Лекция 2:
“Итерационные методы решения систем”

Кирилл Терехов

ИВМ РАН, ВТМГБ ФПМИ МФТИ, ВТМ ВМК МГУ, Университет Сириус

https://dodo.inm.ras.ru/terekhov/lect1/02/presentation.pdf

15 февраля 2023 г.

Содержание

Лекция 2:
Итерационные методы решения систем

Система линейных уравнений

- Решаем систему вида $A\mathrm{x} = b$
- $A > 0 \in \mathcal{R}^{N \times N}$ - вещественная положительно определенная симметричная матрица размера $N \times N$
- $\mathrm{x}, \mathrm{b} \in \mathcal{R}^{N}$ - вектор неизвестных и правой части размера N

Свойства систем

Определение (симметричная матрица)

Матрица *симметричная*, если $A = A^T$ или $\forall i, j : a_{ij} = a_{ji}$

Определение (положительно определенная)

Симметричная матрица называется *положительно определенной* ($A > 0$), если $x^T A x > 0, \forall x \neq 0 \in \mathcal{R}^N$

Свойство (положительно определенная)

Если $A = E \Lambda E^T$ *- спектральное разложение,* $\Lambda = diag(\lambda_1, \lambda_2, \dots, \lambda_N)$ *- собственные значения, то* $\lambda_1 > \lambda_2 > \dots > \lambda_N > 0 \in \mathcal{R}$

Свойство (Число обусловленности)

Для некоторой матричной нормы число обусловленности определяется по формуле $\mu(A) = \|A^{-1}\| \|A\| \geq 1$. *Для* $A > 0 : \mu(A) = \lambda_1 / \lambda_N$.

Метод простой итерации (итерации Ричардсона)

$$\frac{1}{\tau}\left(\mathrm{x}_{k+1}-\mathrm{x}_k\right)+A\mathrm{x}_k=\mathrm{b} \quad \rightarrow \quad \mathrm{x}_{k+1}=\mathrm{x}_k+\tau\left(\mathrm{b}-A\mathrm{x}_k\right)=\mathrm{x}_k+\tau\mathrm{r}_k$$

Определение (невязка)

$\mathrm{r}_k=\mathrm{b}-A\mathrm{x}_k$ - *невязка* на k-ой итерации.

Метод сходится при $0<\tau<2/\lambda_1$, где λ_1 - наибольшее собственное значение.
Сходится оптимально при $\tau=(\lambda_1+\lambda_N)/2$ (итерация Чебышева), скорость сходимости оценивается по норме (чем меньше, тем лучше сходится):

$$\|\mathbb{I}-\tau A\|=\frac{\lambda_1-\lambda_N}{\lambda_1+\lambda_N}=\frac{\mu(A)-1}{\mu(A)+1}$$

Персоналии

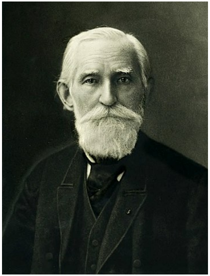

Рис.: Пафнутий Львович Чебышев 1821–1894.

Рис.: Алексей Николаевич Крылов 1863–1945.

Рис.: Борис Григорьевич Галеркин 1871–1945.

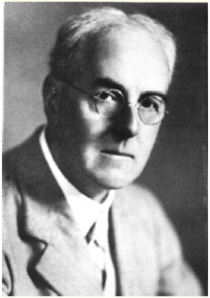

Рис.: Льюис Фрай Ричардсон 1881–1953.

Проекционные методы

Введем $\mathcal{K}, \mathcal{L}$ - два m-мерных подпространства $\mathcal{R}^N$. Пусть x_k - приближение на k-ой итерации, а x_{k+1} - следующее приближение.

Проекционный метод выбирает $\mathrm{x}_{k+1} = \mathrm{x}_k + \delta\mathrm{x}$ такое, что $\delta\mathrm{x} \in \mathcal{K}$, а невязка $\mathrm{r}_{k+1} = \mathrm{b} - A\mathrm{x}_{k+1} \perp \mathcal{L}$ - ортогональна пространству $\mathcal{L}$. Все это звучит довольно странно, пока мы не ввели $\mathcal{K}$ и $\mathcal{L}$.

Проекционные методы

Теорема

Для произвольного вектора $\mathrm{z} \in \mathcal{R}^N$ *вектор* $\mathrm{k}^\star$ *является решением задачи*

$$\mathrm{k}^\star = \arg\min_{\mathrm{k} \in \mathcal{K}} \langle \mathrm{z} - \mathrm{k}, \mathrm{z} - \mathrm{k} \rangle$$

тогда и только тогда, когда

$$\langle \mathrm{z} - \mathrm{k}^\star, \mathrm{y} \rangle = 0, \quad \forall \mathrm{y} \in \mathcal{K}.$$

Доказательство.

Пусть P - ортогональный проектор на $\mathcal{K}$: $P\mathrm{z} \in \mathcal{K}$, $\mathrm{z} - P\mathrm{z} \in \mathcal{K}^\perp$. Тогда

$$\langle \mathrm{z} - \mathrm{k}, \mathrm{z} - \mathrm{k} \rangle = \|\mathrm{z} - \mathrm{k}\|^2 = \|\mathrm{z} - P\mathrm{z} + P\mathrm{z} - \mathrm{k}\|^2 = \|\mathrm{z} - P\mathrm{z}\|^2 + \|P\mathrm{z} - \mathrm{k}\|^2$$

$\Longrightarrow$ $\|\mathrm{z} - \mathrm{k}\|^2 \geq \|\mathrm{z} - P\mathrm{z}\|^2$ а равенство при $\mathrm{k}^\star = P\mathrm{z}$ $\square$

Проекционные методы

Задача (Минимизация функционала)

Пусть $\mathcal{L} = \mathcal{K}$. *Задача минимизации энергетической нормы ошибки:*

$$\forall \mathrm{x} \in \mathcal{R}^N : \mathrm{x}_{k+1} = \underset{\tilde{\mathrm{x}} - \mathrm{x}_k \in \mathcal{K}}{\arg\min} \langle \mathrm{x} - \tilde{\mathrm{x}}, \mathrm{x} - \tilde{\mathrm{x}} \rangle_A$$

Определение (матричное скалярное произведение)

Для $A > 0$: $\langle \mathrm{x}, \mathrm{y} \rangle_A = \langle A\mathrm{x}, \mathrm{y} \rangle = \langle \mathrm{x}, A\mathrm{y} \rangle$.

Определение (метод Галеркина)

По теореме перейдем к проекционной задаче $\forall \mathrm{y} \in \mathcal{K} : \langle \mathrm{z} - \mathrm{k}^\star, \mathrm{y} \rangle_A = 0$ с $z = \mathrm{x} - \mathrm{x}_k \in \mathcal{K}$ и $\mathrm{k}^\star = \mathrm{x}_{k+1} - \mathrm{x}_k (= \tilde{\mathrm{x}})$:

$$\langle \mathrm{z} - \mathrm{k}^\star, \mathrm{y} \rangle_A = \langle A\left(\mathrm{x} - \mathrm{x}_k - \mathrm{k}^\star\right), \mathrm{y} \rangle = \langle \mathrm{b} - A\mathrm{x}_{k+1}, \mathrm{y} \rangle$$

$\implies \mathrm{x}_{k+1}$ - результат проекционного алгоритма.

Метод сопряженных градиентов: CG

Определение (Подпространство Крылова)

Подпространство векторов $\mathcal{K}^m = \{\mathrm{r}_0, A\mathrm{r}_0, \ldots, A^{m+1}\mathrm{r}_0\}$ называется *подпространством Крылова* порядка m, где x_0 - начальное приближение и $\mathrm{r}_0 = \mathrm{b} - A\mathrm{x}_0$ - начальная невязка.

Метод сопряженных градиентов - проекционный метод с $\mathcal{L} = \mathcal{K} = \mathcal{K}^m$ и $A > 0$. Каждое следующее приближение имеет вид

$$\mathrm{x}_m = \mathrm{x}_0 + \sum_{i=1}^{m} c_i A^{i-1}\mathrm{r}_0,$$

где коэффициенты c_i определяются условием $\mathrm{r}_m = \mathrm{b} - A\mathrm{x}_m \perp \mathcal{K}^m$.

Метод сопряженных градиентов: CG

Допустим, что в $\mathcal{K}^m$ известен A-ортогональный базис:

$$\{\mathrm{k}_1, \ldots, \mathrm{k}_m\} \in \mathcal{K}^m, \quad \mathrm{k}_1 = \mathrm{r}_0, \quad \forall i \neq j : \langle \mathrm{k}_i, \mathrm{k}_j \rangle_A = 0 \tag{1}$$

Тогда $\exists \alpha_1, \ldots, \alpha_m$:

$$\mathrm{x}_m = \mathrm{x}_0 + \sum_{i=1}^{m} \alpha_i \mathrm{k}_i, \quad \mathrm{r}_m = \mathrm{b} + A\left(\mathrm{x}_0 + \sum_{i=1}^{m} \alpha_i \mathrm{k}_i\right) \tag{2}$$

Из условия ортогональности $r_m \perp \mathcal{K}^m$ или $\langle \mathrm{r}_{m-1} - \alpha_m A \mathrm{k}_m, \mathrm{k}_i \rangle = 0$ находим α_i:

$$\alpha_i = \frac{\langle \mathrm{r}_{i-1}, \mathrm{k}_i \rangle}{\langle A\mathrm{k}_i, \mathrm{k}_i \rangle}, \quad i \in \{1, m\}$$

Метод сопряженных градиентов: CG

Дополнение пространства $\mathcal{K}^{n+1}$:

$$\mathcal{K}^{n+1} = \left\{\left\{r_0, Ar_0, \ldots, A^{m-1} r_0\right\}, A^m \mathrm{r}_0\right\} = \left\{\left\{\mathrm{k}_1, \ldots, \mathrm{k}_m\right\}, \mathrm{k}_{m+1}\right\}$$

Что означает $\mathrm{k}_{m+1} = A^m \mathrm{r}_0 + \sum_{i=1}^{m} \gamma_i \mathrm{k}_i = \mathrm{r}_m + \sum_{i=1}^{m} \beta_i \mathrm{k}_i$. За счет A-ортогональности $\beta_i = 0, i < m$. Получим:

$$\mathrm{k}_{m+1} = \mathrm{r}_m + \beta_m \mathrm{k}_m, \quad \beta_m = \frac{\langle \mathrm{r}_m, A\mathrm{k}_m \rangle}{\langle A\mathrm{k}_m, \mathrm{k}_m \rangle}.$$

Выражения для α_m, β_m можно упростить:

$$\alpha_m = \frac{\mathrm{r}_{m-1}, \mathrm{r}_{m-1}}{\langle A\mathrm{k}_m, \mathrm{k}_m \rangle}, \quad \beta_m = \frac{\langle \mathrm{r}_m, \mathrm{r}_m \rangle}{\langle \mathrm{r}_{m-1}, \mathrm{r}_{m-1} \rangle}$$

Все готово для алгоритма!

Алгоритм CG

Алгоритм Алгоритм CG

```
задаем начальное приближение x_0
r_0 = b − A x_0
k_1 = r_0
m = 1
цикл пока «не выполнен критерий остановки» выполним
    α_m = ⟨r_{m−1}, r_{m−1}⟩ / ⟨A k_m, k_m⟩
    x_m = x_{m−1} + α_m k_m
    r_m = r_{m−1} − α_m A k_m
    β_m = ⟨r_m, r_m⟩ / ⟨r_{m−1}, r_{m−1}⟩
    k_{m+1} = r_m + β_m k_m
    m = m + 1
завершим цикл пока
```

Спектральная теория сходимости

Для определения скорости сходимости сравним m итераций метода CG и метода Чебышева. Вспомним, что CG минимизирует энергитическую норму ошибки:

$$\mathrm{x}_{m+1} = \underset{\tilde{\mathrm{x}}-\mathrm{x}^m \in \mathcal{K}}{\arg\min} \left\langle A\left(\mathrm{x}-\tilde{\mathrm{x}}\right), \mathrm{x}-\tilde{\mathrm{x}}\right\rangle .$$

Итерации метода Чебышева:

$$\mathrm{x}_{m+1} = \mathrm{x}_m - \tau_{m+1}\left(A\mathrm{x}_m - \mathrm{b}\right),$$

Чебышевские параметры выбираются исходя из минимизации

$$P_m(A) = \prod_{i=1}^{m}\left(\mathbb{I} - \tau_i A\right)$$

связывающего ошибки $e_0 = \mathrm{x} - \mathrm{x}_0$ и $\mathrm{e}_m = \mathrm{x} - \mathrm{x}_m$:

$$\mathrm{e}_m = P_N(A)\mathrm{e}_0 \implies \|\mathrm{e}_m\| \leq \|P_m(A)\|\|\mathrm{e}_0\|_A$$

Многочлен Чебышева

Спектральная норма $\|P_m(A)\|$:

$$\|P_m(A)\| = \sup_{t \in Sp(A)} |P_m(t)| \leq \max_{t \in [\lambda_1, \lambda_N]} |P_m(t)|,$$

где λ_1, λ_N - наибольшее и наименьшее собственное значение A.
Многочлен Чебышева от параметра t:

$$P_m(t) = \prod_{i=1}^{m} (1 - \tau_i t)\,, \quad P_m(0) = 1.$$

Минимизирующий многочлен:

$$P_m(t) = T_m\left(\frac{\lambda_1 + \lambda_N}{\lambda_1 - \lambda_N}\right)^{-1} T_m\left(\frac{\lambda_1 + \lambda_N - 2t}{\lambda_1 - \lambda_N}\right),$$

Многочлен Чебышева первого рода:

$$T_m(x) = \frac{1}{2}\left(\left(x + \sqrt{x^2 - 1}\right)^m + \left(x + \sqrt{x^2 - 1}\right)^{-m}\right)$$

Многочлен Чебышева

Корни минимизирующего многочлена:

$$\tau_i = 2\left(\lambda_1 + \lambda_N + (\lambda_1 - \lambda_N)\cos\left(\frac{\pi(2i+1)}{2m}\right)\right)^{-1}, \quad i \in \{1, m\}$$

При $x \in [0,1] : T_m(x) \in [-1,1]$:

$$\max_{t\in[\lambda_1,\lambda_N]} |P_m(t)| = T_m\left(\frac{\lambda_1+\lambda_N}{\lambda_1-\lambda_N}\right)^{-1} \le \frac{2}{q^m + q^{-m}}$$

где

$$q = \frac{\sqrt{\lambda_1} - \sqrt{\lambda_N}}{\sqrt{\lambda_1} + \sqrt{\lambda_N}} = \frac{\sqrt{\mu(A)} - 1}{\sqrt{\mu(A)} + 1} \in [0,1]$$

$\implies \|e_m\| \le \frac{2\|e_0\|}{q^m+q^{-m}}$ - падение (A-)нормы ошибки за m итераций.

Оценки

После m итераций метода CG для ошибки $\tilde{e}_m = x - x_m$ имеем:

$$\|\tilde{e}_m\| = \langle A\tilde{e}_m, \tilde{e}_m\rangle^{\frac{1}{2}} = \min_{x_m - x_0 \in \mathcal{K}} \|e_m\|_A \le \frac{2\|e_0\|_A}{q^m + q^{-m}}.$$

В случай плохо обусловленных матриц $\mu(A) \gg 1$:

$$q \approx 1 - \frac{2}{\sqrt{\mu(A)}}$$

Оценки Капорина

На практике метод CG за итерацию убирает крайние собственные значения системы: λ_1 и λ_N. Если собственные значения сильно разбросаны, то оценка очень пессиместична.
Оценка Капорина:

$$\|\mathrm{r}_m\| \leq \left(K(A)^{\frac{1}{m}} - 1\right)^{\frac{m}{2}} \|\mathrm{r}_0\|, \quad K(A) = \frac{1}{\det(A)} \left(\frac{1}{N}\mathrm{trace}(A)\right)^N$$

Алгоритм PCG

Алгоритм Алгоритм CG с **предобуславливателем** B

задаем начальное приближение x_0

$\mathrm{r}_0 = \mathrm{b} - A\mathrm{x}_0$

$\mathrm{k}_1 = B^{-1} r_0$

$m = 1$

цикл пока «не выполнен критерий остановки» **выполним**

$$\alpha_m = \frac{\left\langle \mathrm{r}_{m-1}, B^{-1}\mathrm{r}_{m-1} \right\rangle}{\langle A\mathrm{k}_m, \mathrm{k}_m \rangle}$$

$\quad \mathrm{x}_m = \mathrm{x}_{m-1} + \alpha_m \mathrm{k}_m$

$\quad \mathrm{r}_m = \mathrm{r}_{m-1} - \alpha_m A\mathrm{k}_m$

$$\beta_m = \frac{\left\langle \mathrm{r}_m, B^{-1}\mathrm{r}_m \right\rangle}{\langle \mathrm{r}_{m-1}, B^{-1}\mathrm{r}_{m-1} \rangle}$$

$\quad \mathrm{k}_{m+1} = B^{-1}\mathrm{r}_m + \beta_m \mathrm{k}_m$

$\quad m = m + 1$

завершим цикл пока

Предобуславливатель Чебышева

Алгоритм Итерации Чебышева

оцениваем λ_1, λ_N по кругам Гершгорина
$d = (\lambda_1 + \lambda_N)/2, c = (\lambda_1 - \lambda_N)/2$
$\mathrm{x} = \mathrm{x}_0$
$m = 1$
цикл пока «не выполнен критерий остановки» **выполним**
  если $m <= 2$ **то**
    $\alpha = 2d/(2d^2 - (m-1)c^2)$
  иначе
    $\alpha = 1/(d - \alpha c^2/4)$
  завершим если
  $\mathrm{p} = \alpha(\mathrm{b} - A\mathrm{x}) + (\alpha d - 1)\mathrm{p}$
  $\mathrm{x} = \mathrm{x} + \mathrm{p}$
  $m = m + 1$
завершим цикл пока

Предобуславливатель Якоби и Гаусса-Зейделя

Пусть $A = L + D + U$, где D - диагональная, а L, U^T - нижние треугольные без диагонали.
Итерация методом Якоби:

$$x_{k+1} = D^{-1}\left(b - (L + U)x_k\right) \tag{3}$$

Итерация методом Гаусса-Зейделя:

$$x_{k+1} = (D + L)^{-1}\left(b - Ux_k\right) \tag{4}$$

Симметричные итерации методом Гаусса-Зейделя:

$$\begin{aligned} x_{k+1/2} &= (D + L)^{-1}\left(b - Ux_k\right) \\ x_{k+1} &= (D + U)^{-1}\left(b - Lx_{k+1/2}\right) \end{aligned} \tag{5}$$

Несимметричные системы

Если $A \neq A^T$, то существует ряд методов: BiCGSTAB, GMRES, TFQMR

Конец