Téma 5: Úplný problém vl. čísel

Úloha: $A \in \mathbb{C}^{n\times n}$, najdi aproximaci všech vl. čísel a vl. vektorů A

$\rightarrow$ chci převést A do tvaru, ze kterého budou vl. č. viditelná

Jaké transformace zachovávají spektrum?

Lemma: Nechť $A \in \mathbb{C}^{n\times n}$ a $S \in \mathbb{C}^{n\times n}$ regulární. Nechť $B = S^{-1}AS$, tj. B vznikne podobnostní transformací matice A. Pak A a B mají stejná vl. čísla.

D: $A = V . J_A . V^{-1}$, J_A - Jord. tvar A $\Rightarrow$

$B = (S^{-1}V) . J_A . (S^{-1}V)^{-1} \Rightarrow J_A$ je Jord. tvar matice B

$\square$

Pozn: vl. vektory se transformují S^{-1}

PROBLÉM: $A = \hat{A} + E_A$ ... z praktické úlohy

PŘESNÁ DATA — CHYBY MĚŘENÍ

$\Rightarrow S^{-1}AS = S^{-1}\hat{A}S + S^{-1}E_AS$

$B = \hat{B} + E_B$, kde $E_B = S^{-1}E_AS$

CHYBA PO PODOB. TRANSFORMACI

$\|E_B\| \leq \|S^{-1}\| . \|S\| . \|E_A\| = \underbrace{\varkappa(S)}_{\geq 1} . \|E_A\|$

$\Rightarrow$ je-li $\varkappa(S) \gg 1$, pak transf. může zvětšit chyby v datech důsl.: chci unitární podobnostní transformaci A, která vyjeví spektrum

5.1 Schurova věta

Věta (Schurova): Nechť $A \in \mathbb{C}^{n\times n}$. Pak $\exists\, U \in \mathbb{C}^{n\times n}$ unitární taková, že $R := U^*AU$ je horní trojúhelníková matice. Navíc R má na diagonále vl. čísla A v libovolném předem předepsaném pořadí.

def: Rozklad $A = URU^*$ nazveme Schurův rozklad matice A.

D: indukcí dle n, $n=1$ triv.

$n \to n+1$: $A \in \mathbb{C}^{(n+1)\times(n+1)}$, $\lambda :=$ první vl. č. v požadovaném pořadí $\Rightarrow \exists x: Ax = \lambda x, \|x\| = 1$ (vl. vektor)

vol. $X \in \mathbb{C}^{(n+1)\times n}$: $H := [x, X] \in \mathbb{C}^{(n+1)\times(n+1)}$ unitární

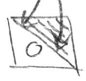
doplním x na unit. matici

$$\Rightarrow H^*AH = \begin{array}{cc} \begin{array}{cc} 1 & n \end{array} & \\ \left[\begin{array}{c|c} x^*Ax & x^*AX \\ \hline X^*Ax & X^*AX \end{array}\right] & \begin{array}{l} \}1 \\ \}n \end{array} \end{array} = \left[\begin{array}{c|c} \lambda & b^* \\ \hline 0 & C \end{array}\right] \begin{array}{l} \}1 \\ \}n \end{array} \Rightarrow$$

$$0 = \lambda X^*x = X^*\lambda x = X^*Ax$$

$$\Rightarrow sp(A) = sp(H^*AH) = \lambda \cup sp(C)$$

na diagonále vlastní čísla C

$$C \in \mathbb{C}^{n\times n} \xRightarrow[\text{předpoklad}]{\text{indukční}} \exists V\text{-unit}: V^*CV = \tilde{R} =$$ 

$$\Rightarrow \text{položíme } U := [x, XV] = \underbrace{H \cdot \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & V \end{array}\right]}_{\text{součin unit. matic}}$$

$$\text{pak } R = U^*AU = \begin{bmatrix} 1 & 0 \\ 0 & V^* \end{bmatrix} \begin{bmatrix} \lambda & b^* \\ 0 & C \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} = \begin{bmatrix} \lambda & b^*V \\ 0 & V^*CV \end{bmatrix} =$$ 

□

Pozn.: Sch. rozklad není vůbec jednoznačný

Př: $A = URU^* = \overbrace{(UD)}^{\tilde{U}}\,\overbrace{D^*RD}^{\tilde{R}}\,\overbrace{(D^*U)}^{\tilde{U}^*}$ $\forall D$ - diagonální a unitární, třeba $D = \begin{bmatrix}1 & \\ & \ddots \\ & & 1\end{bmatrix}$

→ důsledkem Sch. věty je věta o unitární diagonalizovatelnosti (BYLO V LA)

Věta: $A \in \mathbb{C}^{n\times n}$ je normální, tj. $A^*A = AA^*$ $\iff$ $\exists\, U \in \mathbb{C}^{n\times n}$ - unit. taková, že $D = U^*AU$ je diagonální matice. Pak $U = [u_1, .., u_n]$ vl. vektory

$A \in \mathbb{C}^{n\times n}$ obecná: $R = U^*AU =$

$A \in \mathbb{C}^{n\times n}$ normální: $D = U^*AU =$

$A \in \mathbb{C}^{n\times n}$ unitární: $D = U^*AU$, $|d_{ii}| = 1$, $i = 1, .., n$

$A \in \mathbb{C}^{n\times n}$ hermitovská: $D = U^*AU$, $d_{ii} \in \mathbb{R}$, $i = 1, .., n$

$A \in \mathbb{C}^{n\times n}$ diagonalizovatelná:
$\exists\, V \in \mathbb{C}^{n\times n}$ - regul.: $D = V^{-1}AV =$, $V = [v_1, .., v_n]$ vl. vekt.

souvislost: $V = Q_V R_V \Rightarrow R_V D R_V^{-1} = Q_V^* A Q_V$ je Sch. faktor

→ Sch. věta má teoretický zajímavý důsledek

Věta: Množina všech (diagonalizovatelných) matic s navzájem různými vl. čísly je hustá v $\mathbb{C}^{n\times n}$

D: $A \in \mathbb{C}^{n\times n}$ - nediagonaliz. $\Rightarrow \exists$ aspoň jedno vícenásobné vl. č. A

$A = URU^*$ - Sch. rozklad A

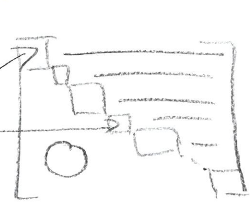
vol. $D_\varepsilon \in \mathbb{C}^{n\times n}$ diagonální tak, aby $\|D_\varepsilon\| < \varepsilon$ a $R_\varepsilon = R + D_\varepsilon$ měla na diagonále navz. různá čísla $\Rightarrow$

$$\Rightarrow \|A - \underbrace{U R_\varepsilon U^*}_{=: A_\varepsilon}\| = \|U D_\varepsilon U^*\| = \|D_\varepsilon\| < \varepsilon$$

$\Rightarrow \forall \varepsilon > 0$ dost malé $\exists A_\varepsilon \in \mathbb{C}^{n\times n}$ diagonaliz. taková, že $\|A - A_\varepsilon\| < \varepsilon$ □

Pozn.: Z věty neplyne, že se při návrhu a analýze metod lze omezit na diagonaliz. matice, neboť perturbace v A_ε mění vlastnosti A!

reálný Sch. rozklad: $A \in \mathbb{R}^{n\times n}$

- je-li A symetrická, pak existuje její reálný Sch. rozklad $D = U^T A U$
- obecně nelze, neboť A může mít komplexní vl. čísla, ale existuje rozklad $T = U^T A U$, $T =$ 

bloky 1x1 a 2x2,

kde 1x1 je pro reálné vl. č. a 2x2 pro kompl. sdružený pár vl. č.

Věta (Abel & Galois) $\Rightarrow$ kořeny polynomu st. $n \geq 5$ nelze počítat kon. algoritmem $\Rightarrow$ vl. čísla nelze počítat kon. algoritmem $\Rightarrow$ Sch. rozkl. lze jen aproximovat

5.2 QR algoritmus

= řeší úplný problém vl. č. pomocí iterační aproximace Sch. rozkladu

KROK 1: PREPROCESING

cíl: přímou unit. podobn. transformací převést A do tvaru blízkého $\begin{bmatrix} \diagdown & \\ 0 & \diagdown \end{bmatrix}$

$$A=\begin{bmatrix} x & \cdots & x\\ x & \cdots & x\\ \vdots & & \vdots\\ x & \cdots & x\end{bmatrix} \xrightarrow[\text{Hous. refl.}]{H_1=\begin{bmatrix}1 & 0\cdots 0\\ 0 & \\ \vdots & \tilde{H}_1 \\ 0 & \end{bmatrix}\in\mathbb{C}^{n\times n}} H_1A=\begin{bmatrix} x & x & \cdots & x\\ x & x & \cdots & x\\ 0 & x & \cdots & x\\ \vdots & & & \vdots \\ 0 & x & \cdots & x\end{bmatrix} \xrightarrow{\cdot H_1^*=\begin{bmatrix}1 & 0\cdots 0\\ 0 & \\ \vdots & \tilde{H}_1^* \\ 0 & \end{bmatrix}} H_1AH_1^*=\begin{bmatrix} x & x & \cdots & x\\ x & x & \cdots & x\\ 0 & x & \cdots & x\\ \vdots & & & \vdots \\ 0 & x & \cdots & x\end{bmatrix} \longrightarrow$$

nezmění se

$$\xrightarrow{H_2=\begin{bmatrix}1 & \\ & 1 \\ & & \tilde{H}_2\end{bmatrix}\in\mathbb{C}^{n\times n}} \cdots \longrightarrow \underbrace{H_{n-1}\cdots H_1}_{=H^*} A \underbrace{H_1^*\cdots H_{n-1}^*}_{=H} = \begin{bmatrix} \diagdown & \diagdown & \diagdown \\ \diagdown & \diagdown & \\ 0 & \diagdown & \diagdown \end{bmatrix} =: A_0$$

máme: $\| \; A_0 = H^*AH$, H - unitární, A_0 - horní Hessen., $sp(A_0)=sp(A)$

vlastnosti: výp. nár. $O(n^3)$
výp. stabilní (OG transformace)

$\Rightarrow$ zbývá se zbavit poddiagonály A_0, což lze pouze iteračně

Pozn: je-li A - hermitovská, pak A_0 vyjde dokonce tridiagonální $\begin{bmatrix} \diagdown & \diagdown & \\ \diagdown & \diagdown & \diagdown \\ & \diagdown & \diagdown \end{bmatrix}$

KROK 2: ITERACE

máme $A_0 = \in \mathbb{C}^{n \times n}$

→ uvažujme posl. matic konstruovaných iteracemi:

$$A_k = Q_k R_k$$ SPOČTU QR ROZKLAD

PŘENÁSOBÍM

$$A_{k+1} = R_k Q_k, \quad k = 0, 1, \dots$$

Lemma: Nechť A je regulární, pak:

1, A_k je horní Hessenb., $k = 0, 1, \dots$

2, A_k jsou si unitárně podobné, přičemž $A_{k+1} = Q_k^* A_k Q_k = P_k^* A_0 P_k$ a $P_k = Q_0 \dots Q_k$ je unitární

3, A_k mají stejné spektrum, $k = 0, 1, \dots$

D(1): A_0 regul. $\Rightarrow A_k, Q_k, R_k$ regul., $k = 0, 1, \dots$

indukcí: $k = 0$ – z konstrukce, $k \mapsto k+1$

$A_k = Q_k R_k \Rightarrow Q_k = A_k R_k^{-1} =$ je horní Hess.

inverze horní △-ové je horní △-ová matice

$A_{k+1} = R_k Q_k =$

D(2): $A_k = Q_k R_k \Rightarrow R_k = Q_k^* A_k$

$$A_{k+1} = R_k Q_k = Q_k^* A_k Q_k = \underbrace{Q_k^* \dots Q_0^*}_{= P_k^*} A_0 \underbrace{Q_0 \dots Q_k}_{= P_k}$$

$$P_k^* P_k = P_k P_k^* = I$$

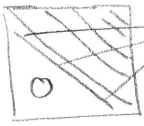
D(3): plyne z 2, □

důsledek: $A_{k+1} = (HP_k)^* \cdot A (HP_k)$, kde

HP_k je unitární, $sp(A_{k+1}) = sp(A)$, $k = 0, 1, \dots$

Věta (*): Nechť $A \in \mathbb{C}^{n \times n}$ a pro její vl. čísla platí, že $|\lambda_1| > |\lambda_2| > \ldots > |\lambda_n| > 0$. Pak $A_k \xrightarrow{k \to \infty} R =$ , $HP_k \xrightarrow{k \to \infty} U$, kde $R = U^* A U$ je Sch. rozklad matice A.

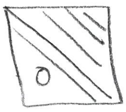
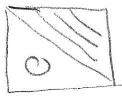
Pozn.: konvergence lze dok., že poddiagonální prvek $|a^{(k)}_{i+1,i}| \xrightarrow{k \to \infty} 0$ rychlostí $\left|\frac{\lambda_{i+1}}{\lambda_i}\right| < 1$, $i = 1, \ldots, n-1$

$A_k =$  $\to |a^{(k)}_{i+1,i}| \xrightarrow{k \to \infty} 0$, ale různě rychle

algoritmus: vstup: A, m = max # iterací
výstup: $A_{m+1} \approx R$ ← ze Sch. rozkladu

$A_0 = H^* A H$ ... preprocesing Hous. reflexemi, $O(n^3)$
for $k = 0, 1, \ldots, m$ do
  $A_k = Q_k R_k$ ... QR rozklad Giv. rotacemi, $O(n^2)$
  $A_{k+1} = R_k Q_k$ ... efektivní násobení, $O(n^2)$
end

výpočet QR rozkladu $A_k =$ :

$A_k =$  $\xrightarrow[\text{rotací}]{(n-1)\text{ Giv.}} \Gamma_k \cdot A_k =$  $= R_k$ ...
↑ složená rotace

$\Rightarrow O(n^2)$ operací

Pozn.: pokud chceme jen vl. čísla A, nemusíme ukládat H a Q_k, stačí přepisovat matici A na A_k

celkové výp. nákl.: obvykle je třeba cca n iterací $\Rightarrow O(n^3)$ s větší konstantou u n^3

pozn: je-li A-herm., pak $A_k \to D = \begin{bmatrix} \ddots & 0 \\ 0 & \ddots \end{bmatrix}$ a náklady jsou o řád menší

pozn: alg. lze vylepšit, aby konvergoval $\forall A \in \mathbb{C}^{n\times n}$ (shifty, rekurze, deflace)

Proč QR alg. konverguje?

- důkaz konv. věty (*) je těžký
- vysvětlíme, že za alg. je mocninná metoda nad celým $\mathbb{C}^n$

Lemma: Ozn. $P_k := Q_0 \cdots Q_k$, $R_k \cdots R_0 =: S_k$. Pak $A_0^{k+1} = P_k S_k$ je QR rozklad matice A_0^{k+1}, tj. P_k je unitární a S_k je horní $\triangle$-ová.

D: P_k - součin unit. matic

S_k - || - horních $\triangle$-ových $\Rightarrow S_k = \begin{bmatrix} \ddots & * \\ 0 & \ddots \end{bmatrix}$

indukcí dle k:

$k=0$: $A_0 = P_0 S_0 = Q_0 R_0$ ✓

$k \mapsto k+1$

$$P_k S_k = P_{k-1} Q_k R_k S_{k-1} = P_{k-1} A_k S_{k-1} \overset{\text{Lemma (o)}}{=}$$

$$= \underbrace{P_{k-1} \cdot (P_{k-1}^*}_{=I} A_0 P_{k-1}) S_{k-1} = A_0 P_{k-1} S_{k-1} =$$

$$\overset{\text{ind. předpoklad}}{=} A_0 \cdot A_0^{k} = A_0^{k+1}$$

□

Víme: $U^* A U = R = \begin{bmatrix} \lambda_1 & * \\ 0 & \ddots \end{bmatrix}$, $U = [u_1, u_2, \dots, u_n]$,

kde $A u_1 = \lambda_1 u_1$

→ místo věty (*) dokážeme slabší ~~tvr~~ tvrzení

Lemma: Nechť platí předpoklady věty (*) a (λ_1, u_1) je dom. vl. pár A. Pak
$a_{11}^{(k)} := e_1^* A_k e_1 \xrightarrow{k\to\infty} \lambda_1$, $H P_k e_1 \xrightarrow{k\to\infty} u_1$
lineárně s rychlostí $\left|\frac{\lambda_2}{\lambda_1}\right|$.

D: $A_0^{k+1} e_1 = P_k S_k e_1 = p_1^{(k)} \cdot s_{11}^{(k)}$, kde

$s_{11}^{(k)} := e_1^* S_k e_1 \in \mathbb{C}$, $p_1^{(k)} := P_k e_1 \in \mathbb{C}^n$

$\Rightarrow p_1^{(k)} = \frac{1}{s_{11}^{(k)}} \cdot A_0^{k+1} e_1$, $\|p_1^{(k)}\| = 1$

$\neq 0$ neboť A je regulární, tj. R_i je regul. $i = 0, \dots, k$

P_k unitární

$\Rightarrow p_1^{(k)}$ je vektor z mocninné metody na matici A_0 se start. vektorem e_1

$\overset{A_0 = H^* A H}{\Longrightarrow} H p_1^{(k)}$ je —— || —— na matici A,

kde $|\lambda_1| > |\lambda_2| \overset{\|H p_1^{(k)}\| = 1}{\Longrightarrow} H p_1^{(k)} \xrightarrow{k\to\infty} u_1$

lineárně s rychlostí $\left|\frac{\lambda_2}{\lambda_1}\right|$

Ray. podíl z $H r_1^{(k)}$ a matice A je :

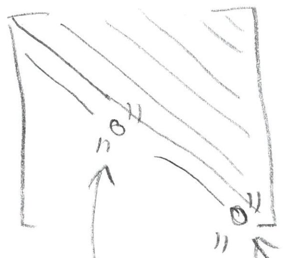
$$(H r_1^{(k)})^* A (H r_1^{(k)}) = r_1^{(k)*} \cdot H^* A H r_1^{(k)} =$$

$$= r_1^{(k)*} A_0 r_1^{(k)} \overset{\substack{\text{Lemma} \\ (0)}}{=} r_1^{(k)*} \cdot P_k A_{k+1} \underbrace{P_k^*}_{\text{unitární}} \underbrace{r_1^{(k)}}_{= P_k e_1} =$$

$$= e_1^* A_{k+1} e_1 = a_{11}^{(k)}$$

$\Rightarrow a_{11}^{(k)} \xrightarrow{k \to \infty} \lambda_1$ dle V. o konv. mocninné metody lineárně s rychl. $\left|\frac{\lambda_2}{\lambda_1}\right|$ □

Implementační aspekty:

$A_k =$ 

– v průběhu iterací může být nějaký poddiag. prvek $a_{i+1,i}^{(k)} \approx 0$

→ pokud $a_{n,n-1}^{(k)} \approx 0$, pak zahodíme posl. řádek a sloupec A_k a voláme QR alg. rekurzivně na menší matici $\tilde{A}_k \in \mathbb{C}^{(n-1)\times(n-1)}$ ... Redukce rozměru

→ pokud $a_{i+1,i}^{(k)} \approx 0$, $i = 1, \dots, n-2$, pak voláme rekurzivně QR alg. na 2 podmatice $A_k =$  $\hat{A}_k$ a $\hat{A}_k$ ... deflace

→ ZVÝŠENÍ EFEKTIVITY VÝPOČTU

5.3 QR algoritmus pro SVD

(11)

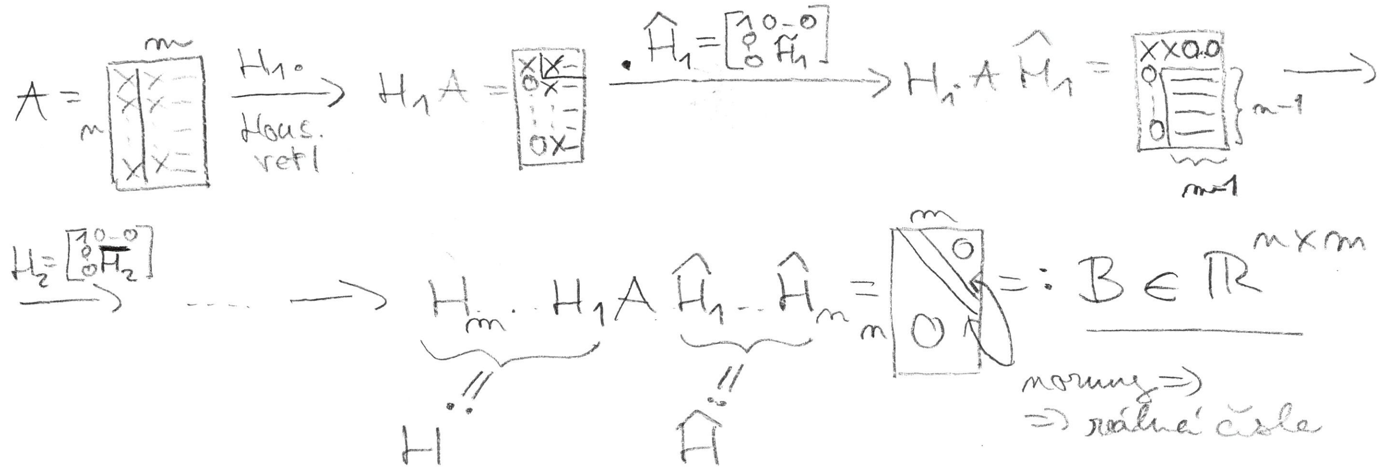
→ QR alg. lze upravit pro aproximaci SVD

Úloha: ‖ $A \in \mathbb{C}^{n\times m}$, $A \stackrel{?}{=} U\Sigma V^*$ ← unitární, $\Sigma = \begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & & 0\end{bmatrix} \in \mathbb{R}^{n\times m}$

$\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_r > 0$, $r = \mathrm{rank}(A)$

KROK 1: PREPROCESING

cíl: převést A do jednoduššího tvaru unitární transformací zleva a zprava (nemusí být stejné!)



máme: ‖ $H \cdot A \cdot \hat H = B$, $H, \hat H$ - unitární

⇒ A, B mají stejná sing. čísla

- výp. náklady $O(n^3)$, zpětně stabilní

KROK 2: ITERACE

cíl: iteračně dopočítat sing. č. B, tj. vl. čísla $B^T B = \; \in \mathbb{R}^{m\times m}$ symetr.

⇒ aplikujeme QR-algoritmus na SPSD tridiagonální matici $B^T B$ (bez preprocessingu)

– nejp. náklady $O(n^2)$, zpětná stabilita

pozn: alg. lze upravit, aby pracoval přímo s B, $B^T B$ se nekonstruuje

pozn: sing. vektory lze dopočítat z $H, \hat{H}$ a matic Q_k z QR-alg.