Тема 2. Урахування похибок наближених обчислень. Класифікація похибок. Абсолютна та відносна похибки. Точні десяткові знаки. Похибка функції. Похибка математичних операцій. Зумовленість матриць і систем лінійних алгебричних рівнянь.

При чисельному розв'язуванні майже будь-яких математичних і прикладних задач отримуємо не точний розв'язок, а результат з тією чи іншою мірою точності. Це пояснюється тим, що під час розв'язування задачі виникає низка похибок, пов'язаних з неточністю вхідних даних, виконанням арифметичних операцій не над точними дійсними числами, а над їхніми наближеннями, які одержують в результаті заокруглень, способом розв'язування поставленої математичної задачі. При цьому ***повну похибку*** результату розв'язування задачі складають: ***похибка задачі*** - пов'язана з наближеним характером вихідної моделі(неусувна), ***похибка методу***(усувна), ***похибка заокруглень***.

Абсолютна та відносна похибки

Точні десяткові знаки.

Нехай А – довільне додатнє число, а а – деяке наближення числа. Тоді А може бути зображене у вигляді нескінченного десяткового дробу

$A = k_1 10^m + k_2 10^{m-1} + ... + k_n 10^{m-n+1} + ...$, де $k_1,\ k_2, ..., k_n$ - десяткові знаки числа А, n - кількість десяткових знаків, $k_1 \neq 0$.

Однак у разі розв'язання задач на комп'ютерах ми можемо використовувати лише числа зі скінченною і цілком визначеною кількістю розрядів. Припустимо, що для зображення чисел використовують n розрядів. Тоді наближене число a, що замінює точне число А, набуде такого вигляду

$a = k_1 10^m + k_2 10^{m-1} + ... + k_n 10^{m-n+1}$, де $k_1 \neq 0$, n - кількість десяткових знаків.

Значущою цифрою вважатимемо будь-яку з цифр 1, 2, …, 9, а також цифру 0, якщо вона є проміжною або знаходиться в кінці числа і є результатом вимірювань або обчислень.

Приклад:

$a_1 = 0,000105$ - три значущі цифри

$a_2 = 20$ - дві значущі цифри

$a_3 = 30,00$ - чотири значущі цифри

$a_4 = 3,0 \cdot 10^3$ або $30 \cdot 10^2$ - дві значущі цифри

Означення

Величина $\Delta a = |A - a|$ називається ***абсолютною похибкою*** наближеного числа a, а $\delta a = \dfrac{\Delta a}{|a|}$ - його ***відносною похибкою***.

На практиці найчастіше число А є невідомим. Тому в цьому випадку ми не зможемо визначити абсолютну похибку. Тоді користуються так званою граничною абсолютною похибкою.

Гранична абсолютна похибка наближеного числа a - це будь-яке число, не менше за абсолютну похибку цього числа: $\Delta_a \geq \Delta a$.

Якщо відома гранична абсолютна похибка числа a, то точне число А лежить у межах $a - \Delta_a \leq A \leq a + \Delta_a$.

Цю формулу скорочено записують у вигляді: $A = a \pm \Delta_a$.

Приклад.

Якщо $A = \pi$, $a = 3{,}14$, то оскільки $3{,}14 < \pi < 3{,}15$, $|\pi - a| < 0{,}01$. Тому можемо прийняти $\Delta_a = 0{,}01$. За умови $3{,}14 < \pi < 3{,}142$, маємо $|\pi - a| < 0{,}002$. Тому одержуємо кращу граничну абсолютну похибку $\Delta_a = 0{,}002$.

Зауваження.

Сформульоване поняття граничної абсолютної похибки досить широке, оскільки будь-яке число Δ_a, яке задовольняє нерівність $a - \Delta_a \leq A \leq a + \Delta_a$, можна прийняти за граничну абсолютну похибку. На практиці за Δ_a треба брати якомога менше число.

Означення.

Гранична відносна похибка δ_a наближеного числа a називається будь-яке число, котре є не менше за відносну похибку цього наближеного числа.

Згідно з означення маємо $\delta \leq \delta_a$.

Приклад.

Якщо $A = \pi$, $a = 3{,}14$, то, оскільки $\delta = \dfrac{\Delta}{\pi} < \dfrac{0{,}002}{3{,}14} < 0{,}0007$, за граничну відносну ***похибку наближеного числа*** $a = 3{,}14$ ***можна прийняти*** $\delta_a = 0{,}0007$.

Означення

Число a ***має*** n ***точних десяткових знак***ів, якщо його абсолютна похибка не перевищує одиниці n-го десяткового знака, тобто $\Delta a \leq 10^{m-n+1}$.

Приклад.

$a = 8{,}2415$, $\Delta a = 0{,}0015$.

Тут $a = 8{,}2415 = 8 \cdot 10^0 + 2 \cdot 10^{-1} + 4 \cdot 10^{-2} + \ldots$Звідси $m = 0$.

Тоді $n = 3$.

Теорема

Якщо число a має n точних десяткових знаків, то його відносна похибка не перевищує величини $\delta a \le \dfrac{1}{k_1 \cdot 10^{n-1}}$.

Теорема

Якщо відносна похибка числа a задовольняє умову $\delta a = \dfrac{1}{10^n}$, то число a має n точних десяткових знаків.

Похибка функцій.

Розглянемо функцію $y = f(x_1, x_2, ..., x_n)$, неперервно диференційована в деякій області зміни аргументу. Нехай $\overline{x}_1, \overline{x}_2, ..., \overline{x}_n$ - наближені значення аргументу, причому їхні абсолютні похибки дорівнюють, відповідно, $\Delta x_1, \Delta x_2, ..., \Delta x_n$ і малі порівняно з величинами $\overline{x}_1, \overline{x}_2, ..., \overline{x}_n$. Тоді абсолютна похибка функції $y = f(x_1, x_2, ..., x_n)$

$$\Delta y = \left| f(\overline{x}_1, \overline{x}_2, ..., \overline{x}_n) - f(x_1, x_2, ..., x_n) \right|.$$

На практиці Δx_i - малі значення. Тому обмежимось лінійною частиною приросту Δy, якою є повний диференціал dy. Тоді

$$\Delta y \approx |dy| = |df(x_1, x_2, ..., x_n)| = \left| \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i \right| \le \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i} \right| \Delta x_i .$$

Звідси позначивши через Δ_{x_i} $(i = 1, 2, ..., n)$ граничні абсолютні похибки аргументів x_i, а через Δ_y граничну похибку функції y для малих Δx_i отримаємо

$$\Delta_y = \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i} \right| \Delta_{x_i} .$$

Легко бачити, що $\delta y = \dfrac{\Delta y}{|y|} \le \displaystyle\sum_{i=1}^{n} \left| \frac{\frac{\partial f}{\partial x_i}}{y} \right| \Delta x_i = \sum_{i=1}^{n} \left| \frac{\partial \ln f(x_1, x_2, ..., x_n)}{\partial x_i} \right| \Delta x_i$.

Тоді за граничну відносну похибку функції y можна взяти

$$\delta_y = \sum_{i=1}^{n} \left| \frac{\partial}{\partial x_i} \ln y \right| \Delta_{x_i} .$$

Похибка арифметичних операцій.

1. ***Похибки суми.***

   Теорема.

   Абсолютна похибка алгебраїчної суми декількох наближених чисел не перевищує суми абсолютних похибок цих чисел.

Нехай $u = \pm x_1 \pm x_2 \pm \ldots \pm x_n$, тоді $|\Delta u| \le |\Delta x_1| + |\Delta x_2| + \ldots + |\Delta x_n|$.

Наслідок.

За граничну абсолютну похибку алгебраїчної суми декількох наближених чисел можна прийняти суму граничних абсолютних похибок цих чисел, тобто $\Delta_u = \Delta_{x_1} + \Delta_{x_2} + \ldots + \Delta_{x_n}$.

Теорема

Гранична відносна похибка суми декількох наближених чисел одного й того ж знака не перевищує найбільшу з граничних відносних похибок цих чисел.

Нехай $u = x_1 + x_2 + \ldots + x_n$, тоді $\delta_u \le \max\limits_{1 \le i \le n} \delta_{x_i}$.

2. ***Похибка різниці***

   Розглянемо різницю двох наближених чисел x_1 та x_2: $u = x_1 - x_2$. Тоді

   $\Delta_u = \Delta_{x_1} + \Delta_{x_2}$, $\delta_u = \dfrac{\Delta_{x_1} + \Delta_{x_2}}{|A|}$.

   З останньої формули випливає, що для близьких чисел x_1 та x_2 гранична відносна похибка буде досить велика. Тому в обчислювальних алгоритмах бажано уникати віднімання досить близьких чисел.

3. ***Похибка добутку і частки***

   Нехай $u = x_1 \cdot x_2 \cdot \ldots \cdot x_n$, де всі множники можемо вважати додатними, тоді

   $$\delta_{\prod x_i} = \sum_{i=1}^{n} \delta_{x_i}.$$

   Якщо $y = \dfrac{x_1}{x_2}$, де $x_1, x_2 > 0$, то $\delta_{\frac{x_1}{x_2}} = \delta_{x_1} + \delta_{x_2}$.

   Отже, граничні відносні похибки під час множення та ділення наближених чисел додаються.

   Найчастіше використовують **норми вектора**:

   $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i|$ - l_1 норма, або норма найменших абсолютних різниць;

   $\|x\|_2 = \left(\sum\limits_{i=1}^{n} x_i^2 \right)^{(1/2)}$ - l_2 норма, або евклідова норма чи норма найменших квадратів;

   $\|x\|_\infty = \max\limits_{1 \le i \le n} |x_i|$ - l_∞ норма, або sup-норма чи норма Чебишева і підпорядковані їй норми матриць

   $\|A\|_1 = \max\limits_{1 \le k \le n} \sum\limits_{i=1}^{n} |a_{ik}|$;

$\|A\|_2 = \sqrt{\lambda_{\max}\left(A^T A\right)}$;

$\|A\|_\infty = \max\limits_{1\le i\le n} \sum\limits_{k=1}^{n} |a_{ik}|$, відповідно, де $\lambda_{\max}\left(A^T A\right)$ - максимальне власне число матриці $A^T A$.

Величину $\operatorname{cond} A = \|A\| \|A^{-1}\|$ називають числом зумовленості матриці A.

Якщо $\operatorname{cond} A$ - невелике, то матриця A СЛАР називається добре зумовленою, якщо ж $\operatorname{cond} A$ - велике, то матриця A називається погано зумовленою.

Контрольні запитання та завдання

1. Вкажіть, що включає повна похибка результату розв'язування задачі.
2. Заокруглюючи числа до трьох значущих цифр, визначити абсолютну та відносну похибки наближених чисел:

   а) 5,678945; б) -0,0039765; в) -123,456; г) 2478,3.
3. Визначити кількість точних десяткових знаків у числі x, якщо відома його абсолютна похибка:

   а) $x = 349.172$, $\Delta x = 0.1$; б) $x = 13.04342$, $\Delta x = 0.1 \times 10^{-3}$.
4. Визначити кількість точних десяткових знаків у числі x, якщо відома його віднона похибка:

   а) $x = 1.7246$, $\delta x = 0.1 \times 10^{-2}$; $x = 23.732$, $\delta x = 0.2$.
5. Визначити суму наближених чисел і зазначити їхні похибки:

   а) 0.145+321+78.2 (всі знаки точні);

   б) 1.27569-1.27531 (відомо, що чотири знаки точні).
6. Оцінити похибку обчислення функції:

   $f(x, y, z) = \dfrac{x^2 z}{y^3}$, $x = 0.15 \pm 0.005$; $y = 2.13 \pm 0.01$; $z = 1.14 \pm 0.007$.
7. Коли матриця СЛАР є добре обумовленою?
8. Яка задача теорії похибок називається прямою?