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

Скалько Юрий Иванович
Цыбулин Иван

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

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. На практике, иногда требуется переставлять строки или столбцы матрицы, чтобы избежать деления на ноль. Также, переставляя стоки и столбцы можно улучшить устойчивость метода, то есть чувствительность к ошибкам округления

Метод Гаусса

Рассмотрим систему

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

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

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

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

Полученное решение значительно отличается от точного.

Метод Гаусса с выделением 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 выделением главного элемента.

Применим его к той же системе

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

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

Обусловенность

Исследуем, насколько плохо обусловлена матрица этой системы

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

Обусловенность

Исследуем, насколько плохо обусловлена матрица этой системы

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

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

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 неустранимой ошибке

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 *устойчив*

Существует вариант метода Гаусса для важных в приложениях *трехдиагональных* систем линейных уравнений

$$
\begin{cases}
b_1x_1 + c_1x_2 & = & f_1 \\
a_2x_1 + b_2x_2 + c_2x_3 & = & f_2 \\
\ddots & & \\
a_{n-1}x_{n-2} + b_{n-1}x_{n-1} + c_{n-1}x_n & = & f_{n-1} \\
a_nx_{n-1} + b_nx_n & = & f_n
\end{cases}
$$

Вычисление прогоночных коэффициентов

Решение ищется в виде прогоночного соотношения

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

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

$$a_k(P_k x_k + Q_k) + b_k x_k + c_k x_{k+1} = f_k$$
$$(a_k P_k + b_k)x_k = -c_k x_{k+1} + f_k - a_k Q_k$$

Записывая это уравнение в виде прогоночного соотношения

$$x_k = \frac{-c_k}{a_k P_k + b_k} x_{k+1} + \frac{f_k - a_k Q_k}{a_k P_k + b_k},$$

получаем рекуррентное соотношение для P_k, Q_k

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

Вычисление прогоночных коэффициентов

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

P_2 и Q_2 можно найти из первого уравнения $b_1 x_1 + c_1 x_2 = f_1$

$$P_2 = \frac{-c_1}{b_1}, \quad Q_2 = \frac{f_1}{b_1}$$

Заметим, что $P_{n+1} = 0$ и $x_n = Q_{n+1}$. Остальные неизвестные находятся с помощью прогоночного соотношения

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

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 факторизация

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

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

решается в два этапа

$$\mathbf{L}\mathbf{U}\mathbf{x} = \mathbf{b} \Rightarrow \begin{cases} \mathbf{L}\mathbf{y} & = & \mathbf{b} \\ \mathbf{U}\mathbf{x} & = & \mathbf{y} \end{cases}$$

На каждом этапе решается система с треугольной матрицей.

LU факторизация

Рассмотрим, как строится **LU** разложение заданной матрицы. Пусть матрица **A** имеет вид

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

Тогда для матрицы **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 $\mathbf{A}' - \mathbf{v}\mathbf{w}^T/a_{11}$

Рассмотрим как стоится **LU** факторизация на примере

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

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**

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

Симметризация

Предположим, имеется эффективный метод решения системы с симметричной положительно определенной матрицей. Что делать если система $\mathbf{Ax} = \mathbf{b}$ не такая?

Симметризация

Предположим, имеется эффективный метод решения системы с симметричной положительно определенной матрицей. Что делать если система $\mathbf{A}\mathbf{x} = \mathbf{b}$ не такая?
Умножим эту систему слева на $\mathbf{A}^T$

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

Полученная система имеет уже положительную матрицу $\mathbf{A}^T\mathbf{A}$.

Симметризация

Предположим, имеется эффективный метод решения системы с симметричной положительно определенной матрицей. Что делать если система $\mathbf{A}\mathbf{x} = \mathbf{b}$ не такая?
Умножим эту систему слева на $\mathbf{A}^T$

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

Полученная система имеет уже положительную матрицу $\mathbf{A}^T\mathbf{A}$. Но число обусловленности такой матрицы

$$\mu(\mathbf{A}^T\mathbf{A}) \lesssim \mu(\mathbf{A})^2$$

То есть такая операция сильно ухудшает обусловленность матрицы.

Предобуславливание

Пусть имеется система $\mathbf{Ax} = \mathbf{b}$ с плохо обусловленной матрицей $\mathbf{A}$. Плохо обусловленная матрица может стать препятствием при решении системы численным метдом. Поэтому имеет смысл попытаться снизить обусловленность системы.

Предобуславливание

Пусть имеется система $\mathbf{Ax} = \mathbf{b}$ с плохо обусловленной матрицей $\mathbf{A}$. Плохо обусловленная матрица может стать препятствием при решении системы численным метдом. Поэтому имеет смысл попытаться снизить обусловленность системы. Предположим, что имеется матрица $\mathbf{C} \approx \mathbf{A}^{-1}$. Домножим систему на $\mathbf{C}$

$$\mathbf{CAx} = \mathbf{Cb}$$

Поскольку $\mathbf{CA} \approx \mathbf{E}$, число обусловленности $\mu(\mathbf{CA}) \approx \mu(\mathbf{E}) = 1$. Полученную систему можно решать численным методом. Такая матрица $\mathbf{C}$ называется предобуславливателем, а сама операция домножения на $\mathbf{C}$ — предобуславливанием системы.

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 $\mathbf{x}^*$

Абстрактный процесс простой итерации

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

Рассмотрим невязку $\mathbf{r}_k = \mathbf{x}_k - \mathbf{x}^*$.

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

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

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

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

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

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

Абстрактный процесс простой итерации

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

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

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

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 τ

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

Ели данный процесс сходится, то он сходится к решению системы

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 **B** связаны с собственными значениями матрицы **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 **A** должны лежать в круге с центром в точке τ^{-1} и радиуса τ^{-1}

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

$$\mathbf{D}^{-1}\mathbf{A}\mathbf{x} = \mathbf{D}^{-1}\mathbf{b}$$

Матрица $\mathbf{D}^{-1}$ выступает в роли предобуславливателя системы. Запишем метод простой итерации с $\tau = 1$ для этой системы

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

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 $\mathbf{A} = \mathbf{L} + \mathbf{U}$. 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

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

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

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

$$\begin{array}{ccccccccc}
a_{11}\boxed{x_{k+1,1}} & + & a_{12}x_{k,2} & + & \dots & + & a_{1n}x_{k,n} & = & b_1 \\
a_{21}x_{k+1,1} & + & a_{22}\boxed{x_{k+1,2}} & + & \dots & + & a_{2n}x_{k,n} & = & b_2 \\
a_{31}x_{k+1,1} & + & a_{32}x_{k+1,2} & + & \dots & + & a_{1n}x_{k,n} & = & b_3 \\
\vdots & & & & & & & & \\
a_{n1}x_{k+1,1} & + & a_{n2}x_{k+1,2} & + & \dots & + & a_{nn}\boxed{x_{k+1,n}} & = & b_n
\end{array}$$

Если решать уравнения сверху вниз, то в каждом уравнении неизвестно ровно одно значение $x_{k+1,j}$.

Сходимость метода Зейделя

Запишем итерационный процесс метода Зейделя в форме $\mathbf{x}_{k+1} = \mathbf{B}\mathbf{x}_k + \mathbf{f}$

$$\mathbf{x}_{k+1} = \mathbf{L}^{-1}(-\mathbf{U}\mathbf{x}_k + \mathbf{b}) = -\mathbf{L}^{-1}\mathbf{U}\mathbf{x}_k + \mathbf{L}^{-1}\mathbf{b}$$

Пусть матрица $\mathbf{A} = \mathbf{A}^T > 0$. Тогда $\mathbf{L} = \mathbf{U}^T + \mathbf{D}$ Пусть $\lambda, \mathbf{x}$ - собственное число и соответствующий вектор матрицы $\mathbf{B} = -\mathbf{L}^{-1}\mathbf{U}$

$$\mathbf{B}\mathbf{x} = \lambda\mathbf{x}$$

$$-\mathbf{L}^{-1}\mathbf{U}\mathbf{x} = \lambda\mathbf{x}, \quad -\mathbf{U}\mathbf{x} = \lambda\mathbf{L}\mathbf{x}$$

Умножим это равенство слева на $\mathbf{x}^T$

$$-\mathbf{x}^T\mathbf{U}\mathbf{x} = \lambda(\mathbf{x}^T\mathbf{D}\mathbf{x} + \mathbf{x}^T\mathbf{U}^T\mathbf{x}) = \lambda(\mathbf{x}^T\mathbf{D}\mathbf{x} + \mathbf{x}^T\mathbf{U}\mathbf{x}) \Rightarrow |\lambda| < 1$$

Все собственные числа $\mathbf{B}$ лежат в единичном круге.

Спасибо за внимание!

tsybulin@crec.mipt.ru