Computing Interior Eigenvalues of Large Sparse Symmetric Matrices

Achiya Dax¹

Accepted: 8 October 2021 / Published online: 28 November 2021
© The Author(s) 2021

Abstract

In this paper we present new restarted Krylov methods for calculating interior eigenvalues of large sparse symmetric matrices. The proposed methods are compact versions of the Heart iteration which are modified to retain the monotonicity property. Numerical experiments illustrate the usefulness of the proposed approach.

Keywords Computing interior eigenvalues · Monotonicity · Large sparse symmetric matrices · Restarted Krylov methods · Semi-compact Heart iteration

Mathematics Subject Classification 65F15 · 67F25 · 65F50

Introduction

The Heart iteration is a new type of Restarted Krylov methods. Given a symmetric matrix $G \in \mathbb{R}^{n \times n}$, the method is aimed at calculating a cluster of k exterior eigenvalues of G . As other Krylov methods it is best suited for large sparse matrices in which a matrix-vector product needs only $O(n)$ flops. Another underlying assumption is that k^2 is considerably smaller than n . The use of restarted Krylov methods for solving large eigenvalue problems was considered by several authors, e.g., [1]–[22]. Most of the methods are based on implicitly restarted Lanczos tridiagonalization algorithms in which the starting vector is determined by “polynomial filtering” or “thick restart”. The Heart iteration is not using these tools. It is based on Gram-Schmidt orthogonalizations and a simple explicit choice of the starting vector. This results in a simple iteration that enjoys rapid convergence. Former implementations of the Heart iteration concentrate on the computation of exterior eigenvalues, e.g., [4]–[8]. In this case the method possesses an important monotonicity property: The computed Ritz values proceed monotonically toward the desired eigenvalues of G . However, when moving to compute interior eigenvalues the monotonicity property is not guaranteed, and the Heart iteration needs some amendments. These amendments are derived in this paper.

✉ Achiya Dax
dax20@water.gov.il ; bela.dax@gmail.com

¹ Hydrological Service, P.O.B. 36118, 91360 Jerusalem, Israel

Note that although we are speaking on clusters of eigenvalues, the algorithm is carried out by computing the related invariant subspace of G . That is, a subspace that is spanned by the corresponding eigenvectors.

The shift-and-invert technique is a useful tool for calculating a cluster of eigenvalues. Let λ_i , $i = 1, \dots, n$, denote the eigenvalues of G and assume that we want to compute k eigenvalues of G that are close to some interior point ν . The term “interior” means that at least one eigenvalue is smaller than ν , and at least one eigenvalue is larger than ν . If $\nu \neq \lambda_j$ for $j = 1, \dots, n$, then the matrix $G - \nu I$ is invertible. The inverse of this matrix has the eigenvalues $1/(\lambda_j - \nu)$, $j = 1, \dots, n$, and an eigenvector of G that corresponds to λ_j is an eigenvector of $(G - \nu I)^{-1}$ that corresponds to $1/(\lambda_j - \nu)$, and vice versa. This relation turns the smallest eigenvalues of $G - \nu I$ to be the largest eigenvalues of $(G - \nu I)^{-1}$. Therefore, applying a restarted Krylov method on $(G - \nu I)^{-1}$ is likely to achieve fast convergence toward the desired invariant subspace.

After the shift we need to compute the k smallest eigenvalues of a given symmetric indefinite matrix, and this is the problem that we actually solve. The term “indefinite” means that the matrix has at least one negative eigenvalue, and at least one positive eigenvalue. This feature is the reason for losing the monotonicity property.

The algorithms proposed in this paper are aimed at solving the following problem. Let $G \in \mathbb{R}^{n \times n}$ be a given large sparse symmetric indefinite matrix whose eigenvalues are sorted to satisfy

$$|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_n|. \quad (1)$$

Then it is desired to compute

$$\lambda_j, \quad j = 1, \dots, k, \quad (2)$$

the k smallest eigenvalues of G . As before, if $|\lambda_1| > 0$ then applying a restarted Krylov method on G^{-1} is likely to achieve rapid convergence. However, since G is a large matrix, it is expensive to compute G^{-1} , and matrix-vector products of the form $G^{-1}\mathbf{b}$ are obtained by solving the linear system

$$G\mathbf{x} = \mathbf{b}. \quad (3)$$

If G is factorizable then we can use a direct solution method. Otherwise the linear system is solved by some iterative method. An important advantage of the Heart iteration is that it allows the use of inexact inversions. That is, there is no need to compute an accurate solution for the linear system, e.g., [4]. Some of the linear solvers are restricted to positive definite matrices. In this case one may solve the linear system

$$G^T G \mathbf{x} = \mathbf{b} \quad (4)$$

instead of (3). The algorithms described in this paper rely on the assumption that we have effective methods for solving these systems.

A straightforward way to extract Ritz vectors is by using the spectral decomposition of the related Rayleigh quotient matrix. Yet, as we shall see, this option lacks the monotonicity property. One way to retain monotonicity is by moving to compute the largest eigenvalues of G^{-1} . This approach has an effective “compact” version that avoids direct computation of G^{-1} but requires accurate solutions of (3).

A second way to achieve monotonicity is by moving from eigenvalues computation to singular values computation. Here the algorithm computes the k smallest singular values of G and the related $n \times k$ orthonormal matrix, V , whose columns are the corresponding right

singular vectors. Once V is computed, the desired eigenpairs of G are retrieved from the Rayleigh quotient matrix $V^T G V$. This method can be viewed as modified version of the basic Heart iteration that is used in [5]. The main improvements are the introduction of a semi-compact expansion scheme and the use of classical Gram-Schmidt orthogonalization.

The plan of the paper is as follows. The SVD iteration is presented in the next section. It computes the k smallest singular values of G and the corresponding right singular vectors. Hence it is necessary to convert the singular vectors into appropriate eigenvectors. This issue is resolved in Sect. 3. Then, in Sect. 4, we establish the monotonicity property. A semi-compact spectral iteration is presented in Sect. 5. This iteration is more intuitive but lacks the monotonicity property. The compact iteration that is described in Sect. 6 gains monotonicity by moving to compute the largest eigenvalues of G^{-1} . This iteration becomes attractive whenever we have an efficient algorithm for accurate solution of (3). The paper ends with numerical experiments that illustrate the behavior of the proposed methods.

A Semi-compact SVD Iteration

The algorithm described below is aimed at computing the k smallest singular values of G and the corresponding right singular vectors. The q th iteration, $q = 0, 1, 2, \dots$, has two parts. It starts with a matrix $V_q \in \mathbb{R}^{n \times k}$ that contains “old” information on the target space, a matrix $Y_q \in \mathbb{R}^{n \times \ell}$ that contains “new” information, and a matrix $X_q = [V_q, Y_q] \in \mathbb{R}^{n \times (k+\ell)}$ that includes all the known information. The matrix X_q has $p = k + \ell$ orthonormal columns. That is

$$X_q^T X_q = I \in \mathbb{R}^{p \times p}. \quad (5)$$

(A typical value for ℓ is $\ell = k$.)

Part 1. Information extraction (Subspace contraction). First compute the matrix

$$H_q = G X_q. \quad (6)$$

Then compute the k smallest singular values of H_q and the corresponding k right singular vectors. The singular vectors are assembled into a matrix

$$U_q \in \mathbb{R}^{p \times k}, \quad U_q^T U_q = I \in \mathbb{R}^{k \times k}, \quad (7)$$

which is used to compute the related matrix of Ritz vectors,

$$V_{q+1} = X_q U_q. \quad (8)$$

Since both X_q and U_q have orthonormal columns the matrix V_{q+1} inherits this property.

Part 2: Information retrieval (Subspace expansion). First set the starting matrix

$$X = V_{q+1} \quad (9)$$

and the starting vector

$$\mathbf{z} = C(V_{q+1} \mathbf{e}), \quad (10)$$

where $\mathbf{e} = (1, 1, \dots, 1)^T \in \mathbb{R}^k$ and C denotes the inverting operator (see below). Then for $j = k + 1, \dots, k + \ell$, the j th column of X is computed as follows:

- Set $\mathbf{z} = \mathbf{z} - X(X^T \mathbf{z})$.
- Set $\mathbf{z} = \mathbf{z} - X(X^T \mathbf{z})$.

- c) Normalize $\mathbf{z}$.
- d) Expand X to be $[X, \mathbf{z}]$.
- e) Set $\mathbf{z} = C\mathbf{z}$.

Explanatory Comments.

Comment 1: The expansion process starts with the matrix $X = V_{q+1} \in \mathbb{R}^{n \times k}$ and ends with the matrix $X = X_{q+1} \in \mathbb{R}^{n \times p}$ that satisfies

$$X_{q+1}^T X_{q+1} = I \in \mathbb{R}^{p \times p}. \quad (11)$$

Note that the first k columns of X_{q+1} are the columns of V_{q+1} . This feature stands behind the monotonicity property. (The description of the expansion process omits indices to ease the notations.)

Comment 2: Steps a) and b) achieve Gram-Schmidt orthogonalizations of $\mathbf{z}$ against the former columns of X . (One orthogonalization is not sufficient, but twice is enough ...) Here we use classical G-S orthogonalization while in earlier versions of the Heart iteration we have used modified G-S orthogonalization, e.g., [4]–[6]. The classical G-S is somewhat simpler and may save time.

Comment 3: The normalization of $\mathbf{z}$ is carried out by replacing $\mathbf{z}$ with $\mathbf{z}/\|\mathbf{z}\|_2$, where $\|\cdot\|_2$ denotes the Euclidean vector norm.

Comment 4: The matrix C is a symbol for an inverting operator. That is, matrix-vector products of the form $C\mathbf{z}$ are computed by solving a suitable linear system such as (3) or (4).

If the linear system is solved by some iterative method, then there is no need to compute a highly accurate solution. In other words, it is possible to use “inexact inversions”. If the linear solver requires a factorization of G (or $G^T G$), then the factorization is done only once, before starting the iterative process. Another alternative is to define C as a suitable polynomial filter, e.g., [13].

Comment 5: The expansion process builds a Krylov subspace whose starting vector is defined by (10). In earlier explicit methods there was no orthogonalization against the columns of V_{q+1} , which causes certain difficulties, e.g., [14, 16].

Comment 6: The above iteration is a “**semi-compact Heart iteration**”. The name “**Heart iteration**” comes from the similarity to the heart’s systole-diastole cardiac cycle: Step 1 achieves subspace contraction, while in Step 2 the subspace is expanded. Former versions of the Heart iteration are proposed in [4]–[8]. It is a “**semi-compact**” scheme since the building of $G X_q$ is separated from the expansion process. In compact versions the Rayleigh-quotient matrix is built during the expansion process, see Sect. 6.

Retrieving the Smallest Eigenpairs

At the end of the iterative process we obtain the k smallest singular values of G and an orthonormal matrix $V \in \mathbb{R}^{n \times k}$ whose columns are the corresponding right singular vectors of G . Hence it is needed to convert this information into the k smallest eigenvalues of G and the corresponding eigenvectors. Usually a right singular vector of G is also an eigenvector of G . Yet this rule can be violated when G has a pair of nonzero eigenvalues, λ_j and λ_{j+1} say, such that $\lambda_{j+1} = -\lambda_j$. In this case it is possible to have a right singular vector that corresponds to $|\lambda_j| = |\lambda_{j+1}|$ which is not an eigenvector of G . However, as explained below, it is possible to ensure that V will have the “**retrieving ability**”, which means that the columns of V provide an orthonormal basis for the desired invariant subspace of G . (An invariant subspace that corresponds to the smallest k eigenvalues of G .) In this case the

related eigenpairs are obtained as follows. First compute the Rayleigh quotient matrix

$$R = V^T G V, \quad (12)$$

then achieve a spectral decomposition of R to obtain the related Ritz values and Ritz vectors. More precisely, let $Q \in \mathbb{R}^{k \times k}$ be an orthonormal matrix such that

$$Q^T V^T G V Q = \text{diag}\{\rho_1, \dots, \rho_k\}, \quad (13)$$

then $\rho_1, \dots, \rho_k$ are the smallest eigenvalues of G , and the columns of the matrix VQ are the corresponding eigenvectors of G .

The question of whether V has the retrieving ability is related to the following property. Let λ_j be a given eigenvalue of G . Then λ_j has the “**sign property**” if any other eigenvalue of G , say λ_i , that satisfies $|\lambda_i| = |\lambda_j|$ also satisfies $\lambda_i = \lambda_j$. In other words, an eigenvalue λ_j has the “sign property” if all the other eigenvalues that have the same modulus as λ_j also have the same sign as λ_j . The importance of this feature comes from the following observations.

Theorem 1 *Let λ_j be an eigenvalue of G that has the sign property. Let $\mathbb{E}_j$ denote the eigenspace of G that corresponds to λ_j . Let $\mathbb{W}_j$ denote the eigenspace of G^2 that corresponds to λ_j^2 . In other words, $\mathbb{W}_j$ spans all the right singular vectors of G that correspond to $|\lambda_j|$. Then*

$$\mathbb{W}_j = \mathbb{E}_j. \quad (14)$$

That is, a right singular vector of G that corresponds to $|\lambda_j|$ is also an eigenvector of G that corresponds to λ_j .

Proof Let $\mathbf{v}_j \in \mathbb{R}^n$ be an eigenvector of G that corresponds to λ_j . The equality $G^2 \mathbf{v}_j = \lambda_j^2 \mathbf{v}_j$ means that $\mathbf{v}_j$ is also a singular vector of G that corresponds to $|\lambda_j|$. This shows that $\mathbb{E}_j$ is contained in $\mathbb{W}_j$. Therefore, since both subspaces have the same dimension, $\mathbb{W}_j = \mathbb{E}_j$. $\square$

Theorem 2 *Let λ_j be an eigenvalue of G that has not the sign property. In this case there exist singular vectors of G that correspond to $|\lambda_j|$ which are not eigenvectors of G , but we have the following relations. Let $\mathbb{W}_j$ be as above. Let $\mathbb{E}_j^+$ and $\mathbb{E}_j^-$ denote eigenspaces of G that correspond to $|\lambda_j|$ and $-|\lambda_j|$, respectively. Then*

$$\mathbb{W}_j = \mathbb{E}_j^+ \oplus \mathbb{E}_j^-. \quad (15)$$

That is, $\mathbb{W}_j$ is a direct sum of the eigenspaces $\mathbb{E}_j^+$ and $\mathbb{E}_j^-$.

Proof Here both $\mathbb{E}_j^+$ and $\mathbb{E}_j^-$ are contained in $\mathbb{W}_j$. Consequently the direct sum $\mathbb{E}_j^+ \oplus \mathbb{E}_j^-$ is also contained in $\mathbb{W}_j$. On the other hand, since $\mathbb{E}_j^+$ and $\mathbb{E}_j^-$ are mutually orthogonal,

$$\dim(\mathbb{E}_j^+ \oplus \mathbb{E}_j^-) = \dim(\mathbb{E}_j^+) + \dim(\mathbb{E}_j^-). \quad (16)$$

This shows that the direct sum has the same dimension as $\mathbb{W}_j$, which proves (15). $\square$

As before it is possible to assume that the eigenvalues of G satisfy (1). Then, since G is symmetric, the singular values of G are $|\lambda_j|$, $j = 1, \dots, n$, and the k smallest singular values that we compute are

$$|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_k|. \quad (17)$$

This means that $|\lambda_k|$ is the largest singular value that we compute, while the next singular value in size, which is not computed, is $|\lambda_{k+1}|$. It is also possible to assume that the columns of V are ordered in a similar manner,

$$V = [\mathbf{v}_1, \dots, \mathbf{v}_k] \quad (18)$$

where $\mathbf{v}_j$, $j = 1, \dots, k$, is a right singular vector of G that corresponds to $|\lambda_j|$. With these notations and assumptions Theorems 1-2 lead to the following conclusions.

Theorem 3 *The Rayleigh quotient matrix (12) is a block diagonal matrix. Eigenvalues that have the sign property result in 1×1 blocks. The larger blocks correspond to eigenvalues without this property, where each block refers to a different singular value.* $\square$

Proof The j th column of $V^T G V$ has the form $V^T G \mathbf{v}_j$, $j = 1, \dots, k$, and the entries of this column are $\mathbf{v}_i^T G \mathbf{v}_j$, $i = 1, \dots, k$. If λ_j has the sign property then $G \mathbf{v}_j = \lambda_j \mathbf{v}_j$ and $V^T G \mathbf{v}_j = \lambda_j \mathbf{e}_j$, where $\mathbf{e}_j$ denotes the j th column of the $k \times k$ identity matrix. Otherwise, when λ_j has not the sign property, there exist indices $i_1 < i_2$ for which the set λ_i , $i = i_1, \dots, i_2$, includes all the eigenvalues λ_i that satisfy $|\lambda_i| = |\lambda_j|$. In this case (15) tells us that $\mathbf{v}_j$ is a linear combination of eigenvectors whose eigenvalues come from this set. Consequently the vector $G \mathbf{v}_j$ is also a linear combination of these eigenvectors, and the vector $V^T G \mathbf{v}_j$ is a linear combination of the vectors $\mathbf{e}_i$, $i = i_1, \dots, i_2$. $\square$

Corollary 4 *For $j = 1, \dots, k$, if*

$$|\lambda_j| < |\lambda_{k+1}| \quad (19)$$

then $\mathbb{E}_j$ is contained in $\text{Range}(V)$. $\square$

The last observation indicates that the key for determining whether V has the retrieving ability lies in the properties of λ_k .

Corollary 5 *If $|\lambda_k| < |\lambda_{k+1}|$ then V has the retrieving ability.* $\square$

Corollary 6 *Assume that $|\lambda_k| = |\lambda_{k+1}|$. If λ_k has the sign property then V has the retrieving ability. Otherwise, when λ_k does not have the sign property, the retrieving ability is not ensured.* $\square$

If V has the retrieving ability then the absolute values of the computed Ritz values should equal the corresponding singular values (which have been computed together with V). This gives a simple practical test to verify that the retrieving ability indeed holds. Otherwise, when this test fails, the situation described in Corollary 6 holds. That is, $|\lambda_k| = |\lambda_{k+1}|$ and the set of all the eigenvalues of G whose modulus equals $|\lambda_k|$ includes both a negative eigenvalue and a positive eigenvalue. Consequently a tiny shift of G , say τ , is likely to split this set into two separate sets: In one set all the eigenvalues have the same positive value, in the other set all the eigenvalues have the same negative value, while the corresponding absolute values differ by the amount of $|2\tau|$. Therefore, if the retrieving test fails then Step 1 is repeated with a slightly different shift, which ensures the retrieving ability.

The Monotonicity Property

The monotonicity property of the computed singular values is based on the following extensions of Cauchy interlace theorem and Poincaré separation theorem. (The original theorems

of Cauchy and Poincaré refer to eigenvalues of symmetric matrices. See, for example, [11] and [15].)

Theorem 7 *Let G be an $n \times n$ matrix, with singular values*

$$0 \leq \sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n.$$

Let the $n \times k$ matrix F be obtained from G by deleting $n - k$ columns of G and let,

$$0 \leq \eta_1 \leq \eta_2 \leq \cdots \leq \eta_k$$

denote the singular values of F . Then

$$\sigma_j \leq \eta_j \text{ for } j = 1, \dots, k, \quad (20)$$

and

$$\eta_{k+1-j} \leq \sigma_{n+1-j} \text{ for } j = 1, \dots, k. \quad (21)$$

In particular, for $k = n - 1$ we obtain the interlacing relations

$$\sigma_1 \leq \eta_1 \leq \sigma_2 \leq \eta_2 \leq \sigma_3 \leq \cdots \leq \sigma_{n-1} \leq \eta_{n-1} \leq \sigma_n. \quad (22)$$

Proof The proof is concluded by applying Cauchy interlace theorem on the matrices $G^T G$ and $F^T F$. $\square$

Corollary 8 *Let the $n \times p$ matrix X have orthonormal columns, and let*

$$0 \leq \xi_1 \leq \xi_2 \leq \cdots \leq \xi_p$$

denote the singular values of the $n \times p$ matrix GX . Let the $n \times k$ matrix V be obtained from X by deleting $\ell = p - k$ columns of X , and let

$$\eta_1 \leq \eta_2 \leq \cdots \leq \eta_k$$

denote the singular values of the $m \times k$ matrix GV . Then

$$\sigma_j \leq \xi_j \leq \eta_j \text{ for } j = 1, \dots, k, \quad (23)$$

and

$$\eta_{k+1-j} \leq \xi_{p+1-j} \leq \sigma_{n+1-j} \text{ for } j = 1, \dots, k. \quad (24)$$

Proof Let $Q \in \mathbb{R}^{n \times n}$ be an orthonormal matrix whose first p columns are those of X . Then the matrix GQ has the same singular values as G . Now the inequalities (23) and (24) are concluded by following the proof of Theorem 7. $\square$

Observe that the proofs are not using the symmetry of G . In fact, similar arguments show that the above bounds hold for any $m \times n$ matrix with $m \geq n$. Moreover, the upper bounds (21) and (24) remain valid when $m < n$. Yet the lower bounds (20) and (23) are not guaranteed when $m < n$.

Theorem 9 (Monotonicity) *Consider the q th iteration of the new method, $q = 1, 2, \dots$, and let*

$$0 \leq \sigma_1^{(q)} \leq \sigma_2^{(q)} \leq \cdots \leq \sigma_p^{(q)} \quad (25)$$

denote the singular values of the matrix $H_q = GX_q$. Then the k smallest singular values of these matrices satisfy

$$\sigma_j \leq \sigma_j^{(q)} \leq \sigma_j^{(q-1)} \quad (26)$$

for $j = 1, \dots, k$ and $q = 1, 2, 3, \dots$

Proof The k Ritz singular values which are computed at Part 1 are the smaller ones,

$$\sigma_1^{(q)} \leq \sigma_2^{(q)} \leq \dots \leq \sigma_k^{(q)}.$$

Similarly, the former iteration has computed

$$\sigma_1^{(q-1)} \leq \sigma_2^{(q-1)} \leq \dots \leq \sigma_k^{(q-1)},$$

which are the singular values of the matrix GV_q . Therefore, since the first k columns of X_q are the columns of V_q , these matrices satisfy the conditions of Corollary 8, and (26) is a direct corollary of (23). $\square$

The proofs of Corollary 8 and Theorem 9 emphasize the importance of the orthonormality relations, and provide the motivation behind the proposed Heart iteration. The monotonicity property (26) ensures that $\sigma_j^{(q)}$ moves closer to σ_j at each iteration. That is, each iteration improves the computed estimates of the smallest singular values. (In eigenvalues terminology the computed estimates are called Ritz values.)

A Semi-compact Spectral Iteration

The traditional way to achieve extraction is by using the spectral decomposition of the related Rayleigh-quotient matrix. Former restarted Krylov methods compute this matrix via the Lanczos tridiagonalization algorithm, e.g., [3, 13, 17, 21, 22]. Yet the Heart iteration avoids the Lanczos process and builds the Rayleigh-quotient matrix directly. This takes Part 1 into the following form, while Part 2 remains unchanged.

Part 1: Spectral extraction. First compute the Rayleigh-quotient matrix

$$S_q = X_q^T G X_q. \quad (27)$$

Then compute the k smallest eigenvalues of S_q . The corresponding k eigenvectors of S_q are assembled into a matrix

$$U_q \in \mathbb{R}^{p \times k}, \quad U_q^T U_q = I \in \mathbb{R}^{k \times k}, \quad (28)$$

which is used to compute the related matrix of Ritz vectors

$$V_{q+1} = X_q U_q. \quad (29)$$

Indeed, this approach works very well when calculating small eigenvalues of positive definite matrices, e.g., [4]. However, when G turns to be indefinite the monotonicity property is not guaranteed, while loss of monotonicity may prevent the algorithm from achieving fast convergence. See Table 4.

A second weakness of this iteration is that C is restricted to (3), since solving (4) may bring inappropriate information that retards convergence. The reasons lie in the difficulties mentioned in Sect. 3.

A Compact Spectral Iteration

This iteration is aimed at computing the k largest eigenvalues of G^{-1} and the related eigenvectors, which are the actual eigenvectors that we seek. Therefore, the extraction stage is carried out by computing the largest k eigenpairs of the Rayleigh quotient matrix

$$S_q = X_q^T G_q^{-1} X_q. \quad (30)$$

As before, G^{-1} is not available and matrix-vector products of the form $G^{-1}\mathbf{b}$ are computed by solving (3). But here S_q is computed “on the fly”, during the expansion process. This can be done by noting that the columns of S_q have the form

$$X_q^T (G_q^{-1} \mathbf{x}_j), \quad j = 1, \dots, p, \quad (31)$$

where $\mathbf{x}_j$ denotes the j th column of X_q . Hence the vector $\mathbf{z} = G^{-1}\mathbf{x}_j$ can be used both to construct S_q and to expand the Krylov subspace. The resulting “compact” iteration has the following form. (We omit the iteration indices for simplicity.)

Part I: Contraction

Compute the k largest eigenvalues of S and the corresponding eigenvectors. The eigenvalues construct a diagonal matrix, $D \in \mathbb{R}^{k \times k}$. The eigenvectors are assembled into the matrix

$$U \in \mathbb{R}^{p \times k}, \quad U^T U = I, \quad (32)$$

which is used to compute the related matrix of Ritz vectors

$$V = XU \in \mathbb{R}^{n \times k}. \quad (33)$$

Since both X and U have orthonormal columns the matrix V inherits this property.

Part II: Expansion

First set

$$\begin{aligned} X &= V, \\ S &= D, \\ \mathbf{z} &= G^{-1}(V\mathbf{e}), \end{aligned}$$

and

$$\mathbf{r} = X^T \mathbf{z}.$$

Then for $j = k + 1, \dots, k + \ell$, do as follows.

- Set $\mathbf{z} = \mathbf{z} - X\mathbf{r}$.
- Set $\mathbf{z} = \mathbf{z} - X(X^T \mathbf{z})$.
- Normalize $\mathbf{z}$
- Expand X to be $[X, \mathbf{z}]$.
- Set $\mathbf{z} = G^{-1}\mathbf{z}$.
- Expand S by computing the vector

$$\mathbf{r} = (r_1, \dots, r_j)^T = X^T \mathbf{z}$$

and setting the new entries of S to be

$$s_{ij} = s_{ji} = r_i \quad \text{for } i = 1, \dots, j.$$

Table 1 The test matrices

Matrix name	Eigenvalues $\lambda_j, j = 1, \dots, n$	Matrix size n	Number of non-zeros
Harmonic Roots	$\lambda_j = (-1)^{j+1}/(n+1-j)^{1/2},$ $j = 1, \dots, n$	10,000	586,840
Harmonic	$\lambda_j = (-1)^{j+1}/j, j = 1, \dots, n$	20,000	629,180
Geometric	$\lambda_j = (-1)^{j+1}(0.99995)^j$ $j = 1, \dots, n$	50,000	673,310,
Normal Dist	Normally distributed random numbers with mean 0 and variance 1	100,000	338,402
Uniform Dist	Uniformly distributed random numbers in $[-1, 1]$	300,000	458,802
Uniform Dist	Uniformly distributed random numbers in $[-1, 1]$	500,000	1,995,004

Table 2 Computing the k smallest eigenvalues with a semi-compact SVD iteration, $l = \max\{40, k\}$

Matrix name	Number of iterations					
	$k = 6$	$k = 10$	$k = 20$	$k = 40$	$k = 100$	$k = 200$
Harmonic Roots	64	64	69	76	26	15
Harmonic	62	65	70	77	27	15
Geometric	60	64	69	77	27	15
Normal Dist	0	0	0	0	0	0
Uniform Dist	0	0	0	0	0	0

The above iteration is a variant of the compact Heart algorithm [7]. Here the compact scheme is implicitly applied on G^{-1} and computes the k largest eigenpairs of this matrix. Assume for a moment that η is an eigenvalue of G^{-1} that has the largest modulus, and let $\eta^{(q)}$ denote the related Ritz value at the q th iteration, $q = 0, 1, 2, \dots$. Then $1/\eta^{(q)}$ is the corresponding estimate for the smallest eigenvalue of G , while the monotonicity property for large eigenvalues [7] implies the relations

$$\eta^{(q)} \leq \eta^{(q+1)} \leq \eta \text{ when } \eta > 0, \quad (34)$$

and

$$\eta \leq \eta^{(q+1)} \leq \eta^{(q)} \text{ when } \eta < 0. \quad (35)$$

Similar relations hold for the other Ritz values, forcing them to proceed monotonically toward the desired eigenvalues of G^{-1} .

The compact spectral iteration becomes attractive when we have an effective method for solving (3) accurately. The accuracy is needed in order to ensure that the computed Rayleigh-quotient matrix indeed equals (30). Thus, clearly, in this approach it is not possible to used inexact solutions.

Table 3 Computing the k smallest eigenvalues with a compact Spectral iteration, $l = \max\{40, k\}$

Matrix name	Number of iterations					
	$k = 6$	$k = 10$	$k = 20$	$k = 40$	$k = 100$	$k = 200$
Harmonic Roots	74	79	85	90	34	19
Harmonic	77	80	85	98	32	18
Geometric	75	79	85	101	34	19
Normal Dist	0	0	0	1	0	0
Uniform Dist	0	0	0	1	0	0

Table 4 Lack of monotonicity in the semi-compact Spectral iteration: Computing the $k = 10$ smallest eigenvalues of the Harmonic Roots matrix, $l = 50$

Iteration number q	Eigenvalues error Ψ_q	Iteration number q	Eigenvalues error Ψ_q	Iteration number q	Eigenvalues error Ψ_q
0	1.2E-2	15	1.2E-1	30	2.4E-10
1	3.6E-3	16	9.6E-2	31	1.4E-10
2	1.6E-3	17	7.2E-2	32	7.9E-11
3	8.8E-4	18	5.0E-2	33	4.5E-11
4	5.3E-4	19	3.1E-2	34	2.5E-11
5	3.3E-4	20	1.6E-2	35	1.4E-11
6	2.1E-4	21	5.5E-3	36	8.2E-12
7	1.3E-4	22	7.8E-4	37	4.6E-12
8	8.6E-5	23	1.1E-5	38	2.6E-12
9	5.5E-5	24	3.0E-8	39	2.0E-1
10	2.0E-1	25	4.0E-9	40	2.0E-1
11	1.9E-1	26	2.3E-9	41	1.8E-1
12	1.8E-1	27	1.3E-9	42	1.5E-1
13	1.6E-1	28	7.6E-10	43	1.1E-1
14	1.4E-1	29	4.3E-10	44	7.7E-2

The Initial Orthonormal Matrix

To start the algorithm we need an “initial” orthonormal matrix $X_0 \in \mathbb{R}^{n \times p}$ where $p = k + \ell$. In our experiments X_0 is obtained by starting the expansion process with

$$\mathbf{x}_1 = C\mathbf{e}/\|C\mathbf{e}\|_2, \quad (36)$$

$$X = [\mathbf{x}_1] \in \mathbb{R}^{n \times 1}, \quad (37)$$

and

$$\mathbf{z} = C\mathbf{x}_1. \quad (38)$$

As before, $\mathbf{e} = (1, 1, \dots, 1)^T \in \mathbb{R}^n$ and $C \in \mathbb{R}^{n \times n}$ symbolizes an inverting operator.

Table 5 Computation times and number of iterations for a Uniform Dist matrix, $n = 500,000$

Number of computed eigenvalues k	Computation time (in sec.) and number of iterations (in brackets) for various values of k and l				eigs
	$l = \max\{5, k/10\}$	$l = \max\{10, k/5\}$	$l = \max\{20, k/2\}$	$l = \max\{40, k\}$	
$k=10$	9.5 (2)	7.9 (1)	7.6 (0)	9.8 (0)	3.8
$k=20$	39.0 (16)	19.0 (5)	11.7 (1)	11.2 (0)	6.6
$k=40$	53.9 (13)	25.4 (4)	17.0 (1)	14.1 (0)	10.0
$k=60$	45.6 (7)	28.8 (3)	24.4 (1)	21.9 (0)	14.2
$k=80$	96.3 (12)	56.2 (5)	34.6 (1)	32.4 (0)	22.4
$k=100$	355.0 (38)	144.6 (12)	48.9 (1)	46.3 (0)	32.2
$k=200$	179.9 (5)	122.6 (2)	142.9 (1)	138.4 (0)	78.1
$k=300$	443.4 (8)	318.9 (3)	297.4 (1)	304.2 (0)	152.3

Numerical Experiments

In this section we describe some experiments with the proposed methods. The **semi-compact SVD** iteration is the algorithm described in Sect. 2 with C defined by solving (4). The **semi-compact Spectral** iteration is the algorithm of Sect. 5, while the **compact Spectral** iteration is the algorithm described in Sect. 6.

The test matrices are large sparse symmetric indefinite matrices that have the form

$$G = HDH^T \in \mathbb{R}^{n \times n} \quad (39)$$

where $D \in \mathbb{R}^{n \times n}$ is a diagonal matrix,

$$D = \text{diag}\{\lambda_1, \dots, \lambda_n\}, \quad (40)$$

and $H \in \mathbb{R}^{n \times n}$ is a sparse orthonormal matrix,

$$H = \prod_{i=1}^4 H_i. \quad (41)$$

The matrices H_i , $i = 1, \dots, 4$, are $n \times n$ sparse random Householder matrices,

$$H_i = I - 2\mathbf{h}_i \mathbf{h}_i^T / \mathbf{h}_i^T \mathbf{h}_i, \quad (42)$$

where $\mathbf{h}_i \in \mathbb{R}^n$ is a sparse random vector. To generate these vectors we have used MATLAB's command $\mathbf{h} = \text{sprand}(n, 1, \text{density})$ with $\text{density} = 200/n$. Since H is a product of Householder matrices, it is orthonormal matrix, and the diagonal entries of D are the eigenvalues of G . The details of the test matrices are specified in Table 1.

The figures in Tables 2–3 provide the number of iterations that were needed to satisfy the stopping condition

$$\Psi_q = \left(\sum_{j=1}^k |\lambda_j - \lambda_j^{(q)}| \right) / \left(\sum_{j=1}^k |\lambda_j| \right) \leq 10^{-13}. \quad (43)$$

As before, the eigenvalues of G are assumed to satisfy (1), and $\lambda_j^{(q)}$, $j = 1, \dots, k$, are the Ritz values at the q th iteration, $q = 0, 1, 2, \dots$ (In the SVD algorithm the Ritz values were computed from the spectral decomposition of the related Rayleigh quotient matrix (12).) If the number of iterations equals 0 it means that the stopping condition is satisfied at the starting point, which is computed as described in Sect. 7.

The experiments in Table 2 and 3 were carried out with $l = \max\{40, k\}$. Comparing these tables shows that the tested methods behave in a similar way, regarding the number of iterations. The first three test problems are difficult to solve and require several iterations. A similar difficulty is shared by MATLAB's "eigs" function. For example, using "eigs" to compute the $k = 10$ smallest eigenvalues of these matrices ends with the warning "0 of the 10 requested eigenvalues converged". Yet in the other test matrices we observe fast convergence.

The semi-compact Spectral method of Sect. 5 lacks the monotonicity property. This deficiency is illustrated in Table 4, which tracks the behavior of the eigenvalues error, Ψ_q , which is defined in (43). The results of Table 4 were obtained when computing the $k = 10$ smallest eigenvalues of the Harmonic Roots matrix. A similar lack of monotonicity was noticed in the Normal Dist matrix and the Uniform Dist matrix. These examples illustrate the importance of having the monotonicity property.

Given k , the number of eigenvalues that we want to compute, it is desired to use an "optimal" value of l , which minimizes the overall computation time that is needed to achieve

our stopping condition. Clearly, increasing l is likely to reduce the number of iterations. Yet at the same time it increases the computational effort per iteration. Hence, finding an optimal value for l is not that straightforward. The experiments in Table 5 illustrate the way k and l effect the number of iterations and the overall computation time. The results are obtained by applying the semi-compact SVD algorithm on a huge “Uniform Dist” matrix with $n = 500,000$. To make this algorithm more effective we have used MATLAB’s function “ichol” to compute an incomplete Cholesky factorization of $G^T G$. (This was done before starting the iterative process, and the timing results include the factorization time.)

The last column of Table 5 includes timing results for MATLAB’s “eigs” function. This enables us to compare SVD with “eigs”. We see that for $l = k$ the computation time of the semi-compact SVD algorithm is nearly twice of “eigs” time. This is quite encouraging since the SVD algorithm was programmed (in MATLAB) exactly as described in Sect. 2. Yet, as in other restarted Krylov methods, the basic iteration can be improved in several ways. For example, more effective orthogonalization schemes, locking of converged eigenvectors, and an improved rule for the choice of l .

Concluding Remarks

The monotonicity property ensures that each iteration provides an improved set of Ritz values. This property holds when using the Heart iteration to calculate a cluster of exterior eigenvalues. Yet, as Table 4 shows, it does not hold when calculating interior eigenvalues, which raises the question of how to overcome this difficulty.

The compact spectral method of Sect. 6 attains monotonicity by replacing G with G^{-1} . The replacement requires an effective algorithm for calculating accurate solutions of (3). But such algorithm is not always available when solving large indefinite linear systems. A second way that ensures monotonicity is the SVD iteration of Sect. 2. This option has the advantage that it is able to use inexact solutions of (3) or (4). Both iterations are very simple and enjoy fast rate of convergence. The experiments that we have done are quite encouraging.

Acknowledgements Not applicable

Author Contributions This work has one author only, A.D.

Funding No funding

Availability of data and material Not applicable

Declaration

Conflict of interest The authors declare that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Baglama, J., Calvetti, D., Reichel, L.: IRBL: an implicitly restarted block Lanczos method for large-scale Hermitian eigenproblems. *SIAM J. Sci. Comput.* **24**, 1650–1677 (2003)
2. Bai, A., Demmel, J., Dongarra, J., Ruhe, A., van der Vorst, H.: *Templates for the solution of Algebraic Eigenvalue problems: a practical guide*. SIAM, Philadelphia, PA (1999)
3. Calvetti, D., Reichel, L., Sorenson, D.C.: An implicitly restarted Lanczos method for large symmetric eigenvalue problems. *Elec. Trans. Numer. Anal.* **2**, 1–21 (1994)
4. Dax, A.: A restarted Krylov method with inexact inversions. *Numerical Linear Algebra with Applications* **26**, (2019)
5. A. Dax, *Computing the smallest singular triplets of a large matrix*, Results in Applied Mathematics **3** (2019)
6. Dax, A.: A cross-product approach for low-rank approximations of large matrices. *Journal of Computational and Applied Mathematics* (2019)
7. Dax, A.: A compact Heart iteration for large eigenvalues problems. Tech. Report, Hydrological Service of Israel (2021)
8. Dax, A.: A compact Heart iteration for low-rank approximations of large matrices. Tech. Report, Hydrological Service of Israel (2021)
9. Demmel, J.W.: *Applied Numerical Linear Algebra*. SIAM, Philadelphia (1997)
10. Golub, G.H., Van Loan, C.F.: *Matrix Computations*, 4th edn. Johns Hopkins Univ. Press, USA (2013)
11. Horn, R.A., Johnson, C.R.: *Matrix Analysis*. Cambridge University Press, Cambridge (1985)
12. R.M. Larsen, *Combining implicit restarts and partial reorthogonalization in Lanczos bidiagonalization*, <http://soi.stanford.edu/rmunk/PROPACK/>
13. Li, R., Xi, Y., Vecharynski, E., Yang, C., Saad, Y.: A thick-restart Lanczos algorithm with polynomial filtering for hermitian eigenvalue problems. *SIAM J. Sci. Comput.* **38**, 2512–2534 (2016)
14. Morgan, R.B.: On restarting the Arnoldi method for large non-symmetric eigenvalues problems. *Math. Comput.* **65**, 1213–1230 (1996)
15. Parlett, B.N.: *The Symmetric Eigenvalue Problem*. Prentice-Hall, Englewood Cliffs, NJ (1980)
16. Saad, Y.: *Numerical methods for large Eigenvalue problems*. Revised SIAM, Philadelphia (2011)
17. Sorensen, D.C.: Implicit application of polynomial filters in a k-step Arnoldi method. *SIAM J. Matrix Anal. Appl.* **13**, 357–385 (1992)
18. Stewart, G.W.: *Matrix Algorithms. Eigensystems*, SIAM, Philadelphia (2001)
19. Trefethen, L.N., Bau, D.: *Numerical Linear Algebra*. SIAM, Philadelphia (1997)
20. Watkins, D.S.: *The matrix eigenvalue problem: GR and Krylov subspace methods*. SIAM, Philadelphia (2007)
21. Wu, K., Simon, H.: Thick-restarted Lanczos method for large symmetric eigenvalue problems. *SIAM J. Matrix Anal. Appl.* **22**, 602–616 (2000)
22. Yamazaki, I., Bai, Z., Simon, H., Wang, L., Wu, K.: Adaptive projection subspace dimension for the thick-restart Lanczos method. *ACM Trans. Math. Softw.* **47**, 27 (2010)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.