Základné algoritmy pre výpočet vlastných čísiel a vektorov

Ing. Gabriel Okša, CSc.

Matematický ústav
Slovenská akadémia vied
Bratislava

Stavebná fakulta STU

Obsah

Definícia vlastných čísiel a vektorov

- Nech A je matica rádu $n \times n$. Potom λ je vlastné číslo matice A, ak existuje nenulový vektor x taký, že platí:

  $$Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0.$$

  Množina vlastných čísiel sa nazýva spektrum matice A.
  Vektor $x \neq 0$ sa nazýva (pravý) vlastný vektor matice A prislúchajúci k vlastnému číslu λ.
  Vektor $y \neq 0$, ktorý vyhovuje rovnici

  $$y^T A = \lambda y^T$$

  sa nazýva ľavý vlastný vektor matice A prislúchajúci k vlastnému číslu λ.
- Homogénny systém rovníc $(A - \lambda I)x = 0$ má nenulové riešenie práve vtedy, ak

  $$\det(A - \lambda I) = 0.$$

Základné fakty - 1

- Polynóm premennej λ definovaný ako

$$P_\lambda = \det(A - \lambda I)$$

  má stupeň n a nazýva sa charakteristický polynóm matice A. Takže vlastné čísla matice A sú *korene jej charakteristického polynómu.*
- Regulárna matica A má nenulové vlastné čísla.
  Ak (λ, x) je vlastný pár regulárnej matice A, potom $(1/\lambda, x)$ je korešpondujúci vlastný pár matice A^{-1}, $(\lambda - \sigma, x)$ je vlastný pár matice $A - \sigma I$ a (λ^k, x) je vlastný pár matice A^k.
- Vlastné čísla trojuholníkovej matice sú jej diagonálne prvky.
- Nech T je regulárna matica. Potom matice A a TAT^{-1} majú rovnaké vlastné čísla (matica TAT^{-1} je podobná matici A).

Základné fakty - 2

- Vlastné vektory prislúchajúce navzájom rôznym vlastným číslam sú lineárne nezávislé.
- Matica A je podobná diagonálnej matici D práve vtedy, keď má úplnú množinu n lineárne nezávislých vektorov.
- Nech A je hermitovská matica, t.j. $A = A^H$, v reálnom prípade $A = A^T$ (symetrická). Potom platí:
  1. Existuje unitárna matica U tak, že $U^H AU = D$, kde D je diagonálna matica. Prvky D sú vlastné čísla a stĺpce U sú vlastné vektory matice A.
  2. Vlastné čísla matice A sú reálne. Ak je navyše matica A pozitívne definitná, potom vlastné čísla sú kladné.
  3. Vlastné vektory matice A sú ortonormálne.

Lokalizácia - Geršgorinove disky

- V niektorých aplikáciach (napr. stabilita dynamických systémov, konvergencia iteračných metód a pod.) nie sú dôležité hodnoty vlastných čísiel, ale ich lokalizácia v komplexnej rovine ($\mathrm{Re}(\lambda) < 0,\ |\lambda| < 1$).
- Geršgorin (1931): Nech A je matica rádu n. Položme

$$r_i = \sum_{j=i, j\neq i}^{n} |a_{ij}|,\ i = 1, 2, \ldots, n.$$

Potom každé vlastné číslo λ spĺňa aspoň jednu z nasledujúcich nerovností:

$$|\lambda - a_{ii}| \leq r_i,\ i = 1, 2, \ldots, n,$$

t.j. spektrum matice A leží v zjednotení n Geršgorinových diskov $\{z \in \mathbb{C} : |\lambda - a_{ii}| \leq r_i\},\ i = 1, 2, \ldots, n.$

Disjunktné Geršgorinove disky

- Nech r Geršgorinových diskov je disjunktných vzhľadom na zostávajúcich $n - r$ diskov ($r < n$). Potom zjednotenie týchto r diskov obsahuje presne r vlastných čísiel matice A.
- **Príklad**:

$$A = \begin{pmatrix} 1 & 0.1 & 0.2 \\ 0.2 & 4 & 0.3 \\ 0.4 & 0.5 & 8 \end{pmatrix}.$$

Geršgorinove disky: $\mathcal{D}_1 = \{z : |z - 1| \leq 0.3\}$, $\mathcal{D}_2 = \{z : |z - 4| \leq 0.5\}$, $\mathcal{D}_3 = \{z : |z - 8| \leq 0.9\}$.

Všetky 3 disky sú *navzájom* disjunktné, takže každý z nich obsahuje práve jedno vlastné číslo ($\lambda_1 = 0.9834$, $\lambda_2 = 3.9671$, $\lambda_3 = 8.0495$).

Mocninová metóda (MM)

- Je veľa aplikácií, kde potrebujeme poznať iba niekoľko vybraných vlastných čísiel a vektorov. Mocninová metóda je určená na výpočet dominantného VČ a VV.
- Nech A má VČ: $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \cdots \geq |\lambda_n|$ a nech v_1 je VV prislúchajúci k λ_1. Nech je A diagonalizovateľná.

1. Zvoľ $x_0 \neq 0, \mathrm{tol},\ m = 0$.
2. **do** $\{$
3. $\quad m = m + 1;$
4. $\quad \hat{x}_m = Ax_{m-1};$
5. $\quad$ Nájdi α_m tak, že: $|\alpha_m| = \max_{1 \leq i \leq n}\{|\hat{x}_m^i|\};$
   (definujme: $\alpha_m = \max(\hat{x}_m)$;)
6. $\quad x_m = (\alpha_m)^{-1} \cdot \hat{x}_m;$
7. $\quad x_m = x_m / \|x_m\|$; //normovanie
8. $\quad r_m = Ax_m - \alpha_m x_m;$
9. $\}$ **while** $(\|r_m\| >= \mathrm{tol});$

MM - konvergencia - 1/3

Konvergencia: $\alpha_m \to \lambda_1$ a $x_m \to c v_1$, kde c je konštanta.
Dôkaz:
A je diagonalizovateľná $\Rightarrow$ $x_0 = \sum_{i=1}^{n} \beta_i v_i$, $\|v_i\| = 1$, vektory v_i sú VV matice A a tvoria bázu $\mathbb{R}^n$. Potom $x_m = \frac{A^m x_0}{\max(A^m x_0)}$, kde:

$$
\begin{aligned}
A^m x_0 &= A^m \left(\sum_{i=1}^{n} \beta_i v_i \right) = \sum_{i=1}^{n} \beta_i \lambda_i^m v_i \\
&= \lambda_1^m \left[\beta_1 v_1 + \beta_2 \left(\frac{\lambda_2}{\lambda_1} \right)^m v_2 + \cdots + \beta_n \left(\frac{\lambda_n}{\lambda_1} \right)^m v_n \right], \\
\max(A^m x_0) &= \lambda_1^m \left[\beta_1 v_1^{j_m} + \beta_2 \left(\frac{\lambda_2}{\lambda_1} \right)^m v_2^{j_m} + \cdots + \beta_n \left(\frac{\lambda_n}{\lambda_1} \right)^m v_n^{j_m} \right], \\
v_i^{j_m} &= \text{komponenta } j_m \text{ vektora } v_i.
\end{aligned}
$$

MM - konvergencia - 2/3

- $(\lambda_k/\lambda_1)^m \to 0$, $|v_k^i| \le 1 \ \Rightarrow$ existuje m_0 tak, že $\forall m > m_0$:

$$\max(A^m x_0) = \lambda_1^m \left[\beta_1 v_1^\ell + \beta_2 \left(\frac{\lambda_2}{\lambda_1}\right)^m v_2^\ell + \cdots + \beta_n \left(\frac{\lambda_n}{\lambda_1}\right)^m v_n^\ell\right],$$

kde $|v_1^\ell| = \max_{1\le k\le n}\{|v_1^k|\}$.

- Potom $\forall m > m_0$:

$$x_m = \frac{\lambda_1^m \left[\beta_1 v_1 + \beta_2 \left(\frac{\lambda_2}{\lambda_1}\right)^m v_2 + \cdots + \beta_n \left(\frac{\lambda_n}{\lambda_1}\right)^m v_n\right]}{\lambda_1^m \left[\beta_1 v_1^\ell + \beta_2 \left(\frac{\lambda_2}{\lambda_1}\right)^m v_2^\ell + \cdots + \beta_n \left(\frac{\lambda_n}{\lambda_1}\right)^m v_n^\ell\right]},$$

takže $x_m \to (v_1^\ell)^{-1} \cdot v_1 \equiv w_1$.

MM - konvergencia - 3/3

- Vieme: $x_m = Ax_{m-1}/\alpha_m \to w_1$, takže aj $x_{m-1} \to w_1$, a teda $Ax_{m-1} \to Aw_1 = \lambda_1 w_1$. Nech $\alpha_m \to \alpha$. Potom:

$$x_m - w_1 = \frac{Ax_{m-1}}{\alpha_m} - w_1 \to \frac{(\lambda_1 - \alpha)w_1}{\alpha} = 0,$$

  takže $\alpha = \lambda_1$, a teda $\alpha_m \to \lambda_1$. $\square$
- *Rýchlosť konvergencie*: je daná podielom $|\lambda_2|/|\lambda_1|$, ktorý môže byť blízky k 1 $\Rightarrow$ pomalá konvergencia.
- Použitie posuvu σ: MM sa aplikuje na $A - \sigma I$, ktorá má dominantné VČ $\lambda_1 - \sigma$ a rovnaké VV ako A. Potom rýchlosť konvergencie je daná podielom $|\lambda_2 - \sigma|/|\lambda_1 - \sigma|$, čo môže byť podstatne menšie ako $|\lambda_2|/|\lambda_1|$, takže sa docieli podstatne rýchlejšia konvergencia. Pre voľbu vhodného posuvu σ je dôležitá informácia o lokalizácii λ_1 a λ_2 $\Rightarrow$ použitie Geršgorinových diskov.

MM aplikovaná na A^{-1}

- Ak $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_{n-1}| > |\lambda_n| > 0$, potom A^{-1} má VČ $1/\lambda_i$ a platí:

  $$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| \geq \left|\frac{1}{\lambda_{n-2}}\right| \geq \cdots \geq \left|\frac{1}{\lambda_1}\right| > 0,$$

  t.j. $1/\lambda_n$ je dominantné VČ matice A^{-1}. Vlastné vektory matice A^{-1} sú rovnaké ako VV matice A.
- Takže MM aplikovaná na A^{-1} vypočíta vlastný pár $(1/\lambda_n, x)$, čiže získame odhad najmenšieho VČ λ_n matice A spolu s VV.
- Kroky č. 4 a 8: vektory $\hat{x}_m = A^{-1}x_{m-1}$ resp. $y_m = A^{-1}x_m$ získame riešením rovníc $A\hat{x}_m = x_{m-1}$ resp. $Ay_m = x_m$ (priama alebo iteračná metóda).

Metóda inverzných iterácií (MII)

- Ak je k dispozícii kvalitná aproximácia σ dominantného VČ λ_1, potom MII efektívne počíta prislúchajúci VV (veľmi rýchla konvergencia).
  Kvalitná aproximácia: $|\lambda_1 - \sigma| \ll |\lambda_i - \sigma|, \ \forall i \neq 1$. Presnosť počítača: konštanta ϵ; pri dvojitej presnosti je $\epsilon \approx 1.11 \times 10^{-16}$.

1. Zvoľ $x_0 \neq 0, \ \sigma, \ m = 0$.
2. **do** $\{$
3. $\quad m = m + 1;$
4. $\quad$ Nájdi $\hat{x}_m$ vyriešením sústavy: $(A - \sigma I)\hat{x}_m = x_{m-1};$
5. $\quad \alpha_m = \max(\hat{x}_m);$
6. $\quad x_m = (\alpha_m)^{-1} \cdot \hat{x}_m;$
7. $\quad x_m = x_m / \|x_m\|;$ //normovanie
8. $\quad r_m = Ax_m - \sigma x_m;$
9. $\}$ **while** $(\|r_m\|_\infty >= \|A\|_\infty \, \epsilon);$

MII - poznámky - 1/2

- MII je vlastne mocninová metóda aplikovaná na maticu $(A - \sigma I)^{-1}$, pričom je k dispozícii kvalitná aproximácia σ VČ λ_1, ktorá sa počas iterácií nemení.
- Riešenie LS $(A - \sigma I)\hat{x}_m = x_{m-1}$: Gaussova eliminácia s parciálnou pivotáciou, t.j. na začiatku: $P(A - \sigma I) = LU$, kde P je permutačná matica ($P^{-1} = P^T$), L je dolná a U horná troj., a potom 3 kroky:
  $z_{m-1} = Px_{m-1},\ Ly_m = z_{m-1},\ U\hat{x}_m = y_m$.
- Konvergencia: postupnosť $\{x_m\}$ konverguje k násobku VV prislúchajúcemu k VČ λ_1, pretože:

$$\hat{x}_m = \frac{1}{(\lambda_1 - \sigma)^m} \left[\beta_1 v_1 + \beta_2 \left(\frac{\lambda_1 - \sigma}{\lambda_2 - \sigma} \right)^m v_2 + \cdots + \right.$$
$$\left. + \beta_n \left(\frac{\lambda_1 - \sigma}{\lambda_n - \sigma} \right)^m v_n \right].$$

MII - poznámky - 2/2

- Voľba x_0 (Wilkinson 1965): Nech $e = (1, 1, \dots, 1)^T$ a $P(A - \sigma I) = LU$. Potom zvoľ: $x_0 = P^T L e$.
- Iná možnosť: Vypočítať M iterácií pomocou MM, a potom vziať aproximácie $x_0 = x_M$ a $\sigma = \alpha_M$ ako vstupy do MII. Toto je kombinácia MM a MII, kde MII slúži na spresnenie odhadu VV. Obvykle totiž MM dáva omnoho presnejší odhad VČ než korešpondujúceho VV.
- Potenciál MII: Ak máme k dispozícii kvalitný odhad σ ľubovoľného VČ λ_j (t.j. $|\lambda_j - \sigma| \ll |\lambda_k - \sigma|, \forall k \neq j$), potom MII spravidla veľmi rýchlo konverguje k VV prislúchajúcemu k λ_j.

Rayleighov kvocient

Nech A je symetrická matica, $A = A^T$. Nech vektor x je kvalitná aproximácia VV matice A. Potom tzv. Rayleighov kvocient

$$\sigma = \frac{x^T A x}{x^T x} = \min_{\mu} \|(A - \mu I)x\|$$

je dobrá aproximácia VČ príslušného k VV x.

Dôkaz: A je symetrická $\Rightarrow$ má ON systém vlastných vektorov $\{v_i\}$. Potom $x = \sum_{i=1}^{n} c_i v_i$ a nech x dobre aproximuje v_1, t.j. $c_1 \gg c_j$, $\forall j \neq 1$. Potom $Ax = \sum_{i=1}^{n} \lambda_i c_i v_i$ a

$$\sigma = \frac{x^T A x}{x^T x} = \frac{\lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2}{c_1^2 + c_2^2 + \cdots + c_n^2} =$$

$$= \lambda_1 \left[\frac{1 + \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{c_2}{c_1}\right)^2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)\left(\frac{c_n}{c_1}\right)^2}{1 + \left(\frac{c_2}{c_1}\right)^2 + \cdots + \left(\frac{c_n}{c_1}\right)^2} \right] \approx \lambda_1.$$

Metóda Rayleighových iterácií (MRI)

Je to kombinácia Rayleighovho kvocientu a metódy inverzných iterácií.

1. Zvoľ $x_0 \neq 0, \ \|x_0\| = 1, \ \mathrm{tol}, \ m = -1, \ \mathrm{maxiter}.$
2. **do** $\{$
3. $\quad m = m + 1;$
4. $\quad \sigma_m = x_m^T A x_m;$ //Rayleighov kvocient
5. $\quad r_m = A x_m - \sigma_m x_m;$
6. $\quad$ **if** $(\|r_m\| < \mathrm{tol})$ **exit**;
7. $\quad$ Nájdi $\hat{x}_{m+1}$: $(A - \sigma_m I)\hat{x}_{m+1} = x_m;$
8. $\quad \alpha_{m+1} = \max(\hat{x}_{m+1});$
9. $\quad x_{m+1} = (\alpha_{m+1})^{-1} \cdot \hat{x}_{m+1};$
10. $\quad x_{m+1} = x_{m+1} / \|x_{m+1}\|;$ //normovanie
11. $\}$ **while** $(m < \mathrm{maxiter});$

MRI - poznámky

- V každom iteračnom kroku sa rieši LS s rôznou maticou sústavy (σ_m sa mení!). Preto je táto metóda je náročná, pokiaľ sa predtým symetrická matica A nezredukuje na sym. trojdiagonálnu: $T = Q^T AQ$, kde Q je OG (Householderove reflexie). Potom T má rovnaké VČ ako A; ak y je VV matice T, potom $x = Qy$ je VV matice A.
  LU faktorizácia matice T v 7. kroku MRI je rýchla, oba faktory sú bidiagonálne (pozri D-Lanczosov algoritmus pre LS) a pre 7. krok MRI treba iba $4n - 4$ aritmetických operácií.
- Konvergencia MRI je kubická - t.j. počet presných číslic v riešení sa strojnásobí v každej iterácii.
- Výber x_0: odporúča sa absolvovať niekoľko iterácií MM a výstup použiť ako vstup do MRI.

Princíp deflácie - 1/3

- Deflácia je metóda na výpočet subdominantného vlastného čísla (VČ) a vlastného vektora (VV): $|\lambda_1| > |\lambda_2| > |\lambda_3| \geq \cdots \geq |\lambda_n|$. Pritom sa predpokladá, že k dispozícii je kvalitný odhad dominantného páru $(\lambda_1,\, x_1)$.
- Householderova matica H (elementárna reflexia): Nech $e_1 = (1, 0, \cdots, 0)^T$ a nech $x = (x_1, x_2, \cdots, x_n)^T \neq \alpha e_1$. Potom existuje vektor u,

$$u = x + \operatorname{sgn}(x_1)\, \|x\|_2\, e_1$$

tak, že

$$H = I - 2\frac{uu^T}{u^T u}, \quad Hx = -\operatorname{sgn}(x_1)\, \|x\|_2\, e_1 .$$

H je symetrická, ortogonálna, $H^2 = I$, $H^{-1} = H$.

Princíp deflácie - 2/3

- Nech $(\lambda_1,\ v_1)$ je dominantný vlastný pár matice A a nech H je taká HM, že $Hv_1 = \alpha e_1$. Potom

$$A_1 = H^{-1}AH = HAH = \begin{pmatrix} \lambda_1 & c^T \\ 0 & A_2 \end{pmatrix}$$

kde A_2 je rádu $n-1$ a má tie isté vlastné čísla ako A okrem λ_1.

Dôkaz: $Av_1 = \lambda_1 v_1 \Rightarrow HAv_1 = \lambda_1 Hv_1 \Rightarrow HAHHv_1 = \lambda_1 Hv_1$ (lebo $H^2 = I$), takže:
$HAH(\alpha e_1) = \lambda_1(\alpha e_1) \Rightarrow HAHe_1 = \lambda_1 e_1$, čiže prvý stĺpec matice HAH je $(\lambda_1, 0, \cdots, 0)^T$.

Druhé tvrdenie: $\det(HAH - \lambda I) = (\lambda_1 - \lambda)\det(A_2 - \lambda I)$. $\square$

- Ak teda $|\lambda_1| > |\lambda_2| > |\lambda_3| \geq \cdots \geq |\lambda_n|$, potom po deflačnom kroku $A_1 = HAH$ je λ_2 dominantné VČ matice A_2.

Princíp deflácie - 3/3

Na výpočet niekoľkých najväčších resp. najmenších vlastných párov matice A sa používa nasledujúca kombinácia MM, MII a deflácie:

1. Použi mocninovú metódu aplikovanú na A (resp. mocninovú metódu aplikovanú na A^{-1}) na výpočet dobrej aproximácie v absolútnej hodnote najväčšieho (resp. najmenšieho) VČ a príslušného VV.
2. Aplikuj metódu inverzných iterácií s aproximovaným VČ (zostáva nemenné) a s aproximovaným VV z kroku 1 ako počiatočným vektorom. Takto sa získa spravidla veľmi spresnený odhad VV.
3. Aplikuj defláciu na výpočet ďalšieho vlastného páru.
4. Opakuj kroky 1-3 pre požadovaný počet vlastných párov.

Metóda QR iterácií (MQRI)

- Je určená na výpočet VČ obecnej, hustej matice A. Patrí v súčasnosti k najčastejšie používaným metódam.
- Je založená na špeciálnych unitárnych resp. ortogonálnych podobnostných transformáciach originálnej matice A, ktoré konvergujú k hornej trojuholníkovej resp. k hornej kvázi-trojuholníkovej matici (tzv. Schurova triangularizácia).
- Schurova triangularizácia: Ak A je matica rádu n, potom existuje unitárna matica U tak, že $U^H AU = T$, kde T je horná trojuholníková. VČ matice T sú rovnaké ako VČ matice A a sú umiestnené na diagonále matice T.
- Aj keď je A reálna, matice U a T sú obecne komplexné! Je to preto, že VČ reálnej matice môže byť komplexné. Ak $\lambda \in \mathbb{C}$ je VČ matice A, potom aj $\bar{\lambda}$ je v spektre A.

MQRI - reálna Schurova forma (RSF)

- Nech A je reálna matica. Potom existuje ortogonálna matica Q tak, že

$$Q^T A Q = R = \begin{pmatrix} R_{11} & R_{12} & \cdots & R_{1k} \\ 0 & R_{22} & \cdots & R_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{kk} \end{pmatrix},$$

kde každé R_{ii} je buď skalár alebo blok rádu 2×2. Skalárne diagonálne prvky odpovedajú reálnym VČ a bloky 2×2 na diagonále odpovedajú párom komplexne združených VČ.
- Dôležité: pri podobnostnej transformácii $Q^T A Q$ sa nevyžaduje komplexná aritmetika. Počet blokov rádu 2×2 na diagonále definuje počet komplexne združených párov VČ matice A. Komplexná aritmetika nastupuje až pri výpočte koreňov charakteristických polynómov týchto blokov.

MQRI - RSF - bloky 2×2

- Diagonálne bloky majú špeciálne charakteristické polynómy $p(\lambda)$:

$$\begin{aligned} p(\lambda) &= \det \begin{pmatrix} r_{11} - \lambda & r_{12} \\ r_{21} & r_{22} - \lambda \end{pmatrix} = (r_{11} - \lambda)(r_{22} - \lambda) - r_{12}\, r_{21} \\ &= \lambda^2 - (r_{11} + r_{22})\lambda + r_{11}\, r_{22} - r_{12}\, r_{21}. \end{aligned}$$

Na druhej strane:

$$p(\lambda) = (\lambda - \lambda_i)(\lambda - \bar{\lambda}_i) = \lambda^2 - (\lambda_i + \bar{\lambda}_i)\lambda + \lambda_i \bar{\lambda}_i,$$

takže:

$$\begin{aligned} r_{11} + r_{22} &= \lambda_i + \bar{\lambda}_i = 2\,\mathrm{Re}(\lambda_i), \\ r_{11}\, r_{22} - r_{12}\, r_{21} &= \lambda_i \bar{\lambda}_i = \left[\mathrm{Re}(\lambda_i)^2 + \mathrm{Im}(\lambda_i)^2 \right]. \end{aligned}$$

MQRI - základná forma

- MQRI je definovaná postupnosťou dvoch základných krokov: QR rozklad a súčin faktorov v opačnom poradí.
  1. Polož $A_0 = A$ a vypočítaj QR rozklad: $A_0 = Q_0 R_0$.
  2. Zameň poradie faktorov a vypočítaj: $A_1 = R_0 Q_0$.
  3. Vypočítaj QR rozklad: $A_1 = Q_1 R_1$ a následne: $A_2 = R_1 Q_1$.
  4. Krok k: $A_k = R_{k-1} Q_{k-1} = Q_k R_k$.
- Takto vznikne postupnosť matíc $\{A_k\}$, ktoré sú ortogonálne podobné matici A:
  $$A_k = R_{k-1} Q_{k-1} = Q_{k-1}^T A_{k-1} Q_{k-1} = (Q_{k-1}^T Q_{k-2}^T) A_{k-2} (Q_{k-2} Q_{k-1}) = \cdots = (Q_0 Q_1 \cdots Q_{k-1})^T A_0 (Q_0 Q_1 \cdots Q_{k-1}).$$
  Takže každá A_k má rovnaké VČ ako A.
- Ak by postupnosť $\{A_k\}$ konvergovala k trojuholníkovej matici resp. k RSF, potom by sme dospeli ku všetkým VČ matice A.

MQRI - konvergencia - 1/3

- Wilkinson, 1965: Nech je matica A diagonalizovateľná: $X^{-1}AX = \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, kde $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$ a nech existuje LU dekompozícia matice X^{-1} . Potom postupnosť $\{A_k\}$ konverguje k hornej trojuholníkovej matici.
- Základná verzia MQRI je príliš nákladná, pretože v každom kroku sa musí počítať QR rozklad plnej matice A_k, čo je rádovo n^3 operácií.

MQRI - konvergencia - 2/3

- Trik: A_0 sa zredukuje ortogonálnou podobnostnou transformáciou na hornú Hessenbergovu (pomocou Householderových matíc).
  Potom sa vypočíta jej QR rozklad (rádovo n^2 operácií) pomocou Givensových rotácií: $Q_0^T A_0 = R_0$, kde:

  $$Q_0^T = G(n-1, n, \theta_{n-1}) \cdots G(2, 3, \theta_2)\, G(1, 2, \theta_1).$$
- Matica Q_0 je horná Hessenbergova (lebo $G(i, i+1, \theta_i)$ má nenulové mimodiag. prvky iba na pozíciach $(i+1, i)$ a $(i, i+1)$) a matica R_0 je horná trojuholníková $\Rightarrow$ $A_1 = R_0 Q_0$ je horná Hessenbergova.
- Matematickou indukciou dokážeme, že všetky matice A_k sú horné Hessenbergove - tzv. Hessenbergova forma QR iterácií.

MQRI - konvergencia - 3/3

- Nech A je redukovaná na hornú Hessenbergovu maticu: $H = P^T AP$, kde P je súčin Householderových reflexií. Potom QR iterácie produkujú postupnosť H_k Hessenbergových matíc. Nech $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$.
- Rýchlosť konvergencie: Prvok $h_{i,i-1}^{(k)}$ matice H_k konverguje k nule rýchlosťou danou geometrickou postupnosťou $\left|\frac{\lambda_i}{\lambda_{i-1}}\right|^k$, čo môže byť veľmi pomalé, ak $\lambda_i \approx \lambda_{i-1}$.
- Trik $=$ posuv: Nech $\hat{\lambda}_i$ je aproximácia VČ λ_i. Potom MQRI sa aplikuje na $H - \hat{\lambda}_i I$, takže geometrická postupnosť má tvar $\left|\frac{\lambda_i - \hat{\lambda}_i}{\lambda_{i-1} - \hat{\lambda}_i}\right|^k$.
  Príklad: $\lambda_i = 0.99$, $\lambda_{i-1} = 1.1$, $\hat{\lambda}_i = 1.0$. Potom $|\lambda_i||/||\lambda_{i-1}| = 0.9$, ale $|\lambda_i - \hat{\lambda}_i|/|\lambda_{i-1} - \hat{\lambda}_i| = 0.1$.

MQRI - jednoduchý posuv

- V praxi nie je odhad $\hat{\lambda}_i$ na začiatku iterácií k dispozícii $\Rightarrow$ berie sa $h_{nn}^{(0)}$ (prvok (n, n) matice $H = H_0$). Tento posun sa obnovuje v každom "dvojkroku" MQRI.

  1. Transformuj OG podobnostnou transformáciou A na hornú Hessenbergovu maticu: $H = P^T AP$.
  2. $H_0 = H$; $m = 0$;
  3. **do** $\{$
  4. $\quad H_k - h_{nn}^{(k)} I = Q_k R_k$;
  5. $\quad H_{k+1} = R_k Q_k + h_{nn}^{(k)} I$;
  6. $\}$ **while** $(h_{n,n-1} \neq 0)$;

- Kedy je $h_{n,n-1}^{(k)} \approx 0$? Beresford Parlett (USA) navrhuje:

  $$|h_{n,n-1}^{(k)}| \leq \mathrm{tol} \cdot \left(|h_{nn}^{(k)}| + |h_{n-1,n-1}^{(k)}|\right),$$

  kde $\mathrm{tol} \ll 1$.

MQRI - dvojitý posuv

- Jednoduchý posuv funguje dobre, ak má matica A iba reálne VČ. Ak má však reálna matica A aj komplexné VČ (vo forme komplexne združených párov), potom:
  1. Od určitého štádia MQRI bude jasné, že napr. $h_{n,n-1}^{(k)}$ nekonverguje k nule, ale spodný blok o rozmere 2×2 korešponduje komplexne združenému páru.
  2. V takom prípade reálny prvok $h_{nn}^{(k)}$ nemôže byť kvalitnou aproximáciou komplexného VČ.
  3. Potom je prirodzené použiť obe VČ spodného 2×2 bloku ako posuv $\Rightarrow$ MQRI s dvojitým posuvom.
- Nech k_1 a $k_2 = \bar{k}_1$ sú komplexne združené VČ spodného 2×2 bloku. Potom jeden iteračný krok MQRI s dvojitým posuvom má tvar:

$$H_s - k_1 I = Q_s R_s; \quad H_{s+1} = R_s Q_s + k_1 I;$$
$$H_{s+1} - k_2 I = Q_{s+1} R_{s+1}; \quad H_{s+2} = R_{s+1} Q_{s+1} + k_2 I;$$

Dvojitý posuv s reálnou aritmetikou 1/3

- Pretože k_1 a k_2 sú komplexné čísla, implementácia dvojitého posuvu by vyžadovala komplexnú aritmetiku aj pre reálnu východziu maticu (a následné matice H_k by už boli komplexné!). Dá sa to obísť, t.j. pracovať stále s reálnou aritmetikou.
- Pretože $k_1 = \bar{k}_2$, je matica $N = (H_s - k_2 I)(H_s - k_1 I) = = H_s^2 - (k_1 + k_2)H_s + k_1 k_2 I$ reálna.
- $(Q_s Q_{s+1})(R_{s+1} R_s)$ je QR faktorizácia matice N:
$$
\begin{aligned}
N &= (H_s - k_2 I)(H_s - k_1 I) = (H_s - k_2 I) Q_s R_s \\
&= Q_s Q_s^H (H_s - k_2 I)\, Q_s R_s = Q_s\, (Q_s^H H_s Q_s - k_2 I)\, R_s \\
&= Q_s\, [(R_s + k_1 Q_s^H) Q_s - k_2 I]\, R_s \\
&= Q_s\, [H_{s+1} - k_1 I + k_1 I - k_2 I]\, R_s = (Q_s Q_{s+1})(R_{s+1} R_s).
\end{aligned}
$$

Dvojitý posuv s reálnou aritmetikou 2/3

- Pretože N je reálna, $(Q_s Q_{s+1})$ je tiež reálna a ortogonálna.
- Ďalší krok: H_{s+2} je OG podobná k H_s práve cez $(Q_s Q_{s+1})$:

$$\begin{aligned}
H_{s+2} &= R_{s+1} Q_{s+1} + k_2 I = Q_{s+1}^H (H_{s+1} - k_2 I) Q_{s+1} + k_2 I \\
&= Q_{s+1}^H \left[R_s Q_s + (k_1 - k_2) I \right] Q_{s+1} + k_2 I \\
&= Q_{s+1}^H \left[Q_s^H (H_s - k_1 I)\, Q_s + (k_1 - k_2) I \right] Q_{s+1} + k_2 I \\
&= Q_{s+1}^H Q_s^H H_s Q_s Q_{s+1} = (Q_s Q_{s+1})^T H_s (Q_s Q_{s+1}).
\end{aligned}$$

  Matice $(Q_s Q_{s+1})$ a H_s sú reálne $\Rightarrow$ H_{s+2} je reálna a "preskočili" sme H_{s+1}!
- Nakoniec: Pri formovaní $N = H_s^2 - (k_1 + k_2) H_s + k_1 k_2 I$ netreba počítať komplexné korene k_1, k_2, lebo potrebujeme iba reálne čísla:

  $t = k_1 + k_2 = \mathrm{trace}(2 \times 2 \text{ bloku}) = h_{n-1,n-1}^{(s)} + h_{n,n}^{(s)}$,

  $d = k_1 k_2 = \det(2 \times 2 \text{ bloku}) = h_{n-1,n-1}^{(s)}\, h_{n,n}^{(s)} - h_{n,n-1}^{(s)}\, h_{n-1,n}^{(s)}$.

Dvojitý posuv s reálnou aritmetikou 3/3

- Konečná podoba jedného iteračného kroku MQRI s dvojitým posuvom v reálnej aritmetike:
  1. $t = h^{(s)}_{n-1,n-1} + h^{(s)}_{n,n}$; $d = h^{(s)}_{n-1,n-1}\, h^{(s)}_{n,n} - h^{(s)}_{n,n-1}\, h^{(s)}_{n-1,n}$;
  2. $N = H_s^2 - tH_s + dI$; // N nie je horná Hessenbergova!
  3. $N = QR$;
  4. $H_{s+2} = Q^T H_s Q$.
- Toto sa nazýva explicitná QR iterácia s dvojitým posuvom v reálnej aritmetike, pretože sa explicitne počíta matica N a jej QR rozklad $\Rightarrow$ jeden iteračný krok vyžaduje rádove n^3 aritmetických operácií.
- Existuje aj implicitná QR iterácia s dvojitým posuvom v reálnej aritmetike, kde sa používa iba prvý stĺpec matice N $\Rightarrow$ jeden iteračný krok vyžaduje rádove n^2 aritmetických operácií.