Tomáš Oberhuber

Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University in Prague

Tomáš Oberhuber

Video na Youtube

Iterační metody pro nelineární rovnice

Buď f reálná funkce jedné reálné proměnné. Budeme hledat řešení rovnice

$$f(x) = 0.$$

Řešení se skládá ze dvou kroků:

- separace kořenů
  - metody nekonvergují globálně, tj. pro libovolné x_0, jako metody pro linerání soustavy – nejprve tedy musíme najít interval nebo intervaly, tak že obsahují vždy jen jeden kořen rovnice
- výpočet kořene se zadanou přesností

Kořen rovnice budeme značit jako α.

Separace kořenů

- separaci kořenů dokážeme provést jen pro algebraické rovnice např. pomocí **Sturmových posloupností** – viz Wikipedie, Numerical Recipes.
- obecně to ale nejde a musíme mít nějaký předběžný odhad podle řešené úlohy
  - bez něj často nejsme schopní tyto rovnice řešit

To, že separaci lze provést a za jakých podmínek, říká následující věta:

Separace kořenů

Theorem 1

Nechť f je reálná funkce jedné reálné proměnné, spojitá na intervalu $< a, b >$ *a nechť* $f(a)f(b) < 0$. *Potom rovnice* $f(x) = 0$ *má na* (a, b) *alespoň jeden kořen a pokud navíc* $f'(x)$ *na* (a, b) *nemění znaménko, pak je tento kořen jediný.*

- budeme konstruovat posloupnost $\{x_k\}_{k=1}^{\infty}$, která bude konvergovat ke kořenu rovnice

Výpočet kořene - metoda bisekcí

- nejjednodušší metoda pro výpočet kořene
- je pomalá, ale poměrně dost robustní
- konstrujeme posloupnost krajních mezí intervalů $< l_k, r_k >$, ve kterém se nachází kořen
- každý následující interval má poloviční délku toho předchozího

Výpočet kořene - metoda bisekcí

- volíme $l_0 = a$ a $r_0 = b$
- v každé iteraci pak počítáme

$$x_k = \frac{l_k + r_k}{2},$$

$$l_{k+1} := \begin{cases} x_k \text{ pokud sign } f(x_k) = \text{sign } f(l_k), \\ l_k \text{ jinak} \end{cases}$$

$$r_{k+1} := \begin{cases} x_k \text{ pokud sign } f(x_k) = \text{sign } f(r_k), \\ r_k \text{ jinak} \end{cases}$$

Výpočet kořene - metoda bisekcí

- označíme-li $\mathcal{I}_k =< l_k, r_k >$, pak vždy platí, že $\alpha \in \mathcal{I}_k$ a

$$|\alpha - x_k| \leq |\mathcal{I}_k| = \frac{b-a}{2^k} \to 0$$

Výpočet kořene - metoda bisekcí

Example 2

Použijte metodu bisekcí pro nalezení kořene polynomu $x^3 + 4x^2 - 10$ v intervalu $< 1, 2 >$ ($\alpha \approx 1.3652$).

k	l_k	r_k	$f(l_k)$	$f(r_k)$	x_k	$f(x_k)$
1	1.0	2.0	-5.0	14.0	1.5	2.375
2	1.0	1.5	-5.0	2.375	1.25	-1.796875
3	1.25	1.5	-1.796875	2.375	1.375	0.162109375
4	1.25	1.375	-1.796875	0.162109375	1.3125	-0.848388671875

Výpočet kořene - metoda bisekcí

Remark 3

Metoda bisekcí koverguje i pro úlohu

$$\tan x = 0,$$

pro $l_0 = 1$, $r_0 = 3$. *Posloupnost* x_k *konverguje k* $\frac{\pi}{2}$, *což ale samozřejmě není kořen. V této úloze jsme ale porušili předpoklad spojitosti funkce* $f(x)$ *na odseparovaném intervalu.*

Iterativní metody pro hledání kořenů

- abychom urychlili konvergenci, využijeme více informací o funkci f, než jen její znaménko v bodě x_k
- máme-li určité x_k, ptáme se, jak nejlépe ho změnit, aby se nové x_{k+1} co nejlépe přiblížilo k α
- pokud je f diferencovatelná, pak platí

$$
\begin{aligned}
0 = f(\alpha) &= f(x_k) + f'(\xi)(\alpha - x_k),\\
\alpha - x_k &= \frac{-f(x_k)}{f'(\xi)},\\
\alpha &= x_k - \frac{f(x_k)}{f'(\xi)}
\end{aligned}
$$

- $f'(\xi)$ ale neznáme, lze ho napočítat jen přibližně

Iterativní metody pro hledání kořenů

- označme aproximaci $f'(\xi)$ jako q_k
- dostáváme obecný předpis

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

- vidíme, že platí $\varphi(\alpha) = \alpha$
- otázka je, za jaké podmínky bude metoda s tímto předpisem konvergovat

Iterativní metody pro hledání kořenů

Theorem 4

Nechť platí $\varphi(\alpha) = \alpha$, φ *je diferencovatelná na okolí*

$$V = \{x \in \mathbb{R} \mid |x - \alpha| \leq r\} \equiv < \alpha - r, \alpha + r >,$$

a nechť pro všechna $x \in V$ *je*

$$\left|\varphi'(x)\right| \leq K < 1.$$

Potom posloupnost $\{x_k\}_{k=1}^{\infty}$ *daná vztahem*

$$x_{k+1} = \varphi(x_k),$$

konverguje k α *pro libovolné* $x_0 \in V$.

Důkaz.

Video na Youtube

Iterativní metody pro hledání kořenů

Remark 5

Je-li $\varphi'(x)$ *spojitá na nějakém okolí* α*, pak posloupnost* $\{x_k\}_{k=1}^{\infty}$ *konverguje, je-li* $\varphi'(\alpha) < 1$*.*

Definition 6

Řekneme, že iterativní metoda daná vztahem $x_{k+1} = \varphi(x_k)$ má řád konvergence m, jestliže platí

$$|x_{k+1} - \alpha| \leq C\,|x_k - \alpha|^m .$$

Iterativní metody pro hledání kořenů

Remark 7

Nechť φ má na okolí α spojité derivace do řádu m včetně (tj. $\varphi \in C^{(m)}(H_\alpha)$) a platí

$$\varphi'(\alpha) = \varphi''(\alpha) = \ldots = \varphi^{(m-1)}(\alpha) = 0,$$

a

$$\varphi^{(m)}(x) \neq 0,$$

na H_α. Pak pro libovolne $x_k \in H_\alpha$ platí

$$\begin{aligned} x_{k+1} - \alpha &= \varphi(x_k) - \varphi(\alpha) = \\ &= \varphi'(\alpha)(x_k - \alpha) + \ldots + \frac{\varphi^{(m-1)}(\alpha)}{(m-1)!}(x_k - \alpha)^{m-1} \\ &\quad + \frac{\varphi^{(m)}(\xi)}{m!}(x_k - \alpha)^m \\ &= \frac{\varphi^{(m)}(\xi)}{m!}(x_k - \alpha)^m . \end{aligned}$$

Iterativní metody pro hledání kořenů

Remark 8

Buď funkce $f \in C^{(1)}(\mathbb{R})$ a nechť $f'(\alpha) \neq 0$. Pak platí:

$$f(x_k) = \underbrace{f(\alpha)}_{=0} + f'(\xi)(x_k - \alpha)$$

a tedy

$$|x_k - \alpha| \leq \frac{|f(x_k)|}{|f'(\xi)|} \leq C\,|f(x_k)|\,.$$

Jako odhad chyby a zároveň kritérium pro zastavení výpočtu proto bereme $|f(x_k)|$.

Metoda regula falsi

- u této metody definujeme $x_0 = l_0 = a$, $x'_0 = r_0 = b$
- vytvoříme úsečku mezi body $(x_0, f(x_0))$ a $\left(x'_0, f\left(x'_0\right)\right)$
- protože $f(x_0)$ a $f\left(x'_0\right)$ mají různá znaménka, úsečka musí protnout osu x, tento průsečík označíme jako x_1
- dále definujeme

$$l_1 := \begin{cases} x_1 & \text{pokud sign } f(x_1) = \text{sign } f(l_0), \\ l_0 & \text{jinak} \end{cases}$$

$$r_1 := \begin{cases} x_1 & \text{pokud sign } f(x_1) = \text{sign } f(r_0), \\ r_0 & \text{jinak} \end{cases}$$

$$x'_1 := \begin{cases} l_1 & \text{pokud sign } f(x_1) = \text{sign } f(r_1), \\ r_1 & \text{jinak} \end{cases}$$

- celý postup pak iterativně opakujeme

Metoda regula falsi

- rovnice úsečky mezi $(x_k, f(x_k))$ a $\left(x'_k, f\left(x'_k\right)\right)$ má tvar

$$y(x) = f(x_k) + \frac{f(x_k) - f\left(x'_k\right)}{x_k - x'_k}(x - x_k)$$

- položíme $y(x) = 0$ a řešíme rovnici

$$0 = f(x_k) + \frac{f(x_k) - f\left(x'_k\right)}{x_k - x'_k}(x_{k+1} - x_k)$$

- výsledkem je

$$x_{k+1} = x_k - \frac{x_k - x'_k}{f(x_k) - f\left(x'_k\right)} f(x_k)$$

- zobrazení φ má tvar

$$\varphi(x) = x - \frac{x - x'_k}{f(x) - f\left(x'_k\right)} f(x)$$

Metoda regula falsi

- definujeme

$$l_{k+1} := \begin{cases} x_{k+1} \text{ pokud sign } f(x_{k+1}) = \text{sign } f(l_k), \\ l_k \text{ jinak} \end{cases}$$

$$r_{k+1} := \begin{cases} x_{k+1} \text{ pokud sign } f(x_{k+1}) = \text{sign } f(r_k), \\ r_k \text{ jinak} \end{cases}$$

$$x'_{k+1} := \begin{cases} l_{k+1} \text{ pokud sign } f(x_{k+1}) = \text{sign } f(r_{k+1}), \\ r_{k+1} \text{ jinak} \end{cases}$$

- jinými slovy je x'_k poslední napočítaný člen s opačným znaménkem než má x_k

Metoda regula falsi

Example 9

Použijte metodu regula falsi pro nalezení kořene polynomu $x^3 + 4x^2 - 10$ v intervalu $< 1, 2 >$ ($\alpha \approx 1.3652$).

k	l_k	r_k	$f(l_k)$	$f(r_k)$	x_k	x'_k
0	1.0	2.0	-5.0	14.0	1.0	2.0
1	1.26315789474	2.0	-1.60227438397	14.0	1.26315789474	2.0
2	1.26315789474	1.42906437248	-1.60227438397	1.08737088326	1.42906437248	1.26315789474
3	1.36199163414	1.42906437248	-0.0533917829598	1.08737088326	1.36199163414	1.42906437248
4	1.36513087878	1.42906437248	-0.00163697021424	1.08737088326	1.36513087878	1.42906437248

Metoda regula falsi

Theorem 10

Nechť platí

- *existuje otevřené okolí H_α bodu α takové, že $f \in C^{(1)}(H_\alpha)$ tj. na H_α existuje derivace f' a je zde spojitá,*
- $f'(\alpha) \neq 0$.

Pak existuje r takové, že

$$V_\alpha \equiv \{x \in \mathbb{R} \mid |x - \alpha| < r\} \subset H_\alpha.$$

a metoda regula falsi konverguje, pro libovolné $x_0 \in V_\alpha$ s rychlostí prvního řádu.

Důkaz.

Video na Youtube

$\square$

Newtonova metoda

- u Newtonovy metody volíme $x_0 = a$
- pro $k = 0, 1, \ldots$ konstruujeme tečny ke grafu funkce v bodě x_k a tam, kde se tečna protne s osou x definujeme nový bod x_{k+1}
- rovnice tečny má tvar

$$y(x) = f(x_k) + f'(x_k)(x - x_k)$$

- pro $y(x) = 0$ dostáváme

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

Newtonova metoda

- tedy platí

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

- zobrazení φ má tvar

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

Newtonova metoda

Example 11

Použijte Newtonovu metodu pro nalezení kořene polynomu $x^3 + 4x^2 - 10$ v intervalu $< 1, 2 >$ ($\alpha \approx 1.3652$).

k	x_k	$f(x_k)$	$f'(x_k)$
0	1.0	-5.0	11.0
1	1.45454545455	1.54019534193	19.4381678093
2	1.37530983126	0.16727562303	16.676910046
3	1.36527945857	0.000816527384453	16.5141996685
4	1.36523001461	$1.97484268938 \cdot 10^{-8}$	16.5133990953

Newtonova metoda

Theorem 12

Nechť platí

- *existuje otevřené okolí* H_α *bodu* α *takové, že* $f \in C^{(1)}(H_\alpha)$ *tj. na* H_α *existuje derivace* f' *a je zde spojitá,*
- $f'(\alpha) \neq 0$.

Pak existuje r *takové, že*

$$V_\alpha \equiv \{x \in \mathbb{R} \mid |x - \alpha| < r\} \subset H_\alpha.$$

a Newtonova metoda konverguje, pro libovolné $x_0 \in V_\alpha$ *s rychlostí prvního řádu.*

Důkaz.

Video na Youtube

$\square$

Newtonova metoda

Theorem 13

Nechť platí

- *existuje otevřené okolí H_α bodu α takové, že $f \in C^{(2)}(H_\alpha)$ tj. na H_α existuje první a druhá derivace f' a f'' a jsou zde spojité,*
- $f'(\alpha) \neq 0$.

Pak existuje r takové, že

$$V_\alpha \equiv \{x \in \mathbb{R} \mid |x - \alpha| < r\} \subset H_\alpha.$$

a Newtonova metoda konverguje, pro libovolné $x_0 \in V_\alpha$ s rychlostí druhého řádu.

Důkaz.

Video na Youtube

$\square$

Newtonova metoda

Remark 14

Newtonova metoda tedy za vhodných podmínek konverguje rychleji, než metoda Regula falsi nebo metoda bisekcí. Vyžaduje ale přesnější počáteční odhad řešení x_0*, tj.* V_α *může být výrazně menší než u metody Regula falsi.*

Newtonova metoda

Example 15

Bude Newtonova metoda konvergovat v případě těchto dvou úloh?

1. $x^3 = 0$
2. $x^2 = 0$

Čebyševova metoda

Remark 16

Pomocí Čebyševovy metody lze odvodit metody ještě vyšších řádů. Ty ale v praxi nejsou příliš užitečné, neboť vyžadují ještě přesnější odhad kořene rovnice a navíc se v něm vyskytují vyšší derivace funkce f. Kvadratická konvergence Newtonovy metody je v praxi naprosto postačující.

Globálně konvergující metody

Definition 17

Pod pojmem *globálně konvergující metody* pro řešení nelineárních rovnic budeme rozumět metody, které pro libovolnou volbu počátečního odhadu kořene x_0 buď konvergují nebo algoritmicky selžou.

Remark 18

Metody, které jsme si ukázaly, často nekonvergují z jednoho důvodu. Pohybují se sice ve správném směru, ale dělají příliš veliký skok. Tím pádem kořen α přeskočí a získají nový odhad, který je horší, než ten předchozí. Řešením je, zamyslet se lépe nad velikostí kroku mezi x_k a x_{k+1}.

Globálně konvergující metody

Line search algoritmus

Algorithm 1 Line search

$r_0 = |f(x)|$
while $r_0 > \epsilon$ **do**
  if $f'(x) = 0$ **then**
    return EXIT_FAILURE;
  end if
  $d = -f(x)/f'(x)$
  $x_t = x + d$
  while $|f(x_t)| \geq |f(x)|$ **do**
    $d = d/2$
    $x_t = x + d$
  end while
  $x = x_t$
  $r_0 = |f(x)|$
end while

Metody pro řešení soustav nelineárních rovnic

Iterativní metody pro řešení soustav nelineárních algebraických a transcendentních rovnic

Remark 19

Mějme reálné funkce reálných proměnných $f_1, \dots, f_n : \mathbb{R}^n \to \mathbb{R}$. Hledáme řešení soustavy rovnic:

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0, \\ f_2(x_1, \dots, x_n) &= 0, \\ \vdots \qquad & \quad \vdots \\ f_n(x_1, \dots, x_n) &= 0. \end{aligned}$$

Řešení budeme značit $\vec{a} = (a_1, \dots, a_n)^T$. Jeho separace je v tomto případě často ještě obtížnější. Přesto pro naše úvahy budeme předpokládat existenci oblasti $H \subset \mathbb{R}^n$ (=otevřená souvislá množina) takové, že $\vec{a} \in H$.

Metody pro řešení soustav nelineárních rovnic

Remark 20

Provedeme formální zobecnění Newtonovy metody:

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - \left[\mathbb{J}_{\vec{f}}\left(\vec{x}^{(k)}\right)\right]^{-1} \vec{f}\left(\vec{x}^{(k)}\right),$$

kde $\mathbb{J}_{\vec{f}}\left(\vec{x}^{(k)}\right)$ *je Jacobiho matice tj.*

$$\left(\mathbb{J}_{\vec{f}}\left(\vec{x}^{(k)}\right)\right)_{ij} = \frac{\partial f_i\left(\vec{x}^{(k)}\right)}{\partial x_j}.$$

resp.

$$\mathbb{J}_{\vec{f}}\left(\vec{x}^{(k)}\right) = \begin{pmatrix} \frac{\partial f_1\left(\vec{x}^{(k)}\right)}{\partial x_1} & \dots & \frac{\partial f_1\left(\vec{x}^{(k)}\right)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n\left(\vec{x}^{(k)}\right)}{\partial x_1} & \dots & \frac{\partial f_n\left(\vec{x}^{(k)}\right)}{\partial x_n} \end{pmatrix}$$

Metody pro řešení soustav nelineárních rovnic

Remark 21

Pro praktické počítání je výhodnější následující tvar

$$\mathbb{J}_{\vec{f}}\left(\vec{x}^{(k)}\right)\left(\vec{x}^{(k+1)}-\vec{x}^{(k)}\right)=-\vec{f}\left(\vec{x}^{(k)}\right).$$

Označíme-li $\Delta\vec{x}^{(k)}=\vec{x}^{(k+1)}-\vec{x}^{(k)}$, *lze přepsat metodu do dvou kroků:*

Algorithm 2 Newtonova metoda pro systémy

1: Vyřeš lineární systém: $\mathbb{J}_{\vec{f}}\left(\vec{x}^{(k)}\right)\Delta\vec{x}^{(k)}=-\vec{f}\left(\vec{x}^{(k)}\right)$
2: Iteruj: $\vec{x}^{(k+1)}=\vec{x}^{(k)}+\Delta\vec{x}^{(k)}$

Metody pro řešení soustav nelineárních rovnic

Remark 22

Výhodou je, že jsme výpočet inverzní matice převedli na řešení lineárního systému. K tomu lze použít i některou z iteračních metod. Jelikož systém v prvním kroku nemusí být nutně řešený s maximální přesností, stačí často udělat jen pár iterací a výpočet metody se tím značně urychlí oproti výpočtu inverze matice pomocí Gaussovy eliminace.

Tomáš Oberhuber

Metody pro řešení soustav nelineárních rovnic

Theorem 23

Buď funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^{(1)}(H)$, kde H je konvexní oblast. Pak pro každé $\vec{u}, \vec{v} \in H$ existuje $\vec{\xi} \in H$ takové, že

$$f(\vec{u}) - f(\vec{v}) = \nabla f(\vec{\xi})(\vec{u} - \vec{v}).$$

Důkaz.

Video na Youtube □

Metody pro řešení soustav nelineárních rovnic

Theorem 24

Nechť platí

- *existuje otevřené okolí $H_{\vec{a}}$ bodu $\vec{a}$ takové, že $\vec{f} \in C^{(1)}\left(H_{\vec{a}}\right)$ tj. na $H_{\vec{a}}$ existují všechny parciální derivace prvního řádu a jsou zde spojité,*
- *$\mathbb{J}_{\vec{f}}\left(\vec{a}\right)$ je regulární.*

Pak existuje r takové, že

$$V_{\vec{a}} \equiv \left\{\vec{x} \in \mathbb{R}^n \mid \left|\vec{x} - \vec{a}\right| < r\right\} \subset H_{\vec{a}},$$

a Newtonova metoda konverguje, pro libovolné $\vec{x}^{(0)} \in V_{\vec{a}}$ s rychlostí prvního řádu.

Důkaz.

Video na Youtube

$\square$

Metody pro řešení soustav nelineárních rovnic

Theorem 25

Nechť platí

- *existuje otevřené okolí $H_{\vec{a}}$ bodu $\vec{a}$ takové, že $\vec{f} \in C^{(2)}\left(H_{\vec{a}}\right)$ tj. na $H_{\vec{a}}$ existují všechny parciální derivace prvního a druhého řádu a jsou zde spojité,*
- *$\mathbb{J}_{\vec{f}}\left(\vec{a}\right)$ je regulární.*

Pak existuje r takové, že

$$V_{\vec{a}} \equiv \left\{ \vec{x} \in \mathbb{R}^n \mid \left| \vec{x} - \vec{a} \right| < r \right\} \subset H_{\vec{a}},$$

a Newtonova metoda konverguje, pro libovolné $\vec{x}^{(0)} \in V_{\vec{a}}$ s rychlostí druhého řádu.

Metody pro řešení soustav nelineárních rovnic

Remark 26

I na Newtonovu metodu pro systémy nelineárních rovnic lze aplikovat line search algoritmus. Získáme tak globálně konvergující metodu.

Metody pro řešení soustav nelineárních rovnic

Example 27

Pomocí Newtonovy metody řešte soustavu rovnic:

$$\begin{aligned} f_1(x,y) = x^3 + 4y^2 + x - y - 4 &= 0, \\ f_2(x,y) = x^4 + 2y^3 - x + y + 2 &= 0, \end{aligned}$$

s počátečním odhadem $\vec{x}^{(0)} = (-1, 1)^T$.

Řešení: (napočítáme jen jednu iteraci)

$$\vec{f}(\vec{x}^{(0)}) = (-3, 7)^T,$$

$$\mathbb{J}_{\vec{f}} = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2+1 & 8y-1 \\ 4x^3-1 & 6y+1 \end{pmatrix},$$

$$\mathbb{J}_{\vec{f}}(\vec{x}^{(0)}) = \begin{pmatrix} 4 & 7 \\ -5 & 7 \end{pmatrix},$$

Řešíme tedy úlohu $\mathbb{J}_{\vec{f}}(\vec{x}^{(0)})\Delta\vec{x}^{(0)} = -\vec{f}(\vec{x}^{(0)})$, tj.:

$$\begin{pmatrix} 4 & 7 \\ -5 & 7 \end{pmatrix} \begin{pmatrix} \Delta x^{(0)} \\ \Delta y^{(0)} \end{pmatrix} = \begin{pmatrix} 3 \\ -7 \end{pmatrix},$$

Dostáváme $\Delta\vec{x}^{(0)} = \frac{1}{18}(-2, 17)^T$ a tedy
$\vec{x}^{(1)} = \vec{x}^{(0)} + \Delta\vec{x}^{(0)} = \frac{5}{18}(-4, 7)^T$. Potom je

$$\vec{f}(\vec{x}^{(1)}) = (11.66, 19.06)^T.$$

Metody pro řešení soustav nelineárních rovnic

Shrnutí a otázky:

- metoda bisekcí
- metoda regula falsi
- Newtonova metoda
- jak se při praktických výpočtech liší podmínky konvergence metody regula falsi a Newtonovy metody?
- **jaké jsou výhody a nevýhody metod vyšších řádů?**
- globálně konvergující metody
- Newtonova metoda pro systémy nelineárních rovnic
  - **jak se dá efektivně vyhnout napočítávání inverzní matice?**