Přednáška 13 – Newtonova metoda

opáčko – Newtonova metoda:

- x_{k+1} získám jako řešení $J_{x_k} \cdot x_{k+1} = J_{x_k} \cdot x_k - F(x_k)$

  kde $[J_{x_k}]_{ij} = \frac{\partial F_i}{\partial x_j}(x_k)$ je Jacobiho matice F v bodě x_k

- „pokud konverguje, pak $\frac{\|x^* - x_{k+1}\|}{\|x^* - x_k\|^2} \xrightarrow{k \to +\infty} C \neq 0$"

  ... tzv. kvadratická konv.

Problém 1: co když nekonverguje?

→ Newton je odvozený pomocí Taylora ⇒ „funguje jen lokálně"
tj. můžeme zaručit konvergenci pokud začneme v x_0 dostatečně blízko přesného řešení x^*.

- **Řešení 1 – „začni blízko"**
  → můžeme začít z mnoha bodů (a doufat, že nějaký zkonverguje) ⇝ pokud nějaký funguje, uvidíme to velmi rychle díky kvadratické konv.
  → můžeme použít jinou metodu pro „počáteční fázi" a postupně zkoušet pouštět Newtona ⇝ jak odvodit „jiné metody"?
  pomocí reformulace na hledání tzv. pevného bodu.

$$\tilde{x} \text{ je pevný bod } F(x) \overset{def.}{\iff} F(\tilde{x}) = \tilde{x}$$

Tedy např.: x^* je kořen $F(x) \iff \underbrace{x^* \text{ je pevný bod fce } F(x)+x}_{\text{tj. } x^* = F(x^*) + x^*}$

..... a na hledání pevných bodů máme hezkou větu:

Banachova věta o pevném bodě

Nechť $M \subseteq \mathbb{R}^n$ je uzavřená a $G: M \to M$ je zobrazení pro které platí:

$$\exists q < 1 \text{ t. že } \forall x, y \in M \quad \|G(x) - G(y)\| \leq q \cdot \|x - y\| \quad (*)$$

($\equiv$ říkáme, že G je tzv. kontrakce na M)

Pak $\exists! \tilde{x}$ t. že $G(\tilde{x}) = \tilde{x}$ a navíc pro libovolné x_0 platí, že posloupnost $x_0, \underbrace{G(x_0)}_{x_1}, \underbrace{G(G(x_0))}_{x_2}, \underbrace{G(G(G(x_0)))}_{x_3}, \ldots$ konverguje k $\tilde{x}$ lineárně s rychlostí q, tj. $\lim\limits_{k \to +\infty} \dfrac{\|x_{k+1} - \tilde{x}\|}{\|x_k - \tilde{x}\|} = q$.

Poznámka: pokud $\forall a \in M$ platí $\left\| \left[\frac{d}{dx} G(a)\right] \right\| \leq q < 1$, pak platí i $(*)$.

příklad:

$F(x) = x^2 + x - 2$ ..... $F: \mathbb{R} \to \mathbb{R}$ & $F(x) = 0 \iff x^2 = 2 - x \iff$

- $x = \pm\sqrt{2-x}$ .... $G_1(x)$
- $x = -1 + \frac{2}{x}$ .... $G_2(x)$

$G_1(x) := \sqrt{2-x}$

- $G_1'(a) = \dfrac{-1}{2\sqrt{2-a}} \rightsquigarrow |G_1'(a)| = \dfrac{1}{2\sqrt{2-a}} \rightsquigarrow |G_1'(a)| < 1 \iff \frac{1}{4} < 2 - a \iff a < 7/4$
- $(c, d] \subset (-\infty, 7/4) \Rightarrow G_1((c,d)) = (G_1(d), G_1(c)) = (\sqrt{2-d}, \sqrt{2-c}) \rightsquigarrow d < 7/4 \;\&\; \sqrt{2-c} < d \;\&\; \sqrt{2-d} \geq c$. (stačí brát $c<0$)
  $\Rightarrow$ vezmeme libovolné $d < 7/4$ ... např. $d = 3/2$. Pak $\underline{c}$ t. že $\sqrt{2-c} = d$, tj. $c = -1/4$.
  Pak $G_1([-1/4, 3/2]) = [\sqrt{1/2}, 3/2] = [\sqrt{2}/2, 3/2] \subset [-1/4, 3/2] \subset (-\infty, 7/4)$
- Tudíž lze použít větu a posloupnost $x_0, G_1(x_0), G_1(G_1(x_0)), \ldots$ konverguje ke kořenu F pro lib. $x_0 \in [-1/4, 3/2]$

$G_2(x) = -1 + 2/x$

- $G_2'(a) = -2/a^2 \rightsquigarrow |G_2'(a)| < 1 \iff 2/a^2 < 1 \iff a \in (-\infty, -\sqrt{2}) \cup (\sqrt{2}, +\infty)$
- $(c, d) \subset (-\infty, -\sqrt{2}) \Rightarrow G_2((c,d)) = (2/d - 1, 2/c - 1) \rightsquigarrow d < -\sqrt{2} \;\&\; 2/c - 1 \leq d \;\&\; 2/d - 1 \geq c \iff \frac{2}{d+1} \leq c \leq \frac{2}{d} - 1 \;\&\; d < -\sqrt{2}$
  Např. $d = -3/2$ & $c \in [-4, -7/3]$. Pak $G_2([-4, -3/2]) = [-7/3, -3/2] \subset [-4, -3/2] \subset (-\infty, -\sqrt{2})$
- $(c, d) \subset (\sqrt{2}, +\infty) \Rightarrow G_2((c,d)) = (2/d - 1, 2/c - 1) \rightsquigarrow c > \sqrt{2} \;\&\; 2/d - 1 \geq c \;\&\; 2/c - 1 \leq d \iff c > \sqrt{2} \;\&\; 2/c - 1 \leq d \leq \frac{2}{c+1}$
  Pak nutně $2/c - 1 < \frac{2}{c+1} \iff c \in (-2, 1)$ .... tedy nelze použít větu o pevném bodě $\rightsquigarrow$ buď máme $|G_2'| < 1$ nebo $G_2: M \to M$, ale nikdy obojí pro $M \subset (\sqrt{2}, +\infty)$.
- Tudíž lze použít větu a posloupnost $x_0, G_2(x_0), G_2(G_2(x_0)), \ldots$ konverguje ke kořenu F pro lib. $x_0 \in [-4, -3/2]$

Python demos: konvergence & iterace funkcí (Julia & & spojitost s Newtonem v $\mathbb{C}$)

Řešení 2 – „směr dobrý, stačí nepřestřelit"

→ klasickým problémem je situace z obrázku z minulé přednášky: začali jsme se přibližovat zprava ke kořenu, ale pak jsme „přestřelili"
(a díky štěstíčku jsme zkonvergovali k druhému kořenu).

→ místo toho lze postupovat „uměřeněji": spočítáme řešení $\boxed{J_{x_k} d_k = F(x_k)}$ a položíme $\boxed{x_{k+1} = x_k + \tau_k d_k}$

τ_k je délka kroku ve směru d_k

d_k je tzv. směr Newtonovy metody

Jedna možnost (tzv. Armijo rule): $\tau = 2^{-m}$ kde $m \in \mathbb{N}$ je nejmenší přirozené číslo pro které $\| F(x_k + 2^{-m} d_k) \| \leq (1 - \alpha 2^{-m}) \| F(x_k) \|$

volný parametr, standardní volba $\sim 10^{-4}$

Problém 2: jak získat J_{x_k}?

→ v praxi pro hodně problémů nemám explicitní formulka pro F ⇝ mám pouze program, který spočítá $x \longmapsto F(x)$
(a i toto vyhodnocení může trvat dlouho a být náročné).

⇝ nemám přístup k $\frac{\partial F_i}{\partial x_j}$ ⟹ musím aproximovat.

$$\frac{\partial F_i}{\partial x_j}(a) = \lim_{h \to 0} \frac{F_i(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - F_i(a_1, \dots, a_{j-1}, a_j - h, a_{j+1}, \dots, a_n)}{2h}$$

$$\approx \frac{F_i(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - F_i(a_1, \dots, a_{j-1}, a_j - h, a_{j+1}, \dots, a_n)}{2h} =: FD_j(F_i, a, h)$$

Finite difference (konečná diference) pro funkci $F_i(x)$ v bodě a a v j-té proměnné

⟹ matici J_{x_k} lze aproximovat maticí $\tilde{J}_{x_k}$ definovanou jako $[\tilde{J}_{x_k}]_{ij} := FD_j(F_i, x_k, h)$,
pro nějaké malé $h > 0$.

⇝ k sestavení $\tilde{J}_{x_k}$ potřebujeme $2n$ vyčíslení $F(x)$ na 1 krok → to může být příliš výpočetně náročné ⇝ existují další aproximační varianty, které za cenu dalších aproximací dále sníží potřebný počet vyčíslení $F(x)$ pro 1 iteraci ⇝ tzv. Quasi-Newton metody.