УДК 004.272:519.61

Сухарський С. С., аспірант,
Малашонок Г. І., д.ф-м.н., професор
Інститут Програмних Систем НАНУ
Національний Університет «Києво-Могилянська академія»

ТРИДІАГОНАЛІЗАЦІЯ СТРІЧКОВИХ СИМЕТРИЧНИХ МАТРИЦЬ

Задача зведення симетричної (ермітової) матриці до тридіагональної форми є однією з базових задач чисельної лінійної алгебри, що лежить в основі алгоритмів обчислення власних значень і власних векторів [1, 2]. Попереднє приведення матриці до тридіагонального вигляду значною мірою визначає ефективність подальших обчислень. У класичних підходах для розв'язання цієї задачі використовуються перетворення Хаусхолдера, які характеризуються високою чисельною стійкістю, але мають значну кількість залежностей між операціями та низький рівень паралелізму на сучасних гібридних кластерних вузлах. Основним обмежуючим фактором у таких системах є не кількість арифметичних операцій, а обсяг і структура обміну даними, що обумовлює необхідність розробки алгоритмів із низькою комунікаційною складністю [3].

У роботі запропоновано алгоритм Ribbon для зведення стрічкових симетричних матриць до тридіагональної форми. Алгоритм застосовується до стрічкових матриць і є етапом переходу від стрічкової до тридіагональної форми у загальному обчислювальному процесі. Алгоритм базується на ідеї послідовного зменшення ширини стрічки матриці вдвічі на кожному кроці з використанням блоково-рекурсивних QR та QP розкладів [4]. Обчислення організовано у вигляді блокової структури, що дозволяє звести основні операції до матричних множень. У результаті формується унітарний розклад вигляду $A = U^T T U$, де T — тридіагональна матриця, а U — ортогональна (унітарна) матриця, яка накопичується у процесі виконання алгоритму. Ключовою особливістю алгоритму є його природна придатність до паралельної реалізації. Обчислення організовано у вигляді дворівневого конвеєра. На першому рівні паралелізм досягається за рахунок незалежної обробки груп блоків (паралелограмів) в межах одного рівня стрічки. На другому рівні забезпечується міжмасштабний паралелізм: обчислення для наступного рівня з меншою шириною стрічки можуть розпочинатися до повного завершення попереднього рівня.

Асимптотичний аналіз показує, що складність обчислення тридіагональної форми в алгоритмі Ribbon має той самий порядок, що й матричне множення, тобто $O(N^{\beta})$, де β визначається алгоритмом множення матриць. Це дозволяє розглядати запропонований підхід як ефективну альтернативу класичним методам у контексті високопродуктивних обчислень. Водночас встановлено, що побудова унітарного множника U має кубічну складність $O(N^3)$ та є домінуючим етапом на великих розмірах матриць. При цьому для повної факторизації спостерігається збільшення константи при кубічному члені порівняно з методом Хаусхолдера, що відображає компроміс між зменшенням комунікаційних витрат і обчислювальною складністю.

Програмну реалізацію алгоритму виконано у середовищі блоково-рекурсивних паралельних обчислень DAP із використанням MPI. Обчислювальний процес представлено у вигляді графа задач, що дозволяє ефективно реалізувати паралельну обробку та масштабування. Експериментальні дослідження проведено системі з використанням CPU та GPU. Отримані результати демонструють узгодження з теоретичними оцінками: час виконання алгоритму зростає кубічно зі збільшенням розміру матриці, а абсолютна похибка реконструкції становить порядку 10^{-13}, що відповідає машинній точності для подвійної арифметики. Показано, що зі зростанням розміру задачі частка часу, витрачена на побудову матриці U, зростає та визначає загальну продуктивність алгоритму.

Таким чином, запропонований алгоритм Ribbon забезпечує поєднання блоково-рекурсивної структури, низької комунікаційної складності та високого рівня паралелізму. Це робить його перспективним для використання у високопродуктивних обчисленнях, зокрема на гібридних CPU–GPU системах і багатовузлових кластерах. Подальші дослідження можуть бути спрямовані на оптимізацію етапу побудови унітарного множника та підвищення ефективності реалізації на розподілених системах.

Список використаних джерел:

1. Golub G. H., Van Loan C. F. Matrix Computations. 4th ed. Baltimore: Johns Hopkins University Press, 2013. 756 p.

2. Trefethen L. N., Bau D. Numerical Linear Algebra. Philadelphia: SIAM, 1997.

3. Ballard G., Demmel J., Holtz O., Schwartz O. Minimizing communication in numerical linear algebra // SIAM Journal on Matrix Analysis and Applications. 2011. Vol. 32, No. 3. P. 866–901.

4. Малашонок Г., Сухарський С. Блоково-рекурсивний підхід до унітарного розкладу матриць // Проблеми програмування. 2026. № 1 (прийнято до друку).