УДК 519.612

В. С. Абрамчук*, канд. фіз.-мат. наук,
І. В. Абрамчук,** старший викладач
*Вінницький державний педагогічний університет
імені Михайла Коцюбинського, м. Вінниця,
**Вінницький національний технічний університет, м. Вінниця

ІТЕРАЦІЙНІ МЕТОДИ РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ РІВНЯНЬ З ДОВІЛЬНИМИ НЕВИРОДЖЕНИМИ МАТРИЦЯМИ

Обґрунтовано, що метод напрямленого пошуку розв'язування систем $A\vec{x}=\vec{b}$, $A\in M_n(R)$, володіє основними факторами ефективності ітераційних методів.

Ключові слова: *метод напрямленого пошуку, підпростір Крилова, максимізація відношення Релея.*

У статті, на основі необхідних і достатніх умов прискорення збіжності методу напрямленого пошуку, пропонуються швидкозбіжні стійкі ітераційні методи розв'язування систем $A\vec{x}=\vec{b}$ з довільними дійсними невиродженими квадратними матрицями $A\in M_n(R)$, $\operatorname{rank} A=n$, що дозволяє обґрунтувати ефективність методів, запропонованих в [1—3].

За метод розв'язування систем

$$A\vec{x}=\vec{b},\ A\in M_n(R), \vec{b}\in im(A), \operatorname{rank} A=n \tag{1}$$

виберемо метод напрямленого пошуку, який подамо у вигляді такої процедури

$$\vec{x}^{(k+1)}=F_k\left(A,\vec{b},\vec{x}^{(k)},\vec{v}_1^{(k)},...,\vec{v}_m^{(k)}\right), \tag{2}$$

де F_k — послідовність операторів; $\left\{\vec{v}_1^{(k)},...,\vec{v}_m^{(k)}\right\}$, $m\ll n$ — система лінійно-незалежних векторів, яка утворює базис напрямного підпростору $P_m^{(k)}$.

Постановка проблеми. 1. З'ясувати причини, що обмежують застосовність існуючих ітераційних методів до розв'язування систем (1) з довільними матрицями. 2 Обґрунтувати, що метод (2) задовольняє основним факторам ефективності ітераційних методів (збігається для систем з довільними матрицями (1); дозволяє будувати швидкозбіжні процедури прискорення; стійкий до похибок заокруглення; одночасно мінімізує норму похибки $\left\|\vec{x}^{(k+1)}-\vec{x}^*\right\|_2$ і максимізує

© В. С. Абрамчук, І. В. Абрамчук, 2012

$\cos^2\left\{A^T\vec{c}, \vec{x}^{(k)} - \vec{x}^*\right\}$ між напрямним вектором $A^T\vec{c}$ і напрямком на розв'язок $\vec{x}^{(k)} - \vec{x}^*$; потребує мінімальних комп'ютерних ресурсів).

Головним фактором ефективності ітераційних методів розв'язування систем (1) є збіжність ітераційного процесу для систем з матрицями широких класів (матриця A вважається «чорним ящиком» [4; 5]). Моделювання складних фізичних процесів приводить до ускладнення структур дискретних задач і вимагає переходу від скінченних різниць до методів лінійних та квадратичних скінченних елементів, застосування граничних елементів, переходу від методу розщеплення для фізичних і просторових змінних до «цілісних» дискретизацій систем рівнянь, які описують тривимірні об'єкти [6—10]. Матриці відповідних систем мають великі розміри (порядків $10^3 - 10^6$), достатньо щільно заповнені (до сотень і навіть тисяч елементів в кожному рядку) і можуть не належати до жодного класу матриць, які дозволяють прискорити збіжність «класичних» методів (наприклад, не мають діагонального переважання, не є монотонними, симетричними, додатно визначеними тощо).

Класичні результати про збіжність ітераційного методу

$$\vec{x}^{(k+1)} = H\vec{x}^{(k)} + \vec{d}, \quad k = 0,1,\dots, \tag{3}$$

узгодженого з системою (1), наведені в [6]:

- узгоджений метод (3) збігається до єдиного розв'язку системи (1) тоді і лише тоді, якщо $\rho(H) < 1$ ($\rho(H) = \max|\lambda_i|$ — спектральний радіус матриці H; метод (3) узгоджений з системою (1), якщо $\vec{x}^* = H\vec{x}^* + \vec{d}$, $\vec{x}^* = A^{-1}\vec{b}$);
- якщо $H \in M_n(R)$, то $\rho(H) < 1$ тоді і лише тоді, якщо існує така симетрична додатно визначена матриця B, що матриця $B - H^T BH$ додатно визначена (теорема Стейна).

Узгоджені ітераційні методи пов'язані з розщепленням матриці $A = P - Q, \quad \det P \neq 0, \quad H = P^{-1}Q$. Основні результати про збіжність методу (3) отримані для P — регулярного розщеплення [5; 6] для випадків, якщо матриця A симетрична і додатно визначена, або має діагональне переважання, або є M — матрицею тощо). Недоліком методів (3) є те, що для довільних матриць A, $\operatorname{rank} A = n$, задача вибору матриці $H = P^{-1}Q$, $\rho(H) < 1$, є проблемною задачею [6].

Другим сімейством ітераційних методів розв'язування великих розріджених систем (1) є три основні групи методів в підпросторах Крилова [8]. До першої групи відносяться методи, які трансформують матрицю A до симетричної форми (до неї відносяться проекційні методи спряжених напрямів, застосовних до матриць AA^T, A^TA), недоліком їх є сильне зростання числа обумовленості трансформованої матриці. Друга група — узагальнені методи спряжених напрямів, які вимагають великого об'єму оперативної пам'яті ЕОМ. До третьої групи відносяться методи біспряжених градієнтів і нев'язок (їх узагальнення — методи здвоєних та стабілізованих біспряжених напрямів). Основним недоліком всіх трьох груп методів є те, що вектори підпростору Крилова утворюють погано обумовлену систему, отже можна використовувати лише варіанти криловських ітерацій з періодичними рестартами [8], що знижує їх швидкість збіжності. Збіжність цих методів обґрунтована в умовах абсолютно точних обчислень і для додатно визначених матриць.

Важливе значення для практичної реалізації ітераційного методу має вибір критерію збіжності. Проілюструємо недолік вибору векторної норми збіжності. Змоделюємо процес методу проекцій на гіперплощини, теоретично збіжний в абсолютно точній арифметиці для систем з довільними матрицями $A \in M_n(R)$, $\operatorname{rank} A = n$, на прикладі погано обумовленої симетричної матриці Гільберта. За критерій оцінки збіжності на k-му кроці виберемо норми $\left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2$, $\left\|\vec{r}^{(k)}\right\|_2 = \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_{A^TA}$, $\vec{x}^* = A^{-1}b$. Якщо $\left\|\vec{r}^{(k)}\right\|_2$ достатньо великі числа, то процес збіжності монотонно спадний як для норм нев'язок так і для норм похибок. Якщо $\left\|\vec{r}^{(k)}\right\|_2$ стає малою величиною, то процес проекцій в силу похибок заокруглення, втрачає монотонність. Причина цього явища пов'язана з наявністю у матриці Гільберта малих власних значень (це явище спостерігається і для різницевих рівнянь великих порядків).

Для кожного ітераційного методу є області найповільнішої швидкості збіжності, які лежать в околі власних векторів $\vec{x} - \vec{x}^*$, що відповідають найменшим власним значенням матриці A^TA [3] (під околом вектора розуміється конус з вершиною в точці $\vec{x}^*$). Залежність швидкості збіжності від щільності власних значень та їх розподілу в спектрі матриці A досліджувалась в [4].

Таким чином, побудова ефективних ітераційних методів наближеного розв'язування систем (1) з довільними дійсними, квадратними, невиродженими, погано обумовленими матрицями вимагає розробки методів, що збігаються за напрямом — максимізують на кожному кроці функціонал $\cos^2\left\{\vec{p}^{(k)}, \vec{x}^{(k)} - \vec{x}^*\right\}$, де $\vec{p}^{(k)}$ — вектор з напрямного підпростору $P_m^{(k)}$.

Виклад основного матеріалу. Метод напрямленого пошуку (МНП). МНП з одновимірним напрямним підпростором $P_1^{(k)} = = \boldsymbol{span}\left\{A^T\vec{c}^{(k)}\right\}$, $\left(\vec{c}^{(k)}, \vec{r}^{(k)}\right) \neq 0$, описується процедурою [1]: для всіх $k = 0,1,\ldots$,

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} + \alpha_k A^T\vec{c}^{(k)}, \quad \alpha_k = -\left(\vec{r}^{(k)}, \vec{c}^{(k)}\right) \Big/ \left\| A^T\vec{c}^{(k)} \right\|_2^2 . \qquad (4)$$

Швидкість збіжності визначається формулою

$$\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2^2 = \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2^2 \left(1 - \cos^2\left\{A^T\vec{c}^{(k)}, \vec{x}^{(k)} - \vec{x}^*\right\}\right). \qquad (5)$$

Якщо напрямний підпростір на k-му кроці задається системою взаємно ортогональних векторів $\vec{p}^{(k,i)} = A^T\vec{c}^{(k,i)}$, $i = 1,\ldots,m$, (псевдоортогональних векторів з матрицею AA^T), то швидкість збіжності

$$\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2^2 = \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2^2 \left(1 - \sum_{i=1}^{m} \cos^2\left\{A^T\vec{c}^{(k,i)}, \vec{x}^{(k)} - \vec{x}^*\right\}\right). \qquad (6)$$

Необхідні умови прискорення швидкості збіжності МНП. Формули (5), (6) можна переписати так: для одновимірного напрямного підпростору

$$\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2^2 = \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2^2 - \left(\vec{r}^{(k)}, \vec{c}^{(k)}\right)^2 \Big/ \left\|A^T\vec{c}^{(k)}\right\|_2^2 ; \qquad (7)$$

для багатовимірного напрямного підпростору

$$\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2^2 = \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2^2 - \sum_{i=1}^{m}\left(\vec{r}^{(k)}, \vec{c}^{(k,i)}\right)^2 \Big/ \left\|A^T\vec{c}^{(k,i)}\right\|_2^2 . \qquad (8)$$

Теорема 1. Для прискорення збіжності ітераційного методу (4) за k кроків необхідно, щоб напрямні вектори $\vec{c}^{(j)}$, $j = 0,1,\ldots,k$, були псевдоортогональними з матрицею AA^T і задовольняли умову

$$\forall j = 0,1,\ldots,k : \left(\vec{r}^{(j)}, \vec{c}^{(j)}\right) \neq 0 . \qquad (9)$$

Умова (9) випливає з (7) та (8). Необхідність умови ортогональності векторів $A^T\vec{c}^{(j)}$, $j = 0,1,\ldots k$, для прискорення збіжності випливає з геометричних міркувань. Для $k = 0$ наближенням до розв'язку є вектор $\vec{x}^{(1)}$, що належить гіперплощині $W^{(0)} = \left\{\vec{x} : \left(\vec{x} - \vec{x}^*, A^T\vec{c}^{(0)}\right) = 0\right\}$. Оскільки вектори $\vec{x}^{(1)}, \vec{x}^*$ належать гіперплощині $W^{(0)}$, то найкраще наближення $\vec{x}^{(2)}$ необхідно шукати в $W^{(0)}$, а це означає, що напрямний вектор $A^T\vec{c}^{(1)}$ повинен бути ортогональним до вектора $A^T\vec{c}^{(0)}$: $\vec{x}^{(2)} \in W^{(1)} = \left\{\vec{x} : \left(\vec{x} - \vec{x}^*, A^T\vec{c}^{(1)}\right) = 0\right\}$, отже $\vec{x}^{(2)} \in W^{(0)} \cap W^{(1)}$. Припустимо, що $\vec{x}^{(k)} \in W^{(0)} \cap W^{(1)} \cap \ldots \cap W^{(k-1)}$. Оскільки вектори $\vec{x}^{(k)}, \vec{x}^*$ належать одночасно всім гіперплощинам $W^{(j)}$, $j = 0,1,\ldots,k-1$, то найкраще наближення повинно належати $\bigcap_{j=0}^{k-1} W^{(j)}$, а це означає, що напрямний вектор $A^T\vec{c}^{(k)}$ повинен бути ортогональним усім векторам нормалей $A^T\vec{c}^{(j)}$, $j = 0,1,\ldots,k-1$.

Підпростір з векторів нормалей за k кроків будемо позначати $V^{(k)} = \boldsymbol{span}\left\{A^T\vec{c}^{(0)}, A^T\vec{c}^{(1)}, \ldots, A^T\vec{c}^{(k)}\right\}$.

Достатня умова прискорення збіжності МНП.

Теорема 2. Оптимальний вектор $\vec{c}^{(k)} \in D^{(k)} \subset R^n$ на k-му кроці методу (4), що мінімізує $\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2$, доставляє максимум функціоналу

$$\varphi(\vec{c}) = \left(\vec{r}^{(k)}, \vec{c}\right)^2 \Big/ \left\|A^T\vec{c}\right\|_2^2 . \tag{10}$$

Оптимальним напрямком на розв’язок $\vec{x}^*$ системи $A\vec{x} = \vec{b}$ з довільної точки $\vec{x}^{(k)} \neq \vec{x}^*$ у просторі R^n є напрямний вектор $A^T\vec{c}$, колінеарний вектору $\vec{x}^{(k)} - \vec{x}^*$: $A^T\vec{c} = \lambda\left(\vec{x}^{(k)} - \vec{x}^*\right)$ або $AA^T\vec{c} = \lambda\vec{r}^{(k)}$, $\lambda \neq 0$. Функціонал $\varphi(\vec{c})$ можна переписати так:

$$\varphi(\vec{c}) = \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2^2 \cos^2\left\{A^T\vec{c}, \vec{x}^{(k)} - \vec{x}^*\right\}. \tag{11}$$

Оскільки множник $\left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2^2$ у формулі (11) є константою, то вибір оптимального вектора $A^T\vec{c}$ означає максимізацію функціонала $\cos^2\left\{A^T\vec{c}, \vec{x}^{(k)} - \vec{x}^*\right\}$, $\vec{c} \in R^n$.

Підпростір Крилова. За базис напрямного підпростору $K_m^{(k)}(A)$ виберемо систему векторів $\vec{r}^{(k)}, A\vec{r}^{(k)}, \ldots, A^{m-1}\vec{r}^{(k)}$, яка задовольняє необхідну умову $\left(\vec{r}^{(k)}, A^i\vec{r}^{(k)}\right) \neq 0$, $i = 0,1,\ldots,m-1$ [3].

Алгоритм 1 (наближеного розв'язування системи (1)).

Вибрати початковий вектор $\vec{x}^{(0)}$, обчислити $\vec{r}^{(0)}$, покласти $\vec{c}^{(0)}=\vec{r}^{(0)}$, $V^{(0)}=\operatorname{span}\left\{A^T\vec{c}^{(0)}\right\}$. Для всіх $k=0,1,...$ знайти наближення $\vec{x}^{(k+1)}$ за формулою (4), обчислити $\vec{r}^{(k+1)}=A\vec{x}^{(k+1)}-\vec{b}$. Застосувати критерій збіжності $1-\cos^2\left\{AA^T\vec{c}^{(k)},\vec{r}^{(k)}\right\}\le\varepsilon$, де ε — задана похибка. Якщо він не виконується, то продовжити процес:

- сформувати базис підпростору Крилова $K_m^{(k+1)}(A)$;
- з векторів підпростору $K_m^{(k+1)}(A)$ сформувати вектори $\vec{c}^{(k+1,j)}$ псевдоортогональні векторам підпростору $V^{(k)}$ (необхідний критерій прискорення збіжності): $\forall\ j=0,1,...,m-1$,

$$\vec{c}^{(k+1,j)}=A^j\vec{r}^{(k+1)}+\sum_{s=0}^{k}\beta_s^{(k)}\vec{c}^{(s)},\quad \beta_s^{(k)}=-\frac{\left(A^TA^j\vec{r}^{(k+1)},A^T\vec{c}^{(s)}\right)}{\left\|A^T\vec{c}^{(s)}\right\|_2^2};$$

- у підпросторі Крилова $K_m^{(k+1)}(A)$ з новим базисом $\left\{\vec{c}^{(k+1,j)},\ j=0,1,...,m-1\right\}$ знайти вектор $\vec{c}^{(k+1)}$, що максимізує (10) (мінімізує $\left\|\vec{x}^{(k+1)}-\vec{x}^*\right\|_2$);

$$\vec{c}^{(k+1)}=\vec{c}^{(k+1,0)}+\sum_{j=0}^{m-1}\alpha_j^{(k+1)}\vec{c}^{(k+1,j)};$$

- розширити підпростір $V^{(k)}$: $V^{(k+1)}=\operatorname{span}\left\{A^T\vec{c}^{(0)},,...,A^T\vec{c}^{(k+1)}\right\}$.

Повторити процес.

Задача максимізації функціонала (10) досліджувалась у [2].

Відношення Релея, вибір початкового напрямного вектора і початкової точки. Розв'язування систем $A\vec{x}=\vec{b}$ великих порядків.

Однією з проблем є розробка швидкозбіжних методів розв'язування різницевих та сіткових рівнянь великих порядків. Актуальність досліджень цієї проблеми полягає у тому, що з ростом розмірності неортогональної матриці зростає її обумовленість і необхідно забезпечити стійкість процесу обчислень (проблема, яку називають «прокляттям розмірності» [12]). Оптимальним підходом для прискорення швидкості збіжності ітераційного процесу наближеного розв'язування системи $A\vec{x}=\vec{b}$ є пошук довільної точки на довільній сингулярній прямій $\vec{x}=\vec{x}^{*}+t\cdot\vec{p}_{\lambda}$, $t\in R$, $\vec{p}_{\lambda}$ — власний вектор матриці $A^{T}A$, що відповідає власному значенню $\lambda\left(A^{T}A\right)$. Проблема розмірності матриці стає несуттєвою, оскільки матриця $A^{T}A$ простої структури [5]. Задача наближення до сингулярної прямої полягає у максимізації функціонала $\cos^{2}\left\{A^{T}A\left(\vec{x}-\vec{x}^{*}\right),\vec{x}-\vec{x}^{*}\right\}$ (або $\cos^{2}\left\{A^{T}A\vec{r},\vec{r}\right\}$, $\vec{r}=A\vec{x}-\vec{b}$) на сфері Ω_0 або в кулі Ω_1:

$$\Omega_0=\left\{\vec{x}:\left\|\vec{x}-\vec{x}^{*}\right\|_2=\left\|\vec{x}^{(0)}-\vec{x}^{*}\right\|_2\right\},$$

$$\Omega_1=\left\{\vec{x}:\left\|\vec{x}-\vec{x}^{*}\right\|_2\le\left\|\vec{x}^{(0)}-\vec{x}^{*}\right\|_2\right\},$$

де $\vec{x}^{(0)}$ — довільний вектор. Оскільки ця задача є задачею нелінійного програмування, то побудуємо часткові алгоритми наближення до власних векторів, що відповідають $\lambda_{\max}\left(A^{T}A\right)$, $\lambda_{\min}\left(A^{T}A\right)$. Для цього модифікуємо ітераційний процес (4):

$$\vec{x}^{(k+1)}=\vec{x}^{(k)}+\alpha_k A^{T}\vec{c}^{(k)},\quad \vec{r}^{(k+1)}=\vec{r}^{(k)}+\alpha_k AA^{T}\vec{c}^{(k)}. \tag{12}$$

Позначимо

$$\beta_k=-2\left(\vec{r}^{(k)},\vec{c}\right)\Big/\left\|A^{T}\vec{c}\right\|_2^{2},\ \gamma_k=-2\left(AA^{T}\vec{r}^{(k)},\vec{c}\right)\Big/\left\|AA^{T}\vec{c}\right\|_2^{2}. \tag{13}$$

Процес наближення до власних значень $\lambda_{\max}\left(A^{T}A\right)$, $\lambda_{\min}\left(A^{T}A\right)$ здійснимо на основі максимізації (мінімізації) відношення Релея $\rho\left(A^{T}A,\vec{x}-\vec{x}^{*}\right)=\left\|\vec{r}\right\|_2^{2}\Big/\left\|\vec{x}-\vec{x}^{*}\right\|_2$ на сфері Ω_0 або в кулі Ω_1. Зазначимо, що відомі ітераційні процедури розв'язання симетричної проблеми власних значень для матриць великих порядків [4; 12], не можна безпосередньо застосувати до пошуку наближень до сингулярних прямих, які проходять через невідому точку $\vec{x}^{*}$ — розв'язок системи $A\vec{x}=\vec{b}$, оскільки вони побудовані на принципі оптимізації без обмежень і приводять до розгону норми вектора похибки.

Теорема 3. Якщо вектор $\vec{r}^{(k)}=A\vec{x}^{(k)}-\vec{b}$ не є власним вектором матриці AA^{T} і якщо а) $\operatorname{sgn}\beta_k=\operatorname{sgn}\gamma_k$, $\left|\gamma_k\right|<\left|\beta_k\right|$ або б) $\operatorname{sgn}\beta_k=-\operatorname{sgn}\gamma_k$, то ітераційний процес (12) максимізує відношення Релея в кулі Ω_1. Якщо в) $\operatorname{sgn}\beta_k = \operatorname{sgn}\gamma_k$, $|\beta_k| < |\gamma_k|$, то мінімізується відношення Релея в кулі Ω_1. Оптимальне значення параметра α_k шукається шляхом дослідження функції

$$\varphi(\alpha_k) = \cos^2\left\{AA^T\vec{r}^{(k)}, \vec{r}^{(k)}\right\} = \left\|A^T\vec{r}^{(k)}\right\|_2^4 \Big/ \left(\left\|AA^T\vec{r}^{(k)}\right\|_2^2 \left\|\vec{r}^{(k)}\right\|_2^2\right)$$

на найбільше (найменше) значення на відрізку $\left[|\gamma_k|; |\beta_k|\right]$ ($\left[|\beta_k|; |\gamma_k|\right]$).

Доведення висновків теореми 4 випливає з того, що параметр $\alpha_k = \beta_k$ забезпечує умову $\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2 = \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2$, а параметр $\alpha_k = \gamma_k$ — умову $\left\|\vec{r}^{(k+1)}\right\|_2 = \left\|\vec{r}^{(k)}\right\|_2$. На проміжках $\left[0; |\beta_k|\right]$, $\left[0; |\gamma_k|\right]$ для опуклих функцій $\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2$, $\left\|\vec{r}^{(k+1)}\right\|_2$ відповідно виконуються нерівності $\left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2 \le \left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2$, $\left\|\vec{r}^{(k+1)}\right\|_2 \le \left\|\vec{r}^{(k)}\right\|_2$. Тому за умови а) відношення $\left\|\vec{r}^{(k+1)}\right\|_2 \Big/ \left\|\vec{x}^{(k+1)} - \vec{x}^*\right\|_2$ максимізується, за умови в) мінімізується, а за умови б) максимізація відбувається найшвидше.

Для забезпечення стійкості обчислювального ітераційного процесу розв'язування систем та проблеми власних значень для погано обумовленої матриці необхідно наближення будувати в області максимальних нев'язок (в околі сингулярної прямої, що визначається власним вектором, відповідним $\lambda_{\max}\left(A^TA\right)$ [3]).

Достатні умови прискорення збіжності до власного значення $\lambda_{\max}\left(A^TA\right)$ сформулюємо як висновок теореми 4.

Теорема 4. Якщо вектор $\vec{r}^{(k)}$ не є власним вектором матриці AA^T і виконується необхідна умова б), то для прискорення збіжності послідовності Релея $\rho_k = \rho\left(A^TA, \vec{x}^{(k)} - \vec{x}^*\right)$ достатньо максимізувати функціонал (10) в підпросторі Крилова $K_m^{(k)}(A) = \operatorname{span}\left(A^i\vec{r}^{(k)}\right)_{i=0}^{m-1}$ (або $K_m^{(k)}\left(AA^T\right)$).

Дійсно, підставивши $\alpha_k = \frac{1}{2}\beta_k = -\frac{1}{2}\gamma_k$ у формулу (12), дістанемо відношення Релея

$$\rho_{k+1} = \frac{\left\|\vec{r}^{(k)}\right\|_2^2 + 2\varphi\left(\vec{c}^{(k)}\right) + \varphi\left(\vec{c}^{(k)}\right)\left\|AA^T\vec{c}^{(k)}\right\|_2^2 \Big/ \left\|A^T\vec{c}^{(k)}\right\|_2^2}{\left\|\vec{x}^{(k)} - \vec{x}^*\right\|_2^2 - \varphi\left(\vec{c}^{(k)}\right)}, \quad (14)$$

яке максимально зросте за рахунок максимізації додатного функціонала (10). Наведемо формули для максимізації функціонала (14) у підпросторі $K_3^{(k)}\left(A A^T\right) = \operatorname{span}\left\{\vec{r}^{(k)}, AA^T\vec{r}^{(k)}, \left(AA^T\right)^2\vec{r}^{(k)}\right\}$.

Алгоритм 2. (максимізації відношення Релея в кулі Ω_1).

Вибрати довільний вектор $\vec{x}^{(0)} \in R^n$, обчислити $\vec{r}^{(0)} = A\vec{x}^{(0)} - \vec{b}$. Якщо $\vec{r}^{(0)}$ — власний вектор матриці AA^T, то процес закінчити, оскільки $\vec{p}^{(0)} = A^T\vec{r}^{(0)}$ — оптимальний напрямок на розв'язок. У протилежному випадку для всіх $k = 0,1,...$ сформувати вектори $\vec{r}^{(k)}, AA^T\vec{r}^{(k)}, \left(AA^T\right)^2\vec{r}^{(k)}$. Утворити вектор $\vec{c}^{(k)} = \left(AA^T\right)^2\vec{r}^{(k)} + {}$ ${}+\beta_{1,k}AA^T\vec{r}^{(k)} + \beta_{0,k}\vec{r}^{(k)}$, що задовольняє умову $\left(AA^T\vec{r}^{(k)}, \vec{c}^{(k)}\right) = {}$ ${}= -\left(\vec{r}^{(k)}, \vec{c}^{(k)}\right)$, де $\beta_{0,k} = -a_{0,k} - a_{1,k}\beta_{1,k}$,

$$\gamma_k = \left\|A^T\vec{r}^{(k)}\right\|_2^2 + \left\|\vec{r}^{(k)}\right\|_2^2, \quad \vec{c}^{(k)} = \vec{p}_{1,k} + \beta_{1,k}\vec{p}_{0,k},$$

$$a_{0,k} = \left(\left\|A^T AA^T\vec{r}^{(k)}\right\|_2^2 + \left\|AA^T\vec{r}^{(k)}\right\|_2^2\right)\Big/\gamma_k,$$

$$a_{1,k} = \left(\left\|AA^T\vec{r}^{(k)}\right\|_2^2 + \left\|A^T\vec{r}^{(k)}\right\|_2^2\right)\Big/\gamma_k,$$

$$\vec{p}_{0,k} = AA^T\vec{r}^{(k)} - a_{1,k}\vec{r}^{(k)}, \quad \vec{p}_{1,k} = \left(AA^T\right)^2\vec{r}^{(k)} - a_{0,k}\vec{r}^{(k)}.$$

Параметр $\beta_{1,k}$ визначимо з умови максимізації $\varphi\left(\vec{c}\right)$:

$$\beta_{1,k} = \left(u_1 v_0 - u_0 v_1\right)/\left(u_1 v_1 - u_2 v_0\right),$$

де $v_0 = \left(\vec{p}_{1,k}, \vec{r}^{(k)}\right)$, $v_1 = \left(\vec{p}_{0,k}, \vec{r}^{(k)}\right)$, $u_0 = \left\|A^T\vec{p}_{1,k}\right\|_2^2$,

$$u_1 = \left(A^T\vec{p}_{1,k}, A^T\vec{p}_{0,k}\right), \quad u_2 = \left\|A^T\vec{p}_{0,k}\right\|_2^2.$$

Обчислити $\vec{x}^{(k+1)} = \vec{x}^{(k)} - \alpha_k A^T\vec{c}^{(k)}$, де $\alpha_k = \left(\vec{r}^{(k)}, \vec{c}^{(k)}\right) \Big/ \left\| A^T \vec{c}^{(k)} \right\|_2^2$, $\vec{r}^{(k+1)} = A\vec{x}^{(k+1)} - \vec{b}$.

Перевірити критерій зупинки $1 - \cos^2 \left\{ AA^T \vec{r}^{(k)}, \vec{r}^{(k)} \right\} \le \varepsilon$. Якщо критерій не виконується, повторити процес.

Після завершення обчислювального процесу вектор $\vec{x}^{(k+1)}$ приймається за початкову точку ітераційного процесу (4), а вектор $A^T \vec{r}^{(k+1)}$ за початковий напрямний вектор.

Розв'язування різницевих еліптичних рівнянь. В околі довільного власного вектора матриці $A A^T$ система векторів підпростору Крилова $K_m\left(A A^T\right)$ стає погано обумовленою, а тому порядок m обирається в алгоритмі 2 малим. Щоб прискорити збіжність до оптимального напрямного вектора бажано використовувати крім підпростору Крилова допоміжний напрямний підпростір $P_s\left(A\right)$, розмірність якого в оптимальному варіанті збігалася б з розмірністю матриці A і обробка базису якого при оптимізації функціоналу (10) (або $\left\| \vec{x}^{(k+1)} - \vec{x}^* \right\|_2$) вимагала б мінімальних обчислювальних затрат. Обґрунтуємо вибір базису $P_s\left(A\right)$ для розв'язування різницевих еліптичних рівнянь, що виникають при дискретизації диференціальних рівнянь гідродинаміки і тепломасоперенесення на шаблонах з $2m+1$ вузлами, $m = 2,3$:

$$\sum_{j=1}^{m} \frac{\partial \rho u_j \Phi}{\partial x_j} = \sum_{j=1}^{m} \frac{\partial}{\partial x_j} \left(\tau \frac{\partial \Phi}{\partial x_j} \right) + S_p \Phi + S_c, \ \left(x_1, ..., x_m\right) \in G,$$

з крайовими умовами

$$q_1 \frac{\partial \Phi}{\partial \vec{n}} = q_2 \Phi + q_3, \ \left(x_1, ..., x_m\right) \in \partial G,$$

де u_j — компоненти поля швидкостей; τ — коефіцієнт дифузії; ρ — густина середовища; $S_p < 0, S_c$ — джерела зародження Φ; $\vec{n}$ — нормаль до поверхні ∂G; q_1, q_2, q_3 — кусково-неперервні функції від $\left(x_1, ..., x_m\right)$ [9]. В роботі [14] показано, що матрицю вузлів сітки можна розфарбувати мінімальним числом фарб так, щоб кожен вузол $2m+1$ точкового шаблону був іншого кольору, не залежно від розмірності і форми сіткової області. Це еквівалентно розбиттю матриці лінійної системи на $2m+1$ ортогональну підсистему векторів нормалей $A^T\vec{e}_i$, $\vec{e}_i$ — одиничні вектори, $i\in[1:n]$. Позначимо через $I(j)$, $j\in[1:2m+1]$, $m=2$ (або 3), множину вузлів j-го кольору, тоді задача мінімізації норми похибки на k-му кроці, еквівалентна знаходженню вектора $A^T\vec{c}^{(k)}$ в підпросторі $B_n(A)=\boldsymbol{span}\left\{A^T\vec{e}_i\right\}_{i=1}^{n}$, який максимізує функціонал (10) або мінімізує $\left\|\vec{x}^{(k+1)}-\vec{x}^*\right\|_2$. Оптимізацію виконаємо за $2m+1$ кроків.

Покласти $\vec{c}^{(k,0)}=\vec{r}^{(k)}$, для всіх $j=1,...,2m+1$ утворити вектори

$$\vec{c}^{(k,j)}=\alpha_{0,j}\vec{c}^{(k,j-1)}+\sum_{i\in I(j)}\alpha_{i,j}A^T\vec{e}_i\ , \qquad (15)$$

параметри $\alpha_{0,j},\alpha_{i,j}$, знайти з умови мінімізації $\left\|\vec{x}^{(k+1)}-\vec{x}^*\right\|_2$,

$$\vec{x}^{(k+1)}=\vec{x}^{(k)}+\alpha_{0,j}A^T\vec{c}^{(k,j-1)}+\sum_{i\in I(j)}\alpha_{i,j}A^T\vec{e}_i\ .$$

Процедуру повторити $2m+1$ раз. Вектор $\vec{c}^{(k,2m+1)}$ позначити через $\vec{c}^{(k)}$.

Теорема 5. Параметри $\alpha_{0,j},\alpha_{i,j}$, $i\in I(j)$, $j=1,...,2m+1$ знаходяться за формулами

$$\alpha_{0,j}=-\frac{\left(\vec{r}^{(k)},\vec{c}^{(k,j-1)}\right)-\sum\limits_{i\in I(j)}\left\|\vec{r}^{(k)}\right\|_2^2\Big/\left\|A^T\vec{e}_i\right\|_2^2}{\left\|A^T\vec{c}^{(k,j-1)}\right\|_2^2-\sum\limits_{i\in I(j)}\left(\vec{r}^{(k)},\vec{c}^{(k,j-1)}\right)AA^T\vec{c}_i^{(k,j-1)}\Big/\left\|A^T\vec{e}_i\right\|_2^2},$$

$$\alpha_{i,j}=-\left(\vec{r}_i^{(k)}+\alpha_{0,j}\ AA^T\vec{c}_i^{(k,j-1)}\right)\Big/\left\|A^T\vec{e}_i\right\|_2^2\ . \qquad (16)$$

Формули (16) є наслідком мінімізації норми похибки $\left\|\vec{x}^{(k+1)}-\vec{x}^*\right\|_2$ за умови ортогональності векторів $A^T\vec{e}_i$, $i\in I(j)$. Зазначимо, що норми векторів $A^T\vec{e}_i$ обчислюються один раз при формуванні сіткової матриці. У формулі (16) позначено: $\vec{r}_i^{(k)}$, $AA^T\vec{c}_i^{(k,j-1)}$ — i - ті компоненти векторів $\vec{r}^{(k)}$, $AA^T\vec{c}^{(k,j-1)}$.

Висновок. Незалежно від розмірності сіткової матриці за $2m+1$ кроків (для $m=2$ матимемо 5 кроків, для $m=3$ — 7 кроків) оптимізується напрямний вектор на розв'язок системи $A\vec{x}=\vec{b}$ з матрицею розмірності $N\times N$ або $N\times N\times N$, де N — кількість вузлів сітки.

Послідовне застосування алгоритмів 2, 1 дає стійкий швидкозбіжний алгоритм розв’язування систем з довільними дійсними невиродженими погано обумовленими матрицями.

Розробка критерію порівняння на ефективність ітераційних методів розв’язування СЛАР.

За критерій порівняння ітераційних методів розв’язування СЛАР на ефективність вибираються умови: швидкість збіжності, що визначається відношенням норм нев’язок в залежності від номера ітерації; сумарний час необхідний для збіжності із заданою точністю; обсяг необхідних комп’ютерних ресурсів; складність алгоритму тощо [4; 5; 8; 13]. Швидкість збіжності ітераційних методів розв’язування СЛАР залежить від повної характеристики матриці: розрідженості матриці, власних значень та їх розподілу в спектрі, власних векторів, обумовленості, структури матриці та характеристики розв’язку (функції $\left(i, \vec{x}_i^{\,*}\right)$, її періодичності, осциляції, тощо).

Виконаємо тестування методу спряжених градієнтів [4; 5] для визначення ефективності цього методу на симетричних додатно визначених погано обумовлених матрицях: матриці Гільберта порядків 10 — 100 ($\chi\left(A\right) \approx 1.5\cdot 10^{10} - e^{350}$ [11]) і матриці Пуассона порядку 10^4 ($\chi\left(A\right) \approx 4134$), яка отримана при дискретизації двомірного рівняння Пуассона на шаблоні «хрест»; для точних розв’язків: $\vec{u}^{\,*}$ — слабко збуреної дискретної функції $\operatorname{tg}^2\left(\vec{u}_{i+1}^{\,*} - \vec{u}_i^{\,*}\right) \ll 1$ та $\vec{v}^{\,*}$ — періодичної дискретної функції. Метод спряжених градієнтів швидко збігається до розв’язку $\vec{u}^{\,*}$ і не збігається до розв’язку $\vec{v}^{\,*}$ для тестового прикладу з матрицею Гільберта і, навпаки, швидко збігається до розв’язку $\vec{v}^{\,*}$ і повільно до розв’язку $\vec{u}^{\,*}$ для прикладу з матрицею Пуассона.

Таким чином, без ґрунтовного теоретичного аналізу не можна визначити ефективність ітераційного методу лише за окремим тестом. Це одна з причин, другою і найважливішою причиною необ’єктивності тестового порівняння алгоритмів є те, що довільний ітераційний процес на будь-якому кроці можна подати у вигляді процедури (4). Вдало або випадково підібравши початковий вектор $\vec{x}^{(0)}$ і напрямний вектор $\vec{p}^{(0)} = A^T\vec{c}^{(0)}$ колінеарний («майже колінеарний») вектору $\vec{x}^{(0)} - \vec{x}^{\,*}$, можна розв’язок системи $A\vec{x} = \vec{b}$ з довільною невиродженою матрицею A і довільним вектором $\vec{b} \in im\left(A\right)$ отримати за один крок (за незначне число кроків) для одного методу і «поставити в тупик» інший метод, обравши $\vec{p}^{(0)} = A^T \vec{c}^{(0)}$ з умови $\min\limits_{\vec{c} \in B_m^{(0)}} \cos^2 \left\{ A^T \vec{c}, \vec{x}^{(0)} - \vec{x}^* \right\}$.

Висновки. Метод напрямленого пошуку (2), заснований на алгоритмах 1, 2, має високу швидкість збіжності, оскільки для прискорення використовує наближення до розв'язків першої та другої спряжених задач з задачею розв'язування системи $A\vec{x} = \vec{b}$ [15]. Він має спільну основу з методом спряжених напрямів і методом проекцій — для прискорення збіжності за напрямні підпростори обираються підпростори Крилова, будуються спряжені вектори з матрицею $A^T A$.

Метод (2) має принципову відмінність: на кожному k-му кроці знаходить проекцію $\vec{x}^{(k+1)}$ наближення $\vec{x}^{(k)}$ на гіперплощину, що проходить через розв'язок $\vec{x}^*$ системи $A\vec{x} = \vec{b}$. Тому метод (2) збігається для систем з довільними невиродженими матрицями A, оскільки $\left\| \vec{x}^{(k+1)} - \vec{x}^* \right\|_2 < \left\| \vec{x}^{(k)} - \vec{x}^* \right\|_2$. Для прискорення збіжності метод (2) мінімізує кут між напрямним вектором $A^T \vec{c}$ і оптимальним напрямком на розв'язок — вектором $\vec{x}^{(k)} - \vec{x}^*$, не використовує ні матриці переходу ні переобумовлення матриці A (отже, мінімізує потреби комп'ютерних ресурсів).

На початкових кроках, які можна циклічно повторити через m кроків алгоритму 1, будується наближення за алгоритмом 2 до сингулярних прямих з напрямними власними векторами $\vec{x} - \vec{x}^*$ матриці $A^T A$, що дозволяє прискорити ітераційний процес.

Метод (2) стійкий до похибок заокруглення, оскільки обчислення здійснюється в околі власного вектора, що відповідає $\lambda_{\max}\left(A^T A\right)$ (окіл максимальних нев'язок), у той час як ітераційні методи, засновані на мінімізації норми нев'язки, є погано обумовленими в околі розв'язку.

Список використаних джерел:

1. Абрамчук В. С. О перспективности методов направленного поиска решения систем $Ax = f$ с плохо обусловленными матрицами / В. С. Абрамчук // Доп. НАН України. Сер. Б. — 1995. — № 2. — С. 5–7.
2. Абрамчук В. Ефективні ітераційні методи розв'язування систем лінійних рівнянь / В. С. Абрамчук, І. В. Абрамчук, А. Вешемірський // Вісник Львівського університету. Сер. Прикладна математика та інформатика. — 2007. — Випуск 12. — С. 5–12.

3. Абрамчук В. С. Обоснование эффективности методов направленного поиска / В. С. Абрамчук // Доп. НАН України. — 1997. — № 11. — С. 7–12.
4. Деммель Дж. Вычислительная линейная алгебра / Дж. Деммель. — М. : Мир, 2001. — 429 с.
5. Воеводин В. В. Матрицы и вычисления / В. В. Воеводин, Ю. А Кузнецов. — М. : Наука, 1984, — 320 с.
6. Ортега Дж. Введение в параллельные и векторные методы решения линейных систем / Дж. Ортега. — М. : Мир, 1991. — 364 с.
7. Капорин И. Е. О предобуславливании метода сопряженных градиентов при решении дискретных аналогов дифференциальных задач / И. Е Капорин // Дифференц. уравнения. — 1990. — Т. 26. — № 7. — С. 1225–1236.
8. Ильин В. П. Методы бисопряженных направлений в подпространствах Крылова / В. П. Ильин // Сибирский журнал индустриальной математики. — 2008. — Т. XI. — № 4 (36). — С.47–60.
9. Зверев В. Г. Модифицированный полинейный метод решения разностных эллиптических уравнений / В. Г. Зверев // ЖВМ и МФ. — 1998. — Т. 38. — № 9. — С. 1553–1562.
10. Yousef Saad. Iterative Methods for Sparse Linear Systems. — N.Y. : PWS Publ., 1996. — 460 p.
11. Хорн Р. Матричный анализ / Р. Хорн, Ч. Джонсон. — М. : Мир, 1989. — 655 с.
12. Парлетт Б. Симметричная проблема собственных значений. Численные методы / Б. Парлетт. — М. : Мир, 1983. — 384 с.
13. Фомин А. А. Сравнение эффективности высокоскоростных методов решения разностных эллиптических СЛАУ / А. А. Фомин, Л. Н. Фомина // Вестник Томского университета. Сер. Математика и механика. — 2009. — № 2. — С. 71–77.
14. Абрамчук В. С. Итерационные методы направленного поиска решения систем $Ax = f$ с сингулярно-естественным упорядочением переменных / В. С. Абрамчук // Доклады АН Украины. — 1996. — № 8. — С. 4–8.
15. Абрамчук В. С. Сопряженные задачи с задачей решения системы $Ax = f$ / В. С. Абрамчук // Доклады АН Украины. — 1993. — № 1. — С. 5–9.

It is proved that the method of guided search for solving linear systems $A\vec{x} = \vec{b}$, $A \in M_n(R)$ has basic factors of iterative methods efficiency.

Key words: *method of guided search, Krylov's subspace, maximization of Relay's relation.*

Отримано: 15.03.2012