ВВЕДЕНИЕ

Предмет вычислительной математики – методы доведения до числового результата основных задач математического анализа, алгебры и геометрии и пути использования для этой цели современных вычислительных средств.

МЕТОД ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ

Круг задач, с которыми приходится сталкиваться в вычислительной математике, очень широк. Разнообразны и методы, применяемые для решения этих задач. Однако можно заметить одну общую идею этих методов. Эта идея отчетливее всего выражается в терминах функционального анализа. Поэтому введем предварительно некоторые важнейшие понятия функционального анализа.

1. Функциональные метрические пространства

Основным предметом исследования в классическом математическом анализе является *числовая функция*.

С появлением понятия функции одной и нескольких переменных, функции точки в евклидовом пространстве начался современный этап развития математики. Начиная с работ Ньютона и Лейбница и до конца XIX века подавляющее большинство математических исследований так или иначе было связано с этим понятием.

Главным предметом изучения были числовые функции и их системы, заданные в n-мерной области, т.е. на некотором множестве n-мерного евклидова пространства.

XX век внес много нового в эту картину. Особо важную роль начинает играть понятие о *функциональном множестве*, *функциональных пространствах* и *функциональных операторах*, т.е. о функциях, аргументами которых также являются элементы функциональных пространств.

Вместо евклидовых пространств рассматриваются абстрактные пространства, элементы которых могут иметь самую различную природу.

Так, например, вводится понятие метрического пространства R как абстрактного множества, для любых двух элементов x и y которого определено понятие расстояния $\rho(x, y)$ удовлетворяющее следующим условиям:

1. $\rho(x, y) \geq 0$, причем $\rho(x, y) = 0$ тогда и только тогда, когда x совпадает с y.

2. $\rho(x, y) = \rho(y, x)$.

3. $\rho(x, y) \leq \rho(x, z) + \rho(x, y)$ для любых трех элементов x, y, z, принадлежащих R (аксиома треугольника).

Евклидовы пространства с обычным определением расстояния в них удовлетворяют всем этим условиям. Но могут быть и другие метрические пространства.

Рассмотрим два примера метрических пространств.

1. Метрическое пространство C. Рассмотрим множество всевозможных непрерывных функций, заданных на отрезке $[a, b]$. Для любых двух таких функций $x(t)$ и $y(t)$ определим расстояние $\rho(x, y)$ равенством

$$\rho(x,\ y) = \max_{t\in[a,b]} |x(t) - y(t)|.$$

Это определение расстояния удовлетворяет всем трем поставленным выше условиям.

2. Метрическое пространства L_p (Здесь p – действительное число ≥ 1). Расстояние $\rho(x, y)$ в L_p определяется следующим образом:

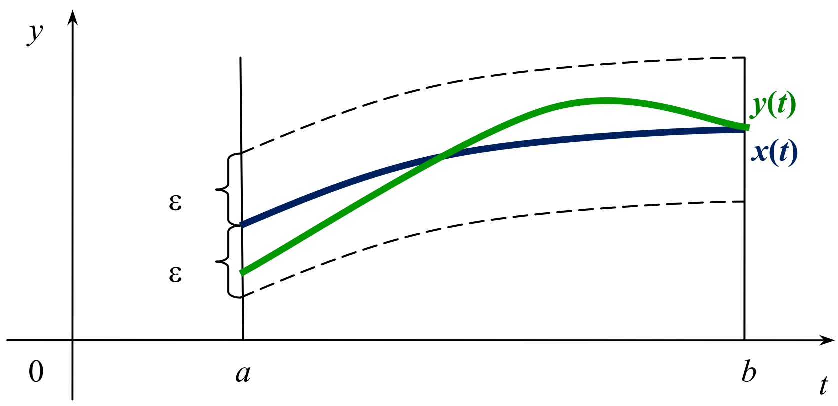
$$\rho(x, y) = \left[\int_a^b |x(t) - y(t)|^p dt\right]^{\frac{1}{p}}.$$

Определенное таким образом расстояние удовлетворяет трем поставленным выше условиям.

В каждом метрическом пространстве можно говорить об окрестности данной точки. Назовем ε-окрестностью точки x некоторого метрического пространства R совокупность его точек y, для которых выполняется неравенство

$$\rho(x, y) < \varepsilon.$$

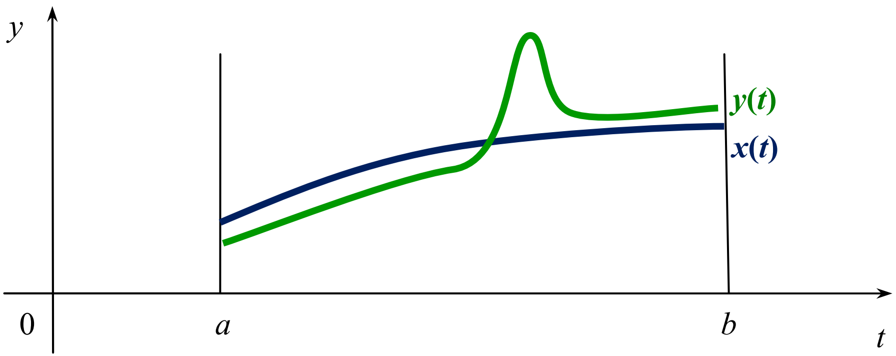
В пространстве C это будет совокупность всех непрерывных на $[a, b]$ функций, лежащих в полосе $x(t) \pm \varepsilon$.



В пространстве L_p это будет совокупность всех функций, принадлежащих L_p, для которых

$$\int_a^b |x(t) - y(t)|^p dt < \varepsilon^p.$$

При этом в отдельных точках отклонение $y(t)$ от $x(t)$ может быть очень большим, а зато в других точках будет очень малым (рис. 2).



В вычислительной математике часто приходится заменять одну функцию $x(t)$ другой функцией, более удобной для вычислительных целей и в каком-то смысле близкой к первой.

Обычно эту вторую функцию берут в некоторой ε-окрестности первой.

Если ε-окрестность берется в пространстве C, то говорят о *равномерном приближении функции* $x(t)$.

Если ε-окрестность берут в пространстве L_p, то говорят *о приближении в среднем*. В частности, при $p = 2$ говорят *о среднеквадратичном приближении*.

2. Функции, определенные на функциональных пространствах

Точно так же, как в классическом математическом анализе, можно ввести понятие функции, аргументом и значением которой будут элементы абстрактных пространств.

Пусть даны два абстрактных пространства R_1 и R_2. Пусть каждому элементу $x \in R_1$ поставлен в соответствие элемент $y \in R_2$. Тогда говорят, что задана функция

$$y = A(x)$$

с областью определения R_1 и областью значений, принадлежащей R_2.

Если R_2 является областью действительных или комплексных чисел, то $A(x)$ называется *функционалом*.

Простейшим примером функционала в пространстве C будет являться

$$I(x) = \int_a^b x(t)dt.$$

Пространство может совпадать с пространством R_1 и тогда $A(x)$ называется *оператором*.

Область математики, изучающая свойства функциональных пространств и заданных на них функций, и носит название *функционального анализа*.

3. Метод вычислительной математики

Большинство математических задач может быть записано в виде

$$y = A(x) \tag{1}$$

где x и y принадлежат заданным пространствам R_1 и R_2 и $A(x)$ – некоторая заданная функция.

Задача состоит

либо в отыскании y по заданному x,

либо в отыскании x по заданному y.

Далеко не всегда с помощью средств современной математики удается точно решить эти задачи, применяя конечное число шагов. В этих случаях и прибегают к вычислительной математике.

Иногда задача и может быть решена точно, но методы классической математики дают ответ после громоздких и трудоемких вычислений. Поэтому и задачи вычислительной математики входит также разработка приемов и методов наиболее рационального решения конкретных задач.

Основной метод, при помощи которого в вычислительной математике решают поставленные выше задачи:

замена пространств R_1 *и* R_2 *и функции* A *некоторыми другими пространствами* $\bar{R}_1$ *и* $\bar{R}_2$ *функцией* $\bar{A}$*, более удобными для вычислительных целей.*

Иногда бывает достаточно произвести замену пространств R_1 и R_2 или даже одного из них.

Иногда достаточно заменить только функцию A.

Замена должна быть сделана так, чтобы решение новой задачи

$$\bar{y} = \bar{A}(\bar{x}), \tag{2}$$

$\bar{x} \in \bar{R}_1$, $\bar{y} \in \bar{R}_2$ – было в каком-то смысле близким к точному решению исходной задачи (1) и его возможно было бы практически отыскать с сравнительно небольшими трудностями.

Например, пусть необходимо вычислить интеграл

$$y = \int_a^b f(x)dx.$$

где $f(x)$ непрерывная функция.

Чтобы получить достаточно точное приближенное значение интеграла, можно идти двумя путями.

1. Заменим функцию $f(x)$ алгебраическим многочленом $P(x)$ равномерно приближающим функцию $f(x)$ на отрезке $[a, b]$ с необходимой степенью точности.

Вместо интеграла

$$y = \int_a^b f(x)dx$$

будем находить интеграл

$$\bar{y} = \int_a^b P(x)dx$$

вычисление которого не составляет труда.

При этом, функционал $A(f) \equiv \int_a^b f(x)dx$ не меняется, заменяется лишь пространство C, которому принадлежит $f(x)$, пространством многочленов и вместо функции $f(x)$ рассматривается многочлен $P(x)$ из некоторой ее ε-окрестности.

2. Из определения интеграла $\int_a^b f(x)dx$ следует, что всегда можно построить интегральную сумму

$$\sum_{i=1}^{n} f(x_i)\Delta x_i,$$

которая будет достаточно близка к значению интеграла.

Следовательно, вместо вычисления интеграла

$$y = \int_a^b f(x)dx$$

можно решать другую задачу – задачу вычисления конечной суммы

$$y = \sum_{i=1}^{n} f(x_i)\Delta x_i.$$

В этом случае функция

$$A(f) \equiv \int_a^b f(x)dx$$

заменяется новой функцией

$$\bar{A} = \sum_{i=1}^{n} f(x_i)\Delta x_i.$$

Для успешного применения указанного выше метода вычислительной математики необходимо в первую очередь иметь рациональные способы замены пространства R другим пространством $\bar{R}$.

Методы вычислений как раздел вычислительной математики

В рамках одной дисциплины невозможно изложить или хотя бы кратко затронуть весь круг вопросов современной вычислительной математики, поэтому мы ограничились кругом вопросов, относящихся к одному разделу вычислительной математики – методам вычислений.

Чтобы более ясно охарактеризовать вопросы, относящиеся к этому разделу вычислительной математики, рассмотрим процесс решения любой математической задачи, если ее решение необходимо довести до числового результата, используя наличные вычислительные средства.

Этот процесс можно разбить на два крупных этапа.

Первый этап – выбор численного метода решения задачи или, как мы говорили ранее, замена задачи $y = A(x)$, где x и y принадлежат к некоторым функциональным пространствам R_1 и R_2 и $A(x)$ – функция, определенная на R_1, задачей $\bar{y} = \bar{A}(\bar{x})$, более удобной для вычислительных целей, но решение которой в некотором смысле близко к решению исходной задачи.

Второй этап – составление вычислительной схемы (алгоритма) или программы решения задачи $\bar{y} = \bar{A}(\bar{x})$ и сам процесс счета.

Для первого этапа необходимо наличие разработанных методов численного решения основных математических задач и должен быть известен сравнительный анализ различных методов решения одной и той же задачи с точки зрения их точности, границ применимости и целесообразности их реализации на вычислительных машинах.

Разработка и анализ этих методов и составляют *предмет методов вычислений*.

Итак, при изучении предмета «Методы вычислений» мы будем изучать численные методы, которые получают решения математических задач в числовой форме. При этом для многих задач известно только о существовании решения, но не существует конечной формулы, представляющей ее решение (пример с интегралом). Даже при наличии такой формулы ее использование для получения отдельных значений решения может оказаться неэффективным (пример с матрицей). Наконец, всегда существует необходимость решать и такие математические задачи, для которых строгие доказательства существования решения на данный момент отсутствуют.

Методы численного решения математических задач всегда составляли неотъемлемую часть математики и неизменно входили в содержание естественно-математического и инженерного образования.

Крупнейшие представители прошлого сочетали в своих исследованиях изучение явлений природы, получение их математического описания, как иногда говорят математической модели явления, и его исследование.

При анализе усложненных моделей потребовалось создание специальных, обычно численных методов – методы Ньютона, Эйлера, Лобачевского, Гаусса, Чебышева, Эрмита – свидетельствуют о том, что их разработкой занимались крупнейшие ученые своего времени.

Как самостоятельная математическая дисциплина вычислительная математика оформилась в начале XX века.

К этому времени в основном были разработаны разнообразные, достаточно эффективные и надежные алгоритмы приближенного решения широкого круга математических задач, включающего стандартный набор задач из алгебры, математического анализа и дифференциальных уравнений.

Прогресс в развитии численных методов способствовал постоянному расширению сферы применения математики в других научных дисциплинах и прикладных разработках, откуда в свою очередь поступали запросы на решение новых проблем, стимулируя дальнейшее развитие вычислительной математики.

Метод математического моделирования, основанный на построении и исследовании математических моделей различных объектов, процессов и явлений и получении информации о них из решения связанных с этими моделями математических задач, стал одним из основных способов исследования в так называемых точных науках.

Параллельно с развитием численных методов шла разработка средств вычислений, представлявших собой различные механические, а затем электромеханические устройства для выполнения арифметических операций.

Причем прогресс в области инструментальных средств не оказывал заметного влияния на ход развития методов вычислений.

Принципиальным образом ситуация изменилась с середины нашего столетия, когда было осуществлено изобретение электронных вычислительных машин.

В результате появления ЭВМ скорость выполнения вычислительных операций выросла в миллионы раз, что позволило решить широкий круг бывших до этого практически не решаемых математических задач.

Широкое внедрение ЭВМ в практику научных и технических расчетов потребовало интенсивного развития методов численного решения самых разных математических задач, причем методов, рассчитанных на реализацию их именно на ЭВМ.

Это связано с тем, что часть из них ранее использовавшихся алгоритмов численного решения неэффективна при реализации на ЭВМ, а некоторые просто непригодны для такого использования.

Основные этапы решения инженерной задачи с применением ЭВМ. Вычислительный эксперимент

Современной формой метода математического моделирования, базирующейся на мощной вычислительной базе в виде ЭВМ и программного обеспечения, реализующего алгоритмы численного решения (методов вычислений), является вычислительный эксперимент. Этот теоретический метод включает существенные черты методологии экспериментального исследования, но эксперименты выполняются не над реальным объектом, а над его математической моделью, и экспериментальной установкой является ЭВМ.

Технологическая цепочка вычислительного эксперимента включает в себя следующие этапы:

- построение математической модели исследуемого объекта (в том числе и анализ модели, выяснение корректности поставленной математической модели);
- построение вычислительного алгоритма – метода приближенного решения поставленной задачи и его обоснование;
- программирование алгоритма на ЭВМ и его тестирование;
- проведение серии расчетов с варьированием определяющих параметров исходной задачи и алгоритма;
- анализ полученных результатов.

Каждый их этих этапов допускает возврат к любому из предыдущих с целью его уточнения и корректировки.

Следует отметить, что вычислительный эксперимент – это, как правило, не разовый счет по стандартным формулам, а прежде всего расчет серии вариантов для различных математических моделей.

Алгоритм численных вычислений, составленный для реализации на ЭВМ, должен отличаться такими свойствами, как ввод, вывод, конечность, определенность, понятность, эффективность.

Ввод. Алгоритм предполагает наличие одного или нескольких исходных данных, задаваемых до начала выполнения алгоритма. Они должны принадлежать допустимому для данного алгоритма классу и быть сформулированы на языке исходных данных.

Вывод. Алгоритм имеет один или несколько искомых результатов, принадлежащих выводу.

Конечность. Алгоритм всегда должен завершаться после конечного, предельно малого, разумного числа шагов.

Определенность. Каждый шаг алгоритма должен быть точно определен, и его выполнение не должно зависеть от фантазии или сообразительности исполнителя. Применяя алгоритм к одному и тому же исходному данному несколько раз, мы всегда будем получать один и тот же результат. В этом состоит свойство однозначности алгоритма. Действия, выполняемые на каждом шаге алгоритма, однозначно независимо от исходных данных, если последние принадлежат классу допустимых. Такая особенность алгоритма называется детерминированностью.

Определенность алгоритма заключается в его однозначности и детерминированности.

Понятность. Исполнитель должен однозначно понимать текст алгоритма и знать, как его выполнять. Для этого необходимо, чтобы исполнителю были известны правила работы с предложенным текстом. Таким образом, свойство понятности предполагает, что имеется алгоритм, определяющий процесс выполнения алгоритма, заданного в виде текста.

Эффективность. Действия, предписываемые каждым шагом алгоритма, должны быть простыми, с тем, чтобы наилучшим образом использовать возможности машины и уменьшить время решения задачи.

Содержание курса

Основы теории погрешностей

Численные методы решения систем линейных алгебраических уравнений

Численные методы решения нелинейных уравнений

Численные методы решения систем нелинейных уравнений

Интерполяция и аппроксимация функций

Численное интегрирование функций

Численное дифференцирование функций

Численное решение обыкновенных дифференциальных уравнений

1. ОСНОВЫ ЭЛЕМЕНТАРНОЙ ТЕОРИИ ПОГРЕШНОСТЕЙ

1.1. Погрешности вычислений

Источники погрешностей. Рассмотрим источники погрешностей на отдельных этапах решения задачи.

Математическая модель, принятая для описания данного процесса или явления, может внести существенные погрешности, если в ней не учтены какие-либо важные черты рассматриваемой задачи. В частности, математическая модель может прекрасно работать в одних условиях и быть совершенно неприемлемой в других, поэтому важно правильно учитывать область ее применимости.

Исходные данные задачи часто являются основным источником погрешностей. Это так называемые **неустранимые погрешности**, поскольку они не могут быть уменьшены вычислением ни до начала решения задачи, ни в процессе ее решения.

Численный метод также является источником погрешностей. Это связано, например, с заменой интеграла суммой, усечением рядов при вычислениях значений функций, интерполированием табличных данных и т.п. Такая погрешность называется **погрешностью метода**. Она, как правило, регулируема, т.е. она может быть уменьшена до любого разумного значения путем изменения некоторого параметра (например, шага интерполирования, числа членов усеченного ряда и т.п.).

Погрешность метода обычно стараются довести до величины, в несколько раз меньшую погрешности исходных данных. Дальнейшее снижение погрешности не приведет к повышению точности результатов, а лишь увеличит стоимость расчетов из-за необоснованного увеличения объема вычислений.

При вычислениях с помощью ЭВМ неизбежны **погрешности округления**, связанные с ограниченностью разрядной сетки машины. Эта погрешность может нарастать в процессе вычислений.

1.2. Абсолютная и относительная погрешности вычислений

Под *ошибкой* или *погрешностью* приближенного числа a^* обычно понимается разность между соответствующим точным числом a и данным приближением.

Во многих случаях знак этой ошибки неизвестен, поэтому целесообразней использовать абсолютную погрешность приближенного числа.

Если a – точное значение некоторой величины, а a^* – известное приближение к нему, то *абсолютной погрешностью* приближения a^* называют обычно некоторую величину $\Delta(a^*)$, про которую известно, что она удовлетворяет неравенству:

$$\left|a^* - a\right| \le \Delta(a^*).$$

Абсолютная погрешность недостаточна для характеристики точности измерения или вычисления. Так, например, если при измерении длин двух стержней получены результаты

$$l_1 = 105{,}3 \text{ см} \pm 0{,}1 \text{ см} \quad \text{и} \quad l_2 = 4{,}8 \text{ см} \pm 0{,}1 \text{ см},$$

то, несмотря на совпадение предельных абсолютных погрешностей, качество первого измерения выше, чем второго. Для точности данных измерений существенна абсолютная погрешность, приходящаяся на единицу длины, которая носит название относительной погрешности.

Относительной погрешностью называют некоторую величину $\delta(a^*)$, про которую известно, что она удовлетворяет неравенству:

$$\left|\frac{a^* - a}{a^*}\right| \le \delta(a^*).$$

Например, для рассмотренного выше примера мы получаем, что при измерении первого стержня относительная погрешность составила

$$0{,}1:105{,}3 = 0{,}00095,$$

а при измерении второго –

$$0{,}1:4{,}8 = 0{,}21.$$

Отсюда уже видно какое измерение было точнее.

Относительную погрешность часто выражают в процентах (0,095% и 2,1% для рассмотренных в примере измерений).

Она дает более точное представление о величине ошибки, содержащейся в некоторой величине.

1.3. Погрешности арифметических операций

Погрешность вычисления значений функции

Пусть $y = f(x_1, x_2, ..., x_n)$ непрерывно дифференцируемая функция, x_i^* – приближенное значение ее аргументов, для которых $|x^* - x| \le \Delta(x^*)$ – известные абсолютные погрешности.

Для погрешности приближенного значения функции $y^* = f(x_1^*, x_2^*, ..., x_n^*)$ по формуле Лагранжа получаем

$$y - y^* = \sum_{i=1}^{n} a_i(\theta)(x_i - x_i^*),$$

где $a_i(\theta) = \dfrac{\partial f}{\partial x_i}(\theta)$ (θ – некоторая внутренняя точка между точками x и x^*).

Заменяя $a_i(\theta)$ на $a_i^* = \dfrac{\partial f}{\partial x_i}(x_1^*, x_2^*, ..., x_n^*)$, получаем

$$y - y^* \approx \sum_{i=1}^{n} a_i^* (x_i - x_i^*) .$$

Оценка погрешности соответственно:

$$\left| y - y^* \right| \le \Delta(y^*) = \sum_{i=1}^{n} A_i \Delta(x_i^*) ,$$

где $A_i = \sup \left| \dfrac{\partial f}{\partial x_i}(x_1, x_2, ..., x_n) \right|$

или

$$\left| y - y^* \right| \approx \sum_{i=1}^{n} A_i^* \Delta(x_i^*) , \tag{1.1}$$

где $A_i = \left| \dfrac{\partial f}{\partial x_i}(x_1^*, x_2^*, ..., x_n^*) \right|$.

Погрешность суммы

$$\left|y - y^*\right| \approx \sum_{i=1}^{n} A_i^* \Delta(x_i^*), \qquad (1.1)$$

где $A_i^* = \left|\dfrac{\partial f}{\partial x_i}(x_1^*, x_2^*, ..., x_n^*)\right|$.

Пусть задана функция

$$y = f(x_1, x_2) = x_1 + x_2 .$$

Тогда $A_i^* = \left|\dfrac{\partial f}{\partial x_i}\right| = 1, \quad i = 1, 2.$

Следовательно, по формуле (1.1) получаем выражение для оценки абсолютной погрешности суммы:

$$\Delta(y^*) = A_1^* \Delta(x_1^*) + A_2^* \Delta(x_2^*) = \Delta(x_1^*) + \Delta(x_2^*) . \qquad (1.2)$$

Относительная погрешность

$$\delta(y^*) = \frac{\Delta(x_1^*) + \Delta(x_2^*)}{x_1^* + x_2^*} = \frac{\dfrac{\Delta(x_1^*)}{x_1^*} \cdot x_1^* + \dfrac{\Delta(x_2^*)}{x_2^*} \cdot x_2^*}{x_1^* + x_2^*} = \frac{\delta(x_1^*) \cdot x_1^* + \delta(x_2^*) \cdot x_2^*}{x_1^* + x_2^*} . \quad (1.3)$$

Пусть $m \le \delta(x_1^*) \le \delta(x_2^*) \le M$, тогда $m \le \delta(y^*) \le M$, т.е. *при сложении приближенных величин относительная погрешность не возрастает*.

Погрешность разности

$$\left| y - y^* \right| \approx \sum_{i=1}^{n} A_i^* \Delta(x_i^*), \qquad (1.1)$$

где $A_i^* = \left| \dfrac{\partial f}{\partial x_i}(x_1^*, x_2^*, ..., x_n^*) \right|$.

Пусть задана функция

$$y = f(x_1, x_2) = x_1 - x_2 .$$

Имеем: $A_i^* = \left| \dfrac{\partial f}{\partial x_i} \right| = 1, \quad i = 1, 2.$

Следовательно, по формуле (1.1) получаем выражение для оценки абсолютной погрешности суммы: $\Delta(y^*) = A_1^* \Delta(x_1^*) + A_2^* \Delta(x_2^*) = \Delta(x_1^*) + \Delta(x_2^*)$.

Тогда аналогично предыдущему примеру имеем формулу

$$\delta(y^*) = \frac{\Delta(x_1^*) + \Delta(x_2^*)}{x_1^* - x_2^*} = \frac{\dfrac{\Delta(x_1^*)}{x_1^*} \cdot x_1^* + \dfrac{\Delta(x_2^*)}{x_2^*} \cdot x_2^*}{x_1^* - x_2^*} = \frac{\delta(x_1^*) \cdot x_1^* + \delta(x_2^*) \cdot x_2^*}{x_1^* - x_2^*}. \qquad (1.3)$$

Отсюда следует, что если приближенные значения x_1^* и x_2^* близки друг к другу, то относительная погрешность их разности $\delta(y^*)$ может оказаться намного больше $\delta(x_1^*)$ и $\delta(x_2^*)$.

Погрешность произведения

Пусть задана функция

$$y = f(x_1, x_2) = x_1 \cdot x_2 .$$

Имеем:

$$A_1^* = \left|\frac{\partial f}{\partial x_1}(x_1^*, x_2^*)\right| = \left|x_2^*\right|,$$

$$A_2^* = \left|\frac{\partial f}{\partial x_2}(x_1^*, x_2^*)\right| = \left|x_1^*\right|.$$

Тогда абсолютная погрешность

$$\Delta(y^*) = \left|x_2^*\right| \Delta(x_1^*) + \left|x_1^*\right| \Delta(x_2^*) . \qquad (1.4)$$

Относительная погрешность

$$\delta(y^*) = \frac{\Delta(y^*)}{\left|x_1^* \cdot x_2^*\right|} = \frac{\Delta(x_1^*)}{\left|x_1^*\right|} + \frac{\Delta(x_2^*)}{\left|x_2^*\right|} = \delta(x_1^*) + \delta(x_2^*) . \qquad (1.5)$$

$$\left|y - y^*\right| \approx \sum_{i=1}^{n} A_i^* \Delta(x_i^*) , \qquad (1.1)$$

где $A_i^* = \left|\frac{\partial f}{\partial x_i}(x_1^*, x_2^*, ..., x_n^*)\right|$.

Погрешность частного

$$\left|y - y^*\right| \approx \sum_{i=1}^{n} A_i^* \Delta(x_i^*), \qquad (1.1)$$

где $A_i^* = \left|\frac{\partial f}{\partial x_i}(x_1^*, x_2^*, ..., x_n^*)\right|$.

Пусть задана функция $y = f(x_1, x_2) = \frac{x_1}{x_2}$.

$$A_1^* = \left|\frac{\partial f}{\partial x_1}(x_1^*, x_2^*)\right| = \left|\frac{1}{x_2^*}\right|,$$

$$A_2^* = \left|\frac{\partial f}{\partial x_2}(x_1^*, x_2^*)\right| = \left|-\frac{x_1^*}{\left(x_2^*\right)^2}\right| = \left|\frac{x_1^*}{\left(x_2^*\right)^2}\right|.$$

$$\Delta(y^*) = A_1^*\Delta(x_1^*) + A_2^*\Delta(x_2^*) = \left|\frac{1}{x_2^*}\right|\Delta(x_1^*) + \left|\frac{x_1^*}{\left(x_2^*\right)^2}\right|\Delta(x_2^*) = \frac{\left|x_2^*\right|\Delta(x_1^*) + \left|x_1^*\right|\Delta(x_2^*)}{\left|x_2^*\right|^2}. \qquad (1.6)$$

$$\delta(y^*) = \frac{\Delta(y^*)}{\left|y^*\right|} = \frac{\dfrac{\left|x_2^*\right|\Delta(x_1^*) + \left|x_1^*\right|\Delta(x_2^*)}{\left|x_2^*\right|^2}}{\left|\dfrac{x_1^*}{x_2^*}\right|} = \frac{\Delta(x_1^*)}{\left|x_1^*\right|} + \frac{\Delta(x_2^*)}{\left|x_2^*\right|} = \delta(x_1^*) + \delta(x_2^*). \quad (1.7)$$

Пример

Вычислить и определить погрешности результатов $X = \frac{m^2 n^3}{\sqrt{k}}$, где m=28,3(±0,02), n=7,45(±0,01), k=0,678(±0,003);

$$\delta(x_1^* \times x_2^*) = \delta(x_1^*) + \delta(x_2^*)$$

$$\delta\left(\left(x_1^*\right)^k\right) = k\delta(x_1^*)$$

$$\delta\left(\frac{x_1^*}{x_2^*}\right) = \delta(x_1^*) + \delta(x_2^*)$$

1. Находим точное значение функции:

$$m^2 = 800{,}9;\ n^3 = 413{,}5;\ \sqrt{k} = 0{,}8234;\ X = \frac{800{,}9 \cdot 413{,}5}{0{,}8234} = 402200 = 4{,}02 \cdot 10^5.$$

2. Находим относительные погрешности аргументов

$$\delta_m = 0{,}02/28{,}3 = 0{,}00071;\ \delta_n = 0{,}01/7{,}45 = 0{,}00134;\ \delta_k = 0{,}003/0{,}678 = 0{,}00443.$$

3. Находим относительную погрешность вычисления функции

$$\delta_X = 2\delta_m + 3\delta_n + 0{,}5\delta_k = 0{,}0077 = 0{,}77\%.$$

4. Находим абсолютную погрешность вычисления функции

$$\Delta_X = 4{,}02 \cdot 10^5 \cdot 0{,}0077 = 3{,}1 \cdot 10^3.$$

Вычислить и определить погрешности результатов $N = \dfrac{(n-1)(m+n)}{(m-n)^2}$, где $n = 3{,}057(\pm 0{,}0004)$, $m = 5{,}72(\pm 0{,}02)$.

1. Находим точное значение функции и погрешности вычсилений:

$n - 1 = 2{,}057(\pm 0{,}0004)$;

$m + n = 5{,}72(\pm 0{,}02) + 3{,}057(\pm 0{,}0004) = 8{,}777(\pm 0{,}0204)$;

$m - n = 5{,}72(\pm 0{,}02) - 3{,}057(\pm 0{,}0004) = 2{,}663(\pm 0{,}0204)$;

$$N = \frac{2{,}0567 \cdot 8{,}777}{2{,}663^2} = \frac{2{,}0567 \cdot 8{,}777}{7{,}092} = 2{,}545 \approx 2{,}55\,;$$

$$\delta_N = \frac{0{,}0004}{2{,}0567} + \frac{0{,}0204}{8{,}777} + 2\frac{0{,}0204}{2{,}663} = 0{,}000149 + 0{,}00233 + 2 \cdot 0{,}00766 = 0{,}0178 = 1{,}78\%$$

$$\Delta_N = 2{,}55 \cdot 0{,}0178 = 0{,}045\,.$$

Ответ: $N \approx 2{,}55(\pm 0{,}045)\,; \delta_N = 1{,}78\%$.

Обратная задача оценки погрешности

Иногда возникает задача *определения допустимой погрешности аргументов, при которой погрешность значений функций будет не более заданной величины* ε.

Используем ранее полученное неравенство

$$\left|y - y^*\right| \le \sum_{i=1}^{n} A_i \Delta(x_i^*), \; A_i = \sup\left|\frac{\partial f}{\partial x_i}\right|.$$

Должно быть $\sum_{i=1}^{n} A_i \Delta(x_i^*) \le \varepsilon$.

$$\sum_{i=1}^{n} A_i \Delta(x_i^*) \le \varepsilon$$

При n = 1 вопрос решается однозначно:

$$\Delta(x_1^*) \le \frac{\varepsilon}{A_1}.$$

При n > 1 возможны разные подходы:

1. Считать погрешности всех аргументов одинаковыми

$$\Delta(x_1^*) = \Delta(x_2^*) = \ldots = \Delta(x_n^*) = \Delta.$$

Тогда получаем $\Delta \sum_{i=1}^{n} A_i \le \varepsilon$, следовательно $\Delta \le \frac{\varepsilon}{\sum_{i=1}^{n} A_i}$.

2. Считать, что вклад погрешности каждого аргумента в погрешность результатов одинаков: $A_1 \Delta(x_1^*) = A_2 \Delta(x_2^*) = \ldots = A_n \Delta(x_n^*) = \frac{\varepsilon}{n}$, тогда $\Delta(x_i^*) = \frac{\varepsilon}{A_i n}$.

Если для разных аргументов достижение определенной точности их задания существенно различается, то можно ввести функцию стоимости $F(\Delta(x_1^*),...,\Delta(x_n^*))$ затрат на задание точки $x_1^*,...,x_n^*$ с заданными абсолютными погрешностями $\Delta(x_1^*),...,\Delta(x_n^*)$ и искать ее минимум в области

$$\sum A_i \Delta(x_i^*) \le \varepsilon, \quad \Delta(x_i^*) \ge 0.$$

Операции с матрицами

Система mn чисел, расположенных в прямоугольной таблице из m строк и n столбцов,

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \qquad (1)$$

называется *матрицей*.

Числа a_{ij} ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$), составляющие данную матрицу, называются ее *элементами*. Здесь первый индекс i обозначает номер строки элемента, а второй j – номер его столбца.

Для матрицы (1) часто употребляется сокращенная запись $A=[a_{ij}]$ ($i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$) или $A=[a_{ij}]_{m,n}$.

Если $m = n$, матрица (1) называется *квадратной порядка* n.

Квадратная матрица вида

$$A = \begin{bmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \alpha_n \end{bmatrix}, \qquad (2)$$

называется *диагональной*.

В случае если $\alpha_i = 1$ $(i = 1, 2, \ldots, n)$, то матрица (2) называется *единичной* и обозначается обычно буквой E, т.е.

$$E = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 \end{bmatrix},$$

Введя *символ Кронекера*

$$\delta_{ij} = \begin{cases} 0, \text{если } i \neq j, \\ 1, \text{если } i = j \end{cases}$$

можно записать $E = [\delta_{ij}]$.

Матрица, все элементы которой равны нулю, называется *нулевой* и обозначается 0.

С квадратной матрицей $A = [a_{ij}]_{n,n}$ связан *определитель* (*детерминант*)

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix} = \sum_{(\alpha_1, \alpha_2, \dots, \alpha_n)} (-1)^{\chi} a_{1\alpha_1} a_{2\alpha_2} \dots a_{n\alpha_n},$$

где сумма распространяется на всевозможные перестановки $(\alpha_1, \alpha_2, \dots, \alpha_n)$ элементов 1, 2, …, n и, следовательно, содержит $n!$ слагаемых, причем $\chi = 0$, если перестановка четная, и $\chi = 1$, если перестановка нечетная.

Примечание: Говорят, что в данной перестановке числа i и j составляют инверсию, если $i > j$, но i стоит в этой перестановке раньше j. Перестановка называется четной, если ее символы составляют четное число инверсий, и нечетной – в противоположном случае. Например, 451 362 ($n = 6$) – 8 инверсий, 38 524 671 ($n = 8$) – 15 инверсий.

Сумма и разность матриц

Суммой двух матриц $A = [a_{ij}]$ и $B = [b_{ij}]$ одинакового типа называется матрица $C = [c_{ij}]$ того же типа, элементы которой c_{ij} равны суммам соответствующих элементов a_{ij} и b_{ij} матриц A и B, т.е. $c_{ij} = a_{ij}+b_{ij}$. Таким образом,

$$C = A + B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}.$$

Из определения суммы матриц непосредственно вытекают следующие ее свойства:

1) $A+(B+C)=(A+B)+C$;

2) $A+B=B+A$;

3) $A+0=A$.

Аналогично определяется *разность матриц*:

$$A - B = \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \dots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \dots & a_{2n} - b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \dots & a_{mn} - b_{mn} \end{bmatrix}.$$

Умножение матриц на число

Произведением матрицы $A = [a_{ij}]$ на число α называется матрица, элементы которой получены умножением всех элементов матрицы A на число α, т.е.

$$A\alpha = \alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2n} \\ \dots & \dots & \dots & \dots \\ \alpha a_{m1} & \alpha a_{m2} & \dots & \alpha a_{mn} \end{bmatrix}.$$

Из определения произведения числа на матрицу непосредственно вытекают следующие его свойства:

1) $1 \cdot A = A$;
3) $\alpha(\beta A) = (\alpha\beta)A$;
5) $\alpha(A + B) = \alpha A + \alpha A$,
2) $0 \cdot A = 0$;
4) $(\alpha + \beta)A = \alpha A + \beta A$;

где A и B – матрицы одинакового порядка, α и β – числа.

Заметим, что если матрица A квадратная порядка n, то $\det(\alpha A) = \alpha^n \det A$.

Матрица $-A = (-1)A$ называется *противоположной*.

Перемножение матриц

Пусть

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \text{ и } B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ \dots & \dots & \dots & \dots \\ b_{p1} & b_{p2} & \dots & b_{pq} \end{bmatrix}$$

матрицы типов соответственно $m \times n$ и $p \times q$. Если число столбцов матрицы A равно числу строк матрицы B, т.е. $n = p$,

то для этих матриц определена матрица C типа $m \times q$, называемая их *произведением*:

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1q} \\ c_{21} & c_{22} & \dots & c_{2q} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mq} \end{bmatrix}, \text{ где } c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, \quad (i = \overline{1,m};\ j = \overline{1,q}).$$

Из определения вытекает следующее правило умножения матриц: *чтобы получить элемент, стоящий в i-й строке и j-м столбце произведения двух матриц, нужно элементы i-й строки первой матрицы умножить на соответствующие элементы j-го столбца второй и полученные произведения сложить*.

Матричное произведение обладает следующими свойствами:

1) $A(BC)=(AB)C$;

2) $\alpha(AB)= (\alpha A)B$;

3) $(A+B)C=AC+BC$;

4) $C(A+B)=CA+CB$,

где A, B и C – матрицы, α – число.

Произведение двух матриц не обладает переместительным свойством, т.е., вообще говоря,

$$AB \neq BA.$$

В тех случаях, когда $AB \neq BA$, матрица A и B называются *перестановочными* (коммутативными).

Так, например, единичная матрица E перестановочна с любой квадратной матрицей того же порядка, причем

$$AE + EA = A.$$

Если A и B – квадратные матрицы одного и того же порядка, то

$$\det(AB) = \det(BA) = \det A \cdot \det B.$$

Транспонированная матрица

Заменив в матрице

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

типа $m \times n$ строки соответственно столбцами, получим так называемую *транспонированную матрицу*

$$A' = A^{\text{T}} = \begin{bmatrix} a_{11} & a_{21} & \dots & an1 \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix},$$

типа $n \times m$.

Транспонированная матрица обладает следующими свойствами:

1) дважды транспонированная матрица совпадает с исходной, т.е. $A'' = (A')' = A$;
2) транспонированная матрица суммы равна сумме транспонированных матриц слагаемых, т.е. $(A + B)' = A' + B'$;
3) транспонированная матрица произведения равна произведению транспонированных матриц сомножителей, взятому в обратном порядке, т.е. $(AB)' = B'A'$.

Если матрица A – квадратная, то очевидно, $\det A' = \det A$. Матрица A=[a_{ij}] называется *симметрическо*й, если она совпадает со своей транспонированной матрицей, т.е. если $A' = A$.

Обратная матрица

Обратной матрицей по отношению к данной называется матрица, которая, будучи умноженной как справа, так и слева на данную матрицу, дает единичную матрицу

$$AA^{-1} = A^{-1}A = E .$$

Некоторые основные свойства обратной матрицы.

1. $\det A^{-1} = \dfrac{1}{\det A}$.

2. $(AB)^{-1} = B^{-1}A^{-1}$.

3. $(A^{-1})' = (A')^{-1}$.

Элементарные преобразования матриц

Следующие преобразования матриц носят названия *элементарных*:

1) перестановка двух строк или столбцов;
2) умножение всех элементов какой-либо строки (столбца) на одно и то же число, отличное от нуля;
3) прибавление к элементам какой-либо строки (столбца) соответствующих элементов другой строки (столбца), умноженных на одно и то же число.

Две матрицы называются эквивалентными, если одна получается из другой с помощью конечного числа элементарных преобразований.