

Застосування паралельних обчислень при наближеному розв'язуванні однобічних матричних рівнянь

Анастасія Недашковська¹, Іван Гусак²

¹к. фіз.-мат. наук, доцент кафедри обчислювальної математики Львівського національного університету імені Івана Франка, м. Львів, вул. Університетська, 1, anastasiya.nedashkovska@lnu.edu.ua

²студент кафедри обчислювальної математики Львівського національного університету імені Івана Франка, м. Львів, вул. Університетська, 1, ivan.husak@lnu.edu.ua

Найчастіше матричні рівняння використовуються у задачах оптимізації систем керування. Переважно це рівняння Рікатті та Сильвестра, однак зустрічаються й інші типи матричних рівнянь. Оскільки загального підходу до розв'язування задач такого типу не існує, то розробка схем наближеного розв'язування є актуальною. У одній із попередніх публікацій було узагальнено метод Хованського для побудови ітераційної схеми розв'язування однобічних матричних рівнянь вигляду $X^n A_n + X^{n-1} A_{n-1} + \dots + X^2 A_2 + X A_1 + A_0 = 0$. У публікації наведена рекурентна формула для обчислення наближеного розв'язку рівнянь у вигляді ланцюгового матричного дробу, досліджено збіжність запропонованого методу. У даній роботі запропонована модифікація розглянутої раніше схеми. Застосовані паралельні алгоритми множення та додавання матриць, а також знаходження оберненої матриці.

Ключові слова: ітераційний метод, матричні рівняння, однобічні поліноміальні матричні рівняння, метод Хованського.

Вступ. У задачах оптимізації систем керування [4], [5] часто виникає потреба у розв'язуванні матричних рівнянь. Однак загального підходу до розв'язування задач цього класу не існує.

У роботі [7] було узагальнено метод Хованського [6] для побудови ітераційної схеми розв'язування поліноміальних матричних рівнянь. Схожу схему було застосовано до наближеного розв'язування однобічних матричних рівнянь у випадку, коли коефіцієнти рівняння розташовані праворуч [2].

1. Обчислювальна схема методу

Розглянемо однобічне поліноміальне матричне рівняння вигляду

$$X^n A_n + X^{n-1} A_{n-1} + X^{n-2} A_{n-2} + \dots + X^2 A_2 + X A_1 + A_0 = 0. \quad (1)$$

Тут матриці $A_n, A_{n-1}, A_{n-2}, \dots, A_0 \in R^{m \times m}$ – задані коефіцієнти, $X \in R^{m \times m}$ – невідомий розв'язок.

У роботі [2] була запропонована схема наближеного розв'язування рівняння (1). Розглянемо її коротко.

Припустимо, що X – невироджена $m \times m$ матриця. Введемо позначення $Y_0 = X^{-1}$. Помножимо рівність (1) зліва на X^{-1} і отримаємо:

$$X^{n-1}A_n + X^{n-2}A_{n-1} + X^{n-3}A_{n-2} + \dots + XA_2 + A_1 + Y_0A_0 = 0. \quad (2)$$

Далі введемо позначення $Y_1 = X^{-1}Y_0 = (X^{-1})^2$ і помножимо рівняння (2) зліва на X^{-1} :

$$X^{n-2}A_n + X^{n-3}A_{n-1} + X^{n-4}A_{n-2} + \dots + A_2 + Y_0A_1 + Y_1A_0 = 0.$$

Після $(n-2)$ множень зліва на X^{-1} рівняння (1) отримаємо

$$X^2A_n + XA_{n-1} + A_{n-2} + Y_0A_{n-3} + \dots + Y_{n-5}A_2 + Y_{n-4}A_1 + Y_{n-3}A_0 = 0, \quad (3)$$

де

$$Y_0 = (X^{-1})^1, Y_1 = (X^{-1})^2, \dots, Y_{n-3} = (X^{-1})^{n-2}.$$

Введемо параметри, деякі невироджені матриці $L, K \in R^{m \times m}$. Легко бачити, що рівняння (3) еквівалентне

$$X^2A_nL + XA_{n-1}L + XK - XK + A_{n-2}L + Y_0A_{n-3}L + \dots + Y_{n-5}A_2L + Y_{n-4}A_1L + Y_{n-3}A_0L = 0.$$

Звідси, припускаючи, що $\det(XA_nL + K) \neq 0$, отримуємо:

$$X = [X(K - A_{n-1}L) - A_{n-2}L - \dots - Y_{n-4}A_1L - Y_{n-3}A_0L](XA_nL + K)^{-1}. \quad (4)$$

Можна показати [2], що

$$\begin{aligned} Y_0 &= (A_nL + Y_0K)(XA_nL + K)^{-1}; \\ Y_1 &= (Y_0A_nL + Y_1K)(XA_nL + K)^{-1}; \\ &\vdots \\ Y_{n-3} &= (Y_{n-2}A_nL + Y_{n-3}K)(XA_nL + K)^{-1}. \end{aligned} \quad (5)$$

Далі, використовуючи формули (4) і (5), отримуємо алгоритм для обчислення розв'язків поліноміального матричного рівняння (1):

1. Задати похибку $\varepsilon > 0$;
2. Задати початкове наближення, невироджену матрицю $X^{(0)} \in R^{m \times m}$;
3. Встановити лічильник $n = 1$;
4. Обчислити

$$Y_0^{(0)} = \left((X^{(0)})^{-1} \right)^1, Y_1^{(0)} = \left((X^{(0)})^{-1} \right)^2, \dots, Y_{n-3}^{(0)} = \left((X^{(0)})^{-1} \right)^{n-2};$$

5. Обчислити

$$\begin{aligned} Y_0^{(n)} &= (A_n L + Y_0^{(n-1)} K) (X^{(n-1)} A_n L + K)^{-1}; \\ Y_1^{(n)} &= (Y_0^{(n)} A_n L + Y_1^{(n-1)} K) (X^{(n-1)} A_n L + K)^{-1}; \\ &\vdots \\ Y_{n-3}^{(n)} &= (Y_{n-2}^{(n)} A_n L + Y_{n-3}^{(n-1)} K) (X^{(n-1)} A_n L + K)^{-1}. \end{aligned} \quad (6)$$

6. Обчислити $\tilde{A}_0 = A_{n-2} + Y_0 A_{n-3} + \dots + Y_{n-5} A_2 + Y_{n-4} A_1 + Y_{n-3} A_0$;

7. Нехай $\tilde{A}_2 = A_n$, $\tilde{A}_1 = A_{n-1}$, тоді

$$X^{(n)} = (X^{(n-1)} K - \tilde{A}_0 L) (X^{(n-1)} \tilde{A}_2 L + \tilde{A}_1 L + K)^{-1}. \quad (7)$$

Перевірити умову $\|X^{(n)} - X^{(n-1)}\| < \varepsilon$. Якщо умова не задовольняється, встановити лічильник $n = n + 1$ і перейти на крок 5, інакше повернути $X^{(n)}$.

2. Про збіжність алгоритму

Із використанням узагальненої в монографії [1] достатньої умови збіжності Ворпійського, у роботі [2] показано, що достатньою умовою збіжності ітераційного процесу (7) буде виконання нерівності

$$\left\| \left(\tilde{A}_1 \tilde{A}_2^{-1} + \frac{k}{l} \tilde{A}_2^{-1} + E \right)^{-2} \left(\frac{k}{l} \tilde{A}_1 \tilde{A}_2^{-2} + \frac{k^2}{l^2} \tilde{A}_2^{-2} + \tilde{A}_0 \tilde{A}_2^{-1} \right) \right\| \leq \frac{1}{4},$$

де коефіцієнти

$$\tilde{A}_0 = A_{n-2} + Y_0 A_{n-3} + \dots + Y_{n-5} A_2 + Y_{n-4} A_1 + Y_{n-3} A_0, \dots, \tilde{A}_2 = A_n, \tilde{A}_1 = A_{n-1}.$$

3. Деякі критерії для оцінки паралельних алгоритмів

Для оцінки якості паралельних алгоритмів використовуються такі критерії [3], як коефіцієнт прискорення $S_p = \frac{T_1}{T_p}$ і коефіцієнт ефективності $E_p = \frac{S_p}{p}$, де p – кількість процесів, що використовуються для розв'язування задачі на MIMD-комп'ютері; T_p – час розв'язування задачі на MIMD-комп'ютері із використанням p процесів; T_1 – час розв'язування тієї ж задачі на гіпотетичному однопроцесорному комп'ютері із швидкодією одного процесора і оперативною пам'яттю, що рівна сумарній пам'яті, що використовується p процесами.

Нехай t – середній час виконання однієї операції; t_o – час, необхідний для обміну одним машинним словом між двома процесами; t_{36} – час, необхідний для встановлення зв'язків між двома процесами. Окрім того введемо наступні позначення:

$$\tau_o = \frac{t_o}{t}; \tau_{36} = \frac{t_{36}}{t}.$$

Тоді

$$T_1 = O_1 t, \text{ а } T_p = (O_p + O_{36} \tau_{36} + O_o \tau_o) t,$$

де O_1 – кількість операцій необхідна для розв’язування задачі однопроцесорним комп’ютером, O_p – кількість операцій, що виконується кожним із p процесів, O_{36} – кількість взаємодій одного із p процесів із іншими, O_o – обмінів одним машинним словом для процесу, що розглядається.

4. Паралельна модель алгоритму наближеного розв’язування однобічних рівнянь

Ітераційний процес (7) було реалізовано у Visual Studio 2022. При цьому було використано систему MIMD (множинний потік команд, множинний потік даних) організації архітектури обчислювального процесу.

При реалізації запропонованого алгоритму основний об’єм роботи виконується на етапі обчислення кроків 4-7. Для множення матриць на кроках 4-7 використовувався паралельний алгоритм скалярних добутків, а при знаходженні обернених матриць на кроках 4,5 та 7 ijk – форма LU розкладу [3].

Відомо [3], що загальна кількість арифметичних операцій необхідна для знаходження оберненої матриці розмірності $m \times m$ оцінюється величиною $O_{1,об} = \frac{2}{3}m^3 + O(m^2)$. Відповідно кількість операцій, що виконує один із p процесорів при знаходженні оберненої матриці буде оцінюватись, як $O_{p,об} = \frac{2m^3}{3p} + O(\frac{m^2}{p})$ [3]. Для виконання множення двох матриць розмірності $m \times m$ необхідно $O_{1,мн} = m^3 + O(m^2)$ арифметичних операцій множення та додавання. А, отже, кількість операцій, що виконує один із p процесорів при знаходженні дабутку матриць, наприклад, методом скалярних добутків буде оцінюватись, як $O_{p,мн} = \frac{m^3}{p} + O(\frac{m^2}{p})$. Кількість арифметичних операцій, що використовується при додаванні двох матриць розмірності $m \times m$ необхідно $O_{1,дод} = m^2$. Тоді для одного із p процесів така кількість становитиме $O_{p,дод} = \frac{m^2}{p}$. Також будемо вважати, що пошук максимуму для знаходження норми матриці $\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^m |a_{ij}|$ буде виконуватися з допомогою найшвидшого алгоритму і потребуватиме m арифметичних операцій. Тож для знаходження норми потрібно $O_{1,норм} = O(m)$ арифметичних операцій. При використанні p процесів для кожного із них кількість арифметичних операцій теж буде становити $O_{p,норм} = O(m)$.

Провівши деякі нескладні обчислення, отримуємо:

$$T_1 = O\left((4n+1)m^3 + (2n+1)m^2\right)t$$

та

$$T_p = \left(O\left(\frac{(4n+1)m^3 + (2n+1)m^2}{p}\right) + O_{36}\tau_{36} + O_o\tau_o\right)t.$$

А отже

$$S_p = \frac{T_1}{T_p} \approx p, \quad E_p = \frac{S_p}{p} \approx 1.$$

Таким чином, для даного алгоритму на обчислювальній системі типу MIMD вдається досягти практично тих же оцінок S_p та E_p , що і для концепції необмеженого паралелізму, що свідчить про доцільність використання такого підходу при наближеному розв'язуванні рівняння (1).

Висновки. У роботі запропонована модифікація розглянутої раніше схеми розв'язування однобічних матричних рівнянь. Застосовані паралельні алгоритми множення та додавання матриць, а також знаходження оберненої матриці. Обчислені коефіцієнти прискорення та ефективності отриманого паралельного алгоритму. Встановлено, що для даної схеми на обчислювальній системі типу MIMD вдається досягти оцінок прискорення і ефективності близьких до оцінок при концепції необмеженого паралелізму, що свідчить про доцільність використання такого підходу при наближеному розв'язуванні однобічного матричного рівняння.

Література

- [1] Боднар Д.И. Ветвящиеся цепные дроби. – К.: Наукова думка, 1986. – 176 с.
- [2] Недашковська А. Узагальнення методу Хованського для наближеного розв'язування однобічних поліноміальних матричних рівнянь/ Анастасія Недашковська // Вісн. Львів. ун-ту ім. І.Франка. Сер. прикл. матем. та інформ. – 2022. – Вип. 30. – С. 60–68.
- [3] Химич А.Н., Молчанов И.Н., Попов А.В., Чистякова Т.В., Яковлев М.Ф. Параллельные алгоритмы решения задач вычислительной математики: монография. – Київ: Наук.думка, 2008. – 248 с.
- [4] Beavers A.N., Denman E.D. A New Solution Method for the Lyapunov Matrix Equation// SIAM Journal on Applied Mathematics. – Vol. 29, No. 3, – 1975. – P. 416–421.
- [5] Boichuk A.A., Krivosheya S.A. Criterion of solvability of matrix equations of Lyapunov type// Ukrainian Mathematical Journal, 1998. – 50, №8. – P.1162–1169.
- [6] Khovanskii A.N. The Application of continued fractions and their generalizations to problems in approximation theory. P. Noordhoff, 1963. – 212 p.
- [7] Nedashkovska A. Generalization of the Khovanskii's method for solving matrix polynomial equations// Журнал обчислювальної та прикладної математики. – 2015. – № 2. – С. 42–49.

Using of parallel calculations in the approximate solving of one-sided matrix equations

Anastasija Nedashkovska, Ivan Gusak

Matrix equations are widely used in optimization problems of control systems. Most often, these are the Ricatti and Sylvester equations, but there are other types of matrix equations. There is no general approach to solving problems of this type, so the development of approximate solution schemes is actual. In one of the previous publications, Khovanskiy's method was generalized. An iterative scheme for solving one-sided matrix equations of the form $X^n A_n + X^{n-1} A_{n-1} + \dots + X^2 A_2 + X A_1 + A_0 = 0$ is obtained. The publication presented a recurrent formula for calculating an approximate solution of equations in the form of a continued matrix fraction. Also, the convergence of the proposed method was investigated. In this article, a modification of the previously considered scheme is proposed. Parallel algorithms are applied for matrix multiplication, matrix addition operations and for finding the inverse matrix.

Отримано 31.03.23