NM3 - Iterační metody pro spd soustavy

5.11.'15

1

Soustava lin. rovnic se symetrickou poz. def. (spd) maticí

$\boxed{A\cdot\bar{u}=\bar{b}}$, $A\in\mathbb{R}^{n\times n}$, $\bar{b}\in\mathbb{R}^{n}$

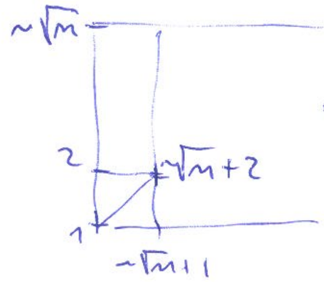
Říkáme, že A je spd řídkou maticí (z řídky), pokud v každém řádku/sloupci počet nenulových prvků je $O(1)$ nezávisle na n.

Matice A a M z 1d, 2d, 3d MKP jsou řídké.

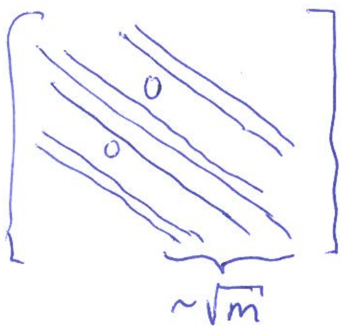
1d stupně volnosti

A, M jsou třídiagonální

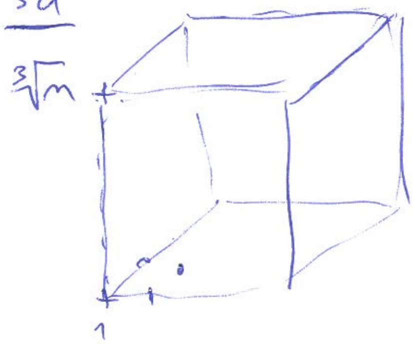
2d



A, M =



3d



A, M =

Gaussova eliminace MKP-matic

Zkušenost

1d : $O(n\cdot 1)$ operací / paměti

2d : $O(n\cdot(\sqrt{n})^2)=O(n^2)$ operací/paměti $\rightsquigarrow n<10^6$ na PC

3d : $O(n\cdot(\sqrt[3]{n^2})^2)=O(n^{7/3})$ operací/paměti $\rightsquigarrow n<10^5$ na PC

Násobení MKP-matice × vektor : $O(n)$ operací/paměti

Iterační metody by neměly mít více než $O(n)$ resp. $O(n^{4/3})$ ideál.

Optimální iterační metody

Multigrid : $k = O(1 \cdot |\log \varepsilon|)$ iterací

DDM : $k = O\left(\left(1 + \log \frac{n}{N}\right)^2 \cdot |\log \varepsilon|\right)$ iterací

kde $\varepsilon = \frac{\|\bar{r}^k\|}{\|\bar{r}^0\|}$, $\boxed{\bar{r}^k := \bar{b} - A\bar{u}^k}$, je rel. přesnost výsledku $\bar{u}^k$ a N je počet paralelních procesů.

Spektrum spd matic

$$\mathbb{A} \cdot \bar{v}_i = \lambda_i \bar{v}_i, \quad \|\bar{v}_i\| = 1$$

přičemž

$$0 < \lambda_1 \le \lambda_2 \le \dots \le \lambda_n \quad \text{... vlastní čísla } \mathbb{A}$$

$$\bar{v}_1 \quad \bar{v}_2 \quad \bar{v}_n \quad \text{... vlastní vektory}$$

tvoří ortonormální bázi $\mathbb{R}^n$

$$\mathbb{A} = \underbrace{[\bar{v}_1, \dots, \bar{v}_n]}_{=:V} \cdot \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_n \end{bmatrix}}_{=:D} \cdot V^T \quad \text{... spektrální rozklad}$$

Richardsonova (Picardova) metoda

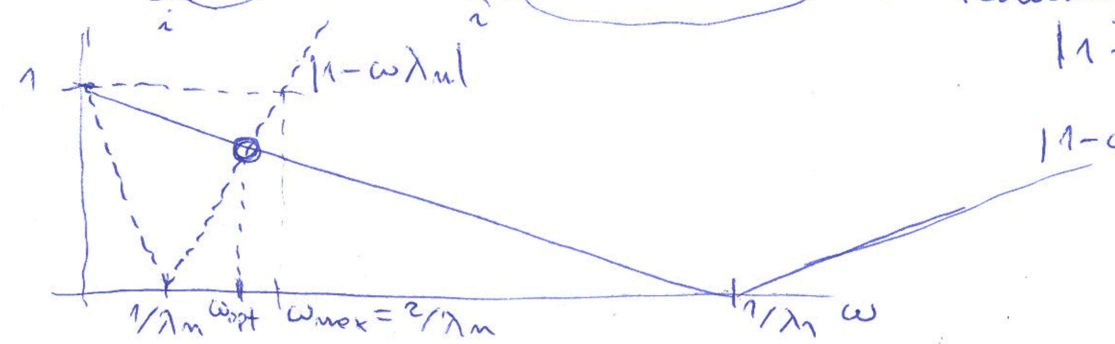
$\omega > 0$ $\qquad \boxed{\bar{u}^{k+1} := \bar{u}^k + \omega \bar{r}^k}$

$$\text{Chyba} \quad \bar{e}^{k+1} := \bar{u} - \bar{u}^{k+1} = \underbrace{(\bar{u} - \bar{u}^k)}_{=:\bar{e}^k} - \omega(\underbrace{\bar{b}}_{A \cdot \bar{u}} - A\bar{u}^k) = \overbrace{(I - \omega A)}^{\text{iterační polynom}} \bar{e}^k$$

$$\bar{e}^{k+1} = \sum_{i=1}^{n} \alpha_i^{k+1} \bar{v}_i, \qquad \bar{e}^k = \sum_{i=1}^{n} \alpha_i^k \bar{v}_i$$

$$\sum_i \alpha_i^{k+1} \bar{v}_i = \sum_i (1 - \omega\lambda_i)\alpha_i^k \bar{v}_i$$

i-tá složka chyby se redukuje faktorem $|1 - \omega\lambda_i|$



NM3 – Iterační metody pro spd soustavy

Richardsonova metoda (pokrač.)

Konvergence je zaručena pro $\omega \in \left(0, \frac{2}{\lambda_n}\right)$

Optimální konvergence: $\omega_{opt}\lambda_n - 1 = 1 - \omega_{opt}\lambda_1$

$$\boxed{\omega_{opt} = \frac{2}{\lambda_1 + \lambda_n}}$$

$$\omega_{opt} = \arg\min_{\omega} \max_{j} |1 - \omega\lambda_j|$$

konvergenční faktor $\boxed{\rho_{opt} = \omega_{opt}\lambda_n - 1 = \frac{\kappa(A)-1}{\kappa(A)+1}}$,

kde $\kappa(A) = \frac{\lambda_n}{\lambda_1}$ je číslo podmíněnosti matice A

Dosažení relativní přesnosti ε lze zajistit v k-iteracích:

$$\frac{\|\bar{e}^k\|}{\|\bar{e}^0\|} \leq \underbrace{(\rho_{\cancel{opt}})^k \overset{!}{\leq} \varepsilon}, \quad 0<\rho<1, \quad 0<\varepsilon<1$$

$$\boxed{k \geq \frac{\log \varepsilon}{\log \rho}}$$

Pozn. Jelikož $\kappa(\mathbb{M})$ nezávisí na n (na h), Richardsonova metoda je optimální pro soustavy $\mathbb{M}\cdot\bar{u} = \bar{b}$, ne však pro soustavy $\mathbb{A}\cdot\bar{u} = \bar{b}$, neboť $\kappa(A) = O(h^{-2})$.

Čebyševova semi-iterační metoda

$\omega^k > 0$ $\quad \boxed{\bar{u}^{k+1} := \bar{u}^k + \omega^k \bar{r}^k}$, $i = 1, \dots, k$

Chyba: $\bar{e}^k = (I - \omega^k A)(I - \omega^{k-1}A)\cdots(I - \omega^1 A)\,\bar{e}^0$

j-tá složka chyby se redukuje faktorem $\left|\prod_{i=1}^{k}(1 - \omega^i\lambda_j)\right|$

4

Hledáme $\omega^1_{opt}, \dots, \omega^k_{opt}$ tak, že

$$\bar{\omega}_{opt} = \arg\min_{\bar{\omega}} \underbrace{\max_j \left| \prod_{i=1}^{k} (1-\omega^i \lambda_j) \right|}$$

polynom $p_k(\lambda)$: $p_k(0) = 1$

$$\leq \max_{\lambda \in \langle \lambda_1, \lambda_n \rangle} \left| \prod_{i=1}^{k} (1-\omega^i \lambda) \right|$$

(jednoznačným)

Řešením min-max úlohy na intervalu je Čebyševův polynom

$$\frac{\chi_k(\lambda)}{\chi_k(0)}, \quad \text{kde} \quad \chi_k(\lambda) := \begin{cases} \cos\left(k \cdot \operatorname{acos} \overbrace{\left(\frac{2\lambda - \lambda_n - \lambda_1}{\lambda_n - \lambda_1}\right)}^{=:t}\right), & |t| \leq 1 \\ \cosh(k \cdot \operatorname{acosh} t), & t \geq 1 \\ (-1)^k \cosh(k \cdot \operatorname{acosh}(-t)), & t \leq 1 \end{cases}$$

Váhy $\bar{\omega}_{opt}$ jsou ~~kořeny~~ převrácenými hodnotami kořenů $\chi_k(\lambda)$:

$$\omega^i_{opt} := \left(\frac{\lambda_n - \lambda_1}{2} \cdot x^i + \frac{\lambda_1 + \lambda_n}{2} \right)^{-1},$$

$$\text{kde} \quad x_i = \cos \frac{\pi(k - 1/2)}{n}$$

(λ_1, λ_n lze nahradit odhady $\lambda \leq \lambda_1$, $\Lambda \geq \lambda_n$)

Nevýhoda

Předem musíme určit počet iterací, proto „semi-iter."

Konvergenční faktor lze odhadnout takto:

$$\boxed{(\rho_{opt})^k \leq \max_{\lambda \in \langle \lambda_1, \lambda_n \rangle} \left| \frac{\chi_k(\lambda)}{\chi_k(0)} \right|} \leq \frac{1}{\left| \cosh\left(k \cdot \operatorname{acosh}\left(\underbrace{-\frac{\lambda_1 + \lambda_n}{\lambda_n - \lambda_1}}_{-\frac{K+1}{K-1} = t:\ |t| > 1}\right)\right) \right|}$$

$$= \frac{2}{e^{k \operatorname{acosh}(t)} + e^{-k \operatorname{acosh}(t)}} \leq 2 \cdot e^{-k \operatorname{acosh}(t)}$$

$$= 2 \cdot e^{-k \ln\left(t + \sqrt{t^2 - 1}\right)} = 2 \cdot \left(\frac{1}{\frac{K+1}{K-1} + \sqrt{\left(\frac{K+1}{K-1}\right)^2 - 1}} \right)^k = \boxed{2 \cdot \left(\frac{\sqrt{K(A)} - 1}{\sqrt{K(A)} + 1} \right)^k}$$

To je nejlepší známá (asymptotická) konvergence iteračních metod.

NM3 – Iterační metody pro spd soustavy

5. 11. '15

Metoda největšího spádu

$A\bar{u} = \bar{b}$, A spd $\iff \bar{u} = \arg\min_{\bar{v}} \varphi(\bar{v})$,

kde $\varphi(\bar{v}) := \frac{1}{2}\bar{v}\cdot A\cdot\bar{v} - \bar{b}\cdot\bar{v}$

$\nabla\varphi(\bar{u}^k) = A\bar{u}^k - \bar{b} = -\bar{r}^k$

tj. $\bar{r}^k$ je směr nejv. spádu φ v $\bar{u}^k$

$$\boxed{\bar{u}^{k+1} := \bar{u}^k + \omega^k \bar{r}^k}$$

kde $\boxed{\omega^k := \arg\min_{\omega} \varphi(\bar{u}^k + \omega\bar{r}^k)} = \boxed{\frac{\bar{r}^k\cdot\bar{r}^k}{\bar{r}^k\cdot A\cdot\bar{r}^k}}$

Konvergence může být pomalá, zvolme lepší směry.

Metoda sdružených gradientů (CG)

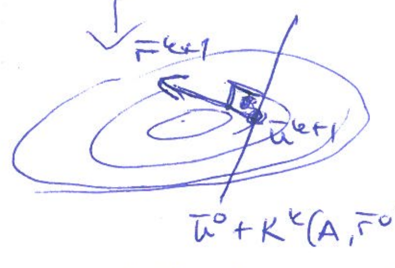
Krylovovský prostor $K^k(A,\bar{r}^0) := \langle \bar{r}^0, A\bar{r}^0, \dots, A^{k-1}\bar{r}^0 \rangle$

$$\boxed{\bar{u}^{k+1} := \arg\min_{\bar{z}\in\bar{u}^0 + K^k(A,\bar{r}^0)} \|\bar{u}-\bar{z}\|_A = \arg\min_{\bar{z}\in\bar{u}^0 + K^k(A,\bar{r}^0)} \varphi(\bar{z})}$$

Pozorování

- $\bar{r}^{k+1} = -\nabla\varphi(\bar{u}^{k+1}) \perp K^k(A,\bar{r}^0)$

Důkaz.



- Důsledek: $\bar{r}^{k+1} \perp_A K^{k-1}(A,\bar{r}^0)$
- Důsledek: Je-li $(\bar{p}_0, \dots, \bar{p}_k)$ A-ortogonální báze $K^k(A,\bar{r}^0)$,

pak $\cancel{\bar{r}^{k+1}}$ $\bar{p}^{k+1} := \bar{r}^{k+1} + \beta^k\bar{p}^k \perp_A K^{k-1}(A,\bar{r}^0)$.

Stačí tedy A-ortogonalizovat $\bar{p}^{k+1} \perp_A \bar{p}^k \rightsquigarrow \boxed{\beta^k = -\frac{\bar{r}^{k+1}A\bar{p}^k}{\bar{p}^k A\bar{p}^k}}$

Konvergence: $\|\bar{e}^k\|_A = \min_{p_k(\lambda)\in P^k:\, p_k(0)=1} \|p_k(A)\bar{e}^0\|_A \overset{\text{lib. inf.}}{\leq} 2\cdot\left(\frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1}\right)^k \cdot \|\bar{e}^0\|_A$

Předpodmínění

je nenáročná ekvivalentní úprava tak, aby výsledná ~~soustava~~ matice soustavy měla ~~spektrum blíže k {1}~~ menší $K(A)$.

Mějme spd matici $B =: \hat{A}^{-1} \approx A^{-1}$, spektr. rozkl. $B = V_B D_B V_B^T$

Místo $\underset{B^{1/2}\bar{v}}{\overset{\parallel}{A \cdot \bar{u}}} = \bar{b}$ řešíme $\boxed{\underbrace{B^{1/2} A B^{1/2}}_{\text{spd}} \bar{v} = B^{1/2} \bar{b}.}$

kde $B^{1/2} = V_B \sqrt{D_B} \cdot V_B^T$

nebo $B \cdot A \bar{u} = B \bar{b}$

Poznámka. $\sigma(BA) = \sigma(B^{1/2} A B^{1/2})$

Důkaz. ~~$BA\bar{v}_B \Rightarrow$~~ $BA\bar{u} = \lambda \bar{u}$, kde $\bar{u} = B^{1/2}\bar{v}$ $\iff \underbrace{B^{1/2} B^{1/2}}_{B} A B^{1/2} \bar{v} = \lambda B^{1/2} \bar{v}$

$\Updownarrow$

$B^{1/2} A B^{1/2} \bar{v} = \lambda \bar{v}$ □

Lemma. $\exists \lambda, \Lambda > 0 \ \forall \bar{w} \in \mathbb{R}^n: \boxed{\lambda \bar{w} \hat{A} \bar{w} \leq \bar{w} A \bar{w} \leq \Lambda \bar{w} \hat{A} \bar{w}}$

$\Updownarrow$

$\sigma(BA) \subset \langle \lambda, \Lambda \rangle$

Idea důkazu.

$\forall i: \ A \bar{u}_i = \lambda_i \hat{A} \bar{u}_i$

$\Updownarrow \checkmark \quad \bar{w} = \sum_{i=1}^{m} w_i \bar{u}_i$

$\forall \bar{w}: \ A \cdot \sum_i w_i \bar{u}_i = \hat{A} \sum \lambda_i w_i \bar{u}_i$

$\Downarrow \checkmark \qquad \Uparrow$ cvičení

$\forall \bar{w}: \ \lambda_1 \bar{w} \hat{A} \bar{w} \leq \bar{w} A \bar{w} \leq \lambda_n \bar{w} \hat{A} \bar{w}$

Celý Důkaz. „$\Updownarrow$" $(\bar{u}, \bar{v})_{B^{-1}} := \bar{u} B^{-1} \bar{v}$

$\lambda \leq \lambda_{min}(BA) = \min_{\bar{w}} \frac{(BA\bar{w}, \bar{w})_{B^{-1}}}{(\bar{w}, \bar{w})_{B^{-1}}} \leq \frac{(A\bar{w}, \bar{w})}{(\hat{A}\bar{w}, \bar{w})}$

$\Lambda \geq \lambda_{max}(BA) = \max_{\bar{w}} \frac{(BA\bar{w}, \bar{w})_{B^{-1}}}{(\bar{w}, \bar{w})_{B^{-1}}} \geq \frac{(A\bar{w}, \bar{w})}{(\hat{A}\bar{w}, \bar{w})}$ □

V implementaci stačí umět aplikovat B, $B^{1/2}$ nepotřebujeme.

Např. v Richardsonově metodě:

$$\bar{u}^{k+1} = B^{1/2}\left[\bar{v}^k + \omega\left(B^{1/2}\bar{b} - B^{1/2} A \underbrace{B^{1/2}\bar{v}^k}_{=\bar{u}^k}\right)\right] = \underbrace{B^{1/2}\left[\bar{v}^k\right.}_{\bar{u}^k} + \omega B^{1/2} \underbrace{(\bar{b} - A\bar{u}^k)}_{\bar{r}^k}$$

tj. $\boxed{\bar{u}^{k+1} = \bar{u}^k + \omega B \bar{r}^k}$

Podobně v PCG.