УДК 512.8
Г.Яковлев
Інститут підготовки кадрів Держслужбі зайнятості України

БЛОКОВИЙ МЕТОД ВИРІШЕННЯ СИСТЕМИ ЛІНІЙНИХ АЛГЕБРАЇЧНИХ РІВНЯНЬ

Отриманий чисельний метод вирішення одного класу лінійних зворотних завдань, співвідношення якого справедливі над полем комплексних чисел. Побудовано декілька варіантів обчислювальних схем на базі різних модифікацій цього методу, показаний зв'язок з відомими методами. Експериментальна перевірка підтвердила ефективність і чисельну стійкість методу.

Ключові слова: *лінійні зворотні завдання, недовизначеність системи.*

Побудуємо прямий метод рішення системи лінійних алгебраїчних рівнянь

$$AX = B, \tag{1}$$

де A – квадратна матриця порядку n системи, X та B – відповідно матриці рішень та правих частин розміру $n \times s$. Одержані співвідношення дозволяють реалізувати паралельні обчислення і далі вони будуть узагальнені на випадок прямокутної матриці A.

Хай матриці A, X та B представлені в блоковій формі

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix},\ A_1 = \left(A_{11} \mid A_{12}\right),\ A_2 = \left(A_{21} \mid A_{22}\right),\ X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix},\ B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}.$$

Тут A_{11} – квадратна матриця порядку k $(0 \le k \le n)$, X_1 та B_1 - матриці розміру $k \times s$. Передбачається такий вибір підматриці A_{11}, що $rank(A_{11}) = k$.

Знайдемо *загальне рішення* підсистеми лінійних рівнянь

$$A_1 X^* = B_1 \tag{2}$$

Вочевидь, простір загальних рішень X^* *містить і рішення* X системи рівнянь (1). Представимо X^* у вигляді

$$X^* = X_a^* C^* + X_b^*. \tag{3}$$

Тут X_a^* – матриця розміру $n \times (n-k)$ фундаментальних рішень приведеної системи рівнянь, одержаної з (2)

$$A_1 X_a^* = O_1, \tag{4}$$

C^* - довільна неособлива матриця розміру $(n-k) \times s$, O_1 – нульова $k \times (n-k)$ -матриця.

Матриця X_a^* задає фундаментальну систему рішень системи рівнянь *(4),* матриця X_b^* є приватним рішенням системи рівнянь (2), визначеним виразом

$$A_1 X_b^* = B_1 \tag{5}$$

Знайдемо таку матрицю C з безлічі довільних матриць C^*, щоб відповідне до неї значення X з простору рішень X^* визначало б і рішення системи рівнянь

$$A_2 X = B_2. \tag{6}$$

Це значення X і є рішенням системи рівнянь (1).

Із співвідношення (6) з врахуванням (3) одержимо систему лінійних рівнянь для визначення C

$$DC = G\,, \tag{7}$$

де $D = A_2 X_a^*$, $G = B_2 - A_2 X_b^*$.

Повне рішення системи (1) відповідно до виразу (3) має вигляд

$$X = X_a^* C + X_b^*. \tag{8}$$

Матриці X_a^* та X_b^* у виразах (3) та (8) визначаються неоднозначно, представимо їх у формі, яка дозволить спростити одержані далі вирази:

$$X_a^* = \begin{pmatrix} X_a \\ E_1 \end{pmatrix},\ X_b^* = \begin{pmatrix} X_b \\ O_2 \end{pmatrix}, \tag{9}$$

де E_1 – одинична матриця порядку $n–k$, O_2 – нульова $(n\text{-}k)\times s$-матриця.

Визначена таким чином матриця X_a^* задає нормовану фундаментальну систему рішень, яка може бути одержана почерговим привласненням кожній з вільних змінних $x_{k+1}, x_{k+2}, \ldots, x_n$ системи (4), значення, рівного 1 при нульових значеннях решти вільних змінних.

Тоді, з виразу (4) з врахуванням (9) та з виразу (5) маємо відповідно

$$A_{11}X_a = -A_{12},\ \ A_{11}\, X_b = B_1.$$

Останні два рівняння можна об'єднати у таке:

$$A_{11} X_{ab} = F\ , \tag{10}$$

де X_{ab}=(X_a | X_b) та F=($-A_{12}$ | B_1).

Відповідно до виразу (10) для отримання загального рішення системи рівнянь (2) необхідно вирішити $n–k+s$ систем рівнянь порядку k з матрицею A_{11}, що може бути економно реалізоване її трикутним розкладанням.

Вираз (7) з врахуванням (9) приймає вигляд

$$DC = G\,,\ D = A_{21} X_a + A_{22}\ ,\ G = B_2 \text{-} A_{21} X_b. \tag{11}$$

Відповідно до виразу (10) для отримання загального рішення системи (2) необхідно вирішити $n–k+s$ систем рівнянь порядку k с матрицею A_{11}, що може бути економно реализовано її трикутною декомпозицією. З (11) витікає, що для визначення матриці C необхідно вирішити s систем рівнянь з матрицею D порядку n-k.

Відповідно до виразів (8) та (9) повне рішення X системи (1) може бути представлено у вигляді

$$X_2 = C \quad \text{та} \quad X_1 = X_a C + X_b = X_a X_2 + X_b. \tag{12}$$

Вираз для X_1 в такій формі реалізує модифіковану форму зворотнього ходу для блокової версії методу Гауса, оскільки з нього витікає очевидна рівність

$$A_{11} X_1 = B_1 - A_{12} X_2.$$

Використання останнього виразу вимагає в порівнянні з (12) додатково $s\times k^2$ операцій для вирішення системи рівнянь з матрицею A_{11}, якщо вона була факторизована при рішенні системи рівнянь (10).

Таким чином, отримані співвідношення для вирішення системи рівнянь (1), які приводять до вирішення двох систем рівнянь (10) та (11) меншої розмірності, причому система (11) будується на базі вирішення системи (10).

З матриць X_a, X_b та *C*, отриманих рішенням цих систем, формується загальне рішення системи рівнянь (1) у формі (12), що в цілому вимагає точно стільки ж операцій, скільки і економна реалізація методу Гауса. При цьому якщо s=1, то на останньому кроці при *s=1* для випадку системи рівнянь з квадратною матрицею *A* буде одержано єдине рішення системи рівнянь (1) – вектор $x=X$.

У наведених співвідношеннях X_a та X_b можуть обчислюватись незалежно, що підвищує можливу ступінь паралелизму алгоритму відносно методу Гауса. Можна відзначити також, що в цьому алгоритмі рядкові перетворення класичного методу Гауса сполучаються зі стовпчиковими перетвореннями з матрицею перестановок *C*, таким чином цей метод включає комбінацію двох відомих методів.

Перевагами описаного методу є також можливість різного вибору підматриць, що у деяких випадках може покращити характеристики обчислювального процесу і можливість краще чи з меншими затратами оцінити чисельну стійкість обчислювального процесу при вирішенні багатьох підсистем рівнянь відповідно до схеми метода .

Відзначимо, що *D* – неособлива матриця, якщо *A* та A_{11} – неособливі. Дійсно, з рівнянь (10) та (11) одержимо

$$X = MB,$$

де $M = \begin{pmatrix} A_{11}^{-1} + KL & -K \\ -D^{-1}L & D^{-1} \end{pmatrix}$, $K = A_{11}^{-1}A_{12}D^{-1}$, $L = A_{21}A_{11}^{-1}$.

Матриця *M* , таким чином, є зворотньою до *A* і одержаний вираз представляє її в першій формі Фробеніуса, для якої має місце співвідношення [1]

$$|M| = |A_{11}|\,|D|.$$

Можна також показати що *D* – ермітова матриця, якщо *A* – ермітова та *D* – позитивно-визначена, якщо такою є *A*.

Узагальнення методу

Перевизначена системи лінійних рівнянь, для якої число рядків матриці A більше число стовпців, приводиться до випадку, коли вони рівні, якщо система сумісна (тільки в цьому випадку можна ставити питання про єдине рішення).

Випадок недовизначеної системи, коли число рядків матриці системи менше число стовпців, заслуговує детальнішого дослідження, тим більш, що вираз (2) використовує цей випадок.

В цьому випадку матриця системи рівнянь має вигляд

$$\tilde{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix} = (A \mid A_{*3}), \qquad (13)$$

де A_{*3} - третій стовпець блокової матриці системи та *A* – її квадратна підматриця, що співпадає з матрицею системи (1).

Система рівнянь з розширеною матрицею $\tilde{A}$ зводиться до раніше розглянутої системи з квадратною матрицею *A* перенесенням блоку стовпчиків A_{*3} матриці системи в праву частину (тобто розширенням матриці *B*)

$$\tilde{B} = (B \mid -A_{*3}),$$

Відповідно B_1 та B_2 в даному випадку замінюються матрицями вигляду

$$\tilde{B}_1 = (B_1 \mid -A_{13}) \text{ та } \tilde{B}_2 = (B_2 \mid -A_{23}). \qquad (14)$$

Таким чином, рішення системи рівнянь з матрицею (13) замінюється рішенням системи

$$A\tilde{X} = \tilde{B}.$$

Загальне рішення $\tilde{X}$ можна представити у вигляді

$$\tilde{X} = \left(X \mid X_3\right),$$

де X – рішення системи (1) з квадратною матрицею A та X_3 – рішення підсистеми

$$AX_3 = -A_{*3}$$

що є частиною нормованого фундаментального рішення недовизначеної системи рівнянь з матрицею (13) .

Зіставляючи останні три вирази з виразом (10) бачимо, що даний випадок можна інтерпретувати, як застосування формули (10), яка призначена для визначення загального рішення недовизначеної системи рівнянь, до всієї недовизначеної системи рівнянь з матрицею (13). Цей висновок важливий для побудови модифікованого методу, який розглянутий нижче.

Використання $\tilde{B}_1$ замість B_1 еквівалентно заміні виразу (10) на

$$A_{11}\tilde{X}_{ab} = \tilde{F} \tag{10'}$$

Тут $\tilde{F}$ – розширена форма матриці F відносно виразу (10)

$$\tilde{F} = (F \mid -A_{13})$$

а $\tilde{X}_{ab}$ – розширена форма матриці X_{ab} , яка визначена виразом

$$\tilde{X}_{ab} = \left(X_a \mid \tilde{X}_b\right),$$

де $\tilde{X}_b = (X_b \mid X_c)$ та X_c визначаються виразами

$$A_{11}\tilde{X}_b = \tilde{B}_1,\; A_{11}X_c = -A_{13}.$$

Вираз (11) стосовно даного випадку зберігає структуру, якщо замінити матрицю G розширеною матрицею

$$\tilde{G} = \tilde{B}_2 - A_{21}\tilde{X}_b$$

зберігаючи вигляд (11) для матриці D, або зберігаючи матрицю G і замінюючи матрицю D розширеною:

$$\tilde{D} = A_{21}\tilde{X}_a + \tilde{A}_{22} \tag{11''}$$

де $\tilde{X}_a = (X_a \mid X_c)$ та $\tilde{A}_{22} = (A_{22} \mid A_{23})$

Обидві інтерпретації – (11') та (11'') можуть бути одержані з виразу (9) розширенням субматриць E_2 і O_2 на випадок системи рівнянь з матрицею (13), при цьому побудова загального рішення системи рівнянь (11'') з недовизначеною матрицею $\tilde{D}$ приводить до рішення перетвореної системи рівнянь з початковою матрицею D і правою частиною $\tilde{G}$. Таким чином, питання вибору інтерпретації системи рівнянь з матрицею (13) – приєднання X_c до X_a (тобто розширення блоку A_{*2}) або до X_b (тобто – розширення блоку B) – має формальний характер і приводить до одного результату.

Вираз (12) відповідно також може бути представлений в двох варіантах

$$X_1 = X_aC + \tilde{X}_b \text{ та } X_2 = C \tag{12'}$$

або

$$X_1 = \tilde{X}_a\tilde{C} + X_b \text{ та } X_2 = \tilde{C} \qquad (12'')$$

де $\tilde{C}$ – рішення системи (11") з розширеною матрицею $\tilde{D}$. Порівняння виразів (10)-(12) з одного боку і (10')-(12') або (10'), (11"),(12") з іншого боку показує, що вони співпадають, якщо замінити в них A_{*2} на $\tilde{A}_{*2} = (A_{*2} \mid \tilde{A}_{*3})$ або B на $\tilde{B}$. У обох випадках йдеться про єдиний метод, що використовує для вирішення системи лінійних рівнянь (1) блок стовпців A_{*1} та блок решти стовпців матриці A (A_{*2} – для квадратної матриці A і $\tilde{A}_{*2}$ – для розширеної $\tilde{A}$) або відповідним чином задану матрицю правої частини. Тому надалі не робитимемо відмінностей для цих випадків з матрицею A та $\tilde{A}$.

Модифікації методу для покрокового вирішення системи рівнянь

Модифікуємо алгоритм (10)-(12) для вирішення системи рівнянь (1) таким чином, що етап 1 – обчислення по формулі (10) – виконується без змін і лише 1 раз, тобто для всієї субматриці A_1, а етап (11), (12) повторюється при послідовному збільшенні з деяким кроком h_i числа рядків n_i вирішуваної підсистеми рівнянь до тих пір, поки не буде досягнута межа n, внаслідок чого буде одержано повне рішення системи рівнянь (1). При цьому на кожному кроці, окрім останнього кроку доведеться вирішувати недовизначену систему рівнянь, використовуючи співвідношення (10)-(12) і приведені вище міркування.

Дві властивості, що створюють передумови для багатокрокової реалізації методу (10)-(12), були виявлені при розгляді алгоритму рішення системи рівнянь з матрицею вигляду (13):

1) за допомогою виразів (10)-(12) можна вирішувати недовизначену систему рівнянь, що доводиться робити на кожному кроці модифікованого алгоритму (окрім останнього кроку, де буде отримано єдине рішення якщо матриця А – квадратна та неособлива матриця та s=1);

2) алгоритм (10)-(12) дає те ж рішення, що і вираз (10), якщо в останньому використовувати розширений блок A_{11}. Це дає можливість циклічно повторювати кроки (11), (12). При цьому збільшується порядок отримуваного рішення до отримання повного вирішення системи (1). Знайдене на кожному кроці значення X використовується для формування X_a із збільшеним числом рядків. При цьому число стовпців X на кожному кроці скорочується. З властивості 2 витікає, що після закінчення кожного кроку i виконання (11), (12) буде отриманий такий результат, неначебто матриця рішень $X^{(i)}$ кроку була визначена безпосередньо виразом (10) для $k=n_i$, тому після виконання (11), (12) для чергового кроку i збільшується значення k до досягнутого значення n_i і відповідним чином змінюються межі інших блоків. При цьому необхідне для продовження процесу рішення узгодження розмірів векторів, що беруть участь в операціях обчислювального процесу, забезпечується автоматично.Алгоритм цієї модифікації виглядає таким чином:

Этап 1. Рішення системи (10) для $n_0= k$;

Этап 2. При $n_1= k+1$, n_2., $n_s=n$, обчислюються вирази (11), (12),

де X_a= X(1.. n_i+1; 1..h), X_b= X(1.. n_i+1; h+1.. n_i-h+1), A_{21}= A(n_i+1..n_i+h_i; 1.. n_i), A_{22}= A(n_i+1..n_i+h_i; n_i+1..n_i+h_i) та B_2 замінюється виразом (-A_{23} | B_2), де A_{23}= A(n_i+1..n_i+h_i; n_i+h_i+1.. n). Тут A(p..q; r..s) - субматриця А, яка включає рядки p, p+1..., q та стовпчики r, r+1..., s.

Розглянута модифікація вимагає такого ж числа операцій типу множення, як і найбільш економна реалізації методу Гауса. Працездатна демореалізація даного алгоритму в системі програмування Matlab із узгодженими з (10)-(12) позначеннями приведена нижче.

```
kk=1:k; A(:,n+1)=-b;  %k,h,n - параметри процесу (зовнішні зміні)
X(kk, 1:n-k+1)=A(kk, kk)\-A(kk, k1:n+1);      %Етап 1
for i=k1:h:n-h+1                              % Етап 2
   i_s=i:i+h-1;
   Xa=X (:, 1:h); Xb=X (:, h+1:n-i+2);
   A21=A(i_s, 1:i-1); A22=A(i_s, i_s);
   D=A21*Xa+A22;  G=-A21*Xb-A(i_s, i+h:n+1);
        C=D\G;  X=[Xa*C+Xb; C];
end
```

Отриманий метод вирішення системи рівнянь виконує поетапне обнуління нев'язки – спочатку для перших k рівнянь - вирішенням системи рівнянь по формулі (10) і побудовою рішення по формулі (8), потім циклічно з кроком h_i для інших рівнянь – рішенням (11), (12). На таке ж значення h_i збільшується розміри векторів $X^{(i)}$ отриманих часткових рішень і зменшується їх кількість. Для початкового алгоритму цей процес відбувається за два кроки. Різновид використовуваної модифікації може бути одержаний з використанням безпосередньо виразів (7), (8) без додаткової умови (9):

$$D = A_2 X_a^* \text{ та } G = B_2 - A_2 X_b^*$$

Працездатна демореалізація даного алгоритму для системи Matlab при його додатковій модифікації використанням розширеної матриці системи рівнянь (A | B) має вигляд:

```
kk=1:k; A(:,n+1)=-b;
X(kk, 1:n-k+1)=A(kk, kk)\-A(kk, k1:n+1);
X=[X;eye(n-k+1)];   %приєднання до X одиничної матриці
for i=k1:h:n-h+1
  ih=i+h-1;
  A2=A(i:ih,:);
  Y=X(:,1:h); Z=X(:,h+1:n-i+2);
  D=A2*Y; G=-A2*Z;
        C=D\G;  X(:,1:n-h-i+2)=Y*C+Z;
end
```

У окремому випадку – при $k=0$ та h =1 одержуємо метод, що нагадує метод Перселла [2]. Дійсно, якщо $k=0$, то в останній модифікації на етапі 1 будується тільки одинична матриця, як в згаданому методі, оскільки в цьому випадку $(X_a^* \mid X_b^*)$ вироджується в одиничну матрицю E порядку $n+1$, яка і буде рішенням, внаслідок того, що $X_b^* = 0$, де 0 – вектор розміру n (ми прийняли, що s =1, як в методі Перселла виключно для того, щоб підкреслити схожість методів в даному окремому випадку, взагалі одержаний нами метод не вимагає такого обмеження).
Матриці A_{22}, а також відповідно D та C в даному випадку – скалярні значення, а часткове рішення кожного кроку представляється у вигляді

$$x = Xc + x_1,$$

де x та x_1 - вектори, c – скаляр, X=X_a.. Алгоритмічні відмінності методу Перселла від даного окремого випадку полягають в тому, що в першому вибирається суміжні пари векторів часткових рішень X для побудови наступного сімейства рішень, а в пропонованому методі – пари векторів з фіксованої і змінної позицій, розрахункові формули при цьому співпадають з точністю до такого вибору. Перселла – цей спосіб економний за обчислювальними витратами тільки при $k=1$ і вибір одного з векторів кожної пари з фіксованої, а іншого - із змінної позиції, як в даному методі – цей метод економний і при $k>1$ (зрозуміло, при цьому йдеться вже про вибір групи з k векторів). У методі Перселла позиція підматриці системи (10) змінюється на кожному кроці, тому він був би неефективний за обчислювальними витратами, якби був узагальнений на випадок $k>1$. Пропоновані модифікації є узагальненням не тільки на випадок $k>1$, але і на випадок $h>1$.Алгоритм методу Перселла записується компактно в математичній символіці, але для економної за обчислювальними витратами програмної реалізації вимагає додаткових логічних операцій для виключення тривіальних дій з очевидним результатом, наприклад, множення на 1, до того ж він вимагає більше пам'ять для зберігання матриць X_a^* та X_b^*, чим алгоритм (10)-(12). Для методу (10)-(12) і його першій покроковій модифікації такі тривіальні операції не входять в розрахункові формули.

1. Ф.Р. Гантмахер. Теорія матриць. М., Наука, (1988), с. 56.
2. Д.К. Фаддєєв, В.Н.Фаддєєва. Обчислювальні методи лінійної алгебри. М., Фізматгиз, (1963) , с. 195.