

О.М. ХІМІЧІнститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: *khimich505@gmail.com*.**О.В. ПОПОВ**Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: *alex50popov@gmail.com*.

РОЗВ'ЯЗУВАННЯ НЕКОРЕКТНИХ ЗАДАЧ ТЕОРІЇ ПРУЖНОСТІ НА ВИСОКОПРОДУКТИВНИХ ОБЧИСЛЮВАЛЬНИХ СИСТЕМАХ

Анотація. Запропоновано методику ефективного дослідження і розв'язування умовно коректних задач, які мають єдиний розв'язок на підпросторі. Обґрунтовано використання дискретної скінченно-елементної моделі на всьому просторі для отримання єдиного розв'язку на підпросторі вихідної варіаційної задачі. Для знаходження нормального псевдорозв'язку дискретної задачі (системи лінійних алгебраїчних рівнянь з розрідженою симетричною напіввизначеною матрицею) запропоновано метод триетапної регуляризації. Цей метод дає змогу отримати наближення до цих розв'язків з наперед заданою точністю. Розроблено ефективні адаптивні високопродуктивні алгоритми зазначеного методу для розв'язування на сучасних комп'ютерах з паралельною організацією обчислень систем лінійних алгебраїчних рівнянь з розрідженими симетричними напіввизначеними матрицями.

Ключові слова: високопродуктивні обчислення, змінне комп'ютерне середовище, метод скінченних елементів, метод триетапної регуляризації, перша основна задача теорії пружності.

ВСТУП

Математичне моделювання з використанням комп'ютерів (комп'ютерне моделювання) є універсальною, інваріантною стосовно предметної галузі методологією дослідження нових процесів, об'єктів та явищ засобами обчислювального експерименту і відіграє значну роль у наукових дослідженнях для потреб розвитку різних галузей науки та інженерії, зокрема, для потреб національної безпеки і оборони України. У процесі вибору чисельних методів математичного моделювання важливим є отримання достовірних розв'язків з оцінками точності результатів.

У багатьох випадках основою для дослідження математичної моделі є розв'язання задач лінійної алгебри, зокрема, систем лінійних алгебраїчних рівнянь (СЛАР). Останні виникають під час дискретизації крайових задач проєкційно-різницею методом (скінченних елементів, скінченних різниць тощо (див. наприклад, [1–3])). Важливою особливістю СЛАР, що виникають у процесі математичного моделювання в різних галузях, є розріджена структура та високі порядки (можуть перевищувати 10^6) їхніх матриць. З урахуванням високих порядків цих СЛАР і, як наслідок, великої кількості арифметичних операцій для дослідження та розв'язування таких задач доцільно розробляти і використовувати ефективні високопродуктивні алгоритми для комп'ютерів з паралельною організацією обчислень різної архітектури, зокрема, гібридної.

В Інституті кібернетики наприкінці 70-х – початку 80-х років минулого сторіччя розпочались роботи зі створення багатопроекторних суперкомп'ютерів на основі ідеї академіка В.М. Глушкова. У результаті було побудовано макроконвеєрний обчислювальний комплекс (МОК) ЕС-1766 — багатопроекторна обчислювальна система з розподіленою пам'яттю. Під час розроблення паралельних алгоритмів для розв'язування задач математичного моделювання було запропоновано декілька математичних інновацій, зокрема, циклічна схема роз-

поділу, збереження та оброблення даних на розподілених обчислювальних системах. Вона дала змогу позбутися обмеження на ефективність унаслідок ефекту Гайдна та була використана в світовій практиці розроблення збалансованих паралельних обчислювальних алгоритмів.

Натепер в Інституті кібернетики НАН України створено інфраструктуру для комп'ютерного моделювання на основі високопродуктивних обчислень — Центр колективного користування обладнанням суперкомп'ютерного комплексу СКІТ НАН України. Сьогодні СКІТ — одна із найпотужніших кластерних систем в Україні (його сегмент СКІТ-5 — 8-вузловий кластер з 192-ядерними вузлами на AMD EPYC 7642 — має пікову продуктивність понад 100 ТФлопс).

Інший напрям розроблень засобів високопродуктивних обчислень — комп'ютери середньої ланки, які мають заповнити нішу між персональними комп'ютерами з порівняно невеликим обчислювальним ресурсом та суперкомп'ютерами. Вченими Інституту кібернетики спільно з колективом ДП «Електронмаш» розроблено концепцію та створено сім'ю інтелектуальних паралельних комп'ютерів Інпарком різноманітних архітектур. Ці комп'ютери призначені для дослідження і розв'язування задач із різних галузей науки та інженерії. На відміну від традиційних інтелектуальних комп'ютерів забезпечують дослідження властивостей задач та адаптивне автоматичне налаштування алгоритмів, програм та архітектури комп'ютера на властивості задач.

Під час дослідження математичних властивостей та розв'язування розрахункових задач математичного моделювання використання змінної розрядності дає змогу підвищити ефективність алгоритмів, зберігаючи достовірність результатів. А засоби багаторозрядної арифметики (апаратні або програмні) надають можливість ідентифікувати в змінному комп'ютерному середовищі вироджену чи погано обумовлену матрицю, близькі чи кратні власні значення, що не можна дослідити в фіксованому комп'ютерному середовищі. Арифметика довільної розрядності була характерною особливістю створеної в 60–70-і роки минулого сторіччя в Інституті кібернетики під керівництвом академіка В.М. Глушкова серії машин для інженерних розрахунків «МИР».

Під час чисельного розв'язання некоректних задач, наприклад у процесі аналізу міцності незакріплених конструкцій — першої основної задачі теорії пружності [4], матриця відповідної розрахункової задачі (СЛАР) виявляється напіввизначеною. Такі СЛАР з напіввизначеною матрицею мають єдиний розв'язок на підпросторі (нормальний псевдорозв'язок), який можна знайти методом найменших квадратів [4, 5].

У цій роботі для першої основної задачі теорії пружності обґрунтовано методику дослідження та розрахунку математичних моделей, що описуються задачами з єдиним розв'язком на підпросторі (умовно коректних задач), з використанням для цього нормального псевдорозв'язку СЛАР дискретної скінченно-елементної моделі. Для знаходження нормального псевдорозв'язку використано паралельні реалізації методу триетапної регуляризації [6–11].

ПЕРША ОСНОВНА ЗАДАЧА ТЕОРІЇ ПРУЖНОСТІ

Неперервна модель. Нехай $\Omega \subset \mathbb{R}^3$ — скінченна область з достатньо гладкою границею Γ , $U_0(\Omega) = (W_2^1(\Omega))^3$ — простір Соболева вектор-функцій $w = (w_1, w_2, w_3)^T$, $u \in U_0(\Omega)$ — вектор-функція пружних зміщень. Позначимо $\varepsilon = (\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, \varepsilon_{12}, \varepsilon_{23}, \varepsilon_{31})^T$, $\sigma = (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{23}, \sigma_{31})^T$ вектори деформацій і напружень відповідно, $f_\Omega = (f_{\Omega,1}, f_{\Omega,2}, f_{\Omega,3})^T \in (L_2(\Omega))^3$, $f_\Gamma = (f_{\Gamma,1}, f_{\Gamma,2}, f_{\Gamma,3})^T \in (L_2(\Gamma))^3$ — вектори розподілених об'ємних і поверхневих сил відповідно ($\varepsilon_{lm}(u) = \varepsilon_{ml}(u)$, $\sigma_{ik}(u) = \sigma_{ki}(u)$).

Розглядається випадок, коли виконується закон Гука: $\sigma = C(x)\varepsilon$, причому матриця 6-го порядку $C(x)$ залежить від поля температур і не залежить від часу, а коефіцієнти пружності $c_{jm}(x) = c_{mj}(x)$, $j, m = 1, \dots, 6$, задовольняють умову

$$\varepsilon^T \sigma = \varepsilon^T C \varepsilon \geq \mu_0 \varepsilon^T \varepsilon, \quad \mu_0 = \text{const} > 0. \quad (1)$$

Для вектор-функцій $u, v \in U_0(\Omega)$ визначимо білінійну і лінійну форми

$$a(u, v) = \int_{\Omega} (\varepsilon(u))^T C(x) \varepsilon(v) dx, \quad l(f, v) = \int_{\Omega} f_{\Omega}^T v dx + \int_{\Gamma} f_{\Gamma}^T v d\gamma.$$

Розглянемо у $\overline{\Omega} = \Omega \cup \Gamma$ крайову задачу теорії пружності [12] у слабкій постановці: знайти такий вектор зміщень $u \in U_0(\Omega)$, щоб виконувалася інтегральна тотожність

$$a(u, v) = l(f, v) \quad \forall v \in U_0(\Omega). \quad (2)$$

Якщо виконується умова

$$\int_{\Omega} f_{\Omega} dx + \int_{\Gamma} f_{\Gamma} d\gamma = 0, \quad \int_{\Omega} (r \times f_{\Omega}) dx + \int_{\Gamma} (r \times f_{\Gamma}) d\gamma = 0, \quad (3)$$

то розв'язок задачі (1)–(3) існує (тут і далі r — радіус-вектор точки в області Ω , а знаком $\times$ позначено векторний добуток).

У цій постановці розв'язок задачі щодо зміщень неєдиний (хоча компоненти вектора напружень можуть бути визначені однозначно), оскільки можливе переміщення тіла (поступальний перенос і жорсткий поворот) як твердого. Єдиний розв'язок виокремлюється двома умовами:

$$\int_{\Omega} u dx = 0, \quad \int_{\Omega} (r \times u) dx = 0. \quad (4)$$

Умови (4) визначають підпростір, в якому шукається єдиний розв'язок

$$V_0(\Omega) = (V_2^1(\Omega))^3 = \left\{ v \mid v \in U_0(\Omega), \int_{\Omega} v dx = 0, \int_{\Omega} r \times v dx = 0 \right\} \subset U_0(\Omega).$$

Варіаційна задача (2) має єдиний розв'язок на підпросторі $V_0(\Omega)$ [13], що випливає з доведених у [4] додатної визначеності функціонала $a(u, u) \geq M_0 \|u\|_{0, \Omega}^2 \quad \forall u \in (L_2(\Omega))^3$ та його V -еліптичності $a(u, u) \geq M_1 \|u\|_{1, \Omega}^2 \quad \forall u \in V_0(\Omega)$ на лінеалі вектор-функцій, що задовольняють умови (4) (сталі M_0 і M_1 не залежать від u , а норми визначаються формулами $\|u\|_{0, \Omega}^2 \equiv \|u\|_{(L_2(\Omega))^3}^2$, $\|u\|_{k, \Omega}^2 \equiv \|u\|_{(W_2^k(\Omega))^3}^2$).

У [14] доведено коефіцієнтну стійкість першої основної крайової задачі теорії пружності, тобто справедливості оцінки

$$\|\tilde{u} - u\|_{1, \Omega}^2 \leq M_2 (\delta (\|f_{\Omega}\|_{0, \Omega} + \|f_{\Gamma}\|_{0, \Gamma}) + \|\tilde{f}_{\Omega} - f_{\Omega}\|_{0, \Omega} + \|\tilde{f}_{\Gamma} - f_{\Gamma}\|_{0, \Gamma}),$$

де M_2 — незалежна від u та $\tilde{u}$ стала, $\tilde{u}$ — єдиний розв'язок (тобто $\int_{\Omega} \tilde{u} dx = 0$, $\int_{\Omega} (r \times \tilde{u}) dx = 0$) задачі $\tilde{a}(\tilde{u}, v) = l(\tilde{f}, v) \quad \forall v \in V_0(\Omega)$ зі збуреними матрицею коефіцієнтів $\tilde{C}(x)$ та векторами $\tilde{f}_{\Omega}(x)$ і $\tilde{f}_{\Gamma}(x)$ розподілених об'ємних та поверхневих сил, $\delta = \max_{i, k} \max_x |c_{ik}(x) - \tilde{c}_{ik}(x)|$.

Отже, перша основна задача теорії пружності (1)–(3) є умовно коректною: її розв'язки існують на множині $U_0(\Omega)$, а єдиний розв'язок визначається в $V_0(\Omega)$.

Дискретна модель. Для чисельного розв'язання першої основної задачі теорії пружності використаємо метод скінченних елементів (МСЕ) (див., на-

приклад, [1–3, 15]), який за своєю суттю полягає в заміні нескінченновимірному простору допустимих функцій його скінченновимірним підпростором.

Нехай $V^h \subset W^h$ — скінченновимірні простори МСЕ, а саме $W^h \subset U_0(\Omega)$ — простір МСЕ з базисом $\{\psi_j\}, j = 1, 2, \dots, N$, а $V^h \subset V_0(\Omega)$ — простір МСЕ з базисом $\{\varphi_i\}, i = 1, \dots, N - n_0$ (n_0 — розмірність нуль-простору оператора вихідної задачі). Ці базиси зазвичай складають вектор-функції, компоненти яких є кусковими поліномами, відмінними від нуля лише на кількох скінченних елементах.

Розглянемо абстрактну дискретну задачу МСЕ: знайти таку функцію $u^h \in V^h$, що

$$a(u^h, v^h) = l(f, v^h) \quad \forall v^h \in V^h. \quad (5)$$

Існування і єдиність розв'язку дискретної задачі (5) впливає з леми Лакса–Мільграма [15], а збіжність і точність розв'язку впливає із загальної теорії МСЕ, оскільки на підпросторі V^h виконано всі потрібні для цього умови.

Однак умовна коректність першої основної задачі теорії пружності призводить до певних труднощів під час її чисельного розв'язання — для побудови скінченно-елементного базису $\{\varphi_i\}$ простору V^h потрібне знання нуль-простору оператора вихідної задачі в явному вигляді. Цей нуль-простір визначається умовами (4) і є натягнутим на n_0 вектор-функцій. Для тривимірних моделей це $u_{0,1} = (1, 0, 0)^T$, $u_{0,2} = (0, 1, 0)^T$, $u_{0,3} = (0, 0, 1)^T$, $u_{0,4} = (0, -x_3, x_2)^T$, $u_{0,5} = (-x_2, x_1, 0)^T$, $u_{0,6} = (0, -x_3, x_2)^T$. Однак алгоритм побудови базису простору V^h у загальному випадку для задач з неєдиним розв'язком на всьому просторі наразі невідомий.

Розглянемо також таку дискретну задачу:

$$a(u^h, v^h) = l(f, v^h) \quad \forall v^h \in W^h. \quad (6)$$

Її розв'язок неєдиний. Покладаючи $u^h = \sum_{i=1}^N y_i \psi_i$, з (6) отримуємо СЛАР

$$Ay = b \quad (7)$$

з розрідженою додатно напіввизначеною матрицею A . Позначимо $y^\#$ нормальний псевдорозв'язок СЛАР (7) та визначимо функцію $u_\#^h = \sum_{i=1}^N y_i^\# \psi_i$.

Для функції $u_\#^h$ справедлива наведена далі теорема (див. [8, 9]).

Теорема 1. Якщо $y^\#$ — нормальний псевдорозв'язок СЛАР (7) з розрідженою напіввизначеною матрицею, а $\{\psi_i\}$ ($i = 1, 2, \dots, N$) — базис скінченновимірному простору $W^h \subset (W_2^1(\Omega))^3$, то функція $u_\#^h = \sum_{i=1}^N y_i^\# \psi_i$ буде розв'язком задачі (5).

РОЗВ'ЯЗУВАННЯ СЛАР З МАТРИЦЯМИ НЕПОВНОГО РАНГУ

У роботах [6, 7] для отримання з гарантованою точністю наближення до нормального псевдорозв'язку СЛАР (7) з розрідженою симетричною напіввизначеною матрицею запропоновано метод триетапної регуляризації.

Розглянемо задачу знаходження єдиного нормального псевдорозв'язку СЛАР з розрідженою симетричною додатно напіввизначеною матрицею порядку n і рангу $k = n - n_0$

$$Ax = b, \quad A = A^T, \quad A \geq 0. \quad (8)$$

Розглянемо також СЛАР зі збуреною правою частиною $\tilde{b} = b + \Delta b$:

$$A\tilde{x} = \tilde{b}, \quad \|\Delta b\| \leq \varepsilon_b \|b\| \leq \delta. \quad (9)$$

Позначимо $x, \tilde{x}$ нормальні псевдорозв'язки задач (8) і (9) відповідно, $A^\#$ — псевдообернену матрицю, I_n — одиничну матрицю порядку n , $h(A) = \|A\| \|A^\#\|$ — число обумовленості матриці A , $b_k = Ax \in \text{Im } A$, $\tilde{b}_k = A\tilde{x} \in \text{Im } A$, $\varepsilon_b = \|\tilde{b}_k - b_k\| / \|b_k\|$. Для похибки нормального псевдорозв'язку задачі (9) має місце оцінка [16, 17]

$$\|x - \tilde{x}\| / \|x\| \leq h(A) \varepsilon_b,$$

з якої випливає неперервна залежність розв'язку від збурення правої частини.

Метод триетапної регуляризації. Для отримання нормального псевдорозв'язку задачі (8) з розрідженою матрицею зазвичай використовують різні методи регуляризації. Для розв'язання задачі найменших квадратів у [18] запропоновано метод двоетапної регуляризації, а в [19] — метод дискретної регуляризації. У [6–8] запропоновано метод триетапної регуляризації, в якому розв'язана проблема вибору параметра регуляризації і який дає змогу обчислити наближення до нормального псевдорозв'язку з наперед заданою точністю.

Наближення u до нормального псевдорозв'язку x задачі (8) отримуємо, розв'язуючи за довільно вибраного параметра $\alpha > 0$ (наприклад, $\alpha = 0.01$) такі дві СЛАР:

$$(A + \alpha I_n)z = \tilde{b}_k, \quad (A + \alpha I_n)u = Az. \quad (10)$$

Нехай $n_1 = n_0 + 1$, λ_{n_1} — мінімальне відмінне від нуля власне значення матриці A (власні значення пронумеровано в порядку зростання). Тоді, виконавши один крок степеневого методу розв'язування задачі на власні значення, наближення до максимального за модулем власного значення матриці $(A + \alpha I_n)^{-1}$ можна оцінити величиною

$$\mu = \max_i |w_i|, \quad (11)$$

де w — розв'язок системи $(A + \alpha I_n)w = u_H$, $u_H \equiv u / \max_i |u_i|$.

Для похибки наближеного розв'язку u справедлива (див. [6, 5]) така теорема.

Теорема 2. Якщо $h(A)_{\varepsilon_b} < \varepsilon < 1$, то з точністю до величин другого порядку малості для відносної похибки наближеного псевдорозв'язку СЛАР з симетричною додатно напіввизначеною матрицею справедлива оцінка

$$\frac{\|x - u\|}{\|x\|} \leq \alpha \frac{2\lambda_{n_1} + \alpha}{(\lambda_{n_1} + \alpha)^2} + \frac{\lambda_n \lambda_{n_1} \varepsilon_b}{(\lambda_{n_1} + \alpha)^2}. \quad (12)$$

Тоді згідно з (11) і (12) вибір параметра α , що задовольняє умови

$$\alpha \leq (\varepsilon - \|A\| \mu \varepsilon_b) / (2\mu \sqrt{1 - \varepsilon}), \quad \alpha < \lambda_{n_1}, \quad (13)$$

гарантуватиме задану точність ε нормального псевдорозв'язку.

Метод триетапної регуляризації полягає у виконанні таких кроків:

1) послідовне розв'язування двох СЛАР (10) з однією тою самою додатно визначеною матрицею $A + \alpha I_n$;

2) обчислення наближення μ до $(\lambda_{n_1} + \alpha)^{-1}$ згідно з (11);

3) перевірка виконання нерівності $(2\alpha + \|A\| \varepsilon_b) \mu \leq \varepsilon$.

Якщо нерівність виконується, то потрібна точність ε досягнута за вибраного значення $\alpha = \alpha_0$; якщо нерівність не виконується, то з оцінки (13) визна-

чається нове значення зсуву $\alpha = \alpha_1 \leq 0.5(\varepsilon / \mu - \|A\| \varepsilon_b)$ і повторюється крок 1 з матрицею $A + \alpha_1 I_n$. Нове наближення u забезпечує потрібну точність ε нормального псевдорозв'язку.

У [8] запропоновано алгоритм методу триетапної регуляризації для розв'язання задачі зважених найменших квадратів з симетричною додатно визначеною ваговою матрицею.

АДАПТИВНІ ВИСОКОПРОДУКТИВНІ АЛГОРИТМИ ДЛЯ РОЗВ'ЯЗУВАННЯ НЕКОРЕКТНИХ СЛАР З НАБЛИЖЕНИМИ ДАНИМИ РОЗРІДЖЕНОЇ СТРУКТУРИ

Під час розв'язування практичних умовно коректних задач, наприклад аналізу міцності складних незакріплених конструкцій, для отримання достовірних результатів використовують дискретні моделі з великою кількістю степенів свободи (може перевищувати 10^6). Це приводить до високих порядків матриць розрахункових задач — СЛАР з розрідженими напіввизначеними матрицями. Для ефективного розв'язання цих задач доцільно використовувати сучасні високопродуктивні комп'ютери з паралельною організацією обчислень у змінному комп'ютерному середовищі.

Аналіз сучасних високопродуктивних комп'ютерних систем показав, що це переважно гетерогенні комп'ютери, які складаються з різних типів обчислювальних пристроїв і поєднують MIMD- та SIMD-моделі обчислень, використовують розподілену та спільну пам'ять, довільну розрядність тощо. Наприклад, у комп'ютерах гібридної архітектури багатоядерні процесори доповнюються співпроцесорами-прискорювачами (графічними процесорами, GPU) [10, 11, 20–23].

Під час розв'язання розрахункових задач математичного моделювання переважна більшість обчислень проводиться з подвійною розрядністю. Водночас використання одинарної розрядності замість подвійної (або за потреби ще вищої) розрядності дає змогу скоротити час розв'язування різних підзадач. Крім того, використання даних задачі у форматі з одинарною точністю надає змогу зберігати вдвічі більший обсяг даних на кожному рівні ієрархії пам'яті. Застосування змішаної розрядності дає можливість підвищувати ефективність обчислень і скорочувати час для розв'язування задачі, не втрачаючи точності отримуваних результатів [21, 22].

Більшість прямих методів розв'язування СЛАР засновано на ідеї послідовних еквівалентних перетворень заданої системи для виключення невідомих з частини рівнянь і, по суті, на розвиненні матриці A в добуток матриць, наприклад трикутних [9, 10]. У цьому разі в задачі розв'язування СЛАР можна виокремити такі підзадачі: розвинення матриці системи $A = LR$, розв'язування двох СЛАР із трикутними матрицями — з нижньою $Ly = b$ та з верхньою $Rx = y$.

Оскільки розвинення матриці потребує значно більше арифметичних операцій ніж інші макрооперації, ефективність факторизації розріджених матриць відіграє ключову роль у процесі розроблення ефективних паралельних алгоритмів прямих методів розв'язування СЛАР. При цьому декомпозиція та схема розподілу даних кожного алгоритму розвинення визначає декомпозицію та розподіл даних всієї задачі. Тому далі розглянемо алгоритми розвинення розріджених матриць.

Для розвинення дійсних симетричних матриць, врахувавши можливу знакове визначеність вихідної матриці СЛАР (8), доцільно використати LDL^T -варіант методу Холецкого, який зводиться до обчислення розвинення матриці $A = LDL^T$ або $A = U^T D U$, де $L = U^T$ — нижня трикутна матриця, що має одиниці на головній діагоналі, D — діагональна матриця. Розглянемо LDL^T -розвинення блочно-розрідженої матриці A порядку n , яку розбито на блоки розміру $s \times s$ (тут і далі, не втрачаючи загальності міркувань, вважаємо, що $n = Ns$).

Після K ($K=1, \dots, N$) кроків модифікована матриця

$$A^{(K)} = (U^{(K)})^T D^{(K)} U^{(K)},$$

де $U^{(K)} = \begin{pmatrix} U_1^{(K)} & U_2^{(K)} \\ 0 & I_R \end{pmatrix}$, $D^{(K)} = \begin{pmatrix} D_1^{(K)} & 0 \\ 0 & A_R^{(K)} \end{pmatrix}$, $U_1^{(K)}$ — верхня трикутна ($U^{(N)} \equiv U$), а $D_1^{(K)}$ — діагональна ($D^{(N)} \equiv D$) підматриці порядку Ks , I_R — одинична підматриця порядку $R = n - Ks$, $U_2^{(K)}$ — прямокутна підматриця розміру $Ks \times R$, $A_R^{(K)}$ — квадратна підматриця порядку R ($A_R^{(0)} \equiv A$).

Позначимо A_{11} квадратний блок порядку s (провідний блок K -го кроку), A_{12} — прямокутний блок розміру $s \times R$, A_{22} — квадратний блок порядку R (тут і далі опущено номер кроку). На K -му кроці виконується розвинення підматриці

$$A_R^{(K-1)} \equiv \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix} = \begin{pmatrix} U_{11}^T & 0 \\ U_{12}^T & I_r \end{pmatrix} \begin{pmatrix} D_{11} & 0 \\ 0 & A_R^{(K)} \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} \\ 0 & I_r \end{pmatrix}.$$

Обчислення проводяться за такої послідовності: розвинення блока

$$A_{11} = U_{11}^T D_{11} U_{11}; \quad (14)$$

обчислення блоків U_{12} і L_{21}^T (розв'язування матричних СЛАР)

$$U_{11}^T L_{21}^T = A_{12}, \quad D_{11} U_{12} = L_{21}^T; \quad (15)$$

s -рангова модифікація підматриці A_{22} (обчислення підматриці $A_R^{(K)}$):

$$A_R^{(K)} = A_{22} - L_{21} U_{12}. \quad (16)$$

У випадку розрідженої матриці системи A матриці розвинення L або U також залишаються розрідженими, хоча в загальному випадку кількість ненульових елементів збільшується. Тому під час виконання розвинення підматриці $A_R^{(K-1)}$

доцільно проводити обчислення тільки з ненульовими блоками — блок з індексами i та j з підматриці A_{22} модифікується згідно з (16) тільки в тому випадку, коли скалярний добуток i -го рядка блоків з L_{21} та j -го стовпчика блоків з U_{12} не дорівнює тотожно нулю. Тому s -рангова модифікація (16) виконується тільки з підматрицею матриці $L_{21} U_{12}$, яка складається з її ненульових блоків.

Розглянемо адаптивні паралельні блочні алгоритми LDL^T -розвинення розрідженої симетричної матриці. Вибір LDL^T -розвинення зумовлений тим, що таке розвинення може використовуватися для більшої множини матриць, ніж LL^T -розвинення, зокрема, для знаковизначених матриць.

Обчислювальний ресурс поділено на p груп (часто одна група зв'язана з одним вузлом), загальна пам'ять є розподіленою, а пам'ять всередині групи є спільною. На комп'ютерах гібридної архітектури кожна група може використовувати один графічний прискорювач (GPU). У цьому випадку зазвичай група — це одне процесорне ядро (CPU). У кожній групі виокремлюють один процесорний елемент (одне ядро), який є головним у групі. На цьому ядрі використовується один MPI-процес (головний потік). Головні потоки насамперед забезпечують організацію паралельного виконання макрооперацій та комунікації між собою і між процесорними пристроями своєї групи, що паралельно реалізують переважну більшість арифметичних операцій для виконання макрооперацій.

Розглянемо найбільш загальний варіант — розвинення симетричної матриці блочно-хмарочосної структури (блочно-стрічкові та блочно-профільні структури є окремими випадками цієї структури) [3, 10]. Зауважимо, що використовується розріджена структура трикутного розвинення вихідної матриці, яку отримують, виконавши відповідне символічне розвинення.

Розріджену симетричну матрицю A розділено на квадратні блоки порядку s (як і раніше, вважатимемо, що $n = Ns$). Рядки блоків верхнього трикутника матриці циклічно розподіляються між головними потоками, щоб кожен мав ненульові блоки (що модифікуються згідно з (14), (15) на цьому кроці) хоча б одного рядка.

Розташовані поряд ненульові блоки відповідного рядка блоків об'єднуються у прямокутні блоки, розмір яких має бути оптимальним з точки зору кешування даних. У випадку матриці хмарочосної структури за такого об'єднання можна отримати, крім діагонального блока $A_{K,K}$, декілька (M_K) ненульових прямокутних блоків $A_{K,K+L}$ ($L = 1, \dots, M_K$), між якими розташовуються прямокутні нульові блоки. На гібридному комп'ютері у пам'яті GPU на кожному кроці алгоритму достатньо зберігати тільки (також циклічно розподілені) ненульові блоки підматриці, які обробляються. Головний потік, який містить на K -му кроці алгоритму K -й рядок блоків, далі називатимемо провідним CPU (як і пов'язаний з ним GPU у випадку гібридного алгоритму).

Уведемо позначення: D_K — діагональна матриця з LDL^T -розвинення діагонального блока $A_{11}^{(K)}$, $U_{K,I}$ — верхній трикутний блок порядку s підматриці $U_{12}^{(K)}$, $L_{I,K} = U_{K,I}^T D_K$ ($I = K, \dots, N$, $K = 1, \dots, N$). Тоді для $K = 1, \dots, N$ виконуються такі мікрооперації:

1) за умов $K > 1$, $U_{K-1,K} \neq 0$ обчислення верхнього трикутника діагонального блока $A_{11}^{(K)} = A_{K,K} - L_{K,K-1} U_{K-1,K}$ з використанням s -рангової модифікації симетричної матриці (на комп'ютері гібридної архітектури також пересилання результату провідному CPU);

2) LDL^T -розвинення діагонального блока $A_{11}^{(K)} = (U_{11}^{(K)})^T D_K U_{11}^{(K)}$;

3) за умов $1 < K < N - 1$, $L_{I,K-1}$, $U_{K-1,J} \neq 0$ s -рангова модифікація $A_{I,J}^{(K)} = A_{I,J}^{(K-1)} - L_{I,K-1} U_{K-1,J}$ ненульових блоків $A_{I,J}$ підматриці $A_{22}^{(K)}$;

4) за умови $K < N - 1$ розсилання обчислених блоків D_K , $U_{11}^{(K)}$ усім процесорним пристроям, які використовуються, виконується під час обчислень макрооперації 3;

5) за умови $K < N - 1$ обчислення ненульових блоків $L_{I,K} \equiv G_{K,I}^T$ ($K + 1 \leq I \leq N$) — розв'язування матричних СЛАР з нижньою трикутною матрицею $(U_{11}^{(K)})^T G_{K,I} = A_{K,I}$ та розсилання результатів усім процесорним пристроям, які використовуються;

6) за умови $K < N - 1$ обчислення ненульових блоків $U_{K,J} = D_K^{-1} G_{K,J}$ ($K + 1 \leq J \leq N$).

Аналіз макрооперацій 1–3 і 5, 6 показав, що більшість обчислень можна реалізувати, використовуючи програмні модулі для матрично-матричних (або матрично-векторних) операцій. Макрооперація 3 може виконуватись на непровідних процесорних пристроях K -го кроку асинхронно з виконанням макрооперацій 1, 2 на провідних процесорних пристроях, макрооперація 2 виконується на провідному CPU, а макрооперації 5 — на провідних процесорних пристроях.

Можна використовувати тільки одну групу процесорних пристроїв з одним головним потоком, але потрібно забезпечити ефективне розпаралелення обчислень між всіма потоками.

Розглянемо алгоритм LDL^T -розвинення блочно-діагональної матриці з обрамленням. Перші паралельні алгоритми трикутного розвинення розрідженої матриці було розроблено для блочно-діагональних матриць з обрамленням. Матрицю такого вигляду можна отримати, застосувавши до вихідної матриці метод паралельних перерізів (див., наприклад, [24]). Передбачається, що діагональні підматриці і підматриці обрамлення мають блочно-розріджену структуру. Тому потрібно оптимізувати структуру діагональних блоків, переупорядкувавши їх до однієї з блочно-розріджених структур.

З урахуванням симетричності вихідної матриці розподіляють і зберігають лише ненульові блоки її верхнього трикутника включно з елементами головної діагоналі. Між p головними потоками використовується такий розподіл даних: потоку з порядковим номером q ($q = 1, \dots, p$) розподіляються ненульові блоки діагональної підматриці $A_{q,q}$ і підматриці обрамлення $A_{p+1,q}$, на місці яких далі можуть розміщуватися відповідні ненульові блоки матриці U , і виділяється пам'ять для ненульових блоків верхнього трикутника (включно з головною діагоналлю) квадратної матриці $G_{q,q} = -L_{p+1,q}^T U_{q,p+1}$ та (за потреби) матриці $L_{p+1,q}$. Підматриця $A_{p+1,p+1}$ розподіляється одному з головних потоків, наприклад з порядковим номером p .

Для LDL^T -розвинення симетричної блочно-діагональної матриці з обрамленням виконуються такі макрооперації ($q = 1, \dots, p$):

- 1) LDL^T -розвинення діагональних підматриць $A_{q,q} = U_{q,q}^T D_{qi} U_{q,q}$;
- 2) обчислення позадіагональних підматриць обрамлення $U_{q,p+1}$ — розв'язування матричних СЛАР $U_{q,q}^T G_{q,p+1} = A_{q,p+1}$ та $D_q U_{q,p+1} = G_{q,p+1}$;
- 3) обчислення квадратної підматриці $G_{q,q} = -G_{q,p+1}^T U_{q,p+1}$;
- 4) модифікація діагональної підматриці $A_{p+1,p+1}^{(1)} = A_{p+1,p+1}^{(0)} + \sum_{q=1}^p G_{q,q}$ виконується головними потоками з використанням операції мультитизбирання блоків $G_{q,q}$;

- 5) LDL^T -розвинення діагонального блока

$$A_{p+1,p+1}^{(1)} = U_{p+1,p+1}^T D_{p+1} U_{p+1,p+1}.$$

Кожна з макрооперацій 1–3 алгоритму для одної пари підматриць не залежить від аналогічних операцій з іншими підматрицями і виконується паралельно, а розвинення 5 виконується потоком, якому розподілено $A_{p+1,p+1}$. Для макрооперацій 1–3 і 5 доцільно застосувати для кожного $q = 1, \dots, p$ та для $p+1$ відповідний структурі діагональних підматриць варіант дрібно-плиткового алгоритму [23], наприклад однопоточковий варіант (використовується один головний потік і декілька пов'язаних з ним потоків або один GPU) описаного раніше блочно-циклічного алгоритму для матриць блочно-хмарочосної структури.

Ефективність алгоритмів LDL^T -розвинення. Для дослідження ефективності паралельних алгоритмів використано методологію з [20].

Ефективність блочно-циклічного алгоритму LDL^T -розвинення симетричної матриці порядку n блочно-хмарочосної структури, якщо використовується p головних потоків та по n_{2L} потоків, пов'язаних з одним головним потоком, визна-

чається оцінками коефіцієнтів прискорення S_p і ефективності E_p :

$$S_p \approx p \left(1 - \frac{(p-1)sm + \tau}{(m + ps)m + \tau} \right), \quad E_p = \frac{S_p}{p} \approx 1 - \frac{(p-1)sm + \tau}{(m + ps)m + \tau}, \quad (17)$$

де

$$\tau = \begin{cases} sm, & \text{якщо } \frac{n_{2L}(t_{CG} + t_{CC} \log_2 p)}{t_G} \leq \frac{s}{p}, \\ pm \frac{n_{2L}(t_{CG} + t_{CC} \log_2 p)}{t_G}, & \text{якщо } \frac{n_{2L}(t_{CG} + t_{CC} \log_2 p)}{t_G} \leq m, \end{cases}$$

за умов $2ps \leq m$ (а у разі використання GPU ще і за умов $t_{CG} \leq \frac{ps^2}{3m} t_C$,

$t_{CG} + \frac{s}{3} t_C \leq \frac{m^2 - ps^2}{ps} \frac{t_G}{n_{2L}}$), а також якщо для обмінів даними між головними

потокми використовується алгоритм «дерева». Оцінки (17) свідчать, що ефективність алгоритму не залежить від порядку матриці, а за певних умов визначається кількістю ненульових блоків у рядку блоків і розмірами цих блоків. У (17) позначено: m — середня кількість ненульових наддіагональних елементів у одному рядку верхньої трикутної матриці U , t_C — середній час виконання однієї арифметичної операції з плаваючою комою одним потоком, t_{CC} — середній час обміну одним машинним словом між двома головними потоками. Якщо використовується GPU, то позначено: n_{2L} — кількість однорідних арифметичних операцій, які породжено однією операцією на CPU та які одночасно виконуються на GPU, t_G — середній час виконання однієї арифметичної операції з плаваючою комою на GPU, t_{CG} — середній час обміну одним машинним словом між CPU і GPU. Якщо GPU не використовується, то $t_G = t_C$, $t_{CG} = 0$.

Ефективність алгоритму можна підвищити, якщо в залежності від математичних властивостей діагональних блоків $A_{11}^{(K)}$ використовувати змінну розмірність — одинарну, подвійну або ще більшу, забезпечуючи при цьому потрібну точність розв'язку.

Розглянемо ефективність блочного алгоритму LDL^T -розвинення симетричної блочно-діагональної матриці з обрамленням порядку n , діагональні підматриці якої є стрічковими з напівшириною стрічки m . Якщо $n + m \gg 2pm$ і для обмінів даними між головними потоками використовується алгоритм «дерева», то можна отримати такі оцінки (використано введені раніше позначення):

$$S_p \approx p \left(1 - \frac{(\alpha + 2)m + \tau}{4n + \alpha m + \tau} \right), \quad E_p \approx 1 - \frac{(\alpha + 2)m + \tau}{4n + \alpha m + \tau}, \quad (18)$$

де $\alpha = \frac{7p^2 - 25p + 12}{3}$, $\tau = 2p \log_2 p \frac{n_{2L} t_{CC}}{t_G}$.

Оцінки (18) справедливі, якщо порядок діагонального блока обрамлення набагато менше порядків інших діагональних блоків. Якщо порівнювати з алгоритмом розвинення стрічкової (неперепорядкованої) матриці, то через зростання щільності підматриць обрамлення у розвиненні ефективність алгоритму зменшується приблизно в 4 рази. Тому цей алгоритм доцільно використовувати за умови $p > 4$.

ВИСНОВКИ

В роботі розглянуто методи та алгоритми розв'язування умовно коректних задач з еліптичними операторами, які мають єдиний розв'язок на підпросторі.

Для розв'язання таких задач, наприклад для задачі аналізу міцності незакріплених конструкцій (першої основної задачі теорії пружності), розроблено та обґрунтовано методологію обчислення наближення до єдиного розв'язку на підпросторі. Тобто обґрунтовано застосування нормального псевдорозв'язку розрахункової задачі (СЛАР) для отримання наближеного розв'язку дискретної задачі на підпросторі. Для практичного застосування такого підходу не потрібно знання в явному вигляді нуль-простору оператора вихідної задачі.

Для обчислення з гарантованою точністю нормального псевдорозв'язку (та зваженого нормального псевдорозв'язку) СЛАР з розрідженою симетричною напіввизначеною матрицею розроблено та обґрунтовано метод триетапної регуляризації. Отримано оцінки точності наближених розв'язків у випадку збурення правої частини СЛАР. Розроблено економічні паралельні алгоритми для методу триетапної регуляризації для комп'ютерів з паралельною організацією обчислень різної архітектури. Отримано оцінки прискорення та ефективності запропонованих алгоритмів. Метод триетапної регуляризації є альтернативою методу регуляризуючого функціонала (переваги — краща обумовленість задач (10) та збереження розрідженої структури матриці) і методу дискретної регуляризації (перевага — очевидно менша кількість арифметичних операцій).

Тестування показало, що метод триетапної регуляризації за часовими витратами, використуваній пам'яті, достовірності отриманих результатів є найбільш ефективним методом для розв'язання на комп'ютерах з паралельною організацією обчислень СЛАР, які виникають у результаті дискретизації задач аналізу міцності складних незакріплених конструкцій.

Запропонована методологія інваріантна для широкого класу задач з умовами Неймана на границі. Її можна застосувати для чисельного розв'язання інших задач теорії пружності, що мають єдиний розв'язок на підпросторі, наприклад для одно- і двовимірних постановок, змішаних крайових задач тощо, а також для чисельного розв'язання крайових задач, що виникають під час математичного моделювання в інших галузях науки і техніки.

СПИСОК ЛІТЕРАТУРИ

1. Strang G., Fix G.J. An analysis of the finite element method. New York: Prentice-Hall, 1973. 306 p.
2. Zienkiewicz O.C., Morgan K. Finite Elements and Approximation. New York: A Wiley-Interscience Publication, 1983. 320 p.
3. Городецкий А.С., Евзеров И.Д. Компьютерные модели конструкций. Киев: ФАКТ, 2007. 394 с.
4. Молчанов И.Н., Галба Е.Ф. Вариационные постановки статической задачи теории упругости при заданных внешних силах. *Укр. матем. журн.* 1991. Т. 43, № 2. С. 158–161.
5. Молчанов И.Н., Левченко И.С., Федонюк Н.Н., Химич А.Н., Чистякова Т.В. Численное моделирование концентрации напряжений в упругом полупространстве с двухслойным включением. *Прикладная механика.* 2002. Т. 38, № 3. С. 65–71.
6. Химич А.Н., Яковлев М.Ф. О решении систем с матрицами неполного ранга. *Компьютерная математика. Сб. науч. трудов.* 2003. Вып. 1. С. 1–15.
7. Попов А.В., Химич А.Н. Исследование и решение первой основной задачи теории упругости. *Компьютерная математика. Сб. науч. трудов.* 2003. Вып. 2. С. 105–114.
8. Попов О.В. Про ефективний метод розв'язування некоректних задач з розрідженими матрицями. *Теорія оптимальних рішень. Зб. наук. праць.* 2013. № 12. С. 77–81.

9. Химич А.Н., Молчанов И.Н., Попов А.В., Чистякова Т.В., Яковлев М.Ф. Параллельные алгоритмы решения задач вычислительной математики. Київ: Наук. думка, 2008. 248 с.
10. Khimich A.N., Popov A.V., Polyanko V.V. Algorithms of parallel computations for linear algebra problems with irregularly structured matrices. *Cybernetics and Systems Analysis*. 2011. Vol. 47, Iss. 6. P. 973–985. <https://doi.org/10.1007/s10559-011-9377-4>.
11. Попов О.В. Про паралельні алгоритми факторизації розріджених матриць. *Комп'ютерна математика. Сб. науч. трудов*. 2013. Вып. 2. С. 115–124. URI: <http://dspace.nbuv.gov.ua/handle/123456789/84755>.
12. Timoshenko S.P., Goodier J.N. Theory of Elasticity. New York: McGraw-Hill, 1970.
13. Михлин С.Г. Вариационные методы в математической физике. Москва: Наука, 1970. 454 с.
14. Галба Е.Ф., Гладкий А.В., Химич А.Н., Яковлев М.Ф. О корректности первой основной задачи теории упругости на подпространстве. *Комп'ютерна математика. Сб. науч. трудов*. 2002. Вып. 1. С. 54–62.
15. Siarlet P.G. The Finite Element Method for Elliptic Problems. North Holland, 1978. 529 p.
16. Khimich A.N. Perturbation bounds for the least squares problem. *Cybernetics and Systems Analysis*. 1996. Vol. 32, Iss. 3. P. 434–436. <https://doi.org/10.1007/BF02366509>.
17. Химич А.Н. Оценки полной погрешности решения систем линейных алгебраических уравнений для матриц произвольного ранга. *Комп'ютерна математика. Сб. науч. трудов*. 2002. Вып. 2. С. 41–49.
18. Морозов В.А. Методы регуляризации неустойчивых задач. Москва: Изд-во Московского ун-та, 1987. 217 с.
19. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. Москва: Наука, 1984. 318 с.
20. Попов А.В., Чистяков О.В. Про ефективність алгоритмів з багаторівневим паралелізмом. *Фізико-математичне моделювання та інформаційні технології*. 2021. Вип. 33. С. 133–137. <https://doi.org/10.15407/fmmit2021.33.133>.
21. Хіміч О.М., Сидорук В.А. Використання мішаної розрядності у математичному моделюванні. *Математичне та комп'ютерне моделювання. Серія: Фізико-математичні науки*. 2019. Вип. 19. С. 180–187.
22. Khimich O.M., Chistyakova T.V., Sidoruk V.A., Ershov 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. <https://doi.org/10.1007/s10559-021-00424-z>.
23. Хіміч О.М., Сидорук В.А. Плитковий гібридний алгоритм факторизації структурно-симетричних матриць. *Теорія оптимальних рішень. Зб. наук. праць*. 2017. № 2017. С. 125–132.
24. Джордж А., Дж. Лю. Численное решение больших разреженных систем уравнений. Москва: Мир, 1984. 334 с.

A.N. Khimich, A.V. Popov

SOLVING ILL-POSED PROBLEMS OF THE THEORY OF ELASTICITY USING HIGH-PERFORMANCE COMPUTER SYSTEMS

Abstract. A method of efficient analysis and solution of conditionally correct problems, which have a unique solution in the subspace, is proposed. The use of a discrete finite-element model on the entire space to obtain a unique solution on the subspace of the initial variational problem is justified. To find a normal pseudo-solution (or a weighted normal pseudo-solution) of a discrete problem (a system of linear algebraic equations with a sparse symmetric semidefinite matrix), a three-stage regularization method is proposed. This method makes it possible to obtain approximations of these solutions with a predetermined accuracy. Efficient adaptive high-performance algorithms of the specified method have been developed for solving systems of linear algebraic equations with sparse symmetric semi-definite matrices using modern computers with parallel computing.

Keywords: high-performance computing, finite element method, three-stage regularization method, sparse symmetric semidefinite matrix, system of linear algebraic equations, first fundamental problem of elasticity theory, variable computer environment.

Надійшла до редакції 09.05.2023