

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Факультет прикладной математики – процессов управления

А. П. ИВАНОВ, Л. Т. ПОЗНЯК

**ПРАКТИКУМ ПО ЧИСЛЕННЫМ МЕТОДАМ
ВЫЧИСЛЕНИЕ ФУНКЦИЙ**

Методические указания

Санкт-Петербург
2016

ГЛАВА 1. ОСНОВЫ ТЕОРИИ ПОГРЕШНОСТЕЙ.

На погрешность результата приближенного решения задачи влияют следующие причины:

- а) *неточность информации о решаемой задаче.* Ошибки в начальных данных дают ту часть погрешности в решении, которая не зависит от математической стороны решения задачи и называется *неустранимой погрешностью*.
- б) *Погрешность аппроксимации (методическая погрешность).* При решении задачи численными методами необходимо считаться с тем, что неизбежно придётся иметь дело только с конечным количеством чисел, и с ними можно выполнить только конечное число операций. Поэтому вместо точного решения задачи приходится прибегать к приближенному методу.
- в) *Погрешность округления.* Всякое положительное число a может быть представлено в виде конечной или бесконечной десятичной дроби

$$a = \alpha_m 10^m + \alpha_{m-1} 10^{m-1} + \dots + \alpha_{m-n+1} 10^{m-n+1} + \dots, \quad (1)$$

где α_i – цифры числа a , причём старшая цифра $\alpha_m \neq 0$, а m – некоторое число (старший десятичный разряд числа a). Например,

$$3141,59\dots = 3 \cdot 10^3 + 1 \cdot 10^2 + 4 \cdot 10^1 + 1 \cdot 10^0 + 5 \cdot 10^{-1} + 9 \cdot 10^{-2} + \dots$$

На практике имеют дело с приближёнными числами, представляющими собой конечные десятичные дроби

$$b^* = \beta_m 10^m + \beta_{m-1} 10^{m-1} + \dots + \beta_{m-n+1} 10^{m-n+1}, \quad \beta_m \neq 0.$$

Определение 1. Цифра β_k в изображении числа b^* называется *верной*, если имеет место неравенство $|b - b^*| \leq \omega 10^k$, $\omega \leq 1$, чаще всего, $\omega = 0.5$. (Здесь b – точное значение величины, представленной приближённой записью через b^* .)

Очевидно, что если цифра β_k верная, то и все цифры в записи числа b , расположенные левее неё, тоже верны.

Определение 2. *Значащей цифрой* числа называется всякая его цифра в десятичном изображении, кроме нулей, стоящих слева в записи числа до первой ненулевой цифры.

Число, являющееся решением конкретной задачи, принято записывать только с *верными значащими цифрами*.

Например, в числе 0.002080 *первые три нуля* не являются значащими цифрами, так как они служат только для установления десятичных разрядов других цифр. Остальные два нуля являются значащими. В случае, если в данном числе 0.002080 последняя цифра не является верной, *то её не следует использовать в записи числа*.

Определение 3. Число $|a - a^*|$ называется абсолютной погрешностью приближённого значения a^* .

Определение 4. Число $\Delta(a^*)$ (другое обозначение – Δ_{a^*}), удовлетворяющее неравенству $|a - a^*| \leq \Delta(a^*)$, называется верхней границей (оценкой) погрешности приближённого значения a^* .

Определение 5. Число $\frac{|a - a^*|}{a^*}$ называется относительной погрешностью приближенного значения a^* .

Определение 6. Число $\delta(a^*) = \delta_{a^*}$ при $a^* \neq 0$ удовлетворяющее неравенству $\frac{|a - a^*|}{a^*} \leq \delta(a^*)$, называется верхней границей (оценкой) относительной погрешности приближённого значения a^* .

Часто в определениях 4 и 6 слово “верхняя” опускают для краткости. Если известна граница абсолютной погрешности $\Delta(a^*)$, то в качестве $\delta(a^*)$, очевидно, можно взять $\frac{\Delta(a^*)}{|a^*|}$. Если же известна верхняя граница относительной погрешности $\delta(a^*)$, то за $\Delta(a^*)$ можно взять $\delta(a^*) \cdot |a^*|$. Эту связь между $\Delta(a^*)$ и $\delta(a^*)$ выражают формулой $\Delta(a^*) = \delta(a^*) \cdot |a^*|$.

Замечание 1. Для $a^* = 0$ относительная погрешность не определена.

§1. Прямая задача теории погрешностей

В дальнейшем изложении будем считать, что погрешность округлений пренебрежимо мала по сравнению с методической погрешностью.

Пусть в выпуклой области $G \in \mathbb{R}^n$ рассматривается непрерывно дифференцируемая функция $y = f(\cdot)$. Предположим, что в точке $x = (x_1, x_2, \dots, x_n)$ области G нужно вычислить значение $y = f(x)$. Пусть нам известны лишь приближённые значения $x_1^*, x_2^*, \dots, x_n^*$ такие, что точка $x^* = (x_1^*, x_2^*, \dots, x_n^*) \in G$. Необходимо найти оценку погрешности приближённого значения функции $y^* = f(x^*)$, обусловленную погрешностями аргументов. Через погрешности $\varepsilon_i = x_i - x_i^*$ аргументов она выражается следующим образом:

$$\varepsilon = f(x_1^* + \varepsilon_1, x_2^* + \varepsilon_2, \dots, x_n^* + \varepsilon_n) - f(x_1^*, x_2^*, \dots, x_n^*),$$

или, если воспользоваться формулой Лагранжа,

$$\varepsilon = \sum_{i=1}^n \frac{\partial}{\partial x_i} f(x_1^* + \theta \varepsilon_1, x_2^* + \theta \varepsilon_2, \dots, x_n^* + \theta \varepsilon_n) \varepsilon_i, \quad 0 < \theta < 1.$$

Отсюда получается оценка для границы погрешности вычисления функции, порождённую погрешностью её аргументов:

$$|\varepsilon| \leq \Delta = \sum_{i=1}^n B_i \Delta_i, \quad (1.1)$$

где $|\varepsilon_i| \leq \Delta_i$, $B_i = \max_{\theta \in (0,1)} \left| \frac{\partial}{\partial x_i} f(x_1^* + \theta \varepsilon_1, x_2^* + \theta \varepsilon_2, \dots, x_n^* + \theta \varepsilon_n) \right|$. Таким образом решается прямая задача теории погрешности: *известны погрешности некоторой системы величин. Требуется определить погрешность вычисления заданной функции $f(\cdot)$ этих величин*, порождённую их погрешностями.

Здесь учтена лишь неустранимая погрешность вычисления функции, порождённая погрешностями её аргументов. Если же считать, что значение функция $f(x)$ не может быть вычислено точно (например, $\sqrt{2}$) и его вычисление заменяется вычислением

другой функции $f^*(x)$ (например, отрезка ряда Тейлора), то возникает и методическая погрешность вычисления функции. Для совокупной (полной) погрешности (без учёта ошибок округления) имеем:

$$|f(x) - f^*(x^*)| \leq |f(x) - f(x^*)| + |f(x^*) - f^*(x^*)| \leq |\varepsilon| + \Delta_{f^*} . \quad (1.2)$$

Таким образом для погрешности ε вычисления функции $f(x)$ следует написать оценку:

$$|\varepsilon| \leq \Delta = \Delta_{f^*} + \sum_{i=1}^n B_i \Delta_i . \quad (1.3)$$

Неравенство (1.3) назовём решением полной задачи теории погрешностей.

§2. Полная обратная задача теории погрешностей

Рассмотрим вопрос: каковы должны быть абсолютные погрешности аргументов функции и методическая погрешность функции, чтобы абсолютная величина полной погрешности вычисления функции не превышала заданной величины?

Эта задача математически неопределена, так как заданную предельную погрешность (верхнюю границу заданной абсолютной погрешности Δ) можно обеспечить, устанавливая по-разному предельные абсолютные погрешности Δ_i её аргументов и вычисления функции Δ_{f^*} , лишь бы они удовлетворяли условию:

$$\Delta = \Delta_{f^*} + \sum_{i=1}^n B_i \Delta_i , \quad \Delta_i, \Delta_{f^*} > 0 .$$

Для единообразия обозначений примем в дальнейшем соглашение $\Delta_{f^*} = \Delta_0$, $B_0 = 1$ и тогда формула (1.3) запишется формально в виде (1.1):

$$|\varepsilon| \leq \Delta = \sum_{i=0}^n B_i \Delta_i . \quad (2.1)$$

Простейшее решение обратной задачи даётся так называемым принципом **равных влияний**. Предполагается, что все слагаемые

$B_i \Delta_i$, $i = \overline{0, n}$ в правой части формулы (2.1) имеют одинаковую величину. Тогда

$$\Delta_i = \frac{\Delta}{(n+1)B_i}, \quad i = \overline{0, n}.$$

Другой столь же простой способ носит название **принципа равных погрешностей**: считается, что $\Delta_i = \Delta_j$, и тогда из той же формулы (2.1) немедленно получаем:

$$\Delta_j = \frac{\Delta}{\sum_{i=0}^n B_i}, \quad j = \overline{0, n}.$$

Исходя из особенностей задачи и вычисляемой функции можно выставлять и другие требования к уровню погрешностей аргументов: задать погрешности для части аргументов, а погрешности остальных найти из условия выполнения равенства (2.1) с учётом положительности искомых величин $\Delta_j, j = \overline{0, n}$.

Пример 2.1. Требуется определить массу металла с погрешностью $\Delta(M) = 1$ г, потребного для изготовления диска толщиной $h \approx 1$ см и радиусом $r \approx 10$ см (плотность металла ρ). Для определенности будем считать, что $\rho = 10$ г/см³, $h = 1 \pm 0.1$ см, $\pi = 3.14 \pm 0.002$, $r = 10 \pm 0.1$ см.

Решение. Поскольку масса диска (цилиндра) вычисляется по формуле $M = \rho h \pi r^2$, то поставленный вопрос эквивалентен вопросу: с какой погрешностью должны быть измерены радиус и толщина диска, а также сколько следует взять знаков в числе π , чтобы выполнить условия по точности вычисления массы диска (считая, что плотность ρ – величина точная)?

Согласно формуле (2.1), имеем:

$$\Delta_M = (B_h \Delta_h + B_\pi \Delta_\pi + B_r \Delta_r),$$

где

$$B_h = \max_{\pi, r} (\rho \pi r^2) = 3205.52;$$

$$B_\pi = \max_{h, r} (\rho h r^2) = 1122.11;$$

$$B_r = \max_{h, \pi, r} (2r \rho h \pi) = 698.15.$$

Применяя принцип равных погрешностей, получаем:

$$\Delta_h = \Delta_\pi = \Delta_r \approx 1/5000 = 2 \cdot 10^{-4}.$$

ГЛАВА 2. ОБРАТНАЯ ЗАДАЧА ДЛЯ ВЫЧИСЛЕНИЯ ЭЛЕМЕНТАРНЫХ ФУНКЦИЙ

§1. Общие положения

Пусть $z(x) = f(\varphi(x), g(x))$, $x \in [a, b]$, и требуется построить таблицу значений этой функции для узлов $x = x_i$, $i = \overline{1, k}$. Здесь x – скалярный аргумент ($x_i < x_{i+1}$). Предполагается, что каждая из трёх функций $f(\cdot)$, $\varphi(\cdot)$, $g(\cdot)$ не может быть вычислена точно и вычисляется приближённо:

$$\varphi(x) \approx \varphi^*(x), \quad g(x) \approx g^*(x), \quad f(u, v) \approx f^*(u, v).$$

Таким образом, реально вместо искомых точных значений функции $z_i = f(\varphi(x_i), g(x_i))$ будут получены (вычислены) приближённые значения $z_i^* = f^*(\varphi^*(x_i), g^*(x_i))$. Требуется определить с какой точностью должны быть вычислены $\varphi^*(x)$, $g^*(x)$ и $f^*(x)$, чтобы обеспечивалась заданная точность приближённых значений z_i^* , т.е. чтобы $|z_i - z_i^*| \leq \varepsilon$. Как нетрудно убедиться, здесь мы имеем дело с рассмотренной ранее обратной задачей теории погрешностей. В самом деле, нам требуется вычислить значения функции двух переменных $f(u, v)$ при некоторых значениях аргументов u, v , которые известны нам не точно, но могут быть найдены их приближённые значения u^* , v^* с требуемой точностью.

Пусть $u_* \leq \varphi(x) \leq u^*$, $g_* \leq g(x) \leq g^*$ $\forall x \in [x_1, x_k]$. В таком случае область G есть прямоугольник $[u_*, u^*] \times [g_*, g^*]$. Считаем далее, что производные $\frac{\partial f}{\partial u}$, $\frac{\partial f}{\partial v}$ мало изменяются в G и что

$$\left| \frac{\partial f}{\partial u}(u, v) \right| \leq c_1, \quad \left| \frac{\partial f}{\partial v}(u, v) \right| \leq c_2 \quad \forall (u, v) \in G.$$

Из предыдущих рассуждений о решении обратной задачи вытекает, что требуемая точность ε приближённых табличных значений z_i^* обеспечивается тогда, когда приближённые значения аргументов $u^* = \varphi^*(x)$ и $v^* = g^*(x)$ удовлетворяют неравенствам

$$|\varphi(x_i) - \varphi^*(x_i)| \leq \frac{\varepsilon}{3c_1}, \quad |g(x_i) - g^*(x_i)| \leq \frac{\varepsilon}{3c_2},$$

а приближённо вычисленное значение $z_i^* = f^*(u^*, v^*)$ удовлетворяет неравенству

$$|f(u^*, v^*) - f^*(u^*, v^*)| \leq \frac{\varepsilon}{3}.$$

Предполагаем, что мы умеем оценивать методическую погрешность вычисления функций $\varphi(\cdot)$, $g(\cdot)$ и $f(\cdot)$, т.е. можем дать оценки:

$$\begin{aligned} |\varphi(x) - \varphi^*(x)| &\leq \Delta_{\varphi^*}, \\ |g(x) - g^*(x)| &\leq \Delta_{g^*}, \\ |f(u^*, v^*) - f^*(u^*, v^*)| &\leq \Delta_{f^*}. \end{aligned}$$

Поставленная задача будет решена, когда будет обеспечено выполнение следующих неравенств:

$$\Delta_{\varphi^*} \leq \frac{\varepsilon}{3c_1}, \quad \Delta_{g^*} \leq \frac{\varepsilon}{3c_2}, \quad \Delta_{f^*} \leq \frac{\varepsilon}{3}.$$

§2. Пример решения общей обратной задачи теории погрешностей

Рассмотрим задачу определения погрешностей вычисления функций $\varphi(x)$ и $\psi(x)$ при заданной погрешности Δ вычисления функции $F(x)$, $x \in [a, b]$:

$$F(x) = u(\varphi(x)) \cdot v(\psi(x)).$$

Поскольку при вычислении произведения uv отсутствует методическая погрешность, то погрешности вычисления u и v (обозначенные здесь Δ_u и Δ_v) и порождаемые методическими погрешностями функций $\varphi(x)$ и $\psi(x)$, определяются по указанной выше формуле (2.1), где следует положить $\Delta_0 = 0$:

$$B_u = \sup_{u,v} \left| \frac{\partial F}{\partial u} \right| = \bar{v}, \quad B_v = \sup_{u,v} \left| \frac{\partial F}{\partial v} \right| = \bar{u},$$

$$\bar{v}\Delta_u + \bar{u}\Delta_v = \Delta.$$

Здесь Δ – оценка требуемой погрешности вычисления заданной функции $F(x)$. Найдя Δ_u и Δ_v , используя метод равных погрешностей или метод равных влияний, находим и погрешности вычисления функций $\varphi(x)$ и $\psi(x)$: $\Delta_u = \Delta_{\varphi^*}$, $\Delta_v = \Delta_{\psi^*}$

§3. Пример построения таблицы значений функции

Рассмотрим конкретный пример. Пусть требуется построить таблицу значений функции

$$z(x) = \frac{\sqrt{\sin(0.9x + 0.51)}}{xe^{x+0.3}}$$

для $x = 0.5(0.01)0.6$, т.е. для $x \in [0.5; 0.6]$ с шагом 0.01 с заданной точностью $\varepsilon = 10^{-6}$.

Положим

$$\varphi(x) = \sin(0.9x + 0.51), \quad g(x) = xe^{x+0.3}, \quad f(u, v) = \frac{\sqrt{u}}{v}.$$

Тогда $z(x) = f(\varphi(x), g(x))$. Найдём пределы изменения величин u, v при $x \in [0.5; 0.6]$. Поскольку функции $\varphi(\cdot)$ и $g(\cdot)$ монотонны на $[0.5; 0.6]$, то $\sin 0.96 < u < \sin 1.05$, $0.5e^{0.8} < v < 0.6e^{0.9}$.

Интервалы изменения u, v можно расширить, чтобы не вычислять верхние и нижние границы изменения этих функций с большей точностью. Положим $\sin 0.96 \approx 0.8$, $\exp(0.8) \approx 2.2$ (с недостатком); $\sin 1.05 \approx 0.9$, $\exp(0.9) \approx 2.5$ (с избытком) (использованные значения функций взяты из таблиц).

Таким образом, можно положить

$$G = \{(u, v) \mid 0.8 \leq u \leq 0.9, \quad 1.1 \leq v \leq 1.5\}.$$

Оценим в G частные производные

$$\begin{aligned} \frac{\partial f(u, v)}{\partial u} &= \frac{1}{2v\sqrt{u}}, & \frac{\partial f(u, v)}{\partial v} &= \frac{-\sqrt{u}}{v^2} : \\ \left| \frac{\partial f(u, v)}{\partial u} \right| &\leq \frac{1}{2 \cdot 1.1\sqrt{0.8}} < 0.6, & \left| \frac{\partial f(u, v)}{\partial v} \right| &\leq \frac{\sqrt{0.9}}{1.21} < 0.9 \end{aligned}$$

Итак, в данном примере $c_1 = 0.6$, $c_2 = 0.9$ и, следовательно, функцию $\varphi(x)$ нужно вычислять с точностью $\varepsilon_1 = \frac{10^{-6}}{1.8}$, функцию $g(x)$ — с точностью $\varepsilon_2 = \frac{10^{-6}}{2.7}$, а функцию $f(u, v)$ — с точностью $\varepsilon_3 = \frac{10^{-6}}{3}$.

Функции $\varphi(x)$, $g(x)$ предлагается вычислять, разлагая функции $\cos y$ и e^t в ряд Маклорена по аргументам $y = \pi/2 - (0.9x + 0.51)$ и $t = x + 0.3$ (при этом будет $y \in (0, \pi/4) \in [0, 1]$ и ряд для $\cos y$ станет лейбницевым). Функцию $f(u, v)$ предлагается вычислять, определяя приближённое значение функции $\sqrt{u^*}$ по формуле Герона (частного случая формулы Ньютона):

$$w_{k+1} = \frac{1}{2} \left(w_k + \frac{u^*}{w_k} \right),$$

w_0 – приближённое значение $\sqrt{u^*}$ с избытком. Можно, к примеру, взять в данном случае $w_0 = 1$. Далее полагаем $f^*(u^*, v^*) = \frac{\sqrt{u^*}}{v^*}$.

Для всех трёх функций мы умеем оценивать абсолютную величину методической погрешности (с учётом того, что элементарные функции вычисляются с помощью разложения в ряд Маклорена):

$$\begin{aligned} |\varphi(x) - \varphi^*(x)| &< |y^{2n}/(2n)!| = \Delta_{\varphi^*} = \varepsilon_1, \\ |g^*(x) - g(x)| &< xt^p/p! = \Delta_{g^*} = \varepsilon_2, \\ |f(u^*, v^*) - f^*(u^*, v^*)| &< |w_k - w_{k-1}|/v^* = \Delta_{f^*} = \varepsilon_3. \end{aligned}$$

Следовательно, требуемая точность табличных значений функции $z(x)$ будет обеспечена тогда, когда номера n, p, k будут удовлетворять неравенствам

$$\begin{aligned} |y^{2n}/(2n)!| &\leq \frac{10^{-6}}{1.8}, \\ xt^p/p! &\leq \frac{10^{-6}}{2.7}, \\ |w_k - w_{k-1}|/v^* &\leq \frac{10^{-6}}{3}. \end{aligned}$$

Здесь

$$v^* = g^*(x) = x \sum_{i=0}^p t^i/i!, \quad t = x + 0.3.$$

ЛИТЕРАТУРА

1. *Н. Бахвалов, Н. Жидков, Г. Кобельков.* Численные методы. — М.: Изд. Физматлит, 2006.
<http://www.studfiles.ru/preview/393510/>
(дата обращения 01.02.2016 г.)
2. *В. М. Вержбицкий.* Численные методы. Линейная алгебра и нелинейные уравнения. — М.: Изд. Высшая школа, 2000.
<http://padabum.com/d.php?id=146968>
(дата обращения 9.02.2016)
3. *Б. П. Демидович, И. А. Марон* Основы вычислительной математики. — М.: Изд. Наука, 1966.
<http://bookfi.net/book/509165> (дата обращения 1.02.2016)