Ortogonálne vektory a matice.
QR faktorizácia matice.

Ing. Gabriel Okša, CSc.

Matematický ústav
Slovenská akadémia vied
Bratislava

Stavebná fakulta STU

Obsah

Ortogonálne matice a vektory

- Nech $x, y \in \mathbb{C}^n$ sú dva nenulové vektory a nech je na $\mathbb{C}^n$ definovaný skalárny súčin $(x, y) = y^\star x = \sum_{i=1}^n \bar{y}_i\, x_i$. Potom vektor x je ortogonálny na vektor y, ak $(x, y) = 0$; hovoríme, že x je kolmý na y: $x \perp y$.
- Ak je k vektorov $x_1,\ x_2, \ldots, x_k$ navzájom ortogonálnych (t.j., každá dvojica je ortogonálna), potom sú lineárne nezávislé. Naopak to neplatí!
- Ak $x_1 \perp x_2$ a zároveň $\|x_1\| = \|x_2\| = 1$, potom sa takéto vektory nazývajú ortonormálne.
- Matica $Q \in \mathbb{C}^{n \times n}$ je unitárna, ak $Q^\star\, Q = Q\, Q^\star = I$ $(Q^\star = (\bar{Q})^T)$. Unitárna reálna matica sa nazýva ortogonálna. Ortogonálna matica, ktorej všetky stĺpce majú jednotkovú normu, sa nazýva ortonormálna.
- Unitárne matice sú regulárne a ich inverzia je jednoduchá: $Q^{-1} = Q^\star$. Súčin dvoch unitárnych matíc je unitárna matica.

Klasická Grammova-Schmidtova ortogonalizácia (CGS)

Nech je daných k LN vektorov $x_1, x_2, \ldots, x_k \in \mathbb{R}^n$. Potom CGS je algoritmus, ktorý z týchto k LN vektorov vypočíta k vzájomne ortonormálnych vektorov $q_1, q_2, \ldots, q_k$.

1. **for** $(j = 1;\ j <= k;\ j++)\ \{$
2. **for** $(i = 1;\ i <= j-1;\ i++) \quad \alpha_{ij} = q_i^T x_j;$
3. $q_j = x_j - \sum_{i=1}^{j-1} \alpha_{ij}\, q_i;$
4. $\beta = \|q_j\|;$
5. $q_j = q_j / \beta; \quad \}$

CGS je numericky nestabilný: vypočítané vektory q_j nemusia byť presne ortonormálne.

Modifikovaná Grammova-Schmidtova ortogonalizácia (MGS)

Vstup: $x_1, x_2, \ldots, x_k$, LN.
Výstup: $q_1, q_2, \ldots, q_k$, ON.

1. **for** $(j = 1;\ j <= k;\ j++)$ $\{$
2. $\quad q_j = x_j;$
3. $\quad$ **for** $(i = 1;\ i <= j - 1;\ i++)$ $\{$
4. $\qquad \alpha_{ij} = q_i^T q_j;$
5. $\qquad q_j = q_j - \alpha_{ij}\, q_i; \quad \}$
6. $\quad \beta = \|q_j\|;$
7. $\quad q_j = q_j / \beta; \quad \}$

Táto modifikácia (použitie α_{ij} okamžite po výpočte) vedie k lepším numerickým vlastnostiam MGS oproti CGS.

CGS versus MGS: numerika

- Zoberme iba dva vektory $x_1,\ x_2 \in \mathbb{R}^n$ a nech $\|x_1\| = 1$, takže $q_1 = x_1$ (t.j. q_1 sa nevypočíta). Potom CGS sa redukuje na:
  $\tilde{\alpha}_{12} = \mathrm{fl}(q_1^T x_2);\ \tilde{q}_2 = \mathrm{fl}(x_2 - \mathrm{fl}(\tilde{\alpha}_{12}\, q_1))$ (neuvažujeme normalizáciu $\tilde{q}_2$).
- Björck (1994) ukázal: $|q_1^T\, \tilde{q}_2| < 1.06\,(2n+3)\,\|x_2\|\,\mu_M$. Takže pri CGS nemusí byť vypočítaný vektor $\tilde{q}_2$ OG ku q_1 (ak je napr. $\|x_2\| \gg 1$).
- MGS: Nech $\tilde{Q} = (\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_k)$ sú vypočítané vektory. Potom Björck (1994) ukázal, že strata ortogonality závisí od čísla podmienenosti matice $X = (x_1, x_2, \ldots, x_k)$:

$$\|I - \tilde{Q}^T \tilde{Q}\| \leq \frac{c_1\,\kappa(X)\,\mu_M}{1 - c_2\,\kappa(X)\,\mu_M},$$

  kde c_1 a c_2 sú malé konštanty. Takže pre $\kappa(X) \gg 1$ môže byť strata ortogonality neprijateľná.

Householderova transformácia (HT)

- Nech je daný vektor $u \in \mathbb{R}^n$, $u \neq 0$. Potom matica
$$H = I - \frac{2\,u\,u^T}{u^T u}$$
sa nazýva Householderova matica (tiež elementárna reflexia). Pomenované podľa amerického numerika Alstona Householdera (1904 - 1993).
- Je to symetrická a ortogonálna matica: $H^T = H$, $H^T H = HH^T = H^2 = I$.
- Pôsobenie H na vektor $x \in \mathbb{R}^n$: $y = Hx = x - \frac{2u^T x}{u^T u}\, u =$ reflexia vektora x v rovine kolmej na u a prechádzajúcej bodom 0. Pritom $\|Hx\|_2 = \|x\|_2$ (OGT nemení dĺžku vek.).
- Nech sú dané dva nenulové vektory $x, y \in \mathbb{R}^n$ s rovnakou Euklidovou normou: $\|x\|_2 = \|y\|_2$. Nech $u = (x - y)/\|x - y\|_2$. Potom $H = I - 2uu^T$ je reflexia x na y a naopak: $Hx = y$, $Hy = x$.

Nulovanie zložiek vektora pomocou HT

Nech je daný vektor $x = (x_1, x_2, \ldots, x_n)^T$. Hľadáme vektor u tak, aby Hx bol násobok vektora $e_1 = (1, 0, 0, \ldots, 0)^T$.
Dá sa ukázať, že:

$$u = x + \mathrm{sign}(x_1)\, \|x\|_2\, e_1;\;\; H\,x = (-sign(x_1)\, \|x\|, 0, 0, \ldots, 0)^T.$$

Vstup: vektor $x = (x_1, x_2, \ldots, x_n)$.
Výstup: vektor u a skalár σ také, že:
$Hx = (I - 2uu^T/(u^T u))x = (\sigma, 0, ..., 0)^T$

1. $m = \max(|x_i|),\; i = 1, 2, \ldots, n$;
2. **for** $(i = 1;\; i <= n;\; i++)$ $u_i = x_i/m$;
3. $\sigma = \mathrm{sign}(u_1)\sqrt{\sum_{i=1}^{n} u_i^2}$;
4. $u_1 = u_1 + \sigma$;
5. $\sigma = -m\,\sigma$;

Násobenie $B = H\,A$

- Householderova matica je definovaná Householderovým vektorom u. Pri násobení $B = H\,A$, kde $H \in \mathbb{R}^{m \times m}$ a $A \in \mathbb{R}^{m \times n}$, $m \geq n$, netreba sformovať H explicitne - stačí poznať vektor u.
- Označme matice po stĺpcoch: $A = (a_1, a_2, \ldots, a_n)$, $B = H\,A = (b_1, b_2, \ldots, b_n)$. Potom maticu B vypočítame v cykle po stĺpcoch:

$$b_i = \left(I - \frac{2uu^T}{\|u\|^2} \right) a_i = a_i - \frac{2u^T a_i}{\|u\|^2}\, u, \quad 1 \leq i \leq n.$$

Takže v cykle stačí počítať skalárne súčiny $u^T a_i$.

Givensova transformácia (GT)

- Nech sú dané dve reálne čísla c, s, kde $c^2 + s^2 = 1$ – t.j., $c = \cos\theta$, $s = \sin\theta$ pre nejaký uhol θ. Potom Givensova matica $G(i,j,\theta) = (g_{kl})$ s $i < j$ rádu n je matica, ktorá sa od identity I_n líši iba v prvkoch:
  $g_{ii} = g_{jj} = c,\ g_{ij} = s,\ g_{ji} = -s$. Nazvaná je podľa amerického numerika Wallacea Givensa (1910 - 1993).
- Givensova matica nie je symetrická, ale je ortogonálna: $G(i,j,\theta)\,G(i,j,\theta)^T = G(i,j,\theta)^T\,G(i,j,\theta) = I$.
- Geometrická interpretácia: $G(i,j,\theta)$ rotuje všetky vektory v rovine (i,j) o uhol $(-\theta)$ (t.j. v smere hodinových ručičiek) – preto sa $G(i,j,\theta)$ nazýva aj elementárna rotácia.
- Operácia $G(i,j,\theta)\,x$ má vplyv iba na zložky x_i a x_j vektora x; ostatné zložky zostanú nezmenené. Takže:
  $x_i' = c\,x_i + s\,x_j,\ x_j' = -s\,x_i + c\,x_j,\ x_k' = x_k$ pre $k \neq i,j$.

Nulovanie zložiek vektora pomocou GT

- Nech $x = (x_1, x_2)^T$, $x_2 \neq 0$. Potom $G(1, 2, \theta)$, kde:
  $c = x_1/\sqrt{x_1^2 + x_2^2}$; $s = x_2/\sqrt{x_1^2 + x_2^2}$;
  nuluje druhú zložku vektora x:
  $$G(1, 2, \theta)\, x = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \star \\ 0 \end{pmatrix}.$$
- Stabilný výpočet parametrov s, t:

  1. **if** $(|x_2| \geq |x_1|)$
  2. $\quad t = x_1/x_2$; $s = 1/\sqrt{1 + t^2}$; $c = s\, t$;
  3. **else if** $(|x_2| < |x_1|)$
  4. $\quad t = x_2/x_1$; $c = 1/\sqrt{1 + t^2}$; $s = c\, t$;
- Zovšeobecnenie: vo vek. $x = (x_1, x_2, \ldots, x_n)^T$ nulujeme jeho k-tu zložku x_k pomocou násobenia $G(i, k, \theta)\, x$, kde $i < k$. Najprv nájdeme 2×2 rotáciu (t.j. c a s) tak, aby:
  $\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x_i \\ x_k \end{pmatrix} = \begin{pmatrix} \star \\ 0 \end{pmatrix}$, a potom "vnoríme" c a s do I_n.

Násobenie $G(i, j, \theta)\, A$

- Násobenie $G(i, j, \theta)\, A$ mení iba riadky i a j matice $A = (a_{ij})$:

  1. **for** $(k = 1;\ k <= n;\ k++)\ \{$
  2. $\quad a = a_{ik};\ b = a_{jk}$
  3. $\quad a_{ik} = ac + bs;\ a_{jk} = -as + bc;\ \}$

  Dôležité: matica $G(i, j, \theta)$ nie je sformovaná explicitne!

- GT sa používa aj na nulovanie daného prvku v matici. Napr. chceme pomocou $G(i, j, \theta)$, $\mathbf{i} < \mathbf{j}$, vynulovať prvok $\mathbf{a_{ji}}$ (v dolnom trojuholníku A):

  1. Nájdi c, s tak, že:
  $$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a_{ii} \\ a_{ji} \end{pmatrix} = \begin{pmatrix} \star \\ 0 \end{pmatrix}.$$

  2. Vypočítaj $G(i, j, \theta)\, A$ pomocou predošlého algoritmu.

- Matica $G(i, j, \theta)$ opäť nie je sformovaná explicitne!

QRF pomocou HT - 1/2

Veta: Nech je daná matica $A \in \mathbb{R}^{m \times n}$, $m \geq n$. Potom existuje ortogonálna matica $Q \in \mathbb{R}^{m \times m}$ a horná trojuholníková matica $R \in \mathbb{R}^{n \times n}$ tak, že:

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix} = (Q_1, Q_2) \begin{pmatrix} R \\ 0 \end{pmatrix} = Q_1\, R, \text{ kde } 0 \in \mathbb{R}^{(m-n) \times n} \text{ je}$$

nulový blok a Q_1 má iba n ON stĺpcov. $\square$

Výpočet QR pomocou HT: Nech $s = \min\{m - 1, n\}$ (pre $m \geq n$ je $s = n - 1$ alebo $s = n$). Potom algoritmus počíta s "vnorenými" Householderovými maticami $\tilde{H}_1, \tilde{H}_2, \ldots, \tilde{H}_s$ tak, aby:

$\tilde{H}_s\, \tilde{H}_{s-1} \cdots \tilde{H}_2\, \tilde{H}_1\, A = Q^T A = (R^T, 0^T)^T$.

Takže: $\tilde{H}_1$ nuluje prvý stĺpec matice A pod a_{11}, vznikne $A^{(1)}$.
$\tilde{H}_2$ nuluje druhý stĺpec matice $A^{(1)}$ pod $a_{22}^{(1)}$, vznikne $A^{(2)}$.
Atď, až $\tilde{H}_s$ nuluje stĺpec s matice $A^{(s-1)}$ pod $a_{ss}^{(s-1)}$, vznikne $(R^T, 0^T)^T$.

QRF pomocou HT - 2/2

- Matica $\tilde{H}_k$, $1 \le k \le s$, je vnorená HT: $\tilde{H}_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & H_k \end{pmatrix}$, kde $H_k = I_{m-k+1} - 2u_{m-k+1}u_{m-k+1}^T / \|u_{m-k+1}\|^2$.
- Vnorenie zaručuje, že pri násobení $\tilde{H}_k A^{(k-1)}$ zostane prvých $k-1$ riadkov a prvých $k-1$ stĺpcov matice A bez zmeny. V praxi sa nikdy neformujú $\tilde{H}_k$ ani H_k explicitne - pracuje sa iba s vektormi u_{m-k+1}.
- **Spätná stabilita QRF(HT)** (Wilkinson, 1965):
  Nech R je vypočítaný R-faktor pomocou HT v pohyblivej rádovej čiarke s jednotkou zaokrúhlenia μ. Potom existuje presne ortogonálna matica $\hat{Q}$ taká, že:

  $$A + E = \hat{Q} \begin{pmatrix} R \\ 0 \end{pmatrix}, \quad \|E\|_{\mathrm{F}} \le \Phi(m)\, \|A\|_{\mathrm{F}}\, \mu,$$

  kde $\Phi(m)$ je pomaly rastúca funkcia m.

QRF pomocou GT

Postup je rovnaký ako pri HT (aj tu je $s = \min\{m-1, n\}$):

1. Nájdeme OG maticu:
   $Q_1 = G(1, m, \theta)\, G(1, m-1, \theta) \cdots G(1, 3, \theta) G(1, 2, \theta)$ tak, aby $A^{(1)} = Q_1\, A$ mala nuly v 1. stĺpci pod $a_{11}^{(1)}$.
2. $Q_2 = G(2, m, \theta)\, G(2, m-1, \theta) \cdots G(2, 3, \theta)$ tak, aby $A^{(2)} = Q_2\, A^{(1)}$ mala nuly v 2. stĺpci pod $a_{22}^{(2)}$.

s. Atď., až v kroku s nájdeme OG maticu:
   $Q_s = G(s, m, \theta)\, G(s, m-1, \theta) \cdots G(s, s+1, \theta)$ tak, aby $A^{(s)} = Q_s\, A^{(s-1)} = R$ mala nuly v stĺpci s pod $a_{ss}^{(s)}$.

Potom $A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$, kde $R = A^{(s)}$ a $Q^T = Q_s\, Q_{s-1} \cdots Q_2\, Q_1$.

Porovnanie výpočtovej zložitosti pre maticu $A \in \mathbb{R}^{m \times n}$, $m \geq n$:
QR(HT): $n^2\,(m - n/3)$ flops ('floating point operations') - bez explicitnej formácie Q;
QR(GT): $2n^2\,(m - n/3)$ flops - bez Q.

Jednoznačnosť QRF

Veta: Nech $A \in \mathbb{R}^{m \times n}$, $m \geq n$, má lineárne nezávislé stĺpce (t.j. hodnosť A je n). Potom existuje práve jedna matica $Q \in \mathbb{R}^{m \times n}$ s n ortonormálnymi stĺpcami a práve jedna horná trojuholníková matica s kladnými diagonálnymi prvkami tak, že $A = Q\,R$. $\square$

To znamená, že dve QR faktorizácie takejto matice A, vypočítané dvomi rôznymi spôsobmi (napr. HT a GT), sa môžu líšiť v znamienkach riadkov R a príslušných stĺpcov Q.

Cvičenie

1. Naprogramujte v C++ QR faktorizáciu pomocou Householderovej OG transformácie pre Vami zadanú maticu A.
2. Naprogramujte v C++ QR faktorizáciu pomocou Householderovej OG transformácie pre Vami zadanú maticu A.