

Лекция № 1

Системы линейных алгебраических уравнений

Лектор: Н.Ю. Золотых

Записал: В. Замараев

?? сентября 2008

Содержание

1. Постановка задачи	1
2. Метод Гаусса	1
3. Целочисленная модификация метода Гаусса	2
4. Метод Бариеса	3
5. Заключение	6

1. Постановка задачи

Дана система уравнений:

$$Ax = b \quad (1)$$

где $A \in \mathbb{Z}^{n \times n}$, $b \in \mathbb{Z}^n$, $\det A \neq 0$. Необходимо найти решение данной системы.

Замечание 1. Условия $A \in \mathbb{Z}^{n \times n}$ и $\det A \neq 0$ приняты для простоты изложения. Рассматриваемый метод без труда распространяется на случай прямоугольной матрицы с коэффициентами из произвольной области целостности.

Очевидно, что если x является решением системы (1), то $x \in Q^n$. Пожалуй, самый простой способ решить систему (1) — это метод Гаусса.

2. Метод Гаусса

Обычный метод Гаусса состоит из двух этапов:

1. преобразование расширенной матрицы;
2. обратный ход.

На первом этапе производятся строчечные преобразования над исходной расширенной матрицей:

$$A^{(0)} = [a_{ij}^{(0)}] = [A|b], 1 \leq i \leq n, 1 \leq j \leq n+1,$$

где $a_{i,n+1}^{(0)} = b_i$. В результате последовательности преобразований

$$A^{(0)} \rightarrow A^{(1)} \rightarrow \dots \rightarrow A^{(n-1)}$$

получается матрица $A^{(n-1)}$, первые n столбцов которой образуют верхнетреугольную матрицу. На втором этапе из матрицы $A^{(n-1)}$ обратным ходом находится решение исходной системы

(1). В обычном методе Гаусса элементы матрицы $A^{(k)}$ последовательно вычисляются по формулам:

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}} a_{kj}^{(k-1)}, k+1 \leq i \leq n, k+1 \leq j \leq n+1 \quad (2)$$

Следующий пример демонстрирует выполнение первого этапа метода Гаусса [1].

Пример 2.

$$\begin{cases} 3x_1 + 4x_2 - 2x_3 + x_4 = -2 \\ x_1 - x_2 + 2x_3 + 2x_4 = 7 \\ 4x_1 - 3x_2 + 4x_3 - 3x_4 = 2 \\ -x_1 + x_2 + 6x_3 - x_4 = 1 \end{cases} \quad (3)$$

Последовательность преобразований расширенной матрицы для данной системы выглядит следующим образом:

$$\begin{aligned} A^{(0)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 1 & -1 & 2 & 2 & 7 \\ 4 & -3 & 4 & -3 & 2 \\ -1 & 1 & 6 & -1 & 1 \end{bmatrix}; & A^{(1)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -\frac{7}{3} & \frac{8}{3} & \frac{5}{3} & \frac{23}{3} \\ 0 & -\frac{25}{3} & \frac{20}{3} & -\frac{13}{3} & \frac{14}{3} \\ 0 & \frac{7}{3} & \frac{16}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}; \\ A^{(2)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -\frac{7}{3} & \frac{8}{3} & \frac{5}{3} & \frac{23}{3} \\ 0 & 0 & -\frac{20}{7} & -\frac{72}{7} & -\frac{159}{7} \\ 0 & 0 & 8 & 1 & 8 \end{bmatrix}; & A^{(3)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -\frac{7}{3} & \frac{8}{3} & \frac{5}{3} & \frac{23}{3} \\ 0 & 0 & -\frac{20}{7} & -\frac{72}{7} & -\frac{159}{7} \\ 0 & 0 & 0 & -\frac{139}{5} & -\frac{278}{5} \end{bmatrix}. \end{aligned}$$

Как видно из примера, на промежуточных шагах первого этапа метода Гаусса могут возникать матрицы с дробными элементами. При этом, дроби каждый раз необходимо сокращать (в целях экономии памяти), что приводит к дополнительным накладным расходам при вычислении НОДа. Ясно, что при нахождении решения системы (1) можно избежать работы с дробями, например, используя правило Крамера. Однако, из-за слишком высокой вычислительной сложности ($O(n!)$) данный метод неприемлем для решения СЛАУ на компьютере.

На самом деле можно изменить метод Гаусса таким образом, чтобы на промежуточных шагах не возникали дробные числа.

3. Целочисленная модификация метода Гаусса

Как вычислять промежуточные матрицы в методе Гаусса, чтобы они были целочисленными? Если еще раз посмотреть на формулы (2), то нетрудно понять, как ответить на этот вопрос. Действительно, если перед тем, как прибавлять k -ю строчку к i -ой, домножить i -ю строчку на элемент a_{kk} , то вместо формул (2) получим формулы:

$$a_{ij}^{(k)} = a_{kk}^{(k-1)} a_{ij}^{(k-1)} - a_{ik}^{(k-1)} a_{kj}^{(k-1)}, k+1 \leq i \leq n, k+1 \leq j \leq n+1 \quad (4)$$

В данных формулах уже нет операции деления!

Пример 3. Рассмотрим систему (3). Выполним первый этап метода Гаусса для данной системы, используя формулы (4). Последовательность преобразований будет следующей:

$$\begin{aligned} A^{(0)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 1 & -1 & 2 & 2 & 7 \\ 4 & -3 & 4 & -3 & 2 \\ -1 & 1 & 6 & -1 & 1 \end{bmatrix}; & A^{(1)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & -25 & 20 & -13 & 14 \\ 0 & 7 & 16 & -2 & 1 \end{bmatrix}; \\ A^{(2)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & 0 & 60 & 216 & 477 \\ 0 & 0 & -168 & -21 & -168 \end{bmatrix}; & A^{(3)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & 0 & 60 & 216 & 477 \\ 0 & 0 & 0 & 35028 & 70056 \end{bmatrix}. \end{aligned}$$

Элементы во всех промежуточных матрицах теперь целочисленные, однако, в данном примере растут очень быстро. На самом деле, если элементы $a_{ij}^{(0)}$ δ -разрядные числа, то элементы $a_{ij}^{(k)}$ могут быть $2^k \delta$ -разрядными числами.

Таким образом, мы избавились от одной проблемы (работы с дробными числами) и получили другую проблему (экспоненциальный рост памяти, необходимой для хранения промежуточных матриц).

4. Метод Бариеса

Бариес предложил метод (1968 г.), который исключает работу с дробными числами и заботится о том, чтобы элементы промежуточных матриц не росли слишком быстро. В методе Бариеса вместо формул (4) используются следующие формулы (которые, согласно Бариесу, были известны еще Жордану [1]):

$$\begin{aligned} a_{00}^{(-1)} &= 1 \\ a_{ij}^{(k)} &= (a_{kk}^{(k-1)} a_{ij}^{(k-1)} - a_{ik}^{(k-1)} a_{kj}^{(k-1)}) / a_{k-1, k-1}^{(k-2)}, \\ k+1 \leq i \leq n, k+1 \leq j \leq n+1 \end{aligned} \quad (5)$$

Пример 4. Воспользуемся формулами (5) для выполнения первого этапа метода Гаусса для системы (3). Последовательность преобразований будет следующей:

$$\begin{aligned} A^{(0)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 1 & -1 & 2 & 2 & 7 \\ 4 & -3 & 4 & -3 & 2 \\ -1 & 1 & 6 & -1 & 1 \end{bmatrix}; & A^{(1)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & -25 & 20 & -13 & 14 \\ 0 & 7 & 16 & -2 & 1 \end{bmatrix}; \\ A^{(2)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & 0 & 20 & 72 & 159 \\ 0 & 0 & -56 & -7 & -56 \end{bmatrix}; & A^{(3)} &= \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & 0 & 20 & 72 & 159 \\ 0 & 0 & 0 & -556 & -1112 \end{bmatrix}. \end{aligned}$$

Все промежуточные матрицы целочисленные (хотя в формулах (5) есть операция деления), и элементы промежуточных матриц растут медленнее, чем в предыдущем примере. Случайно ли это? Оказывается, нет. Для того, чтобы это показать, рассмотрим лемму. Однако, перед этим введем некоторые обозначения.

Через $A \begin{pmatrix} a_1 & a_2 & \dots & a_p \\ b_1 & b_2 & \dots & b_p \end{pmatrix}$ обозначим минор p -го порядка матрицы A , состоящий из элементов матрицы A , стоящих на пересечении строчек с номерами $a_1, a_2, \dots, a_p$ и столбцов с номерами $b_1, b_2, \dots, b_p$. Также введем в рассмотрение окаймляющие минор $A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}$ определители.

$$b_{ik}^{(p)} = A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} \quad (i, k = p+1, \dots, n). \quad (6)$$

Матрицу, составленную из этих определителей, обозначим через

$$B^{(p)} = \left\| b_{ik}^{(p)} \right\|_{p+1}^n.$$

При данных обозначениях для исходной матрицы A справедлива следующая лемма [2]

Лемма 5 (Детерминантное тождество Сильвестра [2]).

$$|B^{(p)}| = \left[A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{n-p-1} |A|$$

для $1 \leq p \leq n-1$

Доказательство. Рассмотрим обычный метод Гаусса (т.е. тот, в котором элементы промежуточных матриц вычисляются по формулам (2)). На p -ом шаге исходная матрица $A^{(0)}$ будет преобразована к матрице

$$A^{(p)} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} & a_{1,p+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2p}^{(1)} & a_{2,p+1}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{pp}^{(p-1)} & a_{p,p+1}^{(p-1)} & \dots & a_{pn}^{(p-1)} \\ 0 & 0 & \dots & 0 & a_{p+1,p+1}^{(p)} & \dots & a_{p+1,n}^{(p)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{n,p+1}^{(p)} & \dots & a_{n,n}^{(p)} \end{pmatrix} \quad (7)$$

При переходе от матрицы $A^{(0)}$ к матрице $A^{(p)}$ к каждой строке матрицы $A^{(0)}$, начиная со 2-й и кончая n -й, последовательно прибавлялись какие-то предыдущие строчки (из числа первых p), помноженные на некоторые коэффициенты. Поэтому у матрицы $A^{(0)}$ и $A^{(p)}$ одинаковы все миноры p -го порядка, содержащиеся в первых p строках, а также все миноры $p+1$ -го порядка, содержащиеся в строках с номерами $1, 2, \dots, p, i$ ($i > p$):

$$\left. \begin{aligned} A^{(0)} \begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} &= A^{(p)} \begin{pmatrix} 1 & 2 & \dots & p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \\ (1 \leq k_1 < k_2 < \dots < k_p \leq n). \\ A^{(0)} \begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix} &= A^{(p)} \begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix} \\ (1 \leq k_1 < k_2 < \dots < k_{p+1} \leq n). \end{aligned} \right\} \quad (8)$$

Из (7) и (8) находим:

$$A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = a_{11} a_{22}^{(1)} \dots a_{pp}^{(p-1)} \quad (9)$$

$$A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} = a_{11} a_{22}^{(1)} \dots a_{pp}^{(p-1)} a_{ik}^{(p)} \quad (10)$$

Деля почленно (10) на (9), получим формулы:

$$a_{ik}^{(p)} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}} \quad (i, k = p+1, \dots, n). \quad (11)$$

Из (8) и (9) следует:

$$|A| = A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \begin{vmatrix} a_{p+1,p+1}^{(p)} & \dots & a_{p+1,n}^{(p)} \\ \dots & \dots & \dots \\ a_{n,p+1}^{(p)} & \dots & a_{n,n}^{(p)} \end{vmatrix} \quad (12)$$

Согласно формулам (11):

$$\begin{vmatrix} a_{p+1,p+1}^{(p)} & \dots & a_{p+1,n}^{(p)} \\ \dots & \dots & \dots \\ a_{n,p+1}^{(p)} & \dots & a_{n,n}^{(p)} \end{vmatrix} = \frac{\begin{vmatrix} b_{p+1,p+1}^{(p)} & \dots & b_{p+1,n}^{(p)} \\ \dots & \dots & \dots \\ b_{n,p+1}^{(p)} & \dots & b_{n,n}^{(p)} \end{vmatrix}}{\left[A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{n-p}} = \frac{|B^{(p)}|}{\left[A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{n-p}}. \quad (13)$$

Поэтому равенство (12) может быть переписано так:

$$|B^{(p)}| = \left[A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{n-p-1} |A| \quad (14)$$

□

Введем следующие обозначения:

$$a_{00}^{(-1)} = 1, [a_{ij}^{(0)}] = A^{(0)},$$

$$a_{ij}^{(k)} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1k} & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2k} & a_{2j} \\ \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} & a_{kj} \\ a_{i1} & a_{i2} & \dots & a_{ik} & a_{ij} \end{vmatrix}, k+1 \leq i \leq n, k+1 \leq j \leq n+1,$$

и применим лемму к матрице

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} & a_{1j} \\ a_{21} & a_{22} & \dots & a_{2k} & a_{2j} \\ \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kk} & a_{kj} \\ a_{i1} & a_{i2} & \dots & a_{ik} & a_{ij} \end{pmatrix}.$$

Получим следующее выражение:

$$a_{ij}^{(k)} [a_{pp}^{(p-1)}]^{k-p} = \begin{vmatrix} a_{p+1,p+1}^{(p)} & \dots & a_{p+1,k}^{(p)} & a_{p+1,j}^{(p)} \\ \dots & \dots & \dots & \dots \\ a_{k,p+1}^{(p)} & \dots & a_{kk}^{(p)} & a_{kj}^{(p)} \\ a_{i,p+1}^{(p)} & \dots & a_{ik}^{(p)} & a_{ij}^{(p)} \end{vmatrix}, \quad (15)$$

для $1 \leq p \leq k-1$. Отсюда видно, что правая часть (15) делится на множитель $[a_{pp}^{(p-1)}]^{k-p}$. Если в (15) положить $p = k-1$, то получим:

$$a_{ij}^{(k)} = \frac{1}{[a_{k-1,k-1}^{(k-2)}]} \begin{vmatrix} a_{kk}^{(k-1)} & a_{k,j}^{(k-1)} \\ a_{i,k}^{(k-1)} & a_{ij}^{(k-1)} \end{vmatrix}, \quad (16)$$

что в точности совпадает с (5).

Бариев [3] показал, как (15) может быть использована для исключения $k-p$ переменных за раз. Наиболее интересен случай при $p = k-2$:

$$a_{ij}^{(k)} = \frac{1}{[a_{k-2,k-2}^{(k-3)}]^2} \begin{vmatrix} a_{k-1,k-1}^{(k-2)} & a_{k-1,k}^{(k-2)} & a_{k-1,j}^{(k-2)} \\ a_{k,k-1}^{(k-2)} & a_{kk}^{(k-2)} & a_{k,j}^{(k-2)} \\ a_{i,k-1}^{(k-2)} & a_{ik}^{(k-2)} & a_{ij}^{(k-2)} \end{vmatrix} \quad (17)$$

При этом используются следующие формулы:

$$a_{00}^{(-1)} = 1, [a_{ij}^{(0)}] = A^{(0)}, \quad (18)$$

$$c_0^{(k-2)} = (a_{k-1,k-1}^{(k-2)} a_{kk}^{(k-2)} - a_{k-1,k}^{(k-2)} a_{k,k-1}^{(k-2)}) / a_{k-2,k-2}^{(k-3)}, \quad (19)$$

$$c_{i1}^{(k-2)} = (a_{k-1,k}^{(k-2)} a_{i,k-1}^{(k-2)} - a_{k-1,k-1}^{(k-2)} a_{i,k}^{(k-2)}) / a_{k-2,k-2}^{(k-3)}, \quad (20)$$

$$c_{i2}^{(k-2)} = (a_{k,k-1}^{(k-2)} a_{ik}^{(k-2)} - a_{kk}^{(k-2)} a_{i,k-1}^{(k-2)}) / a_{k-2,k-2}^{(k-3)}, \quad (21)$$

$$a_{ij}^{(k)} = c_0^{(k-2)} a_{ij}^{(k-2)} + c_{i1}^{(k-2)} a_{kj}^{(k-2)} + c_{i2}^{(k-2)} a_{k-1,j}^{(k-2)} / a_{k-2,k-2}^{(k-3)}, \quad (22)$$

для $k+1 \leq i \leq n, k+1 \leq j \leq n+1$;

$$a_{kk}^{(k-1)} = c_0^{(k-2)}, \quad (23)$$

$$a_{kl}^{(k-1)} = a_{kl}^{(k)} = (a_{k-1,k-1}^{(k-2)} a_{kl}^{(k-2)} - a_{k-1,l}^{(k-2)} a_{k,k-1}^{(k-2)}) / a_{k-2,k-2}^{(k-3)}, \quad (24)$$

для $k+1 \leq l \leq n+1$, которые применяются для $k = 2, 4, \dots, 2 \lfloor \frac{n-1}{2} \rfloor$.

В случае, когда n четное, последний шаг выполняется по формулам (5).

Пример 6. Выполним первый этап метода Гаусса для системы (3), используя формулы (18)–(24). Последовательность преобразований будет следующей:

$$A^{(0)} = \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 1 & -1 & 2 & 2 & 7 \\ 4 & -3 & 4 & -3 & 2 \\ -1 & 1 & 6 & -1 & 1 \end{bmatrix}; \quad A^{(2)} = \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & 0 & 20 & 72 & 159 \\ 0 & 0 & -56 & -7 & -56 \end{bmatrix};$$

$$A^{(3)} = \begin{bmatrix} 3 & 4 & -2 & 1 & -2 \\ 0 & -7 & 8 & 5 & 23 \\ 0 & 0 & 20 & 72 & 159 \\ 0 & 0 & 0 & -556 & -1112 \end{bmatrix}.$$

Анализ относительной эффективности одношагового и двушагового методов [4] показывает, что последний асимптотически быстрее на 50%.

Алгоритм

Вход: $A \in \mathbb{Z}^{n \times n}$, $b \in \mathbb{Z}^n$, $\det A \neq 0$;

Выход: Расширенная матрица, преобразованная по формулам (5);

```

1:  $d := 1$ ; // инициализация
2: for  $k = 1, \dots, n$  do
3:   if  $a_{kk} = 0$  then
4:      $i = k + 1$ ;
5:     while  $(i \leq n) \& (a_{ik} = 0)$  do
6:        $i = i + 1$ ;
7:     if  $i \leq n$  then
8:       переставить  $i$ -ю и  $k$ -ю строчки;
9:     else
10:      вырожденность;
11:   for  $i = k + 1, \dots, n$  do
12:     for  $j = k + 1, \dots, n$  do
13:        $a_{ij} = \frac{1}{d}(a_{ij}a_{kk} - a_{kj}a_{ik})$ ;
14:      $b_i = \frac{1}{d}(b_i a_{kk} - b_k a_{ik})$ ;
15:      $a_{ik} = 0$ ;
16:      $a_{ii} = a_{kk}$ ;
17:      $i = i + 1$ ;
18:    $d = a_{kk}$ ;
```

5. Заключение

Хотя алгоритм Бариеса хороший и довольно просто реализуется, существуют и другие методы решения систем линейных уравнений. Данные методы хорошо приспособлены к конкретным типам задач (например, системы уравнений, коэффициентами которых являются полиномы от нескольких переменных). Познакомиться с данными методами можно в [1].

Список литературы

- [1] Geddes K.O., Czapor S.R., Labahn G. Algorithms for Computer Algebra, 1992.
- [2] Гантмахер Ф.Р. Теория матриц, 1967.
- [3] Bareiss E.H. ‘Sylvester’s Identity and Multistep Integer-Preserving Gaussian Elimination // Math. Comp., 1968, 22 (103), P. 565–578.
- [4] Bareiss E.H. Computational Solutions of Matrix Problems Over an Integral Domain // J. Inst. Maths Applcs, 1972, 10, P. 68–104.