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 по лабораторной работе №1

на тему:

"Прямые методы решения систем линейных алгебраических уравнений"

Студент	ФН2-51Б (Группа)	(Подпись, дата)	М. А. Каган (И. О. Фамилия)
Студент	ФН2-51Б (Группа)	(Подпись, дата)	И. А. Яковлев (И. О. Фамилия)
Проверил		(Подпись, дата)	(И. О. Фамилия)

2024 г.

Оглавление

1. Исходные данные

Система линейных уравнений №1:

$$\begin{cases} 0.2910x_1 & + & 1.8100x_2 & + & 9.3110x_3 & + & 9.1100x_4 & = & 4.2280 \\ 1.4500x_1 & + & 8.5790x_2 & + & 44.1950x_3 & + & 42.9950x_4 & = & 20.4290 \\ -0.2900x_1 & - & 1.7980x_2 & - & 9.2500x_3 & - & 9.0500x_4 & = & -4.2080 \\ 0.0000x_1 & + & 0.0820x_2 & + & 0.4100x_3 & + & 0.4500x_4 & = & 0.1220 \end{cases}$$

Система линейных уравнений №2:

$$\begin{cases} -106.4000x_1 & - & 7.0000x_2 & - & 4.9900x_3 & + & 0.2600x_4 & = & 1040.8100 \\ 3.6100x_1 & + & 22.2000x_2 & - & 8.5900x_3 & - & 8.9200x_4 & = & 615.4100 \\ 2.2800x_1 & + & 7.7500x_2 & + & 52.2000x_3 & + & 9.6500x_4 & = & 427.5400 \\ -9.0000x_1 & + & 5.8100x_2 & - & 0.0900x_3 & + & 136.8000x_4 & = & -265.3500 \end{cases}$$

2. Краткие сведения

Пусть A — невырожденная матрица $n \times n$, b — ненулевой n-мерный вектор. Необходимо найти такой n-мерный вектор x, чтобы он удовлетворял уравнению

$$Ax = b. \tag{1}$$

2.1. Алгоритм Гаусса

Метод Гаусса заключается в последовательном обнулении элементов, находящихся под главной диагональю матрицы A (приведение системы к верхнетреугольному виду), и последовательное вычисление элементов неизвестного вектора x решением линейных уравнений. Вычисление элементов матрицы происходит по формуле $a_{ij}^{(k)} = a_{ij}^{(k-1)} - c_{ik}a_{kj}^{(k-1)}$, где k - номер итерации, $c_{ik} = a_{ik}^{(k)} \,/\, a_{kk}^{(k)}$.

2.2. Метод QR-разложения

Метод QR-разложения заключается в том, чтобы матрицу A представить как произведение ортогональной матрицы Q и верхней треугольной матрицы R.

Рассмотрим алгоритм метод вращений, для нахождения матрицы $T = Q^{\mathrm{T}}$: пусть

$T_{i\,j}$ — матрица $n \times n$ вида :

$$T_{i\,j}(A) = \begin{pmatrix} 1 & 0 & \multicolumn{5}{c}{\dots\dots\dots\dots\dots\dots} & 0 \\ 0 & 1 & \multicolumn{5}{c}{\dots\dots\dots\dots\dots\dots} & 0 \\ \multicolumn{8}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ 0 & 0 & \dots & c & \dots & s & \dots & 0 \\ \multicolumn{8}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ 0 & 0 & \dots & -s & \dots & c & \dots & 0 \\ \multicolumn{8}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\ 0 & 0 & \multicolumn{5}{c}{\dots\dots\dots\dots\dots\dots} & 1 \end{pmatrix}, \quad \text{где } 1 \leqslant i < j \leqslant n.$$

Коэффициенты c и s вычисляются по следующим формулам:

$$c = \frac{a_{i\,i}}{\sqrt{a_{i\,i}^2 + a_{j\,i}^2}}, \quad s = \frac{a_{i\,j}}{\sqrt{a_{i\,i}^2 + a_{j\,i}^2}},$$

При умножении $T_{i\,j}(A) \cdot A = A^{(1)}$ элемент $a_{j\,i}^{(1)}$ равняется нулю, следовательно

$$T = T_{n-1\,n} \cdot \ldots \cdot T_{2\,4} T_{2\,3} T_{1\,n} \ldots \cdot T_{1\,3} T_{1\,2}.$$

где $T_{i\,j} = T_{i\,j}(A^{(l)}),\ l = (i-1)n + j - i$, а верхняя треугольная матрица $R = TA$. Таким образом, уравнение (1) сводится к двум более простым уравнениям:

$$\begin{cases} Rx = b^*, \\ Qb^* = b. \end{cases}$$

Контрольные вопросы

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

   Ответ:

   Метод Гаусса без выбора ведущего элемента может быть применен только в том случае, если в процессе преобразования невырожденной матрицы, на главной диагонали не возникает нулевых элементов, поскольку данный алгоритм предполагает деление на элементы $a_{i\,i}^{(i-1)}$, где $i = 1, \ldots, n-1$. Таким образом, чтобы расширить класс матриц, над которым применим метод Гаусса, достаточно предварительно «выбирать» неизвестное с ненулевым коэффициентом. Выбор происходит посредством перестановок строк и столбцов матрицы.

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.

Ответ:

Пусть при выборе главного элемента в одном из столбцов среди элементов, лежащих не выше главной диагонали нет ни одного отличного от нуля элемента. Тогда, после приведения матрицы к верхнетреугольному виду, определитель матрицы $\det A$ будет равен произведению элементов на главной диагонали; так как один из элементов равен нулю, $\det A = 0$, что противоречит исходному предположению.

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

Ответ:

Чтобы восстановить изначальный порядок неизвестных, можно хранить их индексы в отдельном массиве, каждый раз меняя их местами вместе со столбцами. При окончании работы алгоритма необходимо отсортировать решение по переставленным индексам.

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

Ответ:

Для вычисления констант c и s необходимо произвести минимум две операции умножения и две операции деления. Для умножения матрицы A на матрицу T_{ij} достаточно заменить строки i и j на их линейные комбинации, для чего требуется умножить каждый элемент i-й и j-й строки сначала на c и s, затем на $-s$ и c соответственно: в сумме получается $4n$ операции. Для приведения матрицы A к верхнему треугольному виду необходимо обнулить $\frac{1}{2}n(n-1)$ элементов, а для обнуления одного элемента необходимо произвести $4n+4$ операций, следовательно в сумме получается $2n(n+1)(n-1)$ операций. Для решения полученной системы необходимо произвести еще приблизительно $n(n-1)$ операций. В конечном итоге получаем: $2n^3 + n^2 - 3n \sim 2n^3$.

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?

Ответ:

(a) Рассмотрим линейную систему уравнений заданных с погрешностью:

$$A(x + \Delta x) = b + \Delta b.$$

Оценим погрешности Δx и Δb:

$$\|\Delta x\| \leqslant \|A^{-1}\|\|\Delta b\| \tag{2}$$

Теперь поделим на $\|x\|$ и $\frac{\|b\|}{\|A\|}$ левую и правую часть соответственно:

$$\frac{\|\Delta x\|}{\|x\|} \leqslant \|A\|\|A^{-1}\|\frac{\|\Delta b\|}{\|b\|},$$

$$\frac{\|\Delta x\|}{\|x\|} \leqslant \frac{\|\Delta b\|}{\|b\|}\operatorname{cond} A, \tag{3}$$

образом была получена связь между относительными погрешностями правой и левой части. Значение $\|A\|\|A^{-1}\|$ будем называть числом обусловленности матрицы A, которое характеризует насколько ошибки входных данных влияют на ошибку результата и наоборот.

(b) Между числом обусловленности и определителем матрицы нет прямой связи: с одной стороны, чем меньше $\det A$, тем больше $\det A^-1$, из-за чего $\|A^{-1}\|$ также становится больше, следовательно, из оценки (2) следует большее влияние погрешностей левой части на правую. С другой стороны, если собственные значения матрицы не сильно различаются, то и значение $\operatorname{cond} A$ не будет слишком велико, например, пусть матрица:

(c) Выбор нормы матрицы на прямую влияет на оценку числа обусловленности — это следует из оценки (3). Приведем пример: рассмотрим матрицу

$$A = \begin{pmatrix} 0 & 100 & 0 \\ 34 & 34 & 34 \\ 1 & 0 & 0 \end{pmatrix}, \quad A^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{100} & 0 & 0 \\ \frac{1}{100} & \frac{1}{34} & -1 \end{pmatrix}.$$

Тогда число обусловленности посчитанное по кубической норме будет равно $\operatorname{cond}_\infty A \approx 106$, а по октаэдрической норме будет равно $\operatorname{cond}_1 A \approx 268$, т.е. разница больше чем в два раза.

6. **Как упрощается оценка числа обусловленности, если матрица является:**
   (a) **диагональной;**
   (b) **симметричной;**
   (c) **ортогональной;**
   (d) **положительно определенной;**
   (e) **треугольной;**

   Ответ:

   (a) У диагональной матрицы A собственные значения являются диагональными элементами этой матрицы $a_{ii}, 1 \le i \le n$. Обратная к диагональной матрице A^{-1} тоже является диагональной, причем её диагональными элементами будут $\frac{1}{a_{ii}}, 1 \le i \le n$; тогда число обусловленности
   $$cond A \ge \frac{\max\limits_{1\le i\le n}|a_{ii}|}{\min\limits_{1\le i\le n}|a_{ii}|}.$$

   (b) Для вычисления числа обусловленности необходимо вычислить нормы матрицы $\|A\|$ и $\|A^{-1}\|$; ввиду симметрии, требуемые для нахождения норм $\|A\|_\infty = \max\limits_i \sum\limits_{j=1}^{n}|a_{ij}|$, $\|A\|_1 = \max\limits_j \sum\limits_{i=1}^{n}|a_{ij}|$ и $\|A\|_2 = \Big(\sum\limits_{i,j=1}^{n} a_{ij}^2\Big)^{1/2}$ вычисления будут в половину раз проще.

   (c) Ортогональная матрица A имеет следующее свойство: $A^T A = E$, где E – единичная матрица. Таким образом, $A^{-1} = A^T$, и её число обусловленности $\mathrm{cond}A \ge \|A\|\|A^{-1}\| = \|A\|\|A^T\|$, что упрощает вычисления.

   (d) Положительно определенная матрица имеет все собственные значения положительные, что позволяет получить оценку для числа обусловленности через их соотношение:
   $$\mathrm{cond}A = \frac{\lambda_{max}}{\lambda_{min}}.$$

   (e) Для треугольной матрицы (верхней или нижней) оценку числа обусловленности также можно получить через её диагональные элементы, так как собственные значения треугольной матрицы находятся на её диагонали:
   $$\mathrm{cond}A = \frac{\max|\lambda_i|}{\min|\lambda_i|}.$$

7. **Применимо ли понятие числа обусловленности к вырожденным матрицам?**

Ответ:

Нет, поскольку по определению обусловленности:

$$\mathrm{cond}A = \|A\|\|A^{-1}\|,$$

то есть обязана существовать обратная матрица.

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

Ответ:

Методы разложения требуют значительно больше операций, чем метод Гаусса, однако их удобно использовать, если требуется решить большое количество уравнений с одной и той же матрицей, более того, такие вычисления будут более устойчивы, то есть их можно использовать для плохо обусловленных матриц.

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

Ответ:

Прямой ход метода Гаусса находит верхнюю треугольную матрицу U для некоторой квадратной матрицы A. Следовательно, если использовать алгоритм над транспонированной матрицей A^{T}, то он найдет нижнюю треугольную матрицу L для матрицы A. Таким образом, для полноценной реализации метода Гаусса достаточно дважды запустить процедуру прямого хода: сначала для матрицы A, затем для матрицы U^{T}. При таком подходе в разы упрощается реализация метода Гаусса, при этом происходит повторный проход по нулевым элементам матрицы U^{T} и возникает необходимость в транспонировании U, что может существенно увеличить время исполнения алгоритма .

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

Ответ:

Кубическая норма $\|x\|_\infty = \max\limits_k |x_k|$ называется так потому, что множество $\|x\|_\infty \leq 1$ представляет собой куб со стороной длиной 2. Аналогично, для октаэдрической нормы $\|x\|_1 = \sum\limits_{j=1}^{n} |x_i|$ множество $\|x\|_1 \leq 1$ представляет собой октаэдр, а для евклидовой (шаровой) нормы $\|x\|_2 = \left(\sum\limits_{i=1}^{n} x_i^2\right)^{1/2}$ множество $\|x\|_2 \leq 1$ – шар радиусом 1 в декартовых координатах.

Дополнительные вопросы

1. **Посчитать** $\det A$ **при** $n = 4$ **через миноры.**

   Ответ:

   Пусть дана матрица A:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

   Вычислить определитель этой матрицы можно вычислить с помощью разложения по строке или столбцу. Запишем разложение по первой строке:

$$\det A = a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + \\ + a_{13} \cdot \begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14} \cdot \begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}.$$

2. **Расписать вопрос №3 более подробно: привести пример (небольшой порядок).**

   Ответ:

   Пусть дана система уравнений (1):

$$\begin{cases} x_1 + 5x_2 + 3x_3 = 8; \\ 3x_1 + x_2 + 8x_3 = 3; \\ 6x_1 + 2x_2 + x_3 = 16. \end{cases} \tag{4}$$

   Матрицы A и b соответственно:

$$A = \begin{pmatrix} 1 & 5 & 3 \\ 3 & 1 & 8 \\ 6 & 2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ 3 \\ 16 \end{pmatrix}.$$

   Создадим два массива перестановок, для строк и для столбцов:

$$\text{rows}_{str} = \big(1,\, 2,\, 3\big), \quad \text{rows}_{col} = \big(1,\, 2,\, 3\big).$$

Главным элементом для позиции (1, 1) будет элемент A_{23}. Сделаем соответствующую перестановку строк и столбцов:

$$A^{new} = \begin{pmatrix} 8 & 1 & 3 \\ 3 & 5 & 1 \\ 1 & 2 & 6 \end{pmatrix}, \quad b^{new} = \begin{pmatrix} 3 \\ 8 \\ 16 \end{pmatrix},$$

массивы перестановок при этом имеют вид:

$$rows_{str} = \big[3,\, 2,\, 1\big], \quad rows_{col} = \big[2,\, 1,\, 3\big].$$

При окончании работы алгоритма с помощью данных массивов можно восстановить исходный порядок переменных и строк в матрице A.

3. Число обусловленности cond$A = 100$. Хорошо или плохо?

Ответ:

Число обусловленности cond$A$ матрицы A связано с оценкой того, как ошибка в входных данных δb влияет на ошибку в решении δx системы уравнений $Ax = b$:

$$\frac{\|\delta x\|}{\|x\|} \le \mathrm{cond} A \cdot \frac{\|\delta b\|}{\|b\|}.$$

Скажем, если число обусловленности равна 100, то и ошибка может быть увеличена в 100 раз, и в различных задачах это будет иметь разную степень приемлемости, которая, в частности, будет зависеть от ошибки входных данных b. Например, если ошибка составляет 0.001, т.е. $\frac{\|\delta b\|}{\|b\|} = 0.001$, то верхняя оценка для ошибки в решении будет:

$$\frac{\|\delta x\|}{\|x\|} \le 100 \cdot 0.001 = 0.1.$$

4. Посчитать число операций в методе Гаусса и сравнить с числом операций QR-алгоритма.

Ответ:

Деление вектора b на главный элемент строки – n операций, деление элементов строки на главный элемент строки – $\frac{n(n-1)}{2}$ операций, вычитание из вектора b одной из предыдущих строк, умноженной на главный элемент текущей строки – $\frac{n(n-1)}{2}$ операций, вычисление остальных коэффициентов матрицы – сумма квадратов от 1 до $n-1$, $\frac{1}{6}(n-1)\,n\,(2n-1)$ операций, обратный ход – $\frac{n(n-1)}{2}$ операций. Итого:

$$n + 3\frac{n(n-1)}{2} + \frac{1}{6}(n-1)\,n\,(2n-1) = \frac{n}{3}\left(n^2 + 3n - 1\right) \sim \frac{n^3}{3}.$$

Таким образом, метод Гаусса в среднем работает приблизительно в 6 раз быстрее, чем метод QR-разложения.

5. **Количество операций в методе Гаусса с полным выбором элемента.**

Ответ:

К прямому и обратному ходу в классическом методе Гаусса добавляется поиск максимального элемента и перестановка строк и столбцов на каждой итерации.

Поиск максимального элемента: на каждом шаге поиск происходит по $(n-k+1)\times(n-k+1)$ элементам, $k=\overline{1,\, n-1}$. Количество операций – сумма квадратов от 1 до $n-1$, равная $\frac{1}{6}\,(n-1)\,n\,(2n-1)$.

Перестановка строк и столбцов - $O(n)$ операций на каждом шаге, то есть за $n-1$ шагов потребуется приблизительно n^2 операций.

При добавлении данных операций к классическому методу Гаусса, общее количество операций оценивается как $\frac{2n^3}{3}$, что есть в два раза больше, чем в классическом методе.

6. **Нарисовать картинки к 10 вопросу, определение эквивалентных норм.**

Ответ:

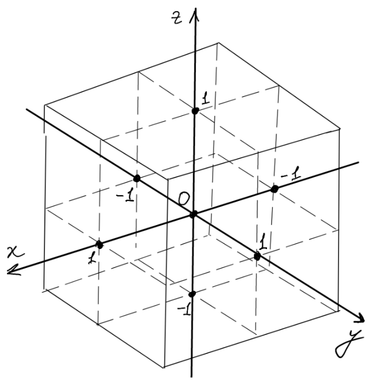

Рис. 1. Множество $\|x\|_\infty \leq 1$

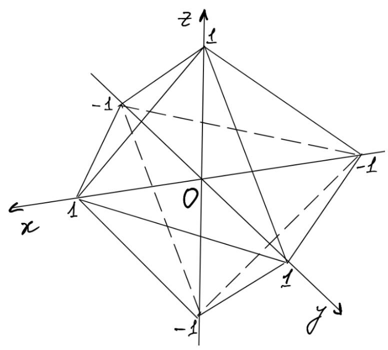

Рис. 2. Множество $\|x\|_1 \leq 1$

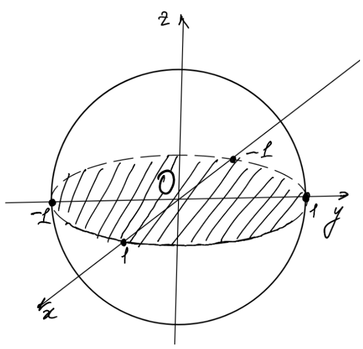

Рис. 3. Множество $\|x\|_2 \leq 1$

Определение. Пусть $\|\cdot\|_a$ и $\|\cdot\|_b$ – нормы на векторном пространстве V, тогда они называются **эквивалентными**, если $\exists$ положительные постоянные $C_1 > 0$ и $C_2 > 0$ такие, что для всех векторов $v \in V$ выполнены неравенства:

$$C_1\|v\|_a \leq \|v\|_b \leq C_2\|v\|_a.$$

7. Привести пример системы, решить которую получится только методом Гаусса с выбором главного элемента:

$$A = \begin{pmatrix} 3 & 6 & 9 \\ 2 & 6 & 7 \\ 1 & 1 & 1 \end{pmatrix}$$

8. (a) Приведите пример матрицы, у которой число обусловленности велико, а определитель мал:

$$A = \begin{pmatrix} 853.55 & -500 & -146.45 \\ 0.14 & 0.5 & -0.85 \\ 0.0005 & 0.0007 & 0.0005 \end{pmatrix},$$

$$\text{cond}A \sim 1.06 \cdot 10^6, \ \det A \sim 0.99$$

(b) число обусловленности мало, а определитель велик:

$$A = \begin{pmatrix} 0 & 100 & 0 \\ 34 & 34 & 34 \\ 99 & 0 & 0 \end{pmatrix}$$

$$\text{cond}A \sim 3.68, \ \det A = 336600$$