Lineární algebra III – 8. přednáška 2.12.2013

Metody Krylovových podprostorů a metoda konjugovaných gradientů (podle přednášky Z. Strakoše, rozšířeno).

↳ skvělá kniha Liesen, Strakoš: Krylov Subspace Methods

Nechť A je čtvercová regulární, dva fundamentální problémy jsou $Ax=b$ a $Ax=\lambda x$.

Už jsme se zabývali problémem, jak spočítat vlastní číslo.

- charakteristický polynom, ← nepoužitelné
- nějaký iterativní algoritmus.

Například: Mocninná metoda ze 7. přednášky.

Hledáme vlastní vektor v Krylovově podprostoru

$$\mathcal{K}_n(A,v)=\langle v, A\cdot v, A^2 v, \dots, A^{n-1} v\rangle$$

pro nějaký vektor v. Vektory postupně svírají malé úhly, proto chceme nalézt ortogonální bázi.

Arnoldiho algoritmus: Generování $\mathcal{K}_n(A,v)$ s Gram-Schmidtovou ortogonalizací, získáme ortogonální bázi $v_1,\dots,v_d$, když podprostor přestane růst po d krocích.

$v_1 = v/\|v\|$.

Gram-Schmidtova ortogonalizace klasická, alternativně modifikovaná

pro $n=1,2,\dots$

$$\begin{aligned}
\widehat{v_{n+1}} &= Av_n - \sum_{i=1}^{n} h_{i,n} v_i, \quad \text{kde } h_{i,n}=v_i^T A v_n.\\
h_{n+1,n} &= \|\widehat{v_{n+1}}\|. \longrightarrow \text{Zastavit pro } =0.\\
v_{n+1} &= \widehat{v_{n+1}}/h_{n+1,n}.
\end{aligned}$$

Lze elegantně popsat v maticovém zápisu:

horní Hessenbergova matice

$$V_n=\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \quad \text{a} \quad H_n=\begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,n} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,n} \\ & \ddots & \ddots & \vdots \\ 0 & & h_{n,n-1} & h_{n,n} \end{pmatrix}.$$

Potom po n krocích algoritmu platí

$$AV_n = V_n H_n + h_{n+1,n} v_{n+1} e_n^T,$$

$$AV_n = \begin{pmatrix} Av_1, \dots, Av_n \end{pmatrix}$$

Vyjádření vůči vektorům $v_1, \dots, v_n$

a vůči v_{n+1} v posledním sloupci.

Rovnice popisuje vztah mezi vektory $v_1, \dots, v_{n+1}$ a $Av_1, \dots, Av_n$.

a v posledním kroku $AV_d = V_d H_d$.

$\mathcal{K}_d(A,v)$ je invariantní podprostor zobrazení A, neboť platí $A\mathcal{K}_d(A,v) \subseteq \mathcal{K}_d(A,v)$. Potom H_d je matice zobrazení $A \restriction \mathcal{K}_d(A,v)$ vůči ortogonální bázi $V_d = [v_1, \dots, v_d]$:

$$V_d^* \underbrace{(V_d V_d^* A)}_{} V_d = V_d^* A V_d = H_d.$$

Ortogonální projekce na $\mathcal{K}_d(A,v)$ vůči bázi $v_1, \dots, v_d$.

Nyní se omezme na symetrické matice $A = A^*$:

→ v řeči skalárního součinu $\langle Ax | y \rangle = \langle x | Ay \rangle$,
tedy $\forall x, y$: $y^* A x = y^* A^* x \iff A = A^*$.

Existuje spektrální rozklad $A = Q \Lambda Q^*$ na systém ortogonálních 1-dimenzionálních podprostorů, na které můžeme nahlížet zvlášť.

V případě Arnoldiho algoritmu je H_d Hessenbergova symetrická, tedy tridiagonální T_d. Proto v_{n+1} závisí pouze na v_{n-1} a v_n.

krátká rekurence

$$T_d = \begin{pmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_n \\ & & \beta_n & \alpha_n \end{pmatrix}$$

Lanczosův algoritmus pro $A = A^*$.

$$AV_n = V_n T_n + \beta_{n+1} v_{n+1} e_n^T.$$

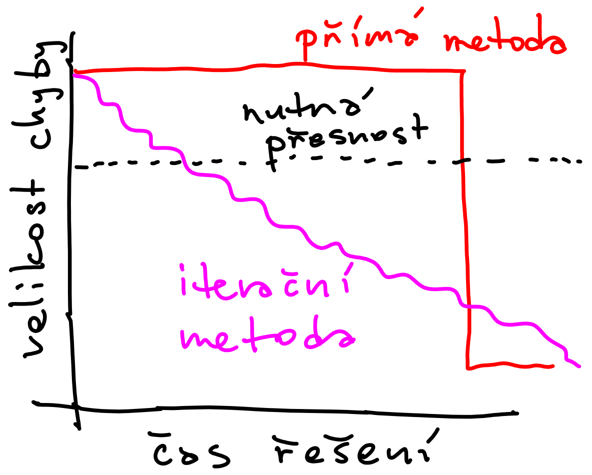
$$Av_n = \beta_n v_{n-1} + \alpha_n v_n + \beta_{n+1} v_{n+1}.$$

Při zaokrouhlování se může ortogonalita ztratit.

Řešení soustavy $Ax=b$

→ přímé metody – třeba Choleského rozklad
$A = LL^*$; $LL^* x = b$ se řeší $Ly = b$, $L^* x = y$.

→ iterační metody – konstruujeme posloupnost
$x_0, x_1, \dots, x_n \to x$, označme chybu $e_n = x - x_n$
a residuum $r_n = b - Ax_n = Ae_n$.



Příklad iterační metody: stacionární/štěpící metody

$A = M - N$, kde M je snadno invertovatelná matice.

$$Ax = b \Rightarrow Mx = Nx + b.$$
$$x = M^{-1}Nx + M^{-1}b.$$

Poslední rovnost iterujeme $x_n = M^{-1}N x_{n-1} + M^{-1}b$.
Pokud metoda konverguje, jde $e_n \to 0$. Konvergence a její rychlost závisí na spektrálním radiu $\rho(M^{-1}N)$.

Bekas, Curioni, Klavík: Metodu lze interpretovat jako aproximaci matice A pomocí M.

Metoda konjugovaných gradientů (CG):

A je symetrická, pozitivně definitní. Jedno z možných odvození:
Chceme $x_n \in x_0 + \mathcal{K}_n(A, r_0)$, tedy $x_n = x_0 + z_n$, $z_n \in \mathcal{K}_n(A, r_0)$.
To volíme tak, aby $r_n \perp \mathcal{K}_n(A, r_0)$. Protože $r_n = A e_n$, ekvivalentně $e_n \perp_A \mathcal{K}_n(A, r_0)$, tedy x_n minimalizuje ze všech vektorů $\mathcal{K}_n(A, r_0)$ normu $\|x - x_n\|_A = \|e_n\|_A$.

V maticovém zápisu podmínky ortogonality:

$$0 = V_n^* r_n = V_n^*(b - Ax_n) = V_n^*(\underbrace{b - Ax_0}_{r_0} - Az_n).$$

Protože $v_1 = r_0/\|r_0\|$, platí $V_n^* r_0 = \|r_0\| \cdot e_1$. Dále z_n leží v $\mathcal{K}_n(A, r_0)$, lze vyjádřit $z_n = V_n \cdot t_n$. Protože $V_n^* A V_n = T_n$,
$0 = \|r_0\| \cdot e_1 - V_n^* A V_n t_n = \|r_0\| \cdot e_1 - T_n t_n$. Pak $x_n = x_0 + V_n t_n$.

> Metoda konjugovaných gradientů je redukce $N \times N$ systému $Ax = b$ na $n \times n$ systém $T_n t_n = \|r_0\| \cdot e_1$.

Metoda má vysoce nelineární chování, protože aproximuje řešení $A^{-1}b$ pomocí polynomu $\langle b, Ab, A^2 b, \dots, A^{n-1}b \rangle$.

Hrubý obecný odhad konvergence:

$$\frac{\|x - x_n\|_A}{\|x - x_0\|_A} \le 2\left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1}\right)^n.$$

Odhad založen na spektru.
Ale 1) ignorujeme strukturu spektra,
2) ignorujeme pravou stranu.

Metoda konstruuje nejlepší možný polynom φ stupně n, že $\varphi(0) = 1$ a minimalizujeme maximum $\varphi(\lambda_i)$ přes vl. čísla A.