Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Московский государственный технический университет имени Н. Э. Баумана (национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ Фундаментальные науки

КАФЕДРА Прикладная математика

ОТЧЕТ К ЛАБОРАТОРНОЙ РАБОТЕ

НА ТЕМУ:

Прямые методы решения систем линейных алгебраических уравнений

Студент ФН2-51Б (Группа) ________ (Подпись, дата) А. С. Киселева (И. О. Фамилия)

Проверил ________ (Подпись, дата) ________ (И. О. Фамилия)

2021 г.

Оглавление

1. Описание алгоритмов

Метод Гаусса. Рассмотрим два прямых метода решения СЛАУ $Ax = b$: метод Гаусса и метод QR -разложения.
Запишем систему уравнений в координатной форме:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, ..., n.$$

Метод Гаусса (метод последовательного исключения неизвестных), состоит в том, что неизвестные $x_j, j = 1, 2, ..., n-1$, последовательно исключаются из системы и в результате она преобразуется к эквивалентной системе с треугольной матрицей:

$$\begin{cases} a_{11}^{(0)}x_1 + a_{12}^{(0)}x_2 + a_{13}^{(0)}x_3 + a_{14}^{(0)}x_4 + ... + a_{1n}^{(0)}x_n = b_1^{(0)}, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + a_{24}^{(1)}x_4 + ... + a_{2n}^{(1)}x_n = b_2^{(1)}, \\ a_{33}^{(2)}x_3 + a_{34}^{(2)}x_4 + ... + a_{3n}^{(2)}x_n = b_3^{(1)}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{ii}^{(i-1)}x_{-} + ... + a_{in}^{(i-1)}x_n = b_i^{(i-1)}, \\ \dots\dots\dots\dots\dots\dots \\ a_{nn}^{(n-1)}x_n = b_n^{(n-1)}. \end{cases}$$

Коэффициенты $a_{ij}^{(k)}$ и компоненты правой части $b_i^{(k)}$ записанной системывычисляются по формулам

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - c_{ik}a_{kj}^{(k-1)},$$
$$b_i^{(k)} = b_i^{(k-1)} - c_{ik}b_k^{(k-1)},$$

где

$$c_{ik} = a_i^{(k-1)}/a_{kk}^{(k-1)},$$
$$k = 1, 2, \dots, (n-1), \quad j = k, k+1, \dots, n, \quad i = k+1, k+1, \dots, n,$$

причем

$$a_{ij}^{(0)} = a_{ij}, \quad b_i^{(0)} = b_i.$$

Вычисления по записанным формулам называются прямым ходом метода Гаусса. Затем неизвестные x_i последовательно, начиная с x_n, определяются из системы по формулам:

$$x_i = \left(b_i^{(i-1)} - \sum_{j=i+1}^{n} a_{ij}^{(i-1)} x_j \right) / a_{ii}^{(i-1)}, \quad i = n, n-1, \ldots, 1.$$

Вычисления по данным формулам называется обратным ходом метода Гаусса.

Для реализации прямого хода метода Гаусса требуется порядка $O(n^3/3)$ операций умножения и деления чисел с плавающей точкой, для обратного — порядка $O(n^2/2)$.

Метод Гаусса применим в том случае, когда все угловые миноры матрицы A ненулевые, что равносильно требованию $a_{ij}^{(i-1)} \neq 0$ для всех значений $i = 1, 2, \ldots, n$.

Метод QR-разложения. Метод QR-разложения основан на представлении матрицы системы в виде произведения ортогональной матрицы Q и верхней треугольной матрицы R. Один из способов получения такого разложения — метод вращений.

Сначала неизвестное x_1 исключается из всех уравнений, кроме первого. Это производится при помощи следующего алгоритма. Для исключения x_1 из второго уравнения вычисляются коэффициенты

$$c_{12} = \frac{a_{11}}{\sqrt{a_{11}^2 + a_{21}^2}}, \quad s_{12} = \frac{a_{21}}{\sqrt{a_{11}^2 + a_{21}^2}}$$

затем первое уравнение системы заменяется линейной комбинацией первого и второго уравнений с коэффициентами c_{12} и s_{12}, а второе уравнение — линейной комбинацией тех же уравнений, но уже с коэффициентами $-s_{12}$ и c_{12}. Так

как $-s_{12}a_{11} + c_{12}a_{11} = 0$, коэффициент во втором уравнении при x_i обратится в нуль.

В итоге исходная система будет приведена к виду:

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + a_{13}^{(1)}x_3 + ... + a_{1n}^{(1)}x_n = b_1^{(1)}, \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + ... + a_{2n}^{(1)}x_n = b_2^{(1)}, \\ a_{31}^{(1)}x_1 + a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 + ... + a_{3n}^{(1)}x_n = b_3^{(1)}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{n1}^{(1)}x_1 + a_{n2}^{(1)}x_2 + a_{n3}^{(1)}x_3 + ... + a_{nn}^{(1)}x_n = b_n^{(1)}, \end{cases}$$

Это преобразование эквивалентно умножению матрицы системы уравнений и вектора правой части слева на ортогональную матрицу T_{12}, имеющую вид

$$T_{12} = \begin{pmatrix} c_{12} & s_{12} & 0 & 0 & \dots & 0 \\ -s_{12} & c_{12} & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Так как коэффициенты c_{12} и s_{12} подобраны таким образом, что $c_{12}^2 + s_{12}^2 = 1$ то можно считать, что

$$c_{12} = \cos\varphi, \quad s_{12} = \sin\varphi.$$

Следовательно, матрица T_{12}— это матрица поворота на угол 3 по часовой стрелке в плоскости (x_1, x_2).

Для исключения x_1 из третьего уравнения, используются коэффициенты c_{13} и s_{13}:

$$c_{13} = \frac{a_{11}^{(1)}}{\sqrt{(a_{11}^{(1)})^2 + (a_{31}^{(1)})^2}}, \quad c_{13} = \frac{a_{31}^{(1)}}{\sqrt{(a_{11}^{(1)})^2 + (a_{31}^{(1)})^2}}$$

Далее первое и третье уравнение заменяются своими линейными комбинациями.

Эта операция равносильна умножению слева матрицы $A^{(1)} = T_{12}A$ и вектора правой части $b^{(1)} = T_{12}b$ на ортогональную матрицу, имеющую вид

$$T_{13} = \begin{pmatrix} c_{13} & 0 & s_{13} & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ -s_{13} & 0 & c_{13} & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Аналогично неизвестная x_1 исключается из остальных уравнений, затем x_2 — из всех уравнений, кроме первого и второго, при этом используются матрицы $T_{23}, T_{24}, \dots, T_{2n}$ и так далее. Процесс продолжается, пока система не будет приведена к верхней треугольной форме. То есть $T = T_{n-1,n} \cdot \ldots \cdot T_{24} \cdot T_{23} \cdot T_{1n} \cdot \ldots \cdot T_{13} \cdot T_{12}$. Причём, $R = TA$, где R — полученная верхнетреугольная матрица и $Q = T^{-1} = T^T$.

2. Исходные данные

13 вариант

Первая система:

$$A = \begin{pmatrix} -190.3270 & 189.7600 & -18.0160 & 72.0640 \\ -194.5200 & 193.9530 & -18.4320 & 73.7280 \\ -919.4800 & 919.4800 & -99.2470 & 398.6800 \\ -219.3900 & 219.3900 & -23.7720 & 95.5110 \end{pmatrix}, \quad b = \begin{pmatrix} 1414.2810 \\ 1446.7620 \\ 7705.5000 \\ 1845.1920 \end{pmatrix}$$

Вторая система:

$$A = \begin{pmatrix} -38.4000 & 4.0300 & 8.3800 & 3.5300 \\ -8.3200 & -81.2000 & -8.0900 & -3.6700 \\ 4.3300 & 7.2100 & -110.8000 & -2.6300 \\ 4.2200 & 4.2200 & 6.1500 & 73.8000 \end{pmatrix}, \quad b = \begin{pmatrix} 292.5900 \\ -504.3200 \\ -185.6600 \\ -430.5000 \end{pmatrix}$$

26 вариант

Первая система:

$$A = \begin{pmatrix} 208.594 & -37.974 & 69.564 & 0.568 \\ -626.574 & 114.052 & -208.95 & -1.804 \\ -625.785 & 113.922 & 680.556 & -1.692 \\ 104.424 & -19.008 & -113.552 & 0.304 \end{pmatrix}, \quad b = \begin{pmatrix} 0.568 \\ -1.804 \\ -1.692 \\ 0.304 \end{pmatrix}$$

Вторая система:

$$A = \begin{pmatrix} 86.0000 & -8.9300 & -9.5900 & -3.9100 \\ 4.0500 & -100.0000 & -9.1000 & -8.1400 \\ 0.2600 & 3.6100 & -71.8000 & -4.2800 \\ -4.0300 & -6.8800 & 6.5700 & -198.6000 \end{pmatrix}, \quad b = \begin{pmatrix} 818.5800 \\ 898.7400 \\ -912.2200 \\ -687.0600 \end{pmatrix}$$

3. Результаты расчетов

Результат для 13 варианта первой системы для типа double:
Метод Гаусса

$$x = (1 \quad 2 \quad 20 \quad 22)^T$$

$$||b - Ax|| = 9.37486 * 10^{-13}$$

$$||b - Ax|| = 1.7211 * 10^{-3}(\text{для float})$$

QR **- метод**

$$x = (1 \quad 2 \quad 20 \quad 22)^T$$

$$||b - Ax|| = 5.56949 * 10^{-13}$$

$$||b - Ax|| = 5.59396 * 10^{-4}(\text{для float})$$

Внесение возмущения в вектор b:

$$b = (1414.28 \quad 1446.76 \quad 7705.5 \quad 1845.19)^T$$

Решение для возмущенной правой части:

$$x = (91.831 \quad 99.1341 \quad 1288.45 \quad 323.23)^T$$

Точная оценка числа обусловленности:

$$\mathrm{cond}_1 A = 5.35508 * 10^{10}, \qquad \mathrm{cond}_\infty A = 1.48185 * 10^{11}$$

Оценка снизу: $\mathrm{cond}_1 A \geq 8.88 \times 10^9$ и $\mathrm{cond}_\infty A \geq 1.48 \times 10^{11}$

Результат для 26 варианта первой системы для типа double:
Метод Гаусса

$$x = (1 \quad 5 \quad 3 \quad 1)^T$$

$$||b - Ax|| = 3.00424 * 10^{-14}$$

$$||b - Ax|| = 1.56322 * 10^{-4}(\text{для float})$$

QR **- метод**

$$x = (1 \quad 5 \quad 3 \quad 1)^T$$

$$||b - Ax|| = 1.30439 * 10^{-13}$$

$$||b - Ax|| = 1.12621 * 10^{-4} (\text{для float})$$

Внесение возмущения в вектор b:

$$b = (0.5679 \quad -1.8039 \quad -1.6921 \quad 0.3039)^T$$

Решение для возмущенной правой части:

$$x = (149.669 \quad -418.856 \quad -434.256 \quad 74.5718)^T$$

Точная оценка числа обусловленности:

$$\text{cond}_1 A = 4.07616 * 10^{10}, \qquad \text{cond}_\infty A = 3.86299 * 10^{11}$$

Оценка снизу: $\text{cond}_1 A \geq 3.50 \times 10^6$ и $\text{cond}_\infty A \geq 3.87 \times 10^{10}$

4. Ответы на контрольные вопросы

1. Каковы условия применимости метода Гаусса без выбора и с выбором ведущего элемента?

В методе Гаусса вычисления возможны, если ведущие элементы матрицы A не равны нулю, то есть $a_{kk}^{k-1} \neq 0$ для $k = 1, 2, \ldots, n$. Но из условия невырожденности матрицы ($\det A \neq 0$) не следует, что в ходе приведения матрицы A к треугольному виду на диагонали не возникнет элементов, равных нулю или малых по абсолютной величине (поскольку это приводит к дополнительным ошибкам округления в вычислениях). В таких случаях метод Гаусса неприменим, поэтому на практике обычно используется вариант алгоритма Гаусса с частичным либо полным выбором главного элемента. Метод Гаусса применим тогда и только тогда, когда все угловые миноры матрицы не нулевые, что равносильно тому, что главные элементы $a_{kk}^{k-1} \neq 0$.

2. Докажите, что если det $A \neq 0$, то при выборе главного элемента в столбце среди элементов, лежащих не выше главной диагонали, всегда найдется хотя бы один элемент, отличный от нуля.

Допустим, что на i шаге была получена матриц, такая, что все элементы i-го столбца не выше главной диагонали нулевые. Выберем главный элемент в столбце, т.е. максимальный по модулю элемент в данном столбце, лежащий на главной диагонали или ниже нее. Определитель матрицы находится через миноры, если мы поменяем местами главный элемент с элементом на главной диагонали, определитель может только поменять свой знак. Тогда, если det $A \neq 0$, то и новый определитель не будет равняться нулю.

3. В методе Гаусса с полным выбором ведущего элемента приходится не только переставлять уравнения, но и менять нумерацию неизвестных. Предложить алгоритм, позволяющий восстановить первоначальный порядок неизвестных.

Использовать дополнительный массив, в котором будут храниться номера строк и столбцов.

4. Оцените количество арифметических операций, требуемых для QR-разложения произвольной матрицы A размера $n \times n$.

Рассмотрев алгоритм метода $QR-$разложения, получим, что для получения верхнетреугольной матрицы R требуется порядка $\frac{4}{3}n^3$, а для получения ортогональной мтрицы Q - $2n^3$. Получается $\frac{10}{3}n^3$. Для реализации метода $QR-$разложения умножим матрицу A на матрицу T_{12}, имеющую вид:

$$T_{12} = \begin{pmatrix} c_{12} & s_{12} & 0 & 0 & \dots & 0 \\ -s_{12} & c_{12} & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Для этого понадобится $4n$ умножений. Затем x_1 исключается из всех уравнений, кроме первого, потребуетмя $n-1$ операций. Далее матрицу T_{23} умножаем на матрицу $A^{(1)}$ и исключаем x_2 $n-2$ раз. В итоге получим сумму:

$$S_R = 4n(n-1) + 4(n-1)(n-2) + \ldots \approx 4\sum_{i=1}^{n-1} i^2$$

Пусть $i^2 = \frac{1}{3}((i+1)^3 - i^3 - 3i - 1)$ тогда,

$$4\sum_{i=1}^{n-1} i^2 = \frac{4}{3}\sum_{i=1}^{n-1}((i+1)^3 - i^3 - 3i - 1) \approx \frac{4}{3}n^3.$$

Матрица $Q = T^{-1} = T^T$, T — матрица результирующего вращения, $S_Q = 2n^3$.

$$S_R + S_Q = \frac{4}{3}n^3 + 2n^3 = \frac{10}{3}n^3.$$

5. Что такое число обусловленности и что оно характеризует? Имеется ли связь между обусловленностью и величиной определителя матрицы? Как влияет выбор нормы матрицы на оценку числа обусловленности?

Число $M_A = \|A^{-1}\|\|A\|$ называют числом обусловленности матрицы . Оно характеризует степень звисимости относительной погрешности решения от относительной погрешности правой части. Если мы умножим матрицу A на костанту $\alpha \neq 0$, то обусловленность не изменится, так как обратная матрица умножится на величину α^{-1}. Поэтому определитель и число обусловленности не связаны. Выбор нормы матрицы хоть и влияет на число обусловленности, но при любом выборе сможем корректно отразить порядок этого числа. Если матрица хорошо обусловлена, то ее число обусловленности будет мало для любой нормы, а если плохо обусловлена, то число обусловленности будет большим для любой нормы. Норму можно корректировать. Подсчет определителя – дорогая операция, и определитель не несет смысловой нагрузги. Число обусловленности показывет степень вырожденности матрицы коэффициентов и позволяет определить степень чувствительности численных решений к ошибкам, то есть, число обусловленности входит в оценку. Рассмотрим, например, единичную матрицу. Определитель ее равен 1, как и число обусловленности. Если мы домножим коэффициенты матрицы на очень малую константу, то число обусловленности не изменится, в отличие от определителя.

6. Как упрощается оценка обусловленности, если матрица является:

а) диагональной;

б) симметричной;

в) ортогональной;

г) положительно определённой;

д) треугольной?

а. $condD = \frac{\max|d_{ii}|}{\min|d_{ii}|}$, где D — диагональная матрица, а $\max|d_{ii}|$ — максимальный элемент и $\min|d_{ii}|$- минимальный элемент.

б. $condA = \frac{a_{max}}{a_{min}}$, a_{max} — максимальное собственное число матрицы, a_{min} — минимальное собственное число матрицы.

в. Для оценки нормы используют тот факт, что для ортогональной матрицы $A^{-1} = A^T$, тогда $condA = \|A\|^2$.

г. Собственные числа положительно определенной матрицы являются вещественными числами, поэтому в этом случае можно считать число обусловленности через собственные значения, используя формулу $\text{cond } D = \frac{|a_{max}|}{|a_{min}|}$.

д. $condA = \frac{a_{max}}{a_{min}}$, a_{max} — максимальный элемент на диагонали, a_{min} — минимальный элемент на диагонали.

7. Применимо ли понятие числа обусловленности к вырожденным матрицам?

Если матрица вырожденная, то обратной матрицы не существует. Значит, понятие числа обусловленности не применимо к варожденным матрицам.

8. В каких случаях целесообразно использовать метод Гаусса, а в каких — методы, основанные на факторизации матрицы?

Метод Гаусса и методы, основанные на факторизации матриц, практически не отличаются по числу операций. Для общего случая метод Гаусса является более эффективным во временном значении и требует меньшее количество действий. Алгоритм $LU-$факторизации целесообразен для матрицы с одинаковой системой коэффициентов, но разными правыми частями.

9. Как можно объединить в одну процедуру прямой и обратный ход метода Гаусса? В чём достоинства и недостатки такого подхода?

Их можно объединить, если после приведения матрицы к верхнетреуглльному виду пойти обратно вверх и привести матрицу с диагональной, но на это потребуется порядка $\frac{n^3}{3}$ операций, а это невыгодно.

10. Объясните, почему, говоря о векторах, норму $\|x\|_1$ часто называют октаэдрической, норму $\|x\|_2$ — шаровой, а норму $\|x\|_\infty$ — кубической.

а. $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i|$ называется октаэдрической. Рассмотрим, как будут выглядеть множество всех единичных векторов, норма котрых равна единице, если ввести эту норму на плоскости $|x_1| + |x_2| = 1$.

Ромб. В трехмерном случае, соответственно, октаэдр.

б. $\|x\|_1 = \sqrt{\sum\limits_{i=1}^{n} |x_i|^2}$. В этом случае геометрическое место на плоскости точек будет задаваться уравнением $x_1^2+x_2^2 = 1$. Окружность. В трехмерном случае, соответственно, шар.

в. $||x||_\infty = \max |x_i|$, на плоскости множество всех единичных векторов, норма которых равна единице, будет выглядеть как квадрат. В трехмерном случае, соответственно, куб.