УДК 519.612

В. С. Абрамчук*, канд. фіз.- мат. наук,
І. В. Абрамчук,** старший викладач

*Вінницький державний педагогічний університет
імені М. Коцюбинського, м. Вінниця,
**Вінницький національний технічний університет, м. Вінниця

ПРОБЛЕМИ, МЕТОДИ, АЛГОРИТМИ РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ РІВНЯНЬ З ПОГАНО ЗУМОВЛЕНИМИ МАТРИЦЯМИ

Запропоновано ітераційні методи розв'язування систем $A\vec{x}=\vec{b}$ з погано зумовленими матрицями, базиси напрямних підпросторів яких формуються на принципах оптимальності.

Ключові слова: *погано зумовлені матриці, принципи оптимальності, метод похибок, метод нев'язок.*

Вступ. Не розв'язані проблеми. В ідеалі робота програми розв'язування системи $A\vec{x}=\vec{b}$ повинна завершуватись повідомленням: 1) — система розв'язана з допустимою похибкою, або повідомленням 2) — матриця вироджена [1]. Але між виродженими і невиродженими матрицями існують неперервні варіації [1, 2]. Зазначимо, що в умовах скінчено-розрядної арифметики не можна за беззаперечні критерії оцінки точності розв'язку системи обирати число зумовленості і значення визначника матиці. Наприклад, для матриці Гільберта двадцять першого порядку пакет BLAS/LINPACK видає значення: $\operatorname{cond} A = 1.5\cdot 10^{18}$, $\det A = -3.2\cdot 10^{-125}$.

Виникають дві проблеми: перша, чи необхідно розв'язувати системи з погано зумовленими матрицями, розрізняючи матрицю близьку до патологічно виродженої [1] від погано зумовленої, для якої $\operatorname{cond} A$ — велике але обмежене число. Крім того, немає чіткого поняття, що означає «матриця близька до патологічно виродженої». Результати наведені в роботах [1–4] лише показують, що проблема розв'язування систем $A\vec{x}=\vec{b}$ залежить від $\operatorname{cond} A$, точніше від розподілу власних значень симетричної додатно визначеної матриці A (і якщо A не симетрична, то, напевно, від розподілу власних значень матриці $A^T A$) і від малості $\left|\det A\right|$, що пов'язано із скупченням малих власних значень на периферії.

Крайові задачі, що описують явища розсіювання механічної енергії, вибуху, швидких хімічних реакцій, деформацію конструкцій при перевантаженнях тощо, пов'язані з погано зумовленими розв'язками [3, 5–7]. Тому різницеві рівняння для таких задач мають погано зумовлені матриці.

Друга проблема полягає у тому, яким методам можна довіряти більше: прямим чи ітераційним. На основі аналізу $\operatorname{cond} A$, в роботі [1] зроблений висновок, що якщо точність розв'язання системи прямими методами мала, то ітераційний метод може не дати більш точного результату. Проте більшість існуючих прямих методів не розв'язують систему з матрицею Гільберта порядку, вище сотого. Однак ітераційні методи дають певну інформацію щодо розв'язку такої системи. Крім того, для правих частин, сформованих на основі достатньо гладких розв'язків, ітераційні методи можуть давати наближення до розв'язку з високою точністю.

Третьою проблемою є нестабільність розв'язків системи $A\vec{x} = \vec{b}$ з погано зумовленими матрицями і правими частинами, сформованими на основі векторів з сильно осцилюючими компонентами. Тоді виникає інша проблема: що вважати за розв'язок у скінчено-розрядній арифметиці, оскільки іншого критерію, ніж малість нев'язки немає.

Постановка задачі.

1. Розробити критерії оптимальності і на їх основі побудувати методи формування базисів напрямних підпросторів ітераційних процесів: як локальних на кожній ітерації, так і по сукупності s кроків.
2. Використати алгоритм декомпозиції «розділяй і володій» і розділити задачу розв'язування системи $A\vec{x} = \vec{b}$ на підзадачі малих розмірностей.
3. Обґрунтувати, що ітераційні методи розв'язування систем з погано зумовленими матрицями є більш перспективними, ніж прямі.

Основна частина. 1. Переваги ітераційних методів перед прямими при розв'язуванні систем $A\vec{x} = \vec{b}$ з погано зумовленими матрицями. Застосуємо обернений аналіз для обґрунтування даної гіпотези. Нехай в умовах наближених обчислень і перетворень системи допущені похибки у коефіцієнтах матриці і правої частини:

$$\left(A + \delta A_{об} + \delta A_n\right)\left(\vec{x} + \delta\vec{x}\right) = \vec{b} + \delta\vec{b}_{об} + \delta\vec{b}_n \Rightarrow B\,\vec{y} = \vec{c}$$

і система $B\,\vec{y} = \vec{c}$ розв'язується точно. Запишемо розв'язок системи за правилом Крамера: $y_i = D_i / D_0$, де $D_0 = D_0\left(\delta A_{об}, \delta A_n\right)$, $D_i = D_i\left(\delta A_{об}, \delta A_n, \delta\vec{b}_{об}, \delta\vec{b}_n\right)$, $i \in [1:n]$, $\delta A_{об}$, $\delta\vec{b}_{об}$ — похибки обчислень, δA_n, $\delta\vec{b}_n$ — похибки перетворень. З цих формул випливає, що розв'язок збуреної (зашумленої) системи є дробово-раціональною функцією відносно похибок, причому похибки, що допускаються в коефіцієнтах матриці визначають полюси збуреного розв'язку. Отже, малі похибки, допущені при розв'язуванні системи з погано зумовленими матрицями можуть привести до значних змін компонент розв'язку (особливо за умови $\operatorname{cond} A$ — велике число, $|\det A|$ — мале число). Ітераційні методи, що не вимагають перетворень матриці, можуть забезпечити більш високу точність розв'язку. Крім того в ітераційних методах можна використовувати властивість наближення послідовності похибок $\vec{\varepsilon}^{(k)} = \vec{x}^{(k)} - \vec{x}^{*}$ до сингулярних прямих, що не лише прискорює збіжність ітераційного процесу а й точність розв'язку [9].

2. Ітераційні методи розв'язування систем $A\vec{x} = \vec{b}$ з довільними дійсними невиродженими матрицями

2.1. Принципи оптимальності вибору базису ітераційного процесу. В основу покладені ітераційні процедури, засновані на принципах оптимальності:

1) для довільного вектора наближення $\vec{x}^{(0)}$ існує оптимальний напрямний вектор $A^T\vec{c} = \vec{x}^{(0)} - \vec{x}^{*}$, $A\vec{x}^{*} = \vec{b}$, такий, що розв'язок досягається за один крок: $\vec{x} = \vec{x}^{(0)} + \alpha A^T\vec{c}$, $\alpha = -\dfrac{\left(\vec{r}^{(0)}, \vec{c}\right)}{\left\|A^T\vec{c}\right\|_2^2}$, $\left\|\vec{x} - \vec{x}^{*}\right\|_2^2 = \left\|\vec{x}^{(0)} - \vec{x}^{*}\right\|_2^2 \left(1 - \cos^2\left\{A^T\vec{c}, \vec{x}^{(0)} - \vec{x}^{*}\right\}\right) = 0$. Пошук такого вектора $\vec{c}$ визначається з умови максимізації функції $\cos^2\left\{A^T\vec{c}, \vec{x}^{(0)} - \vec{x}^{*}\right\}$, що еквівалентно максимізації функції $\varphi(\vec{c}) = \left(\vec{r}^{(0)}, \vec{c}\right)^2 \Big/ \left\|A^T\vec{c}\right\|_2^2$;
2) у просторі $\mathbb{R}^n$ довільної скінченої розмірності існує повний базис власних векторів матриці A^TA: $A^T\vec{r}^{(i)} = A^TA\left(\vec{z}^{(i)} - \vec{x}^{*}\right) = \lambda_i\left(\vec{z}^{(i)} - \vec{x}^{*}\right)$, $i \in [1:n]$, кожний з яких визначає оптимальний напрямок на розв'язок з точок $\vec{z}^{(i)} \neq \vec{x}^{*}$, що належать сингулярним прямим.
3) для довільного вектора $\vec{x}^{(0)} \neq \vec{x}^{*}$, $A\vec{x}^{*} = \vec{b}$, і довільного вектора $\vec{p}^{(1)} = A^T\vec{c}^{(1)}$, $\left(\vec{r}^{(0)}, \vec{c}^{(1)}\right) \neq 0$, існує єдина гіперплощина H_1 з векто-

ром нормалі $\vec{p}^{(1)}$, якій належить розв'язок системи $A\vec{x}=\vec{b}$, така, що проекція $\vec{x}^{(0)}$ на неї визначає найближчу в евклідовій метриці точку $\vec{z}^{(1)}$ до $\vec{x}^*$ у напрямку $\vec{p}^{(1)}$: $\left\|\vec{z}^{(1)}-\vec{x}^*\right\|_2=\min\limits_{\vec{x}\in H_1}\left\|\vec{x}-\vec{x}^*\right\|_2$ [8, 9].

4) якщо проекція $\vec{z}^{(1)}$ визначена, то найкраще наближення необхідно шукати у підпросторі, який є перетином гіперплощин $H_1\cap H_2$, де вектор нормалі $\vec{p}^{(2)}=A^T\vec{c}^{(2)}$ гіперплощини H_2 ортогональний до $\vec{p}^{(1)}$, $\left(A^T\vec{c}^{(1)},A^T\vec{c}^{(2)}\right)=0$, $\vec{x}^*\in H_1\cap H_2$. Тобто послідовність напрямних векторів $\left\{A^T\vec{c}^{(i)}\right\}_{i=1}^{s}$ повинна бути ортогональною (псевдо ортогональною з матрицею AA^T) по сукупності s кроків. Позначимо підпростір ортогональних напрямних векторів через H_s — це базис ітераційного процесу за s кроків (розмірність H_s вибирається невеликою).

У локальний базис на кожному k-му кроці необхідно, на основі критеріїв 1), 2), ввести вектор $\vec{p}^{(k,1)}=A^T\vec{r}^{(k)}\in K$. Проте такий вектор може бути далеким від оптимального, оскільки значення функції $\varphi_{k,1}=\cos^2\left\{\vec{p}^{(k,1)},\vec{\varepsilon}^{(k)}\right\}$, $\vec{\varepsilon}^{(k)}=\vec{x}^{(k)}-\vec{x}^*$, може бути малою величиною. Щоб сформувати базис напрямного підпростору, що максимізує функцією φ_k необхідно ввести нові вектори за умови $\varphi_{k,i+1}>\varphi_{k,i}$. Виконаємо обернений аналіз у припущенні, що такий базис побудовано при виконанні умови $\varphi_{k,i+1}>\varphi_{k,i}$. Тоді для такої послідовності напрямних векторів матимемо:

$$\vec{x}^{(k,2)}=\vec{x}^{(k,1)}+\alpha_{k,2}A^T\vec{r}^{(k,1)}=\vec{x}^{(k)}+\alpha_{k,1}A^T\vec{r}^{(k)}+\alpha_{k,2}A^T\left(\vec{r}^{(k)}+\right.$$
$$\left.+\alpha_{k,1}AA^T\vec{r}^{(k)}\right)=\vec{x}^{(k)}+\gamma_1A^T\vec{r}^{(k)}+\gamma_2A^T\left(AA^T\right)\vec{r}^{(k)}.$$

Продовжуючи цей процес, сформуємо локальний базис підпростору Крилова $K_m=span\left\{\vec{r}^{(k)},AA^T\vec{r}^{(k)},\ldots,\left(AA^T\right)^{m-1}\vec{r}^{(k)}\right\}$. Параметри γ_i, $i\in[1:m]$, обчислюються з умови максимізації функції φ_k або мінімізації квадрата норми похибки $\vec{\varepsilon}^{(k+1)}=\vec{\varepsilon}^{(k)}+\sum_{i=0}^{m-1}\gamma_iA^T\vec{u}_i^{(k)}$, де

$\vec{u}_i^{(k)} = \left(AA^T\right)^i \vec{r}^{(k)}$, $i \in \{0,1,...,m-1\}$ (виконання умови $\varphi_{k,i+1} > \varphi_{k,i}$ випливає з того, що $\forall i \in \{0,1,...,m-1\}$ виконується умова $\left(\vec{u}_i^{(k)}, \vec{r}^{(k)}\right) \neq 0$).

Отже з критеріїв оптимальності побудови швидкозбіжного ітераційного процесу розв'язування систем $A\vec{x} = \vec{b}$ з довільними невиродженими матрицями випливає, що необхідно формувати два базиси: локальний K_m на кожному k-му кроці, що визначає наближення до оптимального напрямного вектора $\vec{\varepsilon}^{(k)} = \vec{x}^{(k)} - \vec{x}^*$, і базис H_s (розмірності m і s вибираються невеликими і в процесі збіжності обновлюються).

2.2. Метод похибок. Метод похибок складається з двох процедур: 1) формування локального базису K_m — на виході дістанемо вектор $\vec{c}^{(k)}$, що максимізує функцію φ_k (або еквівалентно мінімізує норму вектора похибки $\vec{\varepsilon}^{(k)}$); 2) формування з векторів $\vec{c}^{(k)}$ ортогонального базису підпростору $H_s = span\left\{A^T\vec{c}^{(k)}\right\}_{k=1}^{s}$.

Обґрунтуємо процедуру збіжності. Оскільки похибка $\vec{\varepsilon}^{(k+1)}$ є лінійною формою відносно вектора параметрів $\vec{\gamma} = \left(\gamma_0, \gamma_1, ..., \gamma_{m-1}\right)$, то квадрат норми вектора похибки є опуклою квадратичною формою, тому мінімум існує і єдиний. Задача пошуку мінімуму еквівалентна розв'язуванню системи рівнянь $B\vec{\gamma} = \vec{d}$ з симетричною додатно визначеною матрицею $B = \left(B_{i,j}\right)_{i,j=0}^{m-1} = \left(A^T\vec{u}_i^{(k)}, A^T\vec{u}_j^{(k)}\right)_{i,j=0}^{m-1}$ і правою частиною $\vec{d} = \left(d_i\right)_{i=0}^{m-1} = -\left(\vec{r}^{(k)}, \vec{u}_i^{(k)}\right)_{i=0}^{m-1}$. Система $B\vec{\gamma} = \vec{d}$ розв'язується методом Холецького. Оскільки базис K_m погано зумовлений, то його розмірність контролюється двома умовами: $\forall i \in [1:m]$: $\cos^2\left\{\vec{c}^{(i)}, \vec{c}^{(i-1)}\right\} < 1 - \varepsilon_1$ (наприклад $\varepsilon_1 = 10^{-12}$) і $L_{i,i} > \varepsilon_2$ (наприклад $\varepsilon_2 = 10^{-10}$), $L_{i,i}$ — діагональні множники Холецького розкладу матриці B. На виході задачі мінімізації дістаємо вектор $A^T\vec{c}^{(k)} = \sum_{i=0}^{m-1} \gamma_i^* A^T\vec{u}_i^{(k)}$, який ортогоналізуємо до векторів підпростору H_s, дістаємо вектор $\vec{v}^{(k)}$: $\forall i \in [1:s]$ $\left(A^T\vec{p}^{(i)}, A^T\vec{v}^{(k)}\right) = 0$, де $A^T\vec{p}^{(i)}$ — базисні вектори підпростору H_s. Наближення до розв'язку системи $A\vec{x} = \vec{b}$ шукаємо у формі

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

Має місце теорема.

Теорема 1. Метод похибок в абсолютно точній арифметиці збігається до розв'язку системи $A\vec{x} = \vec{b}$.

Оскільки алгоритм методу похибок справджується для довільного базису підпростору H_s, то його можна модифікувати і для інших базисів. Для різницевих рівнянь еліптичного типу, для яких існують укрупнені ортогональні підсистеми векторів нормалей $A^T\vec{e}_i$, $i \in I_m$ ($m = 5$ для двовимірних задач, $m = 7$ для тривимірних задач [8]) алгоритм методу похибок запишеться:

$$\min \left\| \vec{x}^{(0)} - \vec{x}^* + \sum\nolimits_{i=1}^{n} \alpha_i A^T \vec{e}_i \right\|_2^2 = \left\| \vec{x}^{(0)} - \vec{x}^* + A^T \vec{\alpha} \right\|_2^2 =$$

$$= \left\| \vec{\varepsilon}^{(0)} \right\|_2^2 + 2\left(\vec{\alpha}, \vec{r}^{(0)}\right) + \vec{\alpha}^T A A^T \vec{\alpha} .$$

Задача знаходження розв'язку системи $A\vec{x} = \vec{b}$, $A \in M_n(\mathbb{R})$, ***rank*** $A = n$, еквівалентна розв'язуванню системи рівнянь $AA^T\vec{\alpha} = -\vec{r}^{(0)}$. Для розв'язування системи можна використовувати метод спряжених градієнтів з симетричною додатно визначеною матрицею AA^T. Проте цей підхід по швидкості збіжності значно поступається алгоритму заснованому на базисах оптимальності (не через те, що матриця AA^T — погано зумовлена, як це трактується в дослідженнях [2, 4], а тому, що базис $\left\{A^T\vec{e}_i\right\}_{i=1}^{n}$ не задовольняє принципам оптимальності).

2.3. Метод нев'язок. Нехай заданий або вибраний за початкове наближення до розв'язку системи $A\vec{x} = \vec{b}$ вектор $\vec{x}^{(0)}$. Обчислимо нев'язку $\vec{r}^{(0)} = A\vec{x}^{(0)} - \vec{b}$. Уточнимо наближений розв'язок $\vec{x}^{(0)}$ введенням поправок (похибок) $\delta\vec{x}$ шляхом процедури $\vec{r}^{(1)} = \vec{r}^{(0)} + \sum\nolimits_{i \in I_m} A^{\langle i \rangle} \delta\vec{x}_i$, де I_m — підмножина множини індексів $\{1, 2, ..., n\}$, розбита на m підмножин, $1 \le m < n$, n — розмірність A, $A^{\langle i \rangle}$ — i-й стовпець матриці A.

Мінімізуємо квадрат норми вектора нев'язки

$$\left\|\vec{r}^{(1)}\right\|_2^2 = \left\|\vec{r}^{(0)} + \sum\nolimits_{i\in I_m} \delta\vec{x}_i A^{\langle i\rangle}\right\|_2^2 = \left\|\vec{r}^{(1)}\right\|_2^2 + 2\sum\nolimits_{i\in I_m} \delta\vec{x}_i \left(\vec{r}^{(0)}, A^{\langle i\rangle}\right) +$$
$$+\sum\nolimits_{i\in I_m} \sum\nolimits_{j\in I_m} \delta\vec{x}_i \delta\vec{x}_j \left(A^{\langle i\rangle}, A^{\langle j\rangle}\right).$$

Оскільки вектор нев'язки є лінійною формою, а квадрат норми вектора нев'язки є опуклою квадратичною формою, то мінімум існує і єдиний, проблема зводиться до розв'язування системи лінійних алгебричних рівнянь з симетричною додатно визначеною матрицею $C = \left(C_{i,j}\right)_{i,j\in I_m} = \left(A^{\langle i\rangle}, A^{\langle j\rangle}\right)_{i,j\in I_m}$ і вектором правих частин $\vec{d} = \left(d_i\right)_{i\in I_m} = -\left(\vec{r}^{(0)}, A^{\langle i\rangle}\right)_{i\in I_m}$.

Після мінімізації дістанемо уточнення наближення до розв'язку системи $A\vec{x} = \vec{b}$: $\vec{x}_i = \vec{x}_i + \delta\vec{x}_i$, $i \in I_m$. Застосувавши алгоритм нев'язок до всіх m підсистем розбиття множини індексів, дістанемо уточнений розв'язок.

Теорема 2. В абсолютно точній арифметиці метод нев'язок збігається до розв'язку системи $A\vec{x} = \vec{b}$.

Як і для методу похибок, так і для методу нев'язок отримується строго спадна послідовність квадратів норм похибок (нев'язок), обмежена знизу, тому при $k \to \infty$ $\left\|\vec{\varepsilon}^{(k)}\right\|_2^2 \to 0$ (аналогічно, $\left\|\vec{r}^{(k)}\right\|_2^2 \to 0$).

Методом нев'язок можна уточнити одночасно всі координати вектора наближення $\vec{x}^{(0)}$: $\vec{x} = \vec{x}^{(0)} + \sum\nolimits_{i=1}^{n} \delta\vec{x}_i \vec{e}_i \Rightarrow \vec{r} = \vec{r}^{(0)} + \sum\nolimits_{i=1}^{n} \delta x_i A\vec{e}_i = = \vec{r}^{(0)} + A\cdot\overrightarrow{\delta x}$, $\overrightarrow{\delta x} = \left(\delta x_1,\ldots,\delta x_n\right)^T$.

Задача мінімізації квадрата норми нев'язки

$$\|\vec{r}\|_2^2 = \left\|\vec{r}^{(0)}\right\|_2^2 + 2\left(A^T\vec{r}^{(0)}, \overrightarrow{\delta x}\right) + \overrightarrow{\delta x}^T A^T A\,\overrightarrow{\delta x}$$

зведеться до розв'язування системи $A^T A\,\overrightarrow{\delta x} = -A^T\vec{r}^{(0)}$ з симетричною додатно визначеною матрицею $A^T A$. Зазначимо, що системи $AA^T\vec{\alpha} = -\vec{r}^{(0)}$ і $A^T A\,\overrightarrow{\delta x} = -A^T\vec{r}^{(0)}$ взаємозв'язані, тому метод нев'язок необхідно використовувати лише для подавлення компонент нев'язок.

2.4. Алгоритм наближення до власного вектора матриці $A^T A$. Уточнення наближення до розв'язку системи $A\vec{x} = \vec{b}$. Нехай відоме найменше власне значення λ матриці $A^T A$ (у протилежному випадку за λ приймемо $\lambda = \left\| A^T \vec{r}^{(0)} \right\|_2^2 \Big/ \left\| \vec{r}^{(0)} \right\|_2^2$, де $\vec{r}^{(0)} = A\vec{x}^{(0)} - \vec{b}$, $\vec{x}^{(0)}$ — вектор наближення до розв'язку $\vec{x}^*$ системи $A\vec{x} = \vec{b}$, отриманий методом похибок). В силу мінімізації норми вектора похибки наближення $\vec{x}^{(0)}$ до розв'язку системи $A\vec{x} = \vec{b}$ знаходиться в околі сингулярної прямої , що відповідає деякому малому власному значенню матриці $A^T A$.

Наближення до власного вектора, що відповідає λ будемо шукати у вигляді

$$\vec{z} - \vec{x}^* = \vec{x}^{(0)} - \vec{x}^* + \sum\nolimits_{i=1}^{m} \alpha_i A^T \vec{c}^{(i)} .$$

Параметри α_i знайдемо з умови мінімізації квадрата норми вектора похибки

$$\vec{\rho} = \left(A^T A - \lambda I\right)\left(\vec{z} - \vec{x}^*\right) = \left(A^T A - \lambda I\right)\left(\vec{x}^{(0)} - \vec{x}^* + \sum\nolimits_{i=1}^{m} \alpha_i A^T \vec{c}^{(i)}\right) =$$

$$= \left(A^T A\left(\vec{z} - \vec{x}^*\right) - \lambda\left(\vec{z} - \vec{x}^*\right)\right) + \sum\nolimits_{i=1}^{m} \alpha_i A^T \left(AA^T \vec{c}^{(i)} - \lambda \vec{c}^{(i)}\right).$$

Позначимо: $\vec{\rho}^{(0)} = A\left(A^T A\left(\vec{z} - \vec{x}^*\right) - \lambda\left(\vec{z} - \vec{x}^*\right)\right) = AA^T \vec{r}^{(0)} - \lambda \vec{r}^{(0)}$, $\vec{\rho}^{(i)} = AA^T \vec{c}^{(i)} - \lambda \vec{c}^{(i)}$. За $\vec{c}^{(i)}$ виберемо вектори з підпростору Крилова $\left\{\vec{r}^{(0)}, AA^T \vec{r}^{(0)}, \ldots, \left(AA^T\right)^{m-1} \vec{r}^{(0)}\right\}$.

Задача мінімізації $\left\| \vec{z} - \vec{x}^* \right\|_2^2$ зведеться до задачі: мінімізувати квадратичну форму

$$\left\| \vec{z} - \vec{x}^* \right\|_2^2 = \left\| A^T A\, \vec{\varepsilon}^{(0)} - \lambda \vec{\varepsilon}^{(0)} \right\|_2^2 + 2\sum_{i=1}^{m} \alpha_i \left(\vec{p}^{(0)}, \vec{p}^{(i)}\right) + \sum_{i=1}^{m}\sum_{j=1}^{m} \left(A^T \vec{p}^{(i)}, A^T \vec{p}^{(j)}\right).$$

що еквівалентно розв'язуванню методом Холецького системи $B\vec{\alpha} = \vec{d}$ з симетричною додатно визначеною матрицею $B = \left(A^T \vec{p}^{(i)}, A^T \vec{p}^{(j)}\right)_{i,j=1}^{m}$ і правою частиною $\vec{d} = -\left(\vec{p}^{(0)}, \vec{p}^{(i)}\right)_{i=1}^{m}$.

Наближенням до власного вектора є вектор $\vec{z} = \vec{x}^{(0)} + \sum\nolimits_{i=1}^{m} \alpha_i A^T \vec{c}^{(i)} = \vec{x}^{(0)} + A^T \vec{c}$, наближенням до розв'язку системи $A\vec{x} = \vec{b}$ буде вектор $\vec{x} = \vec{z} - \dfrac{\left\|\vec{r}\right\|_2^2}{\left\|A^T \vec{r}\right\|_2^2} A^T \vec{r}$, де $\vec{r} = A\vec{z} - \vec{b}$.

2.5. Тестові погано зумовлені матриці та їх характеристики. На основі нижче наведених, погано зумовлених матриць і правих частин, що формувались на основі векторів з слабко і сильно осцилюючими компонентами, проводилась відладка програмних модулів, порівняння запропонованих ітераційних методів з відомими прямими та ітераційними методами [1–4, 10]. Тестово підтверджено, що метод похибок з базисами оптимальності є коректним, швидко збіжним методом. Метод нев'язок можна застосувати після методу похибок, для уточнення.

1. Матриця $\left(A_{i,j} = (2i-1)/(i+j-1)\right)_{i,j=1}^{n}$, — матриця Гільберта, масштабована з метою, щоб діагональні елементи стали одиничними, а матриця не симетричною, $n \in [10:500]$.

$$n = 21:\ \operatorname{cond} A = 1.5 \cdot 10^{18},\ \det A = -1.2 \cdot 10^{-185};$$
$$n = 30:\ \operatorname{cond} A = 3.3 \cdot 10^{18},\ \det A = 0 \text{ (машинний нуль);}$$

Для $n \geq 50$ пакет BLAS/LINPACK ідентифікує матрицю A як сингулярну.

2. Матриця з випадковими значеннями компонент у стовпцях, розподіленими у заданому інтервалі $A^{\langle i \rangle} = rundom(n, -2i, 2i)$.

$$n = 50:\ \operatorname{cond} A = 454.0,\ \det A = 3.1 \cdot 10^{-11};$$
$$n = 100:\ \operatorname{cond} A = 2.4 \cdot 10^{4},\ \det A = -5.1 \cdot 10^{-23};$$
$$n = 500:\ \operatorname{cond} A = 3.0 \cdot 10^{4},\ \det A = 1.3 \cdot 10^{-109}.$$

3. Матриця з одиничними елементами за виключенням правої головної діагоналі $\vec{d} = rundom(n, -20, 20)$.

$$n = 50:\ \operatorname{cond} A = 597.5,\ \det A = 6.9 \cdot 10^{-15};$$
$$n = 100:\ \operatorname{cond} A = 653.3,\ \det A = -2.4 \cdot 10^{-29};$$
$$n = 500:\ \operatorname{cond} A = 3.2 \cdot 10^{4},\ \det A = -3.5 \cdot 10^{-268}.$$

4. Матриця $\left(A_{i,j} = 2^{-i \cdot j}\right)_{i,j=1}^{n}$.

$$n = 10:\ \operatorname{cond} A = 1.5 \cdot 10^{15},\ \det A = -7.1 \cdot 10^{-55};$$
$$n = 20:\ \operatorname{cond} A \text{ — сингулярна матриця, } \det A = 5.4 \cdot 10^{-254}.$$

5. Матриця $A_{i,j=i-1} = 1 - 10^{-3}$ і $A_{i,j} = 1$, при $j \neq i-1$.

$$n = 20:\ \operatorname{cond} A = 4.1 \cdot 10^{5},\ \det A = 9.8 \cdot 10^{-71};$$
$$n = 50:\ \operatorname{cond} A = 2.5 \cdot 10^{6},\ \det A = 3.4 \cdot 10^{-190};$$
$$n = 100:\ \operatorname{cond} A = 1.0 \cdot 10^{7},\ \det A = 0 \text{ (машинний нуль).}$$

Висновки. 1. Запропоновано і теоретично обґрунтовано ітераційні методи розв’язування систем з довільними погано зумовленими матрицями. Тестово підтверджена коректність і ефективність ітераційних методів з базисами, сформованими на принципах оптимальності, їх перевага над прямими методами.

2. Запропоновані методи налаштовані на розв’язування різницевих рівнянь великих порядків, оскільки підпростори K_m, H_s формуються з базисів малих розмірностей.

Список використаних джерел:

1. Райс. Дж. Матричные вычисления и математическое обеспечение / Дж. Райс ; пер. с англ. — М. : Мир, 1984. — 264 с.
2. Воеводин В. В. Матрицы и вычисления / В. В. Воеводин, Ю. А Кузнецов. — М. : Наука, 1984. — 320 с.
3. Деммель Дж. Вычислительная линейная алгебра / Дж. Деммель. — М. : Мир, 2001. — 429 с.
4. Хейгеман Л. Прикладные итерационные методы / Л. Хейгеман, Д. Янг ; пер. с англ. — М. : Мир, 1986. — 448 с.
5. Ватульян А. О. Обратные задачи в механике деформируемого твердого тела / А. О. Ватульян. — М. : Физматлит, 2007. — 224 с.
6. Гилмор Р. Прикладная теория катастроф / Р. Гилмор. — М. : Мир, 1984. — 350 с. (кн.1), 282 с. (кн.2).
7. Mozgovoy A. V. Methods of constructing basis in solving inverse problems / A. V. Mozgovoy, V. S. Abramchuk, I. V. Abramchuk // Functional materials 21. — 2014. — № 4. — P. 457–462.
8. Абрамчук В. С. Комбінований метод розв’язування еліптичних рівнянь / В. С. Абрамчук, І. В. Абрамчук // Математичне та комп'ютерне моделювання. Серія: Фізико-математичні науки : зб. наук. пр. — Кам’янець-Подільський : Кам’янець-Подільський національний ун-т ім. І .Огієнка, 2014. — Вип. 10. — С. 5–16.
9. Абрамчук В. С. Итерационные методы направленного поиска решения систем $Ax = f$ с сингулярно-естественным упорядочением переменных / В. С. Абрамчук // Доп. НАН Украины. — 1996. — № 8. — С. 4–8.
10. Зверев В. Г. Модифицированный полилинейный метод решения разностных эллиптических уравнений / В. Г. Зверев // ЖВМ и МФ. — 1998. — Т. 38. — № 9. — С. 1553–1562.

An iterative method for solving systems $A\vec{x} = \vec{b}$ with ill-conditioned matrices was proposed, the bases of guides subspaces of which are formed on the principle of optimality.

Key words: *ill-conditioned matrices, principle of optimality, error method, residual method.*

Отримано: 22.03.2016