Лекция 18

Основы матричных вычислений

Рахуба М. В.
07.08.21

Методы решения полной задачи на собственные знач.

1 QR - алгоритм

$$A v^{(i)} = \lambda^{(i)} v^{(i)}, \quad A \in \mathbb{R}^{n \times n}$$

а) Vanilla версия

$$A_1 = A$$

for $k = 1, 2, \dots$

$$A_k = Q_k R_k \quad (QR\text{ - разлож.})$$

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

A_k сходится к блочно верхнетреуг.

Св-во

1) $A_{k+1} = R_k Q_k \overset{R_k = Q_k^T A_k}{=} Q_k^T A_k Q_k = \dots =$

$$= (Q_1 \dots Q_k)^T A (Q_1 \dots Q_k)$$

2) $A^k = (Q_1 R_1)^k = Q_1 \underbrace{R_1 Q_1 R_1 \dots Q_1}_{(R_1 Q_1)^{k-1}} R_1 =$

$$= Q_1 (R_1 Q_1)^{k-1} R_1 = Q_1 A_2^{k-1} R_1 =$$

$$= Q_1 Q_2 A_3^{k-2} R_2 R_1 \dots = (Q_1 \dots Q_k)(R_k \dots R_1)$$

Блочный степенной метод ($p = n$)

$$Y_k = A X_k$$

$$Y_k = X_{k+1} R_{k+1} \quad - \quad \left(\begin{matrix}\text{вычисление}\\ \text{QR-разлож.}\\ \text{матрицы } Y_k\end{matrix}\right)$$

$$A^k = A A^{k-1} = \underbrace{A X_{k-1}}_{Y_{k-1} = X_k R_k} X_{k-1}^T A^{k-1} = X_k R_k X_{k-1}^T A^{k-1} =$$

$$= X_k R_k X_{k-1}^T A \underset{X_{k-2} X_{k-2}^T}{\uparrow} A^{k-2} = X_k R_k \underbrace{X_{k-1}^T X_{k-1}}_{I} R_{k-1} X_{k-2}^T A^{k-2} =$$

$$= X_k R_k \dots R_1 \underset{I}{\underset{\parallel}{X_0^T}} \underbrace{A^0}_{I}$$

$$\Downarrow$$

QR-алгоритм $\Leftrightarrow$ блочно степ. метод. при $X_0 = I$, $A = A^T$

б) Сходимость

$\boxed{\text{Теор}}$ Пусть $\exists\, p$:

1) $\exists\, S$: $A = S \Lambda S^{-1}$ и $\det(S[:p,:p]) \neq 0$:

$$\Lambda = \begin{array}{c c} & \begin{array}{cc} p & q \end{array} \\ \begin{array}{c} p \\ q \end{array} & \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} \end{array}$$

2) $\underbrace{|\lambda_1| \geq \ldots \geq |\lambda_p|}_{\in \lambda(\Lambda_1)} > \underbrace{|\lambda_{p+1}| \geq \ldots \geq |\lambda_{p+q}|}_{\in \lambda(\Lambda_2)} > 0$

Тогда в
$$A_k = \begin{array}{c c} & \begin{array}{cc} p & \end{array} \\ \begin{array}{c} p \\ \end{array} & \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A_{22}^{(k)} \end{bmatrix} \end{array}$$

$$A_{21}^{(k)} \to 0, \quad k \to \infty$$

и
$$\| A_{21}^{(k)} \|_2 \leq C(\varepsilon) \left(\left| \frac{\lambda_{p+1}}{\lambda_p} \right| + \varepsilon \right)^k \quad \forall \varepsilon > 0$$

Если Λ – диаг, то
$$\| A_{21}^{(k)} \|_2 \leq C \left| \frac{\lambda_{p+1}}{\lambda_p} \right|^k$$

$\square\ \lambda_1 > \lambda_2 > \ldots > \lambda_n > 0, \quad A = A^T$ (для простоты)

$$A^k = (Q_1 \ldots Q_k)(R_1 \ldots R_k)$$

$$A^k = (Q \Lambda Q^T)^k = Q \Lambda^k Q^T \equiv Q D \Lambda^k D^T Q^T, \quad D = \begin{pmatrix} \pm 1 & & \\ & \ddots & \\ & & \pm 1 \end{pmatrix}$$

$$Q^T = LU, \quad u_{kk} > 0$$

$$Q \Lambda^k L U = (Q_1 \ldots Q_k)(R_1 \ldots R_k)$$

$$Q \Lambda^k L = (Q_1 \ldots Q_k)(R_1 \ldots R_k) U^{-1}$$

$$Q \underbrace{\Lambda^k L \Lambda^{-k}}_{\begin{pmatrix} 1 & \ddots & 0 \\ l_{ij}\left(\frac{\lambda_i}{\lambda_j}\right)^k & & 1 \end{pmatrix} \to I} = \underbrace{(Q_1 \ldots Q_k)}_{\text{ортог.}} \underbrace{(R_1 \ldots R_k) U^{-1} \Lambda^{-k}}_{\text{верхнетреуг}}$$

$$Q = UR$$

$$\Downarrow$$

$$Q_1 \ldots Q_k \to Q$$

$$(Q_1 \ldots Q_k)^T A (Q_1 \ldots Q_k) \to Q^T A Q = \Lambda$$

$\blacksquare$

в) Сдвиги:

Выбрать s: $\left|\frac{\lambda_{p+1} - s}{\lambda_p - s}\right| \ll \left|\frac{\lambda_{p+1}}{\lambda_p}\right|$

$A_1 = A$

for $k = 1, 2, \ldots$

$A_k - s_k I = Q_k R_k$

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

•) Сначала $s_k = A_k[-1, -1]$. Когда $\|A_k[-1, :]\|$ мала, то выбираем $s_k = A_k[-2, -2]$

•) На практике используют мультисдвиги

$f(A) = (A_k - s_k^{(1)} I)(A_k - s_k^{(2)} I) \ldots$ - характ. полином $q \times q$ блока $A_{22}^{(k)}$

2) Уменьшение сложности

QR: $O(n^3)$ => сложность $O(n^4)$

как уменьшить?

$UAU^T = H$ — верхн. хесс.

$$U_1 \begin{pmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix} U_1^T$$

$$U_1 = \begin{pmatrix} 1 & \\ & I - 2uu^* \end{pmatrix}$$

$\to 0$

Запуск. QR-алгоритм для H. Тогда каждая итерация $O(n^2)$ (верхнехесс. структ. сохраняется)

Если $A = A^T \Rightarrow H$ – трехдиаг. $\Rightarrow O(n)$

② Алгоритм для SVD

поиск SVD можно свести к диагонализ.

A^TA или AA^T или $\begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix}$ (семинар)

1) Для поиска нескольких с.ч. можно запустить Ланцоша (или другой итерац.) для $\begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix}$

scipy.sparse.linalg.svds

2) Полное SVD

$$UAV = B = \begin{bmatrix} \ddots & \ddots & 0 \\ 0 & & \ddots \end{bmatrix}$$ – бидиагонал. Голуба-Кахана

$$U_1 \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_1 \end{bmatrix}$$

$$O(mn \min(m,n))$$

Запускается QR алгоритм для $B^T B = T$

Если $\underset{\| \, B^T B}{T} - sI = QR$

$$\underset{\| \, B_1^T B_1}{T_1} = RQ + sI$$

B_1 можно получить явно из B. Если считать через $B^T B$, то потеря точности

3 Теория возмущений

$\boxed{\text{Опр}}$ A – строгое строч. диаг. преобл.

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|$$

Аналогич. столбцовое

Теор (Леви – Деспланка)

матрица, обладающ. строгим строчным (столбц.) диаг. преобл. явл. невырожд.

если $\|S\| < 1$, то $I-S$ – невырожд
(Теор с лекции про ряд Неймана)

□
$$A = \mathrm{diag}(A)\big(I + \overbrace{\mathrm{diag}(A)^{-1}(A - \mathrm{diag}(A))}\big)$$

$$\|\mathrm{diag}(A)^{-1}(A - \mathrm{diag}(A))\|_\infty < 1$$

$$\| \quad \max_i \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}| \big/ |a_{ii}| < 1$$

∎

Теор (1-я теорема Гершгорина)

Пусть $A \in \mathbb{C}^{n\times n}$. Тогда собств. знач. находятся внутри

$$D = D_1 \cup \ldots \cup D_n$$

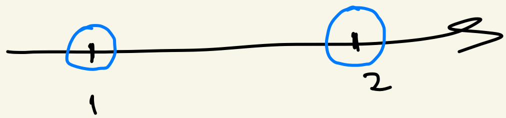
$$D_k = \{ z \in \mathbb{C} : |a_{kk} - z| \le \sum_{i\neq k} |a_{ki}| \} \text{ –}$$

Круги Гершгорина

□ $\lambda \notin D \Rightarrow A - \lambda I$ – строгое строчное диаг. преобл.

$\Rightarrow$ $A - \lambda I$ — невырожд. $\blacksquare$

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad A + E = \begin{pmatrix} 1 & \varepsilon_1 \\ \varepsilon_2 & 2 \end{pmatrix}$$



$$Q \underbrace{\Lambda^k L \Lambda^{-k}}_{\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ l_{ij}\left(\frac{\lambda_i}{\lambda_j}\right)^k & & 1 \end{pmatrix} \to I} = \underbrace{(Q_1 \ldots Q_k)}_{\text{ортог.}} \underbrace{(R_1 \ldots R_k) U^{-1} \Lambda^{-k}}_{\text{верхнетреуг}}$$

$$Q = UR$$

$$Q_1 \ldots Q_k \to Q$$

$$(Q_1 \ldots Q_k)^T A (Q_1 \ldots Q_k) \to Q^T A Q = \Lambda \quad \blacksquare$$

$$(Q_1 \ldots Q_k)^T Q (\Lambda^k L \Lambda^{-k}) = (R_1 \ldots R_k) U^{-1} \Lambda^{-k}$$

$$(\Lambda^k L \Lambda^{-k})^T (\Lambda^k L \Lambda^{-k}) =$$

$$= \underbrace{\left[(R_1 \cdots R_k) M^{-1} \Lambda^{-k} \right]^T}_{M_k} \left[(R_1 \cdots R_k) M^{-1} \Lambda^{-k} \right]$$

$$M_k^T M_k \Rightarrow I = M^T M$$

$$\begin{pmatrix} m_{11}^{(k)} & 0 & \cdots & 0 \\ m_{21}^{(k)} & m_{22}^{(k)} & & \\ & & \ddots & \end{pmatrix} \begin{pmatrix} m_{11}^{(k)} & m_{21}^{(k)} & \cdots & - \\ & \ddots & & \\ 0 & & & - \end{pmatrix}$$