

Основы матричных вычислений, Экзамен (Теория)

Версия от 27.06.2021 13:06

Содержание

Связь прямой и обратной ошибок через число обусловленности.	2
Критерий сходимости ряда Неймана.	3
Существование и единственность LU и LDL разложений.	4
Теорема о сходимости градиентного спуска для линейной системы с симметричной положительно определенной матрицей.	5
Оценка сходимости метода сопряженных градиентов для линейной системы с произвольной симметричной положительно определенной матрицей. Случай $\lambda_1 \gg \lambda_2$.	6
Сходимость степенного метода для диагонализуемых матриц.	8
Вывод двух основных свойств QR алгоритма.	9
Теорема Леви-Деспланка и первая теорема Гершгорина.	10
Теорема Леви-Деспланка.	10
Теорема Гершгорина.	10

Связь прямой и обратной ошибок через число обусловленности.

Критерий сходимости ряда Неймана.

Существование и единственность LU и LDL разложений.

Теорема о сходимости градиентного спуска для линейной системы с симметричной положительно определенной матрицей.

Оценка сходимости метода сопряженных градиентов для линейной системы с произвольной симметричной положительно определенной матрицей. Случай $\lambda_1 \gg \lambda_2$.

Предложение. Для $A = A^T > 0$ и любого многочлена $h(\lambda) : h(0) = 1$ степени k верно:

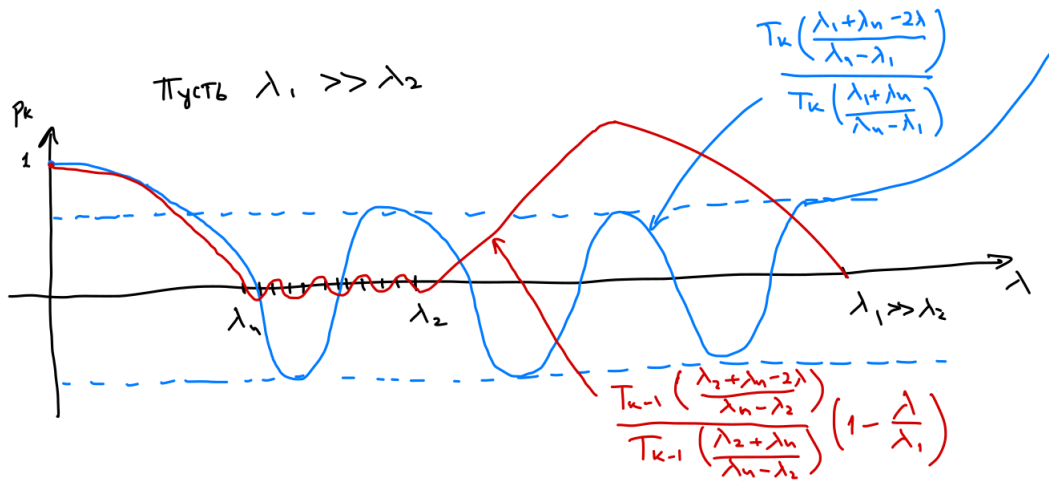
$$\|e_k\|_A \leq \max_i |h(\lambda_i)| \|e_0\|_A$$

В общем случае мы использовали многочлен вида:

$$t_k(\lambda) = \frac{T_k\left(\frac{\lambda_1 + \lambda_n - 2\lambda}{\lambda_1 - \lambda_n}\right)}{T_k\left(\frac{\lambda_1 + \lambda_n}{\lambda_1 - \lambda_n}\right)}$$

Где T_k - многочлен Чебышева. Такой t_k меньше всего отклоняется от 0 на отрезке $[\lambda_n, \lambda_1]$, однако чем больше отрезок тем больше мы отклоняемся. Вообще говоря нас интересует отклонение только в точках $\lambda \in \{\lambda_1, \lambda_2, \dots, \lambda_n\}$. Давайте рассмотрим другой многочлен:

$$p_k(\lambda) = \frac{T_{k-1}\left(\frac{\lambda_2 + \lambda_n - 2\lambda}{\lambda_2 - \lambda_n}\right)}{T_{k-1}\left(\frac{\lambda_2 + \lambda_n}{\lambda_2 - \lambda_n}\right)} \left(1 - \frac{\lambda}{\lambda_1}\right)$$



(Заметка: на лекции путаница с знаком знаменателя аргумента T_i)

То есть мы уменьшили отрезок, а также пожертвовав степенью многочлена Чебышева, мы добавили множитель который обращается в 0 при $\lambda = \lambda_1$

Нам нужно оценить $\max_i |p_k(\lambda_i)|$, заметим, что:

$$\max_i |p_k(\lambda_i)| \leq \max_{\lambda \in [\lambda_n, \lambda_2] \cup \{\lambda_1\}} |p_k(\lambda)| = \max_{\lambda \in [\lambda_n, \lambda_2]} |p_k(\lambda)|$$

Предложение. С лекции 14 нам известно:

$$\frac{T_k\left(\frac{a+b-2\lambda}{a-b}\right)}{T_k\left(\frac{a+b}{a-b}\right)} \leq 2 \left(\frac{\sqrt{\frac{a}{b}} - 1}{\sqrt{\frac{a}{b}} + 1}\right)^k$$

Теперь оценим p_k на множестве $[\lambda_n, \lambda_2]$:

$$p_k(\lambda) = \frac{T_{k-1}\left(\frac{\lambda_2 + \lambda_n - 2\lambda}{\lambda_2 - \lambda_n}\right)}{T_{k-1}\left(\frac{\lambda_2 + \lambda_n}{\lambda_2 - \lambda_n}\right)} \left(1 - \frac{\lambda}{\lambda_1}\right) \leq 2 \left(1 - \frac{\lambda}{\lambda_1}\right) \left(\frac{\sqrt{\frac{\lambda_2}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_2}{\lambda_n}} + 1}\right)^{k-1} \leq 2 \left(1 - \frac{\lambda_n}{\lambda_1}\right) \left(\frac{\sqrt{\frac{\lambda_2}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_2}{\lambda_n}} + 1}\right)^{k-1}$$

Осталось заметить, что $\left|1 - \frac{\lambda_n}{\lambda_1}\right| \leq 1$, тогда:

$$p_k(\lambda) \leq 2 \left(1 - \frac{\lambda_n}{\lambda_1}\right) \left(\frac{\sqrt{\frac{\lambda_2}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_2}{\lambda_n}} + 1}\right)^{k-1} \leq 2 \left(\frac{\sqrt{\frac{\lambda_2}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_2}{\lambda_n}} + 1}\right)^{k-1}$$

Заметим, что:

$$0 \leq 2 \left(\frac{\sqrt{\frac{\lambda_2}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_2}{\lambda_n}} + 1} \right)^{k-1} \implies |p_k(\lambda)| \leq 2 \left(\frac{\sqrt{\frac{\lambda_2}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_2}{\lambda_n}} + 1} \right)^{k-1}$$

Используя первое предложение можно сделать вывод:

$$\|e_k\|_A \leq 2 \left(\frac{\sqrt{\frac{\lambda_2}{\lambda_n}} - 1}{\sqrt{\frac{\lambda_2}{\lambda_n}} + 1} \right)^{k-1} \|e_0\|_A$$

Сходимость степенного метода для диагонализуемых матриц.

Вывод двух основных свойств QR алгоритма.

Теорема Леви-Деспланка и первая теорема Гершгорина.

Теорема Леви-Деспланка.

Определение. Матрица A обладает строгим строчным диагональным преобладанием, если $\forall i |a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}|$.

Определение. Матрица A обладает строгим столбцовым диагональным преобладанием, если $\forall j |a_{jj}| > \sum_{i=1, i \neq j}^n |a_{ij}|$.

Теорема (Леви-Деспланка). Матрица, обладающая строгим строчным (столбцовым) диагональным преобладанием является невырожденной.

Доказательство.

Докажем для строгого строчного, для столбцового аналогично.

Представим A в следующем виде: $A = \text{diag}(A)(I + \text{diag}(A)^{-1})(A - \text{diag}(A))$, где матрица $\text{diag}(A)$ – это диагональная матрица, у которой на диагонали стоят диагональные элементы матрицы A . Раскрыв скобки, можно проверить, что равенство действительно выполняется.

Вспомним немного из курса линала:

- Матрица A – обратима \Leftrightarrow матрица A – невырожденна
- Если $A = BC$, то $\det(A) = \det(B) \cdot \det(C)$

Таким образом, нам необходимо и достаточно доказать обратимость матриц $\text{diag}(A)$ и $(I + \text{diag}(A)^{-1})(A - \text{diag}(A))$.

Обратимость первой почти очевидна: если бы на диагонали могли стоять нулевые элементы, то матрица A не обладала бы строгим строчным диагональным преобладанием.

Теперь заметим, что $(I + \text{diag}(A)^{-1})(A - \text{diag}(A))$ обратима $\Leftrightarrow \exists \sum_{k=0}^{\infty} (-\text{diag}(A)^{-1})(A - \text{diag}(A))^k$ (вспоминаем про ряды Неймана).

Осталось доказать, что такой ряд Неймана сходится. Не будем использовать критерий, а используем признак: докажем, что $\|\text{diag}(A)^{-1}(A - \text{diag}(A))\| < 1$ для некоторой нормы.

Рассмотрим бесконечную норму: $\|\text{diag}(A)^{-1}(A - \text{diag}(A))\|_{\infty} < 1$. Пусть это неравенство не выполняется, тогда: $\max_i \frac{\sum_{j=1, j \neq i}^n |a_{ij}|}{|a_{ii}|} \geq 1$ (это я просто руками записала бесконечную норму для матрицы). Но тогда выходит, что A не обладает строгим строчным диагональным преобладанием – противоречие $\Rightarrow \|\text{diag}(A)^{-1}(A - \text{diag}(A))\|_{\infty} < 1 \Rightarrow \sum_{k=0}^{\infty} (-\text{diag}(A)^{-1})(A - \text{diag}(A))^k$ – сходится $\Rightarrow (I + \text{diag}(A)^{-1})(A - \text{diag}(A))$ обратима.

Таким образом, A представляет собой произведение обратимых матриц $\Rightarrow A$ и сама обратима, то есть, невырожденна. ■

Теорема Гершгорина.

Теорема (1-я теорема Гершгорина). Пусть $A \in \mathbb{C}$, тогда собственные значения матрицы A находятся внутри

$$D = D_1 \cup D_2 \cup \dots \cup D_n$$

где

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

Доказательство.

Пусть $\lambda \notin D \Rightarrow A - \lambda I$ обладает строгим строчным диагональным преобладанием (просто посмотрите на то как мы определяем D_k , на то, что условие в D_k для данной λ не выполняются, и на строение матрицы A) \Rightarrow (по предыдущей теореме) $A - \lambda I$ – невырожденна, но тогда λ не может являться собственным значением A (вспоминаем курс линала: характеристический многочлен и его корни). Получили противоречие. ■

Определение. Матрица D_k – круги Гершгорина.

Теорема (2-я теорема Гершгорина без доказательства). Если есть m кругов Гершгорина, образующих область G , то в G находится ровно m собственных значений.