

## Особливості застосування багаторозрядної арифметики в математичному моделюванні

Олександр Хіміч<sup>1</sup>, Олена Ніколаєвська<sup>2</sup>, Ігор Баранов<sup>3</sup>

<sup>1</sup> Академік, д. фіз.-мат. н., професор, Інститут кібернетики ім. В.М. Глушкова НАН України, Проспект Академіка Глушкова, 40, 03187, Київ, e-mail: khimich505@gmail.com

<sup>2</sup> к. фіз.-мат. н., Інститут кібернетики ім. В.М. Глушкова НАН України, Проспект Академіка Глушкова, 40, 03187, Київ, e-mail: elena\_nea@ukr.net

<sup>3</sup> Інститут кібернетики ім. В.М. Глушкова НАН України, Проспект Академіка Глушкова, 40, 03187, Київ, e-mail: vlasov@ukr.net

*В роботі досліджуються особливості використання багаторозрядної арифметики в математичному моделюванні. Досліджуються засоби підвищення точності комп'ютерних розв'язків погано обумовлених систем лінійних алгебраїчних рівнянь за допомогою бібліотеки багаторозрядної арифметики GMP.*

**Ключові слова:** багаторозрядна арифметика, бібліотека GMP, , наближено задані вихідні дані, достовірність розв'язків, похибки заокруглення, машинна точність.

**Вступ.** Для багатьох обчислень, в тому числі тих, що використовують емпіричні дані, IEEE 32-бітна арифметика з плаваючою комою є достатньою та бажаною, оскільки суттєво економиться пам'ять та час розв'язування задачі. Але для більшості задач необхідна підвищена точність обчислень. Більше того, досвід показує, що в деяких випадках використання навіть 64-бітної арифметики не достатньо для отримання достовірних розв'язків.

Одним з найважливіших напрямків використання комп'ютерів, які забезпечують науково-технічний прогрес, є їх застосування для комп'ютерного моделювання процесів, більшість з яких зводиться або мають своїм проміжним етапом розв'язування задач обчислювальної математики. Використання комп'ютерів паралельної архітектури дає можливість набагато збільшити розмірності дискретних моделей. Але, незважаючи на досить повне теоретичне обґрунтування методів розв'язування цих задач, іноді отримують комп'ютерні розв'язки, що не мають фізичного сенсу.

Причинами цього можуть бути: по-перше, наближений характер вихідних даних та похибки, які виникають внаслідок їх представлення в пам'яті комп'ютера, по-друге, похибки заокруглення в процесі обчислень, обумовлені скінченною величиною розрядної сітки комп'ютера; по-третє, похибки.

Слід також зазначити, що методи класичної математики можуть бути непридатні для комп'ютерної реалізації. Так, наприклад, класичні методи визначення такого фундаментального поняття лінійної алгебри як ранг матриці

непридатні з обчислювальної точки зору для реалізації з обмеженою розрядністю обчислень. У цьому випадку слід використовувати визначення рангу матриці, яка враховує наближений характер машинних моделей задач [1], а для його обчислення – стійкі з обчислювальної точки зору алгоритми, зокрема алгоритм сингулярного розвинування матриць.

Інший приклад пов'язаний з виродженістю матриць. Реалізація класичного визначення виродженості матриці – рівність нулю її визначника – в комп'ютерній арифметиці очевидно є неприйнятною (великий обсяг комп'ютерних обчислень для матриць, які виникають на практиці призводить до неправильних результатів через похибки обчислень). У цьому випадку доцільно перевіряти матрицю на виродженість в межах машинної точності, використовуючи в арифметиці з плаваючою комою співвідношення

$$1.0 + r(\text{cond}A) \neq 1.0, \quad (1)$$

де  $r(\text{cond}A) = 1/\text{cond}A$ ,  $\text{cond}A$  – оцінка числа обумовленості матриці [1, 2]. Виконання цієї умови, яка виконується в арифметиці з плаваючою комою, означає, що матриця має повний ранг у межах машинної точності. У випадку невиконання цього співвідношення для забезпечення достовірності обчислень необхідно продовжити дослідження з використанням підвищеної розрядності обчислень для ідентифікації виродженості чи поганої обумовленості матриці [3].

Слід зазначити, що аналогічні рішення щодо ідентифікації властивостей задач і адаптації комп'ютерного середовища до їх властивостей у паталогічних випадках мають місце і для інших класів задач (близьких і кратних власних значень, близьких і кратних коренів поліномів, нелінійних рівнянь).

Отже, однією з актуальних проблем отримання достовірних комп'ютерних результатів математичного моделювання є створення засобів розв'язування задач з довільною комп'ютерною розрядністю. В цій роботі акцентується увага на вирішенні проблеми отримання достовірного розв'язку програмно-алгоритмічними методами в умовах багаторозрядної арифметики.

## 1. Деякі факти про багаторозрядну арифметику з плаваючою комою

Проблеми підвищення точності комп'ютерних обчислень стали досліджувати ще в середині минулого століття. В Інституті кібернетики ім. В.М. Глушкова в 1965 році вперше в світі на комп'ютерах серії МІР було реалізовано виконання обчислень з довільною розрядністю. Причому оператор «розрядність» можна було використовувати в будь-якому місці програми [4].

З появою комп'ютерів паралельних архітектур стало можливим розв'язувати задачі математичного моделювання надвеликих розмірів і проблеми використання підвищеної комп'ютерної розрядності стали ще більш актуальними.

У 2008 року в напрямку підвищення розрядності обчислень був зроблений важливий крок, пов'язаний з публікацією стандарту IEEE 754-2008, який замінив раніше діючий стандарт обчислень з плаваючою комою IEEE 754-1985. Стандарт 1985 року передбачав 2 типи чисел з плаваючою комою: одинарної (32-розрядні)

і подвійної (64-розрядні) точності. У стандарті IEEE 754-2008 однією з перших апаратних реалізацій став арифметичний співпроцесор Intel 8087, в якому додатково використовується внутрішній «розширений» 80-розрядний формат з 64-розрядною мантисою. Проте на більшості сучасних робочих станціях та суперкомп'ютерів використовуються стандартні масові процесори, які не підтримують довільну розрядність. Це спонукало до створення апаратних та програмних засобів для реалізації обчислень з підвищеною точністю [5].

## 2. Особливості застосування багаторозрядної арифметики для розв'язування СЛАР

Чисельні методи розв'язків систем лінійних алгебраїчних рівнянь мають одну загальну властивість. Саме реально обчислений розв'язок (псевдорозв'язок) є точним відповідно до зворотного аналізу помилок [6] для деякої збуреної задачі. Ці збурення дуже малі і нерідко їх можна порівняти з похибками заокруглення вхідних даних. Якщо вхідні дані задані з похибкою (вимірювань, розрахунків тощо), зазвичай вони вже містять значно більші похибки, ніж похибки округлень. У цьому випадку будь-яка спроба покращити машинний розв'язок (псевдорозв'язок) без залучення додаткових відомостей про точну задачу або похибки вхідних даних виявиться неспроможною. Положення суттєво змінюється, якщо розглядається математична задача з точними вихідними даними. Тепер критерій поганої чи хорошої обумовленості машинної моделі задачі залежить від математичних властивостей комп'ютерної моделі задачі (числа обумовленості матриці задачі) та математичних властивостей процесора (довжини машинного слова) і з'являється принципова можливість досягти будь-якої заданої точності комп'ютерного розв'язку.

В цьому випадку уточнити комп'ютерний розв'язок можна, використавши ітераційне уточнення розв'язку або розв'язання системи з підвищеною розрядністю, зокрема, використовуючи бібліотеку GMP [7] для реалізації обчислень із довільною розрядністю.

Далі, для прогнозу довжини мантиси (машинного слова), яка забезпечує задану точність для розв'язку (сумісних систем) можна користуватися наступним емпіричним правилом: кількість вірних десяткових значущих цифр у комп'ютерному розв'язку оцінюється величиною

$$\mu - \alpha, \quad (2)$$

де  $\mu$  - десятковий порядок мантиси числа з плаваючою точкою,  $\alpha$  - десятковий порядок числа обумовленості. Таким чином, знаючи обумовленість матриці системи та точність обчислень на комп'ютері, можна визначити необхідну розрядність для отримання достовірного розв'язку.

Одним з інструментів для підвищення розрядності є бібліотека GMP, яка має великий набір функцій і дозволяє організувати процес обчислення з різною розрядністю. Бібліотека GMP використовується для роботи над цілими, раціональними числами та числами з плаваючою точкою. Головна особливість бібліотеки – розрядність чисел (precision) практично необмежена. Тому основна

сфера застосування - комп'ютерні алгебраїчні обчислення, криптографія та ін. Функції бібліотеки GMP дозволяють не тільки задавати розрядність на початку програми і виконувати обчислення з цією розрядністю, але і змінювати розрядність при необхідності в процесі обчислень, тобто. різні фрагменти алгоритму виконувати із різною розрядністю. Для того, щоб проводити обчислення з бажаною розрядністю, необхідно в програмах Сі або С++, вставити звернення на відповідні функції з бібліотеки GMP для перетворення типів даних та арифметичних дій над ними, а також підключити файл gmp.h, де описані прототипи цих функцій .

Можливості бібліотеки були апробовані для дослідження розв'язків вироджених і погано обумовлених систем. В роботах [8, 9] було проведено експерименти по застосуванню бібліотеки GMP, а саме, було досліджено питання застосування бібліотеки GMP для отримання уточненого розв'язку СЛАР з погано обумовленими матрицями. Знаючи обумовленість матриці системи та точність обчислень на комп'ютері, можна визначити необхідну розрядність для отримання достовірного рішення.

В табл. 1 наведено значення *macheps* (найбільше число з плаваючою точкою, для якого виконується  $1.0 + macheps = 1.0$ ) і яке характеризує точність плаваючої арифметики для різної розрядності:

Таблиця 1

Значення *macheps* для розрядності double, 64, 128 на процесорах Intel

Мова програмування	Довжина мантиси	<i>macheps</i>
C++	double (53)	1.110e-16
	long double(64)	2.71050543121376108502e-20
C++ з використанням GMP	128	1.46936793852785938496092e-39
	256	4.31808427754722231269317e-78

Приклад. Розглянемо систему лінійних рівнянь алгебри  $Ax = b$ , де  $A$  і  $b$  мають такий вигляд:

$$A = \begin{bmatrix} 0.1348531574394464 & 0.1878970588235294 & 0.1909117647058824 & 0.1779264705882353 \\ 0.1878970588235294 & 0.262 & 0.265 & 0.247 \\ 0.1909117647058824 & 0.265 & 0.281 & 0.266 \\ 0.1779264705882353 & 0.247 & 0.266 & 0.255 \end{bmatrix}$$
$$b = \begin{bmatrix} 0.3516 \\ 0.4887 \\ 0.5105 \\ 0.4818 \end{bmatrix}.$$

«Точний» розв'язок системи (розв'язок, порахований на розрядності 100) має вигляд:

$$x = \begin{bmatrix} 0.6662162162161606738798064490430572 \dots e13 \\ -0.4016891891890723506952166298245477 \dots e13 \\ -0.1665540540539970051894018639784241 \dots e13 \\ 0.9797297297302797072574940696301115 \dots e12 \end{bmatrix}, \quad condA = 2.089 \dots e^{16} \dots$$

В табл. 2 представлені результати розв'язків СЛАР, які отримані з подвоєною розрядністю за допомогою програми C++, а також з підвищеною розрядністю за допомогою програми C++ з використанням функцій бібліотеки GMP.

Таблиця 2.

Розв'язок СЛАР з розрядністю double та підвищеною розрядністю на Інпарком-256

Мова програмування	Довжина мантиси	Комп'ютерний розв'язок системи
C++	Double (53)	$x_1=3.60239e+12$ $x_2=-2.17203e+12$ $x_3=-9.00599e+11$ $x_4=5.29764e+11$
C++ з використанням GMP	64	$x_1=0.666199107066943565109e13$ $x_2=-0.40167887337851497738e13$ $x_3=-0.166549776766692724716e13$ $x_4=0.979704569216725111902e12$
	128	$x_1=0.6662162162161606738798067836677102584254e13$ $x_2=-0.4016891891890723506952168315835297097735e13$ $x_3=-0.1665540540539970051894019476345873995266e13$ $x_4=0.9797297297302797072574945617251937251362e12$

Експерименти з дослідження часу розв'язання задач великих порядків показали, що, наприклад, час розв'язання системи лінійних алгебраїчних рівнянь з однією і тією ж матрицею з розрядністю 64 за допомогою програми C++ з використанням функцій GMP значно збільшується порівняно з часом розв'язання з подвійною розрядністю за допомогою програми C++ без використання функцій GMP (табл. 3), т.к. потрібен додатковий час на виклик та ініціалізацію функцій GMP. Проте із збільшенням розмірів завдання ця різниця зменшується, так як. час викликів функцій GMP виявляється значно меншим, ніж основний час виконання арифметичних операцій. Подальше збільшення розрядності у програмі C++ з використанням функцій GMP ненабагато збільшує час розв'язання задачі.

Таблиця 3

Порівняльна характеристика часу розв'язування СЛАР на Інпарком-256

Порядок матриці	C++	C++, GMP		
	double (53)	64	128	256
200	0,04 сек	0,75 сек	0,84 сек	1,04 сек
1000	6,55 сек	96.44 сек	14.72 сек	107.7 сек
4500	688,94 сек	7320,06 сек	8050,12 сек	10104,56 сек

**Висновки.** Таким чином, для задач математичного моделювання змінна багаторозрядна арифметика по-перше дозволяє реалізувати адаптивні алгоритми

ідентифікації властивостей задач, що неможливо у фіксованому комп'ютерному середовищі, по-друге, забезпечує обчислення комп'ютерного розв'язку з заданою точністю .

### Література

- [1] Молчанов. И. Н. Машинные методы решения прикладных задач. Алгебра, приближение функций, обыкновенные дифференциальные уравнения. – Киев: Наук. думка, 2007. – 550 с.
- [2] Химич А.Н., Молчанов И.Н., Попов О.В. и др. Параллельные алгоритмы решения задач вычислительной математики. – Київ: Наук. думка, 2008. – 247 с.
- [3] Khimich O. M., Chistyakova T.V, Sidoruk V.A., Yershov P.S. Adaptive computer technologies for solving problems of computational and applied mathematics. Cybernetics and Systems Analysis. 2021. Vol. 57, N 6. P. 990–997. DOI: <https://doi.org/10.1007/s10559-021-00424-z>
- [4] Глушков В.М., Молчанов, Б.Н. Брусникин и др Программное обеспечение ЭВМ МИР-1 и МИР-2. Численные методы - Киев: Наук. думка, 1976. – Т. 1. – 280 с.
- [5] Bailey, D. H.; Hida, Y.; Li, Y.; Thompson, B. ARPREC: An arbitrary precision computation package, 2002. <http://www.davidhbailey.com/dhbpapers/arpred.pdf>
- [6] Уилкинсон Дж.Х., Дж.Х., Райни К. Справочник алгоритмов на языке АЛГОЛ. Линейная алгебра – М.: Машиностроение, 1976. – 389 с.
- [7] The GNU multiple precision library. <https://gmplib.org>
- [8] Nikolaevskaya, E.A., Chistyakova, T.V. Program-algorithmic methods to improve the accuracy of computer solutions. *Cybern Syst Anal* **45**, 1004–1007 (2009). <https://doi.org/10.1007/s10559-009-9164-7>
- [9] В. К. Задірака, А. М. Терещенко Комп'ютерна арифметика багаторозрядних чисел у послідовній та паралельній моделях обчислень. – Київ: Наукова думка, 2021. – 135 с.

## Peculiarities of using multiple precision arithmetic in mathematical modeling

Oleksandr Khimich, Olena Nikolaeivska, Igor Baranov

*In the paper the peculiarities of using multi-bit arithmetic in mathematical modeling are investigated. An experimental study of means of improving the accuracy of computer results of solving poorly ill-conditioned systems of linear algebraic equations using the GMP library is described to improve the accuracy of calculations.*

Отримано 31.03.23