

Беляев В.Я.

Введение в численные методы

электронное издание

*методическое пособие
для студентов изучающих прикладную математику*



Мурманск
2025

УДК 519.6

ББК В192

Беляев, В.Я. Введение в численные методы – Мурманск : электронное издание, 2025. – 107 с.

В учебном пособии рассматриваются наиболее важные с точки зрения автора разделы дисциплины “Методы вычислений”. Среди них: решение нелинейных уравнений, решение систем линейных уравнений, вычисление обратных матриц, нахождение собственных чисел матриц, численное интегрирование и некоторые другие. Особое внимание уделяется разъяснению сути применяемых методов. Автор делает попытку если не доказать, то хотя бы показать действие глубоких математических законов, лежащих в основе приводимых алгоритмов.

Составитель – Владимир Яковлевич Беляев, канд. физико-математических наук, доцент кафедры Высшей математики и физики
Мурманского Арктического университета

Печатается в авторской редакции
Компьютерная верстка

Оглавление

Введение	4
Тема 1. Абсолютная и относительная погрешности	6
Тема 2. Численное решение нелинейных уравнений	8
2.1. Метод деления отрезка пополам	8
2.2. Метод касательных Ньютона	9
2.3. Метод хорд.....	12
2.4. Метод хорд и касательных	14
Тема 3. Численное решение систем линейных уравнений.....	15
3.1. Метод Гаусса	16
3.2. Метод Гаусса-Жордана с выбором главного элемента.....	20
3.3. Метод вращений.....	22
3.4. Метод простой итерации	29
3.5. Метод наискорейшего спуска	34
3.6. Градиентный метод	38
Тема 4. Вычисление обратной матрицы	40
4.1. Вычисление обратной матрицы линейными преобразованиями (Жордана)	40
4.2. Вычисление обратной матрицы методом окаймления	41
Тема 5. Метод наименьших квадратов и первая трансформация Гаусса	44
5.1. Метод наименьших квадратов	44
5.2. Первая трансформация Гаусса.....	45
5.3. Пример	46
Тема 6. Вычисление собственных чисел матриц	47
6.0. Собственные числа матриц.....	47
6.1. Собственные числа симметрических матриц.....	49
6.2. Число обусловленности матрицы	51
6.3. Собственное число неотрицательной матрицы	55
6.4. QR-алгоритм	58
6.5. Метод Гивенса	62
Тема 7. Задача одномерной безусловной оптимизации	66
7.1. Метод половинного деления	67
7.2. Метод золотого сечения	69
7.3. Метод Ньютона-Рафсона	72
Тема 8. Задача многомерной безусловной оптимизации	73
8.1. Метод сопряженных направлений.....	73
8.2. Алгоритм метода сопряженных направлений.....	77
Тема 9. Численное интегрирование	81
9.1. Формула трапеций	81
9.2. Формула Симпсона	83
9.3. Правило Рунге	84
9.4. Формула Гаусса	87
Тема 10. Численное решение дифференциальных уравнений.....	90
10.1. Формулы Эйлера (метод Эйлера).....	90
10.2. Формулы Рунге-Кутты 2-го, 3-го, 4-го порядков	91
10.3. Явные схемы (методы) Адамса 2-го, 3-го, 4-го, 5-го порядков	93
10.4. Неявные схемы (методы) Адамса	94
10.5. Сравнение методов на примере	94
ПРИЛОЖЕНИЕ	98
Лабораторная 1. Решение нелинейных уравнений	98
Лабораторная 2. Решение систем линейных уравнений прямыми методами.....	98
Лабораторная 3. Решение систем линейных уравнений итерационными методами.....	99
Лабораторная 4. Вычисление обратной матрицы	100

Лабораторная 5. Метод наименьших квадратов и первая трансформация Гаусса	100
Лабораторная 6. Собственные числа симметрической и неотрицательной матриц.....	100
Лабораторная 7. QR-алгоритм поиска собственных чисел матрицы	104
Лабораторная 8. Задача одномерной безусловной оптимизации	105
Лабораторная 9. Задача многомерной безусловной оптимизации	105
Лабораторная 10. Вычисление интегралов	106
Лабораторная 11. Численное решение обыкновенных дифференциальных уравнений	107
ЛИТЕРАТУРА	107

Введение

Численные методы или методы вычислений – одна из обширных областей прикладной математики, по которой существует весьма богатая литература. Источники [2 – 8], приводимые в конце пособия, составляют только малую ее часть. Настоящее пособие задумано, как небольшое введение в численные методы и рассчитано на студентов провинциальных вузов. Мой скромный опыт преподавания этой дисциплины показывает, что для более, чем 90% студентов читать серьезные учебники, написанные ведущими учеными нашей страны и мира – не намного понятнее, чем читать роман “Война и мир” на китайском языке. И эта ситуация усугубляется с каждым годом. Конечно, сейчас появляется много учебных пособий, которые, несмотря на низкую математическую культуру, помогают справиться с трудностями изучения предмета. Но они сильно зависят от вкусов преподавателя и ориентированы на студентов какого-то конкретного вуза. Примерно такого же плана и настоящее пособие. Его темы – это часть разделов из руководства [1] В.И. Середы, благодаря которому мне посчастливилось окунуться в мир вычислительной математики. Я выбрал и переработал на свой вкус примерно половину этого материала - пропорционально произшедшему с той поры сокращению часов на изучение как численных методов, так и всей математики.

Цель пособия – по возможности повысить процент обучающихся, способных с пониманием прочесть его материал. Но мне часто приходится опираться на понятия и факты из алгебры и математического анализа, которые составляют основу численных методов и понимания которых студентам хронически не хватает. Так что человек, представляющий, что такое предел, непрерывность, обратная матрица, полином и т.д., функция, наконец, имеет шанс по прочтении пособия повысить свои общематематические знания. Всем прочим останется только более-менее бессмысленное умение что-то считать и как-то “спихнуть” экзамен по предмету.

К особенностям настоящего пособия я бы отнес примеры вычислений. Они всегда “не детские”. Оперируемые числа содержат, как правило, от 4-х до 8-ми знаков после запятой. С помощью бумаги и ручки с ними работать просто невозможно. Но в настоящее время каждый человек имеет в своем распоряжении массу мощных вычислительных средств. Большая

часть примеров пособия вычислялась в MS Excel. Но - далеко не всегда. Для того, чтобы повторить вычисления с матрицами надо или программировать в Excel, либо использовать специальные матричные калькуляторы. А для поиска собственных чисел с помощью QR-алгоритма и этого не хватит. Для этих целей автор использовал собственные программы. Их особенность не в том, что они автоматически вычисляют решение той или иной задачи, а в том, что в “полуавтоматическом” режиме позволяют студенту производить шаги алгоритма, избавляя его от рутинных и громоздких вычислений. Ссылку для их скачивания я готов предоставить в ответ на содержательный отзыв или на полезные замечания по содержанию пособия. Наконец, варианты заданий в приложении представлены весьма скромно – их только 5 по каждой лабораторной. По этой ссылке можно будет получить возможность для скачивания нескольких десятков вариантов.

Июнь, 2025.
Беляев Владимир, (bvy1@mail.ru)

Тема 1. Абсолютная и относительная погрешности

Целью алгоритмов, рассматриваемых в данном пособии, является нахождение числовых значений тех или иных интересующих нас величин. Точное их значение остается, как правило, неизвестным. Но некоторой связью между этим точным значением и результатом вычисления является любое число, оценивающее их разницу. Пусть a – некоторая числовая величина и a^* – ее примерное значение. *Абсолютной погрешностью* можно назвать любое число Δa , для которого

$$|a - a^*| \leq \Delta a.$$

Это неравенство, как известно, равносильно тому, что a находится в интервале $(a^* - \Delta a, a^* + \Delta a)$. Поэтому чем меньше Δa , тем более точную информацию величина a^* дает о неизвестной a . Все численные методы в идеале должны давать не только примерное значение, но и погрешность. Далее абсолютной погрешностью мы будем называть величину $\Delta a = |a - a^*|$. Это неточное определение всего лишь означает, что наши методы дают значение a с точностью до Δa , и ничего лучше в данный момент мы придумать не можем.

Абсолютная погрешность в общем случае является результатом действия нескольких факторов. В частности, она складывается из погрешности, которую дает метод и из погрешности, которую дают вычислительные ошибки округления. Учет ошибок округления имел особенно важное значение в те времена, когда вычисления проводились вручную или с помощью простейших вычислительных средств таких, как счеты, логарифмические линейки и арифмометры. Для высокотехнологических областей это остается актуальным и сейчас, не смотря на то, что компьютерные вычисления способны в большинстве случаев дать заведомо достаточную точность. Поэтому в этом параграфе мы даем лишь некоторые способы, позволяющие получить информацию о погрешностях, связанных с ошибками округлений. Для абсолютной погрешности это соотношения:

1. $\Delta(a + b) = \Delta a + \Delta b;$
2. $\Delta(a \cdot b) = \Delta a \cdot b^* + \Delta b \cdot a^*;$
3. $\Delta\left(\frac{a}{b}\right) = \frac{\Delta a \cdot b^* + \Delta b \cdot a^*}{(b^*)^2}$

Обоснованием этих правил являются такие весьма примерные выкладки. Т.к. $|a-a^*| \leq \Delta a$ и $|b-b^*| \leq \Delta b$, то

1. $|a+b-(a+b)^*| = |a+b-(a^*+b^*)| = |a-a^*+b-b^*| \leq |a-a^*| + |b-b^*| = \Delta a + \Delta b.$
2. $|ab-(ab)^*| = |ab-a^*b^*| = |ab-a^*b+a^*b-a^*b^*| = |(ab-a^*b)+(a^*b-a^*b^*)| = |(a-a^*)b+a^*(b-b^*)| \leq |a-a^*| \cdot |b| + |a^*| \cdot |b-b^*| \approx \Delta a \cdot |b^*| + |a^*| \cdot \Delta b.$
3. $\Delta\left(\frac{a}{b}\right) = \left|\frac{a}{b} - \frac{a^*}{b^*}\right| = \left|\frac{ab^*-a^*b}{bb^*}\right| = \left|\frac{ab^*-a^*b^*+a^*b^*-a^*b}{bb^*}\right| \approx \frac{\Delta ab^* + a^* \Delta b}{b^* b^*}.$

Для величины a и ее примерного значения a^* **относительной погрешностью** называется отношение

$$\delta(a) = \frac{\Delta a}{a^*}.$$

Относительная погрешность в некотором смысле более интересна и информативна, чем абсолютная. Пусть, например, $a = 299792458$. По современным оценкам это скорость света в м/сек. Но обычно ее принимают равной 300 000 км/сек, т.е. принимают $a^* = 300000000$. Тогда за Δa можно взять $207542 \approx 210000$. При этом $\delta(a) = \frac{21}{30000} = 0.0007$. Таким образом скорость света равна 300 000 км/сек с относительной погрешностью 0.0007. Но что это значит? При компьютерном представлении очень больших или очень маленьких чисел используется их запись в виде произведения мантиссы на степень числа 10. Так величина $a = 2.99792458 \cdot 10^8$. Тогда и $\Delta a = 0.0021 \cdot 10^8$, а $\delta(a) = \frac{0.0021}{3.0}$. Т.е. относительная погрешность не зависит от степени числа 10. Она характеризует мантиссу величины a , указывая именно ее точность. Можно сказать, что чем меньше относительная погрешность, тем больше значащих цифр мантиссы нам известно. Например, результат измерения скорости света французским физиком Леоном Фуко в 1862 году дал результат $a^* = 298000000$ м/сек. Тогда получится $\delta(a) = \frac{0.002}{2.98} = 0.00067$. К тому же относительная погрешность не зависит от единиц, в которых делается измерение или расчет.

Для относительной погрешности важными и несколько неожиданными являются следующие соотношения:

1. $\delta(a \cdot b) = \delta(a) + \delta(b);$
2. $\delta\left(\frac{a}{b}\right) = \delta(a) + \delta(b).$

Действительно, $\delta(a \cdot b) = \frac{\Delta(a \cdot b)}{a^* b^*} = \frac{|a \cdot b - a^* b^*|}{a^* b^*} = \frac{|a \cdot b - a^* b + a^* b - a^* b^*|}{a^* b^*} = \frac{|(a-a^*)b + a^*(b-b^*)|}{a^* b^*} \leq \frac{|(a-a^*)b|}{a^* b^*} + \frac{|a^*(b-b^*)|}{a^* b^*} \approx \frac{|(a-a^*)|}{a^*} + \frac{|(b-b^*)|}{b^*} = \delta(a) + \delta(b).$

$$\delta\left(\frac{a}{b}\right) = \frac{\Delta\left(\frac{a}{b}\right)}{\left(\frac{a^*}{b^*}\right)} = \frac{\left|\frac{a}{b} - \frac{a^*}{b^*}\right|}{\left(\frac{a^*}{b^*}\right)} = \frac{\left|\frac{ab^* - ba^*}{bb^*}\right|}{\left(\frac{a^*}{b^*}\right)} = \frac{\left|ab^* - ba^*\right|}{ba^*} = \frac{\left|ab^* - a^*b^* + a^*b^* - ba^*\right|}{ba^*} \leq \frac{\left|ab^* - a^*b^*\right|}{ba^*} +$$

$$\frac{\left|a^*b^* - ba^*\right|}{ba^*} \approx \frac{\left|(a - a^*)\right|}{a^*} + \frac{\left|(b - b^*)\right|}{b^*} = \delta(a) + \delta(b).$$

Тема 2. Численное решение нелинейных уравнений

В данном разделе речь идет о решении уравнений вида

$$f(x) = 0 \quad (2.1)$$

где $y=f(x)$ – числовая функция. При этом с точки зрения численных методов интересным является случай, когда уравнение (2.1) не может быть решено простыми “школьными методами”, в частности, не является линейным или квадратным. Однако предполагается, что функция $f(x)$ может быть вычислена в любой точке x своей области определения. В зависимости от применяемых методов к функции $f(x)$ могут предъявляться дополнительные требования. Будем, наконец, считать, что интересующие нас решения как-то локализованы, т.е. известен промежуток $[a,b]$, в котором ищется корень уравнения (2.1).

2.1. Метод деления отрезка пополам. Данна функция $f:[a,b] \rightarrow R$, которая является непрерывной на $[a,b]$ и для которой значения $f(a)$ и $f(b)$ – функции f на концах отрезка $[a,b]$ – имеют разные знаки. Последнее означает, что $f(a) \cdot f(b) < 0$. Тогда по известной теореме Коши из курса математического анализа о промежуточном значении существует такой $x^* \in [a,b]$, что $f(x^*)=0$. Задача заключается в вычислении с заданной точностью хотя бы одного такого x^* .

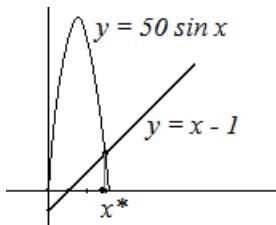
Алгоритм решения задачи заключается в выполнении последовательности шагов, где каждый шаг уменьшает длину промежутка $[a,b]$ ровно в два раза, сохраняя условие $f(a) \cdot f(b) < 0$.

Описание одного шага. Вычисляем $c = \frac{a+b}{2}$ и $f(c)$. Если окажется, что $f(c)=0$, то $x^* = c$. Если $f(a)$ и $f(c)$ разных знаков, то новые границы: $a:=a$, $b:=c$. Если $f(c)$ и $f(b)$ разных знаков, то новые границы: $a:=c$, $b:=b$.

Допустим, что требуется найти решение x^* уравнения (2.1) с абсолютной погрешностью $\varepsilon > 0$, т.е. с точностью до $\pm\varepsilon$. Тогда достаточно вы-

полнять описанный шаг вычисления до тех пор, пока не окажется, что $b-a \leq 2\varepsilon$. После этого можно взять $x^* = \frac{a+b}{2}$. Если, к примеру, начальная длина промежутка $[a,b]$ равна 100, а $\varepsilon=0.0001$, то понадобится всего 19 шагов.

Пример. Найдем с точностью до $\varepsilon = 0.0001$ наименьший положительный корень уравнения $x-1=50\cdot\sin x$. По схематическому рисунку



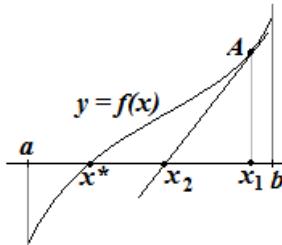
можно определить, что искомый корень находится в промежутке $[0,4]$. Приводим уравнение к виду (1). Тогда $f(x) = x - 1 - 50\cdot\sin x$. $f(0) = -1$, $f(4) = 3 - 50\cdot\sin 4 \approx 3+50\cdot0.76 = 41$. Итак, перед первым шагом $a=0$, $b=4$. Дальнейшие вычисления делаем в MS Excel и оформляем их в виде таблицы

Шаг	a	b	c	$f(a)$	$f(b)$	$f(c)$	$b-a-0.0002$
0	0	4	2	-1	40,84012	-44,4649	3,999800
1	2	4	3	-44,4649	40,84012	-5,056	1,999800
2	3	4	3,5	-5,056	40,84012	20,03916	0,999800
3	3	3,5	3,25	-5,056	20,03916	7,659757	0,499800
4	3	3,25	3,125	-5,056	7,659757	1,295405	0,249800
5	3	3,125	3,0625	-5,056	1,295405	-1,88801	0,124800
6	3,0625	3,125	3,09375	-1,88801	1,295405	-0,29747	0,062300
7	3,09375	3,125	3,109375	-0,29747	1,295405	0,498771	0,031050
8	3,09375	3,109375	3,101563	-0,29747	0,498771	0,100589	0,015425
9	3,09375	3,101563	3,097656	-0,29747	0,100589	-0,09846	0,007613
10	3,097656	3,101563	3,099609	-0,09846	0,100589	0,001062	0,003706
11	3,097656	3,099609	3,098633	-0,09846	0,001062	-0,0487	0,001753
12	3,098633	3,099609	3,099121	-0,0487	0,001062	-0,02382	0,000777
13	3,099121	3,099609	3,099365	-0,02382	0,001062	-0,01138	0,000288
14	3,099365	3,099609	3,099487	-0,01138	0,001062	-0,00516	0,000044
15	3,099487	3,099609	3,099548	-0,00516	0,001062	-0,00205	-0,000078

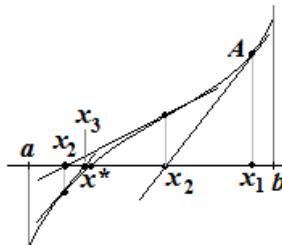
После шага 15 длина промежутка становится меньше 2ε . Поэтому окончательно $x^* = \frac{a+b}{2} = \frac{3,099487+3,099609}{2} = 3,099548$.

2.2. Метод касательных Ньютона. Данна функция $f:[a,b] \rightarrow R$, которая является непрерывной на $[a,b]$ и, кроме того, во всех точках интервала

(a,b) имеет производную, которую можно вычислять. Допустим, что уравнение (2.1) имеет внутри промежутка корни и нам надо найти какой-нибудь из них. Геометрический смысл и идея метода касательных заключается в следующем рассуждении. Предположим, что на промежутке (a,b) функция ведет себя “прилично”, т.е. ее график пересекает ось иксов в только в одной точке x^* и при этом не делает резких перегибов и колебаний (как на рисунке).



Выберем в (a,b) произвольную точку x_1 , возьмем соответствующую ей точку A на графике функции и проведем касательную к графику в этой точке A . Касательная пересечет от OX в некоторой точке x_2 . И скорее всего эта x_2 окажется к неизвестной, но нужной нам x^* ближе, чем x_1 . А что будет, если эту процедуру повторить уже от точки x_2 и сделать это несколько раз? Следующий рисунок иллюстрирует этот процесс.



Получается последовательность точек x_1, x_2, x_3, \dots , которая очень быстро приближается к x^* .

Найти аналитическую формулу, дающую x_2 по x_1 очень просто. Точка A имеет координаты $(x_1, f(x_1))$. Угловой коэффициент касательной в точке A равен $f'(x_1)$ (производной функции f в точке x_1). Отсюда уравнение касательной есть $y = f'(x_1)(x - x_1) + f(x_1)$. Точка x_2 – это тот x , при котором $y=0$. Следовательно $f'(x_1)(x_2 - x_1) + f(x_1) = 0$ и $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$. В общем случае:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n=1,2,\dots) \quad (2.2)$$

В качестве примера возьмем уравнение их предыдущего метода, где $f(x) = x - 1 - 50 \cdot \sin x$. Тогда $f'(x) = 1 - 50 \cdot \cos x$ и формула (2.2) дает

$$x_{n+1} = x_n - \frac{x - 1 - 50 \cdot \sin x}{1 - 50 \cdot \cos x} \quad (n=1,2,\dots)$$

Возьмем $x_0 = 4$ и составим в MS Excel таблицу

n	x
1	4
2	2,787485742
3	3,112138371
4	3,09959113
5	3,099588532
6	3,099588532

Видим, что уже на 6-й итерации получается предельно точное значение корня.

Естественно, что при использовании MS Excel для вычисления x при $n=2$ в соответствующую ячейку заносится полученная выше формула, которая автоматически растягивается на ячейки для $n=2,3,4,5,\dots$ (см. PrnScr ниже).

	B3	C	D	E	F	G
	A	B				
1	n	x				
2	1	4				
3	2	2,787486				
4	3	3,112138				
5	4	3,099591				
6	5	3,099589				
7	6	3,099589				

Главным недостатком метода касательных является то обстоятельство, что его итерации не всегда приводят к решению уравнения или приводят не к тому решению, которое мы ожидаем. К тому же в отличие от метода деления пополам он требует существования первой производной. Поэтому важно понимать, при каких условиях можно гарантировать сходимость метода касательных. Здесь стоит подчеркнуть, что, во-первых, начальное значение x_1 должно быть взято по возможности ближе к исскомому корню и, во-вторых, располагаться слева или справа от него в зависимости от поведения функции вблизи этого корня. Соответствующие случаи иллюстрируются на рисунках:

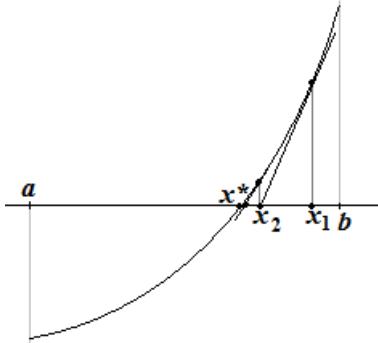


Рис. 2.1 Возрастание, выпуклость вниз

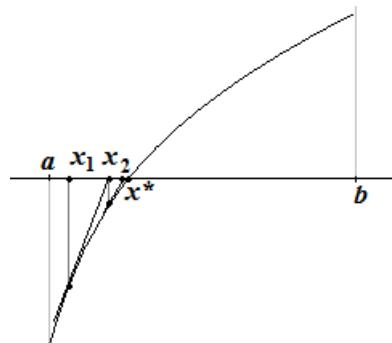


Рис. 2.2 Возрастание, выпуклость вверх

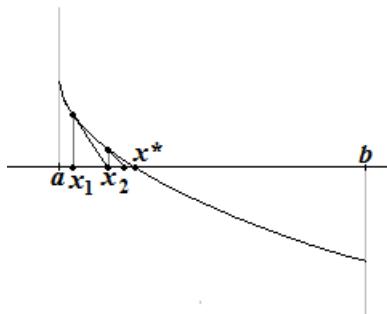


Рис. 2.3 Убывание, выпуклость вниз

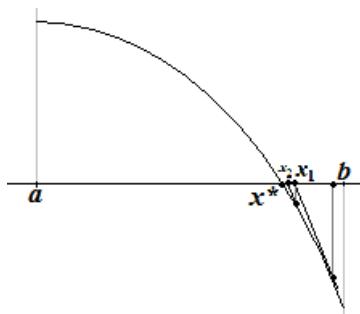


Рис. 2.4 Убывание, выпуклость вверх

Видим, что в случаях на рис. 2.1 и 2.4 точку x_1 следует выбирать справа от x^* , а в случаях 2.2 и 2.4 – слева от x^* . Определить, какой из этих случаев имеет место, не имея возможности достаточно точно построить график, можно с помощью знака первой и второй производной функции $f(x)$. Именно, если $f'(x)>0$ на (a,b) , то $f(x)$ возрастает, если $f'(x)<0$ на (a,b) , то $f(x)$ убывает. Если $f''(x)>0$ на (a,b) , то график выпуклый вниз ($f(x)$ “ускоряется”), если $f''(x)<0$ на (a,b) , то график выпуклый вверх ($f(x)$ “замедляется”).

Например, для функции $f(x) = x - 1 - 50 \cdot \sin x$ из рассмотренного раньше примера производная $f'(x) = 1 - 50 \cdot \cos x$ вблизи точки x^* положительна, и вторая производная $f''(x) = 50 \cdot \sin x$ также положительна (пока $x < \pi \approx 3.14$). Поэтому на промежутке $(3, 3.14)$ имеет место случай на рис. 2.1.

2.3. Метод хорд. Данна функция $f : [a,b] \rightarrow R$, которая является непрерывной на $[a,b]$ и для которой значения $f(a)$ и $f(b)$ – функции f на концах отрезка $[a,b]$ имеют разные знаки (как и в случае 1). Метод хорд чем-то похож на метод деления пополам, но вместо того, чтобы в качестве точки c брать середину отрезка (a,b) берут точку пересечения оси OX с отрезком,

соединяющим концы графика $y=f(x)$, т.е. точки $(a, f(a))$ и $(b, f(b))$). Этот отрезок и называется *хордой*.

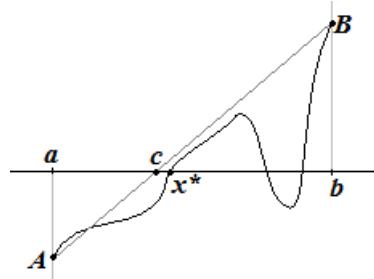


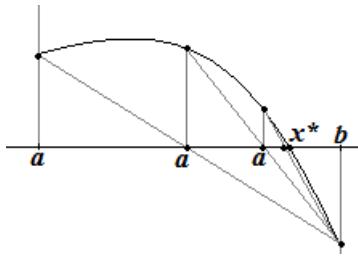
Рис. 2.5 Метод хорд

Прямая, проходящая через концы A и B графика, имеет, очевидно, уравнение $y = \frac{f(b)-f(a)}{b-a}(x - a) + f(a)$. Точка c — это тот x , при котором $y=0$. Отсюда выводим $c = a - \frac{f(a)(b-a)}{f(b)-f(a)} = b - \frac{f(b)(b-a)}{f(b)-f(a)}$. Теперь надо вычислить $f(c)$ и, в зависимости от знака полученной величины переопределить отрезок $[a,b]$, взяв в его качестве или $[a,c]$ или $[c,b]$. Например, в ситуации на рис.5 надо взять $a:=c$, $b:=b$. В этом заключается один шаг метода. Указанный шаг надо повторять до тех пор, пока величина $f(c)$ не превратится в машинный ноль или просто не перестанет уменьшаться.

Таблица вычислений для нашего случая:

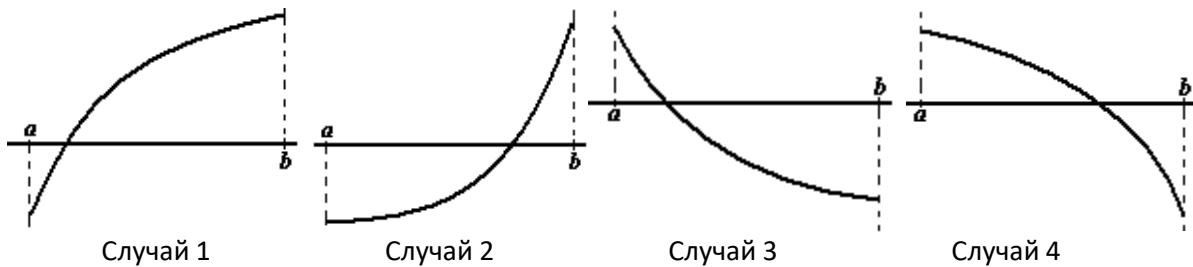
Шаг	a	b	c	$f(a)$	$f(b)$	$f(c)$
0	0	4	0,095602	-1	40,84012	-5,67722
1	0,095602	4	0,572115	-5,67722	40,84012	-27,4985
2	0,572115	4	1,951447	-27,4985	40,84012	-45,4697
3	1,951447	4	3,030665	-45,4697	40,84012	-3,50436
4	3,030665	4	3,107267	-3,50436	40,84012	0,391339
5	3,030665	3,107267	3,099572	-3,50436	0,391339	-0,00083
6	3,099572	3,107267	3,099589	-0,00083	0,391339	-1,2E-07
7	3,099589	3,107267	3,099589	-1,2E-07	0,391339	-1,8E-11
8	3,099589	3,107267	3,099589	-1,8E-11	0,391339	0

Метод хорд не выглядит слишком элегантным. Один из его недостатков заключается в том, что если на рассматриваемом промежутке функция не имеет точек перегиба, т.е. выпуклость всегда вверх или всегда вниз, то будет меняться только один из концов промежутка — или a или b . Этот конец будет приближаться к x^* . Иллюстрацией метода является рисунок:



Ценность метода хорд заключается в возможности его объединения с методом касательных, что дает действительно красивый и очень эффективный метод.

2.4. Метод хорд и касательных. Данна функция $f:[a,b] \rightarrow R$, которая непрерывна на $[a,b]$, дифференцируема на (a,b) , у которой значения $f(a)$ и $f(b)$ имеют разные знаки и выпуклость графика на (a,b) всегда направлена вниз или всегда - вверх. Тогда возможны следующие 4 случая.



Как мы уже знаем, в случаях 1 и 3 метод касательных надо применять, стартуя из точки левее решения, а для случаев 2 и 4 – правее решения. Поэтому для случаев 1 и 3 пересчитываем методом касательных точку a по формуле

$$a := a - \frac{f(a)}{f'(a)}$$

В этих же случаях 1 и 3 метод хорд обязательно дает точку справа от решения. Поэтому новое значение b пересчитываем по формуле метода хорд

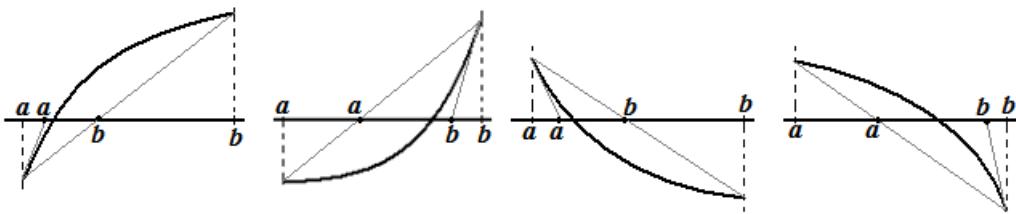
$$b := b - \frac{f(b)(b-a)}{f(b)-f(a)}$$

Аналогично в случаях 2 и 4

$$a := a - \frac{f(a)(b-a)}{f(b)-f(a)},$$

$$b := b - \frac{f(b)}{f'(b)}$$

Иллюстрация:



Метод хорд и касательных на каждом шаге гарантирует, что нужная нам точка лежит между числами a и b .

Для нашего примера на промежутке $[2,5, 3,14]$, который подпадает под случай 2, расчет дает

Шаг	a	b	$f(a)$	$f(b)$	$f'(b)$
	2,5	3,14	-28,4236	2,060367	50,99994
1	3,096743	3,099601	-0,14497	0,000614	50,95592
2	3,099589	3,099589	-3,7E-08	1,53E-10	50,9559
3	3,099589	3,099589	0	0	50,9559

За 3 шага получается практически точный результат.

Тема 3. Численное решение систем линейных уравнений

В данном разделе речь идет о численном решении систем линейных уравнений вида

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \dots \dots \dots \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \quad (3.1)$$

Как видим, в этой системе имеется n уравнений и n неизвестных x_1, \dots, x_n . Кроме того предполагается, что система является крамеровской, т.е. имеющей единственное решение. Это равносильно тому, что определитель матрицы

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix},$$

которая называется **матрицей системы**, не равен нулю. Мы рассматриваем только случай, когда все a_{ij} и b_j являются вещественными числами. a_{ij} – это коэффициенты при неизвестных, b_j – свободные члены. Разумеется, известные формулы Крамера – это последнее, что стоит использовать для решения этой задачи.

Матрица

$$\mathbf{B} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ \dots & \dots & \dots & \dots & | & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & | & b_n \end{pmatrix},$$

называется *расширенной матрицей системы* (3.1). Символически ее можно записать в виде

$$\mathbf{B} = (\mathbf{A}/\mathbf{b}),$$

где \mathbf{b} – столбец из свободных членов. Если еще обозначить $\mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, то

система (3.1) станет эквивалентна матричному уравнению

$$\mathbf{AX} = \mathbf{b} \quad (3.2)$$

Существует много способов решения систем (3.1) гораздо более быстрых и точных, чем формулы Крамера. В пособии приводится лишь малая часть этих методов. Но какой бы метод мы не применяли, при выполнении арифметических операций неизбежно происходят ошибки округления, которые в конечном итоге приводят к приближенному решению. Погрешность этого решения оценить весьма непросто. Поэтому при решении систем линейных уравнений принято оценивать эту точность по тому, насколько найденное решение подходит в уравнения системы. Если мы каким-то методом нашли решение $\mathbf{X}^* = (x_1^*, x_2^*, \dots, x_n^*)$, то можно подставить его в уравнения системы (3.1) и посмотреть, насколько точно совпадут левые и правые части равенств. Разница между ними – это n чисел

$$\begin{aligned} r_1 &= b_1 - a_{11}x_1^* - a_{12}x_2^* - \dots - a_{1n}x_n^* \\ r_2 &= b_2 - a_{11}x_1^* - a_{12}x_2^* - \dots - a_{1n}x_n^* \\ &\dots \\ r_n &= b_n - a_{n1}x_1^* - a_{n2}x_2^* - \dots - a_{nn}x_n^* \end{aligned} \quad (3.3)$$

которые составляют n -мерный вектор $\mathbf{r} = (r_1, \dots, r_n)$. Этот вектор называется *вектором невязки*. Понятно, что чем меньше числа r_1, \dots, r_n , тем ближе $x_1^*, x_2^*, \dots, x_n^*$ к точному решению системы (3.1). Этую близость можно охарактеризовать и одним числом, равным евклидовой норме вектора \mathbf{r} . Это число $\|\mathbf{r}\| = \sqrt{r_1^2 + r_2^2 + \dots + r_n^2}$ называется *невязкой* найденного решения. Для хорошего метода невязка должна быть близка к 0.

3.1. Метод Гаусса. По существу это обычный школьный метод исключения неизвестных. Но, т.к. все действия производятся с матрицей си-

стемы, никак не касаясь неизвестных, узнать в нем простые действия – выразить неизвестное из одного уравнения и подставить в остальные – может не всякий студент.

Итак – с расширенной матрицей системы можно делать следующие преобразования, которые называются *линейными преобразованиями строк*:

- 1) *Переставить местами любые две строки.*
- 2) *Любую строку умножить или разделить на любое ненулевое число.*

3) *К любой строке прибавить любую другую строку.*

Надо понимать, что каждое из этих преобразований не меняет решений системы (3.1). Т.е. система, внешне меняясь, остается эквивалентной первоначальной. Но это в теории. На практике всегда получается ненулевая невязка.

Кроме того к преобразованиям 1) – 3) стоит добавить как следствие преобразования:

- 4) *От любой строки можно отнять любую другую строку.*
- 5) *К любой строке можно прибавить любую другую строку с любым числовым коэффициентом.*
- 6) *От любой строки можно отнять любую другую строку с любым числовым коэффициентом.*

При использовании метода Гаусса делаются однотипные шаги, которые цепочкой линейных преобразований строк сначала приводят расширенную матрицу системы к треугольному виду:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ 0 & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} & b_n \end{array} \right)$$

у которого на главной диагонали, т.е. среди чисел $a_{11}, a_{22}, \dots, a_{nn}$ нет нулей, а под главной диагональю – только нули. Затем, опять с помощью линейных преобразований строк матрицу приводят к виду

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & b_1 \\ 0 & 1 & \dots & 0 & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & b_n \end{array} \right)$$

После этого свободные члены b_1, \dots, b_n будут искомыми решениями. На практике уже на первом этапе мы получим $a_{11} = a_{22} = \dots = a_{nn} = 1$. На втором этапе, который называется обратным ходом метода Гаусса, обнуляются числа над главной диагональю. Приведем пример для $n = 5$. Матрица системы имеет вид

$$\begin{array}{rrrrrr|r} 0.3032 & 8.2869 & -5.7604 & -9.9827 & 8.6236 & 136.1535 \\ 3.4133 & -6.0824 & 9.4853 & 9.4349 & 3.1861 & -139.6078 \\ -4.2811 & 6.9505 & 0.2575 & -7.5744 & 6.5649 & 91.3667 \\ -1.3291 & -5.3959 & -8.0335 & -9.6603 & -8.8668 & 152.3082 \\ -2.4302 & 4.2187 & 3.3620 & -2.2898 & -1.1729 & -29.2527 \end{array} \quad (3.4)$$

Здесь $a_{11} = 0.3032$. Делим всю первую строку на a_{11} . Получим

$$\begin{array}{rrrrrr|r} 1 & 27,33146 & -18,99868 & -32,92447 & 28,44195 & 449,05508 \\ 3.4133 & -6.0824 & 9.4853 & 9.4349 & 3.1861 & -139.6078 \\ -4.2811 & 6.9505 & 0.2575 & -7.5744 & 6.5649 & 91.3667 \\ -1.3291 & -5.3959 & -8.0335 & -9.6603 & -8.8668 & 152.3082 \\ -2.4302 & 4.2187 & 3.3620 & -2.2898 & -1.1729 & -29.2527 \end{array}$$

Теперь из второй строки вычитаем первую с коэффициентом $a_{21} = 3.4133$. Получим

$$\begin{array}{rrrrrr|r} 1 & 27,33146 & -18,99868 & -32,92447 & 28,44195 & 449,05508 \\ 0 & -99,37289 & 74,3335 & 121,816 & -93,89482 & -1672,3675 \\ -4.2811 & 6.9505 & 0.2575 & -7.5744 & 6.5649 & 91.3667 \\ -1.3291 & -5.3959 & -8.0335 & -9.6603 & -8.8668 & 152.3082 \\ -2.4302 & 4.2187 & 3.3620 & -2.2898 & -1.1729 & -29.2527 \end{array}$$

Тоже самое делаем с третьей строкой – вычитаем из нее первую с коэффициентом $a_{31} = -4.2811$. Затем – с четвертой и пятой. Получим

$$\begin{array}{rrrrrr|r} 1 & 27,33146 & -18,99868 & -32,92447 & 28,44195 & 449,05508 \\ 0 & -99,37289 & 74,3335 & 121,816 & -93,89482 & -1672,3675 \\ 0 & 123,95923 & -81,07775 & -148,52736 & 128,32774 & 2013,8164 \\ 0 & 30,93035 & -33,28465 & -53,42022 & 28,9354 & 749,14731 \\ 0 & 70,63962 & -42,80859 & -82,30285 & 67,94673 & 1062,04095 \end{array}$$

Проделанные преобразования – это тоже самое, что выражение из первого уравнения неизвестного x_1 через x_2, x_3, x_4, x_5 с последующей подстановкой этих выражений во 2-е, 3-е, 4-е и 5-е уравнения. После приведения подобных членов эти уравнения дадут систему из 4-х уравнений с 4-мя неизвестными x_2, x_3, x_4, x_5 . Расширенная матрица этой системы – это подматрица последней, полученной нами матрицы, образованная строками – со 2-й по 5-ю и столбцами – со 2-го по 6-й.

Теперь в этой системе надо также выразить из второго уравнения x_2 , подставить его в 3-е, 4-е и 5-е уравнения, исключив, тем самым, из них переменную x_2 . Реально, мы делим 2-ю строку матрицы на a_{22} . Получим

$$\begin{array}{rrrrrr|r} 1 & 27,33146 & -18,99868 & -32,92447 & 28,44195 & 449,05508 \\ 0 & 1 & -0,74803 & -1,22585 & 0,94487 & 16,82921 \\ 0 & 123,95923 & -81,07775 & -148,52736 & 128,32774 & 2013,8164 \\ 0 & 30,93035 & -33,28465 & -53,42022 & 28,9354 & 749,14731 \\ 0 & 70,63962 & -42,80859 & -82,30285 & 67,94673 & 1062,04095 \end{array}$$

Далее, из 3-го уравнения вычитаем 2-е с коэффициентом $a_{32} = 123,95923$, затем - из 4-го 2-е с коэффициентом $a_{42} = 30,93035$, затем - из 5-го 2-е с коэффициентом $a_{52} = 70,63962$. Получим

$$\begin{array}{rrrrrr|r} 1 & 27,33146 & -18,99868 & -32,92447 & 28,44195 & 449,05508 \\ 0 & 1 & -0,74803 & -1,22585 & 0,94487 & 16,82921 \\ 0 & 0 & 11,64697 & 3,42774 & 11,20194 & -72,31984 \\ 0 & 0 & -10,14795 & -15,50433 & -0,28987 & 228,61387 \\ 0 & 0 & 10,03168 & 4,29055 & 1,20122 & -126,76824 \end{array}$$

Проделав аналогично такую же процедуру со строками 3,4,5, затем – со строками 4,5, и, наконец, со строкой 5 (просто разделив ее на a_{55}) получим матрицу треугольного вида

$$\begin{array}{rrrrrr|r} 1 & 27,33146 & -18,99868 & -32,92447 & 28,44195 & 449,05508 \\ 0 & 1 & -0,74803 & -1,22585 & 0,94487 & 16,82921 \\ 0 & 0 & 1 & 0,2943 & 0,96179 & -6,20933 \\ 0 & 0 & 0 & 1 & -0,75655 & -13,22936 \\ 0 & 0 & 0 & 0 & 1 & 6,29139 \end{array}$$

Дальше делаем *обратный ход* метода Гаусса. Из матрицы видно, что $x_5 = 6,29139$. Подставив его в 4-е уравнение можно найти x_4 . Подставив x_4 , x_5 в 3-е уравнение можно найти x_3 . Затем из 2-го найти x_2 и из 1-го – x_1 .

Реально, надо опять делать линейные преобразования строк матрицы, но в таком порядке: вычитаем из 4-го уравнения 5-е с коэффициентом $a_{45} = -0,75655$, затем из 3-го 5-е с коэффициентом $a_{35} = 0,96179$, затем из 2-го 5-е с коэффициентом $a_{25} = 0,94487$, затем из 1-го 5-е с коэффициентом $a_{15} = 28,44195$. Получим

$$\begin{array}{rrrrrr|r} 1 & 27,33146 & -18,99868 & -32,92447 & 0 & 270,11568 \\ 0 & 1 & -0,74803 & -1,22585 & 0 & 10,88466 \\ 0 & 0 & 1 & 0,2943 & 0 & -12,26033 \\ 0 & 0 & 0 & 1 & 0 & -8,46961 \\ 0 & 0 & 0 & 0 & 1 & 6,29139 \end{array}$$

Также обнуляем выше диагонали элементы 4-го столбца, затем 3-го и, наконец, 2-го. Получим

$$\left| \begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right| \begin{array}{l} -8,34244 \\ -6,80436 \\ -9,76772 \\ -8,46961 \\ 6,29139 \end{array}$$

Значит,

$$x_1^* = -8,34244, x_2^* = -6,80436, x_3^* = -9,76772, x_4^* = -8,46961, x_5^* = 6,29139.$$

Заметим, что в данном примере вычисления намеренно делались не с самой максимальной точностью, на которую способно применяемое вычислительное средство. Каждый раз при получении нового числа это число округлялось до 5-ти знаков после запятой. В результате найденное решение – не самое лучшее, из того, что могло бы быть. Далее, вычисляем вектор невязки и невязку полученного решения. На этот раз используется более высокая точность вычислений – до 8-ми знаков после запятой. По формулам (3) получим вектор невязки

$$\mathbf{r} = (-0.00000215, -0.00470859, 0,006212, 0.00116162, 0.00358884).$$

и невязку $\|\mathbf{r}\| = 0.00865961$.

3.2. Метод Гаусса-Жордана с выбором главного элемента. Методом Гаусса-Жордана называют практически тот же метод Гаусса, при котором обратный ход делают сразу. Первый шаг точно такой же – получаем в первом столбце число 1 на первом месте, после чего обнуляем все остальные элементы первого столбца. Дальше, как и раньше делим второе уравнение на диагональный элемент a_{22} . Но потом обнуляем числа второго столбца не только под, но и над диагональю. Аналогично с третьим, четвертым и т.д. столбцами. В методе Гаусса или в методе Гаусса-Жордана числа, стоящие на диагонали сначала превращаются в 1, а затем используются как обнулители остальных чисел соответствующего столбца. Иногда такие элементы называют ведущими. Легко сообразить, что роль ведущих элементов совсем не обязательно должны выполнять числа на диагонали. Можно взять любое ненулевое число a_{ij} матрицы (в i -той строке, j -том столбце) и использовать его в качестве ведущего. Идея выбора главного элемента состоит в следующем. Возьмем в качестве первого ведущего элемента тот элемент a_{ij} матрицы системы, который является максимальным по абсолютной величине. Как и полагается, делим i -тую строку на a_{ij} . Это деление, как и все арифметические операции, приведет к некоторой ошибке округления. Ошибка тем больше, чем ближе к нулю наш делитель.

Выбирая его максимальным (по абсолютной величине) из возможных мы эту ошибку надеемся уменьшить (см. формулу 3. темы 1, оценивающую абсолютную погрешность отношения).

Опять рассмотрим наш пример с матрицей (3.1). Максимальным по абсолютной величине числом в матрице A является $a_{14}=-9.9827$. Делим первую строку на это число. Получим

$$\left| \begin{array}{ccccc|c} -0,03037 & -0,83013 & 0,57704 & 1 & -0,86385 & -13,639 \\ 3,69986 & 1,74976 & 4,041 & 0 & 11,33648 & -10,9257 \\ -4,51115 & 0,66279 & 4,62822 & 0 & 0,02172 & -11,9401 \\ -1,62251 & -13,4152 & -2,45914 & 0 & -17,2119 & 20,5519 \\ -2,49975 & 2,31788 & 4,6833 & 0 & -3,15095 & -60,4832 \end{array} \right|$$

Сделанное преобразование – это то же самое, что выразить из первого уравнения переменную x_4 и подставить – в уравнения 2,3,4,5. Делаем то же самое с системой из уравнений 2,3,4,5. Это значит, что, исключив в последней матрице первую строку и 4-й столбец, среди оставшихся коэффициентов при неизвестных мы отыскиваем максимальное по абсолютной величине число – это $a_{45} = -17,2119$. Делим на него 4-ю строку и привычными линейными преобразованиями строк обнуляем 5-й столбец, оставляя только единицу на 4-м месте. Получим

$$\left| \begin{array}{ccccc|c} 0,05106 & -0,15683 & 0,70046 & 1 & 0 & -14,6704 \\ 2,63121 & -7,08604 & 2,42131 & 0 & 0 & 2,61064 \\ -4,5132 & 0,64586 & 4,62512 & 0 & 0 & -11,9142 \\ 0,09427 & 0,77941 & 0,14287 & 0 & 1 & -1,19405 \\ -2,20272 & 4,77377 & 5,13349 & 0 & 0 & -64,2456 \end{array} \right|$$

Далее, в столбцах 1,2,3 с исключенными строками 1 и 4 находим максимальный по модулю элемент $a_{22}=-7,08604$. Делим на него вторую строку и обнуляем второй столбец. Получим

$$\left| \begin{array}{ccccc|c} -0,00717 & 0 & 0,64687 & 1 & 0 & -14,7282 \\ -0,37132 & 1 & -0,3417 & 0 & 0 & -0,36842 \\ -4,27338 & 0 & 4,84581 & 0 & 0 & -11,6763 \\ 0,38368 & 0 & 0,4092 & 0 & 1 & -0,9069 \\ -0,43011 & 0 & 6,76469 & 0 & 0 & -62,4868 \end{array} \right|$$

В оставшихся столбцах 1 и 3, исключая строки 1,2,4 находим максимальное по модулю $a_{53}=6,76469$. Делим на него строку 5 и обнуляем столбец 3. Получим

0,03396	0	0	1	0	-8,75294
-0,39305	1	0	0	0	-3,52477
-3,96528	0	0	0	0	33,08548
0,4097	0	0	0	1	2,87296
-0,06358	0	1	0	0	-9,2372

Наконец, в первом столбце берем единственный оставшийся вариант $a_{31} = -3,96528$. Делим на него строку 3 и обнуляем столбец 1. Получим

0	0	0	1	0	-8,46958
0	1	0	0	0	-6,8043
1	0	0	0	0	-8,34379
0	0	0	0	1	6,29141
0	0	1	0	0	-9,7677

Остается лишь переставить строки этой матрицы так, чтобы матрица системы стала единичной.

1	0	0	0	0	-8,34379
0	1	0	0	0	-6,8043
0	0	1	0	0	-9,7677
0	0	0	1	0	-8,46958
0	0	0	0	1	6,29141

Последний столбец из свободных членов и есть искомое решение. Как и раньше, считаем невязку, которая с точностью до 8-ми знаков после запятой на этот раз равна $\|r\| = 0,0004654$. Это заметно лучше невязки $\|r\| = 0,00865961$ решения, найденного методом Гаусса.

Справедливости ради надо заметить, что решение, полученное методом Гаусса-Жордана с выбором главного элемента, далеко не всегда оказывается лучше решения, найденного методом Гаусса.

3.3. Метод вращений. Прежде, чем объяснить метод вращения, уделим внимание матричным преобразованиям, которые могут дать решение системы, записанной в матричном виде $AX = b$. Это, прежде всего, формула $X = A^{-1} \cdot b$. Здесь A^{-1} – матрица, обратная к A . Указанное решение получается в результате умножения слева обеих сторон исходного матричного уравнения на матрицу A^{-1} . К сожалению, от этой формулы пользы немного. Беда в том, что вычисление обратной матрицы не проще, чем все остальные методы решения систем линейных уравнений. По существу метод

Гаусса или Гаусса-Жордана – это серия умножений слева столбца b на различные матрицы. Метод вращений похож на метод Гаусса. Но прямой ход для приведения к треугольному виду делается серией умножений слева на специальные матрицы, которые называются матрицами вращения.

Основную идею метода можно объяснить простыми вычислениями, которые проходят в школьной математике. Возьмем точку A на плоскости с координатами (a,b) . Ей соответствует вектор r , для которого эти числа являются его координатами (радиус-вектор точки A). Пусть φ – угол между осью OX и вектором r , r – расстояние от начала координат до точки A (длина или модуль вектора r) (см. рис. 3.1)

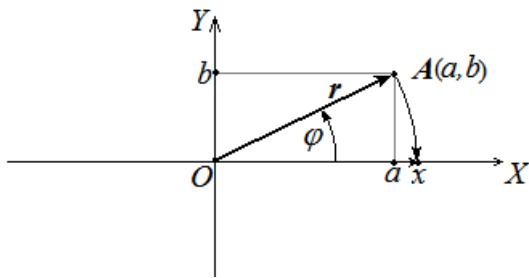
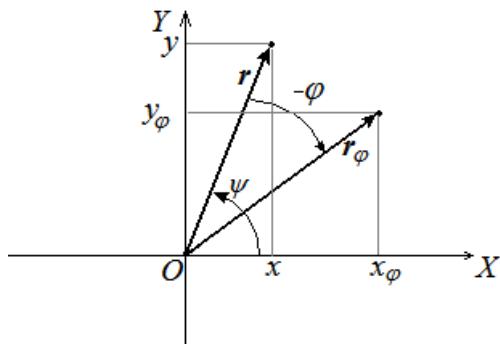


Рис. 3.1. Поворот

Как всем хорошо известно, $a=r\cos \varphi$, $b=r\sin \varphi$. Зададим себе простой вопрос – что произойдет с вектором r , если его повернуть, как стрелку часов, и именно по часовой стрелке до совмещения с осью OX ? Понятно, что координата Y станет нулем. А чему равна координата x ? Очевидно, $x=r$. Из всех написанных равенств заключаем, что $x=r=\sqrt{a^2 + b^2}$, т.е. вектор (a, b) превратится в $(0, \sqrt{a^2 + b^2})$. А как преобразуются остальные векторы плоскости при таком же повороте ?



Нас интересует, как координаты (x_φ, y_φ) вектора, полученного из произвольно взятого $r = (x, y)$ поворотом на угол φ по часовой стрелке, выражаются через эти x, y и угол φ . Эти формулы можно найти в курсе линейной алгебры. Но их нетрудно вывести и непосредственно. Имеем:

$$x = r \cdot \cos \psi, y = r \cdot \sin \psi$$

$$x_\varphi = r \cdot \cos(\psi - \varphi), y_\varphi = r \cdot \sin(\psi - \varphi)$$

Значит, $x_\varphi = r \cdot \cos \psi \cdot \cos \varphi + r \cdot \sin \psi \cdot \sin \varphi = x \cdot \cos \varphi + y \cdot \sin \varphi$,

$$y_\varphi = r \cdot \sin \psi \cdot \cos \varphi - r \cdot \cos \psi \cdot \sin \varphi = y \cdot \cos \varphi - x \cdot \sin \varphi = -x \cdot \sin \varphi + y \cdot \cos \varphi.$$

Эти равенства можно переписать в виде одного матричного равенства

$$\begin{pmatrix} x_\varphi \\ y_\varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3.5)$$

Матрица $\begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}$ – это и есть *матрица вращения*. Частным случаем равенства (3.5) является

$$\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

для взятого нами в начале параграфа вектора $\mathbf{r} = (a, b)$. При этом из $a = r \cdot \cos \varphi$, $b = r \cdot \sin \varphi$ заключаем, что

$$\cos \varphi = \frac{a}{r} = \frac{a}{\sqrt{a^2+b^2}}, \sin \varphi = \frac{b}{r} = \frac{b}{\sqrt{a^2+b^2}}.$$

Выводом из всего выше сказанного является равенство

$$\begin{pmatrix} x & a'_1 & \dots & a'_n \\ 0 & b'_1 & \dots & b'_n \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} a & a_2 & \dots & a_n \\ b & b_2 & \dots & b_n \end{pmatrix}$$

Оно означает, что путем умножения слева произвольной матрицы из двух строк на подходящую матрицу вращения мы можем обнулить первый элемент второй строки. Если пару $\begin{pmatrix} a \\ b \end{pmatrix}$ вставить в качестве второго столбца, то обнулится второй элемент второй строки, если – в качестве третьего столбца, - третий элемент второй строки и т.д. Эти равенства можно обобщить на n -мерный случай. Пусть дана квадратная матрица A порядка n , $A = (a_{ij})_{(i=1,\dots,n;j=1,\dots,n)}$. Пусть

$$T_{ij} = \begin{pmatrix} & i & & j \\ 1 & & & \\ & 1 & & \\ i & & c & s \\ & & 1 & \\ & & & 1 \\ T_{ij} = & j & -s & c \\ & & & 1 \\ & & & & 1 \\ & & & & & 1 \\ & & & & & & 1 \end{pmatrix}$$

где

$$c = \frac{a_{ik}}{\sqrt{a_{ik}^2 + a_{jk}^2}}; \quad s = \frac{a_{jk}}{\sqrt{a_{ik}^2 + a_{jk}^2}},$$

а все пустые места заполнены нулями. Нетрудно видеть, что у произведения матриц $T_{ij}A$ элемент в j -той строке и k -том столбце будет нулевым. В этом можно убедиться и непосредственным вычислением этого элемента в произведении:

$$-sa_{ik} + ca_{jk} = \frac{-a_{jk}a_{ik} + a_{ik}a_{jk}}{\sqrt{a_{ik}^2 + a_{jk}^2}} = 0$$

Кроме того, при умножении A на T_{ij} слева в произведении получится почти та же матрица A . Изменятся только i -тая и j -тая строки, которые станут линейными комбинациями i -той и j -той строк матрицы A с коэффициентами s и $-s$, для i -той строки и $-s$, c для j -той.

Матрицы T_{ij} называются *n*-мерными **матрицами вращения**. Такая матрица соответствует повороту двумерной плоскости, проходящей через i -тую и j -тую координатные оси в векторном пространстве *n*-мерных числовых строк. Эта матрица зависит от индексов $i < j$ и двух чисел a и b .

Матрицы вращения обладают одним важным свойством: $T_{ij} \cdot T_{ij}^T = E$ – единичной матрице. Это говорит о том, что обратной матрицей к матрице вращения T_{ij} является транспонированная к T_{ij} . Вообще матрица Q называется **ортогональной**, если $Q^{-1} = Q^T$. Матрицы вращения – это частный случай ортогональных матриц.

Теперь можно, наконец, объяснить, в чем состоит метод вращений.

Сделаем это на примере той же системы с матрицей (3.4). Это матрица $A =$

0.3032	8.2869	-5.7604	-9.9827	8.6236	136.1535
3.4133	-6.0824	9.4853	9.4349	3.1861	-139.6078
-4.2811	6.9505	0.2575	-7.5744	6.5649	91.3667
-1.3291	-5.3959	-8.0335	-9.6603	-8.8668	152.3082
-2.4302	4.2187	3.3620	-2.2898	-1.1729	-29.2527

Делаются следующие шаги. Вычисляем матрицу вращения T_{12} , взяв в качестве a и b числа $a = a_{11} = 0.3032$ и $b = a_{21} = 3.4133$. Получим

$$T_{12} = \begin{bmatrix} 0.08848 & 0.99608 & 0 & 0 & 0 \\ -0.99608 & 0.08848 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{12} \cdot (\mathbf{A}/\mathbf{b})$ получим $(\mathbf{A}_1/\mathbf{b}_1) = (\mathbf{T}_{12} \cdot \mathbf{A}/\mathbf{T}_{12} \cdot \mathbf{b}) =$

$$\begin{array}{rrrrrr} 3.42675 & -5.32533 & 8.93844 & 8.51465 & 3.93663 & -127.01368 \\ 0 & -8.79259 & 6.57708 & 10.77837 & -8.30789 & -147.97228 \\ -4.2811 & 6.9505 & 0.2575 & -7.5744 & 6.5649 & 91.3667 \\ -1.3291 & -5.3959 & -8.0335 & -9.6603 & -8.8668 & 152.3082 \\ -2.4302 & 4.2187 & 3.362 & -2.2898 & -1.1729 & -29.2527 \end{array}$$

Теперь для полученной матрицы \mathbf{A}_1 вычисляем матрицу вращения \mathbf{T}_{13} , взяв в качестве a и b числа $a = a_{11} = 3.42675$ и $b = a_{31} = -4.2811$

$$\mathbf{T}_{13} = \begin{bmatrix} 0.6249 & 0 & -0.7807 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0.7807 & 0 & 0.6249 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{13} \cdot (\mathbf{A}_1/\mathbf{b}_1)$ получим $(\mathbf{A}_2/\mathbf{b}_2) = (\mathbf{T}_{13} \cdot \mathbf{A}_1/\mathbf{T}_{13} \cdot \mathbf{b}_1) =$

$$\begin{array}{rrrrrr} 5.48363 & -8.75405 & 5.3846 & 11.23414 & -2.66522 & -150.70083 \\ 0 & -8.79259 & 6.57708 & 10.77837 & -8.30789 & -147.97228 \\ 0 & 0.18588 & 7.13915 & 1.91414 & 7.17573 & -42.06453 \\ -1.3291 & -5.3959 & -8.0335 & -9.6603 & -8.8668 & 152.3082 \\ -2.4302 & 4.2187 & 3.362 & -2.2898 & -1.1729 & -29.2527 \end{array}$$

Продолжая в том же духе серией умножений на матрицы вращения приводим матрицу системы к треугольному виду:

Вычисляем \mathbf{T}_{14} , взяв в качестве a и b числа $a = a_{11} = 5.48363$ и $b = a_{41} = -1.3291$

$$\mathbf{T}_{14} = \begin{bmatrix} 0,97186 & 0 & 0 & -0,23556 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0,23556 & 0 & 0 & 0,97186 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{14} \cdot (\mathbf{A}_2/\mathbf{b}_2)$ получим $(\mathbf{A}_3/\mathbf{b}_3) = (\mathbf{T}_{14} \cdot \mathbf{A}_2/\mathbf{T}_{14} \cdot \mathbf{b}_2) =$

$$\begin{array}{rrrrrr} 5,6424 & -7,23665 & 7,12545 & 13,19359 & -0,50156 & -182,33783 \\ 0 & -8,79259 & 6,57708 & 10,77837 & -8,30789 & -147,97228 \\ 0 & 0,18588 & 7,13915 & 1,91414 & 7,17573 & -42,06453 \\ 0 & -7,30616 & -6,53904 & -6,74215 & -9,24511 & 112,52316 \\ -2,4302 & 4,2187 & 3,362 & -2,2898 & -1,1729 & -29,2527 \end{array}$$

Вычисляем \mathbf{T}_{15} , взяв в качестве a и b числа $a = a_{11} = 5,6424$ и $b = a_{41} = -2,4302$

$$\mathbf{T}_{15} = \begin{bmatrix} 0,91843 & 0 & 0 & 0 & -0,39557 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0,39557 & 0 & 0 & 0 & 0,91843 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{15} \cdot (\mathbf{A}_3/\mathbf{b}_3)$ получим $(\mathbf{A}_4/\mathbf{b}_4) = (\mathbf{T}_{15} \cdot \mathbf{A}_3/\mathbf{T}_{15} \cdot \mathbf{b}_3) =$

$$\begin{array}{rrrrrr} 6,14346 & -8,31515 & 5,21432 & 13,02317 & 0,00332 & -155,89304 \\ 0 & -8,79259 & 6,57708 & 10,77837 & -8,30789 & -147,97228 \\ 0 & 0,18588 & 7,13915 & 1,91414 & 7,17573 & -42,06453 \\ 0 & -7,30616 & -6,53904 & -6,74215 & -9,24511 & 112,52316 \end{array}$$

$$0 \quad 1,01198 \quad 5,90638 \quad 3,11597 \quad -1,27563 \quad | \quad -98,99393$$

Вычисляем \mathbf{T}_{23} , взяв в качестве a и b числа $a = a_{22} = -8,79259$ и $b = a_{32} = 0,18588$

$$\mathbf{T}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -0,99978 & 0,02114 & 0 & 0 \\ 0 & -0,02114 & -0,99978 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{23} \cdot (\mathbf{A}_4 / \mathbf{b}_4)$ получим $(\mathbf{A}_5 / \mathbf{b}_5) = (\mathbf{T}_{23} \cdot \mathbf{A}_4 / \mathbf{T}_{23} \cdot \mathbf{b}_4) =$

$$\begin{array}{cccccc|c} 6,14346 & -8,31515 & 5,21432 & 13,02317 & 0,00332 & & -155,89304 \\ 0 & 8,79459 & -6,42471 & -10,73553 & 8,45776 & & 147,05048 \\ 0 & 0 & -7,27662 & -2,14157 & -6,99852 & & 45,18341 \\ 0 & -7,30616 & -6,53904 & -6,74215 & -9,24511 & & 112,52316 \\ 0 & 1,01198 & 5,90638 & 3,11597 & -1,27563 & & -98,99393 \end{array}$$

Вычисляем \mathbf{T}_{24} , взяв в качестве a и b числа $a = a_{22} = 8,79459$ и $b = a_{42} = -7,30616$

$$\mathbf{T}_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0,7692 & 0 & -0,63901 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0,63901 & 0 & 0,7692 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{24} \cdot (\mathbf{A}_5 / \mathbf{b}_5)$ получим $(\mathbf{A}_6 / \mathbf{b}_6) = (\mathbf{T}_{24} \cdot \mathbf{A}_5 / \mathbf{T}_{24} \cdot \mathbf{b}_5) =$

$$\begin{array}{cccccc|c} 6,14346 & -8,31515 & 5,21432 & 13,02317 & 0,00332 & & -155,89304 \\ 0 & 11,43351 & -0,76337 & -3,94947 & 12,41343 & & 41,2078 \\ 0 & 0 & -7,27662 & -2,14157 & -6,99852 & & 45,18341 \\ 0 & 0 & -9,13528 & -12,04617 & -1,70675 & & 180,51954 \\ 0 & 1,01198 & 5,90638 & 3,11597 & -1,27563 & & -98,99393 \end{array}$$

Вычисляем \mathbf{T}_{25} , взяв в качестве a и b числа $a = a_{22} = 11,43351$ и $b = a_{52} = 1,01198$

$$\mathbf{T}_{25} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0,99611 & 0 & 0 & 0,08817 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -0,08817 & 0 & 0 & 0,99611 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{25} \cdot (\mathbf{A}_6 / \mathbf{b}_6)$ получим $(\mathbf{A}_7 / \mathbf{b}_7) = (\mathbf{T}_{25} \cdot \mathbf{A}_6 / \mathbf{T}_{25} \cdot \mathbf{b}_6) =$

$$\begin{array}{cccccc|c} 6,14346 & -8,31515 & 5,21432 & 13,02317 & 0,00332 & & -155,89304 \\ 0 & 11,47826 & -0,23963 & -3,65937 & 12,25267 & & 32,31921 \\ 0 & 0 & -7,27662 & -2,14157 & -6,99852 & & 45,18341 \\ 0 & 0 & -9,13528 & -12,04617 & -1,70675 & & 180,51954 \\ 0 & 0 & 5,95071 & 3,45207 & -2,36516 & & -102,24214 \end{array}$$

Вычисляем \mathbf{T}_{34} , взяв в качестве a и b числа $a = a_{33} = -7,27662$ и $b = a_{43} = -9,13528$

$$\mathbf{T}_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -0,62304 & -0,78219 & 0 \\ 0 & 0 & 0,78219 & -0,62304 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{34} \cdot (\mathbf{A}_7 / \mathbf{b}_7)$ получим $(\mathbf{A}_8 / \mathbf{b}_8) = (\mathbf{T}_{34} \cdot \mathbf{A}_7 / \mathbf{T}_{34} \cdot \mathbf{b}_7) =$

$$\begin{array}{cccccc|c} 6,14346 & -8,31515 & 5,21432 & 13,02317 & 0,00332 & -155,89304 \\ 0 & 11,47826 & -0,23963 & -3,65937 & 12,25267 & 32,31921 \\ 0 & 0 & 11,67915 & 10,75668 & 5,69536 & -169,35165 \\ 0 & 0 & 0 & 5,83013 & -4,4108 & -77,12888 \\ 0 & 0 & 5,95071 & 3,45207 & -2,36516 & -102,24214 \end{array}$$

Вычисляем \mathbf{T}_{35} , взяв в качестве a и b числа $a = a_{33} = 11,67915$ и $b = a_{53} = 5,95071$

$$\mathbf{T}_{35} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0,89101 & 0 & 0,45398 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -0,45398 & 0 & 0,89101 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{35} \cdot (\mathbf{A}_8 / \mathbf{b}_8)$ получим $(\mathbf{A}_9 / \mathbf{b}_9) = (\mathbf{T}_{35} \cdot \mathbf{A}_8 / \mathbf{T}_{35} \cdot \mathbf{b}_8) =$

$$\begin{array}{cccccc|c} 6,14346 & -8,31515 & 5,21432 & 13,02317 & 0,00332 & -155,89304 \\ 0 & 11,47826 & -0,23963 & -3,65937 & 12,25267 & 32,31921 \\ 0 & 0 & 13,10774 & 11,15148 & 4,00089 & -197,3099 \\ 0 & 0 & 0 & 5,83013 & -4,4108 & -77,12888 \\ 0 & 0 & 0 & -1,80749 & -4,69296 & -14,21651 \end{array}$$

Вычисляем \mathbf{T}_{45} , взяв в качестве a и b числа $a = a_{44} = 5,83013$ и $b = a_{54} = -1,80749$

$$\mathbf{T}_{45} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0,95515 & -0,29612 \\ 0 & 0 & 0 & 0,29612 & 0,95515 \end{bmatrix}$$

Вычисляя произведение $\mathbf{T}_{45} \cdot (\mathbf{A}_9 / \mathbf{b}_9)$ получим $(\mathbf{A}_{10} / \mathbf{b}_{10}) = (\mathbf{T}_{45} \cdot \mathbf{A}_9 / \mathbf{T}_{45} \cdot \mathbf{b}_9) =$

$$\begin{array}{cccccc|c} 6,14346 & -8,31515 & 5,21432 & 13,02317 & 0,00332 & -155,89304 \\ 0 & 11,47826 & -0,23963 & -3,65937 & 12,25267 & 32,31921 \\ 0 & 0 & 13,10774 & 11,15148 & 4,00089 & -197,3099 \\ 0 & 0 & 0 & 6,10388 & -2,8233 & -69,45986 \\ 0 & 0 & 0 & 0 & -5,78861 & -36,4183 \end{array}$$

Мы привели матрицу системы к треугольному виду. Осталось сделать обратный ход метода Гаусса точно так же, как в п. 3.1. Приводим окончательный результат

$$\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & -8,34372 \\ 0 & 1 & 0 & 0 & 0 & -6,80425 \\ 0 & 0 & 1 & 0 & 0 & -9,76769 \\ 0 & 0 & 0 & 1 & 0 & -8,4696 \\ 0 & 0 & 0 & 0 & 1 & 6,29137 \end{array}$$

Последний столбец и есть решение. Невязка на этот раз оказывается равна $\|\mathbf{r}\| = 0,00029675$ и это значительно лучше, чем во всех предыдущих методах.

3.4. Метод простой итерации. Методы решения систем линейных уравнений, рассмотренные выше, относятся к так-называемым прямым методам нахождения решений. Далее будут приведены несколько так-называемых итерационных методов. Любой из этих методов осуществляется по внешне простой схеме. Исходная система $AX=b$ каким-то способом приводится к равносильному уравнению $X=F(X)$. Здесь $X=(x_1, \dots, x_n)$ - вектор из неизвестных, который мы иногда записываем в виде столбца. F - это какая-то функция $F : R^n \rightarrow R^n$. Здесь R - это множество вещественных чисел, R^n - множество n -мерных векторов, т.е. строк вида (x_1, \dots, x_n) из элементов R . Функция F возникает в результате какого-то преобразования уравнения $AX=b$. Как бы она не возникала, должна сохраняться возможность вычисления $F(X)$ в любой точке X пространства R^n . И это вычисление должно быть по возможности простым (для того, чтобы метод имел вычислительную ценность). Основная идея метода – простой вопрос: “Что будет, если взять произвольную точку $X_0 \in R^n$ и последовательно вычислить вектора: $X_1 = F(X_0)$, $X_2 = F(X_1)$, $X_3 = F(X_2)$, ... ?” Каждый шаг $X_{n+1} = F(X_n)$ этого вычисления называется *итерацией*.

Оказывается, что весьма высока вероятность того, что получаемая последовательность векторов $X_0, X_1, X_2, X_3, \dots$ имеет предел – какую-то точку X^* из R^n . Если ко всему этому функция F непрерывна, то переходя к пределу при $n \rightarrow \infty$ в равенстве $X_{n+1} = F(X_n)$ мы получим $X^* = F(X^*)$. Т.е. этот предел X^* будет решением системы.

Переделать данную систему в равенство $X = F(X)$ можно множеством различных способов. К сожалению, из них невозможно выбрать самый наилучший. В этом параграфе рассматривается, вероятно, самый простой. Но его недостатком будет то обстоятельство, что последовательность X_0, X_1, X_2, \dots далеко не всегда будет сходиться.

Итак, равенство $AX=b$ запишем в эквивалентном виде как $X=X-AX+b$, что равносильно

$$X = (E - A)X + b$$

Здесь E – единичная матрица порядка n . Таким образом, в *методе простых итераций* $F(X) = (E - A)X + b$. Сходимость или расходимость метода целиком зависит от матрицы $E - A$. Поэтому хотелось бы до начала

вычислений как-то изучить эту матрицу и получить из ее анализа информацию о сходимости.

На этот счет в математике имеется развитая и нетривиальная теория. Центральным моментом в ней является так-называемый *принцип сжимающих отображений*. Он гласит, что если в некотором полном метрическом пространстве P с метрикой $\rho(x,y)$ имеется функция $f: P \rightarrow P$, у которой при некотором числе $q < 1$ для любых $x, y \in P$ выполняется $\rho(f(x), f(y)) \leq q \cdot \rho(x, y)$, то уравнение $x = f(x)$ имеет в P единственное решение. Если отображение f имеет указанное выше свойство, то оно называется *сжимающим* отображением. А решение уравнения $x = f(x)$ называется *неподвижной точкой* функции f . Таким образом, принцип сжимающих отображений – это теорема, которая гласит, что в полном метрическом пространстве любое сжимающее отображение имеет и единственную неподвижную точку.

Что такое метрическое пространство и метрика? Это любое непустое множество P и функция $\rho: P^2 \rightarrow R$, для которой выполняются определенные условия (некие свойства расстояния) – аксиомы метрики.

В нашем случае роль полного метрического пространства выполняет множество R^n . Метрику на нем можно выбирать разными способами (которых бесконечно много). Но для сходимости метода простых итераций достаточно, чтобы существовала хотя бы одна метрика, при которой функция $F(X) = (E - A)X + b$ являлась бы сжимающей.

Для того, чтобы найти какие-нибудь достаточные и легко проверяемые условия на матрицу $(E - A)$, при которых такая метрика существует, надо обратиться к некоторым разделам линейной алгебры. Дело в том, что у нас $F(X) = (E - A)X + b = H \cdot X + b$. Основную роль здесь играет линейное преобразование $G(X) = H \cdot X$ пространства R^n . Напомним, что $H = E - A$. На роль метрики назначим метрику, определяемую подходящей нормой в векторном пространстве R^n . **Норма** – это числовая характеристика векторов – нечто вроде длины вектора, для которой должны выполняться определенные условия – аксиомы нормы. Норму вектора $x \in R^n$ обычно обозначают как $\|x\|$, а метрика $\rho(x,y)$ через норму определяется равенством $\rho(x,y) = \|x-y\|$.

Оказывается, любая норма в R^n для любого линейного преобразования G определяет некую норму $\|G\|$ и этого преобразования. Именно, $\|G\| = \max\{\|G(x)\| : x \in R^n, \|x\|=1\}$. Можно сказать, $\|G\|$ – это максимальная

длина образов векторов из единичной сферы. Норма G заодно называется и **нормой матрицы H** т.е. $\|H\|=\|G\|$. Оказывается, что для любого $x \in R^n$ $\|G(x)\| \leq \|G\| \cdot \|x\|$, т.е. $\|G(x)\| \leq \|H\| \cdot \|x\|$. Поэтому, если $\|H\| < 1$, то отображение G будет сжимающим. А вместе с ним сжимающим будет и отображение F .

Таким образом, для того, чтобы метод простой итерации сходился достаточно существования такой нормы в R^n , для которой норма матрицы $E-A$ была бы меньше 1.

Самой популярной и важной в R^n является **евклидова норма**, которая определяется формулой $\|x\| = \|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$. К сожалению для этой нормы очень трудно вычислять нормы матриц. Для этого надо находить собственные числа, а эта задача, как мы еще увидим, сама по себе превосходит по сложности все методы решения систем линейных уравнений. По счастью имеются две нормы в R^n , которые просто вычисляются и для которых просто вычисляются нормы матриц. Это:

- 1) $\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$ – максимум модулей координат вектора.
- 2) $\|x\|_1 = |x_1| + \dots + |x_n|$ – сумма модулей координат.

Для этих норм соответствующие нормы матрицы $A = (a_{ij})_{i,j=1,\dots,n}$ вычисляются по формулам:

- 1) $\|A\|_\infty = \max\{|a_{11}| + |a_{12}| + \dots + |a_{1n}|, |a_{21}| + |a_{22}| + \dots + |a_{2n}|, \dots, |a_{n1}| + |a_{n2}| + \dots + |a_{nn}|\}$.
- 2) $\|A\|_1 = \max\{|a_{11}| + |a_{21}| + \dots + |a_{n1}|, |a_{12}| + |a_{22}| + \dots + |a_{n2}|, \dots, |a_{1n}| + |a_{2n}| + \dots + |a_{nn}|\}$.

В случае 1) надо по каждой строке матрицы вычислить сумму абсолютных величин элементов этих строк и среди получившихся n чисел выбрать наибольшее. В случае 2) надо то же самое проделать со столбцами матрицы A .

Подводя итог этим теоретическим сведениям, сделаем простой вывод: *Если для матрицы A хотя бы одно из чисел $\|A\|_\infty$ или $\|A\|_1$ строго меньше единицы, то метод простой итерации для системы $AX=b$ сходится к решению этой системы. Причем - для любого b и стартия с любого $x_0 \in R^n$.*

Приведем теперь пример решения системы этим методом. Если попытаться его применить к системе (3.4) с расширенной матрицей

0.3032	8.2869	-5.7604	-9.9827	8.6236	136.1535
3.4133	-6.0824	9.4853	9.4349	3.1861	-139.6078
-4.2811	6.9505	0.2575	-7.5744	6.5649	91.3667
-1.3291	-5.3959	-8.0335	-9.6603	-8.8668	152.3082
-2.4302	4.2187	3.3620	-2.2898	-1.1729	-29.2527

то стоит предварительно посмотреть на нормы $\| \cdot \|_\infty$ и $\| \cdot \|_1$. И без всякого вычисления видно, что обе нормы матрицы $(E-A)$ значительно больше

единицы. Поэтому метод даже не стоит пробовать – он наверняка не будет сходиться. Чтобы уменьшить эти нормы, можно попробовать разделить каждую строку матрицы на какое-нибудь большое число – например, на 100. Но тогда в матрице $(E-A)$ на диагонали будут стоять числа, близкие к 1, и некоторые из них будут больше 1. Например, во второй строке и втором столбце у $E-A$ будет стоять 1.060824. Можно попытаться поменять знак в тех строках системы, где диагональный коэффициент отрицателен. Тогда на диагонали везде будут стоять числа немного меньшие 1. Например, вместо 1.060824 в позиции $_{22}$ будет стоять 0.939176. Но это мало поможет. При суммировании с модулями остальных чисел второй строки или второго столбца мы получим число, большее 1. Поэтому для произвольно взятой системы метод простой итерации обычно не годится.

Но, оказывается, есть класс систем, довольно часто возникающих в практических прикладных задачах, где этот метод прекрасно работает. Это системы, в матрице которой наблюдается явное преобладание диагональных элементов – числа на диагонали по модулю значительно больше всех остальных. Точное определение таково: “Матрица $A=(a_{ij})_{i,j=1,\dots,n}$ обладает **свойством строгого преобладания**, если для каждого $i=1,\dots,n$ сумма $\sum_{j,j \neq i}^n |a_{ij}| < |a_{ii}|$ ”. Оказывается, что системы, матрица которых имеет это свойство, довольно часто встречаются на практике. Иногда достаточно переставить местами уравнения системы, чтобы ее матрица приобрела свойство диагонального преобладания. Имея такую матрицу, надо разделить каждое i -тое уравнение системы на диагональный элемент a_{ii} . Тогда коэффициенты на диагонали превратятся в единицы, а у матрицы $(E-A)$ станут нулями. Легко видеть, что теперь $\|E-A\|_\infty < 1$ и процесс простой итерации будет сходиться.

Приведем пример. Пусть расширенная матрица системы имеет вид

-9,4368	0,31229	0,12008	-0,10491	-0,63085	48,8116
0,63864	3,94688	-0,28872	-0,12941	0,17357	-27,569
0,05316	0,37559	5,41936	-0,53773	0,08506	16,3095
0,17163	0,88415	0,71594	-4,60339	-0,47063	9,91787
-0,89624	0,17969	-0,03469	-0,91332	-9,81284	77,4314

Матрица системы, очевидно, имеет свойство строгого диагонального преобладания. Делим первую строку на $a_{11}=-9,4368$, вторую строку – на $a_{22}=3,94688$, третью – на $a_{33}=5,41936$, четвертую – $a_{44}=-4,60339$, пятую - на $a_{55}=-9,81284$. Получим

1	-0,03309	-0,01272	0,01112	0,06685	-5,1725
0,16181	1	-0,07315	-0,03279	0,04398	-6,9849
0,00981	0,06931	1	-0,09922	0,0157	3,00948
-0,03728	-0,19206	-0,15552	1	0,10224	-2,1545
0,09133	-0,01831	0,00354	0,09307	1	-7,8908

Теперь $E-A=$

0	-0,03309	-0,01272	0,01112	0,06685
0,16181	0	-0,07315	-0,03279	0,04398
0,00981	0,06931	0	-0,09922	0,0157
-0,03728	-0,19206	-0,15552	0	0,10224
0,09133	-0,01831	0,00354	0,09307	0

и $\|E-A\|_\infty = 0.4871$. Это сумма модулей чисел четвертой строки.

Для того, чтобы проделать процесс простой итерации, воспользуемся программой MS Excel. Для этого надо матрицу $E-A$ и столбец b занести в лист Excel и ввести в каких-нибудь подходящих ячейках формулу

$$X_1 = (E - A)X_0 + b,$$

вычисляющую X_1 по X_0 . В качестве X_1 взять любые 5 чисел. Затем сделать соответствующую протяжку и получить таблицу, показанную на картинке.

14	$E-A$					b				
15	0	0,03309	0,01272	-0,01112	-0,06685	-0,0322	-5,17247			
16	-0,1618	0	0,07315	0,03279	-0,04398	-0,0999	-6,98492			
17	-0,0098	-0,06931	0	0,09922	-0,0157	0,0044	3,00948			
18	0,03728	0,19206	0,15552	0	-0,10224	0,28262	-2,15447			
19	-0,0913	0,01831	-0,00354	-0,09307	0	-0,1696	-7,89083			
20										
21	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
22	1	-5,20463	-4,80891	-4,79398	-4,80464	-4,8019	-4,80224	-4,8022	-4,8022	-4,8022
23	1	-7,08477	-5,62917	-5,70591	-5,71264	-5,7086	-5,70923	-5,70913	-5,70914	-5,70914
24	1	3,01388	3,492407	3,322952	3,357204	3,35382	3,354103	3,354152	3,354143	3,354145
25	1	-1,87185	-2,41638	-2,11704	-2,16162	-2,1549	-2,15505	-2,15507	-2,15504	-2,15504
26	1	-8,06046	-7,38167	-7,34217	-7,3722	-7,3673	-7,36811	-7,36808	-7,36808	-7,36808
27										

Ниже показаны формулы для вычисления элементов столбца X_1 .

Буфер обмена	Шрифт	Выравнивание	Число
C22	f _x	=\\$B\$15*B\$22+\$C\$15*B\$23+\$D\$15*B\$24+\$E\$15*B\$25+\$F\$15*B\$26+\$H\$15	
Буфер обмена	Шрифт	Выравнивание	Число
C23	f _x	=\\$B\$16*B\$22+\$C\$16*B\$23+\$D\$16*B\$24+\$E\$16*B\$25+\$F\$16*B\$26+\$H\$16	
Буфер обмена	Шрифт	Выравнивание	Число
C24	f _x	=\\$B\$17*B\$22+\$C\$17*B\$23+\$D\$17*B\$24+\$E\$17*B\$25+\$F\$17*B\$26+\$H\$17	
Буфер обмена	Шрифт	Выравнивание	Число
C25	f _x	=\\$B\$18*B\$22+\$C\$18*B\$23+\$D\$18*B\$24+\$E\$18*B\$25+\$F\$18*B\$26+\$H\$18	
Буфер обмена	Шрифт	Выравнивание	Число
C26	f _x	=\\$B\$19*B\$22+\$C\$19*B\$23+\$D\$19*B\$24+\$E\$19*B\$25+\$F\$19*B\$26+\$H\$19	

После X_9 вектора X_n практически не меняются. Потому можно считать вектор X_9 решением. Невязка для него оказывается меньше 0,0000008. Это объясняется не только тем, что Excel делает вычисления с предельно возможной точностью (без специальных округлений), но и тем, что итера-

ционные методы сходятся к решению с любого места. Поэтому вычислительные погрешности не мешают векторам X_n приближаться к точному решению. В этом одно из преимуществ итерационных методов.

3.5. Метод наискорейшего спуска. Как и раньше имеем систему линейных уравнений $AX=b$ с невырожденной квадратной матрицей A размера $n \times n$. От этой системы перейдем к эквивалентной $(A^T A)X=A^T b$, полученной умножением слева обеих частей уравнения $AX=b$ на матрицу A^T (транспонированную матрицу A). Эта система равносильна исходной, т.к. из нее можно получить первоначальное матричное равенство умножением слева на матрицу $(A^T)^{-1}$. Поэтому в этом и в следующем параграфе будем считать, что матрица A – это и есть $A^T A$. Такая замена теоретически не изменяет решение, но из-за того, что переход от A к $A^T A$ требует довольно большого количества операций умножения и сложения, страдает точность. Но это – неизбежное зло метода. Чем же новая матрица A лучше прежней? Оказывается, что она обладает двумя очень важными свойствами:

1) Новая A является **симметрической**, т.е. у нее всегда $a_{ij} = a_{ji}$. Это то же самое, что $A^T = A$ или, что на симметричных относительно главной диагонали местах стоят одинаковые числа.

2) Новая A является **положительно определенной**. Это означает, что квадратичная форма с матрицей A при ненулевых значениях переменных всегда принимает строго положительные значения.

Разумеется, если первоначальная матрица A сразу имеет свойства 1) и 2), то переходить от нее к $A^T A$ не нужно.

Квадратичная форма с симметрической матрицей A может пониматься как функция от числовых векторов $x = (x_1, \dots, x_n)$, действующая по правилу $x \rightarrow xAx^T$. В координатах эта квадратичная форма имеет вид

$$xAx^T = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

Рассмотрим теперь функцию $F(x) = -\frac{1}{2} xAx^T + (b, x)$. Здесь $(b, x) = b_1 x_1 + \dots + b_n x_n$ – скалярное произведение векторов b и x в пространстве \mathbb{R}^n . Оказывается, что функция $F(x)$ в пространстве \mathbb{R}^n достигает максимально возможного значения в единственной точке и эта точка как раз и является решением системы $AX=b$. Поэтому задача решения этой системы сводится к отысканию максимального значения функции $F(x)$.

Из курса математического анализа известно, что дифференцируемая функция от n переменных в каждой точке пространства \mathbf{R}^n быстрее всего растет в направлении градиента в этой точке. А градиент – это вектор, составленный из частных производных. В каждой точке $\mathbf{x}^* \in \mathbf{R}^n$ его нетрудно найти по формуле $\nabla F(\mathbf{x}^*) = \mathbf{b} - A\mathbf{x}^*$. А это – ни что иное, как вектор невязки \mathbf{r}^* в точке \mathbf{x}^* . (Здесь и далее строки из \mathbf{R}^n считаем, если надо, столбцами, чтобы не загромождать обозначения символами транспонирования T).

Метод наискорейшего спуска (в данном случае, правда, мы поднимаемся) заключается в последовательном выполнении следующих шагов. Стартуем с произвольной точки $\mathbf{x}_0 \in \mathbf{R}^n$. В направлении градиента $\nabla F(\mathbf{x}_0)$ от этой точки ищем максимум функции $F(\mathbf{x})$. Получаем точку $\mathbf{x}_1 \in \mathbf{R}^n$. В направлении градиента $\nabla F(\mathbf{x}_1)$ от этой точки ищем максимум функции $F(\mathbf{x})$. Получаем точку $\mathbf{x}_2 \in \mathbf{R}^n$. И т.д. пока длина градиента, т.е. невязка не станет достаточно мала.

Для окончательного оформления метода остается понять, как отыскивать максимум функции $F(\mathbf{x})$ в данном направлении \mathbf{e} от данной точки \mathbf{x}^* . Этот максимум – это максимум функции $g(t) = F(\mathbf{x}^* + t\mathbf{e})$. Методам одномерной оптимизации посвящен один из последующих разделов. Но в данном случае можно обойтись простым школьным методом. Точка t максимума функции $g(t)$ – это точка, в которой производная $g'(t) = 0$. Несложное вычисление дает уравнение $(\mathbf{r}^*, \mathbf{e}) - t(A\mathbf{e}, \mathbf{e}) = 0$ для переменной t . Значит максимум $g(t)$ достигается при

$$t = \frac{(\mathbf{x}^*, \mathbf{e})}{(A\mathbf{e}, \mathbf{e})}. \quad (3.6)$$

В нашем случае $\mathbf{e} = \mathbf{r}^*$.

Таким образом, метод наискорейшего спуска состоит в следующем:

Берем любой $\mathbf{x}_0 \in \mathbf{R}^n$. Вычисляем

$$\begin{aligned} \mathbf{r}_k &= \mathbf{b} - A\mathbf{x}_k \\ \alpha_k &= \frac{(\mathbf{r}_k, \mathbf{r}_k)}{(A\mathbf{r}_k, \mathbf{r}_k)} \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \alpha_k \mathbf{r}_k \\ (k &= 0, 1, 2, \dots) \end{aligned} \quad (3.7)$$

Вычисление продолжается до тех пор, пока невязка $\|\mathbf{r}_k\|$ не станет меньше заранее заданного числа.

Главным недостатком метода является его медленная сходимость. Следующий рисунок демонстрирует работу метода для $n=2$.

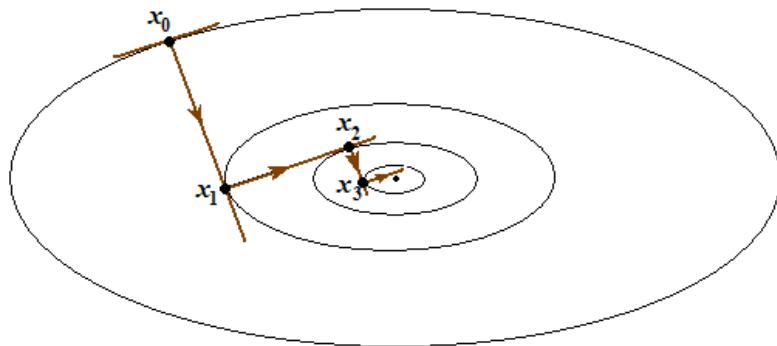


Рис 3.1. Метод наискорейшего спуска

Эллипсы – это линии уровня функции F . С каждым шагом мы приближаемся к решению – центру всех эллипсов. Но чем ближе к этому центру, тем мельче следующий шаг. Чтобы достичь требуемой точности для некоторых ничем не примечательных систем от пяти переменных приходится делать несколько тысяч шагов. При этом, естественно, должны накапливаться ошибки округлений. Но, т.к. метод от любой точки сходится к решению, их роль невелика.

Приведем пример. Возьмем ту же систему из п. 3.1.

0.3032	8.2869	-5.7604	-9.9827	8.6236	136.1535
3.4133	-6.0824	9.4853	9.4349	3.1861	-139.6078
-4.2811	6.9505	0.2575	-7.5744	6.5649	91.3667
-1.3291	-5.3959	-8.0335	-9.6603	-8.8668	152.3082
-2.4302	4.2187	3.3620	-2.2898	-1.1729	-29.2527

После умножения на A^T матрица A примет вид

37.7427	-51.0848	32.0342	80.0083	0.02	-957.7345
-51.0848	200.8909	-46.1083	-150.2923	140.6094	1667.237
32.0342	-46.1083	199.0596	214.9545	49.5243	-3406.9091
80.0083	-150.2923	214.9545	344.6078	-17.4098	-4772.7732
0.02	140.6094	49.5243	-17.4098	207.6115	12.9663

Вычисления по формулам (3.7) можно сделать в MS Excel в виде

	A	B	C	D	E	F	G	H
1								
Система с симметрической положительной матрицей								
3		37,74	-51,08	32,03	80,01	0,02	-957,73	
4		-51,08	200,89	-46,11	-150,29	140,61	1667,24	
5		32,03	-46,11	199,06	214,95	49,52	-3406,91	
6		80,01	-150,29	214,95	344,61	-17,41	-4772,77	
7		0,02	140,61	49,52	-17,41	207,61	12,97	
8								
9	k	0	1	2	3	4	5	6
10	x_k	0,0000	-1,6756	-1,2636	-1,0733	-0,8546	-0,8335	-0,7103
11		0,0000	2,9170	0,9672	0,6756	-0,0569	-0,4425	-0,9414
12		0,0000	-5,9607	-6,8256	-7,2858	-7,1033	-7,5453	-7,3380
13		0,0000	-8,3504	-8,4993	-9,2282	-8,8172	-9,4574	-9,1430
14		0,0000	0,0227	-0,8918	0,3696	0,4714	1,0290	1,2047
15	r_k	-957,7345	113,5663	38,0467	89,0059	4,5999	49,4746	-12,4598
16		1667,2370	-537,3803	-58,3048	-298,1468	-83,9427	-200,4104	-61,7933
17		-3406,9091	-238,3764	-92,0163	74,2709	-96,2228	83,2983	-61,2106
18		-4772,7732	-41,0344	-145,7388	167,2959	-139,3840	126,3077	-108,3829
19		12,9663	-252,0439	252,2081	41,4176	121,4035	70,5902	99,4580
20	Ar_k	-612316,7728	20813,9307	-10188,4641	34355,1888	-9770,0571	24880,7045	-7943,9431
21		1260079,4490	-132037,7709	47952,4532	-87185,9182	25357,1623	-55686,2149	21318,9469
22		-1811018,6149	-40338,2457	-33246,3809	69394,7800	-39085,0476	58053,1343	-28106,3283
23		-2704490,6477	28857,5363	-62586,1275	124825,6293	-57845,9587	94281,5364	-43948,4674
24		151470,2310	-138976,6797	42144,1020	-32556,0776	11062,9942	-11596,9748	10815,1545
25	(r_k, r_k)	38083496,3459	423708,1235	98162,7536	132033,0674	50493,0577	70487,3004	29359,1502
26	(Ar_k, r_k)	21767148269,3763	116778011,3221	19626025,4990	53740448,3609	10993266,7390	28316643,4502	6340933,5930
27	α_k	0,0017	0,0036	0,0050	0,0025	0,0046	0,0025	0,0046
28								

В область B3 – G7 заносится расширенная матрица системы. В столбец B10 – B14 записываем любые числа – стартовый вектор x_0 . В данном примере он состоит из нулей. В область B15 – B27 заносятся формулы, вычисляющие r_0 , Ar_0 , скалярные произведения (r_0, r_0) , (Ar_0, r_0) и их отношение α_0 . Формулы должны вноситься с тем расчетом, чтобы при растягивании ячеек B15 – B27 не модифицировались ссылки на элементы матрицы A – были “заморожены долларами”. Затем в C10 – C14 заносим формулы, вычисляющие x_1 . Теперь осталось “потянуть” B15 – B27 вправо на 1 позицию и следом – весь столбец C10 – C27 вправо до тех пор, пока вектор невязки r_k постепенно приближается к нулю, но делает это весьма лениво. Для нашей матрицы только начиная с итерации 798 все координаты вектора r_k устойчиво имеют 4 нуля после запятой.

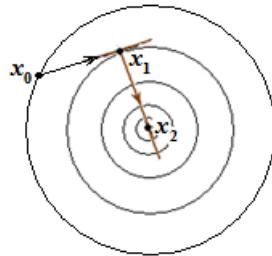
Таким образом, в качестве решения можем принять

$$x^* = (-8,3438, -6,8043, -9,7677, -8,4696, 6,2914)$$

Вычисление невязки (по отношению к исходной системе) для этого решения даст 0.000056.

3.6. Градиентный метод. Внешне метод похож на метод наискорейшего спуска. Им он и является, но по отношению к другому скалярному произведению и связанными с ним понятием ортогональности и расстояния. Этот же метод работает и в более общей ситуации для решения задач многомерной безусловной минимизации. С ним мы познакомимся в параграфе “Метод сопряженных направлений”. Имеем систему линейных уравнений $AX=b$ с невырожденной квадратной матрицей A размера $n \times n$.

Основная идея градиентного метода заключается в таком наблюдении. Если бы линии (поверхности и т.д.) функции $F(\mathbf{x})$ из предыдущего пункта были бы не эллипсами (эллипсоидами и т.д.) а окружностями (сферами, n -мерными сферами), то метод наискорейшего спуска пришел бы в точку максимума за n итераций. Например, при $n=2$ это выглядит так



Первая итерация делается в направлении градиента, который не ортогоначен линиям уровня в новой метрике. А второе направление, сопряженное с первым, сразу дает центр окружности (в старой метрике эти окружности - эллипсы).

Скалярным произведением, превращающим эллипсы в окружности, является билинейная форма, определяемая той же симметрической положительно определенной матрицей A :

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i y_j$$

для $\mathbf{x}=(x_1, \dots, x_n)$, $\mathbf{y}=(y_1, \dots, y_n)$. Вектора, ортогональные в смысле этого скалярного произведения, называются *сопряженными*. Градиентный метод работает также, как метод наискорейшего спуска, но на каждом шаге одномерный максимум ищется не в направлении градиента (который ортогонален поверхности уровня в старой метрике), а в направлении, сопряженном с поверхностью уровня (т.е. ортогональной ей в новой метрике). Если направление известно, то оптимальное t находится по уже известной нам формуле (3.6). Вся проблема в нахождении сопряженных направлений. По существу здесь работает известная в линейной алгебре процедура ортогонализации. После довольно сложных выкладок можно получить формулы:

x_0 – произвольный вектор.

$$\begin{aligned} r_0 &= b - Ax_0 \\ S_1 &= r_0 \end{aligned}$$

S_1 – это первое сопряженное направление. Далее по индукции:

$$\begin{aligned} \alpha_k &= \frac{(r_{k-1}, r_{k-1})}{(S_k, AS_k)} \\ x_k &= x_{k-1} + \alpha_k \cdot S_k \\ r_k &= b - Ax_k \\ \beta_k &= \frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})} \\ S_{k+1} &= r_k + \beta_k \cdot S_k \\ (k &= 1, 2, \dots) \end{aligned} \tag{3.8}$$

В качестве примера рассмотрим ту же систему с расширенной матрицей, что и в предыдущем пункте. Также используем MS Excel.

	A	B	C	D	E	F	G	
1								
2								
3		37,74	-51,08	32,03	80,01	0,02	-957,73	
4		-51,08	200,89	-46,11	-150,29	140,61	1667,24	
5		32,03	-46,11	199,06	214,95	49,52	-3406,91	
6		80,01	-150,29	214,95	344,61	-17,41	-4772,77	
7		0,02	140,61	49,52	-17,41	207,61	12,97	
8								
9	k	0	1	2	3	4	5	
10	x_k	1	-1,417017	-1,052928	0,107435	-2,341687	-8,343839	
11		2	4,497201	1,267391	-3,227295	-5,067543	-6,804306	
12		3	-6,323239	-7,690370	-8,469697	-6,875640	-9,767722	
13		4	-7,975683	-8,101143	-9,745601	-11,074770	-8,469581	
14		5	2,443226	-0,487600	3,478569	4,164931	6,291407	
15	r_k	-1309,543400	166,118005	41,273511	-75,675747	-21,981523	0,000000	
16		1352,987100	-1142,375847	-144,726939	-23,272209	-1,469418	0,000000	
17		-5051,345000	-302,046520	-18,374998	49,405730	-22,582437	0,000000	
18		-6488,442600	166,721847	-61,743525	-26,835737	19,880818	0,000000	
19		-1385,263700	-952,297006	175,832688	-5,652047	8,571500	0,000000	
20	S_k		-1309,5434	123,8552603	44,30443562	-68,41292892	-32,5674997	
21			1352,9871	-1098,71105	-171,6140523	-51,40486764	-9,423626752	
22			-5051,345	-465,068	29,75591216	44,52784938	-15,69235373	
23			-6488,4426	-42,6789198	-62,78794288	-37,12855224	14,13567616	
24			-1385,2637	-997,003463	151,434513	19,17258395	11,53819846	
25	AS_k		-799515,5424	42469,37692	4465,303841	-1499,876132	-119,2710981	
26			1351991,354	-339378,424	-4637,329773	-609,0317336	-7,973016672	
27			-2573176,975	-96498,8717	-2587,973228	2010,892912	-122,5316428	
28			-3605777,989	77718,94223	-1332,833455	-1304,964311	107,8727381	
29			-234581,9984	-383765,144	6929,368363	-397,3157031	46,50870772	
30	(r_k, r_k)	73080407,21	2358515,629	57716,42745	9461,44292	1464,030545	1,80217E-15	
31	(S_k, AS_k)		39595073248	802315634,2	2203702,619	264292,7185	7943,786858	
32	α_k		0,001845694	0,002939636	0,026190661	0,035799106	0,18429882	
33	β_k		0,032272886	0,024471505	0,163929809	0,154736498	1,23096E-18	

Как видим, после пятой итерации получается точное решение.

Тема 4. Вычисление обратной матрицы

Для квадратной матрицы A *обратной* к A называется такая матрица B того же размера, для которой $A \cdot B = B \cdot A = E$. В курсе алгебры доказывается, что если для A такая B существует, то она одна единственная. Поэтому корректно ввести обозначение $B = A^{-1}$. Доказано также, что A^{-1} существует тогда и только тогда, когда матрица A невырождена. Умение эффективно находить обратную матрицу имеет большое значение для решения различных прикладных задач. Ее вычисление с помощью известной формулы $A^{-1} = \frac{1}{|A|} \tilde{A}$, где матрица \tilde{A} составляется из алгебраических дополнений A , к эффективным не относится.

4.1. Вычисление обратной матрицы линейными преобразованиями (Жордана). Для данной A рассмотрим матрицу (A/E) , - к матрице A справа дописывается единичная матрица. Делаем с данной матрицей серию линейных преобразований строк с целью получить на месте матрицы A матрицу E . Тогда на месте матрицы E появится матрица A^{-1} .

Допустимыми преобразованиями являются уже нам знакомые:

- a) перестановки местами любых двух строк;
- b) умножение или деление любой строки на любое ненулевое число;
- c) прибавление (или вычитание) к любой строке любой другой строки с любым коэффициентом.

Продемонстрируем этот простой метод на примере. В качестве матрицы A возьмем матрицу уже рассматривавшейся системы линейных уравнений

$$\begin{array}{cccccc|cccccc} 0,3032 & 8,2869 & -5,7604 & -9,9827 & 8,6236 & & 1 & 0 & 0 & 0 & 0 \\ 3,4133 & -6,0824 & 9,4853 & 9,4349 & 3,1861 & & 0 & 1 & 0 & 0 & 0 \\ -4,2811 & 6,9505 & 0,2575 & -7,5744 & 6,5649 & & 0 & 0 & 1 & 0 & 0 \\ -1,3291 & -5,3959 & -8,0335 & -9,6603 & -8,8668 & & 0 & 0 & 0 & 1 & 0 \\ -2,4302 & 4,2187 & 3,362 & -2,2898 & -1,1729 & & 0 & 0 & 0 & 0 & 1 \end{array}$$

Делим первую строку матрицы на 0,3032. Получим

$$\begin{array}{cccccc|cccccc} 1 & 27.3315 & -18.9987 & -32.9245 & 28.442 & & 3.2982 & 0 & 0 & 0 & 0 \\ 3,4133 & -6,0824 & 9,4853 & 9,4349 & 3,1861 & & 0 & 1 & 0 & 0 & 0 \\ -4,2811 & 6,9505 & 0,2575 & -7,5744 & 6,5649 & & 0 & 0 & 1 & 0 & 0 \\ -1,3291 & -5,3959 & -8,0335 & -9,6603 & -8,8668 & & 0 & 0 & 0 & 1 & 0 \\ -2,4302 & 4,2187 & 3,362 & -2,2898 & -1,1729 & & 0 & 0 & 0 & 0 & 1 \end{array}$$

Теперь от второй строки отнимаем первую с коэффициентом 3,4133. Затем – от третьей первую с коэффициентом -4,2811. Также поступаем с четвертой и пятой строками. Получим

$$\left| \begin{array}{cccccc|ccccc} 1 & 27,3315 & -18,9987 & -32,9245 & 28,442 & & 3,2982 & 0 & 0 & 0 & 0 \\ 0 & -99,373 & 74,3336 & 121,8161 & -93,895 & & -11,2577 & 1 & 0 & 0 & 0 \\ 0 & 123,9594 & -81,0778 & -148,5275 & 128,3279 & & 14,1199 & 0 & 1 & 0 & 0 \\ 0 & 30,9304 & -33,2847 & -53,4203 & 28,9355 & & 4,3836 & 0 & 0 & 1 & 0 \\ 0 & 70,6397 & -42,8086 & -82,3029 & 67,9468 & & 8,0153 & 0 & 0 & 0 & 1 \end{array} \right|$$

Теперь вторую строку делим на -99,373 и обнуляем числа первого столбца. Т.е. действуем точно так же, как в методе Гаусса-Жордана при решении систем линейных уравнений. В итоге получим.

$$\left| \begin{array}{ccccc|ccccc} 1 & 0 & 0 & 0 & 0 & 0.2123 & 0.099 & -0.2523 & 0.0316 & 0.1806 \\ 0 & 1 & 0 & 0 & 0 & 0.048 & -0.0686 & -0.0991 & -0.0674 & 0.1215 \\ 0 & 0 & 1 & 0 & 0 & 0.0363 & 0.1062 & -0.016 & 0.0405 & 0.1592 \\ 0 & 0 & 0 & 1 & 0 & -0.0803 & -0.0902 & 0.0086 & -0.0908 & -0.1017 \\ 0 & 0 & 0 & 0 & 1 & -0.0064 & 0,0289 & 0,1034 & -0,0144 & -0,1345 \end{array} \right|$$

Таким образом, обратной к A является матрица $A^{-1} =$

$$\begin{matrix} 0.2123 & 0.099 & -0.2523 & 0.0316 & 0.1806 \\ 0.048 & -0.0686 & -0.0991 & -0.0674 & 0.1215 \\ 0.0363 & 0.1062 & -0.016 & 0.0405 & 0.1592 \\ -0.0803 & -0.0902 & 0.0086 & -0.0908 & -0.1017 \\ -0.0064 & 0,0289 & 0,1034 & -0,0144 & -0,1345 \end{matrix} \quad (4.1)$$

Для проверки вычислим произведение AA^{-1} . Оно равно

$$\begin{matrix} 0.9995 & -0.0006 & 0.0003 & 0 & -0.0001 \\ -0.001 & 1.0036 & 0.0004 & -0.0006 & -0.0006 \\ 0.0003 & -0.0004 & 1.0009 & -0.0001 & -0.0003 \\ -0.0003 & 0.0005 & -0.0013 & 1.0012 & 0.0005 \\ 0 & -0.0003 & 0.0003 & -0.0002 & 0.9995 \end{matrix}$$

Учитывая, что вычисления велись с точностью до четырех знаков после запятой, это вполне похоже на единичную матрицу.

4.2. Вычисление обратной матрицы методом окаймления. Идея метода заключается в “постепенном” вычислении обратной матрицы. Допустим, что для матрицы A нам известна ей обратная A^{-1} . Добавим теперь к матрице A столбец справа и строку снизу. Это и есть матрица, окаймляющая A :

$$A_o = \begin{pmatrix} A & u \\ v & a \end{pmatrix}$$

Если размер A равен $k \times k$, то здесь u – столбец размера $k \times 1$, v – строка размера $1 \times k$, a – просто число. Оказывается, что обратную к A_o , зная A^{-1} ,

можно вычислить сравнительно небольшим количеством арифметических операций. Это можно сделать по следующей формуле

$$\mathbf{A}_o^{-1} = \begin{pmatrix} \mathbf{A}^{-1} + \frac{\mathbf{A}^{-1} \mathbf{v} \mathbf{v} \mathbf{A}^{-1}}{\alpha} & -\frac{\mathbf{A}^{-1} \mathbf{u}}{\alpha} \\ -\frac{\mathbf{v} \mathbf{A}^{-1}}{\alpha} & -\frac{1}{\alpha} \end{pmatrix} \quad (4.2)$$

где $\alpha = a - \mathbf{v} \mathbf{A}^{-1} \mathbf{u}$. Эту формулу можно вывести, используя правила умножения матриц, записанных в виде блоков – подматриц прямоугольной формы. Если же надо этим способом найти обратную матрицу для матрицы \mathbf{A} размера $n \times n$, то можно действовать следующим образом.

Сначала находим обратную к (a_{11}) - подматрице \mathbf{A}_1 матрицы \mathbf{A} размера 1×1 . Ей будет $\mathbf{A}_1^{-1} = \left(\frac{1}{a_{11}} \right)$. Теперь матрицу \mathbf{A}_1 надо окаймить элементами из второй строки и второго столбца матрицы \mathbf{A} , т.е. рассмотреть

$$\mathbf{A}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

и найти к ней обратную по формуле (4.2). Получим \mathbf{A}_2^{-1} . Теперь окаймляем \mathbf{A}_2 , т.е. рассматриваем

$$\mathbf{A}_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} & \mathbf{A}_2 & a_{13} \\ & a_{31} & a_{32} & a_{33} \end{pmatrix}$$

и снова по формуле (4.2) вычисляем \mathbf{A}_3^{-1} . Аналогично считаются $\mathbf{A}_4^{-1}, \mathbf{A}_5^{-1}$, и т.д. до \mathbf{A}_n^{-1} . Последняя и будет \mathbf{A}^{-1} .

Продемонстрируем метод на матрице из предыдущего параграфа

$$\begin{array}{ccccc} 0,3032 & 8,2869 & -5,7604 & -9,9827 & 8,6236 \\ 3,4133 & -6,0824 & 9,4853 & 9,4349 & 3,1861 \\ -4,2811 & 6,9505 & 0,2575 & -7,5744 & 6,5649 \\ -1,3291 & -5,3959 & -8,0335 & -9,6603 & -8,8668 \\ -2,4302 & 4,2187 & 3,362 & -2,2898 & -1,1729 \end{array}$$

Здесь $\mathbf{A}_1 = (0,3032)$ и $\mathbf{A}_1^{-1} = (3,2982)$. $\mathbf{A}_2 = \begin{pmatrix} 0,3032 & 8,2869 \\ 3,4133 & -6,0824 \end{pmatrix}$. В формуле (4.2)

соответственно $\mathbf{u} = (8,2869)$, $\mathbf{v} = (3,4133)$, $a = -6,0824$. Поэтапно считаем:

$$\mathbf{A}_1^{-1} \mathbf{u} = (27,3319),$$

$$\mathbf{v} \mathbf{A}_1^{-1} = (11,2577),$$

$$\mathbf{v} \mathbf{A}_1^{-1} \mathbf{u} = (93,292),$$

$$\alpha = a - \mathbf{v} \mathbf{A}_1^{-1} \mathbf{u} = -99,3744,$$

$$\mathbf{B} = (\mathbf{A}_1^{-1} \mathbf{u})(\mathbf{v} \mathbf{A}_1^{-1}) = (307,6943),$$

$$\mathbf{B}/\alpha = (-3,0963),$$

$$\mathbf{A}_1^{-1} + \mathbf{B}/\alpha = (0,2019),$$

$$\begin{aligned} -A_1^{-1} u/\alpha &= (0,275), \\ -v A_1^{-1} / \alpha &= (0,1133), \\ 1/\alpha &= -0,0101. \end{aligned}$$

Теперь составляем обратную матрицу $A_2^{-1} = \begin{pmatrix} 0,2019 & 0,275 \\ 0,1133 & -0,0101 \end{pmatrix}$ и делаем проверку, вычисляя $A_2 A_2^{-1} = \begin{pmatrix} 1,0001 & -0,0003 \\ 0 & 1,0001 \end{pmatrix}$. Т.к. вычисления велись с точностью до 4-х знаков после запятой, результат вполне удовлетворительный.

Вычисляем A_3^{-1} . Теперь $u = \begin{pmatrix} -5,7604 \\ 9,4853 \end{pmatrix}$, $v = (-4,2811, 6,9505)$, $a = 0,2575$.

$$A_2^{-1} u = \begin{pmatrix} 1,4454 \\ -0,7485 \end{pmatrix},$$

$$v A_2^{-1} u = (-11,3904),$$

$$v A_2^{-1} = (-0,0769 \quad -1,2475),$$

$$\alpha = a - v A_2^{-1} u = 11,6479,$$

$$B = (A_2^{-1} u)(v A_2^{-1}) = \begin{pmatrix} -0,1112 & -1,8031 \\ 0,0576 & 0,9338 \end{pmatrix},$$

$$B/\alpha = \begin{pmatrix} -0,0095 & -0,1548 \\ 0,0049 & 0,0802 \end{pmatrix},$$

$$A_2^{-1} + B/\alpha = \begin{pmatrix} 0,1924 & 0,1202 \\ 0,1182 & 0,0701 \end{pmatrix},$$

$$-A_2^{-1} u/\alpha = \begin{pmatrix} -0,1241 \\ 0,0643 \end{pmatrix},$$

$$-v A_2^{-1} / \alpha = (0,0066 \quad 0,1071),$$

$$1/\alpha = 0,0859.$$

Составляем обратную матрицу $A_3^{-1} = \begin{pmatrix} 0,1924 & 0,1202 & -0,1241 \\ 0,1182 & 0,0701 & 0,0643 \\ 0,0066 & 0,1071 & 0,0859 \end{pmatrix}$ и делаем проверку, вычисляя $A_3 A_3^{-1} = \begin{pmatrix} 0,9998 & 0,0004 & 0,0004 \\ 0,0004 & 0,9998 & 0,0001 \\ -0,0004 & 0,0002 & 1,0003 \end{pmatrix}$.

Подобным образом вычисляются

$$A_4^{-1} = \begin{pmatrix} 0,204 & 0,1373 & -0,1134 & 0,0123 \\ 0,0422 & -0,0422 & -0,0057 & -0,0803 \\ 0,0288 & 0,1399 & 0,1064 & 0,0235 \\ -0,0756 & -0,1117 & -0,0697 & -0,0799 \end{pmatrix} \text{ и}$$

$$A_5^{-1} = \begin{pmatrix} 0,2126 & 0,0986 & -0,2523 & 0,0316 & 0,1806 \\ 0,048 & -0,0683 & -0,0992 & -0,0673 & 0,1216 \\ 0,0364 & 0,1058 & -0,016 & 0,0405 & 0,1592 \\ -0,0805 & -0,0899 & 0,0086 & -0,0908 & -0,1018 \\ -0,0064 & 0,0288 & 0,1034 & -0,0143 & -0,1345 \end{pmatrix} = A^{-1}.$$

Метод окаймления не производит впечатление более удобного метода, чем метод Жордана. Если проследить количество требуемых для него арифметических операций, то оно окажется примерно таким же, как в методе Жордана – около $4n^3$ умножений. Тем не менее, возможна ситуация, когда дана матрица и известна ей обратная, а требуется найти обратную для окаймляющей. В этом случае метод окаймления может быть особенно полезен.

Тема 5. Метод наименьших квадратов и первая трансформация Гаусса

5.1. Метод наименьших квадратов. Допустим, что имеется множество функций вида

$$\{y = \varphi(x, C_1, C_2, \dots, C_m) \mid C_1, C_2, \dots, C_m \in \mathbf{R}\}.$$

Здесь φ – функция от $(m+1)$ -й переменной задаваемая, как правило, каким-то алгебраическим выражением. Как следствие, эта функция определена при всех значениях своих аргументов и имеет по всем аргументам непрерывные частные производные. При фиксировании аргументов (параметров) C_1, C_2, \dots, C_m получается функция только от x . О множестве (классе, семействе) всех этих функций идет речь. Примерами таких семейств являются $\{y=kx+b\}$ – класс всех линейных функций, $\{y=Ca^x\}$ – класс всех показательных функций, $\{y=A \cdot \sin \omega x + B \cdot \cos \omega x\}$ – важный класс тригонометрических функций.

Естественно, очень часто областью определения $\varphi(x, C_1, C_2, \dots, C_m)$ является не все множество \mathbf{R}^{m+1} , а его часть вида $[a, b] \times D$, где $D \subseteq \mathbf{R}^m$.

Допустим, наконец, что имеется некоторая функция

$$f: \mathbf{R} \rightarrow \mathbf{R} \text{ или } f: [a, b] \rightarrow \mathbf{R}.$$

При этом точно, что это за функция, неизвестно. Имеются только ее значения в некоторых точках $x_1, \dots, x_n \in \mathbf{R}$ или $x_1, \dots, x_n \in [a, b]$. Требуется по ее значениям $y_1 = f(x_1), \dots, y_n = f(x_n)$ в этих точках найти функцию данного семейства, лучше всего похожую на функцию f . Это типичная ситуация, когда числа y_1, \dots, y_n получены в результате измерений (температуры, атмосферного давления, численности населения, финансовых доходов и пр.) и надо как-то угадать закономерность изменения этих чисел.

Критерий этой “похожести” может быть самый разный. В методе наименьших квадратов наиболее подходящей считается та, для которой достигается минимум величины

$$F(C_1, \dots, C_m) = (y_1 - \varphi(x_1, C_1, \dots, C_m))^2 + \dots + (y_n - \varphi(x_n, C_1, \dots, C_m))^2 \quad (5.1)$$

по всем возможным значениям параметров C_1, \dots, C_m .

Поиск нужных C_1, \dots, C_m также может делаться разными способами. Например, по известной теореме математического анализа вектор (C_1, \dots, C_m) должен быть точкой \mathbf{R}^m , в которой градиент функции F является нулевым вектором, т.е. в которой все частные производные этой функции равны нулю. Другими словами, C_1, \dots, C_m ищутся как решение системы

$$\begin{cases} \frac{\partial F}{\partial C_1}(C_1, \dots, C_m) = 0 \\ \dots \\ \frac{\partial F}{\partial C_m}(C_1, \dots, C_m) = 0 \end{cases}$$

m уравнений с m неизвестными C_1, \dots, C_m .

5.2. Первая трансформация Гаусса. Предположим, что семейство функций φ предыдущего пункта имеет вид

$$\varphi(x, C_0, C_1, \dots, C_m) = C_0 + C_1 x + C_2 x^2 + \dots + C_m x^m$$

и y_1, \dots, y_n – значения данной, но неизвестной функции в узлах сетки x_1, \dots, x_n . Требуется методом наименьших квадратов найти C_0, C_1, \dots, C_m , при которых значения $\varphi(x_i, C_0, C_1, \dots, C_m)$ наилучшим образом приближаются к y_n ($i = 1, \dots, n$). Другими словами, ищется многочлен степени не выше m наилучшим образом приближающий неизвестную функцию.

Составим систему из n линейных уравнений с неизвестными C_0, C_1, \dots, C_m

$$\begin{cases} C_0 + C_1 x_1 + C_2 x_1^2 + \dots + C_m x_1^m = y_1 \\ C_0 + C_1 x_2 + C_2 x_2^2 + \dots + C_m x_2^m = y_2 \\ \dots \\ C_0 + C_1 x_n + C_2 x_n^2 + \dots + C_m x_n^m = y_n \end{cases} \quad (5.2)$$

Без всяких частных производных понятно, что минимум суммы квадратов (5.1) – это минимум невязки системы (5.2). Если $m+1 \geq n$, то система имеет точное решение. Но в обычной ситуации $m+1 < n$ и система (5.2) является, как говорят, *переопределенной*. Поэтому бессмысленно пытаться решать ее методом Гаусса или каким-то другим из раздела 2. Система просто не имеет ни одного точного решения. Именно такой случай дальше рассматривается.

На практике такие системы могут появиться как результат многократного наблюдения за каким-нибудь явлением, определенные характеристики которого неизвестны. Каждое наблюдение дает одно уравнение относительно этих характеристик. Но вместе они не совместны из-за неизбежных округлений и неточностей измерений. В этом случае возникает задача отыскания неточного, но в некотором смысле наилучшего решения. Для линейных систем критерием – какое решение лучше – является минимум

невязки. Таким образом, для переопределенной системы требуется найти решение с минимальной невязкой.

Еще раз почеркнем, что неизвестными в системе (5.2) являются C_0, C_1, \dots, C_m , а все x_i^k и y_i – известные числа. Поэтому расширенной матрицей системы (5.2) является

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m & y_1 \\ 1 & x_2 & x_2^2 & \dots & x_2^m & y_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^m & y_n \end{pmatrix} \quad (5.3)$$

Метод решения переопределенных систем с помощью первой трансформации Гаусса заключается в следующем.

Пусть дана система линейных уравнений

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (5.4)$$

в которой $m > n$. При этом пусть столбцы матрицы $A = (a_{ij})_{(i=1,\dots,m; j=1,\dots,n)}$ линейно независимы. Пусть $(A \mid b)$ – расширенная матрица системы (5.4). В матричном виде система (5.4) записывается как

$$AX = b \quad (5.5)$$

Первой трансформацией Гаусса является переход от системы (5.5) к системе

$$(A^T A)X = A^T b \quad (5.5)$$

где A^T – результат транспонирования A - матрица размера $n \times m$. Матрица $(A^T A)$ системы (5.6) имеет размер $n \times n$ и является невырожденной. Поэтому система (5.6) имеет единственное решение. Справедлива

Теорема 5.1. *Решением системы (5.6) является вектор, в котором достигается минимум вектора невязки системы (5.5).*

5.3. Пример. Данна неизвестная функция $f: [-3,5] \rightarrow R$. и таблица ее значений в 6-ти точках промежутка $[-3,5]$.

x_i	y_i
-5	-3,8081
-3,4	2,6682
-1,8	-1,6816
-0,2	0,8567
1,4	-0,2010
3	-0,2794

Требуется найти многочлен степени не выше 3, наилучшим образом приближающий эту функцию в сетке $x_0 < x_1 < x_2 < x_3 < x_4 < x_5$.

Решение. Ищем многочлен в виде

$$\varphi(x, C_0, C_1, C_2, C_3) = C_0 + C_1 x + C_2 x^2 + C_3 x^3.$$

Составляем матрицу (5.3)

$$\begin{pmatrix} 1 & -5 & 25 & -125 & -3,8081 \\ 1 & -3,4 & 11,56 & -39,304 & 2,6682 \\ 1 & -1,8 & 3,24 & -5,832 & -1,6816 \\ 1 & -0,2 & 0,04 & -0,008 & 0,8567 \\ 1 & 1,4 & 1,96 & 2,744 & -0,2010 \\ 1 & 3 & 9 & 27 & -0,2794 \end{pmatrix}$$

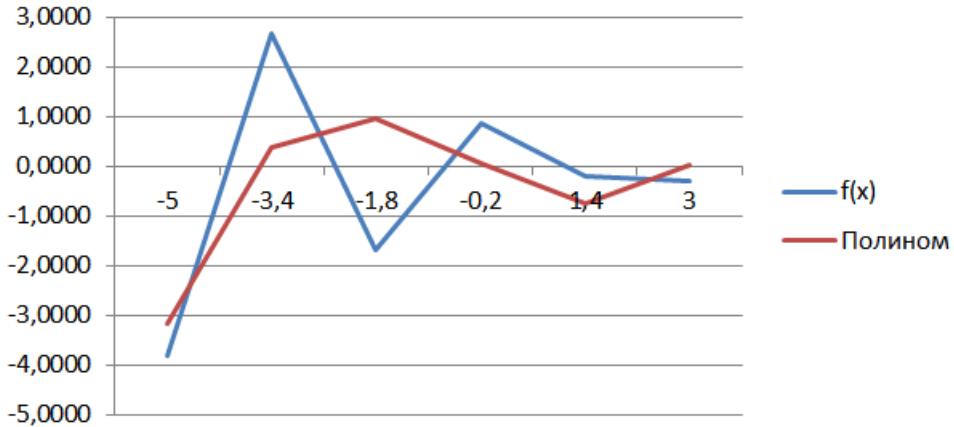
$$\text{Умножаем ее слева на } A^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ -5 & -3,4 & -1,8 & -0,2 & 1,4 & 3 \\ 25 & 11,56 & 3,24 & 0,04 & 1,96 & 9 \\ -125 & -39,304 & -5,832 & -0,008 & 2,744 & 27 \end{pmatrix}$$

Получим матрицу

$$\left| \begin{array}{cccc|c} 6 & -6 & 50.8 & -140.4 & -2.4452 \\ -6 & 50.8 & -140.4 & 853.9744 & 11.70456 \\ 50.8 & -140.4 & 853.9744 & -3349.872 & -72.68078 \\ -140.4 & 853.9744 & -3349.872 & 17940.34624 & 372.84646 \end{array} \right|$$

системы с неизвестными C_0, C_1, C_2, C_3 . Любым методом находим решение

$C_0 = -0.05994, C_1 = -0.68565, C_2 = 0.05023, C_3 = 0.06233$. Таким образом, искомый многочлен есть: $-0.05994 - 0.68565x + 0.05023x^2 + 0.06233x^3$. Ниже на графике показано его поведение в сравнении с функцией $f(x)$.



Тема 6. Вычисление собственных чисел матриц

6.0. Собственные числа матриц. Пусть A – квадратная матрица порядка n . Ненулевой вектор-столбец X размера $n \times 1$ называется **собственным вектором** матрицы A , если для некоторого числа λ выполняется матричное равенство $A \cdot X = \lambda \cdot X$. Число λ при этом называется **собственным числом** матрицы A . Говорят также, что собственное число λ **принадлежит** вектору X , а собственный вектор X **принадлежит** числу λ . В курсе алгебры доказывается, что число λ тогда и только тогда является собственным числом матрицы A , когда выполняется равенство $|A - \lambda E| = 0$. Здесь $| \cdot |$ – определитель матрицы, E – единичная матрица порядка n . Та-

ким образом выражение

$$|A - \lambda E| = 0 \quad (6.1)$$

является уравнением, с помощью которого можно пытаться находить собственные числа матрицы A . Левая часть этого уравнения, если вычислить определитель, превратится в многочлен степени n относительно λ . Он называется **характеристическим многочленом** матрицы, а уравнение (6.1) – характеристическим уравнением. Это объясняет, почему у матрицы A не может быть больше, чем n , собственных чисел. Если же мы имеем дело с матрицей из комплексных чисел и ищем все комплексные собственные числа, то в силу алгебраической замкнутости поля комплексных чисел (основная теорема алгебры) этих чисел ровно n .

Собственные числа матрицы – одна из важнейших ее характеристик. Они имеют колоссальное значение, как в теоретических, так и в практических приложениях алгебры. Поэтому нахождение собственных чисел – одна из важнейших задач вычислительной математики. Эта задача благодаря уравнению (6.1) не кажется особенно сложной. Действительно, если нам удается найти коэффициенты характеристического многочлена, то численно решить уравнение (6.1) несложно. К сожалению, вычисление определителя $|A - \lambda E|$, в котором λ – не конкретное число, а переменная – задача совсем непростая. Потому основную роль в алгоритмах нахождения собственных чисел играют различные итерационные методы, никак не использующие уравнение (6.1).

Тем не менее, для простых матриц уравнение (6.1) весьма полезно. К таким простым матрицам, в частности, относятся **треугольные матрицы**, т.е. матрицы вида

$$A = \begin{pmatrix} a_{11} & a_{11} & a_{11} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & \dots & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Легко видеть, что уравнение (6.1) для такой матрицы превращается в

$$(\lambda - a_{11})(\lambda - a_{22}) \dots (\lambda - a_{nn}) = 0$$

Поэтому собственными числами матрицы являются числа $a_{11}, a_{22}, \dots, a_{nn}$ – все ее диагональные элементы.

Простым случаем является также случай, когда матрица A имеет **блочно-диагональный** вид

где по диагонали стоят блоки размеров 1×1 или 2×2 . Блоки 1×1 – это “готовые” собственные числа. Остальные – это собственные числа блоков 2×2 , которые находятся как решение квадратного уравнения. Эти уравнения – характеристические уравнения соответствующих блоков, и для их решения вполне годится обычная школьная формула для поиска корней квадратных уравнений.

Есть еще один важный факт из алгебры, который используется во всех итерационных методах. Две квадратные матрицы A и B одного размера называются *сопряженными*, если существует такая невырожденная квадратная матрица T того же размера, что $B = T^{-1} \cdot A \cdot T$. Легко доказывается, что у сопряженных матриц одни и те же собственные числа. Это дает основную идею вычисления собственных чисел. Надо от данной матрицы A путем соответствующего подбора матриц T переходить к сопряженной A и делать это многократно, постепенно приближаясь к треугольному или блочно-треугольному виду.

6.1. Собственные числа симметрических матриц. Допустим, что имеется квадратная симметрическая матрица A порядка n , состоящая из действительных чисел. В алгебре доказано, что для симметрических, т.е. симметричных относительно главной диагонали вещественных матриц все их собственные числа являются вещественными. Эти собственные числа весьма эффективно могут быть найдены с помощью следующего метода Якоби.

Основной шаг. Вне диагонали матрицы выбираем любой ненулевой элемент a_{ij} (для ускорения сходимости рекомендуется брать наибольший по абсолютной величине). Тогда $i \neq j$. Строим матрицу вращения T_{ij} , у которой

$$c = \cos \theta, s = \sin \theta$$

а угол θ находится по формуле

$$\theta = \frac{1}{2} \operatorname{arctg} \frac{2a_{ij}}{a_{ii}-a_{jj}} \text{ (если } a_{ii} = a_{jj}, \text{ то полагаем } \theta = \pi/4).$$

Пересчитываем матрицу A , полагая

$$A = T_{ij}^{-1} A T_{ij}$$

Основной шаг повторяем до тех пор, пока элементы матрицы вне диагонали не станут достаточно маленькими.

Проделывая вычисления по указанной формуле мы можем воспользоваться замечательной особенностью матриц вращения – обратной в T_{ij} является в T_{ij}^T . Другими словами, для того, чтобы найти обратную к матрице вращения, надо ее просто транспонировать. Таким свойством обладает класс всех ортогональных матриц – матрицы вращения являются их частным случаем.

Обоснованием алгоритма является тот факт, что:

1. В новой матрице $a_{ij} = a_{ji} = 0$.
2. Сумма квадратов всех недиагональных и отличных от a_{ij} и a_{ji} элементов сохраняется.
3. Сумма квадратов **всех** элементов матрицы сохраняется.

Благодаря этим трем обстоятельствам и происходит постепенная “перекачка” элементов матрицы на диагональ. Каждый шаг сумму квадратов диагональных элементов увеличивает на $2a_{ij}^2$, а сумму квадратов остальных элементов - соответственно уменьшает.

Пример применения метода Якоби. Пусть дана матрица

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 2 & -3 & -4 \\ 3 & -3 & 4 & 3 \\ 4 & -4 & 3 & 5 \end{pmatrix}$$

и требуется найти ее собственные числа. Делаем первый шаг. Выбираем в матрице не на диагонали любой наибольший по абсолютной величине элемент. Пусть это $a_{24} = -4$. Тогда $\theta = \frac{1}{2} \operatorname{arctg} \frac{2(-4)}{5-2} = \frac{1}{2} \operatorname{arctg} (-\frac{8}{3}) = -0,60601$. Значит, $c = \cos(-0,60601) = 0,82193$, $s = \sin(-0,60601) = -0,56959$. Соответственно матрица вращения T_{24} и ей обратная есть

$$T_{24} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0,82193 & 0 & -0,56959 \\ 0 & 0 & 1 & 0 \\ 0 & 0,56959 & 0 & 0,82193 \end{pmatrix}, T_{24}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0,82193 & 0 & 0,56959 \\ 0 & 0 & 1 & 0 \\ 0 & -0,56959 & 0 & 0,82193 \end{pmatrix}$$

Вычисляя $T_{24}^{-1} A T_{24}$ получим

$$\begin{pmatrix} 1 & 3.92222 & 3 & 2.14854 \\ 3.92222 & -0.772 & -0.75702 & 0 \\ 3 & -0.75702 & 4 & 4.17456 \\ 2.14854 & 0 & 4.17456 & 7.77202 \end{pmatrix}$$

Это новая матрица A , но с теми же самыми собственными числами. Повторяем такой шаг для $a_{12} = 3.92222$ (хотя это число и не самое большое по модулю среди не диагональных, никакой беды не произойдет). Получим

$$\begin{pmatrix} 4,1351 & 0 & 1,8708 & 2,14854 \\ 0 & -3,9071 & -2,4644 & -1,34145 \\ 3 & -2,4644 & 4 & 4,17456 \\ 2,14854 & -1,34145 & 4,17456 & 7,77202 \end{pmatrix}$$

Стоит обратить внимание на то, что элемент a_{24} , который на предыдущие шаге был обнулен, опять не равен нулю. Тем не менее, сумма квадратов элементов вне диагонали уменьшилась. Поэтому процесс надо повторять, до тех пор, пока вся матрица не “перетечет” на диагональ.

Примерно после 17-ти шагов получим

$$\begin{pmatrix} 3,5153 & 0 & 0 & 0 \\ 0 & -4,6472 & 0 & 0 \\ 0 & 0 & 1,4757 & 0 \\ 0 & 0 & 0 & 11,6562 \end{pmatrix}$$

На диагонали этой матрицы выстроились все собственные числа исходной матрицы A .

В приложении “Лабораторная 6” приводится макрос для MS Excel, с помощью которого проводились вычисления.

6.2. Число обусловленности матрицы. Пусть A – квадратная матрица порядка n , состоящая из действительных чисел. С матрицей связываем линейный оператор, который действует на вектора $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ по правилу

$$x \rightarrow Ax$$

Здесь Ax означает матричное произведение матрицы A на вектор x , записанный в виде столбца

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Точнее надо было бы использовать знак транспонирования и писать

$$x \rightarrow (Ax)^T$$

но постараемся обойтись без несущественных формальностей. Заметим, что матрица Ax – это левая часть системы линейных уравнений (3.2), которые мы рассматривали в части 3. Теперь же мы смотрим на Ax как на функцию $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

В п 3.4 мы рассматривали понятие нормы вектора и нормы матрицы. Напомним, что это такое. Норма вектора – это его числовая характеристика, которую можно понимать как длину вектора. Она должна подчиняться определенным условиям – аксиомам нормы. В пространстве R^n особенно популярны нормы

- 1) $\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}$ – максимум модулей координат вектора.
- 2) $\|x\|_1 = |x_1| + \dots + |x_n|$ – сумма модулей координат.
- 3) $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$ – евклидова норма.

Любая норма $\|x\|$ векторов в R^n (а их существует бесконечно много) определяет норму любой квадратной матрицы A порядка n по формуле

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

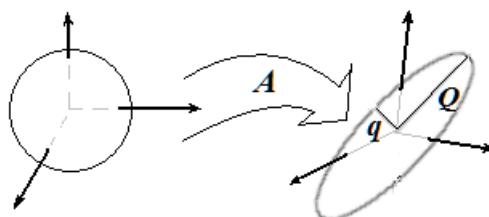
Легко видеть, что $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$. Из этого вытекает одно из важнейших свойств нормы: $\|Ax\| \leq \|A\| \cdot \|x\|$.

Для произвольно взятой нормы $\|x\|$ пусть

$$Q = \|A\| = \max_{\|x\|=1} \|Ax\|$$

$$q = \min_{\|x\|=1} \|Ax\|$$

Число Q (оно же – норма матрицы) показывает, во сколько раз максимально может увеличиться длина вектора единичной сферы под действием преобразования $x \rightarrow Ax$. Величина q , напротив, показывает, во сколько раз длина этих образов может уменьшиться.



Если матрица A невырождена (т.е. имеет обратную), то $q \neq 0$. В этом случае $0 < q \leq Q < \infty$ и величина

$$\text{cond}(A) = \frac{Q}{q}$$

называется **числом обусловленности** матрицы A . Подчеркнем, что число обусловленности зависит от выбранной нормы. Тем не менее, при любой норме оно является важной характеристикой матрицы. Матрица с вычислительной точки зрения тем лучше, чем ближе к единице ее число обусловленности. Особенно это важно для евклидовой нормы. Близость ее к

единице означает отсутствие “деформации” пространства под действием преобразования $x \rightarrow Ax$. Для матриц вращения и вообще для ортогональных матриц это свойство всегда выполняется для евклидовой нормы.

Теорема 6.1. *Если A – квадратная невырожденная матрица и A^{-1} – её обратная, то число обусловленности относительно произвольной нормы может быть найдено по формуле*

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|.$$

Матрицы с большим числом обусловленности называются **плохо обусловленными**. Примером такой матрицы является

$$\begin{pmatrix} 29.063 & 60.2198 & 116.0227 \\ -0.7329 & -2.2905 & -5.6796 \\ 17.9221 & 37.9025 & 74.2386 \end{pmatrix} \quad (6.2)$$

Используя теорему 6.1 найдем число обусловленности матрицы (6.2) в нормах $\|\cdot\|_\infty$ и $\|\cdot\|_1$. Сначала надо вычислить обратную матрицу. Она равна

$$\begin{pmatrix} 44.6382093989406 & -72.1309800323837 & -75.2805237682607 \\ -46.763514514135 & 77.2067125096874 & 78.9903427568564 \\ 13.0988954021661 & -22.0045742196114 & -22.141398023597 \end{pmatrix}$$

$\|A\|_\infty$ как максимум сумм модулей по строкам A равна 205.3055. Для обратной $\|A^{-1}\|_\infty = 202.960569780697$. Окончательно, $\text{cond}(A) = 41668.9$.

Если использовать другую норму $\|\cdot\|_1$, то получится не намного лучше $\text{cond}(A) = 34566.4$. В этих нормах единичная сфера – это многомерный ромб или куб. Наиболее же естественной будет сфера в евклидовой норме. А для поиска числа обусловленности здесь помогают собственные числа матрицы.

Теорема 6.2. *Если A – квадратная невырожденная матрица, то относительно евклидовой нормы*

$$\text{cond}(A) = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}},$$

где λ_{\max} и λ_{\min} – максимальное и минимальное собственные числа матрицы $A^T A$.

Если вычислить $A^T A$, а затем – собственные числа этой симметрической матрицы (например, методом Якоби), получим три числа 25238.406506, 1.154912, 0.000082. Делим большее на меньшее и извлекаем

квадратный корень. Получим $\|A\|_2=17579.1$. Во всех нормах матрица оказывается плохо обусловленной.

Число обусловленности влияет на погрешности различных вычислений, связанных с этой матрицей. Примером такой связи является

Теорема 6.3. Рассматривается система линейных уравнений $Ax = b$ с невырожденной квадратной матрицей A . Пусть относительной погрешности $\delta(b)$ нормы вектора свободных членов соответствует относительная погрешность $\delta(x)$ решения этой системы. Тогда

$$\delta(x) \leq \text{cond}(A)\delta(b)$$

Для матрицы (6.2) с ее огромным числом обусловленности нет, как видим, никакой гарантии устойчивости решения системы $Ax = b$. И действительно, возьмем, к примеру в качестве свободных членов вектор

$$b = \begin{pmatrix} 11.223 \\ -45.4526 \\ 33.2223 \end{pmatrix}$$

Если решить систему $Ax = b$ методом Гаусса-Жордана, делая округления до четвертого знака после запятой, то получим

$$x = (1033.711, -1153.3613, 339.6905)$$

Норма вектора невязки при этом получится подозрительно большой $\|r\| = 12.54$. Если пересчитать решение без округлений с максимально возможной точностью, получим

$$x = (1278.52306231796, -1409.83188023886, 411.585845713479)$$

с невязкой $\|r\| = 2.25 \cdot 10^{-11}$. Т.е. приказалось бы небольшом изменении точности вычислений невязка изменилась почти в триллион раз! Такое поведение системы никак не назовешь хорошим. Ведь вести вычисления с какими-то разумными округлениями - обычное дело. А здесь такое округление дает колоссальную ошибку.

Сделаем еще один эксперимент. Изменим немного свободные члены нашей системы и опять найдем решение, используя максимальную точность вычислений. Пусть у новой системы свободные члены есть

$$b = \begin{pmatrix} 11.3 \\ -45.4 \\ 33.3 \end{pmatrix}$$

Разница между старым и новым столбцами свободных членов составит по норме 0.121. Решением же новой системы будет

$$x = (1272.31681819517, -1403.23404814622, 409.716633429055)$$

Норма разности со старым стала 9.249. Это в 76 раз больше разницы между столбцами свободных членов! Таким образом, наша матрица обладает еще одной неприятной особенностью. Небольшие ошибки в свободных членах очень сильно влияют на ее решения.

6.3. Собственное число неотрицательной матрицы. Пусть A – квадратная матрица порядка n , состоящая из неотрицательных действительных чисел. Такие матрицы называются **неотрицательными**. Оказывается, что неотрицательные матрицы всегда имеют неотрицательное собственное число и у этого числа имеются собственные векторы, состоящие из неотрицательных вещественных чисел. Если, кроме того, матрица A *не разложима*, то такое собственное число является единственным положительным собственным числом (другие собственные числа – комплексны) и оно имеет максимальный модуль среди всех остальных собственных чисел этой матрицы. Эти факты были доказаны Перроном и Фробениусом еще в самом начале 20-го века. **Разложимость** матрицы означает, что перестановкой ее строк с одновременной и такой же перестановкой столбцов матрицу можно привести к матрице вида

$$\begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & A_2 \end{pmatrix}$$

где A_1 и A_2 – непустые матрицы. Такие перестановки не меняют собственных чисел матрицы. Поэтому в случае разложимости задача сводится к поиску собственных чисел матриц меньшего размера.

Существует довольно простой алгоритм нахождения положительного собственного числа и положительного собственного вектора неразложимой неотрицательной матрицы.

Алгоритм. Собственный вектор \mathbf{x}^* этой матрицы вычисляется как предел последовательности векторов $\mathbf{x}^{(m)}$ ($m = 0, 1, 2, \dots$), где

$$\mathbf{x}^{(0)} = (1, 1, \dots, 1)^T;$$

Цикл по $m = 0, 1, 2, \dots$

$$\mathbf{y} = (y_1, \dots, y_n)^T = \mathbf{A}\mathbf{x}^{(m)};$$

$$s = (y_1 + \dots + y_n)/n; \quad (\text{среднее арифметическое чисел столбца } \mathbf{y})$$

$$\mathbf{x}^{(m+1)} = \mathbf{y}/s; \quad (\text{делает среднее арифметическое столбца } \mathbf{x}^{(m+1)} \text{ равным 1})$$

Конец цикла

Цикл надо продолжать, пока все отношения $y_1/x_1^{(m)}, y_2/x_2^{(m)}, \dots, y_n/x_n^{(m)}$, не станут одинаковыми с некоторой удовлетворительной точностью (например, до 5 знаков после запятой). Эти отношения и есть искомое собственное число. Здесь $x_1^{(m)}, x_2^{(m)}, \dots, x_n^{(m)}$, – координаты вектора $\mathbf{x}^{(m)}$.

С практической точки зрения положительное собственное число неотрицательной матрицы может быть полезно для поиска сбалансированного общего мнения разных экспертов в так-называемом методе парных сравнений. В качестве иллюстрации таких парных сравнений рассмотрим итоговую таблицу какого-нибудь спортивного соревнования. Соревноваться должны несколько участников, причем каждый должен встретиться с каждым, результаты чего должны быть отражены в таблице.

В качестве примера возьмем итоговую таблицу шахматного турнира, проходившего в Голландии в 1938 г. с участием великих шахматистов того времени (Шахматный словарь. Москва, Физкультура и спорт, 1964).

		1	2	3	4	5	6	7	8	Очки	Место
1	Керес		1,5	1	1	1,5	1	1,5	1	8,5	1-2
2	Файн	0,5		1,5	2	1	1	1	1,5	8,5	1-2
3	Ботвинник	1	0,5		1,5	1,5	0,5	1,5	1	7,5	3
4	Алехин	1	0	0,5		1	1,5	1,5	1,5	7	4-6
5	Решевский	0,5	1	0,5	1		1,5	1	1,5	7	4-6
6	Эйве	1	1	1,5	0,5	0,5		1	1,5	7	4-6
7	Капабланка	0,5	1	0,5	0,5	1	1		1,5	6	7
8	Флор	1	0,5	1	0,5	0,5	0,5	0,5		4,5	8

Пусть A – матрица размера 8×8 с числами – очками. Каждая клетка таблицы, т.е. этой матрицы – результат парного сравнения участников между собой. Например, Керес против Файна имеет счет 1.5 на 0.5 в свою пользу. В итоге Керес и Файн делят 1-е и 2-е места, а Алехин Решевский и Эйве делят 4-е, 5-е и 6-е места. Вроде бы все логично и справедливо. Но посмотрим, как конкретно сыграли, например Керес и Файн. Между собой – в пользу Кереса. Это может служить поводом именно Кересу отдать первое место. Но, в то же время Файн против других наиболее успешных игроков – Ботвинника и Алехина – набрал 3.5 очков, а Керес – только 2. Попытка определить – кто же все таки из них лучше, опирается на следующую логику.

В данной таблице приводятся результаты игр – выигрыши, ничьи, проигрыши. Итоговые очки – механическая сумма этих результатов. Но если мы согласимся с тем, что выигрыш у более сильного игрока стоит больше, чем выигрыш у слабого, то следует эти результаты суммировать с некоторыми весовыми коэффициентами. А откуда их взять? В данной таблице эти коэффициенты равны единице. Но раз у Кереса результат оказался почти в два раза лучше, чем у последнего – Флора, то и коэффициент за

выигрыш у Кереса должен быть в два раза больше, чем коэффициент выигрыша у Флора. Поэтому возникает мысль пересчитать результаты заново с коэффициентами, пропорциональными набранным очкам в первом подсчете.

Такими коэффициентами являются числа

$$1,2143; 1,2143; 1,0714; 1; 1; 1; 0,8571; 0,6429 \quad (6.3)$$

Они, как и единицы, дают в сумме 8 и в точности пропорциональны набранным очкам. Пересчитаем результаты заново с этими коэффициентами. Получим, например, у Кереса:

$$0*1,2143 + 1,5*1,2143 + 1*1,0714 + 1*1 + 1,5*1 + 1*1 + 1,5*0,8571 + 1*0,6429 = 8,3214$$

а у Флора

$$1*1,2143 + 0,5*1,2143 + 1*1,0714 + 0,5*1 + 0,5*1 + 0,5*0,8571 + 0*0,6429 = 4,8214$$

Разница между Кересом и Флором стала меньше. Значит, с точки зрения новых результатов подсчета, коэффициенты в строке (1) не совсем верны. Было бы идеально, если бы они оказались пропорциональны новым очкам. А это означало бы, что строка коэффициентов (1) образовывала бы собственный вектор матрицы A . А коэффициент пропорциональности был бы собственным числом матрицы A . Ведь то, как мы считаем очки с коэффициентами – ни что иное, как умножение матрицы A на столбец этих коэффициентов!

Алгоритм в начале этого параграфа дает способ найти такие коэффициенты и им соответствующее собственное число. Он говорит, что надо еще раз повторить пересчет коэффициентов используя вновь полученные очки. Сосчитать уже в третий раз очки с новыми коэффициентами. Затем продолжать в том же духе до тех пор, пока очки и коэффициенты не перестанут меняться. С помощью MS Excel это делается очень просто.

Вот какой результат получится в нашем примере

	0	1	2	...	10								
	Кф	Очки	λ	Кф	Очки	λ	Кф	Очки	λ	...	Кф	Очки	λ
1	8,5	8,5	1,21	8,32	6,85	1,23	8,29	6,758	...	1,21839	8,2979	6,81054	
1	8,5	8,5	1,21	8,04	6,62	1,18	8,05	6,798	...	1,18353	8,06051	6,81054	
1	7,5	7,5	1,07	7,25	6,77	1,07	7,23	6,771	...	1,06414	7,24734	6,81054	
1	7	7	1	6,5	6,5	0,96	6,61	6,904	...	0,96954	6,60308	6,81054	
1	7	7	1	6,68	6,68	0,98	6,73	6,837	...	0,98795	6,72847	6,81054	
1	7	7	1	6,86	6,86	1,01	6,91	6,836	...	1,01209	6,89287	6,81054	
1	6	6	0,86	5,82	6,79	0,86	5,87	6,844	...	0,86109	5,86451	6,81054	
1	4,5	4,5	0,64	4,82	7,5	0,71	4,79	6,744	...	0,70327	4,78963	6,81054	

После 10-й итерации коэффициенты и очки станут практически пропорциональными. Коэффициент пропорциональности λ будет собственным числом. Очки в 10-м блоке - полностью сбалансированные результаты участников. И, как видим, среди них уже нет набравших одно и то же количество очков. А в споре за первое место между Кересом и Файном метод отдает предпочтение Кересу.

Скажем в заключение несколько слов об условии неразложимости матрицы, которое требуется для справедливости теоремы Фробениуса - Перрона. В различных турнирах часто происходит так, что какие-то участники по каким-то причинам между собой не встречаются. В таком случае в турнирной таблице логично проставлять им нули. Разложимость матрицы означала бы, что таких нулей окажется настолько много, что все участники турнира фактически разбиваются на несколько групп так, что турнир распадается на несколько независимых турниров между членами этих групп. Ясно, что практически это очень маловероятно.

6.4. QR-алгоритм. Алгоритмы поиска собственных чисел, описанные в параграфах 6.1 и 6.2 относятся к частным случаям. В этом пункте рассматривается алгоритм решения этой задачи в общем случае, т.е. для произвольных вещественных и, даже, комплексных матриц. Алгоритмов решения этой задачи в настоящее время существует достаточно много. Но QR-алгоритм является наиболее известным и весьма эффективным.

Пусть A – произвольная квадратная вещественная матрица размера $n \times n$. В нескольких словах QR-алгоритм поиска ее собственных чисел описывается очень просто. Алгоритм итерационный и каждая его итерация, т.е. шаг заключается в следующем. Матрица A раскладывается в произведение двух квадратных матриц

$$A = Q \cdot R \quad (6.4)$$

где матрица Q – ортогональная, а матрица R – треугольная. После этого матрица A заменяется на матрицу $A = R \cdot Q$ (множители меняются местами!) Оказывается, что:

- 1) У старой и у новой A одни и те же собственные числа.
- 2) Описанные итерации приближают A к блочно-диагональному виду.

Напомним, что квадратная матрица Q называется *ортогональной*, если $Q \cdot Q^T = E$ или, что то же самое, $Q^{-1} = Q^T$. Ортогональными являются все матрицы вращения. Произведение ортогональных матриц является ортогональной матрицей. Обратная к ортогональной матрице является ортогональной матрицей.

Для того, чтобы использовать QR-алгоритм, надо научиться вычислять разложение (6.4). Мы, фактически, это уже делали при решении систем линейных уравнений методом вращений.

Покажем ход вычислений на примере матрицы порядка $n=5$. Как и в п. 3.3 вычисляем матрицы вращения T_{ij} и приводим данную матрицу A к треугольному виду

$$T_{12}, B_1 = T_{12} A;$$

$$T_{13}, B_2 = T_{13} B_1 \text{ (матрица } T_{13} \text{ ищется по } B_1\text{);}$$

$$T_{14}, B_3 = T_{14} B_2 \text{ (матрица } T_{14} \text{ ищется по } B_2\text{);}$$

$$T_{15}, B_4 = T_{15} B_3 \text{ (матрица } T_{15} \text{ ищется по } B_3\text{);}$$

В результате в матрице B_4 в первом столбце начиная со второго места будут стоять нули. Далее

$$T_{23}, B_5 = T_{23} B_4 \text{ (матрица } T_{23} \text{ ищется по } B_4\text{);}$$

$$T_{24}, B_6 = T_{24} B_5 \text{ (матрица } T_{24} \text{ ищется по } B_5\text{);}$$

$$T_{25}, B_7 = T_{25} B_6 \text{ (матрица } T_{25} \text{ ищется по } B_6\text{);}$$

- обнулили второй столбец под диагональю.

$$T_{34}, B_8 = T_{34} B_7 \text{ (матрица } T_{34} \text{ ищется по } B_7\text{);}$$

$$T_{35}, B_9 = T_{35} B_8 \text{ (матрица } T_{35} \text{ ищется по } B_8\text{);}$$

- обнулили третий столбец под диагональю.

$$T_{45}, B_{10} = T_{45} B_9 \text{ (матрица } T_{45} \text{ ищется по } B_9\text{);}$$

- обнулили четвертый столбец под диагональю и получили треугольный вид.

Полученная треугольная матрица есть произведение

$$T_{45} T_{35} T_{34} T_{24} T_{23} T_{15} T_{14} T_{13} T_{12} A = (T_{45} T_{35} T_{34} T_{24} T_{23} T_{15} T_{14} T_{13} T_{12}) A$$

которую обозначим за R . Матрица в скобках является ортогональной. Ее можно взять за Q . Но для того, чтобы получить равенство (6.4) в качестве

Q возьмем матрицу, обратную к матрице в скобках (для чего ее надо только транспонировать). Итак,

$$Q = (T_{45} T_{35} T_{34} T_{24} T_{23} T_{15} T_{14} T_{13} T_{12})^T$$

Из имеющегося у нас равенства $R = Q^T A$ умножением слева на Q получаем $A = QR$. После этого матрица $A_1 = RQ$ - это новая A . Из написанных равенств очевидно, что $A_1 = Q^{-1}AQ$. Значит, матрицы A и A_1 сопряжены. Поэтому у них одни и те же собственные числа.

Запоминаем $Q = T_{45} T_{35} T_{34} T_{25} T_{24} T_{23} T_{15} T_{14} T_{13} T_{12}$ и $R = QA$. Результат первой итерации $A_1 = RQ^{-1}$, где $Q^{-1} = Q^T$ (транспонирование).

Рассмотрим один шаг QR-алгоритма на примере. Данна матрица

$$A = \begin{bmatrix} 4.10648 & 1.86684 & -0.75596 & 0.58101 & 0.72084 \\ -4.02714 & 6.77701 & -3.65456 & 0.55055 & 5.66449 \\ -16.78011 & -5.14376 & 9.2334 & -1.42796 & -3.10835 \\ 16.71219 & 6.14757 & -10.45068 & 2.39451 & 4.13491 \\ -7.91369 & -9.28729 & 8.33393 & -3.77944 & -2.9014 \end{bmatrix}$$

Вычисления матриц вращения и приведение к треугольному виду дают:

$$T_{12} = \begin{bmatrix} 0,71397 & -0,70018 & 0 & 0 & 0 \\ 0,70018 & 0,71397 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, B_1 = T_{12}A = \begin{bmatrix} 5,75163 & -3,41226 & 2,01912 & 0,02934 & -3,4515 \\ 0 & 6,14571 & -3,13855 & 0,79989 & 4,54899 \\ -16,78011 & -5,14376 & 9,2334 & -1,42796 & -3,10835 \\ 16,71219 & 6,14757 & -10,45068 & 2,39451 & 4,13491 \\ -7,91369 & -9,28729 & 8,33393 & -3,77944 & -2,9014 \end{bmatrix}$$

$$T_{13} = \begin{bmatrix} 0,32425 & 0 & -0,94597 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0,94597 & 0 & 0,32425 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, B_2 = T_{13}B_1 = \begin{bmatrix} 17,73845 & 3,75942 & -8,07982 & 1,36032 & 1,82126 \\ 0 & 6,14571 & -3,13855 & 0,79989 & 4,54899 \\ 0 & -4,89576 & 4,90396 & -0,43526 & -4,2729 \\ 16,71219 & 6,14757 & -10,45068 & 2,39451 & 4,13491 \\ -7,91369 & -9,28729 & 8,33393 & -3,77944 & -2,9014 \end{bmatrix}$$

$$T_{14} = \begin{bmatrix} 0,72785 & 0 & 0 & 0,68574 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -0,68574 & 0 & 0 & 0,72785 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, B_3 = T_{14}B_2 = \begin{bmatrix} 24,37115 & 6,95193 & -13,04735 & 2,63212 & 4,16108 \\ 0 & 6,14571 & -3,13855 & 0,79989 & 4,54899 \\ 0 & -4,89576 & 4,90396 & -0,43526 & -4,2729 \\ 0 & 1,89652 & -2,06587 & 0,81002 & 1,76068 \\ -7,91369 & -9,28729 & 8,33393 & -3,77944 & -2,9014 \end{bmatrix}$$

$$T_{15} = \begin{bmatrix} 0,95111 & 0 & 0 & 0 & -0,30884 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0,30884 & 0 & 0 & 0 & 0,95111 \end{bmatrix}, B_4 = T_{15}B_3 = \begin{bmatrix} 25,62371 & 9,48034 & -14,98332 & 3,67068 & 4,85371 \\ 0 & 6,14571 & -3,13855 & 0,79989 & 4,54899 \\ 0 & -4,89576 & 4,90396 & -0,43526 & -4,2729 \\ 0 & 1,89652 & -2,06587 & 0,81002 & 1,76068 \\ 0 & -6,6862 & 3,89694 & -2,78176 & -1,47444 \end{bmatrix}$$

$$\mathbf{T}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0,78216 & -0,62308 & 0 & 0 \\ 0 & 0,62308 & 0,78216 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_5 = \mathbf{T}_{23}\mathbf{B}_4 = \begin{bmatrix} 25,62371 & 9,48034 & -14,98332 & 3,67068 & 4,85371 \\ 0 & 7,85738 & -5,51041 & 0,89684 & 6,2204 \\ 0 & 0 & 1,88011 & 0,15795 & -0,50771 \\ 0 & 1,89652 & -2,06587 & 0,81002 & 1,76068 \\ 0 & -6,6862 & 3,89694 & -2,78176 & -1,47444 \end{bmatrix}$$

$$\mathbf{T}_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0,97208 & 0 & 0,23463 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -0,23463 & 0 & 0,97208 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_6 = \mathbf{T}_{24}\mathbf{B}_5 = \begin{bmatrix} 25,62371 & 9,48034 & -14,98332 & 3,67068 & 4,85371 \\ 0 & 8,08298 & -5,84127 & 1,06186 & 6,45983 \\ 0 & 0 & 1,88011 & 0,15795 & -0,50771 \\ 0 & 0 & -0,71528 & 0,57698 & 0,25203 \\ 0 & -6,6862 & 3,89694 & -2,78176 & -1,47444 \end{bmatrix}$$

$$\mathbf{T}_{25} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0,77054 & 0 & 0 & -0,63739 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0,63739 & 0 & 0 & 0,77054 \end{bmatrix}, \mathbf{B}_7 = \mathbf{T}_{25}\mathbf{B}_6 = \begin{bmatrix} 25,62371 & 9,48034 & -14,98332 & 3,67068 & 4,85371 \\ 0 & 10,48998 & -6,9848 & 2,59127 & 5,91735 \\ 0 & 0 & 1,88011 & 0,15795 & -0,50771 \\ 0 & 0 & -0,71528 & 0,57698 & 0,25203 \\ 0 & 0 & -0,72042 & -1,46664 & 2,98132 \end{bmatrix}$$

$$\mathbf{T}_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0,93465 & -0,35558 & 0 \\ 0 & 0 & 0,35558 & 0,93465 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_8 = \mathbf{T}_{34}\mathbf{B}_7 = \begin{bmatrix} 25,62371 & 9,48034 & -14,98332 & 3,67068 & 4,85371 \\ 0 & 10,48998 & -6,9848 & 2,59127 & 5,91735 \\ 0 & 0 & 2,01158 & -0,05753 & -0,56415 \\ 0 & 0 & 0 & 0,59544 & 0,05503 \\ 0 & 0 & -0,72042 & -1,46664 & 2,98132 \end{bmatrix}$$

$$\mathbf{T}_{35} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0,94145 & 0 & -0,33717 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0,33717 & 0 & 0,94145 \end{bmatrix}, \mathbf{B}_9 = \mathbf{T}_{35}\mathbf{B}_8 = \begin{bmatrix} 25,62371 & 9,48034 & -14,98332 & 3,67068 & 4,85371 \\ 0 & 10,48998 & -6,9848 & 2,59127 & 5,91735 \\ 0 & 0 & 2,13671 & 0,44035 & -1,53633 \\ 0 & 0 & 0 & 0,59544 & 0,05503 \\ 0 & 0 & 0 & -1,40017 & 2,61655 \end{bmatrix}$$

$$\mathbf{T}_{45} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0,39135 & -0,92024 \\ 0 & 0 & 0 & 0,92024 & 0,39135 \end{bmatrix}, \mathbf{B}_{10} = \mathbf{T}_{45}\mathbf{B}_9 = \begin{bmatrix} 25,62371 & 9,48034 & -14,98332 & 3,67068 & 4,85371 \\ 0 & 10,48998 & -6,9848 & 2,59127 & 5,91735 \\ 0 & 0 & 2,13671 & 0,44035 & -1,53633 \\ 0 & 0 & 0 & 1,52152 & -2,38632 \\ 0 & 0 & 0 & 0 & 1,07463 \end{bmatrix}$$

$$Q = \mathbf{T}_{45}\mathbf{T}_{35}\mathbf{T}_{34}\mathbf{T}_{25}\mathbf{T}_{24}\mathbf{T}_{23}\mathbf{T}_{15}\mathbf{T}_{14}\mathbf{T}_{13}\mathbf{T}_{12} = \begin{bmatrix} 0,16026 & -0,15717 & -0,65487 & 0,65221 & -0,30884 \\ 0,03312 & 0,78809 & 0,10148 & -0,0034 & -0,60623 \\ 0,87833 & -0,23627 & 0,06094 & -0,32858 & -0,2471 \\ -0,31539 & -0,53279 & 0,45089 & 0,10117 & -0,63493 \\ 0,31986 & 0,12058 & 0,59484 & 0,67558 & 0,27002 \end{bmatrix}$$

$$R = QA = \begin{bmatrix} 25,62376 & 9,48032 & -14,98333 & 3,67068 & 4,8537 \\ 0 & 10,49007 & -6,9849 & 2,59129 & 5,91742 \\ 0 & 0 & 2,13673 & 0,44033 & -1,53635 \\ 0 & 0 & 0 & 1,52151 & -2,38634 \\ 0 & 0 & 0 & 0 & 1,07465 \end{bmatrix}$$

$$A_1 = RQ^{-1} = \boxed{\begin{array}{ccccc} 13.32361 & 3.84456 & 16.94766 & -22.59873 & 4.2169 \\ 2.78802 & 3.96217 & -5.21769 & -12.23345 & 0.45848 \\ -0.63761 & 1.14669 & 0.36512 & 1.98349 & 1.15362 \\ 1.72931 & 1.44146 & 0.08965 & 1.66916 & 0.38352 \\ -0.3319 & -0.65148 & -0.26555 & -0.68234 & 0.29015 \end{array}}$$

Подобную итерацию надо повторить многократно, добиваясь того, чтобы матрица приобрела блочно-треугольный вид. Такие вычисления вряд ли можно выполнить вручную даже с помощью MS Excel. Требуется программирование. С помощью специальной программы QR.exe через 30 итераций получим

$$\boxed{\begin{array}{ccccc} 13.355949 & -10.907626 & -23.168357 & 12.133438 & 1.482597 \\ 4.321555 & 1.484058 & -1.865268 & 11.167593 & 4.815430 \\ 0 & 0 & 3.490374 & -6.590705 & -1.785529 \\ 0 & 0 & 2.493538 & 0.509624 & 1.772343 \\ 0 & 0 & 0 & 0 & 0.770002 \end{array}}$$

Теперь видим, что одно из собственных чисел есть 0.77. Четыре остальных скорее всего – комплексные и являются собственными числами матриц $\begin{pmatrix} 13.355949 & -10.907626 \\ 4.321555 & 1.484058 \end{pmatrix}$ и $\begin{pmatrix} 3.490374 & -6.590705 \\ 2.493538 & 0.509624 \end{pmatrix}$. Их можно найти как решения уравнения $|A - \lambda E| = 0$. Для данных матриц эти уравнения превращаются в

$$\begin{vmatrix} 13.355949 - \lambda & -10.907626 \\ 4.321555 & 1.484058 - \lambda \end{vmatrix} = 0 \quad \text{и} \quad \begin{vmatrix} 3.490374 - \lambda & -6.590705 \\ 2.493538 & 0.509624 - \lambda \end{vmatrix} = 0$$

Это квадратные уравнения

$$\lambda^2 - 14.84\lambda + 66.959 = 0 \quad \text{и} \quad \lambda^2 - 4\lambda + 18.213 = 0$$

с отрицательными дискриминантами. Их корни по обычной школьной формуле, но с учетом того, что для отрицательного $D \sqrt{D} = \sqrt{-D}i$,

$$\lambda_1 = 7.42 + 3.45i, \lambda_2 = 7.42 - 3.45i, \lambda_3 = 2 + 3.77i, \lambda_4 = 2 - 3.77i.$$

Это две пары сопряженных комплексных чисел. Пятое собственное число $\lambda_5 = 0.77$ уже найдено и является вещественным.

6.5. Метод Гивенса. Метод Гивенса – это одно из усовершенствований QR-алгоритма. Он основан на таком простом факте. При вычислении произведения $T_{ij} \cdot A$, где T_{ij} – матрица вращения, у матрицы A изменяются только i -тая и j -тая строки. А при вычислении произведения $A \cdot T_{ij}$ у матрицы A меняются только i -тый и j -тый столбцы. Допустим теперь, что $1 < k < i < j \leq n$. Рассмотрим произведение $T_{ij} \cdot A \cdot T_{ij}^{-1}$, где в качестве чисел a и b берутся $a = a_{ik}$, $b = a_{jk}$. В произведении $T_{ij} \cdot A$, как мы знаем (см. п. 3.3) эле-

мент b (в j -той строке и k -том столбце) обнуляется. И он останется нулем в произведении $\mathbf{T}_{ij} \cdot \mathbf{A} \cdot \mathbf{T}_{ij}^{-1}$, т.к. находится в k -том столбце, а $k < i < j$.

Это обстоятельство позволяет привести нашу матрицу к почти треугольному виду, сохраняя ее собственные числа. Говорят, что матрица находится в почти треугольном виде, если у нее $a_{ij} = 0$ для всех i, j при $i > j + 1$. Это значит, что в матрице все числа нули под побочной диагональю, стоящей непосредственно под главной. Приведение матрицы к такому виду – подготовительный этап метода Гивенса.

Рассмотрим этот этап более детально для матрицы A_0 размера 5×5 .

Этап 1.

Вычисляем матрицу вращения \mathbf{T}_{23} , взяв $a = a_{21}$, $b = a_{31}$. Затем вычисляем $\mathbf{A}_1 = \mathbf{T}_{23} \mathbf{A} \mathbf{T}_{23}^{-1}$. В этой матрице в первом столбце на третьем месте должен быть 0.

Вычисляем матрицу вращения \mathbf{T}_{24} , взяв в \mathbf{A}_1 $a = a_{21}$, $b = a_{41}$. Затем вычисляем $\mathbf{A}_2 = \mathbf{T}_{24} \mathbf{A}_1 \mathbf{T}_{24}^{-1}$. В этой матрице в первом столбце на третьем и четвертом месте должны быть нули.

Вычисляем матрицу вращения \mathbf{T}_{25} , взяв в \mathbf{A}_2 $a = a_{21}$, $b = a_{51}$. Затем вычисляем $\mathbf{A}_3 = \mathbf{T}_{25} \mathbf{A}_2 \mathbf{T}_{25}^{-1}$. В этой матрице в первом столбике на третьем, четвертом и пятом местах должны быть нули.

Вычисляем матрицу вращения \mathbf{T}_{34} , взяв в \mathbf{A}_3 $a = a_{32}$, $b = a_{42}$. Затем вычисляем $\mathbf{A}_4 = \mathbf{T}_{34} \mathbf{A}_3 \mathbf{T}_{34}^{-1}$. В этой матрице во втором столбце на четвертом месте должен быть 0. Нули в первом столбце на 3-ем, 4-м и 5-м местах сохраняются, т.к. в произведении $\mathbf{T}_{34} \mathbf{A}_3 \mathbf{T}_{34}^{-1}$ первый столбец не меняется.

Вычисляем матрицу вращения \mathbf{T}_{35} , взяв в \mathbf{A}_4 $a = a_{32}$, $b = a_{52}$. Затем вычисляем $\mathbf{A}_5 = \mathbf{T}_{35} \mathbf{A}_4 \mathbf{T}_{35}^{-1}$. В этой матрице в первом и втором столбце под побочной диагональю должны быть нули.

Вычисляем матрицу вращения \mathbf{T}_{45} , взяв в \mathbf{A}_4 $a = a_{43}$, $b = a_{53}$. Затем вычисляем $\mathbf{A}_5 = \mathbf{T}_{45} \mathbf{A}_4 \mathbf{T}_{45}^{-1}$. Эта матрица будет почти треугольная.

Итак, первый этап метода Гивенса сводит задачу нахождения собственных чисел к случаю почти треугольной матрицы

$$\left(\begin{array}{ccccc} \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} & \text{---} \end{array} \right) \quad \left(\begin{array}{c} 0 \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right)$$

Этап 2. Как в QR-алгоритме в п. 6.4 серией итераций \mathbf{A} приводится к блочно-диагональному виду. Одна итерация, как и раньше, начинается с умножений на матрицы вращения слева и приведения \mathbf{A} к треугольному виду. Но на этот раз понадобится только $(n-1)$ вращение. Для $n=5$ это будут:

$$\mathbf{T}_{12}, \mathbf{B}_1 = \mathbf{T}_{12} \mathbf{A};$$

$$\mathbf{T}_{23}, \mathbf{B}_2 = \mathbf{T}_{23} \mathbf{B}_1 \text{ (матрица } \mathbf{T}_{23} \text{ ищется по } \mathbf{B}_1\text{);}$$

$$\mathbf{T}_{34}, \mathbf{B}_3 = \mathbf{T}_{34} \mathbf{B}_2 \text{ (матрица } \mathbf{T}_{34} \text{ ищется по } \mathbf{B}_2\text{);}$$

$$\mathbf{T}_{45}, \mathbf{B}_4 = \mathbf{T}_{45} \mathbf{B}_3 \text{ (матрица } \mathbf{T}_{45} \text{ ищется по } \mathbf{B}_3\text{).}$$

Запоминаем $\mathbf{Q} = \mathbf{T}_{45} \mathbf{T}_{34} \mathbf{T}_{23} \mathbf{T}_{12}$. Получили $\mathbf{R} = \mathbf{QA}$ - треугольная матрица. Значит, $\mathbf{A} = \mathbf{Q}^{-1} \mathbf{R}$. Полагаем

$$\mathbf{A}_1 = \mathbf{R} \mathbf{Q}^{-1} = \mathbf{Q} \mathbf{A} \mathbf{Q}^{-1}.$$

Конец итерации.

Собственные числа у \mathbf{A} и \mathbf{A}_1 одни и те же. Важно, что новая матрица \mathbf{A}_1 как и исходная \mathbf{A} – *почти треугольная*. Повторяем эту итерацию достаточное число раз.

Покажем работу QR-алгоритма с использование метода Гивенса на примере матрицы из п. 6.4.

$$\mathbf{A} = \begin{bmatrix} 4.10648 & 1.86684 & -0.75596 & 0.58101 & 0.72084 \\ -4.02714 & 6.77701 & -3.65456 & 0.55055 & 5.66449 \\ -16.78011 & -5.14376 & 9.2334 & -1.42796 & -3.10835 \\ 16.71219 & 6.14757 & -10.45068 & 2.39451 & 4.13491 \\ -7.91369 & -9.28729 & 8.33393 & -3.77944 & -2.9014 \end{bmatrix}$$

Вычисления матриц вращения и приведение к почти треугольному виду дают:

$$\mathbf{T}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -0.233368 & -0.972388 & 0 & 0 \\ 0 & 0.972388 & -0.233368 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{A}_1 = \mathbf{T}_{23} \mathbf{A} = \begin{bmatrix} 4.106480 & 0.299426 & 1.991711 & 0.581010 & 0.720840 \\ 17.256591 & 7.103068 & 5.222012 & 1.260051 & 1.700612 \\ 0 & 3.732812 & 8.907342 & 0.868589 & 6.233475 \\ 16.712190 & 8.727473 & 8.416683 & 2.394510 & 4.134910 \\ -7.913690 & -5.936459 & -10.975728 & -3.779440 & -2.901400 \end{bmatrix}$$

$$\mathbf{T}_{24} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.718347 & 0 & 0.695685 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -0.695685 & 0 & 0.718347 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{B}_2 = \mathbf{T}_{24} \mathbf{B}_1 = \begin{bmatrix} 4.106480 & 0.619292 & 1.991711 & 0.209061 & 0.720840 \\ 24.022640 & 9.815429 & 9.606577 & -5.926757 & 4.098224 \\ 0 & 3.285719 & 8.907342 & -1.972913 & 6.233475 \\ 0 & 1.540665 & 2.413224 & -0.317851 & 1.787210 \\ -7.913690 & -6.893737 & -10.975728 & 1.414956 & -2.901400 \end{bmatrix}$$

$$\mathbf{T}_{25} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.949791 & 0 & 0 & -0.312886 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0.312886 & 0 & 0 & 0.949791 \end{bmatrix}, \quad \mathbf{B}_3 = \mathbf{T}_{25} \mathbf{B}_2 = \begin{bmatrix} 4.106480 & 0.362657 & 1.991711 & 0.209061 & 0.878415 \\ 25.292563 & 9.401241 & 12.558389 & -6.071899 & 8.151038 \\ 0 & 1.170378 & 8.907342 & -1.972913 & 6.948552 \\ 0 & 0.904116 & 2.413224 & -0.317851 & 2.179528 \\ 0 & -2.840923 & -7.418880 & -0.510488 & -2.487212 \end{bmatrix}$$

$$\mathbf{T}_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.791372 & 0.611334 & 0 \\ 0 & 0 & -0.611334 & 0.791372 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_4 = \mathbf{T}_{34}\mathbf{B}_3 = \begin{bmatrix} 4.106480 & 0.362657 & 1.703991 & -1.052156 & 0.878415 \\ 25.292563 & 9.401241 & 6.226401 & -12.482509 & 8.151038 \\ 0 & 1.478922 & 5.672632 & -6.600556 & 6.831312 \\ 0 & 0 & -2.214419 & 2.916859 & -2.523071 \\ 0 & -2.840923 & -6.183175 & 4.131431 & -2.487212 \end{bmatrix}$$

$$\mathbf{T}_{35} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.461756 & 0 & -0.887007 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0.887007 & 0 & 0.461756 \end{bmatrix}, \mathbf{B}_5 = \mathbf{T}_{35}\mathbf{B}_4 = \begin{bmatrix} 4.106480 & 0.362657 & 0.007668 & -1.052156 & 1.917065 \\ 25.292563 & 9.401241 & -4.354947 & -12.482509 & 9.286653 \\ 0 & 3.202820 & -1.012844 & -6.712456 & 9.663487 \\ 0 & 0 & 1.215459 & 2.916859 & -3.129249 \\ 0 & 0 & -3.351000 & -3.947024 & 4.198264 \end{bmatrix}$$

$$\mathbf{T}_{45} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.340978 & -0.940071 \\ 0 & 0 & 0 & 0.940071 & 0.340978 \end{bmatrix}, \mathbf{B}_6 = \mathbf{T}_{45}\mathbf{B}_5 = \begin{bmatrix} 4.106480 & 0.362657 & 0.007668 & -2.160940 & -0.335424 \\ 25.292563 & 9.401241 & -4.354947 & -12.986379 & -8.567901 \\ 0 & 3.202820 & -1.012844 & -11.373167 & -3.015148 \\ 0 & 0 & 3.564624 & 6.317535 & 2.713547 \\ 0 & 0 & 0 & 1.895772 & 0.797588 \end{bmatrix}$$

Подготовительный этап завершен. Сделаем одну итерацию этапа 2.
Вычисляем:

$$\mathbf{T}_{12}, \mathbf{B}_1 = \mathbf{T}_{12}\mathbf{A};$$

$$\mathbf{T}_{12} = \begin{bmatrix} 0.160261 & 0.987075 & 0 & 0 & 0 \\ -0.987075 & 0.160261 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_1 = \mathbf{T}_{12}\mathbf{A} = \begin{bmatrix} 25.623765 & 9.337850 & -4.297430 & -13.164844 & -8.510916 \\ 0 & 1.148683 & -0.705497 & 0.051800 & -1.042012 \\ 0 & 3.202820 & -1.012844 & -11.373167 & -3.015148 \\ 0 & 0 & 3.564624 & 6.317535 & 2.713547 \\ 0 & 0 & 0.000000 & 1.895772 & 0.797588 \end{bmatrix}$$

$$\mathbf{T}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.337592 & 0.941293 & 0 & 0 \\ 0 & -0.941293 & 0.337592 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_2 = \mathbf{T}_{23}\mathbf{B}_1 = \begin{bmatrix} 25.623765 & 9.337850 & -4.297430 & -13.164844 & -8.510916 \\ 0 & 3.402578 & -1.191553 & -10.687995 & -3.189913 \\ 0 & 0 & 0.322151 & -3.888249 & -0.037051 \\ 0 & 0 & 3.564624 & 6.317535 & 2.713547 \\ 0 & 0 & 0 & 1.895772 & 0.797588 \end{bmatrix}$$

$$\mathbf{T}_{34} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.090008 & 0.995941 & 0 \\ 0 & 0 & -0.995941 & 0.090008 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{B}_3 = \mathbf{T}_{34}\mathbf{B}_2 = \begin{bmatrix} 25.623765 & 9.337850 & -4.297430 & -13.164844 & -8.510916 \\ 0 & 3.402578 & -1.191553 & -10.687995 & -3.189913 \\ 0 & 0 & 3.579151 & 5.941919 & 2.699198 \\ 0 & 0 & 0 & 4.441095 & 0.281142 \\ 0 & 0 & 0 & 1.895772 & 0.797588 \end{bmatrix}$$

$$\mathbf{T}_{45} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.919711 & 0.392597 \\ 0 & 0 & 0 & -0.392597 & 0.919711 \end{bmatrix}, \mathbf{B}_4 = \mathbf{T}_{45}\mathbf{B}_3 = \begin{bmatrix} 25.623765 & 9.337850 & -4.297430 & -13.164844 & -8.510916 \\ 0 & 3.402578 & -1.191553 & -10.687995 & -3.189913 \\ 0 & 0 & 3.579151 & 5.941919 & 2.699198 \\ 0 & 0 & 0 & 4.828798 & 0.571700 \\ 0 & 0 & 0 & 0 & 0.623175 \end{bmatrix}$$

Теперь $Q = T_{45}T_{34}T_{23}T_{12} =$

$$\begin{bmatrix} 0.160261 & 0.987075 & 0 & 0 & 0 \\ -0.333228 & 0.054103 & 0.941293 & 0 & 0 \\ 0.083628 & -0.013578 & 0.030386 & 0.995941 & 0 \\ -0.851059 & 0.138177 & -0.309226 & 0.082781 & 0.392598 \\ 0.363292 & -0.058984 & 0.132000 & -0.035337 & 0.919710 \end{bmatrix},$$

$R = QA =$

$$\begin{bmatrix} 25.623757 & 9.337847 & -4.297429 & -13.164840 & -8.510914 \\ 0 & 3.402576 & -1.191552 & -10.687992 & -3.189910 \\ 0 & 0 & 3.579151 & 5.941921 & 2.699198 \\ 0 & 0 & 0 & 4.828789 & 0.571700 \\ 0 & 0 & 0 & 0 & 0.623174 \end{bmatrix}$$

$A_1 = RQ^{-1} =$

$$\begin{bmatrix} 13.323633 & -12.078484 & -11.225900 & -23.619328 & 0.828503 \\ 3.358596 & -0.937511 & -10.727016 & -1.298496 & -2.914095 \\ 0 & 3.369030 & 6.026559 & 0.444811 & 2.744958 \\ 0 & 0 & 4.809190 & 0.624179 & 0.355164 \\ 0 & 0 & 0 & 0.244657 & 0.573140 \end{bmatrix}$$

- результат первой итерации. Матрица, как и полагается, получилась почти треугольной. Это можно объяснить тем, что матрица A_1 получается в результате умножения треугольной матрицы справа сначала на T_{12}^{-1} . Такое умножение “портит” только верхний 0 в первом столбце. Затем умножение на T_{23}^{-1} “портит” верхний 0 во втором столбце, но не более того. Также для T_{34}^{-1} и T_{45}^{-1} .

В итоге через 22 итерации получится матрица

$$\begin{bmatrix} 5.020246 & -13.964457 & 14.677870 & 20.191325 & -4.270747 \\ 1.264729 & 9.819760 & 1.782299 & -13.633839 & -2.673423 \\ 0 & 0 & 0.157795 & -2.803097 & -1.599600 \\ 0 & 0 & 6.281144 & 3.842200 & -1.941799 \\ 0 & 0 & 0 & 0 & 0.770001 \end{bmatrix}$$

блочно-диагонального вида. Как и в п. 6.4 она дает те же самые собственные числа.

Тема 7. Задача одномерной безусловной оптимизации

В общем случае в этой задаче требуется найти наименьшее или наибольшее значение некоторой числовой функции $y=f(x)$. О функции f всегда что-то известно. В зависимости от этой информации алгоритмы поиска ее максимума или минимума могут сильно различаться. Мы рассмотрим лишь некоторые случаи. Фактически – это один случай.

Предполагается, что функция f определена на некотором числовом промежутке $[a,b]$ и является унимодальной. **Унимодальность** означает, что существует такое число $x^* \in (a,b)$, что на промежутке $[a, x^*]$ функция f строго убывает, а на промежутке $[x^*, b]$ – строго возрастает. Тогда в точке

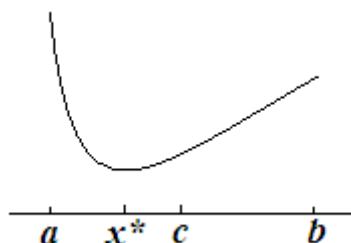
x^* значение f минимально и это единственная точка минимума на $[a,b]$. Вот эту точку и надо найти. Аналогично можно рассматривать функции, которые сначала возрастают, а затем – убывают, и требуется найти их точку максимума. Но такая задача сводится к предыдущей заменой функции $y=f(x)$ на $y=-f(x)$.

Дополнительно к унимодальности мы предполагаем, что умеем вычислять (или как-то получать) значения $f(x)$ для любого $x \in [a,b]$.

7.1. Метод половинного деления. Ищем точку минимума функции $f(x)$ с точностью до некоторого небольшого числа $\varepsilon > 0$. Если область определения $f(x)$ – вся числовая ось или какое-то иное множество, то сначала, исходя из какой-то дополнительной информации, надо найти промежуток $[a,b]$, на котором $f(x)$ унимодальна. Например, сделать серию пробных вычислений и найти три числа $a < c < b$, для которых $f(c) < f(a)$ и $f(c) < f(b)$. Как то проверить (или поверить), что интересующий нас минимум находится внутри $[a,b]$ и что на $[a,b]$ функция унимодальна.

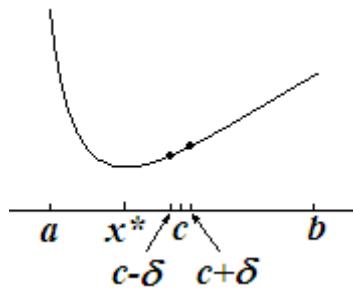
Выбираем число $\delta > 0$. Оно должно быть настолько мало, насколько позволяет точность вычислений. Во всяком случае $\delta < \varepsilon$.

Суть метода. Разделив данный промежуток на две равные части и каким-нибудь способом определив, в какой из этих частей находится x^* , мы сужаем область поиска почти в два раза.



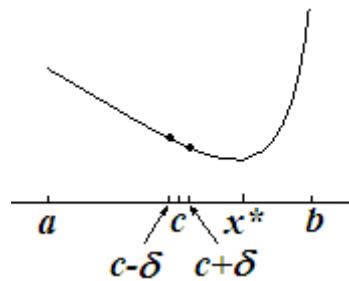
К сожалению, нет способа, который позволил бы, вычислив функцию в середине – точке $c = \frac{a+b}{2}$, выбрать нужную половину (как, например, происходит при решении уравнений методом деления отрезка пополам). Здесь задача немного сложнее. Надо вычислить значения функции не в середине, а в двух точках – немного левее и немного правее точки c , т.е. найти $f(c-\delta)$ и $f(c+\delta)$.

1. Случай $f(c-\delta) < f(c+\delta)$.



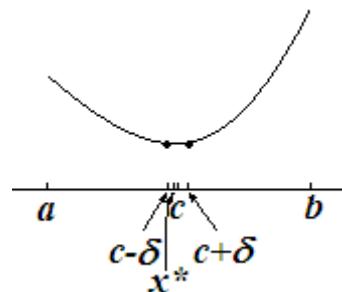
Такое расположение точек гарантирует, что x^* находится в левой половине. Точнее – от a до $c+\delta$. Это почти половина.

2. Случай $f(c-\delta) > f(c+\delta)$.



Такое расположение точек гарантирует, что x^* находится в правой половине. Точнее – от $c-\delta$ до b .

3. Случай $f(c-\delta) = f(c+\delta)$. Самый маловероятный.



Такое расположение точек гарантирует, что x^* находится между $c-\delta$ и $c+\delta$.

Алгоритм.

1. Если $b - a \leq 2\epsilon \Rightarrow$ перейти к выполнению п. 7.
2. Вычислить $c = (a + b)/2$.
3. Задать $x_1 = c - \delta$ и $x_2 = c + \delta$.
4. Вычислить $f(x_1)$ и $f(x_2)$.
5. Если $f(x_1) > f(x_2)$ положить $a := x_1$, число b не менять.

Если $f(x_1) < f(x_2)$ положить $b := x_2$, число a не менять.

Если $f(x_1) = f(x_2)$ положить $a := x_1$; $b := x_2$;

6. Перейти к выполнению п. 1.

7. Положить $x^* = (a + b)/2$. Процесс завершен.

Пример. Пусть $f(x) = x^4 - 6x^2 + 16x - 10$. Несложный анализ показывает, что это унимодальная функция. Требуется найти минимум этой функции с точностью до $\varepsilon=0,0001$. Пробуя вычислять $f(x)$ в различных точках можно найти: $a=-4$, $c=0$, $b=1$, для которых $f(a) = 256 - 96 - 64 - 10 = 86$, $f(c)=-10$, $f(b) = 1$. Значит. точка минимума лежит между $a=-4$ и $b=1$. Выбираем $\delta=0,00001$.

Применяем алгоритм, стартуя с этих a и b и результаты заносим в таблицу.

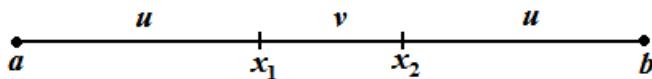
	a	b	$c=(a+b)/2$	$x_1=c-d$	$x_2=c+d$	$f(x_1)$	$f(x_2)$	$b-a-2\varepsilon$
1	-4,000000	1,000000	-1,500000	-1,500010	-1,499990	-42,437705	-42,437295	4,999800
2	-4,000000	-1,499990	-2,749995	-2,750005	-2,749985	-42,183423	-42,184107	2,499810
3	-2,750005	-1,499990	-2,124998	-2,125008	-2,124988	-50,702904	-50,702842	1,249815
4	-2,750005	-2,124988	-2,437496	-2,437506	-2,437486	-49,348050	-49,348304	0,624818
5	-2,437506	-2,124988	-2,281247	-2,281257	-2,281237	-50,641908	-50,641990	0,312319
6	-2,281257	-2,124988	-2,203122	-2,203132	-2,203112	-50,813572	-50,813579	0,156069
7	-2,203132	-2,124988	-2,164060	-2,164070	-2,164050	-50,791961	-50,791932	0,077945
8	-2,203132	-2,164050	-2,183591	-2,183601	-2,183581	-50,811391	-50,811380	0,038882
9	-2,203132	-2,183581	-2,193357	-2,193367	-2,193347	-50,814662	-50,814660	0,019351
10	-2,203132	-2,193347	-2,198239	-2,198249	-2,198229	-50,814666	-50,814668	0,009586
11	-2,198249	-2,193347	-2,195798	-2,195808	-2,195788	-50,814801	-50,814801	0,004703
12	-2,198249	-2,195788	-2,197019	-2,197029	-2,197009	-50,814767	-50,814768	0,002261
13	-2,197029	-2,195788	-2,196408	-2,196418	-2,196398	-50,814793	-50,814793	0,001041
14	-2,196418	-2,195788	-2,196103	-2,196113	-2,196093	-50,814799	-50,814799	0,000430
15	-2,196113	-2,195788	-2,195951	-2,195961	-2,195941	-50,814800	-50,814800	0,000125
16	-2,195961	-2,195788	-2,195874	-2,195884	-2,195864	-50,814801	-50,814801	-0,000027

Только после 16-й итерации длина промежутка $(b-a)$ стала меньше ε . значит, точка минимума есть $x^* = (a+b)/2 = -2,19587$.

Замечание. Вместо выбора не очень понятного числа δ и зависящих от него точек $x_1=c-\delta$ и $x_2=c+\delta$ можно эти x_1 и x_2 вычислять так: $x_1=b-(b-a)\cdot\lambda$, $x_2=a+(b-a)\cdot\lambda$. Здесь λ - какое-то постоянное число, чуть большее, чем 0.5. Например, $\lambda=0.51$. Тогда рассмотренный выше процесс с каждым шагом будет уменьшать длину промежутка (a,b) на 51%.

7.2. Метод золотого сечения. Понятие золотого сечения часто встречается в математике и, особенно, в популярной математике. Говорят, что отрезок длины l разбивается на две части длины l_1 и l_2 в отношении золотого сечения, если $\frac{l}{l_1} = \frac{l_1}{l_2}$. Если это отношение обозначить за τ , то относи-

тельно τ получается уравнение $1 + \frac{1}{\tau} = \tau$, т.е. уравнение $\tau^2 - \tau - 1 = 0$. Положительным корнем этого квадратного уравнения является число $\tau = \frac{\sqrt{5}+1}{2} = 1,618033989$. Это число и также обратное к нему $\frac{1}{\tau} = 0,618033989$ также называют золотым сечением. Главная особенность числа τ , используемая в алгоритме поиска минимума унимодальной функции, состоит в следующем. Возьмем какой-нибудь отрезок, расположим его горизонтально и разобьем в отношении золотого сечения – сначала так, что правая часть будет больше левой, а затем – наоборот. Точки разбиения – это x_1 и x_2 . Пусть u, v – длины этих частей (см. рис.)



Тогда $\frac{2u+v}{u+v} = \frac{u+v}{u}$. Отсюда $1 + \frac{u}{u+v} = 1 + \frac{v}{u}$ и $\frac{u}{u+v} = \frac{v}{u}$ или $\frac{u+v}{u} = \frac{u}{v}$. Это значит, что точка x_1 делит отрезок $[a, x_2]$ в отношении золотого сечения, а точка x_2 делит отрезок $[x_1, b]$ в отношении золотого сечения. Только в первом случае больший отрезок – слева, а во втором случае – справа.

Один шаг метода золотого сечения заключается в переходе от имеющегося промежутка $[a,b]$, на котором функция $f(x)$ унимодальна, к промежутку, который в $\frac{\sqrt{5}+1}{2}$ раз короче. Это сокращает длину промежутка не так быстро, как в случае половинного деления. Но зато позволяет каждый раз (кроме первого шага) значение функции вычислять только в **одной** новой точке.

Итак, разбиваем наш отрезок (a,b) по принципу золотого сечения точкой x_1 . Пусть (a,x_1) – меньшая часть, и (x_1,b) – большая. Разбиваем отрезок (a,b) точкой x_2 таким же образом на две части (x_2,b) – меньшая часть и (a,x_2) – большая. Вычисляем $f(x_1)$ и $f(x_2)$. Далее шаг за шагом уменьшаем отрезок (a,b) в $\tau = 1,6180339887498948482045868343656$ раз. Каждый раз вычисляются новые x_1 и x_2 а также значения функции f в этих точках. При этом точка минимума функции f каждый раз оказывается внутри нового интервала (a,b) .

Новым интервалом (a,b) каждый раз становится либо (a,x_2) , либо (x_1,b) . Какой именно – зависит от соотношения между $f(x_1)$ и $f(x_2)$. См. рис.

При $f(x_1) < f(x_2)$

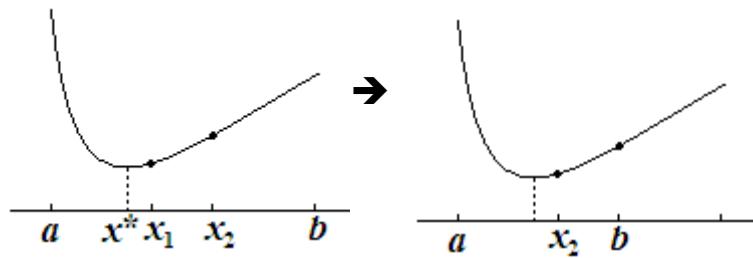


Рис. 7.1. x_2 превращается в b , x_1 превращается в x_2 , новый x_1 и $f(x_1)$ вычисляются.

При $f(x_1) > f(x_2)$

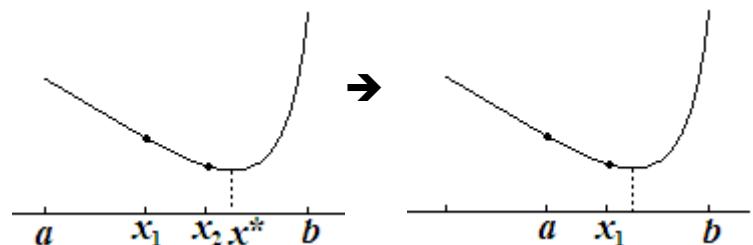


Рис. 7.1. x_1 превращается в a , x_2 превращается в x_1 , новый x_2 и $f(x_2)$ вычисляются.

При этом все время сохраняется пропорция, в которой x_1 разбивает (a,b) , она же – пропорция, в которой x_2 разбивает (x_1,b) и она же – пропорция, в которой x_1 разбивает (a,x_2) . Это позволяет каждый раз, с какой бы стороны интервал не сжимался, пересчитывать только одну новую точку и, соответственно, одно значение функции.

Алгоритм

Пусть $\tau \approx 0.618034$ и задано $\varepsilon > 0$ – требуемая точность определения точки искомого минимума $x^* \in (a, b)$.

1. Если $b - a \leq 2\varepsilon \Rightarrow$ перейти к выполнению п. 7.
2. Вычислить $c = \tau(b - a)$.
3. Вычислить $x_1 = b - c; x_2 = a + c$.
4. Вычислить значения $f(x_1)$ и $f(x_2)$.
5. Если $f(x_1) > f(x_2)$ положить:

$a := x_1; b := b; x_1 := x_2; c = \tau(b - a), x_2 = a + c; f(x_1) := f(x_2)$ и вычислить $f(x_2)$;

иначе,

если $f(x_1) < f(x_2)$ положить:

$a := a; b := x_2; x_2 := x_1; c = \tau(b - a), x_1 = b - c; f(x_2) := f(x_1)$ и вычислить $f(x_1)$;

иначе (при $f(x_1) = f(x_2)$ – что почти невозможно) перейти к п. 7.

6. Перейти к п. 5.

7. Положить $x^* = (a + b)/2$. Процесс завершен.

Продемонстрируем этот алгоритм на примере предыдущего пункта.

Ищем минимум $f(x) = x^4 - 6x^2 + 16x - 10$ на промежутке (-4,1). Заполняем таблицу:

	a	b	$c = \tau(b-a)$	$x_1 = b-c$	$x_2 = a+c$	$f(x_1)$	$f(x_2)$	$b-a-2\varepsilon$
1	-4,000000	1,000000	3,090170	-2,090170	-0,909830	-50,569077	-28,838786	4,999800
2	-4,000000	-0,909830	1,909830	-2,819660	-2,090170	-39,607278	-50,569077	3,089970
3	-2,819660	-0,909830	1,180340	-2,090170	-1,639320	-50,569077	-45,131386	1,909630
4	-2,819660	-1,639320	0,729490	-2,368810	-2,090170	-50,082273	-50,569077	1,180140
5	-2,368810	-1,639320	0,450850	-2,090170	-1,917961	-50,569077	-49,226909	0,729290
6	-2,368810	-1,917961	0,278640	-2,196601	-2,090170	-50,814787	-50,569077	0,450650
7	-2,368810	-2,090170	0,172209	-2,262379	-2,196601	-50,710620	-50,814787	0,278440
8	-2,262379	-2,090170	0,106431	-2,196601	-2,155948	-50,814787	-50,778896	0,172009
9	-2,262379	-2,155948	0,065778	-2,221726	-2,196601	-50,799263	-50,814787	0,106231
10	-2,221726	-2,155948	0,040653	-2,196601	-2,181073	-50,814787	-50,809840	0,065578
11	-2,221726	-2,181073	0,025125	-2,206198	-2,196601	-50,812323	-50,814787	0,040453
12	-2,206198	-2,181073	0,015528	-2,196601	-2,190670	-50,814787	-50,814193	0,024925
13	-2,206198	-2,190670	0,009597	-2,200267	-2,196601	-50,814347	-50,814787	0,015328
14	-2,200267	-2,190670	0,005931	-2,196601	-2,194336	-50,814787	-50,814750	0,009397
15	-2,200267	-2,194336	0,003666	-2,198001	-2,196601	-50,814692	-50,814787	0,005731
16	-2,198001	-2,194336	0,002266	-2,196601	-2,195736	-50,814787	-50,814800	0,003466
17	-2,196601	-2,194336	0,001400	-2,195736	-2,195201	-50,814800	-50,814792	0,002066
18	-2,196601	-2,195201	0,000865	-2,196066	-2,195736	-50,814799	-50,814800	0,001200
19	-2,196066	-2,195201	0,000535	-2,195736	-2,195531	-50,814800	-50,814799	0,000665
20	-2,196066	-2,195531	0,000331	-2,195862	-2,195736	-50,814801	-50,814800	0,000335
21	-2,196066	-2,195736	0,000204	-2,195940	-2,195862	-50,814800	-50,814801	0,000131
22	-2,195940	-2,195736	0,000126	-2,195862	-2,195814	-50,814801	-50,814801	0,000004
23	-2,195862	-2,195736	0,000078	-2,195814	-2,195784	-50,814801	-50,814801	-0,000074

После 23-й итерации длина промежутка $(b-a)$ стала меньше 2ε . значит, точка минимума есть $x^* = (a+b)/2 = -2,195799$.

7.3. Метод Ньютона-Рафсона. Этот метод применим только тогда, когда функция f имеет непрерывные первую и вторую производную и они доступны для вычисления. Тогда минимум унимодальной функции достигается в точке, где ее производная обращается в ноль. Поэтому для нахождения минимума надо решить уравнение

$$f'(x) = 0$$

а его решение часто находится методом касательных (п. 2.2). Отличие только в том, что роль функции $f(x)$ играет ее производная $f'(x)$.

Поэтому надо взять x_0 - любую точку и итерационно вычислять

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)} \quad (n = 0, 1, 2, \dots)$$

до тех пор, пока эти числа не сойдутся к точке минимума. В нашем примере для $f(x) = x^4 - 6x^2 + 16x - 10$ имеем:

$$f'(x) = 4x^3 - 12x + 16, \quad f''(x) = 12x^2 - 12.$$

Поэтому

$$x_{n+1} = x_n - \frac{4x_n^3 - 12x_n + 16}{12x_n^2 - 12} \quad (n = 0, 1, 2, \dots)$$

Вычисления, начиная с $x_0 = -3$ с помощью MS Excel дают

	x_n	$f'(x_n)$	$f''(x_n)$
	-3	-56	96
1	-2,4166667	-11,456	58,08333
2	-2,2194325	-1,09745	47,11057
3	-2,1961373	-0,0144	45,87623
4	-2,1958234	-2,6E-06	45,85968
5	-2,1958233	-7,5E-14	45,85968
6	-2,1958233	0	45,85968
7	-2,1958233	0	45,85968

Результат $x^* = -2,1958233$ получается всего за 6 итераций.

Справедливо ради надо заметить, что при неудачном выборе точки x_0 процесс может не сходиться.

Тема 8. Задача многомерной безусловной оптимизации

В этой задаче требуется найти наименьшее или наибольшее значение некоторой числовой функции от n переменных $y=f(x_1, \dots, x_n)$. Если эта функция определена на всем \mathbf{R}^n , имеет единственную точку минимума и дифференцируема во всех точках, то этот минимум можно искать как решение системы $\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = 0$ ($i=1, \dots, n$). Но кроме трудностей с решением системы надо уметь вычислять эти частные производные, что далеко не всегда возможно. В этом разделе излагается только один метод, который не всегда дает результат, но испробовать его можно почти всегда. Требуется только возможность вычислять значения функции $f(x_1, \dots, x_n)$ в любой точке $(x_1, \dots, x_n) \in \mathbf{R}^n$. Это **метод сопряженных направлений**.

8.1. Метод сопряженных направлений. Для начала допустим, что функция $f(x_1, \dots, x_n)$ имеет следующий “простой” вид

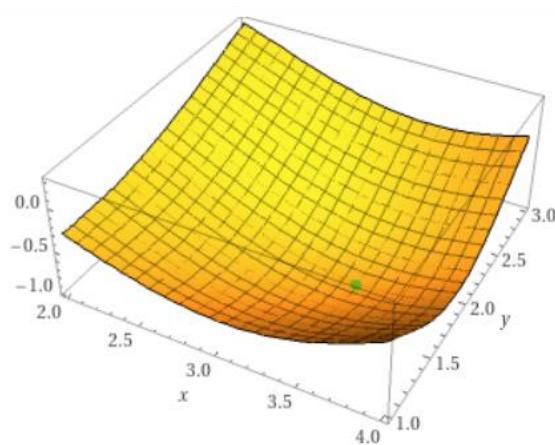
$$f(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c \quad (8.1)$$

Кроме этого считаем, что первая сумма в данном выражении – некая квадратичная форма – является положительно определенной (т.е. ее значения при любых ненулевых x_1, \dots, x_n строго положительны). Тогда функция $f(x_1, \dots, x_n)$ имеет единственную точку минимума, которую можно было бы искать, как решение системы из частных производных f_i , приравненных к нулю. Но, не смотря на то, что эта система является линейной, у нас может не быть никакой информации о значениях ее коэффициентов, т.е. о числах a_{ij} , b_i . Единственное, что мы можем делать – получать (может быть - измерением, а не вычислением) значения функции.

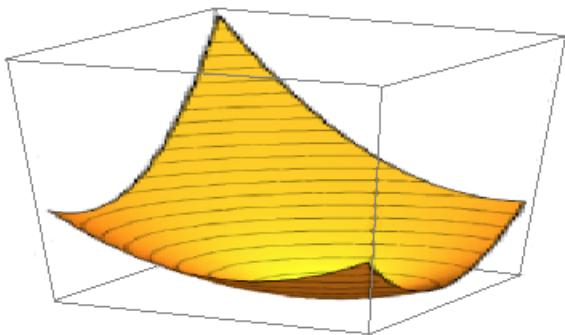
Ниже описывается итерационный алгоритм, который за $\frac{n(n+1)}{2}$ шагов находит точный (если пренебречь ошибками округлений) минимум функции (8.1). Причем каждый шаг – решение некоторой задачи одномерной минимизации.

Этот алгоритм можно испытывать на любой функции. Основанием к этому является тот факт, что достаточно “гладкая функция” в окрестности почти любой точки может быть приближена по формуле Тейлора функцией вида (8.1). Поэтому те же итерации для такой функции ожидаемо будут приближаться к минимуму этой функции.

Некоторой иллюстрацией к сказанному является следующий **пример**: Рассмотрим функцию $f(x,y) = \sin\left(\frac{xy}{4}\right)\cos\left(\frac{x(2y-11)}{8}\right)$ в квадрате $[2,4] \times [1,3]$. Программа WolframAlpha показывает, что у этой функции есть точка минимума $(\frac{12\pi}{11}, \frac{11}{6}) \approx (3.4271919857343196, 1.83333333333)$, в которой функция принимает значение -1. График функции в указанной области изображен на картинке.



Если бы мы могли применить формулу Тейлора для этой функции в точке минимума, то график функции выглядел бы как параболоид



Предположим, что мы не знаем формулы этого параболоида, но откуда-то свыше можем получать информацию о “его” значениях в произвольных точках плоскости. Метод сопряженных направлений в этом случае очень быстро – буквально за одну итерацию – позволяет точно найти точку минимума параболоида. При этом вычисление начинается с любой точки плоскости XY. Используя, но весьма хитро, метод наискорейшего спуска в один прыжок мы приходим в точку минимума параболоида.

На самом деле мы не знаем формулы этого параболоида и не умеем считать его значения. Но мы умеем вычислять значения исходной функции, а вблизи точки минимума они примерно такие же. Если использовать их вместо точек параболоида, то мы должны прийти к чему-то близкому к точке минимума. Потом надо процесс повторить, стартуя с полученной точки. Затем повторить еще раз и т.д., пока не перестанем куда-то сдвигаться. При достижении этого предела мы с большой вероятностью и получим точку минимума исходной функции.

Остается понять, как методом наискорейшего спуска за один прыжок попасть в точку минимума? Стандартный метод наискорейшего спуска, о котором мы говорили еще в разделе 3 (см. рис. 3.1) сходится к минимуму или максимуму очень медленно. Как было сказано в этом разделе надо так изменить метрику пространства, что эллипсы станут окружностями. При этом понятие ортогональности (т.е. перпендикулярности) заменится понятием *сопряженности*.

Остается понять, как вычислять эти сопряженные направления. Вспомним, что все, что мы пока умеем делать – это получать значения функций в различных точках. Теперь нам придется предположить, что наша функция очень близка к параболоиду и, значит, вдоль любой прямой имеет единственную точку минимума. Эти минимумы на прямых надо уметь искать. Как именно? Любым методом одномерной минимизации. Например, методом золотого сечения. Аналитически это выглядит так.

Пусть имеется прямая на плоскости, проходящая через точку x_0 в направлении (т.е. параллельно) вектора e . Тогда параметрическое уравнение этой прямой имеет вид

$$r = \mathbf{x}_0 + t \cdot \mathbf{e} \quad (t - \text{произвольное число})$$

Отыскание минимума данной функции $f(x,y)$ на этой прямой – это поиск минимума функции одной переменной

$$g(t) = f(x^{(0)} + t \cdot m, y^{(0)} + t \cdot l)$$

где $(x^{(0)}, y^{(0)})$ – координаты вектора \mathbf{x}_0 , а (m, l) – координаты вектора \mathbf{e} .

Сопряженные направления на плоскости позволяет находить

Теорема 1. *Если точки z_0 и z_1 – это точки минимума параболоида на двух параллельных и не совпадающих прямых, то вектор $z_1 - z_0$ будет сопряжен с направлением этих прямых.*

Теперь поиск минимума выглядит так

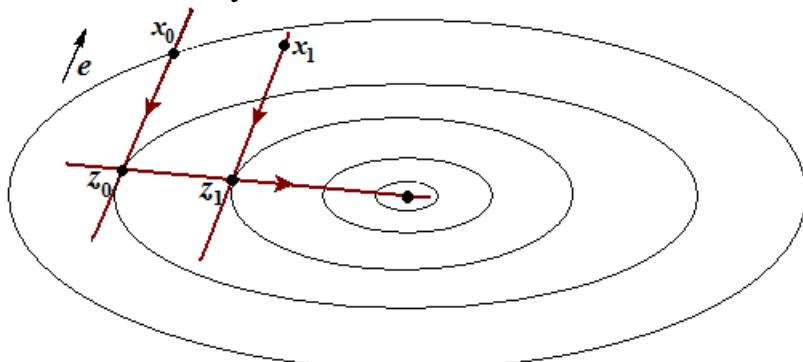


Рис. 8.1. Сопряженные направления на плоскости

Здесь \mathbf{e} – любой ненулевой вектор, $\mathbf{x}_0, \mathbf{x}_1$ – любые две точки, не лежащие на прямой, параллельной \mathbf{e} . Ищем z_0 – минимум на прямой $\mathbf{x}_0 + t \cdot \mathbf{e}$. Затем z_1 – минимум на прямой $\mathbf{x}_1 + t \cdot \mathbf{e}$. И, наконец, минимум на прямой $z_0 + t \cdot (z_1 - z_0)$. Это будет центр эллипсов.

Теперь рассмотрим случай функции от 3-х и более переменных. Для функции от переменных x, y, z мы не в силах представить или изобразить ее график. Это какое-то трехмерное многообразие в 4-х мерном пространстве. Однако легко представить ее поверхности уровня – точки, где эта функция принимает одно и то же значение. Вблизи точки минимума это замкнутые поверхности, близкие к эллипсоиду. В соответствующей метрике – это сферы. Точка минимума – центр этих сфер. Если взять произвольные точки \mathbf{x}_0 и \mathbf{x}_1 в пространстве, произвольное направление s_1 (надо только побеспокоиться, чтобы \mathbf{x}_0 и \mathbf{x}_1 не лежали на прямой, параллельной s_1) и найти одномерные минимумы функции на прямых $\mathbf{x}_0 + t \cdot s_1$ и $\mathbf{x}_1 + t \cdot s_1$, то получим направление $s_2 = z_1 - z_0$ сопряженное с s_1 . Минимум в направлении s_2 от точки z_0 (или от z_1) будет минимумом функции на всей плоскости T , проходящей через параллельные прямые $\mathbf{x}_0 + t \cdot s_1$ и $\mathbf{x}_1 + t \cdot s_1$. Пусть это точка \mathbf{x}_2 .

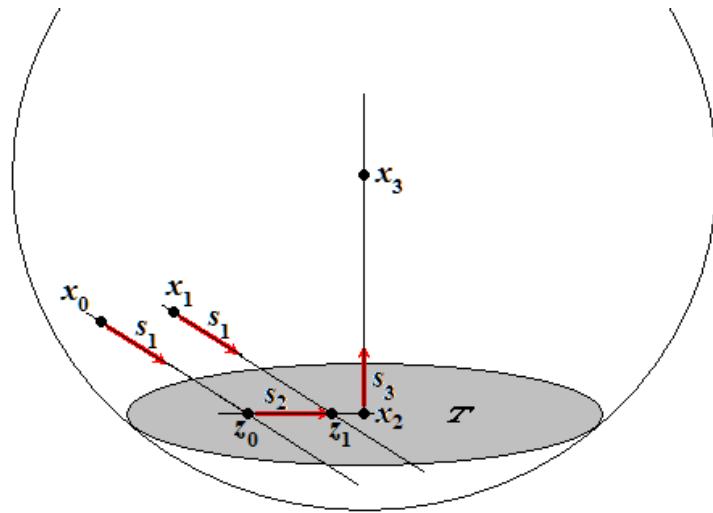


Рис. 8.2. Евклидова метрика

От точки x_2 надо искать одномерный минимум в направлении s_3 , которое сопряжено со всей плоскостью T . Как его найти? Надо повторить поиск минимума еще в одной плоскости, параллельной T . Вектор, соединяющий эти два минимума, и будет вектором s_3 . Минимум на прямой $x_2 + t \cdot s_3$ и будет глобальным минимумом. Здесь мы опираемся на следующее обобщение теоремы 1.

Теорема 2. *Если точки z_0 и z_1 – это точки минимума n -мерного параболоида на двух параллельных и не совпадающих линейных многообразиях T_1 и T_2 , то вектор $z_1 - z_0$ будет сопряжен с этими многообразиями. Более того, одномерный поиск минимума от z_0 (или от z_1) в направлении $z_1 - z_0$ даст минимум в многообразии, порожденном T_1 и T_2 .*

Для большей строгости скажем, что *n -мерный параболоид* – это график функции вида (8.1), где квадратичная форма является положительно определенной. *Линейное многообразие* – это множество векторов $x_0 + L$, где x_0 – один n -мерный вектор, а L – линейное подпространство n -мерного пространства R^n . *Параллельными* называются линейные многообразия вида $x_0 + L$ и $x_1 + L$. L – это направляющее подпространство линейного многообразия $x_0 + L$. Многообразие, *порожденное* двумя не совпадающими многообразиями $x_0 + L$ и $x_1 + L$ – это многообразие $x_0 + L_1$, L_1 – подпространство, порожденное векторами из L и вектором $x_1 - x_0$. Его размерность на единицу больше, чем у L . Вектор считается *сопряженным* с $x_0 + L$, если он сопряжен с каждым вектором из L .

8.2. Алгоритм метода сопряженных направлений. Приведем описание метода сопряженных направлений для минимизации квадратичной функции $f(x)$, $x \in R^n$, т.е. функции вида (8.1) с положительно

определенной квадратичной формой. У такой функции минимум существует и единственный. Алгоритм, описанный ниже, состоит из серии решений задач одномерной минимизации следующего вида:

Найти $t \in \mathbf{R}$, при котором достигается минимум функции $g(t) = f(\mathbf{x}_0 + t\mathbf{e})$. Здесь $\mathbf{x}_0, \mathbf{e} \in \mathbf{R}^n$, причем \mathbf{e} – ненулевой вектор. Проще говоря – это поиск минимума функции f начиная с \mathbf{x}_0 в направлении \mathbf{e} . Решением такой задачи следует считать не число t , а n -мерный вектор (точку) $\mathbf{x}_1 = \mathbf{x}_0 + t\mathbf{e}$. Будем писать для краткости:

$$\mathbf{x}_0 \rightarrow_{\mathbf{e}} \mathbf{x}_1$$

Алгоритм

1. Взять любые две различные точки

$$\mathbf{x}^{(0)} \in \mathbf{R}^n, \mathbf{y}^{(0)} \in \mathbf{R}^n.$$

и любое случайное направление $s^{(1)}$ (ненулевой вектор), не параллельное вектору $\mathbf{y}^{(0)} - \mathbf{x}^{(0)}$. Например, $\mathbf{x}^{(0)} = (0, 0, \dots, 0)$, $\mathbf{y}^{(0)} = (0, 1, \dots, 0)$.

$$s^{(1)} = (1, 0, \dots, 0)$$

2. Осуществить одномерные поиски

$$\mathbf{x}^{(0)} \rightarrow_{s^{(1)}} \mathbf{x}^{(1)}$$

$$\mathbf{y}^{(0)} \rightarrow_{s^{(1)}} \mathbf{y}^{(1)}$$

3. Положить $s^{(2)} = \mathbf{y}^{(1)} - \mathbf{x}^{(1)}$. Этот вектор будет сопряжен с $s^{(1)}$.

4. Продолжить

$$\mathbf{x}^{(0)} \rightarrow_{s^{(1)}} \mathbf{x}^{(1)} \rightarrow_{s^{(2)}} \mathbf{x}^{(2)}$$

5. Выбрать **новый** $\mathbf{y}^{(0)}$, не лежащий в плоскости $\mathbf{x}^{(0)} + L(s^{(1)}, s^{(2)})$. Например, $\mathbf{y}^{(0)} = \mathbf{x}^{(0)} + (0, 0, 1, 0, \dots, 0)$ и построить одномерные поиски

$$\mathbf{y}^{(0)} \rightarrow_{s^{(1)}} \mathbf{y}^{(1)} \rightarrow_{s^{(2)}} \mathbf{y}^{(2)}$$

в результате которых найдутся **новый** $\mathbf{y}^{(1)}$ и какой-то $\mathbf{y}^{(2)}$.

6. Положить $s^{(3)} = \mathbf{y}^{(2)} - \mathbf{x}^{(2)}$. Этот вектор будет сопряжен с $s^{(1)}$ и с $s^{(2)}$.

7. Продолжить

$$\mathbf{x}^{(0)} \rightarrow_{s^{(1)}} \mathbf{x}^{(1)} \rightarrow_{s^{(2)}} \mathbf{x}^{(2)} \rightarrow_{s^{(3)}} \mathbf{x}^{(3)}$$

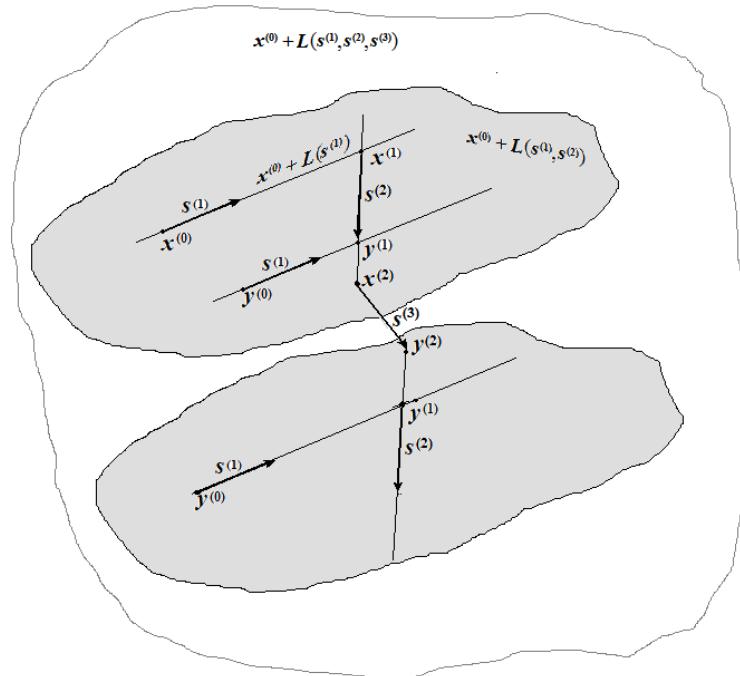


Рис. 8.3. Сопряженные направления

8. Выбрать **новый** $y^{(0)}$, не лежащий в многообразии $x^{(0)}+L(s^{(1)}, s^{(2)}, s^{(3)})$. Например, $y^{(0)}=x^{(0)}+(0,0,0,1,0,\dots,0)$ и построить одномерные поиски

$$y^{(0)} \rightarrow_{s^{(1)}} y^{(1)} \rightarrow_{s^{(2)}} y^{(2)} \rightarrow_{s^{(3)}} y^{(3)}$$
9. Положить $s^{(4)} = y^{(3)} - x^{(3)}$.
-
- пока не построим $s^{(n)}$.
10. Последний одномерный поиск $x^{(n-1)} \rightarrow_{s^{(n)}} x^{(n)}$ даст исковую точку минимума.

Указанный алгоритм для квадратичной функции даст точный минимум за $\frac{n(n+1)}{2}$ одномерных минимизаций. Действительно, количество одномерных поисков при нахождении $s^{(2)}$ равно 2, при нахождении $s^{(k)}$ для $2 < k \leq n$ равно $1 + (k-1) = k$ и еще один поиск при нахождении $x^{(n)}$. Всего $2 + 3 + 4 + \dots + n + 1 = \frac{n(n+1)}{2}$.

Если функция не является квадратичной, то поиск следует продолжить. Для этого надо переиндексировать найденные сопряженные направления $s^{(2)}, s^{(3)}, \dots, s^{(n)}$ в $s^{(1)}, s^{(2)}, \dots, s^{(n-1)}$, точки $x^{(1)}, x^{(2)}, x^{(3)}, \dots, x^{(n)}$ в $x^{(0)}, x^{(1)}, x^{(2)}, \dots, x^{(n-1)}$ и искать новые $x^{(n)}, y^{(0)}, y^{(1)}, \dots, y^{(n)}$ и $s^{(n)}$ также, как и раньше. Новую точку $y^{(0)}$ надо при этом выбрать вне линейного многообразия $x^{(0)}+L(s^{(1)}, s^{(2)}, \dots, s^{(n-1)})$ — например, взяв $y^{(0)}=x^{(0)}+s^{(1)}$ для старого $s^{(1)}$.

Вернемся к рассмотренному выше примеру функции $f(x,y) = \sin(\frac{xy}{4})\cos(\frac{x(2y-11)}{8})$ в квадрате $[2,4] \times [1,3]$. Ищем минимум по описанному алгоритму. Стартуем с точки $x^{(0)}=(3.5, 2)$. В этой точке $f(x^{(0)})=-0.980909816187$. Выберем начальное направления

$$s^{(1)} = (1, 0)$$

Далее :

$$\mathbf{x}^{(0)} \rightarrow_{\mathbf{s}^{(1)}} \mathbf{x}^{(1)} = (3.479382438172137, 2), f(\mathbf{x}^{(1)}) = -0.9811244964667073.$$

Здесь ищется минимум функции $g(t) = f(\mathbf{x}^{(0)} + t \mathbf{s}^{(1)}) = \sin\left(\frac{(3.5+t)^2}{4}\right) \cos\left(\frac{(3.5+t)(4-11)}{8}\right)$ для $t \in [-1.5, 0.5]$. Это те значения t , для которых точка $\mathbf{x}^{(0)} + t \mathbf{s}^{(1)}$ находится в квадрате $[2,4] \times [1,3]$.

Выберем $\mathbf{y}^{(0)} = \mathbf{x}^{(0)} + (0, 1) = (3.5, 3)$ и найдем

$$\mathbf{y}^{(0)} \rightarrow_{\mathbf{s}^{(1)}} \mathbf{y}^{(1)} = (3.3667758989943617, 3), f(\mathbf{y}^{(1)}) = -0.2940089069228037$$

Вычисляем

$$\mathbf{s}^{(2)} = \mathbf{y}^{(1)} - \mathbf{x}^{(1)} = (-0.11261, 1)$$

и находим

$$\mathbf{x}^{(1)} \rightarrow_{\mathbf{s}^{(2)}} \mathbf{x}^{(2)} = (3.495609240060312, 1.85590265617463), f(\mathbf{x}^{(2)}) = -0.9977728074495391.$$

Если бы рассматриваемая функция $f(x,y)$ была квадратичной, найденная точка была бы точным минимумом. Но и сейчас он близок к точке минимума $(3.4271919857343196, 1.8333333333)$. Однако продолжим поиск. Берем за новую стартовую точку $\mathbf{x}^{(0)}$ ранее найденную $\mathbf{x}^{(1)}$. Направление $\mathbf{s}^{(1)}$ – это прежнее $\mathbf{s}^{(2)}$. Новая $\mathbf{x}^{(1)}$ – это прежняя $\mathbf{x}^{(2)}$. Таким образом, имеем

$$\mathbf{x}^{(0)} = (3.479382438172137, 2), \mathbf{x}^{(1)} = (3.495609240060312, 1.85590265617463), \mathbf{s}^{(1)} = (-0.11261, 1)$$

Выбираем **новый** $\mathbf{y}^{(0)} = \mathbf{x}^{(0)} + 0.1 \cdot (1, 0) = (3.579382438172137, 2)$. Коэффициент 0.1 здесь взят для того, чтобы точка $\mathbf{y}^{(0)}$ лежала внутри квадрата $[2,4] \times [1,3]$, а вектор $(1,0)$ – это старое $\mathbf{s}^{(1)}$. Делаем поиски (см. рис. 8.3)

$$\mathbf{y}^{(0)} \rightarrow_{\mathbf{s}^{(1)}} \mathbf{y}^{(1)} = (3.592224424887968, 1.8859605122473075), f(\mathbf{y}^{(1)}) = -0.9870928647813406$$

$$\mathbf{s}^{(2)} = \mathbf{y}^{(1)} - \mathbf{x}^{(1)} = (0.09662, 0.03006)$$

$$\mathbf{x}^{(1)} \rightarrow_{\mathbf{s}^{(2)}} \mathbf{x}^{(2)} = (3.4271045220511747, 1.834589762122108), f(\mathbf{x}^{(2)}) = -0.9999987940243094.$$

Это гораздо ближе к точному минимуму.

Рассмотрим еще один пример. Допустим, что имеется система из двух линейных уравнений с пятью неизвестными. Такие системы скорее всего имеют бесконечно много решений и называются недоопределенными. В этом бесконечном множестве требуется найти вектор минимальной длины. Пусть матрица системы есть

$$\begin{array}{ccccc|c} -3.801 & 5.908 & -0.343 & -5.257 & -6.846 & | & 40.497 \\ 0.789 & 6.540 & 0.732 & 9.630 & 1.770 & | & 69.011 \end{array}$$

Надо найти такие значения неизвестных x_1, x_2, x_3, x_4, x_5 , которые подходят в эту систему и для которых сумма $x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2$ минимальна. Это задача отыскания условного минимума. Для решения задач условного экстремума существует много методов, но в настоящем пособии они не рассматриваются. Поэтому сведем эту задачу к задаче безусловного минимума. Для этого из равенств системы выразим x_4 и x_5 через x_1, x_2, x_3 . Получим

$$x_4 = 9.60985 + 0.02343x_1 - 0.97541x_2 - 0.07778x_3$$

$$x_5 = -13,2948 - 0.57321x_1 + 1.612x_2 + 0.00962x_3$$

Следовательно, речь идет о минимизации функции

$$\begin{aligned} f(x_1, x_2, x_3) &= x_1^2 + x_2^2 + x_3^2 + \\ (9.60985 + 0.02343x_1 - 0.97541x_2 - 0.07778x_3)^2 &+ \\ (-13,2948 - 0.57321x_1 + 1.612x_2 + 0.00962x_3)^2 \end{aligned}$$

Это квадратичная функция. Найдем ее минимум методом сопряженных направлений.

Стартуем с точки $\mathbf{x}^{(0)} = (1,1,1)$ в направлении $\mathbf{s}^{(1)} = (1,0,0)$. Находим

$$\mathbf{x}^{(0)} \rightarrow_{\mathbf{s}^{(1)}} \mathbf{x}^{(1)} = (-5.18514, 1, 1), f(\mathbf{x}^{(1)}) = 175.74531$$

Берем $\mathbf{y}^{(0)} = \mathbf{x}^{(0)} + (0,1,0) = (1,2,1)$ и находим

$$\mathbf{y}^{(0)} \rightarrow_{\mathbf{s}^{(1)}} \mathbf{y}^{(1)} = (-4.47274, 1, 1), f(\mathbf{y}^{(1)}) = 137.11318$$

Берем $\mathbf{s}^{(2)} = \mathbf{y}^{(1)} - \mathbf{x}^{(1)} = (0.7124, 1, 0)$. и находим

$$\mathbf{x}^{(1)} \rightarrow_{\mathbf{s}^{(2)}} \mathbf{x}^{(2)} = (-1.27816, 6.48425, 1), f(\mathbf{x}^{(2)}) = 59.1842$$

Выбираем новый $\mathbf{y}^{(0)} = \mathbf{x}^{(0)} + (0, 0, 1) = (1, 1, 2)$ и находим новый $\mathbf{y}^{(1)}$:

$$\mathbf{y}^{(0)} \rightarrow_{\mathbf{s}^{(1)}} \mathbf{y}^{(1)} = (-5.17962, 1, 2), f(\mathbf{y}^{(1)}) = 177.27183$$

и следующий

$$\mathbf{y}^{(1)} \rightarrow_{\mathbf{s}^{(2)}} \mathbf{y}^{(2)} = (-1.28847, 6.46203, 2), f(\mathbf{y}^{(1)}) = 177.27183$$

Вычисляем последнее сопряженное направление $\mathbf{s}^{(3)} = \mathbf{y}^{(2)} - \mathbf{x}^{(2)} = (-0.01031, -0.02222, 1)$

и находим глобальный минимум

$$\mathbf{x}^{(2)} \rightarrow_{\mathbf{s}^{(3)}} \mathbf{x}^{(3)} = (-1.27064, 6.50046, 0.27039), f(\mathbf{x}^{(3)}) = 58.64964$$

Тема 9. Численное интегрирование

В этом пункте речь идет о вычислении интегралов функции одной переменной, когда нет возможности использовать “школьную” формулу Ньютона-Лейбница. Согласно определению интеграла функции $f(x)$ на промежутке $[a,b]$ – это предел интегральных сумм. Большинство формул формулы численного интегрирования основаны на этом определении. Благодаря ему мы знаем, что интеграл примерно равен соответствующей интегральной сумме. Выглядит такая интегральная сумма как $a_1f(x_1) + a_2f(x_2) + \dots + a_nf(x_n)$, где $a \leq x_1 < x_2 < \dots < x_n \leq b$, а a_1, a_2, \dots, a_n – какие-то числа, не связанные непосредственно со значениями функции f в точках x_1, x_2, \dots, x_n . Подобного рода формулы (а их существует немало) называются **квадратурными**. Как правило, одни из них проще, но не слишком точные. Другие – сложнее, но более точные. Ниже будут рассмотрены три вида таких формул.

9.1. Формула трапеций. Промежуток $[a,b]$ разбивается на n одинаковых отрезков: $a \leq x_0 < x_1 < x_2 < \dots < x_n = b$. Пусть

$$h = x_1 - x_0 = x_2 - x_1 = \dots = x_n - x_{n-1}.$$

Числа x_1, x_2, \dots, x_{n-1} разбивают $[a, b]$ на n отрезков одинаковой длины h . Поэтому говорят, что числа $x_0, x_1, x_2, \dots, x_n$ образуют **равномерную сетку** на $[a, b]$.

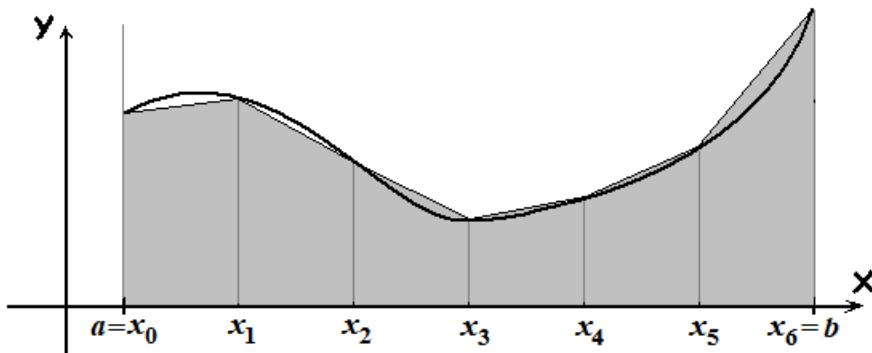


Рис. 9.1. Метод трапеций

На рисунке изображен график функции $y = f(x)$ при $n = 6$. Геометрически интеграл – площадь криволинейной трапеции между графиком, осью ОХ и вертикалями $x = a$ и $x = b$. Суть метода трапеций в приближенном вычислении этой площади путем замены ее на сумму площадей изображенных на рисунке трапеций. Это прямоугольные трапеции (расположенные “вертикально”) с одинаковой высотой h . Боковые стороны i -той трапеции равны $f(x_{i-1})$ и $f(x_i)$. Поэтому ее площадь равна $\frac{f(x_{i-1})+f(x_i)}{2}h$. Следовательно, примерное значение интеграла I – сумма этих площадей есть

$$I \approx \frac{h}{2}(f(x_0) + 2f(x_1) + 2f(x_2) + 2f(x_3) + 2f(x_4) + 2f(x_5) + f(x_6)).$$

В общем случае эту формулу записывают как

$$I = h \left(\frac{f(x_0)}{2} + f(x_1) + \dots + f(x_{n-1}) + \frac{f(x_n)}{2} \right) + R.$$

Здесь R – разница между точным значением интеграла и значением, которое дает первая часть формулы. Если ее обозначить за I_h , то

$$I = I_h + R, \text{ где}$$

$$I_h = h \left(\frac{f(x_0)}{2} + f(x_1) + \dots + f(x_{n-1}) + \frac{f(x_n)}{2} \right).$$

Величина ошибки R во многих случаях может быть оценена. Если потребовать, чтобы функция $f(x)$ имела непрерывную вторую производную на $[a, b]$, и $M_2 = \max_{x \in [a, b]} |f''(x)|$, то справедливо неравенство

$$|R| \leq \frac{M_2(b-a)}{12} h^2 \quad (9.1)$$

По этой причине говорят, что квадратурная формула трапеций имеет **второй порядок точности по h** . Приведенное неравенство говорит о том,

что чем меньше h (т.е. чем больше точек в сетке разбиения $[a,b]$), тем точнее результат, который дает формула. При этом уменьшение h в (допустим) 10 раз уменьшает ошибку в 100 раз. Неточности при вычислениях мы не учитываем.

9.2. Формула Симпсона. Промежуток $[a,b]$ разбивается на $2n$ одинаковых отрезков: $a \leq x_0 < x_1 < x_2 < \dots < x_{2n} = b$ длины $h = \frac{(b-a)}{2n}$. Суть метода в том, что на каждом промежутке $[x_{2k-2}, x_{2k}]$ ($k=1, \dots, n$) функция заменяется квадратичной параболой, проходящей через точки

$$(x_{2k-2}, f(x_{2k-2})), (x_{2k-1}, f(x_{2k-1})), (x_{2k}, f(x_{2k}))$$

Как известно, для любых трех точек (x_1, y_1) , (x_2, y_2) , (x_3, y_3) координатной плоскости при попарно различных x_1, x_2, x_3 существует и единственная функция вида $y(x) = ax^2 + bx + c$, для которой $y(x_i) = y_i$. График этой функции гораздо лучше приближает график $f(x)$. А в случае, когда $x_3 - x_2 = x_2 - x_1 = h$, можно доказать, что

$$\int_{x_1}^{x_3} y(x) dx = \frac{h}{3} (y_1 + 4y_2 + y_3)$$

Это значит, что на каждом промежутке $[x_{2k-2}, x_{2k}]$ интеграл нашей функции приближается величиной $\frac{h}{3} (f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k}))$. В сумме по всем $k=1, \dots, n$ это даст

$$I_h = \frac{h}{3} (f(x_0) + 4f(x_1) + 2f(x_2) + 4f(x_3) + 2f(x_4) + \dots + 4f(x_{2n-1}) + f(x_{2n})).$$

Это и есть знаменитая *формула Симпсона*. На этот раз разница R между точным значением интеграла и приближенным I_h оценивается как

$$|R| \leq \frac{M_4(b-a)}{180} h^4 \quad (9.2)$$

Здесь $M_4 = \max_{x \in [a,b]} |f^{(IV)}(x)|$ - максимум абсолютной величины четвертой производной. Разумеется, для справедливости этой оценки требуется существование непрерывной четвертой производной. Таким образом, формула Симпсона имеет четвертый порядок точности по h . Заметим, что четвертая производная от полиномов степеней не выше трех равна 0. Значит, приведенная оценка, в частности, говорит, что для функций, которые можно определить полиномами до третьей степени включительно формула Симпсона дает точный результат.

9.3. Правило Рунге. При вычислении интеграла с помощью какой-нибудь квадратурной формулы у нас далеко не всегда есть возможность оценить точность вычисления с помощью формул типа (9.1) или (9.2). Но имеется правило (Рунге), которое при известном порядке точности по h применяемой формулы позволяет оценить точность вычисления. Логика здесь следующая. Допустим, что некая формула дает величину I_h с точностью порядка m . Тогда можно предположить, что точное значение интеграла $I \approx I_h + ch^m$, где c – константа, не зависящая от h . Если уменьшить h в два раза, то с такой же легкостью можно допустить, что $I \approx I_{\frac{h}{2}} + c\left(\frac{h}{2}\right)^m$.

Значит, $(I - I_h) \approx ch^m$ и $(I - I_{\frac{h}{2}}) \approx c\left(\frac{h}{2}\right)^m = \frac{ch^m}{2^m}$. Вычитая из первого равенства второе найдем, что

$$(I_{\frac{h}{2}} - I_h) \approx ch^m \left(1 - \frac{1}{2^m}\right) = \frac{ch^m}{2^m} (2^m - 1) \approx (I - I_h)(2^m - 1). \text{ Отсюда}$$

$$(I - I_{\frac{h}{2}}) \approx \frac{(I_h - I_{\frac{h}{2}})}{2^{\frac{h}{2}}} \quad (9.3)$$

Равенство (9.3) и есть правило Рунге. Для его применения требуется вычислить I_h для некоторого h и затем повторить вычисление, уменьшив шаг h в два раза. Разница между полученными $I_{\frac{h}{2}}$ и I_h да еще разделенная на 2^{m-1} даст примерную величину ошибки. Не смотря на такое нестрогое обоснование, правило Рунге неплохо работает на практике.

Пример. Вычислим $\int_0^1 e^{-\frac{x^2}{2}} dx$

Теперь $f(x) = e^{-\frac{x^2}{2}}$ на промежутке $[0,1]$.

Метод трапеций. Возьмем $n=4$, т.е. разобьем отрезок $[0,1]$ на 4 части. Тогда $h=0.25$. С помощью MS Excel произведем подсчет, представленный в таблице

Номер промежутка i	x_i	$f(x_i)$	Коэффициент k	$k f(x_i)$
0	0	1	0,5	0,5
1	0,25	0,969233	1	0,969233
2	0,5	0,882497	1	0,882497
3	0,75	0,75484	1	0,75484
4	1	0,606531	0,5	0,303265

Сумма чисел последнего столбца, умноженная на h , и есть $I_h = 0.852459$.

Увеличим теперь n в два раза. Теперь $n=8, h=0.125$.

Соответствующий подсчет представлен в таблице

Номер промежутка i	x_i	$f(x_i)$	Коэффициент k	$k f(x_i)$
0	0	1	0,5	0,5
1	0,125	0,992218	1	0,992218
2	0,25	0,969233	1	0,969233
3	0,375	0,932102	1	0,932102
4	0,5	0,882497	1	0,882497
5	0,625	0,822578	1	0,822578
6	0,75	0,75484	1	0,75484
7	0,875	0,681941	1	0,681941
8	1	0,606531	0,5	0,303265

На этот раз $I_h = 0.854834$. Применим правило Рунге. Т.к. для метода трапеций $m=3$, то получим $\frac{(I_h - I_h)}{3^{\frac{2}{2}}} = 0,000792$. Это говорит о том, что значение $I_h = 0.854834$ с точностью до третьего знака после запятой скорее всего является правильным.

Увеличим n еще в два раза: $n=16, h=0.0625$.

Номер промежутка i	x_i	$f(x_i)$	Коэффициент k	$k f(x_i)$
0	0	1	0,5	0,5
1	0,0625	0,998049	1	0,998049
2	0,125	0,992218	1	0,992218
3	0,1875	0,982575	1	0,982575
4	0,25	0,969233	1	0,969233
5	0,3125	0,952345	1	0,952345
6	0,375	0,932102	1	0,932102
7	0,4375	0,908734	1	0,908734
8	0,5	0,882497	1	0,882497
9	0,5625	0,853676	1	0,853676
10	0,625	0,822578	1	0,822578
11	0,6875	0,789522	1	0,789522
12	0,75	0,75484	1	0,75484
13	0,8125	0,718868	1	0,718868
14	0,875	0,681941	1	0,681941
15	0,9375	0,644389	1	0,644389
16	1	0,606531	0,5	0,303265

Теперь $I_h = 0.855427$, то получим $\frac{(I_h - I_h)}{3^{\frac{2}{2}}} = 0,000198$.

И еще в два раза: $n=32, h=0.03125$.

Номер промежутка i	x_i	$f(x_i)$	Коэффициент k	$k \cdot f(x_i)$
0	0	1	0,5	0,5
1	0,03125	0,999512	1	0,999512
2	0,0625	0,998049	1	0,998049
3	0,09375	0,995615	1	0,995615
4	0,125	0,992218	1	0,992218
5	0,15625	0,987867	1	0,987867
6	0,1875	0,982575	1	0,982575
7	0,21875	0,976358	1	0,976358
8	0,25	0,969233	1	0,969233
9	0,28125	0,961221	1	0,961221
10	0,3125	0,952345	1	0,952345
11	0,34375	0,942629	1	0,942629
12	0,375	0,932102	1	0,932102
13	0,40625	0,920793	1	0,920793
14	0,4375	0,908734	1	0,908734
15	0,46875	0,895957	1	0,895957
16	0,5	0,882497	1	0,882497
17	0,53125	0,868391	1	0,868391
18	0,5625	0,853676	1	0,853676
19	0,59375	0,838392	1	0,838392
20	0,625	0,822578	1	0,822578
21	0,65625	0,806274	1	0,806274
22	0,6875	0,789522	1	0,789522
23	0,71875	0,772363	1	0,772363
24	0,75	0,75484	1	0,75484
25	0,78125	0,736994	1	0,736994
26	0,8125	0,718868	1	0,718868
27	0,84375	0,700503	1	0,700503
28	0,875	0,681941	1	0,681941
29	0,90625	0,663223	1	0,663223
30	0,9375	0,644389	1	0,644389
31	0,96875	0,625479	1	0,625479
32	1	0,606531	0,5	0,303265

$$I_h = 0,855575, \text{ то получим } \frac{(I_h - I_h)}{\frac{2}{3}} = 0,000049.$$

Формула Симпсона. Берем $n=4$, $h=0.25$. Производим подсчет

Номер промежутка i	x_i	$f(x_i)$	Коэффициент k	$k \cdot f(x_i)$
0	0	1	1	1
1	0,25	0,969233	4	3,876933
2	0,5	0,882497	2	1,764994
3	0,75	0,75484	4	3,019358
4	1	0,606531	1	0,606531

Сумма чисел последнего столбца, умноженная на $\frac{h}{3}$, есть $I_h = 0,855651$.

Увеличим n в два раза: $n=8$, $h=0.125$.

Номер промежутка <i>i</i>	<i>x_i</i>	<i>f(x_i)</i>	Коэффициент <i>k</i>	<i>k f(x_i)</i>
0	0	1	1	1
1	0,125	0,992218	4	3,968872
2	0,25	0,969233	2	1,938466
3	0,375	0,932102	4	3,72841
4	0,5	0,882497	2	1,764994
5	0,625	0,822578	4	3,29031
6	0,75	0,75484	2	1,509679
7	0,875	0,681941	4	2,727763
8	1	0,606531	1	0,606531

 $(I_h - I_{\bar{h}})$

Теперь $I_h = 0,855626$, $\frac{(I_h - I_{\bar{h}})}{3} = -0,000025$. Последнее вычисление, во-первых, демонстрирует превосходство метода Симпсона над методом трапеций, а, во-вторых показывает, что с точностью до четырех знаков после запятой наш интеграл равен 0,8556.

9.4. Формула Гаусса. Формула Гаусса наивысшей алгебраической степени точности способна при выборе в промежутке $[a,b]$ n узлов давать такую же точность, как интерполяционные полиномы степени $2n-1$. Это, в частности, значит, что если интегрируемая функция сама есть полином степени $2n-1$ (или ниже), то формула Гаусса даст абсолютно точный результат. В то же время можно доказать, что по n узлам для многочленов степени $2n$ и выше абсолютно точной квадратурной формулы не существует. В этом и заключается смысл выражения “наивысшая алгебраическая степень точности”.

Формула основана на свойствах так-называемых многочленов Лежандра. **Полиномы Лежандра** – это полиномы, используемые во многих разделах математики. По определению - это

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (n = 0, 1, 2, \dots)$$

Практически по такой формуле вычислять значения этих многочленов и, тем более, их корни почти невозможно. Однако можно воспользоваться следующими итерационными формулами

$$\begin{cases} (n+1)P_{n+1}(x) - (2n+1)x \cdot P_n(x) + n \cdot P_{n-1}(x) = 0 \\ P_0(x) = 1 \\ P_1(x) = x \end{cases}$$

Кубатурные формулы Гаусса применяются после выбора числа n – количества узлов. Это же n есть степень используемого полинома Лежандра. Нужны только корни многочленов Лежандра и некие числа C_k ($k=1,\dots,n$). Они считаются по формулам:

$$C_k = \frac{2(1-x_k^2)}{n^2 \cdot P_{n-1}^2(x_k)}$$

но, как правило, берутся из каких-нибудь справочников

Наконец корни многочленов тоже берутся из справочника или вычисляются. Например, их можно найти методом касательных Ньютона, используя итерационную формулу для производной

$$(n+1)P'_{n+1}(x) - (2n+1)(P_n(x) + xP'_n(x)) + nP'_{n-1}(x) = 0$$

При работе с корнями полиномов Лежандра следует учитывать следующую информацию:

1. Многочлен $P_n(x)$ имеет ровно n различных действительных корней, которые лежат в промежутке $(-1, 1)$.
2. Корни $P_n(x)$ разбиваются на пары симметричных относительно начала координат. Таким образом, при четном n половина из корней положительна и половина – отрицательна. При нечетном n , один из корней равен 0, а остальные опять делятся пополам – на положительные и отрицательные. Следовательно, для $n=7$ надо найти только 3 положительных корня, а при $n=8$ – только 4 положительных корня.

Но на этот раз сетка разбиения $[a,b]$ **не равномерна**. Формула Гаусса есть:

$$I \approx \frac{b-a}{2} (C_1 f(u_1) + C_2 f(u_2) + \dots + C_n f(u_n)) \quad (9.3)$$

Сетка $a < u_1 < u_2 < \dots < u_n < b$ получается из корней
 $-1 < x_1 < x_2 < \dots < x_n < 1$

полинома Лежандра степени n как бы накладыванием отрезка $[-1, 1]$ на $[a,b]$ с соответствующим растяжением. Т.е.

$$u_i = (a+b)/2 + (b-a)x_i/2 \quad (i = 1, 2, \dots, n) \quad (9.4)$$

Оценка точности формулы Гаусса имеет вид

$$|R| \leq M_{2n} k (b-a)^{2n+1} \quad (9.5)$$

Здесь M_{2n} – максимум на $[a,b]$ модуля производной порядка $2n$ функции f . Число k – константа, зависящая только от n и очень быстро стремящаяся к нулю с увеличением n . Например, при $n=8$ справедливо, что

$k < \frac{1}{5 \cdot 10^{22}}$. Неравенство (9.5) показывает, что для многочленов степени не больше $2n-1$ формула Гаусса дает точный результат – производная порядка $2n$ любого такого многочлена равна нулю.

Приведем значения корней полинома Лежандра степени $n=8$ и им соответствующие числа C_i ($i=1,\dots,8$).

i	Корни x_i	C_i
1	-0,96028985649754	0,101228536290376
2	-0,79666647741363	0,222381034453374
3	-0,52553240991633	0,313706645877887
4	-0,18343464249565	0,362683783378362
5	0,18343464249565	0,362683783378362
6	0,52553240991633	0,313706645877887
7	0,79666647741363	0,222381034453374
8	0,96028985649754	0,101228536290376

Вернемся к вычислению интеграла $\int_0^1 e^{-\frac{x^2}{2}} dx$. Составляя таблицу значений для корней, сетки u_i , которая вычисляется по формулам (9.4), функции $f(u_i)$, констант C_i и вычисляя получим

i	Корни x_i	Сетка u_i	$f(x_i)$	C_i	$C_i f(x_i)$
1	-0,96028985649754	0,019855071751230	0,999802907488158	0,101228536290376	0,101208584903888
2	-0,79666647741363	0,101666761293185	0,994845266297047	0,222381034453374	0,221234719440180
3	-0,52553240991633	0,237233795041835	0,972252303448975	0,313706645877887	0,305002009062027
4	-0,18343464249565	0,408282678752175	0,920031497709339	0,362683783378362	0,333680504416484
5	0,18343464249565	0,591717321247825	0,839402714582268	0,362683783378362	0,304437752302764
6	0,52553240991633	0,762766204958165	0,747585841851707	0,313706645877887	0,234522646953096
7	0,79666647741363	0,898333238706815	0,667977155834539	0,222381034453374	0,148545450905707
8	0,96028985649754	0,980144928248770	0,618571779211848	0,101228536290376	0,062617115800149

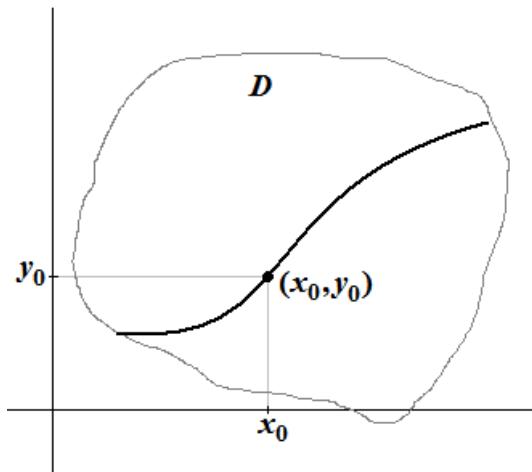
Сумма чисел последнего столбца, умноженная на $\frac{b-a}{2} = 0.5$ даст результат: $I = 0,855624391892148$. Это гораздо точнее, чем можно получить по формуле Симпсона даже при $n=16$.

Тема 10. Численное решение дифференциальных уравнений

В этом пункте рассматриваются некоторые методы численного решения задачи Коши для обыкновенных дифференциальных уравнений 1-го порядка. Предполагаем, что это уравнение имеет вид

$$y' = f(x,y) \quad (10.1)$$

где $f(x,y)$ – функция вида $f : D \rightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^2$ – область полоскости. Напомним, что **задача Коши** – это задача нахождения для данной точки $(x_0, y_0) \in D$ такого частного решения $y = y(x)$ уравнения (10.1), для которого $y(x_0) = y_0$. По-другому говорят, что это задача отыскания решения дифференциального уравнения с *начальными данными* $y(x_0) = y_0$.



Из общей теории дифференциальных уравнений известно, что при сравнительно небольших ограничениях на функцию $f(x,y)$ и область D задача Коши для любых $(x_0, y_0) \in D$ имеет и единственное решение. Например, достаточно, чтобы множество D было открытым, функция $f(x,y)$ была на нем непрерывна и имела непрерывную частную производную по y .

Существует множество способов в числах находить (т.е. строить графики и вычислять значения $y(x)$ на сетках оси OX) такое решение. Но все они основаны на простой идеи: в любой точке (x,y) области D данной графика решения должен идти так, чтобы касательная к нему в этой точке имела угловой коэффициент, равный $y'(x)$.

Самая простая реализация идеи это

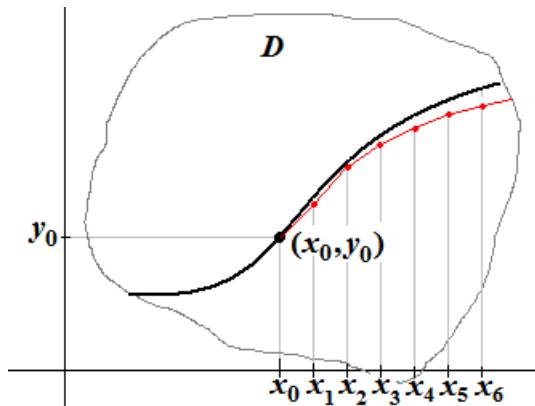
10.1. Формулы Эйлера (метод Эйлера). На оси OX выбирается равномерная сетка с шагом h .

$$x_0, x_1, \dots, x_m$$

шаг h может быть и отрицательным. Для простоты считаем, что $h > 0$.
Даны x_0 и y_0 . Тогда

$$y_{j+1} = y_j + h f(x_j, y_j) \quad (j = 0, 1, \dots, m-1).$$

На рисунке это может выглядеть как красная ломаная



Метод Эйлера самый простой и самый неточный. Он “предсказывает” значение $y(x)$ в следующей точке x_j только на основании ее значения в предыдущей и направления графика в этой предыдущей. Но можно пытаться учитывать поведение функции в нескольких предыдущих или в точках между узлами и какую-то еще информацию. Можно брать разное количество этих дополнительных точек. Ниже приводятся еще несколько схем расчета.

Исправленный метод Эйлера при $n=1$

$$y_{j+1} = y_j + h f(x_j, y_j) + \frac{h^2}{2} \left(\frac{\partial f}{\partial x}(x_j, y_j) + \frac{\partial f}{\partial y}(x_j, y_j) f(x_j, y_j) \right) \quad (j = 0, 1, \dots, m-1).$$

Эти формулы используют частные производные функции $f(x, y)$, и он опирается на известную формулу Тейлора. Точность метода гораздо лучше, чем у простой схемы Эйлера.

10.2. Формулы Рунге-Кутты 2-го, 3-го, 4-го порядков.

2-го порядка

$$\begin{aligned} k_1 &= h f(x_j, y_j) \\ k_2 &= h f(x_j + h, y_j + k_1) \\ y_{j+1} &= y_j + (k_1 + k_2)/2 \quad (j = 0, 1, \dots, m-1). \end{aligned}$$

3-го порядка

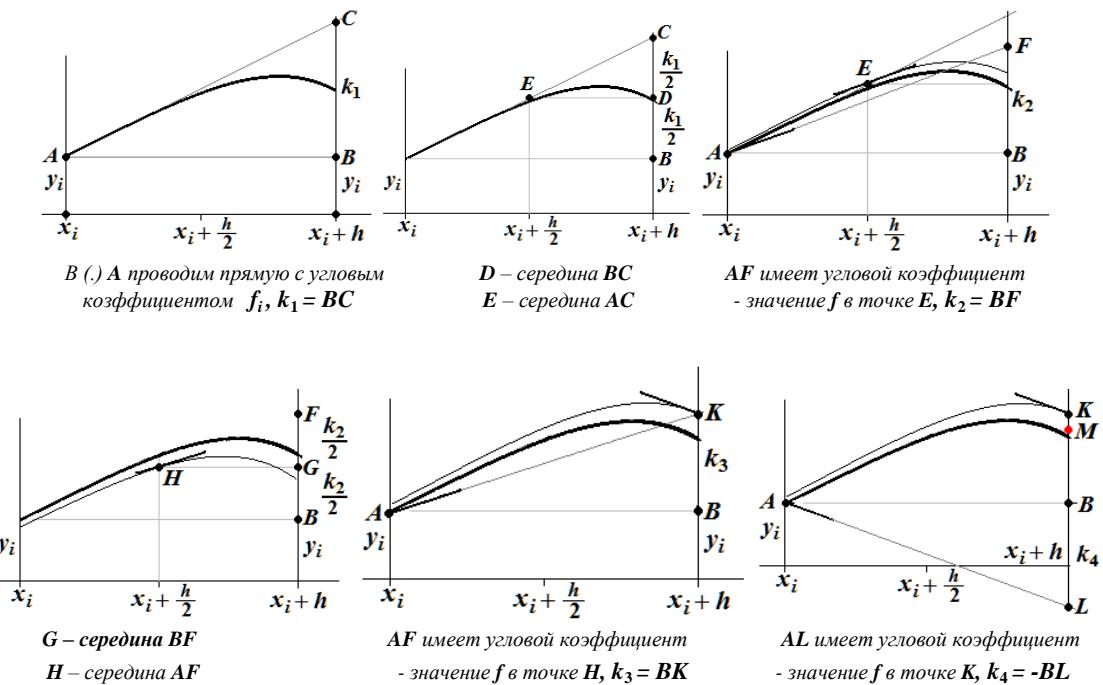
$$\begin{aligned}
 k_1 &= h \cdot f(x_j, y_j) \\
 k_2 &= h \cdot f(x_j + h/2, y_j + k_1/2) \\
 k_3 &= h \cdot f(x_j + h, y_j - k_1 + 2 \cdot k_2) \\
 y_{j+1} &= y_j + (k_1 + 4 \cdot k_2 + k_3)/6 \\
 (j &= 0, 1, \dots, m-1).
 \end{aligned}$$

4-го порядка

$$\begin{aligned}
 k_1 &= h \cdot f(x_j, y_j) \\
 k_2 &= h \cdot f(x_j + h/2, y_j + k_1/2) \\
 k_3 &= h \cdot f(x_j + h/2, y_j + k_2/2) \\
 k_4 &= h \cdot f(x_j + h, y_j + k_3) \\
 y_{j+1} &= y_j + (k_1 + 2 \cdot k_2 + 2 \cdot k_3 + k_4)/6 \\
 (j &= 0, 1, \dots, m-1).
 \end{aligned}$$

Эти последние формулы, вероятно, самые популярные не только среди методов Рунге-Кутты, но и вообще среди методов численного решения дифференциальных уравнений. Если по этим формулам вычислять значения решения задачи Коши на промежутке $[a, b]$ по равномерной сетке с шагом h , то погрешность этих формул оценивается как $k(b-a)h^4$, где константа k зависит только от функции $f(x, y)$ (которая должна быть достаточно гладкой). Поэтому говорят, что метод имеет 4-й порядок точности. Это обстоятельство позволяет контролировать точность вычислений так же, как при вычислении интегралов - с помощью правила Рунге.

Приглядевшись к формуле для вычисления y_{j+1} можно заметить, что она немного похожа на формулу метода Эйлера. Ведь $y_{j+1} = y_j + h \frac{f_1 + 2f_2 + 2f_3 + f_4}{6}$, где f_1, f_2, f_3, f_4 – значения функции $f(x, y)$ в 4-х точках, взятых недалеко от точки (x_i, y_i) . Вся дробь – это средневзвешенное этих величин. При этом числа f_2, f_3 “весят” в два раза больше, чем f_1 и f_4 . Ниже в серии рисунков приводится наглядная демонстрация этих вычислений. Жирная кривая изображает точное, но неизвестное решение, проходящее через точку A . Тонкие кривые – также неизвестные решения, проходящие через точки E, H, K . У этих кривых известны только угловые коэффициенты в точках E, H, K , т.е. касательные линии.



Если теперь найти $(k_1 + 2 \cdot k_2 + 2 \cdot k_3 + k_4)/6$ и отложить такой отрезок от точки B , то получим точку M , предсказывающую значение решения при $x=x_i+h/2$.

В заключение параграфа заметим, что для всех порядков найдено бесконечно много вариантов для формул Рунге-Кутты, аналогичным выше приведенным.

10.3. Явные схемы (методы) Адамса 2-го, 3-го, 4-го, 5-го порядков. Во всех следующих формулах $f_j = f(x_j, y_j)$.

2- го порядка

y_1 вычисляется по формулам Рунге-Кутты или еще как-нибудь

$$y_{j+1} = y_j + h \cdot (3f_j - f_{j-1})/2 \quad (j = 1, \dots, m-1)$$

3- го порядка

y_1, y_2 вычисляется по формулам Рунге-Кутты или еще как-нибудь

$$y_{j+1} = y_j + h \cdot (23f_j - 16 \cdot f_{j-1} + 5 \cdot f_{j-2})/12 \quad (j = 2, \dots, m-1).$$

4- го порядка

y_1, y_2, y_3 вычисляется по формулам Рунге-Кутты или еще как-нибудь

$$y_{j+1} = y_j + h \cdot (55f_j - 59 \cdot f_{j-1} + 37 \cdot f_{j-2} - 9 \cdot f_{j-3})/24 \quad (j = 3, \dots, m-1).$$

5- го порядка

y_1, y_2, y_3, y_4 вычисляется по формулам Рунге-Кутты или еще как-нибудь

$$y_{j+1} = y_j + h \cdot (1901 \cdot f_j - 2774 \cdot f_{j-1} + 2616 \cdot f_{j-2} - 1274 \cdot f_{j-3} + 251 \cdot f_{j-4}) / 720 \quad (j = 4, \dots, m-1).$$

10.4. Неявные схемы (методы) Адамса.

2-го порядка

$$y_{j+1} = y_j + h \cdot (f_{j+1} + f_j) / 2 \quad (j = 0, \dots, m-1)$$

Т.к. $f_{j+1} = f(x_{j+1}, y_{j+1})$, то y_{j+1} находится из этого равенства как решение уравнения (например, методом касательных).

3-го порядка

y_1 вычисляется по формулам Рунге-Кутты или еще как-нибудь.

$$y_{j+1} = y_j + h \cdot (5 \cdot f_{j+1} + 8 \cdot f_j - f_{j-1}) / 12 \quad (j = 1, \dots, m-1)$$

Т.к. $f_{j+1} = f(x_{j+1}, y_{j+1})$, то и здесь y_{j+1} находится из этого равенства как решение уравнения.

4-го порядка

y_1, y_2 вычисляются по формулам Рунге-Кутты или еще как-нибудь.

$$y_{j+1} = y_j + h \cdot (9 \cdot f_{j+1} + 19 \cdot f_j - 5 \cdot f_{j-1} + f_{j-2}) / 24 \quad (j = 2, \dots, m-1)$$

y_{j+1} находится из этого равенства как решение уравнения.

5-го порядка

y_1, y_2, y_3 вычисляется по формулам Рунге-Кутты или еще как-нибудь.

$$y_{j+1} = y_j + h \cdot (251 \cdot f_{j+1} + 646 \cdot f_j - 264 \cdot f_{j-1} + 106 \cdot f_{j-2} - 19 \cdot f_{j-3}) / 720 \quad (j = 3, \dots, m-1)$$

y_{j+1} находится из этого равенства как решение уравнения.

10.5. Сравнение методов на примере. Рассмотрим дифференциальное уравнение $y' = xy + \sin x$. Если сделать небольшой компьютерный эксперимент – на плоскости рассеять случайным образом множество точек и для каждой такой точки (x, y) провести через нее короткий отрезок с угловым коэффициентом, равным $f(x, y)$, то получим примерную картину решений нашего дифференциального уравнения.

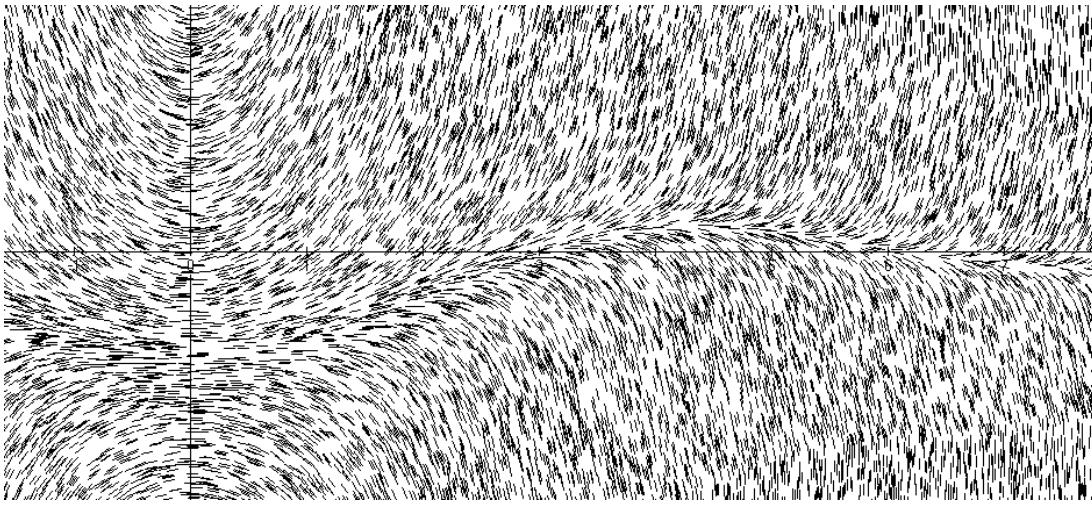


Рис. 10.1 Интегральные кривые

Рисунок показывает, что почти все решения $y(x)$ с увеличением x уходят в $+\infty$ или $-\infty$. Но есть одно решение, которое является ограниченным и с небольшими колебаниями приближается к нулю. Рассматриваемое уравнение является простым линейным уравнением первого порядка и все его решения нетрудно выразить аналитически. Оказывается, что единственное его ограниченное решение – это

$$y(x) = -e^{\frac{x^2}{2}} \int_x^\infty e^{-\frac{t^2}{2}} \sin t dt$$

Для него $y(0) = - \int_0^\infty e^{-\frac{t^2}{2}} \sin t dt \approx -0,724778459007076$.

Построим теперь численное решение задачи Коши на промежутке $[0, 5.7]$ с начальными данными $y(0) = -0,724778459007076$. Используем сетку с шагом $h = 0.1$ и испытываем 5 методов. Вычисления, сделанные в MS Excel, представим в таблице

Эйлера	Явные Адамса			Рунге-Кутты 4-го порядка					
	2-го пор.	3-го пор.	4-го пор.	k_1	k_2	k_3	k_4	y_i	
x_i	y_i	y_i	y_i	y_i	k_1	k_2	k_3	k_4	y_i
0	-0,7248	-0,7248	-0,7248	-0,7248	0	0,00137	0,00138	0,00275	-0,7248
0,1	-0,7248	-0,7234	-0,7234	-0,7234	0,00275	0,00411	0,00412	0,00548	-0,7234
0,2	-0,722	-0,7193	-0,7193	-0,7193	0,00548	0,00683	0,00684	0,00818	-0,7193
0,3	-0,7166	-0,7124	-0,7124	-0,7125	0,00818	0,0095	0,00952	0,01082	-0,7125
0,4	-0,7086	-0,7029	-0,7029	-0,7029	0,01082	0,01211	0,01214	0,0134	-0,7029
0,5	-0,698	-0,6908	-0,6908	-0,6908	0,0134	0,01464	0,01468	0,0159	-0,6908
0,6	-0,6849	-0,6761	-0,6761	-0,6762	0,01589	0,01708	0,01712	0,01829	-0,6762
0,7	-0,6696	-0,6589	-0,6589	-0,6591	0,01829	0,01942	0,01946	0,02057	-0,6591
0,8	-0,652	-0,6394	-0,6394	-0,6396	0,02056	0,02163	0,02168	0,02272	-0,6396
0,9	-0,6324	-0,6177	-0,6179	-0,618	0,02271	0,02371	0,02376	0,02472	-0,618
1	-0,611	-0,5939	-0,5942	-0,5943	0,02472	0,02564	0,02569	0,02658	-0,5943
1,1	-0,588	-0,5681	-0,5685	-0,5686	0,02658	0,02742	0,02746	0,02827	-0,5686
1,2	-0,5635	-0,5405	-0,5411	-0,5412	0,02826	0,02902	0,02907	0,02978	-0,5412
1,3	-0,5379	-0,5113	-0,512	-0,5121	0,02978	0,03045	0,03049	0,03112	-0,5121
1,4	-0,5115	-0,4807	-0,4815	-0,4817	0,03111	0,03169	0,03173	0,03226	-0,4817
1,5	-0,4846	-0,4488	-0,4498	-0,45	0,03226	0,03273	0,03277	0,03321	-0,45
1,6	-0,4575	-0,4157	-0,417	-0,4172	0,0332	0,03359	0,03362	0,03395	-0,4172
1,7	-0,4308	-0,3818	-0,3834	-0,3836	0,03395	0,03424	0,03426	0,0345	-0,3836
1,8	-0,4048	-0,3471	-0,3491	-0,3494	0,0345	0,03468	0,0347	0,03484	-0,3494

1,9	-0,3803	-0,3119	-0,3144	-0,3147	0,03484	0,03493	0,03494	0,03498	-0,3147
2	-0,3579	-0,2763	-0,2794	-0,2798	0,03497	0,03497	0,03497	0,03491	-0,2798
2,1	-0,3386	-0,2405	-0,2444	-0,2448	0,03491	0,03481	0,0348	0,03465	-0,2448
2,2	-0,3234	-0,2046	-0,2095	-0,2101	0,03464	0,03445	0,03443	0,03418	-0,21
2,3	-0,3137	-0,1687	-0,1749	-0,1756	0,03418	0,0339	0,03387	0,03353	-0,1756
2,4	-0,3112	-0,133	-0,1409	-0,1418	0,03353	0,03316	0,03312	0,0327	-0,1417
2,5	-0,3184	-0,0974	-0,1075	-0,1086	0,0327	0,03225	0,03219	0,03169	-0,1086
2,6	-0,3381	-0,062	-0,075	-0,0765	0,03169	0,03116	0,03109	0,03051	-0,0764
2,7	-0,3745	-0,0266	-0,0434	-0,0454	0,03052	0,02992	0,02984	0,02918	-0,0453
2,8	-0,4329	0,00907	-0,013	-0,0155	0,02919	0,02852	0,02843	0,02771	-0,0154
2,9	-0,5206	0,04534	0,01621	0,01289	0,02772	0,02699	0,02688	0,0261	0,01308
3	-0,6476	0,08293	0,0442	0,03975	0,02611	0,02533	0,02521	0,02438	0,04001
3,1	-0,8278	0,12288	0,07093	0,06492	0,02439	0,02356	0,02343	0,02255	0,06527
3,2	-1,0802	0,16676	0,09648	0,08828	0,02256	0,02169	0,02155	0,02063	0,08876
3,3	-1,4317	0,21692	0,121	0,10971	0,02065	0,01974	0,01959	0,01863	0,11037
3,4	-1,9199	0,27687	0,14482	0,12911	0,01866	0,01772	0,01756	0,01658	0,13003
3,5	-2,5983	0,35184	0,16843	0,14636	0,0166	0,01565	0,01548	0,01448	0,14766
3,6	-3,5428	0,44965	0,19268	0,16137	0,01451	0,01355	0,01337	0,01236	0,16322
3,7	-4,8624	0,58205	0,21885	0,17401	0,01239	0,01142	0,01124	0,01022	0,17667
3,8	-6,7145	0,7668	0,24896	0,18413	0,01025	0,00929	0,0091	0,00809	0,18799
3,9	-9,3271	1,03091	0,28615	0,19151	0,00812	0,00717	0,00698	0,00599	0,19718
4	-13,034	1,41573	0,33533	0,19586	0,00602	0,00509	0,0049	0,00392	0,20425
4,1	-18,323	1,98501	0,40419	0,19669	0,00396	0,00306	0,00287	0,00193	0,20924
4,2	-25,917	2,83775	0,50489	0,19325	0,00196	0,0011	0,00092	2E-05	0,21219
4,3	-36,889	4,12878	0,65678	0,18432	5,7E-05	-0,0008	-0,0009	-0,0018	0,21319
4,4	-52,843	6,10207	0,8909	0,16791	-0,0017	-0,0025	-0,0026	-0,0034	0,21235
4,5	-76,189	9,14481	1,25748	0,14072	-0,0033	-0,004	-0,0041	-0,0048	0,2098
4,6	-110,57	13,8761	1,83842	0,09734	-0,0047	-0,0052	-0,0054	-0,0058	0,20574
4,7	-161,53	21,2928	2,76836	0,02869	-0,0058	-0,0061	-0,0062	-0,0064	0,20045
4,8	-237,55	33,0124	4,27052	-0,0804	-0,0063	-0,0063	-0,0063	-0,0061	0,19435
4,9	-351,68	51,6781	6,71791	-0,2558	-0,0061	-0,0056	-0,0055	-0,0046	0,18805
5	-524,1	81,6409	10,739	-0,5413	-0,0046	-0,0033	-0,003	-0,001	0,18256
5,1	-786,25	130,116	17,4013	-1,0123	-0,001	0,00161	0,00229	0,00619	0,17951
5,2	-1187,3	209,153	28,5328	-1,7993	0,00613	0,01109	0,0124	0,01963	0,18167
5,3	-1804,8	339,027	47,2891	-3,1305	0,01949	0,02854	0,03096	0,0441	0,1938
5,4	-2761,5	554,093	79,1616	-5,4073	0,04381	0,06014	0,06458	0,08829	0,22423
5,5	-4252,7	912,991	133,784	-9,3429	0,08775	0,11717	0,12533	0,16824	0,28782
5,6	-6591,8	1516,54	228,195	-16,215	0,16721	0,22046	0,2355	0,31362	0,41132
5,7	-10283	2539,3	392,768	-28,329	0,31169	0,40876	0,43667	0,58	0,64344

Далее приводятся графики всех 5-ти расчетов.

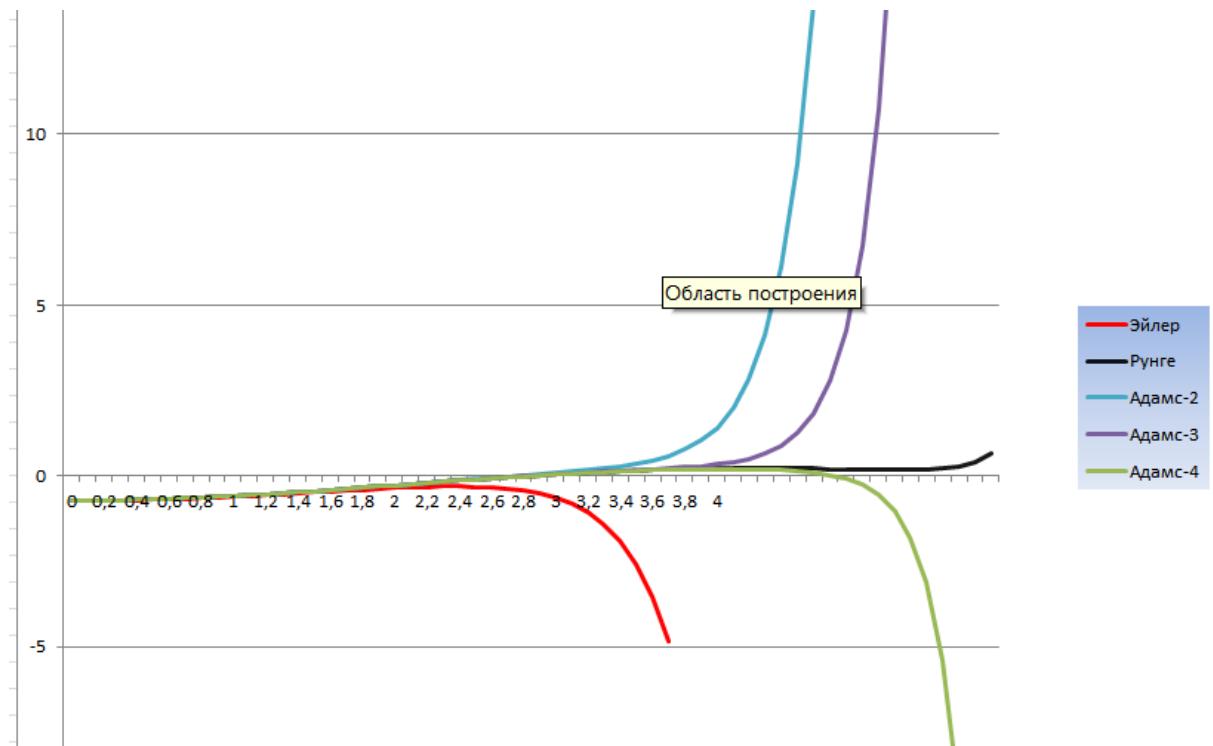


Рис. 10.2

Из-за ошибок округления каждый метод рано или поздно “уходит” с правильной, но неустойчивой траектории и в соответствии с общей картиной (рис. 10.2) выдает числа, стремящиеся или к $+\infty$ или к $-\infty$. Хорошо видно, что метод Рунге-Кутты 4-го порядка держится дольше остальных.

Приложение

Примерные задания для лабораторных работ

Лабораторная 1. Решение нелинейных уравнений. Для данного уравнения найти его наименьшее положительное решение. При этом:

1. Найти решение графически и определить примерный промежуток $[a,b]$, внутри которого находится это решение.
2. Уточнить значение решения с точностью по крайней мере до $\varepsilon=0,0001$:
 - 2.1. методом деления отрезка пополам;
 - 2.2. методом Ньютона;
 - 2.3. методом хорд;
 - 2.4. методом хорд и касательных.
3. Результаты расчетов оформить таблично с кратким описанием каждого использованного метода: расчетные формулы, выбор начального приближения, критерий остановки и пр.

Варианты уравнений:

$$1. \cos(e^x) = 2x$$

$$2. e^x - 2 = \cos x$$

$$3. \lg x - \frac{4}{2x+1} = 0$$

$$4. \operatorname{arctg} x = e^{x-2}$$

$$5. x^3 = 3x^2 + e^{-x}$$

Лабораторная 2. Решение систем линейных уравнений прямыми методами.

Дана система 5-ти линейных уравнений с 5-ю неизвестными. Система задана расширенной матрицей. Найти x^* – решение системы линейных уравнений:

- a) методом Гаусса;
- b) методом Гаусса-Жордана с выбором главного элемента;
- c) методом вращений;

Расчеты вести с 5-ю знаками после запятой.

1. Результаты расчетов представить в удобной для анализа табличной форме.
2. По окончании работы метода найти вектор невязки $r = b - Ax^*$ и вычислить евклидову норму этого вектора (невязку). Этот пункт выполнять при **максимально возможной точности вычислений**.
3. Сравнить между собой невязки каждого из методов.

Варианты систем:

Вариант 1:

-7.87675	-9.17241	4.45850	-5.22250	-3.29586		-93.18340
-2.60621	-1.41728	7.78684	-0.99298	2.94770		0.93978
7.71142	-9.99195	-1.98855	8.39102	-0.11791		76.45125
-8.79212	0.25464	-5.29388	-1.83741	-8.55142		-67.98564
-9.11803	2.83187	-1.17358	5.43015	-9.16936		32.51328

Вариант 2:

-6.82284	-2.20142	1.03684	-4.02060	-4.18905		-6.34299
-5.66021	-0.22402	1.62147	6.58156	-2.43889		-62.96164
-8.23866	-5.91936	7.20223	9.38497	-7.03845		-42.34490
-2.87302	4.01648	-2.27926	9.94610	-0.14568		-96.40896
-0.84328	5.12315	5.64647	-0.57853	1.08819		-20.02157

Вариант 3:

5.41436	-8.29403	7.04051	1.76204	-0.06322		-31.25722
-5.08431	6.37861	-7.98968	8.00909	-0.96880		-27.11019
0.51851	2.85510	-7.87976	-3.42702	0.99878		13.59534
-3.56017	-5.31375	6.94491	4.08592	7.07920		42.27960
3.17893	5.85891	-2.96326	-3.81426	6.02081		56.78426

Вариант 4:

1.82890	-9.93993	-6.48700	-7.87187	7.05981		56.84582
2.36412	-8.54921	-1.46411	5.83528	-5.92770		-2.98713
3.94254	5.88257	7.62338	-6.27039	2.12934		18.56542
-7.35372	-1.51571	-2.11810	0.71794	7.22446		-59.50463
1.65415	2.49347	8.27974	1.78199	-8.57820		-7.05032

Вариант 5:

9.70869	-0.86937	-4.52173	4.16932	1.04572		7.11293
-1.00060	2.06168	-3.66019	9.89036	-8.00916		-37.45934
5.30261	2.56820	-6.77493	5.83074	8.92284		42.29409
-4.71349	-9.16622	-9.30521	-4.09957	-3.74384		125.56277
-6.09132	-5.98015	2.68756	0.62926	5.70233		53.01942

Лабораторная 3. Решение систем линейных уравнений итерационными методами.

1. Решить систему линейных уравнений с данной матрицей методом простых итераций.

Вариант 1:

8.80991	-0.95483	-0.69204	-0.40260	0.78086	67.94025
-0.74103	-2.08391	-0.32177	-0.42902	-0.38894	-13.55912
0.36903	-0.98447	5.35096	0.33293	0.47175	40.57542
-0.88799	0.76721	0.58170	5.78438	0.97218	24.92250
0.58465	-0.62230	-0.55053	-0.49441	6.44276	-14.98188

Вариант 2:

5.30548	-0.29884	-0.02343	0.48054	-0.75836	-27.63183
0.14728	4.43494	0.03761	-0.81619	-0.46126	22.66297
0.27071	0.79186	2.17424	-0.81959	-0.45102	-6.24065
0.71384	-0.23892	0.06848	-5.36651	0.20790	-17.12048
0.43644	0.26386	-0.64131	0.78894	8.40966	66.62392

Вариант 3:

8.51350	0.39446	-0.66490	0.45445	-0.34795	-60.26314
0.25212	-6.50428	-0.73626	-0.18772	-0.57960	4.86183
0.04948	-0.77218	-9.94521	0.51737	0.88010	-90.00371
-0.99866	0.85415	-0.43183	3.11937	-0.92642	12.14180
-0.30340	0.95270	-0.59038	-0.51717	2.37937	-20.51856

Вариант 4:

-8.22139	0.35358	-0.38384	0.54674	-0.76397	-20.96216
0.73905	-8.98091	0.65759	0.45494	0.79855	36.75401
-0.22692	-0.08985	-3.26622	0.10768	-0.81036	32.16616
0.60294	0.36461	-0.08039	-4.46351	0.88587	-19.52075
-0.29583	-0.89043	0.80726	-0.10095	-6.64256	-9.19065

Вариант 5:

5.47089	0.18875	0.52167	0.57087	-0.16997	54.77033
-0.92418	-6.56520	-0.07969	0.10413	0.87469	-27.24432
0.56052	-0.55459	6.64986	0.04636	-0.09637	-25.71957

-0.00680	0.02390	0.88105	7.32699
0.78092	-0.77935	0.28713	0.35608
			42.67659
			0.59491
			-2.73635
			15.76978

2. Решить систему из лабораторной 2:

- a) методом наискорейшего спуска;
- b) градиентным методом.

Лабораторная 4. Вычисление обратной матрицы. Данна квадратная матрица A 5-го порядка. (A – матрица коэффициентов при неизвестных в **лабораторной работе № 2**). Необходимо:

- a) Найти A^{-1} – обратную к A матрицу методом линейных преобразований и методом окаймления.
- б) Для обоих методов проверить выполнение матричных равенств: $AA^{-1}=E$ и $A^{-1}A=E$, где E – единичная матрица соответствующего порядка.

Расчеты вести с 5-ю знаками после запятой.

- в) Для обоих вариантов A^{-1} вычислить решение системы из лабораторной 2 по формуле $X = A^{-1} b$. Вычислить невязки этих решений и сравнить с невязками из лабораторной 2.

Лабораторная 5. Метод наименьших квадратов и первая трансформация Гаусса. Данна функция на промежутке $[a,b]$. Требуется:

1. Взять одномерную равномерную сетку
 $\eta_x = \{x_i / x_i = x_{i-1} + h, i = 1, 2, 3, 4, 5; x_0 = a, x_5 = b, h = (b-a)/5\}$
2. Вычислить значения данной функции в узлах этой сетки и составить соответствующую таблицу.
3. Методом наименьших квадратов найти кубический многочлен, наилучшим образом совпадающий с данной функцией в узлах сетки.
4. Вычислить значения :
 - а) исходной функции;
 - б) кубического многочлена из п. 3;

в точках z_i ($i=0,1,\dots,50$) данного промежутка с шагом $\Delta=(b-a)/50$ и по этим данным нарисовать два графика (в одной и той же системе координат, но разными цветами) – исходной функции и кубического многочлена.

Варианты функций:

1. $f(x) = e^{\frac{x}{3}} \sin x$ на $[2,6]$;
2. $f(x) = \sqrt{e^{-x} + e^x} \cos x$ на $[-2,6]$;
3. $f(x) = \frac{1}{2}x \sin(3x + 1)$ на $[0.5, 2.5]$;
4. $f(x) = \ln(x^2 + 1) \sin(x + 1)$ на $[0.5, 2]$;
5. $f(x) = (x + 1) \sin(1 + x^2)$ на $[-1.5, 1.5]$;

Лабораторная 6. Собственные числа симметрической и неотрицательной матриц.

1. Симметрические матрицы. Пусть A - квадратная матрица 5-го порядка – матрица системы линейных уравнений из **лабораторной 2**. Найти собственные чис-

ла матрицы $A^T A$. Использовать классический метод Якоби. После вычисления собственных чисел найти **число обусловленности** матрицы A относительно Евклидовой нормы. Вычисления рекомендуется производить в MS Excel с использованием макроса:

```

Sub Jacobi()
    Dim n, R1, C1, i, j, k As Integer
    Dim teta, c, s As Double
    Dim Aii, Ajj, Aij As Double
    Dim Bii, Bjj, Bij As Double
    Dim Aik, Ajk, Bik, Bjk As Double

    n = Cells(1, 2).Value      ' порядок матрицы
    R1 = Cells(2, 2).Value      ' координаты первой клетки исходной матрицы
    C1 = Cells(3, 2).Value      ' на текущем листе
    i = Selection.Row - R1 + 1    ' индексы текущей
    j = Selection.Column - C1 + 1    ' клетки
    If i < 1 Or i > n Then Exit Sub
    If j < 1 Or j > n Then Exit Sub
    If i = j Then Exit Sub
    Aij = Cells(R1 + i - 1, C1 + j - 1).Value
    Aii = Cells(R1 + i - 1, C1 + i - 1).Value
    Ajj = Cells(R1 + j - 1, C1 + j - 1).Value
    If Aii = Ajj Then
        teta = Application.Pi / 4
    Else
        teta = Math.Atn(2 * Aij / (Ajj - Aii)) / 2
    End If
    c = Cos(teta)
    s = Sin(teta)
    Bii = c * c * Aii - 2 * s * c * Aij + s * s * Ajj
    Bjj = s * s * Aii + 2 * s * c * Aij + c * c * Ajj
    Bij = (c * c - s * s) * Aij + s * c * (Aii - Ajj)  ' Здесь всегда должен получаться 0
    For k = 1 To n
        If k <> i And k <> j Then
            Aik = Cells(R1 + i - 1, C1 + k - 1).Value
            Ajk = Cells(R1 + j - 1, C1 + k - 1).Value
            Cells(R1 + i - 1, C1 + k - 1).Value = c * Aik - s * Ajk
            Cells(R1 + j - 1, C1 + k - 1).Value = s * Aik + c * Ajk
            Cells(R1 + k - 1, C1 + i - 1).Value = Cells(R1 + i - 1, C1 + k - 1).Value
            Cells(R1 + k - 1, C1 + j - 1).Value = Cells(R1 + j - 1, C1 + k - 1).Value
        Else
            If k = i Then
                Cells(R1 + i - 1, C1 + k - 1).Value = Bii
                Cells(R1 + j - 1, C1 + k - 1).Value = Bij
                Cells(R1 + k - 1, C1 + j - 1).Value = Bij
            End If
            If k = j Then
                Cells(R1 + j - 1, C1 + j - 1).Value = Bjj
            End If
        End If
    Next k
End Sub

```

Краткая инструкция по использованию макроса

1. Открываем MS Excel.
2. Сохраняем документ как книгу с поддержкой макросов
3. В макросах создаем процедуру описанного выше содержания.
4. Назначаем ей комбинацию клавиш (например, Ctrl-z)
5. На любом листе в ячейках B1,B2,B3 записываем порядок матрицы, и координаты ее верхнего левого угла.
6. Заносим на этом же листе в соответствующее место саму матрицу.

	A	B	C	D	E	F	G
1	Порядок матрицы	5					
2	Строка	2	3,120990	-0,062560	0,493110	-1,130480	-0,280940
3	Колонка	3	-0,062560	2,056520	-1,925300	1,911630	1,782140
4			0,493110	-1,925300	4,539030	1,704880	2,251040
5				-1,130480	1,911630	1,704880	1,456140
6					-3,661260	3,537330	
7							

7. Для осуществления 1 шага метода Якоби ищем наибольший по модулю (или какой угодно) не диагональный элемент матрицы, щелкаем по нему, чтобы сделать текущим, и запускаем макрос клавишами Ctrl-z. Делаем так, пока матрица не станет диагональной.

	A	B	C	D	E	F	G	
1	Порядок матрицы	5						
2	Строка	2	2,680004	0,000000	0,000000	0,000000	0,000000	
3	Колонка	3	0,000000	3,890005	0,000000	0,000000	0,000000	
4			0,000000	0,000000	5,450003	0,000000	0,000000	
5			0,000000	0,000000	0,000000	-4,160000	0,000000	
6			0,000000	0,000000	0,000000	0,000000	6,849999	
7								

2. Неотрицательные матрицы. В качестве матрицы взять турнирную таблицу для вашего варианта. Диагональные элементы и возможные пустые места заполнить нулями. Вычислить неотрицательные собственный вектор и собственное число матрицы. На основании коэффициентов – чисел из собственного вектора – пересчитать результаты участников и расставить их по местам согласно набранным очкам. Проанализировать полученные результаты с точки зрения здравого смысла.

Вариант 1.

ИТОГОВАЯ ТАБЛИЦА ШАХМАТНОГО ТУРНИРА

Класс 6-й «Б» школы № 44

№ п/п	УЧАСТНИКИ	1	2	3	4	5	6	7	8	9	10	11	12	Кол-во очков	Место
1	КОРОЛЕВ Борис	●	0	1	1/2	1/2	1	1	1	1/2	1	1/2	1	8	I
2	ПАНОВА Галина	1	●	0	1	1	1	1/2	1/2	0	1	1/2	1	7 1/2	II
3	ГРАЧЕВ Николай	0	1	●	1/2	1/2	1	1	1/2	1/2	0	1	1	7	III
4	КАРПОВ Иван	1/2	0	1/2	●	1	0	1	1/2	1/2	1	1	1/2	6 1/2	IV
5	ПТИЧКИНА Тамара	1/2	0	1/2	0	●	1	0	1/2	1	1	1	1/2	6	V
6	САМОЙЛОВ Лев	0	0	0	1	0	●	1	1/2	1	1/2	1	1/2	5 1/2	VI
7	БЫКОВА Валентина	0	1/2	0	0	1	0	●	1	0	1/2	1	1	5	VII
8	БРУСИЛОВСКИЙ Давид	0	1/2	1/2	1/2	1/2	1/2	0	●	1	0	1/2	1/2	4 1/2	VIII-IX
9	СИДОРОВ Петр	1/2	1	1/2	1/2	0	0	1	0	●	0	1/2	1/2	4 1/2	
10	БАРАНОВА Наталья	0	0	1	0	0	1/2	1/2	1	1	●	0	0	4	X-XI
11	ИВАНОВ Олег	1/2	1/2	0	0	0	0	0	1/2	1/2	1	●	1	4	
12	СЕМЕНОВА Анна	0	0	0	1/2	1/2	1/2	0	1/2	1/2	1	0	●	3 1/2	XII

Вариант 2.

Турнирная таблица

Открытого Первенства Крымского района среди юношеских команд 2005 – 2006 гг.р

Группа А

№ п/п	Команда	1	2	3	4	5	6	ЗМ	О	М
1	«Витязь-(05)»		0:0	9:1	7:0	11:0	2:0	29-1	13	II
2	«Новоукраинский»	0:0		3:2	13:0	10:0	5:0	31-2	13	I
3	«Троицкая»	1:9	2:3		3:1	6:0	4:3	16-16	9	III
4	«Витязь(06)-П»	0:7	0:13	1:3		5:1	0:4	6-28	3	V
5	«Мечта»	0:11	0:10	0:6	1:5		0:7	1-39	0	VI
6	«Нижнебаканская»	0:2	0:5	3:4	4:0	7:0		11-11	6	IV

Вариант 3.

78-й чемпионат Беларуси по шахматам среди мужчин
(высшая лига)

РЦОП по шахматам и шашкам, г. Минск

12-22 января 2012 г.

№	УЧАСТНИК	РЕЙТ.	М	1	2	3	4	5	6	7	8	9	10	11	12	ОЧКИ	МЕСТО
1	Ступак Кирилл	2515	МГ		½	½	1	½	1	1	½	½	½	½	½	7	IV
2	Ломако Павел	2385	ММ	½		0	½	0	½	½	0	½	½	½	0	3,5	XI
3	Ковалев Владислав	2453	МС	½	1		0	0	1	1	½	1	1	½	1	7,5	III
4	Горовец Андрей	2428	ММ	0	½	1		½	½	1	0	0	0	0	1	5,5	VI
5	Жигалко Сергей	2658	МГ	½	1	1	½		1	½	½	1	½	1	½	8	I
6	Жук Сергей	2341	МС	0	½	0	½	0		½	½	½	0	½	½	3,5	XII
7	Кочеев Александр	2353	КМ	0	½	0	0	½	½		½	0	½	½	1	4	IX
8	Жигалко Андрей	2576	МГ	½	1	½	1	½	½	½		1	½	1	½	7,5	II
9	Буланов Дмитрий	2387	МС	½	½	0	1	0	½	1	0		0	½	½	4,5	VIII
10	Федоров Алексей	2579	МГ	½	½	0	1	½	1	½	½	1		½	1	7	V
11	Гриб Роман	2303	МФ	½	½	½	0	0	½	½	0	½	½		0	3,5	X
12	Тетерев Виталий	2538	МГ	½	1	0	0	½	½	0	½	½	0	1		4,5	VII

Главный судья

Д. Лыбин, МА

Главный секретарь

В. Якушев, ВНК

Вариант 4.

Итоговая таблица
открытого интернет-турнира ГрГМУ по шахматам

№	ФИО участника	1	2	3	4	5	6	7	Очки	Место
1	Горелов Юрий		0,5	1	0	1	1	1	4,5	II
2	Наумов Антон	0,5		0	0	0	1	1	2,5	5
3	Шаланда Илья	0	1		1	0	0	1	3	4
4	Лихачев Михаил	1	1	0		0	0,5	1	3,5	III
5	Голотин Вадим	0	1	1	1		1	1	5	I
6	Ярмолич Максим	0	0	1	0,5	0		1	2,5	6
7	Мохсунзаде Алиакпер	0	0	0	0	0	0		0	7

Вариант 5.

Начало: 18 марта 2019 года

ТУРНИРНАЯ ТАБЛИЦА

Окончание: 3 апреля 2019 года

ОТКРЫТОЕ ЛИЧНОЕ ПЕРВЕНСТВО РФЯЦ-ВНИИЭФ ПО ШАШКАМ-64 2019 года

№№ п/п	Фамилия, Разряд, Год Рождения			1	2	3	4	5	6	7	8	9	10	11	12	Очки	Место	Коф.
1	Сушков Н.	М	1960		1	2	2	2	0	2	0	2	1	2	1	15	4	10
2	Периков Н.	К	1942	1		2	2	2	0	1	1	1	1	2	0	13	6	-1
3	Тихонов В.	К	1966	0	0		2	2	1	0	1	2	1	0	1	10	7	-28
4	Быстрицкий В.	2	1957	0	0	0		2	0	1	0	1	0	1	1	6	11	-81
5	Бордашов Л.	1	1937	0	0	0	0		0	0	0	2	0	2	0	4	12	-100
6	Калинин А.	К	1965	2	2	1	2	2		1	1	0	1	1	1	14	5	32
7	Клюев В.	К	1960	0	1	2	1	2	1		0	1	0	0	0	8	9	-57
8	Грачев В.	К	1964	2	1	1	2	2	1	2		2	1	2	2	18	1	62
9	Забелина М.	К	1963	0	1	0	1	0	2	1	0		0	1	0	6	10	-63
10	Веселов Р.	К	1973	1	1	1	2	2	1	2	1	2		1	1	15	2	24
11	Робкин Л.	1	1938	0	0	2	1	0	1	2	0	1	1		0	8	8	-47
12	Зиновьев Е.	К	1968	1	2	1	1	2	1	2	0	2	1	2		15	3	21

Главный судья Демин Д.А. / _____ / Коэф. турнира - 0,33 Разрядные нормы: 1 - 9

Секретарь Аношин Ю.М. / _____ / 2 - 5
КМС - 12

Лабораторная 7. QR-алгоритм поиска собственных чисел матрицы.

Дана A - квадратная матрица 5-го порядка. Найти приближенные значения собственных чисел этой матрицы двумя способами:

1. С помощью QR-алгоритма.
2. С помощью QR-алгоритма, используя метод Гивенса.

Вариант 1.

-1.1487 -21.0653 18.1302 28.6915 -11.4996
-1.5311 -4.2798 20.1932 10.7398 -13.4015
4.9964 6.9438 -4.6944 -9.2739 10.4689
-8.6885 -19.1231 17.1324 25.8498 -17.3786

7.6094	13.8471	-20.8953	-32.0010	10.4931
--------	---------	----------	----------	---------

Вариант 2.

2.9037	2.1752	2.6861	2.5673	-2.8480
3.5047	3.6549	1.7207	0.2288	-6.1424
0.5252	-7.2551	3.1176	-3.7103	4.4693
6.9199	6.3500	-0.3184	2.2722	-5.3208
-7.0468	0.2363	-3.7909	0.5683	-1.0183

Вариант 3.

5.1386	4.2887	0.4370	-0.6305	-2.7216
-0.3279	1.9246	-0.6061	1.9368	-1.1963
0.8400	-5.7913	5.3710	3.0570	6.0899
-3.0713	1.4957	-2.8300	2.6506	0.2756
2.4814	-4.4991	3.3529	-3.9351	2.2551

Вариант 4.

6.1256	-4.2546	5.1051	2.7711	2.4107
-4.3329	5.5557	0.1397	5.2104	0.0607
5.0526	0.1104	3.9439	-1.9546	4.1527
5.1070	2.5357	6.0631	-3.5967	-3.1725
-3.5778	-0.1705	-0.5140	-1.1002	1.3114

Вариант 5.

-1.9000	8.0311	-2.7397	4.6043	-1.0171
-2.6682	-4.2377	3.4009	1.4424	2.0327
9.4031	-12.1602	1.8117	-12.6970	1.1919
-3.0446	15.4255	-4.3457	4.6601	-0.4726
4.0072	-13.1586	4.5869	-5.1222	0.3558

Лабораторная 8. Задача одномерной безусловной оптимизации.

Унимодальная на всей числовой оси функция задана формулой

$$f(x) = \frac{2(n+k)}{n^2+1} \left(\frac{k+1}{2k+20} \left(x - n - \frac{n}{k} \right) - \frac{k+10}{k+1} \cos\left(\frac{k+1}{k+10}\left(x - n - \frac{n}{k}\right)\right) \right)$$

Требуется найти ее точку минимума и минимальное значение:

- 1) методом половинного деления;
- 2) методом золотого сечения;

Вариант 1. $n=6, k=11$;

Вариант 2. $n=7, k=12$;

Вариант 3. $n=5, k=19$;

Вариант 4. $n=6, k=18$;

Вариант 5. $n=7, k=17$;

Лабораторная 9. Задача многомерной безусловной оптимизации. Дано расширенная матрица системы двух уравнений с пятью неизвестными. системы. Требуется найти такое решение $(x_1, x_2, x_3, x_4, x_5)$ системы, для которого величина

$$D = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2$$

минимальна. Сначала - привести задачу к задаче безусловной оптимизации. Для этого нужно выразить из равенств системы неизвестные x_4, x_5 через x_1, x_2, x_3 . Подставив их в выражение для D получим функцию $F(x_1, x_2, x_3)$, которую нужно минимизировать с помощью метода сопряженных направлений.

Вариант 1:

$$\begin{array}{cccccc|c} 0.1115 & -4.1716 & 4.4212 & -4.0110 & 3.3712 & | & 44.4143 \\ -4.7557 & 5.3207 & 3.9126 & 2.5842 & 1.1880 & | & -11.3978 \end{array}$$

Вариант 2:

$$\begin{array}{cccccc|c} 2.0084 & 5.0014 & 2.7403 & -4.1050 & 8.5919 & | & 41.7688 \\ -8.4284 & -3.3859 & 7.2838 & 0.3470 & -9.9042 & | & -50.6511 \end{array}$$

Вариант 3:

$$\begin{array}{cccccc|c} 4.6121 & -2.3639 & 9.4271 & -7.2109 & -0.7937 & | & -14.5270 \\ 4.1903 & -5.4093 & -5.9082 & 1.6574 & 8.4122 & | & 34.7463 \end{array}$$

Вариант 4:

$$\begin{array}{cccccc|c} 9.8015 & 4.4701 & 9.5468 & 1.3571 & -5.9697 & | & -214.9842 \\ 8.3447 & 9.9503 & -6.3799 & -5.9750 & 2.9703 & | & -111.5829 \end{array}$$

Вариант 5:

$$\begin{array}{cccccc|c} -7.9219 & -9.8583 & -0.4185 & 7.4381 & -5.0447 & | & -127.5052 \\ 3.8386 & 0.1328 & -0.9765 & -3.7675 & 9.4208 & | & 85.1366 \end{array}$$

Лабораторная 10. Вычисление интегралов. Дан интеграл

$$I = \int_a^b f(x) dx$$

Задание

1. Вычислить I по формуле трапеций, разбивая промежуток интегрирования на k частей при $k = 5$.
2. Вычислить интеграл по формуле Симпсона, разбивая промежуток интегрирования на $2k$ частей при $k = 5$.
3. Оценить точность вычисления в п.2 с помощью правила Рунге.
4. Добиться точности **до 5 знаков** после запятой, производя пересчеты по формуле Симпсона, увеличивая при необходимости число узлов в сетке.
5. Вычислить интеграл на всем промежутке $[a, b]$ с помощью формулы Гаусса наивысшей алгебраической степени точности. Использовать для этого полиномы Лежандра степени 8 (значения его корней и соответствующие коэффициенты).
6. Сравнить результаты вычисления по формулам Гаусса и по формуле Симпсона.

Вариант 1. $f(x) = \arctan x + \frac{x}{x^2+1}$ на $[-2, 1]$

Вариант 2. $f(x) = 2x \sin x + (x^2 + 1) \cos x$ на $[-1, 2]$

Вариант 3. $f(x) = e^x (\cos x + x \cos x - x \sin x)$ на $[-1, 1]$

Вариант 4. $f(x) = \frac{(1-3x^2)}{2(\sqrt{x}(x^2+1)^2)}$ на $[1, 2]$

Вариант 5. $f(x) = \frac{x}{\sqrt{x^2+1}}$ на $[0, 1.25]$

Лабораторная 11. Численное решение обыкновенных дифференциальных уравнений. Для данного дифференциального уравнения решить задачу Коши следующими способами:

1. Методом Эйлера.
2. Методом Рунге-Кутты 4-го порядка.
3. Явным методом Адамса 4-го порядка.

Решения изобразить на графике, проанализировать и сделать соответствующие выводы.

Вариант 1. $y'(x^2 - 1) = e^x y$, на $[2, 4]$ с шагом $h=0.1$, $y(2) = -1$.

Вариант 2. $y'(\sin^2 x + 1) = e^x y$, на $[0, 3]$ с шагом $h=0.2$, $y(0) = 0$.

Вариант 3. $y'x - y \ln x = x^2$ на $[2, 4]$ с шагом $h=0.1$, $y(2) = -1$.

Вариант 4. $y'(x^2 - 1) = e^x y$, на $[2, 4]$ с шагом $h=0.1$, $y(2) = -1$.

Вариант 5. $y'e^x + xy = x$, на $[-1, 1]$ с шагом $h=0.2$, $y(-1) = 0$.

Литература

1. Середа А.-В.И. Вычислительная математика : учебное пособие по дисциплине "Вычислительная математика" для студентов технических специальностей (полный курс обучения), Мурманск : Изд-во МГТУ, 2007.
2. Боглаев Н.С., Жидков Н.П., Кобельков Г.П. Численные методы, 6-е издание, Москва, БИНОМ, Лаборатория знаний, 2008. – 636 с.
3. Вержбицкий В.М. Основы численных методов: Учебник для вузов. М.: Высш.шк., 2002. – 840 с.
4. Гантмахер Ф.Р. Теория матриц. – 5-е изд., - М., ФИЗМАТЛИТ, 2010. – 560 с.
5. Шарый С.П. Курс вычислительных методов. Новосибирск – 2024. - 771 с.
6. Самарский А.А., Гулин А.В. Численные методы: Учеб. пособие. – М.: Наука. Гл.ред. физ-мат. лит., 1989. – 432 с.
7. Березин И.С., Жидков Н.П. Методы вычислений. т. 1, М.: Гл.ред. физ-мат. лит., 1962. – 464 с.
8. Березин И.С., Жидков Н.П. Методы вычислений. т. 2, М.: Гл.ред. физ-мат. лит., 1959. – 620 с.