Numerikus módszerek 2. Nemlineáris egyenletek közelítő megoldása

Egyenletmegoldás intervallumfelezéssel

Legyen $f:[a,b]\to\mathbf{R}$ folytonos, $f(a)<0,\quad f(b)>0$, és keressük az

$$f(x)=0$$

egyenlet egy $[a,b]$ -beli megoldását.

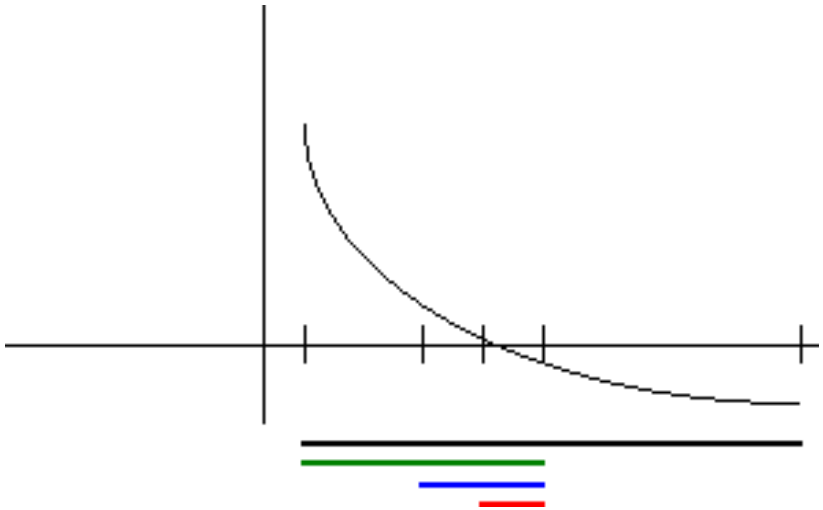
Bolzano tétele: Legyen *f* folytonos a véges, zárt $[a,b]$ intervallumon.
Tegyük fel, hogy $f(a)$ és $f(b)$ különböző előjelűek, pl. $f(a)<0,\quad f(b)>0$.
Akkor *f*-nek van (legalább egy) zérushelye ebben az intervallumban.

Felezzük szisztematikusan az $[a,b]$ intervallumot úgy, hogy a két fél-intervallumból mindig azt tekintjük, melyre a végpontokon felvett függvényértékek különböző előjelűek. Jelölje x_n az *n*-edik lépésben nyert részintervallum középpontját.

Akkor az így definiált sorozat a fenti egyenlet (egyik) gyökéhez konvergál. A konvergencia sebessége legalább egy 1/2 kvóciensű mértani sorozat konvergenciájának sebessége.

Egyenletmegoldás intervallumfelezéssel

Az intervallumfelezés algoritmusa szemléletesen:



Hibabecslés: Legyen $x_0 := \frac{a+b}{2}$, akkor nyilván

$$\left| x_n - x^* \right| \leq \frac{b-a}{2^n}$$

A Banach-fixpont-iterációs módszer

Legyen X Banach-tér, $f : X \to X$ egy leképezés, és keressük az

$$x = f(x)$$

egyenlet egy megoldását (ezt az f függvény egy **fixpontjának** nevezzük).

Banach-féle fixponttétel: Legyen X Banach-tér, $f : X \to X$ **kontrakció** X-en, azaz alkalmas $0 < q < 1$ számra $\| f(x) - f(y) \| \le q \cdot \| x - y \|$ teljesül minden $x, y \in X$ mellett.
Akkor f-nek egyetlenegy $x \in X$ fixpontja van,
és ez előáll az alábbi *iterációs sorozat* limeszeként:

$$x_0 \in X \text{ tetsz.}, \qquad x_{n+1} := f(x_n) \quad (n = 0,1,2,\ldots)$$

A valós függvények speciális esete. Ha $f : \mathbf{R} \to \mathbf{R}$ olyan, hogy $\max | f'(x) | < 1$, akkor f kontrakció, mert a *Lagrange-középértéktétel* miatt

$$| f(x) - f(y) | = | f'(\xi) | \cdot | x - y | \le (\max | f' |) \cdot | x - y |$$

A fixponttétel bizonyítása: Két egymást követő tag eltérése:

$$\| x_{n+1} - x_n \| = \| f(x_n) - f(x_{n-1}) \| \le q \| x_n - x_{n-1} \| = q \| f(x_{n-1}) - f(x_{n-2}) \| \le q^2 \| x_{n-1} - x_{n-2} \| \le$$
$$\ldots \le q^n \| x_1 - x_0 \|$$

Ezt felhasználva megmutatjuk, hogy az (x_n) sorozat Cauchy-sorozat *X*-ben:

$$\| x_{n+k} - x_n \| = \| x_{n+k} - x_{n+k-1} + x_{n+k-1} - x_{n+k-2} + \ldots + x_{n+1} - x_n \| \le$$
$$\le \| x_{n+k} - x_{n+k-1} \| + \| x_{n+k-1} - x_{n+k-2} \| + \ldots + \| x_{n+1} - x_n \| \le$$
$$\le (q^{n+k-1} + q^{n+k-2} + \ldots + q^n) \cdot \| x_1 - x_0 \| \le$$
$$\le (q^n + q^{n+1} + q^{n+2} + \ldots) \cdot \| x_1 - x_0 \| = \frac{q^n}{1-q} \cdot \| x_1 - x_0 \| \to 0$$

Ezért a sorozat konvergens, $x_n \to x \in X$. Megmutatjuk, hogy a limesz fixpontja *f*-nek. A rekurzív definíció szerint: $x_{n+1} = f(x_n)$ A bal oldal nyilván *x*-hez tart. A jobb oldal *f folytonossága miatt f* (*x*)-hez. Innen *x* = *f* (*x*).

Végül igazoljuk, hogy *csak egy* fixpont van. Ha *x*, *y* két *különböző* fixpont volna, akkor

$$0 < \| x - y \| = \| f(x) - f(y) \| \le q \cdot \| x - y \| < \| x - y \|$$

volna, ami lehetetlen.

A Banach-fixpont-iterációs módszer, példák:

1. Oldjuk meg az $x = \frac{1}{2}\cos x$ egyenletet.

Az $f(x) := \frac{1}{2}\cos x$ függvény kontrakció, mert $|f'(x)| = \frac{1}{2}|\sin x| \le \frac{1}{2}$. Ezért egyetlenegy fixpont létezik, és pl. az $x_0 := 0$, $x_{n+1} := \frac{1}{2}\cos x_n$ sorozat ide konvergál. A sorozat első néhány eleme (4 tizedesjegy pontossággal):

0.0000, 0.5000, 0.4387, 0.4526, 0.4496, 0.4502, 0.4501, 0.4501, 0.4501, ...

2. Legyenek $B \in \mathbf{M}_{N\times N}$, $g \in \mathbf{R}^N$ adottak, és oldjuk meg az $x = Bx + g$ egyenletet.

Ha $\| B \| < 1$, akkor az $f(x) := Bx + g$ leképezés kontrakció, mert

$$\| f(x) - f(y) \| = \| Bx + g - By - g \| \le \| B \| \cdot \| x - y \|$$

Ezért ekkor egyetlenegy fixpont létezik, és pl. az $x_0 := \mathbf{0}$, $x_{n+1} := Bx_n + g$ vektorsorozat ide konvergál.

A Newton-módszer egyváltozós függvényekre

Legyen $f:(a,b)\to\mathbf{R}$ adott függvény. Keressük az

$$f(x)=0$$

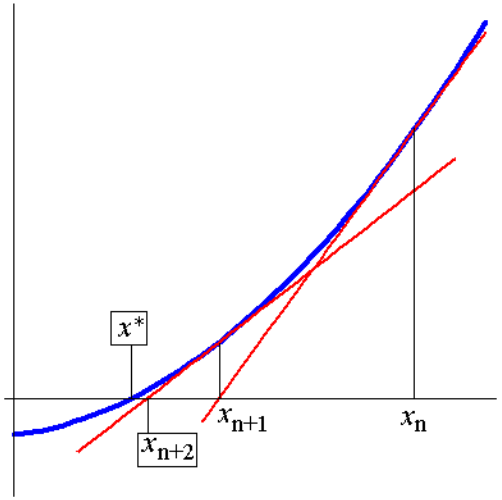
egyenlet egy (a,b)-beli megoldását.

A Newton-módszer: Ha x_n a megoldás egy közelítése, akkor legyen a következő közelítés az x_n-beli érintő egyenes zérushelye. Az érintő egyenlete: $y=f(x_n)+f'(x_n)\cdot(x-x_n)$, innen:

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}\quad (n=0,1,2,\ldots)\qquad\qquad (x_0\in(a,b)\text{: kezdeti közelítés})$$

A Newton-módszer egyváltozós függvényekre

A módszer szemléletes jelentése (x^*: pontos megoldás):



Ha f kétszer folytonosan differenciálható, f-nek van gyöke (a,b)-n, és erre $f'(x^*) \neq 0$ teljesül, akkor a Newton-módszer **kvadratikusan konvergál** minden, az x^* pontos megoldáshoz elég közeli x_0 kezdeti közelítés esetén, azaz alkalmas $C > 0$ szám mellett:

$$|x_{n+1} - x^*| \leq C \cdot |x_n - x^*|^2$$

Bizonyítás: A *Lagrange-középértéktételt* használjuk (kétszer is):

$$x_{n+1} - x^* = x_n - x^* - \frac{f(x_n) - f(x^*)}{f'(x_n)} = x_n - x^* - \frac{f'(t)(x_n - x^*)}{f'(x_n)} =$$

$$= \frac{f'(x_n) - f'(t)}{f'(x_n)} \cdot (x_n - x^*) = \frac{f''(s)}{f'(x_n)} \cdot (x_n - t) \cdot (x_n - x^*)$$

Mivel $f'(x^*) \neq 0$, azért a deriváltfüggvény x^* *egy egész* (*zárt*) *környezetében* is zérustól különböző marad. Itt pedig:

$$|x_{n+1} - x^*| \leq \frac{\max|f''|}{\min|f'|} \cdot |x_n - t| \cdot |x_n - x^*| \leq \frac{\max|f''|}{\min|f'|} \cdot |x_n - x^*|^2 = C \cdot |x_n - x^*|^2$$

A Newton-módszer, példa:

Legyen A rögzített pozitív szám, és $f(x) := x^2 - A.$ $(\Rightarrow \quad f'(x) = 2x)$
Ekkor az

$$f(x) = 0$$

egyenlet egyetlen (pozitív) megoldása: $x = \sqrt{A}$.

Kiindulva egy tetszőleges $x_0 > 0$ kezdeti közelítésből (pl. $x_0 := A$), a Newton módszer:

$$x_{n+1} := x_n - \frac{x_n^2 - A}{2x_n} = \frac{1}{2} \cdot \left(x_n + \frac{A}{x_n} \right)$$

A sorozat *nagyon gyorsan* $\sqrt{A}$-hoz tart.

Megjegyzés: hasonlóan lehet bármely (egész kitevős) gyökvonás közelítésére a Newton-módszert használni.

A Newton-módszer változatai

Probléma: a derivált számítása.

A szelőmódszer: itt $f'(x_n) \approx \dfrac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$. Legyenek tehát x_0, x_1 kezdő közelítések, és

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

Ha f kétszer folytonosan differenciálható, f-nek van gyöke (a,b)-n, és erre $f'(x^*) \neq 0$ teljesül, akkor a szelőmódszer minden, az x^* pontos megoldáshoz elég közeli x_0, x_1 kezdeti közelítések esetén legalább egy mértani sorozat sebességével konvergál,
azaz alkalmas $C > 0$, $0 < q < 1$ mellett:

$$|x_n - x^*| \leq C \cdot q^n$$

A Newton-módszer változatai

A Steffensen-módszer:

Legyen az $f:\mathbf{R}\to\mathbf{R}$ kétszer folytonosan differenciálható függvénynek pontosan egy x^* zérushelye. Tegyük fel, hogy $f'(x^*)\neq 0$. Akkor minden, x^*-hoz elég közel levő valós kezdeti x_0 közelítésből kiindulva, az

$$x_{n+1} := x_n - \frac{f(x_n)^2}{f(x_n + f(x_n)) - f(x_n)} \qquad (n = 0,1,2,\ldots)$$

sorozat kvadratikusan konvergál x^*-hoz.

Bizonyítás: A *Lagrange-középértéktételt* használjuk:

$$f(x_n + f(x_n)) - f(x_n) = f'(t)\cdot(x_n + f(x_n) - x_n) = f'(t)\cdot f(x_n)$$

ezért

$$x_{n+1} - x^* = x_n - x^* - \frac{f(x_n)}{f'(t)} = x_n - x^* - \frac{f(x_n) - f(x^*)}{f'(t)} = x_n - x^* - \frac{f'(s)}{f'(t)}(x_n - x^*) =$$
$$= \frac{f'(t) - f'(s)}{f'(t)}\cdot(x_n - x^*) = \frac{f''(w)}{f'(t)}\cdot(t-s)\cdot(x_n - x^*)$$

Mivel $f'(x^*) \neq 0$, azért a deriváltfüggvény x^* *egy egész* (*zárt*) *környezetében* is zérustól különböző marad. Itt pedig:

$$|x_{n+1} - x^*| \leq \frac{\max|f''|}{\min|f'|}\cdot|t-s|\cdot|x_n - x^*| \leq \frac{\max|f''|}{\min|f'|}\cdot|x_n - x^*|^2 = C\cdot|x_n - x^*|^2$$

Differenciálás Banach-terek között értelmezett leképezésekre

Legyenek X,Y Banach-terek. Az $F : X \to Y$ leképezés az $x \in X$ pontban **differenciálható**, deriváltja pedig az $A: X \to Y$ **folytonos lineáris leképezés**, ha a **0** egy alkalmas környezetéből választott h vektorokra teljesül, hogy

$$F(x+h) = F(x) + Ah + o(h)$$

ahol $o(h)$ olyan kifejezés, hogy $\dfrac{o(h)}{\| h \|} \to 0 \quad (h \to \mathbf{0})$.

A derivált lineáris leképezést még így is jelöljük: $F'(x)$ vagy $DF(x)$.

Példa: $F : \mathbf{R}^N \to \mathbf{R}$, $F(x) := \langle Ax, x \rangle$ (ahol $A \in \mathbf{M}_{N \times N}$ önadjungált mátrix). Akkor:

$$F(x+h) = \langle A(x+h), x+h \rangle = \langle Ax, x \rangle + 2\langle Ax, h \rangle + \langle Ah, h \rangle = F(x) + \langle 2Ax, h \rangle + O(h^2),$$

ezért $F'(x) = 2Ax \in \mathbf{R}^N$.

Általánosított Newton-módszer

A Newton-módszer az $F(x)=0$ egyenletre:

$$x_{n+1} = x_n - (DF(x_n))^{-1} F(x_n) \quad (n = 0,1,2,\ldots)$$

Másképpen felírva: $x_{n+1} = x_n - w_n \quad (n = 0,1,2,\ldots)$

ahol a w_n javító tag az alábbi *lineáris* egyenlet megoldása:

$$DF(x_n) w_n = F(x_n)$$

Ha F kétszer folytonosan differenciálható, F-nek van gyöke X-ben, és $DF(x^*)$ *reguláris* (azaz invertálható, és az inverze is folytonos), akkor a Newton-módszer kvadratikusan konvergál minden, az x^* pontos megoldáshoz elég közeli x_0 kezdeti közelítés esetén, tehát alkalmas $C > 0$ mellett:

$$\| x_{n+1} - x^* \| \le C \cdot \| x_n - x^* \|^2$$

Általánosított Newton-módszer, példa

Mátrixinverzió. Legyen $A \in \mathbf{M}_{N\times N}$ egy reguláris mátrix. Tetszőleges $X \in \mathbf{M}_{N\times N}$ reguláris mátrixra jelölje:

$$F(X) := X^{-1} - A$$

Akkor $F : \mathbf{M}_{N\times N} \to \mathbf{M}_{N\times N}$, és az $F(X) = 0$ egyenlet egyetlen megoldása: $X = A^{-1}$.

Alkalmazzuk a mátrixegyenletre a Newton-módszert. Az F leképezés deriváltját számítva:

$$F(X+H) = (X+H)^{-1} - A = (X(I + X^{-1}H))^{-1} - A = (I + X^{-1}H)^{-1}X^{-1} - A$$

Ha most H elég kis normájú (tetszőleges mátrixnormában), akkor $\| X^{-1}H \| \le \| X^{-1} \| \cdot \| H \| < 1$. Felhasználva a $\| B \| < 1$ esetén fennálló $(I - B)^{-1} = I + B + B^2 + B^3 + B^4 + \ldots$ egyenlőséget, (melyből $(I + B)^{-1} = I - B + (\| B \|^2)$ következik):

$$F(X+H) = (I + X^{-1}H)^{-1}X^{-1} - A = (I - X^{-1}H + o(H))X^{-1} - A =$$
$$= X^{-1} - X^{-1}HX^{-1} + o(H) - A = F(X) - X^{-1}HX^{-1} + o(H)$$

ahonnan

$$DF(X)H = -X^{-1}HX^{-1} \qquad \Rightarrow \qquad DF(X)^{-1}W = -XWX$$

Általánosított Newton-módszer, példa

Így a Newton-módszer algoritmusa:

$$X_{n+1} := X_n - (DF(x_n))^{-1}(X_n^{-1} - A) = X_n + X_n(X_n^{-1} - A)X_n = \\ = X_n(2I - AX_n).$$

A közelítés hibájára: $\| A^{-1} - X_n \| = \| A^{-1}(I - AX_n) \| \le \| A^{-1} \| \cdot \| I - AX_n \|$. Ugyanakkor $\| I - AX_n \|$ nagyon gyorsan tart 0-hoz (ha a kezdeti közelítés elég jó volt), mert:

$$I - AX_{n+1} = I - AX_n(2I - AX_n) = I - 2AX_n + AX_nAX_n = (I - AX_n)^2,$$

ahonnan

$$\| I - AX_{n+1} \| \le \| I - AX_n \|^2$$