J. Appl. Math. & Informatics Vol. **30**(2012), No. 1 - 2, pp. 201 - 210
Website: http://www.kcam.biz

MODIFIED MULTIPLICATIVE UPDATE ALGORITHMS FOR COMPUTING THE NEAREST CORRELATION MATRIX†

JUN-FENG YIN* AND YUMEI HUANG

Abstract. A modified multiplicative update algorithms is presented for computing the nearest correlation matrix. The convergence property is analyzed in details and a sufficient condition is given to guarantee that the proposed approach will not breakdown. A number of numerical experiments show that the modified multiplicative updating algorithm is efficient, and comparable with the existing algorithms.

AMS Mathematics Subject Classification : 65F10, 65F35, 65F50.
Key words and phrases : Multiplicative update algorithm, nonnegative matrix factorization, nearest correlation matrix.

1. Introduction

Correlation matrix plays a very important role in a large number of areas in finance and risk management, e.g., in the applications of the BGM interestrate option models, credit-derivative pricing or credit risk management, stress-testing and scenario analysis for market risk management purpose [7].

Since the (i, j) entry of the correlation matrix represents the correlation coefficient between two financial products, the matrix should be symmetric positive semidefinite with unit diagonal. However, the correlation matrix obtained from practical finance applications is real symmetric, but no more positive semidefinite due to some constraints or limitations of information. The existence of the negative eigenvalue will lead to the breakdown of standard approaches in financial analysis.

Thus, in practical applications from finance and risk management, assume that $A \in \mathbb{R}^{n\times n}$ is a real symmetric matrix, it is interesting to look for a nearest

---

Received April 18, 2011. Revised July 15, 2011. Accepted July 16, 2011. *Corresponding author. †This work was supported by NSFC (Grant No. 10801106), the Fundamental Research Funds for the Central Universities (No. lzujbky-2009-49) and the Fundamental Research Fund for Physics and Mathematics of Lanzhou University (No. LZULL200909).

correlation matrix $H \in \mathbb{R}^{n\times n}$ to minimize

$$f(H) = \frac{1}{2}\big\|A - H\big\|_F^2, \tag{1}$$

where H is a symmetric positive semidefinite matrix with unit diagonal. Here, $\|\cdot\|_F$ represents the Frobenius norm of the corresponding matrix.

This problem attracted great interest of a number of researchers in past decade [1, 4, 5, 6, 9, 11]. A general methodology for a valid correlation matrix was firstly presented in [6], which include the hypersphere decomposition method and the spectral decomposition method. These methods were further improved in [9]. In addition, Higham presented a modified alternating projections method and analyzed its convergence in detail in [5]. For the details of more algorithms, we refer readers to [10, 11].

It is well known that H has a decomposition of the form $H = BB^T$ with $\text{rank}(B) = \text{rank}(H)$, since H is symmetric semi-definite. From the viewpoint of matrix factorization, an equivalent problem is considered instead of problem (1) in this paper as follow: given a symmetric matrix $A \in \mathbb{R}^{n\times n}$, we are looking for the matrices $B \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{n\times n}$ to minimize the cost function

$$f(B, C) = \frac{1}{2}\big\|A - BC\big\|_F^2, \tag{2}$$

with the constraints $B = C^T$ and BC has unit diagonal. Then, $H = BC = BB^T$ is symmetric positive semidefinite with unit diagonal, and thus a solution of the original problem (1).

By simply deduction, the gradients of the function $f(B, C)$ with respect to B and C are given by

$$\text{grad}(B) := \frac{\partial f(B, C)}{\partial B} = (BC - A)C^T, \tag{3}$$

$$\text{grad}(C) := \frac{\partial f(B, C)}{\partial C} = B^T(BC - A). \tag{4}$$

By the Karush-Kuhn-Tucker (KKT) optimality condition, it is clear that the cost function $f(B, C)$ is minimized in the case of $\text{grad}(B) = 0$ and $\text{grad}(C) = 0$.

In order to minimize the cost function (2) with respect to B and C, a general method alternatively updating B and C respectively with their corresponding gradients is given as follow: given an arbitrary initial guesses $B^{(0)}$ and $C^{(0)}$, compute

$$\begin{cases} B_k = B_{k-1} - \zeta \text{grad}(B_{k-1}), \\ C_k = C_{k-1} - \eta \text{grad}(C_{k-1}), \end{cases} \quad k = 1, 2, \ldots \tag{5}$$

until converges where $\text{grad}(B)$ and $\text{grad}(C)$ are the gradients defined by (3) and (4), ζ and η are positive parameters. It is obvious that as long as ζ and η are sufficiently small, the update should reduce the object function $\frac{1}{2}\|A - BC\|_F^2$ and finally obtain the approximate optimal solution of the original problem (1). This idea of alternatively updating between two directions can also be found in the Hermitian and skew-Hermitian splitting iteration method for solving non-Hermitian positive-definite systems of linear equation, see [2, 3] for the detail.

As a special case, we design a modified multiplicative updating (MMU) algorithm for the solution of problem (2) based on multiplicative updating (MU) algorithm proposed by Lee and Seung [8]. The convergence performance of the new algorithm is discussed in detail. We also give a sufficient condition to guarantee that the new algorithm will not breakdown. Numerical experiments further show that modified multiplicative updating algorithm is efficient, and comparable with the existing algorithms.

This paper is organized as follow. In the next section, we present the modified multiplicative updating algorithm and analyze its properties. Numerical experiments for some model matrices are presented in Section 3 and finally, the conclusion and future work are presented in Section 4.

2. Modified multiplicative updating algorithm

In this section, a modified multiplicative updating algorithm is proposed for computing the nearest correlation matrix to minimize the cost function (2), and then its properties are discussed in detail.

Before going on, we first briefly review the multiplicative updating algorithm proposed by Lee and Seung [8] and its properties. Multiplicative updating algorithm is a fast and effective method for computing the non-negative matrix factorization, and thus widely used in image processing, text mining, spectral data analysis and air quality analysis.

Given a nonnegative $A \in \mathbb{R}^{n\times n}$, the multiplicative updating algorithm can minimize $\|A - BC\|_F^2$ in terms of two nonnegative matrices B and C by the following update rules:

$$B_{ij} \leftarrow B_{ij}\frac{(AC^T)_{ij}}{(BCC^T)_{ij}}, \quad C_{ij} \leftarrow C_{ij}\frac{(B^TA)_{ij}}{(B^TBC)_{ij}}. \tag{6}$$

When A, B and C are nonnegative matrices, it is obvious from (6) that the elements of B and C will increase when $(BC)_{ij} < A_{ij}$ and decrease when $(BC)_{ij} > A_{ij}$ for $1 \le i, j \le n$. It was proved in [8] that the cost function $\frac{1}{2}\|A - BC\|_F^2$ is nonincreasing under this update rules (6), and it is invariant under these updates if and only if B and C are at a stationary point of the distance.

The multiplicative updating algorithm can be described as follow.

Algorithm 2.1. The multiplicative updating algorithm

1. Choose B_0 and C_0.
2. For $j = 1, 2, \ldots,$ until convergence, Do*:*
3. Compute $W_k = \left(AC_{k-1}^T\right) ./ \left(B_{k-1}C_{k-1}C_{k-1}^T\right)$*;*
4. Compute $B_k = B_{k-1}. * W_k$*;*
5. Compute $V_k = \left(B_k^T A\right) ./ \left(B_k^T B_k C_{k-1}\right)$*;*
6. Compute $C_k = C_{k-1}. * V_k$*;*

7. EndDo

Denote $.*$ and $./$ the elementwise multiplication and elementwise division between two matrices respectively, the multiplicative updating rule (6) can be written as follow:

$$\begin{aligned} B &= B.*\left(AC^T\right)./\left(BCC^T\right) = B.*\left(BCC^T - \mathrm{grad}(B)\right)./\left(BCC^T\right) \\ &= B - B./\left(BCC^T\right).*\mathrm{grad}(B) \end{aligned}$$

and

$$\begin{aligned} C &= C.*\left(B^TA\right)./\left(B^TBC\right) = C.*\left(B^TBC - \mathrm{grad}(B)\right)./\left(B^TBC\right) \\ &= C - C./\left(B^TBC\right).*\mathrm{grad}(C). \end{aligned}$$

Thus, the multiplicative updating algorithm actually alternatively update B and C with their corresponding gradients respectively at each iteration step as follow:

$$\begin{cases} B_{ij} = B_{ij} - \zeta_{ij}[\mathrm{grad}(B)]_{ij}, \\ C_{ij} = C_{ij} - \eta_{ij}[\mathrm{grad}(C)]_{ij}, \end{cases} \tag{7}$$

where $\zeta_{ij} = B_{ij}/(BCC^T)_{ij}$ and $\eta_{ij} = C_{ij}/(B^TBC)_{ij}$ are the chosen parameters, $1 \le i,j \le n$. From the convergence analysis in [8], such a choice for these parameters can guarantee the convergence of multiplicative updating algorithm when A, B and C are nonnegative.

In the following, we present a modified multiplicative update algorithm for computing the nearest correlation matrix H in terms of B and C to minimize the cost function of problem (2), combined with two constraints $B = C^T$ and BC has unit diagonal.

In order to satisfy the first constraint $B = C^T$, we define

$$B = C^T = \frac{B + C^T}{2},$$

after the multiplicative updating (7), so that the result matrix $H = BC = BB^T$ is symmetric semidefinite.

Let D be a diagonal matrix whose diagonal is the same as BB^T. It is obvious that all the elements are positive. Then, by updating

$$B_{\mathrm{new}} = D^{-1/2}B, \quad C_{\mathrm{new}} = CD^{-1/2},$$

where the elements of $D^{1/2}$ are the square root of those of D, the result matrix

$$H_{\mathrm{new}} = B_{\mathrm{new}}C_{\mathrm{new}} = D^{-1/2}BB^TD^{-1/2},$$

has unit diagonal. Note that in this case, $B_{\mathrm{new}} = C_{\mathrm{new}}^T$ is still satisfied.

More precisely, the modified multiplicative updating algorithm can be described in detail as follow.

Algorithm 2.2. The modified multiplicative updating algorithm

1. Choose B_0 and C_0.
2. For $k = 1, 2, \ldots,$ until convergence, Do*:*
3. Compute $W_k = (AC_{k-1}^T)./(B_{k-1}C_{k-1}C_{k-1}^T);$
4. Compute $B_k = B_{k-1}. * W_k;$
5. Compute $V_k = (B_k^T A)./(B_k^T B_k C_{k-1});$
6. Compute $C_k = C_{k-1}. * V_k;$
7. Set $B_k = (B_k + C_k^T)/2;$
8. Compute $D = \mathrm{diag}(B_k B_k^T);$
9. Set $B_k = D^{-1/2} B_k$
10. Set $C_k = B_k^T;$
11. EndDo

It is seen that even the initial matrices B_0 and C_0 are chosen randomly, the condition $B_k = C_k^T$ is still satisfied after the 10th step of the algorithm 2.2. Consequently, the result matrix $H_k = B_k C_k$ is a symmetric positive semidefinite matrix with unit diagonal, and thus a valid correlation matrix.

It is clear that the algorithm 2.2 is also an alternative multiplication updating method based on the gradients in terms of B and C respectively. Thus, the convergence property of algorithm 2.2 is given as follow.

Proposition 2.1. *Assume that $A \in \mathbb{R}^{n\times n}$ is a real symmetric matrix, $H^* = B^*C^* \in \mathbb{R}^{n\times n}$ is the result matrix given by the algorithm 2.2. Then, H^* a correlation matrix, and satisfied*

$$\frac{1}{2}\|A - H^*\|_F^2 = \min_{H\in\mathcal{H}} \frac{1}{2}\|A - H\|_F^2.$$

where $\mathcal{H}$ is the set of all correlation matrix.

Proof. From the algorithm 2.2, it is clear that

$$H^* = B^*C^* = B^*(B^*)^T,$$

is symmetric semidefinite with unit diagonal, thus it is a correlation matrix.

Note that the algorithm 2.2 is an alternative multiplication updating method with the gradients defined in (3) and (4), the algorithm converges if and only if both the conditions $\mathrm{grad}(B) = 0$ and $\mathrm{grad}(C) = 0$ are satisfied. In this case, we have

$$\frac{\partial \mathrm{grad}(B)}{\partial B} = CC^T \geq 0, \quad \frac{\partial \mathrm{grad}(C)}{\partial C} = BB^T \geq 0.$$

Thus, the cost function $f(B, C)$ is minimized and H^* is the nearest correlation matrix. □

Then, we consider the conditions to guarantee that the modified multiplicative updating algorithm will not breakdown when A, B and C are nonnegative. Here and in the following, $X \geq 0$ ($X > 0$) means all the elements of X are nonnegative (positive).

Assume that A, B and C are nonnegative, some properties of the matrix generated in modified multiplicative updating algorithm 2.2 are presented as follow.

Lemma 2.1. *Assume that $A \geq 0$, $B_0 \geq 0$, $C_0 \geq 0$ and both B_0 and C_0 have full rank, then*

(1) *both B_k and C_k in modified multiplicative updating algorithm 2.2 are nonnegative and have full rank, $k = 1, 2, \ldots, n$;*
(2) *both $B_k B_k^T$ and $B_k^T B_k$ are positive and full rank, $k = 1, 2, \ldots, n$.*

Proof. From line 3 of algorithm 2.2, it is seen that W_k is nonnegative when A, B_{k-1} and C_{k-1} are nonnegative, which results in B_k is nonnegative. Similarly, both V_k and C_k are nonnegative for $k = 1, 2, \ldots, n$. It is obvious that both B_k and C_k are nonnegative after the updating of line 7 and line 8 in algorithm 2.2.

It is also seen that both B_k and C_k have full rank if both B_{k-1} and C_{k-1} have full rank from line 3 and 5 of algorithm 2.2.

On the other hand, since B_k has full rank, B_k has not zero row or column. Thus, we have

$$(B_k B_k^T)_{ij} > 0, \quad 1 \leq i, j \leq n.$$

So does for $B_k^T B_k$. Since the determinant of $B_k B_k^T$ and $B_k^T B_k$ are the square of that of B_k, and thus not equal to 0. This indicate that both $B_k B_k^T$ and $B_k^T B_k$ have full rank, $k = 1, 2, \ldots, n$. □

Denote $H_k = B_k B_k^T$, $k = 1, 2, \ldots, n$. Under the assumptions of lemma 2.1, it also can be seen that H_k are full rank and nonnegative for $k = 1, 2, \ldots, n$.

It is clear that the modified multiplicative updating algorithm 2.2 will breakdown when $B_k C_k C_k^T$ or $B_k^T B_k C_{k-1}$ has zero elements. The following proposition gives a sufficient condition to guarantee that the modified multiplicative updating algorithm 2.2 will not breakdown.

Proposition 2.2. *If $A \geq 0$, $B_0 \geq 0$ and $C_0 \geq 0$ and both B_0 and C_0 have full rank, then the modified multiplicative updating algorithm 2.2 will not breakdown.*

Proof. From lemma 2.1, it is clear that B_k and C_k are full rank and nonnegative, $k = 1, 2, \ldots, n$. Since $C_k = B_k^T$, we have $B_k C_k$ is positive and full rank from lemma 2.1. It follows that $B_k C_k C_k^T$ is positive, and the modified multiplicative updating algorithm 2.2 does not breakdown in line 3.

Since B_k is nonnegative and has full rank, then $B_k^T B_k$ is positive and full rank from lemma 2.1. It follows that $B_k^T B_k C_{k-1}$ is positive, and the modified multiplicative updating algorithm 2.2 does not breakdown in line 5. □

If the matrices A, B and C are not nonnegative, it could happen that $B_k C_k C_k^T$ or $B_k^T B_k C_{k-1}$ has some zero elements. In this case, we can introduce a positive parameter $\epsilon > 0$, so that the update rules in line 3–6 of algorithm 2.2 can be replaced by

$$B_{ij} \leftarrow B_{ij} \frac{(AC^T)_{ij}}{(BCC^T)_{ij} + \epsilon}, \quad C_{ij} \leftarrow C_{ij} \frac{(B^T A)_{ij}}{(B^T BC)_{ij} + \epsilon}.$$

However, the choice of ϵ should be investigated from practical application and it is difficult to find the optimal parameter theoretically.

3. Numerical Results

In this section, a number of numerical experiments are presented to show the convergence behavior of modified multiplicative updating algorithm.

All experiment were performed on a PC-Pentium(R) using Matlab 6.5. The initial matrix B_0 and C_0 are chosen randomly and nonnegative. Denote $r_k = \|A - B_k C_k\|_F$, the stopping criterion is

$$\|r_k - r_{k-1}\| < \varepsilon$$

where $\varepsilon = 10^{-4}$, or the number of maximal iteration, e.g., 100, is reached.

The first example $A \in \mathbb{R}^{4\times 4}$ is taken from [4], where the modified alternatively project (MAP) method is proposed.

Denote H_{MAP} and H_{MMU} be the nearest matrix obtained by modified alternatively project method [4] and the modified multiplicative updating algorithm respectively. The original matrix A and the computed correlation matrices are listed in Table 1, as well as the corresponding Frobenius norms. The eigenvalues of A, H_{MAP} and H_{MMU} are illustrated in Table 2.

TABLE 1. The matrices A, H_{MAP} and H_{MMU}.

A			H_{MAP}			H_{MMU}		
1	1	0	1.0000	0.7607	0.1573	1.0000	0.7410	0.0978
1	1	1	0.7607	1.0000	0.7607	0.7410	1.0000	0.7408
0	1	1	0.1573	0.7607	1.0000	0.0978	0.7408	1.0000
			$\|A - H_{\text{MAP}}\|_F = 0.5278$			$\|A - H_{\text{MMU}}\|_F = 0.5363$		

TABLE 2. The eigenvalues of A, H_{MAP} and H_{MMU}.

A	H_{MAP}	H_{MMU}
-0.4142	-0.000013	0.00000031
1.0000	0.8427	0.9021
2.4142	2.1573	2.0978

From Tables 1 and 2, it is observed that H_{MMU} is symmetric semidefinite, which shows that the modified multiplicative updating algorithm is efficient, and can find a nonnegative correlation matrix.

From Table 1, it is seen that H_{MAP} is relatively closer to A than H_{MMU} from the view point of the F-norm. However, it is found that H_{MAP} still has small negative eigenvalue while all the eigenvalues of H_{MMU} are positive. Thus, H_{MMU} is better than H_{MAP} in the sense of nonnegative property.

Further, we depict the curve of Frobenius norm between A and computed approximate matrix versus the number of iteration in Figure 1.

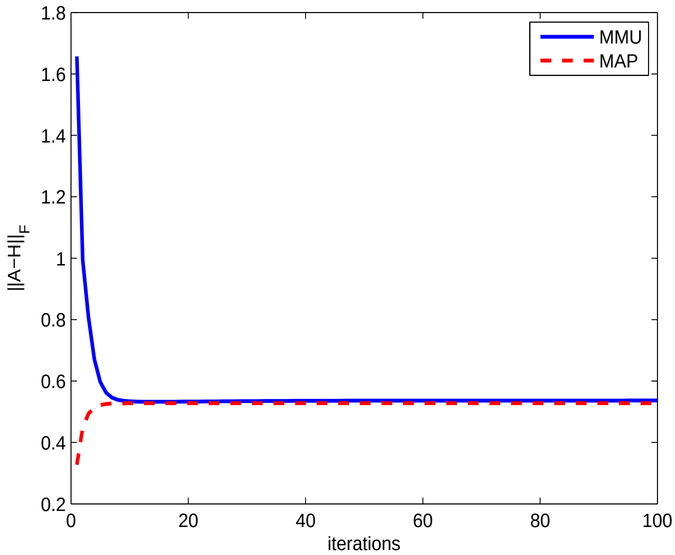

FIGURE 1. The curves of $\|A - B_k C_k\|_F$ versus the number of iteration of MMU and MAP method

From Figure 1, it is seen that the modified multiplicative updating algorithm converges monotonically, e.g., $\|A - B_k C_k\|_F$ is decreasing monotonically and converges to a unique solution. On the other hand, it is observed that $\|A-H_k\|_F$ of MAP method proposed by Higham is increasing monotonically and converges to a unique solution.

The second example

$$A = \begin{pmatrix} 1.00 & 0.50 & 0.50 & 0 & 0 \\ 0.50 & 1.00 & 0.84 & 0.84 & 0.84 \\ 0.50 & 0.84 & 1.00 & 0.84 & 0.84 \\ 0 & 0.84 & 0.84 & 1.00 & 0.84 \\ 0 & 0.84 & 0.84 & 0.84 & 1.00 \end{pmatrix},$$

is taken from [1] while the obtained nearest matrix in this paper is

$$H = \begin{pmatrix} 1 & 0.4964 & 0.5008 & 0.0011 & 0.0050 \\ 0.4964 & 1 & 0.8819 & 0.7317 & 0.7363 \\ 0.5008 & 0.8819 & 1 & 0.7272 & 0.7305 \\ 0.0011 & 0.7317 & 0.7272 & 1 & 0.8432 \\ 0.0050 & 0.7363 & 0.7305 & 0.8432 & 1 \end{pmatrix},$$

with $\|A - H\|_F = 0.3131$.

Using modified multiplicative updating algorithm, the computed approximate nearest matrix is

$$H_{\mathrm{MMU}} = \begin{pmatrix} 1.0000 & 0.4741 & 0.4731 & 0.0165 & 0.0240 \\ 0.4741 & 1.0000 & 0.8414 & 0.8085 & 0.8105 \\ 0.4731 & 0.8414 & 1.0000 & 0.8072 & 0.8094 \\ 0.0165 & 0.8085 & 0.8072 & 1.0000 & 0.8433 \\ 0.0240 & 0.8105 & 0.8094 & 0.8433 & 1.0000 \end{pmatrix},$$

with $\|A - H_{\mathrm{MMU}}\|_F = 0.1107$. This further verified that the MMU algorithm is efficient, since the obtained object function is smaller than that of the method in [1].

Moreover, the eigenvalues of these matrices are listed Table 3.

TABLE 3. The eigenvalues of A, H and H_{MMU}.

A	H	H_{MMU}
-0.0859	0.1031	0.0026
0.1600	0.1182	0.1567
0.1600	0.1568	0.1586
1.1477	1.1881	1.1243
3.6182	3.4338	3.5578

From Table 3, it is observed that H_{MMU} is symmetric definite (nearly semi-definite), and thus a valid nonnegative correlation matrix.

4. Conclusion

In this paper, a modified multiplicative updating algorithm for computing the nearest correlation matrix was proposed and the convergence property of the new algorithm was analyzed. Numerical experiments examined the efficiency of proposed algorithm. Future work on how to accelerate the convergence speed will be further studied.

Acknowledgments

We would like to thank the referees for valuable remarks, which helped greatly to improve the paper.

REFERENCES

1. A. A. Al-Subaihi, *Simulating Correlated Multivariate Pseudorandom Numbers*, J. Stat. Soft., vol. **9**(2004), 1–20.
2. Z.-Z. Bai, G. H. Golub, L.-Z. Lu, J.-F. Yin, *Block triangular and skew-Hermitian splitting methods for positive-definite linear systems*, SIAM Journal on Scientific Computing, **26**(2005), 844–863.
3. Z.-Z. Bai, G. H. Golub, M. K. Ng, *Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems*, SIAM Journal on Matrix Analysis and Applications, **24**(2003), 603–626.

4. N. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra and its Applications, **103**(1988), 103–118.
5. N. Higham, *Computing the nearest correlation matrix–a problem from finance*. IMA Journal of Numerical Analysis, **22**(2002), 329–343.
6. P. Jackel and R. Rebonato, *The most general methodology for creating a valid correlation matrix for risk management and option pricing purposes*, Journal of Risk, **2**(1999), 17–28.
7. P. H. Kupiec, *Stress testing in a value at risk framework*. Journal of Derivatives, **6**(1998), 7–24.
8. D. D. Lee and H. S. Seung, *Algorithms for non-negative matrix factorization*. In T. K. Leen, T. G. Dietterich, and V. Tresp, editors, Advances in Neural Information Processing Systems, **13**(2001), 556–562.
9. K. Schottle and R. Werner, *Improving "the most general methodology to create a valid correlation matrix"*, Risk Analysis IV, Management Information Systems, **9**(2004), 701–712.
10. P. Sonneveld, J.J.I.M. van Kan, X. Huang and C. W. Oosterlee, *Nonnegative matrix factorization of a correlation matrix*, Linear Algebra and its Applications, **431**(2009), 334–349.
11. D.-X. Xie, Y.-P. Sheng and Z.-Z. Zhang, *Computing the nearest bisymmetric positive semidefinite matrix under the spectral restriction*, Numerical Mathematics: A Journal of Chinese Universities English Series, **12**(2003), 71–82.

Junfeng Yin received Ph.D at China Academy of Science. Since 2008 he has been an associate professor at Tongji University. His research interests include numerical linear algebra, iterative method, parallel computation, financial and biohealth computation.

Department of Mathematics, Tongji University, Shanghai 200092, P. R. China.
e-mail: `yinjf@tongji.edu.cn`

Yumei Huang received Ph.D. from Hong Kong Bapist University in 2008. She is currently an associate professor at Lanzhou University. Her research interests are numerical linear algebra, iterative methods and image processing.

School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu, 730000, P. R. China.
e-mail: `ymhuang08@gmail.com`