Ю. П. Петров

Как получать НАДЕЖНЫЕ РЕШЕНИЯ систем уравнений

bhv®

Ю. П. Петров

Как получать НАДЕЖНЫЕ РЕШЕНИЯ систем уравнений

Санкт-Петербург
«БХВ-Петербург»
2009

УДК 681.3.06
ББК 32.973
П30

Петров Ю. П.

П30 Как получать надежные решения систем уравнений. — СПб.: БХВ-Петербург, 2009. — 176 с.: ил.

ISBN 978-5-9775-0450-8

Необходимость вычислять решения систем алгебраических уравнений встречается во многих задачах техники и физики, и без точных оценок возможной погрешности решения не надежны. В книге изложены методы и алгоритмы, впервые позволяющие дать точную оценку погрешности каждой из составляющей вектора решений системы линейных алгебраических уравнений, тогда как ранее были известны только приближенные оценки.

Для студентов, аспирантов, инженеров, научных работников и специалистов, выполняющих расчеты, включающие системы алгебраических уравнений

УДК 681.3.06
ББК 32.973

Группа подготовки издания:

Главный редактор	*Екатерина Кондукова*
Зам. главного редактора	*Татьяна Лапина*
Зав. редакцией	*Григорий Добин*
Компьютерная верстка	*Ольги Сергиенко*
Корректор	*Зинаида Дмитриева*
Дизайн обложки	*Елены Беляевой*
Зав. производством	*Николай Тверских*

Рецензенты:

Блехман И. И., д. ф.-м. н., профессор НИИ "Механобр";
Игнатьев М. Б., д. т. н., профессор Санкт-Петербургского университета аэрокосмического приборостроения, лауреат Государственной премии;
Ушаков А. В., д. т. н., профессор Санкт-Петербургского университета информационных технологий, механики и оптики

Лицензия ИД № 02429 от 24.07.00. Подписано в печать 16.04.09.
Формат $60\times90^1/_{16}$. Печать офсетная. Усл. печ. л. 11.
Тираж 1500 экз. Заказ №
"БХВ-Петербург", 190005, Санкт-Петербург, Измайловский пр., 29.

Санитарно-эпидемиологическое заключение на продукцию
№ 77.99.60.953.Д.003650.04.08 от 14.04.2008 г. выдано Федеральной службой по надзору в сфере защиты прав потребителей и благополучия человека.

Отпечатано с готовых диапозитивов
в ГУП "Типография "Наука"
199034, Санкт-Петербург, 9 линия, 12

ISBN 978-5-9775-0450-8

Оглавление

Список примеров

Предисловие

В данном учебном пособии приведены — в наиболее простой и доступной форме — результаты исследований автора в области надежности расчета объектов физики и техники, описываемых системами линейных алгебраических уравнений[1].

Проведенное исследование показало, что традиционные и повсеместно используемые методы расчета не всегда и не для всех объектов дают надежные результаты. Результаты расчета могут не соответствовать реальному поведению рассчитываемых объектов, и это служит причиной многих аварий и даже катастроф.

Одной из причин этих аварий и катастроф является недостаточно полное исследование проблемы зависимости решений уравнений от неизбежной на практике неточности в задании их коэффициентов. Для реальных объектов эти коэффициенты не могут идеально точно соответствовать значениям, заданным при проектировании и расчете. Отклонения неизбежны. Кроме того, в ходе эксплуатации параметры объекта могут испытывать изменения (вариации), которые влияют на решения, могут существенно изменить их и привести к авариям.

В данной работе влияние происходящих по разным причинам вариаций коэффициентов и параметров на решения уравнений, на поведение рассчитываемых объектов было подвергнуто дополнительному исследованию, которое привело к новым (иногда — неожиданным) результатам.

Данное учебное пособие предназначено для студентов, а также для инженеров и специалистов, желающих повысить свою квалификацию. В нем предложены новые методы оценки влияния вариаций параметров на поведение исследуемых объектов, позволяющие повысить надежность результатов расчета и тем самым уменьшить вероятность аварий.

---

[1] Ранее, в [2; 4] подобное исследование было проведено относительно объектов, описываемых дифференциальными уравнениями.

Отметим сразу, что вариации параметров не являются, разумеется, единственной причиной погрешностей расчета. Кроме них есть и другие причины: погрешности вычислительных методов, влияние ошибок округления и т. п. Такие погрешности могут быть уменьшены, и методам их уменьшения посвящена довольно богатая литература.

Предлагаемое учебное пособие посвящено другой, совсем другой теме. Оно посвящено неустранимым и неуменьшаемым погрешностям решений, возникающим из-за вариаций исходных данных расчета, данных о параметрах исследуемого объекта, которые в реальных условиях почти всегда являются приближенными и заданы чаще всего лишь интервалом своих возможных значений. К этой исходной неточности знаний о параметрах объекта и о коэффициентах его математической модели добавляются вариации коэффициентов и параметров, возникающие в ходе эксплуатации объекта и увеличивающие интервал значений коэффициентов. Каков будет интервал, внутри которого лежат решения системы, если известен интервал изменения ее коэффициентов и параметров? — вот основная проблема, рассмотренная в данной книге. Этой проблеме в последние годы заслуженно уделяется все больше и больше внимания, поскольку стало ясно, что без учета вариаций коэффициентов и параметров, без точной оценки величины интервала, на границах которого может оказаться решение, нельзя говорить о надежности результатов расчета. В данном учебном пособии приведен и проиллюстрирован примерами разработанный автором алгоритм вычисления точной оценки погрешности решения, порожденной вариациями коэффициентов.

Эта работа продолжает линию предыдущих исследований в области прикладной математики, учитывающих неточное знание коэффициентов уравнений или законов их изменения, учитывающих возможность их вариаций, и т. п. Публикаций по этому направлению исследований накопилось уже достаточно много (см. [1; 2; 4; 7; 8; 10; 11; 15; 16; 17; 18] и многие другие). По-видимому, настало время назвать это направление исследований "Математикой-2", в отличие от "Математики-1", исследующей математические модели с точно заданными коэффициентами и законами их изменения.

"Математика-2" отличается от "Математики-1" как предметом исследования, так и его методами. Главное (недавно обнаруженное) отличие: в привычной нам "Математике-1" очень широко используются эквивалентные (равносильные) преобразования, которые не изменяют решений, но упрощают исследование (примеры таких преобразований: умножение или деление всех членов уравнения на число, не равное нулю, замена члена уравнения на равный ему и т. п.). В работе [2] было показано, что эквивалентные преобразования, не изменяя самих решений как таковых, могут изменять важные свойства решений — в том числе степень их зависимости от вариаций коэффициентов и параметров. Поэтому в "Математике-2" эквивалентными преобразованиями можно пользоваться лишь с большой осторожностью, что и отличает ее от "Математики-1".

Автор благодарен И. А. Петрову за помощь в подготовке рукописи.

Пожелания, замечания и возражения по тексту учебного пособия можно направлять на адрес: *petrov*1930@*mail.ru*.

§ 1. Примеры расчета

Расчет многих технических объектов сводится к составлению и решению систем линейных алгебраических уравнений — систем вида:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= b_n, \end{aligned}\right\} \qquad (1)$$

где x_1; x_2; ...; x_n — неизвестные, подлежащие вычислению, a_{ij} — коэффициенты системы, b_1; b_2; ...; b_n — коэффициенты правой части.

Решением системы (1) является n-мерный вектор X — т. е. набор n чисел x_1; x_2; ...; x_n, который при подстановке его в левую часть системы (1) обращает все уравнения системы в тождества.

В дальнейшем основное внимание будет уделяться не решению X в целом, а наиболее важной для практики задаче — оценке погрешности вычисления каждой из составляющих решения, каждого из чисел x_1; x_2; ...; x_n. Для сокращения письма каждую из составляющих часто будем коротко называть "решением" вместо полного названия — "составляющая решения X". Надеемся, что такая небольшая вольность в терминологии не приведет к путанице.

Рассмотрим простой пример (**пример № 1**): пусть балка AB лежит на двух опорах, как показано на рис. 1.

Длина балки левее левой опоры равна l_1 метров с удельной нагрузкой на единицу длины q_1 кН/м (килоньютонов на метр). Длина участка между опорами равна l_2 метров с удельной нагрузкой q_2 кН/м. Длина балки правее правой опоры равна l_3 метров с удельной нагрузкой q_3 кН/м. Нагрузку на левую опору обозначим через x_1, на правую — через x_2. Опоры предполагаем неудерживающими. Это означает, что если, например, нагрузка x_1 оказа-

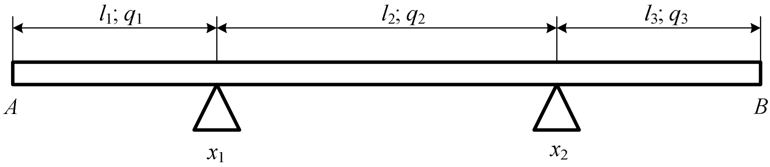

Рис. 1

лась отрицательной, то балка потеряет статическую устойчивость, соскользнет с правой опоры и упадет. Значения $x_1 = 0$ или $x_2 = 0$ означают, что балка лежит на границе статической устойчивости. Для отыскания x_1 и x_2 можно составить уравнения равновесия моментов сил относительно точек A и B (левого и правого концов балки).

Относительно правого конца уравнение равновесия моментов имеет вид:

$$(l_2 + l_3)x_1 + l_3 x_2 = (l_1 q_1 + l_2 q_2 + l_3 q_3) \cdot (l_1 + l_2 + l_3) \cdot \frac{1}{2} \tag{2}$$

Относительно левого конца уравнение равенства моментов сил имеет вид:

$$l_1 x_1 + (l_1 + l_2)x_2 = (l_1 q_1 + l_2 q_2 + l_3 q_3) \cdot (l_1 + l_2 + l_3) \cdot \frac{1}{2} \tag{3}$$

В уравнениях (2) и (3) величины l_1; l_2; l_3; q_1; q_2; q_3 — это *параметры* объекта (в данном случае — балки).

Мы убеждаемся, что коэффициенты системы уравнений (1) зависят от параметров, являются функциями от них.

В данном примере система (1) является системой второго порядка ($n = 2$), и в ней:

$$a_{11} = l_2 + l_3;$$

$$a_{12} = l_3;\ a_{21} = l_1;$$

$$a_{22} = l_1 + l_2;$$

$$b_1 = b_2 = (l_1 q_1 + l_2 q_2 + l_3 q_3) \cdot (l_1 + l_2 + l_3) \cdot \frac{1}{2}.$$

Размерности левой и правой частей уравнений (2) и (3) одинаковы и равны размерности момента, т. е. произведению килоньютон на метр.

Другим примером объекта, поведение которого описывается системой уравнений (1), может служить разветвленная электрическая цепь. Как известно, на основании второго закона Кирхгофа для цепи, состоящей из n контуров с сопротивлениями, не завясящими от токов и напряжений, можно составить систему из n уравнений:

$$\left.\begin{array}{l} r_{11}i_1 + r_{12}i_2 + \ldots + r_{1n}i_n = E_1; \\ r_{21}i_1 + r_{22}i_2 + \ldots + r_{2n}i_n = E_2; \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ r_{n1}i_1 + r_{n2}i_2 + \ldots + r_{nn}i_n = E_n, \end{array}\right\} \qquad (4)$$

где i_1; i_2; ...; i_n — контурные токи, r_{ij} — контурные сопротивления, E_1; E_2; ...; E_n — контурные электродвижущие силы. Размерность левых и правых частей уравнений (4) выражается в вольтах.

Помимо объектов, математическими моделями которых с самого начала являются системы линейных алгебраических уравнений — типа систем (2)–(3) и (4), — подобные системы возникают при расчете объектов, математическими моделями которых являются уравнения в частных производных, или интегральные уравнения.

Поясним сказанное на примере интегрального уравнения Фредгольма:

$$\int_a^b K(x;s)y(s)ds = f(x),$$

где $K(x;\ s)$ — ядро уравнения, функция двух переменных x и s, отражающая свойства исследуемого объекта. Переменные x и s изменяются в пределах: $a \le s \le b$; $c \le x \le d$. При этом $y(s)$ — искомая функция, $f(x)$ — заданная функция, отражающая обычно внешние воздействия на объект.

Распространенный метод решения интегральных уравнений — это сведение их к системам вида (1). Для этого вводят равномерные сетки узлов с шагом Δs по переменной s и шагом Δx по переменной x. Заменяя непрерывную функцию $K(x; s)$ ее значениями в узлах, приводим интегральное уравнение к системе алгебраических уравнений:

$$A_{i,j} y_j = f_i ,$$

где $i = 1, 2, ..., m$; $j = 1, 2, ..., n$.

Интегральные уравнения, как известно, описывают очень многие объекты физики и техники (примеры приведены в [8], [9], [32], [33], [34]). Исследование каждого из них приводит к системе вида (1), коэффициенты которой зависят от свойств объекта. Поскольку функции $K(x; s)$ и $f(x)$ почти всегда известны только приближенно и не могут оставаться идеально неизменными с течением времени или в ходе эксплуатации технического объекта, то необходимо учитывать, что и все коэффициенты описывающих исследуемые объекты системы уравнений вида (1) также могут быть известны только приближенно, и это необходимо учитывать при расчетах.

Для объектов, описываемых дифференциальными уравнениями в частных производных, также часто возникает необходимость решать системы уравнений вида (1).

Известны и многие другие проблемы, в которых одним из важных этапов является вычисление решения той или иной системы линейных алгебраических уравнений и оценка погрешности этого решения, происходящей из-за вариаций его коэффициентов. Автор известной публикации [25] обоснованно считает, что в настоящее время "более 70% всех математических расчетов приходится на вычисление решений систем линейных алгебраических уравнений". Это подчеркивает важность вопросов, рассматриваемых в данном учебном пособии.

В дальнейшем мы будем в основном рассматривать примеры из строительной механики — например, уравнения, возникающие при расчете сил и нагрузок в элементах конструкций. Это сделано потому, что примеры из строительной механики нагляднее. Однако необходимость решать системы линейных алгебраических уравнений и оценивать возможную погрешность их решений при вариациях коэффициентов и параметров возникает во многих областях техники и физики. Поэтому значение рассмотренных примеров много шире.

§ 2. Исследование зависимости решений от параметров

Рассмотрим более подробно конкретный пример (**пример № 2**), когда в системе (2)–(3) будет $l_1 = l_2 = 2$ м; $l_3 = 3{,}8$ м; $q_1 = q_2 = q_3 = 1$ кН/м. Система (2)–(3) в этом случае примет вид:

$$\left.\begin{aligned} 5{,}8x_1 + 3{,}8x_2 &= 30{,}42; \\ 2x_1 + 4x_2 &= 30{,}42. \end{aligned}\right\} \tag{5}$$

Будем решать эту систему по известным формулам Крамера:

$$x_1 = \frac{D_1}{D}; x_2 = \frac{D_2}{D}, \tag{6}$$

где D — определитель системы, D_1 — тот же определитель, но в котором первый столбец заменен на столбец коэффициентов правой части системы, D_2 — тот же определитель системы, но в котором теперь уже второй столбец заменен на столбец коэффициентов правой части.

Для системы (5) будет:

$$D = \begin{vmatrix} 5{,}8 & 3{,}8 \\ 2 & 4 \end{vmatrix} = 5{,}8 \cdot 4 - 3{,}8 \cdot 2 = 15{,}6\,;$$

$$D_1 = \begin{vmatrix} 30{,}42 & 3{,}8 \\ 30{,}42 & 4 \end{vmatrix} = 30{,}42 \cdot \begin{vmatrix} 1 & 3{,}8 \\ 1 & 4 \end{vmatrix} = 30{,}42 \cdot (4 - 3{,}8) = 6{,}084\,;$$

$$D_2 = \begin{vmatrix} 5{,}8 & 30{,}42 \\ 2 & 30{,}42 \end{vmatrix} = 30{,}42 \cdot \begin{vmatrix} 5{,}8 & 1 \\ 2 & 1 \end{vmatrix} = 30{,}42 \cdot (5{,}8 - 2) = 111{,}6\,,$$

откуда следует, что $x_1 = \dfrac{6{,}084}{15{,}6} = 0{,}39$ кН; $x_2 = \dfrac{111{,}6}{15{,}6} = 7{,}41$ кН.

Рассмотрим теперь как изменятся решения при изменении параметров системы. Этих параметров — шесть (l_1; l_2; l_3; q_1; q_2; q_3).

Пусть, например, q_3 увеличился на 5% и стал равным 1,05 кН/м. Пользуясь зависимостями, связывающими параметры с коэффициентами системы, т. е. зависимостями:

$$\left.\begin{aligned} &a_{11} = l_2 + l_3; a_{12} = l_3; \\ &a_{21} = l_1; a_{22} = l_1 + l_2; \\ &b_1 = b_2 = \frac{1}{2}(l_1 + l_2 + l_3)(l_1 q_1 + l_2 q_2 + l_3 q_3) \end{aligned}\right\} \quad (7)$$

получим, что система (5) при $q_3 = 1{,}05$ перейдет в систему:

$$\left.\begin{aligned} 5{,}8x_1 + 3{,}8x_2 = 31{,}161; \\ 2x_1 + 4x_2 = 31{,}161 \end{aligned}\right\} \quad (8)$$

с решениями $x_1 = 0{,}395$; $x_2 = 7{,}59$. Таким образом, вариация параметра q_3 на +5% изменила только коэффициенты правой части системы (5). Они выросли на 2,44%. Решения x_1 и x_2 также выросли на 2,44%.

Если параметр l_1 изменился на –5% (т. е. вместо $l_1 = 2$ м стало $l_1 = 1{,}9$ м), то система (5) перейдет в систему:

$$\left.\begin{aligned} 5{,}8x_1 + 3{,}8x_2 = 29{,}645; \\ 1{,}9x_1 + 3{,}9x_2 = 29{,}645 \end{aligned}\right\} \quad (9)$$

с решениями $x_1 = 0{,}195$; $x_2 = 7{,}608$.

Таким образом, в данном случае коэффициенты a_{11} и a_{12} остались без изменения, коэффициент a_{21} изменился на –5%, a_{22} — на –2,5%, правые части изменились на –2,54%, решение x_1 изменилось на –50%, а x_2 изменилось на +2,7%.

Если произойдет изменение сразу двух параметров: l_1 изменится на +5% и l_3 — тоже на +5%, то система (5) перейдет в систему

$$\left.\begin{aligned} 5{,}99x_1 + 3{,}99x_2 = 32{,}724; \\ 2{,}1x_1 + 4{,}1x_2 = 32{,}724 \end{aligned}\right\} \quad (10)$$

с решениями $x_1 = 0{,}225$; $x_2 = 7{,}87$.

В данном случае коэффициент a_{11} изменился на +3,28%, a_{12} — на +5%, a_{21} — на +5%, a_{22} — на +5%, b_1 и b_2 на +7,56%. Решение x_1 изменилось на –42,4%, решение x_2 — на +6,2%.

Изменения параметров могут происходить и в разные стороны. Если, например, l_1 изменилось на –5% (вместо 2 стало l_1 = 1,9 м), а l_3 изменилось на +5%, то система (5) перейдет в систему

$$\left.\begin{aligned} 5{,}99x_1 + 3{,}99x_2 &= 31{,}126; \\ 1{,}9x_1 + 3{,}9x_2 &= 31{,}126 \end{aligned}\right\} \qquad (11)$$

с решениями $x_1 = -0{,}176$; $x_2 = 8{,}06$.

В данном случае при тех же абсолютных величинах вариаций параметров l_1 и l_3 решение x_1 изменилось коренным образом — изменило свой знак. Вместе с этим изменилось и все поведение исследуемого объекта: балка *AB* уже не может устойчиво лежать на двух (неудерживающих) опорах. Она соскользнет с правой опоры и упадет.

Анализ простого примера с системой (5) позволяет сделать общие выводы:

при исследовании влияний вариаций параметров на поведение объекта необходимо исследовать все возможные сочетания положительных и отрицательных вариаций. Если поведение объекта зависит от n параметров, то нужно составить, в общем случае, 2^n систем уравнений, а затем найти и проверить решения каждой из 2^n систем.

Для рассмотренного примера № 1 с балкой будет $n = 6$ и $2^n = 2^6 = 64$. Поведение более сложного объекта может зависеть, например, от 20 параметров, и тогда придется составлять и решать

$$2^{20} = 2^{10} \cdot 2^{10} = 1024 \cdot 1024 = 1048576$$

систем уравнений, что затруднительно даже при использовании быстродействующей вычислительной техники.

Разумеется, в каждом отдельном случае, используя особенности решаемой задачи, можно сократить число необходимых вычислений. Так, в рассматриваемом примере с балкой можно вместо второго уравнения равновесия моментов (3) использовать уравнение равновесия сил:

$$x_1 + x_2 = (l_1 q_1 + l_2 q_2 + l_3 q_3) \tag{12}$$

и решать для определения x_1 и x_2 систему уравнений (2)–(12).

В уравнении (12) коэффициенты при x_1 и x_2 всегда равны единице и не зависят от параметров балки (то есть от l_i и q_i), что сокращает объем вычислений.

Однако в целом сохраняет силу общий и важный вывод:

> решение задачи о поведении того или иного объекта при учете возможных вариаций его параметров является значительно более сложной проблемой (часто на много порядков более сложной проблемой), чем решение при заданных и неизменных значениях параметров.

Этот вывод сохраняет силу и при учете специальных методов, разработанных для учета вариаций параметров (таких, как теория чувствительности, описанная в работах [1], [7] и т. д.). Некоторые из этих методов в дальнейшем подвергнутся критическому разбору и будет показано, что они обладают серьезными, ранее не всегда замечаемыми, недостатками. Будут предложены новые методы.

Полное решение проблемы расчета с учетом вариаций коэффициентов и параметров состоит из двух частей:

- первая часть проблемы: анализируя свойства исследуемого объекта при вариациях параметров, надо оценить: в каких пределах может изменяться каждый из коэффициентов математической модели объекта, какой может быть вариация каждого из коэффициентов;
- вторая часть проблемы: опираясь на исследование вариаций коэффициентов, считая их известными и заданными, оценить вариации решений математической модели.

Первая часть проблемы — это чисто инженерная задача, решение которой требует хорошего знания свойств и особенностей конкретного исследуемого объекта.

Вторая часть проблемы относится уже к прикладной математике. Решение второй части проблемы требует исследования свойств используемого класса математических моделей (например — систем алгебраических уравнений $AX = B$). Решение второй части проблемы значимо для специалистов широкого круга специальностей, для всех тех, кто использует рассматриваемый класс математических моделей.

В данном учебном пособии — как и в других учебниках и руководствах, посвященных методам вычислений, таких, как [4; 5; 6; 24; 25; 26; 27; 28; 32; 33 и др.], — не рассматривается первая часть проблемы. Вариации коэффициентов предполагаются известными и заданными, исследуются зависящие от них вариации решений. При этом исследованию подвергаются самые различные варианты знания о вариациях коэффициентов:

- исследуется вариант, когда абсолютные величины вариаций ограничены сверху единым числом ε_0, а знак каждой из вариаций может быть любым;
- исследуется вариант, когда вариации испытывает только часть коэффициентов;
- исследуется вариант относительных вариаций (в долях от первоначальной величины коэффициента) и вариант абсолютной вариации, когда к коэффициенту a_i прибавляется вариация ε_i, не зависящая от a_i;
- рассматриваются варианты связанных вариаций, когда вариации одних коэффициентов зависят от вариаций других.

Автор надеется, что рассмотрение всех этих вариантов охватит потребности большинства специалистов, работающих в различных областях техники и физики, но одинаково использующих в своих расчетах системы алгебраических уравнений $AX = B$.

§ 3. Проблемы "Математики-2"

На совсем простом примере, рассмотренном в § 1, сразу видно, что расчеты и исследование поведения тех или иных объектов при неизменных параметрах и те же расчеты, но с учетом неизбежных на практике вариаций этих параметров, — задачи совершенно разной степени сложности и требуют — как мы увидим далее — существенно различных методов исследования.

Поэтому полезно выделить две математики: "Математику-1", ведущую исследование математических моделей реальных объектов и систем при известных и заданных значениях коэффициентов или при известных законах их изменения, и "Математику-2", не предполагающую точного знания коэффициентов и параметров, учитывающую вариации коэффициентов, законов их изменения и т. п., т. е. учитывающую вариации, которые на практике всегда неизбежны.

Достаточно очевидно, что результаты, получаемые в "Математике-2", имеют не меньшее (а иногда и большее) значение, чем результаты, получаемые в привычной и хорошо известной "Математике-1".

Но вот количество исследований, объем и глубина этих исследований в "Математике-1" неизмеримо больше, чем в "Математике-2". Поэтому многие проблемы, возникающие в "Математике-2", недостаточно исследованы и не получили пока сколько-нибудь полного и исчерпывающего решения.

Укажем на некоторые из проблем:

- выделение особенно опасных объектов, математические модели которых имеют некорректные решения, — т. е. решения, которые изменяются на конечные величины или даже изменяются коренным образом при сколь угодно малых (и тем самым почти всегда неизбежных на практике) вариациях параметров.

Для рассматриваемых нами объектов, математические модели которых являются системами алгебраических уравнений, эта проблема, как мы увидим далее, имеет простое решение;

- для объектов, математической моделью которых являются системы обыкновенных дифференциальных уравнений, та же проблема решается гораздо сложней, поскольку существуют весьма опасные "особые" объекты, которые имеют некорректные решения, но это их опасное свойство не всегда может быть обнаружено традиционными методами, не учитывающими возможности изменения корректности решения (как и других его свойств) при эквивалентных (равносильных) преобразованиях уравнений. Эта проблема была ранее рассмотрена в [2];
- выделение так называемых *плохо обусловленных* систем, в которых "малым" вариациям коэффициентов соответствуют "большие" изменения решений. Слова "малые" и "большие" в этом определении плохо обусловленных систем не случайно взяты в кавычки. Эти кавычки подчеркивают, что данное определение не является полным и точным. В каждой конкретной задаче на основе анализа ее физического смысла надо дополнительно установить — какую именно вариацию того или иного коэффициента следует считать "малой", имеющей смысл (или даже неизбежной) для данной задачи, и какие именно изменения решений следует считать "большими", ведущими к недопустимым изменениям характеристик исследуемого объекта. Только после этого определение "плохо обусловленной" системы получает точный и определенный смысл. Но и не точное, не полное (а как бы "описательное") определение плохо обусловленной системы в ряде случаев полезно;
- получение количественных оценок, связывающих величины вариаций коэффициентов и параметров, интервалы их возможных значений, с величинами изменений решений, порожденных этими вариациями, с величиной интервала возможных значений решений. Это — важная проблема; решение ее было бы очень и очень полезно для практических приложений. Однако даже для систем линейных алгебраических уравнений эта проблема пока до самого последнего времени не имела решения. Известны — и широко используются — более скромные

результаты, когда для исследуемых изменений решений при вариациях коэффициентов даются лишь их "оценки сверху", оценки в виде неравенств.

Эти оценки в ходе дальнейшего изложения мы подробно рассмотрим. Будут также предложены новые оценки, в том числе — точные оценки наибольших возможных вариаций каждой из составляющих решения, каждой из составляющих вектора X.

Мы перечислили проблемы "Математики-2", особенно существенные для рассматриваемых в данном учебном пособии математических моделей в виде систем линейных алгебраических уравнений. Для других математических моделей возникают свои проблемы.

Так, например, для объектов, поведение которых описывается системами дифференциальных уравнений, важнейшей проблемой является анализ сохранения динамической устойчивости при тех или иных вариациях коэффициентов и параметров. Эта проблема была ранее рассмотрена в [2].

Свои проблемы возникают для объектов, описываемых интегральными уравнениями (они рассмотрены в [8]) и т. д.

Еще раз отметим, что помимо вариаций коэффициентов и параметров на погрешность решения могут влиять, например, ошибки округления при численном решении. В сложных задачах количество вычислительных операций достигает значительных величин, и неизбежные ошибки округления при вычислениях могут существенно влиять на точность решения — для восстановления точности используют, как известно, итерационные методы решения, но и они, — поскольку в реальных вычислениях число итераций конечно — не свободны от погрешности. Впрочем, погрешности вычислительных методов могут быть уменьшены: за счет увеличения числа значащих цифр, за счет увеличения числа итераций и т. п. Теоретически — разумеется, за счет увеличения объема вычислений, — погрешности вычислительных методов могут быть сделаны сколь угодно малыми.

Однако в настоящем учебном пособии — подчеркнем это еще раз — исследуется совсем другая проблема — проблема оценки неизбежной погрешности решения, происходящей от вариаций коэффициентов и параметров. Эта погрешность не может быть уменьшена за счет усовершенствования вычислительных методов. Если известны вариации коэффициентов, если известны интервалы, внутри которых находятся истинные, неизвестные нам значения коэффициентов, если эти интервалы заданы, например, в форме неравенств:

$$a_{ij}(1-\varepsilon_{ij}) \le \overline{a}_{ij} \le a_{ij}(1+\varepsilon_{ij}),$$

где $\overline{a}_{ij}$ — истинное значение, а a_{ij} — значение, принятое при расчете, то можно оценить интервалы, внутри которых находятся составляющие вектора решений, но нельзя эти интервалы уменьшить. Правильная оценка интервалов, внутри которых находятся решения, является важнейшей составляющей надежности всех технических расчетов, и в последующих параграфах будет показано, какими методами можно вычислить эти интервалы и обеспечить надежность расчета.

Само вычисление решения полезно начинать с оценки его неизбежной погрешности. Если, например, из-за вариаций коэффициентов и параметров погрешность не может быть меньше, чем ±0,01, то бесполезно использовать сложные вычислительные методы, обеспечивающие точность вычисления решения до ±0,001, и следует применить методы более простые.

Отметим, что вариации параметров исследуемого объекта и зависящие от них вариации коэффициентов математической модели могут происходить по разным причинам:

- при изготовлении объекта невозможно идеально точно выполнить проектные размеры конструкции;
- свойства материалов, из которых конструкция изготовлена, имеют неизбежный разброс;
- в ходе эксплуатации из-за износа все параметры объекта и их свойства неизбежно изменяются;
- внешние нагрузки на объект могут отличаться от расчетных и т. п.

Мы будем рассматривать вариации коэффициентов и параметров, происходящие от всех этих причин.

§ 4. Вычисление решений через обратные матрицы; "числа обусловленности" решений

Решение систем линейных уравнений излагают чаще всего на языке векторов и матриц. Основную систему уравнений (1) записывают в векторно-матричной форме, и она принимает вид:

$$AX = B, \tag{13}$$

где $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$ — матрица коэффициентов, которая является квадратной матрицей и имеет размер $n \times n$; X — это решение системы, т. е. вектор-столбец чисел x_1; x_2; ...; x_n, которые мы — как уже оговаривалось — для краткости будем называть решениями; B — вектор-столбец коэффициентов правой части, коэффициентов b_1; b_2; ...; b_n.

При вычислении составляющих вектора X и анализе изменений решений при вариациях коэффициентов используются такие понятия, как обратная матрица и нормы матрицы.

Обратная матрица обозначается символом A^{-1} и ее определением служит равенство:

$$A \cdot A^{-1} = E, \tag{14}$$

где E — единичная матрица,

$$E = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}, \tag{15}$$

т. е. в единичной матрице на главной диагонали стоят единицы, а все остальные элементы — нули.

Из равенства (14) выводится формула для элементов обратной матрицы:

$$A^{-1} = \begin{pmatrix} \frac{A_{11}}{D} & \frac{A_{21}}{D} & \dots & \frac{A_{n1}}{D} \\ \frac{A_{12}}{D} & \frac{A_{22}}{D} & \dots & \frac{A_{n2}}{D} \\ \dots & \dots & \dots & \dots \\ \frac{A_{1n}}{D} & \frac{A_{2n}}{D} & \dots & \frac{A_{nn}}{D} \end{pmatrix}, \tag{16}$$

где A_{ij} — алгебраические дополнения элементов a_{ij} матрицы A, а D — определитель матрицы, т. е.

$$D = \det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}. \tag{17}$$

Таким образом, элементами обратной матрицы являются алгебраические дополнения элементов исходной матрицы, деленные на ее определитель. Его можно вынести за знак матрицы и получить более удобную формулу:

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}. \tag{18}$$

Обратная матрица может быть использована для вычисления решений системы по формуле:

$$X = A^{-1} B. \tag{19}$$

Поскольку вычисление обратной матрицы трудоемко, для отыскания решений чаще используют метод последовательного исключения переменных (метод Гаусса) или итерационные методы. Однако формула (19) широко используется в теоретических исследованиях.

Для матрицы второго порядка, размера 2×2:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{20}$$

обратной будет матрица:

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \cdot \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}, \tag{21}$$

т. е. элементы главной диагонали исходной матрицы A переставляются, элементы второй диагонали меняют знаки на противоположные, а затем все делится на $\det A = a_{11}a_{22} - a_{12}a_{21}$.

Для матрицы размера 3×3:

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

обратной будет матрица:

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix}. \tag{23}$$

Пример № 3: для матрицы $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ с определителем $\det A = 6$ обратной будет матрица:

$$A^{-1} = \frac{1}{6}\begin{pmatrix} -1 & -1 & +3 \\ +2 & -7 & +3 \\ -1 & +5 & -3 \end{pmatrix}.$$

Важным понятием является понятие *нормы* вектора и матрицы. Наиболее часто используется *евклидова*, или сферическая норма.

Для вектора B евклидова норма:

$$\|B\| = \sqrt{b_1^2 + b_2^2 + \ldots + b_n^2}\,. \tag{24}$$

Для матрицы A евклидова норма:

$$\|A\| = \sqrt{\sum_{i=1;\, j=1}^{n} a_{ij}^2}\,. \tag{25}$$

Используются и другие нормы. Так, например, *кубическая* норма вектора (следуем терминологии учебника [3]) определяется равенством:

$$\|B\|_{куб} = \max_i |b_i|, \tag{26}$$

а его *октаэдрическая* норма — равенством:

$$\|B\|_{окт} = \sum_{i=1}^{n} |b_i|\,. \tag{27}$$

Примечание

Следует особо отметить, что обозначения матриц и их норм еще не окончательно установились, и в разных книгах можно встретить различные обозначения. Еще раз напомним, что матрицу мы обозначаем вертикальными чертами с закруглением сверху и снизу (формулы (15), (16) и далее), определитель матрицы обозначается знаком det (формула (17), но при использовании формул Крамера используются также обозначения D и D_i — (формула (6)), норма матрицы обозначается двойными вертикальными чертами слева и справа. В названиях евклидовой, кубической и октаэдрической норм (формулы (24)–(27)) мы следуем учебнику [3].

Свойства евклидовой нормы:

$$\|AX\| \le \|A\| \cdot \|X\|, \tag{28}$$

т. е. норма произведения не больше произведения норм.

Аналогично:

$$\|A + X\| \le \|A\| + \|X\|, \tag{29}$$

т. е. норма суммы не больше суммы норм слагаемых. Из свойства (28) следует:

$$\|A\| \cdot \|A^{-1}\| \ge \|E\| = 1. \tag{30}$$

Рассмотрим теперь систему уравнений (13) и предположим, что коэффициенты правой части изменились, и вместо вектора B мы имеем дело с вектором $B + \Delta B$. В результате изменения правой части изменится и решение: вместо вектора решения X появится вектор $X + \Delta X$, и вместо равенства (13) будет иметь место равенство:

$$A(X + \Delta X) = B + \Delta B, \tag{31}$$

или — что то же самое:

$$AX + A\Delta X = B + \Delta B. \tag{32}$$

Вычитая из (32) равенство (13), получим:

$$A\Delta X = \Delta B,$$

откуда следует:

$$\Delta X = A^{-1}\Delta B. \tag{33}$$

Переходя к нормам, имеем:

$$\|\Delta X\| \le \|A^{-1}\| \cdot \|\Delta B\|, \tag{34}$$

откуда следует, что

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta B\|}{\|B\|}. \tag{35}$$

Произведение $\|A\| \cdot \|A^{-1}\|$ называется "числом обусловленности", и из формулы (35) следует, что относительное изменение решений не превышает относительного изменения правой части, умноженного на число обусловленности, которое — согласно свойству (30) — всегда больше единицы.

Если вариация каждого из коэффициентов b_1; b_2; ...; b_n правой части не превышает $\pm\delta b_i$ и выполняются неравенства:

$$b_i(1-\delta) \le \overline{b_i} \le b_i(1+\delta), \tag{36}$$

где $\overline{b_i}$ — истинное, неизвестное нам значение коэффициента, b_i — номинальное значение, принятое при расчете, то в этом случае $\|\Delta B\| \le \delta \|B\|,$ и формула (35) принимает вид:

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\| \cdot \|A^{-1}\| \cdot \delta. \tag{37}$$

Рассмотрим простой **пример 4**. Система:

$$\left.\begin{aligned} 2x_1 + x_2 &= 2; \\ x_1 + x_2 &= 1 \end{aligned}\right\}$$

имеет решение: $x_1 = 1$; $x_2 = 0$. Норма решения $\|X\| = 1.$ Если коэффициент b_1 изменится на одну десятую и станет равным 2,2, и коэффициент b_2 также станет равным 1,1, то решением станут $x_1 = 1{,}1$; $x_2 = 0$. Поскольку для данного примера $\|B\| = \sqrt{4+1} = \sqrt{5};$

$\|\Delta B\| = \sqrt{0,05}$; $\|A\| = \sqrt{4+1+1+1} = \sqrt{7}$; $\det A = 1$; $A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ и $\|A^{-1}\| = \sqrt{7}$, то из формулы (36) следует:

$$0,1 \le 7 \cdot \frac{\sqrt{0,05}}{\sqrt{5}} = 0,7, \text{ что верно.}$$

Если бы мы не вычисляли нового решения системы (37) при $b_1 = 2,2$ и $b_2 = 1,1$, то, используя формулу (35), могли бы утверждать, что относительное изменение его нормы — величины $\frac{\|\Delta X\|}{\|X\|}$ — не превысит значения 0,7. Действительное относительное изменение нормы решения равно 0,1.

Хотя в формуле (35) стоит знак "≤", означающий, что возможно равенство левой и правой частей формулы (35), но в действительности знак равенства может появиться лишь для некоторых, исключительно редко встречающихся матриц A и векторов B. Для подавляющего большинства векторов и матриц, встречающихся в реальных задачах, левая часть формулы (35) много меньше ее правой части и "число обусловленности" чаще всего дает очень грубую, сильно завышенную оценку изменения нормы решения при вариациях коэффициентов.

Если необходимо учитывать как вариации коэффициентов правой части, так и вариации коэффициентов матрицы A системы (13), то вместо формулы (35) используется формула:

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\| \cdot \|A^{-1}\| \cdot \left(\frac{\|\Delta B\|}{\|B\|} + \frac{\|\Delta A\|}{\|A\|} \right). \tag{38}$$

Если вариации коэффициентов матрицы A системы (13) не превышают $\pm\varepsilon a_{ij}$ и выполняются неравенства:

$$a_{ij}(1-\varepsilon) \le \overline{a}_{ij} \le a_{ij}(1+\varepsilon), \tag{39}$$

где $\overline{a}_{ij}$ — истинное значение, а a_{ij} — номинальное значение, используемое при расчете, и εa_{ij} — вариация коэффициента a_{ij}, то:

$$\frac{\|\Delta A\|}{\|A\|} \le \varepsilon,$$

и если выполняются неравенства (36), то кроме формулы (38) может быть использована формула:

$$\frac{\|\Delta X\|}{\|X\|} \le \|A\| \cdot \|A^{-1}\| \cdot (\delta + \varepsilon). \tag{40}$$

Формулы (35) и (38) (вместе с их выводом) представлены в публикациях [4; 5; 6; 7; 8].

Отметим, что использование соотношений (36) и (39) означает, что рассматриваются относительные вариации коэффициентов b_i и a_{ij}, а не абсолютные. Если $b_i = 0$ и $a_{ij} = 0$, то и вариации $\delta b_i = 0$ и $\varepsilon a_{ij} = 0$, т. е. "нуль не варьируется" — нулевой коэффициент остается нулевым, что чаще всего (но не всегда!) соответствует физическому смыслу. Существуют, разумеется, и задачи, в которых необходимо учитывать возможность превращения нулевых коэффициентов в ненулевые, но такие задачи требуют другого подхода. Более подробно этот вопрос рассмотрен в [10], стр. 18–19 и [2], стр. 66–68.

Можно рассматривать не относительные, а "абсолютные" вариации, когда вместо формул (36) и (39) учитываются неравенства:

$$b_i - \delta_i \le \overline{b}_i \le b_i + \delta_i;$$

$$a_{ij} - \varepsilon_{ij} \le \overline{a}_{ij} \le a_{ij} + \varepsilon_{ij},$$

где δ_i и ε_{ij} — числа, не зависящие от b_i и a_{ij}. Целесообразность учета относительных или абсолютных вариаций зависит от свойств той или иной конкретной задачи.

§ 5. Некорректные и плохо обусловленные задачи в "Математике-1" и в "Математике-2"

Как известно, до последних лет все математические задачи делили на два класса: класс корректных и класс некорректных задач. К корректным задачам (их еще называют "корректно поставленными" задачами) относили задачи, удовлетворяющие трем условиям:

- решение существует;
- решение единственно;
- сколь угодно малым изменениям исходных данных соответствуют сколь угодно малые изменения решений.

Задачи, не удовлетворяющие хотя бы одному из этих трех условий, относили к классу некорректных.

В авторитетном учебнике [9] дается и более формализованное определение класса корректных задач. Сперва вводится понятие метрических пространств F и U, к которым относятся элементы исходных данных u и решений z. В [9] задача называется корректно поставленной на паре метрических пространств F и U, если выполняются все три условия:

- для всякого элемента из пространства U существует решение z из пространства F;
- решение определяется однозначно;
- задача устойчива на пространствах F и U.

По сути дела, это определение равнозначно ранее приведенному.

В работах [2; 11] эти определения корректных задач были подвергнуты критике. Так, уже невыполнение первого условия растворяет действительно некорректные задачи в необъятном море задач, не имеющих решений.

Уже простейшее уравнение:

$$x^2 + 1 = 0$$

не имеет решений в поле действительных чисел, но задачу исследования и решения этого уравнения вряд ли можно отнести к некорректным.

Для того чтобы не входить в противоречие с ранее введенными определениями, в [2] и [11] было предложено опираться на определения корректного и некорректного решения. *Корректным решением* назовем то, которое при сколь угодно малых изменениях исходных данных изменяется сколь угодно мало. *Некорректное решение* при сколь угодно малых изменениях исходных данных может измениться на конечную величину или может даже измениться коренным образом (например, при номинальных коэффициентах решения не существует, а при сколь угодно малых изменениях этих коэффициентов решение существует).

Кроме того, в работах [2]; [8]; [10]; [11] было показано, что корректность задачи может изменяться при эквивалентных преобразованиях, не изменяющих решений, использованных при ее решении. Само решение при этом не изменяется, но его корректность может измениться. Поэтому известное деление математических задач на два класса — класс корректных и класс некорректных задач — недостаточно. Необходимо ввести важный третий класс — класс задач, меняющих свою корректность в ходе эквивалентных преобразований, используемых при их решении. Неучет существования третьего класса может привести к ошибочным заключениям о корректности решения, а подобные ошибки уже не раз становились (как показано в [2], [10]) причиной аварий и катастроф. В работе [10] было показано также, что степень обусловленности решений может меняться при эквивалентных преобразованиях, не изменяющих самих решений и поэтому широко используемых при вычислении решений.

Для рассматриваемых нами систем линейных алгебраических уравнений (13) в рамках "Математики-1" некорректные решения у системы (13) будут при $\det A = 0$. При тех значениях коэффициентов, для которых $\det A = 0$, решений x_i либо не существует (если

$D_i \neq 0$), либо (если $D_i = 0$) имеется целое семейство решений, зависящее от произвольных постоянных. Все это хорошо известно. В то же время, в области "Математики-2", где учитывается, что, по крайней мере, достаточно малые изменения коэффициентов на практике почти всегда неизбежны, о точном равенстве $\det A = 0$ (и, следовательно, о некорректных задачах в области линейных алгебраических уравнений) говорить уже трудно. Здесь лучше ввести понятие об "очень плохо обусловленных системах уравнений" — т. е. таких системах, у которых внутри возможных интервалов изменения одного или нескольких коэффициентов матрицы A лежит значение $\det A = 0$. Решения таких систем при вариациях коэффициентов могут изменяться очень сильно, на сколь угодно большие величины.

Простой пример (**пример № 5**). Рассматривается система:

$$\left.\begin{aligned} 1{,}009x_1 + x_2 &= 1{,}009; \\ (1+\varepsilon)x_1 + x_2 &= 1. \end{aligned}\right\} \qquad (41)$$

При $\varepsilon = 0$ она имеет решение: $x_1 = 1$; $x_2 = 0$. Пусть известно, что коэффициент при x_1 во втором уравнении системы изменяется в пределах от 0,99 до 1,01, т. е. число ε может изменяться от $\varepsilon = -0{,}01$ до $\varepsilon = +0{,}01$. Поскольку для любого ε определитель системы:

$$D = \begin{vmatrix} 1{,}009 & 1 \\ 1+\varepsilon & 1 \end{vmatrix} = 0{,}009 - \varepsilon, \qquad (42)$$

а определители D_1 и D_2 равны:

$$D_1 = \begin{vmatrix} 1{,}009 & 1 \\ 1 & 1 \end{vmatrix} = 0{,}009; D_2 = \begin{vmatrix} 1{,}009 & 1{,}009 \\ 1+\varepsilon & 1 \end{vmatrix} = -1{,}009\varepsilon, \qquad (43)$$

то $x_1 = \dfrac{0{,}009}{0{,}009 - \varepsilon} = \dfrac{9}{9 - 1000\varepsilon}$.

Мы убеждаемся, что определитель системы (41) обращается в нуль внутри возможного интервала изменения коэффициента при x_1 во втором уравнении системы (41).

Для различных ε имеем (табл. 1).

Таблица 1. Значения решений x_1 системы (41) для различных ε

ε	–0,01	0	0,005	0,008	0,0085	0,0089	0,009	0,00901	0,00905	0,01
x_1	0,474	1	2,25	9	18	900	Не существует	–900	–18	–9

Таким образом, решение x_1 системы (41) не является некорректным. При сколь угодно малых ε решение меняется сколь угодно мало. Но при изменении ε внутри возможного интервала изменения коэффициента при x_1 во втором уравнении решение x_1 может принимать сколь угодно большие по абсолютной величине положительные и отрицательные значения.

Такие решения будем называть *очень плохо обусловленными,* а системы уравнений, имеющие такие решения, будем называть *очень плохо обусловленными системами*. Объектов, математическими моделями которых являются такие системы, надо избегать. Они опасны, поскольку их поведение в высшей степени не предсказуемо и ненадежно.

Для опознания и выделения подобных опасных объектов предлагается следующий метод: необходимо проверить, не обращается ли в нуль определитель матрицы A исследуемой системы $AX = B$ с учетом возможных вариаций ее коэффициентов — т. е. не обращается ли в нуль определитель матрицы:

$$A = \begin{pmatrix} a_{11} \pm \varepsilon_{11} a_{11} & a_{12} \pm \varepsilon_{12} a_{12} & \dots & a_{1n} \pm \varepsilon_{1n} a_{1n} \\ a_{21} \pm \varepsilon_{21} a_{21} & a_{22} \pm \varepsilon_{22} a_{22} & \dots & a_{2n1} \pm \varepsilon_{2n} a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} \pm \varepsilon_{n1} a_{n1} & a_{n2} \pm \varepsilon_{n2} a_{n2} & \dots & a_{nn} \pm \varepsilon_{nn} a_{nn} \end{pmatrix} \qquad (44)$$

при всех возможных сочетаниях вариаций $\pm\varepsilon_{ij} a_{ij}$. Метод проверки будет изложен в § 18.

В публикации [10] было показано, что степень обусловленности решений систем алгебраических уравнений может изменяться при эквивалентных преобразованиях уравнений системы, не изменяющих самих решений и поэтому широко использующихся при их вычислении.

Пример № 6. Система уравнений:

$$40x_1 - 6x_2 = 17; \tag{45}$$

$$32x_1 - 12x_2 = 13 \tag{46}$$

имеет решение: $x_1 = \dfrac{7}{16}; x_2 = \dfrac{1}{12}$.

Если сложить почленно уравнения (45) и (46) и поделить потом все члены на общий множитель 6, то после этих эквивалентных преобразований получим уравнение:

$$12x_1 - 3x_2 = 5, \tag{47}$$

а если вычесть из уравнения (46) уравнение (45), то получим:

$$8x_1 + 6x_2 = 4. \tag{48}$$

Система (47)–(48) имеет те же решения $x_1 = \dfrac{7}{16}; x_2 = \dfrac{1}{12}$, что и система (45)–(46). Это подтверждает эквивалентность этих систем. Однако обусловленность систем, чувствительность их к вариациям коэффициентов будет различной. Действительно, все уравнения (45–48) являются уравнениями прямых линий на плоскости с координатными осями $0x_1$ и $0x_2$.

Пользуясь формулами аналитической геометрии, нетрудно вычислить, что тангенс угла между прямыми (45) и (46) равен 0,212, а тангенс угла между прямыми (47) и (48) равен 1,23 (хотя точки пересечения прямых (45) и (46) те же, что у прямых (47) и (48)). В то же время, величина угла между прямыми, описываемыми уравнениями с двумя переменными, является хорошей мерой обусловленности решений этих уравнений. Чем меньше угол между прямыми, тем хуже обусловленность решения.

На рис. 2, *а* показаны прямые, описываемые уравнениями (45)–(46), а на рис. 2, *б* — прямые, описываемые уравнениями (47)–(48). Изменение степени обусловленности при эквивалентных преобразованиях здесь сразу видно.

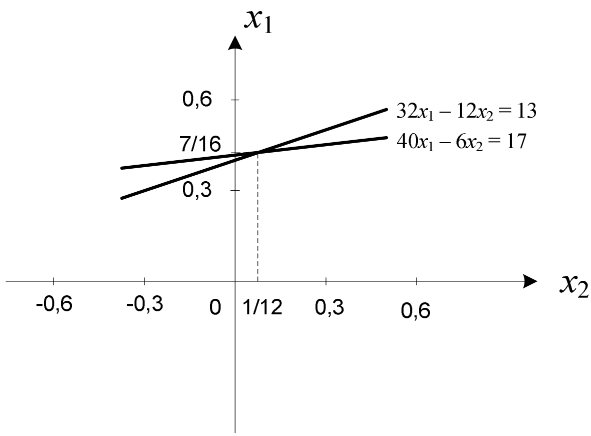

Рис. 2, *а*

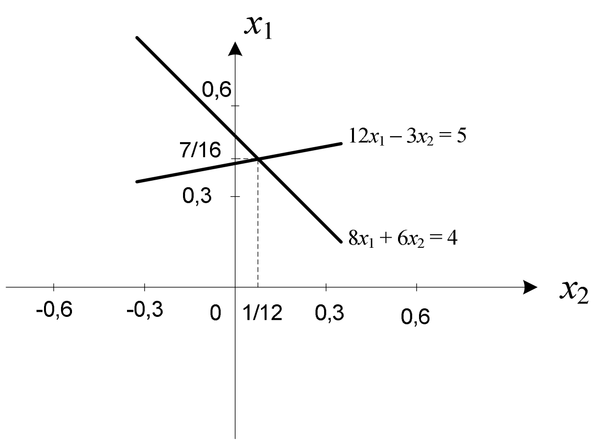

Рис. 2, *б*

Выводы

Эквивалентные преобразования, широко применяемые при вычислении решений систем уравнений, могут изменять корректность решений и степень их обусловленности. Поэтому желательно производить оценку степени обусловленности и величину погрешности решений по исходным, наименее преобразованным уравнениям.

Отметим теперь, что, например, такой распространенный метод вычисления решений системы алгебраических уравнений, как метод Гаусса — метод последовательного исключения неизвестных — использующий много эквивалентных преобразований, не позволяет оценить погрешность решений, обусловленную вариациями коэффициентов исходной системы.

Заметим также, что из утверждения: *эквивалентные преобразования могут изменить степень обусловленности* совсем не следует, что они всегда изменяют обусловленность. Могут изменять и могут не изменять. Чаще даже не изменяют. Но возможность изменения обусловленности надо всегда учитывать, и эквивалентные преобразования следует применять осторожно.

§ 6. Достоинства и недостатки оценки погрешности по "числу обусловленности"

Первый (и хорошо известный) недостаток методики оценки погрешности по "числу обусловленности" — она дает оценку лишь для осредненной нормы вектора решений x_1; x_2; ...; x_n исследуемой системы уравнений. Для наиболее часто используемой эвклидовой нормы дается оценка величины $\|\Delta x\| = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \ldots + \Delta x_n^2}$, в то время как часто нужно иметь оценку погрешности конкретного решения x_i.

Вернемся к уже рассмотренной в § 2 системе уравнений (5). Для этой системы имеем:

$$A = \begin{pmatrix} 5,8 & 3,8 \\ 2 & 4 \end{pmatrix};$$

$$\det A = 15,6;$$

$$\|A\| = \sqrt{5,8^2 + 3,8^2 + 2^2 + 4^2} = \sqrt{68};$$

$$A^{-1} = \frac{1}{15,6}\begin{pmatrix} 4 & -3,8 \\ -2 & 5,8 \end{pmatrix}$$

и "число обусловленности" равно $\|A\| \cdot \|A^{-1}\| = \dfrac{68}{15,6} = 4,36.$

Малая величина "числа обусловленности" говорит, казалось бы, о хорошей обусловленности решений x_1 и x_2. Действительно, если, например, относительная погрешность каждого из элементов матрицы A не превышает $\varepsilon = 0,03$, то погрешность нормы решений будет оценена неравенством:

$$\frac{\|\Delta x\|}{\|x\|} \le 4,36 \cdot 0,03 = 0,1308,$$

которое говорит о том, что норма погрешности решения при $\varepsilon = 0{,}03$ очень невелика.

В то же время, решая систему (5) при учете погрешностей элементов матрицы A и наиболее опасного сочетания их знаков, т. е. решая систему:

$$\left.\begin{aligned} 5{,}8(1-\varepsilon)x_1 + 3{,}8(1+\varepsilon)x_2 &= 30{,}42; \\ 2(1+\varepsilon)x_1 + 4(1-\varepsilon)x_2 &= 30{,}42; \end{aligned}\right\}$$

при $\varepsilon = 0{,}03$, получим: $x_1 = -0{,}0751$; $x_2 = 7{,}88$, в то время как при $\varepsilon = 0$ было: $x_1 = 0{,}39$; $x_2 = 7{,}41$.

В данном случае $\Delta x_1 = 0{,}465$; $\Delta x_2 = 0{,}47$ и $\dfrac{\|\Delta x\|}{\|x\|} = 0{,}089$ — т. е. норма погрешности решения, действительно, очень невелика. В то же время решение x_1 изменилось на 119% и изменило свой знак. Это означает, что решение x_1 плохо обусловлено: уже при $\varepsilon = 0{,}03$ оно может изменить знак, а это приведет к коренному изменению поведения исследуемой системы — балка соскользнет с правой опоры и упадет.

Этот пример подтверждает, что обобщенная оценка нормы всех составляющих вектора погрешности решения Δx_1; Δx_2; ...; Δx_n, которую дает оценка по "числу обусловленности", часто оказывается недостаточной и может приводить к ошибочным заключениям.

Приведем еще один простой и наглядный пример: пусть некоторая система уравнений имеет решения: $x_1 = 10$; $x_2 = 10$; $x_3 = 0{,}4$ и после вариаций решениями стали $x_1 = 9{,}5$; $x_2 = 9{,}5$; $x_3 = -0{,}1$.

В данном случае:

$$\Delta x_1 = \Delta x_2 = \Delta x_3 = 0{,}5,$$

норма изменения решений:

$$\|\Delta x\| = \sqrt{3 \cdot 0{,}25} = \sqrt{0{,}75},$$

норма решений:

$$\|x\| = \sqrt{10^2 + 10^2 + 0{,}4^2} = \sqrt{200{,}16},$$

следовательно, отношение:

$$\frac{\|\Delta x\|}{\|x\|} = \frac{\sqrt{0,75}}{\sqrt{200,16}} = 0,061.$$

Таким образом, если судить по норме, то изменение решений невелико. На самом же деле решение x_3 изменилось на 125% и даже изменило знак! Такое изменение вполне может стать причиной аварии и даже катастрофы. Поэтому оценка нормы изменения решения, которую обеспечивает "число обусловленности", далеко не достаточна.

В дальнейшем, в § 15, будет приведен пример конкретной задачи расчета усилий в элементах одной из конструкций, когда оценка по "числу обусловленности" говорит о малой норме изменения решений, в то время как одна из составляющих вектора решений уже при малых вариациях коэффициентов может изменить знак.

В последующих параграфах будут рассмотрены методы оценки (в том числе и точной оценки) погрешности каждой из составляющих x_i вектора решений X.

Вторым (и тоже хорошо известным) недостатком оценки погрешности по "числу обусловленности" является то, что она дает лишь "оценку сверху", оценку, которая может оказаться довольно грубой и допускающей существенное улучшение с помощью более тонких методов расчета, учитывающих специфику той или иной конкретной задачи.

Уже рассмотренный в § 3 пример 4 показывает, что оценка по "числу обусловленности" позволяет лишь утверждать, что для системы из примера 4 будет:

$$\frac{\|\Delta x\|}{\|x\|} \le 0,627, \tag{49}$$

в то время как на самом деле:

$$\frac{\|\Delta x\|}{\|x\|} = 0,1. \tag{50}$$

Надо, конечно, учитывать, что для получения точного равенства (50) необходимо найти решения системы и для исходных, и для изменившихся коэффициентов, в то время как оценка (49) вычисляется много проще.

Приближенность оценки по "числу обусловленности" является принципиальным и неустранимым недостатком, поскольку этот недостаток связан с самим выводом формул (35) и (38) через систему неравенств.

Однако этот недостаток не мешал и не мешает самому широкому применению оценок, использующих "числа обусловленности". Все перевешивает такое решающее достоинство, как простота и универсальность метода. При его использовании не нужно думать над особенностями конкретной задачи, приведшей к системе уравнений (13). Достаточно вычислить нормы прямой и обратной матриц, вычислить $\|A\|$ и $\|A^{-1}\|$, перемножить их, и сразу получается "оценка сверху" для нормы погрешностей решения.

Особенно удобно стало вычислять "числа обусловленности" после того, как были разработаны программы для быстродействующих вычислительных машин — программы для вычисления норм прямой и обратной матриц. Теперь достаточно набрать коэффициенты матрицы A и вектора B, вызвать соответствующие программы — и на экране компьютера высвечивается число обусловленности и "оценка сверху" для величины $\dfrac{\|\Delta x\|}{\|x\|}$.

То, что эта оценка является "оценкой сверху", не мешает практическому применению расчетов, основанных на числах обусловленности, поскольку основной их целью является проверка надежности рассчитываемого объекта. Если расчет по "числу обусловленности" говорит о том, что норма относительной погрешности расчета не превышает допустимой величины, то можно быть спокойным — реальная погрешность окажется только меньше расчетной. Если же расчет по "числу обусловленности" покажет, что норма относительной погрешности больше допустимой величины, то это еще не означает, что она и на самом деле будет больше допустимой. Это означает только, что в данном случае желательно произвести более точный (но и более трудоемкий) расчет и уже по его результатам принимать решение о надежности или ненадежности рассчитываемого объекта.

Мы подробно остановились на данном вопросе потому, что иногда снова возникает дискуссия: а нужны ли и полезны ли заведомо приближенные, не идеально точные "оценки сверху" величин погрешностей?

Опыт применения оценок по "числу обусловленности", много десятилетий успешно применяемых в практике расчетов и в учебной литературе (см., например, [4, 5, 6, 7, 8]), показывает, что "оценки сверху", пусть даже приближенные, нужны и полезны.

Отметим однако, что в последнее время были вскрыты новые, ранее не замечаемые недостатки методов оценки по "числам обусловленности". О них — в следующем параграфе.

§ 7. Новые результаты в проблеме оценок по "числам обусловленности"

Помимо хорошо известных недостатков методики оценки погрешности по "числу обусловленности", о которых говорилось в предыдущем параграфе и которые не мешали практическому применению методики, в ходе недавних исследований были обнаружены новые, ранее не замечаемые недостатки.

1. Зависимость "числа обусловленности" от эквивалентных преобразований уравнений

Пример № 7. Рассмотрим совсем простую систему:

$$2x_1 + x_2 = 1; \tag{51}$$

$$x_1 + x_2 = 1 \tag{52}$$

с решением $x_1 = 0$; $x_2 = 1$ и умножим все члены второго уравнения на число K. После умножения оно примет вид:

$$Kx_1 + Kx_2 = K \tag{53}$$

и мы будем иметь дело с системой уравнений (51)–(53). Эта система имеет то же решение $x_1 = 0$; $x_2 = 1$, что и система (51)–(52). Так и должно быть, поскольку умножение всех членов на постоянное число K, не равное нулю, является эквивалентным преобразованием, не изменяющим решений. Отметим, что в данном случае это эквивалентное преобразование не изменяет и степень обусловленности решений. Действительно, уравнение (53) при любых остается уравнением все той же прямой на плоскости с осями $0x_1$; $0x_2$. Угол между прямыми (51) и (53) для любых $K \neq 0$ остается одним и тем же, а это означает, что степень обусловленности решений от умножения на $K \neq 0$ не изменяется.

Теперь вычислим для системы (51)–(53) "число обусловленности". Для этого примера матрица A равна:

$$A = \begin{pmatrix} 2 & 1 \\ K & K \end{pmatrix}, \tag{54}$$

определитель матрицы:

$$\det A = \begin{vmatrix} 2 & 1 \\ K & K \end{vmatrix} = 2K - K = K, \tag{55}$$

обратная матрица:

$$A^{-1} = \frac{1}{K} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \tag{56}$$

евклидова норма матрицы A:

$$\|A\| = \sqrt{5 + 2K^2}, \tag{57}$$

та же норма для обратной матрицы:

$$\|A^{-1}\| = \frac{1}{K}\sqrt{5 + 2K^2}, \tag{58}$$

и число обусловленности:

$$\|A\| \cdot \|A^{-1}\| = \frac{5}{K} + 2K \tag{59}$$

оказывается зависящим от K.

Имеем следующую зависимость числа обусловленности от K (табл. 2).

Таблица 2. Зависимость числа обусловленности системы (51)–(53) от K

K	0,1	1	1,5	2	10	100
$\|A\| \cdot \|A^{-1}\|$	50,2	7	6,33	8,5	20,5	200,05

Это означает, что, имея дело с уравнением (53) для различных коэффициентов K, мы должны будем делать совершенно разные заключения об обусловленности решений системы (51)–(53). При $K = 1{,}5$ решения будут казаться хорошо обусловленными, при $K = 100$ они будут казаться плохо обусловленными. На самом же деле, обусловленность решений системы (51)–(53) от K не зависит. Точно так же не будет зависеть обусловленность решений любой системы $AX = B$ от умножения всех членов любого из уравнений системы на любое число $K \neq 0$.

Методика оценки погрешности решений по "числу обусловленности" может вести нас по неверному пути и приводить к неправильным заключениям.

Зависимость произведения $\|A\| \cdot \|A^{-1}\|$ от умножения всех членов любого из уравнений системы на постоянное число была отмечена в [4] на стр. 212. Влияние этой зависимости на правильность оценки погрешности по "числу обусловленности" в [4] не рассматривалась.

Помимо "числа обусловленности" определитель системы уравнений $AX = B$ также сильно зависит от такого эквивалентного преобразования, как умножение всех членов на число $K \neq 0$. Так, определитель системы (51)–(53) равен $\det A = \begin{vmatrix} 2 & 1 \\ K & K \end{vmatrix} = K$, и в зависимости от числа K может стать и большим, и малым. Это подчеркивает ненадежность оценки степени обусловленности решений по величине определителя системы.

Часто приходится встречать утверждение: "если определитель системы мал (по сравнению с нормой матрицы A), то система плохо обусловлена, если определитель не мал, то система хорошо обусловлена". На самом деле такое утверждение очень ненадежно, поскольку $\det A$ зависит от эквивалентных преобразований уравнений. Об этом уже говорилось в [4].

Отметим, что в [2], а ранее в публикациях [12], [13] приводились примеры, когда эквивалентные преобразования, не изменяющие самих решений, действительно изменяли корректность и обусловленность решений, и было показано, что это явление (редко встречающееся, но возможное изменение ряда свойств решений при эквивалентных преобразованиях) играет очень важную роль как среди причин аварий и катастроф, так и в их предотвращении. Теперь мы убеждаемся, что встречается и обратное явление, когда некоторые из методик расчета погрешностей решений приводят к ложным выводам о зависимости этих погрешностей от эквивалентных преобразований.

2. Ложная зависимость "числа обусловленности" от масштабов измерения коэффициентов уравнений

Еще более существенным недостатком "чисел обусловленности" является их зависимость от выбора единиц измерения размерных коэффициентов уравнений.

Вернемся к ранее рассмотренному примеру № 1 и рассмотрим более подробно тот случай, когда нагрузки на опоры x_1 и x_2 определяются из системы уравнений (2) и (12). Если длины измеряются в метрах, удельные нагрузки q_1; q_2; q_3 — в килоньютонах на метр, x_1 и x_2 — в килоньютонах, и $l_1 = l_2 = 2$ метра, $l_3 = 3{,}8$ м, $q_1 = q_2 = q_3 = 1$ кН/м, то уравнения (2) и (12) принимают вид:

$$\left.\begin{aligned} 5{,}8x_1 + 3{,}8x_2 &= 30{,}42; \\ x_1 + x_2 &= 7{,}8. \end{aligned}\right\} \qquad (60)$$

Если же длины измеряются в миллиметрах (как на технических чертежах), а нагрузки соответственно в 0,001 кН/мм, то уравнения (2) и (12) принимают вид:

$$\left.\begin{aligned} 5800x_1 + 3800x_2 &= 30420; \\ x_1 + x_2 &= 7{,}8. \end{aligned}\right\} \qquad (61)$$

Мы убеждаемся, что смена единиц измерения оказалась равнозначна умножению всех членов первого из уравнений (60) на 1000. Поэтому решения x_1 и x_2 у систем (60) и (61) будут одними и теми же, а вот "числа обусловленности" изменятся, поскольку только что было показано, что при умножении всех членов одного из уравнений системы на постоянную величину, "число обусловленности" может измениться. Это обстоятельство говорит, разумеется, о серьезных (хотя — как увидим далее — вполне устранимых) недостатках оценки по "числу обусловленности".

Однако надо отметить, что зависимость "числа обусловленности" от выбора единиц измерения будет возникать лишь при условии, что уравнения исследуемой системы имеют разную размерность. Так, все члены уравнения (2) системы (2)–(13) имеют размерность момента (сила, умноженная на длину), а все члены уравнения (13) имеют размерность силы.

3. Ошибочные суждения о влиянии параметров системы на обусловленность решений

Простота и универсальность вычисления "чисел обусловленности" позволяет применять их для оценки влияния тех или иных параметров объекта на обусловленность решений, а тем самым — и на надежность работы самого объекта. Однако и здесь возможны ошибочные суждения. Это особенно важно иметь в виду в ходе проектирования, когда надо быстро оценить много возможных проектных вариантов и выбрать из них тот, который обеспечит наименьшее изменение решений (а значит, и наименьшее изменение свойств проектируемого объекта) при неизбежных последующих изменениях параметров исследуемого объекта в ходе эксплуатации.

Пример № 8. Рассмотрим простую систему:

$$\left.\begin{aligned}(1+m)x_1 + x_2 &= 2;\\ x_1 + x_2 &= 1\end{aligned}\right\} \tag{62}$$

с решениями $x_1 = \dfrac{1}{m}; x_2 = \dfrac{m-1}{m}$. Предположим, что коэффициенты системы (62) могут испытывать вариации, и проследим зависимость степени обусловленности решений от параметра m.

Для системы (62) имеем:

$$A = \begin{pmatrix} 1+m & 1 \\ 1 & 1 \end{pmatrix};$$

$$\det A = m;$$

$$\|A\| = \sqrt{(1+m)^2 + 3};$$

$$\|A\| \cdot \|A^{-1}\| = \frac{4}{m} + m + 2. \tag{63}$$

Зависимость "числа обусловленности" от m отражена в табл. 3.

Таблица 3. Зависимость "числа обусловленности" системы (62) от m

m	0,1	1	2	10	100
$\|A\| \cdot \|A^{-1}\|$	42,1	7	6	12,4	102,04

Согласно формуле (63) получается, что при возрастании параметра m от $m = 2$ обусловленность системы ухудшается и, например, при изменении m от $m = 2$ до $m = 10$ ухудшается более чем в два раза.

На самом деле, разумеется, ничего подобного не происходит. Действительно, уравнения системы (62) — это уравнения прямых линий на плоскости с осями $0x_1$; $0x_2$. Тангенс угла между прямыми легко вычисляется и равен:

$$\operatorname{tg}\phi = \frac{m}{m+2}. \tag{64}$$

Из формулы (64) следует, что угол между прямыми монотонно возрастает с ростом m, а тем самым и степень обусловленности решений тоже возрастает с увеличением m.

Таким образом, методика, основанная на "числах обусловленности", дает неверный ответ.

Это можно подтвердить и прямым вычислением. Рассмотрим наиболее неблагоприятное сочетание знаков вариаций элементов матрицы коэффициентов системы (62), когда она переходит в систему:

$$\left.\begin{aligned} (1+m)(1-\varepsilon)x_1 + (1+\varepsilon)x_2 &= 2; \\ (1+\varepsilon)x_1 + (1-\varepsilon)x_2 &= 1. \end{aligned}\right\} \tag{65}$$

При $m = 2$ и $\varepsilon = 0$ решениями системы (65) будут $x_1 = x_2 = 0{,}5$, а при $\varepsilon = 0{,}1$ решениями станут $x_1 = 0{,}574$; $x_2 = 0{,}41$. Относительная вариация x_1 равна:

$$\frac{\Delta x_1}{x_1} = \frac{0{,}574 - 0{,}5}{0{,}5} = 0{,}148 = 1{,}48\varepsilon.$$

Для x_2 будет:

$$\frac{\Delta x_2}{x_2} = \frac{0{,}41 - 0{,}5}{0{,}5} = -0{,}18 = -1{,}8\varepsilon.$$

При $m = 4$ получим:

для $\varepsilon = 0$, решением будет $x_1 = 0{,}25$; $x_2 = 0{,}75$;
для $\varepsilon = 0{,}1$ решением станет $x_1 = 0{,}247$; $x_2 = 0{,}81$; следовательно,

$$\frac{\Delta x_1}{x_1} = \frac{0{,}247 - 0{,}25}{0{,}25} = -0{,}12 = -1{,}2\varepsilon; \quad \frac{\Delta x_2}{x_2} = \frac{0{,}81 - 0{,}75}{0{,}75} = 0{,}08 = 0{,}8\varepsilon.$$

Этот расчет подтверждает, что при $m = 4$ влияние вариаций коэффициентов на величину вариаций решений действительно меньше, чем при $m = 2$. Кроме того, расчет показывает, насколько грубое приближение обеспечивает расчет по "числам обусловленности" по отношению к реальным изменениям решений при вариациях коэффициентов системы.

Пример № 9. Рассмотрим систему из трех уравнений:

$$\left.\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1; \\ 2x_1 + x_2 + 3x_3 &= 2; \\ 3x_1 + x_2 + 2x_3 &= 3 \end{aligned}\right\} \tag{66}$$

с решениями $x_1 = 1$; $x_2 = 0$; $x_3 = 0$.

Для системы (66) имеем:

$$A=\begin{pmatrix}1&2&3\\2&1&3\\3&1&2\end{pmatrix};\det A=6;A^{-1}=\frac{1}{6}\begin{pmatrix}-1&-1&3\\2&-7&3\\-1&5&-3\end{pmatrix};$$

$$\|A\|=\sqrt{3(1+4+9)}=\sqrt{42}=6,48;\left\|A^{-1}\right\|=\frac{1}{6}\sqrt{108}=1,732;$$

$$\|A\|\cdot\left\|A^{-1}\right\|=11,225.$$

Добавим к правой части, к вектору B, вариацию δB, причем каждую составляющую вектора B увеличим на $0,1$ от исходного значения. При этом будет $\dfrac{\|\Delta B\|}{\|B\|}=0,1.$

Согласно основной формуле (35), должно выполняться неравенство:

$$\frac{\|\Delta X\|}{\|X\|}\le 11,225\cdot\frac{\|\Delta B\|}{\|B\|}=1,1225.$$

В то же время, решая непосредственно систему (66) с учетом вариаций правой части — т. е. решая систему:

$$\left.\begin{aligned}x_1+2x_2+3x_3&=1,1;\\2x_1+x_2+3x_3&=2,2;\\3x_1+x_2+2x_3&=3,3,\end{aligned}\right\}\tag{67}$$

получим решение: $x_1 = 1,1$; $x_2 = x_3 = 0$; и, следовательно,

$$\frac{\|\Delta X\|}{\|X\|}=0,1.\tag{68}$$

Таким образом, оценка по числу обусловленности оказалась в данном случае в 11,225 раза больше истинной погрешности.

§ 8. Вычисление погрешности решений при вариациях правой части

Наиболее просто вычисляется погрешность решений системы уравнений $AX = B$ в том случае, если изменяются только коэффициенты правой части, коэффициенты вектора B, а коэффициенты матрицы A остаются неизменными, или же их изменения настолько малы, что этими изменениями можно пренебречь. Такие случаи нередко встречаются на практике. В строительной механике это тот случай, когда конструкция остается неизменной, но изменяются нагрузки, при расчете электрических цепей — когда изменяются электродвижущие силы, но не меняются сопротивления и т. п.

Методику вычисления погрешностей решений удобно пояснить на примере системы третьего порядка.

Пример № 10. Рассмотрим систему:

$$\left.\begin{aligned} x_1 + 2x_2 + 3x_3 &= 7; \\ 2x_1 + x_2 + 3x_3 &= 8; \\ 3x_1 + x_2 + 2x_3 &= 9; \end{aligned}\right\} \tag{69}$$

и оценим величину погрешности решений при вариациях правой части, когда система (69) принимает вид:

$$\left.\begin{aligned} x_1 + 2x_2 + 3x_3 &= 7(1 + \delta_1); \\ 2x_1 + x_2 + 3x_3 &= 8(1 + \delta_2); \\ 3x_1 + x_2 + 2x_3 &= 9(1 + \delta_3). \end{aligned}\right\} \tag{70}$$

Величина и знаки чисел δ_1; δ_2; δ_3 могут быть любыми. Понятно, что величина погрешности решений зависит от сочетания знаков вариаций коэффициентов, от сочетания знаков чисел δ_1; δ_2; δ_3. Для системы третьего порядка (70) число возможных сочетаний знаков равно $2^3 = 8$, для системы n-го порядка число сочетаний знаков равно 2^n и очень быстро возрастает с ростом n.

Число необходимых вычислений может быть существенно уменьшено при разумном использовании формул Крамера (формулы (6) из § 2).

Для системы (69) определитель D равен:

$$D = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & 1 & 2 \end{vmatrix} = 6, \qquad (71)$$

определитель D_1 равен:

$$D_1 = \begin{vmatrix} 7 & 2 & 3 \\ 8 & 1 & 3 \\ 9 & 1 & 2 \end{vmatrix} = 7\begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} - 8\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} + 9\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} = \qquad (72)$$

$$= 7 \cdot (-1) - 8 \cdot 1 + 9 \cdot 3 = 12$$

(мы разложили определитель D_1 по первому столбцу) и поэтому:

$$x_1 = \frac{D_1}{D} = 2.$$

Для системы (70) определитель D не зависит от δ_i и остается прежним:
$D = 6$, а определитель D_1 зависит от δ_i и равен:

$$D_{1\delta} = \begin{vmatrix} 7 + 7\delta_1 & 2 & 3 \\ 8 + 8\delta_2 & 1 & 3 \\ 9 + 9\delta_3 & 1 & 2 \end{vmatrix} = (7 + 7\delta_1)\begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} - (8 + 8\delta_2)\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} + (9 + 9\delta_3)\begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix}$$

или

$$D_{1\delta} = 12 + [7\delta_1 \cdot (-1) - 8\delta_2 \cdot 1 + 9\delta_3 \cdot 3]. \qquad (73)$$

Из формулы (73) сразу видно, что изменение определителя D_1 (а значит, и изменение решения x_1) будет наибольшим, если δ_3 положительно, а δ_1 и δ_2 — отрицательны. Если δ_1; δ_2; δ_3 равны друг другу по абсолютной величине, т. е. $|\delta_1| = |\delta_2| = |\delta_3| = \delta,$ то $D_{1\delta} = 12 + 42\delta$, и, следовательно,

$$x_{1\delta} = \frac{D_{1\delta}}{D} = 2 + 7\delta. \tag{74}$$

Если вариации δ_1; δ_2; δ_3 имеют обратные знаки — т. е. если δ_1 и δ_2 — положительны, а δ_3 — отрицательна, то в этом случае при $\left|\delta_i\right| = \delta$ будет $D_{1\delta} = 12 - 42\delta$, и, следовательно,

$$x_{1\delta} = \frac{D_{1\delta}}{D} = 2 - 7\delta.$$

Аналогично вычисляются и определители $D_{2\delta}$ и $D_{3\delta}$, а по ним решения $x_{2\delta}$ и $x_{3\delta}$. Получаем:

$$D_{2\delta} = \begin{vmatrix} 1 & 7+7\delta_1 & 3 \\ 2 & 8+8\delta_2 & 3 \\ 3 & 9+9\delta_3 & 2 \end{vmatrix} = -(7+7\delta_1)\cdot\begin{vmatrix} 2 & 3 \\ 3 & 2 \end{vmatrix} + (8+8\delta_2)\cdot\begin{vmatrix} 1 & 3 \\ 3 & 2 \end{vmatrix} -$$
$$-(9+9\delta_3)\cdot\begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} = 6 + 35\delta_1 - 56\delta_2 + 27\delta_3 \tag{75}$$

и, следовательно, $x_{2\delta} = \dfrac{D_{2\delta}}{D} = x_2 + (\dfrac{35}{6}\delta_1 - \dfrac{56}{6}\delta_2 + \dfrac{27}{6}\delta_3)$.

Здесь вариация решения x_2 будет наибольшей при положительных δ_1 и δ_3 и отрицательном δ_2. Если δ_1, δ_2 и δ_3 равны друг другу по абсолютной величине, т. е. $\left|\delta_1\right| = \left|\delta_2\right| = \left|\delta_3\right| = \delta$, то $x_{2\delta} = 1 + 18\delta$. При обратных знаках вариаций δ, т. е. при отрицательных δ_1 и δ_3 и положительной δ_2 будет $x_{2\delta} = 1 - 18\delta$.

Аналогично получаем:

$$D_{3\delta} = \begin{vmatrix} 1 & 2 & 7+7\delta_1 \\ 2 & 1 & 8+8\delta_2 \\ 3 & 1 & 9+9\delta_3 \end{vmatrix} =$$
$$= (7+7\delta_1)\cdot\begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} - (8+8\delta_2)\cdot\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} + (9+9\delta_3)\cdot\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = \tag{76}$$
$$= 6 - 7\delta_1 + 40\delta_2 - 27\delta_3$$

и, следовательно,

$$x_{3\delta} = \frac{D_{3\delta}}{D} = 1 - \frac{7}{6}\delta_1 + \frac{40}{6}\delta_2 - \frac{27}{6}\delta_3. \tag{77}$$

Если $|\delta_1| = |\delta_2| = |\delta_3| = \delta$, то $x_{3\delta} = 1 + 12{,}33\delta$, а при обратных знаках δ_i будет $x_{3\delta} = 1 - 12{,}33\delta$.

Рассмотренный пример сразу проясняет порядок вычисления погрешности решения для системы уравнений $AX = B$ произвольного порядка, когда вектор-столбец правой части имеет вид:

$$B_\delta = \begin{pmatrix} b_1 + \delta_1 b_1 \\ b_2 + \delta_2 b_2 \\ \dots \\ b_n + \delta_n b_n \end{pmatrix}. \tag{78}$$

В этом случае для вычисления, например, величины x_1 используем формулы Крамера и разложим определитель D_1 по элементам первого столбца — т. е. в данном случае столбца (78). Получим:

$$x_1 = \frac{D_1}{D} = \\ = \frac{(b_1 + \delta_1 b_1)\cdot M_1 - (b_2 + \delta_2 b_2)\cdot M_2 + \dots + (-1)^n (b_n + \delta_n b_n)\cdot M_n}{D}, \tag{79}$$

где M_1; M_2; ...; M_m — миноры, соответствующие элементам первого столбца. Из (79) следует, что:

$$x_1 = x_{1\delta=0} + \Delta x_1, \tag{80}$$

где $x_{1\delta=0}$ — значение x_1, соответствующее $\delta = 0$ и вычисляемое при номинальных значениях коэффициентов b_i, а Δx_1 — это приращение x_1, произошедшее из-за вариаций этих коэффициентов. Из (79) следует, что:

$$\Delta x_1 = \frac{1}{D}\sum_{i=1}^{n} (-1)^i \delta_i b_i M_i. \tag{81}$$

Из формулы (81) сразу видно, что вариация решения x_1 будет наибольшей в том случае, если знаки вариаций $\delta_i b_i$ совпадают со знаками произведений $(-1)^i M_i$. Отметим, что вариации решений линейно зависят от вариаций коэффициентов правой части. Это позволяет — если, например, вычислена вариация решения x_i при $|\delta_i| = 0{,}01$ — легко пересчитать величину вариации решения на любое другое значение $|\delta_i|$.

Пользуясь формулой (81) можно, например, легко оценить наибольшее возможное изменение решения x_1 в системе уравнений (66) из примера № 9 для случая $|\delta| = 0{,}1$. Поскольку для системы (66) имеем:

$$M_1 = \begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} = -1; M_2 = \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} = +1; M_2 = \begin{vmatrix} 2 & 3 \\ 1 & 3 \end{vmatrix} = 3,$$

то наибольшее изменение x_1 будет при $\delta_1 = -0{,}1$; $\delta_2 = -0{,}1$; $\delta_3 = +0{,}1$.

При этом $\Delta x_1 = \dfrac{1}{6}(0{,}1 + 0{,}2 + 0{,}9) = 0{,}2.$

Вычисляя непосредственно определитель D_1 при изменившихся коэффициентах правой части и новую величину x_1:

$$D_{1\delta} = \begin{vmatrix} 1(1-0{,}1) & 2 & 3 \\ 2(1-0{,}1) & 1 & 3 \\ 3(1+0{,}1) & 1 & 2 \end{vmatrix} = 7{,}2;$$

$$x_1 = \frac{D_{1\delta}}{D} = \frac{7{,}2}{6} = 1{,}2,$$

убеждаемся, что изменение x_1 действительно равно 20% первоначального значения. Если все вариации коэффициентов правой части имеют одинаковые знаки, то изменение x_1 — как это уже было вычислено в § 7 — будет при тех же $|\delta_i|$ в два раза меньше.

Мы убеждаемся, что если изменяются только коэффициенты правой части системы уравнений $AX = B$, то изменения решения x_i и максимально возможные изменения вычислить не сложно, и — что самое важное — они вычисляются *точно* (при известных δ_i).

Значительно сложнее вычислить изменения решений при вариациях коэффициентов в матрице A. Здесь могут помочь "модульные определители", предложенные ранее в [10], но они позволяют дать для погрешности решения только ее "оценку сверху". Предварительно покажем, как можно с их помощью выделить очень плохо обусловленные системы.

§ 9. Выделение "очень плохо обусловленных систем" с использованием "модульных определителей"

Рассмотрим поведение решений системы уравнений $AX = B$ при вариациях коэффициентов матрицы A, когда о каждом из коэффициентов известно только то, что он находится внутри некоторого интервала:

$$a_{ij}(1-\varepsilon_{ij}) \le \overline{a}_{ij} \le a_{ij}(1+\varepsilon_{ij}). \qquad (82)$$

Сделаем допущения:

- максимально возможные величины (абсолютные) всех чисел ε_{ij} одинаковы и равны ε;
- знаки чисел ε_{ij} не известны и возможно любое сочетание этих знаков, в том числе и наиболее неблагоприятное.

Будем основываться на формулах Крамера и поставим вопрос: при каких значениях ε определитель системы $AX = B$, равный:

$$D = \begin{vmatrix} a_{11} \pm \varepsilon a_{11} & a_{12} \pm \varepsilon a_{12} & \dots & a_{1n} \pm \varepsilon a_{1n} \\ a_{21} \pm \varepsilon a_{21} & a_{22} \pm \varepsilon a_{22} & \dots & a_{2n} \pm \varepsilon a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} \pm \varepsilon a_{n1} & a_{n2} \pm \varepsilon a_{n2} & \dots & a_{nn} \pm \varepsilon a_{nn} \end{vmatrix}, \qquad (83)$$

может обратиться в нуль?

Ранее (в § 5) мы установили, что если определитель (83) может обратиться в нуль, то решения (при $D_i \ne 0$) для различных значений ε могут быть любыми, а при одновременно: $D_i = 0$ и $D = 0$ система делается неопределенной.

Если при $D \ne 0$ определитель D_i при некоторых значениях ε обращается в нуль и меняет знак, то и решение x_i меняет знак, а это в ряде случаев меняет все поведение исследуемого объекта.

Системы, для которых при $\varepsilon \neq 0$ возможно равенство $D = 0$, мы предложим называть *очень плохо обусловленными*. Рассмотрим метод выделения этих опасных систем и начнем с определителя второго порядка, в котором все члены известны с точностью лишь до $\pm\varepsilon$:

$$\begin{vmatrix} a_{11} \pm \varepsilon a_{11} & a_{12} \pm \varepsilon a_{12} \\ a_{21} \pm \varepsilon a_{21} & a_{22} \pm \varepsilon a_{22} \end{vmatrix}. \tag{84}$$

Номинальное значение определителя (84) при $\varepsilon = 0$ равно:

$$a_{11}a_{22} - a_{12}a_{21}. \tag{85}$$

Обозначим его через $\det_{НОМ}$.

Пренебрегая квадратами малых чисел ε по сравнению с их первыми степенями, получим с точностью до линейных по ε членов:

$$\begin{gathered} \begin{vmatrix} a_{11} \pm \varepsilon a_{11} & a_{12} \pm \varepsilon a_{12} \\ a_{21} \pm \varepsilon a_{21} & a_{22} \pm \varepsilon a_{22} \end{vmatrix} = \\ = a_{11}a_{22} - a_{12}a_{21} \pm \varepsilon a_{11}a_{22} \pm \varepsilon a_{11}a_{22} \mp \varepsilon a_{12}a_{21} \mp \varepsilon a_{12}a_{21} = \\ = a_{11}a_{22} - a_{12}a_{21} \pm 2\varepsilon(a_{11}a_{22} \mp a_{12}a_{21}). \end{gathered} \tag{86}$$

Теперь надо учесть, что не известные нам знаки погрешностей $\pm\varepsilon$ могут комбинироваться так, что все члены, содержащие ε, будут либо все положительны, либо все отрицательны. С учетом этого имеем:

$$\det_{НОМ} - 2\varepsilon\det_{МОД} \leq \overline{\det} \leq \det_{НОМ} + 2\varepsilon\det_{МОД}, \tag{87}$$

где $\overline{\det}$ — это истинное, не известное нам значение определителя (84), а:

$$\det_{МОД} = |a_{11}| \cdot |a_{22}| + |a_{12}| \cdot |a_{21}|. \tag{88}$$

Величину $\det_{МОД}$ мы будем называть *модульным определителем*. Из соотношения (88) видно, что он вычисляется по той же формуле (85), что и обычный определитель, но каждый коэффициент, стоящий в формуле (85), заменяется на его модуль, и между произведениями всегда стоит знак "плюс", а не "минус". Всегда выполняется неравенство:

$$\det_{МОД} \geq \det_{НОМ}\,. \tag{89}$$

Формула (87) может быть обобщена на определители более высоких порядков. Так, определитель третьего порядка состоит, как известно, из шести произведений его элементов, взятых по три:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$\begin{gathered} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - \\ - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{32}. \end{gathered} \tag{90}$$

Оценим теперь пределы, внутри которых будет заключено истинное значение $\overline{\det}$ определителя (90) с учетом вариаций его членов — т. е. оценим величину определителя:

$$\begin{vmatrix} a_{11} \pm \varepsilon a_{11} & a_{12} \pm \varepsilon a_{12} & a_{13} \pm \varepsilon a_{13} \\ a_{21} \pm \varepsilon a_{21} & a_{22} \pm \varepsilon a_{22} & a_{23} \pm \varepsilon a_{23} \\ a_{31} \pm \varepsilon a_{31} & a_{32} \pm \varepsilon a_{32} & a_{33} \pm \varepsilon a_{33} \end{vmatrix}. \tag{91}$$

Пренебрегая снова всеми степенями малого числа ε выше первой степени и допуская, что возможны наиболее неблагоприятные сочетания знаков +ε и –ε в определителе (91), получим оценку:

$$\det_{НОМ} - 3\varepsilon\det_{МОД} \leq \overline{\det} \leq \det_{НОМ} + 3\varepsilon\det_{МОД}, \tag{92}$$

где $\det_{НОМ}$ — номинальное значение определителя, вычисляемое по формуле (90), а:

$$\begin{gathered} \det_{МОД} = |a_{11}a_{22}a_{33}| + |a_{12}a_{23}a_{31}| + \\ + |a_{13}a_{21}a_{32}| + |a_{13}a_{22}a_{31}| + |a_{11}a_{23}a_{32}| + |a_{12}a_{21}a_{33}|. \end{gathered} \tag{93}$$

Сравнивая формулы (90) и (93) убеждаемся, что "модульный определитель" $\det_{МОД}$ вычисляется аналогично обычному определителю, но только вместо каждого из произведений берется его абсолютное значение (модуль) и все модули только складываются (знаков "минус" в формуле (93) нет).

Формула (92) может быть обобщена на определители n-го порядка, и для них она имеет вид:

$$\det_{НОМ} - n\varepsilon \det_{МОД} \le \overline{\det} \le \det_{НОМ} + n\varepsilon \det_{МОД}. \qquad (94)$$

Подчеркнем: при выводе формул (92) и (94) допускалась возможность наиболее неблагоприятного сочетания знаков вариаций элементов определителей. В реальных определителях сочетание знаков вариаций, использованное при выводе формул (92) и (94), часто не может реализоваться, и поэтому эти формулы могут использоваться лишь для "оценок сверху" вариаций определителей, но не для точных оценок, которые будут даны далее.

Приведем примеры применения модульных определителей.

Пример № 11. Вернемся к примеру № 2, рассмотренному в § 2, и установим, при каких вариациях $\pm\varepsilon$ может обратиться в нуль определитель D системы уравнений (5).

Для определителя D имеем:

$$\det_{НОМ} = \begin{vmatrix} 5,8 & 3,8 \\ 2 & 4 \end{vmatrix} = 23,2 - 7,8 = 15,6;$$

$$\det_{МОД} = 23,2 + 7,8 = 31.$$

Согласно формуле (87), определитель D может обратиться в нуль при:

$$\varepsilon = \varepsilon_{КР} = \frac{\det_{НОМ}}{2\det_{МОД}} = 0,252.$$

Тот же результат получим и при прямом вычислении.

В данном случае для обращения определителя в нуль необходимы весьма большие вариации $\pm\varepsilon$ даже при наиболее неблагоприятном сочетании их знаков. Такие большие вариации на практике, разумеется, встречаться не будут, поэтому система уравнений (5), рассмотренная ранее в примере № 2, не является "очень плохо обусловленной".

Заметим, что поскольку формулы (87); (92); (94) дают для возможной величины изменения определителя "оценку сверху", то выполнение неравенства:

$$\det_{НОМ} > n\varepsilon \det_{МОД}$$

означает, что исследуемый определитель заведомо не обратится в нуль, в то время как выполнение противоположного неравенства:

$$\det_{НОМ} < n\varepsilon \det_{МОД}$$

говорит о том, что определитель может обратиться в нуль, но может и не обратиться, и поэтому желательно использовать более точные (но и более сложные) методы оценки для возможной величины вариации определителя. Об этих более точных методах оценки будет рассказано в § 18.

§ 10. Оценка погрешности решений через "модульные определители"

Модульные определители могут быть использованы для "оценки сверху" возможных вариаций решений системы алгебраических уравнений как при вариациях коэффициентов вектора B, стоящего в правой части системы (13), так и при вариациях коэффициентов матрицы A. Для этого необходимо, пользуясь формулами Крамера:

$$x_i = \frac{D_i}{D} \tag{95}$$

оценить вариации определителей D_i и D при условии, что каждый из элементов определителей D_i и D может испытывать относительную вариацию $\pm\varepsilon$. Допускается также, что возможно наиболее неблагоприятное сочетание знаков этих вариаций.

Из формулы (94) следует, что абсолютное значение (модуль) относительной вариации величины определителя n-го порядка будет подчинено неравенству:

$$\left|\varepsilon_{\det}\right| \le n\frac{\det_{МОД}}{\det_{НОМ}}, \tag{96}$$

а наибольшая возможная относительная вариация решения x_i будет равна (согласно формуле (95)) сумме относительных вариаций числителя и знаменателя, сумме вариаций определителей D_i и D.

Пример № 12. Рассмотрим простую систему:

$$\left.\begin{aligned} 2x_1 + x_2 + x_3 &= 1; \\ x_1 + x_2 + x_3 &= 2; \\ x_1 + 2x_2 + x_3 &= 1 \end{aligned}\right\} \tag{97}$$

с решениями $x_1 = -1$; $x_2 = 1$; $x_3 = 2$. Предположим, что все коэффициенты системы (97) известны лишь с точностью до одной сотой (т. е. $\varepsilon = 0{,}01$), и оценим наибольшие возможные погрешности решений.

Для системы (97) определитель D равен:

$$D_{НОМ} = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{vmatrix} = 2+1+2-1-4-1 = 5-6 = -1, \tag{98}$$

а модульный определитель:

$$\det_{МОД} = 5+6 = 11. \tag{99}$$

Таким образом, наибольшая возможная погрешность определителя D с учетом принятых допущений равна:

$$3 \cdot 0{,}01 \cdot \frac{11}{-1} = -0{,}33. \tag{100}$$

Определитель D_1 равен:

$$D_{1НОМ} = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 3 & 2 & 1 \end{vmatrix} = 1+3+4-3-2-2 = 8-7 = 1, \tag{101}$$

а его модульный определитель

$$\det_{1МОД} = 8+7 = 15. \tag{102}$$

Наибольшая возможная погрешность определителя D_1 равна:

$$3 \cdot 0{,}01 \cdot \frac{15}{1} = 0{,}45. \tag{103}$$

Наибольшая возможная погрешность решения x_1, учитывая, что:

$$x_1 = \frac{D_1}{D} = \frac{-1}{1} = -1, \tag{104}$$

в общем случае будет равна (как погрешность частного) сумме абсолютных величин погрешностей D_1 и D, т. е. будет равна:

$$0,33 + 0,45 = 0,78. \tag{105}$$

Аналогично вычисляются и "оценки сверху" погрешностей решений x_2 и x_3.

Отметим важное обстоятельство: поскольку оценка через модульные определители позволяет оценить не только абсолютную величину погрешностей D и D_i, но и их знак, то это может существенно улучшить оценку погрешности любого решения x_i. Так, если бы погрешность D_1 равнялась бы не +0,45, а –0,45, то погрешность $x_1 = \dfrac{D_1}{D}$ можно было бы оценить числом (0,45 – 0,33) = 0,12 (пренебрегая всеми степенями, кроме первой малого в сравнении с единицей числа ε).

Следует обратить внимание на то, что расчет погрешности решения через модульные определители предполагает возможность наиболее неблагоприятного сочетания знаков погрешностей коэффициентов, предполагает возможность суммирования погрешностей D_i и D. Поэтому расчет через модульные определители дает для погрешности решений — как уже говорилось — только "оценку сверху".

Эта оценка за счет использования особенностей конкретной задачи обычно может быть улучшена. Вернемся к примеру № 1, рассмотренному в § 1, и обратим внимание на формулы (2) и (12). Рассмотрим часто встречающийся случай, когда длина балки измерена достаточно точно, но она положена на опоры с некоторым смещением. Если балка смещена вправо, то длина l_3 увеличилась на некоторую величину ε, длина l_1 уменьшилась на ту же величину ε, а длина l_2 и сумма $l_1 + l_2 + l_3$ остались без изменения. Если балка будет смещена влево, то знак ε изменится. Таким образом, в системе уравнений (2)–(12) при неточной установке балки на опоры могут измениться только два коэффициента уравнения (2), а все коэффициенты уравнения (12) останутся без изменения. Это позволяет дать гораздо более точную оценку погрешности решения, произошедшую от смещения балки на неизвестную величину ε, — более точную, чем это обеспечит общая методика "модульных определителей".

Поэтому, разумеется, везде, где особенности той или иной конкретной задачи позволяют выделить вариации или комбинации знаков вариаций, которые в данной задаче возникнуть не могут, такие вариации или комбинации знаков вариаций надо выделять и исключать из рассмотрения. Это позволяет повысить точность оценок, но требует глубокого изучения специфики каждой конкретной задачи.

При оценке погрешностей составляющих вектора решений X (т. е. чисел x_1; x_2; ...; x_n) через "модульные определители" нужно помнить, что они дают только "оценку сверху". Точные (но и более сложно вычисляемые) оценки будут приведены в §§ 13–17.

§ 11. Недостатки и достоинства методики оценки погрешностей решений через "модульные определители"

Основной (и совершенно очевидный) недостаток: при вариациях коэффициентов матрицы A системы $AX = B$ модульные определители дают для исследуемой погрешности "оценку сверху", и не всегда эта оценка оказывается достаточно близкой к реальной величине, которая часто оказывается меньше вычисленной оценки.

Это обстоятельство не мешает применению модульных определителей. Действительно, для практики важнее всего надежная оценка. Если расчет по методике модульных определителей показал, что рассчитанная погрешность решений меньше допустимого уровня, то можно быть спокойным: реальная погрешность может быть только меньше расчетной.

Если же вычисленная погрешность оказалась больше допустимой, то это еще не означает, что исследуемый объект непригоден и плох. Он может быть как плохим, так и хорошим. Просто надо продолжить исследование и прибегнуть к более точным, но и более сложным методам оценки погрешностей решений (о них будет рассказано в § 18), или же к методам, учитывающим специфические особенности исследуемого объекта. Кроме того, всегда остается "в резерве" описанный еще в § 1 метод повторения расчета для различных сочетаний знаков вариаций параметров исследуемого объекта — этот метод требует очень большого объема вычислений, но он надежен.

Отметим, что начинать исследование полезно с использования простых и универсальных методов оценки погрешностей решений, что намного сокращает общую работу. Так же, как ранее широко использовался простой и универсальный метод оценки нормы погрешности по "числу обусловленности", который тоже давал только "оценку сверху", так и теперь не меньшее распространение может получить метод "модульных определителей", свободный от многих недостатков, присущих "числам обусловленности".

Вернемся еще раз коротко к ранее рассмотренным недостаткам оценок по "числам обусловленности" и покажем, что методика, основанная на "модульных определителях", от этих недостатков свободна.

1) В § 7 было показано, что "числа обусловленности" оказываются зависящими от постоянного числа K, на которое в ходе эквивалентных преобразований умножается одно из уравнений системы. При таком эквивалентном преобразовании (умножении на число, не равное нулю) изменяется не реальная, а кажущаяся степень обусловленности, что и приводит к ошибкам.

Вернемся к рассмотренной в § 7 системе уравнений (51)–(52) и покажем, что суждение о погрешности решений при вариациях коэффициентов для этой системы (как и для других систем) при использовании "модульных определителей" не зависит от числа K.

Для системы (51)–(52):

$$D_{НОМ} = \begin{vmatrix} 2 & 1 \\ K & K \end{vmatrix} = 2K - K = K; D_{МОД} = 2K + K = 3K \qquad (106)$$

и отношение $\dfrac{D_{МОД}}{D_{НОМ}} = \dfrac{3K}{K} = 3,$ определяющее погрешность решения, не зависит от K.

То же самое справедливо относительно определителей D_1 и D_2.

Этот простой пример сразу показывает, почему "число обусловленности" зависит от K (что и приводит к ошибочным заключениям об обусловленности решений), а методика, основанная на модульных определителях, дает правильные результаты.

Причина заключается в том, что "число обусловленности" зависит лишь от коэффициентов матрицы A, в то время как определитель номинальный и определитель модульный для любой системы уравнений $AX = B$ в равной мере пропорциональны числу K, на которое умножается правая и левая часть любого из уравнений системы. Поэтому отношение $\dfrac{D_{МОД}}{D_{НОМ}}$ от этого числа не зависит.

Таким образом, методика оценки погрешности на основе модульных определителей свободна от серьезного недостатка, присущего оценке по "числу обусловленности". Свободна она и от рассмотренной в § 7 зависимости "числа обусловленности" от единиц измерения коэффициентов любого из уравнений, поскольку перемена единиц измерения равнозначна умножению на постоянное число K.

2) В § 7 на примере системы уравнений (62) с параметром m было показано, что расчет по "числу обусловленности" может дать неверные ответы относительно зависимости степени обусловленности решений от входящего в уравнения параметра.

Исследуем эту зависимость с помощью модульных определителей. Для системы (62) имеем:

$$\begin{gathered} D_{НОМ} = \begin{vmatrix} 1+m & 1 \\ 1 & 1 \end{vmatrix} = m; D_{МОД} = m+3; \\ D_{1НОМ} = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1; D_{1МОД} = 3; \end{gathered} \tag{107}$$

$$x_1 = \frac{D_{1МОД}}{D_{1НОМ}} = \frac{1}{m}.$$

Наибольшая возможная погрешность решения x_1 при условии, что все коэффициенты известны с относительной точностью до $\pm\varepsilon$, равна сумме погрешностей определителей D_1 и D. Но эти погрешности равны:

$$2\varepsilon\frac{D_{МОД}}{D_{НОМ}} = 2\varepsilon\frac{m+3}{m}; 2\varepsilon\frac{D_{1МОД}}{D_{1НОМ}} = 6\varepsilon. \tag{108}$$

Суммарная погрешность равна:

$$\varepsilon\left(8+\frac{6}{m}\right) \tag{109}$$

и монотонно возрастает с уменьшением коэффициента m. Это полностью соответствует как физическому смыслу, так и более детальному расчету, приведенному в § 7. Расчет погрешности x_2 аналогичен. Она так же уменьшается с ростом m.

Таким образом, методика оценки максимальной погрешности решений на основе модульных определителей свободна и от второго недостатка методики, основанной на "числах обусловленности".

Это связано с тем, что в модульных определителях учитываются вариации как коэффициентов матрицы A системы уравнений, так и вариации коэффициентов правой части системы, в то время как число обусловленности $\|A\|\cdot\|A^{-1}\|$ зависит только от коэффициентов матрицы A.

Основным достоинством методики, основанной на модульных определителях, является ее универсальность и простота вычисления "оценки сверху" для погрешности, поскольку здесь не требуется учета специфики той или иной задачи. Если выписана система уравнений и оценена возможная относительная погрешность коэффициентов, равная $\pm\varepsilon$, то все дальнейшее идет совершенно автоматически: в компьютер вводятся коэффициенты матрицы A и вектора B, вызываются программы вычисления определителей D; D_1; D_2; ... D_n и соответствующих модульных определителей, после чего на экране компьютера высвечиваются "оценки сверху" возможных погрешностей определителей и решений системы.

Расчет легко повторяется для различных сочетаний величин параметров исследуемой системы и позволяет быстро выбрать в ходе проектирования наилучший вариант, с наименьшими погрешностями и вариациями решений.

Отметим, что методика использования "модульных определителей" еще не до конца разработана, и в ней остается место для интересной и плодотворной научной работы. Так, например, давно созданы программы для быстрого и удобного вычисления определителей достаточно высоких порядков (для определителя порядка n они требуют, как показано в [6], примерно $\frac{1}{3}n^3 + n^2$ умножений и делений). Для модульных определителей подобных алгоритмов и программ еще не существует, но, разумеется, они будут созданы достаточно скоро.

§ 12. Возможности улучшения оценок нормы погрешности по "числу обусловленности"

Отмеченная в § 6 зависимость "числа обусловленности" от величины числа, на которое умножается правая и левая части любого из уравнений рассматриваемой системы $AX = B$ в ходе эквивалентных преобразований, может быть использована для увеличения точности оценки погрешности решения при вариациях коэффициентов матрицы A.

Действительно, для любого числа $K \neq 0$, на которое умножается правая и левая часть любого из уравнений системы, оценка по "числу обусловленности" остается "оценкой сверху". Здесь мы можем положиться на доказательства этого, приведенные в [4], [6], [7] и других публикациях. Но раз это так, то можно выбрать то число K, которое приводит к наименьшему "числу обусловленности" и тем самым обеспечивает наилучшую оценку, наиболее близкую к действительной максимальной величине вариации решения при вариациях коэффициентов матрицы A.

Если — как это было сделано при рассмотрении примера № 7 из § 7 — получена аналитическая зависимость "числа обусловленности" от K, то можно продифференцировать эту зависимость и найти значение K, обеспечивающее точный минимум. В примере № 7 точный минимум соответствует $K = 1{,}5$ и числу обусловленности $\|A\| \cdot \|A^{-1}\| = 6{,}33$. Используя его, получаем оценку:

$$\frac{\|\Delta X\|}{\|X\|} \leq 6{,}33\varepsilon, \tag{110}$$

где $\pm\varepsilon$ — наибольшая относительная вариация коэффициентов матрицы A.

В данном примере нетрудно оценить точность оценки. Решая вместо системы (51)–(52) систему:

$$\left.\begin{aligned} 2(1-\varepsilon)x_1+(1+\varepsilon)x_2&=1\\ (1+\varepsilon)x_1+(1-\varepsilon)x_2&=1\end{aligned}\right\};$$

для $\varepsilon = 0{,}1$, получаем: $x_1 = -0{,}328$; $x_2 = 1{,}148$; и, следовательно:

$$\Delta x_1 = -0{,}328;\ \Delta x_2 = 0{,}148;\ \|\Delta X\| = \sqrt{0{,}328^2 + 0{,}148^2} = 0{,}36,$$

откуда следует:

$$\frac{\|\Delta X\|}{\|X\|} = 0{,}36, \tag{111}$$

что не слишком сильно отличается от оценки по "числу обусловленности" при $K = 1{,}5$, когда значению $\varepsilon = 0{,}1$ соответствует оценка $\dfrac{\|\Delta X\|}{\|X\|} \le 0{,}633$. При других числах K оценка оказывается значительно грубее.

Разумеется, аналитическая зависимость числа обусловленности от числа K может быть получена лишь в простейших случаях. При практических вычислениях можно поступать проще: поскольку для вычисления "числа обусловленности" имеются удобные программы, то можно просто повторить расчет для ряда значений числа K и выбрать наилучшее. Поступая таким образом, можно получить существенно улучшенную оценку по "числу обусловленности" для величины $\dfrac{\|\Delta X\|}{\|X\|}$.

Еще одним недостатком оценок по "числу обусловленности", также отмеченным в § 6, является зависимость от единиц измерения коэффициентов исследуемой системы уравнений.

Любопытно отметить, что этот недостаток — при условии, разумеется, что причины зависимости оценки от единиц измерения исследованы и правильно поняты — может быть превращен в достоинство: действительно, изменение единиц измерения коэффициентов в любом из уравнений равнозначно умножению всех его членов на постоянное число, а такое умножение — как мы только что показали — может быть использовано для улучшения точности оценки.

Вернемся к системам (60) и (61) из § 6, отличающимся друг от друга только единицами измерения длины: в метрах для системы (60) и в миллиметрах для системы (61). Для системы (60) имеем:

$$A = \begin{pmatrix} 5,8 & 3,8 \\ 1 & 1 \end{pmatrix}; \det A = 2; \|A\| = \sqrt{5,8^2 + 3,8^2 + 1 + 1} = \sqrt{50,8};$$

$$A^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -3,8 \\ -1 & 5,8 \end{pmatrix}; \|A\| \cdot \left\|A^{-1}\right\| = 25,4,$$

в то время как для системы (61) будет:

$$A = \begin{pmatrix} 5800 & 3800 \\ 1 & 1 \end{pmatrix}; \det A = 2000; \|A\| \cdot \left\|A^{-1}\right\| = 24400,1.$$

Казалось бы, возникает сложный вопрос — какое из двух чисел обусловленности: 25,4 или 24400,1 — отражает истинную обусловленность исследуемого объекта? На этот вопрос легко ответить, если все члены первого из уравнений (60) умножить на число K. При этом система (60) примет вид:

$$\left.\begin{aligned} 5,8Kx_1 + 3,8Kx_2 &= 30,42K; \\ x_1 + x_2 &= 7,8. \end{aligned}\right\} \tag{112}$$

Теперь легко видеть, что переход от измерения длины в метрах к измерению ее в миллиметрах (т. е. переход от системы уравнений (60) к системе (61)) эквивалентен умножению на $K = 1000$.

Для этой системы будет:

$$A = \begin{pmatrix} 5,8K & 3,8K \\ 1 & 1 \end{pmatrix}; \det A = 2K; \|A\| \cdot \left\|A^{-1}\right\| = 24,4K + \frac{1}{K}.$$

Минимум "числа обусловленности" $24,4K + \dfrac{1}{K}$ будет достигаться при $K = \dfrac{1}{4,94}$, и этот минимум равен 5,17.

Таким образом, наилучшей оценкой по "числу обусловленности" для нормы погрешности решений систем (60) и (67) — при условии, что относительные погрешности коэффициентов не превышают ±ε — будет оценка:

$$\frac{\|\Delta X\|}{\|X\|} \le 5{,}17\varepsilon. \qquad (113)$$

Любопытно заметить, что для получения этой наилучшей оценки за единицу измерения длины следовало бы брать не метр, а 4,94 метра. Разумеется, на практике проще не вводить новую единицу измерения длины, а использовать любую, наиболее удобную единицу измерения, а потом умножить все члены уравнения на число K и искать величину этого числа, приводящую к наиболее точной оценке нормы погрешности через "число обусловленности".

§ 13. Точная оценка изменения решений при вариациях коэффициентов системы уравнений

В предыдущих параграфах использование свойств определителей позволило дать "оценку сверху" для погрешностей каждого из решений x_1; x_2; ... x_n системы $AX = B$, а также для изменений решений, происходящих из-за вариаций коэффициентов матрицы A и вектора B.

Та же методика при ее дальнейшем развитии позволяет дать точную оценку изменения решений. В основе получаемой точной оценки лежит, как и ранее, использование формул Крамера (6) и оценка изменений участвующих в них определителей D и D_i.

Рассмотрим определитель n-го порядка:

$$\begin{vmatrix} a_{11}+\varepsilon_{11}a_{11} & a_{12}+\varepsilon_{12}a_{12} & \dots & a_{1n}+\varepsilon_{1n}a_{1n} \\ a_{21}+\varepsilon_{21}a_{21} & a_{22}+\varepsilon_{22}a_{22} & \dots & a_{2n}+\varepsilon_{2n}a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1}+\varepsilon_{n1}a_{n1} & a_{n2}+\varepsilon_{n2}a_{n2} & \dots & a_{nn}+\varepsilon_{nn}a_{nn} \end{vmatrix} \qquad (114)$$

и будем исследовать его изменение, происходящее от вариаций элементов — т. е. исследуем разность между определителем (114) и определителем:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}, \qquad (115)$$

т. е. определителем (114), в котором все $\varepsilon_{ij} = 0$.

Сложность исследования этой разности заключается в том, что она зависит от сочетания знаков (плюс или минус) чисел ε_{ij}, а число этих сочетаний равно 2^{n^2} и очень быстро возрастает с ростом n. Так, при $n = 3$ число сочетаний знаков ε_{ij} равно $2^9 = 512$, при $n = 5$ оно равно $2^{25} > 10^7$, а при $n = 10$ возрастает до $2^{100} > 10^{33}$.

Однако, опираясь на результаты, приведенные в [10], можно установить наиболее неблагоприятное сочетание знаков чисел ε_{ij}, приводящее к наибольшей величине разности определителей (114) и (115), после чего вычислить эту разность уже не трудно.

Пример № 13. Рассмотрим простой определитель:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 3 & 4 & 5 \end{vmatrix} = 8, \tag{116}$$

для которого наиболее неблагоприятное сочетание знаков ε_{ij} соответствует следующей "таблице знаков":

$$\begin{vmatrix} - & - & + \\ + & - & + \\ + & + & - \end{vmatrix}. \tag{117}$$

Способ построения "таблицы знаков" для определителя любого порядка будет приведен далее, в § 16.

Если для всех i и j будет $|\varepsilon_{ij}| = 0{,}01$, то при наиболее неблагоприятном сочетании знаков ε_{ij} определитель (116) перейдет в:

$$\begin{vmatrix} 0{,}99 & 1{,}98 & 3{,}03 \\ 4{,}04 & 0{,}99 & 2{,}02 \\ 3{,}03 & 4{,}04 & 4{,}95 \end{vmatrix} = 9{,}6604. \tag{118}$$

Таким образом вариации элементов определителя (116), величиной каждая ±1%, при наиболее неблагоприятном сочетании их знаков могут привести к изменению величины определителя на $\Delta_+ = 1{,}6604$, или на +20,76%. Можно вычислить и наибольшее возможное изменение величины определителя в отрицательную сторону. Для этого нужно в "таблице знаков" (117) поменять знаки плюс на минус и минус на плюс. Тогда — если для всех i и j будет $|\varepsilon_{ij}| = 0{,}01$ — определитель (116) перейдет в:

$$\begin{vmatrix} 1{,}01 & 2{,}02 & 2{,}97 \\ 3{,}96 & 1{,}01 & 1{,}98 \\ 2{,}97 & 3{,}96 & 5{,}05 \end{vmatrix} = 6{,}3804\,,$$

в этом случае определитель уменьшился на $\Delta_- = -1{,}6196$, или на –20,24%. Небольшое расхождение между $|\Delta_+|$ и $|\Delta_-|$ объясняется нелинейным характером зависимости величины определителя от ε_{ij}.

Таким образом, возможные изменения определителя (116) — обозначим их $\Delta_{\text{опр}}$ — при вариациях его элементов на ±0,01 подчинены неравенству:

$$1{,}6196 \le \Delta_{\text{опр}} \le 1{,}6604, \tag{119}$$

или — в относительных единицах:

$$0{,}2024 \le \frac{\Delta_{\text{опр}}}{\det} \le 0{,}2076,$$

причем оценка (119) является точной — т. е. существуют такие комбинации вариаций элементов, при которых неравенства (119) превращаются в точные равенства.

Таким образом, в данном примере наибольшая вариация определителя (116) оказывается в 20,76 раза больше, чем наибольшая вариация каждого из его элементов.

Наличие точной оценки позволяет оценить степень точности полученной ранее приближенной оценки величины $\Delta_{\text{опр}}$ с помощью "модульных определителей" (см. § 10). Для определителя (116) модульный определитель $\det_{МОД} = 122$ и поэтому при $|\varepsilon_{ij}| = 0{,}01$ можно утверждать, что наибольшее изменение величины определителя не превысит $0{,}01 \cdot 3 \cdot 122 = 3{,}66$. Неравенства (119) дают, естественно, более точные оценки, хотя для получения их требуется существенно больший объем вычислений.

Изложенная методика может быть использована и в тех случаях, когда некоторые из элементов определителя известны точно (их вариации $\varepsilon_{ij}a_{ij}$ равны нулю), или когда известны знаки вариаций некоторых элементов. В этом случае используется только часть "таблицы знаков". Пусть, например, относительно определителя (116) известно, что только вариации элементов его первой строки отличны от нуля, и для первой строки $|\varepsilon_{ij}| = 0{,}01$, а для остальных строк $\varepsilon_{ij} = 0$. В этом случае используем только первую строку "таблицы знаков" (116). При наиболее неблагоприятном сочетании знаков вариаций элементов первой строки определитель принимает вид:

$$\begin{vmatrix} 0{,}99 & 1{,}98 & 3{,}03 \\ 4 & 1 & 2 \\ 3 & 4 & 5 \end{vmatrix} = 8{,}7,$$

т. е. в данном случае вариации трех элементов определителя изменяют его величину на 0,7 или на 8,75%.

Изложенная методика может быть использована при исследовании влияния не только относительных вариаций элементов определителя, но и вариаций абсолютных — т. е. когда после вариации элемент a_{ij} переходит в элемент $a_{ij} + \varepsilon_{ij}$, причем знак числа ε_{ij} может быть любым.

Так, определитель общего вида (115) перейдет после подобных вариаций в определитель:

$$\begin{vmatrix} a_{11}+\varepsilon_{11} & a_{12}+\varepsilon_{12} & \dots & a_{1n}+\varepsilon_{1n} \\ a_{21}+\varepsilon_{21} & a_{22}+\varepsilon_{22} & \dots & a_{2n}+\varepsilon_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1}+\varepsilon_{n1} & a_{n2}+\varepsilon_{n2} & \dots & a_{nn}+\varepsilon_{nn} \end{vmatrix}, \qquad (120)$$

а определитель (116) перейдет в:

$$\begin{vmatrix} 1+\varepsilon_{11} & 2+\varepsilon_{12} & 3+\varepsilon_{13} \\ 4+\varepsilon_{21} & 1+\varepsilon_{22} & 2+\varepsilon_{23} \\ 3+\varepsilon_{31} & 4+\varepsilon_{32} & 5+\varepsilon_{33} \end{vmatrix}. \qquad (121)$$

Первым этапом в исследовании вариаций ε_{ij} является, как и ранее, составление "таблицы знаков". Для определителя (121) "таблица знаков" сохраняет вид (117). Если для всех ε_{ij} будет $|\varepsilon_{ij}| = 0{,}01$, то при наиболее неблагоприятном сочетании знаков определитель (120) примет вид:

$$\begin{vmatrix} 0{,}99 & 1{,}99 & 3{,}01 \\ 4{,}01 & 0{,}99 & 2{,}01 \\ 3{,}01 & 4{,}01 & 4{,}99 \end{vmatrix} = 8{,}5628,$$

т. е. величина определителя увеличится на 0,5628 или на 7,06%.

Здесь также не трудно вычислить и наибольшее отклонение определителя в отрицательную сторону. Если заменить в "таблице знаков" (117) "плюсы" на "минусы", то придем к "таблице знаков":

$$\begin{vmatrix} + & + & - \\ - & + & - \\ - & - & + \end{vmatrix}$$

и к определителю:

$$\begin{vmatrix} 1{,}01 & 2{,}01 & 2{,}99 \\ 3{,}99 & 1{,}01 & 1{,}99 \\ 2{,}99 & 3{,}99 & 5{,}01 \end{vmatrix} = 7{,}4428,$$

т. е. определитель (116) уменьшился на 0,5572 или на 6,96%. Для определителя (120) при $|\varepsilon_{ij}| \le 0{,}01$ выполняются неравенства:

$$\det_{HOM} - 0{,}5572 \le \det_{\text{ист}} \le \det_{HOM} + 0{,}5628.$$

Важно отметить, что в теории определителей и в теории линейных алгебраических уравнений можно по единой методике рассматривать как вариации отличных от нуля коэффициентов, так и вариации (абсолютные вариации) коэффициентов, номинальное значение которых равно нулю, т. е. рассматривать "вариации нуля". Напомним, что в теории дифференциальных уравнений исследование "вариаций нуля" старших коэффициентов требует особого математического аппарата, разрабатываемого в теории "сингулярно возмущенных уравнений" (смотри также [2], стр. 67–68 и [10], стр. 18–19).

Для систем алгебраических уравнений все проще, хотя и здесь "вариации нуля" могут приводить к разным последствиям. Иногда малая "вариация нуля" может приводить к большим изменениям решений.

Рассмотрим простую систему уравнений:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 1; \\ a_{21}x_1 + a_{22}x_2 &= 1; \end{aligned}\right\}$$

при $a_{11} = 1$; $a_{12} = 0$; $a_{21} = 0$; $a_{22} = 0$. В этом случае $x_1 = 1$, а x_2 не существует. Пусть нулевой коэффициент a_{22} испытал вариацию и стал равным $a_{22} = \varepsilon$, где $0 \le \varepsilon \le 0{,}01$. В этом случае

$$D = \begin{vmatrix} 1 & 0 \\ 0 & \varepsilon \end{vmatrix} = \varepsilon; D_1 = \begin{vmatrix} 1 & 0 \\ 1 & \varepsilon \end{vmatrix} = \varepsilon; D_2 = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1; x_1 = \frac{\varepsilon}{\varepsilon} = 1; x_2 = \frac{1}{\varepsilon}.$$

Таким образом, "вариации нуля" привели в данном случае к тому, что определители D и D_1 изменились на малые величины (от значения нуль до значения ε), решение x_1 не изменилось, а решение x_2 стало теперь заключено в пределах от $x_2 = 1/\varepsilon$ до $x_2 \to +\infty$, или — иными словами — решение x_2 стало теперь заключено в неограниченном открытом интервале $(1/\varepsilon; +\infty)$.

Значительно чаще малые изменения нулевого элемента определителя не приводят к значительным изменениям его величины. Так, например, определитель:

$$\begin{vmatrix} 2 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1$$

при вариации его нулевого элемента переходит в определитель:

$$\begin{vmatrix} 2 & \varepsilon & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1 - \varepsilon$$

и вариация величины определителя при $\varepsilon \to 0$ сколь угодно мала.

Если произведена оценка вариаций определителей, входящих в формулы Крамера, то уже несложно — как это было показано в § 10 — оценить возможные вариации каждого из решений x_1; x_2; ...; x_n.

Появление методики точной оценки возможной вариации каждого из решений системы $AX = B$, происходящих из-за вариаций коэффициентов, точной оценки интервала, в котором может оказаться решение из-за вариаций коэффициентов, — является большим шагом вперед. Теперь, располагая, разумеется, оценками точности коэффициентов, оценками их возможных вариаций, можно точно оценить: у каких систем уравнений решения надежны и у каких — заведомо не надежны, причем не надежны при любых методах расчета.

Дело в том, что вариации коэффициентов, погрешности в их определении не являются, конечно, единственной причиной погрешности результатов расчета. Существуют еще погрешности методов вычислений — такие, как погрешности от округления промежуточных результатов, погрешности от конечного числа итераций в итерационных методах — и т. п. Часто именно уменьшению этих погрешностей уделяют наибольшее внимание, не замечая того, что погрешность, происходящая от вариаций параметров, не устранима при любом совершенствовании методов вычислений. Если, например, при $|\varepsilon_{ij}| \leq 0{,}01$ у определителя (116) его возможная погрешность, происходящая только из-за вариаций коэффициентов, подчинена неравенствам (119), то улучшить эти неравенства нельзя при любом совершенствовании методов вычислений.

Если, например, из-за вариаций коэффициентов интервал, внутри которого заключено решение x_1 системы $AX = B$, равен $\pm 0{,}1$, то нерационально вычислять решение с существенно большей точностью.

Отметим теперь, что при тех же оценках на относительные (или абсолютные) вариации коэффициентов: $|\varepsilon_{ij}| \leq \varepsilon_0$, вариация величины определителя возрастает с ростом его порядка.

Это обстоятельство удобно пояснить на примере так называемых "треугольных определителей":

$$\begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}, \tag{122}$$

у которых все элементы, лежащие выше главной диагонали (или наоборот — ниже главной диагонали), равны нулю. Такие определители равны произведению своих диагональных элементов. Пренебрегая всеми степенями выше первой малого в сравнении с единицей числа ε_0, получаем оценку:

$$\det\big|_{|\varepsilon_{ij}|=\varepsilon_0} - \det\big|_{|\varepsilon_{ij}|=0} = (a_{11} + \varepsilon_0 a_{11})(a_{22} + \varepsilon_0 a_{22}) \cdot \ldots \cdot (a_{nn} + \varepsilon_0 a_{nn}) -$$

$$- a_{11} a_{22} \cdot \ldots \cdot a_{nn} > n\varepsilon_0 \det\big|_{|\varepsilon_{ij}|=0},$$

которая растет с ростом порядка определителя n.

Возникает интересный вопрос: наибольшая возможная вариация величины определителя, а с ней и вариации решений x_i системы уравнений $AX = B$ при наиболее неблагоприятном сочетании знаков вариаций элементов определителя растут с ростом n. Вместе с тем вероятность реализации именно наиболее неблагоприятного сочетания знаков вариаций элементов определителя мала: она равна $\dfrac{1}{2^{n^2}}$ и очень быстро убывает с ростом n. Очевидно, что нужно учитывать величину вариации определителя не только при единственном и очень мало вероятном сочетании знаков вариаций элементов, но и при всех тех сочетаниях знаков, которые приводят к вариациям величины определителя, близким к максимальным значениям. Этот вопрос будет далее обсужден в § 14.

§ 14. Результаты численного эксперимента

Помимо вычисления наибольших вариаций определителей, происходящих из-за вариаций их элементов, представляет интерес оценить — с какой вероятностью может появиться та или иная вариация.

Для ориентации в вероятностях различных вариаций удобно прибегнуть к численному эксперименту, который выполнил М. В. Волошин. Определитель третьего порядка содержит 9 элементов, поэтому для него возможны $2^9 = 512$ сочетаний положительных и отрицательных вариаций $\pm\varepsilon_{ij}$. Положив $|\varepsilon_{ij}| = 0{,}01$, вычислялись все 512 определителей вида (114) для $a_{11} = 1$; $a_{12} = 2$; $a_{13} = 3$; $a_{21} = 4$; $a_{22} = 1$; $a_{23} = 2$; $a_{31} = 3$; $a_{32} = 4$; $a_{33} = 5$ и всех возможных сочетаний знаков чисел ε_{ij}. При $\varepsilon_{ij} = 0$ будет $\det_{НОМ} = 8$; $\det_{МОД} = 122$. Для данного определителя наиболее неблагоприятным окажется следующее сочетание знаков вариаций: $\varepsilon_{11} = +0{,}01$; $\varepsilon_{12} = -0{,}01$; $\varepsilon_{13} = +0{,}01$; $\varepsilon_{21} = +0{,}01$; $\varepsilon_{22} = -0{,}01$; $\varepsilon_{23} = +0{,}01$; $\varepsilon_{31} = +0{,}01$; $\varepsilon_{32} = +0{,}01$; $\varepsilon_{33} = -0{,}01$. При этом сочетании знаков вариаций элементов исследуемый определитель принимает вид:

$$\det = \begin{vmatrix} 0{,}99 & 1{,}98 & 3{,}03 \\ 4{,}04 & 0{,}99 & 2{,}02 \\ 3{,}03 & 4{,}04 & 4{,}95 \end{vmatrix}. \tag{123}$$

Если элементы с положительной вариацией обозначить условно через "+", а элементы с отрицательной вариацией — через "–", то определитель можно условно обозначить как:

$$\begin{vmatrix} - & - & + \\ + & - & + \\ + & + & - \end{vmatrix}. \tag{124}$$

Такое изображение особенно наглядно показывает, какое сочетание знаков вариаций в данном случае наиболее опасно.

Прямой расчет подтверждает, что наихудшее сочетание знаков вариаций элементов действительно соответствует определителю (123), а наибольшая величина вариации исходного определителя равна 1,6604.

Анализ всех 512 вычисленных вариаций (т. е. всех возможных вариантов) показывает, что только 16 из них (т. е. 3,125%) заключены в пределах от $\Delta_{\max}$ до $0{,}9\Delta_{\max}$.

Для другого определителя — для определителя:

$$\det_{\varepsilon=0} = \begin{vmatrix} 1 & 2 & 4 \\ 1 & 3 & 4 \\ 2 & 3 & 5 \end{vmatrix} = -3; \det_{\varepsilon} = \begin{vmatrix} 1+\varepsilon_{11} & 2+2\varepsilon_{12} & 4+4\varepsilon_{13} \\ 1+\varepsilon_{21} & 3+3\varepsilon_{22} & 4+4\varepsilon_{23} \\ 2+2\varepsilon_{31} & 3+3\varepsilon_{32} & 5+5\varepsilon_{33} \end{vmatrix} \quad (125)$$

аналогичный численный эксперимент показал, что при $|\varepsilon_{ij}| = 0{,}01$ наибольшая вариация определителя равна $|\Delta_{\max}| = 0{,}503$, и она достигается при следующем сочетании знаков вариаций элементов определителя:

$$\begin{vmatrix} - & - & + \\ - & + & - \\ + & + & - \end{vmatrix}.$$

На этот раз только у двух определителей из 512 их вариации заключены в пределах от $\Delta_{\max}$ до $0{,}9\Delta_{\max}$ и у восьми — в пределах от $\Delta_{\max}$ до $0{,}8\Delta_{\max}$.

Учитывая, что далеко не всегда вариация любого элемента определителя достигает своего наибольшего (по модулю) значения, следует признать, что вариация определителя очень редко достигает значений, близких к максимальным. Однако если такое редкое сочетание величин и знаков вариаций элементов определителя все же возможно, то его необходимо учитывать.

Из указанного обстоятельства вытекают важные следствия: пусть, например, усилия в каком-либо узле выпускаемого объекта численно равны решению x_1 системы из трех уравнений, и каждый из коэффициентов этой системы измеряется с точностью до $\pm 0{,}01$. Пусть выпущено 999 объектов, и все они прекрасно работают. К сожалению, это не гарантирует, что тысячный выпущенный объект не сломается и не приведет к аварии — причем он может сломаться не сразу, не на испытаниях, а после некоторого периода эксплуатации, в ходе которой параметры объекта и коэффициенты его математической модели испытывают неизбежные малые вариации. Причиной аварии может оказаться то, что на предыдущих выпущенных 999 объектах еще ни разу не реализовывалось опасное сочетание величин и знаков вариаций, а на тысячном оно реализовалось.

Гарантию от опаснейших поломок и аварий, гарантию от гибели людей в этих авариях может дать только хороший, надежный и достоверный расчет.

Для иллюстрации приведем некоторые из вычисленных значений определителя:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 3 & 4 & 5 \end{vmatrix} = 8$$

при вариациях его элементов на одну сотую от номинальных значений — т. е. при $|\varepsilon_{ij}| = 0{,}01$. Приводим "таблицы знаков" и соответствующие этим "таблицам" значения определителей:

$$\begin{vmatrix} + & + & + \\ + & + & + \\ + & + & + \end{vmatrix} \to 8{,}242408; \quad \begin{vmatrix} + & + & + \\ + & + & - \\ + & + & + \end{vmatrix} \to 8{,}1608; \quad \begin{vmatrix} + & + & + \\ + & - & + \\ + & + & + \end{vmatrix} \to 8{,}32401;$$

$$\begin{vmatrix} + & + & + \\ + & - & + \\ + & + & - \end{vmatrix} \to 9{,}0401; \quad \begin{vmatrix} + & + & + \\ + & - & + \\ - & - & + \end{vmatrix} \to 7{,}44309; \quad \begin{vmatrix} + & + & + \\ + & + & - \\ + & - & + \end{vmatrix} \to 7{,}28513;$$

$$\begin{vmatrix} + & + & + \\ - & + & + \\ - & + & + \end{vmatrix} \to 8{,}01798; \quad \begin{vmatrix} + & + & + \\ - & + & - \\ + & + & + \end{vmatrix} \to 7{,}99758; \quad \begin{vmatrix} + & + & - \\ + & + & + \\ + & - & + \end{vmatrix} \to 6{,}65004;$$

$$\begin{vmatrix} + & + & + \\ - & - & - \\ - & - & + \end{vmatrix} \to 7,21928; \quad \begin{vmatrix} + & + & - \\ - & + & + \\ + & + & - \end{vmatrix} \to 8,00081; \quad \begin{vmatrix} + & + & - \\ - & + & - \\ - & - & + \end{vmatrix} \to 6,38039;$$

$$\begin{vmatrix} + & - & + \\ + & - & + \\ + & + & - \end{vmatrix} \to 9,5952; \quad \begin{vmatrix} + & + & - \\ - & - & - \\ - & - & + \end{vmatrix} \to 6,4548; \quad \begin{vmatrix} + & - & + \\ + & - & + \\ - & - & + \end{vmatrix} \to 8,0192;$$

$$\begin{vmatrix} + & - & + \\ - & - & + \\ + & + & - \end{vmatrix} \to 9,4; \quad \begin{vmatrix} + & - & - \\ + & - & + \\ + & - & + \end{vmatrix} \to 7,29927; \quad \begin{vmatrix} + & - & - \\ + & - & + \\ + & - & - \end{vmatrix} \to 7,9992;$$

$$\begin{vmatrix} - & + & - \\ - & + & + \\ - & + & - \end{vmatrix} \to 7,9992; \quad \begin{vmatrix} - & + & - \\ - & + & + \\ - & - & + \end{vmatrix} \to 6,51835; \quad \begin{vmatrix} - & - & + \\ + & - & + \\ + & + & - \end{vmatrix} \to 9,6604.$$

Эти примеры показывают, как сильно может изменяться величина определителя при совсем небольших изменениях в "таблице знаков" — т. е. при изменениях знаков вариаций всего нескольких элементов определителя.

§ 15. Рассмотрение расчета одной из конструкций

Рассмотрим пример (**пример № 14**) расчета усилий в одной из простых конструкций, приведенный в хорошо известном учебнике [14] на стр. 205. Там исследуется нагруженная рама, показанная на рис. 3, где $l_1 = l_2 = l_3 = l$.

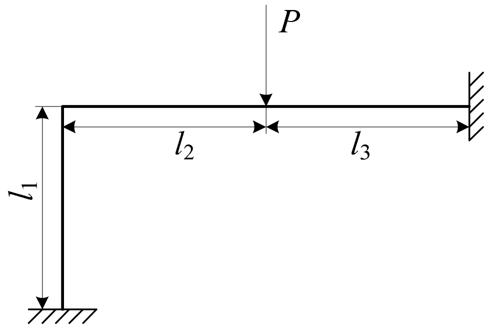

Рис. 3

Концы рамы заделаны, в середине горизонтального участка приложена сила P. Требуется рассчитать горизонтальную силу x_1, действующую в нижней заделке, вертикальную силу x_2 и изгибающий момент x_3.

Мы выбрали этот пример потому, что он приведен в известном, выдержавшем много изданий учебнике [14], приведен, безусловно, потому, что его автор и многочисленные преподаватели, пользовавшиеся учебником, считали расчет усилий x_1 и x_2 и момента x_3 достаточно показательным и достоверным.

В учебнике [14] на стр. 205–206 для определения x_1; x_2; x_3 была составлена система из трех уравнений:

$$\left.\begin{aligned} 14lx_1 + 12lx_2 + 15x_3 &= 3Pl; \\ 12lx_1 + 16lx_2 + 12x_3 &= 5Pl; \\ 5lx_1 + 4lx_2 + 6x_3 &= Pl, \end{aligned}\right\} \tag{126}$$

имеющая решение $x_1 = -\frac{1}{4}P; x_2 = \frac{7}{16}P; x_3 = \frac{1}{12}Pl.$

Достоверность и надежность этого решения в [14] не проверялась, хотя все коэффициенты, входящие в систему (126), могут испытывать малые изменения: из-за неточного соответствия реальных длин l_1; l_2; l_3 проектному чертежу, из-за неточного знания модуля упругости на различных участках рамы и т. д.

Оценка надежности решений системы (126) по "числу обусловленности" не заставляет насторожиться: для системы (126) имеем:

$$A = \begin{pmatrix} 14 & 12 & 15 \\ 12 & 16 & 12 \\ 5 & 4 & 6 \end{pmatrix}; \det A = 48; A^{-1} = \frac{1}{48}\begin{pmatrix} 48 & -12 & -96 \\ -12 & 9 & 12 \\ -32 & 4 & 80 \end{pmatrix}.$$

Евклидова норма матрицы A равна $\|A\| = 34{,}438$; та же норма обратной матрицы $\|A^{-1}\| = 2{,}907$ и число обусловленности $\|A\|\cdot\|A^{-1}\| = 100{,}111$ говорит о том, что система (126) достаточно хорошо обусловлена.

Проведем проверку с использованием модульных определителей. Для системы уравнений (110) имеем:

$$D_{НОМ} = \begin{vmatrix} 14 & 12 & 15 \\ 12 & 16 & 12 \\ 5 & 4 & 6 \end{vmatrix} = 2784 - 2736 = 48; D_{МОД} = 2784 + 2736 = 5520,$$

их отношение:

$$\frac{D_{МОД}}{D_{НОМ}} = \frac{5520}{48} = 115.$$

Аналогично:

$$D_{1НОМ} = \begin{vmatrix} 3 & 12 & 15 \\ 5 & 16 & 12 \\ 1 & 4 & 6 \end{vmatrix} = 732 - 744 = -12;\ D_{1МОД} = 1476$$

и их отношение:

$$\frac{D_{1МОД}}{D_{1НОМ}} = -123,$$

точно так же:

$$D_{2НОМ} = \begin{vmatrix} 14 & 3 & 15 \\ 12 & 5 & 12 \\ 5 & 1 & 6 \end{vmatrix} = 780 - 759 = 21;\ D_{2МОД} = 1539$$

и их отношение:

$$\frac{D_{2МОД}}{D_{2НОМ}} = 73,286,$$

в то время как:

$$D_{3НОМ} = \begin{vmatrix} 14 & 12 & 3 \\ 12 & 16 & 5 \\ 5 & 4 & 1 \end{vmatrix} = 668 - 664 = 4; D_{3МОД} = 1232$$

и их отношение:

$$\frac{D_{3МОД}}{D_{3НОМ}} = 308.$$

Таким образом, опираясь на формулы (92), устанавливаем, что "оценка сверху" для погрешности решения x_1 составит 714ε (суммируются относительные погрешности числителя и знаменателя). Для решения x_2 эта оценка составит 365ε, а для решения x_3 эта оценка равна 1269ε.

Эти "оценки сверху" сразу говорят о не слишком большой надежности решений x_1 и x_2 и о совсем плохой надежности решения

x_3. Встретившееся в данном примере сочетание величин решений (x_1 и x_2 много больше x_3) и их погрешностей как раз и приводит к тому, что норма относительных погрешностей всех решений (x_1; x_2 и x_3) сравнительно невелика, в то время как у решения x_3 погрешность большая. Мы уже говорили о том, что "число обусловленности" позволяет оценить лишь норму всех составляющих: от x_1 до x_n — вектора решений, а не погрешность конкретного решения x_i. Именно поэтому проверка по "числу обусловленности" не позволила в данном примере выявить плохую обусловленность решения x_3.

Используя для анализа определителя D_3 методику, изложенную в § 13, убеждаемся, что в данном случае при наиболее неблагоприятном сочетании знаков погрешностей, условно изображенном как:

$$\begin{vmatrix} - & + & - \\ + & - & + \\ + & - & + \end{vmatrix},$$

вариация определителя (в линейном приближении) будет равна $\Delta_{max} = 620\varepsilon_0$, а это означает, что уже при $|\varepsilon_0| \geq \dfrac{4}{620} = 0{,}0065$ определитель D_3, а с ним и решение x_3 могут изменить знак. То, что знак решения x_3 может измениться при столь малых вариациях коэффициентов системы, говорит о большой ненадежности решения x_3.

Исследование, проведенное в работе [10], показало, что даже если изменятся не все двенадцать коэффициентов системы уравнений (126), а только три коэффициента определителя D_3, и он примет вид:

$$D_3 = \begin{vmatrix} 14(1-\varepsilon) & 12(1+\varepsilon) & 3(1-\varepsilon) \\ 12 & 16 & 5 \\ 5 & 4 & 1 \end{vmatrix} = 4 - 308\varepsilon,$$

то уже при $\varepsilon > 1{,}3\%$ определитель D_3, а с ним и момент x_3 могут измениться коренным образом, могут изменить знак, стать отрицательными. Решение x_3 особенно не надежно и практического смысла не имеет. Неизбежные в ходе эксплуатации малые изменения параметров конструкции могут привести к ее разрушению при изменении знака момента x_3.

Любопытно, что учебник [14] только к 1979 году выдержал восемь изданий большими тиражами. Рассмотренный пример с простой конструкцией, показанной на рис. 3, приведенный в [14] на стр. 205, читали и решали десятки тысяч студентов. Тысячи преподавателей проверяли их решения. Тем не менее, никто не обратил внимания на то, что решение x_3, приведенное в [14], не надежно.

§ 16. Обоснование построения "таблиц знаков" и точной оценки вариаций определителей

В предыдущих параграфах была показана достижимость точной оценки изменений (вариаций) любой из составляющих x_1; x_2; ...; x_n вектора решений X системы $AX = B$ при вариациях коэффициентов матрицы A и вектора B. Точность оценки была подтверждена численным экспериментом.

Рассмотрим теперь вывод этой оценки и методику ее вычисления, а также методику построения "таблицы знаков" вариаций коэффициентов, о которой говорилось в § 13. Для вычисления изменения (вариации) всех составляющих x_1; x_2; ...; x_n вектора решений вычислим сначала вариации определителей D и D_i, на основе которых по формулам Крамера можно вычислить величины x_i, учитывая, что $x_i = \frac{D_i}{D}$.

Вернемся к определителю (114) и будем рассматривать его как функцию n^2 переменных величин ε_{ij}. Выпишем главную, линейную часть приращения этой функции (ее полный дифференциал), для чего сперва вычислим частные производные определителя по переменным ε_{ij}. Для достижения этой цели выполним разложение определителя по минорам той строки, в которую входит элемент $a_{ij} + \varepsilon_{ij}a_{ij}$. В это разложение будет входить член:

$$(a_{ij} + \varepsilon_{ij} a_{ij}) \cdot A_{ij}, \tag{127}$$

где A_{ij} — алгебраическое дополнение элемента a_{ij}, а во все остальные члены разложения определителя по минорам переменная ε_{ij} входить не будет. Поэтому частная производная определителя (114) по переменной ε_{ij} будет равна:

$$a_{ij} A_{ij}, \tag{128}$$

а главная, линейная часть приращения определителя (его полный дифференциал) будет равна:

$$\Delta_{лин} = \sum_{i=1;\,j=1}^{i=n;\,j=n} a_{ij} A_{ij} \Delta\varepsilon_{ij}, \tag{129}$$

где $\Delta\varepsilon_{ij}$ — приращение переменной ε_{ij}.

Предположим, что каждая из переменных ε_{ij} может изменяться независимо одна от другой и все они подчинены неравенствам:

$$|\varepsilon_{ij}| \le \varepsilon_0.$$

Тогда изменение определителя в направлении его возрастания будет наибольшим в том случае, если для всех i и j будет $|\varepsilon_{ij}| = \varepsilon_0$, а знак величины ε_{ij} будет совпадать со знаком произведения $a_{ij} A_{ij}$ — т. е. будет совпадать со знаком произведения элемента a_{ij}, на его алгебраическое дополнение. Наибольшее возможное приращение определителя в линейном приближении равно:

$$\Delta_{лин.макс} = \sum_{i=1;\,j=1}^{i=n;\,j=n} |\, a_{ij} A_{ij} \,| \cdot \varepsilon_0. \tag{130}$$

Отсюда следует очень простое правило построения "таблицы знаков", о которой говорилось в §§ 13 и 14: каждый элемент "таблицы знаков" ("плюс" или "минус") совпадает со знаком произведения $a_{ij} A_{ij}$ и равен нулю, если это произведение равно нулю.

Вернемся к рассмотренному в § 13 определителю (116). Для него:

$$A_{11} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = 1 \cdot 5 - 4 \cdot 2 = -3; A_{12} = -\begin{vmatrix} 4 & 2 \\ 3 & 5 \end{vmatrix} = -14 \tag{131}$$

и аналогично:

$$A_{13} = 13;\ A_{21} = 2;\ A_{22} = -1;\ A_{23} = 2;\ A_{31} = 1;\ A_{32} = 10;\ A_{33} = -7. \tag{132}$$

Из равенств (131) и (132) сразу следует "таблица знаков" (117), ранее приведенная в § 13. Используя таблицу (117) и задавшись $\varepsilon_0 = 0{,}01$, был вычислен определитель (118), равный 9,6604. Приращение определителя при вариациях его элементов оказалось равным 1,6604.

Воспользовавшись формулой (129), вычислим главную, линейную часть приращения определителя (116) при $\varepsilon_0 = 0{,}01$. Получим:

$$\begin{aligned}\Delta_{лин.макс} &= 0{,}01(1\cdot 3+2\cdot 14+3\cdot 13+4\cdot 2+\\ &+1\cdot 1+2\cdot 2+3\cdot 1+4\cdot 10+5\cdot 7)=1{,}61\end{aligned} \qquad (133)$$

и отличается от полного приращения (учитывающего нелинейности) на 0,0504 или на 3,13%.

Отметим, что в линейном приближении изменения определителя в положительную и в отрицательную сторону по абсолютной величине одинаковы. С учетом нелинейного характера зависимости величины определителя от ε_{ij} эти изменения уже не равны друг другу. Для вычисления наибольшего изменения определителя в отрицательную сторону достаточно, как это было указано в § 13, в уже полученной "таблице знаков" изменить все знаки на противоположные — т. е. заменить, например, для определителя (116) "таблицу знаков" (117) на таблицу

$$\begin{vmatrix} + & + & - \\ - & + & - \\ - & - & + \end{vmatrix}, \qquad (134)$$

и — с учетом новой таблицы, которую мы будем называть "обратной таблицей знаков", — вычислить определитель с измененными элементами. Для определителя (116) это вычисление было проведено в § 13, и оно показало, что при "таблице знаков" (134) определитель (116) при $\varepsilon_0 = 0{,}01$ уменьшился на 1,6196, что отличается от расчета в линейном приближении всего на 0,6%.

Перейдем теперь к обоснованию вычисления "таблицы знаков" для случая не относительных, а "абсолютных" вариаций, когда каждый из элементов определителя a_{ij} переходит в элемент $a_{ij} + \varepsilon_{ij}$, причем знаки величин ε_{ij} — рассматриваемых в дальнейшем как переменные — могут быть любыми.

При таких вариациях определитель общего вида (125) перейдет в определитель (120). Выделяя, как и ранее, конкретный элемент $a_{ij} + \varepsilon_{ij}$ и разлагая определитель по элементам той строки, в которую этот элемент входит, получим, что в этом случае частная производная определителя (120) по переменной ε_{ij} будет равна

A_{ij} — т. е. будет равна алгебраическому дополнению элемента a_{ij}. Для построения "таблицы знаков" достаточно вычислить алгебраические дополнения A_{ij}. Главная, линейная часть приращения определителя будет в данном случае равна:

$$\Delta_{лин} = \sum_{i=1;\,j=1}^{i=n;\,j=n} A_{ij}\Delta\varepsilon_{ij}, \tag{135}$$

где $\Delta\varepsilon_{ij}$ — приращение переменной ε_{ij}. Полагая, как и ранее, что для всех i и j соблюдается неравенство $|\varepsilon_{ij}| \le \varepsilon_0$, получим, что наибольшее приращение определителя в линейном приближении будет при $|\varepsilon_{ij}| = \varepsilon_0$, и оно равно:

$$\Delta_{лин.макс} = \sum_{i=1;\,j=1}^{i=n;\,j=n} \left|A_{ij}\right| \cdot \varepsilon_0. \tag{136}$$

Для вычисления точной оценки наибольшей возможной вариации определителя при вариациях его элементов надо составить "таблицу знаков"; с учетом этой таблицы и таблицы "обратной к ней" (с заменой "плюсов" на "минусы" и "минусов" на "плюсы") можно вычислить наибольшее возможное и наименьшее возможное значение определителя.

Для определителя (116), который после вариации всех своих элементов перейдет в определитель (121), алгебраические дополнения уже вычислялись (формулы (131) и (132)), с учетом которых "таблица знаков" сохраняет вид (117)). С учетом этой таблицы для $|\varepsilon_{ij}| = \varepsilon_0 = 0{,}01$ в § 13 были вычислены наибольшее изменение определителя в положительном направлении $\Delta_+ = 0{,}5628$ и в отрицательном направлении $\Delta_- = -0{,}5572$.

В линейном приближении максимальное изменение определителя, рассчитываемое по формуле (136), равно:

$$\Delta_{лин.макс} = \sum_{i=1;\,j=1}^{i=3;\,j=3} 0{,}01 \cdot \left|A_{ij}\right| = 0{,}53.$$

Таким образом, опять оно не сильно отличается от точных значений Δ_+ и Δ_- (от Δ_+ оно отличается на 6%, от Δ_- — на 5,13%).

Отметим, что если абсолютные или относительные значения вариаций всех элементов определителя n-го порядка равны между собой и равны ε_0, то вариации величины определителя при любой "таблице знаков" являются полиномами n-й степени от переменной ε_0. Это сразу следует из понимания определителя как суммы $n!$ произведений, составленных каждое из n элементов. Поскольку в каждый элемент определителей общего вида (114) и (120) входит ε_0 в первой степени, то каждое из $n!$ произведений — а значит, и их сумма — будет полиномом n-й степени от ε_0. При ε_0, малых в сравнении с единицей (для относительных вариаций, соответствующих определителю (114)), или при ε_0, малых по сравнению с a_{ij}, соответствующих определителю (120) — в вариации величины определителя будет доминировать член с первой степенью ε_0, и поэтому она будет примерно пропорциональна ε_0. Это позволяет, вычислив, например, вариацию определителя при $\varepsilon_0 = 0{,}01$, легко оценивать величину вариации определителя при $\varepsilon_0 = 0{,}001$, $\varepsilon_0 = 0{,}005$ и т. д.

Для иллюстрации на рис. 4 показана зависимость от ε_0 наибольших и наименьших значений определителя (116) при наиболее неблагоприятном сочетании знаков вариаций его элементов. Мы убеждаемся, что вплоть до $\varepsilon_0 = 0{,}07$ эти зависимости практически неотличимы от линейных. На рис. 5 эти же зависимости показаны для сравнительно больших ε_0, когда нелинейность уже существенна.

Приведем еще "таблицу знаков" для следующего определителя четвертого порядка:

$$\begin{vmatrix} 1{,}8 & -3{,}8 & 0{,}7 & -3{,}7 \\ 0{,}7 & 2{,}1 & -2{,}6 & -2{,}8 \\ 7{,}3 & 8{,}1 & 1{,}7 & -4{,}9 \\ 1{,}9 & -4{,}3 & -4{,}9 & -4{,}7 \end{vmatrix} = 616{,}9496. \qquad (137)$$

Она имеет вид

$$\begin{vmatrix} - & + & + & + \\ - & + & - & + \\ + & + & - & - \\ + & + & + & - \end{vmatrix}.$$

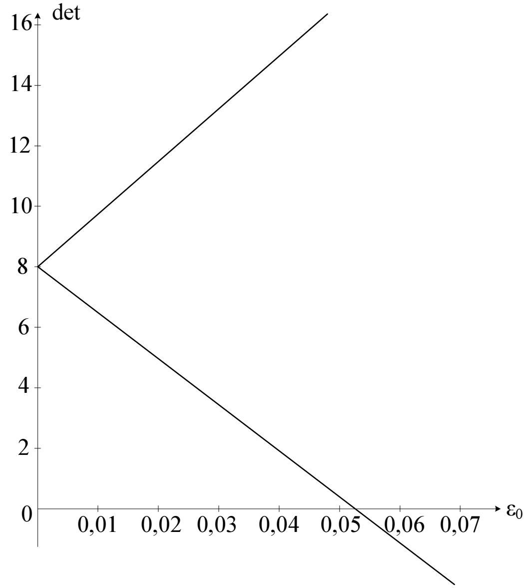
det
16
14
12
10
8
6
4
2
0
0,01
0,02
0,03
0,04
0,05
0,06
0,07
ε_0

Рис. 4

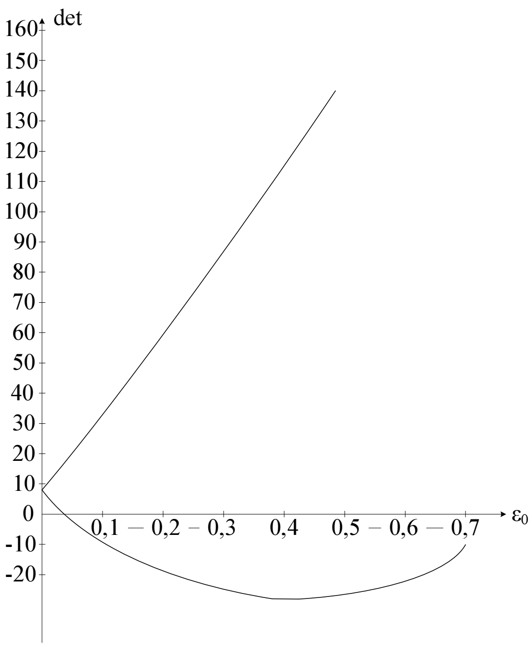
det
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
0,1
0,2
0,3
0,4
0,5
0,6
0,7
ε_0

Рис. 5

Если для всех i и j будет $|\varepsilon_{ij}| = 0{,}01$, то определитель (137) при наиболее неблагоприятном сочетании знаков вариаций его элементов перейдет в определитель

$$\begin{vmatrix} 1{,}8(1-0{,}01) & -3{,}8(1+0{,}01) & 0{,}7(1+0{,}01) & -3{,}7(1+0{,}01) \\ 0{,}7(1-0{,}01) & 2{,}1(1+0{,}01) & -2{,}6(1-0{,}01) & -2{,}8(1+0{,}01) \\ 7{,}3(1+0{,}01) & 8{,}1(1+0{,}01) & 1{,}7(1-0{,}01) & -4{,}9(1-0{,}01) \\ 1{,}9(1+0{,}01) & -4{,}3(1+0{,}01) & -4{,}9(1+0{,}01) & -4{,}7(1-0{,}01) \end{vmatrix} = 666{,}9333,$$

или 108,1% от значения определителя при $\varepsilon_{ij} = 0$.

§ 17. Рассмотрение особых частных случаев

При построении "таблицы знаков" особым частным случаем является равенство нулю одного или нескольких алгебраических дополнений A_{ij}. В этом частном случае методика, описанная в предыдущем параграфе, не позволяет выбрать в соответствующем месте "таблицы знаков" знак "плюс" или "минус".

Примером (**пример № 15**) может служить простой определитель:

$$\begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1, \tag{138}$$

у которого $A_{11} = 3$; $A_{12} = 0$; $A_{13} = -1$; $A_{21} = -1$; $A_{22} = 1$; $A_{23} = -1$; $A_{31} = -1$; $A_{32} = -1$; $A_{33} = 3$.

Условимся в "таблице знаков" на места, соответствующие нулевым алгебраическим дополнениям, ставить пока цифру "нуль". "Таблица знаков" для определителя (138) примет вид:

$$\begin{vmatrix} + & 0 & - \\ - & + & - \\ - & - & + \end{vmatrix}. \tag{139}$$

Главная, линейная часть приращения определителя (138) не будет зависеть от величины вариации элемента a_{12} и от ее знака, поскольку в данном случае (рассматриваем "абсолютные" вариации):

$$\begin{gathered} \Delta_{лин} = \varepsilon_{11} \cdot 3 + \varepsilon_{12} \cdot 0 + \varepsilon_{13} \cdot (-1) + \varepsilon_{21} \cdot (-1) + \\ + \varepsilon_{22} \cdot 1 + \varepsilon_{23} \cdot (-1) + \varepsilon_{31} \cdot (-1) + \varepsilon_{32} \cdot (-1) + \varepsilon_{33} \cdot 3. \end{gathered} \tag{140}$$

Если все $|\varepsilon_{ij}| \leq 0{,}01$, то наибольшее приращение определителя (139) в линейном приближении равно:

$$\Delta_{лин.макс} = 0{,}01(3+0+1+1+1+1+1+1+3) = 0{,}12 \qquad (141)$$

и не зависит от ε_{12}. Однако полное приращение определителя за счет нелинейных эффектов может зависеть от ε_{12}. Если в таблице знаков (139) вместо нуля поставить "плюс", то при $|\varepsilon_{ij}| = 0{,}01$ вместо определителя (138) получим определитель

$$\begin{vmatrix} 2{,}01 & 1{,}01 & 0{,}99 \\ 0{,}99 & 2{,}01 & 0{,}99 \\ 0{,}99 & 0{,}99 & 1{,}01 \end{vmatrix} = 1{,}1308, \qquad (142)$$

и, следовательно, $\Delta_{+} = 0{,}1308$, что на 9% больше $\Delta_{лин.макс}$, а если в "таблице знаков" (139) вместо нуля поставить знак "минус", то получим определитель

$$\begin{vmatrix} 2{,}01 & 0{,}99 & 0{,}99 \\ 0{,}99 & 2{,}01 & 0{,}99 \\ 0{,}99 & 0{,}99 & 1{,}01 \end{vmatrix} = 1{,}1204, \qquad (143)$$

тогда $\Delta_{+} = 0{,}1204$ или на 3,82% больше $\Delta_{лин.макс}$.

Таким образом, с учетом нелинейностей, изменение определителя все же зависит от знака ε_{ij}, соответствующего $A_{ij} = 0$, и в "таблице знаков" вместо нуля более правильно ставить двойной знак "плюс/минус"— т. е., например, таблицу (139) записывать в виде:

$$\begin{vmatrix} + & \pm & - \\ - & + & - \\ - & - & + \end{vmatrix} \qquad (144)$$

и вычислять определитель как для одного, так и для другого знака, стоящего в таблице.

Пример № 16. Рассмотрим определитель:

$$\begin{vmatrix} 1 & 2 & 3 \\ -2 & -4 & -5 \\ 3 & 5 & 6 \end{vmatrix} = 1, \tag{145}$$

у которого $A_{11} = 1$; $A_{12} = -3$; $A_{13} = 2$; $A_{21} = 3$; $A_{22} = -3$; $A_{23} = 1$; $A_{31} = 2$; $A_{32} = -1$; $A_{33} = 0$.

Если для всех i и j будет $|\varepsilon_{ij}| \le 0{,}01$, то наибольшее приращение определителя в линейном приближении:

$$\begin{gathered} \Delta_{лин.макс} = 0{,}01(1 + 2\cdot 3 + 3\cdot 2 + \\ + 2\cdot 3 + 4\cdot 3 + 5\cdot 1 + 3\cdot 2 + 5\cdot 1 + 6\cdot 0) = 0{,}47. \end{gathered} \tag{146}$$

Таблица знаков для определителя (145) имеет вид:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & \pm \end{vmatrix}. \tag{147}$$

Вычислим проварьированный определитель при выборе знака "плюс" в "таблице знаков" (147), когда она принимает вид:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}, \tag{148}$$

а проварьированный определитель равен:

$$\begin{vmatrix} 1{,}01 & 1{,}98 & 3{,}03 \\ -1{,}98 & -4{,}04 & -4{,}95 \\ 3{,}03 & 4{,}95 & 6{,}06 \end{vmatrix} = 1{,}4749.$$

Убеждаемся, что приращение определителя в этом случае достигает 0,4749, или 49,34% от номинального. Если же в "таблице знаков" (147) выбран знак "минус" и она приняла вид:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & - \end{vmatrix}, \tag{149}$$

то определитель (145) станет равен:

$$\begin{vmatrix} 1,01 & 1,98 & 3,03 \\ -1,98 & -4,04 & -4,95 \\ 3,03 & 4,95 & 5,94 \end{vmatrix} = 1,4934$$

и приращение станет равно 0,4934, или 49,34% его номинального значения, и 105% от максимального изменения в линейном приближении.

К наибольшему изменению определителя приводит в данном случае выбор знака "минус" в "таблице знаков" (147), и вариация определителя оказывается в этом случае в 49,34 раза больше наибольшей вариации каждого из его элементов.

При других сочетаниях знаков вариаций элементов определителя, его вариация может быть много меньше. Так, при $|\varepsilon_{ij}| = 0,01$ и сочетании знаков вариаций, соответствующему "таблице знаков":

$$\begin{vmatrix} + & - & + \\ + & - & + \\ + & - & - \end{vmatrix}, \tag{150}$$

когда определитель (145) переходит в:

$$\begin{vmatrix} 1,01 & 1,98 & 3,03 \\ -2,02 & -3,96 & -5,05 \\ 3,03 & 4,95 & 5,94 \end{vmatrix} = 1,099, \tag{151}$$

изменение определителя составит уже всего 9,9% от его номинального значения.

Отметим, что если в определителе много алгебраических дополнений, равных нулю (или близких к нулю), то величина вариации будет, в общем и целом, меньше, чем при не равных нулю алгебраических дополнениях.

Пример № 17. Рассмотрим в качестве примера определитель:

$$\begin{vmatrix} 2 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -1, \tag{152}$$

у которого $A_{11} = \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 0;\ \ A_{12} = -\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = -1$ и, аналогично, $A_{13} = 1$; $A_{21} = -1$; $A_{22} = 1$; $A_{23} = 0$; $A_{31} = 1$; $A_{32} = 0$; $A_{33} = -2$.

Если для всех i и j будет $\varepsilon_{ij} = \pm\varepsilon_0 = \pm 0{,}01$, то наибольшая вариация определителя в линейном приближении будет равна

$$\Delta_{лин.макс} = 0{,}01(0+2+1+2+1+0+1+0+2) = 0{,}09, \tag{153}$$

т. е. вариация в линейном приближении будет всего в 9 раз больше вариаций его элементов.

"Таблица знаков" с возможными вариантами в данном случае имеет вид:

$$\begin{vmatrix} \pm & - & + \\ - & + & \pm \\ + & \pm & - \end{vmatrix}. \tag{154}$$

Непосредственное вычисление определителя для всех $2^9 = 512$ сочетаний положительных и отрицательных вариаций его девяти элементов приводит к любопытному результату: к одинаковой (с точностью до шестого знака) величине определителя, равной –0,911897 (и, тем самым, к наибольшей вариации определителя в направлении возрастания $\Delta_+ = 0{,}088103$), приводят разные "таблицы знаков", а именно:

$$\text{первая: } \begin{vmatrix} - & - & + \\ - & + & - \\ + & + & - \end{vmatrix};\quad \text{вторая: } \begin{vmatrix} - & - & + \\ - & + & + \\ + & - & - \end{vmatrix};\quad \text{третья: } \begin{vmatrix} - & - & + \\ - & + & + \\ + & + & - \end{vmatrix};$$

$$\text{четвертая: } \begin{vmatrix} + & - & + \\ - & + & + \\ + & - & - \end{vmatrix};\ \text{пятая: } \begin{vmatrix} + & - & + \\ - & + & - \\ + & - & - \end{vmatrix};\ \text{шестая: } \begin{vmatrix} + & - & + \\ - & + & - \\ + & + & - \end{vmatrix}. \tag{155}$$

Все они являются вариантами таблицы (154) и приводят к одному и тому же значению определителя: –0,911897.

Любопытно, что к наибольшему отрицательному значению определителя (к значению: –1,092695) и, следовательно, к наибольшей вариации в направлении убывания (к Δ_{-} = –0,092695) приводит всего одна "таблица знаков", а именно таблица:

$$\begin{vmatrix} - & + & - \\ + & - & - \\ - & - & + \end{vmatrix}, \tag{156}$$

которая, как и следовало ожидать, является обратной по отношению к одному из вариантов таблицы (154).

Для определителя (152) при $|\varepsilon_{ij}| \le 0{,}01$ наибольшая вариация определителя в 9,2695 раз больше вариации каждого из его элементов.

Для других "таблиц знаков", отличных от таблиц (155) и (156), вариации определителя будут меньше.

Так, для "таблицы знаков"

$$\begin{vmatrix} + & + & + \\ + & + & + \\ + & + & + \end{vmatrix}$$

будет det = –1,030301 и вариация определителя равна 0,030301.

Для таблиц:

$$\begin{vmatrix} + & + & + \\ + & - & + \\ - & + & + \end{vmatrix} \text{ и } \begin{vmatrix} + & + & + \\ + & - & + \\ - & - & + \end{vmatrix}$$

будет одинаково det = –1,071509.

Для таблицы:

$$\begin{vmatrix} - & - & - \\ + & + & + \\ + & + & + \end{vmatrix}$$

имеем det = –1,009899 — т. е. вариация определителя будет в этом случае очень малой — даже меньшей, чем вариация каждого из элементов определителя. Еще меньшей будет вариация определителя при "таблицах знаков":

$$\begin{vmatrix} - & - & - \\ + & - & + \\ - & + & - \end{vmatrix} \text{ и } \begin{vmatrix} - & - & - \\ + & - & + \\ - & - & - \end{vmatrix},$$

для которых одинаково det = –1,0095.

Таблицам

$$\begin{vmatrix} - & - & - \\ - & + & + \\ + & - & + \end{vmatrix} \text{ и } \begin{vmatrix} - & - & - \\ - & + & - \\ + & - & + \end{vmatrix}$$

соответствует det = –0,98109 и одинаковая малая вариация определителя в направлении возрастания, а именно $\Delta_+ = 0{,}010891$.

Пример № 18. Рассмотрим теперь наиболее экзотический объект — определитель:

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 0, \tag{157}$$

у которого *все* алгебраические дополнения равны нулю. У подобных определителей при вариациях их элементов главная, линейная часть приращения определителя равна нулю, и величина приращения будет зависеть только от членов высшего порядка.

Для такого экзотического объекта, как определитель (157), общая теория построения "таблицы знаков" для вычисления наибольшего отклонения определителя от его номинального значения не работает. Прямой перебор показал, что в данном случае существует не одна, а несколько "таблиц знаков", приводящих (если для всех i и j $|\varepsilon_{ij}| = 0{,}01$) к одинаковым с точностью до четвертого знака максимальным вариациям определителя.

Так, к наибольшему отклонению в направлении возрастания, равному с точностью до четвертого знака 0,0012, приводят "таблицы знаков":

$$\begin{vmatrix} + & + & - \\ - & + & + \\ + & - & + \end{vmatrix} \text{ и } \begin{vmatrix} + & - & + \\ + & + & - \\ - & + & + \end{vmatrix}, \tag{158}$$

а к наибольшему отклонению в направлении убывания, равному –0,0012, приводят "таблицы знаков":

$$\begin{vmatrix} + & - & + \\ - & + & + \\ + & + & - \end{vmatrix} \text{ и } \begin{vmatrix} + & + & - \\ + & - & + \\ - & + & + \end{vmatrix}.$$

Таким образом, для определителя (157) его наибольшая вариация не превысит всего 0,12 от наибольшей вариации каждого из элементов определителя.

Рассмотрим первую из "таблиц знаков" (158), когда определитель (157) при $|\varepsilon_{ij}| = \varepsilon_0$ принимает вид:

$$\begin{vmatrix} 1+\varepsilon_0 & 1+\varepsilon_0 & 1-\varepsilon_0 \\ 1-\varepsilon_0 & 1+\varepsilon_0 & 1+\varepsilon_0 \\ 1+\varepsilon_0 & 1-\varepsilon_0 & 1+\varepsilon_0 \end{vmatrix} = 12\varepsilon_0^2 + 4\varepsilon_0^3.$$

Эта формула особенно наглядно показывает, что вариация определителя целиком зависит от нелинейных членов.

Рассмотрим вторую из "таблиц знаков", когда определитель (157) после вариаций его элементов принимает вид:

$$\begin{vmatrix} 1+\varepsilon_0 & 1-\varepsilon_0 & 1+\varepsilon_0 \\ 1+\varepsilon_0 & 1+\varepsilon_0 & 1-\varepsilon_0 \\ 1-\varepsilon_0 & 1+\varepsilon_0 & 1+\varepsilon_0 \end{vmatrix} = 12\varepsilon_0^2 + 4\varepsilon_0^3.$$

Это показывает, что и для первой и для второй из "таблиц знаков" (158) вариации определителей в точности равны друг другу. Зависимость величины определителей от ε_0 показана на рис. 6.

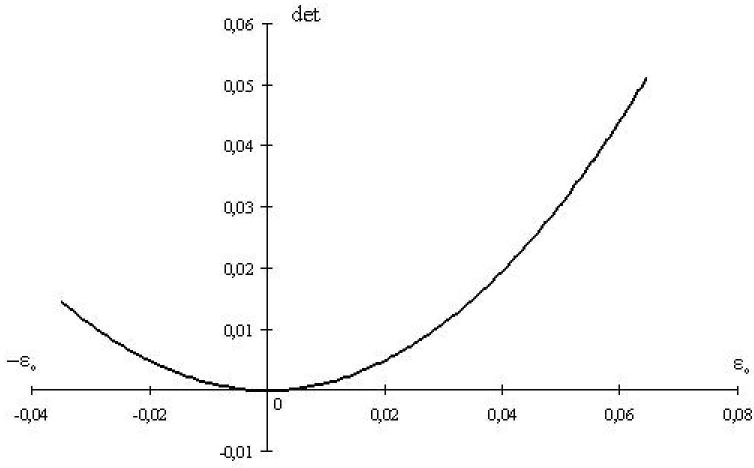

Рис. 6

Аналогично, составляя вариации элементов определителя (157) в соответствии с первой и второй "таблицами знаков" (158), убеждаемся, что

$$\begin{vmatrix} 1+\varepsilon_0 & 1-\varepsilon_0 & 1+\varepsilon_0 \\ 1-\varepsilon_0 & 1+\varepsilon_0 & 1+\varepsilon_0 \\ 1+\varepsilon_0 & 1+\varepsilon_0 & 1-\varepsilon_0 \end{vmatrix} = \begin{vmatrix} 1+\varepsilon_0 & 1+\varepsilon_0 & 1-\varepsilon_0 \\ 1+\varepsilon_0 & 1-\varepsilon_0 & 1+\varepsilon_0 \\ 1-\varepsilon_0 & 1+\varepsilon_0 & 1+\varepsilon_0 \end{vmatrix} = -12\varepsilon_0^2 - 4\varepsilon_0^3.$$

Разумеется, определители, у которых все алгебраические дополнения равны нулю, на практике почти никогда не встречаются, но как особый экзотический объект они интересны.

Изложенный материал позволяет дать дополнительное обоснование достоверности оценки вариации определителя при малых, но конечных вариациях его элементов. Строго говоря, "таблица знаков" строится для сколь угодно малых вариаций ε_{ij}. Не изменится ли она, если вариации ε_{ij} малы, но конечны? Знак, стоящий в "таблице знаков" на пересечении i-й строки и j-го столбца, может измениться в том случае, если алгебраическое дополнение A_{ij} очень мало и может изменить знак при переходе от номинальных значений элементов определителя к проварьированным. Для получения точной величины максимальной вариации определителя надо провести двойной расчет — как для знака в "таблице знаков", соответствующего малому алгебраическому дополнению, так и для противоположного знака.

Пример № 19. Рассмотрим вместо определителя (145) близкий к нему определитель:

$$\begin{vmatrix} 1 & 2,0001 & 3 \\ -2 & -4 & -5 \\ 3 & 5 & 6 \end{vmatrix} = 1,$$

у которого $A_{33} = \begin{vmatrix} 1 & 2,0001 \\ -2 & -4 \end{vmatrix} = 0,0002$ — т. е. A_{33} положительно, но очень мало. В этом случае, хотя "таблица знаков" определителя будет соответствовать таблице (148), вычисление проварьированного определителя надо проводить как для таблицы (148), так и для таблицы (149). Изменение определителя в данном случае будет наибольшим при выборе таблицы (149).

Правило:

если алгебраическое дополнение A_{ij} равно нулю или очень мало, то вычисление проварьированного определителя нужно проводить два раза: и для знака "плюс", и для знака "минус" в соответствующем месте "таблицы знаков". Это гарантирует точность оценки максимальной величины вариации определителя.

§ 18. Вычисление точных значений вариаций каждой из составляющих вектора решений

Методика вычисления вариаций определителей, изложенная в предыдущих параграфах, позволяет — основываясь на формулах Крамера:

$$x_i = \frac{D_i}{D} (i = 1, 2, \ldots, n)$$

получить точные значения вариаций всех составляющих x_1; x_2; ...; x_n вектора решений. Однако здесь есть некоторые тонкости.

В общем и целом вариация x_i будет наибольшей в том случае, если определитель D получит наибольшую возможную вариацию в направлении уменьшения величины определителя, а определитель D_i испытает наибольшую возможную вариацию в направлении увеличения.

В первом приближении вариация x_i, т. е. вариация частного от деления определителя D_i на определитель D, является суммой вариаций D и D_i. Однако, на самом деле, вариация x_i меньше суммы вариаций D и D_i, поскольку нужно учитывать, что вариации элементов определителей D и D_i не являются независимыми, поскольку $n - 1$ столбцов у определителей D и D_i совпадают.

Пример № 20. Возникающие проблемы поясним на простом примере системы:

$$\left.\begin{aligned} 2x_1 - x_2 &= 1; \\ x_1 + x_2 &= 2. \end{aligned}\right\} \qquad (159)$$

Для этой системы:

$$D = \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = 3; D_1 = \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} = 3; D_2 = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3\,, \qquad (160)$$

и, следовательно, $x_1 = \dfrac{3}{3} = 1; x_2 = \dfrac{3}{3} = 1$.

Остановимся на вычислении вариации x_2, предполагая, что относительные вариации всех шести коэффициентов системы (159) не превышают по модулю величину 0,01, но знаки их могут быть любыми.

Для определителя D наибольшая вариация в направлении убывания будет тогда, когда знаки вариаций элементов будут определяться его обратной "таблицей знаков", которая для определителя D имеет вид:

$$\begin{vmatrix} - & + \\ - & - \end{vmatrix}, \qquad (161)$$

а сам определитель при этом переходит в определитель:

$$\begin{vmatrix} 2(1-1,01) & -1(1+0,01) \\ 1(1-0,01) & 1(1-0,01) \end{vmatrix} = 2,96 \qquad (162)$$

(здесь и в дальнейшем вычисления проводились с точностью до третьего знака).

Для определителя D_2 наибольшая вариация в направлении возрастания будет при "таблице знаков":

$$\begin{vmatrix} + & - \\ - & + \end{vmatrix}, \qquad (163)$$

когда определитель D_2 примет вид:

$$\begin{vmatrix} 2(1+1,01) & 1(1-0,01) \\ 1(1-0,01) & 2(1+0,01) \end{vmatrix} = 3,06. \qquad (164)$$

Вариация определителя D_2 равна 0,06, или 2% от его номинального значения.

Вариация x_2 в первом приближении равна сумме вариаций D_2 и D и равна — в первом приближении — трем процентам.

Однако вариации элементов первого столбца определителей D и D_2 одинаковы (это вариации элементов a_{11} и a_{12}; они одинаковы и у D, и у D_2). Поэтому возможны два варианта расчета:

- *первый вариант* — ориентируемся в основном на "таблицу знаков" определителя D. В нем, а также в первом столбце определителя D_2 выбираем знаки в соответствии с "таблицей знаков" (161), и только во втором столбце определителя D_2 расставляем знаки вариаций в соответствии с его "таблицей знаков", т. е. в соответствии с таблицей (163). Получаем:

$$x_2 = \frac{\begin{vmatrix} 2(1-1,01) & 1(1-0,01) \\ 1(1-0,01) & 2(1+0,01) \end{vmatrix}}{\begin{vmatrix} 2(1-1,01) & -1(1+0,01) \\ 1(1-0,01) & 1(1-0,01) \end{vmatrix}} = \frac{3,02}{2,96} = 1,02, \qquad (165)$$

  т. е. вариация x_2 в направлении возрастания равна 0,02;

- *второй вариант* — ориентируемся на числитель, на определитель D_2. В нем, а также в первом столбце определителя D выбираем знаки в соответствии с "таблицей знаков" для определителя D_2 — т. е. с таблицей (163). И только во втором столбце определителя D выбираем знаки в соответствии с обратной "таблицей знаков" определителя D — т. е. в соответствии с таблицей (161).

  Получаем:

$$x_2 = \frac{\begin{vmatrix} 2(1+1,01) & 1(1-0,01) \\ 1(1-0,01) & 2(1+0,01) \end{vmatrix}}{\begin{vmatrix} 2(1+1,01) & -1(1+0,01) \\ 1(1-0,01) & 1(1-0,01) \end{vmatrix}} = \frac{3,06}{3} = 1,02, \qquad (166)$$

т. е. вариация x_2 при втором варианте комбинаций знаков вариаций элементов матрицы A и вектора B системы (159) равна 0,02, или 2% от его номинального значения.

В данном случае первый и второй варианты расчета привели к практически одинаковому результату.

Поясним: в первом варианте мы принимаем в расчет наибольшую возможную (в направлении убывания) вариацию определителя D. Вариация определителя D_i при этом варианте расчета может оказаться не максимально возможной.

Во втором варианте мы учитываем в расчете максимально возможную (в направлении возрастания) вариацию определителя D_2; при этом вариация определителя D может оказаться не максимально возможной. Теперь можно сформулировать общее правило вычисления наибольшей вариации любой составляющей x_i вектора решений X.

Общее правило

Исходным материалом для расчета служат оценки максимальных абсолютных величин вариаций элементов a_i и b_i в виде неравенств:

$$\left|\varepsilon_{ij}\right| \le \varepsilon_{ij0}; \left|b_i\right| \le b_{i0}$$

и вычисленные "таблицы знаков" определителей D и D_i в формулах Крамера.

Выполняем два варианта расчета:

- *первый вариант* ориентирован на определитель D в формуле $x_i = \dfrac{D_i}{D}$. Вычисляем наименьшее возможное значение определителя D по его "обратной таблице знаков". После этого вычисляем значение определителя D_i с учетом вариаций его элементов. При этом знаки вариаций элементов всех столбцов определителя, кроме i-го столбца определителя (этот столбец, напомним, совпадает со столбцом b_1; b_2; ...; b_n правой части), выбираем в соответствии с "таблицей знаков" определителя

D_i. Далее делим получившиеся с учетом вариаций значения определителей D и D_i друг на друга и получаем значение x_i;

- *второй вариант* ориентирован на определитель D_i. Знаки вариаций всех элементов определителя D_i выбираются в соответствии с его "таблицей знаков". После этого вычисляем определитель D_i и получаем его наибольшее возможное значение (при данной величине вариаций коэффициентов). После этого приступаем к вычислению определителя D с учетом вариаций его элементов. При этом знаки вариаций элементов всех столбцов определителя D, кроме i-го столбца, выбираем в соответствии с "таблицей знаков" определителя D_i. После того как вычислены определители D_i и D, вычисляем $x_i = \dfrac{D_i}{D}$ и сравниваем с результатом расчета x_i по первому варианту.

Заметим, что, используя приведенное правило, мы вычисляем наибольшую возможную вариацию x_i в положительном направлении — в направлении возрастания. Если необходимо вычислить наибольшую возможную вариацию x_i в отрицательном направлении, в направлении убывания, то нужно учитывать, что эта вариация будет наибольшей при условии, что вариация определителя D_i будет наибольшей в направлении убывания (т. е. знаки вариаций элементов будут соответствовать "обратной таблице знаков" определителя D_i), а вариация определителя D будет наибольшей в направлении возрастания (т. е. знаки вариаций элементов будут соответствовать прямой "таблице знаков" определителя D). Поскольку определители D и D_i имеют $n-1$ общих столбцов, то совместить эти противоречивые требования невозможно — и нужно, как и ранее, использовать два варианта расчета:

- *первый вариант* — как и ранее, он ориентирован на определитель D. В определителе D выбираем знаки всех его вариаций в соответствии с его "таблицей знаков", после чего вычисляем определитель с учетом вариаций. Затем в определителе D_i знаки вариаций элементов всех столбцов, кроме i-го столбца, расставляем в соответствии с "таблицей знаков" определителя D, а знаки вариаций i-го столбца расставляем в соответствии с "обратной таблицей знаков" определителя D_i. Потом вычисляем определитель D_i с учетом вариаций его элементов и делим D_i на D;

- *второй вариант* — как и ранее, он ориентирован на определитель D_i, знаки вариаций всех элементов которого выбираем в соответствии с его "обратной таблицей знаков", после чего вычисляем определитель D_i с учетом вариаций его элементов. В определителе D знаки вариаций всех его столбцов, кроме i-го, расставляем в соответствии с "обратной таблицей знаков" определителя D_i, а знаки вариаций i-го столбца расставляем в соответствии с "таблицей знаков" определителя D. Затем вычисляем определитель D с учетом вариаций его элементов и делим D_i на D. Далее — сравниваем в каком варианте расчета вариация x_i больше.

В целом общее правило вычисления вариаций x_i в направлении возрастания и в направлении убывания с учетом двух вариантов расчета получилось довольно громоздким. Для реального использования этого правила необходимо составить программу для вычислений на ЭВМ.

Продолжение примера № 20. Рассмотрим ту же самую систему (159) и рассчитаем наибольшую возможную вариацию x_2 в направлении ее убывания:

- *первый вариант* — учитывая "таблицы знаков" (161) и (162) получим, что для первого варианта, когда за основу берется "таблица знаков" определителя и только во втором столбце определителя D_2 знаки вариаций проставляются в соответствии с "обратной таблицей знаков" определителя D_2, будет:

$$x_2 = \frac{\begin{vmatrix} 2(1+1,01) & 1(1+0,01) \\ 1(1+0,01) & 2(1-0,01) \end{vmatrix}}{\begin{vmatrix} 2(1+1,01) & -1(1-0,01) \\ 1(1+0,01) & 1(1+0,01) \end{vmatrix}} = \frac{2,98}{3,03} = 0,983; \qquad (167)$$

- *второй вариант* — здесь за основу берется определитель D_2, знаки вариаций в котором расставляются в соответствии с его "обратной таблицей знаков". В соответствии с той же таблицей приходится расставлять знаки вариаций и в первом столбце определителя D. И только во втором его столбце знаки вариаций расставляются в соответствии с "таблицей знаков" определителя D, т. е. таблицей:

$$\begin{vmatrix} + & - \\ + & + \end{vmatrix}. \tag{168}$$

Таким образом,

$$x_2 = \frac{\begin{vmatrix} 2(1-1,01) & 1(1+0,01) \\ 1(1+0,01) & 2(1-0,01) \end{vmatrix}}{\begin{vmatrix} 2(1-1,01) & -1(1-0,01) \\ 1(1+0,01) & 1(1+0,01) \end{vmatrix}} = 0,967. \tag{169}$$

В данном случае к наибольшей величине вариации x_2 в направлении убывания приводит второй вариант и сочетание знаков вариаций элементов a_{11}; a_{12}; a_{21}; a_{22}; b_1; b_2, показанное в формуле (169) — т. е. вариация a_{11} имеет знак "минус", a_{12} — знак "минус", a_{21} — знак "плюс", a_{22} — знак "плюс", b_1 — знак "плюс", b_2 — знак "минус".

Окончательно устанавливаем: составляющая x_2 вектора решений простой системы уравнений (159), вследствие вариаций коэффициентов системы, не превышающих 0,01 от их номинальных значений, подчинена неравенствам:

$$0,967 \le x_2 \le 1,02, \tag{170}$$

причем оценка (170) является точной, поскольку можно указать такое конкретное сочетание знаков вариаций коэффициентов системы (159), при которых неравенства (170) превращаются в точные равенства.

Приведенный пример показывает, что точная оценка погрешности решений системы линейных алгебраических уравнений является значительно более сложной задачей, чем вычисление самого решения.

Пример № 21. Вернемся еще раз к рассмотренной в § 15 (пример № 14) системе уравнений (126). Ранее было показано, что составляющая решения x_3 совершенно не надежна. Это выявлялось уже на стадии исследования определителя D_3, поскольку при очень малых вариациях коэффициентов определителя D_3, при $\varepsilon_0 \geq 0{,}0065$, определитель D_3, а с ним и x_3 (т. е. момент, приложенный к концу рамы) меняет знак. Исследуем теперь x_1 и x_2.

Составляющие решения x_1 и x_2, определяемые по формулам Крамера через определители D_1; D_2 и D, равны:

$$x_{1HOM} = \frac{D_1}{D} = \frac{\begin{vmatrix} 3 & 12 & 15 \\ 5 & 16 & 12 \\ 1 & 4 & 6 \end{vmatrix}}{\begin{vmatrix} 14 & 12 & 15 \\ 12 & 16 & 12 \\ 5 & 4 & 6 \end{vmatrix}} = -\frac{12}{48} = -\frac{1}{4};$$

$$x_{2HOM} = \frac{D_2}{D} = \frac{\begin{vmatrix} 14 & 3 & 15 \\ 12 & 5 & 12 \\ 5 & 1 & 6 \end{vmatrix}}{\begin{vmatrix} 14 & 12 & 15 \\ 12 & 16 & 12 \\ 5 & 4 & 6 \end{vmatrix}} = \frac{21}{48} = 0,4375. \qquad (171)$$

Первоначальная простая проверка надежности их вычисления через модульные определители проводилась в § 15. Она показала, что если ориентироваться на "оценку сверху", которую обеспечивают модульные определители, то значения x_1 и x_2 нельзя признать надежными. Однако это — как уже ранее говорилось — не означает, что x_1 и x_2 обязательно ненадежны, а говорит лишь о необходимости дополнительной более тщательной (и более трудоемкой!) проверки.

Проведем эту проверку для x_2, причем — в отличие от примера № 20 — начнем с оценки вариаций в линейном приближении.

Вычисляя алгебраические дополнения для определителя D, получаем:

$$A_{11} = \begin{vmatrix} 16 & 12 \\ 4 & 6 \end{vmatrix} = 48; A_{12} = -\begin{vmatrix} 12 & 12 \\ 5 & 6 \end{vmatrix} = -12; A_{13} = \begin{vmatrix} 12 & 16 \\ 5 & 4 \end{vmatrix} = -32; \quad (172)$$

и, аналогично, $A_{21} = -12$; $A_{22} = 9$; $A_{23} = 4$; $A_{31} = -96$; $A_{32} = 4$; $A_{33} = 80$.

"Таблица знаков" определителя D принимает вид:

$$\begin{vmatrix} + & - & - \\ - & + & + \\ - & + & + \end{vmatrix}. \quad (173)$$

Если абсолютные величины вариаций всех элементов определителя D равны $|\varepsilon_{ij}| \le \varepsilon_0$, то наибольшая вариация определителя при наиболее неблагоприятном сочетании знаков вариаций его элементов будет (в линейном приближении) равна:

$$\begin{gathered} \Delta_{лин} = \varepsilon_0(14 \cdot 48 + 12 \cdot 12 + 15 \cdot 32 + 12 \cdot 12 + \\ +16 \cdot 9 + 12 \cdot 4 + 5 \cdot 96 + 4 \cdot 12 + 6 \cdot 80) = 2630\varepsilon_0 \end{gathered} \quad (174)$$

и определитель D удовлетворяет (в линейном приближении) неравенствам:

$$D_{HOM} - 2630\varepsilon_0 \le D \le D_{HOM} + 2630\varepsilon_0,$$

или — в относительных единицах:

$$1 - 54,79\varepsilon_0 \le \frac{D}{D_{HOM}} \le 1 + 54,79\varepsilon_0, \quad (175)$$

где D_{HOM} — номинальное значение определителя D.

Для учета слабой нелинейной зависимости вариации определителя от вариаций его элементов следует вычислить определитель D с проварьированными элементами, причем знаки вариаций соответствуют "таблице знаков" (173).

Если для всех i и j будет $|\varepsilon_{ij}| = 0{,}01$, то определитель D переходит в:

$$D_{\max} = \begin{vmatrix} 14(1+0{,}01) & 12(1-0{,}01) & 15(1-0{,}01) \\ 12(1-0{,}01) & 16(1+0{,}01) & 12(1+0{,}01) \\ 5(1-0{,}01) & 4(1+0{,}01) & 6(1+0{,}01) \end{vmatrix} =$$
$$= \begin{vmatrix} 14{,}14 & 11{,}88 & 14{,}85 \\ 11{,}88 & 16{,}16 & 12{,}12 \\ 4{,}95 & 4{,}04 & 6{,}06 \end{vmatrix} = 74{,}665 = D_{НОМ} + 26{,}665. \tag{176}$$

Формула (176) показывает, что для определителя D будет $\Delta_+ = 26{,}665$ и очень мало отличается от $\Delta_{лин}$, поскольку при $\varepsilon_0 = 0{,}01$ будет $\Delta_{лин} = 26{,}3$.

Для вычисления наибольшей вариации определителя D в направлении убывания знаки вариаций элементов следует выбирать в соответствии с "таблицей знаков", обратной по отношению к таблице (173), — т. е. с таблицей:

$$\begin{vmatrix} - & + & - \\ + & - & - \\ + & - & - \end{vmatrix}, \tag{177}$$

когда определитель D переходит в определитель:

$$D_{\min} = \begin{vmatrix} 14(1-0{,}01) & 12(1+0{,}01) & 15(1+0{,}01) \\ 12(1+0{,}01) & 16(1-0{,}01) & 12(1-0{,}01) \\ 5(1+0{,}01) & 4(1-0{,}01) & 6(1-0{,}01) \end{vmatrix} =$$
$$= \begin{vmatrix} 13{,}86 & 12{,}12 & 15{,}55 \\ 12{,}12 & 15{,}84 & 11{,}88 \\ 5{,}05 & 3{,}96 & 5{,}94 \end{vmatrix} = 21{,}863. \tag{178}$$

Снова убеждаемся, что $\Delta_- = 48 - 21{,}863 = 27{,}137$ мало отличается от $\Delta_{лин}$.

Те же вычисления проделаем для определителя D_2, для которого:

$$A_{11}=\begin{vmatrix}5 & 12\\ 1 & 6\end{vmatrix}=18;\ A_{12}=-\begin{vmatrix}12 & 12\\ 5 & 6\end{vmatrix}=-12;\ A_{13}=\begin{vmatrix}12 & 5\\ 5 & 1\end{vmatrix}=-13;$$
$$A_{21}=-\begin{vmatrix}3 & 15\\ 1 & 6\end{vmatrix}=-3;\ A_{22}=\begin{vmatrix}14 & 15\\ 5 & 6\end{vmatrix}=9;\ A_{23}=-\begin{vmatrix}14 & 3\\ 5 & 1\end{vmatrix}=1; \qquad (179)$$
$$A_{31}=\begin{vmatrix}3 & 15\\ 5 & 12\end{vmatrix}=-39;\ A_{32}=-\begin{vmatrix}14 & 15\\ 12 & 12\end{vmatrix}=12;\ A_{33}=\begin{vmatrix}14 & 3\\ 12 & 5\end{vmatrix}=34,$$

"таблица знаков" имеет вид:

$$\begin{vmatrix}+ & - & -\\ - & + & +\\ - & + & +\end{vmatrix}$$

и, следовательно, если для всех i и j будет $|\varepsilon_{ij}| \le \varepsilon_0$, то наибольшая возможная вариация определителя D_2 в линейном приближении равна:

$$\begin{gathered}\Delta_{лин}=\varepsilon_0(14\cdot 18+3\cdot 12+15\cdot 13+12\cdot 3+\\ +5\cdot 9+12\cdot 1+5\cdot 39+1\cdot 9+6\cdot 34)=984\varepsilon_0\end{gathered} \qquad (180)$$

и, следовательно, определитель D_2 в линейном приближении удовлетворяет неравенствам:

$$D_{2НОМ}-984\varepsilon_0 \le D_2 \le D_{2НОМ}+984\varepsilon_0, \qquad (181)$$

или, в относительных единицах:

$$1-47\varepsilon_0 \le \frac{D_2}{D_{2НОМ}} \le 1+47\varepsilon_0. \qquad (182)$$

Из формул (175) и (182) следует простая "оценка сверху" в линейном приближении для вариации x_2:

$$\begin{gathered}1-96{,}7\varepsilon_0=\\ =1-47\varepsilon_0-49{,}7\varepsilon_0 \le \frac{x_2}{x_{2НОМ}} \le 1+47\varepsilon_0+49{,}7\varepsilon_0=\\ =1+96{,}7\varepsilon_0,\end{gathered} \qquad (183)$$

но достигать верхней и нижней граней этой оценки отношение $\dfrac{x_2}{x_{2НОМ}}$ могло бы лишь в случае независимости вариаций определителей D и D_2. На самом деле эти вариации зависимы, и это позволяет дать более точную оценку.

Неравенства (175) и (182) показывают, что исследование определителя D_2 еще не позволяет вынести окончательное заключение о надежности или ненадежности составляющей решения x_2, и необходимо более детальное исследование вариации x_2 с учетом возможных сочетаний вариаций определителей D и D_2.

Вариация x_2 в направлении возрастания будет наибольшей в том случае, если определитель D_2, стоящий в числителе, в наибольшей степени возрастет, а определитель D в знаменателе в наибольшей степени уменьшится. Величина составляющей решения x_2 в наибольшей степени уменьшится, если определитель D_2 как можно более уменьшится, а определитель D как можно более возрастет при вариациях своих элементов. Однако вариации D и D_2 не являются независимыми и поэтому необходимо — как уже указывалось ранее — учитывать два варианта сочетания вариаций элементов определителей D и D_2.

Первый вариант расчета

Ориентируемся на знаменатель, на наибольшую (в направлении возрастания и направлении убывания) вариацию определителя D. В линейном приближении она — как показывает формула (175) — равна $\pm 2630\varepsilon_0$ и соответствует либо прямой, либо обратной "таблице знаков" определителя D, а именно, либо таблице (173), либо обратной к ней таблице (177).

В определителе D_2 первый и третий столбцы при первом варианте расчета совпадают с соответствующими столбцами определителя D, и поэтому знаки их вариаций не произвольны, а обязаны соответствовать "таблице знаков" определителя D. Знаки вариаций элементов второго столбца определителя D_2 произвольны, и к наибольшему возрастанию D_2 приведут вариации, знаки которых соответствуют "таблице знаков" (180). В целом "таблица знаков" определителя D_2, обеспечивающая его наибольшее возрастание, при выборе первого варианта расчета принимает вид:

$$\begin{vmatrix} - & - & + \\ + & + & - \\ + & + & - \end{vmatrix}. \tag{184}$$

Вычисляя определитель D_2 при $|\varepsilon_{ij}| = 0{,}01$ и "таблице знаков" (184), получаем, что в данном случае:

$$D_2' = \begin{vmatrix} 13{,}86 & 2{,}97 & 15{,}15 \\ 12{,}12 & 5{,}05 & 11{,}88 \\ 5{,}05 & 1{,}01 & 5{,}94 \end{vmatrix} = 14{,}919. \tag{185}$$

Поделив D_2' на определитель (178), получаем

$$x_{2\max} = \frac{14{,}919}{21{,}863} = 0{,}6824. \tag{186}$$

Это — максимальное значение x_2, достигаемое при комбинации знаков вариаций, соответствующей первому варианту расчета.

Теперь приступим к вычислению изменения x_2 в направлении убывания. Это изменение будет наибольшим, если определитель D_2 станет как можно меньше, а определитель D, стоящий в знаменателе, как можно больше. Таким образом, изменение x_2 в направлении убывания будет наибольшим, если знаки вариаций элементов определителя D, а также первого и третьего столбцов определителя D_2 будут соответствовать "таблице знаков" (173), а знаки вариаций второго столбца определителя D_2 будут соответствовать "обратной таблице знаков" для D_2. В целом "таблицы знаков" для дроби:

$$x_2 = \frac{D_2}{D}, \tag{187}$$

обеспечивающие наибольшее изменение x_2 в направлении убывания, имеют вид:

$$x_2 \to \frac{\begin{vmatrix} + & + & - \\ - & - & + \\ - & - & + \end{vmatrix}}{\begin{vmatrix} + & - & - \\ - & + & + \\ - & + & + \end{vmatrix}}. \tag{188}$$

Вычисляя определители D и D_2 при вариациях $|\varepsilon_{ij}| = 0{,}01$ и "таблицах знаков" вариаций, указанных в формуле (188), получаем:

$$x_{2\min} = \frac{\begin{vmatrix} 14{,}14 & 3{,}03 & 14{,}85 \\ 11{,}88 & 4{,}95 & 12{,}12 \\ 4{,}95 & 0{,}99 & 6{,}06 \end{vmatrix}}{\begin{vmatrix} 14{,}14 & 11{,}88 & 14{,}85 \\ 11{,}88 & 16{,}16 & 12{,}12 \\ 4{,}95 & 4{,}04 & 6{,}06 \end{vmatrix}} = \frac{28{,}93}{74{,}665} = 0{,}38746. \tag{189}$$

Это — минимальное значение x_2, достигаемое при комбинации знаков вариаций, соответствующих той, что показана в формуле (188). Перейдем теперь ко второму варианту расчета.

Второй вариант расчета

В этом варианте ориентируемся, прежде всего, на числитель, на определитель D_2, и учитываем, что вариация x_2 в направлении возрастания будет наибольшей тогда, когда D_2 примет наибольшее значение, а определитель D — наименьшее из возможных. Определитель D_2 будет наибольшим при знаках вариаций его элементов, соответствующим "таблице знаков" (180). В определителе D знаки вариаций первого и третьего столбца должны соответствовать той же таблице (180), и только во втором столбце знаки вариаций не связаны этим условием, и к минимальной величине D приводят знаки, соответствующие "обратной таблице знаков" для D, т. е. таблице (177).

В целом "таблицы знаков" для вычисления $x_{2\max}$ можно записать в виде:

$$x_2 \to \frac{\begin{vmatrix} + & - & - \\ - & + & + \\ - & + & + \end{vmatrix}}{\begin{vmatrix} + & + & - \\ - & - & + \\ - & - & + \end{vmatrix}}, \tag{190}$$

и тогда,

$$x_{2\max} = \frac{\begin{vmatrix} 14{,}14 & 2{,}97 & 14{,}85 \\ 11{,}88 & 5{,}05 & 12{,}12 \\ 4{,}95 & 1{,}01 & 6{,}06 \end{vmatrix}}{\begin{vmatrix} 14{,}14 & 12{,}12 & 14{,}85 \\ 11{,}88 & 15{,}84 & 12{,}12 \\ 4{,}95 & 3{,}96 & 6{,}06 \end{vmatrix}}; \tag{191}$$

вычислив определитель, получим:

$$x_{2\max} = \frac{30{,}878}{67{,}484} = 0{,}45756. \tag{192}$$

Переходя к вычислению наименьшего возможного при втором варианте расчета значения x_2, отметим, что это наименьшее значение будет достигаться, если числитель D_2 будет минимален, знаменатель D — максимален. Поскольку во втором варианте расчета начинаем с числителя, с D_2, то для D_2 учитываем его "обратную таблицу знаков", т. е. таблицу:

$$\begin{vmatrix} - & + & + \\ + & - & - \\ + & - & - \end{vmatrix}, \tag{193}$$

обратную к таблице (180).

В "таблице знаков" для знаменателя, для определителя D, первый и третий столбцы совпадают с соответствующими столбцами таблицы (193), и только второй столбец может быть выбран как обеспечивающий наибольшее возможное значение D. Это значение обеспечит столбец, совпадающий со вторым столбцом таблицы (173). В целом "таблица знаков" для знаменателя, для определителя D, имеет вид:

$$\begin{vmatrix} - & - & + \\ + & + & - \\ + & + & - \end{vmatrix}. \tag{194}$$

Вычисляя определители D и D_2 с вариациями, у которых $|\varepsilon_{ij}| = 0{,}01$ и имеющими знаки, соответствующие таблицам (193) и (194), получаем

$$x_{2\min} = \frac{\begin{vmatrix} 13{,}86 & 3{,}03 & 15{,}15 \\ 12{,}12 & 4{,}95 & 11{,}88 \\ 5{,}05 & 0{,}99 & 5{,}94 \end{vmatrix}}{\begin{vmatrix} 13{,}86 & 11{,}88 & 15{,}15 \\ 12{,}12 & 16{,}16 & 11{,}88 \\ 5{,}05 & 4{,}04 & 5{,}94 \end{vmatrix}} = \frac{11{,}23}{28{,}124} = 0{,}39766. \tag{195}$$

Сопоставляя неравенства (186), (189), (192) и (195) окончательно находим интервал, внутри которого может находиться составляющая x_2 вектора решений X системы уравнений (126):

$$0{,}38746 \le x_2 \le 0{,}6824, \tag{196}$$

или в относительных единицах

$$0{,}886 \le \frac{x_2}{x_{2НОМ}} \le 1{,}56. \tag{197}$$

При этом оценки (196) и (197) являются оценками точными — т. е. всегда можно указать такую комбинацию вариаций коэффициентов исследуемой системы уравнений, при которой левое или правое неравенство вида (196) выполняется со знаком равенства.

Так, например, равенство:

$$x_2 = 0,38746 \tag{198}$$

выполняется, если вариации матрицы коэффициентов системы уравнений (126) (при $|\varepsilon_{ij}| = 0,01$) соответствуют "таблице знаков":

$$\begin{vmatrix} + & - & - \\ - & + & + \\ - & + & + \end{vmatrix}, \tag{199}$$

а вариации коэффициентов вектора-столбца правой части удовлетворяют "таблице знаков"

$$\begin{vmatrix} + \\ - \\ - \end{vmatrix}. \tag{200}$$

§ 19. Общий алгоритм точной оценки погрешностей каждой из составляющих вектора решений

Материал предыдущего раздела показывает, что точная оценка возможной погрешности каждой из составляющих x_1; x_2; ...; x_n вектора X требует довольно громоздких вычислений.

Сперва для построения таблицы знаков необходимо вычислить алгебраические дополнения определителей D и D_i в формулах Крамера. Каждое из алгебраических дополнений у системы, состоящей из n уравнений, является определителем $n - 1$ порядка, и его вычисление требует, как известно [4; 5; 6], примерно $(n - 1)^3$ умножений. Всего вычисление всех n^2 алгебраических дополнений требует примерно n^5 умножений. Однако для построения "таблицы знаков" можно обойтись и без непосредственного вычисления алгебраических дополнений, а использовать обратную матрицу A^{-1}, для вычисления которой существуют удобные и хорошо разработанные программы, поскольку каждый из элементов обратной матрицы является частным от деления соответствующего алгебраического дополнения на определитель матрицы.

Кроме того, далеко не во всех случаях необходимо использовать все вычисления, изложенные в предыдущем параграфе. Начинать можно с простого — с оценки вариации решения "сверху", через модульные определители, которые были рассмотрены в § 10. Если эта оценка не указывает на подозрения в плохой обусловленности, то исследуемую систему можно с полным основанием считать хорошо обусловленной и дальнейших вычислений не проводить. Пример — система (159). Для нее

$$D_{НОМ} = \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = 2 + 1 = 3; \quad D_{МОД} = 2 + 1 = 3 \qquad (201)$$

и, аналогично, $D_{1НОМ} = 3$; $D_{1МОД} = 3$; $D_{2НОМ} = 3$; $D_{2МОД} = 5$. Таким образом, для системы (159) модульные определители мало отличаются (или даже совсем не отличаются) от номинальных значений определителей.

С учетом результатов, приведенных в § 10, это сразу говорит о том, что система уравнений (159) очень хорошо обусловлена, вариации коэффициентов системы оказывают малое влияние на ее решение, на x_1 и x_2. Подробное исследование, проведенное в качестве примера в § 18, разумеется, это подтвердило.

Еще один пример простой оценки: если в ходе вычисления вариаций определителей D_1; D_2; ...; D_n выявляется, что для интересующей нас составляющей решения x_i определитель D_i при тех вариациях его элементов, которые реально могут встретиться при его эксплуатации, меняет свой знак, а определитель D знака не меняет, то и x_i будет менять знак, а это сразу говорит о ненадежности решения, и вычисления на этом можно закончить. Так, в § 15 при рассмотрении примера № 14 уже исследование вариации определителя D_3 сразу показало, что вычисление составляющей x_3 вектора решения X системы уравнений (126) заведомо ненадежно.

Если же исследование определителей D_i и сравнение их с модульными определителями не говорит сразу о надежности или ненадежности решения, то необходимо провести исследование точных интервалов, внутри которых находятся интересующие нас составляющие x_1; x_2; ...; x_n вектора решения X, используя методику, приведенную в § 18.

Отметим, что хотя в примерах, рассмотренных в § 18, мы ограничивались вычислением вариаций составляющих решения X для $\varepsilon_0 = 0{,}01$, это не снижает общности исследования. Действительно, в § 16 было показано, что при малых ε_0 зависимость вариаций определителя от ε_0 с очень хорошей степенью точности близка к линейной (иллюстрацией могут служить рис. 4 и 5). Поэтому, когда, например, в § 18 в примере № 21 было вычислено, что при $\varepsilon_0 = 0$ будет $x_2 = 0{,}4375$, а при $\varepsilon_0 = 0{,}01$ имеем $x_2 = 0{,}6824 = 0{,}4375 + 0{,}2449$, то можно утверждать, что с очень хорошей точностью при $\varepsilon_0 = 0{,}001$ будет $x_2 \le 0{,}4375 + 0{,}02449 = 0{,}462$ и т. д.

Далее отметим, что хотя в приведенных примерах при вариациях коэффициентов a_{ij}, удовлетворяющих неравенствам $|\varepsilon_{ij}| \le \varepsilon_0$, расчет проводился для предельного, наихудшего варианта, когда для всех i и j было $\varepsilon_{ij} = \pm\varepsilon_0$, не представляет труда вычислить вариации определителей для любых конкретных значений ε_{ij}, разумеется, если эти значения нам известны. Главная трудность заключается в построении "таблицы знаков". Если она построена, то вариация определителя при любых ε_{ij} вычисляется легко.

Вернемся к примеру № 20 из § 18, где рассматривалась система (159) с определителем:

$$D = \begin{vmatrix} 2 & -1 \\ 1 & 1 \end{vmatrix} = 3, \qquad (202)$$

для которого "таблица знаков" имеет вид (161). Для определителя (202) и $|\varepsilon_{ij}| = 0{,}01$ было вычислено по формуле (162) его наименьшее возможное значение, равное 2,97. Если известно, что для элемента a_{11} будет $|\varepsilon_{11}| = 0$, а для остальных i и j остается $|\varepsilon_{ij}| = 0{,}01$, то наименьшее возможное значение определителя D можно вычислить по формуле (аналогу формулы (162)):

$$D_\varepsilon = \begin{vmatrix} 2 & -1{,}01 \\ 0{,}99 & 0{,}99 \end{vmatrix} = 2{,}98. \qquad (203)$$

Известные дополнительные сложности вносит наличие согласованных знаков вариаций некоторых коэффициентов. Вернемся к примеру № 1, рассмотренному в § 1. Предположим, что балка сместилась немного вправо по сравнению с ее проектным размещением. В этом случае параметр l_3 (длина балки правее правой опоры) получит вариацию положительного знака, но тогда параметр l_1 (длина балки слева от левой опоры) будет иметь вариацию непременно отрицательного знака.

Наличие согласованных вариаций уменьшает величину вариаций определителей и составляющих x_1; x_2; ...; x_n вектора решений X по сравнению со случаем полностью независимых вариаций.

Вернемся к рассмотренной в примере № 20 системе уравнений (159) и рассмотрим определитель D_2:

$$D_2 = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3, \tag{204}$$

для которого "таблица знаков" имеет вид:

$$\begin{vmatrix} + & - \\ - & + \end{vmatrix}. \tag{205}$$

Если для всех i и j будет $|\varepsilon_{ij}| = 0{,}01$, а знаки вариаций ε_{ij} не зависят друг от друга и могут быть любыми, то наибольшее возможное значение определителя (как уже вычислялось в примере № 20) равно:

$$\begin{vmatrix} 2{,}02 & 0{,}99 \\ 0{,}99 & 2{,}02 \end{vmatrix} = 3{,}06. \tag{206}$$

Теперь предположим, что вариации коэффициентов второго столбца зависимы друг от друга и могут быть либо обе положительными, либо обе отрицательными (напомним, что коэффициенты второго столбца определителя (204) — это коэффициенты правой части системы уравнений (159), так что подобная зависимость между вариациями возможна). Если обе вариации элементов второго столбца положительны, то определитель принимает вид:

$$\begin{vmatrix} 2{,}02 & 1{,}01 \\ 0{,}99 & 2{,}02 \end{vmatrix} = 3{,}04. \tag{207}$$

Если обе вариации отрицательны, то определитель будет равен:

$$\begin{vmatrix} 2{,}02 & 0{,}99 \\ 0{,}99 & 1{,}98 \end{vmatrix} = 3{,}02. \tag{208}$$

В обоих случаях вариации определителя, как и следовало ожидать, меньше, чем при независимых вариациях элементов.

Таким образом, учет зависимости между собой вариаций коэффициентов системы уравнений может уменьшить вариации входящих в формулы Крамера определителей и, в конечном счете, сузить интервалы, внутри которой находятся интервалы исследуемой системы уравнений.

Однако для того, чтобы воспользоваться этой возможностью, нужно доказать, что учет зависимости вариаций непременно изменит "таблицы знаков" определителей, переведет их в другие, новые "таблицы знаков", которые нужно найти. Сделать это чаще всего трудно, поэтому обычно используется расчет, предполагающий независимость вариаций коэффициентов.

Однако существует важный частный случай, заслуживающий отдельного рассмотрения, — это случай симметричных матриц A в системах уравнений $AX = B$. С этим частным случаем приходится нередко встречаться в задачах строительной механики, но он не вносит каких-либо осложнений в общий алгоритм.

Частный случай симметричных матриц

В строительной механике многие задачи расчета сводятся к вычислению решений таких систем алгебраических уравнений $AX = B$, у которых матрица коэффициентов A симметрична, т. е. $a_{ij} = a_{ji}$.

Пример № 22. Система уравнений:

$$\left.\begin{aligned} 3x_1 + x_2 + x_3 &= 4; \\ x_1 + 2x_2 + x_3 &= 2; \\ x_1 + x_2 + 2x_3 &= 1 \end{aligned}\right\}$$

имеет симметричную матрицу, поскольку $a_{12} = a_{21}$; $a_{13} = a_{31}$; $a_{23} = a_{32}$. Определитель системы:

$$D = \begin{vmatrix} 3 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 7,$$

естественно, тоже симметричен, а, например, определитель D_1

$$D_1 = \begin{vmatrix} 4 & 1 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 9$$

уже несимметричен, поскольку $a_{12} \neq a_{21}$.

Составляющая x_1 вектора решений X равна:

$$x_1 = \frac{D_1}{D} = \frac{7}{9}.$$

Для вычисления вариаций x_1, происходящих из-за вариаций коэффициентов a_{ij} и b_i, можно использовать изложенную ранее методику, первым шагом которой является построение "таблиц знаков" для определителей D и D_1. У определителя D построение "таблицы знаков" облегчает симметричность определителя: достаточно вычислить наиболее неблагоприятные знаки вариаций не для девяти, а, в данном случае, только для шести элементов — для элементов, лежащих на главной диагонали и выше нее. Для элементов, лежащих ниже главной диагонали, знаки расставляются по принципу симметрии — действительно, если, например, $a_{12} = a_{21}$, то и $\varepsilon_{12} = \varepsilon_{21}$.

Поскольку у определителя D будет:

$$A_{11} = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3; A_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = -1; A_{13} = \begin{vmatrix} 1 & 2 \\ 1 & 1 \end{vmatrix} = -1;$$

$$A_{22} = \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 5; A_{23} = -\begin{vmatrix} 3 & 1 \\ 1 & 1 \end{vmatrix} = -2;$$

$$A_{33} = \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 5,$$

то его "таблица знаков" имеет вид:

$$D \to \begin{vmatrix} + & - & - \\ - & + & - \\ - & - & + \end{vmatrix}.$$

Для определителя D_1 имеем:

$$A_{11} = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3; A_{12} = -\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = -3; A_{13} = \begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix} = 0;$$

$$A_{21}=-\begin{vmatrix}1&1\\1&2\end{vmatrix}=-1;\ A_{22}=\begin{vmatrix}4&1\\1&2\end{vmatrix}=7;\ A_{23}=-\begin{vmatrix}4&1\\1&1\end{vmatrix}=-3;$$

$$A_{31}=\begin{vmatrix}1&1\\2&1\end{vmatrix}=-1;\ A_{32}=-\begin{vmatrix}4&1\\2&1\end{vmatrix}=-2;\ A_{33}=\begin{vmatrix}4&1\\2&2\end{vmatrix}=6$$

и его "таблица знаков" имеет вид:

$$D_1\to\begin{vmatrix}+&-&\pm\\-&+&-\\-&-&+\end{vmatrix}.$$

Располагая "таблицами знаков" можно — используя уже изложенную методику — найти интервал, внутри которого находится x_1 при тех или иных вариациях ε_{ij} и b_i. Чтобы не повторять довольно громоздких вычислений, ограничимся вычислением вариаций определителей D и D_1 в линейном приближении для $|\varepsilon_{ij}|=\varepsilon_0=0{,}01$. Для определителя D будет

$$\Delta_{лин}=0{,}01(3\cdot 3+1+1+1+2\cdot 5+2+1+2+10)=0{,}37,$$

для D_1 имеем

$$\Delta_{лин}=0{,}01(4\cdot 3+3+2+3+1+2+2\cdot 6)=0{,}49.$$

Этих данных уже достаточно для заключения о том, что вариации x_1 будут небольшими и что рассматриваемая в примере № 22 система уравнений хорошо обусловлена. Об этом же можно судить и по "модульным определителям". Так, для определителя D имеем $D_{НОМ}=7$, а $D_{МОД}=21$. Аналогично, для D_1 будет $D_{1НОМ}=9$, а $D_{1МОД}=19$. Умеренная величина отношений $\dfrac{D_{МОД}}{D_{НОМ}}$ и $\dfrac{D_{1МОД}}{D_{1НОМ}}$ говорит о хорошей обусловленности x_1, столь же хорошо обусловлена и составляющая x_2 решения X исследуемой системы. Можно, разумеется, вычислить точную величину наибольших вариаций x_1 и найти интервал, внутри которого находится x_1 при тех или иных вариациях коэффициентов.

§ 20. Использование оценок вариаций при вычислении решений обыкновенных дифференциальных уравнений

В предыдущих параграфах оценки вариаций решений систем линейных алгебраических уравнений иллюстрировались на примерах из строительной механики, из сопротивления материалов. Здесь все просто: задачи расчета усилий и нагрузок в тех или иных конструкциях непосредственно сводятся к вычислению составляющих x_1; x_2; ...; x_n вектора X решения систем алгебраических уравнений вида $AX = B$.

Однако необходимость решать подобные уравнения возникает и в других задачах физики и техники — возникает уже как один из необходимых этапов решения задачи в целом, и в этих случаях знание возможных вариаций каждого из компонентов вектора X системы вида $AX = B$ является необходимой предпосылкой надежности решения задачи в целом.

Важным примером является проблема вычисления решений обыкновенных дифференциальных уравнений различных порядков или систем таких уравнений. Одним из этапов вычисления решения часто является определение постоянных интегрирования, входящих в общее решение дифференциального уравнения или в решение системы уравнений, а для определения постоянных интегрирования обычно приходится составлять и решать систему алгебраических уравнений.

Пример № 23. Требуется найти решение $x(t)$ дифференциального уравнения:

$$\ddot{x} - 3\dot{x} + 2 = 0, \tag{209}$$

удовлетворяющее начальным условиям: $x(0) = 1;\ \dot{x}(0) = 0$.

Характеристический полином уравнения (209):

$$\lambda^2 - 3\lambda + 2 \tag{210}$$

имеет корни: $\lambda_1 = 1$; $\lambda_2 = 2$, поэтому общее решение имеет вид

$$x(t) = C_1 e^t + C_2 e^{2t} \tag{211}$$

и, следовательно, $\dot{x}(t) = C_1 e^t + 2C_2 e^{2t}$.

Из начального условия $x(0) = 1$ следует, что

$$C_1 + C_2 = 1. \tag{212}$$

Из второго начального условия $\dot{x}(0) = 0$ следует, что:

$$C_1 + 2C_2 = 0. \tag{213}$$

Равенства (212) и (213) образуют систему двух уравнений для определения двух постоянных интегрирования C_1 и C_2.

Для больших систем дифференциальных уравнений порядок системы алгебраических уравнений, которой удовлетворяют постоянные интегрирования C_i, может быть большим.

Так, в учебном пособии [10] на стр. 83 были рассмотрены уравнения системы управления частотой вращения электропривода. Поведение двух переменных x_1 и x_2 (где x_1 — это отклонение частоты вращения от номинальной, x_2 — отклонение вращающего момента от номинального значения) описывается двумя уравнениями, одно из которых третьего, а одно первого (относительно x_2) порядка. Общие решения такой системы, $x_1(t)$ и $x_2(t)$, имеют вид:

$$x_1(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t} + C_3 e^{\lambda_3 t} + C_4 e^{\lambda_4 t}, \tag{214}$$

где λ_1; λ_2; λ_3; λ_4 — корни характеристического полинома — полинома четвертой степени. Аналогичный вид имеет и решение $x_2(t)$.

Для определения произвольных постоянных из начальных условий здесь необходимо найти решение системы, состоящей из четырех уравнений для неизвестных C_1; C_2; C_3; C_4.

При этом необходимо иметь в виду, что корни характеристического полинома любой системы могут быть вычислены лишь с конечной, ограниченной точностью, и уже поэтому все коэффициенты системы уравнений, из которых определяются постоянные интегрирования, несут в себе неизбежные погрешности. Для того, чтобы вычисленное решение системы дифференциальных уравнений было надежным, нужно обязательно установить — как связаны неизбежные погрешности вычисления постоянных интегрирования с погрешностями коэффициентов системы, которой эти постоянные удовлетворяют. Это можно сделать методами, описанными в предыдущих параграфах. Поскольку существуют многочисленные примеры (продемонстрированные в предыдущих параграфах), когда малые погрешности в коэффициентах приводят к большим погрешностям в составляющих решения x_1; x_2; ...; x_n (а значит, и в постоянных интегрирования C_1; C_2; ...; C_n), то для получения надежного решения системы дифференциальных уравнений необходимо использовать методику расчета возможных вариаций составляющих решения, изложенную в предыдущих параграфах.

Отметим, что если исследуемая система приведена к нормальной форме Коши — т. е. к системе из n уравнений первого порядка:

$$\left.\begin{aligned} \dot{x}_1 &= f_1(x_1; x_2; ...; x_n; t); \\ \dot{x}_2 &= f_2(x_1; x_2; ...; x_n; t); \\ &\dots\dots\dots\dots\dots\dots\dots\dots; \\ \dot{x}_n &= f_n(x_1; x_2; ...; x_n; t), \end{aligned}\right\}, \qquad (215)$$

то, на первый взгляд, кажется, что трудностей в определении постоянных интегрирования можно избежать: действительно, если известны начальные условия для всех переменных в системе (215), известны величины $x_i(0)$, то решения системы (215) можно непосредственно получить одним из численных методов, в том числе и с использованием вычислительной техники.

Однако, как это было подробно показано в [2], если первоначальные дифференциальные уравнения математической модели исследуемого объекта не имели формы системы, состоящей из n

уравнений первого порядка, а состояли из системы нескольких уравнений различных порядков, то приведение подобной системы к нормальной форме Коши, к форме уравнений (215), может совершенно исказить реальную зависимость решений математической модели объекта от вариаций коэффициентов, исказить реальную зависимость поведения исследуемого объекта от вариаций его параметров. Поэтому для получения надежного решения, которое правильно отражает реальное поведение исследуемого объекта, к преобразованиям математической модели, в том числе и к преобразованию ее к нормальной форме Коши, следует относиться осторожно, проверять его правомерность, проверять — не исказилось ли влияние вариаций параметров объекта на его истинное поведение. Обо всем этом более подробно рассказано в [2].

Отметим также, что необходимость составлять системы алгебраических уравнений, решать и оценивать погрешности решения возникает при решении краевых задач, когда условия для искомых функций и их производных задаются не в одной, а в нескольких точках (краевые условия).

Необходимость составлять и решать системы алгебраических уравнений возникает и при решении дифференциальных уравнений методами операционного исчисления.

§ 21. Применения к решению интегральных уравнений

Хорошо известно (см., например, [8, 9, 32, 33, 34]), что решение многих проблем техники и физики сводится к вычислению решений интегральных уравнений различных типов (уравнения Фредгольма первого и второго рода, уравнения Вольтерра первого и второго рода, сингулярные уравнения, уравнения с вырожденным ядром и др.).

В интегральных уравнениях искомая функция находится под знаком интеграла. Так, в уравнении Фредгольма второго рода:

$$y(x) - \lambda \int_a^b K(x;s)y(s)ds = f(x) \qquad (216)$$

искомой функцией будет $y(x)$. Функция $f(x)$ — известная функция (правая часть), функция $K(x; s)$ двух переменных x и s называется *ядром*, а постоянное число λ — *параметром* уравнения.

Основной метод решения интегральных уравнений основан на замене интеграла конечной суммой с помощью одной из квадратурных формул (формул приближенного интегрирования), т. е. на замене:

$$\int_a^b F(x)dx \approx \sum_{j=1}^{n} A_j F(x_j), \qquad (217)$$

где x_j — абсциссы точек отрезка $[a, b]$, A_j ($j = 1; 2; ...; n$) — коэффициенты квадратурной формулы (прямоугольников, трапеций или другой). Заменяя приближенно интеграл в уравнении (216) по формуле (217), получим:

$$y_i - \lambda \sum_{j=1}^{n} A_j K_{ij} y_j = f_i, \qquad (218)$$

где $y_i = y(x_i)$; $K_{ij} = K(x_i; y_j)$; $f_i = f(x_i)$ — т. е. получим систему линейных алгебраических уравнений относительно y_i. Решив эту систему, получим таблицу приближенных значений y_i в точках x_i. Это позволит записать приближенное решение уравнения (216) в виде:

$$y(x) = f(x) + \lambda \sum_{j=1}^{n} A_j K(x; x_j) y_j. \qquad (219)$$

Приведем пример (заимствованный из [27], стр. 294): используя квадратурную формулу Симпсона при $n = 2$, найти приближенное решение уравнения:

$$y(x) + \int_0^1 x e^{xs} y(s) ds = e^x. \qquad (220)$$

Решение: поскольку для квадратурной формулы Симпсона $A_0 = A_2 = \frac{1}{6}$; $A_1 = \frac{2}{3}$; $x_0 = 0$; $x_1 = 0{,}5$; $x_2 = 1$, то для уравнения (220) при использовании формулы Симпсона можем записать:

$$y(x) + \frac{1}{6}(xe^{0\cdot x} y_0 + 4xe^{0{,}5\cdot x} y_1 + xe^{1\cdot x} y_2) = e^x. \qquad (221)$$

Полагая в этом равенстве последовательно:

$$x = x_0 = 0;\ x = x_1 = 0{,}5;\ x = x_2 = 1,$$

получаем систему из трех уравнений для y_0; y_1; y_2:

$$y_0 = 1;$$

$$y_1 + \frac{0{,}5}{6}(y_0 + 4e^{0{,}25} y_1 + e^{0{,}5} y_2) = e^{-0{,}5};$$

$$y_2 + \frac{1}{6}(y_0 + 4e^{0{,}5} y_1 + e y_2) = e,$$

решив которую находим: $y_0 = 1$; $y_1 = 1{,}0002$; $y_2 = 0{,}9995$, после чего приближенное решение уравнения (221) записывается в виде:

$$y(x) = e^x - \frac{x}{6}(1 + 4{,}001 e^{\frac{x}{2}} + e^x).$$

Другие примеры сведения решения интегральных уравнений к системам алгебраических уравнений приведены в [27], стр. 295–303.

Надежность решений интегральных уравнений зависит, таким образом, от надежности решений систем алгебраических уравнений, все коэффициенты которых уже вследствие приближенности замены интеграла на конечную сумму известны лишь с конечной, ограниченной точностью. Поскольку мы теперь располагаем методикой точной оценки погрешности (вариации) каждой из составляющих x_1; x_2; ...; x_n вектора решений X системы $AX = B$ в зависимости от погрешностей (вариаций) коэффициентов (методикой, приведенной в § 18), то это позволяет существенно улучшить надежность вычисления решений интегральных уравнений, а тем самым — и надежность решения многочисленных задач техники и физики, сводящихся к интегральным уравнениям.

§ 22. Применения к решению дифференциальных уравнений в частных производных

Многие задачи физики и техники приводят к необходимости вычислять решения дифференциальных уравнений в частных производных — таких, например, как уравнение Лапласа:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \tag{222}$$

уравнение Пуассона:

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x; y) \tag{223}$$

и многих других.

Одним из самых распространенных методов численного решения уравнений с частными производными является метод сеток, или метод конечных разностей, в основе которого лежит идея приближенной замены производных конечно-разностными отношениями.

Если рассматриваются уравнения с двумя независимыми переменными x и y, то решением является функция $u(x; y)$ внутри некоторой области G с границей Γ на плоскости $x0y$ (рис. 7). Искомая функция $u(x; y)$ внутри области удовлетворяет соответствующему уравнению в частных производных (например — уравнению (222) или (223)), а на границе Γ удовлетворяет граничным условиям.

В методе сеток на плоскости $x0y$ строятся два семейства параллельных прямых:

$$x = x_0 \pm ih (i = 0; 1; 2; \ldots);$$

$$y = y_0 \pm kl (k = 0; 1; 2; \ldots).$$

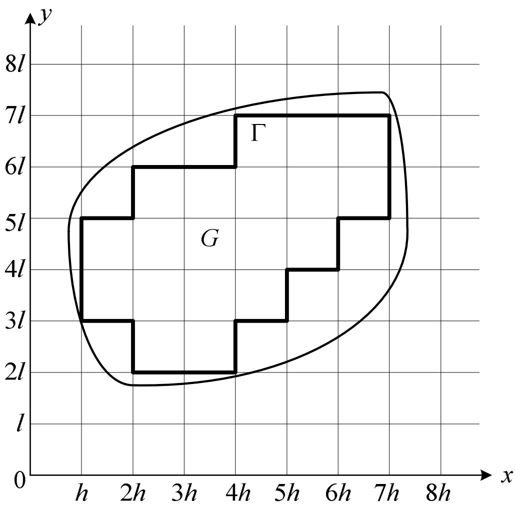

Рис. 7

Точки пересечения этих прямых называются *узлами*. Значения искомой функции в узлах сетки обозначаются через $u_{kl} = u(x_0 + ih;$ $y_0 + kl)$ и в каждом внутреннем узле частные производные приближенно заменяются разностными отношениями:

$$\left(\frac{\partial u}{\partial x}\right)_{ik} \approx \frac{u_{i+1;k} - u_{i-1;k}}{2h};$$

$$\left(\frac{\partial u}{\partial y}\right)_{ik} \approx \frac{u_{i+1;k} - u_{i-1;k}}{2l}.$$

Аналогично заменяются приближенно и производные второго порядка:

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{ik} \approx \frac{u_{i+1;k} - 2u_{ik} + u_{i-1;k}}{h^2};$$

$$\left(\frac{\partial u}{\partial y}\right)_{ik} \approx \frac{u_{i+1;k} - 2u_{ik} + u_{i-1;k}}{l^2}.$$

Аналогичные замены производятся и для узлов, лежащих вблизи границы области.

Для уравнения Лапласа, применив указанные замены, приходим к уравнениям:

$$\frac{u_{i+1;k} - 2u_{ik} + u_{i-1;k}}{h^2} + \frac{u_{i+1;k} - 2u_{ik} + u_{i-1;k}}{l^2} = 0,$$

т. е. к системе алгебраических уравнений для u_{ik}.

Конкретные примеры систем алгебраических уравнений, которые появляются при решении различных уравнений в частных производных методом сеток, приведены в [27], стр. 266–286, в [32] и во многих других учебниках и руководствах.

Погрешность замены дифференциального уравнения разностным, т. е. остаточный член R_{ik} для уравнения Лапласа, оценивается, как известно, неравенством:

$$\left|R_{ik}\right| \le \frac{h^2}{6} M_4, \tag{224}$$

где:

$$M_4 = \max_G \left\{ \left|\frac{\partial^4 u}{\partial x^4}\right|; \left|\frac{\partial^4 u}{\partial y^4}\right| \right\}.$$

Погрешность приближенного решения, полученного методом сеток, складывается из трех погрешностей:

- погрешности замены производной конечной разностью;
- аналогичной погрешности аппроксимации условий на границе области;
- погрешности вычисления решения получившейся большой системы алгебраических уравнений.

Третий источник погрешности можно уменьшить до необходимого на практике уровня за счет усовершенствования методов расчета, увеличения количества вычислительных операций (числа итераций в итерационных методах и т. п.).

Первый и второй источники погрешности не могут быть уменьшены за счет вычислительных методов. Они зависят, в основном, от густоты сетки, от величины интервалов h и l в сеточных методах. Для обеспечения надежности вычислений очень важно оценить погрешность окончательного результата в зависимости от погрешностей коэффициентов промежуточной системы алгебраических уравнений.

Отметим, что хотя порядок системы алгебраических уравнений, которой в методе сеток приближенно заменяются исходные уравнения в частных производных, часто оказывается очень большим, учитывать вариации коэффициентов, стоящих при u_{ik}, обычно необходимо лишь для узлов, лежащих вблизи границы области. Это облегчает вычисление оценок возможных вариаций величин u_{ik}, от которых зависит и надежность расчета в целом.

Теперь, когда разработана методика точной оценки вариаций и погрешностей любой из составляющих вектора решения X, изложенная в § 18, есть возможность существенно увеличить надежность вычисления решений дифференциальных уравнений в частных производных.

Приведенные в §§ 20, 21, 22 примеры не исчерпывают, разумеется, всего многообразия применения решений систем алгебраических уравнений в самых различных прикладных задачах. В публикации [25] указывается, что "более 70% всех математических расчетов приходится на решение систем линейных алгебраических уравнений". Поэтому использование точных оценок погрешностей их решений — оценок, описанных в данной работе — может обеспечить существенное повышение надежности расчетов, снизить число аварий и катастроф, значительная часть которых происходит из-за неточностей и ошибок расчета.

§ 23. Примеры аварий и катастроф. Анализ их причин

Недостаточная разработка проблемы оценки погрешности решений систем линейных алгебраических уравнений, отсутствие до последнего времени надежных методов вычисления вариаций решений, происходящих из-за неизбежных на практике вариаций параметров проектируемого объекта в ходе его эксплуатации, приводили к тому, что эти вариации не рассчитывались. В лучшем случае ограничивались весьма несовершенной оценкой "по числу обусловленности". В результате этого с самыми различными техническими объектами неоднократно случались аварии и катастрофы.

Вот характерный пример: 14 февраля 2004 года в Москве обрушилась крыша аквапарка с водными аттракционами "Трансвааль". Погибло 27 человек, в том числе дети, 113 человек получили различные травмы. Тщательное расследование исключило версии террористического акта, плохого качества строительства или нарушения проектных решений, отступлений от проекта. Специалисты пришли к выводу, что единственной причиной катастрофы явилась ошибка проектирования.

Следует отметить, что здание аквапарка "Трансвааль" проектировалось как здание уникальное — т. е. в точности таких же зданий ранее не было, надежность и безопасность проектных решений нужно было тщательно просчитать. Проектирование и расчет здания аквапарка проводились в последние годы двадцатого века и были поручены известному и опытному архитектурному бюро ЗАО "К", возглавляемому авторитетным и заслуженным архитектором Нодаром Вахтанговичем Канчели, руководившим до того проектированием многих уникальных красивых зданий Москвы.

Здание аквапарка "Трансвааль" оказалось очень красивым, но когда оно, простояв всего около двух лет, рухнуло и погубило много людей, то Н. В. Канчели было предъявлено обвинение сразу по двум грозным статьям Уголовного кодекса РФ, грозящим многолетним тюремным заключением: статья 109, часть 3 ("Причинение смерти по неосторожности двум и более лицам") и статья 118, часть 2 ("Причинение тяжкого вреда здоровью вследствие ненадлежащего исполнения профессиональных обязанностей"). Московская прокуратура возбудила против Н. В. Канчели уголовное дело, которое тянулось долго, но все же до конца, до полного и исчерпывающего выяснения причин катастрофы, его не довели, и 5 августа 2006 года дело прекратили в связи с амнистией, объявленной тогда Государственной Думой, — Н. В. Канчели попал под амнистию как гражданин, достигший пенсионного возраста. Следует помнить, что амнистия — это не реабилитация. Прекращение уголовного дела так и оставило невыясненным — был ли Н. В. Канчели виновен, были или не были допущены ошибки при расчете и проектировании, и в чем они состояли.

Ошибки и неточности при расчете и проектировании отнюдь не являются редкостью. По заключению одного из научно-исследовательских институтов (ЦНИИ ПСК им. Н. П Мельникова; заключение составлялось совместно с московским "Городским центром экспертиз" и опубликовано в 2006 г.), 9,8% обрушений зданий, произошедших за последние годы, произошли из-за неточностей и ошибок проектирования и расчета.

Характерный пример — серия аварий зданий и сооружений в первом квартале 2006 года:

- 3 января 2006 года в Германии в земле Бавария рухнула крыша катка. Погибло 11 человек;
- 27 января 2006 года в Польше в городе Катовице обрушилась крыша универмага. Погибло 67 человек;
- 23 февраля 2006 года в Москве обрушилась крыша Басманного рынка;
- 20 марта 2006 года в Ярославле рухнула крыша недостроенного торгового центра;
- 31 марта 2006 года рухнула крыша незадолго до этого построенного катка "Охта-парк" во Всеволожском районе Ленинградской области.

В качестве главной и основной причины всех этих аварий выдвигалось скопление снега на крышах. Под его тяжестью крыши, якобы, и рухнули. Да, зимой 2005–2006 года снега выпало больше, чем в предыдущие годы. Но "больше" совсем не означает "катастрофически больше". Крыши зданий и сооружений в местностях, где выпадает снег, рассчитываются на экстремальные снеговые нагрузки, на такое количество снега, которое выпадает не чаще, чем один раз в 30–40 лет. А такого необычно большого, чрезвычайного выпадения снега в 2006 году в Европе не было. Снега выпало много, но заведомо в пределах расчетных снеговых нагрузок, тех нагрузок, на которые здания были рассчитаны. А поскольку здания рухнули, то это означает, что расчеты запасов прочности и устойчивости при предстоящих в ходе эксплуатации вариациях нагрузок оказались не верны.

Не исключается, что причины ошибок в расчетах заключались в недавно открытых недостатках (описаны в [10]) широко распространенных оценок по "числу обусловленности" или же в неучете существования "особых" объектов и описывающих их "особых" математических моделей, для которых традиционные методы расчета заведомо не могли дать верных результатов. "Особые" объекты и "особые" математические модели, как и порождающие их опасные свойства изменения корректности при эквивалентных преобразованиях, были подробно описаны лишь в [2; 10] (а ранее — и впервые — в статьях [12], [13]).

Мы подробно остановились на авариях зданий и сооружений только потому, что здесь все сравнительно просто и наглядно, да и скрыть обрушение здания трудно. Гораздо больше аварий происходит там, где используется более сложная техника, например, — в военной авиации. До последнего времени эта область была тщательно засекречена, и только в 2007 году Первый заместитель министра обороны Российской Федерации генерал-полковник Александр Колмаков своим решением раскрыл ряд фактов и цифр (опубликованы в газете "Красная звезда" № 207–208 от 10.11.2007). Оказалось, что за последние 10 лет военная авиация потеряла в авариях и катастрофах 280 воздушных судов, погибло 707 человек. Ежегодный ущерб от потери воздушных судов вместе с затратами на проведение расследований составляет 3 миллиарда рублей в год.

В авиации — как и во многих других областях техники — на первый план выходят не только проблемы оценки погрешностей решений для систем алгебраических уравнений, но и оценки тех же погрешностей для систем дифференциальных уравнений, поскольку математическими моделями очень многих объектов и систем в авиации, в автоматике и других областях техники являются, прежде всего, системы дифференциальных уравнений различных порядков. Для дифференциальных уравнений на первый план выходят проблемы устойчивости решений и сохранения устойчивости при вариациях параметров. Поскольку эти проблемы ранее исследовались, и результаты исследований опубликованы в работах [1; 2; 7; 8; 10; 11; 12; 13] и многих других, ограничимся анализом одного важного вопроса — предвзятости очень многих заключений о причинах аварий и катастроф. В этих "заключениях" вина за аварию чаще всего перекладывается на "человеческий фактор", на ошибки пилотов и операторов, и остается в тени то, что многие аварии и катастрофы имеют причиной несовершенство методов расчета и проектирования, и поэтому аварии и катастрофы (в том числе и с гибелью людей) будут продолжаться до тех пор, пока не начнут использоваться усовершенствованные методы — в том числе и описанные в данной работе.

На первый взгляд кажется, что расследование причин катастроф — особенно в области гражданской авиации — производится тщательно и скрупулезно. Существует специальный Межгосударственный авиационный комитет (МАК), расследующий аварии и катастрофы, с многочисленным штатом специалистов. Каждый самолет имеет два хорошо защищенных бортовых самописца (их часто называют "черными ящиками"), в одном из которых записываются показания приборов, регистрирующих работу различных самолетных систем, а в другом — записываются все переговоры в кабине пилотов, как между собой, так и с диспетчером на земле. Показания бортовых самописцев чаще всего сохраняются, и их обычно удается расшифровать даже после гибели и разрушения самолета. Казалось бы, что имеются все условия для выяснения истинных причин аварий и катастроф.

Однако практика расследований показывает, что истинные причины катастроф в официальных заключениях МАК называются очень редко, почти никогда. Чаще всего в качестве причины называется "человеческий фактор", т. е. ошибки пилотов — особенно если пилоты погибли и возразить не могут.

Дело в том, что в состав Межгосударственного авиационного комитета входят представители могущественных и богатых авиакомпаний, корпораций по проектированию и изготовлению самолетов и авиационной техники. Этим компаниям невыгодно признавать свои ошибки, ошибки при проектировании и изготовлении авиатехники, и поэтому они оказывают давление на МАК, стараясь снять вину с себя и переложить на пилотов. Завязывается интереснейшая — напоминающая детектив — борьба между теми, кто старается раскрыть истинные причины катастрофы, и теми, кто хочет эти причины скрыть (разумеется, такая же борьба происходит и при расследовании аварий и катастроф в строительстве, при обрушениях зданий, при катастрофах в других областях техники).

Перипетии этой захватывающей борьбы изложены, например, в публикации [19], имеющей подзаголовок "научный детектив". Об обстоятельствах расследования ряда знаменитых катастроф: самолета Ту-144 в 1973 году, катастрофы аэробуса А-310 над Междуреченском в 1994 году, погубившей всех пассажиров и экипаж, катастрофы самолета Ту-154 над Донецком в 2006 году, в которой погибли 170 человек, и о ряде других катастроф рассказано также в [4], стр. 27–29, в [10], стр. 88–92 и 139–148.

Один из характерных примеров — расследование катастрофы, произошедшей на аэродроме Иркутска 9 июля 2006 года, когда при посадке разбился и сгорел аэробус А-310 Сибирских авиалиний. Тогда погибло 124 человека — весь экипаж и большая часть пассажиров. Уцелело лишь 20 пассажиров, сидевших в самом хвосте самолета. Они единодушно свидетельствовали, что самолет садился нормально и уже коснулся колесами бетонной посадочной полосы аэродрома. Обычно после касания колес с землей летчик дает команду на реверс тяги двигателей, и самолет нормально тормозится. Однако пассажиры почувствовали — и потом, после катастрофы рассказали об этом — что вместо обычного торможения самолет стал разгоняться, понесся по земле на большой скорости, докатился до ограды аэродрома, пробил ее, врезался в гаражи, стоящие сразу за аэродромной оградой, и загорелся, погубив 124 жизни.

Межгосударственный авиационный комитет (МАК), расследовавший катастрофу, в своем официальном заключении возложил вину на пилотов: якобы они "перепутали рычаги" и, вместо нажатия на "тормоз", нажали на "полный газ". В публикации [19] были проанализированы несообразности и противоречия, присутствовавшие в заключении МАК и заставляющие усомниться в его беспристрастности. Эта публикация, или другие причины, побудили Следственный комитет Российской Федерации провести самостоятельное расследование, показавшее, что в катастрофе виноваты не летчики, а неисправности в системах управления. Летчики подавали правильные команды, компьютерная система управления выполняла команды противоположные. Таким образом, в катастрофе виновны не российские летчики, а франко-германская авиастроительная компания "Airbus" с центром в Тулузе, проектировавшая и изготавливавшая самолет и не обеспечившая должных запасов устойчивости и надежной работы самолетных систем. Не исключено, что все это произошло из-за недостатков использованных методов расчета, тем более, что подобные аварии, с выкатыванием за пределы взлетно-посадочной полосы с самолетами А-310 и А-320 происходили неоднократно. Так, почти точно через год после катастрофы в Иркутске, 18 июля 2007 года самолет А-320, принадлежавший бразильской авиакомпании, при посадке на аэродром города Сан-Паулу не смог затормозить, выкатился за пределы аэродрома, налетел на здание авиакомпании, загорелся и погубил более 200 человек.

Выводы расследования Следственным комитетом Российской Федерации катастрофы в Иркутске опубликованы в газете "Известия" от 11 декабря 2008 года. Там же обращено внимание на то, что "председатель МАК Татьяна Анодина является председателем Международного союза авиапромышленников" и поэтому "для нее любой самолет — священная корова" и беспристрастных заключений от нее ждать не приходится. Впрочем, об этом говорилось (и писалось) ранее. Так, еще 14 июля 2006 года Совет Федерации "рекомендовал Правительству передать функции расследованиями авиационных происшествий независимому органу" ("Известия" от 08.08.2006), но до конца 2008 года ничего сделано не было.

Общий вывод: аварий и катастроф, происходящих из-за несовершенства методов проектирования и расчета, происходит гораздо больше, чем это официально признается. Если мы хотим сократить число аварий и катастроф, хотим спасти жизни людей, погибающих в этих катастрофах, то нужно с особенным вниманием отнестись к совершенствованию методов проектирования и расчета, а главное — использовать (и использовать без длительной "раскачки") эти усовершенствованные методы, некоторые из которых изложены в данной работе.

§ 24. Краткий обзор методов и результатов "Математики-2"

Как уже указывалось, в области вычислительной математики основные усилия исследователей были направлены на разработку методов вычислений с точно заданными числами, методов решения уравнений с точно заданными коэффициентами (или известными законами их изменения в случае переменных коэффициентов). Исследовались погрешности, которые (при точно заданных коэффициентах) происходили из-за несовершенства вычислительных методов: из-за конечности числа цифр, которые можно использовать при вычислениях и возникающих по этой причине погрешностей округления, из-за неизбежной конечности числа итераций в итерационных методах — и т. д. Все эти исследования можно условно отнести к "Математике-1".

К направлению исследований, которые можно условно назвать относящимися к "Математике-2", следует относить исследования, в которых с самого начала предполагается возможность неточного знания чисел, участвующих в вычислениях, неточного знания или непредсказуемого изменения коэффициентов уравнений и т. п.

По размаху и интенсивности исследований, по достигнутым результатам "Математика-2" далеко уступает "Математике-1". Однако и в "Математике-2" получены многие интересные результаты.

Это — в частности — правила приближенных вычислений, правила арифметических действий — сложения, вычитания, умножения и деления — с числами, имеющими различное (но всегда конечное) число цифр. Правила приближенных вычислений были разработаны очень давно и именно с них — с анализа погрешностей, возникающих при арифметических операциях с приближенными числами, — часто начинаются современные курсы вычислительной математики (см., например, [26], [27]).

Значительно более сложной является оценка погрешностей и вариаций решений систем линейных алгебраических уравнений.

Она осуществлялась, в частности, в рамках так называемого "интервального анализа" или "интервальной математики", очень интересного научного направления, которое интенсивно развивается. Первые работы по интервальному анализу появились в шестидесятых годах XX века: в 1966 году вышла первая монография (Moore R. E., Interval analysis), а уже в библиографии книги [31], вышедшей на языке оригинала в 1983 году, содержалось 540 наименований публикаций, а в книге [15], вышедшей в 1986 г., библиография содержала 740 наименований, посвященных, в большинстве своем, интервальному анализу.

Одной из областей применения интервальных методов стала линейная алгебра и, в частности, теория систем линейных алгебраических уравнений. Постановка задачи была совершенно правильной: из тех или иных соображений считались заданными интервалы, внутри которых находится каждый из коэффициентов a_{ij} и b_i (где $i = 1; 2; ...; n$) матрицы A и вектора B системы уравнений $AX = B$ (всего, таким образом, задавалось $n^2 + n$ интервалов), и требовалось найти интервалы, внутри которых заключены составляющие x_1; x_2; ...; x_n решения, составляющие вектора X в уравнении $AX = B$.

Эта задача решалась далее в наиболее общей постановке, когда интервалы, внутри которых заключены коэффициенты a_{ij} и b_i, не предполагались малыми. В такой общей постановке решение задачи оказывалось громоздким, труднодоступным и далеко не всегда выполнимым, тем более что поясняющих решение примеров было мало. Так, например, в большой монографии [15] приведен всего один пример конкретной системы, состоящей из двух уравнений, причем сам процесс отыскания x_1 и x_2 не показан, а приведен лишь конечный результат.

В работе [30] приведены два примера отыскания интервалов, внутри которых заключены составляющие вектора X, но уже для простой, приведенной в [30] системы:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 = b_1; \\ a_{21}x_1 + a_{22}x_2 = b_2, \end{aligned}\right\}, \tag{225}$$

где:

$$\left.\begin{array}{l} 2 \le a_{11} \le 4; -2 \le a_{12} \le 1; -2 \le b_1 \le 2; \\ -1 \le a_{21} \le 2; 2 \le a_{22} \le 4; -2 \le b_2 \le 2 \end{array}\right\} \qquad (226)$$

результаты вычисления сомнительны. В [30] утверждается, что при выполнении неравенств (226) x_1 и x_2 будут заключены в интервалах: $-\frac{1}{3} \le x_1 \le \frac{1}{3}$ и $-\frac{1}{3} \le x_2 \le \frac{1}{3}$.

Однако для системы:

$$\left.\begin{array}{l} 2x_1 + x_2 = -2; \\ 2x_1 + 2x_2 = 2, \end{array}\right\},$$

коэффициенты которой удовлетворяют неравенствам (226), имеем $x_1 = -3$; $x_2 = 4$, а это означает, что вектор X выходит далеко за пределы приведенного в [30] интервала.

Проблема вычисления интервалов, внутри которых находятся решения систем линейных алгебраических уравнений при произвольных, не обязательно малых, вариациях коэффициентов, в полном объеме не решена.

В данном учебном пособии рассматривается другая, более скромная, но зато более практичная задача: вариации составляющих вектора решений X системы уравнений $AX = B$ вычисляются при условии, что вариации коэффициентов a_{ij} матрицы A и коэффициентов b_i вектора B малы по сравнению с номинальными значениями коэффициентов. Это позволяет дать существенно более простое решение. Именно это условие малости вариаций коэффициентов характерно для подавляющего большинства встречающихся на практике задач. В то же время введение этого упрощающего условия, учитывающего практические требования, позволило дать простое решение проблеме оценки вариаций всех составляющих вектора X, получить решение этой проблемы, вполне доступное для студентов, а также инженеров и исследователей, желающих повысить свою квалификацию.

К другому направлению исследований относятся работы по теории чувствительности [1; 7 и много других публикаций]. Основной предпосылкой в этих исследованиях является — подчеркнем это — выполнение гипотезы о непрерывной дифференцируемости по параметру траекторий движения. С опорой на эту гипотезу вычислялись оценки изменения траекторий при вариациях коэффициентов.

Для наиболее распространенных математических моделей в виде систем линейных дифференциальных уравнений с постоянными коэффициентами важнейшей проблемой являлась проверка сохранения устойчивости при малых изменениях коэффициентов и параметров. Об устойчивости судили по корням характеристического полинома системы:

$$a_n\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0. \tag{227}$$

Если все корни характеристического полинома (227) имели отрицательные вещественные части или — что то же самое — лежали на комплексной полуплоскости левее мнимой оси, то такой полином называли *гурвицевым*, а исследуемую систему считали устойчивой. Кроме того, поскольку координаты корней полинома (227) зависят, как известно, от его коэффициентов непрерывно, то считалось, что устойчивость сохранится, по крайней мере, при малых изменениях коэффициентов характеристического полинома или параметров исследуемого объекта.

Публикаций по этому направлению исследований было много, и среди них нужно особо отметить статью [18] молодого сотрудника факультета прикладной математики — процессов управления СПбГУ В. Л. Харитонова. Он впервые поставил вопрос о сохранении устойчивости не только при малых, но и при больших изменениях характеристического полинома, а точнее — о проверке гурвицевости интервального полинома — т. е. такого полинома (227), о котором известны лишь оценки интервалов, внутри которых находятся его коэффициенты a_i. До 1978 года считалось, что для решения этой задачи необходимо проверить гурвицевость 2^n граничных полиномов, что для больших n очень громоздко. В. Л. Харитонов показал, что объем вычислений может быть очень существенно сокращен. Статья [18] получила заслуженную известность и была продолжена другими работами на ту же тему.

Что касается вычисления решений систем дифференциальных уравнений — а вычислений этих решений часто становилось необходимым при анализе тех или иных вопросов, возникающих в ходе проектирования и расчета различных технических объектов, — то при этих вычислениях опирались на известную теорему о непрерывной зависимости решений систем в нормальной форме от параметров. Эта теорема приводится во всех достаточно полных учебниках по дифференциальным уравнениям, и считалось общепризнанным, что эта теорема гарантирует надежность результатов вычислений, по крайней мере, при малых вариациях параметров.

Мы привели очень краткий очерк тех исследований по "Математике-2", которые, казалось бы, гарантировали надежность технических расчетов.

Перелом произошел в последнее десятилетие XX века. Сперва в статьях [12; 13], а затем более подробно в книгах [2; 8; 10; 11] было показано, что на самом деле все гораздо сложней, что многие общепризнанные и казавшиеся надежными методы, используемые в технических расчетах и проектировании, на самом деле ничего не гарантируют, могут стать (и становятся!) причинами аварий и катастроф.

Об авариях и катастрофах — в том числе и тех, причиной которых стали неточности расчета — рассказано в [2], стр. 27–30 четвертого издания, в [10], стр. 139–148, в [19].

Все началось с обнаружения новых свойств у привычных, повсеместно используемых в расчетах эквивалентных (равносильных) преобразований — таких, как умножение всех членов уравнения на число, не равное нулю, замена любого члена на член, равный ему, и т. п. Оказалось, что, не изменяя самих решений как таковых, эти преобразования могут изменять многие важные свойства решений — такие, как корректность, обусловленность и др.

Поскольку эквивалентные преобразования используются в расчетах очень широко, повсеместно, то открытие их новых свойств позволило быстро установить, что многие методы расчета, ранее считавшиеся безукоризненными, требуют совершенствования и поправок. Обнаружилось, что существуют "особые" объекты и соответствующие им "особые" математические модели, для которых традиционные методы расчета дают заведомо неверные результаты, и поэтому каждая неожиданная встреча с "особым" объектом может стать причиной аварии и катастрофы. К счастью для человека, "особые" объекты встречаются не очень часто, поэтому и катастрофы не происходят на каждом шагу, однако были обнаружены (и опубликованы в [2]) многочисленные примеры таких объектов, у которых все корни характеристического полинома лежат в левой полуплоскости комплексного переменного далеко от мнимой оси, выполнены все условия теоремы Харитонова, построена функция Ляпунова — и, тем не менее, их устойчивость теряется при сколь угодно малых, неизбежных на практике вариациях параметров (несколько позже будет приведен пример "особого" объекта).

Были обнаружены серьезные ошибки во многих популярных пакетах прикладных программ — таких, как MATLAB, Mathcad и многих других (опубликовано в [2], стр. 155–170 и в [23]).

Вместе с тем действенность традиционных методов можно возродить, если дополнить их описанной в [2], [10] методикой выявления опасных "особых" объектов и преобразования их в объекты безопасные.

Использование усовершенствованных методов и алгоритмов, описанных в [2], [8], [10], позволяет значительно уменьшить вероятность аварий и катастроф.

Однако публикация результатов этих исследований пришлась на трудные для российской науки годы, когда резкое сокращение тиражей научных журналов и книг нарушило связи между учеными, между инженерами (небольшой пример: тираж журнала "Автоматика и телемеханика" в 1977 году составлял 7 тыс. экз., а в 1994–2008 годах не превышал 400–500 экз.; тираж журнала "Известия ВУЗ. Электромеханика" в 1973 году — 3000 экз., в 1996 г. — 273 экз.). Новые научные результаты, даже опубликованные, но опубликованные очень малыми тиражами, остаются практически мало кому известными и мало используются. Это как раз и произошло со статьями [12; 13] и книгами [2; 8; 10; 11], несмотря на поддержку опубликованных в них результатов со стороны Российской академии наук.

Будем надеяться, что в дальнейшем эти результаты постепенно все же получат известность, и это поможет сократить число аварий и катастроф.

В данном учебном пособии результаты, полученные ранее в [2]; [8] и относящиеся, в основном, к объектам, описываемым дифференциальными уравнениями, распространены на объекты, расчет которых сводится к решению систем линейных алгебраических уравнений.

Приведем пример "особого" объекта и "особой" математической модели, для которой традиционные методы расчета приводят к ошибочным выводам и рекомендациям.

Пример № 24 — пример "особого" объекта и "особой" математической модели.

Переходные процессы в одном из электроприводов, снабженном системой регулирования частоты вращения, описываются системой двух дифференциальных уравнений (где $D = \frac{d}{dt}$):

$$[mD^3 + (2+2m)D^2 + (4+m)D + 3]x_1 - (D^2 + 2D + 1)x_2 = 0; \quad (228)$$

$$(D^2 + 4D + 5)x_1 - (D+1)x_2 = 0. \quad (229)$$

Уравнение (228) — уравнение объекта управления, электропривода, уравнение (229) — уравнение регулятора. Переменная x_1 — это частота вращения электропривода в относительных единицах, x_2 — ток якоря, играющий роль управления. Коэффициент m — это механическая постоянная времени электропривода. Ее номинальное значение равно единице, но в ходе эксплуатации она может испытывать малые вариации.

Рассмотрим сперва процессы в электроприводе при $m = 1$, когда они описываются системой состоящей из уравнения (229) и уравнения:

$$(D^3 + 4D^2 + 5D + 3)x_1 - (D^2 + 2D + 1)x_2 = 0, \quad (230)$$

в которое переходит уравнение (228) при $m = 1$ (более подробно регулируемый электропривод и описывающие его системы уравнений описаны в [2], стр. 64–66, 96–98, в [10], стр. 81–88).

Характеристический полином системы (229)–(230) равен определителю:

$$\begin{vmatrix} \lambda^3 + 4\lambda^2 + 5\lambda + 3 & -(\lambda^2 + 2\lambda + 1) \\ \lambda^2 + 4\lambda + 5 & -(\lambda + 1) \end{vmatrix} = \qquad (231)$$

$$= \lambda^3 + 5\lambda^2 + 7\lambda + 3 = (\lambda + 3)(\lambda + 1)^2,$$

который имеет корни: $\lambda_1 = -3$; $\lambda_2 = \lambda_3 = -1$.

Поскольку все корни лежат в левой полуплоскости комплексного переменного, то в согласии с общепринятыми методами расчета устойчивости необходимо сделать заключение: система уравнений (229)–(230) устойчива и сохранит устойчивость, по крайней мере, при малых вариациях своих коэффициентов. Этот же вывод будет подтвержден при использовании других известных методов расчета устойчивости: с использованием логарифмических амплитудно-частотных характеристик, годографа Найквиста и т. п.

Если — следуя методике В. Л. Харитонова, описанной в [18], — исследовать гурвицевость характеристического полинома (231), то нетрудно убедиться в том, что он остается гурвицевым даже при достаточно больших вариациях своих коэффициентов. Таким образом, все традиционные методики проверки устойчивости по характеристическому полиному говорят об устойчивости системы (229)–(230) и о сохранении этой устойчивости при малых вариациях любых коэффициентов полинома, т. е. говорят о том, что система (229)–(230) параметрически устойчива.

На самом деле все обстоит как раз наоборот. Если исследовать устойчивость системы (228)–(229) при вариациях параметра m — механической постоянной времени — то характеристический полином системы (228)–(229) окажется равным определителю:

$$\begin{vmatrix} m\lambda^3 + (2 + 2m)\lambda^2 + (4 + m)\lambda + 3 & -(\lambda^2 + 2\lambda + 1) \\ \lambda^2 + 4\lambda + 5 & -(\lambda + 1) \end{vmatrix} = \qquad (232)$$

$$= [(1 - m)\lambda^2 + (2 - m)\lambda + 3](\lambda + 1)^2.$$

При $m = 1$ определители (231) и (232) совпадают, и решение x_1 системы (128)–(129) имеет вид:

$$x_1(t) = C_1 e^{-3t} + (C_2 t + C_3) e^{-t}, \tag{233}$$

где C_1; C_2; C_3 — постоянные интегрирования.

Однако если $m = 1 + \varepsilon$, где $\varepsilon > 0$ — сколь угодно малое число, то в решении $x_1(t)$ появляется стремительно растущий четвертый член, приближенно (при малых ε) равный:

$$C_4 e^{+\frac{t}{\varepsilon}} \tag{234}$$

и растущий тем быстрее, чем меньше ε.

Таким образом, система (228)–(229) является "особой" и может терять устойчивость при сколь угодно малых (и, значит, неизбежных на практике) вариациях своих коэффициентов и при сколь угодно малых вариациях параметров исследуемого объекта, хотя все традиционные методы проверки устойчивости говорят об обратном.

Отметим также, что при $\varepsilon < 0$ устойчивость сохранится, поскольку четвертый член в решении $x_1(t)$ будет экспоненциально убывающим.

Заметим теперь, что решения системы (228)–(229) не имеют непрерывной зависимости от параметра m. В точке $m = 1$ значение $x_1(t = t_0)$ для любого $t_0 > 0$ терпит разрыв. Дифференцируемости решения при $m = 1$ нет, и методы теории чувствительности не могут быть применены.

Отметим еще, что для системы (228)–(229) — как для системы линейных уравнений с постоянными коэффициентами — существует и может быть построена функция Ляпунова. Но и существование функции Ляпунова не гарантирует от потери устойчивости при сколь угодно малых вариациях параметра m. Таким образом, и второй метод Ляпунова — как и все остальные методы теории устойчивости — даст неверный ответ о поведении "особых" систем.

Все это связано с недавно открытыми новыми свойствами эквивалентных преобразований. Если с их помощью системы (228)–

(229) привести к нормальной форме, к форме системы из четырех уравнений первого порядка, то решения такой системы будут непрерывно зависеть от параметра m в точке $m = 1$, будут сохранять устойчивость при малых (и даже довольно больших) вариациях m.

Однако физическому смыслу (как это подробно показано в [2] и в [10]) соответствует именно система (228)–(229), и поведение реального объекта (электропривода) будет ей соответствовать: устойчивость объекта действительно может теряться при малых вариациях параметра m, при $m = 1 + \varepsilon$, где $\varepsilon > 0$.

Поскольку при изготовлении реального электропривода его параметры (в том числе и параметр m) будут неизбежно отклоняться от расчетных значений, и знак этих отклонений непредсказуем, то у изготовленного электродвигателя может оказаться и $\varepsilon < 0$, и $\varepsilon > 0$. Если $\varepsilon > 0$, то это еще половина беды: на испытаниях электропривод окажется неустойчивым и будет забракован. Все будет гораздо страшней, если окажется, что $\varepsilon < 0$. В этом случае электропривод будет устойчиво работать, может успешно пройти испытания и оказаться установленным на ответственный объект (например — на самолет). Запас его устойчивости будет очень малым (он будет целиком зависеть от малой погрешности изготовления), и в ходе эксплуатации этот запас устойчивости может оказаться исчерпанным за заранее неизвестное время. Поэтому в совершенно неожиданный момент времени электропривод может потерять устойчивость, "пойти в разнос" и создать аварийную ситуацию, которая может перерасти в аварию и даже катастрофу.

В публикациях [2], [10], [19] рассказано об авариях и катастрофах, происходящих по подобным причинам.

Вместе с тем, надо особо подчеркнуть, что традиционные методы расчета (теория чувствительности, изложенная в [1], методика Харитонова и т. п.) не будут приводить к подобным результатам, если их дополнить анализом исследуемых систем на "особость" и отсеивать опасные "особые" системы. Методы проверки на "особость", методы выделения "особых" систем описаны в [2], в [10], и они помогут избежать многих аварий и катастроф.

Отметим еще раз важное обстоятельство: когда происходит авария или катастрофа, и, особенно, — крупная, с большим числом жертв, то при публикации расследования ее причин, как уже говорилось, чаще всего называют "человеческий фактор", ошибки пилотов или ошибки архитекторов. Но это обусловлено лишь тем большим давлением, которое оказывают крупные компании, авиационные или строительные, на тех, кто производит расследование. Крупным компаниям очень не хочется признавать свои ошибки, свои погрешности в проектировании, в изготовлении и проверке своей продукции, и они стараются переложить вину на других. Часто проходит много лет до того момента, когда все же вскрывается истина, вскрываются истинные причины катастроф. И очень часто оказывается, что истинной причиной служат именно ошибки расчета и проектирования. Они гораздо чаще оказываются причиной аварий и катастроф, чем это кажется по публикуемым итогам пристрастно проводившихся расследований.

Вот дополнительные характерные примеры: 03.06.1973 года во время демонстрационного полета во Франции, на авиационной выставке в Ле-Бурже разбился сверхзвуковой пассажирский самолет Ту-144 — гордость советской авиации. Долгие годы основной причиной называли попадание Ту-144 в "спутный вихрь" от ошибочно пролетевшего слишком близко французского самолета. И только через 34 года специалисты, работавшие в 1973 году в конструкторском бюро Туполева, раскрыли истинную причину — сбой в одной из систем автоматики, которые в те годы — а во многом и до сих пор — рассчитываются по методике, не обеспечивающей верной оценки запасов устойчивости для случая, когда проектируемая система оказалась "особой" (более подробно об этой катастрофе и ее причинах — в [10], на стр. 92).

Именно так же причиной страшной катастрофы самолета А-310 22.03.1994 года над Междуреченском, когда погибли все пассажиры и экипаж, первоначально были названы действия пилота. Лишь после того, как в [2] на стр. 28–29 был опубликован подробный анализ причин аварии, ошибка первоначального заключения была признана. Истинной причиной оказалась потеря устойчивости в автопилоте, запас устойчивости которого был явно меньше расчетного, поскольку расчет велся по методике, не учитывающей существования "особых" систем, которые в те годы, когда проектировался самолет А-310, еще не были известны (подробнее — в [19], стр. 100 и [10], стр. 90–92).

Что касается известной катастрофы самолета Ту-154 над Донецком 22.08.2006 года, когда погибло 170 человек — экипаж и все пассажиры — то до сих пор в официальных "заключениях" вина возлагается на командира воздушного судна, хотя на самом деле истинной причиной этой, как и других подобных катастроф самолетов Ту-154, были ошибки их расчета и проектирования (более подробно об этой катастрофе и ее причинах — в [10], стр. 141–143).

Многие другие катастрофы рассмотрены в [19], и часто их причинами являлись погрешности в недостаточно совершенных методиках проектирования и расчета.

§ 25. Дополнительные пояснения и примеры

Завершив изложение всех аспектов вычисления вариаций решения системы уравнений, зависящих от вариаций коэффициентов, сформулируем теперь наиболее компактно алгоритм вычисления и поясним его дополнительными примерами.

Напомним еще раз, что мы рассматриваем вариации (погрешности) решения, порожденные именно вариациями коэффициентов системы уравнений, а не погрешностями или несовершенствами процесса вычисления решений (конечность числа итераций, погрешности округления и т. п.), которые можно уменьшить за счет усовершенствования вычислительного процесса. Именно такие погрешности, в принципе допускающие уменьшение до приемлемо малой величины, рассматриваются чаще всего в курсах вычислительной математики.

Данная работа посвящена другой теме — точной оценке вариаций (погрешностей) решений, которые невозможно уменьшить, а можно только вычислить. Если испытывают вариации коэффициенты уравнения, то неизбежна и вариация решения, которая полностью определяется величинами вариаций коэффициентов и их знаками и поэтому уменьшена быть не может. Таким образом, точная оценка вариаций решений является важнейшим этапом обеспечения надежности результата решения рассматриваемой технической задачи или физической проблемы.

До последнего времени использовали весьма приближенную оценку погрешности решения через "числа обусловленности", т. е. через произведения норм прямой и обратной матриц коэффициентов. Сейчас есть возможность вычислить точные оценки. Автор надеется, что наличие точных оценок позволит повысить надежность расчетов, уменьшить число аварий и катастроф, порожденных ошибками в расчетах, о которых говорилось в §§ 23–24.

При этом нужно учитывать, что алгоритм вычисления точного значения вариации решения системы успешно работает в том случае, если известны оценки наибольших возможных абсолютных значений вариаций коэффициентов. Мы опускаем этот этап вычислений только потому, что оценка наибольших возможных вариаций коэффициентов может быть выполнена лишь для конкретного объекта и требует хорошего знания специфики этого объекта. Единого рецепта вычисления здесь дать нельзя. Наоборот, алгоритм второго этапа оценки погрешности — т. е. этапа вычисления наибольшей возможной вариации решения при известных вариациях коэффициентов — является универсальным, он может быть применен к любой системе уравнений и поэтому полезно знать этот алгоритм, составить для него программу вычислений и пользоваться ею.

Помимо алгоритма, описанного в § 18, можно использовать алгоритм, основанный на предварительном построении полной "таблицы знаков" для вариации любой из составляющих x_i решения x_1; x_2; ...; x_n — т. е. "таблице знаков" и для числителя, и для знаменателя формул Крамера:

$$x_i = \frac{D_i}{D}.$$

Пример № 25. Рассмотрим простую систему уравнений:

$$\left.\begin{aligned} 2x_1 + x_2 &= 2; \\ x_1 + x_2 &= 1, \end{aligned}\right\}, \tag{235}$$

для которой:

$$D = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1, \tag{236}$$

$$D_1 = \begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} = 1, \tag{237}$$

$$x_1 = \frac{D_1}{D} = \frac{1}{1} = 1; x_2 = 0. \tag{238}$$

Составим "таблицы знаков" для определителей D и D_1. Для определителя D "таблицей знаков" будет:

$$\begin{vmatrix} + & - \\ - & + \end{vmatrix}, \tag{239}$$

"обратной таблицей знаков" будет "таблица":

$$\begin{vmatrix} - & + \\ + & - \end{vmatrix}. \tag{240}$$

Для определителя D_1 "прямая" и "обратная" "таблицы знаков" совпадают с "таблицами" (239) и (240).

Вычислим теперь "таблицу знаков" для приращения (вариации в направлении возрастания) составляющей вектора решения x_1.

Для первого варианта расчета, когда за основу берутся вариации элементов определителя D, "таблица знаков" для x_1 имеет вид:

$$x_1 \to \frac{\begin{vmatrix} + & + \\ - & - \end{vmatrix}}{\begin{vmatrix} - & + \\ + & - \end{vmatrix}} \tag{241}$$

(приращение x_1 будет наибольшим, если определитель D будет возможно меньше, а определитель D_1 возможно больше; поэтому к максимуму приращения приведет "обратная таблица знаков" для определителя D — т. е. "таблица" (240); к максимальному приращению привела бы "прямая таблица знаков" определителя D_1 — т. е. "таблица" (239), но из нее в "таблицу знаков" для x_1 может войти лишь первый столбец, а второй столбец должен совпадать со вторым столбцом "таблицы знаков" определителя D — отсюда и следует таблица (241)).

Таким образом, если $|\varepsilon_{ij}| \le 0{,}01$ и $|b_i| \le 0{,}01$, то:

$$x_1 = \frac{\begin{vmatrix} 2,02 & 1,01 \\ 0,99 & 0,99 \end{vmatrix}}{\begin{vmatrix} 1,98 & 1,01 \\ 1,01 & 0,99 \end{vmatrix}} = 1,0635. \tag{242}$$

Для второго варианта, когда за основу берутся вариации определителя D_1 и его "таблица знаков" (239), для вычисления наибольшего значения x_1 следует использовать "таблицу знаков":

$$x_1 \to \frac{\begin{vmatrix} + & - \\ - & + \end{vmatrix}}{\begin{vmatrix} - & - \\ + & + \end{vmatrix}} \qquad (243)$$

и тогда, при тех же условиях $|\varepsilon_{ij}| \leq 0{,}01$ и $|b_i| \leq 0{,}01$ будет:

$$x_1 = \frac{\begin{vmatrix} 2{,}02 & 0{,}99 \\ 0{,}99 & 1{,}01 \end{vmatrix}}{\begin{vmatrix} 1{,}98 & 0{,}99 \\ 1{,}01 & 1{,}01 \end{vmatrix}} = 1{,}0401. \qquad (244)$$

В данном случае к наибольшему приращению x_1, к наибольшей ее вариации в направлении возрастания приводит первый вариант и "таблица знаков" (241).

Предварительное составление полных "таблиц знаков" для x_i — "таблиц" и для числителя, и для знаменателя — немного упрощает алгоритм и облегчает составление программы вычислений.

Следует иметь в виду, что если испытывают вариации только коэффициенты правой части системы уравнений $AX = B$, то отыскание вариаций любой из составляющих x_1; x_2; ...; x_n вектора X упрощается очень существенно. Об этом уже говорилось в § 8, где приводился алгоритм вычисления вариации x_i с использованием формулы (81) для x_1 и аналогичных формул для любого x_i. Отметим, что в данном случае вариация x_i будет зависеть от вариации $\delta_i b_i$ коэффициентов правой части системы $AX = B$ линейно, и поэтому нет необходимости отдельно вычислять вариацию x_i в направлении возрастания и в направлении убывания — как это оказалось необходимым при расчете вариаций x_i, происходящих от вариаций коэффициентов матрицы A системы $AX = B$. Формула (81) сразу дает точное значение вариации x_i; заметим, что с помощью алгебраических дополнений формула (81) может быть записана в более простом виде:

$$\Delta x_1 = \frac{1}{D}\sum_{i=1}^{n} \delta_i b_i A_i, \qquad (245)$$

где D — определитель матрицы A, b_i — коэффициенты вектора B правой части, A_i — алгебраические дополнения элементов b_i в определителе D_1 в формуле Крамера:

$$x_1 = \frac{D_1}{D}.$$

Из формулы (245) следует, что в данном случае достаточно вычислить n алгебраических дополнений, что требует примерно n^4 умножений, что существенно меньше, чем в общем случае, когда необходимо учитывать вариации элементов и матрицы A, и вектора B.

Для общего случая объем вычислений, необходимых для точной оценки погрешностей, велик и поэтому необходима разработка программы, реализующей алгоритмы, описанные в §§ 18 и 19. Такая программа была разработана В. В. Жегаловой. Расчет по этой программе производится достаточно быстро. Так, для примера, расчет "таблицы знаков" для системы, состоящей из ста уравнений, т. е. для матрицы A, имеющей размер 100×100, вычисления заняли семь секунд.

§ 26. Заключение

В данной работе изложены новые результаты, полученные автором в области оценки погрешностей решений систем алгебраических уравнений. Эти результаты автор стремился изложить наиболее простым и доступным образом — изложить на уровне учебного пособия, доступного для студентов, а также инженеров и специалистов, желающих повысить уровень своей квалификации.

Автор надеется, что использование изложенных здесь результатов позволит повысить надежность расчета и проектирования технических объектов, снизить число аварий и катастроф.

Основной новый результат: методика вычисления *точной* оценки погрешности (вариации) каждой из составляющих x_1; x_2; ...; x_n вектора X систем уравнений вида $AX = B$ в зависимости от погрешностей (вариаций) коэффициентов матрицы A и вектора B. Этот результат изложен в § 18. Помимо этого основного результата (опирающегося на результаты, приведенные в §§ 14, 16, 17), изложен и ряд других положений:

- рассмотрены и проанализированы серьезные недостатки наиболее распространенного метода оценки надежности расчета по величине "числа обусловленности" (исследованы как ранее известные недостатки, так и обнаруженные автором);
- было обнаружено, что "число обусловленности" может давать неверную оценку влияния параметров системы на погрешность решений. "Числа обусловленности" обладают ложной зависимостью от некоторых эквивалентных преобразований и от единиц измерения коэффициентов уравнений, что приводит к неверным заключениям о погрешностях решений;
- некоторые из недостатков "чисел обусловленности" могут быть исправлены и устранены (описаны методы исправления), другие являются принципиально неустранимыми (особенно важно то, что "числа обусловленности" позволяют оценить лишь осредненную относительную норму погрешности всех

составляющих вектора X — от x_1 до x_n — в то время как на практике чаще всего необходима оценка погрешности того или иного конкретного x_i, и эта погрешность может очень сильно отличаться от оценки по относительной норме погрешности всех составляющих вектора X — от x_1 до x_n);

- предложен и описан новый метод "оценки сверху" погрешности решений, происходящих от вариаций коэффициентов системы уравнений — метод оценки через "модульные определители";
- если испытывают вариации только коэффициенты правой части системы уравнений $AX = B$ (коэффициенты b_i), то предложенный в § 8 метод позволяет дать простую и точную оценку вариаций любой составляющей x_i вектора X и оценку ее погрешности;
- если необходимо учитывать вариации коэффициентов матрицы A системы $AX = B$, то в этом случае точная оценка погрешностей x_i требует более сложных вычислений, описанных в §§ 18 и 19.

Все предлагаемые методы расчета пояснены многочисленными примерами.

Показано, что в системах алгебраических уравнений обусловленность решений и погрешности их могут существенно зависеть от эквивалентных преобразований, не изменяющих самих решений (ранее это было показано в [2; 10; 11] для систем дифференциальных уравнений). Это еще раз показывает, что "Математика-2" отличается от "Математики-1" не только предметом исследования, но и методами — и прежде всего значительно более осторожным использованием эквивалентных преобразований.

Автор еще раз напоминает, что готов выслушать замечания и пожелания по настоящему учебному пособию; их можно направлять на адрес: *petrov*1930@*mail.ru*.

Литература

1. Розенвассер Е. Н., Юсупов Р. М. Чувствительность систем управления. М.: Наука, 1981.
2. Петров Ю. П., Петров Л. Ю. Неожиданное в математике и его связь с авариями и катастрофами. Первое издание — 1999, 108 с., второе издание — 2000, третье издание — 2002, четвертое издание, расширенное и дополненное — 2005 г., издательство "БХВ-Петербург", 224 с.
3. Ильин В. А., Позняк Э. Г. Линейная алгебра. М.: Наука, 1978, 302 с.
4. Мышкис А. Д. Прикладная математика для инженеров. Специальные курсы. Издание третье. М.: Физматлит, 2007, 687 с.
5. Зализняк В. Е. Основы научных вычислений. Москва — Ижевск, 2006, 264 с.
6. Шноль Э. Э. Семь лекций по вычислительной математике. М.: URRS, 3-е издание, 2007, 106 с.
7. Никифоров В. О., Ушаков А. В. Управление в условиях неопределенности: чуствительность, адаптация, робастность. СПб.: издательство ИТМО, 2002, 232 с.
8. Петров Ю. П., Сизиков В. С. Корректные, некорректные и промежуточные задачи с приложениями. Учебное пособие для вузов. СПб.: издательство "Политехника", 2003, 261 с. (Перевод на английский язык выпущен в свет в 2005 году издательством "VSP", Бостон —Лейден).
9. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. Издание третье. М.: Наука, 1986, 287 с.
10. Петров Ю. П. Обеспечение достоверности и надежности компьютерных расчетов. Издательство "БХВ-Петербург", 2008, 160 с.
11. Петров Ю. П. Третий класс задач физики и техники — промежуточный между корректными и некорректными. СПб.: ООП СПбГУ, 1998, 30 с.

12. Петров Ю. П. О скрытых опасностях, содержащихся в традиционных методах проверки устойчивости. Известия ВУЗ. Электромеханика, 1991, № 11, стр. 106–108.

13. Петров Ю. П. Устойчивость линейных систем при вариациях параметров. Автоматика и телемеханика. 1994, № 11, стр. 186–189.

14. Феодосьев В. И. Сопротивление материалов. Учебник для вузов. Издание восьмое. М.: Наука, 1979, 559 с.

15. Калмыков С. А., Шокин Ю. И., Юлдашев З. Х. Методы интервального анализа. Новосибирск, Наука, 1986, 224 с.

16. Шокин Ю. И. Интервальный анализ. Новосибирск, Наука, 1981, 112 с.

17. Меньшиков Г. Г. Практические начала интервальных вычислений. Л.: издательство ЛГУ, 1992, 112 с.

18. Харитонов В. Л. Об асимптотической устойчивости положения равновесия семейства линейных дифференциальных уравнений. Дифференциальные уравнения, 1978, № 11.

19. Петров Ю. П. Расследование и предупреждение техногенных катастроф. Издательство "БХВ-Петербург", 2007, 104 с.

20. Петров Ю. П. Новые свойства систем дифференциальных уравнений и их связь с задачами строительной механики. Промышленное и гражданское строительство, 2005, № 11, стр. 51–52.

21. Орлов В. Б., Петров Ю. П. Расширение исследований А. М. Ляпунова на проблему параметрической устойчивости. Труды международного конгресса "Нелинейный динамический анализ-2007", СПб.: издательство СПбГУ, 2007, 402 с.

22. Шароватов В. Т. Обеспечение стабильности показателей качества автоматических систем. Л.: Энергоатомиздат, 1987, 176 с.

23. Шароватов В. Т., Петров Ю. П. "Об ошибках пакета MATLAB" и Чертков К. Г., Петров Ю. П. "Ошибки, обнаружившиеся в пакете MATLAB". Труды Второй Всероссийской конференции "Проектирование научных и инженерных при-

ложений в среде MATLAB". Институт проблем управления Российской академии наук, 2004, стр. 318–323 и 324–327.

24. Лизунова Н. А., Шкрова С. П. Матрицы и системы линейных уравнений. М.: Физматлит, 2007, 349 с.
25. Махмутов М. М. Лекции по численным методам. Институт компьютерных исследований. Москва — Ижевск, 2007, 236 с.
26. Демидович Б. П., Марон И. А. Основы вычислительной математики. Издание шестое. Издательство "Лань", 2007, 664 с.
27. Копченова Н. В., Марон И. А. Вычислительная математика в примерах и задачах. Издание второе. Издательство "Лань", 2008, 367 с.
28. Охорзин В. А. Прикладная математика в системе Mathcad. Издательство "Лань", 2008, 348 с.
29. Уоткинс Д. Основы матричных вычислений. Перевод с английского. М.: Изд. "Бином", 2006, 664 с.
30. Захаров А. В., Шокин Ю. И. Алгебраическое интервальное решение систем линейных интервальных уравнений $AX = B$. Вычислительный центр Сибирского отделения Академии наук СССР. Препринт № 5. Красноярск, 1987, 17 с.
31. Алефельд Г., Херцбергер Ю. Введение в интервальные вычисления. М.: "Мир", 1987, 356 с.
32. Краснов М. Л., Киселев А. И., Макаренко Г. И. Интегральные уравнения. Задачи и примеры с подробными решениями. Издание четвертое. М.: URSS, Ком. Книга. 2007, 190 с.
33. Верлань А. Ф., Сизиков В. С. Интегральные уравнения: методы, алгоритмы, программы. Киев, Наукова думка, 1986, 544 с.
34. Апарцин А. С. Неклассические уравнения Вольтерра первого рода; теория и численные методы. Новосибирск, Наука, 1999, 193 с.
35. Игнатьев М. Б. Активные методы обеспечения надежности алгоритмов и программ. СПб.: изд. "Политехника", 1993, 288 с.
36. Шарый С. П. Интервальный анализ или методы Монте-Карло? Расширенные тезисы докладов Всероссийского совещания по интервальному анализу "Интервал-06" 1–4 июля 2006 г., стр. 140–144.