Přednáška 20 – $A\vec{x}=\vec{b}$ & iterační metody

Lepší/stabilnější algoritmy?

⇝ nejenom, že bychom rádi stabilnější algoritmy –
→ my bychom rádi měli <u>iterační algoritmy</u>,
tj. algoritmy produkující sekvenci aproximací $\vec{x}_1, \vec{x}_2, \ldots$
přesného řešení $\vec{x}$, např. abychom mohli kontrolova
požadovanou přesnost řešení a/nebo její trade-off s
již uběhlým časem výpočtu
⇝ je dobré mít „oba nástroje" – iterační i přímé

Zatím jsme viděli v principu 1 způsob jak
„vytvořit" iterační metodu → $\boxed{\text{reformulace na hledání pevného bodu + Banachova věta o kontrakci}}$

⇝ idea:

<u>Krok 1:</u> „rozštěpíme" matici A, například

$$\text{platí } \underbrace{A = D + N}_{\text{tzv. štěpení matice (matrix splitting)}} \quad \text{kde} \quad \begin{aligned} D &= \operatorname{diag}(A) \\ N &= A - \operatorname{diag}(A) = \text{off-diag}(A) \end{aligned}$$

<u>Krok 2:</u> $A\vec{x}=\vec{b} \iff D\vec{x} = \vec{b} - N\vec{x} \iff$

$\iff \vec{x} = D^{-1}(\vec{b} - N\vec{x})$ .... $\vec{x}$ je pevný bod funkce

$$F: \vec{v} \longmapsto D^{-1}(\vec{b} - N\vec{v})$$

Tedy: $\vec{x}_{k+1} = D^{-1}(\vec{b} - N\vec{x}_k) = D^{-1}(\vec{b} - A\vec{x}_k + D\vec{x}_k) = \vec{x}_k + D^{-1}(\vec{b} - A\vec{x}_k)$

Jak / Kdy konverguje?

$\to$ vezmeme si $\vec{x}^*$ jako přesné řešení, tj. $A\vec{x}^* = \vec{b}$ (to je rovnost, nikoliv rovnice)

$$\vec{x}_{k+1} - \vec{x}^* = \vec{x}_k - \vec{x}^* + D^{-1}(\overset{=A\vec{x}^*}{\vec{b}} - A\vec{x}_k) = (I - D^{-1}A)(\vec{x}_k - \vec{x}^*) =$$

$$= (I - D^{-1}A)^2(\vec{x}_{k-1} - \vec{x}^*) = \dots = (I - D^{-1}A)^k(\vec{x}_0 - \vec{x}^*)$$ .... stejný postup jako u Banachovy věty o kontrakci

$$\Rightarrow \left[\text{konverguje } \forall \vec{x}_0 \in \mathbb{R}^n \iff (I - D^{-1}A)^k \to 0 \quad \text{pro } k \to +\infty\right]$$

Mějme $T = I - D^{-1}A$ diagonalizovatelnou (následující výpočet projde i pro Jordanův tvar, jen to je krapet složitější vidět)

$\to T = S \Lambda S^{-1}$ ... S = matice vl. vekt. $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ matice vl. čísel

pak $T^k = S\Lambda S^{-1} S \Lambda S^{-1} \dots S \Lambda S^{-1} = S \Lambda^k S^{-1}$

a tedy $T^k \xrightarrow{k \to +\infty} 0 \iff \underbrace{\max_{i=1,\dots,n} |\lambda_i|}_{=: \rho(T)} < 1$ ... tzv. spektrální poloměr A

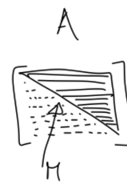
$$\Rightarrow \left[\text{konverguje } \forall \vec{x}_0 \in \mathbb{R}^n \iff \rho(I - D^{-1}A) < 1\right]$$

Obecné štěpící metody: $A = M - \underbrace{(M - A)}_{=: N} \quad \forall M$

$\Rightarrow$ pro libovolnou regulární $M \in \mathbb{R}^{n \times n}$ lze napsat iterační metodu $\to$ $\vec{x}_{k+1} = \vec{x}_k + M^{-1}\underbrace{(\vec{b} - A\vec{x}_k)}_{=: \vec{r}_k \text{ ... tzv. reziduum}}$

Richardson method: $M = I$

Jacobi method: $M = \mathrm{diag}(A)$

Gauss-Seidel method: $M = \text{upper-triangle}(A) =$ 

Pozorování pro Richardson method:

$$\vec{x}_{k+1} = \vec{x}_k + \vec{b} - A\vec{x}_k \iff \vec{x}^* - \vec{x}_{k+1} = \vec{x}^* - \vec{x}_k - A(\vec{x}^* - \vec{x}_k) \implies$$

$$\implies A(\vec{x}^* - \vec{x}_{k+1}) = (I - A)\cdot(A(\vec{x}^* - \vec{x}_k)) \implies \underbrace{\vec{b} - A\vec{x}_{k+1}}_{=:\vec{r}_{k+1}} = (I-A)^k \underbrace{(b - A\vec{x}_0)}_{=:\vec{r}_0}$$

tzv. reziduum v iteraci k+1 nebo 0

$(I-A)^k$ lze rozepsat jako $\alpha_0 I + \alpha_1 A + \dots + \alpha_k A^k$ ⇝ to je polynom v A →

$\to (I-A)^k = \alpha_0 I + \alpha_1 A + \dots + \alpha_k A^k = p^{Rich}(A)$ kde $p^{Rich}(t) = \sum_0^k \alpha_i t^i = (1-t)^k$

→ akorát místo skaláru $\underline{t}$ akorát mocním matici A

$$\implies \vec{r}_k = p^{Rich}(A)\cdot \vec{r}_0$$

→ v praxi $\vec{r}_k$ často používáme jako ukazatel konvergence:
pro $\vec{x}_k = \vec{x}^*$ máme $\vec{r}_k = 0$ a tedy $\|\vec{r}_k\| \approx 0$ často **indikuje** $\vec{x}_k \approx \vec{x}^*$

⇒ jedna možnost, jak zlepšit Richardsonovu metodu je najít polynom $p_k^{opt}(t)$ tak aby:

$$\boxed{\|p_k^{opt}(A)\vec{r}_0\| = \min_{\substack{p(t) \text{ je polynom} \\ \text{stupně} \le k}} \|p(A)\vec{r}_0\|}$$

Tady ale pozor! Protože $\vec{r}_k = A(\vec{x}^* - \vec{x}_k)$ tak na základě přednášky 17 víme, že $\|\vec{x}^* - \vec{x}_k\| \le \varkappa(A)\cdot\|\vec{r}_k\| \implies$

⇒ malé reziduum neznamená vždy přesnou aproximaci – záleží na podmíněnosti matice A

→ tzv. generalized minimal rezidual method (GMRES).

Porovnání s Richardsonem:

- není jasné jak z • dostat předpis pro $\vec{x}_k$
- Richardson používá v každé iteraci stejnou formulku

× GMRES každou iteraci musí konstruovat ten polynom "$p^{opt}(t)$" a na jeho základě správné $\vec{x}_k$

- Richardson velmi často nekonverguje
- Lingebra 2: Věta (Caley-Hamilton) $\forall A \in \mathbb{R}^{n\times n}\ \exists$ polynom $q(t)$ stupně n t. že $A^{-1} = q(A)$

⇒ na papíře dostanu $p_n^{opt}(t) = q(t)$ ⇒ na papíře GMRES zkonverguje v $\underline{n}$ krocích

Metody „typu GMRES" (= založené na optimálních polynomech) se nazývají Krylovovské metody & jsou „prakticky nejpoužívanější".