СТАТТІ ТА ОГЛЯДИ

doi: https://doi.org/10.15407/visn2022.10.053

ЗАДІРАКА
Валерій Костянтинович —
академік НАН України,
завідувач відділу оптимізації чисельних методів Інституту кібернетики ім. В.М. Глушкова НАН України

ШВІДЧЕНКО
Інна Віталіївна —
кандидат фізико-математичних наук, провідний науковий співробітник відділу оптимізації чисельних методів Інституту кібернетики ім. В.М. Глушкова НАН України

ПРО ТОЧНІСТЬ НАБЛИЖЕНОГО РОЗВ'ЯЗКУ ЗАДАЧІ

У статті розглянуто питання якості наближеного розв'язку задачі, комплексний підхід до оцінки точності (повна похибка обчислювального алгоритму), а також оптимальні за точністю обчислювальні алгоритми. Наведено випадки доцільності їх застосування.

Ключові слова: повна похибка обчислювального алгоритму, похибки методу, неусувна похибка, похибка заокруглення, оптимальні за точністю обчислювальні алгоритми.

Характеристики обчислювального алгоритму. Питання, яким присвячена стаття, дуже важливі, особливо на етапі, коли задача вже розв'язана і потрібно дослідити якість її наближеного розв'язку.

У чому ж полягає ця важливість? По-перше, без з'ясування питання щодо якості наближеного розв'язку задачі отриманий наближений розв'язок краще не використовувати. Чому? Тому що отриманий розв'язок задачі надалі застосовують при конструюванні нових зразків техніки, у відповідних комп'ютерних технологіях для вирішення нагальних завдань економіки країни, у військовій справі тощо. Якщо ж ми не знаємо якість цього розв'язку (або не маємо оцінки якості), то його використання може призвести до техногенних катастроф. На жаль, це не гіпотеза, а підтверджені факти, відомі інциденти, які сталися у ядерній енергетиці, морській справі, космічній промисловості, при будівництві гідротехнічних споруд, використанні високоточної зброї тощо.

Для того, щоб проаналізувати, від чого залежить точність, розглянемо комплексний підхід до оцінки точності наближеного розв'язку задачі. Комплексність підходу пов'язана з поняттям повної похибки обчислювального алгоритму [1, 2], яка за нерівністю трикутника не перевищує суми абсолютних похибок методу, неусувної та заокруглень.

Похибка методу виникає внаслідок заміни функції її апроксимацією, інтегралів — інтегральними сумами, похідних — скінченними різницями тощо. Для отримання оцінок похибок методу найчастіше використовують методи теорії апроксимації.

При цьому важливим фактором є якість отриманої оцінки (мажорантна або мажорантна непокращувана). Крім того, оцінки можуть бути апріорні, апостеріорні, мажорантні, асимптотичні, детерміновані та ймовірнісні. Кожен із зазначених видів оцінок похибки методу має свої переваги перед іншими, і використання тієї чи іншої оцінки залежить від постановки задачі і наявних обчислювальних ресурсів, виділених для її розв'язання.

Тепер розглянемо неусувну похибку обчислювального алгоритму, яка виникає внаслідок наближеного задання вхідної інформації про задачу. Слід звернути увагу, що лише для модельних задач можна припустити, що вхідна інформація задана точно, для реальних же задач вона задана наближено. І ще одне зауваження — неусувна похибка не залежить від алгоритму, вона залежить лише від задачі. Її потрібно відняти від точності, з якою необхідно розв'язати задачу, а те, що залишиться, припадає на суму двох похибок — похибки методу і похибки заокруглень. Підбираючи параметри алгоритму, ми маємо задовольнити ці обмеження.

Неточність вхідної інформації породжує так звані некоректно поставлені задачі, які розв'язують методами регуляризації [3].

Оцінки неусувної похибки можуть бути як детерміновані, так і ймовірнісні. Ймовірнісні оцінки більш точні, але справджуються вони не з імовірністю $p = 1$, як детерміновані, а з достатньо високою для практики ймовірністю, близькою до одиниці.

На підтвердження сказаного вище розглянемо детерміновані та ймовірнісні оцінки неусувної похибки обчислення перетворення Фур'є фінітно заданої функції

$$F(\omega) = \int_a^b f(x) e^{-i\omega x} dt$$

у припущенні, що інформація про $f(x)$ задана не більш ніж у N точках рівномірної сітки на $[a, b]$.

Нехай $f(x)$ у вузлах x_i задана наближено з максимально можливою похибкою,

$$\delta : \left| f(x_j) - \overline{f}(x_j) \right| \le \delta , \ j = \overline{0,N} .$$

Розглянемо таку модель вхідних даних. Нехай $\overline{f}_j = f_j + \xi_j$ і можливі похибки ξ_j задання підінтегральної функції у вузлах є взаємно незалежними випадковими величинами, розподіленими:

1) рівномірно на відрізку $[0, \delta]$;

2) за нормальним законом з дисперсією σ^2. Припустимо також, що системні похибки відсутні, тобто справжні значення вимірюваних величин дорівнюють математичним сподіванням можливих значень.

За припущення (1) можна покласти

$$\sigma = \frac{\delta}{\sqrt{3}}$$

і з імовірністю 0,96 отримати оцінку

$$E_H \le 5(b-a)\delta / \sqrt{3N},$$

а в разі припущення (2) —

$$\sigma = \frac{\delta}{3}$$

і з імовірністю 0,98 отримати оцінку

$$E_H \le (b-a)\delta / \sqrt{N}.$$

Детермінована оцінка неусувної похибки має вигляд

$$E_H \le \delta(b-a).$$

Порівнюючи ймовірнісні оцінки з детермінованою оцінкою, бачимо, що імовірнісні істотно точніші.

У випадку, коли $a = -\infty$, $b = \infty$ (класичний випадок) задача обчислення $F(\omega)$ є некоректною, оскільки E_H може бути як завгодно великою.

Для стійкого обчислення $F(\omega)$ необхідно компенсувати вплив великих значень x, помножуючи $f(x)$ на відповідно задану функцію $f(x, \alpha)$, яка визначена для всіх невід'ємних значень параметра α за будь-яких x.

Якщо функцію $f(x, \alpha)$ підпорядкувати певним умовам, то оператор

$$F_\alpha(\omega) = \int_{-\infty}^{\infty} f(x) \cdot f(x,\alpha) e^{-i\omega x} dx$$

буде регуляризуючим для $F(\omega)$.

Для цього достатньо $f(x, \alpha)$ взяти у вигляді

$$f(x,\alpha) = \frac{1}{1 + \alpha \cdot \xi(x)},$$

де $\xi(x)$ — додатна функція, порядок зростання якої не нижчий, ніж $x^{2+\delta}$, $\delta \geq 0$.

Нехай

$$f(x,\alpha)=\frac{1}{1+\alpha(\varepsilon)\left|x\right|^{r}},\ r>1.$$

Тоді при оптимальному виборі параметра регуляризації ($\alpha^{*}(\varepsilon) = c\cdot\varepsilon$, де c — деяка константа) оцінка неусувної похибки обчислення $F(\omega)$ має вигляд

$$E_H \leq c_1 \cdot \varepsilon^{1-1/r},\ r>1$$

(без регуляризації оцінка була $E_H \leq \varepsilon\cdot\infty$).

Розглянемо тепер похибки заокруглення, які виникають при реалізації арифметичних операцій на комп'ютері із заокругленням результату до фіксованої кількості розрядів. Розрізняють два режими роботи комп'ютера: з фіксованою комою і з плаваючою комою.

Похибка заокруглення E_3 залежить від виду правила заокруглення:

- класичне;
- відсічення;
- рандомізоване.

Якщо проаналізувати їх за якістю (накопиченням похибок заокруглення), можна сказати, що найкращим є рандомізоване правило (коли розряди, що відкидаються, становлять половину останнього із залишених розрядів і заокруглення в більшу або меншу сторону відбувається випадково за деяким законом розподілу), далі йде класичне, а найгірше (відбувається найбільше накопичення похибки) — правило відсічення.

На жаль, починаючи від комп'ютера М-220 і дотепер навіть у суперкомп'ютерах закладено саме це нестійке правило відсічення.

Чому так сталося? Цьому є дві причини. Перша — правило відсічення простіше реалізується. Електронники цьому раді, оскільки їх не хвилює точність результату, але математики дуже занепокоєні, адже, коли зросла складність розв'язуваних задач, похибки заокруглення стають більшими від інших і це призводить до того, що, наприклад, для задач математичного моделювання можна отримати комп'ютерні моделі, які не мають нічого спільного з фізичними системами. З року в рік проблема поглиблюється, і уряди деяких країн уже почали створювати організації для дослідження накопичення похибок заокруглення при розв'язанні задач трансобчислювальної складності (наприклад, це стосується накопичення похибки заокруглення в методі скінченних елементів та способів боротьби з цим накопиченням).

Друга причина — поява на початку 1980-х років комп'ютерів серії ЕС ЕОМ, які були орієнтовані на задачі автоматизованої системи управління технологічними процесами (АСУ ТП), тобто вони не використовували серйозні математичні задачі і правило відсічення їх влаштовувало. Проте сьогодні електронники безпідставно продовжують використовувати його, не пропонуючи користувачам альтернативи.

При розв'язанні задач трансобчислювальної складності та високоточних задач слід контролювати накопичення похибки заокруглення і застосовувати прийоми зниження похибки заокруглення (моделювання правила рандомізованого заокруглення, використання багаторозрядної арифметики [4], обчислення зі змінною розрядністю тощо).

Нехай $fl(\cdot)$ — результат обчислення на комп'ютері виразу в дужках, тобто рівність вигляду

$$z = fl\begin{pmatrix} \pm \\ x\cdot y \\ / \end{pmatrix}$$

означає, що z отримано виконанням відповідної операції в режимі плаваючої коми. Нехай використовується класичне правило заокруглення до τ двійкових розрядів у нормалізованих мантис чисел. Похибки заокруглення в цих операціях мають вигляд

$$z = fl\begin{pmatrix} \pm \\ x\cdot y \\ / \end{pmatrix}(1+\varepsilon),$$

де ε — відносна похибка і $\left|\varepsilon\right| \leq 2^{-\tau}$.

Результати, отримані в режимі плаваючої коми, приводять до оцінок вигляду

$$(1-2^{-\tau})^{r} \leq 1+\varepsilon \leq (1+2^{-\tau})^{r},$$

які можна спростити, припустивши, що виконується умова $r \cdot 2^{-\tau} < 0{,}1$ (це цілком виправдано в практичних застосуваннях для будь-якого прийнятого τ).

Тоді

$$(1+2^{-\tau})^r < 1+1{,}06 \cdot r \cdot 2^{-\tau},$$

$$(1-2^{-\tau})^r < 1-1{,}06 \cdot r \cdot 2^{-\tau},$$

і

$$1-1{,}06 \cdot r \cdot 2^{-\tau} \le 1+\varepsilon \le 1+1{,}06 \cdot r \cdot 2^{-\tau},$$

звідки $|\varepsilon| < 1{,}06 \cdot r \cdot 2^{-\tau}$. Останнє співвідношення використовується в усіх наступних оцінках:

1) $fl\left(x_1 \cdot x_2 \cdot \ldots \cdot x_N\right) = \prod_{i=1}^{N} x_i \left(1+E\right),$

де

$$|E| < (N-1) \cdot 1{,}06 \cdot 2^{-\tau};$$

2) $$fl(x_1 + x_2 + \ldots + x_N) \equiv$$
$$\equiv x_1(1+\varepsilon_1) + \ldots + x_N(1+\varepsilon_N),$$

де

$$|\varepsilon_1| \le (N-1) \cdot 1{,}06 \cdot 2^{-\tau}, \ |\varepsilon_r| \le (N-r+1) \cdot 1{,}06 \cdot 2^{-\tau},$$

$$r = \overline{2,N}.$$

Тут передбачалося, що

$$s_2 = fl(x_1 + x_2); \ s_r = fl(s_{r-1} + x_r), \ r = \overline{3,N};$$

3) $$fl(x_1 y_1 + \ldots + x_N y_N) \equiv$$
$$\equiv x_1 y_1 (1+\varepsilon_1) + \ldots + x_N y_N (1+\varepsilon_N),$$

де

$$|\varepsilon_1| \le N \cdot 1{,}06 \cdot 2^{-\tau}, \ |\varepsilon_r| \le (N-r+2) \cdot 1{,}06 \cdot 2^{-\tau},$$

$$r = \overline{2,N};$$

4) $$fl\left(\frac{x_1 \cdot x_2 \cdot \ldots \cdot x_m}{y_1 \cdot y_2 \cdot \ldots \cdot y_n}\right) = \frac{x_1 \cdot x_2 \cdot \ldots \cdot x_m}{y_1 \cdot y_2 \cdot \ldots \cdot y_n}(1+E_1),$$

де

$$|E_1| < (m+n-1) \cdot 1{,}06 \cdot 2^{-\tau}.$$

На основі зазначених результатів можна одержувати мажорантні оцінки похибок заокруглення для багатьох обчислювальних алгоритмів розв'язку задач обчислювальної та прикладної математики.

Маючи оцінки основних джерел похибок (E_H, E_M, E_τ), можна на основі нерівності трикутника отримати оцінку абсолютної повної похибки обчислювального алгоритму $E \le E_H + E_M + E_\tau$.

Саме повна похибка є гарантією якості наближеного розв'язку задачі, оскільки і E_H, і E_M, і E_τ реально супроводжують обчислювальний процес. Неврахування хоча б одного виду похибки гарантію якості дати не може.

Ми розглянули одну з основних характеристик обчислювального алгоритму — точність. У практиці чисельного розв'язку задач розглядають й інші характеристики обчислювального алгоритму, такі як час розв'язку задачі та пам'ять комп'ютера, необхідна для розв'язку задачі. Ми не будемо їх аналізувати, а лише зауважимо, що всі перелічені характеристики взаємопов'язані, і часто покращення однієї з них веде до погіршення інших.

Оптимальні за точністю обчислювальні алгоритми. Одним з основних критеріїв оптимальності наближеного розв'язку задач може бути вимога його максимальної точності (чи мінімальної похибки) при заданих ресурсах, які можна використовувати в процесі розв'язання. Поняття «ресурс» охоплює обсяг і точність вхідних даних задачі, вільну для використання пам'ять комп'ютера, ліміт часу обчислень на певному комп'ютері, наявний запас математичного забезпечення комп'ютера тощо.

На початку дослідження в такій постановці природно розглянути питання про «потенційну спроможність» чисельних методів, тобто про максимальну точність розв'язку, якої можна досягти при заданій вхідній інформації про задачу.

Кожний обчислювальний алгоритм розв'язання конкретної задачі використовує лише скінченний обсяг вхідних даних про задачу і тим самим автоматично є обчислювальним алгоритмом розв'язання класу задач, які мають такі самі вхідні дані. На цій множині задач завжди знайдуться дві задачі, при розв'язанні яких досягаються найгірша і найкраща границя значення характеристики, що оптимізується. Тому кожний, у тому числі оптимальний, обчислювальний алгоритм розв'язання задачі, який нас цікавить, буде мати певну «потенційну спроможність». Якщо, наприклад, існують дві задачі з тими самими вхідними даними, точні розв'язки яких x_1 і x_2 є елементами метричного простору, причому відстань між

ними $\rho(x_1, x_2) \geq d > 0$ то для кожного обчислювального алгоритму їх розв'язання отримується розв'язок x, який має властивість

$$\max_{i=1,2} \rho(x, x_i) \geq \frac{d}{2}.$$

Це означає, що не існує обчислювального алгоритму, який давав би розв'язок розглянутої задачі з гарантованою точністю, меншою за $d/2$. Для того щоб підвищити точність розв'язку задачі, потрібно залучити додаткові відомості про неї. Тоді задача належатиме до нового, більш «вузького» класу задач, обчислювальні алгоритми розв'язання яких матимуть більшу «потенційну спроможність».

Під оптимальним розв'язком задачі будемо розуміти розв'язок з максимально можливою точністю при наявній інформації, а під оптимальним обчислювальним алгоритмом — алгоритм, який дає оптимальний розв'язок.

Існують різні постановки задач (критерії прийняття рішень в умовах невизначеності) побудови оптимальних за точністю алгоритмів. Наголосимо тут на двох методах — методі «капелюхів» та методі граничних функцій, розробленому в Інституті кібернетики ім. В.М. Глушкова НАН України [1, 2]. Ці методи добре зарекомендували себе при розв'язанні задач апроксимації, чисельного інтегрування та мінімізації функції.

В якому ж разі доцільно їх застосовувати? Наведемо кілька прикладів задач, для розв'язання яких ці методи є ефективними:

- задачі, що не піддаються розв'язанню наявними алгоритмами і програмами;
- високоточні задачі;
- задачі інформаційної безпеки [5];
- різного роду задачі оборонного характеру.

Слід зазначити, що поняття повної похибки і оптимальних алгоритмів становлять основу комп'ютерної технології розв'язання задач обчислювальної та прикладної математики із заданими значеннями характеристик якості за точністю та швидкодією [6].

Інститут кібернетики ім. В.М. Глушкова НАН України за зазначеною тематикою проводить міжнародні наукові форуми «Питання оптимізації обчислень». На сьогодні їх уже було проведено 47. Працюють 7 секцій з типових класів задач обчислювальної та прикладної математики. Форуми відбуваються що два роки у вересні.

REFERENCES
[СПИСОК ЛІТЕРАТУРИ]

1. Ivanov V.V. *Metody vychisleniy na EVM (Methods of computing)*. Kyiv: Naukova Dumka, 1986 (in Russian).
   [Иванов В.В. *Методы вычислений на ЭВМ*: справочное пособие. Киев: Наук. думка, 1986.]
2. Zadiraka V.K. *Teoriya vychisleniya preobrazovaniya Fur'e (Theory for computing the Fourier transform)*. Kyiv: Naukova Dumka, 1983 (in Russian).
   [Задирака В.К. *Теория вычисления преобразования Фурье*. Киев: Наук. думка, 1983.]
3. Morozov V.A. *Regulyarnye metody resheniya nekorrektnyh zadach* (*Regular methods for solving ill-posed problems*). Moscow, 1974 (in Russian).
   [Морозов В.А. *Регулярные методы решения некорректных задач*. Москва: Изд-во Моск. ун-та, 1974.]
4. Zadiraka V.K., Tereshchenko A.M. *Kompiuterna aryfmetyka bahatorozriadnykh chysel u poslidovnii ta paralelnii modeliakh obchyslen*. Kyiv: Naukova Dumka, 2021. (in Ukrainian).
   [Задірака В.К., Терещенко А.М. *Комп'ютерна арифметика багаторозрядних чисел у послідовній та паралельній моделях обчислень*. Київ: Наук. думка, 2021.]
5. Zadiraka V.M., Kudin A.M. Analiz stojkosti kriptograficheskih i steganograficheskih sistem na osnove obshchej teorii optimal'nyh algoritmov. *Journal of Qafgaz University. Mathematics and Computer Science*. 2010. **30**: 49—58. (in Russian).
   [Задирака В.М., Кудин А.М. Анализ стойкости криптографических и стеганографических систем на основе общей теории оптимальных алгоритмов. *Journal of Qafgaz University. Mathematics and Computer Science*. 2010. Т. 30. С. 49—58.]
6. Sergienko I.V., Zadiraka V.K., Lytvyn O.M. *Elements of the General Theory of Optimal Algorithms*. Springer, 2021. https://doi.org/10.1007/978-3-030-90908-6

Valery K. Zadiraka

V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine
ORCID: https://orcid.org/0000-0001-9628-0454

Inna V. Shvidchenko

V.M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine
ORCID: https://orcid.org/0000-0002-5434-2845

ON ACCURACY OF THE APPROXIMATE PROBLEM SOLUTION

The article considers the issues of the quality of the approximate problem solution, the comprehensive approach to accuracy assessment (the total error of the computational algorithm), as well as the computational algorithms that are optimal in terms of accuracy, and the cases they should be used in.

Keywords: total error of the computational algorithm, method errors, non-removable, rounding, optimal computational algorithms.