Math Sci (2017) 11:257–265
DOI 10.1007/s40096-017-0223-3

ORIGINAL RESEARCH

New iterative methods for generalized singular-value problems

A. H. Refahi Sheikhani[1] · S. Kordrostami[1]

Received: 12 October 2016 / Accepted: 15 April 2017 / Published online: 27 April 2017

Abstract This paper presents two new iterative methods to compute generalized singular values and vectors of a large sparse matrix. To reach acceleration in the convergence process, we have used a different inner product instead of the common one, Euclidean one. Furthermore, at each restart, a different inner product has been chosen by the researchers. A number of numerical experiments illustrate the performance of the above-mentioned methods.

Keywords Generalized singular value · Krylov subspace · Iterative · Sparse

Mathematics Subject Classification 15A18 · 65F10 · 65L15

Introduction

There are a number of applications for generalized singular-value decomposition (GSVD) in the literature including the computation of the Kronecker form of the matrix pencil $A - \lambda B$ [5], solving linear matrix equations [1], weighted least squares [2], and linear discriminant analysis [6] to name but a few. In a number of applications like the generalized total least squares problem, the matrices A and B are large and sparse, so in such cases, only a few of the generalized singular vectors corresponding to the smallest or largest generalized singular values are needed. There is a kind of close connection between the GSVD problem and two different generalized eigenvalue problems. In fact, there are many efficient numerical methods to solve generalized eigenvalue problems [8–11]. In this paper, we will examine the Jacobi–Davidson-type subspace method which is related to the Jacobi–Davidson for the SVD [5], which in turn is inspired by the Jacobi–Davidson method to solve the eigenvalue problem [4]. The main step in Jacobi–Davidson-type method for the (GSVD) is solving the correction equations in an exact manner requiring the solution of linear systems of original size at each iteration. In general, these systems are considered as large, sparse, and nonsymmetrical. For this matter, we use the weighted Krylov subspace process to solve the correction equations in an exact manner, and we show that our proposed method has the feature of asymptotic quadratic convergence.

The paper is organized as follows. In "Preparations", we will remind the readers of basic definitions of the generalized singular-value decomposition problems and their elementary properties. "A new iterative method for GSVD" introduces our new numerical methods to solve generalized eigenvalue problems together with an analysis of the convergence of these methods. Several numerical examples are presented in "Numerical experiments". Finally, the conclusions are given in the last section.

Preparations

Definition 2.1 Supposes that $A \in R^{m \times n}$ and $B \in R^{p \times n}$. The generalized singular values of the pair (A, B) are presented as

$$\sum(A,B) = \{\sigma | \sigma \geq 0,\ \det(A^{\mathrm{T}}A - \sigma^2 B^{\mathrm{T}}B) = 0\}.$$

✉ A. H. Refahi Sheikhani
ah_refahi@liau.ac.ir; ah_refahi@yahoo.com

1 Department of Applied Mathematics, Faculty of Mathematical Sciences, Lahijan Branch, Islamic Azad University, Lahijan, Iran

Definition 2.2 A generalized singular value is called simple if $\sigma_i \neq \sigma_j$, for all $i \neq j$.

Theorem 2.3 *Suppose $A \in R^{m\times n}$, $B \in R^{p\times n}$, and $m \geq n$. Here, taking the previous theorem into consideration, we see that there are orthogonal matrices $U_{m\times m}$, $V_{p\times p}$ and a nonsingular matrix $X_{n\times n}$, such that*

$$\begin{aligned} U^{\mathrm{T}}AX &= \sum\nolimits_1 = \mathrm{diag}(\alpha_1,\ldots,\alpha_n) \quad \alpha_i \geq 0, \\ V^{\mathrm{T}}BX &= \sum\nolimits_2 = \mathrm{diag}(\beta_1,\ldots,\beta_n) \quad \beta_i \geq 0, \end{aligned} \tag{1}$$

where $q = \min\{p, n\}$, $r = \mathrm{rank}(B)$, and $\beta_1 \geq \cdots \geq \beta_r > \beta_{r+1} = \cdots = \beta_q = 0$. If $\alpha_j = 0$ for any j,$r+1 \leq j \leq n$, then $\sum(A,B) = \{\sigma | \sigma \geq 0\}$. Otherwise, $\sum(A,B) = \left\{\frac{\alpha_i}{\beta_i} | i = 1,\ldots,r\right\}$.

Proof Refer to [3].

Theorem 2.4 *Let $A \in R^{n\times n}$, $B \in R^{n\times n}$ have the GSVD:*

$$U^{\mathrm{T}}AX = \sum\nolimits_1 = \mathrm{diag}(\alpha_i), \quad V^{\mathrm{T}}BX = \sum\nolimits_2 = \mathrm{diag}(\beta_i)\,;$$

furthermore, consider it as nonsingular. Here, then, the matrix pencil

$$\begin{pmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{pmatrix} - \lambda \begin{pmatrix} I & 0 \\ 0 & B^{\mathrm{T}}B \end{pmatrix} \tag{2}$$

has eigenvalues $\lambda_j = \pm\alpha_j/\beta_j$, $j = 1,\ldots,n$ which corresponds to the eigenvectors:

$$\begin{pmatrix} u_j \\ \pm x_j/\beta_j \end{pmatrix}, \quad j = 1,\ldots,n \tag{3}$$

where u_j is the ith column of U and x_j is the ith column of X.

Proof Refer to [3].

Let D be a diagonal matrix, that is, $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$. If u and v are two vectors of R^n, we define the D-scalar product of $(u,v)_D = v^{\mathrm{T}}Du$. which is well defined if and only if the matrix D is positively definite or to say $d_i > 0$, $i = 1,\ldots,n$. The norm associated with this inner product is the D-norm $\|\cdot\|_D$ which is defined as $\|u\|_D = \sqrt{(u,u)_D} = \sqrt{u^{\mathrm{T}}Du}$ $\forall u \in R^n$.

As assumption B has full rank, $(x,y)_{(B^{\mathrm{T}}B)^{-1}} := y^{\mathrm{T}}(B^{\mathrm{T}}B)^{-1}x$ is an inner product, and due to this, the corresponding norm satisfies $\|x\|^2_{(B^{\mathrm{T}}B)^{-1}} := (x,x)_{(B^{\mathrm{T}}B)^{-1}}$. Inspired by the equality $\|Z\|_F^2 = \mathrm{trace}(Z^{\mathrm{T}}Z)$ for a real matrix Z, we define the $(B^{\mathrm{T}}B)^{-1}$-Frobenius norm of Z by

$$\|Z\|^2_{(B^{\mathrm{T}}B)^{-1},F} = \mathrm{trace}(Z^{\mathrm{T}}(B^{\mathrm{T}}B)^{-1}Z). \tag{4}$$

A new iterative method for GSVD

We will advance different extraction methods here which are often more appropriate for small generalized singular values than the standard one from "A new iterative method for GSVD". Before dealing with these new methods, we should refer to our main idea which is developed considering Krylov subspace methods.

Theorem 3.1 *Assume that (σ, u, v) is a generalized singular triple: $Aw = \sigma u$ and $A^{\mathrm{T}}u = \sigma B^{\mathrm{T}}Bw$, where σ is a simple nontrivial generalized singular value, and $\|u\| = \|Bw\| = 1$, and suppose that the correction equations*

$$P = \begin{pmatrix} I - \tilde{u}\tilde{u}^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}B\tilde{w}\tilde{w}^{\mathrm{T}} \end{pmatrix}, \tag{5}$$

are solved exactly in every step. Provided that the initial vectors $(\tilde{u}, \tilde{w})$ are close enough to (u, w) the sequence of approximations $(\tilde{u}, \tilde{w})$ converges quadratically to (u, w).

Proof Refer to [4].

Lemma 3.2 *Having in mind the Theorem 3.1, now suppose that m steps of the weighted Arnoldi process [7] have been performed on the following matrix:*

$$\begin{pmatrix} I - uu^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}Bww^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}. \tag{6}$$

Furthermore, consider the matrix $\widetilde{H}_m$ as the Hessenberg matrix, whose nonzero entries are the scalars $\tilde{h}_{i,j}$, constructed by the Weighted Arnoldi process. Here, we notice that the basis $\widetilde{V}_m = [\tilde{v}_1,\ldots,\tilde{v}_m]$ constructed by this algorithm is D-orthonormal and we have

$$\widetilde{V}_m^{\mathrm{T}} D \widetilde{V}_m = I_m, \tag{7}$$

$$\begin{pmatrix} I - uu^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}Bww^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix} \widetilde{V}_m = \widetilde{V}_{m+1} \begin{pmatrix} \widetilde{H}_m \\ h_{m+1,m}e_m^{\mathrm{T}} \end{pmatrix}. \tag{8}$$

Proof See [4].

We know that similar to Krylov methods, the mth $(m \geq 1)$ iterate $x_m = [s_m, t_m]^t$ of the weighted-FOM and weighted-GMRES methods belong to the affine Krylov subspace:

$$\begin{pmatrix} s_0 \\ t_0 \end{pmatrix} + \kappa_m \left(\begin{pmatrix} I - uu^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}Bww^{\mathrm{T}} \end{pmatrix} \times \begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}, \begin{pmatrix} r_0^{(s)} \\ r_0^{(t)} \end{pmatrix} \right). \tag{9}$$

Now, it is the time to prove our main theorem.

Theorem 3.3 *Considering Theorem* 3.1, m *steps of the weighted Arnoldi process have been run on* (7). *Here, the iterate* $x_m = [s_m, t_m]^t$ *is the exact solution of the correction equation:*

$$P\begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} = -r, \quad s\perp\tilde{u}, \quad t\perp\tilde{w}. \tag{10}$$

Proof The iterate x_m^{WF} of the weighted-FOM method is selected, because its residual is D-orthonormal or

$$\begin{pmatrix} r_m^{(s)} \\ r_m^{(t)} \end{pmatrix}^{\mathrm{WF}} \perp_D \kappa_m \left(\begin{pmatrix} I - uu^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}Bww^{\mathrm{T}} \end{pmatrix} \times \begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}, \begin{pmatrix} r_0^{(s)} \\ r_0^{(t)} \end{pmatrix} \right). \tag{11}$$

The iterate x_m^{WG} of the weighted-GMRES method is selected to lessen the residual D-norm in (9). Here, we notice that it is the solution of the least squares problem:

$$\mathrm{minimize}_{[s,t]^t\in(4.4)} \left\| \begin{pmatrix} A\tilde{w} - \theta\tilde{u} \\ A^{\mathrm{T}}\tilde{u} - \theta B^{\mathrm{T}}B\tilde{w} \end{pmatrix} - P\begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}\begin{pmatrix} s \\ t \end{pmatrix} \right\|_D. \tag{12}$$

In these methods, we use the D-inner product and the D-norm to calculate the solution in the affine subspace (9) and we create a D-orthonormal basis of the Krylov subspace:

$$\kappa_m\left(\begin{pmatrix} I - uu^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}Bww^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}, \begin{pmatrix} r_0^{(s)} \\ r_0^{(t)} \end{pmatrix} \right). \tag{13}$$

by the weighted Arnoldi process. An iterate x_m of these two methods can be transcribed as

$$\begin{pmatrix} s_m \\ t_m \end{pmatrix} = \begin{pmatrix} s_0 \\ t_0 \end{pmatrix} + \widetilde{V}_m \begin{pmatrix} y_m^{(s)} \\ y_m^{(t)} \end{pmatrix},$$

where $y_m \in R^m$.

Therefore, the matching residual $r_m = \left[r_m^{(s)}, r_m^{(t)} \right]^t$ satisfies

$$\begin{aligned} \begin{pmatrix} r_m^{(s)} \\ r_m^{(t)} \end{pmatrix} &= \begin{pmatrix} A\tilde{w} - \theta\tilde{u} \\ A^{\mathrm{T}}\tilde{u} - \theta B^{\mathrm{T}}B\tilde{w} \end{pmatrix} - \begin{pmatrix} I - uu^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}Bww^{T} \end{pmatrix}\begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}\begin{pmatrix} s_m \\ t_m \end{pmatrix} \\ &= \begin{pmatrix} A\tilde{w} - \theta\tilde{u} \\ A^{\mathrm{T}}\tilde{u} - \theta B^{\mathrm{T}}B\tilde{w} \end{pmatrix} - P\begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}\left(\begin{pmatrix} s_0 \\ t_0 \end{pmatrix} + \widetilde{V}_m \begin{pmatrix} y_m^{(s)} \\ y_m^{(t)} \end{pmatrix} \right), \\ &= \begin{pmatrix} r_0^{(s)} \\ r_0^{(t)} \end{pmatrix} - P\begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}\widetilde{V}_m \begin{pmatrix} y_m^{(s)} \\ y_m^{(t)} \end{pmatrix}, \\ &= \widetilde{V}_{m+1}\left(\beta e_1 - \begin{pmatrix} \widetilde{H}_m \\ h_{m+1,m}e_m^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} y_m^{(s)} \\ y_m^{(t)} \end{pmatrix} \right), \end{aligned}$$

where $\beta = \|r_0\|_D$, $r_0 = \left[r_0^{(s)}, r_0^{(t)} \right]^t$, and e_1 is the first vector of the canonical basis.

At this point, the weighted-FOM method entails finding the vector $y_m^{\mathrm{WF}} = \left[y_m^{(s)}, y_m^{(t)} \right]^t$ solution of the problem:

$$\widetilde{V}_m^{\mathrm{T}} D \widetilde{V}_{m+1}(\beta e_1 - \widetilde{H}_m y_m^{\mathrm{WF}}) = 0,$$

which is equal to solve

$$\widetilde{H}_m y_m^{\mathrm{WF}} = \beta e_1. \tag{14}$$

To the extent that the weighted-GMRES method is considered, the matrix $\widetilde{V}_{m+1}$ is D-orthonormal, so we have

$$\|r_m\|_D^2 = \left\| \widetilde{V}_{m+1}(\beta e_1 - \widetilde{H}_m y_m) \right\|_D^2 = \left\| \beta e_1 - \widetilde{H}_m y_m \right\|_2^2,$$

and problem (12) is condensed to find the vector y_m^{WG} solution of the minimization problem:

$$\mathrm{minimize}_{y\in R^m} \left\| \beta e_1 - \widetilde{H}_m y \right\|_2. \tag{15}$$

We can reach the solution of (14) and (15) with the use of the **QR** decomposition of the matrix $\widetilde{H}_m$, as for the FOM and GMRES algorithms.

When m is equal to the degree of the minimal polynomial of

$$\begin{pmatrix} I - uu^{\mathrm{T}} & 0 \\ 0 & I - B^{\mathrm{T}}Bww^{\mathrm{T}} \end{pmatrix}\begin{pmatrix} -\theta I & A \\ A^{\mathrm{T}} & -\theta B^{\mathrm{T}}B \end{pmatrix}$$

for $r_0 = [r_0^{(s)}, r_0^{(t)}]^t$, the Krylov subspace (13) will be invariant. Therefore, the iterate $x_m = [s_m, t_m]^t$ gained by both methods is the exact solution of the correction Eq. (10). ■

It is time to write the main algorithm in this paper now. The following algorithm applies FOM, GMRES, weighted-FOM, and weighted-GMRES processes to solve the correction Eq. (10) and as a final point to solve the generalized singular-value decomposition problem. They are represented as F-JDGSVD, G-JDGSVD, WF-JDGSVD, and WG-JDGSVD.

Algorithm 3.1 F-JDGSVD, G-JDGSVD, WF-JDGSVD and WG-JDGSVD

Input: *Starting vectors* u_1 *and* w_1, *and a tolerance* ε, *parameter* m, *and initial vector* x_0

Output: *An approximate triple* (θ, u, w) *for the largest generalized singular triple satisfying*

$$\left\| \begin{pmatrix} Aw - \theta u \\ A^T u - \theta B^T B w \end{pmatrix} \right\| \le \varepsilon$$

1. $s = u_1, t = w_1, U_0 = [\ ], W_0 = [\ ]$
   for $k = 1, 2, \ldots$ **do**

2. $U_k = RGS(U_{k-1}, s)$
   $W_k = RGS_{B^T B}(W_{k-1}, t)$

3. *Compute the kth column of* AW_k, $A^T U_k$, *and* $B^T B W_k$
   Compute the kth row and column of $H_k = U_k^T A W_k$

4. *Compute the approximately largest generalized singular triple* (θ, c, d),
   of the projected system using the standard extraction technique.

5. $u = U_k c$, $w = W_k d$

6. $r = \begin{pmatrix} Aw - \theta u \\ A^T u - \theta B^T B w \end{pmatrix}$

7. *Stop if* $\|r\| \le \varepsilon$
8. *Compute*
   $$\mathsf{A} = \begin{pmatrix} I - uu^T & 0 \\ 0 & I - B^T B w w^T \end{pmatrix} \begin{pmatrix} -\theta I & A \\ A^T & -\theta B^T B \end{pmatrix}$$

9. *for solving the system* $\mathsf{A} \begin{pmatrix} s \\ t \end{pmatrix} = -r$
   compute $r_0 = b - Ax_0$
   for $l = 1, 2, \ldots$ **do**
   for solving the system by FOM or GMRES methods:
   Compute $\beta = \|r_0\|_2$ *and* $v_1 = r_0 / \beta$.
   Construct the orthonormal basis V_m *by the Arnoldi process, starting with* v_1
   Form the approximate solution:
   FOM: *Solve the system* $H_m y_m = \beta e_1$ *by the QR factorization of* H_m *and set*
   $x_m = x_0 + V_m y_m$, $r_m = b - Ax_m$.

GMRES: *Compute* $y_m = \operatorname{argmin}_{y \in R^m} \|\beta e_1 - H_m y\|_2$, *by the QR factorization of* H_m *and set* $x_m = x_0 + V_m y_m$, $r_m = b - Ax_m$

for solving the system by WFOM or WGMRES methods:
Choose the vector d *such as* $\|d\|_2 = \sqrt{n}$, *and set* $D = diag(d)$
Compute $\tilde{\beta} = \|r_0\|_D$ *and* $\tilde{v}_1 = r_0 / \tilde{\beta}$.
Construct the D *-orthonormal basis* V_m *by the weighted Arnoldi process, starting with* $\tilde{v}_1$
Form the approximate solution:
WFOM: *Solve the system* $H_m y_m = \tilde{\beta} e_1$ *by the QR factorization of* H_m *and set*
$x_m = x_0 + V_m y_m$, $r_m = b - Ax_m$.
WGMRES: *compute* $y_m = \operatorname{argmin}_{y \in R^m} \|\tilde{\beta} e_1 - \overline{H}_m y\|_2$, *by the QR factorization of* $\overline{H}_m$,

set $x_m = x_0 + V_m y_m$, $r_m = b - Ax_m$.
If $\|r_m\|_2 \le \varepsilon_1$ *stop else set* $x_0 = x_m$, $r_0 = r_m$.

As Algorithm 3.1 displays, there are two loops in this algorithm. One of them computes the largest generalized singular value called the outer iteration, and the other called the inner iteration solves the system of linear equation at each iteration. Numerical tests indicate that there is a significant relation between parameter m and the norm of residual vector and the computational time.

Convergence

We will now demonstrate that the method we have proposed has asymptotically quadratic convergence to generalized singular values when the correction equations are solved in an exact manner and tend toward linear convergence when they are solved with a sufficiently small residual reduction.

Theorem 3.4 *Having in mind Theorem* 3.3*, suppose that m steps of the weighted Arnoldi process have been performed on* (6) *and* $x_m = [s_m, t_m]^{\mathrm{T}}$ *is the exact solution of the correction Eq.* (10)*. Provided that he initial vectors* $(\tilde{u}, \tilde{w})$ *are close enough to* (u, w)*, the sequence of approximations* $(\tilde{u}, \tilde{w})$ *converges quadratically to* (u, w).

Proof Suppose

$$\mathrm{A} = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} I & 0 \\ 0 & B^T B \end{pmatrix}$$

and P are like what you have seen in (5). Let $[s_m, t_m]^T$ with $s_m \perp \tilde{u}$ and $t_m \perp \tilde{w}$ be the exact solution to the correction equation:

$$P(\mathrm{A} - \theta \mathrm{B}) \begin{pmatrix} s_m \\ t_m \end{pmatrix} = -r. \tag{16}$$

Besides, let $\alpha u = \tilde{u} + s$, $s \perp \tilde{u}$, and $\beta w = \tilde{w} + t$, $t \perp \tilde{w}$, for certain scalars α and β, satisfy (15); note that these decompositions are possible meanwhile $u^{\mathrm{T}} \tilde{u} \neq 0$ and $w^{\mathrm{T}} \tilde{w} \; / = 0$ because of the assumption that the vectors $(\tilde{u}, \tilde{w})$ are close to (u, w). Projecting (16) yields

$$P(\mathrm{A} - \theta \mathrm{B}) \begin{pmatrix} s \\ t \end{pmatrix} = -r + P \begin{pmatrix} (\mu_1 - \theta) s \\ (\mu_2 - \theta) B^{\mathrm{T}} B t \end{pmatrix}. \tag{17}$$

Subtracting (16) from (17) gives

$$P(\mathrm{A} - \theta \mathrm{B}) \begin{pmatrix} s - s_m \\ t - t_m \end{pmatrix} = P \begin{pmatrix} (\mu_1 - \theta) s \\ (\mu_2 - \theta) B^{\mathrm{T}} B t \end{pmatrix}.$$

Thus for $(\tilde{u}, \tilde{w})$ close enough to (u, w), $P(\mathrm{A} - \theta \mathrm{B})$ is a bijection from $\tilde{u}^{\perp} \times \tilde{w}^{\perp}$ onto itself. Together with

$$\mu_1 = \tilde{u}^{\mathrm{T}} A (\tilde{w} + t) = \theta + O(\|t\|),$$
$$\mu_2 = (\tilde{w} + t)^{\mathrm{T}} A^{\mathrm{T}} (\tilde{u} + s) / \|B(\tilde{w} + t)\|^2 = \theta + O(\|s\| + \|t\|),$$

this implies asymptotic quadratic convergence:

$$\left\| \begin{pmatrix} \alpha u - (\tilde{u} + s_m) \\ \beta w - (\tilde{w} + t_m) \end{pmatrix} \right\| = \left\| \begin{pmatrix} s - s_m \\ t - t_m \end{pmatrix} \right\| = O\left(\left\| \begin{pmatrix} s \\ t \end{pmatrix} \right\|^2 \right).$$

Numerical experiments

In this section, we look for the largest generalized singular value, using the following default options of the proposed method:

Maximum dimension of search spaces	30
Maximum iterations to solve correction equation	10
Fix target until $\|r\| \leq \varepsilon$	0.01
Initial search spaces	Random

Example 4.1 The matrix pair (A, B) is constructed, such that that they are similar to experiments as [7]. We choose two diagonal matrices of dimension $n = 1000$. For $j = 1, 2, \ldots, 1000$

$$C = \mathrm{diag}(c_j), \quad c_j = (n - j + 1)/2n, \quad S = \sqrt{1 - C^2},$$
$$D = \mathrm{diag}(d_j), \quad d_j = \lceil j/250 \rceil + r_j$$

where the r_j uniformly distributed on the interval $(0, 1)$ and $\lceil \cdot \rceil$ denotes the ceil function. We take

$$A = Q_1 C D Q_2, \quad B = Q_1 S D Q_2$$

where Q_1 and Q_2 are two random orthogonal matrices. The estimated condition numbers of A and B are $4.4e2$ and $5.7e0$, respectively (Table 1).

We can see that by increasing the value of m, the number of outer and inner iterations decreases. Therefore, the consuming time also decreases. But not that if m is very large, the number of iterations increases because of loosing the orthogonality property. This example is given to show the improvement brought by the weighted methods WF-JDGSVD and WG-JDGSVD is simultaneously on the relative error and on the computational time (Fig. 1).

From figure one, we can see that the suggested method WG-JDGSVD is more accurate form the other methods.

Table 1 Implementation of Algorithm 3.1 for (A, B) with different values of m

m	F-JDGSVD			G-JDGSVD			WF-JDGSVD			WG-GSVD		
	$\sigma_{\max}$	$\|r\|_2$	Time	$\sigma_{\max}$	$\|r\|_2$	Time	$\sigma_{\max}$	$\|r\|_2$	Time	$\sigma_{\max}$	$\|r\|_2$	Time
4	0.5766	0.0084	23.95	0.5767	0.0062	28.35	0.5773	9.22e−5	31.13	0.5770	8.88e−6	22.08
6	0.5773	0.0052	19.82	0.5770	0.0043	23.32	0.5772	4.82e−5	28.76	0.5768	4.01e−6	17.51
8	0.5773	0.0023	16.10	0.5771	0.0028	19.30	0.5773	7.92e−6	23.66	0.5772	1.00e−6	14.70
10	0.5772	0.0058	14.85	0.5772	0.0014	16.31	0.5773	2.81e−6	17.99	0.5772	9.94e−7	12.04

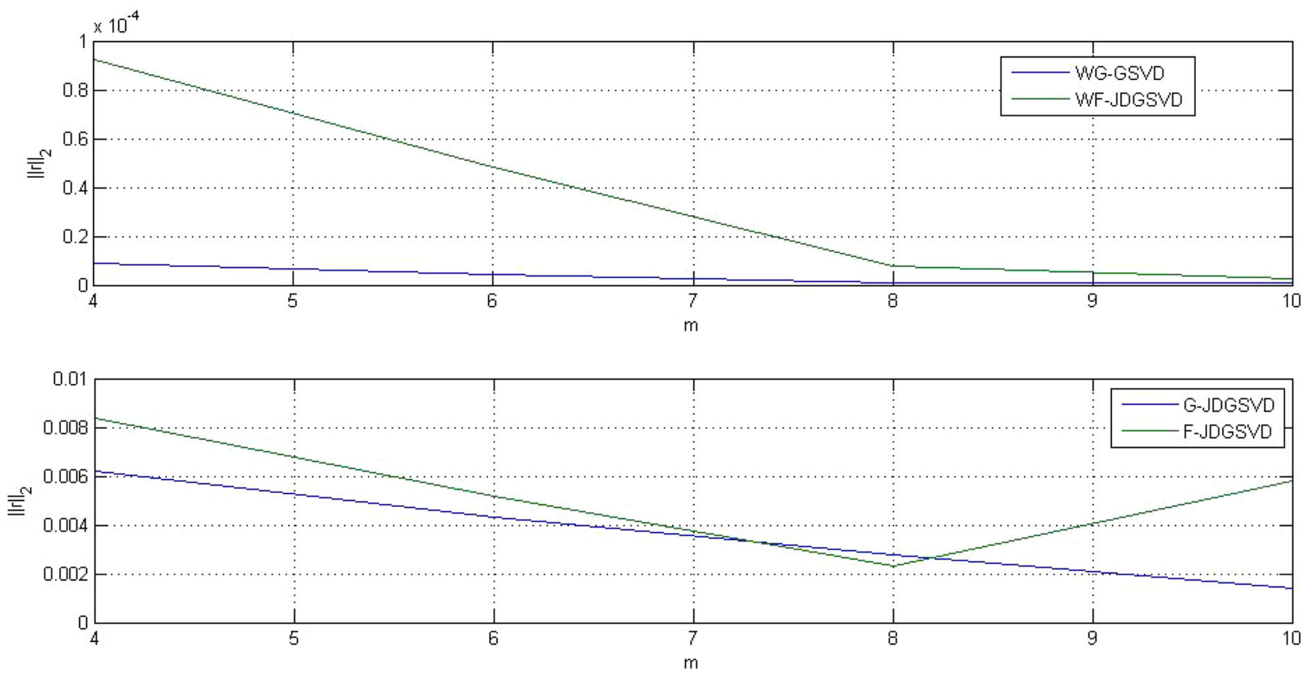


Fig. 1 *Errors plot* created by F-JDGSVD, G-JDGSVD, WF-JDGSVD, and WG-GSVD

Example 4.2 In this experiment, we take $A = CD$ and $B = SD$ of various dimension $n = 400$, 800, 1000, 1200.

This example is given to show the performance of two new methods on the large sparse problems. In this test, we have difficulties in computing the largest singular value for ill-conditioned matrices A and B. We note that in this experiments, due to the ill-conditioning of A and B, it turned out to be advantageous to turn of the Krylov option.

Example 4.3 Consider the matrix pair (A, B), where A is selected from the university of Florida sparse matrix collection [8] as *lp*-ganges. This matrix arises from a linear programming problem. Its size is 1309×1706 and it has a total of $Nz = 6937$ nonzero elements. The estimated condition number is $2.1332e4$, and B is the 1309×1706 identity matrix (Tables 2, 3).

We should mention that, for all considered Krylov subspaces sizes, each weighted method converges in less iterations and less time than its corresponding standard method. The convergence of F-JDGSVD and G-JDGSVD is slow, and we have linear asymptotic convergence. However, the two WF-JDGSVD and WG-JDGSVD methods have quadratic asymptotic convergence, because the correction Eq. (10) is solved exactly.

Remark 4.4 From the above examples and tables, we can see that the two suggested methods are more accurate than G-JDGSVD and F-JDGSVD for the same value m, but its computational times are often a little longer than G-JDGSVD and F-JDGSVD. Therefore, we can use WF-JDGSVD and WG-GSVD if the computational time is less important.

Remark 4.5 The algorithm we have described finds the largest generalized singular triple. We can compute multiple generalized singular triples of the pair (A, B) using a deflation technique. Suppose that $U_f = [u_1, \ldots, u_f]$ and

Table 2 Implementation of Algorithm 3.1 for (A, B) with various dimensions and $m = 6$

n	F-JDGSVD		G-JDGSVD		WF-JDGSVD		WG-GSVD		$\kappa(A)$	$\kappa(B)$
	$\|r\|_2$	Time	$\|r\|_2$	Time	$\|r\|_2$	Time	$\|r\|_2$	Time		
400	8.82e−4	7.03	0.0098	6.08	2.47e−8	11.85	2.14e−9	11.78	3.5e2	3.2e0
800	0.0085	19.59	0.0063	21.89	9.19e−8	26.09	4.44e−8	22.25	3.6e2	5.6e0
1200	0.0034	27.83	0.0073	29.35	6.74e−6	41.18	5.19e−7	42.35	4.8e2	6.6e0
1600	0.0075	38.65	0.0084	35.89	1.19e−5	49.09	4.99e−5	58.17	6.0e2	8.9e0

Table 3 Implementation of Algorithm 3.1 for (A, B) with different values of m

m	F-JDGSVD			G-JDGSVD			WF-JDGSVD			WG-GSVD		
	$\sigma_{\max}$	$\|r\|_2$	Time	$\sigma_{\max}$	$\|r\|_2$	Time	$\sigma_{\max}$	$\|r\|_2$	Time	$\sigma_{\max}$	$\|r\|_2$	Time
10	3.9889	0.0075	52.57	3.9865	0.0079	48.86	2.7297	0.00034	63.59	3.9890	0.00015	55.36
20	3.9907	0.0054	46.63	3.9889	0.0035	42.84	2.7298	0.00098	56.99	3.9890	0.00041	47.39
30	2.7298	0.0016	39.78	3.9889	0.0097	36.08	3.9907	0.00043	48.74	3.9888	0.00040	39.65
40	3.9897	0.0091	33.17	3.9888	0.0052	30.89	2.7298	0.00027	38.37	3.9887	0.00014	32.68

$W_f = [w_1, \ldots, w_f]$ contain the already found generalized singular vectors, where BW_f has orthonormal columns. We can check that the pair of deflated matrices

$$\hat{A} := (I - U_f U_f^{\mathrm{T}})A(I - W_f W_f^{\mathrm{T}} B^{\mathrm{T}} B) \quad \text{and} \\ \hat{B} := B(I - W_f W_f^{\mathrm{T}} B^{\mathrm{T}} B) \tag{18}$$

has the same generalized singular values and vectors as the pair (A, B) (see [3]).

Example 4.6 In generalized singular-value decomposition, if $B = I_n$, the $n \times n$ identity matrix, we get the singular value of A. SVD has important applications in image and data compression. For example, consider the following image.

This image is represented by a 1185×1917 matrix A. Which we can then decompose via the singular-value decomposition as $A = U \sum V^{\mathrm{T}}$ where U is 1185×1185, $\sum$ is 1185×1917, and V is 1917×1917. The matrix A, however, can also be written as a sum of rank 1 matrices $A = \sum_{j=1}^{r} \sigma_j u_j v_j^{\mathrm{T}}$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the r nonzero singular value of A. In digital image processing, any matrix A of order $m \times n (m \geq n)$ generally has a large number of small singular values. Suppose there are $(n - k)$ small singular values of A that can be neglected (Fig. 2).

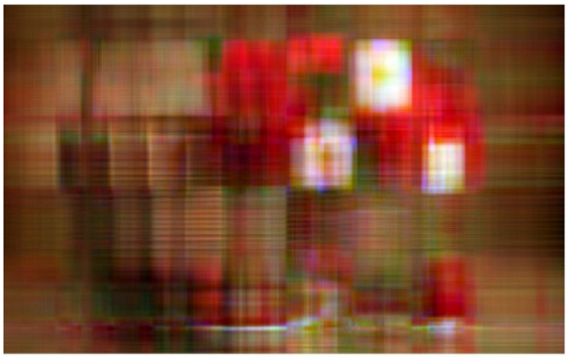
Then, the matrix $A_k = \sigma_1 u_1 v_1^{\mathrm{T}} + \sigma_2 u_2 v_2^{\mathrm{T}} + \cdots + \sigma_k u_k v_k^{\mathrm{T}}$ is a very good approximation of A, and such an approximation can be adequate. Even when k is chosen much less then n, the digital image corresponding to A_k can be very close to the original image. Below are the subsequent approximations using various numbers of singular values.

The observation on those examples, we found when $k \leq 20$, the images are blurry but with the increase of singular values, when their numbers are about 50, we have a good approach to the original image.

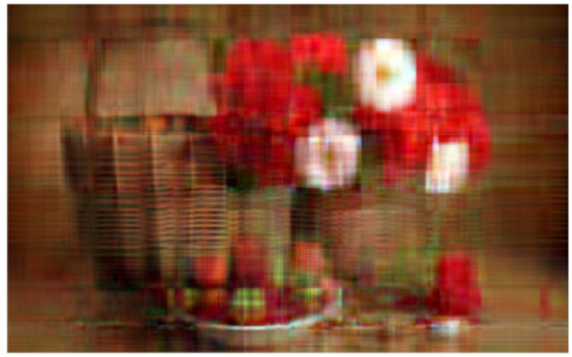
Fig. 2 Original image

Conclusions

In this paper, we have suggested two new iterative methods, namely, WF-JDGSVD and WG-JDGSVD, for the computation of some of the generalized singular values and corresponding vectors. Various examples studied illustrate these methods. To accelerate the convergence, we applied the Krylov subspace method for solving the correction equations in large sparse problems. In our methods, we see the existence of asymptotically quadratic convergence, because the correction equations are solved exactly. In the meantime, the correction equations in F-JDGSVD and G-JDGSVD methods are solved inexactly for large sparse problems, so we have linear convergence.



5 singular values



10 singular values

20 singular values

30 singular values

50 singular values

100 singular values

As the amount of the WF-JDGSVD and WG-JDGSVD methods is not much larger than that of the F-JDGSVD and G-JDGSVD methods, and as the weighted methods need less iterations to convergence, the parallel version of the weighted methods seems very interesting. From the tables and the figures, we see that when m increases, the suggested methods are more accurate than the previous methods; moreover, by increasing the dimension of the matrix, two suggested methods are applicable; this results are supported by convergence theorem which shows the asymptotically quadratic convergence to generalized singular values.

References

1. Betcke, T.: The generalized singular value decomposition and the method of particular solutions. SIAM. Sci. Comput. **30**, 1278–1295 (2008)
2. Hochstenbach, M.E.: Harmonic and refined extraction methods for the singular value problem, with applications in least square problems. BIT **44**, 721–754 (2004)
3. Hochstenbach, M.E.: A Jacobi–Davidson type method for the generalized singular value problem. Linear Algebra Appl. **431**, 471–487 (2009)
4. Hochstenbach, M.E., Sleijpen, G.L.C.: Two-sided and alternating Jacobi–Davidson. Linear Algebra Appl. **358**(1–3), 145–172 (2003)
5. Kagstrom, B.: The generalized singular value decomposition and the general $A - \lambda B$ problem. BIT **24**, 568–583 (1984)
6. Park, C.H., Park, H.: A relationship between linear discriminant analysis and the generalized minimum squared error solution. SIAM J. Matrix Anal. Appl. **27**, 474–492 (2005)
7. Saad, Y.: Krylov subspace methods for solving large unsymmetrical linear systems. Math. Comput. **37**, 105–126 (1981)
8. Saberi Najafi, H., Refahi Sheikhani, A.H: A new restarting method in the Lanczos algorithm for generalized eigenvalue problem. Appl. Math. Comput. **184**, 421–428 (2007)
9. Saberi Najafi, H., Refahi Sheikhani, A.H.: FOM-inverse vector iteration method for computing a few smallest, (largest) eigenvalues of pair (A, B). Appl. Math. Comput. **188**, 641–647 (2007)
10. Saberi Najafi, H., Refahi Sheikhani, A.H., Akbari, M.: Weighted FOM-inverse vector iteration method for computing a few smallest (largest) eigenvalues of pair (A, B). Appl. Math. Comput. **192**, 239–246 (2007)
11. Saberi Najafi, H., Edalatpanah, S.A., Refahi Sheikhani, A.H.: Convergence analysis of modified iterative methods to solve linear systems. Mediterr. J. Math. **11**(3), 1019–1032 (2014)