PA081: Programování numerických výpočtů

5. Systémy nelineárních rovnic

Aleš Křenek

jaro 2012

Řešený problém

- formulace problému

$$F_i(x_1, \dots, x_N) = 0 \qquad \text{pro } i = 1, \dots N$$

Řešený problém

- formulace problému

$$F_i(x_1, \ldots, x_N) = 0 \qquad \text{pro } i = 1, \ldots N$$

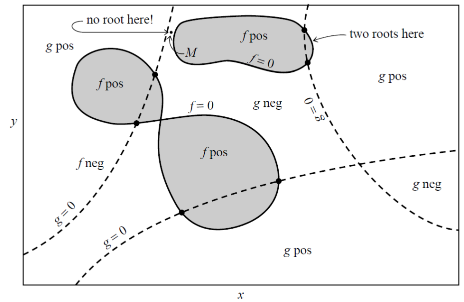
- není to jednoduché, neexistuje univerzální řešení

Řešený problém

- formulace problému

$$F_i(x_1, \ldots, x_N) = 0 \qquad \text{pro } i = 1, \ldots N$$

- není to jednoduché, neexistuje univerzální řešení
- pro dvě dimenze, funkce $f(x, y), g(x, y)$



Neřešitelný problém

- funkce nemají obecně nic společného
- řešení celého systému se objevují v podstatě náhodně
- v obecném případě hledáme společné body hyperpovrchů dimenze $N - 1$
  - oblasti, kde je konkrétní F_i nulová
- konkrétní F_i jich může vygenerovat několik
  - ani nevíme kolik jich je
- bez podrobnější znalosti konkrétního problému se neobejdeme

Metoda prosté iterace

- zobecnění metody pro jednu proměnnou
- problém přeformulujeme do podoby

$$
\begin{aligned}
x_1 &= G_1(\mathbf{x}) \\
x_2 &= G_2(\mathbf{x}) \\
&\vdots \\
x_N &= G_N(\mathbf{x})
\end{aligned}
$$

- potřebujeme nějakou metriku na $\mathbb{R}^N$
- je-li $\mathbf{G}(\mathbf{x})$ kontrakce na nějaké uzavřené omezené podmnožině $\mathbb{R}^N$, pak v ní existuje pevný bod $\mathbf{G}(\xi) = \xi$
- posloupnost $\mathbf{x}_{i+1} = \mathbf{G}(\mathbf{x}_i)$ konverguje k ξ

Metoda prosté iterace

- praktické kritérium konvergence

$$\left|\frac{\partial G_i(\mathbf{x})}{\partial x_j}\right| \leq \frac{q}{N} \qquad \text{pro všechna } i, j = 1, \ldots, N \text{ a } 0 \leq q < 1$$

- resp. slabší

$$\sum_{j=1}^{N}\left|\frac{\partial G_i(\mathbf{x})}{\partial x_j}\right| \leq q < 1 \qquad \text{pro všechna } i = 1, \ldots, N$$

- důkaz uvažuje metriku $\rho(\mathbf{x}, \mathbf{y}) = \max_{1 \leq i \leq N} |x_i - y_i|$
- základem je Taylorův rozvoj $\mathbf{G}(\mathbf{x} + \Delta\mathbf{x})$
  - ve více dimenzích vede právě na sumu parciálních derivací

Metoda prosté iterace – příklad

- řešte systém rovnic

$$(x-1)^2 - y - \frac{1}{2} = 0$$
$$\frac{1}{4}x^2 + y^2 - 1 = 0$$

- v grafu
  - parabola s vrcholem $(1, -\frac{1}{2})$
  - elipsa se středem v 0 a poloosami 2 a 1

Metoda prosté iterace – příklad

- rovnice vyjádříme jako

$$x = G_1(x, y) = \frac{1}{2}(x^2 - y + \frac{1}{2})$$
$$y = G_2(x, y) = \frac{1}{8}(-x^2 - 4y^2 + 8y + 4)$$

- parciální derivace

$$\frac{\partial G_1}{\partial x} = x \qquad \frac{\partial G_1}{\partial y} = \frac{1}{2}$$
$$\frac{\partial G_2}{\partial x} = -\frac{1}{4}x \qquad \frac{\partial G_2}{\partial y} = -y + 1$$

- v okolí $(0, 1)$ jsou podmínky konvergence splněny
- metoda najde kořen cca. $(-0.222, 0.994)$

Metoda prosté iterace – příklad

- druhý kořen je poblíž $(2, 0)$
- metoda diverguje
  - $\partial G_1/\partial x$ a $\partial G_2/\partial y$ porušují podmínky konvergence

Metoda prosté iterace – příklad

- druhý kořen je poblíž $(2, 0)$
- metoda diverguje
  - $\partial G_1/\partial x$ a $\partial G_2/\partial y$ porušují podmínky konvergence
- stačí použít

$$G_1(x, y) = -\frac{1}{4}(x^2 - 6x - y + \frac{1}{2})$$

- metoda už konverguje, i když kritérium není zcela splněno
  - podmínka byla postačující, nikoli nutná
  - při jinak definované metrice už je **G** kontrakce

Newtonova metoda pro systémy rovnic

- Taylorův rozvoj systému

$$F_i(\mathbf{x} + \Delta\mathbf{x}) = F_i(\mathbf{x}) + \sum_{j=1}^{N} \frac{\partial F_i}{\partial x_j} \Delta x_j + O(\Delta\mathbf{x}^2)$$

- v maticovém vyjádření s Jakobiánem $\mathbf{J} = (\partial F_i / \partial x_j)$

$$\mathbf{F}(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{F}(\mathbf{x}) + \mathbf{J}\Delta\mathbf{x} + O(\Delta\mathbf{x}^2)$$

- položíme $\mathbf{F}(\mathbf{x} + \Delta\mathbf{x}) = 0$ a zanedbáním dostáváme

$$\mathbf{J}\Delta\mathbf{x} = -\mathbf{F}(\mathbf{x})$$

- $\Delta\mathbf{x}$ použijeme jako iterační krok
  - získáme řešením systému lineárních rovnic
  - metody probereme v některé z dalších přednášek

Newtonova metoda pro systémy rovnic

- vlastnosti analogické jednodimenzionální verzi
- rychlá (kvadratická) konvergence
  - začne stagnovat při dosažení strojové přesnosti v $\mathbf{x}$ nebo $\mathbf{F}$
  - kritéria kovergence zpravidla

$$\sum |x_i| < \epsilon_x \quad \text{nebo} \quad \sum |F_i| < \epsilon_F$$

  (které nastane dříve)
- špatné globální vlastnosti
  - při nepřesném odhadu může snadno zabloudit
  - nutná kontrola řešení
- potřebné výpočty derivací
  - není tolik jiných metod na výběr
  - v nouzi má smysl i aproximace konečnou diferencí
  - dopředný (recyklace hodnoty funkce) nebo centrální výpočet

Newtonova metoda pro systémy rovnic

- problém $f' \to 0$ „obohacen“ na singularity $\mathbf{J}$
  - mnoho řešení – určitě se netrefíme
  - žádné řešení – ??

Newtonova metoda pro systémy rovnic

- problém $f' \to 0$ „obohacen“ na singularity $\mathbf{J}$
  - mnoho řešení – určitě se netrefíme
  - žádné řešení – ??
- závažnější numerické potíže
  - např. opakované přičtení malé hodnoty by nasměrovalo metodu jinam
  - přetečení a podtečení při násobení
  - maskovány v maticových operacích – špatně odhalitelné

Newtonova metoda pro systémy rovnic

- problém $f' \to 0$ „obohacen“ na singularity $\mathbf{J}$
  - mnoho řešení – určitě se netrefíme
  - žádné řešení – ??
- závažnější numerické potíže
  - např. opakované přičtení malé hodnoty by nasměrovalo metodu jinam
  - přetečení a podtečení při násobení
  - maskovány v maticových operacích – špatně odhalitelné
- správné škálování
  - řádově srovnatelné hodnoty nezávislých proměnných
  - dtto. funkční hodnoty, i vůči proměnným
  - nejlépe v řádu jednotek
  - odpovídající nastavení intervalu pro diference místo derivací
  - když to nejde – řeším ten správný problém?

Převedení na optimalizaci

- optimalizační metody (hledání minima)
  - numericky a algoritmicky příjemnější
  - zdánlivě složitější problém
  - zkusíme řešení rovnic převést na optimalizaci
- definujeme funkci

$$f(\mathbf{x}) = \frac{1}{2} \sum F_i(\mathbf{x})^2$$

- pouze nezáporné hodnoty
- globální minima (0) v kořenech původního systému $\mathbf{F}$

Převedení na optimalizaci

- optimalizační metody (hledání minima)
  - numericky a algoritmicky příjemnější
  - zdánlivě složitější problém
  - zkusíme řešení rovnic převést na optimalizaci
- definujeme funkci

$$f(\mathbf{x}) = \frac{1}{2} \sum F_i(\mathbf{x})^2$$

- pouze nezáporné hodnoty
- globální minima (0) v kořenech původního systému $\mathbf{F}$
- optimalizační metody hledají lokální minimum
  - taková $f(\mathbf{x})$ jich má zpravidla příliš mnoho
  - nutný ještě přesnější odhad řešení
- využití v kombinovaných metodách

Newtonova metoda s hledáním po přímce

- krok Newtonovy metody pro řešení $\mathbf{F}(\mathbf{x}) = 0$
- používá se k minimalizaci funkce $f(\mathbf{x}) = \sum F_i(\mathbf{x})^2$
- v případě potřeby se krok zkrátí
- v reálných případech úspěšnější
  - exaktní specifikace podmínek konvergence je komplikovaná a v praxi téměř neověřitelná
- metoda může uváznout v lokálním minimu f
  - lze snadno detekovat
  - je třeba zkusit jiný počáteční odhad

Newtonova metoda s hledáním po přímce

- Newtonovský krok $\Delta \mathbf{x}$ je směrem dolů vůči f

$$(\nabla f)^T \Delta \mathbf{x} = (\mathbf{F}^T \mathbf{J})(-\mathbf{J}^{-1}\mathbf{F}) = -\mathbf{F}^T \mathbf{F} < 0$$

Newtonova metoda s hledáním po přímce

- Newtonovský krok $\Delta\mathbf{x}$ je směrem dolů vůči f

$$(\nabla f)^T \Delta\mathbf{x} = (\mathbf{F}^T\mathbf{J})(-\mathbf{J}^{-1}\mathbf{F}) = -\mathbf{F}^T\mathbf{F} < 0$$

- směr $\Delta\mathbf{x}$ je správně
  - při dostatečně krátkém kroku hodnota f poklesne
- celý krok může být příliš
  - kvadratická aproximace může být nepřesná

Newtonova metoda s hledáním po přímce

- Newtonovský krok $\Delta\mathbf{x}$ je směrem dolů vůči f

$$(\nabla f)^T \Delta\mathbf{x} = (\mathbf{F}^T\mathbf{J})(-\mathbf{J}^{-1}\mathbf{F}) = -\mathbf{F}^T\mathbf{F} < 0$$

- směr $\Delta\mathbf{x}$ je správně
  - při dostatečně krátkém kroku hodnota f poklesne
- celý krok může být příliš
  - kvadratická aproximace může být nepřesná
- volíme krok $\lambda\Delta\mathbf{x}$ pro vhodné $0 < \lambda \leq 1$
  - zdánlivě nejlepší je minimalizovat $f(\mathbf{x} + \lambda\Delta\mathbf{x})$ vůči λ
  - výpočetně náročné (počet vyhodnocení $\mathbf{F}$)
  - lze ukázat, že to není nezbytné k dosažení rychlé konvergence

Newtonova metoda s hledáním po přímce

- Newtonovský krok $\Delta\mathbf{x}$ je směrem dolů vůči f

$$(\nabla f)^T \Delta\mathbf{x} = (\mathbf{F}^T\mathbf{J})(-\mathbf{J}^{-1}\mathbf{F}) = -\mathbf{F}^T\mathbf{F} < 0$$

- směr $\Delta\mathbf{x}$ je správně
  - při dostatečně krátkém kroku hodnota f poklesne
- celý krok může být příliš
  - kvadratická aproximace může být nepřesná
- volíme krok $\lambda\Delta\mathbf{x}$ pro vhodné $0 < \lambda \leq 1$
  - zdánlivě nejlepší je minimalizovat $f(\mathbf{x} + \lambda\Delta\mathbf{x})$ vůči λ
  - výpočetně náročné (počet vyhodnocení $\mathbf{F}$)
  - lze ukázat, že to není nezbytné k dosažení rychlé konvergence
- prosté kritérium $f(\mathbf{x} + \lambda\Delta\mathbf{x}) < f(\mathbf{x})$ nestačí
  - je třeba $f(\mathbf{x} + \lambda\Delta\mathbf{x}) \leq f(\mathbf{x}) + \alpha\lambda(\nabla f)^T\Delta\mathbf{x}$
  - stačí malé α, např. 10^{-4}

Newtonova metoda s hledáním po přímce

- první pokus – plný krok $\Delta\mathbf{x}$
- ideální by bylo minimalizovat $g(\lambda) = f(\mathbf{x} + \lambda\Delta\mathbf{x})$
- platí $g'(0) = (\nabla f)^T \Delta\mathbf{x}$

Newtonova metoda s hledáním po přímce

- první pokus – plný krok $\Delta\mathbf{x}$
- ideální by bylo minimalizovat $g(\lambda) = f(\mathbf{x} + \lambda\Delta\mathbf{x})$
- platí $g'(0) = (\nabla f)^T \Delta\mathbf{x}$
- aproximujeme $g(\lambda)$ polynomem
- nejprve kvadratický procházející $g(0)$ a $g(1)$

$$a\lambda^2 + g'(0)\lambda + g(0)$$

minimum je v $\lambda_1 = -g'(0)/2a$

Newtonova metoda s hledáním po přímce

- první pokus – plný krok $\Delta\mathbf{x}$
- ideální by bylo minimalizovat $g(\lambda) = f(\mathbf{x} + \lambda\Delta\mathbf{x})$
- platí $g'(0) = (\nabla f)^T \Delta\mathbf{x}$
- aproximujeme $g(\lambda)$ polynomem
- nejprve kvadratický procházející $g(0)$ a $g(1)$

  $$a\lambda^2 + g'(0)\lambda + g(0)$$

  minimum je v $\lambda_1 = -g'(0)/2a$
- další kroky kubický polynom procházející $g(0)$ a dvěma nejčerstvějšími λ_i

  $$a\lambda^3 + b\lambda^2 + g'(0)\lambda + g(0)$$

- opakujeme do dosažení $g(\lambda) \le g(0) + \alpha\lambda g'(0)$

Newtonova metoda – shrnutí

- rychlá konvergence, špatné globální vlastnosti
- vyžaduje derivace
- globální vlastnosti lze vylepšit řízeným zkrácením kroku
- i tak není metoda 100 % spolehlivá
- neobejdeme se bez základní znalosti řešeného problému

Vícerozměrná metoda sečen

- jednorozměrná metoda pracuje s aproximací derivace

$$d_{i+1} = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$$

- analogicky pro více rozměrů

$$\mathbf{B}_{i+1} \Delta \mathbf{x}_i = \Delta \mathbf{F}_i$$

- nedává jednoznačné řešení
- Broydenova metoda
  - rekurentní formule $\mathbf{B}_{i+1}$ na základě $\mathbf{B}_i$
  - minimální změna, která ještě vyhoví kritériu sečny

$$\mathbf{B}_{i+1} = \mathbf{B}_i + \frac{(\Delta \mathbf{F}_i - \mathbf{B}_i \Delta \mathbf{x}_i) \otimes \Delta \mathbf{x}_i}{(\Delta \mathbf{x}_i)^T \Delta \mathbf{x}_i}$$

Vícerozměrná metoda sečen

- neexistuje jednoznačná analogie separace kořenů
  - může zabloudit stejně jako Newtonova metoda
- lze aplikovat stejný princip hledání po přímce
- $\mathbf{B}$ není přesně Jakobián
  - negarantuje, že $\Delta\mathbf{x}$ směřuje dolů vůči $f = \frac{1}{2}\mathbf{F}\mathbf{F}$
  - základní předpoklad hledání po přímce
- v případě selhání náhrada $\mathbf{B}_i$ aproximací $\mathbf{J}$
  - zpravidla dopředné diference

Vícerozměrná metoda sečen

- neexistuje jednoznačná analogie separace kořenů
  - může zabloudit stejně jako Newtonova metoda
- lze aplikovat stejný princip hledání po přímce
- $\mathbf{B}$ není přesně Jakobián
  - negarantuje, že $\Delta\mathbf{x}$ směřuje dolů vůči $f = \frac{1}{2}\mathbf{F}\mathbf{F}$
  - základní předpoklad hledání po přímce
- v případě selhání náhrada $\mathbf{B}_i$ aproximací $\mathbf{J}$
  - zpravidla dopředné diference

- problém singularity $\mathbf{B}$ (resp. $\mathbf{J}$ v Newtonově metodě)
  - nelze určit následující krok
  - řeší se umělou modifikací $\mathbf{B}$ nebo „krokem stranou“
  - de-facto restart z jiného, bližšího bodu

Vícerozměrná metoda sečen

- neexistuje jednoznačná analogie separace kořenů
  - může zabloudit stejně jako Newtonova metoda
- lze aplikovat stejný princip hledání po přímce
- $\mathbf{B}$ není přesně Jakobián
  - negarantuje, že $\Delta\mathbf{x}$ směřuje dolů vůči $f = \frac{1}{2}\mathbf{F}\mathbf{F}$
  - základní předpoklad hledání po přímce
- v případě selhání náhrada $\mathbf{B}_i$ aproximací $\mathbf{J}$
  - zpravidla dopředné diference

- problém singularity $\mathbf{B}$ (resp. $\mathbf{J}$ v Newtonově metodě)
  - nelze určit následující krok
  - řeší se umělou modifikací $\mathbf{B}$ nebo „krokem stranou“
  - de-facto restart z jiného, bližšího bodu

- existují i komplikovanější metody
  - neomezují se jen na hledání po přímce

Shrnutí

- řešení systémů nelineárních rovnic je jeden z nejnepříjemnějších problémů
  - řešení se vyskytují „náhodně“
  - nelze mechanicky separovat kořeny
  - vždy je třeba mít základní představu o charakteristice problému
- metoda prosté iterace
  - dokážeme-li splnit podmínky konvergence
  - rychost konvergence neurčitá
- Newtonova a Broydenova metoda
  - vylepšení globálních vlastností hledáním po po přímce
  - nejsou zcela spolehlivé, mohou uváznout v lokálním minimu pomocné funkce
- detaily a implementace viz W.H.Press, *Numerical Recipes in C*
  - netřeba se učit vzorce zpaměti
  - důležité je znát podstatu problémů a principy metod