Annual of Konstantin Preslavsky University of Shumen
vol. XXV C, 2024, pp. 3 – 14, https://doi.org/10.46687/VZHP2021

ITERATIVE METHODS FOR SOLVING A LINEAR MATRIX EQUATION*

Vejdi I. Hasanov

ABSTRACT: *In this paper we study iterative methods for solving a linear matrix equation. The considered type of equations appear when applying Newton's method for nonlinear matrix equations. The convergence of the studied methods was investigated. Theoretical results are illustrated by numerical examples.*

KEYWORDS: *Linear matrix equation, Iterative method*

2010 Math. Subject Classification: 15A24

ИТЕРАЦИОННИ МЕТОДИ ЗА РЕШАВАНЕ НА ЛИНЕЙНО МАТРИЧНО УРАВНЕНИЕ

Вежди И. Хасанов

АБСТРАКТ: *В тази статия разглеждаме итерационни методи за решаване на едно линейно матрично уравнение. Разглеждания тип уравнения се появяват при прилагане на метода на Нютон за нелинейни матрични уравнения. Изследвана е сходимостта на изследваните методи. Теоретичните резултати са илюстрирани с числени примери.*

КЛЮЧОВИ ДУМИ: *Линейно матрично уравнение, Итерационен метод*

*This paper is partially supported by Scientific Research Grant RD-08-144/02.02.2024 of the Konstantin Preslavsky University of Shumen

1 Въведение

В тази статия изследваме итерационни методи за матричното уравнение

$$X - A^*XA - B^*XB = Q\,, \tag{1}$$

където A, B са $n \times n$ комплексни матрици, Q е положително определена матрица и с A^* е отбелязана спрегнатата матрица на A.

Уравнение 1) е частен случай на по-общото уравнение

$$X - A_1^*XA_1 - A_2^*XA_2 - \cdots - A_m^*XA_m = Q\,, \tag{2}$$

което е изследвано от Ran и Reurings [1]. В [1] са получени достатъчни условия за съществуване на единствено решение на (2), което е и положително определено. Друг частен случай на уравнение (2) и на уравнение (1) е добре познатото уравнение на Stain:

$$X - A^*XA = Q\,, \tag{3}$$

познато още като дискретно уравнение на Ляпунов.

Едно от приложенията на всяко едно горните матрично уравнение е при прилагане на метода на Нютон за решаване на нелинейни матрични уравнения. На всяка итерация при метода на Нютон за уравнението

$$X + B_1^*X^{-1}B_1 + B_2^*X^{-1}B_2 + \cdots + B_m^*X^{-1}B_m = P$$

се налага да се решава уравнение от вида на (2) [2, 4]. В случай на $m = 2$ следва да се решава уравнение от вида (1) [3], а при $m = 1$ уравнение на Sein [5].

Всяко линейно матрично уравнение може да се сведе до $n^2 \times n^2$ система от линейни уравнения чрез използване на кронекерово произведение. Това еквивалентно представяне е полезно за определяне на условия за съществуване на решения, но от изчислителна гледна точка не е за предпочитане. В случай на $m = 1$, т.е. за уравнението на Stain са предложени ефективни директни методи, чрез разлагане на матрицата A (вижте [6]) или прилагане на итерационни методи [7]. Приближенията на решенията на линейни матрични уравнения чрез итерационни методи се използват за модификации на матода на Нютон, наречени неточни методи на Нютон (Inexact Newton's methods) [8].

Мотивирани от изследванията в [3, 7, 8] разглеждаме итерационни методи за решаване на (1). Освен това провеждаме числени експерименти за илюстриране на теоретичните резултати.

В рамките на тази статия $\mathcal{C}^{n\times n}(\mathcal{R}^{n\times n})$ означава множеството на $n \times n$ комплексни (реални) матрици, $\mathcal{C}^n$ – множеството на n мерните комплексни вектори, а $\mathcal{H}^n$ – множеството на $n \times n$ ермитови матрици. $A > 0$ $(A \geq 0)$ означава че A е ермитова положително определена (полуопределена) матрица. $A > B$ $(A \geq B)$ означава $A - B > 0$ $(A - B \geq 0)$. С $\rho(A)$, $\|A\|$, $\|A\|_F$ и $A \otimes B$ бележим съответно спектралния радиус, спектралната норма $\big(\|A\| = \sqrt{\rho(A^*A)}\big)$ и нормата на Фробениус $\big(\|A\|_F = \sqrt{\sum_{i,j}|a_{ij}|^2}\big)$ на матрица A и кронекерово произведение $\big(A \otimes B = (a_{ij}B)\big)$ на матриците A и B. За матрица $A = (a_1, a_2 \ldots, a_n) \in \mathcal{C}^{n\times n}$, където a_i, $i = 1, 2, \ldots, n$ са стълбовете на A, дефинираме функция $\mathrm{vec}(A) = (a_1^T, a_2^T \ldots, a_n^T)^T \in \mathcal{C}^{n^2}$.

2 Предварителни резултати

В тази точка ще представим някои предварителни резултати.

Лема 1. *[9] Нека A и Q са квадратни матрици.*

(а) Ако $\rho(A) < 1$, то уравнение (3) има единствено решение X_+. $X_+ \geq 0$ $(X_+ > 0)$, когато $Q \geq 0$ $(Q > 0)$.

(б) Ако съществува $P > 0$ такава, че $P - A^*PA > 0$ (≥ 0), тогава $\rho(A) < 1$ $(\rho(A) \leq 1)$.

Забележка 1. *Ако* $\rho(A) < 1$, $Q_1 \leq Q_2$ *(*$Q_1 < Q_2$*) и* P_i, $i = 1, 2$*, са единствените решения съответно на уравненията* $P - A^*PA = Q_i$, $i = 1, 2$*, то от лема 1 имаме* $P_1 \leq P_2$ *(*$P_1 < P_2$*).*

Забележка 2. *В случай на* $\rho(A) < 1$ *единственото решение* X_+ *на уравнение (3) има следното представяне*

$$X_+ = Q + \sum_{j=1}^{\infty}(A^*)^j Q A^j.$$

Ran и Reurings [1] обобщават лема 1 за уравнение (2). Тук ще представим съответните резултати при $m = 2$, т.е., за уравнение (1).

Лема 2 ([1])**.** *Нека* A, B *и* Q *са квадратни матрици и*

$$L = A^T \otimes A^* + B^T \otimes B^*.$$

(a) *Ако* $\rho(L) < 1$*, то уравнение (1) има единствено решение* X_+. $X_+ \geq 0$ $(X_+ > 0)$*, когато* $Q \geq 0$ $(Q > 0)$.

(б) *Ако съществува* $P > 0$ *такава, че* $P - A^*PA - B^*PB > 0$ (≥ 0)*, тогава* $\rho(L) < 1$ $(\rho(L) \leq 1)$.

Нека с $(A, B)_k$ бележим произволен едночлен на A и B от степен k и $\mathcal{M}_{AB}^k$ е множеството на всички такива едночлени. Например:

$$\mathcal{M}_{AB}^1 = \{A, B\};$$
$$\mathcal{M}_{AB}^2 = \{A^2, BA, AB, B^2\};$$
$$\mathcal{M}_{AB}^3 = \{A^3, BA^2, ABA, B^2A, A^2B, BAB, AB^2, B^3\};$$

Забележка 3. *Нека* L *е от лема 2. В случай на* $\rho(L) < 1$ *единственото решение* X_+ *на уравнение (1) има следното представяне*

$$X_+ = Q + \sum_{j=1}^{\infty} \sum_{(A,B)_j \in \mathcal{M}_{AB}^j} (A, B)_j^* Q (A, B)_j.$$

3 Итерационни методи

Съгласно лема 2 уравнение (1) има единствено решение, което е положително определено тогава и само тогава, когато $\rho(L) < 1$, където $L = A^T \otimes A^* + B^T \otimes B^*$. По-нататък ще предполагаме, че $\rho(L) < 1$ и X_+ е единственото решение на (1) и $X_+ > 0$.

Първия метод, който ще разгледаме е базиран на принципа на неподвижната точка:

$$\begin{cases} X_0 \geq Q, \\ X_{k+1} = Q + A^* X_k A + B^* X_k B, \quad k = 0, 1, \ldots. \end{cases} \tag{4}$$

По индукция може да се докаже, че при $X_0 = Q$ (вижте [1, лема 2.5])

$$X_{k+1} = Q + \sum_{j=1}^{k} \sum_{(A,B)_j \in \mathcal{M}_{AB}^j} (A,B)_j^* Q (A,B)_j.$$

Освен това в случай на X_0 такова, че $X_0 - A^* X_0 A - B^* X_0 B \leq Q$, редицата $\{X_k\}$ определена с (4) има свойството

$$X_0 \leq X_1 \leq \cdots \leq X_k \leq \cdots \leq X_+, \tag{5}$$

докато за редицата $\{X_k'\}$, определена с (4) при $X_0 = X_0'$ такова, че $X_0' - A^* X_0' A - B^* X_0' B \geq Q$, имаме

$$X_0' \geq X_1' \geq \cdots \geq X_k' \geq \cdots \geq X_+. \tag{6}$$

Горните свойства следват от лема 2 и следствие 2.4 в [1].

Теорема 1. *За всяко $X_0 \geq Q$ и $k \geq 0$ за матричната редица $\{X_k\}$, определена с метод (4) имаме*

$$\|X_{k+1} - X_+\| \leq (\|A\|^2 + \|B\|^2)\|X_k - X_+\|.$$

Освен това

$$\limsup_{k\to\infty} \sqrt[k]{\|X_k - X_+\|_F} \leq \rho(L) < 1.$$

Доказателство: За решението X_+ и X_{k+1} от (4) получаваме

$$X_{k+1} - X_+ = A^*(X_k - X_+)A + B^*(X_k - X_+)B, \tag{7}$$

откъдето

$$\begin{aligned}\|X_{k+1} - X_+\| &\leq \|A^*(X_k - X_+)A\| + \|B^*(X_k - X_+)B\| \\ &\leq \|A\|^2\|X_k - X_+\| + \|B\|^2\|X_k - X_+\| \\ &= (|A\|^2 + \|B\|^2)\|X_k - X_+\|.\end{aligned}$$

Равенство (7) е еквивалентно на

$$\begin{aligned}\mathrm{vec}(X_{k+1} - X_+) &= (A^T \otimes A^* + B^T \otimes B^*)\mathrm{vec}(X_k - X_+) \\ &= L\mathrm{vec}(X_k - X_+).\end{aligned}$$

Следователно $\mathrm{vec}(X_k - X_+) = L^k\mathrm{vec}(X_0 - X_+)$. Относно евклидовата норма $\|x\|$ на вектор $\left(\|x\| = \sqrt{\sum_i |x_i|^2}\right)$ имаме

$$\|\mathrm{vec}(X_k - X_+)\| \leq \|L^k\|\|\mathrm{vec}(X_0 - X_+)\|,$$

откъдето следва $\|X_k - X_+\|_F \leq \|L^k\|\|X_0 - X_+\|_F$. Оттук получаваме

$$\limsup_{k\to\infty} \sqrt[k]{\|X_k - X_+\|_F} \leq \rho(L).$$

С това теоремата е доказана □

Понеже $\rho(L) < 1$, то метод (4) има R-линейна скорост на сходимост.

Втория метод, който ще разгледаме се състои в решаване на две уравнения на Stein с постоянни матрични коефициенти съответно A и B на всяка итерация:

$$\begin{cases} X_0 \geq Q, \\ Y_k - A^*Y_kA = Q + B^*X_kB, \qquad k = 0, 1, \ldots. \\ X_{k+1} - B^*X_{k+1}B = Q + A^*Y_kA. \end{cases} \tag{8}$$

Теорема 2. *Матричните редици $\{Y_k\}$ и $\{X_k\}$ в метод (8) са добре определени и имат следните свойства:*

(а) за X_0 такова, че $X_0 - A^*X_0A - B^*X_0B \leq Q$,

$$X_0 \leq Y_0 \leq X_1 \leq Y_1 \leq \cdots \leq X_+;$$

(б) за X_0, такова, че $X_0 - A^*X_0A - B^*X_0B \geq Q$,

$$X_0 \geq Y_0 \geq X_1 \geq Y_1 \geq \cdots \geq X_+.$$

Доказателство: По предположение в началото на точката уравнение (1) има решение и то е $X_+ > 0$, т.е.

$$X_+ - A^*X_+A - B^*X_+B = Q. \tag{9}$$

Следователно

$$X_+ - A^*X_+A = Q + B^*X_+B > 0,$$
$$X_+ - B^*X_+B = Q + A^*X_+A > 0.$$

Оттук и лема (2) следва $\rho(A) < 1$ и $\rho(B) < 1$.

Следователно $\{Y_k\}$ и $\{X_k\}$ в (8) са добре определени.

(а) Доказателството ще извършим по индукция. Нека $X_0 \geq Q$ съгласно (8) и $X_0 - A^*X_0A - B^*X_0B \leq Q$. Оттук и (8) имаме

$$Y_0 - X_0 - A^*(Y_0 - X_0)A = Q - X_0 + B^*X_0B + A^*X_0A \geq 0,$$

откъдето съгласно лема 1 следва $X_0 \leq Y_0$.

След почленно изваждане на уравненията

$$X_+ - A^*X_+A - B^*X_+B = Q,$$
$$Y_0 - A^*Y_0A = Q + B^*X_0B,$$

получаваме

$$X_+ - Y_0 - A^*(X_+ - Y_0)A - B^*(X_+ - Y_0)B = B^*(Y_0 - X_0)B \geq 0.$$

Съгласно лема 2 последното уравнение има единствено решение, което е положително полуопределено, т.е. $Y_0 \leq X_+$.

Нека за произволно k са изпълнени $X_k \leq Y_k \leq X_+$. Тогава за X_{k+1} и Y_{k+1} от (8) имаме

$$X_{k+1} - B^* X_{k+1} B = Q + A^* Y_k A, \tag{10}$$

$$Y_k - A^* Y_k A = Q + B^* X_k B, \tag{11}$$

$$Y_{k+1} - A^* Y_{k+1} A = Q + B^* X_{k+1} B. \tag{12}$$

След почленно изваждане на уравненията (11) и (12) от уравнение (10), получаваме съответно:

$$\begin{aligned} X_{k+1} - Y_k - B^*(X_{k+1} - Y_k)B &= B^*(Y_k - X_k)B \geq 0, \\ X_{k+1} - Y_{k+1} - A^*(X_{k+1} - Y_{k+1})A &= A^*(Y_k - X_{k+1})A, \end{aligned}$$

откъдето съгласно лема 1 последователно получаваме $Y_k \leq X_{k+1}$ и $X_{k+1} \leq Y_{k+1}$.

Отново след почленно изваждане на уравнение (9) и от (12), получаваме

$$\begin{aligned} X_+ - Y_{k+1} - A^*(X_+ - Y_{k+1})A - B^*(X_+ - Y_{k+1})B \\ = B^*(Y_{k+1} - X_{k+1})B \geq 0, \end{aligned}$$

откъдето съгласно лема 2 имаме $Y_{k+1} \leq X_+$. С това доказателството на *(а)* е завършено. Доказателството на *(б)* е аналогично на *(а)*. □

4 Числени експерименти

В тази точка провеждаме числени експерименти с предложените методи за пресмятане на единственото положително определено решение на уравнение (1).

Нека $res(X) = \|X - A^*XA - B^*XB - Q\|_\infty$, където използваната норма е $\|A\|_\infty = \max_i \sum_j |a_{ij}|$. За стоп-критерий на двата метода (4) и (8) използваме

$$\|X_{k+1} - X_k\|_\infty \leq tol = 10^{-8},$$

където k е броят на итерациите.

За реализация на метод (8) използваме функцията *dlyap*. Отбелязваме, че би могло да се реализира чрез ортогонални методи. Поради факта, че матричните коефициенти са едни и същи на всяка итерация, то тяхната трансформация следва да се изпълни еднократно (за подробности вижте [6]).

Пример 1. *Разглеждаме уравнение (1) със следните матрични коефициенти*

$$A = \frac{1}{7}\begin{pmatrix} 4 & 1 \\ 3 & 5 \end{pmatrix}, \quad B = \frac{1}{9}\begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 7 & 5 \\ 5 & 10 \end{pmatrix}.$$

За матриците A и B от пример 1 имаме $\rho(A)^2 = 0.8107$,

$$\rho(B)^2 = 0.1809, \quad \rho(A^T \otimes A^* + B^T \otimes B^*) = 0.9884.$$

В таблица 1 са представени резултатите от експериментите за пример 1 с методите (4) и (8).

Таблица 1: Резултати от експериментите за пример 1

Метод (4)	Метод (8)
$k = 1826$	$k = 301$
$\|X_k - X_{k-1}\|_\infty = 9.94e-09$	$\|X_k - X_{k-1}\|_\infty = 9.40e-09$
$res(X_{1826}) = 9.8233e-09$	$res(X_{301}) = 1.3723e-09$

Пример 2. *Разглеждаме уравнение (1) със следните матрични коефициенти*

$$A = \frac{1}{120}\begin{pmatrix} 37 & 13 & 12 \\ -10 & 34 & 12 \\ 11 & -17 & 29 \end{pmatrix}, \quad B = \frac{1}{13}\begin{pmatrix} 5 & 2 & 4 \\ 3 & 7 & 3 \\ 3 & 4 & 5 \end{pmatrix},$$

$$Q = \frac{1}{10}\begin{pmatrix} 12 & 3 & 1 \\ 3 & 22 & 2 \\ 1 & 2 & 6 \end{pmatrix}.$$

За матриците A и B от пример 2 имаме $\rho(A)^2 = 0.1415$,

$$\rho(B)^2 = 0.8631, \quad \rho(A^T \otimes A^* + B^T \otimes B^*) = 0.9680.$$

В таблица 2 са представени резултатите от експериментите за пример 2 с методите (4) и (8).

Таблица 2: Резултати от експериментите за пример 2

Метод (4)	Метод (8)
$k = 589$	$k = 72$
$\|X_k - X_{k-1}\|_\infty = 9.77e-09$	$\|X_k - X_{k-1}\|_\infty = 7.41e-09$
$res(X_{589}) = 9.4580e-09$	$res(X_{72}) = 8.6514e-10$

Пример 3. *Разглеждаме уравнение (1) със следните матрични коефициенти*

$$A = (0.499I - S_1)S_2^{-1}A_1, \;\; B = (0.499I - T_2 + \beta(T_2 - T_1))T_2^{-1}B_1$$

и $Q = I$, където

$$A_1 = (a_{ij}'^{(1)}) \in \mathcal{R}^{n\times n}, \;\; a_{ij}'^{(1)} = a_{ji}'^{(1)} = i^2 + j, \;\; i \le j,$$

$$B_1 = (b_{ij}'^{(1)}) \in \mathcal{R}^{n\times n}, \;\; b_{ij}'^{(1)} = b_{ji}'^{(1)} = (i+2)j, \;\; i \le j,$$

$$S_1 = \mathrm{diag}(0, s_2', s_3', \ldots, s_n'), \quad s_i' = \sum_{j=1}^{i-1} a'(i,j), \;\; i = 2, 3, \ldots, n,$$

$$S_2 = \mathrm{diag}(s''_1, s''_2, \dots, s''_n), \quad s''_i = \sum_{j=i}^{n} a'(i,j), \;\; i = 1, 2, \dots, n,$$

$$T_1 = \mathrm{diag}(0, t'_2, t'_3, \dots, t'_n), \quad t'_i = \sum_{j=1}^{i-1} b'(i,j), \;\; i = 2, 3, \dots, n,$$

$$T_2 = \mathrm{diag}(t''_1, t''_2, \dots, t''_n), \quad t''_i = \sum_{j=i}^{n} b'(i,j), \;\; i = 1, 2, \dots, n.$$

Пример 3 е разгледан в два случая на $beta = 0$ и $beta = 1$, като и в двата случая $n = 40$. За матриците A и B при $\beta = 0$ имаме $\rho(A)^2 = 2.0573e - 05$, $\rho(B)^2 = 0.9950$ и $\rho(L) = 0.9901$, а при $\beta = 1 - \rho(A)^2 = 0.7036$, $\rho(B)^2 = 0.7036$ и $\rho(L) = 0.9901$, където $L = A^T \otimes A^* + B^T \otimes B^*$.

В таблица 3 са представени резултатите от експериментите за пример 3 при $n = 40$ и $\beta = 0$ с методите (4) и (8), а в таблица 4 при $n = 40$ и $\beta = 1$.

Таблица 3: Резултати от експериментите за пример 3 при $\beta = 0$

Метод (4)	Метод (8)
$k = 1892$ $\|X_k - X_{k-1}\|_\infty = 9.92e - 09$ $res(X_{1892}) = 99.8217e - 09$	$k = 3$ $\|X_k - X_{k-1}\|_\infty = 1.04e - 13$ $res(X_3) = 2.2093e - 13$

Таблица 4: Резултати от експериментите за пример 3 при $\beta = 1$

Метод (4)	Метод (8)
$k = 1852$ $\|X_k - X_{k-1}\|_\infty = 9.93e - 09$ $res(X_{1852}) = 9.8287e - 09$	$k = 501$ $\|X_k - X_{k-1}\|_\infty = 9.75e - 09$ $res(X_{501}) = 2.3895e - 09$

При проведените експерименти се наблюдава, че с метод (8) са необходими в пъти по-малкия брой на итерации, отколкото с метод (4) за достигане на желаната точност. Предимствата на метод (8) се засилват при малки стойности на $\rho(A)$ или $\rho(B)$, а $\rho(A^T \otimes A^* + B^T \otimes B^*)$ е близко до 1.

ЛИТЕРАТУРА:

[1] A.C.R. Ran, M.C.B. Reurings, The symmetric linear matrix equation, *Electron. J. Linear Al.*, 9, (2002), 93-107.

[2] P. Liu, S. Zhang, Newton's method for solving a class of nonlinear matrix equations, *J. Comput. Appl. Math.*, 256, (2014), 254–267.

[3] V.Hasanov, S. Hakkaev, Newton's method for a nonlinear matrix equation, *C. R. Acad. Bulgare Sci.*, 68(8) (2015), 973-982.

[4] V. Hasanov, S. Hakkaev, Convergence analysis of some iterative methods for a nonlinear matrix equation, Computers and Mathematics with Applications, 72 (2016), 1164-1176.

[5] C.-H. Guo, P. Lancaster, Iterative Solution of Two Matrix Equations, *Math. Comput.*, 68, (1999), 1589-1603.

[6] Kh.D. Ikramov, *Numerical solution of Matrix Equations*, Moscow, Nauka, 1985. (in Russion).

[7] B. Zhou, J. Lam, G.-R. Duan, On Smith-type iterative algorithms for the Stein matrix equation, *Appl. Math. Letters,* 22 (2009), 1038-1044.

[8] Y.-H. Gao, Z.-Z. Bai, On inexact Newton methods based on doubling iteration scheme for non-symmetric algebraic Riccati equations, *Numer. Linear Algebra Appl.*, 18:325-341, 2011.

[9] P. Lancaster, M. Tismenetsky, *The Theory of Matrices*, 2nd ed. San Diego (CA), Academic Press, 1985.

Вежди Исмаилов Хасанов

Шуменски университет "Епископ Константин Преславски"

Факултет по математика и информатика

Лаборатория по математическо моделиране

Шумен, България

v.hasanov@shu.bg