Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(4): 8022–8035.
DOI: 10.3934/math.2025368
Received: 27 November 2024
Revised: 10 March 2025
Accepted: 25 March 2025
Published: 07 April 2025

Research article

Generalized low-rank approximation to the symmetric positive semidefinite matrix

Haixia Chang[1], **Chunmei Li**[2] **and Longsheng Liu**[3,*]

[1] School of Statistics and Mathematics, Shanghai Lixin University of Accounting and Finance, Shanghai 201209, China

[2] College of Mathematics and Computational Science, Guilin University of Electronic Technology, Guilin 541004, China

[3] School of Mathematics and Physics, Anqing Normal University, Anqing 246011, China

* **Correspondence:** Email: 062207@aqnu.edu.cn.

Abstract: In this paper, we consider the generalized low-rank approximation to the symmetric positive semidefinite matrix in the Frobenius norm: $\min_X \sum_{i=1}^m \left\|A_i - B_i X B_i^T\right\|_F^2$, where X is an unknown symmetric positive semidefinite matrix whose rank is less than or equal to a positive integer k. We first characterize the feasible set as $X = YY^T$, where Y has the order $n \times k$, and then convert the generalized low-rank approximation into an unconstrained generalized optimization problem. Finally, we employ the nonlinear conjugate gradient method with an exact line search to solve the generalized optimization problem. We also give numerical examples to exemplify the results.

Keywords: generalized low-rank approximation; symmetric positive semidefinite matrix; generalized optimization; nonlinear conjugate gradient method; feasible set
Mathematics Subject Classification: 15A33, 65F30, 65K10, 68W25

1. Introduction

Throughout this paper, $\mathbb{R}^{m\times n}$ denotes the set of all $m\times n$ matrices over the real number field $\mathbb{R}$. For a real or complex matrix D, the symbols D^T, D^*, $r(D)$, $tr(D)$, $\|D\|_F$, and ∂D stand for the transpose, the conjugate transpose, the rank, the trace, the Frobenius norm, and the differential of D , respectively. A matrix D is called symmetric if $D^T = D$. We write $D \geq 0$ if D is a real symmetric positive semidefinite matrix. For a matrix $D = (d_{ij})_{m\times n} = (d_1, d_2, \ldots, d_n)$, we use the symbol $vec(D)$ to stand for a vector, $vec(D) = (d_1^T, d_2^T, \ldots, d_n^T)^T$. It is known that $\|D\|_F^2 = tr(D^T D) = [vec(D)]^T vec(D)$.

Low-rank matrix approximation is an important topic in matrix theory and has wide applications. It emerges in data analysis, including dimensionality reduction and noise reduction, in machine learning, such as model simplification, recommendation systems, and in signal recovery, see [1–3]. The original low-rank matrix approximation is from E. Schmidt in [4], who solved the problem in the infinite-dimensional context of integral operators in 1907. In 1936, C. Eckart and G. Young developed the theory and gave us the Eckart−Young−Mirsky theorem in [5], which was described as approximating one matrix with another of lower rank, and gave a constructive solution. In 1960, Mirsky [6] used the singular value decomposition (SVD) of a matrix to study the problem

$$\|A - \widehat{X}\|_F = \min_{r(X)=k} \|A - X\|_F ,$$

and obtained the solution. In 1987, Golub et al. [7] gave a generalization of the Eckart−Young−Mirsky matrix approximation theorem and found $\widehat{X_2}$ such that

$$\|(X_1, \widehat{X_2}) - (X_1, X_2)\|_F = \min_{r(X_1,\overline{X}_2)\leq k} \|(X_1, \overline{X}_2) - (X_1, X_2)\|_F, \tag{1.1}$$

where the columns of the initial matrix X_1 remain fixed. By (1.1), the low-rank approximation of the initial matrix with some specified structure is called the structured low-rank matrix approximation, which can be written as

$$\min_{r(X)\leq k, X\in\Omega} \|A - X\|_F , \tag{1.2}$$

where Ω is a structure matrix set. For a complete survey of (1.2), see [8–10]. For the different matrix sets of Ω in (1.2), there are many rich results, such as the symmetric matrix in [11], the symmetric non-negative definite matrix in [12–14], the correlation matrix in [15, 16], the Hankle matrix in [17], the circulant matrix [18], Sylvester matrix in [19–21], and so on. Especially, Duan et al. [22] studied the low-rank approximation (1.2) of the symmetric positive semidefinite matrix, using the property of symmetric positive definite matrix, and the nonlinear conjugate gradient method to get an algorithm and compared it with Cadzow algorithm to show the effectiveness and feasibility.

For the generalized forms of structured low-rank approximation (1.2), there are some significant results. For instance, Zhang et al. [23] in 2003, Wei et al. [24] in 2007, and Chang [25] in 2020 studied the fixed-rank Hermitian non-negative definite solution X for the following fixed-rank matrix approximation least squares problem

$$\min_{r(X)=k} \|A - BXB^*\|_F ,$$

which discussed the ranges of the rank k and derived expressions of the solutions by applying the SVD of the matrix of B. In 2007, Friedland and Torokhti [26] considered

$$\min_{r(X)\leq k} \|A - BXC\|_F ,$$

and applied the SVD of the matrices to derive the explicit solution. For the low-rank approximation in the spectral norm, the authors of [27] and [28] applied the norm-preserving dilation theorem and matrix decomposition to obtain the explicit expression of the solution.

Motivated by the work mentioned above and keeping the applications of and interest in low-rank approximation in view, we consider the generalized low-rank approximation problem of the symmetric positive semidefinite matrix, which generalizes the results of [22]. The problem can be expressed as follows.

Problem 1. *Given the matrices $A_i \in R^{m_i \times m_i}$, $B_i \in R^{m_i \times n}$, $i = 1, 2, \ldots, m$, and a positive integer k, find an $n \times n$ real symmetric positive semidefinite matrix $\widetilde{X}$ with $r(\widetilde{X}) \le k$ such that*

$$\sum_{i=1}^{m} \left\| A_i - B_i \widetilde{X} B_i^T \right\|_F^2 = \min_{X \ge 0, r(X) \le k} \sum_{i=1}^{m} \left\| A_i - B_i X B_i^T \right\|_F^2 . \tag{1.3}$$

The paper is organized as follows. We first use the property of a symmetric positive semidefinite matrix $X = YY^T$, $Y \in R^{n \times k}$ to convert the generalized low-rank approximation into an unconstrained generalized optimization problem. Then we employ the nonlinear conjugate gradient method with an exact line search to solve the generalized optimization problem. Finally, we give numerical examples to exemplify the results.

2. The solution of Problem 1

In this section, we first characterize the feasible set and convert Problem 1 into an unconstrained optimization problem. Then, we apply the nonlinear conjugate gradient method to solve it.

Lemma 2.1. (See [29]) *An $n \times n$ matrix X is real symmetric positive semidefinite with $r(X) \le k$ if and only if an $n \times k$ matrix Y exists such that $X = YY^T$.*

The properties of the trace of the matrices can be found in linear algebra books, e.g., [29].

Lemma 2.2. *For the matrices A, B, C, and D with appropriate sizes, the following properties hold:*

$$\begin{aligned} tr(A^T) &= tr(A), \\ tr(A + B) &= tr(A) + tr(B), \\ tr(AC) &= tr(CA), \\ tr(ACD) &= tr(CDA) = tr(DAC). \end{aligned}$$

Lemma 2.3. (See [30]) *Let A, B, C be the constant matrices with appropriate sizes, and let X, Y be variable matrices. Then the following rules apply for deriving the differential of the matrix expressions:*

$$\begin{aligned} \partial A &= 0, \\ \partial(X + Y) &= \partial(X) + \partial(Y), \\ \partial(tr(X)) &= tr(\partial(X)), \\ \partial(XY) &= \partial(X)Y + X\partial(Y), \\ \frac{\partial}{\partial X} tr(X^T B X) &= BX + B^T X, \\ \frac{\partial}{\partial X} tr(X B X^T) &= X B^T + XB, \\ \frac{\partial}{\partial X} tr(AXBX) &= A^T X^T B^T + B^T X^T A^T, \\ \frac{\partial}{\partial X} tr(B^T X^T C X X^T C X B) &= C X X^T C X B B^T + C^T X B B^T X^T C^T X + C X B B^T X^T C X + C^T X X^T C^T X B B^T . \end{aligned}$$

By Lemma 2.1, we can characterize the feasible set $\{X \mid X \ge 0, r(X) \le k\}$ of Problem 1 as $X = YY^T$, where $Y \in R^{n \times k}$. Thus, Problem 1 can then be reformulated as follows.

Problem 2. *Given the matrices $A_i \in R^{m_i\times m_i}$, $B_i \in R^{m_i\times n}$, $i = 1, 2, \ldots, m$, and a positive integer k, find a matrix $\widetilde{Y} \in R^{n\times k}$ such that*

$$\sum_{i=1}^{m} \left\| A_i - B_i \widetilde{Y}\widetilde{Y}^T B_i^T \right\|_F^2 = \min_{Y\in R^{n\times k}} \sum_{i=1}^{m} \left\| A_i - B_i Y Y^T B_i^T \right\|_F^2 .$$

We observe that the problem

$$\min_{Y\in R^{n\times k}} \sum_{i=1}^{m} \left\| A_i - B_i Y Y^T B_i^T \right\|_F^2 \tag{2.1}$$

is an unconstrained nonlinear generalized matrix optimization problem. That is, if (2.1) has a solution $Y \in R^{n\times k}$, then $X = YY^T$ is the solution of (1.3) in Problem 1. Compared with (1.3), (2.1) has two advantages: (1) the constraints of rank and the positive semidefiniteness of the feasible set are eliminated; (2) for the new variable matrix $Y \in R^{n\times k}$, it only requires us to search a space of dimension nk, which is much lower than the dimension n^2 of $X \in R^{n\times n}$. However, the objective function becomes quadratic instead of linear. We will use the nonlinear conjugate gradient method for (2.1).

Define the function

$$f(Y) = \sum_{i=1}^{m} \left\| A_i - B_i Y Y^T B_i^T \right\|_F^2 , \quad Y \in R^{n\times k}, \tag{2.2}$$

which is a map $R^{n\times k} \to R$. It can be verified that Problem 2 is equivalent to the following problem

$$\min_{Y\in R^{n\times k}} f(Y). \tag{2.3}$$

We now apply the nonlinear conjugate gradient method with an exact line search to solve (2.3). First, we derive the gradient of the objective function $f(Y)$.

Theorem 2.4. *The gradient of the objective function $f(Y)$ in (2.2) is given by*

$$\nabla f(Y) = \sum_{i=1}^{m} (4B_i^T B_i Y Y^T B_i^T B_i Y - 2B_i^T A_i B_i Y - 2B_i^T A_i^T B_i Y). \tag{2.4}$$

Proof. By Lemma 2.2, we have

$$\begin{aligned} \left\| A_i - B_i Y Y^T B_i^T \right\|_F^2 &= tr[(A_i - B_i Y Y^T B_i^T)^T (A_i - B_i Y Y^T B_i^T)] \\ &= tr(B_i Y Y^T B_i^T B_i Y Y^T B_i^T) - tr(B_i Y Y^T B_i^T A_i) - tr(A_i^T B_i Y Y^T B_i^T) + tr(A_i^T A_i). \end{aligned} \tag{2.5}$$

From (2.5), the objective function $f(Y)$ can be expressed as follows:

$$\begin{aligned} f(Y) &= \sum_{i=1}^{m} \left\| A_i - B_i Y Y^T B_i^T \right\|_F^2 \\ &= \sum_{i=1}^{m} [tr(B_i Y Y^T B_i^T B_i Y Y^T B_i^T) - tr(B_i Y Y^T B_i^T A_i) - tr(A_i^T B_i Y Y^T B_i^T) + tr(A_i^T A_i)]. \end{aligned} \tag{2.6}$$

Using Lemma 2.3 and the expression in (2.6), the gradient of $f(Y)$ can be derived as

$$\nabla f(Y) = \frac{\partial}{\partial Y} f(Y)$$
$$= \sum_{i=1}^{m} [\frac{\partial}{\partial Y} tr(B_i Y Y^T B_i^T B_i Y Y^T B_i^T) - \frac{\partial}{\partial Y} tr(B_i Y Y^T B_i^T A_i) - \frac{\partial}{\partial Y} tr(A_i^T B_i Y Y^T B_i^T) + \frac{\partial}{\partial Y} tr(A_i^T A_i)]. \tag{2.7}$$

By Lemma 2.2, we note the following equalities:

$$tr(B_i Y Y^T B_i^T B_i Y Y^T B_i^T) = tr(B_i^T B_i Y Y^T B_i^T B_i Y Y^T) = tr(Y^T B_i^T B_i Y Y^T B_i^T B_i Y),$$
$$tr(B_i Y Y^T B_i^T A_i) = tr(B_i^T A_i B_i Y Y^T) = tr(Y^T B_i^T A_i B_i Y),$$
$$tr(A_i^T B_i Y Y^T B_i^T) = tr(B_i^T A_i^T B_i Y Y^T) = tr(Y^T B_i^T A_i^T B_i Y).$$

Applying Lemma 2.3, we obtain the following partial derivatives:

$$\frac{\partial}{\partial Y} tr(B_i Y Y^T B_i^T B_i Y Y^T B_i^T) = \frac{\partial}{\partial Y} tr(Y^T B_i^T B_i Y Y^T B_i^T B_i Y)$$
$$= 4 B_i^T B_i Y Y^T B_i^T B_i Y, \tag{2.8}$$

$$\frac{\partial}{\partial Y} tr(B_i Y Y^T B_i^T A_i) = \frac{\partial}{\partial Y} tr(Y^T B_i^T A_i B_i Y)$$
$$= B_i^T A_i B_i Y + B_i^T A_i^T B_i Y, \tag{2.9}$$

$$\frac{\partial}{\partial Y} tr(A_i^T B_i Y Y^T B_i^T) = \frac{\partial}{\partial Y} tr(Y^T B_i^T A_i^T B_i Y)$$
$$= B_i^T A_i^T B_i Y + B_i^T A_i B_i Y, \tag{2.10}$$

$$\frac{\partial}{\partial Y} tr(A_i^T A_i) = 0. \tag{2.11}$$

Substituting (2.8), (2.9), (2.10), and (2.11) into (2.7), we verify that the equality (2.4) holds. □

In Theorem 2.4, we derived the gradient of $f(Y)$. We now apply the nonlinear conjugate gradient method with an exact line search to solve the minimization problem (2.3). For details of the nonlinear conjugate gradient method, refer to [31, 32]. Below is the algorithm for solving (2.3).

Algorithm 2.5. (1) *Input: matrices* $A_i \in R^{m_i \times m_i}$, $B_i \in R^{m_i \times n}$, $i = 1, 2, \ldots, m$, *initial matrix* $Y_0 \in R^{n \times k}$, *and tolerant error* $\varepsilon > 0$*;*
(2) *Evaluate* $f_0 = f(Y_0), \nabla f_0 = \nabla f(Y_0)$, $D_0 = -\nabla f(Y_0)$, *and set* $k = 0$*;*
(3) *While* $\| \nabla f_k \|_F > \varepsilon$
Find t_k *such that*
$\sum_{i=1}^{m} \left\| A_i - B_i (Y_k + t_k D_k)(Y_k + t_k D_k)^T B_i^T \right\|_F^2 = \min_{t>0} \sum_{i=1}^{m} \left\| A_i - B_i (Y_k + t D_k)(Y_k + t D_k)^T B_i^T \right\|_F^2$;
$Y_{k+1} = Y_k + t_k D_k$;
$\nabla f_{k+1} = \nabla f(Y_{k+1})$;
$\beta_{k+1} = \frac{\|\nabla f_{k+1}\|_F^2}{\|\nabla f_k\|_F^2}$;
$D_{k+1} = -\nabla f_{k+1} + \beta_{k+1} D_k$*;*
end.

From Algorithm 2.5, we observe that Steps 1−2 are computed only once, while Step 3 constitutes the iterative loop for deriving the solutiom to Problem 1. The primary computational complexity of Algorithm 2.5 lies in computing the gradient and the step length. Specifically,

- The computational cost of the gradient is $O(4n^3m^3k)$;
- The computational cost of the step length is $O(2n^2m_ik + n^2m_i^2)$.

Thus, the total computational cost of Algorithm 2.5 is $O(T(4n^3m^3k + 2n^2m_ik + n^2m_i^2))$, where T is the number of iterations. Note that this estimate does not account for the available structure or sparsity of the matrices.

Remark 2.1. *Algorithm 2.5 is implemented with an exact line search for a step length t_k. We can apply the exact line search method in [33] to calculate the step length t_k, because the univariate function $\phi(\cdot)$ is defined by*

$$\phi(t) = \sum_{i=1}^{m} \left\| A_i - B_i(Y_k + tD_k)(Y_k + tD_k)^T B_i^T \right\|_F^2 = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0,$$

where

$$\begin{aligned} a_4 &= \sum_{i=1}^{m} \left\| B_iD_kD_k^TB_i^T \right\|_F^2, \\ a_3 &= \sum_{i=1}^{m} 2tr[B_iD_kD_k^TB_i^TY_kD_k^T + D_kY_k^T], \\ a_2 &= \sum_{i=1}^{m} -2tr[B_iD_kD_k^TB_i^T(A_i - B_iY_kY_k^TB_i^T)], \\ a_1 &= \sum_{i=1}^{m} -2tr[(A_i - B_iYY_k^TB_i^T)(Y_kD_k^T + D_kY_k^T)], \\ a_0 &= \sum_{i=1}^{m} \left\| A_i - B_iY_kY_k^TB_i^T \right\|_F^2, \end{aligned}$$

are quadratic, which is similar to the metric function for Newton's method with a line search for solving the algebraic Riccati equation in [33].

Remark 2.2. *Note that D_{k+1} in Algorithm 2.5 has a descending direction. In fact, we know that the exact line search always satisfies the following equality:*

$$tr[(\nabla f_{k+1})^T D_k] = [vec(\nabla f_{k+1})]^T vec(D_k) = 0.$$

By applying $vec(\cdot)$ to the equality $D_{k+1} = -\triangledown f_{k+1} + \beta_{k+1}D_k$ and multiplying by $[vec(\triangledown f_{k+1})]^T$, we get

$$[vec(\triangledown f_{k+1})]^T vec(D_{k+1}) = -\| \triangledown f_{k+1})\|_F^2 + \beta_{k+1}[vec(\triangledown f_{k+1})]^T vec(D_k).$$

Therefore, we obtain $[vec(\triangledown f_{k+1})]^T vec(D_{k+1}) < 0$, which implies that D_{k+1} has a descending direction.

According to Page 81 in [31], it is feasible for us to establish the global convergence theorem associated with Algorithm 2.5.

Theorem 2.6. *Assume that the function f is continuously differentiable and bounded from below. If the gradient $\triangledown f$ exhibits Lipschitz continuity, i.e., a constant L exists such that*

$$\|\triangledown f(X) - \triangledown f(Y)\|_F \leq L\,\|X - Y\|_F\,, \quad for\ \forall X, Y \in R^{n\times k},$$

then the sequence $\{Y_k\}$ generated by Algorithm 2.5 satisfies the condition

$$\lim_{k\to\infty} \inf \| \triangledown f(Y_k)\|_F = 0.$$

From the main results of this paper, we can obtain the special cases of Problem 1.

Remark 2.3. *When $B_i s$ are $n \times n$ identity matrices in Problem 1, we can derive the corresponding solution to the problem*

$$\sum_{i=1}^{m} \left\|A_i - \widetilde{X}\right\|_F^2 = \min_{X\geq 0, r(X)\leq k} \sum_{i=1}^{m} \|A_i - X\|_F^2\,.$$

This is the low-rank approximation of the symmetric positve semidefinite matrix X to the given matrices $A_i s$.

Remark 2.4. *When $m = 1$ in Remark 2.3, we can get the result of a low-rank approximation of the symmetric positive semidefinite matrix*

$$\left\|A_i - \widetilde{X}\right\|_F^2 = \min_{X\geq 0, r(X)\leq k} \|A_i - X\|_F^2$$

in [22].

3. Numerical examples

In this section, we first give an example to illustrate that Algorithm 2.5 is able to solve Problem 1, and then compare our algorithm with the nonmonotone spectral projected gradient algorithm [34], which can also be used to solve Problem 1. All experiments are performed in Windows 11 and MATLAB version 23.2.0.2365128 (R2023b) with an AMD Ryzen 7 5800H with Radeon Graphics CPU at 3.20 GHz and 16 GB of memory. We denote the relative residual error as

$$\varepsilon(k) = \frac{\sum_{i=1}^{m} \|A_i - B_i Y_k Y_k^T B_i^T\|_F}{\sum_{i=1}^{m} \|A_i\|_F}$$

and the gradient norm as

$$\|\nabla f_k\|_F = \|\nabla f(Y_k)\|_F,$$

where Y_k is the kth iterative matrix of Algorithm 2.5. The stopping criterion we used is

$$\|\nabla f_k\|_F < 1 \times 10^{-4}.$$

The initial matrix Y_0 is randomly generated by the *rand* function in MATLAB.

Example 3.1. *Consider Problem 1 with $m = 2$ and*

$$A_1 = \begin{bmatrix} 0.6938 & 0.1093 & 0.0503 & 0.8637 \\ 0.9452 & 0.3899 & 0.2287 & 0.0781 \\ 0.7842 & 0.5909 & 0.8342 & 0.6690 \\ 0.7056 & 0.4594 & 0.0156 & 0.5002 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 0.2180 & 0.5996 & 0.0196 & 0.5201 \\ 0.5716 & 0.0560 & 0.4352 & 0.8639 \\ 0.1222 & 0.0563 & 0.8322 & 0.0977 \\ 0.6712 & 0.1523 & 0.6174 & 0.9081 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 0.1080 & 0.0046 & 0.9870 & 0.5078 \\ 0.5170 & 0.7667 & 0.5051 & 0.5856 \\ 0.1432 & 0.8487 & 0.2714 & 0.7629 \\ 0.5594 & 0.9168 & 0.1008 & 0.0830 \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 0.6616 & 0.5905 & 0.4519 & 0.6801 \\ 0.5170 & 0.4406 & 0.8397 & 0.3672 \\ 0.1710 & 0.9419 & 0.5326 & 0.2393 \\ 0.9386 & 0.6559 & 0.5539 & 0.5789 \end{bmatrix}.$$

Case I: Set $k = 2$. We use Algorithm 2.5 with the initial value

$$Y_0 = \begin{bmatrix} 0.8669 & 0.3002 \\ 0.4068 & 0.4014 \\ 0.1126 & 0.8334 \\ 0.4438 & 0.4036 \end{bmatrix}$$

to solve the problem (2.1). We get the solution $\hat{Y}$ as follows:

$$\hat{Y} \approx Y_{242} = \begin{bmatrix} 0.7015 & -1.0397 \\ -0.7793 & 0.3971 \\ -0.5158 & -0.2175 \\ -0.1875 & -0.0296 \end{bmatrix}.$$

Then the solution $\hat{X}$ of Problem 1 is

$$\hat{X} = \hat{Y}\hat{Y}^T = \begin{bmatrix} 1.5731 & -0.9596 & -0.1357 & 0.1008 \\ -0.9596 & 0.7651 & 0.3156 & 0.1344 \\ -0.1357 & 0.3156 & 0.3133 & 0.1032 \\ -0.1008 & 0.1344 & 0.1032 & 0.0360 \end{bmatrix}.$$

The curves of the relative residual error $\varepsilon(k)$ and the gradient norm $\|\nabla f(Y_k)\|_F$ are in Figure 1.

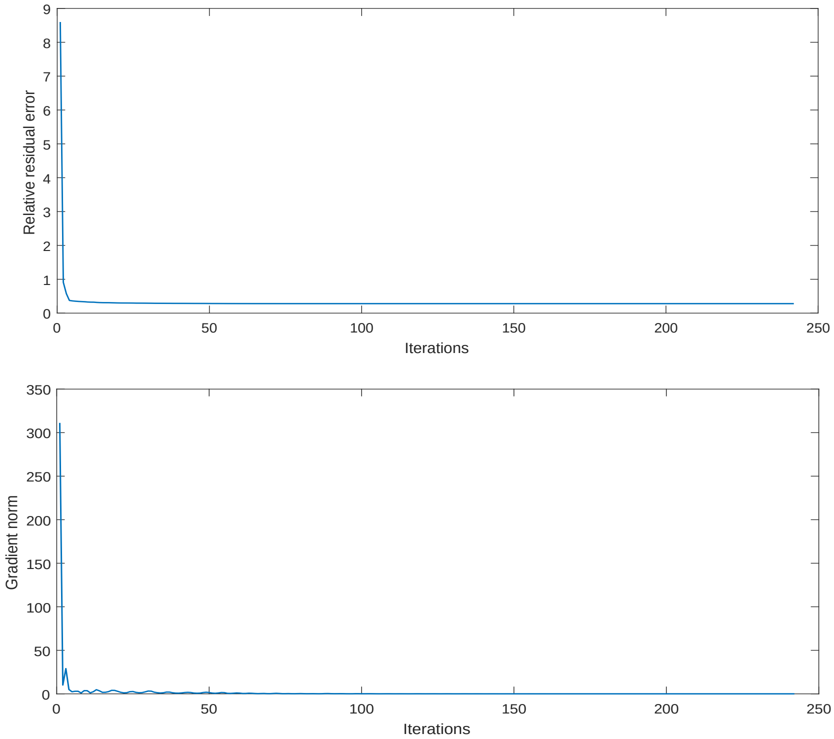

Figure 1. Convergence curves of the relative residual error $\varepsilon(k)$ and the gradient norm $\|\nabla f(Y_k)\|_F$.

Case II: Set k $= 3$*. We use Algorithm 2.5 with the initial value*

$$Y_0 = \begin{bmatrix} 0.5211 & 0.6791 & 0.0377 \\ 0.2316 & 0.3955 & 0.8852 \\ 0.4889 & 0.3674 & 0.9133 \\ 0.6241 & 0.9880 & 0.7962 \end{bmatrix},$$

to solve Problem (2.1). We get the solution

$$\hat{Y} \approx Y_{224} = \begin{bmatrix} 0.4617 & 0.4371 & -1.0811 \\ -0.4498 & -0.5971 & 0.4541 \\ -0.2575 & -0.4660 & -0.1729 \\ -0.0979 & -0.1620 & -0.0141 \end{bmatrix}.$$

Hence, the solution $\hat{X}$ *of Problem 1 is*

$$\hat{X} = \hat{Y}\hat{Y}^T = \begin{bmatrix} 1.5731 & -0.9596 & -0.1357 & 0.1008 \\ -0.9596 & 0.7650 & 0.3156 & 0.1344 \\ -0.1357 & 0.3156 & 0.3133 & 0.1032 \\ -0.1008 & 0.1344 & 0.1032 & 0.0360 \end{bmatrix}.$$

The curves of the relative residual error $\varepsilon(k)$ *and the gradient norm* $\|\nabla f(Y_k)\|_F$ *are in Figure 2.*

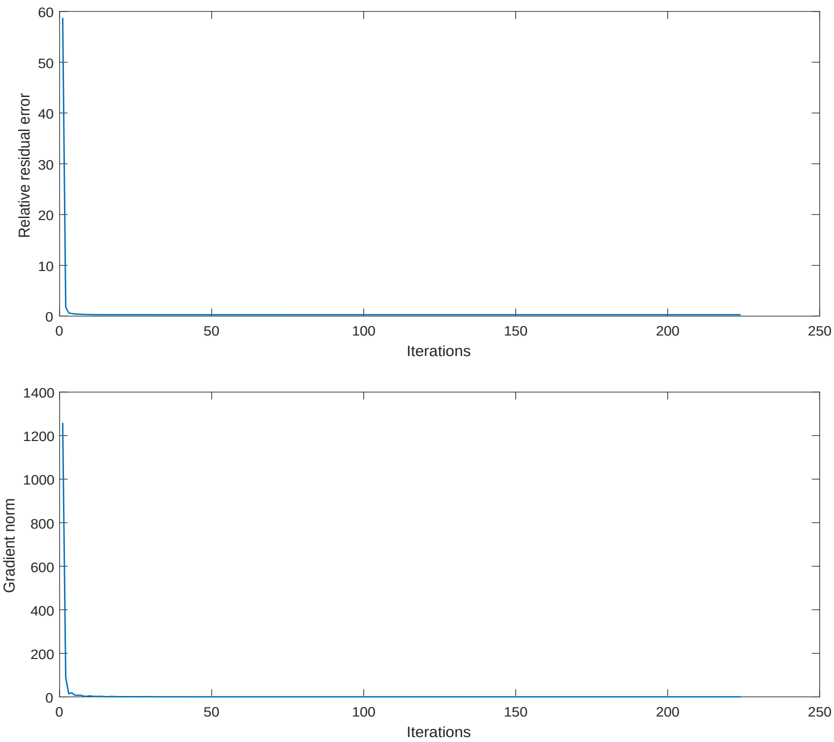

Figure 2. Convergence curves of the relative residual error $\varepsilon(t)$ and the gradient norm $\|\nabla f(Y_t)\|_F$.

By Example 3.1, Algorithm 2.5 is shown to be able to solve Problem 1.

Example 3.2. *Consider Problem 1 with* $A_i = rand(m_i)$, $B_i = rand(m_i, n)$, $i = 1, 2$. *Under the same initial value and stopping criteria, we use Algorithm 2.5 and the nonmonotone spectral projected gradient algorithm (denoted "Algorithm NSPG") in [34] to solve Problem 1 for the coefficient matrices above with different values of* m_i, *n, and k. The experimental results are listed in Table 1, including the number of iterations (denoted "IT"), CPU time (denoted "CPU"), and the residual error (denoted* $\varepsilon(k)$*).*

Table 1. Comparative results of Example 3.2 for different values m_i, n, k.

$m_1 = 20, m_2 = 20, n = 20, k = 10$	Algorithm 2.5	Algorithm NSPG
IT	59	136
CPU	16.7305	55.8277
$\varepsilon(k)$	0.2483	0.2488

$m_1 = 80, m_2 = 80, n = 70, k = 40$	Algorithm 2.5	Algorithm NSPG
IT	79	206
CPU	44.9927	162.7959
$\varepsilon(k)$	0.7131	0.7135

$m_1 = 100, m_2 = 100, n = 80, k = 50$	Algorithm 2.5	Algorithm NSPG
IT	94	278
CPU	59.8921	267.4421
$\varepsilon(k)$	0.0232	0.0230

$m_1 = 130, m_2 = 120, n = 80, k = 60$	Algorithm 2.5	Algorithm NSPG
IT	121	386
CPU	71.3387	295.732
$\varepsilon(k)$	0.1871	0.1876

From Table 1, we can see that the performance of Algorithm 2.5 is slightly better than that of Algorithm NSPG in terms of the iteration steps and computing time. From this table, we also see that the CPU time of Algorithm 2.5 increases very slowly with the increase in the problem's size, so it is more suitable for solving the large-scale problems than Algorithm NSPG.

4. Conclusions

In this paper, we give an iterative method to solve the generalized low-rank approximation of a symmetric positive semidefinite matrix. We convert the generalized low-rank approximation into an unconstraint generalized optimization problem. We apply the nonlinear conjugate gradient method to solve the generalized optimization problem. The numerical examples show that the algorithm is feasible and effective.

Author contributions

Haixia Chang, Chunmei Li, and Longsheng Liu: Conceptualization, formal analysis, investigation, methodology, software, validation, writing original draft, writing review and editing. All authors of this article have contributed equally. All authors have read and agreed to the published version of the manuscript.

Use of Generative-AI tools declaration

The authors declare they have not used artificial intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported by the Natural Science Foundation of Guangxi Province (Grant Nos. 2023GXNSFAA026067, 2024GXNSFAA010521), the Natural Science Foundation of China (Nos. 11601328,12201149), and key scientific research projects of universities in Anhui province (Grant No. 2023AH050476).

Conflict of interest

The authors declare that they have no conflicts of interest.

References

1. W. Jia, M. Sun, J. Lian, S. Hou, Feature dimensionality reduction: A review, *Complex Intell. Syst.*, **8** (2022), 2663–2693. https://doi.org/10.1007/s40747-021-00637-x
2. Y. Yan, D. Cheng, J. Feng, H. Li, J. Yue, Survey on applications of algebraic state space theory of logical systems to finite state machines, *Sci. China Inf. Sci.*, **66** (2023), 111201. https://doi.org/10.1007/s11432-022-3538-4
3. J. Zhang, M. Huang, N. Wan, Z. Deng, Z. He, J. Luo, Missing measurement data recovery methods in structural health monitoring: The state, challenges and case study, *Measurement*, **231** (2024), 114528. https://doi.org/10.1016/j.measurement.2024.114528
4. E. Schmidt, Zur Theorie der linearen und nichtlinearen Integralgleichungen, *Math. Ann.* **63** (1907), 433–476. https://doi.org/10.1007/BF01449770
5. C. Eckart, G. Young, The approximation of one matrix by another of lower rank, *Psychometrika*, **1** (1936), 211–218. https://doi.org/10.1007/BF02288367
6. L. Mirsky, Symmetric gauge functions and unitarily invariant norms, *Quarterly J. Math.*, **11** (1960), 50–59. https://doi.org/10.1093/qmath/11.1.50
7. G. H. Golub, A. Hoffman, G. W. Stewart, A generalization of the Eckart- Young-Mirsky matrix approximation theorm, *Linear Algebra Appl.*, **88-89** (1987), 317–327. https://doi.org/10.1016/0024-3795(87)90114-5
8. M. T. Chu, R. E. Funderlic, R. J. Plemmons, Structured low rank approximation, *Linear Algebra Appl.*, **366** (2003) , 157–172. https://doi.org/10.1016/s0024-3795(02)00505-0
9. I. Markovsky, Structured low rank approximation and its applications, *Automatica*, **44** (2008), 891–909. https://doi.org/10.1016/j.automatica.2007.09.011
10. I. Markovsky, *Low rank approximation: Algorithms, implementation, applications*, Cham: Springer, 2019. https://doi.org/10.1007/978-3-319-89620-5

11. M. T. Chu, R. J. Plemmons, Real-value, low rank, circulant approximation, *SIAM J. Matrix Anal. Appl.*, **24** (2003), 645–659. https://doi.org/10.1137/S0895479801383166
12. F. R. Bach, M. I. Jordan, Kernel independent component analysis, *J. Mach. Learn. Res.*, **3** (2002), 1–48.
13. L. Boman, H. Koch, A. S. de Meras, Method specific Cholesky decomposition: Coulomb and exchange energies, *J. Chem. Phys.*, **129** (2008), 134107. https://doi.org/10.1063/1.2988315
14. H. Harbrecht, M. Peters, R. Schneider, On the low rank approximation by the pivoted Cholesky decomposition, *Appl. Numer. Math.*, **62** (2012), 428–440. https://doi.org/10.1016/j.apnum.2011.10.001
15. Z. Y. Zhang, L. X. Wu, Optimal low rank approximation to a correlation matrix, *Linear Algebra Appl.*, **364** (2003), 161–187. https://doi.org/10.1016/s0024-3795(02)00551-7
16. X. F. Duan, J. C. Bai, M. J. Zhang, X. J. Zhang, On the generalized low rank approximation of the correlation matrices arising in the asset portfolio, *Linear Algebra Appl.*, **461** (2014), 1–17. https://doi.org/10.1016/j.laa.2014.07.026
17. H. Park, L. Zhang, J. B. Rosen, Low rank approximation of a Hankel matrix by structured total least norm, *BIT Numer. Math.*, **39** (1999), 757–779. https://doi.org/10.1023/A:1022347425533
18. U. Helmke, M. A. Shayman, Critical points of matrix least squares distance functions, *Linear Algebra Appl.*, **215** (1995), 1–19. https://doi.org/10.1016/0024-3795(93)00070-g
19. B. Y. Li, Z. Yang, L. H. Zhi, Fast low rank approximation of a Sylvester matrix by structured total least norm, *Japan Soc. Symbolic Algebr. Comput.*, **11** (2005), 165–174.
20. J. Winkler, J. Allan, Structured low rank approximations of the Sylvester resultant matrix for approximation gcds of bernstein basis polynomials, *Electron. Trans. Numer. Anal.*, **31** (2008), 141–155.
21. J. Winkler, J. Allan, Structured total least norm and approximate gcds of inexact polynomials, *J. Comput. Appl. Math.*, **215** (2008), 1–13. https://doi.org/10.1016/j.cam.2007.03.018
22. X. F. Duan, J. F. Li, Q. W. Wang, X. J. Zhang, Low rank approximation of the symmetric positive semidefinite matrix, *J. Comput. Appl. Math.*, **260** (2014), 236–243. https://doi.org/10.1016/j.cam.2013.09.080
23. X. Zhang, M. Y. Zhang, The rank-constrained Hermitian nonnegative-definite and positive - definite solutions to the matrix $AXA^* = B$, *Linear Algebra Appl.*, **370** (2003), 163–174. https://doi.org/10.1016/S0024-3795(03)00385-9
24. M. Wei, Q. Wang, On rank-constrained Hermitian nonnegative-definite least squares solutions to the matrix equation $AXA^H = B$, *Int. J. Comput. Math.*, **84** (2007), 945–952. https://doi.org/10.1080/00207160701458344
25. H. Chang, Constrained low rank approximation of the Hermitian nonnegative-definite matrix, *Adv. Linear Algebra Matrix Theory*, **10** (2020), 22–33. https://doi.org/10.4236/alamt.2020.102003
26. S. Friedland, A. Torokhti, Generalized rank-constrained matrix approximations, *SIAM J. Matrix Anal. Appl.*, **29** (2007), 656–659. https://doi.org/10.1137/06065551
27. M. Wei, D. Shen, Minimum rank solution to the matrix approximation problems in the spectral norm, *SIAM J. Matrix Anal. Appl.*, **33** (2012), 940–957. https://doi.org/10.1137/110851134

28. D. Shen, M. Wei, Y. Liu, Minimum rank (skew)Hermitian solutions to the matrix approximation problem in the spectral norm, *J. Comput. Appl. Math.*, **288** (2015), 351–365. https://doi.org/10.1016/j.cam.2015.04.033

29. F. Z. Zhang, *Matrix theory: Basic results and techniques*, New York: Springer, 2011. https://doi.org/10.1007/978-1-4614-1099-7

30. K. B. Petersen, M. S. Pedersen, *The matrix cookbook*, Technical Report, 2012.

31. D. H. Li, X. J. Tong, Z. Wan, *Numerical optimization*, Beijing: Science Press, 2005.

32. J. Nocedal, S. J. Wright, *Numerical optimization*, New York: Springer, 1999.

33. P. Benner, R. Byers, A exact line searches method for generalized continuous-time algebra Riccati equations, *IEEE Trans. Autom. Control*, **43** (1998), 101–107. https://doi.org/10.1109/9.654908

34. E. G. Birgin, J. M. Martínez, M. Raydan, Nonmonotone spectral projected gradient methods on convex sets, *SIAM J. Optim.*, **10** (2000), 1196–1211. https://doi.org/10.1137/S1052623497330963