Zeros de funcions

Equacions en una variable

Considerem el problema de trobar les solucions d'una equació no lineal $f(x) = 0$. Direm que α és una arrel, o solució, o un zero de f, si $f(\alpha) = 0$. Començarem per equacions en la recta real. Posteriorment considerarem el cas complexe (per exemple en trobar zeros de polinomis), i els sistemes no lineals d'equacions. Assumirem, en principi, tota la regularitat que calgui per a la funció f.

Distingirem tres possibles fases en el procés de trobar zeros

1. Localització. Consisteix en la determinació de regions que continguin els zeros. Es poden fer servir taules de la funció, representacions gràfiques, resultats analitics (per exemple teoremes de localització de zeros de polinomis) , etc.
2. Separació. Per evitar repetir el càlcul d'un mateix zero es convenient tenir intervals o regions al pla complex en els que es tingui un únic zero.
3. Aproximació numèrica. Construcció d'una successió convergent al zero que volem trobar.

Mètodes iteratius d'aproximació de solucions de $f(x) = 0$.

Mètode de bisecció

Suposem conegut un interval $[a_0, b_0]$ tal que $f(a_0)f(b_0) < 0$ i que contingui un únic zero. Considerem la successió d'intervals $I_k = [a_k, b_k]$ definida de la forma següent. $I_0 = [a_0, b_0]$.

Sigui $x_k = (a_k + b_k)/2$. Si $f(x_k) = 0$ ja tenim la solució (molt poc probable numèricament). Si no, definim

$$I_{k+1} = [a_{k+1}, b_{k+1}] = \begin{cases} (a_k, x_k) & \text{si} \quad f(a_k)f(x_k) < 0 \\ (x_k, b_k) & \text{si} \quad f(x_k)f(b_k) < 0. \end{cases}$$

Després de n passos $b_k - a_k = (b_{k-1} - a_{k-1})/2 = \cdots = (b_0 - a_0)/2^k$ i, si prenem x_k com a aproximació de α llavors l'error en el zero, $E = x_k - \alpha$ verifica

$$|E| < (b_0 - a_0)/2^{k+1}.$$

El mètode és convergent però només geomètricament.

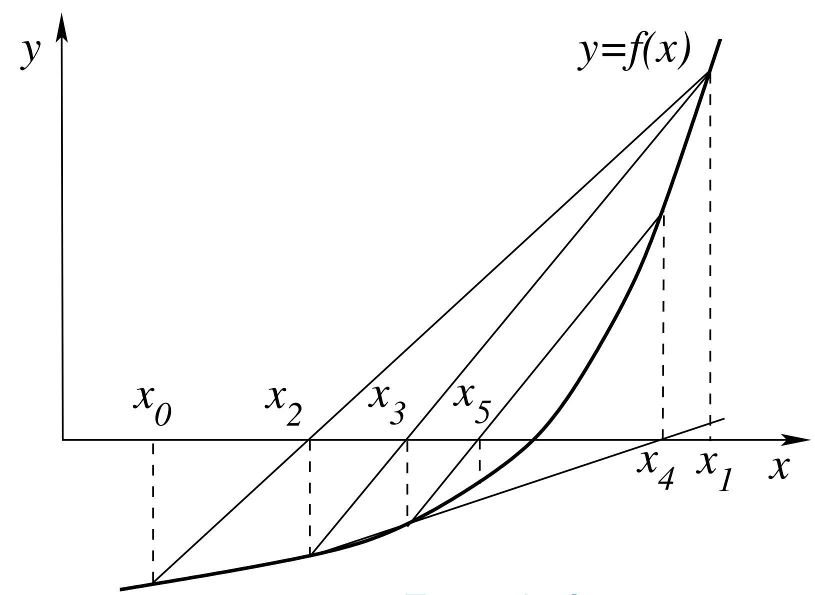
Mètode de la secant

Donats dues parelles $(x_{k-1}, f(x_{k-1}))$ i $(x_k, f(x_k))$ considerem la recta que passa pels dos punts i troben el valor de x per al que es fa zero.

Aquest punt existeix si $f(x_{k-1}) \neq f(x_k)$ i és

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

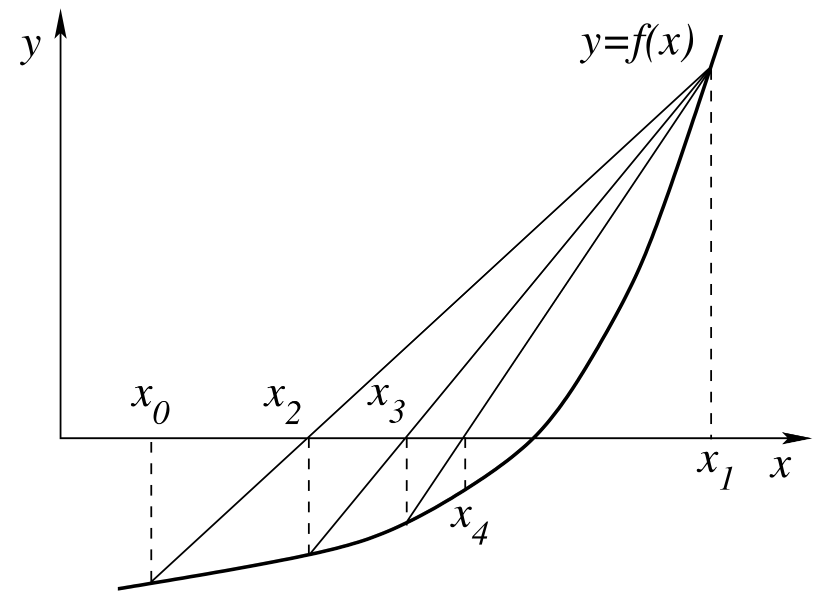
Això defineix un altre mètode iteratiu que no és necessàriament convergent, més ràpid que la bisecció (ho veurem més endavant) i només requereix una avaluació de f per pas al igual que el mètode de bisecció.



Mètode de la *regula falsi*

És una variant de la secant que és convergent, si inicialment $f(x_0)f(x_1) < 0$ amb $x_0 < x_1$. Sigui $I_1 = [x_0, x_1]$. Es calcula x_2 com en el mètode de la secant i es pren, anàlogament al mètode de bisecció,

$$I_2 = \begin{cases} (x_0, x_2) & \text{si} \quad f(x_0)f(x_2) < 0 \\ (x_2, x_1) & \text{si} \quad f(x_1)f(x_2) < 0. \end{cases}$$



En el següent pas es fa servir, per calcular x_3, el nou interval que contingui el zero, i així successivament. En general és més lent que la secant, però és convergent. En la figura la successió d'intervals que conté el zero és $[x_0, x_1]$, $[x_2, x_1]$, $[x_3, x_1]$, $[x_4, x_1]$, etc.

Mètode de Newton

Donada una parella $(x_k, f(x_k))$ considerem la recta que passa per x_k amb pendent $f'(x_k)$ i trobem el seu zero, x_{k+1},

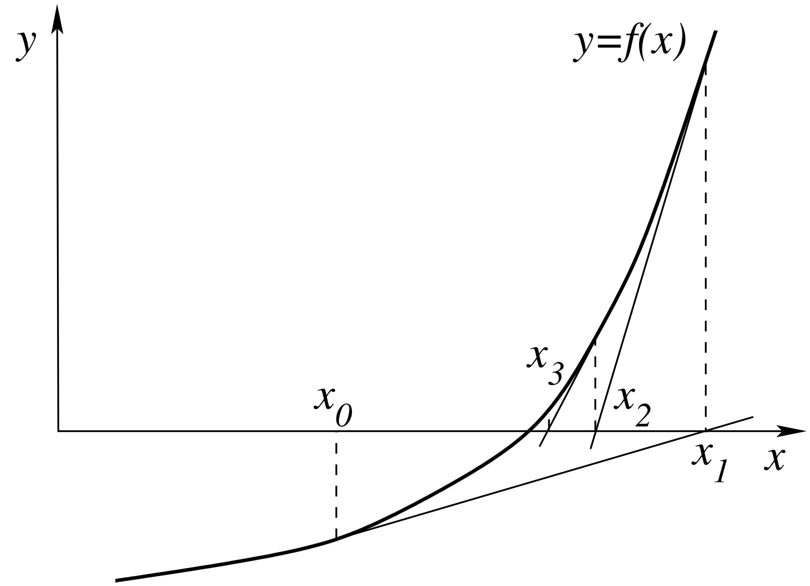
$$0 = f(x_k) + f'(x_k)(x_{k+1} - x_k)$$

llavors

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

A cada iteració cal una avaluació de f i f'.

No sempre és convergent, però és el mètode més ràpid dels que hem vist (ho quantificarem més endavant). Per tenir convergència cal que x_0 estigui prou a prop de l'arrel. El següent teorema proporciona condicions per a la convergència (veure la demostració a Henrici 1964).



Teorema. Sigui $f : [a, b] \to \mathbb{R}$, de classe $\mathcal{C}^2$ tal que

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

ii) $f'(x) \neq 0 \ \forall x \in [a, b]$

iii) $f''(x) \geq 0$ (o $f''(x) \leq 0$) $\forall x \in [a, b]$

iv) Si c és l'extrem de $[a, b]$ en el que $|f'(x)|$ és menor, llavors $|f(c)|/|f'(c)| \leq b - a$

Llavors el mètode de Newton convergeix a la unica solució α de $f(x) = 0$ a $[a, b]$ per a qualsevol condició inicial $x_0 \in [a, b]$.

Les condicions anteriors diuen el següent:

i) Hi ha un zero a l'interval $[a, b]$ (Bolzano).

ii) El zero és únic.

iii) La funció és còncava o convexa a tot l'interval. Això garanteix que les aproximacions s'aproparan a la solució.

iv) Aquesta condició garanteix que la recta tangent en el punt c talla l'eix x dins l'interval $[a, b]$.

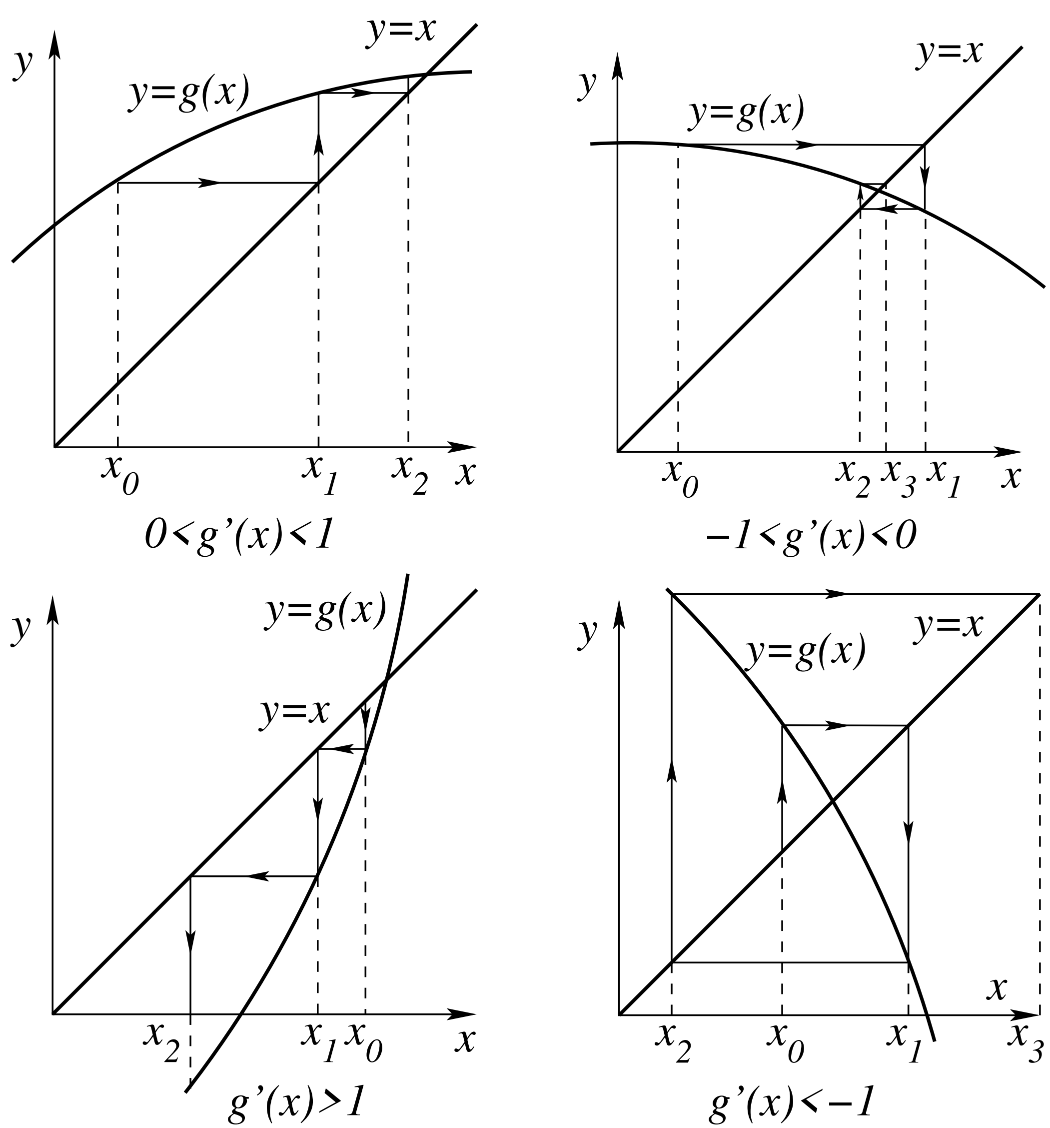
Si $f'(\alpha) \neq 0$ i $f''(\alpha) \neq 0$ sempre hi ha un interval $[a, b]$ en el que se satisfan totes les condicions anteriors.

Exemple: Calcular el zero de $\sin x - x/2 = 0$ a l'interval $[\pi/2, \pi]$ pels mètodes de bisecció, secant i Newton.

Exemple: Calcular les arrels de l'equació $4 \sin x + 1 - x = 0$ pels mètodes de bisecció, secant i Newton.

Teoria d'iteració. Mètodes de punt fix.

Imaginem que transformem el problema de trobar x, satisfent $f(x) = 0$, al de trobar x, satisfent $x = g(x)$ i obtingut, per exemple, aïllant una de les x que apareixen a f. Podem definir una successió $x_k = g(x_{k-1})$ $k = 1, 2, \cdots$ a partir d'un x_0 inicial. Si aquesta successió convergeix a s, $\lim_{k\to\infty} x_k = s$, i g és continua llavors $s = g(s)$.



Teorema del punt fix. Sigui $g : I = [a, b] \to \mathbb{R}$ de classe $\mathcal{C}^1$ tal que

i) $g(I) \subset I$

ii) $|g'(x)| \leq L < 1 \quad \forall x \in I$

Llavors

a) Existeix un únic punt fix s de g en I $(s = g(s))$.

b) Per a qualsevol $x_0 \in I$ la successió $x_k = g(x_{k-1})$ $k = 1, 2, \cdots$ convergeix cap a s.

c)

$$\text{c.1)} \quad |x_k - s| \leq \frac{L^k}{1-L}|x_1 - x_0|, \qquad \text{c.2)} \quad |x_k - s| \leq \frac{L}{1-L}|x_k - x_{k-1}|$$

El resultats c.1 i c.2 donen fites de l'error que es comet en prendre x_k com aproximació de s.

Demostració:

a) Sigui $F(x) = g(x) - x$. Per i) tenim que $g(a) \geq a$ i $g(b) \leq b$. Llavors $F(a) \geq 0$ i $F(b) \leq 0$. Aplicant Bolzano obtenim que F té ún zero a I.

Suposem que hi ha dues solucions s_1 i s_2. Llavors

$$|s_1 - s_2| = |g(s_1) - g(s_2)| = |g'(\xi)||s_1 - s_2| \leq L|s_1 - s_2| < |s_1 - s_2|$$

i arribem a contradicció.

b)

$$|x_k - s| = |g(x_{k-1}) - g(s)| = |g'(\xi_{k-1})|\,|x_{k-1} - s| \leq L|x_{k-1} - s| \leq \cdots \leq L^k|x_0 - s|.$$

Prenent límits tenim $0 \leq \lim_{k\to\infty} |x_k - s| \leq \lim_{k\to\infty} L^k|x_0 - s| = 0$ i $\lim_{k\to\infty} x_k = s$.

c)

$$|x_{k+1} - x_k| = |g(x_k) - g(x_{k-1})| \leq L|x_k - x_{k-1}| \leq \cdots \leq L^k|x_1 - x_0|.$$

Si $m > k$

$$|x_m - x_{m-1}| \leq L|x_{m-1} - x_{m-2}| \leq \cdots \leq L^{m-k}|x_k - x_{k-1}|.$$

c.1) Si $m > k$

$$\begin{aligned} |x_m - x_k| &\leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{k+1} - x_k| \\ &\leq L^{m-1}|x_1 - x_0| + L^{m-2}|x_1 - x_0| + \cdots + L^k|x_1 - x_0| \\ &= L^k|x_1 - x_0|(1 + L + \cdots + L^{m-k-1}) \\ &\leq L^k|x_1 - x_0| \sum_{i=0}^{\infty} L^i = L^k|x_1 - x_0| \frac{1}{1-L} \end{aligned}$$

Fent el límit $m \to \infty$ obtenim, finalment,

$$|s - x_k| \leq \frac{L^k}{1-L}|x_1 - x_0|.$$

c.2)

$$
\begin{aligned}
|x_m - x_k| = \leq |x_m - x_{m-1}| + |x_{m-1} - x_{m-2}| + \cdots + |x_{k+1} - x_k| \\
\leq L^{m-k}|x_k - x_{k-1}| + L^{m-k-1}|x_k - x_{k-1}| + \cdots + L|x_k - x_{k-1}| \\
= L|x_k - x_{k-1}|(1 + L + \cdots + L^{m-k-1}) \\
\leq L|x_k - x_{k-1}| \sum_{i=0}^{\infty} L^i = \frac{L}{1-L}|x_k - x_{k-1}|.
\end{aligned}
$$

Fent el límit $m \to \infty$ obtenim,

$$
|s - x_k| \leq \frac{L}{1-L}|x_k - x_{k-1}|.
$$

En el teorema es pot substituir la condició ii) per

ii') g satisfà la condició de Lipschitz $|g(x) - g(x')| \leq L|x - x'|\ \forall x, x' \in I$ amb $L < 1$.

Es diu llavors que g és contractiva. De fet ii) implica ii').

Exemple: Calculeu el zero de $x - e^{-x} = 0$ amb un error menor que 10^{-4}. Preneu $g(x) = e^{-x}$ i considereu l'interval $I = [e^{-1}, \ln 2]$.

Exemple: Considerem la iteració $x_{k+1} = g(x_k)$ amb $g(x) = \sqrt{1+x^2} - 0.05$. Comproveu que si $x_0 = 9$ i s'atura la iteració quan $|x_k - x_{k-1}| < \varepsilon$ amb $\varepsilon = 10^{-4}$, l'iterat x_k està encara a una distancia més gran que 0.01 de l'arrel (que es pot calcular analíticament i val $s = 9.975$). En canvi si el criteri d'aturada és $(L/(1-L))|x_k - x_{k-1}| < \varepsilon$ s'obté un error per a l'arrel menor que ε. Preneu l'interval $I = [9, 10]$ i $L = |g'(10)|$.

Ordre de convergència

Definició. Direm que una successió $(x_k)_{k\geq 0}$, convergent cap a α, té ordre de convergència almenys p si existeixen $N \geq 0$ i $C > 0$ tals que

$$|x_{k+1} - \alpha| \leq C|x_k - \alpha|^p \quad \forall k \geq N.$$

Si $p = 1$ haurem d'exigir, a més, que $C < 1$.

En particular, si existeix el límit

$$L = \lim_{k\to\infty} \frac{x_{k+1} - \alpha}{(x_k - \alpha)^p}$$

la successió tindrà ordre de convergència almenys p (si $p = 1$ haurem d'exigir, a més, que $L < 1$). Si $L \neq 0$ direm que la successió $(x_k)_{k\geq 0}$ té ordre de convergència p i que L és la constant asimptòtica de l'error. Si $p = 1, 2, 3, \cdots$ parlarem de convergència lineal, quadràtica, cúbica, etc.

Exemple: Veiem que, en general, un mètode iteratiu $x_{k+1} = g(x_k)$ té convergència lineal. Sigui $e_{k+1} = x_{k+1} - s$. Tenim

$$e_{k+1} = g(x_k) - g(s) = g'(c_k)(x_k - s) = g'(c_k)e_k$$

amb $\lim_{k\to\infty} c_k = s$. Llavors $\lim_{k\to\infty} e_{k+1}/e_k = g'(s)$ (amb $|g'(s)| < 1$). Això vol dir que, asimptòticament, $e_{k+1} \approx g'(s)e_k$.

Veiem que cal per tenir ordres de convergència alts.

Proposició. Considerem la iteració $x_{k+1} = g(x_k)$ amb $s = g(s)$ un punt fix. Si $g \in \mathcal{C}^\infty(s-\varepsilon, s+\varepsilon)$, $\varepsilon > 0$, i $g^{(j)}(s) = 0$, $j = 1, \cdots, p-1$, $g^{(p)}(s) \neq 0$, llavors el mètode iteratiu convergeix i té ordre p.

Demostració:
El desenvolupament en serie de Taylor de g al voltant de s és

$$g(x) = g(s) + \frac{g^{(p)}(\xi_x)}{p!}(x-s)^p = s + \frac{g^{(p)}(\xi_x)}{p!}(x-s)^p.$$

Llavors, fent $x = x_k$,

$$e_{k+1} = x_{k+1} - s = \frac{g^{(p)}(\xi_k)}{p!}e_k^p \quad \text{i} \quad \lim_{k\to\infty} \frac{e_{k+1}}{e_k^p} = \frac{g^{(p)}(s)}{p!} \neq 0.$$

Exemple: Mètode babilònic o d'Heró per calcular arrels quadrades. Considereu la iteració $x_{k+1} = g(x_k)$ amb

$$g(x) = \frac{1}{2}\left(x + \frac{c}{x}\right)$$

amb $c > 0$ i $x_0 > 0$. També surt d'aplicar el mètode de Newton a $x^2 - c = 0$. Veieu que $g'(\sqrt{c}) = 0$ i deduïu que, com a mínim a prop de $\sqrt{c}$, la successió convergeix quadràticament a $\sqrt{c}$. De fet és fàcil veure que $x_{k+1} - \sqrt{c} = (x_k - \sqrt{c})^2/2x_k$ i deduir que per a qualsevol $x_0 > 0$ $x_1 \geq x_2 \geq x_3 \geq \cdots \geq \sqrt{c}$ i la convergència és global ($\forall x_0 > 0$).

Convergència del mètode de Newton.

El mètode de Newton per a l'equació $f(x) = 0$ és $x_{k+1} = g(x_k)$ amb

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Siqui s l'arrel ($f(s) = 0$) i supossem que $f'(s) \neq 0$. Per veure que és d'ordre dos només cal veure que $g'(s) = 0$. Tenim

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

que es fa zero a $x = s$. Per tant el mètode és almenys d'ordre 2. En general $g''(s) \neq 0$ i el mètode és d'ordre 2.

Zeros múltiples. Si $f(x) = 0$ té s com a zero de multiplicitat q llavors podem escriure $f(x) = (x - s)^q h(x)$ amb $h(s) \neq 0$. Llavors

$$g(x) = x - \frac{f(x)}{f'(x)} = x - \frac{(x-s)^q h(x)}{q(x-s)^{q-1}h(x) + (x-s)^q h'(x)} = x - \frac{(x-s)h(x)}{qh(x) + (x-s)h'(x)}$$

i

$$g'(x) =$$

$$1 - \frac{[h(x) + (x-s)h'(x)][qh(x) + (x-s)h'(x)] - (x-s)h(x)[(q+1)h'(x) + (x-s)h''(x)]}{[qh(x) + (x-s)h'(x)]^2}.$$

Aleshores

$$g'(s) = 1 - \frac{h(s)qh(s)}{q^2h(s)^2} = 1 - \frac{1}{q}$$

i $g'(s) \neq 0$ si $q \geq 2$. Llavors la convergència és només lineal. Si canviem $g(x)$ a

$$g(x) = x - q\frac{f(x)}{f'(x)}$$

es veu que $g'(s) = 0$. D'aquesta manera podem recuperar la convergència quadràtica si coneixem la multiplicitat de l'arrel. En general no ho sabrem. Una altra possibilitat és aplicar el mètode de Newton a f/f' que té totes les arrels simples. Això implica, però, calcular la derivada segona.

Exemple: Considereu el polinomi $p(x) = x^3 - 4x + 5x - 2 = (x-1)^2(x-2)$. Apliqueu el mètode de Newton amb $q = 1$ o $q = 2$ des de la condició inicial $x_0 = 0.8$ i veieu quantes iteracions calen per tenir $|x_k - 1| \leq 10^{-8}$ en cada cas.

Es pot veure que

- El mètode de bisecció té convergència lineal independentment del comportament de f. Obviament cal que sigui continua.
- Si $f'(s) \neq 0$ el mètode de la secant té ordre de convergència almenys $(1+\sqrt{5})/2$.
- Si $f'(s) \neq 0$ el mètode de la regula falsi té convergència, almenys lineal.

Acceleració de la convergència.

Volem ara trobar procediments que, a partir d'una successió convergent $(x_k)_{k\geq 0}$, proporcionin una altra $(x'_k)_{k\geq 0}$ amb el mateix límit però que convergeixi més ràpidament. Això es pot aplicar a d'altres tipus de problemes, a banda del de trobar zeros.

Mètode d'acceleració d'Aitken

Suposem que $x_k \to s$. Definim $\Delta x_k = x_{k+1} - x_k$. Tenim que $\Delta x_k \to 0$. Considerem la recta que passa per la parella de punts $(x_k, \Delta x_k)$ i $(x_{k+1}, \Delta x_{k+1})$

$$y = \Delta x_k + \frac{\Delta x_{k+1} - \Delta x_k}{x_{k+1} - x_k}(x - x_k) = \Delta x_k + \frac{\Delta^2 x_k}{\Delta x_k}(x - x_k)$$

a on $\Delta^2 x_k = \Delta x_{k+1} - \Delta x_k = x_{k+2} - 2x_{k+1} + x_k$. La x que fa $y = 0$ és la que defineix x'_k

$$x'_k = x_k - \frac{(\Delta x_k)^2}{\Delta^2 x_k}.$$

Proposició. Suposem que $\lim_{k\to\infty} x_k = s$, $x_k \neq s$ $\forall k$ i que existeix C, $|C| < 1$ tal que $x_{k+1} - s = (C + \delta_k)(x_k - s)$ amb $\lim_{k\to\infty} \delta_k = 0$ (convergència almenys lineal).

Llavors la successió anterior $(x'_k)_{k\geq 0}$ està ben definida per a k prou gran i

$$\lim_{k\to\infty} \frac{x'_k - s}{x_k - s} = 0,$$

es a dir, $(x'_k)_{k\geq 0}$ convergeix a s més ràpidament que $(x_k)_{k\geq 0}$.

Exemple: Considerem l'equació $f(x) = 2x^2 - 10x + 10 + \cos(x/9) = 0$ que reescrivim com el problema de punt fix $x = g(x) = [2x^2 + \cos(x/9)]/10 + 1$. Tenim que $g'(x) = [4x - (1/9)\sin(x/9)]/10$, $f(0) > 0$, $f(2) < 0$, $|g'(x)| < 1$ en $(0, 2)$. Llavors $x_{k+1} = g(x_k)$ serà convergent. Compareu la convergència del mètode de punt fix amb el modificat pel procediment d'Aitken.

Mètode d'acceleració de Steffensen

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

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

dòna un mètode de convergència almenys quadràtica cap al mateix s (cal que $g'(s) \neq 1$). Si la convergència de $x_{k+1} = g(x_k)$ és almenys lineal llavors la de $y_{k+1} = G(y_k)$ és almenys quadràtica. Si la convergència de $x_{k+1} = g(x_k)$ és almenys d'ordre $p > 1$ llavors la de $y_{k+1} = G(y_k)$ és almenys d'ordre $2p - 1$. Això últim no és gaire útil ja que $y_{k+1} = g(g(y_k))$ ja té convergència d'ordre p^2.

La diferència entre els dos mètodes és que el d'Aitken treballa sobre la successió inicial, mentre que en el de Steffensen es construeix la successió ja accelerada. En general el segon produeix millors resultats però requereix dues avaluacions de g per iteració.

Sistemes d'equacions no lineals

Considerem els sistema d'n equacions no lineals

$$f_i(x_1, \cdots, x_n) = 0, \quad i = 1, \cdots, n$$

que podem escriure en forma compacta com $\mathbf{f}(\mathbf{x}) = 0$ amb $\mathbf{x} = (x_1, \cdots, x_n)$ i $\mathbf{f} = (f_1, \cdots, f_n)$.

Una primera possible via per calcular les solucions és la iteració de punt fix.

Teorema del punt fix. Sigui $T \subset \mathbb{R}^n$ un conjunt tancat, $\mathbf{G} : T \to \mathbb{R}^n$ tal que

i) $\mathbf{G}(T) \subset T$

ii) Existeix una constant $L < 1$ tal que $\|\mathbf{G}(\mathbf{x}_2) - \mathbf{G}(\mathbf{x}_1)\| \leq L\|\mathbf{x}_2 - \mathbf{x}_1\| \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in T$, amb $L < 1$ ($\mathbf{G}$ és de Lipschitz i contractiva)

Llavors

a) Existeix un únic punt fix $\mathbf{s}$ de $\mathbf{G}$ en T ($\mathbf{s} = \mathbf{G}(\mathbf{s})$).

b) Per a qualsevol $\mathbf{x}_0 \in T$ la successió $\mathbf{x}_k = \mathbf{G}(\mathbf{x}_{k-1})$ $k = 1, 2, \cdots$ convergeix cap a $\mathbf{s}$.

c)

$$\text{c.1)} \quad \|\mathbf{x}_k - \mathbf{s}\| \leq \frac{L^k}{1-L}\|\mathbf{x}_1 - \mathbf{x}_0\|, \qquad \text{c.2)} \quad \|\mathbf{x}_k - \mathbf{s}\| \leq \frac{L}{1-L}\|\mathbf{x}_k - \mathbf{x}_{k-1}\|.$$

La demostració és com la del cas unidimensional (1D), substituint valors absoluts per normes, excepte la part d'existència que es dedueix de forma diferent (no es pot aplicar Bolzano). La condició

$$\|\mathbf{x}_m - \mathbf{x}_k\| \leq \frac{L^k}{1-L}\|\mathbf{x}_1 - \mathbf{x}_0\|$$

diu que la successio$(\mathbf{x}_k)_{k\geq 0}$ és de Cauchy i per tant convergent en T. Si $\mathbf{s}$ és el límit, prenent límits en $\mathbf{x}_{k+1} = \mathbf{G}(\mathbf{x}_k)$ tenim $\mathbf{s} = \mathbf{G}(\mathbf{s})$.

Com en el cas 1D ii) es pot substituir per ii')

ii') $\mathbf{G}$ és de classe $\mathcal{C}^1$ i satisfà $\|D\mathbf{G}(\mathbf{x})\| \leq L < 1$, $\forall \mathbf{x} \in T$,

essent $\|.\|$ la norma matricial consistent amb la norma vectorial considerada. Només cal aplicar el teorema del valor mitjà per veure que ii') implica ii).

Hi ha casos en que el problema original ja està escrit com un problema de punt fix o és fàcil fer-ho. Això no passa en general o no es verifica la contractivitat, i el mètode a aplicar és el de Newton amb alguna variant.

El mètode de Newton. Es dedueix igual que en el cas 1D. Considerem el sistema $\mathbf{f}(\mathbf{x}) = 0$ i desenvolupem en serie de Taylor al voltant de $\mathbf{x}_k$ fins a primer ordre

$$\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\mathbf{x}_k) + D\mathbf{f}(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k).$$

Fent $\mathbf{f}(\mathbf{x}) = 0$ obtenim el següent iterat $\mathbf{x}_{k+1}$. Si $\Delta\mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ llavors $\Delta\mathbf{x}_k$ verifica

$$D\mathbf{f}(\mathbf{x}_k)\Delta\mathbf{x}_k = -\mathbf{f}(\mathbf{x}_k)\,, \quad \text{i} \quad \mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k.$$

Notes:

- La definició d'ordre de convergència és la mateixa que en el cas 1D substituint valors absoluts per normes.
- Es demostra que sota condicions molt generals que el mètode convergeix quadràticament igual que en el cas 1D, si la condició inicial està prou a prop de la solució. Això acostuma a ser un problema important. Quan el problema depèn d'un paràmetre p, $\mathbf{f}(\mathbf{x}, p) = 0$, i es volen estudiar les solucions variant p, s'apliquen mètodes de continuació que proporcionen bones aproximacions inicials (ho veurem més endavant).
- A cada pas cal resoldre un sistema lineal amb una matriu diferent a cada iteració. Això és costós si n és gran. Si en lloc de fer això és fixa la matriu corresponent a la primera iteració, $D\mathbf{f}(\mathbf{x}_0)$ el mètode convergeix només linealment. També es pot actualitzar de tant en tant (cada l iteracions per exemple).
- Si no es poden calcular les derivades $D\mathbf{f}(\mathbf{x}_k)$ fàcilment, es poden aproximar per diferencies finites. Per exemple

$$\partial_{x_j}\mathbf{f}(\mathbf{x}) \approx \frac{\mathbf{f}(\mathbf{x} + h_j\mathbf{e}_j) - \mathbf{f}(\mathbf{x})}{h_j}$$

  amb $\mathbf{e}_j$ els vectors de la base cartesiana i h_j constants que poden ser totes iguals o no. Aquesta aproximació redueix també, en general, l'ordre de convergència del mètode.

- Quan n és molt gran s'acostumen a fer servir mètodes iteratius per resoldre els sistemes lineals a cada iteració del mètode de Newton. Es parla llavors de mètodes de Newton inexactes.
- Hi ha variants que fan més gran la conca d'atracció del mètode de Newton (Newton globalitzat) per tenir convergència a partir d'una condició inicial més lluny de la solució.

Zeros de polinomis

Considerem el cas important de trobar zeros de polinomis. Escriurem un polinomi de grau n en la forma

$$p(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n .$$

La forma $p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, pot ser més intuitiva, fa que en les fòrmules de recurrència que escriurem el indexos hagin d'anar sempre cap enrera. Com que els vectors a MATLAB comencem per l'index 1 els polinomis s'escriuen en aquest llenguatge en la forma $p(x) = a_1 x^n + a_2 x^{n-1} + \cdots + a_n x + a_{n+1}$, guardant únicament els coeficients $(a_1, a_2, \cdots, a_n, a_{n+1})$ en un vector de longitud $n+1$.

Evaluació d'un polinomi i les seves derivades, regla de Horner, i divisió sintètica i deflació.
Si agrupem el polinomi en la forma

$$p(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = (\cdots(a_0 x + a_1)x + a_2)x \cdots + a_{n-2})x + a_{n-1})x + a_n$$

en veu que $p(\alpha)$ es pot avaluar per la recurrència $b_0 = a_0$, $b_i = b_{i-1}\alpha + a_i$, $i = 1, \cdots, n$ amb $p(\alpha) = b_n$. Si no cal aprofitar els b_i es pot reescriure l'algorisme de Horner sense ells.

```
b(0)=a(0)
for i=1:n
  b(i)=b(i-1)*alpha+a(i)
end
p=b(n)
```

```
p=a(0)
for i=1:n
  p=p*alpha+a(i)
end
```

L'algorisme requereix n sumes i n productes. L'avaluació directa n sumes i $2n-1$ productes.

Els coeficients b_i de l'algorisme de Horner proporcionen la divisió sintètica per $(x-\alpha)$. En efecte si

$$p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = (x-\alpha)(b_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-2}x + b_{n-1}) + b_n$$

igualant coeficients tenim també $b_0 = a_0$, $a_i = -b_{i-1}\alpha + b_i$, $i = 1, \cdots, n$ amb $b_n = p(\alpha)$.

Considerem el desenvolupament en serie de Taylor de $p(x)$ al voltant de α

$$p(x) = p(\alpha) + p'(\alpha)(x-\alpha) + \frac{1}{2}p''(\alpha)(x-\alpha)^2 + \frac{1}{3!}p'''(\alpha)(x-\alpha)^3 + \cdots + \frac{1}{n!}p^{(n)}(\alpha)(x-\alpha)^n.$$

Si $q(x) = b_0x^{n-1} + b_1x^{n-2} + \cdots + b_{n-2}x + b_{n-1}$ tenim $p(x) = (x-\alpha)q(x) + p(\alpha)$,

$$q(x) = \frac{p(x)-p(\alpha)}{x-\alpha} = p'(\alpha) + (x-\alpha)[\frac{1}{2}p''(\alpha) + \frac{1}{3!}p'''(\alpha)(x-\alpha) + \cdots]$$

i $p'(\alpha) = q(\alpha)$.

Si $q(x) = (x-\alpha)(c_0x^{n-2} + c_1x^{n-3} + \cdots + c_{n-3}x + c_{n-2}) + c_{n-1} = (x-\alpha)r(x) + c_{n-1}$ amb $c_{n-1} = q(\alpha)$ llavors

$$p'(\alpha) = q(\alpha) = c_{n-1}.$$

Anàlogament

$$r(x) = \frac{q(x) - q(\alpha)}{x - \alpha} = \frac{1}{2}p''(\alpha) + (x - \alpha)[\frac{1}{3!}p'''(\alpha) + \frac{1}{4!}p^{(4)}(\alpha)(x - \alpha) + \cdots]$$

i $p''(\alpha) = 2r(\alpha)$.

Si $r(x) = (x - \alpha)(d_0 x^{n-3} + d_1 x^{n-4} + \cdots + d_{n-4}x + d_{n-3}) + d_{n-2} = (x - \alpha)s(x) + d_{n-2}$ amb $d_{n-2} = r(\alpha)$ llavors

$$p''(\alpha) = 2r(\alpha) = 2d_{n-2}.$$

Es a dir, si fem les iteracions

$$b_0 = a_0, \quad b_i = b_{i-1}\alpha + a_i \quad i = 1, \cdots, n$$

$$c_0 = b_0, \quad c_i = c_{i-1}\alpha + b_i \quad i = 1, \cdots, n-1$$

$$d_0 = c_0, \quad d_i = d_{i-1}\alpha + c_i \quad i = 1, \cdots, n-2$$

$$e_0 = d_0, \quad e_i = e_{i-1}\alpha + d_i \quad i = 1, \cdots, n-3$$

Llavors $p(\alpha) = b_n$, $p'(\alpha) = c_{n-1}$, $p''(\alpha) = 2d_{n-2}$, $p'''(\alpha) = 3!e_{n-3}$.

Si anomenem $d(i+1) = p^{(i)}(x_0)$, $i = 0, \cdots, k$, el codi MATLAB

```
function [d]=eval_pol_der(n,a,x0,k)
  d(1)=a(1);
  d(2:k+1)=0;
  for j=1:n
     for i=1:min(k,n+1-j);
        d(i+1)=d(i)+x0*d(i+1);
     end
     d(1)=a(j+1)+x0*d(1);
  end
  m=1;
  for j=2:k
     m=m*j;
     d(j+1)=m*d(j+1);
  end
end
```

calcula el polinomi $p(x)$ i les seves derivades fins a ordre k al punt x_0. El grau és n, els coeficients són $(a(1), \cdots, a(n+1))$ i k pot ser més gran que $n+1$. Fixem-nos que no es guarden els coeficients dels polinomis q, r, s, etc. Només es fa servir un vector d de longitud $k+1$ que contindrà la sortida.

Deflació. Si hem calculat l'arrel α de $p(x)$, fent la divisió sintètica obtindrem $p(x) = (x - \alpha)q(x) + p(\alpha)$ amb $p(\alpha) \approx 0$. Podem passar llavors a calcular els següent zero, treballant amb $q(x)$. Aquest procés es coneix com a deflació. També s'aplica en el càlcul de valors propis pel mètode de la potència, etc. La propagació dels errors d'arrodoniment es minimitzen si les arrels es troben en ordre creixent de mòdul.

Si tenim arrels complexes i el polinomi té coeficients reals és millor seguir treballant en aritmètica real. Com que les arrels complexes venen en parelles conjugades tenim que si $\alpha \in \mathbb{C}$ és arrel de $p(x)$ també ho serà $\bar{\alpha}$ i $p(x)$ serà divisible per $(x - \alpha)(x - \bar{\alpha}) = x^2 - 2\Re(\alpha)x + |\alpha|^2$ que té coeficients reals.

Veiem com fer la divisió de $p(x)$ per $x^2 + px + q$. Si

$$\begin{aligned} p(x) =& a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \\ =& (x^2 + px + q)(b_0x^{n-2} + b_1x^{n-3} + \cdots + b_{n-3}x + b_{n-2}) + Rx + S \end{aligned}$$

obtenim

$$\begin{aligned} a_0 =& b_0, & b_0 =& a_0, \\ a_1 =& b_1 + pb_0, & b_1 =& a_1 - pb_0, \\ a_i =& b_i + pb_{i-1} + qb_{i-2}, \quad i = 2, \cdots, n-2 & b_i =& a_i - pb_{i-1} - qb_{i-2}, \quad i = 2, \cdots, n-2 \\ a_{n-1} =& R + pb_{n-2} + qb_{n-3}, & R =& a_{n-1} - pb_{n-2} - qb_{n-3}, \\ a_n =& S + qb_{n-2} & S =& a_n - qb_{n-2} \end{aligned}$$

Si definim $b_{-1} = 0$, $b_{n-1} = R$ i $b_n = S - pb_{n-1}$, la recurrència es pot escriure com

$$b_0 = a_0, \quad b_i = a_i - pb_{i-1} - qb_{i-2},\ i = 1, \cdots, n, \quad R = b_{n-1},\ S = b_n + pb_{n-1}$$

ja que $b_n = a_n - pb_{n-1} - qb_{n-2} = S - pb_{n-1}$.

També, sense definir b_{-1},

$$b_0 = a_0, \quad b_1 = a_1 - pb_0, \quad b_i = a_i - pb_{i-1} - qb_{i-2},\ i = 2, \cdots, n, \quad R = b_{n-1},\ S = b_n + pb_{n-1}$$

Si des del principi dividim el polinomi pel coeficient de major grau tindrem un polinomi mònic de la forma $p(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$ i la recurrència serà

$$b_{-1} = 0,\ b_0 = 1,\ b_i = a_i - pb_{i-1} - qb_{i-2},\ i = 1, \cdots, n, \quad R = b_{n-1},\ S = b_n + pb_{n-1}.$$

Una vegada trobades totes les arrels es poden corregir els errors introduits amb la deflació refinant-les amb alguna iteració del mètode de Newton.

Fitació de les arrels d'un polinomi.

Hi ha molts criteris per trobar regions al pla complex que continguin totes les arrels d'un polinomi, o intervals que continguin els zeros reals. Sigui

$$p(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n,$$

amb coeficients reals.

Regla de Laguerre. Si dividim un polinomi $p(x)$ pel monomi $(x - L)$, amb $L > 0$, $p(x) = (x - L)(b_0 x^{n-1} + b_1 x^{n-2} + \cdots + b_{n-2} x + b_{n-1}) + b_n$ i $b_i > 0$, $i = 0, \cdots, n$ (o $b_i < 0$, $i = 0, \cdots, n$) llavors L és una fita superior de les arrels reals de $p(x)$.

Demostració: Si $x > L$ tots els termes a la dreta de la igualtat són estrictament positius (o negatius) i per tant $p(x) > 0$ (o $p(x) < 0$), i no hi ha arrels més grans que L.

Regla de Newton. Si $p(x)$ és un polinomi de grau n, $L \in \mathbb{R}$ i es verifica que $p(L), p'(L), \cdots, p^{(n)}(L)$ són tots positius (o tots negatius) llavors L és una fita superior de les arrels reals de $p(x)$.

Demostració: Considerem el desenvolupament en serie de Taylor de $p(x)$ al voltant de L

$$p(x) = p(L) + p'(L)(x - L) + \frac{1}{2} p''(L)(x - L)^2 + \cdots + \frac{1}{n!} p^{(n)}(L)(x - L)^n.$$

Si $x > L$ i $p(L), p'(L), \cdots, p^{(n)}(L)$ són tots positius llavors $p(x) > 0$.

Considerem el polinomi $p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$, amb coeficients reals o complexes i amb $a_0 \neq 0$. Sigui $z \in \mathbb{C}$ una de les seves arrels. Llavors es verifiquen

Regla de Lagrange.

$$|z| \leq \max\left\{1, \sum_{i=1}^{n} \left|\frac{a_i}{a_0}\right|\right\}$$

Demostracio: Suposem que $|z| > 1$. Si no, 1 ja es una fita superior. Llavors

$$|a_0||z^n| = \left|\sum_{i=1}^{n} a_i z^{n-i}\right| \leq \sum_{i=1}^{n} |a_i||z^{n-i}| \leq \sum_{i=1}^{n} |a_i||z|^{n-1},$$

i dividint per $|a_0||z|^{n-1}$ tenim el resultat.

Regla de Cauchy.

$$|z| \leq 1 + \max\left\{\left|\frac{a_1}{a_0}\right|, \left|\frac{a_2}{a_0}\right|, \cdots, \left|\frac{a_{n-1}}{a_0}\right|\right\}$$

Demostracio: Suposem que $|z| > 1$ igual que abans.

$$|a_0||z^n| \leq \sum_{i=1}^{n} |a_i||z^{n-i}| \leq \max|a_i| \sum_{i=1}^{n} |z|^{n-i} = \frac{|z|^n - 1}{|z| - 1} \max|a_i| \leq \frac{|z|^n}{|z| - 1} \max|a_i|,$$

i $|a_0|(|z| - 1) \leq \max|a_i|$.

Si es considera el polinomi $p(x) = x - M$ amb $M > 1$ es veu que la regla de Lagrange es converteix en una igualtat. El polinomi $p(x) = x^n - M(x^{n-1} + x^{n-2} + \cdots + x + 1)$ té una arrel tant a prop de $M + 1$ com es vulgui si n és prou gran. Això diu que la regla de Cauchy es també ajustada.

Les fites inferiors de les arrels reals positives o dels moduls de les complexes es troben fent el canvi $x \to 1/x$. Per fitar les arrels reals negatives es fa el canvi $x \to -x$.

Els mètodes generals que hem vist anteriorment per trobar zeros, es poden fer servir per les arrels reals. Per a les complexes es poden fer servir els mètodes de la secant o de Newton o d'altres de més específics per a polinomis.

El mètode de Bairstow. Donat un polinomi $P(x)$ de coeficients reals volem obtenir els seus factors quadràtics $x^2 + px + q$ corresponents a parelles d'arrels complexes conjugades. Podem escriure

$$\begin{aligned} P(x) =& x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n \\ =& (x^2 + px + q)(b_0 x^{n-2} + b_1 x^{n-3} + \cdots + b_{n-3}x + b_{n-2}) + Rx + S \end{aligned}$$

amb $b_0 = 1$ i a on hem assumit que el polinomi s'ha dividit pel coeficient de major grau. Volem resoldre, pel mètode de Newton, el sistema

$$\begin{aligned} R(p, q) &= 0 \\ S(p, q) &= 0. \end{aligned}$$

Llavors la convergència serà quadràtica.

En una iteració del mètode de Newton hem de trobar Δp_k, Δq_k, per calcular $p_{k+1} = p_k + \Delta p_k$, $q_{k+1} = q_k + \Delta q_k$, com solucions del sistema

$$\begin{aligned} \frac{\partial R}{\partial p}\Delta p + \frac{\partial R}{\partial q}\Delta q &= -\,R \\ \frac{\partial S}{\partial p}\Delta p + \frac{\partial S}{\partial q}\Delta q &= -\,S \end{aligned}$$

a on, per simplificar, hem eliminat els subindex k i la dependència de R i S amb p_k i q_k. Recordem que aquestes equacions son la linealització de les equacions $R(p+\Delta p, q+\Delta q) = 0$, $S(p+\Delta p, q+\Delta q) = 0$.

El càlcul de R i S es fa amb la deflació del factor x^2+px+q com ja s'ha vist. Si definim $b_{-1} = 0$, $b_0 = 1$, $b_{n-1} = R$ i $b_n = S - pb_{n-1}$ tenim

$$b_i = a_i - pb_{i-1} - qb_{i-2},\ i = 1, \cdots, n, \quad R = b_{n-1},\ S = b_n + pb_{n-1}$$

Amb això el sistema queda

$$\begin{aligned} \frac{\partial b_{n-1}}{\partial p}\Delta p + \frac{\partial b_{n-1}}{\partial q}\Delta q &= -\,b_{n-1} \\ \left(\frac{\partial b_n}{\partial p} + p\frac{\partial b_{n-1}}{\partial p} + b_{n-1}\right)\Delta p + \left(\frac{\partial b_n}{\partial q} + p\frac{\partial b_{n-1}}{\partial q}\right)\Delta q &= -\,b_n - pb_{n-1} \end{aligned}$$

o també com

$$\begin{aligned}\frac{\partial b_{n-1}}{\partial p}\Delta p+\frac{\partial b_{n-1}}{\partial q}\Delta q&=-\,b_{n-1}\\ \left(\frac{\partial b_n}{\partial p}+b_{n-1}\right)\Delta p+\frac{\partial b_n}{\partial q}\Delta q&=-\,b_n.\end{aligned}$$

Derivant $b_{-1}=0$, $b_0=1$ i $b_i=a_i-pb_{i-1}-qb_{i-2}$ tenim

$$\begin{aligned}-\frac{\partial b_i}{\partial p}&=b_{i-1}+p\frac{\partial b_{i-1}}{\partial p}+q\frac{\partial b_{i-2}}{\partial p}, \quad \frac{\partial b_0}{\partial p}=\frac{\partial b_{-1}}{\partial p}=0,\\ -\frac{\partial b_i}{\partial q}&=b_{i-2}+p\frac{\partial b_{i-1}}{\partial q}+q\frac{\partial b_{i-2}}{\partial q}, \quad \frac{\partial b_0}{\partial q}=\frac{\partial b_{-1}}{\partial q}=0,\end{aligned}$$

que es poden escriure, definint $\partial b_i/\partial p=-c_{i-1},\ \partial b_i/\partial q=-d_{i-2}$, com

$$c_{i-1}=b_{i-1}-pc_{i-2}-qc_{i-3},\ \ i=1,\cdots,n, \quad d_{i-2}=b_{i-2}-pd_{i-3}-qd_{i-4},\ \ i=1,\cdots,n$$

amb

$$c_{-1}=c_{-2}=0, \quad d_{-2}=d_{-3}=0$$

i que es resumeixen en una única iteració ja que $d_i=c_i$

$$c_i=b_i-pc_{i-1}-qc_{i-2},\ i=1,\cdots,n-1, \quad c_0=1,\ c_{-1}=0,$$

que és la mateixa que per trobar els b_i. Finalment el sistema queda

$$\begin{aligned}c_{n-2}\Delta p+c_{n-3}\Delta q&=b_{n-1}\\ (c_{n-1}-b_{n-1})\Delta p+c_{n-2}\Delta q&=b_n.\end{aligned}$$

Resumint, si tenim p i q volem trobar els coeficients del sistema

$$\begin{aligned} c_{n-2}\Delta p + c_{n-3}\Delta q =& b_{n-1} \\ (c_{n-1} - b_{n-1})\Delta p + c_{n-2}\Delta q =& b_n. \end{aligned}$$

A partir dels coeficients del polinomi mònic $P(x)$, a_i, $i = 1, \cdots, n$, els b_i i els c_i es calculen per les recurrències

$$b_0 = 1,\ b_{-1} = 0, \quad b_i = a_i - pb_{i-1} - qb_{i-2},\ i = 1, \cdots, n,$$

$$c_0 = 1,\ c_{-1} = 0, \quad c_i = b_i - pc_{i-1} - qc_{i-2},\ i = 1, \cdots, n-1.$$

Mètodes per trobar totes les arrels de $p(z)$ simultàniament.

Matriu acompanyant o de Frobenius. Donat el polinomi mònic $p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n$ la seva matriu de Frobenius és

$$C(p) = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_n \\ 1 & 0 & \dots & 0 & -a_{n-1} \\ 0 & 1 & \dots & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_1 \end{bmatrix}.$$

També es pot trobar definida com la transposta d'aquesta matriu. El seu polinomi característic és $det(C(p) - zI) = (-1)^n p(z)$. Llavors els seus valors propis són els zeros de $p(z)$. Es poden aplicar els mètodes per trobar valors propis per trobar els zeros de $p(z)$. Això és el que fa MATLAB amb la funció `roots`.

Mètode de Durand-Kerner (Weierstrass). Partim d'un polinomi mònic $p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n$. Si z_i, $i = 1, \cdots, n$ són aproximacions dels seus n zeros, volem trobar correccions Δz_i, $i = 1, \cdots, n$, que millorin l'aproximació. Les noves aproximacions seran $z_i' = z_i + \Delta z_i$. Impossant

$$\prod_{i=1}^{n} (z - z_i') = \prod_{i=1}^{n} (z - (z_i + \Delta z_i)) = p(z),$$

i desenvolupant en potències de Δz_i tenim

$$\prod_{k=1}^{n}(z-z_k) - \sum_{j=1}^{n}\Delta z_j \prod_{\substack{k=1\\k\neq j}}^{n}(z-z_k) + \sum_{\substack{j,l=1\\j<l}}^{n}\Delta z_j \Delta z_l \prod_{\substack{k=1\\k\neq j,l}}^{n}(z-z_k) - \cdots = p(z).$$

Negligint les potències més grans que 1 dels Δz_i tenim la definició dels increment en aquest mètode

$$\prod_{k=1}^{n}(z-z_k) - \sum_{j=1}^{n}\Delta z_j \prod_{\substack{k=1\\k\neq j}}^{n}(z-z_k) = p(z).$$

Fent ara $z = z_i$, $i = 1, \cdots, n$, s'obté

$$-\Delta z_i \prod_{\substack{k=1\\k\neq i}}^{n}(z_i - z_k) = p(z_i),$$

i llavors

$$\Delta z_i = -p(z_i)/\prod_{\substack{k=1\\k\neq i}}^{n}(z_i - z_k) \quad \text{i} \quad z_i' = z_i - \frac{p(z_i)}{\prod_{\substack{k=1\\k\neq i}}^{n}(z_i - z_k)}.$$

Hem assumit que les z_i són diferents dos a dos. Qualsevol implementació del algorisme ha evitar que coincideixin. No hi ha problema en que hi hagi zeros múltiples, però, com sempre, la convergència és més lenta.

La convergència del mètode es quadràtica quan cada z_i està a prop d'un zero. Això és conseqüència de que el mètode equival a trobar els zeros d'un sistema no lineal pel mètode de Newton. Recordem que els polinomis simètrics elementals en n variables z_i es defineixen com

$$\begin{aligned}
e_0(z_1,\cdots,z_n) &= 1\\
e_1(z_1,\cdots,z_n) &= \sum_{1\le j\le n} z_j\\
e_2(z_1,\cdots,z_n) &= \sum_{1\le j<k\le n} z_j z_k\\
e_3(z_1,\cdots,z_n) &= \sum_{1\le j<k<l\le n} z_j z_k z_l\\
&\cdots\\
e_n(z_1,\cdots,z_n) &= \prod_{1\le j\le n} z_j
\end{aligned}$$

i que

$$\prod_{1\le j\le n}(z-z_j) = e_0 z^n - e_1 z^{n-1} + e_2 z^{n-1} - \cdots + (-1)^n e_n.$$

Si $p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n$ i les z_i són les arrels de $p(z)$ llavors $p(z) = \prod_{1\le j\le n}(z-z_j)$ i $a_i = (-1)^i e_i$ (Fòrmules de Vieta).

Si apliquem el mètode de Newton al sistema de n equacions

$$\begin{aligned}
a_1 &= \sum_{1\le j\le n} z_j \\
a_2 &= -\sum_{1\le j<k\le n} z_j z_k \\
a_3 &= \sum_{1\le j<k<l\le n} z_j z_k z_l \\
&\cdots \\
a_n &= (-1)^n \prod_{1\le j\le n} z_j
\end{aligned}$$

obtenim el mètode de Durand-Kerner en la formulació que hem presentat, que és equivalent. A falta d'una millor opció es poden agafar les aproximacions inicials com

$$z_j = r\exp(2\pi j\mathrm{i}/n), \quad j = 0, \cdots, n-1$$

amb i la unitat imaginària i r un radi que pot ser $r = 3t$ amb t una fita superior del mòdul de les arrels, o un radi intermedi entre les fites superior i inferior del mòdul de les arrels.

Mètode de Ehrlich-Aberth. Partim del polinomi $p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n$. Si z_i, $i = 1, \cdots, n$ són aproximacions dels seus zeros, volem trobar correccions Δz_i, $i = 1, \cdots, n$, que millorin l'aproximació. En aquest mètode el nous iterats són $z_i' = z_i + \Delta z_i$ amb

$$\Delta z_i = \frac{\frac{p(z_i)}{p'(z_i)}}{1 - \frac{p(z_i)}{p'(z_i)} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{z_i - z_j}} = \frac{p(z_i)}{p'(z_i) - p(z_i) \sum_{\substack{j=1 \\ j \neq i}}^{n} (z_i - z_j)^{-1}}.$$

Una forma d'arribar a aquest mètode és la següent. Suposem, per simplificar, que tots els zeros de $p(z)$ són simples. Considerem les funcións racionals

$$R_i(z) = -\frac{p(z)}{\prod_{\substack{j=1 \\ j \neq i}}^{n} (z - z_j)}, \quad i = 1, \cdots, n,$$

amb z_i una aproximació del zero ξ_i de $p(z)$, $i = 1, \cdots, n$. Tenim llavors que $R_i(\xi_i) = 0$ i que $R_i(z)$ té pols a z_j si $j \neq i$ (assumim que $z_i \approx \xi_i$ però $z_i \neq \xi_i$). $R_i(z) \approx -(z - \xi_i)$ és gaire bé lineal a prop de les ξ_i ($z_i \approx \xi_i$, $i = 1, \cdots, n$) ja que $p(z) = \prod_{j=1}^{n} (z - \xi_j)$.

Apliquem una iteració del mètode de Newton unidimensional a $R_i(z)$ per trobar ξ_i

$$z_i' = z_i - \frac{R_i(z_i)}{R_i'(z_i)}.$$

La derivada de

$$R_i(z) = -\frac{p(z)}{\prod_{\substack{j=1\\j\neq i}}^{n}(z-z_j)}$$

es

$$\begin{aligned} R_i'(z) &= -\frac{p'(z)\prod_{\substack{j=1\\j\neq i}}^{n}(z-z_j) - p(z)\sum_{\substack{k=1\\k\neq i}}^{n}\prod_{\substack{j=1\\j\neq i,k}}^{n}(z-z_j)}{\prod_{\substack{j=1\\j\neq i}}(z-z_j)^2} \\ &= -\frac{p'(z) - p(z)\sum_{\substack{k=1\\k\neq i}}^{n}(z-z_k)^{-1}}{\prod_{\substack{j=1\\j\neq i}}(z-z_j)}, \end{aligned}$$

i

$$z_i' = z_i - \frac{R_i(z_i)}{R_i'(z_i)} = z_i + \frac{p(z_i)}{p'(z_i) - p(z_i)\sum_{\substack{k=1\\j\neq i}}^{n}(z_k-z_j)^{-1}}.$$

Estem iterant simultàniament n mètodes de Newton un per a cada arrel. El mètode es pot implementar tal com s'ha descrit (Jacobi style) o aprofitant cada z_i' ja calculat en les fòrmules per z_j' per $j > i$ (Gauss-Seidel style). Cada iteració de Newton convergiria quadràticament si la resta de z_j no es moguessin. Si es mouen totes a la vegada tal com hem descrit el mètode global és d'ordre 3.

Per veure-ho, escrivim la iteració com una de punt fix

$$z_i' = z_i + \Delta z_i = F_i(z_1, \cdots, z_n), \quad i = 1, \cdots, n,$$

amb $F_i(\xi_1, \cdots, \xi_n) = \xi_i$. Desenvolupant en serie de Taylor al voltant del conjunt d'arrels $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)$ tenim

$$z_i' = F_i(z_1, \cdots, z_n) = \xi_i + \sum_{i=1}^{n} \frac{\partial F_i}{\partial z_i}(\boldsymbol{\xi})(z_i - \xi_i) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 F_i}{\partial z_i \partial z_j}(\boldsymbol{\xi})(z_i - \xi_i)(z_j - \xi_j) + \cdots .$$

i

$$z_i' - \xi_i = \sum_{i=1}^{n} \frac{\partial F_i}{\partial z_i}(\boldsymbol{\xi})(z_i - \xi_i) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 F_i}{\partial z_i \partial z_j}(\boldsymbol{\xi})(z_i - \xi_i)(z_j - \xi_j) + \cdots .$$

El mateix es pot fer amb el mètode de Durand-Kerner. En aquest cas es pot veure que totes les derivades primeres són zero però no, en general, les segones. En el mètode de Ehrlich-Aberth es fan zero les primeres i segones derivades però no, en general, les terceres. Llavors el primer mètode és d'ordre 2 i el segon d'ordre 3.

Exemple: Apliqueu el dos mètodes anteriors per veure amb quina velocitat convergeigen a les arrels n-essimes de 1. Mireu com evoluciona $\max_i |\Delta z_i|$. Feu $n = 10, 20, 30$.

Exemple: Apliqueu el mètode de Durand-Kerner per trobar els zeros del polinomi de Wilkinson pertorbat $p(z) = (z-1)(z-2)\cdots(z-20) + \delta z^{19}$ amb $\delta = 2^{-23} \approx 10^{-7}$ per veure que una petita pertorbació dels coeficients pot modificar molt els zeros, es a dir, que el problema de trobar els zeros d'un polinomi a partir dels seus coeficients està mal condicionat.