

Системы линейных алгебраических уравнений
Часть 2. Прямые и итерационные методы решения

Скалько Юрий Иванович
Цыбулин Иван

Прямые и итерационные методы

Численные методы решения СЛАУ делятся на 2 типа:

- Прямые — решение системы получается за конечное число действий
- Итерационные — решение строится в виде последовательности приближений к решению. При достижении заданной точности итерации прекращают

Метод Гаусса

В основе метода лежит последовательное исключение неизвестных. Метод Гаусса состоит из *прямого хода*, на котором матрица системы элементарными преобразованиями строк преобразуется к верхнетреугольному виду, и *обратного хода*, когда полученная треугольная система решается подстановкой. Метод Гаусса *корректен* (в ходе решения не придется делить на ноль), если все главные миноры матрицы отличны от нуля.

Метод Гаусса

В основе метода лежит последовательное исключение неизвестных. Метод Гаусса состоит из *прямого хода*, на котором матрица системы элементарными преобразованиями строк преобразуется к верхнетреугольному виду, и *обратного хода*, когда полученная треугольная система решается подстановкой. Метод Гаусса *корректен* (в ходе решения не придется делить на ноль), если все главные миноры матрицы отличны от нуля. На практике, иногда требуется переставлять строки или столбцы матрицы, чтобы избежать деления на ноль. Также, переставляя строки и столбцы можно улучшить устойчивость метода, то есть чувствительность к ошибкам округления

Метод Гаусса

Рассмотрим систему

$$\begin{pmatrix} 10^{-3} & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \approx \begin{pmatrix} 2.997003 \\ 1.997003 \end{pmatrix}$$

Будем решать ее методом Гаусса, используя в вычислениях только три значащие цифры

$$\begin{aligned} \left(\begin{array}{cc|c} 10^{-3} & 1 & 2 \\ 1 & -1 & 1 \end{array} \right) &\sim \left(\begin{array}{cc|c} 1 & 10^3 & 2 \cdot 10^3 \\ 1 & -1 & 1 \end{array} \right) \stackrel{!}{\sim} \left(\begin{array}{cc|c} 1 & 10^3 & 2 \cdot 10^3 \\ 0 & -10^3 & -2 \cdot 10^3 \end{array} \right) \sim \\ &\left(\begin{array}{cc|c} 1 & 10^3 & 2 \cdot 10^3 \\ 0 & -10^3 & -2 \cdot 10^3 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 10^3 & 2 \cdot 10^3 \\ 0 & 1 & 2 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 2 \end{array} \right) \end{aligned}$$

Полученное решение значительно отличается от точного.

Метод Гаусса с выбором главного элемента

Данную проблему можно устранить, если в качестве ведущего элемента (того на который делится очередная строка) выбирать наибольший по модулю элемент в оставшейся подматрице. Такой метод называется методом Гаусса с выбором главного элемента. Применим его к той же системе

$$\left(\begin{array}{cc|c} 10^{-3} & 1 & 2 \\ \boxed{1} & -1 & 1 \end{array} \right) \sim \left(\begin{array}{cc|c} 0 & \boxed{1} & 2 \\ 1 & -1 & 1 \end{array} \right) \sim \left(\begin{array}{cc|c} 0 & 1 & 2 \\ 1 & 0 & 3 \end{array} \right) \sim \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \end{array} \right)$$

Теперь отличие от точного решения не превосходит $3 \cdot 10^{-3}$, что оказалось даже меньше погрешности вычислений.

Обусловленность

Исследуем, насколько плохо обусловлена матрица этой системы

$$\mathbf{A} = \begin{pmatrix} 10^{-3} & 1 \\ 1 & -1 \end{pmatrix} \quad \mathbf{A}^{-1} = \frac{1}{1001} \begin{pmatrix} 1000 & 1000 \\ 1000 & -1 \end{pmatrix}$$

Обусловленность

Исследуем, насколько плохо обусловлена матрица этой системы

$$\mathbf{A} = \begin{pmatrix} 10^{-3} & 1 \\ 1 & -1 \end{pmatrix} \quad \mathbf{A}^{-1} = \frac{1}{1001} \begin{pmatrix} 1000 & 1000 \\ 1000 & -1 \end{pmatrix}$$

$$\mu_{\infty}(\mathbf{A}) = \|\mathbf{A}\|_{\infty} \|\mathbf{A}^{-1}\|_{\infty} = 2 \cdot \frac{2000}{1001} \approx 4$$

Число обусловленности матрицы невелико, и хорошим численным методом можно решить данную систему с относительной ошибкой, не превосходящей $4 \frac{\delta \mathbf{b}}{\mathbf{b}}$.

Следовательно, потеря точности при решении методом Гаусса без выбора главного элемента связана с самим методом, а не с плохой постановкой задачи. К тому же, метод Гаусса с выбором главного элемента нашел решение, погрешность которого вполне соответствует неустраняемой ошибке

Вычислительная устойчивость

Проблема метода Гаусса заключается в накоплении ошибок округления. Эта проблема возникает всегда при операциях с числами разных знаков или разных порядков. Метод с выбором главного элемента позволяет уменьшить ошибку, но не всегда этого оказывается достаточно.

Теорема об устойчивости метода Гаусса

Если матрица A имеет диагональное преобладание

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|, \quad i = \overline{1, n},$$

то при решении методом Гаусса погрешность не превосходит неустранимой погрешности, то есть метод Гаусса *вычислительно устойчив*

Вычисление прогоночных коэффициентов

Решение ищется в виде прогоночного соотношения

$$x_i = P_i x_{i+1} + Q_i$$

При подстановке его в уравнение $a_k x_{k-1} + b_k x_k + c_k x_{k+1} = f_k$, получаем

$$\begin{aligned} a_k(P_{k-1}x_k + Q_{k-1}) + b_kx_k + c_kx_{k+1} &= f_k \\ (a_kP_{k-1} + b_k)x_k &= -c_kx_{k+1} + f_k - a_kQ_{k-1} \end{aligned}$$

Записывая это уравнение в виде прогоночного соотношения

$$x_k = \frac{-c_k}{a_kP_{k-1} + b_k}x_{k+1} + \frac{f_k - a_kQ_{k-1}}{a_kP_{k-1} + b_k},$$

получаем рекуррентное соотношение для P_k, Q_k

$$P_k = \frac{-c_k}{a_kP_{k-1} + b_k}, \quad Q_k = \frac{f_k - a_kQ_{k-1}}{a_kP_{k-1} + b_k}$$

Вычисление прогоночных коэффициентов

$$P_k = \frac{-c_k}{a_k P_{k-1} + b_k}, \quad Q_k = \frac{f_k - a_k Q_{k-1}}{a_k P_{k-1} + b_k}$$

P_1 и Q_1 можно найти из первого уравнения $b_1 x_1 + c_1 x_2 = f_1$

$$P_1 = \frac{-c_1}{b_1}, \quad Q_1 = \frac{f_1}{b_1}$$

Заметим, что $P_n = 0$ и $x_n = Q_n$. Остальные неизвестные находятся с помощью прогоночного соотношения

$$x_k = P_k x_{k+1} + Q_k, \quad k = n-1, n-2, \dots, 1$$

Свойства метода прогонки

Как и метод Гаусса, прогонка устойчива (в процессе прогонки не будут встречаться деления на числа близкие к 0) при наличии у системы диагонального преобладания, то есть

$$|b_k| > |a_k| + |c_k|, k = \overline{1, n}.$$

Арифметическая сложность прогонки составляет $O(n)$ действий, в то время как у метода Гаусса — $O(n^3)$. Это одна из причин выделения прогонки как отдельного метода.

LU разложение

В результате применения метода Гаусса для матрица A вычисляется ее LU разложение, то есть представление в виде произведения нижнетреугольной матрицы L на верхнетреугольную матрицу U . При этом система

$$Ax = b$$

решается в два этапа

$$LUx = b \Rightarrow \begin{cases} Ly = b \\ Ux = y \end{cases}$$

На каждом этапе прямой или обратной подстановкой решается система с треугольной матрицей.

Построение LU разложения

Рассмотрим, как строится LU разложение заданной матрицы.
Пусть матрица \mathbf{A} имеет вид

$$\mathbf{A} = \begin{pmatrix} a_{11} & \mathbf{w}^T \\ \mathbf{v} & \mathbf{A}' \end{pmatrix}$$

Тогда для матрицы \mathbf{A} справедливо представление

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ \mathbf{v}/a_{11} & \mathbf{E} \end{pmatrix} \begin{pmatrix} a_{11} & \mathbf{w}^T \\ 0 & \mathbf{A}' - \mathbf{v}\mathbf{w}^T/a_{11} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mathbf{v}/a_{11} & \mathbf{L}' \end{pmatrix} \begin{pmatrix} a_{11} & \mathbf{w}^T \\ 0 & \mathbf{U}' \end{pmatrix}$$

Здесь $\mathbf{L}'\mathbf{U}'$ — LU-разложение матрицы $\mathbf{A}' - \mathbf{v}\mathbf{w}^T/a_{11}$

Пример

Рассмотрим построение **LU** разложения на примере

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 4 \\ 3 & 7 & 10 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 4-2 & 4-2 \\ 0 & 7-3 & 10-3 \end{pmatrix} \equiv$$

Построим разложение для подматрицы 2×2

$$\begin{pmatrix} 4-2 & 4-2 \\ 7-3 & 10-3 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 4 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 4/2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 0 & 7-4 \end{pmatrix}$$

Теперь можно записать разложение для **A**

$$\equiv \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

Абстрактный процесс простой итерации

Рассмотрим абстрактный итерационный процесс

$$\mathbf{x}_{k+1} = \mathbf{B}\mathbf{x}_k + \mathbf{f},$$

где \mathbf{B} — некоторая матрица, \mathbf{f} — некоторый вектор.

Пусть данный процесс сходится к \mathbf{x}^* . Тогда

$$\mathbf{x}^* = \mathbf{B}\mathbf{x}^* + \mathbf{f}$$

Рассмотрим при каких ограничениях этот процесс сходится к \mathbf{x}^*

Достаточное условие сходимости

$$\mathbf{x}_{k+1} = \mathbf{B}\mathbf{x}_k + \mathbf{f},$$

Рассмотрим вектор ошибки $\mathbf{r}_k \equiv \mathbf{x}_k - \mathbf{x}^*$.

$$\mathbf{r}_{k+1} = \mathbf{x}_{k+1} - \mathbf{x}^* = \mathbf{B}\mathbf{x}_k + \mathbf{f} - \mathbf{B}\mathbf{x}^* - \mathbf{f} = \mathbf{B}\mathbf{r}_k,$$

$$\mathbf{r}_{k+1} = \mathbf{B}\mathbf{r}_k$$

Таким образом, вектор ошибки за одну итерацию умножается на \mathbf{B} .
Если какая-то норма $\|\mathbf{B}\|_{\bullet} = q < 1$, то

$$\|\mathbf{r}_{k+1}\|_{\bullet} = \|\mathbf{B}\mathbf{r}_k\|_{\bullet} \leq \|\mathbf{B}\|_{\bullet} \|\mathbf{r}_k\|_{\bullet} = q \|\mathbf{r}_k\|_{\bullet}.$$

Таким образом, норма невязки будет стремиться к нулю со скоростью геометрической прогрессии

$$\|\mathbf{r}_k\|_{\bullet} \leq Cq^k$$

Это — *достаточное* условие сходимости процесса.

Из-за эквивалентности норм в \mathbb{R}^n , для других норм будет справедливо то же самое, но с другой константой C

Критерий сходимости

$$\mathbf{x}_{k+1} = \mathbf{B}\mathbf{x}_k + \mathbf{f},$$

Для того чтобы процесс сходился при любом начальном приближении \mathbf{x}_0 *необходимо и достаточно*, чтобы все собственные значения \mathbf{B} (возможно комплексные) лежали внутри единичного круга

$$|\lambda_i(\mathbf{B})| < 1$$

Если это условие не выполнено, то итерационный процесс может не сойтись при некотором приближении. На практике это будет означать, что почти при всех начальных приближениях он будет расходиться.

Метод простой итерации с параметром τ

Рассмотрим систему $\mathbf{Ax} = \mathbf{b}$. Построим такой итерационный процесс $\mathbf{x}_{k+1} = \mathbf{Bx}_k + \mathbf{f}$, чтобы его предел совпадал с решением системы.

Метод простой итерации с параметром τ

Рассмотрим систему $\mathbf{Ax} = \mathbf{b}$. Построим такой итерационный процесс $\mathbf{x}_{k+1} = \mathbf{Bx}_k + \mathbf{f}$, чтобы его предел совпадал с решением системы.

Рассмотрим процесс с некоторым параметром $\tau \neq 0$

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \tau(\mathbf{Ax}_k - \mathbf{b}) = (\mathbf{E} - \tau\mathbf{A})\mathbf{x}_k + \tau\mathbf{b}$$

Если данный процесс сходится, то он сходится к решению системы

$$\mathbf{x}^* = \mathbf{x}^* - \tau(\mathbf{Ax}^* - \mathbf{b}) \Leftrightarrow \mathbf{Ax}^* = \mathbf{b}$$

Сходимость метода простой итерации с параметром τ

Метод простой итерации соответствует процессу с

$$\mathbf{B} = \mathbf{E} - \tau\mathbf{A}, \mathbf{f} = \tau\mathbf{b}$$

Сходимость метода простой итерации с параметром τ

Метод простой итерации соответствует процессу с

$$\mathbf{B} = \mathbf{E} - \tau\mathbf{A}, \mathbf{f} = \tau\mathbf{b}$$

Собственные значения матрицы \mathbf{B} связаны с собственными значениями матрицы \mathbf{A} соотношением

$$\lambda(\mathbf{B}) = 1 - \tau\lambda(\mathbf{A})$$

Необходимое и достаточное условие сходимости для любого приближения даются условием $|\lambda(\mathbf{B})| < 1$

$$|1 - \tau\lambda(\mathbf{A})| < 1 \Leftrightarrow \left| \lambda(\mathbf{A}) - \frac{1}{\tau} \right| < \frac{1}{|\tau|}$$

Последнее условие означает, что все собственные числа матрицы \mathbf{A} должны лежать в круге с центром в точке τ^{-1} и радиуса $|\tau|^{-1}$

Оптимальный параметр для метода простой итерации

Допустим, что $\mathbf{A} = \mathbf{A}^T > 0$. Пусть собственные числа \mathbf{A} лежат на отрезке $[\lambda_{\min}, \lambda_{\max}]$

Оптимальный параметр для метода простой итерации

Допустим, что $\mathbf{A} = \mathbf{A}^T > 0$. Пусть собственные числа \mathbf{A} лежат на отрезке $[\lambda_{\min}, \lambda_{\max}]$

Итерационный процесс сходится со скоростью геометрической прогрессии

$$\|\mathbf{r}_k\|_E < Cq^k, \quad q = \max |\lambda(\mathbf{B})| = \|\mathbf{B}\|_E$$

Оптимальный параметр для метода простой итерации

Допустим, что $\mathbf{A} = \mathbf{A}^T > 0$. Пусть собственные числа \mathbf{A} лежат на отрезке $[\lambda_{\min}, \lambda_{\max}]$

Итерационный процесс сходится со скоростью геометрической прогрессии

$$\|\mathbf{r}_k\|_E < Cq^k, \quad q = \max |\lambda(\mathbf{B})| = \|\mathbf{B}\|_E$$

Чтобы метод сходился необходимо, чтобы $0 < \tau < \frac{2}{\lambda_{\max}}$. Найдем скорость сходимости при фиксированном τ .

Оптимальный параметр для метода простой итерации

Допустим, что $\mathbf{A} = \mathbf{A}^T > 0$. Пусть собственные числа \mathbf{A} лежат на отрезке $[\lambda_{\min}, \lambda_{\max}]$

Итерационный процесс сходится со скоростью геометрической прогрессии

$$\|\mathbf{r}_k\|_E < Cq^k, \quad q = \max |\lambda(\mathbf{B})| = \|\mathbf{B}\|_E$$

Чтобы метод сходился необходимо, чтобы $0 < \tau < \frac{2}{\lambda_{\max}}$. Найдем скорость сходимости при фиксированном τ .

Если собственные числа матрицы \mathbf{A} лежат на отрезке $[\lambda_{\min}, \lambda_{\max}]$, то собственные числа \mathbf{B} — на отрезке $[1 - \tau\lambda_{\min}, 1 - \tau\lambda_{\max}]$.

$$q = \max(|1 - \tau\lambda_{\max}|, |1 - \tau\lambda_{\min}|)$$

Минимально возможное значение q будет в случае

$$|1 - \tau\lambda_{\min}| = |1 - \tau\lambda_{\max}|$$

$$1 - \tau\lambda_{\min} = \tau\lambda_{\max} - 1, \quad \tau_{\text{опт}} = \frac{2}{\lambda_{\min} + \lambda_{\max}}, \quad q_{\text{опт}} = \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}}$$

Метод Якоби

Пусть $\mathbf{D} = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$, то есть диагональ матрицы \mathbf{A} . Метод Якоби можно трактовать как метод простой итерации для системы

$$\mathbf{D}\mathbf{x}_{k+1} + (\mathbf{A} - \mathbf{D})\mathbf{x}_k = \mathbf{b}$$

В данном случае система легко решается относительно диагональных неизвестных.

$$\begin{array}{rcccccc} a_{11}x_1^{(k+1)} & + & a_{12}x_2^{(k)} & + & \dots & + & a_{1n}x_n^{(k)} & = & b_1 \\ a_{21}x_1^{(k)} & + & a_{22}x_2^{(k+1)} & + & \dots & + & a_{2n}x_n^{(k)} & = & b_2 \\ a_{31}x_1^{(k)} & + & a_{32}x_2^{(k)} & + & \dots & + & a_{3n}x_n^{(k)} & = & b_3 \\ \vdots & & & & & & & & \\ a_{n1}x_1^{(k)} & + & a_{n2}x_2^{(k)} & + & \dots & + & a_{nn}x_n^{(k+1)} & = & b_n \end{array}$$

Критерий сходимости метода Якоби

Стандартный критерий $|\lambda(\mathbf{B})| < 1$ можно в случае метода Якоби записать в следующем виде: корни уравнения

$$\det(\lambda\mathbf{D} + \mathbf{A} - \mathbf{D}) = \begin{vmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & \lambda a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & \lambda a_{nn} \end{vmatrix} = 0$$

должны попасть в единичный круг. Фактически, это уравнение эквивалентно характеристическому уравнению для матрицы \mathbf{B}

$$\det(\mathbf{B} - \lambda\mathbf{E}) = 0.$$

Сходимость метода Якоби

Пусть матрица \mathbf{A} имеет диагональное преобладание

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}| + \delta |a_{ii}|, \quad i = \overline{1, n}$$

Матрица итерационного процесса $\mathbf{B} = \mathbf{E} - \mathbf{D}^{-1}\mathbf{A}$. Найдем ее $\|\cdot\|_\infty$ норму

$$\|\mathbf{B}\|_\infty = \max_i \sum_j |b_{ij}| = \max_i \sum_{j \neq i} \left| -\frac{a_{ij}}{a_{ii}} \right| = \max_i \frac{\sum_{j \neq i} |a_{ij}|}{|a_{ii}|} \leq 1 - \delta < 1$$

Достаточное условие сходимости процесса выполнено

Метод Зейделя

Разобьем матрицу \mathbf{A} на сумму двух матриц $\mathbf{A} = \mathbf{L} + \mathbf{U}$. (Это не матрицы из LU разложения, а просто части матрицы \mathbf{A}). Матрица \mathbf{L} будет содержать все элементы на главной диагонали и под ней, а \mathbf{U} — все элементы выше главной диагонали

$$\mathbf{L} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & & \ddots & 0 \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad \mathbf{U} = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ 0 & 0 & \dots & a_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Итерационный процесс записывается в виде

$$\mathbf{L}\mathbf{x}_{k+1} + \mathbf{U}\mathbf{x}_k = \mathbf{b}$$

Критерий сходимости метода Зейделя

Стандартный критерий $|\lambda(\mathbf{B})| < 1$ можно в случае метода Зейделя записать в следующем виде: корни уравнения

$$\det(\lambda \mathbf{L} + \mathbf{U}) = \begin{vmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn} \end{vmatrix} = 0$$

должны попасть в единичный круг. Фактически, это уравнение эквивалентно характеристическому уравнению для матрицы \mathbf{B}

$$\det(\mathbf{B} - \lambda \mathbf{E}) = 0.$$

Метод Зейделя

$$\mathbf{L}\mathbf{x}_{k+1} + \mathbf{U}\mathbf{x}_k = \mathbf{b}$$

Для вычисления следующего приближения \mathbf{x}_{k+1} необходимо на каждом шаге решать систему с треугольной матрицей \mathbf{L} .

В данном случае треугольная система решается очень просто

$$\begin{aligned} a_{11}x_1^{(k+1)} + a_{12}x_2^{(k)} + \dots + a_{1n}x_n^{(k)} &= b_1 \\ a_{21}x_1^{(k+1)} + a_{22}x_2^{(k+1)} + \dots + a_{2n}x_n^{(k)} &= b_2 \\ a_{31}x_1^{(k+1)} + a_{32}x_2^{(k+1)} + \dots + a_{3n}x_n^{(k)} &= b_3 \\ \vdots & \\ a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} + \dots + a_{nn}x_n^{(k+1)} &= b_n \end{aligned}$$

Если решать уравнения сверху вниз, то в каждом уравнении неизвестно ровно одно значение $x_j^{(k+1)}$.

Спасибо за внимание!

Цыбулин Иван
e-mail: tsybulin@crec.mipt.ru