Лекция 5-6:
“Алгебраический многосеточный метод”

Кирилл Терехов

ИВМ РАН

ВТМГБ МФТИ

14 апреля 2020 г.

Содержание

1. Лекция 5:
   Алгоритмы построения многосеточного метода
   - Введение
   - Оператор сглаживания
   - Оператор интерполяции
   - Алгоритм C/F-разбиения
   - Алгоритм выбора весов
   - Алгоритм построения
2. Лекция 6:
   Теория сходимости многосеточного метода
   - Сходимость метода
   - Свойства оператора сглаживания
   - Свойства оператора интерполяции
   - Литература

Лекция 5:
Алгоритмы построения многосеточного метода

Алгебраический многосеточный метод

Решим систему вида

$$Ax = b, \quad A \in \mathcal{R}^{N \times N}, \quad x, b \in \mathcal{R}^{N}. \tag{1}$$

При помощи многосеточного метода, решая последовательность

$$A^m x^m = b^m, \quad A^m \in \mathcal{R}^{N_m \times N_m}, \quad x^m, b^m \in \mathcal{R}^{N_m}, \tag{2}$$

где $1 \leq m \leq q$ и $N_m < N_{m+1}$.

Операторы

Для многосеточного метода необходимы операторы:

Определение (операторы)

$$I_{m+1}^{m} : \underset{\text{(грубая сетка)}}{\mathcal{R}^{N_{m+1}}} \to \underset{\text{(мелкая сетка)}}{\mathcal{R}^{N_m}}, \quad \text{интерполяция } (\textit{пролонгация})$$

$$I_{m}^{m+1} : \underset{\text{(мелкая сетка)}}{\mathcal{R}^{N_m}} \to \underset{\text{(грубая сетка)}}{\mathcal{R}^{N_{m+1}}}, \quad \text{сужение } (\textit{рестрикция})$$

$$S^{m} : \mathcal{R}^{N_m} \to \mathcal{R}^{N_m}. \quad \text{оператор сглаживания}$$

Операторы

- В стандартном многосеточном методе операторы определяются из дискретизации задачи на последовательности сеток.
- В алгебраическом методе операторы определяются по матрице.

Построение метода

Пусть

- матрица $A^m > 0$ симметричная положительно определенная,
- оператор интерполяции имеет полный ранг $\mathrm{rank}(I_{m+1}^{m}) = N^{m+1}$,

тогда

Пример (построение операторов)

$$I_m^{m+1} = \left(I_{m+1}^m\right)^T, \qquad \text{оператор сужения}$$

$$A^{m+1} = I_m^{m+1} A^m I_{m+1}^m, \qquad \text{система на грубой сетке}$$

$$S^m = I - \left(Q^m\right)^{-1} A^m. \qquad \text{оператор сглаживания}$$

Оператор сглаживания

Определение (сглаживание)

Оператор сглаживания в общем виде $S^m = I - (Q^m)^{-1} A^m$.

Пример (варианты методов сглаживания)

Матрица Q^m в методе:

- Гаусса-Зейделя, Q^m - это A^m без верхней треугольной части,
- Якоби, $Q^m = \mathrm{diag}(A^m)$,
- неполной факторизации, $Q^m = LU \approx A^m$.

Сетка в алгебраическом методе

Определение (сетка)

Ω^m - сетка на уровне m, элементы сетки - действительные числа $[1, N^m]$, связи определяются матрицей системы $A^m = \{a_{ij}^m\}$.

Определение (связи)

Множество связей элемента i, $\mathcal{N}_i^m = \left\{j : a_{ij}^m \neq 0, j \neq i\right\}$.

Определение (разбиение)

C/F-разбиение сетки $\Omega^m = C^m \cup F^m$.

- C^m - элементы из Ω^m на грубой сетке $C^m = \Omega^m \cap \Omega^{m+1}$.
- F^m - элементы только из Ω^m, но не из Ω^{m+1}, $F^m = \Omega^m \setminus \Omega^{m+1}$.

Оператор интерполяции

Определение (интерполяция)

Действие интерполяции $I_{m+1}^{m} : C^m \to F^m \cup C^m$ на вектор

$$e^m = I_{m+1}^{m} e^{m+1} = \begin{cases} e_i^{m+1}, & i \in C^m, \\ \sum\limits_{k \in \mathcal{I}_i^m} w_{ik}^m e_k^{m+1}, & i \in F^m, \end{cases} \tag{3}$$

где

- $\mathcal{I}_i^m \subseteq \mathcal{N}_i^m$ - множество интерполяционных связей,
- $\mathbf{w}_i^m = \{w_{ik}^m \neq 0 : k \in \mathcal{I}_i^m\}$ вектор весов.

Определение (пропущенные связи)

$\mathcal{D}_i^m = \mathcal{N}_i^m \setminus \mathcal{I}_i^m$ - множество пропущенных связей.

Основные определения

Определение

Элемент i связан сильно с j, если $|a_{ij}^m| \geq \theta \max_{l \neq i} |a_{il}^m|$, где $0 < \theta \leq 1$.

Определение

Множество $\mathcal{S}_i^m = \left\{ j : |a_{ij}^m| \geq \theta \max_{l \neq i} |a_{il}^m| \right\}$ сильных связей элемента i.

Определение

Множество $(\mathcal{S}_i^m)^T = \left\{ j : i \in \mathcal{S}_j^m \right\}$ смежных сильных связей элемента i.

Определение

Разбиение множества пропущенных связей $\mathcal{D}_i^m = \mathcal{D}_i^s \cap \mathcal{D}_i^w$:

- $\mathcal{D}_i^s = \mathcal{D}_i^m \cap \mathcal{S}_i^m$ - множество сильных пропущенных связей
- $\mathcal{D}_i^w = \mathcal{D}_i^m \setminus \mathcal{S}_i^m$ - множество слабых пропущенных связей.

C/F - разбиение

C/F-разбиение строится исходя из двух правил:

1. C^m должно быть максимальным подмножеством всех элементов Ω^m, с условием, что никакие два элемента из C^m не являются сильно связанными
$$\left(C^m = \underset{U \subset \Omega^m}{\arg\max} \left| \left\{ \forall i, j \in U, i \neq j : \mathcal{S}_i^m \cap \mathcal{S}_j^m = \emptyset \right\} \right| \right)$$
2. Для каждого i из F^m, каждый элемент j из $\mathcal{S}_i^m$ должен либо принадлежать C^m, либо должен быть сильно связан хотя бы с одним элементом из $\mathcal{I}_i^m$.
$$\left(\forall i \in F^m : \forall j \in \mathcal{S}_i^m : j \in \mathcal{I}_i^m \vee \ \mathcal{S}_j^m \cap \mathcal{I}_i^m \neq \emptyset \right)$$

C/F - разбиение

Алгоритм 1 Выбор C^m

пусть $C^m = F^m = \emptyset$ и $U = \Omega^m$
определим $\lambda_i = \left|\left(\mathcal{S}_i^m\right)^T\right|, \forall i \in \Omega^m$
цикл пока $U \neq \emptyset$ **выполним**
    выберем $i = \arg\max_{j \in U}(\lambda_j)$.
        $\triangleright$ на каждом шаге $\lambda_i = \left|\left(\mathcal{S}_i^m\right)^T \cap U\right| + 2\left|\left(\mathcal{S}_i^m\right)^T \cap F^m\right|$
    добавим $C^m = C^m \cup \{i\}$ и уберем $U = U \setminus \{i\}$
    цикл для всех $j \in \left(\mathcal{S}_i^m\right)^T \cap U$ **выполним**
        добавим $F^m = F^m \cup \{j\}$ и уберем $U = U \setminus \{j\}$
        увеличим $\lambda_l = \lambda_l + 1, \quad \forall l \in \mathcal{S}_j^m \cap U$
    завершим цикл для
    уменьшим $\lambda_j = \lambda_j - 1, \quad \forall j \in \mathcal{S}_i^m \cap U$
завершим цикл пока

C/F - разбиение

Алгоритм 2 Расширение C^m

цикл для всех $i \in F^m$ **выполним**
  пусть $\hat{C}^m = \emptyset$ и $\mathcal{I}_i^m = \mathcal{S}_i^m \cap C^m$ и $\mathcal{D}_i^s = \mathcal{S}_i^m \setminus \mathcal{I}_i^m$
  цикл для всех $j \in \mathcal{D}_i^s$ **выполним**
    если $\mathcal{S}_j^m \cap \left(\mathcal{I}_i^m \cup \hat{C}^m\right) = \emptyset$ **то**
      добавим $\hat{C}^m = \hat{C}^m \cup \{j\}$
    завершим если
  завершим цикл для
  если $|\hat{C}^m| > 1$ **то**
    добавим $C^m = C^m \cup \{i\}$ и уберем $F^m = F^m \setminus \{i\}$
  иначе
    добавим $C^m = C^m \cup \hat{C}^m$ и уберем $F^m = F^m \setminus \hat{C}^m$
  завершим если
завершим цикл для

Расчет весов

Алгоритм 3 Выбор весов

цикл для всех $i \in F^m$ **выполним**
    пусть $\mathcal{I}_i^m = \mathcal{S}_i^m \cap C^m$
    пусть $\mathcal{D}_i^s = \mathcal{S}_i^m \setminus \mathcal{I}_i^m$ и $\mathcal{D}_i^w = \mathcal{N}_i^m \setminus \mathcal{S}_i^m$
    определим $d_i = a_{ii}^m + \sum_{j \in \mathcal{D}_i^w} a_{ij}^m$
    определим $d_j = a_{ij}^m, \quad \forall j \in \mathcal{I}_i^m$
    цикл для всех $j \in \mathcal{D}_i^s$ **выполним**
        посчитаем $s_j = \sum_{k \in \mathcal{I}_i^m \cap \mathcal{S}_j^m} a_{jk}^m$
        добавим $d_k = d_k + a_{ij}^m a_{jk}^m / s_j, \quad \forall k \in \mathcal{I}_i^m \cap \mathcal{S}_j^m$
    завершим цикл для
    определим веса $w_{ik}^m = -d_k / d_i, \quad \forall k \in \mathcal{I}_i^m$
завершим цикл для

Алгоритм 4 Настройка AMG

1: $m = 1$
2: вычислим C/F-разбиение для Ω^m (**Алгоритм 1**)
3: уточним множество C^m (**Алгоритм 2**)
4: вычислим интерполятор I^m_{m+1} (**Алгоритм 3**)
5: зададим оператор сужения $(I^m_{m+1})^T = I^{m+1}_m$
6: вычислим систему на грубой сетке $A^{m+1} = I^{m+1}_m A^m I^m_{m+1}$
7: **если** система A^{m+1} мала **то**
8: $\quad q = m + 1$
9: $\quad$ вычислим $(A^{m+1})^{-1}$
10: **иначе**
11: $\quad m = m + 1$
12: $\quad$ **перейдем к шагу 2**
13: **завершим если**

Конец

Лекция 6:
Теория сходимости многосеточного метода

Основные определения

Определение (коррекция)

Оператор коррекции на грубой сетке $T^m = I - I^m_{m+1}\left(A^{m+1}\right)^{-1} I^{m+1}_m A^m$.

Определение (рекурсивная коррекция)

$\tilde{T}^m$ - оператор рекурсивной коррекции на грубой сетке с использованием V-цикла.

Отличие T^m от $\tilde{T}^m$ заключается в неточном обращении $(A^{m+1})^{-1}$ во втором случае.

Основные определения

Определение (ошибка)

Вектор $e^m \in \mathcal{R}^{N_m}$ называется ошибкой приближенного решения $\tilde{x}^m \in \mathcal{R}^{N_m}$ определяемый по $A^m e^m = b^m - A\tilde{x}^m$.

Определение (скалярное произведение и норма)

Скалярное произведение $\langle u, v \rangle_\alpha$ векторов $u, v \in \mathcal{R}^{N_m}$:

$\langle u, v \rangle_\alpha = \left\langle \left((D^m)^{-1} A^m \right)^{\alpha - 1} u, A^m v \right\rangle$, где $D^m = \operatorname{diag}(A^m)$.

Норма $\|u\|_\alpha = \sqrt{\langle u, u \rangle_\alpha}$.

Определение (скорость сходимости)

Скорость сходимости метода определяется коэффициентом уменьшением нормы ошибки после применения метода.

Сходимость многосеточного метода

Теорема (сходимость с пост-сглаживанием)

Пусть

- $A > 0$ *симметричная положительно определенная,*
- *интерполяция* I_{m+1}^{m} *имеет полный ранг,*
- *сужение определяется из* $I_{m}^{m+1} = \left(I_{m+1}^{m}\right)^{T}$,
- *матрица* $A^{m+1} = I_{m}^{m+1} A^{m} I_{m+1}^{m}$,
- *система с* A^{q} *на последнем уровне решается точно,*
- *предположим* $\exists\, \delta_1 > 0 : \forall e^{m} \in \mathcal{R}^{N_m}$:

$$\|S^{m} e^{m}\|_1^2 \le \|e^{m}\|_1^2 - \delta_1 \|T^{m} e^{m}\|_1^2 . \tag{4}$$

Тогда $\delta_1 \le 1$ *и если оператор сглаживания применяется* после *шага коррекции на грубой сетке* $\tilde{T}^{m}$*, то скорость сходимости V-цикла ограничена сверху* $\|S^{m} \tilde{T}^{m} e^{m}\|_1 \le \sqrt{1 - \delta_1}\|e^{m}\|_1$.

Сходимость многосеточного метода

Теорема (сходимость с пре-сглаживанием)

Если при тех же условиях вместо (4) $\exists\, \delta_2 > 0 : \forall e^m \in \mathcal{R}^{N_m}$*:*

$$\|S^m e^m\|_1^2 \le \|e^m\|_1^2 - \delta_2 \, \|T^m S^m e^m\|_1^2 \,, \tag{5}$$

и оператор сглаживания применяется перед шагом коррекции на грубой сетке $\tilde{T}^m$*, то скорость сходимости V-цикла ограничена сверху* $\|\tilde{T}^m S^m e^m\|_1 \le \|e^m\|_1 / \sqrt{1 + \delta_2}$.

Следствие (сходимость с пре- и пост-сглаживанием)

Если сглаживание применяется и перед и после шага коррекции, то скорость сходимости $\|S^m \tilde{T}^m S^m e^m\|_1 \le \sqrt{1 - \delta_1} / \sqrt{1 + \delta_2} \|e^m\|_1$.

Сходимость многосеточного метода

Условия (4) и (5) эквивалентны

$$\|S^m e^m\|_1^2 \le \|e^m\|_1^2 - \alpha_1 \|e^m\|_2^2 , \qquad \text{свойство пост-сглаживания}$$

$$\|S^m e^m\|_1^2 \le \|e^m\|_1^2 - \alpha_2 \|S^m e^m\|_2^2 , \qquad \text{свойство пре-сглаживания}$$

$$\|T^m e^m\|_1^2 \le \beta \|e^m\|_2^2 , \qquad \text{свойство коррекции}$$

тогда $\delta_1 = \alpha_1/\beta$ и $\delta_2 = \alpha_2/\beta$.

Свойства оператора сглаживания

Лемма (оценка оператора сглаживания)

Пусть оператор сглаживания S для матрицы $A > 0$ имеет общий вид $S = I - Q^{-1}A$, где Q - не вырождена.Тогда неравенства вида

$$\|Se\|_1^2 \leq \|e\|_1^2 - \alpha_1 \|e\|_2^2 , \quad \|Se\|_1^2 \leq \|e\|_1^2 - \alpha_2 \|Se\|_2^2 , \tag{6}$$

эквиваленты неравенствам вида

$$\alpha_1 Q^T D^{-1} Q \leq Q + Q^T - A, \quad \alpha_2 (A - Q^T) D^{-1} (A - Q) \leq Q + Q^T - A, \tag{7}$$

где $D = diag(A)$.

Свойства оператора сглаживания

Теорема (оценка для метода Гаусса-Зейделя)

Пусть

- $A > 0$ *- симметричная положительно определенная,*
- S *- метод релаксаций Гаусса-Зейделя,*
- *для некоторого* $\omega \in \mathcal{R}^N : \omega_i > 0$*:*

$$\gamma_- = \max_{1 \le i \le N} \left(\frac{1}{\omega_i a_{ii}} \sum_{j<i} \omega_j |a_{ij}| \right), \quad \gamma_+ = \max_{1 \le i \le N} \left(\frac{1}{\omega_i a_{ii}} \sum_{j>i} \omega_j |a_{ij}| \right), \tag{8}$$

Тогда

$$\alpha_1 \le \frac{1}{(1+\gamma_-)(1+\gamma_+)}, \quad \alpha_2 \le \frac{1}{\gamma_- \gamma_+}, \tag{9}$$

где α_1, α_2 *параметры* (6).

Оператор интерполяции

Образ интерполятора должен содержать ошибку оператора сглаживания, $im(S^m) \subseteq im(I^m_{m+1})$, тогда в случае равенства для ортогонального проектора T^m выполняется $im(T^m) \perp im(S^m)$.

Оператор интерполяции

Определение (гладкая ошибка)

Назовем ошибку $e^m \in \mathcal{R}^{N_m}$ гладкой, если $\|S^m e^m\| \approx \|e^m\|$.

Аппроксимацией компоненты e_i^m гладкой ошибки в зависимости от значений e_j^m из соседних элементов $j \in \mathcal{N}_i^m$ является выражение

$$r_i^m = a_{ii}^m e_i^m + \sum_{j \in \mathcal{N}_i^m} a_{ij}^m e_j^m \approx 0. \tag{10}$$

Тогда интерполятор можно задать выбрав

Пример (пример интерполятора)

- $\mathcal{I}_i^m = \mathcal{N}_i^m$,
- $w_{ik}^m = -a_{ik}^m / a_{ii}^m$.

Сходимость с пост-сглаживанием

Теорема (зависимость сходимости от свойств интерполятора)

Пусть

- $A^m > 0$ *симметричная положительно определенная,*
- *оператор сглаживания* S^m *удовлетворяет*
$$\|S^m e^m\|_1^2 \leq \|e^m\|_1^2 - \alpha_1 \|e^m\|_2^2, \tag{11}$$
- *интерполятор* I_{m+1}^m *имеет полный ранг,*
- $\exists\, \beta > 0 : \forall e^m$ *выполнено*
$$\min_{e^{m+1}} \|e^m - I_{m+1}^m e^{m+1}\|_0^2 \leq \beta \|e^m\|_1^2. \tag{12}$$

Тогда $\beta \geq \alpha_1$ *и выполняется неравенство*

$$\|S^m T^m e^m\|_1 \leq \sqrt{1 - \alpha_1/\beta}\, \|e^m\|_1. \tag{13}$$

Свойства интерполятора

Теорема (условия для весов интерполятора)

Пусть

- $A^m > 0$ *симметричная положительно определенная,*
- *множество* C^m*-элементов задано произвольно,*
- *веса* I^m_{m+1} *удовлетворяют* $w^m_{ik} \geq 0$ *и* $\sum_{k \in C^m} w^m_{ik} \leq 1$.

Тогда свойство (12) *выполнено, если* $\exists\, \beta > 0 : \forall e^m$ *такое что*

$$\begin{aligned} \sum_{i \in F^m} \sum_{k \in C^m} a^m_{ik} w^m_{ik} (e^m_i - e^m_k)^2 &\leq -\frac{\beta}{2} \sum_{i,j} a^m_{ij} (e^m_i - e^m_j)^2, \\ \sum_{i \in F^m} a^m_{ii} \left(1 - \sum_{k \in C^m} w^m_{ik} \right) (e^m_i)^2 &\leq \beta \sum_i \left(\sum_j a^m_{ij} \right) (e^m_i)^2 . \end{aligned} \tag{14}$$

Применение теоремы

Пусть $\mathcal{N}_i^m = \mathcal{I}_i^m \cup \mathcal{D}_i^m$. Для элементов $k \in \mathcal{I}_i^m$ зададим веса

$$w_{ik}^m = \eta_i |a_{ik}^m|, \quad \forall i \in F^m, k \in \mathcal{I}_i^m, \tag{15}$$

где η_i параметр удовлетворяющий

$$0 \le \eta_i \le \left(\sum_{l \in \mathcal{I}_i^m} |a_{il}^m| \right)^{-1}, \tag{16}$$

введенный для выполнения условия $\sum_{k \in C^m} w_{ik}^m \le 1$ теоремы. Чтобы выполнялось свойство (12) согласно теореме достаточно потребовать $\forall i \in F^m, k \in \mathcal{I}_i^m$:

$$0 \le a_{ii}^m w_{ik}^m \le \beta |a_{ik}^m|, \quad 0 \le a_{ii}^m \left(1 - \sum_{k \in \mathcal{I}_i^m} w_{ik}^m \right) \le \beta \sum_j a_{ij}^m. \tag{17}$$

Свойства системы на грубом уровне

Теорема

Пусть

- A^m *- симметричная слабо диагонально - доминантная M-матрица,*
- *веса интерполяции удовлетворяют* (17) *с* $\beta \leq 2$.

Тогда A^{m+1} *имеет те же свойства, что и* A^m.

Ссылки

John W. Ruge, Klaus Stüben (1987)
Algebraic multigrid
Society for Industrial and Applied Mathmetatics, 1987. pp. 73 – 130

Klaus Stüben (2001)
A review of algebraic multigrid
Numerical Analysis: Historical Developments in the 20th Century, Elsevier, 2001. pp. 331–359

Конец