

On preconditioned eigensolvers and Invert-Lanczos processes

Klaus Neymeyr

*Universität Rostock, Institut für Mathematik, Universitätsplatz 1,
18055 Rostock, Germany. (April 27, 2008)*

SUMMARY

This paper deals with the convergence analysis of various preconditioned iterations to compute the smallest eigenvalue of a discretized self-adjoint and elliptic partial differential operator. For these eigenproblems several preconditioned iterative solvers are known, but unfortunately, the convergence theory for some of these solvers is not very well understood.

The aim is to show that preconditioned eigensolvers (like the preconditioned steepest descent iteration (PSD) and the locally optimal preconditioned conjugate gradient method (LOPCG)) can be interpreted as truncated approximate Krylov subspace iterations. In the limit of preconditioning with the exact inverse of the system matrix (such preconditioning can be approximated by multiple steps of a preconditioned linear solver) the iterations behave like Invert-Lanczos processes for which convergence estimates are derived.

KEY WORDS: Elliptic eigenvalue problem; Preconditioner; Krylov space; Lanczos methods; Rayleigh quotient.

1. Introduction

Eigenvalue problems for elliptic and self-adjoint partial differential operators can be solved numerically by means of preconditioned (gradient type/approximate inverse iteration type) eigensolvers; see Chap. 11 in [2] for a survey. Geometric multigrid preconditioning and, recently, algebraic multigrid preconditioning [1, 4] have been proved useful in order to construct effective preconditioned eigensolvers. At best, numerical approximations of a fixed number of the smallest eigenvalues together with the eigenvectors can be computed with optimal complexity, i.e., with total costs increasing linearly in the number of unknowns.

The discretization of such an operator eigenproblem leads to the generalized eigenvalue problem

$$Ax = \lambda Mx$$

for symmetric positive definite n -by- n matrices A and M . The associated Rayleigh quotient reads

$$\rho(x) = \frac{(x, Ax)}{(x, Mx)}.$$

Here we are interested in iterative methods which work with the gradient of the Rayleigh quotient

$$\nabla\rho(x) = \frac{2}{(x, Mx)}(Ax - \rho(x)Mx).$$

The negative gradient vector is the *direction of correction* underlying the (basic and ineffective) gradient iteration

$$x_{j+1} = x_j - \omega_j \nabla\rho(x_j)$$

where ω_j is a scaling parameter. A significantly faster convergence can be gained by proper preconditioning. D'yakonov [6] suggests to interpret preconditioning as a change of the underlying geometry. Preconditioned gradient type iterations for the eigenproblem work with a *preconditioned gradient vector*. The gradient ∇_B is derived with respect to a Euclidean space whose inner product is induced by a symmetric positive definite matrix B . It is assumed that $B^{-1}A$ is in some sense close to the identity matrix and B^{-1} (sometimes B) is called a *preconditioner*. The B -gradient reads

$$\nabla_B \rho(x) = B^{-1} \nabla \rho(x) = \frac{2}{(x, Mx)} B^{-1} (Ax - \rho(x)Mx).$$

Proper spectral assumptions on the preconditioner B , see Section 1.3, guarantee that the iterates

$$x_{j+1} = x_j - B^{-1}(Ax_j - \rho(x_j)Mx_j) \quad (1)$$

converge to an eigenvector and that the Rayleigh quotients $\rho(x_j)$ form a monotone decreasing sequence converging to the corresponding eigenvalue. Iterations like (1) were first analyzed in 1958 by Samokish [32] and later on by several, mainly Russian, authors; see [2] for a survey.

To consider (1) as a preconditioned gradient type iteration is one, but perhaps not the optimal point of view for analyzing its convergence. Sharp convergence estimates for (1) have been derived by interpreting the method as a preconditioned variant of inverse iteration [18]. The key equation underlying this interpretation is a reformulation of (1) in the form of an *error propagation equation*, i.e.,

$$x_{j+1} = \rho(x_j)A^{-1}Mx_j + (I - B^{-1}A)(x_j - \rho(x_j)A^{-1}Mx_j). \quad (2)$$

In (2) the new iterate x_{j+1} is represented as the result of scaled inverse iteration applied to x_j , that is $\rho(x_j)A^{-1}Mx_j$, plus a perturbation $(I - B^{-1}A)(x_j - \rho(x_j)A^{-1}Mx_j)$ whose magnitude is controlled by the spectral radius of the error propagation matrix $I - B^{-1}A$. A direct way to derive (2) is to consider the linear system

$$Az = \rho(x_j)Mx_j \quad (3)$$

for z . Then $z = \rho(x_j)A^{-1}Mx_j$ results from applying (non-shifted) inverse iteration to x_j . The approximate (or preconditioned) solution of (3) gives (2). If (3) is solved approximately by a number of k steps of a preconditioned iteration, then the error propagation matrix $I - B^{-1}A$ in (2) is substituted by its k th power. This shows how practically the preconditioning with preconditioners close to A^{-1} can be realized.

1.1. Acceleration with the Rayleigh-Ritz procedure

The basic preconditioned iteration (1) can be significantly accelerated by means of the Rayleigh-Ritz procedure. Such improved preconditioned eigensolvers are the well-known *preconditioned steepest descent* (PSD) iteration [2] and the *locally optimal preconditioned conjugate gradients* (LOPCG) scheme [2, 13].

For PSD the optimal step-length is

$$\omega_j = \arg \min_{\omega \in \mathbb{R}} \rho(x_j - \omega B^{-1}(Ax_j - \rho(x_j)Mx_j)),$$

which implicitly defines an optimally scaled preconditioner $\omega_j B^{-1}$. The next PSD-iterate is

$$x_{j+1} = x_j - \omega_j B^{-1}(Ax_j - \rho(x_j)Mx_j). \quad (4)$$

The step length ω_j depends on x_j , A , M and B^{-1} , e.g., see Kantorovich [10]. An equivalent way to compute the PSD-iterate x_{j+1} is to apply the Rayleigh-Ritz procedure to the 2D subspace

$$\mathcal{S}_j^{(2)} = \text{span}\{x_j, B^{-1}(Ax_j - \rho(x_j)Mx_j)\}. \quad (5)$$

The Ritz vector corresponding to the smallest Ritz value is collinear to x_{j+1} in (4). The optimality of the Rayleigh-Ritz approximation shows that estimates for (1) apply as upper estimates for PSD. However, such trivial upper estimates are not sharp; see also [2, 8, 27, 32].

The *locally optimal preconditioned conjugate gradient* (LOPCG) method derives by enlarging the PSD subspace $\mathcal{S}_j^{(2)}$ by the preceding iterate x_{j-1} which yields

$$\mathcal{S}_j^{(3)} = \text{span}\{x_{j-1}, x_j, B^{-1}(Ax_j - \rho(x_j)Mx_j)\}. \quad (6)$$

The new iterate $x_{j+1} \in \mathcal{S}_j^{(3)}$ is a Ritz vector corresponding to the smallest Ritz value and minimizes the Rayleigh quotient in $\mathcal{S}_j^{(3)}$, i.e.

$$x_{j+1} \in \arg \min \rho(\mathcal{S}_j^{(3)}). \quad (7)$$

The LOPCG method and its block variant (LOBPCG), which is used to compute several of the smallest eigenvalues/vector approximations simultaneously, have been introduced by A. Knyazev. He suggested the scheme in [13], see also [2, 14, 16]. The acronym LOPCG contains the term “conjugate gradients”, which alludes to strong relations to the preconditioned conjugate gradient iteration (PCG) for linear systems. Numerical experiments show a striking similarity of the convergence behavior of PCG and LOPCG if a linear system (discretization of a boundary value problem) and a matrix eigenvalue problem (discretization of an operator eigenvalue problem) are considered for the same elliptic and self-adjoint partial differential operator. In such comparative studies both for the linear system (PCG) and for the eigenvalue problem (LOPCG) the same multigrid preconditioner can be applied.

Unfortunately, there is no sound theory available up to now providing sharp convergence estimates for LOPCG (aside from trivial upper bounds as derived for the slower converging PSD scheme or even for the scheme (1)). A partial answer is given in Sections 3 and 4, where upper and lower estimates are derived for some types of best and poorest preconditioning.

1.2. Simplifications

Without restriction of generality, we always make use of a diagonalizing basis, i.e., we transform the generalized eigenvalue problem (A, M) by means of an M -orthogonal basis of eigenvectors to the standard eigenproblem for a symmetric and positive definite matrix, once again denoted by A . All convergence estimates which are derived with respect to this diagonal problem apply without any changes to the original problem. The diagonal eigenproblem is denoted by

$$Ax = \lambda x$$

with

$$A = \text{diag}(\lambda_1, \dots, \lambda_n), \quad (8)$$

so that the eigenvector e_i corresponding to λ_i is just the i -th column of the n -by- n identity matrix. We assume $0 < \lambda_1 < \lambda_2 \leq \dots \leq \lambda_n$. The smallest eigenvalue λ_1 is assumed to be a simple eigenvalue for the sake of a simple representation. Multiple eigenvalues do not add fundamental difficulties to the problem, see Sec. 3 in [18].

In this paper we are mainly interested in the smallest eigenvalue λ_1 and the corresponding eigenvector e_1 . Implicitly all eigeniterations which use the Rayleigh-Ritz procedure provide further eigenvalue approximations/Ritz vectors whose quality is not analyzed here.

1.3. Preconditioning

In general, for any n -by- n symmetric positive definite matrix B^{-1} (the preconditioner) real constants γ_0, γ_1 with $0 < \gamma_0 \leq \gamma_1$ exist so that

$$\gamma_0(x, Bx) \leq (x, Ax) \leq \gamma_1(x, Bx), \quad \forall x \in \mathbb{R}^n. \quad (9)$$

The true importance of (9) is based on cases where A is a finite element discretization of an elliptic and self-adjoint partial differential operator. Then the constants γ_0, γ_1 , at best, do not depend on the mesh parameter h for certain geometric multigrid or multilevel preconditioners B^{-1} . In such cases preconditioned eigensolvers can converge with a *grid-independent convergence rate* and can provide a fixed number of eigenvector and eigenvalue approximations with computational costs which increase linearly in the number of unknowns (case of *optimal complexity*).

The ratio γ_1/γ_0 is the spectral condition number of the preconditioned matrix $B^{-1}A$. If B is ideally scaled, then it can be assumed that

$$\|I - B^{-1}A\|_A \leq \gamma, \quad 0 \leq \gamma < 1. \quad (10)$$

Therein $\|\cdot\|_A$ denotes the A -operator norm which is induced by the A -based vector norm $\|x\|_A = (x, Ax)^{1/2}$.

Inequality (9) follows from (10) with $\gamma_0 = 1 - \gamma$ and $\gamma_1 = 1 + \gamma$. If (9) is assumed to hold, then the optimally scaled preconditioner $2B/(\gamma_0 + \gamma_1)$ in place of B satisfies (10) with $\gamma = (\gamma_1 - \gamma_0)/(\gamma_1 + \gamma_0)$.

Next we always assume (10). This does not restrict the generality of the analysis, as the Rayleigh-Ritz procedure in the following Algorithms 1 and 2 implicitly computes the optimal scaling constant.

1.4. Overview and aim

The core issue of this paper is to show that various *preconditioned iterations* for the solution of *positive definite* (mesh) eigenproblems can be understood as approximate Invert-Lanczos processes. Here the analysis is restricted to eigenproblems for positive definite matrices and, in the same way, the preconditioning is confined to positive definite operators, cf. Section 1.3. This restriction is made on account of our setting of an (*adaptive*) *finite element discretization of a self-adjoint partial differential operator*. Positive definite preconditioning is realized by (one or more cycles of) a multigrid solver. Indefinite preconditioners approximating in some sense $(A - \sigma M)^{-1}$ ($M = I$ on the assumptions of Sect. 1.2) with $\sigma > \lambda_1$ are not within the scope of our method. Typically, symmetric positive definite *multigrid* preconditioners can be realized with only linearly increasing computational costs (optimal complexity) and, at best, convergence rates can be guaranteed which do not depend on the mesh size. In contrast to this, the multigrid preconditioning of indefinite problems is complicated and computationally very expensive.

The restriction to positive definite matrices distinguishes this work from the recent approach by Stathopoulos [34], where the Generalized Davidson (GD) method and its variants have been investigated. In that paper the optimality of various eigensolvers for the hermitian eigenproblem to compute the smallest eigenvalue is considered. An important result is that the $\text{GD}(m_{\min}, m_{\max}) + 1$ scheme appears to be even more effective than the LOBPCG solver. However, in the current paper we pursue a different approach, i.e., we only consider positive definite preconditioners and get somewhat different results. The focus of [34] is the construction and the analysis of optimal solvers including the use of indefinite preconditioners. In such a setting the GD as well as the JDQMR scheme [34] appear to be nearly optimal candidates; cf. [22–24] for recent results on the analysis of GD-like schemes.

In contrast to this, the intention of this paper is not to construct *optimal* Invert-Lanczos processes. Instead, preconditioned gradient type eigensolvers are studied in the limit of "exact-inverse preconditioning" ($B = A$) and the convergence of the resulting Invert-Lanczos processes is

analyzed. According to that, the preconditioned eigensolvers (1), (4) and (7) can be interpreted as *approximate Krylov subspace iterations* within a general hierarchy of preconditioned eigeniterations; see Algorithm 1. Following this point of view we derive lower and upper convergence estimates (corresponding to exact and poorest preconditioning).

Though the results of this paper have a prevailing theoretical character, it is clear how the convergence behavior for $B \approx A$ can be approximated for any (multigrid) preconditioner satisfying (10): If k steps of a preconditioned linear solver are applied to the linear system (3), then $(I - B^{-1}A)^k$ is the resulting error propagator in (2). This amounts to the action of a preconditioner with the spectral radius of the error propagation matrix being γ^k . However, computational experiments show that the additional computational costs for the improved solution of the linear system cannot be justified by the potential acceleration of convergence.

The paper is structured as follows: In Section 2 a hierarchy of preconditioned eigensolvers is introduced, which includes the eigensolvers (1), (4) and (7). The aim of Section 3 is to show that these solvers in the limit $B = A$ are certain Invert-Lanczos processes, and to give a convergence analysis for these Lanczos type solvers. The convergence analysis in the general case $B \neq A$ is still an open problem for LOPCG and more complex schemes. A partial solution is given in Section 4, where bounds are derived on the fastest and slowest possible convergence corresponding to the best and poorest possible preconditioning.

2. A hierarchy of preconditioned eigensolvers

In this section, a unifying framework is suggested for a class of preconditioned gradient type eigensolvers. This framework includes the *Preconditioned gradient iteration* or *Preconditioned inverse iteration* (PINVIT) by Eq. (1), the *Preconditioned steepest descent scheme* (PSD) by Eq. (4), the *Locally optimal preconditioned conjugate gradient iteration* (LOPCG) by Eq. (7) and more general schemes.

Common to all these eigensolvers is the *correction direction*

$$d_j := B^{-1}(Ax_j - \rho(x_j)x_j) = \frac{(x_j, x_j)}{2} \nabla_B \rho(x_j)$$

which is the preconditioned residual of the j th iterate x_j or the B -gradient of the Rayleigh quotient in x_j . The new iterate x_{j+1} is formed from the preconditioned residual d_j by a *suboptimal* linear combination with x_j (in the case of (1)), by an *optimal* linear combination with x_j (in the case of PSD) or by an *optimal* linear combination with x_j and x_{j-1} (in the case of LOPCG). Optimality means that x_{j+1} minimizes the Rayleigh quotient either with respect to the trial subspace $\mathcal{S}_j^{(2)}$ given by (5) or $\mathcal{S}_j^{(3)}$ given by (6). In each of these cases the new iterate is the Ritz vector corresponding to the smallest Ritz value.

The straightforward generalization is to apply the Rayleigh-Ritz procedure to the nested subspaces

$$\mathcal{S}_j^{(k)} := \text{span}\{x_{j-k+2}, \dots, x_j, B^{-1}(Ax_j - \rho(x_j)Mx_j)\}, \quad k \geq 2, \quad j \in \mathbb{N}, \quad (11)$$

which are formed by stepwise expansion in the previous iterates $x_{j-1}, \dots, x_{j-k+2}$. Hence $\mathcal{S}_j^{(2)} \subseteq \mathcal{S}_j^{(3)} \subseteq \dots \subseteq \mathcal{S}_j^{(k)}$. The smallest subspace $\mathcal{S}_j^{(2)}$ is associated with PSD. The Courant-Fischer principles guarantee a monotone decrease of the smallest attainable Ritz value in $\mathcal{S}_j^{(k)}$ for increasing k which shows the stabilizing effect of such subspace enlargements.

The application of the Rayleigh-Ritz procedure to $\mathcal{S}_j^{(k)}$ defines a hierarchy of preconditioned eigensolvers; see Algorithm 1. These can be called PINVIT(k), or briefly, the (k)-scheme since the

k	Eigensolver	Subspace
$k = 1$	Preconditioned gradient iteration/inverse iteration	$[x_j - d_j] \in \mathbb{R}^n$
$k = 2$	Preconditioned steepest descent	$[x_j, d_j] \in \mathbb{R}^{n \times 2}$
$k = 3$	Locally optimal preconditioned conjugate gradient	$[x_{j-1}, x_j, d_j] \in \mathbb{R}^{n \times 3}$
$k \geq 4$	Higher order schemes; practically of minor importance	$\mathcal{S}_j^{(k)} \in \mathbb{R}^{n \times k}$

Table I. (k) -scheme, Algorithm 1, for small k ; d_j is the preconditioned residual of x_j .

schemes derive from the basic iteration (1) whose close relation to inverse iteration has already been pointed out in Section 1.

Algorithm 1 (k) -scheme, PINVIT(k), $k \geq 1$.

Required: Matrix-vector products as subroutines $y \mapsto Ay$, $z \mapsto B^{-1}z$.

Input: $x_1 \in \mathbb{R}^n \setminus \{0\}$ and $k \geq 1$.

- Initialization:** If $k \geq 3$, then compute an initial sequence of $k - 2$ vectors x_2, \dots, x_{k-1} by executing single steps of the (m) -scheme with the initial sequence x_1, \dots, x_{m-1} for $m = 2, \dots, k - 1$.
- Iteration:**
If $k = 1$, then compute for $j = 1, 2, \dots$

$$\tilde{x}_{j+1} = x_j - B^{-1}(Ax_j - \rho(x_j)x_j), \quad x_{j+1} = \tilde{x}_{j+1}/\|\tilde{x}_{j+1}\|.$$

If $k \geq 2$, then for $j = k - 1, k, k + 1, \dots$, until (approximate) convergence do:
Compute

$$d_j := B^{-1}(Ax_j - \rho(x_j)x_j), \tag{12}$$

apply the Rayleigh-Ritz procedure to the column space of

$$\mathcal{S}_j^{(k)} := [x_{j-k+2}, \dots, x_j, d_j] \in \mathbb{R}^{n \times k} \tag{13}$$

and let x_{j+1} be a Ritz vector corresponding to the smallest Ritz value.

For $k \geq 2$ the Rayleigh-Ritz procedure in Algorithm 1 guarantees that the Rayleigh quotients of x_j form a monotone decreasing sequence for any (even indefinite and/or non-symmetric) preconditioner B^{-1} . Thus the method is robust with respect to the choice of the preconditioner. This includes that the scaling of the preconditioner (see also Section 1.3) is of no importance since any (nonzero) multiple of d_j does not change the Ritz approximations from $\mathcal{S}_j^{(k)}$, $k \geq 2$.

Table I summarizes the (k) -schemes for $k = 1, 2, 3$. Numerical experiments as given in [14] or more explicitly in [16] give clear evidence of the minor practical relevance of the (k) -scheme for $k \geq 4$. In [16] multigrid preconditioning has been used for mesh eigenproblems; for $k \geq 4$ no mentionable speedup has been observed compared to LOPCG, but at the same time, the computational costs increase in k . The practical experiences recommend the LOPCG scheme as the optimal choice within the (k) -scheme hierarchy of preconditioned eigensolvers.

A subspace variant of the (k) -scheme for computing an eigenspace corresponding to several of the smallest eigenvalues can be formulated in a self-suggesting way; see [19] for the explicit construction.

3. The limit $B = A$ and the Invert-Lanczos process

Here a convergence analysis is given for Algorithm 1 in the limit $B = A$. Then the preconditioned eigensolvers can be interpreted as *truncated/implicitly restarted Invert-Lanczos* processes in the Krylov subspaces

$$\mathcal{K}_j(A^{-1}, x_1) = \text{span}\{x_1, A^{-1}x_1, \dots, A^{-(j-1)}x_1\}.$$

For very large matrices A these Krylov subspaces based on A^{-1} cannot be used in practice as the solution of the linear systems in A is too expensive. However, the *asymptotic convergence analysis* for $\gamma \rightarrow 0$, see (10), describes the limit of working with the exact inverse $B^{-1} = A^{-1}$. Studying this limit is not an academic question, but within the scope of any practical preconditioner as the action of a high-quality preconditioner with $\gamma \approx 0$ can be emulated by applying several steps of a (poorly) preconditioned iteration to solve the linear system (3). In the following we call $\gamma = 0$ the limit of *exact inverse preconditioning*.

The substitution $B = A$ in Algorithm 1 yields the INVIT(k) iteration, see Algorithm 2. For $k = 1$ the resulting scheme is the basic non-shifted inverse iteration (INVIT). For $k \geq 2$ inverse iteration is combined with the Rayleigh-Ritz procedure.

Algorithm 2 INVIT(k) for $k \in \mathbb{N}$

Input: A s.p.d., $x_1 \in \mathbb{R}^n \setminus \{0\}$ and $k \geq 1$

1. **Initialization:** If $k \geq 3$, then compute an initial sequence of $k - 2$ vectors x_2, \dots, x_{k-1} by executing single steps of INVIT(m) with the initial sequence x_1, \dots, x_{m-1} for $m = 2, \dots, k - 1$.

2. **Iteration:**

If $k = 1$, then solve for $j = 1, 2, \dots$

$$A\tilde{x}_{j+1} = x_j, \quad x_{j+1} = \tilde{x}_{j+1}/\|\tilde{x}_{j+1}\|.$$

If $k \geq 2$, then for $j = k - 1, k, k + 1, \dots$, until (approximate) convergence do:
Solve the linear system

$$Au_j = x_j$$

and apply the Rayleigh-Ritz procedure to the column space of

$$T_j^{(k)} := [x_{j-k+2}, \dots, x_j, u_j] \in \mathbb{R}^{n \times k}. \quad (14)$$

Let x_{j+1} be a Ritz vector corresponding to the smallest Ritz value.

3.1. The Krylov subspace $\mathcal{K}_j(A^{-1}, x_1)$

Algorithm 2 works in the Krylov subspace

$$\mathcal{K}_j(A^{-1}, x_1) = \{p(A^{-1})x_1; p \text{ polynomial with } \deg p \leq j - 1\} \quad (15)$$

as shown in Lemma 3.2. However, k is a truncation parameter which controls the computational costs of the Rayleigh-Ritz projections. For $j \geq k$ the Ritz vector x_j is in the (smaller) column space of $T_j^{(k)}$ being a subspace of $\mathcal{K}_j(A^{-1}, x_1)$.

In this sense Algorithm 2 can be interpreted as an *implicitly restarted Invert-Lanczos process* (for IR-Lanczos see Sec. 4.5 of [2] and for SI-Lanczos see [7],[29]). To make this clear, first note that for

the Krylov subspaces $\mathcal{K}_j(A, x_1)$ the stepwise extraction of Rayleigh-Ritz projections is well-known as the Lanczos process, as expressed by, [26, Ch. 13],

$$\mathcal{K}_j(A, x_1) + \text{Rayleigh-Ritz}(A) \equiv \text{Lanczos}(A, x_1, j). \quad (16)$$

Therein, $\text{Lanczos}(A, x_1, j)$ denotes the j th step of the Lanczos process for A with the starting vector x_1 . In general, one might express this also as

$$\text{Krylov space} + \text{Orthogonalization} \equiv \text{Lanczos},$$

but here we prefer (16) because of its contrast with (17).

The crucial point is that Algorithm 2 works with the inverse matrix A^{-1} . In the initialization phase of $\text{INVIT}(k)$ for $j < k$ the column space of

$$T_j^{(k)} = [x_1, \dots, x_j, A^{-1}x_j] \in \mathbb{R}^{n \times j+1}$$

is the Krylov subspace $\mathcal{K}_{j+1}(A^{-1}, x_1)$. Then the Rayleigh-Ritz procedure for A (and not for A^{-1}) is applied in order to extract the Rayleigh-Ritz approximations. This is an Invert-Lanczos process with a modified starting vector as shown in the next lemma.

Lemma 3.1. *The initialization phase of $\text{INVIT}(k)$ for $j < k$ is an Invert-Lanczos process with the starting vector $A^{1/2}x_1$ which is expressed by*

$$\mathcal{K}_j(A^{-1}, x_1) + \text{Rayleigh-Ritz}(A) \equiv \text{Lanczos}(A^{-1}, A^{1/2}x_1, j). \quad (17)$$

Proof. The Lanczos process for A^{-1} with the initial vector $A^{1/2}x_1$ works in the Krylov subspaces $\mathcal{K}_j(A^{-1}, A^{1/2}x_1)$ with the Krylov matrices $\bar{K}_j = (A^{1/2}x_1, A^{-1/2}x_1, \dots, A^{1-j}A^{1/2}x_1) \in \mathbb{R}^{n \times j}$. Then $K_j = A^{-1/2}\bar{K}_j$ has the column space $\mathcal{K}_j(A^{-1}, x_1)$. The Lanczos process for A^{-1} with \bar{K}_j gets the Rayleigh-Ritz approximations from the matrix pencil $(\bar{K}_j^T A^{-1} \bar{K}_j, \bar{K}_j^T \bar{K}_j)$. Since $(\bar{K}_j^T A^{-1} \bar{K}_j, \bar{K}_j^T \bar{K}_j) = (K_j^T K_j, K_j^T A K_j)$ the Ritz values generated by $\text{Lanczos}(A^{-1}, A^{1/2}x_1, j)$ are the inverses of the Ritz values gained in $\mathcal{K}_j(A^{-1}, x_1)$. \square

Algorithm 2 is a simple Invert-Lanczos and not a Shift-and-Invert (SI) Lanczos process (with a nonzero shift). For appropriate shifts the SI-Lanczos process can converge much faster compared to the non-shifted scheme, and SI-Lanczos allows us to compute eigenvalues in the interior of the spectrum. A substantial drawback is that the indefinite linear systems are to be solved accurately. The solution of these linear systems is usually realized by direct solvers, and restarting is the key to an effective realization of SI-Lanczos [5, 37]. In our setting of very large eigenproblems which derive from the discretization of partial differential operators we cannot treat indefinite systems effectively because of the following reason: First, *direct solvers* cannot be applied to these large indefinite linear systems because of the computational costs. Second, *iterative solvers* can not be effectively applied to these indefinite linear system since effective multigrid preconditioning of indefinite systems is still a non-trivial problem (see, e.g., [21, 36, 38] and [3] for promising new developments). Hence the preconditioning in Algorithm 1 is only applied to s.p.d. linear systems (3) and preconditioning of indefinite systems is not within the scope of these schemes.

To prepare the further analysis note that the dimension of the Krylov subspace $\mathcal{K}_j(A^{-1}, x_1)$ is always less or equal to j . Its maximal dimension is the *grade* with respect to A^{-1} of x_1 denoted by $\text{grade}(A^{-1}, x_1)$. In general, the grade of v with respect to $Y \in \mathbb{R}^{n \times n}$ is the lowest degree of a nonzero (minimal) polynomial $p(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_\mu t^\mu$, $\alpha_\mu \neq 0$, so that $p(Y)v = 0$. If Y is a regular matrix, then $\alpha_0 \neq 0$ as otherwise the factorization $p(t) = t(\alpha_1 + \dots + \alpha_\mu t^{\mu-1})$ would imply $\text{grade}(Y, v) \leq \mu - 1$. Further, for invertible Y the factorization

$$0 = (\alpha_0 I + \dots + \alpha_\mu Y^\mu)v = Y^\mu(\alpha_\mu I + \dots + \alpha_0 Y^{-\mu})v$$

shows that

$$\text{grade}(Y, v) = \text{grade}(Y^{-1}, v). \quad (18)$$

Thus, both $\mathcal{K}_j(A^{-1}, x_1)$ and $\mathcal{K}_j(A, x_1)$ have the maximal dimension $\text{grade}(A, x_1)$.

We always assume that the grade μ is sufficiently large (larger than k in Algorithm 2). A small grade is not a misfortune as $\mathcal{K}_\mu(A^{-1}, x_1)$ with $\mu = \text{grade}(A, x_1)$ is an A -invariant subspace [31, Prop. 6.2]. Random initial vectors x_1 rarely have a small grade ($\mu < n$) since the A -invariant linear subspaces form a set of measure zero within the set of all linear subspaces of the \mathbb{R}^n .

Lemma 3.2. *The INVIT(k)-iterates x_j satisfy*

$$x_j \in \mathcal{K}_j(A^{-1}, x_1), \quad j \geq 1.$$

If $\text{grade}(A, x_1) \geq j$, then any linear expansion of x_j in $x_1, A^{-1}x_1, \dots, A^{-(j-1)}x_1$ has a non-vanishing coefficient in $A^{-(j-1)}x_1$ or the iteration terminates in x_{j-1} being an eigenvector.

Proof. If $x_{j-1} \in \mathcal{K}_{j-1}(A^{-1}, x_1)$, then $u_{j-1} = A^{-1}x_{j-1}$ and for the Ritz vector x_j it holds that $x_j \in \text{span}(T_{j-1}^{(k)}) \subseteq \mathcal{K}_j(A^{-1}, x_1)$.

If x_{j-1} is not an eigenvector of A , then $\rho(A^{-1}x_{j-1}) < \rho(x_{j-1})$ by the strictly monotone decrease of the Rayleigh quotient due to inverse iteration. Then $\min \rho(\mathcal{K}_j(A^{-1}, x_1)) < \min \rho(\mathcal{K}_{j-1}(A^{-1}, x_1))$ so that x_j must have a non-vanishing expansion coefficient in $A^{1-j}x_1$. \square

3.2. Convergence estimates

The following convergence estimates adapt the classical proofs of Kaniel [9], Paige [25] and Saad [30] (KPS-technique) to the Lanczos process $\text{Lanczos}(A^{-1}, A^{1/2}x_1, j)$. Proofs are omitted if classical results concerning A are applied to A^{-1} . If Rayleigh-Ritz approximations for A are extracted from Krylov subspaces generated by A^{-1} , then slight modifications are required in the classical proofs. The latter Rayleigh-Ritz approximations for A in a Krylov space generated by A^{-1} can also be considered as harmonic Ritz extractions for A^{-1} ; c.f. Sec. 3.2 in [2] and the references therein. The following estimates are upper bounds on the eigenvector/value approximations which are computed in the initialization phase of INVIT(k) and, partially, for the iteration phase (see Section 3.3).

Lemma 3.3 provides a representation for the acute angle enclosed by the i th eigenvector e_i of A and the Krylov subspace $\mathcal{K}_m(A^{-1}, x_1)$.

Lemma 3.3. *If $x \in \mathbb{R}^n$ with $x^T e_i \neq 0$, then the acute angle $\varphi(e_i, \mathcal{K}_m(A^{-1}, x))$ enclosed by the eigenvector e_i and the Krylov subspace $\mathcal{K}_m(A^{-1}, x)$ is given by*

$$\tan \varphi(e_i, \mathcal{K}_m(A^{-1}, x)) = \min_{p \in \mathbb{P}_{m-1}, p(\lambda_i^{-1})=1} \|p(A^{-1})z_i\| \tan \varphi(e_i, x) \quad (19)$$

with

$$z_i = \begin{cases} \frac{(I-P_i)x}{\|(I-P_i)x\|}, & \text{if } (I-P_i)x \neq 0, \\ 0 & \text{else.} \end{cases}$$

Therein \mathbb{P}_{m-1} is the set of polynomials with a degree less or equal to $m-1$. Furthermore P_i is the spectral projector on the i th eigenvector, i.e., $P_i x = x|_i e_i$.

The proof follows follows by replacing A by A^{-1} in Lemma 6.1 of [31].

Next the approximation properties of $\mathcal{K}_m(A^{-1}, x)$ are described and the error of the smallest Ritz value is estimated.

Theorem 3.4. *The Krylov subspace $\mathcal{K}_m(A^{-1}, x)$ provides an eigenvector approximation for e_i whose quality is controlled in terms of the Chebyshev polynomial T_{m-i} in the following manner*

$$\tan \varphi(e_i, \mathcal{K}_m(A^{-1}, x)) \leq \frac{\kappa_i}{T_{m-i}(1 + 2\delta_i)} \tan \varphi(e_i, x) \quad (20)$$

with

$$\kappa_1 = 1, \quad \kappa_i = \left| \prod_{j=1}^{i-1} \frac{1/\lambda_j - 1/\lambda_n}{1/\lambda_j - 1/\lambda_i} \right|, \quad \delta_i = \frac{1/\lambda_i - 1/\lambda_{i+1}}{1/\lambda_{i+1} - 1/\lambda_n}. \quad (21)$$

For the smallest Ritz value $\theta_1^{(m)}$ it holds

$$0 \leq \theta_1^{(m)} - \lambda_1 \leq (\lambda_n - \lambda_1) \left(\frac{\tan \varphi(e_1, x)}{T_{m-1}(1 + 2\delta_1)} \right)^2. \quad (22)$$

Proof. The proof follows Thm. 6.4 in [31], but here Rayleigh-Ritz approximations for A are taken with respect to $\mathcal{K}_m(A^{-1}, x)$. First $\|p(A^{-1})z_i\|$ in (19) for $i = 1$ is estimated. The eigenvector expansion $z_1 = \sum_{j=2}^n \eta_j e_j$ with $\|z_1\| = 1$ yields

$$\|p(A^{-1})z_1\|^2 = \sum_{j=2}^n (p(\lambda_j^{-1}))^2 \eta_j^2 \leq \max_{j=2, \dots, n} (p(\lambda_j^{-1}))^2 \leq \max_{\lambda \in [1/\lambda_n, 1/\lambda_2]} |p(\lambda^{-1})|^2.$$

The Chebyshev polynomial T_{m-1} shifted to $[1/\lambda_n, 1/\lambda_2]$ provides the upper bound

$$\begin{aligned} \min_{p \in \mathbb{P}_{m-1}, p(\lambda_1^{-1})=1} \|p(A^{-1})z_1\| &\leq \min_{p \in \mathbb{P}_{m-1}, p(\lambda_1^{-1})=1} \max_{\lambda \in [1/\lambda_n, 1/\lambda_2]} |p(\lambda^{-1})| \\ &\leq \left(T_{m-1} \left(1 + 2 \frac{1/\lambda_1 - 1/\lambda_2}{1/\lambda_2 - 1/\lambda_n} \right) \right)^{-1}. \end{aligned} \quad (23)$$

In the general case $i \neq 1$ the minimization can be restricted to all polynomials having the form

$$p(\lambda) = \frac{(1/\lambda_1 - \lambda) \cdots (1/\lambda_{i-1} - \lambda)}{(1/\lambda_1 - 1/\lambda_i) \cdots (1/\lambda_{i-1} - 1/\lambda_i)} q(\lambda)$$

with $\deg q(\lambda) \leq m - i$ and $q(\lambda_i^{-1}) = 1$. Bounding this with the Chebyshev polynomial T_{m-i} yields (21). From

$$\theta_1^{(m)} - \lambda_1 = \min_{0 \neq p \in \mathbb{P}_{m-1}} \frac{((A - \lambda_1 I)p(A^{-1})x, p(A^{-1})x)}{(p(A^{-1})x, p(A^{-1})x)}$$

an upper estimate using (23) yields (22). \square

3.3. Explicit estimates for INVIT(2)

Theorem 3.4 provides an explicit estimate for the convergence of INVIT(2) toward e_1 (the eigenvector corresponding to λ_1). For $m = 2$ and $i = 1$ it holds

$$\tan \varphi(e_1, \text{span}\{A^{-1}x, x\}) \leq \frac{1}{T_1 \left(1 + 2 \frac{\lambda_1^{-1} - \lambda_2^{-1}}{\lambda_2^{-1} - \lambda_n^{-1}} \right)} \tan \varphi(e_1, x).$$

The convergence factor $\vartheta_L = 1/T_1(1 + 2\delta_1)$ is

$$\vartheta_L = \frac{\lambda_1(\lambda_n - \lambda_2)}{\lambda_2(\lambda_n - \lambda_1) + \lambda_n(\lambda_2 - \lambda_1)}. \quad (24)$$

Thus ϑ_L is smaller than the convergence factor

$$\vartheta = \frac{\lambda_1(\lambda_n - \lambda_2)}{\lambda_2(\lambda_n - \lambda_1)}$$

which has been derived in [19, Thm. 6.3] for the convergence of the Ritz vector corresponding to the smallest Ritz value in $\text{span}\{A^{-1}x, x\}$. There ϑ has been gained by the adaption of convergence estimates on steepest ascent/descent in $\text{span}\{Ax, x\}$ to the Krylov subspace $\text{span}\{A^{-1}x, x\}$ by using mini-dimensional proof techniques. Such convergence estimates on steepest ascent/descent in $\text{span}\{Ax, x\}$ have a long history; see Kantorovich [10, 11] and Hestenes and Karush [8] for classical asymptotic estimates and for non-asymptotic estimates [12, 17, 28, 39].

For mesh eigenproblems with the discretization parameter $h \rightarrow 0$ and so $\lambda_n \rightarrow \infty$ one obtains

$$\lim_{\lambda_n \rightarrow \infty} \vartheta_L = \frac{\lambda_1}{\lambda_2 + (\lambda_2 - \lambda_1)} < \frac{\lambda_1}{\lambda_2} = \lim_{\lambda_n \rightarrow \infty} \vartheta$$

confirming *mesh-independent* convergence for both estimates.

The Ritz value estimate (22), once again for $m = 2$, $i = 1$ and $\vartheta_L = 1/T_1(1 + 2\delta_1)$, with the convergence factor $(\lambda_n - \lambda_1)(\vartheta_L)^2 \tan^2 \varphi(e_1, x)$ is unsatisfying in a PDE context (cf. Section 3.4). Grid-independent estimates for $m = 1, 2$ are well known. For instance Theorem 6.3 in [19] shows for the smallest Ritz value $\theta_1^{(2)}$ generated by the (2)-scheme that

$$\frac{\Delta_{1,2}(\theta_1^{(2)})}{\Delta_{1,2}(\rho(x))} \leq \left(\frac{1 - \xi}{1 + \xi} \right)^2$$

with

$$\xi = \frac{\lambda_2 - \lambda_1}{\lambda_2 - \frac{\lambda_1 \lambda_2}{\lambda_n}} \quad \text{and} \quad \Delta_{1,2}(\kappa) = \frac{\kappa - \lambda_1}{\lambda_2 - \kappa}.$$

3.4. Lanczos(A) vs. Lanczos(A^{-1}) in a PDE context

The Lanczos process simultaneously approximates both the largest and the smallest eigenvalues though the rate of convergence to the smallest eigenvalues may be slower. In the case of PDE eigenproblems we are typically interested in some of the smallest eigenvalues and, sometimes, in eigenvalues which are in the interior lower part of the spectrum. The largest eigenvalues of the discretized problem are only poor approximations of the underlying continuous problem.

Next let us compare the convergence behavior of the Lanczos(A) and the Lanczos(A^{-1}) processes. The decisive term controlling convergence in the classical KPS proofs reads (i.e. the pendant of (23))

$$\min_{p \in \mathbb{P}_{m-1}} \max_{j=2, \dots, n} \left(\frac{p(\lambda_j)}{p(\lambda_1)} \right)^2 \leq \frac{1}{(T_{m-1}(\lambda_1; \lambda_2, \lambda_n))^2}$$

with the shifted Chebyshev polynomial $T_{m-1}(\lambda; a, b) = T_{m-1}((2\lambda - a - b)/(b - a))$. It can be bounded as follows

$$\frac{1}{(T_{m-1}(\lambda_1; \lambda_2, \lambda_n))^2} < 4 \left(e^{-4\sqrt{\tau}} \right)^{m-1} \quad (25)$$

with

$$\tau = \frac{\lambda_2 - \lambda_1}{\lambda_n - \lambda_1}.$$

The quantity τ is called the *gap ratio* of λ_2 with respect to λ_1 , λ_n , see [35, Sec. 4], [33].

For the Lanczos(A^{-1}) process one obtains from (23) instead of (25)

$$\frac{1}{\left(T_{m-1}\left(\frac{1}{\lambda_1}; \frac{1}{\lambda_n}, \frac{1}{\lambda_2}\right)\right)^2} < 4 \left(e^{-4\sqrt{\tilde{\tau}}}\right)^{m-1} \quad (26)$$

with the gap ratio

$$\tilde{\tau} = \frac{\frac{1}{\lambda_1} - \frac{1}{\lambda_2}}{\frac{1}{\lambda_1} - \frac{1}{\lambda_n}}.$$

It is instructive to study the gap ratios for mesh eigenproblems assuming the typical behavior $\lambda_n = \mathcal{O}(h^{-2})$ with the discretization parameter h . First the Lanczos(A) process shows a grid-dependent convergence rate with

$$\lim_{\lambda_n \rightarrow \infty} \tau = 0 \quad \text{and} \quad \tau = \mathcal{O}(h^2).$$

In contrast to this Lanczos(A^{-1}) owns a *grid-independent* upper estimate on the convergence rate (the dependence on λ_n vanishes) as

$$\lim_{\lambda_n \rightarrow \infty} \tilde{\tau} = 1 - \frac{\lambda_1}{\lambda_2} \quad \text{and} \quad \tilde{\tau} = 1 - \mathcal{O}(1).$$

This very different behavior provides a justification for the preconditioning as used in Algorithm 1. There the asymptotic behavior for $B \rightarrow A$ is that of the Lanczos(A^{-1}) process and grid-independent convergence is attained (not only in the limit $B = A$).

3.5. Numerical experiments

The test problem is $n = 10^6$ dimensional with the eigenvalues $\lambda_{l,m} = l^2 + m^2$, $l, m = 1, 2, \dots$, of the (continuous) Laplacian $-\Delta$ on $[0, \pi]^2$. In Figure 1 the convergence history both for the eigenvector and eigenvalue approximations is shown for INVIT(k), $k = 1, \dots, 6$. Moreover, the approximations from the full Krylov subspace $\mathcal{K}_j(A^{-1}, x)$ are plotted versus the iteration index j . The error of the eigenvector approximations is displayed as $\tan \varphi(e_1, V)$ where V is the current approximating subspace. Further the eigenvalue error $\theta_1 - \lambda_1$ is plotted where θ_1 in each step is the smallest Ritz value. For $k \geq 3$, at least for $k \geq 4$, the truncated Krylov subspace iteration INVIT(k) shows a convergence which is very close to the optimal convergence in the full Krylov subspace scheme in $\mathcal{K}_j(A^{-1}, x)$. The slope of the KPS estimates (semilogarithmic plot) is determined by the Chebyshev polynomial, see (26). The slope gained analytically is close to the findings by the numerical experiments. However, the analytical bounds are not very sharp; for the eigenvalue approximations this is an effect of the disturbing factor $\lambda_n - \lambda_1$ in (22).

For the eigenvector approximations averaged convergence factors concerning the convergence of $\tan \varphi(e_1, \mathcal{K}_m(A^{-1}, x))$ have been computed for 2000 random initial vectors (averaged stepwise convergence factors between iterations 3 up to 8). The factors are as follows:

Method	INVIT(1)	INVIT(2)	INVIT(3)	INVIT(4)	INVIT(5)	INVIT(6)	KPS
Conv. factor	0.3875	0.1712	0.1162	0.0861	0.0828	0.0825	0.1270

For INVIT(k), $k \geq 4$, no significant acceleration can be observed. Numerically the limit convergence factor (averaged factor for the same 2000 random initial vectors) of the non-truncated iteration, i.e. all previous iterates form the subspace to which the Rayleigh-Ritz procedure is applied) is about 0.081.

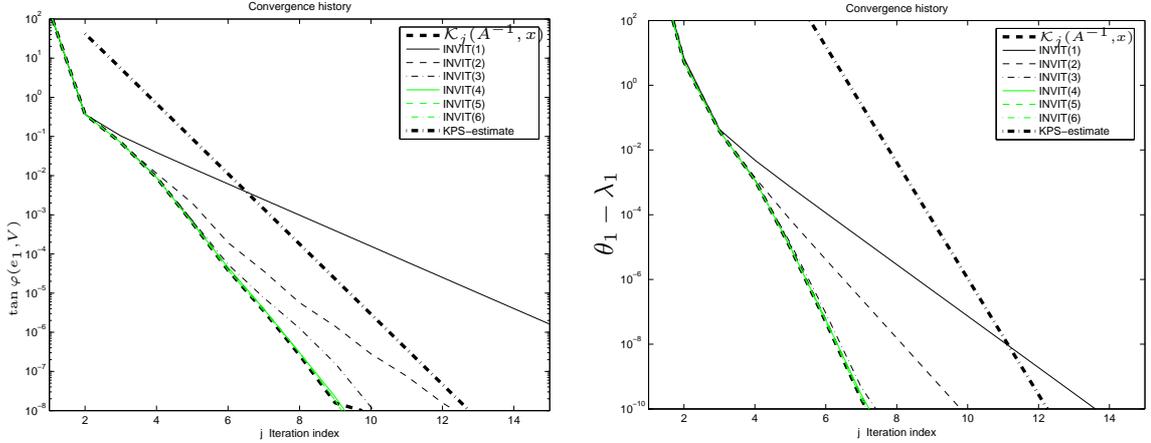


Figure 1. Eigenvector and eigenvalue approximations in $\mathcal{K}_j(A^{-1}, x)$ and $\text{INVIT}(k)$ approximations.

4. Upper and lower convergence estimates for poor preconditioners

The limit $B = A$ of exact inverse preconditioning has been treated in the last section. However, optimality of a preconditioner for a *linear system* does not imply optimality for the *eigenvalue problem*. For linear systems $Ax = b$ exact inverse preconditioning is optimal. It yields the solution $x = A^{-1}b$ within a single step if the preconditioning is applied to the (sometimes so-called) *simple iteration*

$$x_{j+1} = x_j + B^{-1}(b - Ax_j) \Big|_{B=A} = A^{-1}b.$$

Contrastingly, for the eigenvalue problem the inverse matrix A^{-1} is a fairly good, but not the best possible preconditioner. To see this, substitute $B = A$ (and $M = I$) in (1) which amounts to inverse iteration, i.e.,

$$x_{j+1} = \rho(x_j)A^{-1}x_j.$$

However, preconditioners $B \neq A$ exist, which give much more accurate eigenvalue/eigenvector approximations compared to exact inverse preconditioning. See [20] for a systematic analysis.

To summarize, preconditioners which are “poor” for the solution of linear systems can potentially be very suitable for the eigenvalue problem. Those preconditioners for which the control parameter γ by (10) is close to 1 are considered as poor (linear systems) preconditioners as they represent poor approximations to the inverse A^{-1} . However, these preconditioners have the potential of leading to rapid (even one-step) convergence for an iterative eigensolver. This behavior is summarized in Table II.

For the most simple (1)-scheme poorest convergence is estimated by (Thm. 1 in [15])

$$\frac{\rho(x_{j+1}) - \lambda_k}{\lambda_{k+1} - \rho(x_{j+1})} \leq (q(\gamma, \lambda_k, \lambda_{k+1}))^2 \frac{\rho(x_j) - \lambda_k}{\lambda_{k+1} - \rho(x_j)}$$

with the convergence factor

$$q(\gamma, \lambda_k, \lambda_{k+1}) = \gamma + (1 - \gamma) \frac{\lambda_k}{\lambda_{k+1}}. \quad (27)$$

The limit $\gamma \rightarrow 0$ with $q(0, \lambda_k, \lambda_{k+1}) = \lambda_k/\lambda_{k+1}$ (being the rate of convergence of non-shifted inverse iteration) is the topic of Section 3.

	Linear system	Eigenproblem
$\gamma = 0$	One-step convergence	Linear convergence (Inverse Iteration)
$\gamma \rightarrow 1$	Stationarity	1. One-step convergence 2. Stationarity

Table II. Extremal convergence for preconditioners B with $\|I - B^{-1}A\|_A \leq \gamma$ for linear systems and eigenproblems.

Taking the limit $\gamma \rightarrow 1$ in (27) yields

$$\lim_{\gamma \rightarrow 1} q(\gamma, \lambda_k, \lambda_{k+1}) = 1$$

which suggests that stationarity can occur as a case of poorest convergence. And in fact, for each $x \in \mathbb{R}^n \setminus \{0\}$ and any sequence (γ_i) with $\gamma_i \rightarrow 1$ a sequence of symmetric and positive definite preconditioners (B_i) , $\|I - B_i^{-1}A\|_A \leq \gamma_i$, exists so that these preconditioners generate a sequence of iterates

$$x_i := x - B_i^{-1}(Ax - \rho(x)x) = (I - B_i^{-1}(A - \rho(x)I))x$$

which converges to x . Thus stationarity can be attained in the limit $\gamma \rightarrow 1$. In the following the analysis of the fastest and slowest possible convergence in the limit $\gamma \rightarrow 1$ is extended to the general (k) -scheme for $k > 1$.

4.1. The set of admissible preconditioners

For $\gamma \in [0, 1)$ let

$$\mathcal{B}_\gamma = \{B^{-1} \in \mathbb{R}^{n \times n}; B \text{ symmetric positive definite, } \|I - B^{-1}A\|_A \leq \gamma\},$$

be the set of admissible preconditioners containing all symmetric and positive definite preconditioners which satisfy the quality constraint (10). For analytical purposes it can be more convenient to work with the whole set of admissible preconditioners instead of using the constraint (10) only:

Lemma 4.1. *Let $x \in \mathbb{R}^n$, $x \neq 0$, and let*

$$B_\gamma(x) := \{\rho(x)A^{-1}x + y; y \in \mathbb{R}^n, \|y\|_A \leq \gamma\|(I - \rho(x)A^{-1})x\|_A\}, \quad (28)$$

which is a ball with respect to the norm induced by A with the center $\bar{x} = \rho(x)A^{-1}x$. Then the mapping

$$E_x : \mathcal{B}_\gamma \rightarrow B_\gamma(x) : B^{-1} \mapsto x' = x - B^{-1}(Ax - \rho(x)x)$$

is a surjection.

The proof is given by Lemma 2.2 and Lemma 2.3 in [18].

The size of \mathcal{B}_γ is controlled by the parameter $\gamma \in [0, 1)$. The smallest set is $\mathcal{B}_0 = \{A^{-1}\}$. For $0 < \gamma < 1$ the compact ball $\mathcal{B}_\gamma \subset \mathbb{R}^{n \times n}$ contains preconditioners allowing either faster or slower convergence of the eigenvalue solver compared to $\gamma = 0$. The limit set \mathcal{B}_1 is not a closed set, as for the sequence $(\frac{1}{i}A^{-1})_{i \in \mathbb{N}}$

$$\|I - \frac{1}{i}A^{-1}A\|_A = 1 - \frac{1}{i} < 1, \quad i \in \mathbb{N},$$

but $\lim_{i \rightarrow \infty} (1/i)A^{-1}$ is the singular null matrix. Therefore extremal convergence for $\gamma \rightarrow 1$ cannot be analyzed by means of the limit set \mathcal{B}_1 . Instead we work with the interior of \mathcal{B}_1 .

4.2. Fastest possible convergence for $\gamma \rightarrow 1$

Here the question is as follows: does the set \mathcal{B}_γ of admissible preconditioners contain preconditioners which force the (k) -scheme to converge in a single step to an eigenvector of A ?

For the (1)-scheme the situation has been analyzed in [20], Lemma 3. For the (k) -schemes, $k \geq 2$ the Rayleigh-Ritz procedure guarantees one-step convergence if $e_1 \in \mathcal{S}_j^{(k)}$; see (11). Lemma 4.2 shows that one-step convergence to the eigenpair (e_1, λ_1) can occur if γ is sufficiently large.

Lemma 4.2. *Let $x \in \mathbb{R}^n$ with $\|x\| = 1$ be an iterate of the (k) -scheme, $k \geq 2$. If $(e_1, x) \neq 0$, then for all $\gamma \geq \tilde{\gamma}$ with*

$$\tilde{\gamma} = \frac{\min_{(\alpha, \beta) \in \mathbb{R}^2} \|\alpha e_1 + \beta x - \rho(x)A^{-1}x\|_A}{\|x - \rho(x)A^{-1}x\|_A} < 1 \quad (29)$$

the set \mathcal{B}_γ contains a preconditioner which results in one-step convergence to the eigenvector e_1 . The minimum in (29) is attained in

$$\alpha = \frac{\left(\frac{\rho(x)}{\lambda_1} - 1\right)(e_1, x)}{1 - \frac{\lambda_1}{\rho(x)}(e_1, x)^2}, \quad \beta = \frac{1 - (e_1, x)^2}{1 - \frac{\lambda_1}{\rho(x)}(e_1, x)^2}. \quad (30)$$

Proof. If $\alpha e_1 + \beta x$ is contained in the ball $B_\gamma(x)$ given by (28), then a preconditioner B^{-1} exists so that $(1 - \beta)x - B^{-1}(Ax - \rho(x)x) = \alpha e_1$. Hence the Rayleigh-Ritz procedure applied to $\mathcal{S}_j^{(k)} \supseteq \{x, B^{-1}(Ax - \rho(x)x)\}$, for $k \geq 2$, yields the eigenpair (e_1, λ_1) .

The distance of $\alpha e_1 + \beta x$ to the center $\rho(x)A^{-1}x$ of $B_\gamma(x)$ is $R(\alpha, \beta) := \|\alpha e_1 + \beta x - \rho(x)A^{-1}x\|_A$. Since $(e_1, x) \neq 0$, we have

$$\min_{(\alpha, \beta) \in \mathbb{R}^2} \|\alpha e_1 + \beta x - \rho(x)A^{-1}x\|_A < \|x - \rho(x)A^{-1}x\|_A$$

so that $\tilde{\gamma}$, which is the ratio of this smallest distance and the (maximal) radius $\|x - \rho(x)A^{-1}x\|_A$ of $B_1(x)$, is less than 1.

A vanishing gradient $\nabla_{(\alpha, \beta)} R(\alpha, \beta)$ results in a linear system for α, β

$$\alpha \lambda_1 + (\beta \lambda_1 - \rho(x))(e_1, x) = 0, \quad \beta \rho(x) + \alpha \lambda_1 (e_1, x) - \rho(x) = 0.$$

Its solution (30) minimizes $R(\alpha, \beta)$. \square

The critical quantity $\tilde{\gamma}$ is not at all close to 1. For the test problem from Section 3.5 the averaged value of $\tilde{\gamma}$ for 10^3 random vectors x with $\rho(x) < \lambda_{1,2} = 5$ is about 0.19 (for random vectors x with $\rho(x) < \lambda_{1,3} = 10$ the mean value of $\tilde{\gamma}$ is about 0.24).

4.3. Poorest convergence for $\gamma \rightarrow 1$

Next we show that an initial sequence for Algorithm 1 can be constructed (in the case of flexible preconditioning, i.e., the preconditioner may change from step to step) in a way that the Rayleigh-Ritz procedure cannot realize a decrease of the Rayleigh quotient in the limit $\gamma \rightarrow 1$.

Theorem 4.3. *Let $\rho \in \mathbb{R}$ with $\lambda_1 < \rho \leq \lambda_{n-k+1}$. Then a vector $x \in \mathbb{R}^n$ with $\rho = \rho(x)$ can be constructed so that for sequences of preconditioners from \mathcal{B}_γ with $\gamma \rightarrow 1$ the (k) -scheme with flexible preconditioning may attain stationarity in the limit $\gamma \rightarrow 1$. If $k \geq 3$, then the initial sequence spanning $\mathcal{S}_j^{(k)}$ is taken from iterates of the basic (1)-scheme applied to x .*

Proof. We first assume $\rho \in (\lambda_1, \lambda_2)$ and consider $x = (\xi_1, \xi_2, 0, \dots, 0)^T \in \mathbb{R}^n$ with

$$\xi_1 = \left(\frac{\lambda_2 - \rho}{\lambda_2 - \lambda_1}\right)^{1/2}, \quad \xi_2 = \left(\frac{\rho - \lambda_1}{\lambda_2 - \lambda_1}\right)^{1/2}.$$

Then $\|x\| = 1$ and $(x, Ax) = \rho$. Next let $k = 2$. For $\epsilon \in [0, 2(1 - \rho/\lambda_2)/3]$ we define

$$x' := (\xi_1, (1 - 2\epsilon)\xi_2, -\left(\frac{\lambda_2}{\lambda_3}\epsilon(2(1 - \frac{\rho}{\lambda_2}) - \epsilon)\right)^{1/2} \xi_2, 0, \dots, 0)^T. \quad (31)$$

This shows that

$$\|x - \rho A^{-1}x\|_A^2 - \|x' - \rho A^{-1}x\|_A^2 = 2(x', x)\rho - \rho - \|x'\|_A^2 = \epsilon \frac{(\rho - \lambda_1)(2(\lambda_2 - \rho) - 3\epsilon\lambda_2)}{\lambda_2 - \lambda_1} > 0.$$

Thus $x' \in \cup_{\gamma \in [0,1)} B_\gamma(x)$. Hence Lemma 4.1 guarantees the existence of a preconditioner \bar{B}^{-1} in the interior of B_1 so that $x' = x - \bar{B}^{-1}(Ax - \rho(x)x)$.

Next the Rayleigh-Ritz procedure is applied to the two-dimensional subspace $\mathcal{S}^{(2)}$ spanned by x and $x' - x$. The direction of correction

$$x' - x = (0, -2\epsilon\xi_2, -\left(\frac{\lambda_2}{\lambda_3}\epsilon(2(1 - \frac{\rho}{\lambda_2}) - \epsilon)\right)^{1/2} \xi_2, 0, \dots, 0)^T$$

is collinear to

$$d^{(\epsilon)} = (0, 2\sqrt{\epsilon}\xi_2, \left(\frac{\lambda_2}{\lambda_3}(2(1 - \frac{\rho}{\lambda_2}) - \epsilon)\right)^{1/2} \xi_2, 0, \dots, 0)^T$$

and its limit is

$$\lim_{\epsilon \rightarrow 0} d^{(\epsilon)} = d^{(0)} = Ce_3 \quad C \neq 0$$

with e_3 being the eigenvector corresponding to λ_3 , see (8). As $(d^{(0)}, x) = 0$ the Rayleigh-Ritz projections $V^T AV$, $V^T V$ with $V = [x, d^{(0)}]$ are diagonal matrices. The Ritz values are $\rho(x)$ and $\lambda_3 > \rho(x)$. Continuous dependence of the Rayleigh-Ritz approximations (in the case of simple Ritz values) on ϵ proves the stationarity for $\epsilon \rightarrow 0$ (or $\gamma \rightarrow 1$).

The case $k > 2$ is treated similarly. Starting from $x = (\xi_1, \xi_2, 0, \dots, 0)^T \in \mathbb{R}^n$ one first takes

$$x^{(1)} := x' = (\xi_1, \xi_2[\epsilon^{(1)}], \xi_3[\epsilon^{(1)}], 0, \dots, 0)^T$$

where the components $\xi_2[\epsilon^{(1)}]$, $\xi_3[\epsilon^{(1)}]$ are the second and third component in (31). If $\epsilon^{(2)}$ is sufficiently small, then the component construction underlying (31) can be applied to the third and fourth component resulting in

$$x^{(2)} = (\xi_1, \xi_2[\epsilon^{(1)}], \xi_3[\epsilon^{(1)}, \epsilon^{(2)}], \xi_4[\epsilon^{(1)}, \epsilon^{(2)}], 0, \dots, 0)^T$$

are computed by a similar construction as used in (31). Then $x^{(2)} \in \cup_{\gamma \in [0,1)} B_\gamma(x)$ and

$$\|x^{(2)} - \rho A^{-1}x\|_A^2 < \|x^{(1)} - \rho A^{-1}x\|_A^2.$$

One can extend this construction up to $\xi_k[\epsilon^{(1)}, \dots, \epsilon^{(k-1)}]$. All vectors $x, x^{(1)}, \dots, x^{(k-1)}$ are in the interior of $B_1(x)$ and can be constructed by consecutive steps of the (1)-scheme using variable preconditioners. Taking the limits $\epsilon^{(k-1)} \rightarrow 0, \dots, \epsilon^{(1)} \rightarrow 0$ shows that the limit subspace is

$$\text{span}\{x, e_3, \dots, e_{k+1}\}.$$

The Rayleigh-Ritz procedure provides Ritz values converging to $\rho, \lambda_3, \dots, \lambda_{k+1}$, which proves stationarity.

If $\lambda_j \leq \rho < \lambda_{j+1} \leq \lambda_{n-k+1}$, then the starting point is a vector $x = (0, \dots, 0, \xi_j, \xi_{j+1}, 0, \dots, 0)^T$ with $\rho(x) = \rho$. Because of $\lambda_{j+1} \leq \lambda_{n-k+1}$ at least $k-1$ zero components ξ_{j+2}, \dots, ξ_n are available to pursue the construction outlined above. \square

Remark 4.4. The assumption $\rho(x) \leq \lambda_{n-k+1}$ made in Thm. 4.3 cannot be skipped as otherwise the Courant-Fischer principle would enforce non-stationarity. To see this, assume $\rho(x) > \lambda_{n-k+1}$ and a k -dimensional iteration subspace $\mathcal{S}_{k-1}^{(k)}$. Then

$$\min_{z \in \mathcal{S}_{k-1}^{(k)} \setminus \{0\}} \rho(x) \leq \max_{V, \dim V=k} \min_{x \in V \setminus \{0\}} \rho(x) = \lambda_{n-k+1}$$

where the maximum is taken over all k -dimensional subspaces.

5. Conclusion

A link has been presented from the (partially not very well understood) preconditioned gradient type eigensolvers to the (well understood) Invert-Lanczos process. The joining element is the limit of preconditioning with the exact inverse of the system matrix. The analysis of the Invert-Lanczos process is instructive in order to understand the convergence behavior of such practically important preconditioned eigensolvers like the Locally Optimal Preconditioned Conjugate Gradients (LOPCG) scheme. The application of the inverse system matrix for preconditioning purposes is usually impossible and/or too expensive. However, the action of accurate preconditioners can be approximated by multiple steps of a preconditioned linear solver which, for example, can be based on a simple V-cycle in the context of a mesh discretizations of a partial differential operator.

By using standard techniques for the analysis of the Lanczos process (estimates using Chebyshev polynomials) upper and lower convergence estimates are accessible for this limit case. It has been shown that standard assumptions on the quality of the preconditioner, i.e. assumptions which are made for linear systems solvers, allow, on the one hand, extremely fast convergence of the preconditioned eigensolver and, on the other hand, very poor convergence up to stationarity.

REFERENCES

1. P. Arbenz, U.L. Hetmaniuk, R.B. Lehoucq, and R.S. Tuminaro. A comparison of eigensolvers for large-scale 3D modal analysis using AMG-preconditioned iterative methods. *Int. J. Numer. Meth. Engng.*, 64(2):204–236, 2005.
2. Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the solution of algebraic eigenvalue problems: A practical guide*. SIAM, Philadelphia, 2000.
3. M. Bollhöfer and Y. Saad. Multilevel preconditioners constructed from inverse-based ILUs. *SIAM J. Sci. Comput.*, 27(5):1627–1650, 2006.
4. A. Borzi and G. Borzi. Algebraic multigrid methods for solving generalized eigenvalue problems. *Int. J. Numer. Meth. Engng.*, 65(8):1186–1196, 2006.
5. D. Calvetti, L. Reichel, and D.C. Sorensen. An implicit restarted Lanczos method for large symmetric eigenvalue problems. *ETNA, Electron. Trans. Numer. Anal.*, 2:1–21, 1994.
6. E.G. D'yakonov. *Optimization in solving elliptic problems*. CRC Press, Boca Raton, Florida, 1996.
7. T. Ericsson and A. Ruhe. The spectral transformation Lanczos method for the numerical solution of large sparse generalized symmetric eigenvalue problems. *Math. Comput.*, 35:1251–1268, 1980.
8. M.R. Hestenes and W. Karush. A method of gradients for the calculation of the characteristic roots and vectors of a real symmetric matrix. *J. Res. Nat. Bureau Standards*, 47:45–61, 1951.
9. S. Kaniel. Estimates for some computational techniques in linear algebra. *Math. Comp.*, 20:369–378, 1966.
10. L. V. Kantorovich. *Functional analysis and applied mathematics*. U. S. Department of Commerce National Bureau of Standards, Los Angeles, Calif., 1952. Translated by C. D. Benster.
11. L.V. Kantorovich and G.P. Akilov. *Functional analysis in normed spaces*. Pergamon, London, 1964.
12. A.V. Knyazev. Convergence rate estimates for iterative methods for a mesh symmetric eigenvalue problem. *Russian J. Numer. Anal. Math. Modelling*, 2:371–396, 1987.
13. A.V. Knyazev. A preconditioned conjugate gradient method for eigenvalue problems and its implementation in a subspace. In *International Ser. Numerical Mathematics, 96, Eigenwertaufgaben in Natur- und Ingenieurwissenschaften und ihre numerische Behandlung, Oberwolfach.*, pages 143–154, Basel, 1991. Birkhäuser.

14. A.V. Knyazev. Preconditioned eigensolvers—an oxymoron? *Electron. Trans. Numer. Anal.*, 7:104–123, 1998.
15. A.V. Knyazev and K. Neymeyr. A geometric theory for preconditioned inverse iteration. III: A short and sharp convergence estimate for generalized eigenvalue problems. *Linear Algebra Appl.*, 358:95–114, 2003.
16. A.V. Knyazev and K. Neymeyr. Efficient solution of symmetric eigenvalue problems using multigrid preconditioners in the locally optimal block conjugate gradient method. *Electron. Trans. Numer. Anal.*, 15:38–55, 2003.
17. A.V. Knyazev and A.L. Skorokhodov. On exact estimates of the convergence rate of the steepest ascent method in the symmetric eigenvalue problem. *Linear Algebra Appl.*, 154–156:245–257, 1991.
18. K. Neymeyr. A geometric theory for preconditioned inverse iteration. I: Extrema of the Rayleigh quotient. *Linear Algebra Appl.*, 322:61–85, 2001.
19. K. Neymeyr. A hierarchy of preconditioned eigensolvers for elliptic differential operators. Habilitation thesis, Universität Tübingen, Germany, 2001.
20. K. Neymeyr. A geometric theory for preconditioned inverse iteration IV: On the fastest convergence cases. *Linear Algebra Appl.*, 415(1):114–139, 2006.
21. M.G. Neytcheva and P.S. Vassilevski. Preconditioning of indefinite and almost singular finite element elliptic equations. *SIAM J. Sci. Comput.*, 19(5):1471–1485, 1998.
22. Y. Notay. Is jacobi–davidson faster than davidson? *SIAM Journal on Matrix Analysis and Applications*, 26(2):522–543, 2004.
23. E. Ovtchinnikov. Convergence estimates for the generalized davidson method for symmetric eigenvalue problems i: The preconditioning aspect. *SIAM Journal on Numerical Analysis*, 41(1):258–271, 2003.
24. E. Ovtchinnikov. Convergence estimates for the generalized davidson method for symmetric eigenvalue problems ii: The subspace acceleration. *SIAM Journal on Numerical Analysis*, 41(1):272–286, 2003.
25. C.C. Paige. *The computation of eigenvalues and eigenvectors of very large sparse matrices*. PhD thesis, London University, Institute of computer science, 1971.
26. B.N. Parlett. *The symmetric eigenvalue problem*. Prentice Hall, Englewood Cliffs New Jersey, 1980.
27. W.V. Petryshyn. On the eigenvalue problem $Tu - \lambda Su = 0$ with unbounded and non-symmetric operators T and S . *Philos. Trans. Roy. Soc. Math. Phys. Sci.*, 262:413–458, 1968.
28. V.G. Prikazchikov. Strict estimates of the rate of convergence of an iterative method of computing eigenvalues. *USSR J. Comput. Math. and Math. Physics*, 15:235–239, 1975.
29. Axel Ruhe. Rational Krylov sequence methods for eigenvalue computation. *Linear Algebra Appl.*, 58:391–405, 1984.
30. Y. Saad. On the rates of convergence of the Lanczos and the block-Lanczos methods. *SIAM J. Numer. Anal.*, 17(5):687–706, 1980.
31. Y. Saad. *Numerical methods for large eigenvalue problems*. Manchester University Press, Manchester, 1992.
32. B.A. Samokish. The steepest descent method for an eigenvalue problem with semi-bounded operators. *Izv. Vyssh. Uchebn. Zaved. Mat.*, 5:105–114, 1958. (In Russian).
33. G.L.G. Sleijpen and A. van der Sluis. Further results on the convergence behavior of conjugate-gradients and Ritz values. *Linear Algebra Appl.*, 246:233–278, 1996.
34. A. Stathopoulos. Nearly optimal preconditioned methods for hermitian eigenproblems under limited memory. part i: Seeking one eigenvalue. *SIAM Journal on Scientific Computing*, 29(2):481–514, 2007.
35. A. van der Sluis. The convergence behaviour of conjugate gradients and Ritz values in various circumstances. Beauwens, R. (ed.) et al., Iterative methods in linear algebra. Proceedings of the IMACS international symposium, Brussels, Belgium, 2-4 April, 1991. Amsterdam: North-Holland. 49-66 (1992)., 1992.
36. P.S. Vassilevski. Preconditioning nonsymmetric and indefinite finite element matrices. *Numer. Linear Algebra Appl.*, 1:59–76, 1992.
37. K. Wu and H. Simon. Thick-restart Lanczos method for large symmetric eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 22(2):602–616, 2000.
38. H. Yserentant. Preconditioning indefinite discretization matrices. *Numer. Math.*, 54(6):719–734, 1989.
39. P. F. Zhuk and L. N. Bondarenko. Sharp estimates for the rate of convergence of the s -step method of steepest descent in eigenvalue problems. *Ukrain. Mat. Zh.*, 49(12):1694–1699, 1997.