Лекция 17

Основы матричных вычислений

Рахуба М. В.
31.05.21

Методы решения частичной задачи на собств. значения

Частичное EVP:
(eigenvalue problem)

$$A v^{(i)} = \lambda^{(i)} v^{(i)}$$

1 EVP как задача оптимизации

$$\lambda^{(1)} \geq \ldots \geq \lambda^{(n)}$$

$\boxed{\text{Утв.}}$ Пусть $A = A^T$. Тогда

// $R(x)$ – отношение Релея

1) $$\min \lambda^{(i)} = \min_{x \neq 0} \frac{(Ax, x)}{(x, x)}$$

$$\max \lambda^{(i)} = \max_{x \neq 0} \frac{(Ax, x)}{(x, x)}$$

2) $$\lambda^{(k)} = \min_{x \perp v^{(n)}, \ldots, v^{(k+1)}} \frac{(Ax, x)}{(x, x)}, \quad k > 1$$

$$\lambda^{(k)} = \max_{x \perp v^{(1)}, \ldots, v^{(k-1)}} \frac{(Ax, x)}{(x, x)}$$

$\square$ $A = U \Lambda U^T$, $x = Uc = c_1 u_1 + \ldots + c_n u_n$

1)
$$R(x) = R(Uc) = \frac{(U\Lambda \overbrace{U^T U}^{I} c,\ Uc)}{(Uc,\ Uc)} = \frac{(\Lambda c, c)}{(c, c)} =$$

$$= \frac{\lambda^{(1)} c_1^2 + \ldots + \lambda^{(n)} c_n^2}{c_1^2 + \ldots + c_n^2} \leq \lambda^{(1)}, \quad c = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\geq \lambda^{(n)}, \quad c = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

2) $x \perp v^{(n)}, \ldots, v^{(k+1)} \Rightarrow c_n = \ldots = c_{k+1} = 0$

$$R(x) = \frac{\lambda^{(1)} c_1^2 + \ldots + \lambda^{(k)} c_k}{c_1^2 + \ldots + \lambda^{(k)} c_k} \geq \lambda^{(k)}, \quad c = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k$$

$\blacksquare$

$\boxed{\text{Теор}}$ (Куранта - Фишера)

$$\lambda^{(k)} = \max_{\dim(L) = k} \ \min_{\substack{x \in L \\ x \neq 0}} \frac{(Ax, x)}{(x, x)} =$$

$$= \min_{\dim(L) = n-k+1} \ \max_{\substack{x \in L \\ x \neq 0}} \frac{(Ax, x)}{(x, x)}$$

без док-ва.

$\boxed{\text{Замечание}}$

$$R(v^{(i)}) = \frac{(Av^{(i)}, v^{(i)})}{(v^{(i)}, v^{(i)})} = \lambda^{(i)}$$

2 Степенной метод (power iteration)

$$x_{k+1} = \frac{Ax_k}{\|Ax_k\|_2},$$

(если необходимо, то вычисляем $R(x_k) = (Ax_k, x_k)$ на каждой итерации)

Т.е.

$$y_k = Ax_k$$
$$x_{k+1} = y_k / \|y_k\|_2$$
$$\lambda_k = (y_k, x_k)$$

а) Сходимость

Пусть A – диагонализ., и макс. по модулю с.з. – простое

$\Downarrow$

$\exists$ n л.н.з. $v^{(i)}$ и $|\lambda^{(1)}| > |\lambda^{(2)}| \geq \ldots \geq |\lambda^{(n)}|$

предполагаем

$$x_0 = \underset{\neq 0}{\alpha_1} v^{(1)} + \dots + \alpha_n v^{(n)}$$

$$x_k = \frac{A x_{k-1}}{\|A x_{k-1}\|_2} = \dots = \frac{A^k x_0}{\|A^k x_0\|_2} = \frac{\alpha_1 (\lambda^{(1)})^k v^{(1)} + \dots + \alpha_n (\lambda^{(n)})^k v^{(n)}}{\|\alpha_1 (\lambda^{(1)})^k v^{(1)} + \dots + \alpha_n (\lambda^{(n)})^k v^{(n)}\|_2} =$$

$$= \left(\frac{\lambda^{(1)}}{|\lambda^{(1)}|}\right)^k \frac{v^{(1)} + \frac{\alpha_2}{\alpha_1}\left(\frac{\lambda^{(2)}}{\lambda^{(1)}}\right)^k v^{(2)} + \dots + \frac{\alpha_n}{\alpha_1}\left(\frac{\lambda^{(n)}}{\lambda^{(1)}}\right)^k v^{(n)}}{\| v^{(1)} + \dots \|_2} =$$

$$= e^{i\varphi_k} \frac{v^{(1)} + O\left(\left(\frac{\lambda^{(2)}}{\lambda^{(1)}}\right)^k\right)}{\underbrace{\| v^{(1)} + O\left(\left(\frac{\lambda^{(2)}}{\lambda^{(1)}}\right)^k\right)\|_2}_{= 1 + O\left(\left(\frac{\lambda^{(2)}}{\lambda^{(1)}}\right)^k\right)}} =$$

$$= e^{i\varphi_k} \left(v^{(1)} + O\left(\left(\frac{\lambda^{(2)}}{\lambda^{(1)}}\right)^k\right)\right)$$

$$R(x_k) = (A x_k, x_k) = \lambda^{(1)} + O\left(\left(\frac{\lambda^{(2)}}{\lambda^{(1)}}\right)^k\right)$$

Замечание $\Rightarrow$ $A^i x$ становятся близки к лин. завис. (при больших i) в $\mathcal{K}_k(A, x)$ и базис $\{x, Ax, \dots, A^{k-1}x\}$ не подходит для вычислений.

б) Блочная версия

$X_0 \in \mathbb{R}^{n \times p}$, $p << n$

$$Y_k = A X_k$$

$$Y_k = X_{k+1} R_{k+1} \text{ — } QR \text{ разлож. } Y_k$$

$L_k = \mathrm{im}(X_k)$ — сходится к инвар. подпр. A $\quad O\left(\left|\frac{\lambda^{(p+1)}}{\lambda^{(p)}}\right|^k\right)$

в) Обратная итерация

$$A \to A^{-1} \qquad \left(\begin{array}{l} A v^{(i)} = \lambda^{(i)} v^{(i)} \\ \frac{1}{\lambda^{(i)}} v^{(i)} = A^{-1} v^{(i)} \end{array}\right)$$

$$x_{k+1} = \frac{A^{-1} x_k}{\|A^{-1} x_k\|_2}$$

$$R(x_k) \xrightarrow{k \to \infty} \frac{1}{\min\limits_i |\lambda^{(i)}|}$$

$$x_{k+1} = \frac{(A - \sigma I)^{-1} x_k}{\|(A - \sigma I)^{-1} x_k\|_2},$$ σ выбир. близким к желаемому с.з.

Помогает также ускорить сходимость:

$\sigma \approx \lambda^{(1)}$, сходимость $O\left(\left|\frac{\lambda^{(1)}-\sigma}{\lambda^{(2)}-\sigma}\right|^{k}\right)$ $\ll 1$

2) Итерация Релея

$\sigma = \sigma_k = R(x_k)$ – адаптивные сдвиги

$|e_{k+1}| \leq c \cdot |e_k|^{\gamma}$

$\gamma = 3$ при $A = A^T$

$\gamma = 2$ иначе

сверхлинейная сходимость

(для $\sigma = const$ было $\gamma = 1$)

Замечание

1) $(A - \sigma I)$ – почти вырождена, если σ близко к с.з.,
Но правая часть x_k в $(A-\sigma I)y = x_k$ не произвольная.
Систему можно решать не точно (inexact inverse iteration)

2) Итерацию Релея надо начинать с хорошего начального приближения. Иначе может сойтись не туда.

3 Метод Релея – Ритца

Пусть есть подпр-во $V_k \subseteq \mathbb{R}^n$, $\dim(V_k) = k$, в котором мы хотим приблизить с.в. наилучш. образом. Пусть Q_k – ортонорм. базис в V_k.

$\boxed{\text{Теорема}}$ Пусть $Q_k \in \mathbb{R}^{n \times k}$: $Q_k^T Q_k = I$, тогда

$$\underbrace{Q_k^T A Q_k}_{A_k} = \underset{R \in \mathbb{R}^{k \times k}}{\arg\min} \|A Q_k - Q_k R\|_F$$

$\square$ $Q_k^T (A Q_k - Q_k A_k) = \overbrace{Q_k^T A Q_k}^{A_k} - A_k = 0$

$$\|A Q_k - Q_k \overset{A_k + Z}{\overset{\|}{R}}\|_F^2 = \|A Q_k - Q_k A_k - Q_k Z\|_F^2 =$$

$$= \|A Q_k - Q_k A_k\|_F^2 + \|Q_k Z\|_F^2 \to \min_Z$$

$$\Rightarrow \quad Z = 0 \qquad \blacksquare$$

Пусть $A = A^T$.

$$\min_R \|AQ_k - Q_k R\|_F =$$

$$= \|AQ_k - Q_k \overset{S\Lambda S^T}{\overset{\parallel}{A_k}}\|_F = \|AQ_k - Q_k S \Lambda S^T\|_F =$$

$$= \|\underbrace{AQ_k S}_{Z_k} - \underbrace{Q_k S}_{Z_k} \Lambda\|_F = \|AZ_k - Z_k\Lambda\|_F$$

$$\Downarrow$$

Метод Релея - Ритца:

1) $A_k = Q_k^T A Q_k \in \mathbb{R}^{k \times k}$

2) $A_k = S \Lambda S^T$ - собств. разл. A_k

$\Lambda = \operatorname{diag}(\Theta_1, \dots, \Theta_k)$ числа Ритца - прибл. с.з. A

3) $Z_k = Q_k S$

$\parallel$

$[z^{(1)}, \dots, z^{(k)}]$ - векторы Ритца - прибл. с.в. A

4 Методы Ланцоша и Арнольди

В степенном методе считали $A^k x_0$. Почему не попытаться приближ. в пр-ве Крылова $\mathcal{K}_k(A, x_0)$?

Q_k – ортог. базис в $\mathcal{K}_k(A, x_0)$

Из соотн. Арнольди ($AQ_k = Q_{k+1}\hat{H}_{k+1}$)

$$A_k = Q_k^T A Q_k$$

$A = A^T$: трёхдиаг. $\Rightarrow$ **метод Ланцоша**

scipy.sparse.linalg.eigsh

произв.: верхне хесс. $\Rightarrow$ **метод Арнольди**

scipy.sparse.linalg.eigs

LOBPCG (locally optimal block preconditioned CG) – если есть предобусловлив.

PRIMME – библиотека, автоматич. подбирающая солвер к задаче.