'anomena *constant asimptòtica de l'error*. En el cas $p = 1$, demanem, a més, que $C < 1$. Per $p = 1$ parlem de *convergència lineal*, per $p = 2$ de *convergència quadràtica* i, per $p = 3$, *cúbica*.

Si es compleix

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = 0$$

direm que la successió té *ordre de convergència almenys p* .

Observació

Suposem que $x_n \xrightarrow{n} \alpha$ té ordre p amb constant asimptòtica de l'error C . Si pensem x_n com a aproximació de α , llavors

$$C = \lim_{n \rightarrow \infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \lim_{n \rightarrow \infty} \frac{|e_a(x_{n+1}, \alpha)|}{|e_a(x_n, \alpha)|^p}.$$

Per tant, $|e_a(x_{n+1}, \alpha)| \approx C |e_a(x_n, \alpha)|^p$.

Formalment: $\forall \varepsilon > 0 \exists n_0 : \forall n \geq n_0$

$$|e_a(x_{n+1}, \alpha)| = (C + \varepsilon_n) |e_a(x_n, \alpha)|^p$$

amb $|\varepsilon_n| \leq \varepsilon$. Per tant, llevat de la constant C , el nombre de decimals correctes essencialment es multiplica per p a cada iterat (preneu logaritmes). És per això que podem considerar l'ordre de convergència com una mesura de la velocitat de convergència.

Exemple (Comprovació de l'observació anterior)

Prenem $p = 2$ i suposem que $|e_a(x_0, \alpha)| = 10^{-2}$.

$$|e_a(x_1, \alpha)| \approx C |e_a(x_0, \alpha)|^2 = C \cdot 10^{-4}$$

$$|e_a(x_2, \alpha)| \approx C |e_a(x_1, \alpha)|^2 = C \cdot C^2 \cdot 10^{-8}$$

$$|e_a(x_3, \alpha)| \approx C |e_a(x_2, \alpha)|^2 = C \cdot C^2 \cdot C^4 \cdot 10^{-16}$$

Exemple (Ordre de convergència d'algunes successions)

- La successió $x_n := (1/2)^n$ convergeix a 0. Té ordre 1:

$$\lim_{n \rightarrow \infty} \frac{|(1/2)^{n+1} - 0|}{|(1/2)^n - 0|} = \lim_{n \rightarrow \infty} \frac{1}{2} = \frac{1}{2},$$

i per tant la constant asimptòtica de l'error és $1/2$.

- La successió $x_n := (\frac{1}{3})^{2^n}$ convergeix a zero. Té ordre 2:

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - 0|}{|x_n - 0|^2} = \lim_{n \rightarrow \infty} \frac{(\frac{1}{3})^{2^{n+1}}}{((\frac{1}{3})^{2^n})^2} = \lim_{n \rightarrow \infty} \frac{(\frac{1}{3})^{2^{n+1}}}{(\frac{1}{3})^{2^{n+1}}} = 1,$$

i, per tant, la constant asimptòtica de l'error és 1.

Definició (Ordre de convergència d'un mètode iteratiu)

Sigui $g: [a, b] \rightarrow \mathbb{R}$, $\alpha = g(\alpha)$ punt fix. Direm que el mètode iteratiu

x_0 aproximació inicial

$\forall n = 0, 1, 2, \dots$

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

té ordre p si existeix $\varepsilon > 0$ tal que $\forall x_0 \in (\alpha - \varepsilon, \alpha + \varepsilon)$ es compleix que $x_{n+1} = g(x_n) \xrightarrow{n} \alpha$ amb ordre de convergència p .

A continuació anem a obtenir condicions suficients per a que un mètode iteratiu tingui ordre de convergència p .

Seguidament estudiarem l'ordre de convergència d'alguns mètodes concrets.

Condicó suficient d'ordre de convergència p

Teorema (Condicó suficient d'ordre de convergència p)

Sigui $g: \mathbb{R} \rightarrow \mathbb{R}$ de classe \mathcal{C}^p i sigui α un punt fix de g . Suposem que

$$g'(\alpha) = g''(\alpha) = \dots = g^{(p-1)}(\alpha) = 0$$

amb $|g'(\alpha)| < 1$ si $p = 1$. Aleshores, la funció d'iteració g dóna lloc a un mètode iteratiu d'ordre almenys p per trobar α .

Si, a més, $g^{(p)}(\alpha) \neq 0$, l'ordre del mètode iteratiu donat per g és exactament p , amb constant asimptòtica de l'error $C = \frac{|g^{(p)}(\alpha)|}{p!}$.

Observació

Si $0 \neq |g'(\alpha)| < 1$, el mètode $x_{n+1} = g(x_n)$ té ordre de convergència lineal.

Exercici (per lliurament suplementari)

Demostreu que l'ordre de convergència del mètode de la secant és $p = \frac{1+\sqrt{5}}{2}$.

Demostració

En aquestes hipòtesis, pel corol·lari de l'aplicació contractiva, existeix $\varepsilon > 0$ tal que g és contractiva a l'interval $[\alpha - \varepsilon, \alpha + \varepsilon]$. Per altra banda, pel Teorema de Taylor en dimensió 1,

$$\begin{aligned}g(x_n) &= g(\alpha) + g'(\alpha)(x_n - \alpha) + \cdots + \frac{g^{(p-1)}(\alpha)}{(p-1)!}(x_n - \alpha)^{p-1} + \\ &\quad \frac{g^{(p)}(\xi_{x_n})}{p!}(x_n - \alpha)^p \\ &= \alpha + \frac{g^{(p)}(\xi_{x_n})}{p!}(x_n - \alpha)^p\end{aligned}$$

per a tot $x_n \in [\alpha - \varepsilon, \alpha + \varepsilon]$ i ξ_{x_n} entre α i x_n .

Demostració (cont.)

Per tant, com que g és de classe \mathcal{C}^p ,

$$\frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \frac{|g^{(p)}(\xi_{x_n})|}{p!} \xrightarrow{n \rightarrow \infty} \frac{|g^{(p)}(\alpha)|}{p!} := C.$$

Si $|g^{(p)}(\alpha)| = 0$ llavors $C = 0$ i el mètode té ordre almenys p .

Si $|g^{(p)}(\alpha)| \neq 0$ llavors $C \neq 0$ i el mètode té ordre exactament p (en el cas $p = 1$ tenim $C = |g'(\alpha)| < 1$). □

Teorema (Ordre de convergència del mètode de Newton)

Sigui $f: [a, b] \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 , $f(\alpha) = 0$ i $f'(\alpha) \neq 0$ (i.e. α és un zero simple de f). Si, $f''(\alpha) \neq 0$, aleshores el mètode de Newton per trobar α com a zero de $f(x) = 0$ és quadràtic amb constant asimptòtica de l'error $C = \frac{1}{2} \left| \frac{f''(\alpha)}{f'(\alpha)} \right|$. Si, $f''(\alpha) = 0$ el mètode és al menys cúbic.

Si $f \in \mathcal{C}^{m+2}$, $f'(\alpha) = 0$ i α és un zero d' f de multiplicitat $m > 1$, llavors el mètode de Newton és lineal.

Demostració

El mètode de Newton consisteix a iterar

$$g_N(x) := x - \frac{f(x)}{f'(x)},$$

a partir d'una llavor inicial x_0 .

D'altra banda,

$$g'_N(x) = 1 - \frac{(f'(x))^2 - f(x)f''(x)}{(f'(x))^2} = \frac{f(x)f''(x)}{(f'(x))^2} ;$$

$$g''_N(x) = \frac{(f'(x)f''(x) + f(x)f'''(x))(f'(x))^2 - f(x)f''(x)2f'(x)f''(x)}{(f'(x))^4}$$

Demostració (cont.)

Com que $f(\alpha) = 0$, si $f'(\alpha) \neq 0$, tenim $g_N(\alpha) = \alpha$, $g'_N(\alpha) = 0$ i $g''_N(\alpha) = f''(\alpha)/f'(\alpha)$. Si $f''(\alpha) = 0$ llavors, pel **Teorema de la condició suficient d'ordre de convergència p** , el mètode de Newton és al menys cúbic. Si, en canvi, $f''(\alpha) \neq 0$, tenim que $g''_N(\alpha) \neq 0$ i, per tant, el mètode de Newton és exactament quadràtic i té constant asimptòtica de l'error $C = \frac{1}{2} \left| \frac{f''(\alpha)}{f'(\alpha)} \right|$.

Demostració (cont.)

Suposem ara que α és un zero d' f de multiplicitat $m > 1$. Tenim:

$$f(x) = (x - \alpha)^m h(x) \quad \text{amb} \quad h \in \mathcal{C}^2 \text{ i } h(\alpha) \neq 0.$$

Ara,

$$g_N(\alpha) = \alpha - \lim_{x \rightarrow \alpha} \frac{f(x)}{f'(x)}$$

i

$$\begin{aligned} \frac{f(x)}{f'(x)} &= \frac{(x - \alpha)^m h(x)}{m(x - \alpha)^{m-1} h(x) + (x - \alpha)^m h'(x)} \\ &= \frac{(x - \alpha) h(x)}{m h(x) + (x - \alpha) h'(x)} \xrightarrow{x \rightarrow \alpha} \frac{0}{m h(\alpha)} = 0. \end{aligned}$$

Per tant, $g_N(\alpha) = \alpha$.

Demostració (cont.)

Pel que fa a la derivada:

$$g'_N(x) = \frac{f(x)f''(x)}{(f'(x))^2} = \frac{(x - \alpha)^m h(x)f''(x)}{(m(x - \alpha)^{m-1}h(x) + (x - \alpha)^m h'(x))^2} \text{ on}$$
$$f''(x) = m(m - 1)(x - \alpha)^{m-2}h(x) + m(x - \alpha)^{m-1}h'(x) +$$
$$m(x - \alpha)^{m-1}h'(x) + (x - \alpha)^m h''(x).$$

Per tant, dividint el numerador i el denominador de $g'_N(x)$ per $(x - \alpha)^{2m-2}$ i agrupant termes obtenim

$$g'_N(x) = \frac{h(x) [m(m - 1)h(x) + 2m(x - \alpha)h'(x) + (x - \alpha)^2 h''(x)]}{[mh(x) + (x - \alpha)h'(x)]^2}.$$

Demostració (cont.)

Llavors,

$$g'_N(\alpha) = \lim_{x \rightarrow \alpha} g'_N(x) = \frac{m(m-1)h^2(\alpha)}{m^2 h^2(\alpha)} = 1 - \frac{1}{m}.$$

Quan $m > 1$ tenim $\frac{1}{m} \leq \frac{1}{2}$ i $0 < g'_N(\alpha) < 1$. Per tant, el mètode de Newton és lineal. □

En el cas de zeros de multiplicitat $m > 1$ el mètode de Newton es pot modificar per a recuperar la convergència quadràtica.

En les hipòtesis esmentades es defineix el *mètode de Newton modificat* com

$$g_{Nm}(x) = x - m \frac{f(x)}{f'(x)}.$$

Proposició

El mètode de Newton modificat té convergència al menys quadràtica.

Demostració

Observem que

$$g_{Nm}(x) = x - m \frac{f(x)}{f'(x)} = x(1 - m) + m \left(x - \frac{f(x)}{f'(x)} \right)$$

$$= x(1 - m) + mg_N(x) \text{ i}$$

$$g'_{Nm}(x) = (1 - m) + mg'_N(x)$$

Llavors, usant els càlculs anteriors,

$$g_{Nm}(\alpha) = \alpha(1 - m) + mg_N(\alpha) = \alpha(1 - m) + m\alpha = \alpha \text{ i}$$

$$g'_{Nm}(\alpha) = (1 - m) + mg'_N(\alpha) = (1 - m) + m \left(1 - \frac{1}{m} \right) = 0$$

i el **Teorema de la condició suficient d'ordre de convergència** **p** ens assegura que el mètode de Newton modificat té convergència al menys quadràtica. □

Lema Tècnic (Control de l'error en funció de l'ordre)

Suposem que $g: (\alpha - \varepsilon, \alpha + \varepsilon) \rightarrow \mathbb{R}$, i que existeixen $\varepsilon > 0$, $K > 0$ i $p > 1$ tals que

$$|g(x) - \alpha| \leq K|x - \alpha|^p$$

si $|x - \alpha| < \varepsilon$ (és a dir, $g(x) - \alpha = \mathcal{O}(|x - \alpha|^p)$). Aleshores, g dóna lloc a un mètode d'ordre almenys p , i $x_{n+1} = g(x_n)$ convergeix per a $|x_0 - \alpha| < \min(K^{-1/(p-1)}, \varepsilon)$. A més, tenim l'estimació

$$|x_n - \alpha| \leq K^{\frac{-1}{p-1}} (K^{\frac{1}{p-1}} |x_0 - \alpha|)^{p^n}.$$

Demostració

Si $|x - \alpha| < K^{-1/(p-1)} \iff |x - \alpha|^{p-1} < K^{-1}$,

$$|g(x) - \alpha| \leq K|x - \alpha|^p = (K|x - \alpha|^{p-1})|x - \alpha| < |x - \alpha|,$$

d'on, si x_0 és t.q. $|x_0 - \alpha| < \min(\varepsilon, K^{-1/(p-1)})$ i $x_{n+1} = g(x_n)$, llavors $|x_n - \alpha| < \min(\varepsilon, K^{-1/(p-1)}) \forall n$. A més,

$$\begin{aligned} |x_n - \alpha| &\leq K|x_{n-1} - \alpha|^p \\ &\leq K(K|x_{n-2} - \alpha|^p)^p = K^{1+p}|x_{n-2} - \alpha|^{p^2} \\ &\leq K^{1+p}(K|x_{n-3} - \alpha|^p)^{p^2} = K^{1+p+p^2}|x_{n-3} - \alpha|^{p^3} \\ &\leq \dots \leq K^{1+p+\dots+p^{n-1}}|x_0 - \alpha|^{p^n} = K^{\frac{p^n-1}{p-1}}|x_0 - \alpha|^{p^n} \\ &= K^{\frac{-1}{p-1}} \left(K^{\frac{1}{p-1}} |x_0 - \alpha| \right)^{p^n} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

d'on $x_n \xrightarrow{n} \alpha$ amb ordre almenys p . □

Exemple realista: aplicant el mètode de Newton

Imaginem-nos que estem implementant l'aritmètica de punt flotant d'un processador, que ja tenim fetes la suma, el producte i la divisió i volem fer l'arrel quadrada. Podem emprar el mètode de Newton. Anem a provar la convergència del mètode de Newton, a partir d'una aproximació inicial adequada, en un nombre d'iterats "raonable" i independent del número del que busquem l'arrel quadrada.

Suposem que volem trobar \sqrt{c} , on $c = mb^q > 0$, amb m mantissa normalitzada ($m \geq b^{-1}$). Això és equivalent a trobar el zero positiu de $f(x) = x^2 - c$. Els iterats del mètode de Newton són

$$x_{n+1} = g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)} = \frac{1}{2} \left(x_n + \frac{c}{x_n} \right).$$

Exemple realista: aplicant el mètode de Newton (cont.)

Una aproximació inicial de l'arrel que podem trobar sense fer cap operació és

$$x_0 = b^{q/2}$$

donat que, si q és parell $b^{q/2}$ es pot calcular trivialment, i si no, només cal tenir emmagatzemat $b^{1/2}$. Aleshores

$$\alpha = \sqrt{c} = \sqrt{m} b^{q/2} < b^{q/2} = x_0,$$

d'on $x_0 > \alpha$. A més, pel Teorema de Taylor i usant que $g(\alpha) = \alpha$,

$$\forall x > \alpha, \quad g(x) - \alpha = \frac{g''(\xi_x)}{2}(x - \alpha)^2 > 0 \implies g(x) > \alpha,$$

donat que $g''(x) = c/x^3 > 0$ si $x > 0$. A més, com que $g''(\xi)$ és decreixent a $[\alpha, \infty)$ podem prendre com a constant asimptòtica de l'error

$$K = \max_{\xi \in [\alpha, \infty)} \frac{g''(\xi)}{2} = \frac{c}{2\alpha^3} = \frac{1}{2\sqrt{c}}.$$

Exemple realista: aplicant el mètode de Newton 2/5

Ara, d'acord amb el **Lema Tècnic sobre el control de l'error en funció de l'ordre**

$$|x_n - \alpha| \leq K^{-1}(K |x_0 - \alpha|)^{2^n}, \quad (2)$$

però, per tenir convergència, cal que $K |x_0 - \alpha| < 1$:

$$\begin{aligned} K |x_0 - \alpha| &= \frac{1}{2\sqrt{c}}(x_0 - \sqrt{c}) = \frac{1}{2m^{1/2}b^{q/2}}(b^{q/2} - m^{1/2}b^{q/2}) \\ &= \frac{1 - m^{1/2}}{2m^{1/2}} \leq \frac{1 - b^{-1/2}}{2b^{-1/2}} = \frac{\sqrt{b} - 1}{2}. \end{aligned}$$

Per $b = 2$ tenim $(\sqrt{b} - 1)/2 \approx 0.207$ i per tant tindrem convergència. Per a $b = 10$ tindríem $(\sqrt{b} - 1)/2 > 1$, i per tant no podríem assegurar convergència (necessitaríem una condició inicial x_0 més acurada).

Suposem que estem implementant l'aritmètica IEEE doble, per la qual $b = 2$. Substituint el valor de K i $K |x_0 - \alpha|$ en (2), obtenim

$$\frac{|x_n - \sqrt{c}|}{\sqrt{c}} \leq 2 \left(\frac{\sqrt{b} - 1}{2} \right)^{2^n},$$

d'on, si volem $|e_r(x_n, \sqrt{c})| < \varepsilon$, on $\varepsilon > 10^{-16}$ és l'èpsilon-màquina de l'aritmètica IEEE doble, només cal demanar

$$n > \log_2 \left(\frac{-16 - \log_{10}(2)}{\log_{10}((\sqrt{b} - 1)/2)} \right) \approx 4.5.$$

Exemple realista: aplicant el mètode de Newton (cont.)

Per tant, *per a qualsevol número màquina* $c = mb^q$, prenent $x_0 = b^{q/2}$ i fent 5 iterats de Newton obtenim $\text{fl}(\sqrt{c})$ amb error relatiu més petit que l'èpsilon-màquina¹. Cada iterat de Newton només consta d'una divisió i una suma (no comptem la divisió per dos perquè, en una aritmètica binària, només cal decrementar l'exponent en una unitat).

¹Estem suposant que no cometem errors a les operacions. Aquesta suposició és realista si les operacions a nivell intern es fan amb una aritmètica més precisa que la IEEE doble, cosa que passa a la pràctica totalitat dels processadors actuals.

Suposem que $f \in \mathcal{C}^3(\mathbb{R}, \mathbb{R})$ i que α és una arrel simple de f : $f(\alpha) = 0$. Denotem per $h = f^{-1}$, la inversa local d' f al voltant de 0: $\alpha = h(0)$.

Tal com varem fer a la deducció analítica del mètode Newton com a mètode de Taylor invers suposem que x_n és una aproximació d' α . Tenim $x_n = h(y_n)$ on $y_n = f(x_n)$. Expandim $h(0)$ per Taylor fins a ordre 2 amb punt base y_n :

$$\alpha = h(0) = h(y_n) - h'(y_n)y_n + \frac{h''(y_n)}{2}y_n^2 + \mathcal{O}(y_n^3).$$

Per altra banda $f(h(y)) = y$ i, derivant per la regla de la cadena, tenim

$$\begin{aligned}f'(h(y))h'(y) &= 1 \text{ i} \\f''(h(y))(h'(y))^2 + h''(y)f'(h(y)) &= 0.\end{aligned}$$

Per tant, donat que $h(y_n) = x_n$,

$$h'(y_n) = \frac{1}{f'(h(y_n))} = \frac{1}{f'(x_n)} \text{ i}$$
$$h''(y_n) = -\frac{f''(h(y_n))}{(f'(h(y_n)))^3} = -\frac{f''(x_n)}{(f'(x_n))^3}.$$

Llavors, com que $y_n = f(x_n)$,

$$\alpha \approx x_{n+1} := h(y_n) - h'(y_n)y_n + \frac{h''(y_n)}{2}y_n^2$$
$$= x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(x_n)(f(x_n))^2}{2(f'(x_n))^3},$$

que suggereix que x_{n+1} hauria de ser una millor aproximació d' α que x_n .

Proposició

El Mètode de Txebishev $g_T(x) = x - \frac{f(x)}{f'(x)} - \frac{f''(x)(f(x))^2}{2(f'(x))^3}$ té ordre de convergència al menys 3 per a zeros simples.

Demostració

Sigui α un zero simple d' f ($f(\alpha) = 0$ i $f'(\alpha) \neq 0$). Llavors, clarament, $g_T(\alpha) = \alpha$, que és equivalent a resoldre $f(\alpha) = 0$.

Anem a veure que el mètode és convergent. Farem servir el **Corol·lari del Teorema de l'Aplicació Contractiva** que diu que si $|g_T'(\alpha)| < 1$ llavors g_T és localment contractiva.

Demostració (cont.)

Tenim,

$$g'_T(x) = \frac{f''(x)f(x)}{(f'(x))^2} + \frac{f'''(x)(f(x))^2 + 2f''(x)f'(x)f(x)}{2(f'(x))^3} - \frac{3(f''(x))^2(f(x))^2}{2(f'(x))^4}.$$

Donat que els denominadors són diferents de zero al punt α i que els tres numeradors tenen $f(x)$ (que s'anul·la al punt α) multiplicant, és obvi que $g'_T(\alpha) = 0$ i el mètode és convergent.

Ara ens cal veure que té ordre 3. Ho podríem fer demostrant que $g''_T(\alpha) = 0$ però és més fàcil a partir de la definició usant el Teorema de Taylor.

Demostració (cont.)

Precisant el terme d'error al desenvolupament de Taylor anterior tenim

$$\alpha = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(x_n)(f(x_n))^2}{2(f'(x_n))^3} - \frac{h'''(\xi_{y_n})}{6} y_n^3 = x_{n+1} - \frac{h'''(\xi_{y_n})}{6} (f(x_n))^3.$$

Com que $f(\alpha) = 0$, pel Teorema del valor mig,

$$x_{n+1} - \alpha = \frac{h'''(\xi_{y_n})}{6} (f(x_n) - f(\alpha))^3 = \frac{h'''(\xi_{y_n})}{6} (f'(\nu_{x_n}))^3 (x_n - \alpha)^3.$$

Donat que el mètode és convergent, per a tot x_0 prop d' α tenim $x_n = g_T(x_{n-1}) \xrightarrow{n \rightarrow \infty} \alpha$. Per tant, $\nu_{x_n} \xrightarrow{n \rightarrow \infty} \alpha$ i $\xi_{y_n} \xrightarrow{n \rightarrow \infty} 0$. Així, com que $f \in \mathcal{C}^3$,

$$\frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^3} = \left| \frac{h'''(\xi_{y_n})}{6} (f'(\nu_{x_n}))^3 \right| \xrightarrow{n \rightarrow \infty} \frac{1}{6} |h'''(0)(f'(\alpha))^3|,$$

i Txebishev té ordre al menys 3. □

La comparació entre Newton i Txebishev planteja una pregunta natural: *quin mètode és millor?* El d'ordre més alt? La resposta òbvia a la pregunta és: el que convergeixi més ràpid en temps. Si una iteració de Txebishev necessita molt més temps de càlcul que una de Newton, encara que Txebishev necessiti menys iterats, pot ser més lent.

Una manera de formalitzar això és amb la següent:

Definició (eficiència computacional)

L'*eficiència computacional* és una funció $E(t, p)$ on t és el temps que tarda cada iterat del mètode i p és l'ordre del mètode.

La funció $E(t, p)$ ha de complir les següents propietats naturals:

- 1 $E(t, p)$ és creixent en p i decreixent en t .
- 2 $E(t, p) = E(mt, p^m)$ (l'eficiència es conserva a l'escalar el temps i l'orde del mètode de manera homogènia).

Funcions senzilles que compleixen les condicions anteriors són

$$E(t, p) = \frac{\log(p)}{t} \quad \text{i} \quad E(t, p) = p^{\frac{1}{t}}.$$

Prendrem la segona de les dues funcions.

Exemple (Comparem Newton i Txebishev al resoldre $\cos(x) = x$)

Tenim

$$f(x) = \cos(x) - x; \quad f'(x) = -\sin(x) - 1 \quad \text{i} \quad f''(x) = -\cos(x)$$

i les funcions d'iteració de Newton ($p = 2$) i Txebishev ($p = 3$) són:

$$g_N(x) = x - \frac{f(x)}{f'(x)} \quad \text{i} \quad g_T(x) = x - \frac{f(x)}{f'(x)} - \frac{f''(x)(f(x))^2}{2(f'(x))^3}$$

Exemple (Comparem Newton i Txebishev al resoldre $\cos(x) = x$)

Anem a comptar el cost computacional amb la taula següent

	+ -	× /	avaluacions trigonomètriques
f	1	0	1
f'	1	0	1
f''	0	0	1 (ja calculat abans)
g_N	3	1	2
g_T	4	7	2

Per a avaluar el temps normalment no es tenen en compte les sumes i restes (són molt ràpides) i, per simplificar, suposarem que tant els productes i quocients com les avaluacions trigonomètriques costen una unitat de temps (poc realista).

Exemple (Comparem Newton i Txebishev al resoldre $\cos(x) = x$)

Així els cost de Newton és de 3 unitats de temps mentre que el de Txebishev és de 9. Per tant,

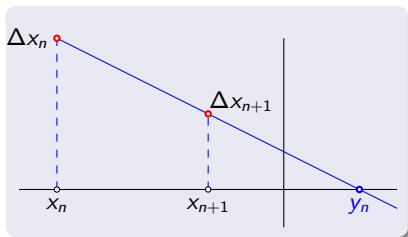
$$E_N = 2^{\frac{1}{3u}} > E_T = 3^{\frac{1}{9u}}.$$

Això explica perquè Newton és tan popular mentre que Txebishev es tan desconegut.

Mètode d'acceleració Aitken

Suposem que tenim un mètode que ens dona una successió $x_n \xrightarrow{n \rightarrow \infty} \alpha$. Volem accelerar la velocitat de convergència.

Denotem $\Delta x_n = x_{n+1} - x_n$. Podem pensar que Δx_n és una mesura de l'error de x_n , donat que $\Delta x_n \xrightarrow{n} 0$. Si considerem que a x_n cometem "error" Δx_n i a x_{n+1} cometem error Δx_{n+1} , podem fer "extrapolació dels errors cap a zero": prenem la recta per $(x_n, \Delta x_n)$ i $(x_{n+1}, \Delta x_{n+1})$ i la tallem amb l'eix x .



La recta esmentada és

$$\begin{aligned}y &= \Delta x_n + \frac{\Delta x_{n+1} - \Delta x_n}{x_{n+1} - x_n} (x - x_n) \\ &= \Delta x_n + \frac{\Delta^2 x_n}{\Delta x_n} (x - x_n),\end{aligned}$$

on hem denotat $\Delta^2 x_n := \Delta x_{n+1} - \Delta x_n = x_{n+2} - 2x_{n+1} + x_n$.

Si imposem $y = 0$, obtenim

$$\begin{aligned}x &= x_n - \Delta x_n \frac{\Delta x_n}{\Delta^2 x_n} \\&= x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n} \\&= \frac{x_n x_{n+2} - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n}\end{aligned}$$

(noteu que el que hem fet equival a fer un iterat del mètode de la secant a partir de $(x_n, \Delta x_n)$ i $(x_{n+1}, \Delta x_{n+1})$). Esperem que x sigui millor aproximació de α que x_n, x_{n+1}, x_{n+2} . D'acord amb això, podem considerar una nova successió “accelerada”:

$$y_n := x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} \quad (3)$$

(és per aquesta fórmula que aquest mètode es coneix com “la Δ^2 d'Aitken”).

Ens preguntem: està ben definida (el denominador es podria anul·lar)? Convergeix a α més ràpid que $\{x_n\}_n$?

Teorema

Suposem $x_n \xrightarrow{n \rightarrow \infty} \alpha$, $x_n \neq \alpha$ per tot n , i que existeix una constant C amb $|C| < 1$ tal que

$$x_{n+1} - \alpha = (C + \delta_n)(x_n - \alpha)$$

amb $\delta_n \xrightarrow{n} 0$. Aleshores la successió $\{y_n\}_n$ de (3) està ben definida per n prou gran, i

$$\lim_{n \rightarrow \infty} \frac{y_n - \alpha}{x_n - \alpha} = 0$$

(i.e., y_n tendeix a α més ràpid que x_n).

Observació

Si $g: [a, b] \rightarrow \mathbb{R}$ és de classe \mathcal{C}^1 , $g(\alpha) = \alpha$, $g'(\alpha) \neq 0$,
 $|g'(\alpha)| < 1$, llavors, per a x_0 prop de α i $x_{n+1} = g(x_n)$, si prenem
 $C = g'(\alpha)$ i

$$\delta_n := \frac{x_{n+1} - \alpha}{x_n - \alpha} - C$$

estem en les hipòtesis del teorema.

Demostració

Tenim

$$y_n = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}. \quad (4)$$

Definim

$$e_n := x_n - \alpha \implies e_{n+1} = (C + \delta_n)e_n.$$

Aleshores

$$\begin{aligned} x_{n+2} - 2x_{n+1} + x_n &= e_{n+2} - 2e_{n+1} + e_n \\ &= e_n \left((C + \delta_n)(C + \delta_{n+1}) - 2(C + \delta_n) + 1 \right) \\ &= e_n \left((C - 1)^2 + \underbrace{C(\delta_n + \delta_{n+1}) + \delta_n \delta_{n+1} - 2\delta_n}_{=: \mu_n \xrightarrow{n} 0} \right) \end{aligned}$$

Demostració (cont.)

Per n prou gran, $|\mu_n| < (C - 1)^2$, d'on $x_{n+2} - 2x_{n+1} + x_n \neq 0$ i, com que $e_n \neq 0$ per hipòtesi, y_n està ben definida.

També

$$x_{n+1} - x_n = e_{n+1} - e_n = e_n((C + \delta_n) - 1) = e_n((C - 1) + \delta_n).$$

Per tant

$$\frac{y_n - \alpha}{x_n - \alpha} = \frac{y_n - \alpha}{e_n} \stackrel{(4)}{=} 1 - \frac{1}{e_n} \frac{(e_n(C - 1) + \delta_n)^2}{e_n(C - 1)^2 + \mu_n} \xrightarrow{n \rightarrow \infty} 0,$$

donat que $\delta_n, \mu_n \xrightarrow{n} 0$. □

Mètode d'acceleració Steffensen

En comptes d'accelerar una successió ja generada, podem emprar la Δ^2 d'Aitken per a accelerar un mètode. Concretament, donat un mètode iteratiu

x_0 aproximació inicial

$$n \geq 0, \quad x_{n+1} = g(x_n)$$

podem considerar una nova iteració

$$y_0 = x_0$$

$$\begin{aligned} n \geq 0, \quad y_{n+1} &= y_n - \frac{(g(y_n) - y_n)^2}{g(g(y_n)) - 2g(y_n) + y_n} \\ &= \frac{y_n g(g(y_n)) - (g(y_n))^2}{g(g(y_n)) - 2g(y_n) + y_n}, \end{aligned}$$

que consisteix a aplicar la Δ^2 d'Aitken a y_n , $g(y_n)$ i $g(g(y_n))$. Aquesta estratègia és coneguda com a *acceleració d'Steffensen*.

Mètode d'acceleració Steffensen (cont.)

Steffensen usa els iterats accelerats mentres que Aitken sempre fa servir els iterats originals:

Exemple (Diferències entre Aitken i Steffensen)

Considerem el mètode d'iteració directa $g(x) = 1 + \frac{1}{x} + \frac{1}{x^2}$, que convergeix a $s = 1.839287$

x_n	g	Aitken		Steffensen
x_0	2.0000000000	2.0000000000	y_0	2.0000000000
x_1	1.7500000000		$g(y_0)$	1.7500000000
x_2	1.8979591837	1.8429487179	y_1	1.8429487179
x_3	1.8044860678	1.8406751916	$g(y_1)$	1.837032892
x_4	1.8612836844	1.8398159694	y_2	1.8392889123
x_5	1.8259158061	1.8394879875	$g(y_2)$	1.839285424
x_6	1.8476132213	1.8393634113	y_3	1.8392867552

Esperem que aquesta estratègia doni millors resultats que aplicar Aitken a una successió ja generada, donat que reciclem quantitats accelerades. Concretament, tenim el següent resultat.

Proposició

Sigui $g: [a, b] \rightarrow \mathbb{R}$ de classe \mathcal{C}^2 , $\alpha \in [a, b]$ amb $g(\alpha) = \alpha$, $g'(\alpha) \neq 1$. Aleshores, la funció d'iteració

$$G(x) = x - \frac{(g(x) - x)^2}{g(g(x)) - 2g(x) + x}$$

dóna un mètode almenys quadràtic.

Demostració

Anem a veure que podem suposar que $\alpha = 0$.

Per a α qualsevol, definim y i $h(y)$ per

$$\begin{aligned}x &=: \alpha + y, \\g(x) &=: \alpha + h(y),\end{aligned}$$

aleshores $h(0) = 0$ i $h'(0) \neq 1$.

Demostració (cont.)

Llavors, la funció d'iteració d'Steffensen que correspon a la funció h és

$$H(y) := y - \frac{(h(y) - y)^2}{h(h(y)) - 2h(y) + y}.$$

Observem que $g(g(x)) = \alpha + h(g(x) - \alpha) = \alpha + h(h(y))$. Per tant,

$$\begin{aligned} G(x) &= x - \frac{(g(x) - x)^2}{g(g(x)) - 2g(x) + x} \\ &= \alpha + y - \frac{(\alpha + h(y) - (\alpha + y))^2}{\alpha + h(h(y)) - 2(\alpha + h(y)) + \alpha + y} \\ &= \alpha + H(y). \end{aligned}$$

Demostració (cont.)

Suposem que hem demostrat que la funció d'iteració $H(y)$ dóna lloc a un mètode almenys quadràtic.

Donat x_0 qualsevol, prenem y_0 amb $x_0 = \alpha + y_0$, llavors, si definim $x_{n+1} = G(x_n)$ i $y_{n+1} = H(y_n)$, de la fórmula anterior, obtenim per inducció que $x_n = \alpha + y_n \forall n$, i, com que y_n té ordre de convergència almenys quadràtic, x_n també.

Per tant, n'hi ha prou de demostrar la proposició en el cas $\alpha = 0$.

Demostració (cont.)

Expandim g per Taylor al voltant de $\alpha = 0$:

$$g(x) = g(0) + g'(0)x + \frac{1}{2}g''(\xi_x)x^2 = Ax + \mathcal{O}(x^2), \quad (5)$$

amb a $A = g'(0) \neq 1$.

Recordem que es defineix

$$\begin{aligned} f(x) &= \mathcal{O}(g(x)) \quad \text{per } x \rightarrow a \\ \Leftrightarrow \exists K, \varepsilon > 0 : |f(x)| &\leq K |g(x)| \quad \text{per } |x - a| \leq \varepsilon. \end{aligned}$$

(en aquesta demostració, totes les \mathcal{O} seran per $x \rightarrow 0$).

Demostració (cont.)

Tenim

$$\begin{aligned}g(g(x)) &= g(Ax + \mathcal{O}(x^2)) \stackrel{(5)}{=} A(Ax + \mathcal{O}(x^2)) + \mathcal{O}(x^2) \\ &= A^2x + \mathcal{O}(x^2),\end{aligned}\tag{6}$$

$$\begin{aligned}(g(x))^2 &= (Ax + \mathcal{O}(x^2))^2 \stackrel{(5)}{=} A^2x^2 + 2Ax\mathcal{O}(x^2) + (\mathcal{O}(x^2))^2, \\ &= A^2x^2 + \mathcal{O}(x^3),\end{aligned}\tag{7}$$

on, en la darrera igualtat, hem usat que $x\mathcal{O}(x^2) = \mathcal{O}(x^3)$ i $(\mathcal{O}(x^2))^2 = \mathcal{O}(x^4) = \mathcal{O}(x^3)$.

Demostració (cont.)

Usant (6) i (7), obtenim (recordem que estem suposant que $\alpha = 0$)

$$\begin{aligned} G(x) &= \frac{xg(g(x)) - (g(x))^2}{g(g(x)) - 2g(x) + x} \\ &= \frac{x(A^2x + \mathcal{O}(x^2)) - A^2x^2 - \mathcal{O}(x^3)}{A^2x + \mathcal{O}(x^2) - 2Ax - 2\mathcal{O}(x^2) + x} \\ &= \frac{\mathcal{O}(x^3)}{(A-1)^2x + \mathcal{O}(x^2)} \\ &= \frac{x\mathcal{O}(x^2)}{(A-1)^2x + x\mathcal{O}(x)} = \frac{\mathcal{O}(x^2)}{(A-1)^2 + \mathcal{O}(x)} \\ &\stackrel{A \neq 1}{=} \mathcal{O}(x^2). \end{aligned}$$

Per tant, d'acord amb el **Lema Tècnic sobre el control de l'error en funció de l'ordre**, la funció d'iteració G dona lloc a un mètode d'ordre almenys 2. □

Observació

- 1 Noteu que, a la proposició anterior, no ens cal demanar $|g'(\alpha)| < 1$. D'aquí se segueix que *el mètode de Steffensen fa convergents esquemes iteratius divergents*.
- 2 Es pot demostrar que, si una funció d'iteració g dóna lloc a un mètode d'ordre $p \geq 2$, aleshores l'acceleració de Steffensen produeix un mètode d'ordre $2p - 1$. En aquest cas no val la pena fer-la, donat que $g \circ g$ dóna lloc a un mètode d'ordre p^2 (comproveu-ho!!), i $g \circ g$ és un càlcul necessari en l'avaluació de G .

Exemple d'acceleració

Suposem que volem resoldre $x = \cos(x)$. Definim

$$f(x) := \cos(x) - x, \quad g(x) := \cos(x).$$

Com que $f(0.73) \approx 0.0152$ i $f(0.74) \approx -0.0015$, per Bolzano tenim que $\exists \alpha \in [0.73, 0.74] =: [a, b]$ tal que $f(\alpha) = 0 \Leftrightarrow g(\alpha) = \alpha$. Com que $g'(x) = -\sin(x)$ és decreixent i negativa a $[a, b]$, tenim que

$$|g'(x)| \leq |g'(0.74)| \leq 0.675 =: K.$$

Per tant, les iteracions amb $g(x) = \cos(x)$ són convergents a $[0.73, 0.74]$.

Anem a estimar el nombre d'iterats necessaris per tenir $|x_n - \alpha| \leq 10^{-8}$, suposant $x_0 = 0.735$ i $x_{n+1} = g(x_n) = \cos(x_n)$.

Exemple d'acceleració (cont.)

Tenim $x_1 \approx 0.7418\ 3076$, d'on $|x_1 - x_0| \approx 0.00683$, i, d'acord amb la fita de l'error a posteriori, volem $\frac{K^n}{1-K} 0.00683 \leq 10^{-8}$ que és equivalent a

$$n \geq \frac{-8 + \log(1 - K) - \log(0.00683)}{\log(K)} = 37.040,$$

i per tant calen 38 iterats.

Generem la successió:

n	x_n	n	x_n
0	0.735	5	0.7396 5084
1	0.7418 3075	\vdots	\vdots
2	0.7372 3287	35	0.7390 8514
3	0.7403 3157	36	0.7390 8513
4	0.7382 4494	37	0.7390 8514
		38	0.7390 8513

Exemple d'acceleració (cont.)

Usem Aitken:

$$y_n := x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n}.$$

Ara surt

n	y_n	n	y_n
0	0.7390 8265	6	0.7390 8511
1	0.7390 8401	7	0.7390 8512
2	0.7390 8462	8	0.7390 8513
3	0.7390 8490	\vdots	\vdots
\vdots	\vdots	36	0.7390 8513

En aquesta taula, els iterats a partir del 8 ja no canvien.

Si usem Steffensen:

$$z_0 = x_0, \quad z_{n+1} = z_n - \frac{(g(z_n) - z_n)^2}{g(g(z_n)) - 2g(z_n) + z_n},$$

surt

n	z_n
0	0.735
1	0.7390 8265
2	0.7390 8513
3	ja no canvia

Exemple d'acceleració (cont.)

Si iterem la funció inversa: $x_0 = 0.735$, $x_{n+1} = \arccos(x_n)$,
obtenim una seqüència divergent:

n	x_n	n	x_n
0	0.735	4	0.7191 3445
1	0.7451 2965	\vdots	\vdots
2	0.7300 6711	11	1.047 70623
3	0.7523 7617	12	complex

mentre que, si fem Steffensen:

$$z_0 = x_0, \quad z_{n+1} = z_n - \frac{(\arccos(z_n) - z_n)^2}{\arccos(\arccos(z_n)) - 2 \arccos(z_n) + z_n}$$

surt

n	z_n
0	0.735
1	0.7390 7308
2	0.7390 8513
3	no canvia

i encara tenim convergència quadràtica.

El càlcul de les arrels d'un polinomi a coeficients reals és un problema específic amb una metodologia estàndard. El Teorema fonamental de l'àlgebra ens diu que el nombre de solucions coincideix amb el grau del polinomi i que si tenim una arrel complexa també tenim la seva conjugada.

La metodologia de càlcul de les arrels consta de tres etapes:

- 1 Localització i comptatge del nombre d'arrels de certes classes (reals, positives, negatives — fitació amb resultats analítics)
- 2 Aïllament/separació de les arrels (Sturm)
- 3 Càlcul aproximat (Regula Falsi + Newton)

Abans de desenvolupar i estudiar aquesta estratègia farem un repàs dels (pocs) casos en que sabem trobar les solucions analíticament. La Teoria de Galois ens diu que no hi ha fórmules per a les solucions d'equacions polinòmiques de grau més gran que 4. Per altra banda la solució per grau 1 és trivial i per grau 2 és ben coneguda. A continuació estudiarem la solució de la cúbica (Tartaglia-Cardano) i de la quàrtica (Ferrari).

Solució de la cúbica. Fórmula de Tartaglia-Cardano

L'equació general de tercer grau

$$Ax^3 + Bx^2 + Cx + D = 0$$

amb $A, B, C, D \in \mathbb{R}$ i $A \neq 0$ es pot convertir a *forma normal* sense canviar les seves solucions dividint per A :

$$x^3 + ax^2 + bx + c = 0.$$

Seguidament, fent una *transformació de Tschirnhaus*, (que consisteix a substituir $x = z - \frac{a}{3}$) s'elimina el terme quadràtic de la forma normal i s'obté la *forma reduïda*:

$$z^3 + pz + q = 0$$

on

$$\begin{cases} p = b - \frac{a^2}{3} \\ q = c - \frac{ab}{3} + \frac{2a^3}{27} \end{cases}$$

En realitat es resol l'equació en forma reduïda, *desfent la substitució* $x = z - \frac{a}{3}$ per a trobar les solucions de l'equació en forma normal.

Solució de la cúbica. Fórmula de Tartaglia-Cardano 1/5

Solució de la cúbica. Fórmula de Tartaglia-Cardano (cont.)

Partint de la forma reduïda $z^3 + pz + q = 0$ es realitza una substitució del tipus $z = u + v$ i s'obté:

$$\begin{aligned}0 &= z^3 + pz + q = (u + v)^3 + p(u + v) + q \\ &= u^3 + 3uv(u + v) + v^3 + p(u + v) + q \\ &= u^3 + v^3 + (3uv + p)(u + v) + q\end{aligned}$$

Observem que si trobem una solució del sistema

$$\begin{cases} u^3 + v^3 + q = 0 \\ 3uv + p = 0 \end{cases} \iff \begin{cases} u^3 + v^3 = -q \\ u^3 v^3 = -\frac{p^3}{27} \end{cases}$$

obtenim una solució de l'equació reduïda. El darrer sistema és equivalent a afirmar que u^3 i v^3 són les solucions de l'equació

$$y^2 + qy - \frac{p^3}{27} = 0.$$

Solució de la cúbica. Fórmula de Tartaglia-Cardano (cont.)

Per tant,

$$u = \sqrt[3]{\frac{-q + \sqrt{\Delta}}{2}} \quad \text{i} \quad v = \sqrt[3]{\frac{-q - \sqrt{\Delta}}{2}}$$

on $\Delta = q^2 + 4\frac{p^3}{27}$ és el *discriminant*, i les solucions de l'equació en forma reduïda són:

$$z_0 = u + v$$

$$z_1 = ju + \bar{j}v \text{ i}$$

$$z_2 = j^2u + \bar{j}^2v = \bar{j}u + jv$$

amb

$$j = e^{i\frac{2\pi}{3}} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \quad (\text{observeu que } j^2 = \bar{j}).$$

En funció del signe del discriminant s'obtidran unes solucions o unes altres:

$\Delta > 0$: u i v són reals i diferents. Per tant z_0 és real i z_1 i z_2 són arrels complexes conjugades.

$\Delta < 0$: En aquest cas $v = \bar{u}$. Per tant obtenim tres solucions reals diferents (encara que l'expressió inicial sigui complexa). Les solucions es poden escriure en forma trigonomètrica:

$$z_k = 2\sqrt{\frac{-p}{3}} \cos\left(\frac{1}{3} \arccos\left(\frac{q}{2}\sqrt{\frac{27}{-p^3}}\right) + \frac{2k\pi}{3}\right) \quad \text{amb } k \in \{0, 1, 2\}.$$

$\Delta = 0 \iff q^2 = -4\frac{p^3}{27}$: De les fórmules anteriors obtenim

$$u = v = \sqrt[3]{\frac{-q}{2}} = \sqrt[3]{\frac{q^3}{2} \frac{27}{4p^3}} = \frac{3q}{2p}, \quad \text{sign}(u) = -\text{sign}(q),$$

$$u^6 = u^3 v^3 = -\frac{p^3}{27} \text{ i } p < 0 \iff u = \text{sign}(u) \sqrt{\frac{-p}{3}} = -\text{sign}(q) \sqrt{\frac{-p}{3}}$$

Llavors, com que $j + \bar{j} = -1$, l'equació té dues solucions reals, una de simple i una de doble:

$$z_0 = 2u = 2\sqrt[3]{\frac{-q}{2}} = -2\text{sign}(q) \sqrt{\frac{-p}{3}} = \frac{3q}{p} \text{ i}$$

$$z_1 = z_2 = -u = -\frac{1}{2}z_0 = \sqrt[3]{\frac{q}{2}} = \text{sign}(q) \sqrt{\frac{-p}{3}} = \frac{-3q}{2p}.$$

Nota:

$$\text{sign}(x) = \begin{cases} 1 & \text{si } x > 0, \\ 0 & \text{si } x = 0, \\ -1 & \text{si } x < 0. \end{cases}$$

Solució de la quàrtica. Fórmula de Ferrari

Comencem amb l'equació en *forma normal*

$$x^4 + ax^3 + bx^2 + cx + d = 0.$$

En aquest cas la *transformació de Tschirnhaus* per a obtenir la *forma reduïda* és $x = z - \frac{a}{4}$. Llavors la forma reduïda queda:

$$z^4 + pz^2 + qz + r = 0$$

on

$$\begin{cases} p = b - \frac{3a^2}{8} \\ q = c - \frac{ab}{2} + \frac{a^3}{8} \\ r = d - \frac{ac}{4} + \frac{ba^2}{16} - \frac{3a^4}{256} \end{cases}$$

Per a resoldre la quàrtica en forma reduïda primer s'escriu com $z^4 + pz^2 = -qz - r$ i es completa el quadrat per obtenir

$$z^4 + 2pz^2 + p^2 = pz^2 - qz - r + p^2 \iff (z^2 + p)^2 = pz^2 - qz - r + p^2.$$

Solució de la quàrtica. Fórmula de Ferrari (cont.)

Ara el truc: per a cada y tenim

$$\begin{aligned}(z^2 + p + y)^2 &= pz^2 - qz - r + p^2 + 2y(z^2 + p) + y^2 \\ &= (p + 2y)z^2 - qz + (p^2 - r + 2py + y^2)\end{aligned}\quad (8)$$

i el terme de la dreta és quadràtic en z . Per tant, podem triar y de manera que sigui un quadrat perfecte. Això es fa igualant el discriminant a zero:

$$\begin{aligned}0 &= (-q)^2 - 4(p + 2y)(p^2 - r + 2py + y^2) \\ &= -8y^3 - 20py^2 + (-16p^2 + 8r)y + (q^2 - 4p^3 + 4pr),\end{aligned}$$

que és una cúbica en y (anomenada *resolvent cúbica de la quàrtica*). Donat que la resolvent té grau 3 sempre té una arrel real α que podem calcular amb la fórmula de Tartaglia-Cardano.

Solució de la quàrtica. Fórmula de Ferrari (cont.)

Prenent $y = \alpha$, el polinomi $(p + 2y)z^2 - qz + (p^2 - r + 2py + y^2)$ té $\frac{q}{2(p+2\alpha)}$ com a arrel doble. Per tant,

$$(p + 2\alpha)z^2 - qz + (p^2 - r + 2p\alpha + \alpha^2) = \mu^2 \left(z - \frac{q}{2\mu^2}\right)^2 = \left(\mu z - \frac{q}{2\mu}\right)^2$$

amb $\mu = \sqrt{p + 2\alpha}$. Llavors, (8) queda

$$(z^2 + p + \alpha)^2 = \left(\mu z - \frac{q}{2\mu}\right)^2,$$

que ens dona les dues equacions quadràtiques següents:

$$\begin{cases} z^2 + p + \alpha = \mu z - \frac{q}{2\mu} \\ z^2 + p + \alpha = -\mu z + \frac{q}{2\mu} \end{cases} \iff z^2 \mp \mu z + \left(p + \alpha \pm \frac{q}{2\mu}\right) = 0$$

Cada una de les dues equacions ens dona dues solucions de l'equació reduïda.

Lema

Sigui $p(z) = a_0 + a_1z + \dots + a_nz^n$ amb $a_n \neq 0$ un polinomi arbitrari a coeficients complexos i sigui

$$\lambda := \max \left\{ \left| \frac{a_0}{a_n} \right|, \left| \frac{a_1}{a_n} \right|, \dots, \left| \frac{a_{n-1}}{a_n} \right| \right\}.$$

Lavors, $|s| \leq \lambda + 1$ per a tota arrel s de $p(z)$.

Demostració

$\lambda = 0$ implica que $p(z) = a_nz^n$, que té una arrel $s = 0$ de multiplicitat n . Clarament $|s| = 0 < 1 = \lambda + 1$ i el lema és cert. Suposem ara que $\lambda > 0$. L'equació $p(s) = 0$ és equivalent a $s^n + \frac{a_{n-1}}{a_n}s^{n-1} + \dots + \frac{a_1}{a_n}s + \frac{a_0}{a_n} = 0$.

Demostració (cont.)

Llavors,

$$\begin{aligned} |s|^n &\leq \left| \frac{a_{n-1}}{a_n} \right| |s|^{n-1} + \dots + \left| \frac{a_1}{a_n} \right| |s| + \left| \frac{a_0}{a_n} \right| \\ &\leq \lambda \left(|s|^{n-1} + \dots + |s| + 1 \right) = \lambda \frac{|s|^n - 1}{|s| - 1}. \end{aligned}$$

Ara suposem $|s| > \lambda + 1$ i arribem a contradicció. En aquest cas tenim $|s| - 1 > \lambda > 0$ i

$$0 < \frac{1}{|s| - 1} < \frac{1}{\lambda}.$$

Per tant,

$$|s|^n \leq \lambda \frac{|s|^n - 1}{|s| - 1} < |s|^n - 1;$$

contradicció. □

L'objectiu d'aquest apartat és *separar les arrels reals* d'una equació $f(x) = 0$. Això és, determinar intervals $I \subset \mathbb{R}$ que continguin una única arrel.

L'eina fonamental per a això seran les *successions de Sturm*.

Definició (Successió de Sturm)

Sigui $f: [a, b] \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 . Una successió finita de funcions contínues

$$f_0, \dots, f_m: [a, b] \rightarrow \mathbb{R}$$

es diu una **successió de Sturm** per f a $[a, b]$ si es compleixen les propietats següents:

- 1 $f_0 = f$,
- 2 $\alpha \in [a, b]$, $f_0(\alpha) = 0 \implies f'(\alpha)f_1(\alpha) > 0$,
- 3 $i = 1 \div m - 1$, $\alpha \in [a, b]$, $f_i(\alpha) = 0 \implies f_{i-1}(\alpha)f_{i+1}(\alpha) < 0$,
- 4 $f_m(x) \neq 0 \forall x \in [a, b]$.

En aquesta situació, per a cada $x \in [a, b]$, denotarem per $V(x)$ el nombre de canvis de signe de la successió

$$\{f_0(x), f_1(x), \dots, f_m(x)\}$$

(sense tenir en compte els valors zero).

L'aplicació de les successions de Sturm ve donada pel

Teorema (de Sturm)

Sigui $f : [a, b] \rightarrow \mathbb{R}$ de classe C^1 amb un nombre finit de zeros. Suposem, a més, que tots els zeros d' f són simples. Sigui f_0, \dots, f_m una successió de Sturm per a $f_0 = f$ a $[a, b]$. Aleshores, el nombre de zeros de f a $[a, b]$ és $V(a) - V(b)$.

Localització

Sigui $p(x)$ un polinomi i $\sigma = \lambda + 1$ on λ és la donada pel teorema de fitació d'arrels anterior, de manera que sabem que l'interval $[-\sigma, \sigma]$ conté totes les arrels reals de $p(x)$. Llavors, pel Teorema de Sturm $V(-\sigma) - V(\sigma)$ és exactament el nombre d'arrels reals de $p(x)$.

La *localització de les arrels de $p(x)$* consisteix a trobar exactament $V(-\sigma) - V(\sigma)$ subinterval·ls de l'interval $[-\sigma, \sigma]$ tals que cada un d'ells conté exactament una arrel. Això es pot fer automàticament usant tècniques de bisecció.

Observació

Com veurem més endavant, en el cas de polinomis i per a localització i càlcul d'arrels, podem esquivar fàcilment els efectes de la hipòtesi que tots els zeros d' f són simples

Calcul efectiu

El càlcul efectiu de cada arrel, donat l'interval de localització donat per l'aplicació del Teorema de Sturm, es fa amb una tècnica mixta. En primer lloc s'usa bisecció o Regula Falsi per a “aprimar” l'interval de localització fins a un interval *prou petit* $[a, b]$ perquè el mètode de Newton pugui convergir.

Seguidament s'usa el mètode de Newton per a aproximar ràpidament l'arrel amb precisió.

Un control de qualitat de que l'interval $[a, b]$ és prou petit, és que el Mètode de Newton convergeixi amb pocs iterats (4 o 5) a una (la única) arrel continguda a $[a, b]$.

Demostració (del Teorema de Sturm)

Suposem que movem x contínuament d' a a b . Cal veure que $V(x)$ decreix en una unitat cada cop que travessem un zero d' f .

Si $x \in (a, b)$ és tal que $f_0(x) \dots f_m(x) \neq 0$, com que $f_0 \dots f_m$ són contínues, tenim que $V(\cdot)$ és constant a un entorn de x .

Per tant, $V(\cdot)$ només pot canviar de valor a un punt \hat{x} tal que $f_i(\hat{x}) = 0$ amb $i \in \{0, \dots, m-1\}$ (no pot ser $f_m(\hat{x}) = 0$ per (4)).

Considerem un tal \hat{x} , i suposem primer que $i \geq 1$.

Demostració (del Teorema de Sturm — cont.)

Les possibilitats per als signes de $f_{i-1}(x)$, $f_i(x)$ i $f_{i+1}(x)$ a un entorn de \hat{x} són:

x	$f_{i-1}(x)$	$f_i(x)$	$f_{i+1}(x)$
$\hat{x} - \varepsilon$	+	?	-
\hat{x}	+	0	-
$\hat{x} + \varepsilon$	+	?	-

x	$f_{i-1}(x)$	$f_i(x)$	$f_{i+1}(x)$
$\hat{x} - \varepsilon$	-	?	+
\hat{x}	-	0	+
$\hat{x} + \varepsilon$	-	?	+

Els signes a la fila $x = \hat{x}$ se segueixen de la propietat (3).

A partir d'aquests, els signes de la primera i tercera columnes se segueixen per continuïtat de f_{i-1} i f_{i+1} , per a ε prou petit.

D'aquestes taules se segueix que $V(x)$ es manté constant quan travessem \hat{x} .

Demostració (del Teorema de Sturm — cont.)

Suposem ara $i = 0$. Aleshores, les possibilitats són:

x	$f_0(x)$	$f_1(x)$
$\hat{x} - \varepsilon$	+	-
\hat{x}	0	-
$\hat{x} + \varepsilon$	-	-

x	$f_0(x)$	$f_1(x)$
$\hat{x} - \varepsilon$	-	+
\hat{x}	0	+
$\hat{x} + \varepsilon$	+	+

Els signes a la columna $f_0(x)$ tenen en compte que els zeros d' f són simples i, per tant, f és localment estrictament monòtona.

Fixats aquestes signes, els signes a la fila $x = \hat{x}$, se segueixen de la propietat (2).

A partir d'aquests, els signes de la tercera columna se segueixen per continuïtat de f_1 per a ε prou petit.

En tots dos casos, $V(x)$ decreix en una unitat al passar per \hat{x} , tal com volíem veure. □

Ara que coneixem la utilitat de les successions de Sturm, necessitem una manera de generar-les. No existeix un procediment general per funcions qualssevol, però sí per polinomis usant l'algorisme de la divisió.

Sigui $p(x)$ polinomi de grau n *amb totes les arrels simples*. Definim

$$p_0(x) := p(x),$$

$$p_1(x) := p'(x),$$

$$p_2(x) : p_0(x) = p_1(x)q_1(x) - c_2p_2(x), \quad c_2 > 0, \quad \text{gr}(p_2) < \text{gr}(p_1),$$

$$p_3(x) : p_1(x) = p_2(x)q_2(x) - c_3p_3(x), \quad c_3 > 0, \quad \text{gr}(p_3) < \text{gr}(p_2),$$

⋮

$$p_m(x) : p_{m-2}(x) = p_{m-1}(x)q_{m-1}(x) - c_m p_m(x), \quad c_m > 0,$$

$$0 = \text{gr}(p_m) < \text{gr}(p_{m-1}) \quad (p_m(x) = \text{constant}),$$

(on $\text{gr}(\cdot)$ denota el grau d'un polinomi).

A la definició anterior, per a obtenir p_2 dividim p_0 per p_1 i prenem com a p_2 el reste de la divisió, però canviat de signe i multiplicat per una constant positiva, que podem triar lliurement. A continuació dividim p_1 per p_2 i prenem p_3 com el reste de la divisió, però canviat de signe i multiplicat per una constant. I així successivament, fins que arribem a $p_m(x) = \text{constant}$.

Notem que $\text{gr}(p_i) \leq n - i$, i per tant $m \leq n$ (i.e., com a molt podem fer n passos).

Notem també que, llevat del canvi de signe de la resta i el producte per la constant, aquest és l'algorisme d'Euclides per trobar el màxim comú divisor de polinomis. En particular,

$$\text{mcd}(p_0, p_1) = \text{mcd}(p, p') = \begin{cases} p_m(x) & \text{si } p_m(x) \neq 0, \\ p_{m-1}(x) & \text{si } p_m(x) = 0. \end{cases}$$

Proposició

Si $p(x)$ té totes les arrels simples, llavors la successió $\{p_0, \dots, p_m\}$ definida abans és una successió de Sturm per a p .

Observació

Suposem que $p(x)$ té una arrel α amb multiplicitat $m > 1$ (és a dir $p(x) = (x - \alpha)^m \tilde{p}(x)$). Clarament, $p'(x) = (x - \alpha)^{m-1} q(x)$ i, per tant, el polinomi $\text{mcd}(p, p')$ conté el factor $(x - \alpha)^{m-1}$. De fet, el $\text{mcd}(p, p')$ està format per tots els factors d'aquesta mena, potser multiplicat per una constant.

Per altra banda, quan $m = 1$ tenim que α no és arrel de p' . En particular, si $p(x)$ té totes les arrels simples llavors $p(x)$ i $p'(x)$ no tenen cap arrel en comú.

Demostració

Hem de comprovar que la successió compleix les propietats de les successions de Sturm. La propietat (1) es compleix per definició.

La hipòtesi que p té totes les arrels simples implica que p i p' no tenen cap arrel en comú (veure la observació anterior). En particular, si $p(\alpha) = 0$ tenim $p'(\alpha) \neq 0$. Llavors (2) és òbvia perquè $p_1 = p'_0$.

Hem vist abans que si $p_m(x) = 0$, llavors $\text{mcd}(p, p') = p_{m-1}(x)$ amb $\text{gr}(p_{m-1}) \geq 1$. Per altra banda, $p_{m-1}(x) = \text{mcd}(p, p')$ divideix $p(x)$ i $p'(x)$. Sigui α tal que $p_{m-1}(\alpha) = 0$. Clarament, $p(\alpha) = p'(\alpha) = 0$, que contradiu que les arrels de p són simples.

Demostració (cont.)

Ara demostrem (3). Suposem que $\alpha \in [a, b]$ és tal que $p_i(\alpha) = 0$ per cert $i \in \{1, \dots, m-1\}$. De la definició tenim que

$$p_{i-1}(\alpha) = p_i(\alpha)q_i(\alpha) - c_{i+1}p_{i+1}(\alpha) = -c_{i+1}p_{i+1}(\alpha), \quad c_{i+1} > 0.$$

Si $p_{i-1}(\alpha) \neq 0$, (3) és cert. Per tant, suposem que $p_{i-1}(\alpha) = 0$ i arribarem a contradicció. Aplicant recurrentment la definició del p_j 's tenim:

$$p_{i-2}(\alpha) = p_{i-1}(\alpha)q_{i-1}(\alpha) - c_i p_i(\alpha) = 0,$$

$$p_{i-3}(\alpha) = p_{i-2}(\alpha)q_{i-2}(\alpha) - c_{i-1}p_{i-1}(\alpha) = 0,$$

\vdots

que acaba amb $p_0(\alpha) = p_1(\alpha) = 0$. Això contradiu novament que les arrels de p són simples. □

Com que no podem saber a priori si $p(x)$ té o no arrels múltiples, el procediment efectiu de construcció d'una successió de Sturm per a localitzar les arrels reals de $p(x)$ és com segueix:

- Apliquem l'algorisme de càlcul dels polinomis p_i descrit abans.
- Si $p_m(x) \neq 0$ el polinomi p té totes les arrels simples i la successió que hem construït és de Sturm pel Teorema anterior.
- Si $p_m(x) = 0$ llavors $p_{m-1}(x) = \text{mcd}(p, p')$ i, per la observació anterior, $q(x) = p(x)/p_{m-1}(x)$ és un nou polinomi que té les mateixes arrels que p però totes elles simples. Apliquem la definició anterior a construir una successió de Sturm q_i pel polinomi q . Això ens permet localitzar les arrels de q que són les mateixes que p . Un cop les hem calculat, la seva multiplicitat la podem saber o bé per deflació o bé avaluant les derivades successives de $p(x)$ a cada una de les arrels.

Exercici (per lliurament suplementari)

Demostreu que si $p(x)$ té arrels múltiples, amb la notació de les planes anteriors, la successió donada per

$$q_0(x) = q(x) = p(x)/p_{m-1}(x) \text{ i}$$

$$q_i(x) = p_i(x)/p_{m-1}(x) \text{ per } i = 1, \dots, m-1$$

és de Sturm.

Exemple I de localització i separació d'arrels

Volem localitzar les arrels de

$$p(x) := x^3 - 3x^2 + x - 2.$$

Prenem

$$p_0(x) := p(x) = x^3 - 3x^2 + x - 2,$$

$$p_1(x) := p'_0(x) = 3x^2 - 6x + 1.$$

Lavors, si dividim p_0 per p_1 , obtenim

$$p_0 = p_1 q_1 + r_2 = p_1 q_1 - c_2 p_2,$$

amb quocient $q_1 = \frac{1}{3}x - \frac{1}{3}$ i resta $r_2 = -\frac{4}{3}x - \frac{5}{3}$. Amb la finalitat que la següent divisió sigui més fàcil, prenem $c_2 = \frac{1}{3}$ i $p_2 = 4x + 5$.

Exemple I de localització i separació d'arrels (cont.)

Dividim p_1 entre p_2 i obtenim

$$p_1 = p_2q_2 + r_3 = p_2q_2 - c_3p_3,$$

amb quocient $q_2 = \frac{3}{4}x - \frac{39}{16}$ i resta $r_3 = \frac{211}{16}$, d'on prenem $c_3 = \frac{211}{16}$ i $p_3 = -1$.

Com que p_3 és constant, diferent de zero, p té totes les arrels simples (si no hagués estat així, haguéssim hagut de tornar a començar amb p/p_3). Per tant, per la proposició anterior,

$$\{p_0, p_1, p_2, p_3\} = \{x^3 - 3x^2 + x - 2, 3x^2 - 6x + 1, 4x + 5, -1\}$$

és una successió de Sturm per p .

Exemple I de localització i separació d'arrels (cont.)

La fita donada pel teorema de fitació d'arrels anterior és $\sigma = \lambda + 1 = 4$ ($\lambda = 3$). Així, l'interval $[-4, 4]$ conté totes les arrels reals de $p(x)$.

Separem les arrels de $p(x)$ a la següent taula (notem que la successió de signes a ± 4 coincideix amb la de $\pm\infty$, més fàcilment calculable):

x	$-\infty$	-4	0	2	3	$+4$	$+\infty$
signe $p_0(x)$	-	-	-	-	+	+	+
signe $p_1(x)$	+	+	+	+	+	+	+
signe $p_2(x)$	-	-	+	+	+	+	+
signe $p_3(x)$	-	-	-	-	-	-	-
$V(x)$	2	2	2	2	1	1	1

Així $V(-\infty) - V(+\infty) = V(-4) - V(4) = 2 - 1 = 1$ i només tenim una arrel real. A més, com que $V(0) = 2$, $V(0) - V(+\infty) = 2 - 1 = 1$ deduïm que l'arrel és positiva.

Finalment, $V(2) - V(3) = 1 \implies$ l'arrel és a l'interval $(2, 3)$.

Exemple II de localització i separació d'arrels

Volem localitzar les arrels de

$$p(x) := x^3 + 2x^2 - 31x + 28$$

(que són -7 , 1 i 4). Prenem:

$$p_0(x) := p(x) = x^3 + 2x^2 - 31x + 28,$$

$$p_1(x) := p'_0(x) = 3x^2 + 4x - 31,$$

$$p_2(x) = x - \frac{157}{97} \text{ i}$$

$$p_3(x) = 1,$$

ja que

$$p_0(x) = p_1(x) \left(\frac{1}{3}x + \frac{2}{9} \right) - \frac{194}{9} \left(x - \frac{157}{97} \right) \text{ i}$$

$$p_1(x) = p_2(x) \left(3x + \frac{859}{97} \right) - \frac{156816}{9409}.$$

Exemple II de localització i separació d'arrels (cont.)

La fita donada pel teorema de fitació d'arrels és $\sigma = \lambda + 1 = 32$ ($\lambda = 31$). Per tant, l'interval $[-32, 32]$ conté totes les arrels reals de $p(x)$.

Separem les seves arrels mitjançant la següent taula:

x	$-\infty$	0	2	$+\infty$
signe $p_0(x)$	-	+	-	+
signe $p_1(x)$	+	-	-	+
signe $p_2(x)$	-	-	+	+
signe $p_3(x)$	+	+	+	+
$V(x)$	3	2	1	0

Així $V(-\infty) - V(+\infty) = 3$ i les tres arrels són reals. A més tenim:

$$V(-\infty) - V(0) = V(-32) - V(0) = 1 \implies 1 \text{ arrel a l'interval } (-32, 0),$$

$$V(0) - V(2) = 1 \implies 1 \text{ arrel a l'interval } (0, 2),$$

$$V(2) - V(+\infty) = V(2) - V(32) = 1 \implies 1 \text{ arrel a l'interval } (2, 32).$$

Exemple III de localització i separació d'arrels

Volem localitzar les arrels de

$$p(x) := x^8 - 2x^7 - x^6 + 4x^5 - 6x^4 + 8x^3 + 4x^2 - 16x + 8$$

(que són 1 —doble, $\pm\sqrt{2}$ —dobles— i $\pm i\sqrt{2}$).

Tenim:

$$p_0(x) := p(x) = x^8 - 2x^7 - x^6 + 4x^5 - 6x^4 + 8x^3 + 4x^2 - 16x + 8,$$

$$p_1(x) := p_0'(x) = 8x^7 - 14x^6 - 6x^5 + 20x^4 - 24x^3 + 24x^2 + 8x - 16,$$

$$p_2(x) = 11x^6 - 21x^5 + 38x^4 - 68x^3 - 60x^2 + 220x - 120,$$

$$p_3(x) = 59x^5 - 123x^4 - 52x^3 + 244x^2 - 132x + 4,$$

$$p_4(x) = -31x^4 + 67x^3 + 26x^2 - 134x + 72,$$

$$p_5(x) = -x^3 + x^2 + 2x - 2, \text{ i}$$

$$p_6(x) = 0,$$

ja que:

Exemple III de localització i separació d'arrels (cont.)

$$p_0(x) = p_1(x) \frac{4x - 1}{32} - \frac{1}{16} (11x^6 - 21x^5 + 38x^4 - 68x^3 - 60x^2 + 220x - 120)$$

$$p_1(x) = p_2(x) \frac{2(44x + 7)}{121} - \frac{64}{121} (59x^5 - 123x^4 - 52x^3 + 244x^2 - 132x + 4)$$

$$p_2(x) = p_3(x) \frac{649x + 114}{3481} - \frac{5808}{3481} (-31x^4 + 67x^3 + 26x^2 - 134x + 72)$$

$$p_3(x) = p_4(x) \left(-\frac{1829x + 140}{961} \right) - \frac{6962}{961} (-x^3 + x^2 + 2x - 2) \quad \text{i}$$

$$p_4(x) = p_5(x) (-(31x - 36))$$

Exemple III de localització i separació d'arrels (cont.)

Per tant,

$$p_5(x) = \text{mcd}\{p_0(x), p_1(x)\}$$

i, consegüentment (com que $p_1(x) = p_0'(x)$),

$$\begin{aligned} q(x) &= \frac{p_0(x)}{-p_5(x)} \\ &= \frac{x^8 - 2x^7 - x^6 + 4x^5 - 6x^4 + 8x^3 + 4x^2 - 16x + 8}{x^3 - x^2 - 2x + 2} \\ &= x^5 - x^4 - 4x + 4 = (x - 1)(x^2 - 2)(x^2 + 2). \end{aligned}$$

té les mateixes arrels que $p(x)$ però totes elles simples.

Exemple III de localització i separació d'arrels (cont.)

Ara, la Successió de Sturm per $q(x)$ és:

$$q_0(x) := q(x) = x^5 - x^4 - 4x + 4$$

$$q_1(x) := q'_0(x) = 5x^4 - 4x^3 - 4$$

$$q_2(x) = x^3 + 20x - 24 \quad \left(q_0(x) = q_1(x) \frac{5x - 1}{25} - \frac{4}{25} (x^3 + 20x - 24) \right)$$

$$q_3(x) = x^2 - 2x + 1 \quad \left(q_1(x) = q_2(x)(5x - 4) - 100 (x^2 - 2x + 1) \right)$$

$$q_4(x) = -23x + 26 \quad \left(q_2(x) = q_3(x)(x + 2) - (-23x + 26) \right)$$

$$q_5(x) = -1 \quad \left(q_3(x) = q_4(x) \left(-\frac{23x - 20}{529} \right) - \frac{9}{259} (-1) \right)$$

Exemple III de localització i separació d'arrels

La fita donada pel teorema de fitació d'arrels és $\sigma = \lambda + 1 = 5$ ($\lambda = 4$). Per tant, l'interval $[-5, 5]$ conté totes les arrels reals de $p(x)$.

Separem les seves arrels mitjançant la següent taula:

x	$-\infty$	-2	0	1.25	2	$+\infty$
signe $q_0(x)$	-	-	+	-	+	+
signe $q_1(x)$	+	+	-	+	+	+
signe $q_2(x)$	-	-	-	+	+	+
signe $q_3(x)$	+	+	+	+	+	+
signe $q_4(x)$	+	+	+	-	-	-
signe $q_5(x)$	-	-	-	-	-	-
$V(x)$	4	4	3	2	1	1

Així $V(-\infty) - V(+\infty) = 3$ i $q(x)$ ($p(x)$) té tres arrels reals. Una arrel és a l'interval $(-2, 0)$; una altra és a l'interval $(0, 1.25)$ i la tercera és a l'interval $(1.25, 2)$.

Definició

Donada una successió de nombres reals $\underline{a} = \{a_0, a_1, \dots, a_n\}$ denotarem per $\#cs(\underline{a})$ el nombre de canvis de signe (sense comptar zeros) de la successió \underline{a} .

Sigui $p(x)$ un polinomi de grau n . Per a cada $x \in \mathbb{R}$, denotarem $\delta_p(x) := \#cs\left(\{p(x), p'(x), p''(x), \dots, p^{(n)}(x)\}\right)$.

Teorema (de Budan-Fourier)

Sigui $p(x)$ un polinomi de grau n i sigui $[a, b]$ un interval tal que $p(a)p(b) \neq 0$. Aleshores, el nombre (comptant multiplicitats) de zeros de $p(x)$ a l'interval $[a, b]$ és

$$\delta_p(a) - \delta_p(b) - 2K \quad \text{amb} \quad K \in \mathbb{Z}^+.$$

Corol·lari (Regla de descartes)

Sigui $p(x) = a_0 + a_1x + \dots + a_nx^n$ un polinomi amb $a_0, a_n \neq 0$ (és a dir, el polinomi té grau n i $p(0) \neq 0$). Aleshores, el nombre (comptant multiplicitats) de zeros **positius** de $p(x)$ és

$$\#\text{cs}(\{a_0, a_1, \dots, a_n\}) - 2K \quad \text{amb } K \in \mathbb{Z}^+.$$

Observació

Fent el canvi $x \mapsto -x$ obtenim el polinomi $p^-(x) = a_0 - a_1x + \dots + (-1)^n a_n x^n$ i les arrels negatives de $p(x)$ són les positives de $p^-(x)$. Per tant, el nombre (comptant multiplicitats) de zeros **negatius** de $p(x)$ és

$$\#\text{cs}(\{a_0, -a_1, a_2, \dots, (-1)^{n-1} a_{n-1}, (-1)^n a_n\}) - 2K \quad \text{amb } K \in \mathbb{Z}^+.$$

Observació (Aplicació de la Regla de Descartes a un interval finit)

Farem com a exemple el cas de l'interval $(0, 1)$ i usarem un canvi de variables adequat. La funció

$$\begin{aligned}\phi: [0, 1) &\longrightarrow [0, +\infty) \\ y &\longmapsto \frac{y}{1-y}\end{aligned}$$

és un isomorfisme. Així, podem fer el canvi $x = \frac{y}{1-y}$ amb $y \in (0, 1)$ o, equivalentment, $y = \frac{x}{x+1}$:

$$p(y) = p\left(\frac{x}{x+1}\right) = \frac{q(x)}{(1+x)^n}$$

amb

$$q(x) = a_0(x+1)^n + a_1x(x+1)^{n-1} + \dots + a_{n-1}x^{n-1}(x+1) + a_nx^n.$$

Llavors el nombre d'arrels reals positives de $q(x)$ coincideix amb el nombre d'arrels reals de $p(y)$ a l'interval $(0, 1)$.

Exemple (d'aplicació de la Regla de Descartes)

Considerem $p(x) = ax^n + bx^m + c$ amb $n > m$ i $a, c \neq 0$. Tenim $\#cs(\{a, b, c\}) \leq 2$ i per tant $p(x)$ té com a màxim dues arrels reals positives.

Anàlogament, $\#cs(\{(-1)^n a, (-1)^m b, c\}) \leq 2$ i $p(x)$ té com a màxim dues arrels reals negatives. Per tant, $p(x)$ té com a màxim 4 arrels reals (comptant multiplicitat).

Demostració (de la Regla de Descartes)

Pel Teorema de Budan-Fourier, el nombre de zeros positius de $p(x)$ és $\delta_p(0) - \delta_p(+\infty) - 2K$ (formalment $p^{(i)}(+\infty) = p^{(i)}(x)$ per x prou gran).

Observem que

$$\delta_p(+\infty) = \#cs\left(\{p(+\infty), p'(+\infty), p''(+\infty), \dots, p^{(n)}(+\infty)\}\right)$$

$$= \#cs\left(\{a_n, na_n, n(n-1)a_n, \dots, (n!)a_n\}\right) = 0 \text{ i}$$

$$\delta_p(0) = \#cs\left(\{p(0), p'(0), p''(0), \dots, p^{(n)}(0)\}\right)$$

$$= \#cs\left(\{a_0, a_1, 2a_2, \dots, (n!)a_n\}\right)$$

$$= \#cs\left(\{a_0, a_1, \dots, a_n\}\right).$$

Per tant, $\delta_p(0) - \delta_p(+\infty) - 2K = \#cs(\{a_0, a_1, \dots, a_n\}) - 2K$.



Demostració (del Teorema de Budan-Fourier)

Com hem fet a la demostració del Teorema de Sturm, comptarem els canvis a $\delta_p(x)$ al moure x contínuament d' a a b .

Concretament veurem que si $\alpha \in (a, b)$ i $\varepsilon > 0$ és prou petit, llavors

$$\delta_p(\alpha - \varepsilon) - \delta_p(\alpha + \varepsilon) = \begin{cases} 0 & \text{si } p^{(i)}(\alpha) \neq 0 \text{ per a tot } i, \\ 2K_\alpha & \text{si } p(\alpha) \neq 0 \text{ i } p^{(i)}(\alpha) = 0 \text{ per algun } i > 0, \\ m + 2K_\alpha & \text{si } \alpha \text{ és un zero de multiplicitat } m \text{ de } p(x), \end{cases} \quad (9)$$

amb $K_\alpha \in \mathbb{Z}^+$.

Demostració (del Teorema de Budan-Fourier — cont.)

Observem que el nombre de zeros de les funcions $p^{(i)}(x)$ és finit. Per tant, de les fórmules (9) deduïm que $\delta_p(a) - \delta_p(b)$ és el nombre (comptant multiplicitats) de zeros de $p(x)$ a l'interval $[a, b]$ augmentat en un nombre $2K$ amb $K \in \mathbb{Z}^+$.

Llavors, per a demostrar el Teorema de Budan Fourier n'hi ha prou de mostrar que (9) és certa.

Ara començarem la demostració de la fórmula (9) per cassos.

Cas: $p^{(i)}(\alpha) \neq 0$ per a tot $i = 0, 1, 2, \dots, n$:

La continuïtat de les funcions $p^{(i)}(x)$ assegura que existeix un entorn $U_\alpha \ni \alpha$ tal que, $p^{(i)}(x)$ no s'anul·la (i per tant no canvia de signe) per a tot $x \in U_\alpha$ i per a tot $i = 0, 1, 2, \dots, n$. Per tant, en aquest cas, $\delta_p(x)$ és constant per a tot $x \in U(\alpha)$.

Demostració (del Teorema de Budan-Fourier — cont.)

Cas: $p^{(i)}(\alpha) = 0$ per algun $i = 1, 2, \dots, n$:

Suposem que $p^{(i-1)}(\alpha) \neq 0$ i que α és un zero de multiplicitat $m \geq 1$ de $p^{(i)}(x)$. És a dir,

$$p^{(i)}(\alpha) = p^{(i+1)}(\alpha) = \dots = p^{(i+m-1)}(\alpha) = 0,$$

$$p^{(i-1)}(\alpha) \neq 0 \text{ i } p^{(i+m)}(\alpha) \neq 0.$$

Farem la demostració en el cas $p^{(i+m)}(\alpha) > 0$. El cas $p^{(i+m)}(\alpha) < 0$ es tracta anàlogament.

Com que α té multiplicitat m el Teorema de Taylor ens dona (veure la transparència 11/175):

$$p^{(i)}(x) = \frac{p^{(i+m)}(\alpha)}{m!} (x - \alpha)^m + \mathcal{O}((x - \alpha)^{m+1})$$

i, derivant, per $k = 1, 2, \dots, m$ obtenim:

$$p^{(i+k)}(x) = \frac{p^{(i+m)}(\alpha)}{(m-k)!} (x - \alpha)^{m-k} + \mathcal{O}((x - \alpha)^{m+1-k}).$$

Demostració (del Teorema de Budan-Fourier — cont.)

Llavors, per a x en un entorn prou petit d' α i $k = 0, 1, 2, \dots, m$,

$$\text{signe}(p^{(i+k)}(x)) = \text{signe}((x - \alpha)^{m-k}) = \begin{cases} (-1)^{m-k} & \text{si } x < \alpha, \\ 1 & \text{si } x > \alpha, \end{cases}$$

donat que hem suposat que $p^{(i+m)}(\alpha) > 0$.

Llavors, la taula de signes de la subsuccessió

$$\{p^{(i-1)}(x), p^{(i)}(x), p^{(i+1)}(x), \dots, p^{(i+m)}(x)\}$$

és:

x	$p^{(i-1)}$	$p^{(i)}$	$p^{(i+1)}$	\dots	$p^{(i+m-2)}$	$p^{(i+m-1)}$	$p^{(i+m)}$
$\alpha - \varepsilon$	σ	$(-1)^m$	$(-1)^{m-1}$	\dots	$+$	$-$	$+$
α	σ	0	0	\dots	0	0	$+$
$\alpha + \varepsilon$	σ	$+$	$+$	\dots	$+$	$+$	$+$

on $\sigma := \text{signe}(p^{(i-1)}(\alpha))$.

Demostració (del Teorema de Budan-Fourier — cont.)

De la taula anterior deduïm:

$$\begin{aligned} \#\text{cs}\{p^{(i-1)}(\alpha - \varepsilon), p^{(i)}(\alpha - \varepsilon), p^{(i+1)}(\alpha - \varepsilon), \dots, p^{(i+m)}(\alpha - \varepsilon)\} \\ = \begin{cases} m + v(\sigma) & \text{si } m \text{ és parell,} \\ m + 1 - v(\sigma) & \text{si } m \text{ és senar,} \end{cases} \quad i \end{aligned}$$

$$\#\text{cs}\{p^{(i-1)}(\alpha + \varepsilon), p^{(i)}(\alpha + \varepsilon), p^{(i+1)}(\alpha + \varepsilon), \dots, p^{(i+m)}(\alpha + \varepsilon)\} = v(\sigma)$$

$$\text{on denotem } v(\sigma) := \begin{cases} 0 & \text{si } \sigma = +, \\ 1 & \text{si } \sigma = -. \end{cases}$$

Per tant,

$$\begin{aligned} \#\text{cs}\{p^{(i-1)}(\alpha - \varepsilon), p^{(i)}(\alpha - \varepsilon), p^{(i+1)}(\alpha - \varepsilon), \dots, p^{(i+m)}(\alpha - \varepsilon)\} - \\ \#\text{cs}\{p^{(i-1)}(\alpha + \varepsilon), p^{(i)}(\alpha + \varepsilon), p^{(i+1)}(\alpha + \varepsilon), \dots, p^{(i+m)}(\alpha + \varepsilon)\} = \\ \left. \begin{cases} m + v(\sigma) - v(\sigma) = 2\frac{m}{2} & \text{si } m \text{ és parell,} \\ m + 1 - v(\sigma) - v(\sigma) = 2\left(\frac{m+1}{2} - v(\sigma)\right) & \text{si } m \text{ és senar,} \end{cases} \right\} = \\ 2K_i(\alpha) \in \mathbb{Z}^+. \end{aligned}$$

Demostració (del Teorema de Budan-Fourier — cont.)

Ara ja podem acabar la demostració de la segona igualtat de (9). La successió $\{p(\alpha), p'(\alpha), p''(\alpha), \dots, p^{(n)}(\alpha)\}$ descompon en unió de subsuccessions com la d'abans i subsuccessions on cap de les funcions $p^{(i)}(\alpha)$ s'anul·la (és a dir, trossos $\{p^{(j)}(\alpha), p^{(j+1)}(\alpha), \dots, p^{(j+k)}(\alpha)\}$ tals que $p^{(j)}(\alpha) \neq 0; p^{(j+1)}(\alpha) \neq 0; \dots; p^{(j+k)}(\alpha) \neq 0$). A més, dues d'aquestes subsuccessions s'intersequen com a màxim al principi i al final.

Com es dedueix de la demostració del cas **$p^{(i)}(\alpha) \neq 0$ per a tot $i = 0, 1, 2, \dots, n$** :

$$\#cs\{p^{(j)}(\alpha - \varepsilon), p^{(j+1)}(\alpha - \varepsilon), \dots, p^{(j+k)}(\alpha - \varepsilon)\} - \\ \#cs\{p^{(j)}(\alpha + \varepsilon), p^{(j+1)}(\alpha + \varepsilon), \dots, p^{(j+k)}(\alpha + \varepsilon)\} = 0$$

per un ε adequat.

Demostració (del Teorema de Budan-Fourier — cont.)

Llavors, ajuntant tots els càlculs anteriors, $\delta_p(\alpha - \varepsilon) - \delta_p(\alpha + \varepsilon)$ és una suma d'enters parells no negatius i, per tant, la segona igualtat de (9) és certa.

Cas: α és un zero de multiplicitat m de $p(x)$:

Com abans farem la demostració en el cas $p^{(m)}(\alpha) > 0$. El cas $p^{(m)}(\alpha) < 0$ es tracta anàlogament.

El fet que α té multiplicitat m ens dóna:

$$p(x) = \frac{p^{(m)}(\alpha)}{m!} (x - \alpha)^m + \mathcal{O}((x - \alpha)^{m+1})$$

i, derivant, per $k = 1, 2, \dots, m$ tenim:

$$p^{(k)}(x) = \frac{p^{(m)}(\alpha)}{(m-k)!} (x - \alpha)^{m-k} + \mathcal{O}((x - \alpha)^{m+1-k}).$$

Demostració (del Teorema de Budan-Fourier — cont.)

Llavors, per a x en un entorn prou petit d' α i $k = 0, 1, 2, \dots, m$,

$$\text{signe}(p^{(k)}(x)) = \text{signe}((x - \alpha)^{m-k}) = \begin{cases} (-1)^{m-k} & \text{si } x < \alpha, \\ 1 & \text{si } x > \alpha, \end{cases}$$

donat que hem suposat que $p^{(m)}(\alpha) > 0$.

Llavors, la taula de signes de la subsuccessió

$$\{p(x), p'(x), p''(x), \dots, p^{(m)}(x)\}$$

és:

x	p	p'	p''	\dots	$p^{(m-2)}$	$p^{(m-1)}$	$p^{(m)}$
$\alpha - \varepsilon$	$(-1)^m$	$(-1)^{m-1}$	$(-1)^{m-2}$	\dots	+	-	+
α	0	0	0	\dots	0	0	+
$\alpha + \varepsilon$	+	+	+	\dots	+	+	+

Demostració (del Teorema de Budan-Fourier — cont.)

De la taula anterior deduïm:

$$\begin{aligned} & \#cs\{p(\alpha - \varepsilon), p'(\alpha - \varepsilon), \dots, p^{(m)}(\alpha - \varepsilon)\} - \\ & \#cs\{p(\alpha + \varepsilon), p'(\alpha + \varepsilon), \dots, p^{(m)}(\alpha + \varepsilon)\} = m - 0 = \\ & m = \text{multiplicitat d}'\alpha. \end{aligned}$$

Ara ja podem demostrar la tercera igualtat de (9).

La successió $\{p(\alpha), p'(\alpha), p''(\alpha), \dots, p^{(n)}(\alpha)\}$ descompon en unió de subsuccessions com la d'abans i com les del cas anterior, on dues d'aquestes subsuccessions s'intersequen com a màxim al principi i al final.

Per tant, ajuntant tots els càlculs anteriors, $\delta_p(\alpha - \varepsilon) - \delta_p(\alpha + \varepsilon)$ és $m + 2K$ amb $K \in \mathbb{Z}^+$.

Això demostra la tercera igualtat de (9) i el Teorema de Budan-Fourier. □

Localització de valors propis: Teorema de Gershgorin

Sigui $A = (a_{ij})$ una matriu complexa $n \times n$ ($a_{ij} \in \mathbb{C}$).

Per a tot $i, j \in \{1, 2, \dots, n\}$ definim:

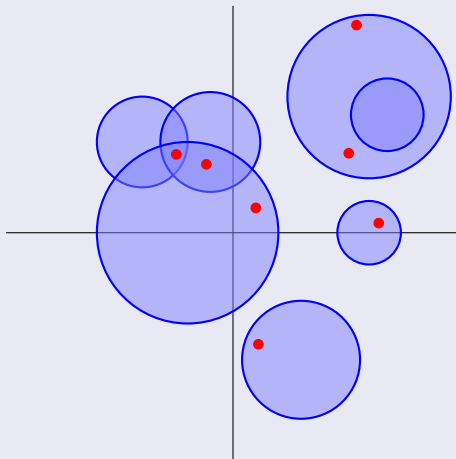
$$r_i := \sum_{j=1; j \neq i}^n |a_{ij}|, \quad R_i := \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\},$$

$$c_j := \sum_{i=1; i \neq j}^n |a_{ij}|, \quad C_j := \{z \in \mathbb{C} : |z - a_{jj}| \leq c_j\}.$$

Teorema (Gershgorin)

Sigui $A = (a_{ij})$ una matriu complexa $n \times n$ i sigui λ un valor propi d' A . Llavors, $\lambda \in \bigcup_{i=1}^n R_i$ i $\lambda \in \bigcup_{j=1}^n C_j$. A més a cada component connexa de $\bigcup_{i=1}^n R_i$ (respectivament $\bigcup_{j=1}^n C_j$) hi ha tants valors propis (comptats amb multiplicitat) com discs R_i (respectivament C_j) formen la component.

El Teorema de Gershgorin per a una possible matriu complexa 7×7



Localització de valors propis: Teorema de Gershgorin 2/15

Exemple (Gershgorin)

Sigui

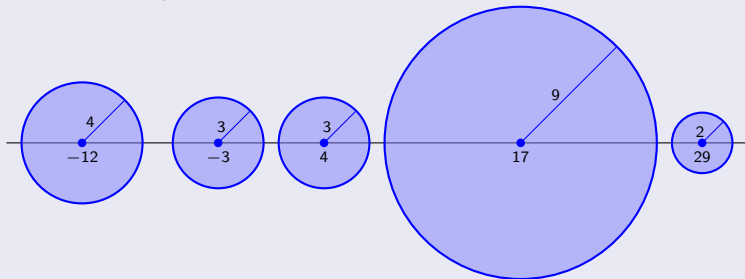
$$A = \begin{pmatrix} 29 & -3 & 4 & 0 & 0 \\ 1 & -12 & 1 & -2 & 0 \\ 1 & 0 & 17 & 1 & 2 \\ 0 & 1 & -2 & -3 & 1 \\ 0 & 0 & -2 & 0 & 4 \end{pmatrix}$$

que té,

$$\begin{array}{ccccc} r_1 = 7; & r_2 = 4; & r_3 = 4; & r_4 = 4; & r_5 = 2; \\ c_1 = 2; & c_2 = 4; & c_3 = 9; & c_4 = 3; & c_5 = 3. \end{array}$$

Exemple (Gershgorin — cont.)

Si usem els discs C_j tenim:



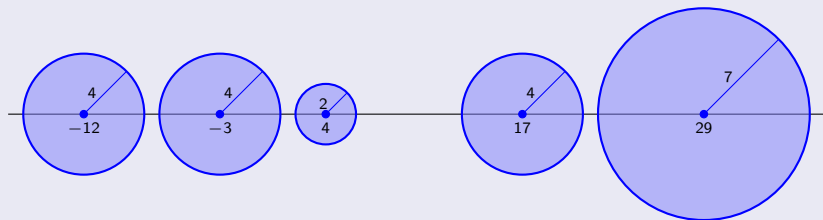
Els discs són disjunts dos a dos. És a dir, cada component connexa està formada per un únic disc. Llavors, cada disc conté un únic valor propi.

Com que la matriu és real, els discs són simètrics respecte de l'eix real i el valor propi contingut a cada disc ha de ser real (altrament, n'hi hauria d'haver 2).

Per tant, els valors propis són als intervals: $[-16, 8]$, $[-6, 0]$, $[1, 7]$, $[8, 26]$ i $[27, 31]$ i que cada un d'aquest interval conté un únic valor propi.

Exemple (Gershgorin — cont.)

Si usem els discs R_j tenim:



La situació és anàloga a la dels discs C_j amb una localització diferent: $[-16, 8]$, $[-7, 1]$, $[2, 6]$, $[13, 21]$ i $[22, 36]$.

Notem que intersecant aquests intervals amb els anteriors tenim una millor localització.

Exemple (Gershgorin — exemple analític)

Volem localitzar les arrels reals i complexes del polinomi

$$z^{n+1} + a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0.$$

Considerem la matriu

$$A = \begin{pmatrix} -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Veurem que el polinomi anterior és el polinomi característic de la matriu A . Amb aquesta finalitat calculem $\det(\lambda \text{Id} - A)$, desenvolupant successivament per la primera columna i usant que el determinant d'una matriu triangular és el producte dels elements de la diagonal:

Exemple (Gershgorin — exemple analític — cont.)

$$\begin{aligned}
 \det(\lambda \text{Id} - A) &= \begin{vmatrix} \lambda + a_n & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ -1 & \lambda & 0 & \cdots & 0 & 0 \\ 0 & -1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & \lambda \end{vmatrix} \\
 &= (\lambda + a_n)\lambda^{n-1} + \begin{vmatrix} a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ -1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & \lambda \end{vmatrix} \\
 &= \lambda^n + a_n\lambda^{n-1} + a_{n-1}\lambda^{n-2} + \begin{vmatrix} a_{n-2} & a_{n-3} & \cdots & a_1 & a_0 \\ -1 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & \lambda \end{vmatrix} \\
 &= \cdots = \lambda^n + a_n\lambda^{n-1} + a_{n-1}\lambda^{n-2} + a_{n-2}\lambda^{n-3} + \cdots + a_1\lambda + a_0.
 \end{aligned}$$

Exemple (Gershgorin — exemple analític — cont.)

Ara, com que el polinomi que volem estudiar és el polinomi característic de la matriu A , podem usar Gershgorin per a localitzar les arrels. Definim

$$c := \max\{|a_0|, |a_1| + 1, |a_2| + 1, \dots, |a_{n-1}| + 1\} \text{ i } r := \sum_{i=0}^{n-1} |a_i|.$$

Així,

$$\text{spec}(A) \subset (B(0, 1) \cup B(-a_n, r)) \cap (B(-a_n, 1) \cup B(0, c)).$$

Demostració (del Teorema de Gershgorin)

Solament farem la demostració de les fites per files. Les fites per columnes es dedueixen de la mateixa manera usant la matriu A^T en lloc d' A ja que A^T té els mateixos valors propis que A .

Sigui $w = (w_1, w_2, \dots, w_n) \neq 0$ el vector propi associat al valor propi λ . Existeix i tal que $|w_i| = \|w\|_\infty > 0$ (és a dir $|w_i| \geq |w_k|$ per $k = 1, 2, \dots, n$).

Considerem la component i de l'equació característica $Aw = \lambda w$:

$$\sum_{j=1}^n a_{ij}w_j = \lambda w_i \iff (\lambda - a_{ii})w_i = \sum_{j=1; j \neq i}^n a_{ij}w_j.$$

Demostració (del Teorema de de Gershgorin — cont.)

Afitant l'expressió anterior tenim

$$|\lambda - a_{ii}| |w_i| \leq \sum_{j=1; j \neq i}^n |a_{ij}| |w_j| \leq \left(\sum_{j=1; j \neq i}^n |a_{ij}| \right) |w_i| = r_i |w_i|$$

i, donat que $w_i \neq 0$,

$$|\lambda - a_{ii}| \leq r_i \iff \lambda \in R_i \subset \bigcup_{k=1}^n R_k.$$

Això acaba la demostració de la primera afirmació del teorema.

Ara demostrarem la segona afirmació.

Demostració (del Teorema de de Gershgorin — cont.)

Sigui D la matriu diagonal que té les entrades de la diagonal d' A en el mateix ordre. Per a cada $t \in [0, 1]$ definim la matriu

$$A_t := tA + (1 - t)D.$$

A_t és una família uniparamètrica i contínua de matrius que comença amb $A_0 = D$ i acaba amb $A_1 = A$. A més la diagonal d' A_t coincideix amb la diagonal d' A per a tot t .

La demostració consisteix a seguir els valors propis i els discs d' A_t començant amb A_0 i variant contínuament el paràmetre t fins A_1 , comprovant així que l'afirmació es compleix per a cada t (veure la figura de la transparència 174/175, que mostra 4 moments successius del procés al variar t):

Demostració (del Teorema de de Gershgorin — cont.)

Per construcció, el nombre $r_i(A_t)$ corresponent a la matriu A_t és tr_i . Denotarem $R_i^t := \{z \in \mathbb{C} : |z - a_{ii}| \leq tr_i\}$.

La primera part del teorema ens diu que el conjunt dels valors propis d' A_t està contingut a $\bigcup_{k=1}^n R_k^t$ per a tot t .

Notem que els valors propis de $A_0 = D$ són $a_{11}, a_{22}, \dots, a_{nn}$ que, al mateix temps són els elements de la diagonal. Per altra banda, per A_0 tenim $r_1 = r_2 = \dots = r_n = 0$. En particular $R_i = \{a_{ii}\}$ per a cada i . Observem que si un element a_{ii} apareix a la diagonal $m \geq 1$ vegades llavors $\{a_{ii}\}$ és una component connexa de $\bigcup_{k=1}^n R_k$ formada per m discs R_k (degenerats a un punt) mentre que aquesta component conté m valors propis d' $A_0 = D$. Això mostra que l'afirmació del teorema és certa per $t = 0$.

Ara anem a veure que aquest fet es conserva al moure la t .

Demostració (del Teorema de de Gershgorin — cont.)

Els discs R_i^t tenen mida tr_i . Es va dir, augmenten de manera *contínua* sense que el seu centre es mogui de lloc. D'altra banda,

- els elements de les matrius A_t depenen contínuament de t ,
- els coeficients del polinomi característic d'una matriu depenen contínuament dels seus elements i , finalment,
- les arrels d'un polinomi depenen contínuament dels seus coeficients.

En conclusió, tots els valors propis d' A_t varien contínuament en t .

Si per un valor de t (inicialment per $t = 0$) a cada component connexa de $\bigcup_{i=1}^n R_i$ hi ha tants valors propis (comptats amb multiplicitat) com discs R_i , aquesta propietat es mantindrà al augmentar t mentre la component connexa no intersequi amb una altra, ja que

- la continuïtat dels valors propis d' A_t respecte de t i el fet que
 - els valors propis d' A_t estan continguts a $\bigcup_{k=1}^n R_k^t$ per a tot t ,
- no permet que aquests saltin a una altra component disjunta.

Demostració (del Teorema de de Gershgorin — cont.)

Quan diverses components connexes intersequen (i passen a ser la mateixa) els valors propis d' A_t es poden canviar de disc de manera contínua dins de la component connexa (veure el cas 3 de la figura de la transparència 174/175).

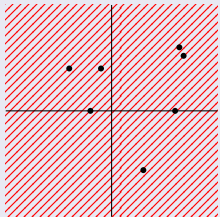
De tota manera, en aquest cas, el nombre de discs i de valors propis d' A_t que hi ha a la nova component connexa és la suma dels nombre de discs i de valors propis d' A_t que hi havia a les components originals (que han quedat integrades a l'actual).

Donat que aquests nombres coincidien per hipòtesi, les seves sumes continuen coincidint i, per tant, la tesi del teorema es compleix per a cada t . □

Localització de valors propis: Teorema de Gershgorin (cont.)

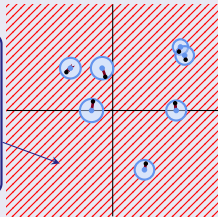
Veure la pel·lícula completa

1: $A_0 = D$: discs = valors propis



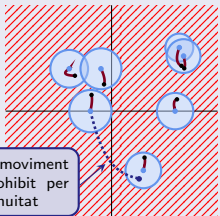
2: Cada VaP és al seu disc

La zona vermella és prohibida pels valors propis: han de ser dins dels discs

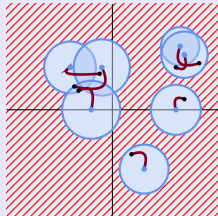


3: Un VaP ha canviat de disc

Aquest moviment està prohibit per la continuïtat



4: La situació per la matriu A_1



Localització de valors propis: Teorema de Gershgorin 15/15

