Содержание

Общая информация

Курс «Вычислительная математика»

Александр Викторович Чикиткин
chikitkin.av@mipt.ru

Кафедра вычислительной физики (603 КПМ)

Лаборатория математического моделирования нелинейных процессов в газовых средах (606 КПМ)

Вычислительная математика.
Введение, обзор курса.

МФТИ

О чём предмет?

Вычислительная математика изучает алгоритмы решения задач непрерывной математики

- **Алгоритм** – последовательность элементарных операций для получения **числового ответа**.
- **Непрерывная математика** – решаемые задачи содержат непрерывные переменные ($\mathbb{R}^n$, $\mathbb{C}^n$)

Примеры задач

- Вычисление стандартных функций:
$$\sqrt{x}, sin(x), 1/x$$
- $f'(x)$, $\int_a^b f(x)\,dx$
- Решение нелинейных уравнений: $f(x) = 0$
- Решение линейных систем $Ax = b$
- Вычисление собственных чисел: $Av = \lambda v$

Примеры задач

- Поиск экстремума функционала: $F(\vec{x}) \to \min$
- Решение нелинейных ОДУ:

$$u'(t) = f(t, u)$$

- Решение уравнений в частных производных:

$$\Delta u(x, y) = f(x, y)$$

Области приложения методов вычислительной математики

- Математическое моделирование
  - Физика
  - Химия
  - Биология
  - Экономика
  - ...
- Математическая статистика
- Анализ данных, машинное обучение
- Обработка сигналов и изображений
- ...

Обзор курса

1. Классификация погрешностей, машинная арифметика
2. Вычислительная линейная алгебра
3. Приближение функций
4. Численное дифференцирование, численное интегрирование
5. Методы решения нелинейных уравнений
6. Методы решения ОДУ

Зачем изучать численные методы?

Многие методы уже реализованы,
есть готовые программы,
зачем это изучать?

Вычисление производной

$$f'(x) = \lim_{dx \to 0} \frac{f(x+dx) - f(x)}{dx}$$

$$\approx \frac{f(x+dx) - f(x)}{dx} = f'(x) + \frac{1}{2} f''(\xi) dx$$

Плохая матрица

- $Ax = b$
- Решаем стандартным методом (методом Гаусса)

1. $A(i,j) = \dfrac{1}{i+j-1}, i,j = 1,\ldots,n$
2. $A(i,j)$ – случайная равномерно распределенная величина из $[0,1]$

Неверная интерполяция

- Дано:
  $x_0, \ldots, x_n$
  $y_0, \ldots, y_n$
- $\exists!$ многочлен степени $\leq n$ $p_n : p_n(x_k) = y_k, k = \overline{0, n}$

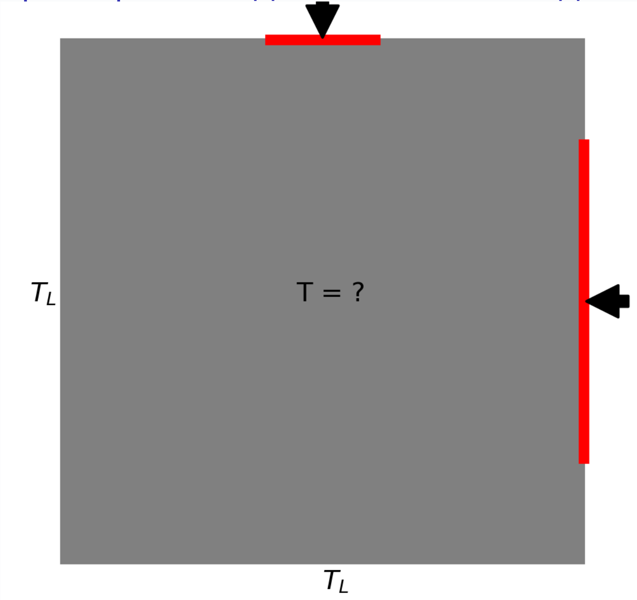
Развитие вычислительных технологий

Современные возможности

- Высокая производительность компьютеров
- Языки программирования высокого уровня
- Готовые реализации основных методов

Можно самостоятельно и быстро
решать сложные задачи!

Пример исследовательской задачи



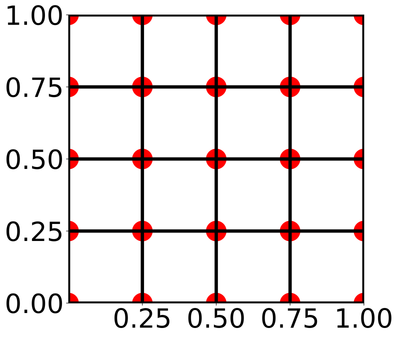
Математическая постановка

Математическая модель (уравнение Лапласа):

$$
\begin{aligned}
&\Delta u = u_{xx} + u_{yy} = 0,\ (x, y) \in [0, 1] \times [0, 1] \\
&u(0, y) = T_L, \\
&u(x, 0) = T_L \\
&u(1, y) = ..., \\
&u(x, 1) = ...
\end{aligned}
$$

Численный метод (1)

1. Введем сетку:



$$x_i = ih,\ \ y_j = jh, i = \overline{0, N}, j = \overline{0, N}, h = 1/N$$

Неизвестные значения

$$u_{ij} \approx u(x_i, y_j), i \neq 0, N,\ \ j \neq 0, N$$

Численный метод (2)

2. Заменим производные конечными разностями:

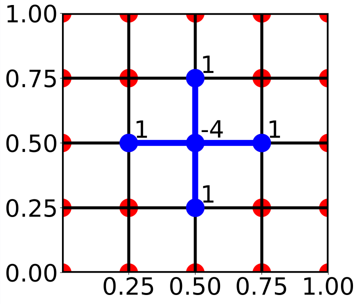
$$
\begin{aligned}
f''(x) \approx & \frac{f'(x+\frac{h}{2}) - f'(x-\frac{h}{2})}{h} \approx \\
& \approx \frac{1}{h}\left(\frac{f(x+h)-f(x)}{h} - \frac{f(x)-f(x-h)}{h}\right) = \\
& = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}
\end{aligned}
$$

Численный метод (3)

3. Приближенные уравнения в каждом внутреннем узле $i, j = \overline{1, N-1}$:

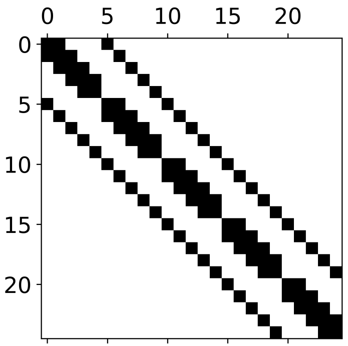
$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h^2} = 0$$

$$u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1} - 4u_{i,j} = 0$$



Численный метод, система линейных уравнений

$$U = \begin{bmatrix} u_{11} \\ u_{21} \\ \vdots \\ u_{N-1,1} \\ u_{12} \\ \vdots \\ u_{N-1,N-1} \end{bmatrix}, \; AU = b$$



Ненулевые элементы матрицы A

Численный метод, система линейных уравнений

- Уравнения образуют систему с матрицей $(N-1)^2 \times (N-1)^2$
- Матрица *разреженная* и имеет особую структуру
- Метод Гаусса $O(n^3) = O(N^6)$
- При $N = 10^3$: 1 год счета на ноутбуке

Итерационный метод

- Последовательность приближенных решений:

$$u_{ij}^{k+1} = (u_{i+1,j}^k + u_{i-1,j}^k + u_{i,j+1}^k + u_{i,j-1}^k)/4$$

  k – номер итерации (приближения)

- Начальное приближение $u_{ij}^0 = 0$
- Критерий остановки: $|u_{ij}^{k+1} - u_{ij}^k| < \epsilon$ (плохой критерий!)

Программная реализация

Особенности метода

- Дискретизация: переход от непрерывных функций/операторов к дискретным
- Большие системы линейных уравнений с матрицами особого вида
- Перекрестное влияние ошибок (дискретизация + итерации)

Типы погрешностей

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} = f'(x) + \frac{h}{2} f''(\xi)$$

- Ошибки входных данных:

$$f(x+h) \to f(x+h) + \Delta_1, f(x) \to f(x) + \Delta_2$$

- Ошибки метода: $E = \dfrac{h}{2} f''(\xi)$
- Ошибки машинной арифметики
  - $x, f(x+h), f(x), h$ округляются
  - Вычитание и деление вычисляются неточно

Машинная арифметика. Числа с плавающей точкой

- Число называется *нормализованным*, если старший разряд мантиссы $\neq 0$.
  - $0.10101_2 \times 2^3$ нормализовано
  - $0.010101_2 \times 2^4$ не нормализовано

Относительная ошибка округления

- $\left|\dfrac{a - a_M}{a_M}\right| \le ?$
- Округление *до ближайшего машинного числа*
- $a = 0.1\ldots d_p d_{p+1}\ldots \times 2^0$
- $|a - a_M| \le \dfrac{1}{2}2^{-p} = 2^{-p-1}$, p – длина мантиссы
- $\left|\dfrac{a - a_M}{a_M}\right| \le \dfrac{2^{-p-1}}{2^{-1}} \le 2^{-p} = \epsilon$

Верхнюю границу относительной ошибки ϵ называют *машинным эпсилон*.

IEEE-стандарт. Одинарная точность

Одинарная точность

1	**8**	**23**
знак	показатель	мантисса

- $a = (-1)^s \times 2^{e-127} \times (1+f)$
- $\epsilon = 2^{-24} \approx 6 \times 10^{-8}$
- Диапазон чисел $[10^{-38}, 10^{38}]$

IEEE-стандарт. Двойная точность

Двойная точность

1	11	52
знак	показатель	мантисса

- $a = (-1)^s \times 2^{e-1023} \times (1+f)$
- $\epsilon = 2^{-53} \approx \times 10^{-16}$
- Диапазон чисел $[10^{-308}, 10^{308}]$

Анализ влияния ошибок округления

- Обозначим $*$ одну из операций $+, -, \times, /$
- $fl(a * b)$ – результат округления $a * b$

Если $a * b$ не выходит за пределы области допустимых показателей, то

$$fl(a * b) = a * b(1 + \delta), \quad |\delta| \leq \epsilon$$

- IEEE-арифметика обеспечивает также

$$fl(\sqrt{a}) = \sqrt{a}(1 + \delta), |\delta| \leq \epsilon$$

Пример. Вычисление скалярного произведения

```
s = 0.
for i = 1,n
    s = s + x[i] * y[i]
```

- $s_1 = 0 + x_1 y_1 (1 + \delta_1) = x_1 y_1 (1 + \delta)$
- $s_2 = (s_1 + x_2 y_2 (1 + \delta_2))(1 + \delta_3) = x_1 y_1 (1 + \delta)^2 + x_2 y_2 (1 + \delta)^2$
- $s_M = (x_1 y_1 + x_2 y_2)(1 + \delta)^n + x_3 y_3 (1 + \delta)^{n-1} + \dots$

Пример. Вычисление скалярного произведения

- При $n\epsilon < 1$:

$$\prod_{i=1}^{n}(1+\delta_i) \le (1+\epsilon)^n = 1 + n\epsilon + O(\epsilon^2)$$

$$\prod_{i=1}^{n}(1+\delta_i) \ge (1-\epsilon)^n \ge 1 - n\epsilon$$

- $s_M = \sum_{i=1}^{n}(1+\Delta_i)x_i y_i$
  где с точностью до $O(\epsilon^2)$ $|\Delta_i| \le n\epsilon$
- $|s_M - s| = |\sum_{i=1}^{n} \Delta_i x_i y_i| \le n\epsilon\langle|x|, |y|\rangle$

Пример. Вычисление производной

$$f'(x) \approx \tilde{f}' = \frac{f(x+h) - f(x)}{h}$$

$$\tilde{f}'_M = \frac{f(x+h)(1+\delta_1) - f(x)(1+\delta_2)}{h} =$$

$$= \frac{f(x+h) - f(x)}{h} + \frac{f(x+h)\delta_1 - f(x)\delta_2}{h}$$

$$|\tilde{f}'_M - f'(x)| \le \underbrace{\left|\frac{f(x+h) - f(x)}{h} - f'(x)\right|}_{\text{ошибка метода}} +$$

$$\underbrace{\left|\frac{f(x+h)\delta_1 - f(x)\delta_2}{h}\right|}_{\text{ошибка округления}}$$

Вычисление производной, ошибки

$$\left|\frac{f(x+h)-f(x)}{h}-f'(x)\right| =$$

$$\left|\frac{f(x)+f'(x)h+f''(\xi)h^2/2-f(x)}{h}-f'(x)\right| =$$

$$= |f''(\xi)\frac{h}{2}| \le \frac{M_2 h}{2}, \quad M_2 = \max_x |f''(x)|$$

$$\left|\frac{f(x+h)\delta_1 - f(x)\delta_2}{h}\right| \le \frac{2M_0\epsilon}{h}, \quad M_0 = \max_x |f(x)|$$

$$E(h) = \frac{M_2 h}{2} + \frac{2M_0\epsilon}{h} \to \min$$

$$E'(h) = \frac{M_2}{2} - \frac{2M_0\epsilon}{h^2} = 0 \Rightarrow h^* = \left(\frac{4M_0\epsilon}{M_2}\right)^{1/2}$$

Заключение

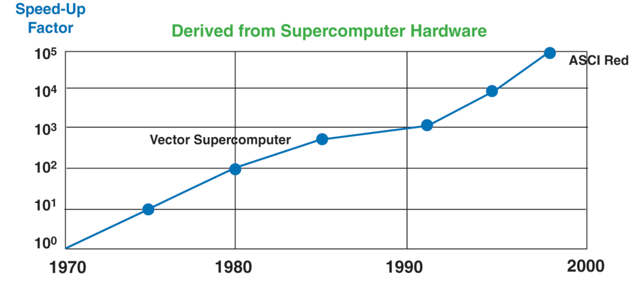
- Задачи, решаемые методами вычислительной математики
- Особенности вычислительных методов
- Типы ошибок
- Машинная арифметика
- Исследование влияния ошибок округления

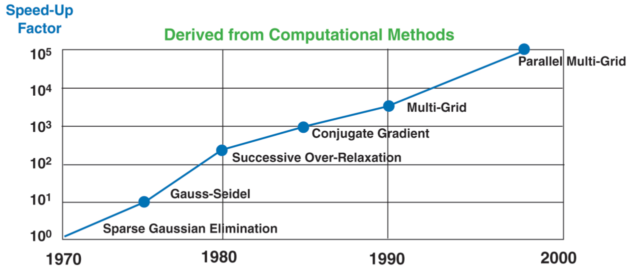
Вычислительная математика.

Нормы векторов и матриц, сингулярное разложение.

МФТИ

Развитие методов вычислительной линейной алгебры





10 важнейших алгоритмов 20-го века

Список, опубликованный в журнале *Computing in Science and Engineering* в 2001 году:

- Вычислительные методы, основанные на разложениях матриц
- QR алгоритм для вычисления собственных значений
- Итерационные методы подпространств Крылова
- Быстрое преобразование Фурье

Векторные нормированные пространства

$Ax = b$, y – приближенное решение. Как оценить ошибку $E(x, y)$?

$x \in V$, $\|x\| : V \to \mathbb{R}$ называется *нормой*, если

1. $\|x\| \geq 0 \ \forall x \in V$; $\|x\| = 0 \Leftrightarrow x = 0$
2. $\|\alpha x\| = |\alpha|\|x\|$, $\forall \alpha \in \mathbb{R}$ (или $\mathbb{C}$)
3. $\|x + y\| \leq \|x\| + \|y\|$ (неравенство треугольника)

Определение (полное пространство)

Векторное пространство называется *полным* или *банаховым*, если любая фундаментальная последовательность в нём сходится.

p - нормы

- В $\mathbb{C}^n$ (или $\mathbb{R}^n$) $\forall p > 1$ нормой является функция

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

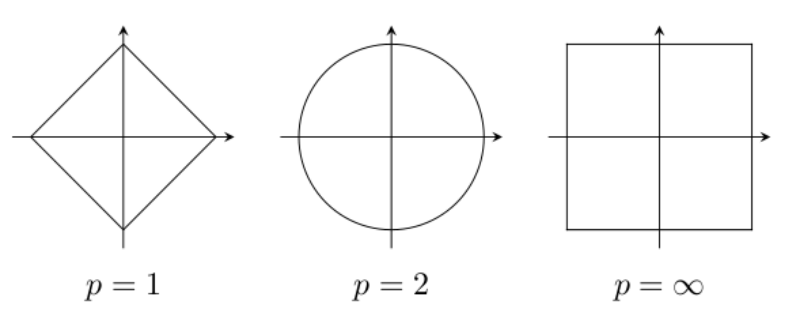
- К p-нормам относят также

$$||x||_\infty = \lim_{p\to\infty} ||x||_p = \max_i |x_i|, \quad \|x\|_1 = \sum_i |x_i|$$

- Неравенство Гёльдера: $\frac{1}{p} + \frac{1}{q} = 1,\ p, q \in [1, \infty]$

$$\sum_{i=1}^{n} |x_i y_i| \le \|x\|_p \|y\|_q, \quad \|y\|_q = \max_{x\neq 0} \frac{|x^T y|}{\|x\|_p}$$

p - нормы, единичные сферы



Эквивалентность норм

Определение (эквивалентные нормы)

Нормы $||\cdot||_{(1)}$ и $||\cdot||_{(2)}$ называются эквивалентными, если $\exists c_1, c_2$:

$$\forall x \in V \; c_1||x||_{(1)} \leq ||x||_{(2)} \leq c_2||x||_{(1)}$$

Теорема об эквивалентности норм

Любые две нормы в конечномерном векторном пространстве эквивалентны.

Скалярное произведение

Пусть на векторном пространстве V для каждой пары векторов x, y определено число (x, y). $(\cdot, \cdot)$ называется *скалярным произведением*, если

1. $(x, x) \geq 0,\ \forall x,\ (x, x) = 0 \Leftrightarrow x = 0$
2. $(x, y) = \overline{(y, x)}$
3. $(\alpha x, y) = \alpha(x, y)$, α - число
4. $(x + y, z) = (x, z) + (y, z)$

- Стандартное с.п. в $\mathbb{C}^n$:

$$(x, y) = \sum_k x_k \overline{y}_k,\ \|x\|_2 = (x, x)^{1/2} = \left(\sum_k |x_k|^2\right)^{1/2}$$

- $|(x, y)| \leq (x, x)^{1/2}(y, y)^{1/2}$ (неравенство К-Б-Ш)

Основные обозначения (1)

Вектор-столбец: $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

Матрица $A = \left[\begin{array}{c|c|c} a_1 & \ldots & a_n \end{array}\right] = \left[\begin{array}{c} s_1^T \\ \hline \ldots \\ \hline s_m^T \end{array}\right]$

$$b = \left[\begin{array}{c|c|c} a_1 & \ldots & a_n \end{array}\right] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_1 \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_n \end{bmatrix}$$

Основные обозначения (2)

$$c^T = [x_1, \ldots, x_m] \begin{bmatrix} s_1^T \\ \hline \ldots \\ \hline s_m^T \end{bmatrix} = x_1 [\quad s_1^T \quad] + \ldots + x_m [\quad s_m^T \quad]$$

$$A^* = \overline{A^T}\ ,\ (x, y) = \sum_k x_k \overline{y_k} = x^T \overline{y} = y^* x \underset{\mathbb{R}^n}{=} y^T x = x^T y$$

$$AB = \begin{bmatrix} s_1^T \\ \hline \ldots \\ \hline s_n^T \end{bmatrix} \left[\begin{array}{c|c|c} b_1 & \ldots & b_n \end{array}\right] = \begin{bmatrix} s_1^T b_1 & & s_1^T b_n \\ & \ddots & \\ s_n^T b_1 & & s_n^T b_n \end{bmatrix}$$

$$AB = \left[\begin{array}{c|c|c} a_1 & \ldots & a_n \end{array}\right] \begin{bmatrix} q_1^T \\ \hline \ldots \\ \hline q_n^T \end{bmatrix} = a_1 q_1^T + \ldots + a_n q_n^T$$

Основные обозначения (3)

Собственные числа и векторы ($A \in \mathbb{C}^{n\times n}$):

$$Au_k = \lambda_k u_k,\ \det(A - \lambda I) = 0$$
$$U = [u_1| \dots |u_n], \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n) \Rightarrow AU = U\Lambda$$

Блочное умножение матриц:

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right], B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] \Rightarrow$$
$$C = AB = \left[\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array}\right]$$

Матричные нормы

- $\mathbb{C}^{m\times n}$ – векторное пространство $\Rightarrow$ можно взять любую векторную норму
- *Матричная норма* удовлетворяет дополнительному свойству

Определение

Пусть каждой матрице A поставлено в соответствие число $||A||$. $||\cdot||$ называется матричной нормой, если:

1. на $\mathbb{C}^{m\times n}$ $||\cdot||$ является векторной нормой
2. для любых матриц A и B допускающих умножение выполняется: $||AB|| \leq ||A||\ ||B||$ (субмультипликативность)

Пример: норма Фробениуса

$$||A||_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

Доказательство субмультипликативности:

$$A = [a_1|, \dots, |a_n] \in \mathbb{C}^{m\times n}, \;\; B = \begin{bmatrix} b_1^T \\ \hline \dots \\ \hline b_n^T \end{bmatrix} \quad AB = \sum_{i=1}^{n} a_i b_i^T$$

$$\|AB\|_F = \| \sum_{i=1}^{n} a_i b_i^T \|_F \leq \sum_i \|a_i b_i^T\|_F = \sum_i \left(\sum_{k,l} |a_{ki} b_{il}|^2 \right)^{1/2}$$

$$= \sum_i \|a_i\|_2 \|b_i\|_2 \leq \left(\sum_i \|a_i\|_2^2 \right)^{1/2} \left(\sum_i \|b_i\|_2^2 \right)^{1/2} = \|A\|_F \|B\|_F$$

Пример: нарушение свойства субмультипликативности

- $\|A\| = \max\limits_{i,j} |a_{ij}|$
- $$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \; AB = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$
$$\|AB\| = 2 > \|A\|\|B\| = 1 \cdot 1 = 1$$

Операторные нормы

- Линейный оператор – отображение $P: V_1 \to V_2$

Определение (норма оператора)

Нормой оператора $P: V_1 \to V_2$ называется

$$||P|| = \sup_{x \neq 0} \frac{||Px||_{(2)}}{||x||_{(1)}} = \sup_{\|x\|_{(1)}=1} ||Px||_{(2)}$$

- $||A|| = \max\limits_{x \neq 0} \dfrac{||Ax||_{(2)}}{||x||_{(1)}} = \max\limits_{\|x\|_{(1)}=1} ||Ax||_{(2)}$

$$\|AB\| = \max_{x \neq 0, Bx \neq 0} \frac{\|ABx\|_{(3)}}{\|x\|_{(1)}} = \max_{x \neq 0, Bx \neq 0} \frac{\|ABx\|_{(3)}}{\|Bx\|_{(2)}} \frac{\|Bx\|_{(2)}}{\|x\|_{(1)}}$$

$$\leq \max_{x \neq 0, Bx \neq 0} \frac{\|A(Bx)\|_{(3)}}{\|Bx\|_{(2)}} \max_{x \neq 0} \frac{\|Bx\|_{(2)}}{\|x\|_{(1)}} \leq \|A\| \|B\|$$

Матричные нормы, порожденные p-нормами

- $||A||_1 = \max\limits_{1 \le j \le n} \sum\limits_{i=1}^{m} |a_{ij}|$
  (максимальная сумма модулей по столбцу)
- $||A||_\infty = \max\limits_{1 \le i \le m} \sum\limits_{j=1}^{n} |a_{ij}|$
  (максимальная сумма модулей по строке)
- $||A||_2 = \max\limits_{k} \sigma_k$
  (максимальное сингулярное число A,
  $\sigma_k = \sqrt{\lambda_k(A^* A)}$).

Изометричные матрицы

Определение

Матрица $Q \in \mathbb{C}^{n\times n}$ *изометрична* относительно нормы $\|\cdot\|$ на $\mathbb{C}^n$, если $\|Qx\| = \|x\|\ \forall x \in \mathbb{C}^n$

- $p, q \in [1, \infty] \quad 1/p + 1/q = 1.$
- $\|Qx\|_p = \|x\|_p \Rightarrow p$-норма столбцов равна 1.
- Неравенство Гёльдера $\left|\sum_{i=1}^n x_i y_i\right| \le \|x\|_p \|y\|_q$:

$$\|Q^T y\|_q \underset{\text{нер-во Г.}}{=} \max_{x\neq 0} \frac{|y^T Qx|}{\|x\|_p} = \max_{x\neq 0} \frac{|y^T Qx|}{\|Qx\|_p} =$$

$$= \max_{z\neq 0} \frac{|y^T z|}{\|z\|_p} \underset{\text{нер-во Г.}}{=} \|y\|_q$$

q-норма строк Q равна 1.

Изометричные матрицы, $p = 2$

- Пусть Q (и Q^T) сохраняют 2-норму:

$$\|Q^T\overline{q}_k\|_2^2 = \|\overline{q}_k\|_2^2 = \|q_k\|_2^2 = 1 = \left\| \begin{bmatrix} q_1^T\overline{q_k} \\ \cdots \\ q_n^T\overline{q_k} \end{bmatrix} \right\|_2^2 =$$

$$= \sum_{i=1}^{n} |(q_k, q_i)|^2 = 1 + \sum_{i=1, i\neq k}^{n} |(q_k, q_i)|^2$$

- $\Rightarrow (q_i, q_j) = \delta_{ij} \Leftrightarrow Q^*Q = I$
- аналогично $QQ^* = I$, $Q^{-1} = Q^*$ (Q^T в $\mathbb{R}^{n\times n}$)

Определение

Матрица $Q \in \mathbb{C}^{n\times n}$: $Q^{-1} = Q^*$ называется *унитарной*.

Унитарные матрицы

- Только унитарные матрицы сохраняют 2-норму
- Обратная к унитарной матрице – унитарная матрица:

$$Q^{-1} = Q^* \Rightarrow (Q^{-1})^{-1} = (Q^*)^{-1} = (Q^{-1})^*$$

- Произведение унитарных матриц - унитарная матрица

$$(Q_1 Q_2)^{-1} = Q_2^{-1} Q_1^{-1} = Q_2^* Q_1^* = (Q_1 Q_2)^*$$

Теорема Шура

Теорема Шура (разложение Шура)

$\forall\ A \in \mathbb{C}^{n\times n}$ с собственными значениями $\lambda_1, \dots, \lambda_n$
$\exists$ унитарная Q такая, что:

$$Q^*AQ = T = \begin{bmatrix} \lambda_1 & \times & \cdots & \times \\ 0 & \lambda_2 & \dots & \times \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}, \quad A = QTQ^*$$

Доказательство теоремы Шура

1. Пусть $Av_1 = \lambda_1 v_1$, $\|v_1\|_2 = 1$
2. Дополним v_1 до ортонормированного базиса:
   $V_1 = [v_1|v_2|\dots|v_n]$ – унитарная.
3. $$V_1^* A V_1 = \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_n^* \end{bmatrix} \left[\lambda v_1 \middle| A v_2 \middle| \dots \right] = \begin{bmatrix} \lambda_1 & \times & \dots & \times \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$
4. $\det(A - \lambda I) = \det(V_1^*)\det(A - \lambda I)\det(V_1) =$
   $\det(V_1^* A V_1 - \lambda I) = (\lambda_1 - \lambda)\det(A_1 - \lambda I_{n-1})$
5. Далее – по индукции для A_1

Нормальные матрицы

Определение (нормальная матрица)

Матрица $A \in \mathbb{C}^{n \times n}$ называется *нормальной*, если $A^*A = AA^*$

Примеры нормальных матриц:

- Эрмитовы (самосопряженные) матрицы: $H^* = H$
- Унитарные матрицы: $U^* = U^{-1}$

Теорема

Матрица $A \in \mathbb{C}^{n \times n}$ нормальная $\iff$ в $\mathbb{C}^n$ $\exists$ ортонормированный базис из ее собственных векторов:

$$AU = U\Lambda, \; U^*AU = \Lambda, \; A = U\Lambda U^*$$

Доказательство теоремы ($\Leftarrow$)

- ($\Leftarrow$) $U = [u_1 | \dots | u_n]$ – матрица с.в.

$$AU = U\Lambda, \;\; U^* = U^{-1} \Rightarrow A = U\Lambda U^*$$

-

$$\begin{aligned}
A^*A &= (U\Lambda U^*)^* \, (U\Lambda U^*) = U\Lambda^* \underline{U^* U}\Lambda U^* \\
&= U\Lambda\Lambda^* U^* = U|\Lambda|^2 U^* \\
AA^* &= (U\Lambda U^*) \, (U\Lambda U^*)^* = U\Lambda \underline{U^* U}\Lambda^* U^* \\
&= U\Lambda^*\Lambda U^* = U|\Lambda|^2 U^* \; \square
\end{aligned}$$

Доказательство теоремы ($\Rightarrow$)

- по теореме Шура $A = UTU^*$
- $A^*A = AA^* \Rightarrow UT^*TU^* = UTT^*U^* \Rightarrow T^*T = TT^*$
- $$\underbrace{\begin{bmatrix} \lambda_1 & t_{12} & \dots & t_{1n} \\ 0 & \ddots & & \\ \vdots & & & \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}}_{T} \quad \underbrace{\begin{bmatrix} \overline{\lambda}_1 & 0 & \dots & 0 \\ \overline{t}_{12} & \ddots & & \\ \vdots & & & \\ \overline{t}_{1n} & 0 & \dots & \overline{\lambda}_n \end{bmatrix}}_{T^*}$$
- $[T^*T]_{11} = |\lambda_1|^2 = [TT^*]_{11} = |\lambda_1|^2 + |t_{12}|^2 + \dots |t_{1n}|^2$
- Внедиагональные элементы равны 0: $T = \Lambda$
- $AU = U\Lambda$ – столбцы унитарной матрицы U являются собственными векторами A $\square$.

Нормальные матрицы (2)

- Степени нормальных матриц:

$$A^2 = U\Lambda U^* U\Lambda U^* = U\Lambda^2 U^*$$

- Нормальная матрица является эрмитовой $\Leftrightarrow$ $\lambda_k \in \mathbb{R},\ k = \overline{1, n}$
- Нормальная матрица является унитарной $\Leftrightarrow$ $|\lambda_k| = 1,\ k = \overline{1, n}$

Знакоопределенные матрицы

- Для любой эрмитовой матрицы $H = H^*$, $\forall x \in \mathbb{C}^n$

$$(Hx, x) = (x, H^*x) = (x, Hx) = \overline{(Hx, x)} \Rightarrow (Hx, x) \in \mathbb{R}$$

Определение (положительно определенная матрица)

Эрмитова матрица A называется *положительно определенной* (*полуопределенной*), если

$$\forall x \in \mathbb{C}^n, x \neq 0: \ (Ax, x) = x^* Ax > 0 \ (\geq 0)$$

- $A > 0 \ (\geq 0) \iff \lambda_k > 0(\geq 0), \ k = \overline{1, n}$

Сингулярное разложение (SVD)(1)

- $\forall A \in C^{m\times n}$ эрмитова матрица $A^*A \geq 0$:
  $(A^*Ax, x) = (Ax, Ax) = \|Ax\|_2^2 \geq 0$
- По теореме о диагонализации нормальной матрицы:
  $\exists V = [v_1|\dots|v_n] \in \mathbb{C}^{n\times n}$:

$$V^*(A^*A)V = diag(\sigma_1^2, \dots, \sigma_n^2),\ \sigma_1 \geq \dots \geq \sigma_n \geq 0$$

- Обозначим r – число ненулевых σ_k, $r \leq \min(n, m)$
  возьмём $V_r = [v_1|, \dots, |v_r]$, $\Sigma_r = diag(\sigma_1, \dots, \sigma_r)$

$$V_r^*A^*AV_r = \Sigma_r^2 \Rightarrow (\Sigma_r^{-1}V_r^*A^*)(AV_r\Sigma_r^{-1}) = I_{r\times r}$$

- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
  $U_r^*U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.

Сингулярное разложение (SVD)(2)

- $U_r = AV_r\Sigma_r^{-1} \in \mathbb{C}^{m\times r}$:
  $U_r^* U_r = I_{r\times r} \Rightarrow$ столбцы U_r - ортонормированные.
- При $1 \le i \le r$: $Av_i = \sigma_i u_i$
- При $i > r$:
  $A^*Av_i = 0 \Rightarrow (v_i^* A^*)(Av_i) = \|Av_i\|_2^2 = 0 \Rightarrow Av_i = 0.$
- Достроим U_r до унитарной матрицы $U \in \mathbb{C}^{m\times m}$:

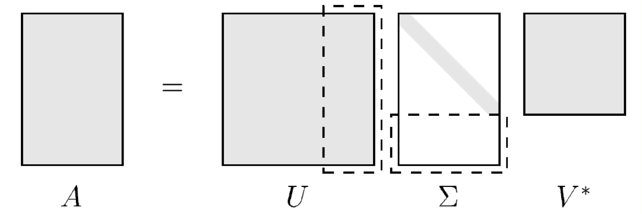
$$AV = U \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow A = U\Sigma_{m\times n}V^*$$

- Т.к. умножение на невырожденные матрицы не меняет ранг, то $rank(A) = rank(\Sigma) = r$

Сингулярное разложение, теорема

Теорема (сингулярное разложение)

$\forall A \in \mathbb{C}^{m\times n}$ ранга r существует разложение:



где $U \in \mathbb{C}^{m\times m}$, $V \in \mathbb{C}^{n\times n}$ - унитарные; $\Sigma \in \mathbb{C}^{m\times n}$:

$$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \\ & & & \end{bmatrix}, \ \sigma_k > 0, k = 1, \dots, r$$

Сингулярное разложение, следствия

- Можно задать матрицу $(m+n)r$ параметрами
- Разделение «переменных» (индексов)

$$A = U\Sigma V^* = \sum_{k=1}^{r} \sigma_k u_k v_k^* \Rightarrow A(i,j) = \sum_{k=1}^{r} \sigma_k u_k(i) v_k^*(j)$$

- $Ax = \sum_{k=1}^{r} \sigma_k u_k (v_k^* x)$ – $O((n+m)r)$ операций
- 4 основные подпространства, связанные с матрицей A:

$$\ker(A) = \mathcal{L}\{v_{r+1}, \ldots, v_n\}$$
$$\operatorname{im}(A) = \mathcal{L}\{u_1, \ldots, u_r\}$$
$$\ker(A^*) = \mathcal{L}\{u_{r+1}, \ldots, u_n\}$$
$$\operatorname{im}(A^*) = \mathcal{L}\{v_1, \ldots, v_r\}$$

SVD и приближение малого ранга

Теорема (Эккарта-Янга)

Пусть $r < \operatorname{rank}(A) = R$, $A_r = U_r \Sigma_r V_r^*$. Тогда

$$\min_{\operatorname{rank}(B)=r} \|A - B\|_2 = \|A - A_r\|_2 = \sigma_{r+1}$$

$$\min_{\operatorname{rank}(B)=r} \|A - B\|_F = \|A - A_r\|_F = \sqrt{\sigma_{r+1}^2 + \dots + \sigma_R^2}$$

Примеры применения SVD

Заключение

- Векторные и матричные нормы
- Унитарные, эрмитовы, нормальные, знакоопределенные матрицы
- Разложение Шура, сингулярное разложение

Вычислительная математика.

Обусловленность системы линейных уравнений, число обусловленности.

МФТИ

Число обусловленности

- Как сильно меняется $f(x)$ при малых изменениях x?
- Пусть f - функция, $f(x+\delta) \approx f(x) + f'\delta$.
  Если $f(x) \neq 0$ и $x \neq 0$:

$$\frac{\|f(x+\delta) - f(x)\|}{\|f(x)\|} \approx \left(\frac{\|f'(x)\|}{\|f(x)\|}\|x\|\right)\frac{\|\delta\|}{\|x\|}$$

- Мерой чувствительности является *число обусловленности* (*condition number*)

$$cond(f(x)) = \frac{\|f'(x)\|}{\|f(x)\|}\|x\|$$

Число обусловленности матрицы

- Чему равно число обусловленности для $f(A) = A^{-1}$?
- Пусть A и $A + \Delta$ - невырожденные матрицы, тогда:

$$(A+\Delta)^{-1} - A^{-1} = \left[I - A^{-1}(A+\Delta)\right](A+\Delta)^{-1} =$$
$$= -A^{-1}\Delta(A+\Delta)^{-1} \approx -A^{-1}\Delta A^{-1} \Rightarrow$$

$$\frac{\|(A+\Delta)^{-1} - A^{-1}\|}{\|A^{-1}\|} \leq\approx \left(\|A^{-1}\|\|A\|\right)\frac{\|\Delta\|}{\|A\|}$$

- Величину $cond(A) = \|A^{-1}\|\|A\|$ называют *числом обусловленности матрицы* A
- при каких Δ матрица $(A + \Delta)$ будет невырожденной?

Матричные ряды

Определение (сходящийся ряд)

Ряд $\sum_{k=0}^{\infty} A_k$, где $A_k \in \mathbb{C}^{n\times n}$ называется сходящимся, если сходится последовательность его частичных сумм.

- $\|B - \sum\limits_{k=0}^{N} A_k\| \to 0, \; N \to \infty$
- Достаточное условие: сходимость ряда $\sum\limits_{k=0}^{\infty} \|A_k\|$:

$$\|\sum_{k=N_1}^{N_2} A_k\| \leq \sum_{k=N_1}^{N_2} \|A_k\| < \epsilon, \; \forall N_1, N_2 > M$$

Степенной ряд

- Ряд $\sum_{k=0}^{\infty} F^k$, $F^0 = I$ называется *рядом Неймана*.
- Достаточное условие сходимости: $\|F\| < 1$.

Определение (спектральный радиус)

Спектральным радиусом матрицы A называется число $\rho(A) = \max_k |\lambda_k(A)|$

- Матрицу $A: \ \rho(A) < 1$ называют *сходящейся* (*convergent*)

Критерий сходимости степенного ряда

Ряд Неймана с матрицей F сходится $\iff F^k \to 0 \iff \rho(F) < 1$ (матрица F – сходящаяся).

Доказательство критерия ($\Rightarrow$)

- Если ряд сходится, то по критерию Коши:

$$\|F^k\| \to 0, \;\; k \to \infty$$

- Пусть $Fu = \lambda u$

$$\|F^k\| \geq \frac{\|F^k u\|}{\|u\|} = \frac{\|\lambda^k u\|}{\|u\|} = |\lambda|^k$$

- если $|\lambda| \geq 1$, то $\|F^k\| \not\to 0$ $\square$

Доказательство критерия ($\Leftarrow$)

- сходимость ряда равносильна сходимости ряда с подобной матрицей $\tilde{F} = S^{-1}FS$:

$$\sum_{k=N_1}^{N_2} \tilde{F}^k = \sum_{k=N_1}^{N_2} S^{-1}F^kS = S^{-1}\left(\sum_{k=N_1}^{N_2} F^k\right)S$$

- По теореме Шура: $T = U^{-1}FU$
- $D_\varepsilon = \operatorname{diag}(1, \varepsilon, \ldots, \varepsilon^{n-1})$, $\tilde{T} = D_\varepsilon^{-1}TD_\varepsilon$: $\tilde{T}_{ij} = \varepsilon^{j-i}t_{ij}$, $i \leq j$
- $\|\tilde{T}\|_1 = \max_j \left(\sum_{i\leq j} |t_{ij}|\varepsilon^{j-i}\right) = \max_j \left(|\lambda_j| + \varepsilon \sum_{i<j} |t_{ij}|\varepsilon^{j-i-1}\right)$
- подбором ε можно сделать $\|\tilde{T}\|_1 < 1$. $\Rightarrow$ сходится ряд для $\tilde{T}$ и для F $\square$.

Сходимость степенного ряда, следствия

- Если $\rho(F) < 1$, то $\exists\ (I - F)^{-1} = \sum_{k=0}^{\infty} F^k$

$$(I - F)\sum_{k=0}^{N} F^k = I - F^{N+1} \to I \ \square$$

- $\|F\| < 1$: $\|(I - F)^{-1}\| \leq \|I\| \sum_{k=0}^{\infty} \|F\|^k = \frac{\|I\|}{1-\|F\|}$

Обратная к возмущенной матрице

Если A - невырожденная и $\|A^{-1}\Delta\| < 1$, то
$A + \Delta = A(I + A^{-1}\Delta)$ – невырожденная и

$$(A + \Delta)^{-1} = \left[\sum_{k=0}^{\infty} (-A^{-1}\Delta)^k\right] A^{-1} = A^{-1}\sum_{k=0}^{\infty}(-\Delta A^{-1})^k$$

$$\frac{\|(A+\Delta)^{-1} - A^{-1}\|}{\|A^{-1}\|} \leq \frac{\|A^{-1}\|\|\Delta\|}{1 - \|A^{-1}\Delta\|} = \boxed{\frac{\|A^{-1}\|\|A\|}{1 - \|A^{-1}\Delta\|}}\frac{\|\Delta\|}{\|A\|}$$

Обусловленность линейной системы (1)

- $Ax = f,\ f \neq 0 \to (A + \Delta A)\tilde{x} = f + \Delta f$
- Как $\Delta x = \tilde{x} - x$ зависит от ΔA и Δf?
- Рассмотрим случай $\Delta A = 0$, $A\Delta x = \Delta f$

$$\left(\frac{\|A^{-1}\Delta f\|}{\|A^{-1}f\|}\right) \Big/ \left(\frac{\|\Delta f\|}{\|f\|}\right) = \frac{\|A^{-1}\Delta f\|}{\|\Delta f\|}\frac{\|f\|}{\|A^{-1}f\|} =$$
$$= \frac{\|A^{-1}\Delta f\|}{\|\Delta f\|}\frac{\|Ax\|}{\|x\|} \leq \|A^{-1}\|\|A\| = cond(A)$$

- переходя к $\leq$, берём max по Δf и x (или f)
- Для фиксированной f:

$$\left(\frac{\|A^{-1}\Delta f\|}{\|A^{-1}f\|}\right) \Big/ \left(\frac{\|\Delta f\|}{\|f\|}\right) \leq \|A^{-1}\|\frac{\|f\|}{\|x\|} = \nu(A, f)$$

Обусловленность линейной системы (2)

- Пусть $\|A^{-1}\Delta A\| < 1$:

$$\Delta x = \tilde{x} - x = (A + \Delta A)^{-1}(f + \Delta f) - A^{-1} f =$$

$$\left[(A + \Delta A)^{-1} - A^{-1}\right] f + (A + \Delta A)^{-1} \Delta f =$$

$$\left[\sum_{k=1}^{\infty} (-A^{-1}\Delta A)^k\right] (A^{-1} f) + \left[\sum_{k=0}^{\infty} (-A^{-1}\Delta A)^k\right] (A^{-1}\Delta f)$$

- Возьмем норму обеих частей

$$\|\Delta x\| \le \frac{\|A^{-1}\Delta A\|}{1 - \|A^{-1}\Delta A\|} \|A^{-1} f\| + \frac{\|A^{-1}\Delta f\|}{1 - \|A^{-1}\Delta A\|} \le$$

$$\|x\| \frac{\|A^{-1}\|\|A\|}{1 - \|A^{-1}\Delta A\|} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta f\|}{\|A\|\|x\|}\right) \Rightarrow$$

$$\boxed{\frac{\|\Delta x\|}{\|x\|} \le \frac{\|A^{-1}\|\|A\|}{1 - \|A^{-1}\Delta A\|} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta f\|}{\|f\|}\right)}$$

Согласованность матрицы и правой части

- В общем случае $\|A\|_2\|A^{-1}\|_2 = \sigma_1/\sigma_n$
- Пусть $A = U\Sigma V^* = \sum_k \sigma_k u_k v_k^*$
- $f = \sum_k a_k u_k$, $\Delta f = \sum_k b_k u_k$
- $x = \sum_k (a_k/\sigma_k) v_k$, $\Delta x = \sum_k (b_k/\sigma_k) v_k$
- Пусть $a_k = b_k = 0$ при $k > r$:

$$\frac{\|\Delta x\|_2}{\|x\|_2} = \frac{\left(\sum\limits_{k=1}^{r} (b_k/\sigma_k)^2\right)^{1/2}}{\left(\sum\limits_{k=1}^{r} (a_k/\sigma_k)^2\right)^{1/2}}, \quad \frac{\|\Delta f\|_2}{\|f\|_2} = \frac{\left(\sum\limits_{k=1}^{r} (b_k)^2\right)^{1/2}}{\left(\sum\limits_{k=1}^{r} (a_k)^2\right)^{1/2}} \Rightarrow$$

$$\boxed{\frac{\|\Delta x\|_2}{\|x\|_2} \le \frac{\sigma_1}{\sigma_r} \frac{\|\Delta f\|_2}{\|f\|_2}}$$

Пример плохо обусловленной матрицы

- $x_0, \ldots, x_n$
- $y_0, \ldots, y_n$
- $p_n : p_n(x_k) = y_k$
- $$\begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \ldots & \ldots & \ddots & \ldots & \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_0 \\ \vdots \\ y_n \end{bmatrix}$$

Обусловленность матрицы Вандермонда

Диагональное преобладание

Диагональное преобладание

Матрица $A \in \mathbb{C}^{n\times n}$ имеет *строчное диагональное преобладание*, если

$$|a_{ii}| > r_i \equiv \sum_{j=1,\ j\neq i}^{n} |a_{ij}|,\ \ i = 1, \ldots, n$$

и *столбцовое диагональное преобладание*, если

$$|a_{jj}| > c_j \equiv \sum_{i=1,\ i\neq j}^{n} |a_{ij}|,\ \ j = 1, \ldots, n$$

Невырожденность матрицы с д.п.

Теорема

Матрица, имеющая диагональное преобладание, является невырожденной.

- $\mathrm{diag}(A) = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ - диагональная часть A
- $\mathrm{off}(A) = A - \mathrm{diag}(A)$ - внедиагональная часть A
- Для строчного д.п. : $\|[\mathrm{diag}(A)]^{-1}\mathrm{off}(A)\|_\infty < 1 \Rightarrow$
- $I + [\mathrm{diag}(A)]^{-1}\mathrm{off}(A)$ - невырожденная
- Домножим на невырожденную матрицу $\mathrm{diag}(A)$:

$$\mathrm{diag}(A)\left(I + [\mathrm{diag}(A)]^{-1}\mathrm{off}(A)\right) = \mathrm{diag}(A) + \mathrm{off}(A) = A\Box$$

Теорема (круги Гершгорина)

Круги Гершгорина

Для $A \in \mathbb{C}^{n \times n}$ рассмотрим круги

$$R_i \equiv \{z \in \mathbb{C} : |a_{ii} - z| \leq r_i\}$$
$$C_i \equiv \{z \in \mathbb{C} : |a_{ii} - z| \leq c_i\}$$
$$r_i \equiv \sum_{j=1,\, j\neq i}^{n} |a_{ij}|, \; c_j \equiv \sum_{i=1,\, i\neq j}^{n} |a_{ij}|, \; i, j = \overline{1, n}$$

Тогда для собственного числа λ матрицы A верно:

$$\lambda \in \bigcup_{i=1}^{n} R_i \text{ и } \lambda \in \bigcup_{i=1}^{n} C_i$$

Доказательство теоремы Гершгорина

- Если $\lambda \notin \bigcup_i R_i$, то

$$|\lambda - a_{ii}| > r_i$$

- $\Rightarrow A - \lambda I$ имеет строчное д. п.
- $\Rightarrow A - \lambda I$ - невырождена – противоречие $\square$.

С помощью кругов Гершгорина можно получить оценку на границы спектра A.

Круги Гершгорина

Вычислительная математика. Прямые методы решения систем линейных уравнений.

МФТИ

Содержание

1. LU-разложение и метод Гаусса
2. Метод Холецкого
3. QR-разложение
4. Метод наименьших квадратов

Метод Гаусса

- Исключение неизвестных:

$$+\left\{\begin{matrix} \times -\frac{a_{21}}{a_{11}} \\ \\ \end{matrix}\right. \quad \left[\begin{array}{ccc} a_{11} & a_{12} & \dots \\ \hline a_{21} & a_{11} & \dots \end{array}\right]$$

- Один шаг исключения – умножение слева на нижнетреугольную матрицу L

$$A_1 = \underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 & \dots \\ & & \ddots & \end{bmatrix}}_{L} \times A$$

LU- разложение

- Основная идея: с помощью операций со строками привести матрицу к верхнетреугольному виду:

$$\underset{A}{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}} \xrightarrow{L_1} \underset{L_1 A}{\begin{bmatrix} \times & \times & \times & \times \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\ \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}} \xrightarrow{L_2}$$

$$\underset{L_2 L_1 A}{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\ & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \end{bmatrix}} \xrightarrow{L_3} \underset{L_3 L_2 L_1 A}{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & \mathbf{0} & \boldsymbol{\times} \end{bmatrix}}$$

- $\underbrace{L_{n-1} \dots L_2 L_1}_{L^{-1}} A = U, \; L = L_1^{-1} L_2^{-1} \dots L_{n-1}^{-1}, \; A = LU$
- Могли столкнуться с делением на 0!

Треугольные матрицы

- $L_1 L_2$ – нижнетреугольная матрица.

$$(L_1 L_2)_{ij} = \sum_{k=1}^{n} l_{ik}^1 l_{k_j}^2 = \sum_{k \le i, k \ge j} l_{ik}^1 l_{k_j}^2 \overset{i<j}{=} 0$$

- L^{-1} – нижнетреугольная матрица.
  По правилу Крамера: $(L^{-1})_{ij} = \frac{\Delta_{ij}}{\Delta}$

$$i < j: \ \Delta_{ij} = \begin{vmatrix} x & & & \\ x & x & & \\ x & x & 0 & \\ x & x & 1 & x \end{vmatrix} = 0$$

Матрицы L_k

$$x_k = \begin{bmatrix} x_{1k} \\ \vdots \\ x_{kk} \\ x_{k+1,k} \\ \vdots \\ x_{nk} \end{bmatrix}, L_k x_k = \begin{bmatrix} x_{1k} \\ \vdots \\ x_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

нужно вычесть из j-й строки k-ю строку, умноженную на

$$l_{jk} = \frac{x_{jk}}{x_{kk}} \quad k < j \le n$$

$$L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nk} & & & 1 \end{bmatrix}$$

Матрицы L_k

$$L_k = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{nk} & & & 1 \end{bmatrix}, \quad l_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ l_{k+1,k} \\ \vdots \\ l_{n,k} \end{bmatrix}$$

- $L_k = I - l_k e_k^T$
- $(I - l_k e_k^T)(I + l_k e_k^T) = I - l_k \cancel{(e_k^T l_k)} e_k^T = I$
- $L_k^{-1} = I + l_k e_k^T$
- $L_k^{-1} L_{k+1}^{-1} = (I + l_k e_k^T)(I + l_{k+1} e_{k+1}^T) =$
  $I + l_k e_k^T + l_{k+1} e_{k+1}^T + l_k \cancel{e_k^T l_{k+1}} e_{k+1}^T = I + l_k e_k^T + l_{k+1} e_{k+1}^T$

Алгоритм LU-разложения без перестановок

$$L = L_1^{-1} \dots L_{n-1}^{-1} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{bmatrix}$$

Алгоритм:

```
U = A
L = I
for k = 1, n-1
    for j = k+1, n
        l[j,k] = u[j,k]/u[k,k]
        u[j,k:n] = u[j,k:n] - l[j,k] * u[k,k:n]
```

Число операций

```
U = A
L = I
for k = 1, n-1
    for j = k+1, n
        l[j,k] = u[j,k]/u[k,k]
        u[j,k:n] = u[j,k:n] - l[j,k] * u[k,k:n]
```

$$\sum_{k=1}^{n-1} \sum_{j=k+1}^{n} 2(n-k+1) = \sum_{k=1}^{n-1} 2(n-k+1)(n-k) \overset{q=n-k}{=}$$

$$\sum_{q=1}^{n-1} 2(q+1)q = 2\sum_{q=1}^{n-1} q^2 + 2\sum_{q=1}^{n-1} q =$$

$$= 2\frac{(n-1)n(2n-1)}{6} + O(n^2) = \frac{2}{3}n^3 + O(n^2)$$

Программа для вычисления LU-разложения без перестановок

Существование LU-разложения

1. Алгоритм построен в предположении, что $u_{kk} \neq 0$
2. $\begin{bmatrix} 0. & 1. \\ 1. & 1. \end{bmatrix}$

Строго регулярная матрица

Матрица называется *строго регулярной*, если все её ведущие подматрицы невырожденные.

***LU*-разложение**

A имеет LU-разложение $\iff$ A строго регулярная.

Свойства LU-разложения

- LU-разложение определяется однозначно:
  Допустим $L_1U_1 = L_2U_2$, тогда
  $L_2^{-1}L_1 = U_2U_1^{-1} = D, \Rightarrow$
  D - диагональная, $D = I$, т.е. $L_1 = L_2$, $U_1 = U_2$ $\square$
- Все ведущие миноры положительны $\Leftrightarrow$ все диагональные элементы $U > 0$

$$A = LU = \begin{bmatrix} \boxed{\begin{matrix} 1 & 0 & \dots \\ x & \ddots & \\ \vdots & & 1 \end{matrix}} & \\ & \ddots \end{bmatrix} \begin{bmatrix} \boxed{\begin{matrix} u_{11} & x & \dots \\ 0 & \ddots & \\ \vdots & & u_{kk} \end{matrix}} & \\ & \ddots \end{bmatrix}$$

$$\det A_{1:k,1:k} = \det L_{1:k,1:k} \det U_{1:k,1:k} = \prod_1^k u_{ii}$$

Свойства LU-разложения

- Если $A = A^*$, то
$$A = LU = L \begin{bmatrix} u_{11} & & \\ & \ddots & \\ & & u_{nn} \end{bmatrix} \tilde{U} = LD\tilde{U} = \tilde{U}^*(D^*L^*)$$
из единственности LU: $\tilde{U}^* = L$, $A = LDL^*$, $u_{ii} \in \mathbb{R}$
- Если $A = A^* > 0$, то существует *разложение Холецкого*:
$$A = LDL^* = (LD^{1/2})(D^{1/2}L^*) = CC^*$$
C - нижнетреугольная, $C_{ii} > 0$.

LU-разложение и ошибки округления

- Пусть t - длина мантиссы

$$A = \begin{bmatrix} 2^{-t} & 1. \\ 1. & 1. \end{bmatrix}, \tilde{L} = \begin{bmatrix} 1. & 0 \\ 2^t & 1. \end{bmatrix} \; \tilde{U} = \begin{bmatrix} 2^{-t} & 1. \\ 0 & \cancel{1} - 2^t \end{bmatrix}$$

$$\tilde{L}\tilde{U} = \begin{bmatrix} 2^{-t} & 1. \\ 1. & 0. \end{bmatrix} = A + \begin{bmatrix} 0 & 0 \\ 0 & -1. \end{bmatrix}$$

- Малая величина диагонального элемента, на который мы делим, приводит к росту элементов в L и U.
- Решение: на каждом шаге выбирать *ведущий элемент* (pivot) из подматрицы: максимальный по модулю элемент в столбце/строке/подматрице.

LU-разложение с выбором ведущего элемента (pivoting)

Выбор по столбцу (partial pivoting):

- выбираем максимальный по модулю элемент в столбце
- меняем его строку с текущей строкой

$$
\begin{bmatrix}
\times & \times & \times & \times \\
0 & \times & \times & \times \\
0 & x_{ik} & \boldsymbol{\times} & \boldsymbol{\times} \\
0 & \times & \times & \times
\end{bmatrix}
\xrightarrow{P}
\begin{bmatrix}
\times & \times & \times & \times \\
0 & x_{ik} & \boldsymbol{\times} & \boldsymbol{\times} \\
0 & \boldsymbol{\times} & \boldsymbol{\times} & \boldsymbol{\times} \\
0 & \times & \times & \times
\end{bmatrix}
$$

$$
\xrightarrow{L}
\begin{bmatrix}
\times & \times & \times & \times \\
0 & x_{ik} & \times & \times \\
0 & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times} \\
0 & \mathbf{0} & \boldsymbol{\times} & \boldsymbol{\times}
\end{bmatrix}
$$

LU-разложение с выбором ведущего элемента (pivoting)

- $L_{n-1}P_{n-1}\cdots L_2P_2L_1P_1A = U$
  P_k - матрица перестановки, $P_k^{-1} = P_k^T$
-

$$L_3P_3L_2P_2L_1P_1 =$$
$$L_3(P_3L_2P_3^{-1})(P_3P_2L_1P_2^{-1}P_3^{-1})P_3P_2P_1 =$$
$$L_3'L_2'L_1'P_3P_2P_1$$

где $L_3' = L_3,\ L_2' = P_3L_2P_3^{-1},\ L_1' = P_3P_2L_1P_2^{-1}P_3^{-1}$

Действие матриц перестановок

$$
P_{k+1} L_k P_{k+1}^{-1} = P_{k+1} \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & x & 1 & & \\ & & \vdots & & \ddots & \\ & & x & & & 1 \end{bmatrix} P_{k+1}^{-1} =
$$

$$
\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \boldsymbol{x} & 0 & \dots & 1 \\ & & \vdots & & & \\ & & \boldsymbol{x} & 1 & \dots & 0 \end{bmatrix} P_{k+1}^{-1} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & x & 1 & & \\ & & \vdots & & \ddots & \\ & & x & & & 1 \end{bmatrix}
$$

LU-разложение с выбором ведущего элемента

- $L'_{n-1} \dots L'_1 P_{n-1} \dots P_1 A = U \Rightarrow PA = LU$
- Выбор главного элемента по столбцу добавляет $O(n^2)$ операций – сложность $\approx \frac{2}{3}n^3$

```
U = A, L = I, P = I
for k = 1, n-1
    i = argmax(|u[k:n,k]|)
    u[k,k:n] <-> u[i,k:n]
    l[k,1:k-1] <-> l[i,1:k-1]
    p[k,:] <-> p[i,:]
    for j = k + 1, n
        l[j,k] = u[j,k]/u[k,k]
        u[j,k:n] = u[j,k:n] - l[j,k] * u[k,k:n]
```

LU-разложение с выбором ведущего элемента

Решение системы с помощью LU разложения

Процесс решения системы $Ax = b$ можно разделить на 3 этапа

1. Вычисление LU разложения $PA = LU$ (далее $A = LU$)
2. Решение системы $Ly = b$ (прямая подстановка)
3. Решение системы $Ux = y$ (обратная подстановка)

Для разных b достаточно вычислить LU-разложение только 1 раз!

Решение системы с треугольными матрицами

Алгоритм прямой подстановки $Ly = b$:

```
y[:] = 0
for k = 1, n
    y[k] = b[k]
    for i = 1, k-1
        y[k] = y[k] - y[i] * l[k,i]
    y[k] = y[k] / l[k,k]
```

С использованием скалярного произведения:

```
y[:] = 0
for k = 1, n
    y[k] = (b[k] - <y[1:k-1], l[k,1:k-1]> / l[k,k]
```

Количество операций: $n^2 + O(n)$

Рост элементов в LU-разложении

- При выборе в столбце возможен рост элементов U:

$$\frac{\max_{i,j}|u_{ij}|}{\max_{i,j}|a_{ij}|} \leq 2^{n-1}$$

  Оценка достигается на специально подобранных матрицах

- На практике сильный рост не встречается!

In fifty years of computing, no matrix problem that excite an explosive instability are known to have arisen under natural circumstances.
Trefethen L. N., Bau III D. Numerical linear algebra. – Siam, 1997

Разложение Холецкого

- $A = A^T > 0$: $A = CC^T$, C – нижнетреугольная с положительной диагональю.
- Для $n = 3$:

$$\begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} c_{11} & 0 & 0 \\ c_{21} & c_{22} & 0 \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ 0 & c_{22} & c_{21} \\ 0 & 0 & c_{33} \end{bmatrix}$$

- Последовательно будем вычислять элементы C:

$$c_{11} = \sqrt{a_{11}} \qquad c_{21} = \frac{a_{21}}{c_{11}} \qquad c_{31} = \frac{a_{31}}{c_{11}}$$

$$c_{22} = \sqrt{a_{22} - c_{21}^2} \qquad c_{32} = \frac{a_{32} - c_{31}c_{21}}{c_{22}}$$

$$c_{33} = \sqrt{a_{33} - c_{31}^2 - c_{32}^2}$$

- Решение линейной системы: $Ax = b$:
$A = CC^T$, $CC^Tx = b \to Cy = b \to C^Tx = y$.

Алгоритм разложения Холецкого

```
C = 0
for k = 1, n
    c[k,k] = (a[k,k] - sum(c[k,1:k-1]^2))^1/2
    for i = k+1, n
        c[i,k] = a[i,k] - sum(c[i,1:k-1] * c[k,1:k-1])
        c[i,k] = c[i,k] / c[k,k]
```

- Количество операций

$$\sum_{k=1}^{n}\left(2k + \sum_{i=k+1}^{n} 2k\right) = \sum_{k=1}^{n}(2k + 2k(n-k)) =$$

$$(n+1)n + n^2(n+1) - 2\frac{n(n+1)(2n+1)}{6} = \frac{1}{3}n^3 + O(n^2)$$

- В методе Холецкого нет роста элементов!

QR-разложение

- Пусть $A \in \mathbb{C}^{m\times n}$, $m \geq n$, $A = [a_1|a_2|\dots|a_n]$
- Применим ортогонализацию Грамма-Шмидта:
  1. $q_1 \equiv a_1/\|a_1\|_2$
  2. $p_2 = a_2 - (a_2, q_1)q_1$, $q_2 = p_2/\|p_2\|_2$ (нормировка)
  3. $p_3 = a_3 - (a_3, q_1)q_1 - (a_3, q_2)q_2, \dots$
- k-й столбец A есть линейная комбинация $q_1, \dots, q_k$.
  В матричном виде:

$$\left[\begin{array}{c|c|c} a_1 & \dots & a_n \end{array}\right] = \underbrace{\left[\begin{array}{c|c|c} q_1 & \dots & q_n \end{array}\right]}_{Q} \underbrace{\begin{bmatrix} r_{11} & \dots & \times \\ 0 & \ddots & \times \\ 0 & \dots & r_{nn} \end{bmatrix}}_{R} = QR$$

Алгоритм Грамма-Шмидта

```
for j = 1, k
    p[:,j] = a[:,j]
    for i = 1, j - 1
        p[:,j] = p[:,j] - q[:,i] * <a[:,j], q[:,i]>
    q[:,j] = p[:,j] / ||p[:,j]||_2
```

- В стандартном алгоритме происходит *потеря ортогональности*
- На практике используют модифицированный алгоритм Грамма-Шмидта:

```
for j = 1, k
    p[:,j] = a[:,j]
    for i = 1, j - 1
        p[:,j] = p[:,j] - q[:,i] * <p[:,j], q[:,i]>
    q[:,j] = p[:,j] / ||p[:,j]||_2
```

Решение линейной системы с помощью QR-разложения

- Можно решить систему $Ax = b$ в 3 шага

$$Q^*(QR)x = Q^*b \Rightarrow Rx = Q^*b = y$$

1. $A = QR$ - вычисляем разложение ($O(n^3)$)
2. вычисляем $y = Q^*b$ ($O(n^2)$)
3. $x = R^{-1}y$ - решаем систему с треугольной матрицей за $O(n^2)$

- В 2 раза больше вычислений, чем LU разложение
- В QR разложении нет проблем, связанных с ростом элементов
- Матрицы Q, R содержат дополнительную информацию

Переопределенные системы, метод наименьших квадратов

- $Ax = b$, $A \in \mathbb{C}^{m \times n}$, $m > n$, $rank(A) = n$
- В общем случае система не имеет решения:
  $Ax = b \iff b \in Lin\{a_1, \ldots, a_n\}$
- Естественное обобщение: $\|r\| = \|Ax - b\| \to \min$
- Введем квадратичный функционал:

$$F(x) = \|Ax - b\|_2^2 = (Ax - b, Ax - b) =$$

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right)^2$$

$$\nabla F(x) = 2(A^* Ax - A^* b) = 0 \iff A^* Ax = A^* b$$

- Систему $(A^* A)x = A^* b$ называют системой *нормальных уравнений*

МНК, геометрическая интерпретация

- Ax - вектор из линейной оболочки столбцов A
- Хотим найти в $Lin\{a_1, ..., a_n\}$ вектор, ближайший к вектору b из $\mathbb{R}^m$, $m > n$.
- Решение получается *ортогональной проекцией* b на $Lin\{a_1, ..., a_n\}$
- $$b = Ax + v, v \perp Lin\{a_1, \ldots, a_n\} \Rightarrow A^* b = A^* Ax + 0$$

Решение переопределенной системы

- Метод Холецкого:
  1. Вычисляем матрицу A^*A и вектор A^*b
  2. Вычисляем разложение Холецкого $A^*A = CC^*$
  3. Решаем 2 системы $Cy = A^*b$, $C^*x = y$
- QR-разложение

$$x = (A^*A)^{-1}A^*b = (R^*Q^*QR)^{-1}R^*Q^*b =$$
$$(R^*R)^{-1}R^*Q^*b = R^{-1}(R^*)^{-1}R^*Q^*b = R^{-1}Q^*b$$

  1. Вычисляем $A = QR$
  2. Вычисляем Q^*b
  3. Решаем систему $Rx = Q^*b$

Пример применения МНК, регрессия

- Даны значения функции $y_1, \dots, y_m$ в точках $x_1, \dots, x_m$
- Хотим приблизить функцию многочленом степени $n-1$, $n \leq m$

$$\begin{bmatrix} 1 & x_1^1 & \cdots & x_1^{n-1} \\ 1 & x_2^1 & \cdots & x_2^{n-1} \\ & \vdots & & \vdots \\ 1 & x_m^1 & \cdots & x_m^{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

- При $n = m$ существует единственное решение

Пример применения МНК, регрессия

Вычислительная математика. Итерационные методы решения систем линейных уравнений.

МФТИ

План

1. Общий вид простых методов
2. Вид $P(x^{k+1} - x^k) = b - Ax^k$
3. Примеры простых методов
4. Общая теорема для $A = A^*$
5. Идея предобуславливания
6. Чебышёвские итерации (минимизация по подпространству)
7. Крыловские методы

Недостатки прямых методов

- Сложность $O(n^3)$ для матриц общего вида
- Нужны элементы матрицы в явном виде

$$F : \mathbb{R}^n \to \mathbb{R}^n;\ J = \frac{\partial F}{\partial x} = [\frac{\partial F_i}{\partial x_j}(x_0)]$$

$$J(x_0)\Delta x = F(x_0 + \Delta x) - F(x_0) + O(\Delta x^2)$$

- Нельзя найти приближение к решению с заданной погрешностью
- Сложно использовать специальную структуру матрицы
  Пример: если A - разреженная, матрицы L и U, Q и R не будут в общем случае разреженными

Потеря разреженности

Общий вид простого метода

$$
\begin{aligned}
Ax = b \iff x &= Sx + f \\
x^k &= Sx^{k-1} + f \\
e^k = (x^k - x) &= S(x^{k-1} - x) = \\
&= S^k(x^0 - x) = S^k e^0
\end{aligned}
$$

- Критерий сходимости:
  $e^k \to 0 \iff S^k e^0 \to 0 \; \forall e^0 \iff \rho(S) < 1$
- Достаточное условие сходимости: $\|S\| = q < 1$

$$\|e^k\| = \|S^k e^0\| \leq \|S\|^k \|e^0\| = q^k \|e^0\|$$

- Нужно задать начальное приближение x^0 и *критерий остановки*, например: $\|Ax^k - b\| \leq \epsilon$

Общий вид с предобуславливателем

- Любой метод можно записать в виде:

  $$P(x^{k+1} - x^k) = b - Ax^k,\ k = 0, 1, \dots,\ P \text{ — невырожденная}$$

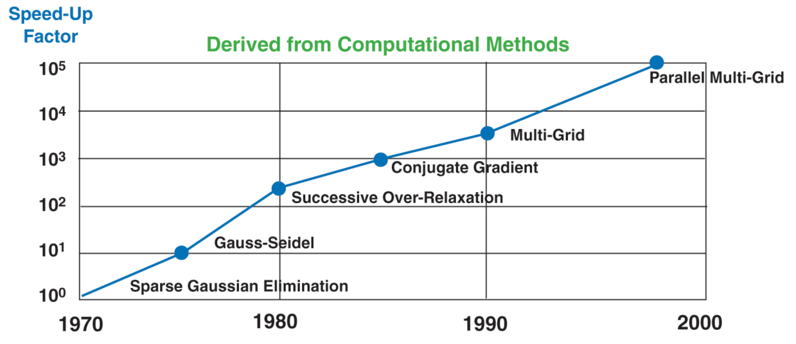
- $x^{k+1} = (I - P^{-1}A)x^k + P^{-1}b, \quad S = (I - P^{-1}A)$
- Требования к P:
  1. $P \approx A^{-1}$
  2. Решение системы $Py = c$ стоит $\ll O(n^3)$
- Матрицу P (иногда P^{-1}) называют *предобуславливатель* (preconditioner)

Преимущества итерационных методов

$$x^k = Sx^{k-1} + f$$

- Сложность $O(Kn^2)$, где K - число итераций.
- $\|e^k\| \le q^k \|e^0\| \le \epsilon \Rightarrow k \ge \dfrac{\log(\epsilon / \|e^0\|)}{\log q}$
- Нужна только процедура для вычисления Sx
- Матрично-векторное умножение легко распараллеливается с высокой эффективностью.
- Для разреженных матриц сложность $O(K \times NNZ(S))$

Развитие итерационных методов



Список 10 важнейших алгоритмов 20-го века, *Computing in Science and Engineering*, 2001 год:

- *итерационные методы* подпространств Крылова

Метод Ричардсона

- Рассмотрим систему $Ax = b$, $A = A^T > 0$:

$$\begin{aligned}
&Ax = b \\
&\tau(Ax - b) = 0 \\
&x - \tau(Ax - b) = x \\
&\underline{x^{k+1} = x^k - \tau(Ax^k - b) = (I - \tau A)x^k + \tau b}
\end{aligned}$$

τ - итерационный параметр. $S = I - \tau A$

- Критерий сходимости:

$$\mathrm{sp}(I - \tau A) = 1 - \tau\ \mathrm{sp}(A),\ \lambda_i(A) > 0$$

$$|1 - \tau\lambda_i| < 1,\ i = 1, \ldots, n \Rightarrow$$

$$(1 - \tau\lambda_i)^2 < 1 \Rightarrow\ \tau\lambda_i(\tau\lambda_i - 2) < 0 \Rightarrow \underline{0 < \tau < \frac{2}{\lambda_{max}}}$$

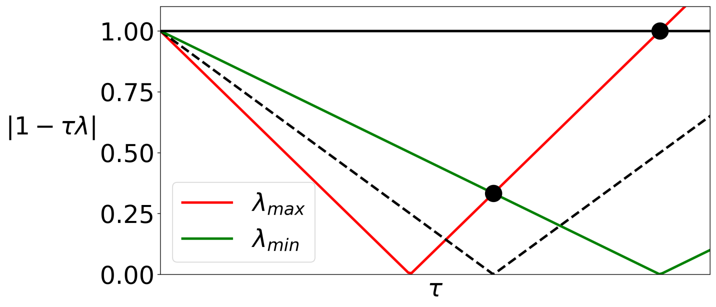
Оптимальный параметр

- $\|e^k\|_2 = \|(I - \tau A)^k e^0\|_2 \leq \|(I - \tau A)\|_2^k \|e^0\|_2$
- $\|(I - \tau A)\|_2 = \max\limits_{\lambda \in sp(A)} |1 - \tau\lambda|$
- Часто, известны только оценки границы спектра:

$$sp(A) \in [m, M]$$

- $\tau^* = \arg\min\limits_{\tau} \max\limits_{\lambda \in [m,M]} |1 - \tau\lambda|$

Оптимальный параметр



Оптимальное значение:

$$1 - \tau m = \tau M - 1 \Rightarrow \tau^* = \frac{2}{M+m} \Rightarrow$$

$$\|(I - \tau^* A)\|_2 = q^* = \frac{M-m}{M+m} = \frac{M/m - 1}{M/m + 1} = \frac{cond_2(A) - 1}{cond_2(A) + 1}$$

Скорость сходимости зависит от обусловленности!

Методы Якоби и Гаусса-Зейделя

- Общая идея: $A = M - N$, M - обратимая и *легко обращается* (система решается за $\ll O(n^3)$):
$Mx = Nx + b \Rightarrow Mx^{k+1} = Nx^k + b,\ x^{k+1} = M^{-1}(Nx^k + b)$
- Легко обращаются:
    - Диагональные матрицы – $O(n)$
    - Треугольные матрицы – $O(n^2)$
- $A = L + D + U$,
L - строго нижнетреугольная часть
D - диагональная часть
U - строго верхнетреугольная часть.
- **Предполагается, что D - невырожденная.**
- Метод Якоби: $Dx^{k+1} = -(L + U)x^k + b$
- Метод Зейделя: $(L + D)x^{k+1} = -Ux^k + b$

Метод Якоби

$$x^{k+1} = -D^{-1}(L+U)x^k + D^{-1}b$$

Достаточное условие сходимости метода Якоби

Если матрица A имеет строчное диагональное преобладание, то метод Якоби сходится.

- докажем, что $\|D^{-1}(L+U)\|_\infty < 1$
- $|a_{ii}| > \sum\limits_{j \neq i} |a_{ij}|, \; i = 1, \ldots, n$ (определение диагонального преобладания) $\Rightarrow$
- $\|D^{-1}(L+U)\|_\infty = \max\limits_i \sum\limits_{j \neq i} \left|\dfrac{a_{ij}}{a_{ii}}\right| = \max\limits_i \dfrac{\sum_{i \neq j} |a_{ij}|}{|a_{ii}|} < 1 \square$

Алгоритм метода Якоби

```
A
b
xkp1[:] = 0 # начальное приближение
xk[:] = 0
while (||dot(A, xkp1) - b|| > tol)
    xk = xkp1
    for i = 1, n
        xkp1[i] = b[i]
        for j = 1, n; j \= i
            xkp1[i] = xkp1[i] - A[i,j] * xk[j]
        xkp1[i] = xkp1[i] / A[i,i]
```

- Метод Якоби редко используется в *чистом виде*
- Acceleration of the Jacobi iterative method by factors exceeding 100 using scheduled relaxation, JCP, 2014

Пример программы для метода Якоби

Метод Зейделя

$$x^{k+1} = -(L+D)^{-1}Ux^k + (L+D)^{-1}b$$

Достаточное условие сходимости метода Зейделя

Если $A \in \mathbb{R}^{n \times n}$, $A = A^T > 0$, то метод Зейделя сходится.

- $A = A^T > 0 \Rightarrow \quad A = L + D + L^T$
- $\langle Ae_k, e_k \rangle = a_{kk} > 0 \Rightarrow D > 0$
- $x^{k+1} = -(L+D)^{-1}L^Tx^k + (L+D)^{-1}b = \boxed{(I - (L+D)^{-1}A)}x^k + (L+D)^{-1}b$
- Можно ввести A-норму: $\|v\|_A^2 \equiv \langle Av, v \rangle = \langle v, v \rangle_A$, где $\langle \cdot, \cdot \rangle$ - обычное скалярное произведение

Теорема о сходимости для $A = A^* > 0$

Метод Зейделя(2)

$$\|e^{k+1}\|_A^2 \equiv \left\langle Ae^{k+1}, e^{k+1} \right\rangle =$$

$$= \left\langle A(I - (L+D)^{-1}A)e^k, (I - (L+D)^{-1}A)e^k \right\rangle =$$

$$= \left\langle Ae^k - A(L+D)^{-1}Ae^k, e^k - (L+D)^{-1}Ae^k \right\rangle \boxed{=}$$

Обозначим $\underline{v = (L+D)^{-1}Ae^k; Ae^k = (L+D)v \; (1)}$

$$\boxed{=} \left\langle Ae^k, e^k \right\rangle - \left\langle Av, e^k \right\rangle - \left\langle Ae^k, v \right\rangle + \left\langle Av, v \right\rangle =$$

$$= \|e^k\|_A^2 - \left(2\left\langle (L+D)v, v \right\rangle - \left\langle (L+D+L^T)v, v \right\rangle \right) =$$

$$= \|e^k\|_A^2 - \left\langle Dv, v \right\rangle \le \underline{\|e^k\|_A^2 - d\|v\|_2^2 \; (2)}, \; d = \min_i D_{ii} = \min_i a_{ii}$$

$$\|e^k\|_A^2 = \left\langle Ae^k, e^k \right\rangle = \left\langle (L+D)v, A^{-1}(L+D)v \right\rangle \le$$

$$\|A^{-1}\|_2 \|(L+D)\|_2^2 \|v\|_2^2 = \lambda_{min}^{-1} \|(L+D)\|_2^2 \|v\|_2^2 \Rightarrow \text{ из } (2):$$

$$\underline{\|e^{k+1}\|_A^2 \le \left(1 - \lambda_{min} d \|L+D\|_2^{-2}\right) \|e^k\|_A^2} \; \square$$

Алгоритм метода Зейделя

$$(L + D)x^{k+1} = -Ux^k + b$$

```
A
b
xkp1[:] = 0 # начальное приближение
xk[:] = 0
while (||dot(A, xkp1) - b|| > tol)
    xk = xkp1
    for i = 1, n
        xkp1[i] = b[i]
        for j = 1, i-1
            xkp1[i] = xkp1[i] - A[i,j] * xkp1[j]
        for j = i+1, n # матрица U
            xkp1[i] = xkp1[i] - A[i,j] * xk[j]
        xkp1[i] = xkp1[i] / A[i,i]
```

Квадратичные функционалы и линейные системы

- Для $A = A^T \in \mathbb{R}^{n \times n}$ рассмотрим:
  $f(x) = \dfrac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle = \dfrac{1}{2}x^T Ax - x^T b$
- Пусть z - решение системы $Az = b$. Если $A > 0$, то:

  $$E(x) \equiv f(x) - f(z) = 0.5\langle A(x-z), x-z\rangle > 0, \forall x \neq z$$

  $\Rightarrow z$ - точка минимума
- Для произвольной A можно ввести *функционал невязки*:
  $$R(x) = \|b - Ax\|_2$$
- Общая идея: $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k r^k$
  $\tau^k = \arg\min\limits_{\tau} F(x^k - \tau r^k),\ F(x) = E(x), R(x)$

Метод наискорейшего спуска

- $F(x) = E(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + f(z)$
- $\nabla E(x^k) = Ax^k - b = r^k$ (градиентный спуск)
- $x^{k+1} = x^k - \tau^k(Ax^k - b) = x^k - \tau^k \nabla f(x^k)$

$$f(x^{k+1}) = \frac{1}{2}\langle A(x^k - \tau r^k), x^k - \tau r^k\rangle - \langle b, x^k - \tau r^k\rangle =$$

$$= \frac{1}{2}\langle Ax^k, x^k\rangle - \underline{\frac{1}{2}\tau\langle Ax^k, r^k\rangle - \frac{1}{2}\tau\langle Ar^k, x^k\rangle} +$$

$$+ \underline{\underline{\frac{1}{2}\tau^2\langle Ar^k, r^k\rangle}} - \langle b, x^k\rangle + \underline{\tau\langle b, r^k\rangle} =$$

$$= \frac{1}{2}\tau^2\langle Ar^k, r^k\rangle - \tau\langle Ax^k - b, r^k\rangle + \ldots =$$

$$= \frac{1}{2}\tau^2\langle Ar^k, r^k\rangle - \tau\langle r^k, r^k\rangle + \ldots = \Phi(\tau)$$

$$\Phi'(\tau) = \tau\langle Ar^k, r^k\rangle - \langle r^k, r^k\rangle = 0 \Rightarrow \tau^k = \frac{\langle r^k, r^k\rangle}{\langle Ar^k, r^k\rangle}$$

Заключение

- Сравнение прямых и итерационных методов
- Метод Ричардсона
- Методы Якоби и Зейделя
- Методы, основанные на минимизации функционала

Чебышёвский набор параметров

- Обобщение метода Ричардсона:
  $x^{k+1} = (I - \tau_k A)x^k + \tau b$
- Можно подобрать $\tau_0, \tau_1, \dots, \tau_{N-1}$ так, чтобы $\|e^N\|$ была минимальна.
- Изменение ошибки:

$$e^N = (I - \tau_{N-1}A)(I - \tau_{N-2}A) \dots (I - \tau_0 A)e^0 = p(A)e^0$$

$$\|p(A)\|_2 \to \min$$

Чебышёвский набор параметров (2)

- $A = A^T > 0 \Rightarrow A$ - нормальная
  $\Rightarrow A = U\Lambda U^*, U^*U = I,$

$$p(A) = (I - \tau_{N-1}A)\dots(I - \tau_0 A) =$$
$$(UIU^* - \tau_{N-1}U\Lambda U^*)\dots(UIU^* - \tau_0 U\Lambda U^*) =$$
$$U\Big[(I - \tau_{N-1}\Lambda)\dots(I - \tau_0\Lambda)\Big]U^*$$
$$\|p(A)\|_2 = \max_{\lambda \in sp(A)} \Big|(1 - \tau_{N-1}\lambda)\dots(1 - \tau_0\lambda)\Big| \leq$$
$$\max_{\lambda \in [m,M]} \Big|(1 - \tau_{N-1}\lambda)\dots(1 - \tau_0\lambda)\Big| = \max_{\lambda \in [m,M]} |p(\lambda)|$$

m, M - границы спектра.

Оптимальный многочлен, условие альтернанса

- Обозначим $\mathcal{P}_N$ семейство многочленов степени N со свойством $q(0) = 1$.

$$p = \arg\min_{q \in \mathcal{P}_N} \max_{\lambda \in [m,M]} |q(\lambda)|$$

Условие альтернанса

Если есть $N+1$ точек $\lambda_k^* \in [m, M]$, $k = \overline{0, N}$, в которых p принимает максимальное по модулю значение $|p(\lambda_k^*)| = \max\limits_{\lambda \in [m,M]} |p(\lambda)|$ с чередующимися знаками, то p имеет наименьшее отклонение от нуля в $\mathcal{P}_N$

Условие альтернанса

Доказательство:

1. Допустим обратное: $\exists q \in \mathcal{P}_N$:
   $\max\limits_{\lambda \in [m,M]} |q(\lambda)| < \max\limits_{\lambda \in [m,M]} |p(\lambda)|$
2. Введем $\Delta = p - q$.
3. В точках λ_k^* многочлен q по модулю меньше, чем p, следовательно Δ имеет тот же знак, что и p.
4. Т.к. в этих точках знаки p чередуются, то на отрезке $[m, M]$ Δ имеет N нулей.
5. Кроме того, $\Delta(0) = p(0) - q(0) = 1 - 1 = 0$. Следовательно, многочлен Δ степени N имеет $N + 1$ ноль, значит $\Delta \equiv 0$ $\square$

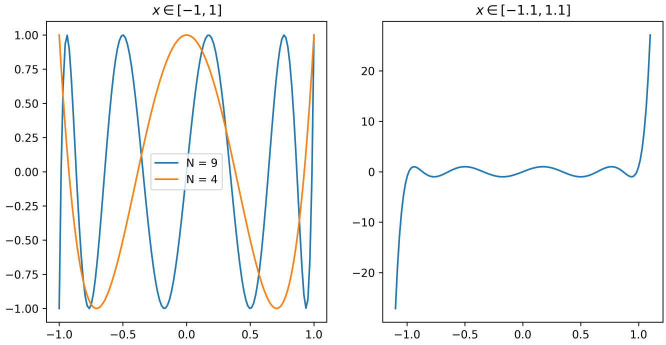
Многочлены Чебышёва

- Многочлен, обладающий условием альтернанса, можно выразить через *многочлены Чебышёва*.
- Многочлены чебышёва можно задать реккурентно:

$$T_{k+1}(t) = 2tT_k(t) - T_{k-1}(t),\ k > 0, \quad P_0(t) = 1,\ P_1(t) = t$$

- Из реккурентного соотношения можно получить явные формулы:
  1. $T_k(t) = \cos(k \arccos t),\ t \in [-1, 1]$
  2. $T_k(t) = \frac{1}{2}\left(t + \sqrt{t^2 - 1}\right)^k + \frac{1}{2}\left(t - \sqrt{t^2 - 1}\right)^k$
- Нули и экстремумы:
  - $t_i = \cos\left(\frac{\pi(2i-1)}{2k}\right), \quad i = 1, \ldots, k$ (корни)
  - $t_i^* = \cos\left(\frac{\pi i}{k}\right),\ i = 0, \ldots, k\ \ T_k(t_j^*) = (-1)^j$ (экстремумы)

Графики



Caption

Оптимальный многочлен

- Введем отображение
  $$\lambda = \frac{M+m}{2} + \frac{M-m}{2}t, \quad t \in [-1,1]$$
- $$p(\lambda) = T_N\left(\frac{\lambda - (M+m)/2}{(M-m)/2}\right) \Bigg/ T_N\left(-\frac{M+m}{M-m}\right)$$
- $p(0) = 1$ и удовлетворяет условию альтернанса
  $$\Rightarrow p = \arg\min_{q \in \mathcal{P}_N} \max_{\lambda \in [m,M]} |q(\lambda)|$$
- Получаем оценку ошибки

$$\|e^{(N)}\|_2 \leq \max_{\lambda \in [m,M]} |p(\lambda)| \|e^{(0)}\|_2 = \frac{1}{\left|T_N\left(-\frac{M+m}{M-m}\right)\right|} \|e^{(0)}\|_2$$

Оценка ошибки

Для оценки $\left| T_N \left(-\frac{M+m}{M-m} \right) \right|$ выразим

$$t_0 = -\frac{M+m}{M-m} = -\frac{1+v}{1-v} \leq -1, \;\; v = \frac{m}{M} \leq 1$$

$$t^2 - 1 = \frac{(1+v)^2 - (1-v)^2}{(1-v)^2} = \frac{4v}{(1-v)^2}$$

$$\sigma = |t| + \sqrt{t^2 - 1} = \frac{1 + v + 2\sqrt{v}}{1-v} = \frac{(1+\sqrt{v})^2}{1-v} = \frac{1+\sqrt{v}}{1-\sqrt{v}}$$

$$2|T_N(t_0)| =$$

$$|0.5(t + \sqrt{t^2-1})^N + 0.5(t - \sqrt{t^2-1})^N| \geq \sigma^N \Rightarrow$$

$$\boxed{\|e^{(N)}\| \leq 2 \left(\frac{1 - \sqrt{m/M}}{1 + \sqrt{m/M}} \right)^N \|e^{(0)}\|_2}$$

Идея Крыловских методов

- Метод наискорейшего спуска: минимизировать функционал в подпространстве $\Delta x^{k+1} = \tau_k r_k$
- Метод с чебышёвскими параметрами: минимизировать функционал в подпространстве:

$$span\{b, Ab, \ldots, A^{N-1}b\}$$

Подпространства Крылова

Подпространства

$$\mathcal{K}_i \equiv \mathcal{K}_i(A, b) \equiv span\{b, Ab, \ldots, A^{i-1}b\}$$

называются *подпространствами Крылова*

- $x^k = \mathrm{argmin}_{x \in \mathcal{K}_k} \|Ax - b\|_2$

Крыловские методы

- $A = A^T > 0$ - метод сопряженных градиентов (conjugate gradient method): направления поиска A-ортогональны $(Ap^k, p^i) = 0, \ k < i$
- $A \neq A^T$ – GMRES (Generilized Minimal Residual)
- Скорость сходимости зависит от обусловленности матрицы.

Предобуславливатель

- Общая форма простых методов:
$$P(x^{k+1} - x^k) = b - Ax^k$$
- $e^{k+1} = (I - P^{-1}A)e^k$
- Нужно $P \approx A$, P - предобуславливатель (preconditioner)
- Метод Ричардсона:
$$P = \frac{1}{\tau}I = \frac{\lambda_{max} + \lambda_{min}}{2}I$$
- Метод Якоби $P = D$
- Метод Гаусса-Зейделя $P = L + D$
- Можно сразу строить методы для предобусловленной системы
$$Ax = b \to P^{-1}Ax = P^{-1}b$$

Предобуславливатель (2)

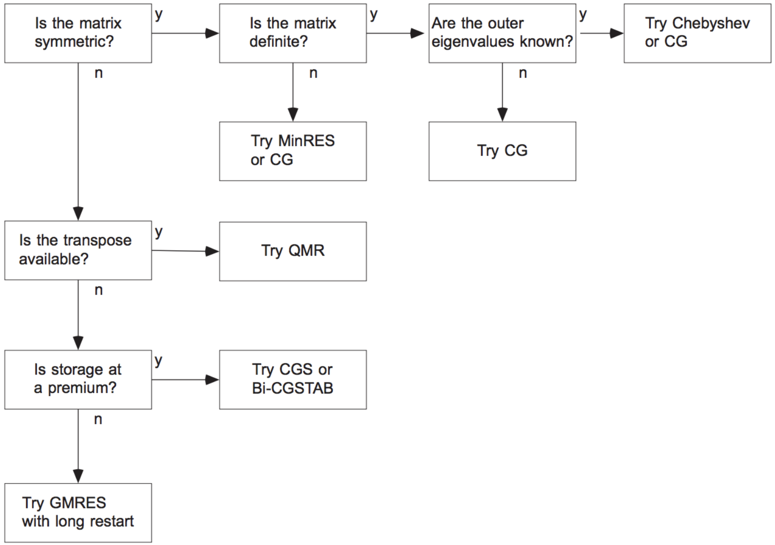
- Скорость сходимости итерационных методов зависит от числа обусловленности:

$$q \sim \frac{1 - 1/c}{1 + 1/c}, \ \frac{1 - 1/\sqrt{c}}{1 + 1/\sqrt{c}}$$

- Для всех методов нужны предобуславливатели!
- Распространены предобуславливатели, основанные на прямых методах
- ILU (Incomplete LU) – для разреженной матрицы вычисляется приближенное разреженное LU разложение
- P^{-1} и $P^{-1}A$ явно не вычисляются. Нужна только процедура (черный ящик) для вычисления $P^{-1}y$

Алгоритм выбора подходящего метода

Взято из https://github.com/oseledets/nla2019



Ссылки

- https://mipt.ru/news/matematiki_nashli_zoloto_v_dannykh – статья про то, как ученые из МФТИ улучшили качество георазведки. Упоминается предобуславливатель.
- gmres – функция из пакета scipy. Позволяет использовать предобуславливатель в явном виде (один из аргументов P^{-1})
- PETSC – открытая библиотека с большим набором линейных солверов для решения больших систем линейных уравнений.
- https://github.com/oseledets/nla2019/blob/master/lectures/lecture13/lecture-13.ipynb

Степенной метод для вычисления собственных значений

- Пусть A имеет базис из нормированных с.в. $z_i,\ Az_i = \lambda_i z_i, \|z_i\| = 1$ и

$$|\lambda_1| > |\lambda_2| \ldots > |\lambda_n|$$

- Возьмем $y_0 = c_1 z_1 + \ldots c_n z_n \Rightarrow$

$$y_k = A^k y_0 = c_1 \lambda_1^k z_1 + \ldots + c_n \lambda_n^k z_n = c_1 \lambda_1^k \left(z_1 + \mathcal{O}\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right) \right)$$

- $$\frac{y_k^* A y_k}{\|y_k\|^2} = \frac{(c_1 \lambda_1^k z_1^*) \lambda_1 (c_1 \lambda_1^k z_1)}{(c_1 \lambda_1^k)^2} + \mathcal{O}\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right) \to \lambda_1$$

Алгоритм степенного метода

- Добавляем нормировки на каждом шаге:

1. Начальное приближение y_0. $x_0 = y_0/\|y_0\|$
2. $y_k = Ax_{k-1}, \; x_k = y_k/\|y_k\|, \quad k = 1, 2, \ldots$
3. $x_k^* A x_k \to \lambda_1$

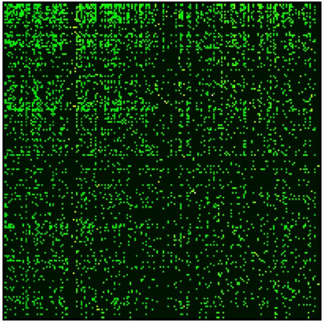
- Наиболее популярный метод вычисления с. ч. матриц общего вида - итерационный QR-алгоритм ($O(n^3)$)

Вычислительная математика.
Матрицы специального вида.

МФТИ

Разреженные матрицы

- $A \in \mathbb{R}^{m \times n}$ – разреженная (sparse), если число ненулевых элементов $NNZ(A) \ll nm$
- Пример: описание графов
  Google-matrix



«Google»-матрица статей Википедии

Структуры данных: координатный формат

- COO (Coordinate format): хранятся тройки (i, j, a_{ij})
  scipy.sparse.coo_matrix
- легко добавлять новые элементы
- сложно реализовать *index slicing* (выделение подматрицы)
- неудобное матрично-векторное умножение

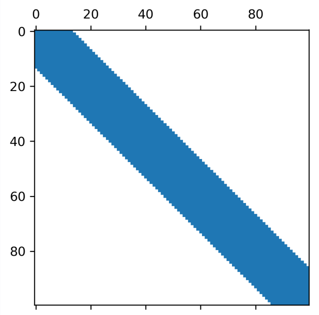
Структуры данных: CSR-формат

- CSR (Compressed Sparse Row)
  1. $A[k], \ k = 1, NNZ(A)$ – элементы матрицы
  2. $IA[0] = 0, \ IA[i] = IA[i-1] + NNZ(A[i,:])$
     $i = 1, m+1,$ – число ненулевых элементов в каждой строке
  3. $JA[k], k = 1, NNZ(A)$ – индексы столбца для каждого элемента
- сложно добавлять новые элементы
- удобно вычислять умножение $b = Ax$

```
for i = 1, n
    b[i] = 0.
    for k = IA[i-1]  + 1, IA[i]
        b[i] = b[i] + A[k] * x[JA[k]]
```

Пример использования координатного формата

Ленточные матрицы (band matrices)



- $A_{ij} = 0$ при $|i - j| > w$
- Матрица задаётся векторами на диагоналях
- LU разложение также будет ленточным $\Rightarrow$
- Систему можно решить за $\mathcal{O}(nw^2)$, где w - ширина ленты, scipy.linalg.solve_banded

Пример задачи

- $u_{xx} = f(x),\ u(0) = u(1) = 0$
- Введем сетку $x_i = ih, i = 0, \dots, n+1, \quad h = 1/(n+1)$
- $u_{xx}(x_i) \approx \dfrac{u_x(x_{i+1/2}) - u_x(x_{i-1/2})}{h} \approx$
$\dfrac{1}{h}\left(\dfrac{u_{i+1} - u_i}{h} - \dfrac{u_i - u_{i-1}}{h}\right) = \dfrac{(u_{i+1} - 2u_i + u_{i-1})}{h^2}$
- Получим линейную систему:

$$\underline{1} \cdot u_{i+1} - \underline{2} \cdot u_i + \underline{1} \cdot u_{i-1} = h^2 f_i,\ i = 1, \dots, n$$

$$\begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = h^2 \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{bmatrix}$$

Трехдиагональная прогонка (Thomas algorithm)

$$\begin{bmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}$$

- Прямой проход:
  $c'_1 = \frac{c_1}{b_1},\ c'_i = \frac{c_i}{b_i - c'_{i-1} a_i},\ d'_1 = \frac{d_1}{b_1},\ d'_i = \frac{d_i - a_i d'_{i-1}}{b_i - c'_{i-1} a_i}$
- Обратный проход: $x_n = d'_n,\ \ x_i = d'_i - c'_i x_{i+1}$
- Одно из достаточных условий устойчивости:
  $|b_i| \geq |a_i| + |c_i|,\ \ |b_1| > |c_1|$

Матрица Фурье

- $\varepsilon = \exp\left(-\frac{2\pi}{n} i\right) = \cos\left(-\frac{2\pi}{n}\right) + i \sin\left(-\frac{2\pi}{n}\right)$
- Построим матрицу Вандермонда для корней из единицы $\varepsilon^0, \varepsilon^1, \ldots, \varepsilon^{n-1}$

$$F_n = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \varepsilon^{1\cdot 1} & \varepsilon^{1\cdot 2} & \ldots & \varepsilon^{1\cdot(n-1)} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & \varepsilon^{(n-2)\cdot 1} & \varepsilon^{(n-2)\cdot 2} & \ldots & \varepsilon^{(n-2)\cdot(n-1)} \\ 1 & \varepsilon^{(n-1)\cdot 1} & \varepsilon^{(n-1)\cdot 2} & \ldots & \varepsilon^{(n-1)\cdot(n-1)} \end{bmatrix}$$

$$(F_n^* F_n)_{lj} = \sum_{k=0}^{n-1} \overline{\varepsilon}^{kl} \varepsilon^{kl} = \sum_{k=0}^{n-1} \varepsilon^{k(j-l)} = \begin{cases} \frac{1-\varepsilon^{(j-l)n}}{1-\varepsilon^{j-l}} = 0, & l \neq j \\ n, & l = j \end{cases}$$

$$F_n^{-1} = \frac{1}{n} F_n^*$$

Быстрое преобразование Фурье

- Прямое дискретное преобразование Фурье:

$$[F_n f]_k = \sum_{j=1}^{n} f_j \exp\left(-i\frac{2\pi}{n}kj\right)$$

- Быстрое преобразование Фурье (БПФ, FFT): умножение на F_n за $O(n \log n)$
- Входит в TOP-10 алгоритмов 20-го века
- Основа цифровой обработки сигналов

Быстрое преобразование Фурье

- $n = 2^L$, $m = n/2$, нумерация индексов $\overline{0, n-1}$
- $\widetilde{F}_n = P_n F_n$ – сначала четные строки, потом нечетные

$$\widetilde{F} = \begin{bmatrix} \left[\varepsilon^{2kl}\right]_{m\times m} & \left[\varepsilon^{2k(m+l)}\right]_{m\times m} \\ \left[\varepsilon^{(2k+1)l}\right]_{m\times m} & \left[\varepsilon^{(2k+1)(m+l)}\right]_{m\times m} \end{bmatrix}$$

$$\left[\varepsilon^{2k(m+l)}\right]_{m\times m} = \left[\varepsilon^{2kl}\right]_{m\times m} = \left[\exp\left(i\frac{2\pi}{(n/2)}kl\right)\right] = F_m$$

$$\left[\varepsilon^{(2k+1)l}\right]_{m\times m} = F_m D_m, \quad \left[\varepsilon^{(2k+1)(m+l)}\right]_{m\times m} = -F_m D_m$$

$$F_n = P_n \begin{bmatrix} F_m & 0 \\ 0 & F_m \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & D_m \end{bmatrix} \begin{bmatrix} I_m & I_m \\ I_m & -I_m \end{bmatrix}$$

Пример применения БПФ

Циркулянтные матрицы и матрица Фурье

- *Циркулянтной* называется матрица вида:

$$A = \begin{bmatrix} a_0 & a_{n-1} & a_{n-2} & \dots & a_2 & a_1 \\ a_1 & a_0 & a_{n-1} & \dots & a_3 & a_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_1 & a_0 \end{bmatrix}$$

- $\xi^n = 1,\ \lambda = \lambda(\xi) \equiv a_0 + \xi a_1 + \ldots + \xi^{n-1} a_{n-1}$
- Умножим обе части на $1, \xi, \xi^2, \ldots, \xi^{n-1}$:

$$\begin{cases} \lambda \cdot 1 & = & a_0 & +\xi a_1 & \dots & +\xi^{n-1} a_{n-1} \\ \lambda \cdot \xi & = & a_{n-1} & +\xi a_0 & \dots & +\xi^{n-1} a_{n-2} \\ \dots \\ \lambda \cdot \xi^{n-1} & = & a_1 & +\xi a_2 & \dots & +\xi^{n-1} a_0 \end{cases}$$

$$\lambda(\xi)[1, \xi, \ldots, \xi^{n-1}] = [1, \xi, \ldots, \xi^{n-1}]A$$

Циркулянтные матрицы и матрица Фурье (2)

- $\lambda(\xi)[1, \xi, \dots, \xi^{n-1}] = [1, \xi, \dots, \xi^{n-1}]A$
- Возьмем $\varepsilon = \exp(-2\pi i/n)$.
- $\xi = 1, \varepsilon, \dots, \varepsilon^{n-1} \Rightarrow \Lambda F_n = F_n A$, где $\Lambda = diag\left(\lambda(1), \lambda(\xi), \dots, \lambda(\xi^{n-1})\right)$

Теорема о циркулянтах

Пусть A - циркулянтная матрица с 1-м столбцом $a = [a_0, \dots, a_{n-1}]^T$, тогда

$$A = \frac{1}{n} F_n^* \Lambda F_n, \quad \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} = F_n \begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix}$$

Тёплицевы матрицы

- *Тёплицевой* называется матрица вида:

$$T_n = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \dots & t_{1-n} \\ t_1 & t_0 & t_{-1} & \dots & t_{2-n} \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ t_{n-1} & t_{n-2} & t_{n-3} & \dots & t_0 \end{bmatrix}$$

- Тёплицеву матрицу можно вложить в циркулянтную матрицу C_{2n}:

$$C_{2n} = \begin{bmatrix} T_n & B_n \\ B_n & T_n \end{bmatrix}, B_n = \begin{bmatrix} 0 & t_{n-1} & \dots & t_2 & t_1 \\ t_{1-n} & 0 & t_{n-1} & \dots & t_2 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ t_{-1} & t_{-2} & \cdots & t_{1-n} & 0 \end{bmatrix}$$

- Умножение на $T_n : O(n \log n)$

Заключение

- Разреженные матрицы
- Ленточные матрицы
- Циркулянтные и тёплицевы матрицы, преобразование Фурье

Вычислительная математика. Приближение функций. Полиномиальная интерполяция.

МФТИ

Приближение функций

Основная задача

Для функции f из класса F найти наилучшее *в некотором смысле* приближение $\phi \approx f$ из класса Φ «простых» функций.

- построение эффективного алгоритма вычисления приближения
- доказательство различных оценок ошибки в зависимости от свойств класса F (гладкость, величина производных и т.п.)

Важная идея

Идея

Приблизим исследуемую функцию «простой» функцией, а для простой функции решим задачу точно.

Примеры:

- Численное дифференцирование
- Численное интегрирование
- Методы оптимизации
- Численное решение дифференциальных уравнений

Общие подходы теории приближений

- F, Φ – часто векторные нормированные пространства

Два общих подхода:

1. **Минимизационный подход**:
   Выбирается норма $\|\cdot\|$ на F и ищется функция $\phi \in \Phi \subset F$, минимизирующая $\|f - \phi\|$
2. **Интерполяционный подход**:
   Выбираются точки (узлы) $x_0, \ldots, x_n$ и ищется функция $\phi \in \Phi$, удовлетворяющая *интерполяционным условиям*:

$$\phi(x_i) = f(x_i), \;\; i = 0, 1, \ldots, n$$

Интерполяция обобщенным многочленом

- Рассмотрим базис в линейном подпространстве $\Phi_n \subset \Phi$: $\phi_0, \dots, \phi_n$:

$$\phi = \sum_{k=0}^{n} c_k \phi_k$$

- Интерполяционные условия образуют систему:

$$\begin{bmatrix} \phi_0(x_0) & \phi_1(x_0) & \dots & \phi_n(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \dots & \phi_n(x_1) \\ \dots & \dots & \ddots & \dots \\ \phi_0(x_n) & \phi_1(x_n) & \dots & \phi_n(x_n) \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix}$$

- Если матрица невырожденная, то ϕ существует и единственен

Интерполяция алгебраическими многочленами

- Φ_n - пространство многочленов степени $\leq n$. Базис: $1, x^1, x^2, \ldots, x^n$:

$$\underbrace{\begin{bmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \ldots & \ldots & \ddots & \ldots & \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{bmatrix}}_{A} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix}$$

- A – матрица Вандермонда: $det(A) = \prod\limits_{0 \leq i < j \leq n} (x_j - x_i)$
- Если все узлы различны, то интерполяционный многочлен существует и единственен.

Обусловленность матрицы Вандермонда

Интерполяционный многочлен в форме Лагранжа

Достаточно иметь способ вычисления значения многочлена в любой точке.

- возьмем многочлены степени n $l_0(x), \ldots, l_n(x)$:

$$l_j(x_i) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad \text{Тогда } L_n(x) = \sum_{j=0}^{n} f(x_j) l_j(x)$$

- l_j – *элементарные многочлены Лагранжа*

$$l_j(x) = \prod_{k=0,\, k\neq j}^{n} \frac{x - x_k}{x_j - x_k} \underset{x\neq x_j}{=} \frac{\prod_{k=0}^{n}(x - x_k)}{(x - x_j)\prod\limits_{k=0, k\neq j}(x_j - x_k)}$$

Графики многочленов Лагранжа

Погрешность интерполяции

Погрешность интерполяции

Пусть $x_0, \dots, x_n \in [a, b]$ и $f \in C^{n+1}[a, b]$. Тогда

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}\omega(x), \;\; \omega(x) = \prod_{k=0}^{n}(x - x_k)$$

где $\xi(x) \in \left[\min\{x, x_0, \dots, x_n\}, \max\{x, x_0, \dots, x_n\}\right]$

Погрешность интерполяции (доказательство)

- $x \neq x_k$, $\omega(x) \neq 0$. Рассмотрим:

$$g(t) \equiv f(t) - L_n(t) - c\omega(t), \quad c \equiv \frac{f(x) - L_n x}{\omega(x)}$$

- $g(t)$ обращается в нуль при $t = x, x_0, \ldots, x_n$
- По теореме Ролля $g^{(1)}$ имеет хотя бы $n+1$ нуль
- $\ldots$ g^{n+1} имеет хотя бы один ноль: $\exists \xi : g^{(n+1)}(\xi) = 0$

$$g^{(n+1)}(t) = f^{(n+1)}(t) - c(n+1)! \Rightarrow$$

$$f^{(n+1)}(\xi(x)) - \frac{f(x) - L_n(x)}{\omega(x)}(n+1)! = 0 \quad \square$$

Разделенные разности

- Значения $f(x_k)$ в узлах будем называть *разделенными разностями порядка 0*
- Для любой пары узлов x_l, x_m введем *разделенные разности порядка 1*:

$$f(x_l; x_m) = \frac{f(x_l) - f(x_m)}{x_l - x_m}$$

- Реккурентно: разделенная разность порядка k:

$$f(x_0; \ldots; x_k) = \frac{f(x_1; \ldots; x_k) - f(x_0; \ldots; x_{k-1})}{x_k - x_0}$$

Разделенные разности (2)

Лемма

$$f(x_0;\ldots;x_k)=\sum_{j=0}^{k}\frac{f(x_j)}{\prod\limits_{\substack{l=0\\ l\neq j}}^{k}(x_j-x_l)}$$

$$f(x_0;\ldots;x_k)=\frac{f(x_1;\ldots;x_k)-f(x_0;\ldots;x_{k-1})}{x_k-x_0}=$$

$$\sum_{j=1}^{k}\frac{f(x_j)}{\prod\limits_{l=1,l\neq j}^{k}(x_j-x_l)(x_k-x_0)}-\sum_{j=0}^{k-1}\frac{f(x_j)}{\prod\limits_{l=0,l\neq j}^{k-1}(x_j-x_l)(x_k-x_0)}=$$

Разделенные разности (3)

$$\sum_{j=1}^{k} \frac{f(x_j)}{\prod\limits_{l=1, l\neq j}^{k} (x_j - x_l)(x_k - x_0)} - \sum_{j=0}^{k-1} \frac{f(x_j)}{\prod\limits_{l=0, l\neq j}^{k-1} (x_j - x_l)(x_k - x_0)} =$$

$$f(x_0) / \prod_{l=0, l\neq 0}^{k} (x_0 - x_l) + f(x_k) / \prod_{l=0, l\neq k}^{k} (x_k - x_l) \quad +$$

$$\sum_{j=1}^{k-1} \frac{f(x_j)}{\prod_{l=1, l\neq j}^{k-1} (x_j - x_l)(x_k - x_0)} \left\{ \frac{1}{x_j - x_k} - \frac{1}{x_j - x_0} \right\} =$$

$$= \frac{f(x_0)}{\prod\limits_{l=0, l\neq 0}^{k} (x_0 - x_l)} + \frac{f(x_k)}{\prod\limits_{l=0, l\neq k}^{k} (x_k - x_l)} + \sum_{j=1}^{k-1} \frac{f(x_j)}{\prod\limits_{l=0, l\neq j}^{k} (x_j - x_l)} \qquad \square$$

Следствия леммы

$$f(x_0;\dots;x_k) = \sum_{j=0}^{k} \frac{f(x_j)}{\prod_{\substack{l=0\\ l\neq j}}^{k}(x_j - x_l)}$$

- Значение не зависит от порядка узлов
- $f(x) - L_n(x) = f(x;x_0;\dots;x_n)\omega(x)$

$$f(x)-L_n(x) = \omega(x)\left(\frac{f(x)}{\omega(x)} + \sum_{j=0}^{n} \frac{f(x_j)}{(x_j - x)\prod\limits_{k=0,k\neq j}^{n}(x_j - x_k)}\right) \quad \square$$

- $f(x) - L_n(x) = \frac{f^{n+1}(\xi)}{(n+1)!}\omega(x) \Rightarrow f(x_0;\dots;x_k) = \frac{f^{(k)}(\xi)}{k!}$

Форма Ньютона

Теорема

$$
\begin{aligned}
L_n(x) = f(x_0) &+ f(x_0; x_1)(x - x_0) \\
&+ f(x_0; x_1; x_2)(x - x_0)(x - x_1) + \dots \\
&+ f(x_0; x_1; \dots; x_n)(x - x_0) \cdots (x - x_{n-1})
\end{aligned}
$$

- $L_n = L_0 + (L_1 - L_0) + (L_2 - L_1) + \ldots + (L_n - L_{n-1})$
- L_{k-1} интерполирует L_k в узлах $x_0, \ldots, x_{k-1}$

  $$L_k(x) - L_{k-1}(x) = L_k(x; x_0; \ldots; x_{k-1})(x - x_0) \ldots (x - x_{k-1})$$
- $L_k(x; x_0; \ldots; x_{k-1}) = L_k^{(k)}(\xi)/k! = a_k$, $L_k = a_k x^k + \ldots$
- Возьмем $x = x_k$: $L_k(x_k; x_0; \ldots; x_{k-1}) = f(x_0; \ldots; x_k)$ $\square$

Таблица разделенных разностей

Для вычисления значений многочлена в форме Ньютона удобно использовать таблицу:

$$
\begin{array}{ccccccc}
f(x_0) & & & & & & \\
& \ddots & & & & & \\
f(x_1) & \dots & f(x_0;x_1) & & & & \\
& \ddots & & \ddots & & & \\
f(x_2) & \dots & f(x_1;x_2) & \dots & f(x_0;x_1;x_2) & & \\
& \ddots & & \ddots & & \ddots & \\
f(x_3) & \dots & f(x_2;x_3) & \dots & f(x_1;x_2;x_3) & \dots & f(x_0;x_1;x_2;x_3)
\end{array}
$$

Связь разных форм с базисом

$$\underbrace{\begin{bmatrix} \phi_0(x_0) & \phi_1(x_0) & \dots & \phi_n(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \dots & \phi_n(x_1) \\ \dots & \dots & \ddots & \dots \\ \phi_0(x_n) & \phi_1(x_n) & \dots & \phi_n(x_n) \end{bmatrix}}_{A} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix}$$

- $\phi_k(x) = x^k$: A - матрица Вандермонда
- $\phi_k(x) = l_k(x)$: $A = I$
- $1, (x - x_0), (x - x_0)(x - x_1), \dots$: A - нижнетреугольная

Выбор узлов интерполяции

- Оценка ошибки:

$$|f(x) - L_n(x)| = \left|\frac{f^{(n+1)}(\xi(x))}{(n+1)!}\prod_{k=0}^{n}(x - x_k)\right| \leq$$

$$\frac{M_{n+1}}{(n+1)!}\boxed{\max_{[a,b]}\left|\prod_{k=0}^{n}(x - x_k)\right|}$$

- Оптимизационная задача:

$$\min_{\{x_0,\dots,x_n\}\subset[a,b]}\max_{[a,b]}\left|\prod_{k=0}^{n}(x - x_k)\right|$$

- Решение: многочлены Чебышёва

Многочлены Чебышёва

- $T_n(t) = \cos(n \arccos t), \ \ t \in [-1, 1]$
- $T_0(t) = 1, \ T_1(t) = t$, при $n \geq 1$

$$T_{n+1}(t) = 2tT_n(t) - T_{n-1}, n = 1, 2, \dots$$

- $T_n(t)$ - многочлен степени n со старшим коэффициентом 2^{n-1}
- Корни многочлена

$$t_{nj} = \cos\left(\frac{\pi}{2n} + \frac{\pi}{n}j\right), \ j = 0, \dots, n-1$$

- **Наименьшее отклонение от нуля**

$$\frac{1}{2^{n-1}} T_n = \arg \min_{p \in \mathcal{P}_n, a_n = 1} \max_{[-1,+1]} |p(x)|$$

Сходимость на разных сетках

Заключение

- Интерполяционный подход
- Формы Лагранжа, Ньютона
- Остаточный член интерполяции
- Выбор узлов интерполяции

Вычислительная математика. Сходимость интерполяционного процесса. Сплайны.

МФТИ

Интерполяционный многочлен

- $L_n \in \mathcal{P}_n,\ L_n(x_k) = f(x_k),\ k = 0, \dots, n$
- Форма Лагранжа: $L_n(x) = \sum\limits_{k=0}^{n} f_k l_k(x)$
- Форма Ньютона:

$$\begin{aligned} L_n(x) =& f(x_0) + f(x_0;x_1)(x-x_0) + \dots \\ &+ f(x_0;\dots;x_n)(x-x_0)\cdots(x-x_{n-1}) \end{aligned}$$

- Остаточный член интерполяции:

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}\omega(x),\ \omega(x) = \prod_{k=0}^{n}(x-x_k)$$

Сходимость интерполяционных многочленов

- $L_n(x) \xrightarrow{?} f(x), \; n \to \infty \; \forall x \in [a, b]$
- $\sup\limits_{x \in [a,b]} |f(x) - L_n(x)| \xrightarrow{?} 0$
- Как сходимость зависит от свойств функции?
- Как сходимость зависит расположения узлов?

Простое достаточное условие

Если $f \in C^\infty[a,b]$ и $\forall n : \sup_x |f^{(n)}(x)| \le M^n$, то для любой последовательности узлов:

$$\|f(x) - L_n(x)\|_{C[a,b]} \le \frac{(M(b-a))^{n+1}}{(n+1)!}$$

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}\prod_{k=0}^{n}(x - x_k) \le$$

$$\frac{(M(b-a))^{n+1}}{(n+1)!} \approx \frac{(M(b-a))^{n+1}}{\sqrt{2\pi(n+1)}((n+1)/e)^{n+1}} \to 0 \quad \square$$

Расходимость на равномерной сетке

Сходимость операторов

- $F = C[a,b]$ - банахово пространство с нормой

$$\|f\|_{C[a,b]} \equiv \max_{x\in[a,b]} |f(x)|$$

- Π_n - пространство полиномов степени $\leq n$. Введем оператор P_n:

$$P_n : F \to \Pi_n, \quad P_n f = L_n$$

- Оператор P_n
  - линейный: $P_n(f_1 + f_2) = P_n f_1 + P_n f_2$
  - непрерывный (ограниченный)
  - является проектором: $P_n^2 = P_n$

При каких условиях $P_n f \to f \quad \forall f \in F$?

Принцип равномерной ограниченности

Принцип равномерной ограниченности

Для любой последовательности непрерывных линейных операторов $\{P_n\}$ верно

$$\sup_n \|P_n f\| \leq c(f) < +\infty \ \forall f \in F \Leftrightarrow \sup_n \|P_n\| < +\infty$$

Теорема Банаха-Штейнгауза

Последовательность $\{P_n f\}$ сходится для всех $f \in F \Leftrightarrow$

1. $\sup_n \|P_n\| \leq M < +\infty$
2. $P_n f$ сходится при $n \to \infty$ на подмножестве $\tilde{F}$, всюду плотном в F

Доказательство теоремы Б-Ш

- $(\Rightarrow)$ - из принципа равномерной ограниченности
- $(\Leftarrow)$ возьмем $f \in F$ и ε-приближение $f_\varepsilon \in \tilde{F}$:

$$\|f - f_\varepsilon\| \leq \varepsilon : \|P_n f - P_m f\| \leq$$
$$\|P_n f - P_n f_\varepsilon\| + \|P_n f_\varepsilon - P_m f_\varepsilon\| + \|P_m f_\varepsilon - P_m f\| \leq$$
$$M\varepsilon + \varepsilon + M\varepsilon$$

- $\Rightarrow \{P_n f\}$ фундаментальная $\Rightarrow$ сходится $\square$

Применение т. Б-Ш к интерполяции

Теорема Банаха-Штейнгауза

Последовательность функций $P_n f$ сходится для всех $f \in F \Leftrightarrow$

1. $\sup_n \|P_n\| \leq M < +\infty$
2. $P_n f$ сходится при $n \to \infty$ на подмножестве $\tilde{F}$, всюду плотном в F

- $P_n f \to f$, если f – многочлен ($f \in \mathcal{P}$)
- По теореме Вейерштрасса $\mathcal{P}$ – всюду плотное множество в $C[a, b]$
- Сходимость зависисит от поведения $\|P_n\|$ при $n \to \infty$

Константы Лебега

$$|[P_n f](x)| = |\sum_j f_j l_j(x)| \le \sum_j |f_j||l_j(x)| \le$$

$$\|f\|_{C[a,b]} \sum_j |l_j(x)| = \|f\|\Lambda(x)$$

$$\|P_n f\|_{C[a,b]} \le \|f\| \max_{x\in[a,b]} \sum_j |l_j(x)| = \|f\|\Lambda$$

- Неравенство достигается на $f_j = \text{sign } l_j(x^*)$
- $\|P_n\| = \sup\limits_{f\neq 0} \dfrac{\|P_n f\|}{\|f\|} = \Lambda = \max\limits_{x\in[a,b]} \sum_j |l_j(x)|$

Обусловленность задачи интерполяции

- Пусть значения в узлах заданы с ошибкой

$$\tilde{f}_j = f_j + \delta f_j, \ |\delta f_j| \le \delta_j \le \delta$$

- Погрешность в значении многочлена:

$$|L_n(x;\tilde{f}) - L_n(x;f)| = |L_n(x;\delta f)| =$$

$$|\sum_j \delta f_j \, l_j(x)| \le \sum_j \delta_j |l_j(x)|$$

$$\|L_n(\tilde{f}) - L_n(f)\| \le \delta \max_{x\in[a,b]} \sum_j |l_j(x)| = \delta \, \Lambda$$

$\|P_n\|$ на равномерных сетках

- Равномерные сетки:

$$t_j = -1 + \frac{2}{n}j,\ j = 0, \dots, n$$

$$t = -1 + \frac{2}{n}\theta,\ 0 < \theta < 1:$$

$$\|P_n\| = \max_{t\in[-1,1]} \sum_{k=0}^{n} \left| \prod_{j=0, j\neq k}^{n} \frac{t - t_j}{t_k - t_j} \right| \geq$$

$$\sum_{k=0}^{n} \left| \prod_{j=0, j\neq k}^{n} \frac{\theta - j}{k - j} \right| \geq \frac{\theta(1-\theta)}{n^2} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} =$$

$$\boxed{\frac{\theta(1-\theta)}{n^2} 2^n}$$

Теорема Бернштейна-Фабера

- Для любой последовательности сеток:

$$\|P_n\| \geq c \ln\, n \to \infty$$

- Чебышёвские сетки асимптотически оптимальны:

$$\|P_n^{cheb}\| = O(\ln\, n)$$

Теорема Бернштейна-Фабера

Для любой последовательности сеток на $[a, b]$ $\exists f \in C[a, b]$, для которой последовательность интерполяционных многочленов не сходится равномерно ни к какой непрерывной функции.

Теоремы для чебышёвских сеток [1]

- $E = \|f - L_n\|_C$:

Условия на f	Убывание ошибки
f - Липшиц непрерывная	$E \to 0$
$\nu \geq 1,\ f \in C^{\nu-1}$, $f^{(\nu)}$ имеет ограниченную вариацию V	$E \leq \dfrac{4V}{\pi\nu(n-\nu)^{\nu}}$
f – аналитическая на $[-1, 1]$	$E \leq C^{-n}, C > 1$

[1] Lloyd N. Trefethen. *Approximation Theory and Approximation Practice*.
SIAM, Philadelphia, PA, USA, 2012

Примеры

Сплайны

Сплайн

Сплайн - это кусочно-многочленная функция S.

- $a = x_0, \ldots, x_n = b$, $S_{[x_{k-1}, x_k]}$ – многочлен.
- Максимальная степень m многочлена называется степенью сплайна.
- Если $S \in C^k[a, b]$, k называется гладкостью сплайна.
- Разность $n - k$ называется дефектом сплайна.

Кубический сплайн

- $a = x_0, \ldots, x_n = b$, $f_k = f(x_k)$

$$\Phi = \{\phi \in C^2[a, b] : \phi(x_k) = f_k, k = 0, 1, \ldots, n\}$$

- $S(x) \in \Phi$ называется *интерполяционным кубическим сплайном*, если $\forall\ [x_{k-1}, x_k], k = \overline{1, n}$, S - многочлен степени ≤ 3.
- $4n - 2$ уравнения для коэффициентов:
  1. $2 + 2(n-1) = 2n$ интерполяционных условия
  2. $(n-1)$ - непрерывность 1-й производной
  3. $(n-1)$ - непрерывность 2-й производной
- Дополнительные граничные условия:
  - $S''(x_0) = S''(x_n) = 0$ - естественный сплайн
  - $S'(x_0) = f'(x_0)$, $S'(x_n) = f'(x_n)$
  - $S'(x_0) = S'(x_0), S''(x_0) = S''(x_n)$ - периодичность

Вариационное свойство е. с.

Естественный сплайн минимизирует *функционал энергии*:

$$E(\phi) = \int_a^b (\phi''(x))^2 \, dx, \; E(\phi) > E(S) \, \forall \phi \in \Phi, \phi \neq S$$

Доказательство

$$(\phi'')^2 - (S'')^2 = (\phi'' - S'')^2 + 2S''(\phi'' - S'') \Rightarrow$$

$$E(\phi) - E(S) = \underbrace{E(\phi - S)}_{>0} + 2\int_a^b S''(\phi'' - S'')\,dx$$

Возьмем интеграл по частям:

$$\int_a^b S''(\phi'' - S'')\,dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} S''(\phi'' - S'')\,dx =$$

$$S''(\phi' - S')\Big|_a^b - \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} S^{(3)}(\phi' - S')\,dx =$$

$$-\sum_{k=1}^{n} S^{(3)}(\phi - S)\Big|_{x_{k-1}}^{x_k} + \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} S^{(4)}(\phi - S)\,dx = 0 \quad \square$$

Вычисление сплайна

- Рассмотрим $S(x)$ на k-ом отрезке $[x_{k-1}, x_k]$:

$$u_k = S''(x_k),\ \ h_k = x_k - x_{k-1}, x = x_{k-1} + th_k$$

- Поскольку S - кубический многочлен при $x \in [x_{k-1}, x_k]$, находим:

$$S''(x) = (1-t)u_{k-1} + tu_k \Rightarrow$$

$$S'(x) = S'(x_{k-1}) + h_k\left(\frac{1}{2} - \frac{(1-t)^2}{2}\right)u_{k-1} + h_k\frac{t^2}{2}u_k$$

$$S(x) = S(x_{k-1}) + h_k tS'(x_{k-1}) +$$

$$h_k^2\left(\frac{t}{2} + \frac{(1-t)^3}{6} - \frac{1}{6}\right)u_{k-1} + h_k^2\frac{t^3}{6}u_k$$

Вычисление сплайна (2)

$$S(x) = S(x_{k-1}) + h_k t S'(x_{k-1}) +$$
$$h_k^2 \left(\frac{t}{2} + \frac{(1-t)^3}{6} - \frac{1}{6} \right) u_{k-1} + h_k^2 \frac{t^3}{6} u_k$$

Обозначим $\delta f_k = (f_k - f_{k-1})/h_k$. При $t = 1$:

$$S'(x_{k-1}) = \delta f_k - \frac{h_k}{3} u_{k-1} - \frac{h_k}{6} u_k$$

Учитывая это, полагаем $t = 1$ в выражении для $S'(x)$:

$$S'(x_k) = \delta f_k + \frac{h_k}{6} u_{k-1} + \frac{h_k}{3} u_k$$

Приравняем последние два выражения:

$$\delta f_k + \frac{h_k}{6} u_{k-1} + \frac{h_k}{3} u_k = \delta f_{k+1} - \frac{h_{k+1}}{3} u_k - \frac{h_{k+1}}{6} u_{k+1} \Rightarrow$$
$$h_k u_{k-1} + 2(h_k + h_{k+1}) u_k + h_{k+1} u_{k+1} = \rho_k \equiv 6(\delta f_{k+1} - \delta f_k)$$

Вычисление сплайна (3)

- $u_0 = u_n = 0$, получаем систему линейных уравнений с трехдиагональной матрицей:

$$T \begin{bmatrix} u_1 \\ \dots \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \dots \\ \rho_{n-1} \end{bmatrix}$$

$$T = \begin{bmatrix} 2(h_1 + h_2) & h_2 & & & \\ h_2 & 2(h_2 + h_3) & h_3 & & \\ & \ddots & \ddots & \ddots & \\ & & & h_{n-1} & 2(h_{n-1} + h_n) \end{bmatrix}$$

- Система невырожденная, т.к. T имеет диагональное преобладание.
- Естественный сплайн существует и единственен.

Погрешность сплайн-интерполяции

Теорема

Пусть $f \in C^j[a, b]$, $1 \leq j \leq 4$, тогда

$$\|f - S_n\|_{C[a,b]} = \mathcal{O}(h^j), \ \ h = \max_k h_k$$

Последовательность сеток называется *квазиравномерной*, если h_{max}/h_{min} равномерно ограничено при $n \to \infty$.

Теорема

Для любой $f \in C[a, b]$ последовательность $\{S_n\}$ на квазиравномерных сетках сходится к f равномерно.

Пример

Свойство квазилокальности

- В обычной интерполяции при изменении значения в точке x_k весь многочлен «перестраивается»
- мера чувствительности $|l_j(x)|$
- Зависимость сплайна на отрезке $[x_{k-1}, x_k]$ от значений f вне этого отрезка осуществляется через коэффициенты u_k.
- Как сильно изменится сплайн «вдали» от возмущения?

Элементы T^{-1}

Свойство квазилокальности

Теорема

Пусть $A = [a_{ij}]$ - невырожденная ленточная матрица порядка n, $a_{ii} \neq 0$, $a_{ij} = 0$ при $|i - j| > L$.
Пусть матричная норма $\|\cdot\|$ такова, что норма любой матрица не может быть меньше модуля каждого из её элементов (например, $\|\cdot\|_1$, $\|\cdot\|_\infty$). Тогда если

$$q \equiv \|(\mathrm{diag} A)^{-1} \mathrm{off} A\| < 1$$

то для элементов обратной матрицы $A^{-1} = [a_{ij}^{(-1)}]$:

$$|a_{ij}^{(-1)}| \leq \|(\mathrm{diag} A)^{-1}\| \frac{q^{|i-j|/L}}{1-q}, \ i, j = 1, \dots, n$$

Доказательство

- $F = -(\text{diag } A)^{-1} \text{ off } A$, $\|F\| = q < 1$
- Рассмотрим ряд Неймана:

$$(I - F)^{-1} = \sum_{k=0}^{\infty} F^k = \left(I + (\text{diag } A)^{-1} \text{ off } A \right)^{-1} =$$

$$\left((\text{diag } A)^{-1} (\text{diag } A + \text{off } A) \right)^{-1} = A^{-1} \text{diag } A$$

- $\|F^k (\text{diag } A)^{-1}\| \leq \|(\text{diag } A)^{-1}\| q^k$
- Ширина ленты $F^k \leq kL$
- $[F^k]_{i,j} \neq 0$ при $k \geq \lceil |i-j|/L \rceil$

$$|a_{ij}^{(-1)}| \leq \|(\text{diag } A)^{-1}\| \sum_{k=\lceil \frac{|i-j|}{L} \rceil}^{\infty} q^k \leq \|(\text{diag } A)^{-1}\| \frac{q^{|i-j|/L}}{1-q} \quad \square$$

Заключение

- Сходимость интерполяционного процесса
- Обусловленность задачи интерполяции
- Сплайн-интерполяция
- Естественный кубический сплайн

Вычислительная математика.
Методы приближения функций,
основанные на минимизации нормы.

МФТИ

Подходы к приближению функций

- Интерполяция: $\phi(x_k) = f(x_k)$
  - Интерполяция многочленами
  - Сплайн-интерполяция
- Минимизация нормы $\|\phi - f\| \to \min$
  - Равномерное приближение $\|f\|_C = \sup\limits_x |f(x)|$
  - Приближение в гильбертовом пространстве $\|f\| = (f, f)^{1/2}$, например

$$\|f\| = \left(\int_a^b |f(x)|^2 \, dx \right)^{1/2}$$

Равномерное приближение, альтернанс

Альтернанс (equioscillation)

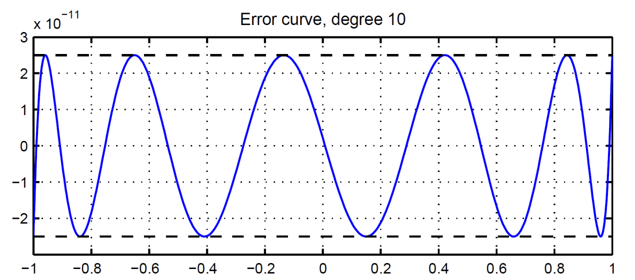
Альтернансом $g \in C[a, b]$ называется множество точек $a \leq x_1 < x_2 < \ldots < x_N \leq b$, такое что

1. $|g(x_k)| = \|g\|_{C[a,b]}$
2. $g(x_k)g(x_{k+1}) < 0$ (знаки чередуются)

Теорема об альтернансе

Теорема об альтернансе

Для $f \in C[-1,1]$ существует единственное наилучшее приближение $p^* \in \mathcal{P}_n$.
$p \in \mathcal{P}_n$ является многочленом наилучшего приближения тогда и только тогда, когда $f - p$ имеет не меньше $n+2$ точек альтернанса на $[-1,1]$.



Доказательство (1)

Существование:

- $\|f - p\|$ – непрерывный функционал от p
- Наилучшее приближение лежит в шаре $\{p \in \mathcal{P}_n : \|f - p\| \leq \|f\|\}$
- По теореме Больцано-Вейерштрасса функционал достигает минимума $\square$

$\Leftarrow$ (из альтернанса следует оптимальность)

- Допустим что $\|f - q\| < \|f - p\|$
- Тогда $p - q$ принимает чередующиеся по знаку ненулевые значения в $x_0, \ldots, x_{n+1}$
- $\Rightarrow p - q$ имеет хотя бы $n + 1$ ноль $\Rightarrow p - q = 0$ $\square$

Доказательство (2)

$\Rightarrow$ (из оптимальности следует альтернанс)

- Предположим, что точек альтернанса $\leq n+1$, $E = \|f - p\|$
- Пусть самый левый экстремум $-E$
- Тогда можно выбрать $-1 < x_1 < \ldots < x_k < 1$, $k \leq n$, такие что:

$$(f-p)(x) < E \text{ при } x \in [-1, x_1] \cup [x_2, x_3] \cup [x_4, x_5] \ldots$$
$$(f-p)(x) > -E \text{ при } x \in [x_1, x_2] \cup [x_3, x_4] \cup \ldots$$

- $\delta p(x) = (x_1 - x)(x_2 - x) \cdots (x_k - x)$
- $(p - \epsilon \delta p)(x)$ будет лучшим приближением при достаточно малом ϵ $\square$

Доказательство(3) (единственность)

- Пусть p – н. п., $x_0, x_1, \ldots, x_{n+1}$ – точки альтернанса
- $\|f - q\| \leq \|f - p\|$, $q \in \mathcal{P}_n$
- $(p - q)(x) \leq 0$ в $x_0, x_2, x_4, \ldots$ и ≥ 0 в $x_1, x_3, \ldots$
- $\Rightarrow$ $(p - q)$ имеет корни в каждом из $n + 1$ отрезков $[x_0, x_1], [x_1, x_2], \ldots, [x_n, x_{n+1}]$
- Докажем, что $p - q$ имеет $\geq k$ корней на $[x_0, x_k]$
- Допустим, $p - q$ имеет j корней на $[x_0, x_j]$, $j \leq k - 1$, но только $k - 1$ корень на $[x_0, x_k]$
- $\Rightarrow$ x_{k-1} – простой корень $\Rightarrow$ $x_{k-2}, \ldots x_1$ – п.к.
- $p - q \neq 0$ в x_0, x_k, знаки совпадают для нечетных k, и отличаются для четных: противоречит условию альтернанса $\square$

Многочлены наименее отклоняющиеся от нуля

Определение

Многочлен $p_n = x^n + a_{n-1}x^{n-1} + \ldots + a_0$ называется наименее уклоняющимся от нуля на $[a, b]$, если он имеет наименьшую норму $C[a, b]$ среди всех многочленов такого вида.

- $\|q_n(x)\|_{C[a,b]} \leq \|x^n - p_{n-1}(x)\|_{C[a,b]} \; \forall p_{n-1} \in \mathcal{P}_{n-1}$
- $\Rightarrow x^n - q_n(x)$ – наилучшее приближение к x^n
- По теореме об альтернансе: q_n наименее отклоняется от нуля $\iff$ $E = q_n$ имеет $(n-1) + 2 = n + 1$ точку альтернанса на $[a, b]$

Многочлены Чебышёва

- $T_0(x) = 1,\ T_1(x) = x,\ T_{n+1}(x) = 2xT_n(x) - T_{n-1}$
- $T_n(x) = \cos(n \arccos(x)),\ x \in [-1, 1]$
- T_n имеет $n+1$ точку альтернанса:
$$x_k = \cos(k\pi/n),\ k = 0, \ldots, n$$
- $\Rightarrow\ T_n/2^{n-1}$ – наименее уклоняющийся от нуля многочлен с $a_n = 1$
- На произвольном отрезке $[a, b]$ с помощью замены
$$x = \frac{a+b}{2} + t\frac{b-a}{2},\ t = \frac{2x-a-b}{b-a}$$
Получаем многочлен
$$Q_n(x) = 2^{1-2n}(b-a)^n T_n\left(\frac{2x-a-b}{b-a}\right)$$

Интерполяция и равномерное приближение

Теорема

Пусть $f \in C[-1,1]$, $E_n(f) = \min\limits_{\phi \in \mathcal{P}_n} \|f - \phi\|_{C[-1,1]}$. Тогда: $\|f - L_n\|_C \le (1 + \|P_n\|)E_n(f)$

- $P_n \phi_n = \phi_n$

$$\|f - L_n\| \le \|f - \phi_n\| + \|P_n \phi_n - P_n f\| = E_n(f) + \|P_n(\phi_n - f)\| \le (1 + \|P_n\|)E_n(f) \quad \square$$

- Для чебышёвских сеток: $\|f - L_n\| \le (1 + c \ln n)E_n(f)$

Алгоритм Ремеза

Возьмем $n+2$ точки $x_1, \ldots, x_{n+2}$

1. Решаем линейную систему:

$$c_0 + c_1 x_i + \ldots c_n x_i^n + (-1)^i E = f(x_i),\ \ i = 1, 2, \ldots, n+2$$

   относительно неизвестных $c_0, \ldots, c_n, E$.
2. Находим точки локального максимума ошибки $|f(x) - p_n|$, $p_n = c_0 + c_1 x + \ldots + c_n x^n$. Если условие альтернанса выполнено с заданной точностью - останавливаемся, иначе
3. Заменяем часть точек на точки локального максимума, так чтобы знак $f - p_n$ чередовался. Переходим к (1)

Пример: аппроксимация $|x|$

Наилучшее приближение в $||\cdot||_\infty$

- Для гладких функций интерполяция по чебышёвским узлам даёт близкую точность
- Примеры использования равномерного приближения:
  - Аппроксимация специальных функций
  - Построение цифровых фильтров в обработке сигналов
- Свойство альтернанса используется для приближения рациональными функциями $\dfrac{p_n(x)}{q_m(x)}$

Приближение в гильбертовом пространстве

- $\forall f$ в гильбертовом пространстве F и любого замкнутого подпространства Φ существует единственное разложение:

$$f = u + \phi, \ \phi \in \Phi, \ u \perp \Phi$$

- $\Phi = \mathrm{span}\{v_1, \ldots, v_n\}$, $\phi = c_1 v_1 + \ldots + c_n v_n$:

$$\begin{bmatrix} (v_1, v_1) & \ldots & (v_1, v_n) \\ \ldots & \ldots & \ldots \\ (v_n, v_1) & \ldots & (v_n, v_n) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} (v_1, f) \\ \vdots \\ (v_n, f) \end{bmatrix}$$

Скалярное произведение

- Для любой неотрицательной функции w с положительным интегралом по $[a, b]$ можно ввести:

$$(f, g) = \int_a^b f(x)g(x)w(x)\,dx$$

- Можно взять $\Phi = \mathcal{P}^n$ и провести ортогонализацию Грамма-Шмидта:

$$(L_i, L_j) = \delta_{ij}$$

- Скалярное произведение определяет набор *ортогональных многочленов* с точностью до множителя ± 1

Трехчленное рекуррентное соотношение

Теорема

Для любых ортогональных многочленов верно трехчленное рекуррентное соотношение:

$$xL_n(x) = \beta_{n-1}L_{n-1}(x) + \alpha_n L_n(x) + \beta_n L_{n+1}(x), \ \ n = 0, 1, \ldots$$

- Разложим многочлен xL_n по базису:

$$xL_n(x) = s_{n0}L_0(x) + \ldots + s_{nn}L_n(x) + s_{nn+1}L_{n+1}(x)$$

- $s_{nj} = (xL_n, L_j) = (L_n, xL_j) = 0$, при $j \leq n-2$
- Обозначим $\alpha_n = s_{nn}, \beta_n = s_{n,n+1} = (xL_n, L_{n+1})$
- $s_{n,n-1} = (xL_n, L_{n-1}) = (xL_{n-1}, L_n) = \beta_{n-1}$ $\square$

Следствия

$$
x\begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_{n-1}(x) \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_0 & \beta_1 & & \\ \beta_1 & \alpha_1 & \beta_2 & \\ & \ddots & \ddots & \ddots \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}}_{T_n} \begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_{n-1}(x) \end{bmatrix} + \beta_n L_n(x) \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}
$$

Теорема

L_n имеет n простых вещественных корней $x_1, \ldots, x_n$, которые являются собственными числами матрицы T_n с собственными векторами:

$$
[L_0(x_j), \ldots, L_{n-1}(x_j)]^T, \ 1 \leq j \leq n
$$

Доказательство

$$x \begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_{n-1}(x) \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_0 & \beta_1 & & \\ \beta_1 & \alpha_1 & \beta_2 & \\ & \ddots & \ddots & \ddots \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}}_{T_n} \begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_{n-1}(x) \end{bmatrix} + \beta_n L_n(x) \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- Если $L_n(\lambda) = 0$ – то λ - с.ч. T_n $(L_0(x) \neq 0)$
- Допустим, что λ - кратный корень: $L_n'(\lambda) = 0$:

$$T_n L'(\lambda) = \lambda L'(\lambda) + L(\lambda), \; T_n L(\lambda) = \lambda L(\lambda)$$

- $(T_n - \lambda I)L' = L, \; (T_n - \lambda I)^2 L' = (T_n - \lambda I)L = 0$
- $\Rightarrow L' \in ker(T_n - \lambda I)^2 = ker(T_n - \lambda I)$ (т.к. T_n - эрмитова $\Rightarrow$ нормальная)

Доказательство (2)

$$x \begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_{n-1}(x) \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha_0 & \beta_1 & & \\ \beta_1 & \alpha_1 & \beta_2 & \\ & \ddots & \ddots & \ddots \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}}_{T_n} \begin{bmatrix} L_0(x) \\ L_1(x) \\ \vdots \\ L_{n-1}(x) \end{bmatrix} + \beta_n L_n(x) \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- $\beta_i \neq 0, \Rightarrow T_n - \lambda I$ имеет отличный от нуля минор порядка $n-1$ (в котором β_i – на диагонали)
- $\dim\ ker(T_n - \lambda I) = 1, \Rightarrow$ для некоторого a

$$\begin{bmatrix} L'_0(\lambda) \\ L'_1(\lambda) \\ \vdots \\ L'_{n-1}(\lambda) \end{bmatrix} = a \begin{bmatrix} L_0(\lambda) \\ L_1(\lambda) \\ \vdots \\ L_{n-1}(\lambda) \end{bmatrix}$$

- $L'_0(\lambda) = 0 \Rightarrow a = 0$, противоречие с $L'_1(\lambda) \neq 0$ $\square$

Расположение корней

Теорема

При $n \geq 1$ все корни L_n вещественны, попарно различны и расположены внутри отрезка $[a, b]$

- Пусть $z_1, \dots, z_m$ - корни полинома L_n внутри $[a, b]$
- $L_n(x) = (x - z_1) \cdots (x - z_m) p_{n-m}(x)$
- $p_{n-m}(x)$ имеет один и тот же знак на $[a, b]$
- Если $m < n$, то $L_n \perp \prod_{k=1}^{m} (x - z_k)$:

$$\int_a^b (x - z_1)^2 \cdots (x - z_m)^2 p_{n-m}(x) w(x)\, dx = 0$$

- Противоречие: $w(x) \geq 0$, p_{n-m} не меняет знак $\square$

Разложение интерполяционного многочлена

$x_1, \ldots, x_n$ - корни ортогонального многочлена L_n. Тогда матрица

$$Q_n = \begin{bmatrix} L_0(x_1) & \cdots & L_{n-1}(x_1) \\ \cdots & \cdots & \cdots \\ L_0(x_n) & \cdots & L_{n-1}(x_n) \end{bmatrix}$$

имеет ортогональные строки, а матрица $D_n^{-1} Q_n$

$$D_n = \operatorname{diag}(d_1, \ldots, d_n), \; d_j = \|Q_n(j,:)\|_2$$

является ортогональной.

Разложение интерполяционного многочлена (2)

Интерполяционный многочлен p_{n-1}, построенный по корням $x_1, \dots, x_n$ ортогонального многочлена L_n и значениям $f_1, \dots, f_n$, представляется в виде

$$p_{n-1}(x) = c_1 L_0(x) + \ldots + c_n L_{n-1}(x)$$

$$\begin{bmatrix} c_1 \\ \cdots \\ c_n \end{bmatrix} = Q_n^T D_n^{-2} \begin{bmatrix} f_1 \\ \cdots \\ f_n \end{bmatrix}$$

$Q_n c = f \Rightarrow D_n(D_n^{-1} Q_n) c = f \Rightarrow c = (D_n^{-1} Q_n)^T D_n^{-1} f = Q_n^T D_n^{-2} f \quad \square$

Примеры

Заключение

- Наилучшее приближение в $C[a, b]$
- Условие альтернанса
- Приближение в гильбертовом пространстве ($L_2[a, b]$)
- Ортогональные многочлены

Вычислительная математика. Численное дифференцирование и численное интегрирование

МФТИ

Численное дифференцирование

Основная задача

- Дана процедура вычисления $f(x)$.
  Нужно вычислить значение d-й производной в точке $f^{(d)}(x_0)$.

- Даны точки $x_1, \dots, x_n$ и значения гладкой функции в этих точках $f_k = f(x_k)$.
  Нужно найти значения $f^{(d)}(x)$ или $f^{(d)}(x_k), k = 1, \dots, n$

Пример

$$f'(x) \approx \hat{f}' = \frac{f(x+h) - f(x)}{h} =$$

$$\frac{f(x) + f'(x)h + f''(x)h^2/2 + O(h^3) - f(x)}{h} =$$

$$f'(x) + f''(x)\frac{h}{2} + \mathcal{O}(h^2) =$$

$$f'(x) + f''(\xi)\frac{h}{2} = f'(x) + \mathcal{O}(h)$$

- В точной арифметике:

$$|f'(x) - \hat{f}'| \leq E(h) = \frac{M_2}{2}h, M_2 = \max_x |f''|$$

Пример (2)

- В машинной арифметике:

$$\hat{f}' = \frac{f(x+h)(1+\epsilon_1) - f(x)(1+\epsilon_2)}{h} =$$

$$f'(x) + f''(\xi)\frac{h}{2} + \frac{\epsilon_1 f(x+h) - \epsilon_2 f(x)}{h}, \quad |\epsilon_k| \leq \epsilon$$

$$|f'(x) - \hat{f}'| \leq \frac{M_2}{2}h + \left|\frac{\epsilon_1 f(x+h) - \epsilon_2 f(x)}{h}\right| \leq$$

$$\frac{M_2}{2}h + \frac{2M_0\epsilon}{h} = E(h)$$

- Точка минимума - оптимальный шаг:

$$E'(h) = 0 = \frac{M_2}{2} - \frac{2M_0\epsilon}{h^2}, \quad h^* = \left(\frac{4M_0\epsilon}{M_2}\right)^{1/2}$$

Пример расчета

Повышение порядка, метод неопределенных коэффициентов

- Даны точки $x_1, \dots, x_n$ и точка x:

$$\hat{f}' = \sum_{k=1}^{n} a_k f(x_k)$$

- Нужно подобрать a_k так, чтобы занулить как можно больше членов в разложении ошибки

$$f'(x) - \hat{f}' =$$

$$f'(x) - \sum_{k=1}^{n} a_k \left(f(x) + f'(x)(x_k - x) + \frac{f''(x)}{2}(x_k - x)^2 + \dots \right)$$

Метод неопределенных коэффициентов

$$f'(x) - \sum_{k=1}^{n} a_k \left(f(x) + f'(x)(x_k - x) + \frac{f''(x)}{2}(x_k - x)^2 + \ldots \right)$$

Занулим коэффициенты при $f(x)$, $f'(x)$, $f''(x), \ldots$:

$$\sum_{k=1}^{n} a_k = 0, \ \sum_{k=1}^{n} a_k(x_k - x) = 1, \ \sum_{k=1}^{n} a_k \frac{(x_k - x)^2}{2!} = 0$$

$$\begin{bmatrix} 1 & 1 & \ldots & 1 \\ \frac{(x_1-x)}{1!} & \frac{(x_2-x)}{1!} & \ldots & \frac{(x_n-x)}{1!} \\ \ldots & & & \\ \frac{(x_1-x)^{n-1}}{(n-1)!} & \frac{(x_2-x)^{n-1}}{(n-1)!} & \ldots & \frac{(x_n-x)^{n-1}}{(n-1)!} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Вид ошибки

- Ошибка имеет вид:

$$E = f'(x) - \hat{f}' = -\sum_{k=1}^{n} a_k \frac{f^{(n)}(x)}{n!}(x_k - x)^n + \ldots =$$

$$-\sum_{k=1}^{n} a_k \frac{f^{(n)}(\xi_k)}{n!}(x_k - x)^n$$

- $h = \max(x_k - x)$: $a_k = O(1/h)$, $E = O(h^{n-1})$,
- Если узлы расположены симметрично относительно x, то $E = O(h^n)$ (следующий член зануляется)
- Для производной d по n узлам максимальный порядок: $n - d$ $(n - d + 1$ в симметричном случае)

Дифференцирование интерполяционного многочлена

Общая идея

Приблизить функцию *простой* функцией, а простую функцию продифференцировать точно.

$$L_{n-1}(x) = \sum_{k=1}^{n} l_k(x) f_k$$

$$L'_{n-1}(x) = \sum_{k=1}^{n} l'_k(x) f_k$$

$$a_k = l'_k(x)$$

Матрицы дифференцирования

- $f = [f_1, \ldots, f_n]$, $w = [w_1, \ldots, w_n] \approx [f'(x_1), \ldots]$
- Можно ввести дискретный аналог $\frac{d}{dx}$: $w = Df$
- Пример:

$$D = \begin{bmatrix} -\frac{1}{h} & \frac{1}{h} & 0 & \cdots \\ -\frac{1}{2h} & 0 & \frac{1}{2h} & \cdots \\ & \ddots & \ddots & \ddots \\ 0 & \cdots & -\frac{1}{h} & \frac{1}{h} \end{bmatrix}$$

- При использовании формул максимального порядка, матрица будет *плотной*: $D = \left[l_j'(x_i)\right]$
- Для чебышевских узлов умножение на D можно вычислять с помощью FFT за $O(n \log n)$

Пример расчета

Численное интегрирование

Задача численного интегрирования

Дана процедура вычисления значений $f(x)$. Найти приближенное значение определенного интеграла

$$S(f) \approx I(f) = \int_a^b f(x)dx$$

Основная идея

Приблизить f *простой* функцией ϕ, а функцию ϕ проинтегрировать точно.

Интерполяционные квадратурные формулы (ф-лы Ньютона-Котса)

- Отобразим стандартный отрезок $[-1, 1]$ на $[a, b]$:

$$x = x(t) = \frac{a+b}{2} + \frac{b-a}{2}t$$

Выберем узлы $t_1, \ldots, t_n \in [-1, 1]$, $x_i = x(t_i)$

- Построим интерполяционный многочлен:

$$L_{n-1}(x) = \sum_{i=1}^{n} f(x_i) l_i(x) = \sum_{i=1}^{n} f(x_i) \prod_{j=1,\, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}$$

- Проинтегрируем по $[a, b]$

$$S(f) = \int_a^b L_{n-1}(x)\, dx = \sum_{i=1}^{n} d_i\, f(x_i),\ d_i = \frac{b-a}{2} \int\limits_{-1}^{1} l_i(t)\, dt$$

Точность квадратурных формул

- Если $f \in C^n[a,b]$, то из формулы для ошибки интерполяции получаем:

$$|I(f) - S(f)| \leq \frac{\|f^{(n)}\|_C}{n!} \left(\frac{b-a}{2}\right)^{n+1} \int\limits_{-1}^{+1} \left| \prod_{j=1}^{n} (t - t_j) \right| dt$$

- Говорят, что $S(f)$ имеет алгебраическую точность m, если она точна для многочленов степени $\leq m$
- Формула S для n узлов имеет алгебраическую точность $m \geq n-1 \iff S$ – интерполяционная квадратурная формула.
  - $\Leftarrow$ По построению
  - $\Rightarrow$ Вместо f подставим l_i, для которых формула точна, получим ту же формулу для d_i $\square$

Примеры квадратурных формул

$$|I(f) - S(f)| \le \frac{\|f^{(n)}\|_C}{n!} \left(\frac{b-a}{2}\right)^{n+1} \int\limits_{-1}^{+1} \left| \prod_{j=1}^{n} (t - t_j) \right| dt$$

- $h = b - a$, $M_m = \|f^{(m)}\|_{C[a,b]}$
- Формула прямоугольников с центральной точкой:

$$n = 1,\ S(f) = f(\frac{a+b}{2})h, \quad |I(f) - S(f)| \le \frac{1}{4} M_1 h^2$$

- Ту же формулу можно получить из эрмитовой интерполяции $t_1 = t_2 = 0$: $H(t) = f(0) + f'(0)t$:

$$f \in C^2 : |I(f) - S(f)| \le \frac{1}{24} M_2 h^3$$

Примеры квадратурных формул (2)

$$|I(f) - S(f)| \le \frac{\|f^{(n)}\|_C}{n!} \left(\frac{b-a}{2}\right)^{n+1} \int\limits_{-1}^{+1} \left| \prod_{j=1}^{n} (t - t_j) \right| dt$$

- Формула трапеций:

  $$n = 2,\ S(f) = \frac{1}{2}(f(a) + f(b))h,\ |I(f) - S(f)| \le \frac{1}{12} M_2 h^3$$

- Формула Симпсона: обычная интерполяция по 3-м точкам $-1, 0, +1$ и эрмитова интерполяция с кратным узлом $-1, 0, 0, +1$ дают один и тот же результат:

  $$S(f) = \frac{h}{6}\left(f(a) + 4f(\frac{a+b}{2}) + f(b)\right),\ E \le Ch^5$$

Составные квадратурные формулы

- отрезок $[a, b]$ разбивается на N отрезков длины $h = \frac{b-a}{N}$
- на каждом отрезке применяется квадратурная формула, результаты складываются
- если квадратурная формула имеет порядок точности p, то для составной формулы получается порядок $p - 1$:

$$E = N\,O(h^p) = \frac{b-a}{h} O(h^p) = O(h^{p-1})$$

Пример расчета по составной формуле

Квадратурные формулы Гаусса-Кристофеля

- Интерполяционные формулы по *любым* n узлам имеют алгебраическую точность $\geq n-1$
- Можно ли подобрать узлы так, чтобы точность была $2n-1$?

$$\sum_{i=1}^{n} d_i x_i^k = \int_{-1}^{1} x^k \, dx \quad k = 0, \ldots, 2n-1$$

Теорема

Для любого числа узлов n существует единственная квадратурная формула с алгебраической точностью $2n-1$.

Доказательство

- $\omega_n = \prod_{j=1}^{n}(x - x_j)$
- Если существует формула точности $2n$, то
$$I(\omega^2) = \sum_i d_i \omega^2(x_i) = 0 \text{ – противоречие } (I(\omega^2) > 0)$$
- Если формула имеет точность $2n - 1$, то
$$I\big(\omega_n(x) r_{n-1}(x)\big) = S\big(\omega_n(x) r_{n-1}(x)\big) = 0, \ \forall\, r_{n-1} \in \mathcal{P}_{n-1}$$
- $\Rightarrow \omega_n$ - n-й ортогональный многочлен с корнями x_k
- Веса вычисляются по формуле: $d_i = \frac{b-a}{2} \int_{-1}^{1} l_i(t)\, dt$
- Докажем, что для $p_{2n-1} = q_{n-1}\omega_n + r_{n-1}$ ф-ла точна:
$$I(p_{2n-1}) = I(q_{n-1}\omega_n) + I(r_{n-1}) =$$
$$= S(r_{n-1}) = S(q_{n-1}\omega_n) + S(r_{n-1}) = S(p_{2n-1}) \quad \square$$

Пример расчета по формуле Гаусса

Формулы Гаусса-Кристофеля с весовой функцией

- Формулы Гаусса можно построить для интегралов вида:

$$I(f) = \int_{-1}^{+1} w(x) f(x)\, dx$$

  где $w(x) \geq 0$
- Весовая функция w порождает скалярное произведение и систему ортогональных многочленов
- Весовая функция может иметь особенность:

$$w(x) = \frac{1}{\sqrt{1 - x^2}}$$

Правило Рунге

- Рассмотрим формулу S_1 алгебраической точности $n-1$ на отрезке длины h, c - середина отрезка
- Разложим $f(x)$ в ряд Тейлора в точке c:

$$I(f) - S_1(f) = \alpha f^{(n)}(c) h^{n+1} + O(h^{n+2})$$

- Пусть S_2 - составная формула по 2-м половинкам того же отрезка, тогда:

$$I(f) - S_2(f) = \alpha f^{(n)}(c) \frac{h^{n+1}}{2^n} + O(h^{n+2})$$

- С точностью до членов $O(h^{n+2})$ получаем:

$$I(f) - S_2(f) \approx \frac{S_2 - S_1}{2^n - 1}$$

Вычисление несобственных интегралов

$$\int_a^b f(x)\,dx, \quad f(x) \to \infty \text{ при } x \to a \qquad \int_0^1 \frac{\cos x}{\sqrt{x}}\,dx$$

- Замена переменной $x = t^2 \Rightarrow I = 2\int_0^1 \cos t^2\,dt$
- Интегрирование по частям

$$I = \int_0^1 \frac{\cos x}{\sqrt{x}}\,dx = 2\sqrt{x}\cos x\Big|_0^1 + 2\int_0^1 \sqrt{x}\sin x\,dx$$

  - Второй интеграл можно вычислить по квадратурной формуле
  - 2-я производная $\sqrt{x}\sin x$ не ограничена, можно ещё раз проинтегрировать по частям

Вычисление несобственных инт-в (2)

- Выделение особенности:

$$I = I_1 + I_2 = \int_0^{\delta} \frac{\cos x}{\sqrt{x}}\,dx + \int_{\delta}^{1} \frac{\cos x}{\sqrt{x}}\,dx$$

$$I_1 = \int_0^{\delta} \frac{\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}}{\sqrt{x}}\,dx = \sum_{k=0}^{n} \frac{(-1)^k \delta^{2k+1/2}}{(2k)!(2k+1/2)} + R$$

$|R|$ не больше последнего члена частичной суммы

- Использование квадратурных формул для интеграла с весом w:

$$I = \int_0^1 w(x)\cos x\,dx, \;\; w(x) = \frac{1}{\sqrt{x}}$$

Заключение

- Численное дифференцирование:
  - Дифференцирование интерполяционного многочлена
  - Метод неопределенных коэффициентов
  - Оптимальный шаг
- Численное интегрирование
  - Интерполяционные квадратурные ф-лы
  - Составные квадратурные ф-лы
  - Квадратурные ф-лы Гаусса-Кристофеля
  - Вычисление несобственных интегралов

Вычислительная математика.
Методы приближения функций многих переменных

МФТИ

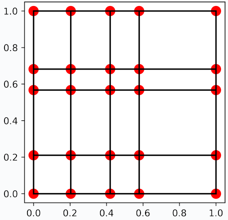
Содержание

Изученные методы приближения

- Лагранжева интерполяция
- Сплайн интерполяция
- Приближение многочленами в L_2, C нормах

Как распространить эти методы на многомерный случай $f(x_1, \ldots, x_d)$?

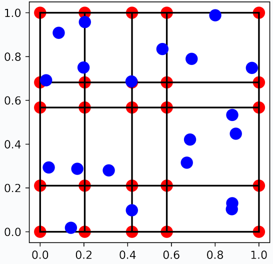
Лагранжева интерполяция в 2D



- Дано $\boldsymbol{x}_k \in \{(x_{1,i}\}_{i=0}^{n} \times \{(x_{2,i}\}_{i=0}^{n}$, $f_{i_1,i_2} = f(x_{1,i_1}, x_{2,i_2})$.
- $L_n(x_1, x_2) = \sum\limits_{i=0}^{n} f_{i_1,i_2} l_{i_1}^{(1)}(x_1) l_{i_2}^{(2)}(x_2)$
- $l_{i_1}^{(1)}$, $l_{i_2}^{(2)}$ – одномерные многочлены Лагранжа.

Проблемы в пространствах высокой размерности

- $||f - f_n|| = O(n^{-r/d})$, r - показатель гладкости, d - размерность, n – количество точек или размер базиса.
- Во многих задачах значения функции известны в точках не на структурированной сетке



Вырожденность интерполяционной матрицы

Теорема (Mairhuber–Curtis theorem)

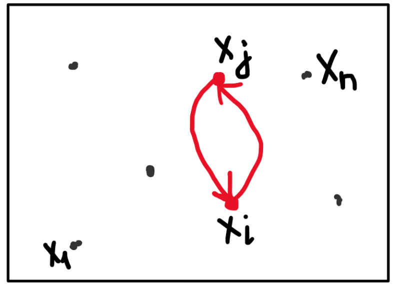
Для любого набора линейно независимых функций $\{\phi_k : \mathbb{R}^2 \to \mathbb{R}\}$, $k = 1, \ldots, n$, $d > 1$ существует набор точек $x_1, \ldots, x_n$, при котором матрица

$$A = \begin{bmatrix} \phi_1(x_1) & \ldots & \phi_n(x_1) \\ \ldots & & \\ \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix}$$

является вырожденной

Вырожденность матрицы (доказательство)

- Возьмем произвольный набор точек, допустим что $det(A) \neq 0$
- Начнём непрерывно менять местами x_i, x_j, $det(A)$ поменяет знак, $\Rightarrow det(A) = 0$ в каком-то промежуточном положении $\square$.



Радиальные базисные функции

- Функция $g : \mathbb{R}^d \to \mathbb{R}$ называется *радиальной*, если её значение зависит только от нормы аргумента:

$$g(\mathbf{x}) = \phi(||\mathbf{x}||) = \phi(r), \ \phi : [0, \infty] \to \mathbb{R}.$$

  Будем называть ϕ *радиальной базисной функцией (РБФ)*

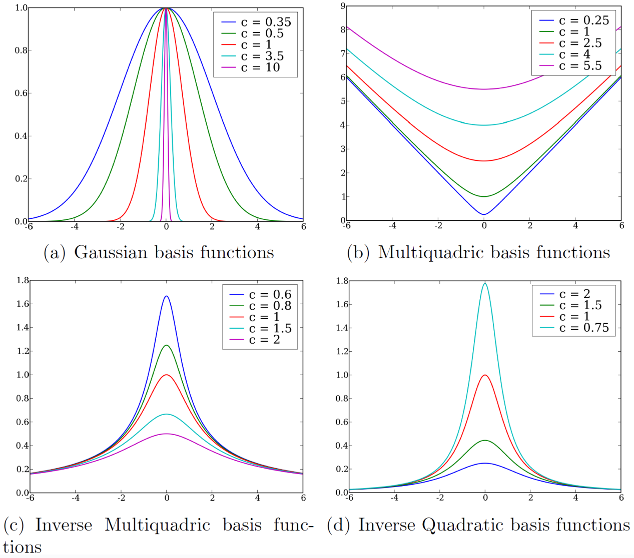
- Интерполяция РБФ: $s(\mathbf{x}) = \sum_{i=1}^{n} c_i \phi(||\mathbf{x} - \mathbf{x}_i)$

$$\begin{bmatrix} \phi(||\mathbf{x}_1 - \mathbf{x}_1||) & \phi(||\mathbf{x}_2 - \mathbf{x}_1||) & \dots & \phi(||\mathbf{x}_n - \mathbf{x}_1||) \\ \vdots & & & \\ \phi_(||\mathbf{x}_1 - \mathbf{x}_n||) & \phi(||\mathbf{x}_2 - \mathbf{x}_n||) & \dots & \phi(||\mathbf{x}_n - \mathbf{x}_n||) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

Примеры РБФ

Radial Basis Function	$\phi(r)$	**parameters**	**order**
Gaussians	$e^{-(cr)^2}$	$c > 0$	0
Polyharmonic Splines	r^{2k-1}	$k \in \mathbb{N}$	$m = k$
	$r^{2k}\log(r)$	$k \in \mathbb{N}$	$m = k+1$
Multiquadrics	$\sqrt{r^2 + c^2}$	$c > 0$	1
Inverse Multiquadrics	$\frac{1}{\sqrt{r^2+c^2}}$	$c > 0$	0
Inverse Quadratics	$\frac{1}{r^2+c^2}$	$c > 0$	0

Графики РБФ



(a) Gaussian basis functions

(b) Multiquadric basis functions

(c) Inverse Multiquadric basis functions

(d) Inverse Quadratic basis functions

Пример интерполяции

Гауссовский процесс

Гауссовский процесс

Случайный процесс f называется гауссовским, если для любого конечного набора $x_1, \dots, x_k$ вектор $f(x_1), \dots, f(x_k)$ имеет многомерное нормальное распределение.

- Чтобы определить г.п. нужно задать:
  - среднее значение $m(x)$
  - ковариационную функцию $k(x_1, x_2) = \mathbb{E}(f(x_1) - m(x_1))(f(x_2) - m(x_2)))$
  - Пример к.ф.

$$k(x, y) = \sigma^2 \exp\left(-\frac{\|x - y\|^2}{2l^2}\right)$$

σ^2 – дисперсия, l - пространственный масштаб.

Регрессия на основе г.п. (кригинг)

- Дано $(X, \mathbf{y}) = \{(x_i, y_i), x_i \in \mathbb{R}^d, y_i \in \mathbb{R}\}_{i=1}^N$
- Предполагается, что $y = f(x) + \varepsilon$, где $f(x)$ - *гауссовский случайный процесс*, $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$
- С помощью Байесовского подхода выводится апостериорное распределение величины $y_* = f(x_*)$:

$$y_* | X, \mathbf{y}, x_* \sim \mathcal{N}(m(x_*), \sigma(x_*))$$

$$m(x_*) = \mathbf{k}^T \mathbf{K}_y^{-1} \mathbf{y} = \sum_{i=1}^N \alpha_i k(x_*, x_i)$$

$$\sigma^2(x_*) = k(x_*, x_*) - \mathbf{k}^T \mathbf{K}_y^{-1} \mathbf{k}$$

$$\mathbf{k} = (k(x_*, x_1), \ldots, k(x_*, x_N))^T, \; \mathbf{K}_y = [k(x_i, x_j)]_{i,j=1}^N + \sigma_n^2 \mathbf{I}$$

Оптимизационный подход к приближению функций

- Возьмём какой-то достаточно богатый параметризованный класс функций $g(\boldsymbol{x}; \boldsymbol{w})$, где $\boldsymbol{w}$ – вектор параметров
- Введём функционал ошибки, например MSE (mean squared error):

$$L(\boldsymbol{w}) = \frac{1}{n} \sum_{i=1}^{n} (f(\boldsymbol{x}_i) - g(\boldsymbol{x}_i; \boldsymbol{w}))^2$$

- Решим задачу оптимизации: $\boldsymbol{w}^* = \arg\min_{\boldsymbol{w}} L(\boldsymbol{w})$

Оптимизационный подход (2)

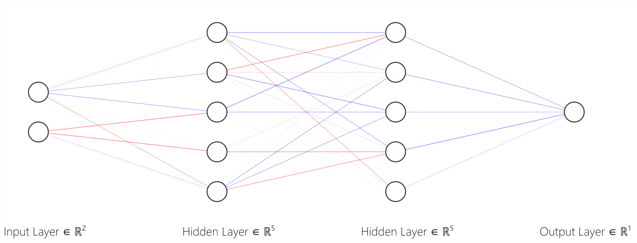
- Во многих классических методах (интерполяция и т.п.) $g(\boldsymbol{x}; w) = \sum_i w_i g_i(\boldsymbol{x})$, т.е. семейство функций это линейная оболочка базовых функций.
- Оптимизационный подход в таких случаях избыточен, т.к. можно вывести аналитическое решение
- Оптимизационный подход позволяет легко добавлять дополнительные ограничения любого вида.
- Оптимизационный подход позволяет использовать нелинейные семейства, например: $f(x) \approx c_1 \sin(c_2 x)$.
- Функционал L можно получать из вероятностных предпосылок (machine learning)

Искусственные нейронные сети

- С точки зрения теории приближений, нейронные сети определенной архитектуры – это нелинейное семейство функций, например:

$$N(\boldsymbol{x}; \boldsymbol{w}) = W_3\sigma(W_2\sigma(W_1\boldsymbol{x} + \boldsymbol{b}_1) + \boldsymbol{b}_2) + \boldsymbol{b}_3$$

$W_k, \boldsymbol{b}_k$ - матрицы и векторы весов, σ - нелинейная функция, например $\sigma(z) = \frac{1}{1+e^{-z}}$



Вычислительная математика. Методы решения нелинейных уравнений.

МФТИ

Основная задача

Найти изолированный корень уравнения $f(x) = 0$

- Если f - непрерывная функция, то z принадлежит отрезку, на концах которого $f(x)$ разного знака
- Простейшие методы: деление пополам, метод золотого сечения
- Итерационные методы: используют значения функции и производных в предыдущих приближениях

Критерий остановки

Теорема

Если f' непрерывна и $\forall y \in [x_k - \epsilon, x_k + \epsilon]$ верно:

$$|f(x_k)/f'(y)| \leq \epsilon$$

то $f(z) = 0$ для некоторого $z \in [x_k - \epsilon, x_k + \epsilon]$

- По теореме Лагранжа $f(x_k + t) = f(x_k) + f'(y(t))t$
- Следовательно

$$\frac{f(x_k + t)}{f'(y)} = \frac{f(x_k)}{f'(y)} + t \geq 0 \text{ при } t = \epsilon, \leq 0 \text{ при } t = -\epsilon$$

- Производная не меняет знак
  $\Rightarrow f(x_k + \epsilon)f(x_k - \epsilon) \leq 0 \quad \square.$

Метод простой итерации

- $f(x) = 0 \iff x = F(x)$: $x_{k+1} = F(x_k), k = 0, 1, \dots$
- Решение z уравнения $z = F(z)$ называется *неподвижной точкой отображения* F.

Теорема о сжимающих отображениях

Пусть M полное метрическое пространство с расстоянием $\rho(\cdot, \cdot)$. $F : M \to M$ – сжимающее:

$$\rho(F(x), F(y)) \le q\rho(x, y), 0 < q < 1 \quad \forall x, y \in M$$

Тогда $x = F(x)$ имеет единственное решение z и $\forall x_0$ МПИ сходится к z:

$$\rho(x_k, z) \le \frac{q^k}{1 - q}\rho(x_1, x_0)$$

Доказательство теоремы

- При $m \geq k$

$$\rho(x_m, x_k) \leq \sum_{i=k}^{m-1} \rho(x_{i+1}, x_i) \leq$$

$$\sum_{i=k}^{m-1} q^i \rho(x_1, x_0) \leq \frac{q^k}{1-q} \rho(x_1, x_0)$$

- $\Rightarrow \{x_k\}$ - фундаментальная $\Rightarrow$ в силу полноты M она сходится к некоторому $z \in M$.
- $0 \leq \rho(x^{k+1}, F(z)) = \rho(F(x^k), F(z)) \leq q\rho(x^k, z) \Rightarrow z = F(z)$
- Если $z_1 = F(z_1) : \rho(z_1, z) = \rho(F(z_1), F(z)) < \rho(z_1, z) \Rightarrow z_1 = z$
- Переход к $m \to \infty$ даёт оценку ошибки $\square$.

Пример применения теоремы

- Задача Коши: $y'(x) = f(y, x),\ y(0) = y_0, x \in [0, x_1]$
- $y(x) = y_0 + \int_0^x f(y(t), t)\, dt$
- $u^{k+1}(x) = y_0 + \int_0^x f(u^k(t), t)\, dt$
- $f(y, \cdot) \in C[0, x_1]$ и $|f(y_2, x) - f(y_1, x)| \le L|y_2 - y_1|$
- $\rho(y_2, y_1) = \max\limits_{0 \le x \le x_1} e^{-L_1 x}|y_2(x) - y_1(x)|$

$$\rho(u^k, u^{k+1}) = \max_x e^{-L_1 x}\left| \int_0^x \Big(f(u^k, t) - f(u^{k-1}, t) \Big)\, dt \right| \le$$

$$L \max_x \int_0^x e^{-L_1(x-t)} e^{-L_1 t}|u^k - u^{k-1}|\, dt \le$$

$$\Big[(1 - e^{-L_1 x_1}) L / L_1 \Big]\, \rho(u^k, u^{k-1})$$

МПИ в $\mathbb{R}^1$

- Пусть $F \in C^1[z-\delta, z+\delta]$, z – единственная неподвижная точка
- Если $|F'(z)| < 1$, то в некоторой δ-окрестности

  $$q = \max_{|x-z|\le\delta} |F'(x)| < 1 \;\Rightarrow |F(x)-F(y)| \le q|x-y| \; \forall x, y \in [z-\delta, z+\delta]$$

- Следовательно, F - сжимающее отображение на $M = [z-\delta, z+\delta]$, МПИ сходится $\forall x_0 \in M$

МПИ в $\mathbb{R}^n$

$$F'(x) = \begin{bmatrix} \frac{\partial F_1(x)}{\partial x_1} & \dots & \frac{\partial F_1(x)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial F_n(x)}{\partial x_1} & \dots & \frac{\partial F_n(x)}{\partial x_n} \end{bmatrix}$$

Достаточное условие сходимости

Пусть отображение $F : \mathbb{R}^n \to \mathbb{R}^n$ имеет единственную неподвижную точку $z = F(z)$ и непрерывно дифференцируемо в некоторой её окрестности.
Тогда, если $\rho(F'(z)) < 1$, то $\forall x_0$ из некоторой окрестности точки z метод простой итерации сходится к z.

Метод простой итерации с параметром

- $F(x) = x - \alpha(x)f(x)$
- Если z изолированный *простой* корень ($f'(z) \neq 0$), то α всегда можно выбрать так, чтобы выполнялось $|F'(z)| < 1$.
- Если известны оценки $|f'(x)|$, то можно выбрать *оптимальное* значение α (как в м-де Ричардсона).
- Нулевое значение $|F'(z)|$ будет при $\alpha = 1/f'(z)$. Можно взять приближение, тогда получим метод Ньютона:

$$\alpha = \frac{1}{f'(x_k)} \approx \frac{1}{f'(z)} \Rightarrow x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Метод Ньютона

- f приближается многочленом 1-й степени (мн-н Тейлора или интерполяционный мн-н Эрмита)

$$f(x) \approx H(x) = f(x_k) + f'(x_k)(x - x_k)$$

- Пусть $f \in C^2$, $f'(z) \neq 0$. Тогда

$$f(z) - H(z) = \frac{f''(\xi_k)}{2}(z - x_k)^2$$
$$H(x_{k+1}) - H(z) = f'(x_k)(x_{k+1} - z)$$

- Т.к. $f(z) = 0, H(x_{k+1}) = 0$, получаем оценку ошибки:

$$\boxed{e_{k+1} = -\frac{f''(\xi_k)}{2f'(x_k)}e_k^2}, \ e_k = z - x_k$$

Метод Ньютона, свойства

- Метод Ньютона в лучшем случае сходится *квадратично*.
- По определению последовательность x_k сходится к z с порядком p, если

$$\lim_{k\to\infty} \sup \left| \frac{e_{k+1}}{e_k^p} \right| \leq c \leq \infty$$

- Условие $f'(z) \neq 0$ означает, что корень является простым. В общем случае, z - корень кратности m, если $f^{(j)}(z) = 0$ при $0 \leq j \leq m-1$ и $f^{(m)}(z) \neq 0$.
- Метод Ньютона может сходиться и для кратного корня, но сходимость может не быть квадратичной. Для $f(x) = x^2$, $e_{k+1} = e_k/2$.

Сходимость метода Ньютона

Теорема

Пусть z - простой корень уравнения $f(x) = 0$ и

1. $f \in C^2[z-\delta, z+\delta]$
2. $f'(x) \neq 0$ при $x \in [z-\delta, z+\delta]$
3. $\gamma \equiv \dfrac{\max\limits_{|x-z|\le\delta} |f''(x)|}{2 \min\limits_{|x-z|\le\delta} |f'(x)|} \neq 0$

Тогда $\forall \epsilon : \quad 0 < \epsilon < \min\{\delta, \gamma^{-1}\}$ метод Ньютона сходится при любом начальном приближении $x_0 \in [z-\epsilon, z+\epsilon]$ и верно:

1. $|e_{k+1}| \le \gamma |e_k|^2$
2. $|e_k| \le \gamma^{-1}(\gamma|e_0|)^{2^k}$

Доказательство

- Пусть $x_k \in [z-\epsilon, z+\epsilon]$. Т.к. $\epsilon < \gamma^{-1}$:

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(x_k)} e_k^2 \Rightarrow$$
$$|e_{k+1}| \leq \gamma |e_k|^2 \leq (\gamma\epsilon)\epsilon \leq \epsilon \Rightarrow x_{k+1} \in [z-\epsilon, z+\epsilon]$$

- Умножим обе части на γ и обозначим $d_k = \gamma |e_k|$, тогда
$$d_{k+1} \leq d_k^2 \Rightarrow d_k \leq d_0^{2^k}$$
- Из условий на начальное приближение ($|x-z| < \gamma^{-1}$) следует, что $d_0 < 1 \Rightarrow d_k \to 0$ $\square$
- Следствие: $\lim\limits_{k\to\infty} \dfrac{e_{k+1}}{e_k^2} = -\dfrac{f''(z)}{2f'(z)}$

Пример: квадратный корень из числа

Метод Ньютона в $\mathbb{R}^n$

$$x_{k+1} = x_k - \left[f'(x_k)\right]^{-1} f(x_k)$$

Теорема

$f : \mathbb{R}^n \to \mathbb{R}^n$, $f(z) = 0$ и в $B_\delta(z) = \{x : \|x - z\|_\infty \leq \delta\}$ f - непрерывно дифференцируемо, якобиан f существует и невырожден, и $\forall x, y \in B_\delta$:

$$\|f'(x) - f'(y)\|_\infty \leq c\|x - y\|_\infty, \; c > 0$$

Пусть $\gamma = c \max\limits_{\|z-x\|_\infty \leq \delta} \|[f'(x)^{-1}]\|_\infty$, $0 < \epsilon < \min\{\delta, \gamma^{-1}\}$. Тогда $\forall x_0 \in B_\epsilon(z)$ метод Ньютона сходится и

$$\|e_{k+1}\|_\infty \leq \gamma \|e_k\|_\infty^2, \quad \|e_k\|_\infty \leq \gamma^{-1}(\gamma\|e_0\|_\infty)^{2^k}$$

Доказательство

- По теореме Лагранжа найдутся точки $\xi_1, \dots, \xi_n$:

$$f(x) - f(y) = J_k(x-y), \quad J_k = \begin{bmatrix} \frac{\partial f_1(\xi_1)}{\partial x_1} & \dots & \frac{\partial f_1(\xi_1)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_n(\xi_n)}{\partial x_1} & \dots & \frac{\partial f_n(\xi_n)}{\partial x_n} \end{bmatrix}$$

- $e_{k+1} = e_k - [f'(x_k)]^{-1}(f(z) - f(x_k)) =$
  $e_k - [f'(x_k)]^{-1} J_k e_k = [f'(x_k)]^{-1}(f'(x_k) - J_k)e_k$
- По условию Липшица

$$\|f'(x_k) - J_k\|_\infty \le c \max_{1 \le j \le n} \|x_k - \xi_j\|_\infty \le c\|x_k - z\|_\infty$$

- $\|e_{k+1}\|_\infty = \|[f'(x_k)]^{-1}(f'(x_k) - J_k)e_k\|_\infty \le \gamma \|e_k\|_\infty^2$

Метод Ньютона, решение линейной системы

- $x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k)$
- $[f'(x_k)]\,\Delta x_{k+1} = -f(x_k), \quad \Delta x_{k+1} = x_{k+1} - x_k$
- При больших n – большая вычислительная сложность
- $[f'(x_k)]$ можно обновлять не на каждой итерации:

$$[f'(x_0)]\,\Delta x_{k+1} = -f(x_k)$$

- Имеет смысл использовать итерационные методы:
  - $\Delta x_{k+1} \approx 0$ – хорошее начальное приближение
  - Не нужно решать систему точно
  - $[f'(x_k)]\,\Delta x \approx f(x_k + \Delta x) - f(x_k)$ – можно не хранить матрицу Якоби

Метод Ньютона и методы оптимизации

- Рассмотрим функционал $f : \mathbb{R}^n \to \mathbb{R}$
- Если $f \in C^2$ имеет единственную точку минимума – градиент в этой точке равен нулю:

$$\operatorname{grad} f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix} = 0$$

- Для решения можно применить метод Ньютона:

$$x_{k+1} = x_k - [f''(x_k)]^{-1} \operatorname{grad} f(x_k)$$

- $f''(x) = H(x) = [\operatorname{grad} f(x_k)]' = \left[\frac{\partial f}{\partial x_i \partial x_j}\right]$ - гессиан
- $H(x) = H(x)^T$, в точке минимума $H(z) > 0$

Метод секущих

Идея методов высокого порядка

Локально приблизим f или f^{-1} многочленом, и для многочлена решим задачу точно.

- По двум последним приближениям x_{k-1}, x_k построим линейный интерполянт:

$$L(x) = f(x_{k-1})\frac{x - x_k}{x_{k-1} - x_k} + f(x_k)\frac{x - x_{k-1}}{x_k - x_{k-1}}$$

- Корень многочлена:

$$x_{k+1} = x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$$

Скорость сходимости метода секущих

$$x_{k+1} = x_k - f(x_k)\frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}$$

- Пусть z - корень, $e_k \equiv z - x_k$, $f \in C^2$, $f'(z) \neq 0$
-

$$\underbrace{f(z)}_{0} - L(z) = \frac{f''(\xi_k)}{2}(z - x_k)(z - x_{k-1})$$

$$\underbrace{L(x_{k+1})}_{0} - L(z) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(x_{k+1} - z) =$$

$$= f'(\zeta_k)(x_{k+1} - z)$$

- $e_{k+1} = -\dfrac{f''(\xi_k)}{2f'(\zeta_k)}e_k e_{k-1}$

Скорость сходимости м-да секущих(2)

$$e_{k+1} = -\frac{f''(\xi_k)}{2f'(\zeta_k)} e_k e_{k-1}$$

- Предположим, что метод сходится
- для некоторого $\gamma > 0$: $|e_{k+1}| \leq \gamma |e_k||e_{k-1}|$
- Введем величины $d_k = \gamma |e_k|$ и предположим, что $d_0 \leq d < 1$, $d_1 \leq d < 1$
- $d_2 \leq d_1 d_0 \leq d^2$
- $d_3 \leq d_2 d_1 \leq d^3 \dots$
- $d_k \leq d^{\phi_k}$, $\phi_0 = \phi_1 = 1$, $\phi_k = \phi_{k-1} + \phi_{k-2}$, $k = 2, \dots$
- $\phi_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^{k+1} - \left(\frac{1-\sqrt{5}}{2} \right)^{k+1} \right) = O(1.618^k)$
- Для метода Ньютона: $d_k \leq d^{2^k}$

Заключение

- Метод простой итерации
- Принцип сжимающих отображений
- Метод Ньютона
- Идея построение методов высокого порядка
- Метод секущих

Вычислительная математика.
Численные методы решения ОДУ.

МФТИ

Основная задача

- Найти приближенное решение краевой задачи, например:

$$u_{xx}(x) = f(x, u)$$
$$u(0) = a$$
$$u(1) = b$$

- Найти приближенное решение задачи Коши, например:

$$u_t = f(t, u)$$
$$u(0) = u_0$$

Пример численного метода

$$u_{xx}(x) = f(x)$$
$$u(0) = a, \; u(1) = b$$

- Вместо непрерывной функции ищем $U_0, \ldots, U_{m+1}$, $U_j \approx u(x_j)$, $x_j = jh$, $h = 1/(m+1)$
- Заменим 2-ю производную конечной разностью:

$$\frac{1}{h^2}(U_{j-1} - 2U_j + U_{j+1}) = f(x_j), \quad j = 1, 2, \ldots, m$$

- 1-е и m-е уравнения содержат известные значения $U_0 = a$, $U_{m+1} = b$

Пример численного метода (2)

Получим систему линейных уравнений вида $AU = F$

$$\underbrace{\frac{1}{h^2}\begin{bmatrix} -2 & 1 & & & & \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \end{bmatrix}}_{A}, \quad \underbrace{\begin{bmatrix} f(x_1) - a/h^2 \\ f(x_2) \\ f(x_3) \\ \vdots \\ f(x_{m-1}) \\ f(x_m) - b/h^2 \end{bmatrix}}_{F}$$

- Насколько точно U приближает u?

Норма ошибки, сходимость

- Вектор ошибки:

$$\hat{U} = \begin{bmatrix} u(x_1) \\ u(x_2) \\ \vdots \\ u(x_m) \end{bmatrix}, \quad E = U - \hat{U}$$

- Нужно оценить норму ошибки, например:

$$\|E\|_\infty = \max_j |E_j|, \quad \|E\|_1 = h \sum_{j=1}^{m} |E_j|$$

Определение (сходимость)

Численное решение *сходится* к точному с порядком p, если $\|E\| = O(h^p)$, $h \to 0$

Невязка, аппроксимация

- КР формула имеет 2-й порядок аппроксимации:

$$\frac{1}{h^2}\left(u(x_{j-1}) - 2u(x_j) + u(x_{j+1})\right) - f(x_j) = r_j =$$
$$\frac{1}{12}h^2 u^{(4)}(x_j) + O(h^4) = O(h^2)$$

- Введем *вектор невязки* r с компонентами r_j:

$$r = A\hat{U} - F, \quad A\hat{U} = F + r, \quad \|r\| = O(h^2)$$

Определение (аппроксимация)

Разностная задача *аппроксимирует* дифференциальную задачу с порядком p, если $\|r\| = O(h^p)$

Связь ошибки и невязки, устойчивость

$$\begin{cases} A_m U = F \\ A_m \hat{U} = F + r \end{cases}, \; A_m E = -r \;, E = A_m^{-1} r \;, \|E\| \le \|A_m^{-1}\| \|r\|$$

- $h \to 0$, $m \to \infty$
- Если $\|A_m^{-1}\| \le C \quad \forall m > m_0$:

$$\|E\| \le \|A_m^{-1}\| \|r\| \le C \|r\| = O(h^2)$$

Определение (устойчивость)

Численный метод называется устойчивым, если $\exists m_0 : \|A_m^{-1}\| \le C \quad \forall m > m_0$, C не зависит от m.

Основная теорема

Теорема Лакса-Рябенького

Если численный метод *аппроксимирует* дифференциальную задачу с порядком p и является *устойчивым*, то численное решение *сходится* к точному решению с порядком p.

- Аппроксимация $+$ Устойчивость $\Rightarrow$ Сходимость
- **Аппроксимацию и устойчивость можно проверить, не зная точное решение**

Пример исследования устойчивости

- $A = A^T$, $\|A\|_2 = \max_p |\lambda_p|$, $\|A^{-1}\| = (\min |\lambda_p|)^{-1}$

$$\lambda_p = \frac{2}{h^2}\left(\cos(p\pi h) - 1\right), \; u_j^p = \sin(p\pi jh), \; j = \overline{1, m}$$

$$(Au^p)_j = \frac{1}{h^2}(u_{j-1}^p - 2u_j^p + u_{j+1}^p) =$$

$$= \frac{1}{h^2}(\sin(p\pi(j-1)h) - 2\sin(p\pi jh) + \sin(p\pi(j+1)h)) =$$

$$// \sin(a \pm b) = \sin a \cos b \pm \cos a \sin b //$$

$$= \frac{1}{h^2}(2\sin(p\pi jh)\cos(p\pi h) - 2\sin(p\pi jh)) = \lambda_p u_j^p$$

$$\lambda_1 = \tfrac{2}{h^2}(\cos(\pi h) - 1) = \tfrac{2}{h^2}(-\tfrac{1}{2}\pi^2 h^2 + O(h^4)) = -\pi^2 + O(h^2)$$

Постановка граничных условий

- Г.у. 2-го рода: $u'(0) = s,\ u(1) = b$
- $\dfrac{U_1 - U_0}{h} = s$ – снизит порядок до 1-го
- Использование главного члена ошибки:

$$\frac{u_1 - u_0}{h} = u'(x_0) + \underline{\frac{h}{2}u''(x_0)} + O(h^2),\ u''(x_0) = f(x_0) \Rightarrow$$

$$\frac{U_1 - U_0}{h} = s + \frac{h}{2}f(x_0)$$ – 2-й порядок

- $\dfrac{3U_0 - 4U_1 + U_2}{2h} = s$

Постановка г.у., вид системы

$$\frac{1}{h^2}\begin{bmatrix} -h & h & 0 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ & & & & 0 & h^2 \end{bmatrix}\begin{bmatrix} U_0 \\ U_1 \\ \vdots \\ U_m \\ U_{m+1} \end{bmatrix} = \begin{bmatrix} s + \frac{f(x_0)}{2}h \\ f(x_1) \\ \vdots \\ f(x_m) \\ b \end{bmatrix}$$

Пример расчета

Нелинейные задачи

- $u_{xx}(x) = f(x, u(x)),\ u(0) = a,\ u(1) = b$
- После КР аппроксимации получается система *нелинейных* уравнений:

$$G(U) = 0,\ G : \mathbb{R}^m \to \mathbb{R}^m$$

- Метод Ньютона:

$$U^{k+1} = U^k - [J(U^k)]^{-1} G(U^k),\ J_{ij}(U) = \frac{\partial}{\partial U_j} G_i(U)$$

$$J = \frac{1}{h^2} \begin{bmatrix} (-2 - h^2 f_u(x_1, U_1)) & 1 & \\ 1 & (-2 - h^2 f_u(x_2, U_2)) & 1 \\ \ddots & \ddots & \ddots \end{bmatrix}$$

Линеаризация исходной задачи

- $u_{xx}(x) = f(x, u(x)),\ u(0) = a,\ u(1) = b$
- Начальное приближение:

$$y^0(x) \approx u(x),\ y^0(0) = a, y^0(1) = b$$

- $u \approx y^k + \delta^k$:

$$y^k_{xx} + \delta^k_{xx} = f(x, y^k(x)) + f_u(x, y^k(x))\delta^k + \cancel{O\left((\delta^k)^2\right)}$$

- Получаем *линейную* задачу с нулевыми г.у. для δ^k:

$$\delta^k_{xx} = f_u(x, y^k)\delta^k + f(x, y^k) - y^k_{xx}$$

Устойчивость в нелинейном случае

$$\begin{cases} G(U) = 0 \\ G(\hat{U}) = r \end{cases}, \Rightarrow G(U) - G(\hat{U}) = -r$$

Определение (устойчивость)

Метод называется устойчивым, если $\forall h < h_0$:

$$\|U - V\| \leq C\|r_1 - r_2\| \text{ при } G(U) = r_1, G(V) = r_2$$

- $G(U) = G(\hat{U}) + J(\hat{U})E + O(\|E\|^2)$
- $J(\hat{U})E = -r + O(\|E\|)^2$
- Аналогичное определение :$\|(\hat{J}^h)^{-1}\| \leq C, \ \forall h < h_0$

Заключение

- Основные понятия
  - Сходимость
  - Аппроксимация
  - Устойчивость
- Пример КР метода
- Постановка г.у. для краевой задачи
- Подходы к решению нелинейных уравнений

Вычислительная математика. Численные методы решения ОДУ, задача Коши.

МФТИ

Численные методы решения задачи Коши

- Задача Коши:

$$
\begin{aligned}
&u_t(t) = f(t,u), \ t \in [0,T] \\
&u(0) = u_0
\end{aligned}
$$

- $t_n = nh, \ n = \overline{0,N}, \ h = T/N, \ U_n \approx u(t_n)$
- Заменим производную конечной-разностью:

$$
\begin{aligned}
&\frac{1}{h}(U_{n+1} - U_n) = f(t_n, U_n), \ \ n = 0, N-1 \\
&U_0 = u_0
\end{aligned}
$$

- $U_{n+1} = U_n + h\,f(t_n, U_n)$ – устойчивость будет определяться свойствами метода на одном шаге .

Метод Эйлера

$$\begin{cases} u'(t) = \lambda u(t) + g(t) \\ u(t_0) = u_0 \end{cases}, \quad u(t) = e^{\lambda(t-t_0)}u_0 + \int_{t_0}^{t} e^{\lambda(t-\tau)}g(\tau)\,d\tau$$

- Метод Эйлера

$$U_{n+1} = U_n + h(\lambda U_n + g(t_n)) = (1 + h\lambda)U_n + hg(t_n)$$

- Ошибка аппроксимации (невязка):

$$r_n = \left(\frac{u(t_{n+1}) - u(t_n)}{h}\right) - (\lambda u(t_n) + g(t_n)) =$$

$$\left(u'(t_n) + \frac{1}{2}hu''(t_n) + O(h^2)\right) - u'(t_n) =$$

$$\frac{1}{2}hu''(t_n) + O(h^2)$$

Устойчивость метода Эйлера

- Связь ошибки и невязки:

$$\begin{cases} U_{n+1} = (1 + h\lambda)U_n + hg(t_n) \\ u(t_{n+1}) = (1 + h\lambda)u(t_n) + hg(t_n) + hr_n \end{cases}$$

$$\Rightarrow \quad E_{n+1} = (1 + h\lambda)E_n - hr_n$$

$$E_n = (1 + h\lambda)^n E_0 - h\sum_{m=1}^{n}(1 + h\lambda)^{n-m} r_{m-1}$$

$$|1 + h\lambda| \le \exp(|\lambda|h) \Rightarrow (1 + h\lambda)^{n-m} \le$$

$$e^{(n-m)h|\lambda|} \le e^{nh|\lambda|} \le e^{|\lambda|T},\ T = Nh$$

$$|E_n| \le e^{|\lambda|T}\left(|E_0| + h\sum_{m=1}^{n}|r_{m-1}|\right) \le$$

$$e^{|\lambda|T}\left(|E_0| + nh\|r\|_\infty\right) = e^{|\lambda|T}\left(|E_0| + T\|r\|_\infty\right)$$

Матричная форма уравнений

$$AU = F$$

$$A = \frac{1}{h}\begin{bmatrix} 1 & & & & \\ -(1+h\lambda) & 1 & & & \\ & & \ddots & & \\ & & -(1+h\lambda) & 1 & \\ & & & -(1+h\lambda) & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_N \end{bmatrix}, \quad F = \begin{bmatrix} (1/h+\lambda)U_0 + g(t_0) \\ g(t_1) \\ \vdots \\ g(t_{N-1}) \end{bmatrix}$$

- $AE = -r$, $\|A^{-1}\| \overset{?}{<} C$

Обратная матрица

$$\underbrace{h\begin{bmatrix} 1 & & & & \\ (1+h\lambda) & 1 & & & \\ (1+h\lambda)^2 & (1+h\lambda) & 1 & & \\ \vdots & & & \ddots & \\ (1+h\lambda)^{N-1} & (1+h\lambda)^{N-2} & \cdots & (1+h\lambda) & 1 \end{bmatrix}}_{A^{-1}}$$

- $\|A^{-1}\|_\infty = h\sum_{m=1}^{N} |(1+h\lambda)^{N-m}|$
- $\|A^{-1}\| \le hNe^{|\lambda|T} = Te^{|\lambda|T}$

Пример неустойчивой схемы

- $u'(t) = \lambda u, u(0) = b,\ t \in [0, 1]$
- Разностная схема:

$$4\frac{U_{n+1} - U_{n-1}}{2h} - 3\frac{U_{n+1} - U_n}{h} = \lambda U_n,\ \ n = 1, \ldots, N-1$$

$$U_0 = b$$

$$U_1 = be^{\lambda h}$$

- 1-й порядок аппроксимации
- Общее решение разностного ур-я: $U_n = c_1 q_1^n + c_2 q_2^n$
- q_1, q_2 - корни уравнения $q^2 + (\lambda h - 3)q + 2 = 0$
- $U_n = \dfrac{U_0}{q_2 - q_1}(q_2 q_1^n - q_1 q_2^n) + \dfrac{U_1}{q_2 - q_1}(-q_1^n + q_2^n)$
- $\max|U_n| \to \infty$ при $h \to 0$

Методы высокого порядка

- *Многошаговые методы*, пример:

$$\frac{U_{n+1} - U_{n-1}}{2h} = f(t_n, U_n)$$
$$U_{n+1} = U_{n-1} + 2hf(t_n, U_n), n = 1, \dots, N$$

- Нужно задать дополнительное условие $U_1 = \dots$
- *Одношаговые многостадийные методы (м-ды Рунге-Кутты)*, пример:

$$k_1 = f(t_n, U_n)$$
$$k_2 = f(t_n + \frac{h}{2}, U_n + \frac{h}{2}k_1)$$
$$U_{n+1} = U_n + hk_2$$

Одношаговые методы (методы Рунге-Кутты)

Общий вид методов Рунге-Кутты:

$$k_i = f(t_n + c_i h, U_n + h \sum_{j=1}^{s} a_{ij} k_j), \; i = 1, \dots, s$$

$$U_{n+1} = U_n + h \sum_{j=1}^{s} b_j k_j$$

Таблица Бутчера:

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ c_2 & a_{21} & a_{22} & \dots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array} = \begin{array}{c|c} \mathbf{c} & A \\ \hline & \mathbf{b}^{\mathsf{T}} \end{array}$$

М-ды Рунге-Кутты, условия порядка

- Упрощающие условия: $\sum_{j=1}^{s} a_{ij} = c_i, i = 1, \dots, s$
- Первый порядок: $\sum_{j=1}^{s} b_j = 1$
- Второй порядок (+ к предыдущим): $\sum_{j=1}^{s} b_j c_j = \frac{1}{2}$
- Третий порядок (+ к предыдущим):

$$\sum_{j=1}^{s} b_j c_j^2 = \frac{1}{3} \quad , \sum_{i=1}^{s} \sum_{j=1}^{s} b_i a_{ij} c_j = \frac{1}{6}$$

Устойчивость методов Р-К

- Пример: $(f(t,u) = f(u))$

$$k_1 = f(U_n),\ k_2 = f(U_n + \frac{h}{2}k_1),\ U_{n+1} = U_n + hk_2$$

$$U_{n+1} = U_n + h\Psi(U_n, t_n, h),\ \Psi(u,t,h) = f(u + \frac{1}{2}hf(u))$$

$$|\Psi(u_1,h) - \Psi(u_2,h)| \le L|u_1 + \frac{1}{2}hf(u_1) - u_2 - \frac{1}{2}hf(u_2)| \le$$

$$L|u_1 - u_2| + \frac{1}{2}hL^2|u_1 - u_2| = (L + \frac{1}{2}hL^2)|u_1 - u_2|$$

- $r_n = (u(t_{n+1}) - u(t_n))/h - \Psi(u(t_n), t_n, h)$
- $u(t_{n+1}) = u(t_n) + h\Psi(u(t_n), t_n, h) - hr_n$
- $|E_{n+1}| = \left|E_n + h\Big(\Psi(U_n, t_n, h) - \Psi(u(t_n), t_n, h)\Big) - hr_n\right| \le$
  $|E_n| + hL'|E_n| + hr_n$

Автоматический выбор шага

- Для начальных данных U_0, t_0 сделаем:
  - Два шага длины h: U_1, U_2
  - Один шаг длины $2h$: $\widetilde{U}_2$
- Для метода порядка p:

$$e_2 = u(t_0 + 2h) - U_2 = 2Ch^{p+1} + O(h^{p+2})$$

$$\widetilde{e}_2 = u(t_0 + 2h) - \widetilde{U}_2 = C(2h)^{p+1} + O(h^{p+2})$$

- $e_2 \approx \dfrac{U_2 - \widetilde{U}_2}{2^p - 1}$
- Чтобы $e_2(h) < \epsilon$: $h_{new} = h \left(\dfrac{\epsilon}{e_2}\right)^{1/(p+1)}$

Подклассы методов Р-К

$$k_i = f(t_n + c_i h, U_n + h\sum_{j=1}^{s} a_{ij}k_j), \ \ i = 1, \dots, s$$

$$U_{n+1} = U_n + h\sum_{j=1}^{s} b_j k_j$$

- Явные методы: $a_{ij} = 0, \ \ j \geq i$
  - Не нужно решать нелинейные уравнения
- Неявные методы: $a_{ij} \neq 0$ при каких-то $j \geq i$
  - Нужно решать систему нелинейных уравнений
- Диагонально неявные $(a_{ij} = 0, \ \ j > i)$ (DIRK)
  - Нужно решать на каждой стадии одну нелинейную систему

Жёсткие (stiff) задачи

- Модельная задача: $u'(t) = \lambda u, \;\; u(0) = a, \; Re(\lambda) \le 0$
- Явный метод Эйлера:

$$\frac{U_{n+1} - U_n}{h} = \lambda U_n \Rightarrow U_{n+1} = (1 + \lambda h) U_n = R(z) U_n$$

- $z = \lambda h$, $R(z)$ - функция устойчивости
- Неявный метод Эйлера

$$\frac{U_{n+1} - U_n}{h} = \lambda U_{n+1} \Rightarrow U_{n+1} = \frac{1}{(1 - \lambda h)} U_n$$

- При каких h $|U_n| \not\to \infty$ при $n \to \infty$?
  - Явный метод: $|R(z)| \le 1 : \;\; |1 + z| < 1$
  - Неявный метод: $\left| \frac{1}{1-z} \right| \le 1$ при любых $Re(z) < 0$

Пример задачи

$$\begin{cases} u'(t) = v(t) \\ v'(t) = -u(t) \end{cases} \quad V = [u, v]^T, V(0) = [1, 0]^T$$

$$V' = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{A} V, \;\; V(t) = [\cos(t), -\sin(t)]^T$$

- Явный метод Эйлера
- Неявный метод Эйлера
- Метод трапеций:

$$\frac{U_{n+1} - U_n}{h} = \frac{1}{2} A(U_n + U_{n+1})$$

Пример расчета