A Note on Inverse Iteration

Klaus Neymeyr

Universität Rostock, Fachbereich Mathematik, Universitätsplatz 1, 18051 Rostock, Germany;

SUMMARY

Inverse iteration, if applied to a symmetric positive definite matrix, is shown to generate a sequence of iterates with monotonously decreasing Rayleigh quotients. We present sharp bounds from above and from below which highlight inverse iteration as a descent scheme for the Rayleigh quotient. Such estimates provide the background for the analysis of the behavior of the Rayleigh quotient in certain approximate variants of inverse iteration.

KEY WORDS: Symmetric eigenvalue problem; Inverse iteration; Rayleigh quotient.

1. Introduction

Inverse iteration is a well-known iterative procedure to compute approximations of eigenfunctions and eigenvalues of linear operators. It was introduced by Wielandt in 1944 in a sequence of five papers, see [1], to treat the matrix eigenvalue problem

$$Ax_i = \lambda_i x_i$$

for a real or complex square matrix A. The scalar λ_i is the ith eigenvalue and the vector x_i denotes a corresponding eigenvector. Given a nonzero starting vector $x^{(0)}$, inverse iteration generates a sequence of iterates $x^{(k)}$ by solving the linear systems

$$(A - \sigma I)x^{(k+1)} = x^{(k)}, \qquad k = 0, 1, 2, \dots, \tag{1}$$

where σ denotes an eigenvalue approximation and I is the identity matrix. In practice, the iterates are normalized after each step. If A is a symmetric matrix, then the iterates $x^{(k)}$ converge to an eigenvector associated with an eigenvalue closest to σ if the starting vector $x^{(0)}$ is not perpendicular to that vector. For non-symmetric matrices the issue of starting vectors is discussed in Sec. 2.6 in [2]. Elementary results on the convergence theory of inverse iteration and of the complementary power method are contained in many monographs on numerical linear algebra, see e.g. Parlett [3], Chatelin [4] or Golub and van Loan [5]. The history of inverse iteration and new results on its convergence have been presented by Ipsen [2, 6].

Convergence of inverse iteration toward an eigenvector can be estimated in terms of the Rayleigh quotients of the iterates. The Rayleigh quotient of a vector x is given by

$$\lambda(x) = \frac{(x, Ax)}{(x, x)}, \tag{2}$$

where $(\cdot,\cdot)$ denotes the Euclidean product. The eigenvectors are the stationary points of $\lambda(\cdot)$ and its absolute extrema are the extremal eigenvalues of A.

2. Convergence estimates for Inverse Iteration

The purpose of this paper is to derive *sharp convergence estimates* for the Rayleigh quotient in the case of inverse iteration being restricted to a symmetric positive definite matrix A. This restrictive assumption is typically fulfilled for an important class of (extremely) large eigenproblems, i.e., discretizations of certain elliptic partial differential operators; see below. These convergence estimates show that inverse iteration for a symmetric positive definite matrix and under a certain assumption on the shift parameter σ is a *descent scheme for the Rayleigh quotient.*

Why is it worthwhile to understand inverse iteration in such a way? Let us first make the point that usually, the convergence theory of inverse iteration is founded on an eigenvector expansion of the initial vector, i.e., applying $(A-\sigma I)^{-1}$ to the actual iteration vector results in a relative amplification of the eigenvector components corresponding to eigenvalues close to σ [3]. Such a convergence analysis does not exploit any properties of the Rayleigh quotient. But there is a different way to look at inverse iteration which is initiated by the demand that today, one is faced with the problem to solve extremely large eigenproblems in which the dimension of A exceeds, say, 10^6 up to 10^9. Such matrix eigenproblems appear for instance as mesh discretizations of self-adjoint, elliptic partial differential operators. Typically, only a few of the smallest eigenvalues together with the eigenvalues are to be computed. Inverse iteration can be applied. Due to several reasons, the associated linear system of equations given by (1) can only be solved approximately by using an approximate inverse (or preconditioner) of the system matrix. See e.g. [7, 10–14].

For these approximate versions of inverse iteration (called "inexact inverse iteration" or "preconditioned inverse iteration") any convergence theory built on an eigenvector expansion of the initial vector breaks down because an approximate solution of (1) may weed out certain eigenvector components and may amplify others in a complex and hardly controllable way. Nevertheless, as it turned out in the convergence analysis of these methods, the *Rayleigh quotient* can serve as a *robust convergence measure* since one can prove its stepwise *monotonous decrease* [10, 13].

This behavior of the Rayleigh quotient motivates a more detailed investigation of inverse iteration, i.e., for an exact solution of (1). The results are summarized in this paper and highlight an interesting property of inverse iteration. Theorem 2.1 provides *sharp* bounds from above and below for the decrease of the Rayleigh quotients of the iterates of (1). The technique of proof is rather unusual: the Lagrange multiplier method is applied to determine constrained extrema of the Rayleigh quotient with respect to its level sets. By doing so we first obtain the justification to restrict the analysis to two-dimensional A-invariant subspaces. In a second step we derive the convergence estimates by means of a *mini-dimensional* (2D) analysis.

The convergence of the Rayleigh quotients is measured in terms of the ratios

$$\Delta_{i,i+1}(\lambda) := \frac{\lambda-\lambda_i}{\lambda_{i+1}-\lambda} \quad \text{and} \quad \Delta_{1,n}(\lambda) := \frac{\lambda-\lambda_1}{\lambda_n-\lambda}. \tag{3}$$

The eigenvalues of A with arbitrary multiplicity are indexed in ascending order, i.e. $0 < \lambda_1 <$

$\lambda_2 < \ldots < \lambda_n$. A small ratio $\Delta_{i,i+1}(\lambda)$, for instance $0 \le \Delta_{i,i+1}(\lambda) \le \epsilon$ with $\epsilon > 0$, is an absolute measure for the closeness of λ to the eigenvalue λ_i, as then $\lambda \le (\lambda_i + \epsilon\lambda_{i+1})/(1+\epsilon) = \lambda_i + \mathcal{O}(\epsilon)$.

In Theorem 2.1 we also need the *convergence factors*

$$\rho_{i,i+1} = \frac{\lambda_i - \sigma}{\lambda_{i+1} - \sigma} \quad \text{and} \quad \rho_{1,n} = \frac{\lambda_1 - \sigma}{\lambda_n - \sigma}, \tag{4}$$

which are less than 1 under our assumptions. See also [8, 9] for comparable estimates based on the quantities (3) and (4) and compare with the results on inexact inverse iteration gained in [10, 13] which in the limit of exact solution result in (6). Here our main intention is to present a condensed form of a convergence theory for inverse iteration which is based on the *Lagrange multiplier* technique.

Theorem 2.1. *Let $A \in \mathbb{R}^{s\times s}$ be a symmetric positive definite matrix with the eigenvalues $0 < \lambda_1 < \ldots < \lambda_n$; the multiplicity of λ_i is denoted by m_i so that $m_1 + \ldots + m_n = s$.*

For any real number $\lambda \in (\lambda_i, \lambda_{i+1})$ let

$$L(\lambda) = \{x \in \mathbb{R}^s;\ \lambda(x) = \lambda\}, \tag{5}$$

which is a level set *of the Rayleigh quotient. Moreover, assume also the shifted matrix $A - \sigma I$ positive definite, i.e. $\sigma \in [0, \lambda_1)$, cf. Remark 2.3.*

Then for any $x \in L(\lambda)$ the next iterate $x' = (A - \sigma I)^{-1}x$ of (1) with the Rayleigh quotient $\lambda(x') = \lambda\left((A - \sigma I)^{-1}x\right)$ satisfies

$$\Delta_{i,i+1}(\lambda(x')) \le (\rho_{i,i+1})^2 \Delta_{i,i+1}(\lambda). \tag{6}$$

Inequality (6) is an estimate on the poorest convergence *of $\lambda(x')$ toward the closest eigenvalue $\lambda_i < \lambda$ in terms of (3) and the* convergence factor *$\rho_{i,i+1}$ which is defined by (4). The right-hand side of (6) does not depend on the choice of x, but only on λ. Estimate (6) is sharp as it is attained in a certain $x \in L(\lambda)$.*

Moreover, the fastest convergence *is described by a sharp estimate from below*

$$(\rho_{1,n})^2 \Delta_{1,n}(\lambda) \le \Delta_{1,n}(\lambda(x')). \tag{7}$$

Once again, there is a certain $x \in L(\lambda)$ in which the lower bound (7) is attained.

For $\sigma = 0$ the following sharp estimate for $\lambda(x')$ results from (6) and (7)

$$\frac{1}{\lambda_1^{-1} + \lambda_n^{-1} - (\lambda_1 + \lambda_n - \lambda)^{-1}} \le \lambda(x') \le \frac{1}{\lambda_i^{-1} + \lambda_{i+1}^{-1} - (\lambda_i + \lambda_{i+1} - \lambda)^{-1}}. \tag{8}$$

Remark 2.2. *If the initial iterate $x^{(0)}$ satisfies $\lambda(x^{(0)}) < \lambda_2$, then (6) can be applied recursively. This yields*

$$\frac{\Delta_{1,2}(\lambda(x^{(k)}))}{\Delta_{1,2}(\lambda(x^{(0)}))} \le \left(\frac{\lambda_1 - \sigma}{\lambda_2 - \sigma}\right)^{2k}, \qquad k = 1, 2, \ldots,$$

and guarantees convergence of $(x^{(k)}/\|x^{(k)}\|, \lambda(x^{(k)}))$ to an eigenpair (x_1, λ_1).

Note that Theorem 2.1 does not refer to the components of an eigenvector expansion of the initial vector $x^{(0)}$. Consequently, as reflected by the estimate (6), inverse iteration starting with $\lambda(x^{(0)}) \in (\lambda_i, \lambda_{i+1})$ in the case of poorest convergence can only be shown to converge to an eigenpair (x_i, λ_i). See also the remarks above on inexact, or preconditioned, inverse iteration for which, typically, no assumptions on eigenvector expansions of the iteration vectors can be made.

Proof. Let

$$U^T A U = \operatorname{diag}(\underbrace{\lambda_1,\ldots,\lambda_1}_{m_1},\underbrace{\lambda_2,\ldots,\lambda_2}_{m_2},\ldots,\underbrace{\lambda_n,\ldots,\lambda_n}_{m_n}) =: \Lambda \in \mathbb{R}^{s\times s}$$

be a diagonal matrix with λ_i being the ith eigenvalue of A with the multiplicity m_i. Then for any $x \in L(\lambda)$ one obtains $v = U^T x$ as the corresponding coefficient vector with respect to the eigenbasis. It holds $\lambda(x) = (v, \Lambda v)/(v,v) =: \lambda(v)$ and

$$\lambda(x') = \lambda((A-\sigma I)^{-1}x) = \frac{(v, \Lambda(\Lambda-\sigma I)^{-2}v)}{(v,(\Lambda-\sigma I)^{-2}v)} =: \lambda((\Lambda-\sigma I)^{-1}v), \tag{9}$$

where we use the same notation $\lambda(\cdot)$ for the Rayleigh quotient with respect to both bases.

Next we give a justification for restricting the analysis to simple eigenvalues only. Therefore, let $\bar{\Lambda} := \operatorname{diag}(\lambda_1,\ldots,\lambda_n) \in \mathbb{R}^{n\times n}$. For any $v \in \mathbb{R}^s$ with $\lambda = \lambda(v)$ define $\bar{v} \in \mathbb{R}^n$ in such a way that

$$\bar{v}_i = \left(\sum_{l=m+1}^{m+m_i} v_l^2\right)^{1/2} \quad \text{with } m = m_1 + m_2 + \ldots + m_{i-1}, \quad i = 1,\ldots,n, \quad m_0 = 0,$$

i.e., all components of v corresponding to λ_i are condensed into the single component $\bar{v}_i$. Then

$$\bar{\lambda}(\bar{v}) := \frac{(\bar{v},\bar{\Lambda}\bar{v})}{(\bar{v},\bar{v})} = \frac{(v,\Lambda v)}{(v,v)} = \lambda$$

and

$$\bar{\lambda}((\bar{\Lambda}-\sigma I_{n\times n})^{-1}\bar{v}) := \frac{(\bar{v},\bar{\Lambda}(\bar{\Lambda}-\sigma I_{n\times n})^{-2}\bar{v})}{(\bar{v},(\bar{\Lambda}-\sigma I_{n\times n})^{-2}\bar{v})} = \frac{(v,\Lambda(\Lambda-\sigma I_{s\times s})^{-2}v)}{(v,(\Lambda-\sigma I_{s\times s})^{-2}v)} = \lambda(x'),$$

which is a representation of the Rayleigh quotient (9) in terms of the reduced matrix $\bar{\Lambda}$ with only simple eigenvalues. This justifies to assume $m_i = 1$, $i = 1,\ldots,n$ in the following. Thus $s = n$ and $\bar{\Lambda} = \Lambda$.

A necessary condition for (9) being an extremum on the level set $L(\lambda)$ can be derived by means of the Lagrange multiplier method. Let us reformulate the non-quadratic constraint $\lambda(v) = \lambda$ as a quadratic normalization condition, i.e. $(v,v) = 1$, and the quadratic constraint $(v,\Lambda v) = \lambda$. Then we consider the Lagrange function

$$\mathcal{L}(v) = \frac{(v,\Lambda(\Lambda-\sigma I)^{-2}v)}{(v,(\Lambda-\sigma I)^{-2}v)} + \mu\left((v,v)-1\right) + \nu\left((v,\Lambda v)-\lambda\right),$$

with μ and ν being the Lagrange multipliers. Any constrained extremum in v has to satisfy the equation

$$\nabla\mathcal{L}(v) = \frac{2}{(v,(\Lambda-\sigma I)^{-2}v)}(\Lambda-\sigma I)^{-2}\left[\Lambda-\lambda' I\right]v + 2\mu v + 2\nu\Lambda v = 0 \tag{10}$$

with $\lambda' := \lambda((\Lambda-\sigma I)^{-1}v)$. Since v is not an eigenvector (as $\lambda \neq \lambda_i$, $i = 1,\ldots,n$), there are at least two nonzero components v_k and v_l with $k \neq l$. Take k as the smallest index with $v_k \neq 0$ and l as the largest index with $v_l \neq 0$. Then $\lambda_k < \lambda'$. We determine the Lagrange multipliers μ and ν from Equation (10) by solving the linear system

$$\begin{pmatrix} 1 & \lambda_k \\ 1 & \lambda_l \end{pmatrix}\begin{pmatrix} \mu \\ \nu \end{pmatrix} = \frac{1}{(v,(\Lambda-\sigma I)^{-2}v)}\begin{pmatrix} (\lambda'-\lambda_k)(\lambda_k-\sigma)^{-2} \\ (\lambda'-\lambda_l)(\lambda_l-\sigma)^{-2} \end{pmatrix} \tag{11}$$

having a non-vanishing determinant. Its solution reads

$$\mu = \left[\sigma^2\lambda' + 2\sigma(\lambda_k\lambda_l - \lambda_k\lambda' - \lambda_l\lambda') + \lambda_l^2(\lambda' - \lambda_k) + \lambda_k^2(\lambda' - \lambda_l) + \lambda_k\lambda_l\lambda'\right]/C,$$
$$\nu = -\left[\sigma^2 - 2\sigma\lambda' + \lambda'(\lambda_k + \lambda_l) - \lambda_k\lambda_l\right]/C,$$

with $C = (v, (\Lambda - \sigma I)^{-2}v)(\lambda_k - \sigma)^2(\lambda_l - \sigma)^2$. To show that v has exactly two nonzero components, i.e., $v_j = 0$ for $j \neq k, l$, we insert μ and ν in the jth component of (10). We write $(\nabla\mathcal{L}(v))_j = \alpha(\sigma)p(\sigma)v_j$, where $\alpha(\sigma) = (\lambda_l - \lambda_j)(\lambda_k - \lambda_j)/(C(\lambda_j - \sigma)^2) \neq 0$ and

$$p(\sigma) = 2\sigma^3 - \sigma^2(\lambda_k + \lambda_l + \lambda_j + 3\lambda') + 2\sigma\lambda'(\lambda_k + \lambda_l + \lambda_j) + \lambda_j\lambda_k\lambda_l - \lambda'(\lambda_k\lambda_l + \lambda_k\lambda_j + \lambda_l\lambda_j).$$

It remains to be shown that $p(\sigma) \neq 0$, where by assumption $\sigma \in [0, \lambda_1)$. First notice that $0 \leq \sigma < \lambda_1 \leq \lambda_k < \lambda' < \lambda_l$ and $\lambda_k < \lambda_j < \lambda_l$ as well as $\lim_{\sigma\to-\infty} p(\sigma) = -\infty$. Moreover, the local extrema of $p(\sigma)$, i.e., $p'(\sigma) = 0$, are taken in λ' and $(\lambda_k + \lambda_l + \lambda_j)/3$ and are both larger than λ_k. Finally, we conclude with

$$p(\lambda_k) = -(\lambda_j - \lambda_k)(\lambda_l - \lambda_k)(\lambda' - \lambda_k) < 0,$$

that it is impossible for the third order polynomial $p(\sigma)$ to take a zero in $[0, \lambda_1)$.

Thus the further ("mini-dimensional"-) analysis can be restricted to the 2D space spanned by the eigenvectors to λ_k and λ_l. The nonzero components v_k and v_l are determined by $(v, v) = 1$ and $(v, \Lambda v) = \lambda$. We obtain

$$v_k^2 = \frac{\lambda_l - \lambda}{\lambda_l - \lambda_k}, \qquad \text{and} \qquad v_l^2 = \frac{\lambda - \lambda_k}{\lambda_l - \lambda_k}. \tag{12}$$

Inserting (12) in $\lambda' = \lambda((\Lambda - \sigma I)^{-1}v)$ results in

$$\lambda' = \lambda'(\lambda_k, \lambda_l, \lambda, \sigma) = \frac{\sigma^2\lambda - 2\sigma\lambda_k\lambda_l + \lambda_k\lambda_l(\lambda_k + \lambda_l - \lambda)}{\sigma^2 - 2\sigma(\lambda_k + \lambda_l - \lambda) + \lambda_k^2 + \lambda_l^2 - \lambda(\lambda_k + \lambda_l) + \lambda_k\lambda_l} \tag{13}$$

The differentiation of λ' with respect to λ_k and λ_l together with $0 < \sigma < \lambda_1 \leq \lambda_k \leq \lambda_i < \lambda < \lambda_{i+1} \leq \lambda_l \leq \lambda_n$ results in

$$\frac{\partial}{\partial\lambda_k}\lambda'(\lambda_k, \lambda_l, \lambda, \sigma) = \frac{[2(\lambda_k - \sigma) + \lambda_l - \lambda](\lambda_l - \lambda)(\lambda_l - \sigma)^2}{(\sigma^2 + 2\sigma(\lambda - \lambda_k - \lambda_l) + \lambda_k^2 - \lambda_k\lambda + \lambda_k\lambda_l - \lambda_l\lambda + \lambda_l^2)^2} > 0$$
$$\frac{\partial}{\partial\lambda_l}\lambda'(\lambda_k, \lambda_l, \lambda, \sigma) = \frac{[2(\sigma - \lambda_l) + \lambda - \lambda_k](\lambda - \lambda_k)(\lambda_k - \sigma)^2}{(\sigma^2 + 2\sigma(\lambda - \lambda_k - \lambda_l) + \lambda_k^2 - \lambda_k\lambda + \lambda_k\lambda_l - \lambda_l\lambda + \lambda_l^2)^2} < 0$$

Hence $\lambda'(\lambda_k, \lambda_l, \lambda, \sigma)$ takes its maximum in $\lambda'(\lambda_i, \lambda_{i+1}, \lambda, \sigma)$, whereas its minimum is taken in $\lambda'(\lambda_1, \lambda_n, \lambda, \sigma)$, i.e.

$$\lambda'(\lambda_1, \lambda_n, \lambda, \sigma) \leq \lambda((\Lambda - \sigma I)^{-1}v) \leq \lambda'(\lambda_i, \lambda_{i+1}, \lambda, \sigma). \tag{14}$$

Reformulation of (14) using (13) yields

$$\frac{\lambda_1 + \lambda_n R_{1,n}(\lambda)}{1 + R_{1,n}(\lambda)} \leq \lambda' = \lambda((\Lambda - \sigma I)^{-1}v) \leq \frac{\lambda_i + \lambda_{i+1}R_{i,i+1}(\lambda)}{1 + R_{i,i+1}(\lambda)} \tag{15}$$

with

$$R_{i,i+1}(\lambda) = \rho_{i,i+1}^2\Delta_{i,i+1}(\lambda) = \left(\frac{\lambda_i - \sigma}{\lambda_{i+1} - \sigma}\right)^2 \frac{\lambda - \lambda_i}{\lambda_{i+1} - \lambda} \tag{16}$$

and

$$R_{1,n}(\lambda) = \rho_{1,n}^2 \Delta_{1,n}(\lambda) = \left(\frac{\lambda_1 - \sigma}{\lambda_n - \sigma}\right)^2 \frac{\lambda - \lambda_1}{\lambda_n - \lambda}. \tag{17}$$

The right inequality of (15) reads

$$\lambda' + \lambda' \rho_{i,i+1}^2 \Delta_{i,i+1}(\lambda) \leq \lambda_i + \lambda_{i+1} \rho_{i,i+1}^2 \Delta_{i,i+1}(\lambda),$$

from which (6) follows immediately. Reformulation of the left-hand inequality of (15) proves (7) analogously.

For $\sigma = 0$ the inequality (14) simply reads

$$\frac{\lambda_1 \lambda_n (\lambda_1 + \lambda_n - \lambda)}{\lambda_n^2 - (\lambda - \lambda_1)(\lambda_1 + \lambda_n)} \leq \lambda' \leq \frac{\lambda_i \lambda_{i+1} (\lambda_i + \lambda_{i+1} - \lambda)}{\lambda_{i+1}^2 - (\lambda - \lambda_i)(\lambda_i + \lambda_{i+1})}, \tag{18}$$

which proves (8).

The estimates (6) and (7) are derived in the 2D invariant subspaces to either λ_i, λ_{i+1} or λ_1, λ_n and they are attained (by construction) exactly in these invariant subspaces. Therefore, (6) and (7) are each attained in a vector whose components are defined by (12) and whose Rayleigh quotients are given by (13). □

Remark 2.3. *Theorem 2.1 does even hold under the assumption* $\sigma \in [0, \frac{\lambda_1+\lambda_2}{2}) \setminus \{\lambda_1\}$*; but here we avoid additional technicalities in the proof of Theorem 2.1. The choice* $\sigma \in (\lambda_1, \frac{\lambda_1+\lambda_2}{2})$ *covers the case of the Rayleigh quotient iteration converging to* λ_1*.*

By Theorem 2.1 the Rayleigh quotients $\lambda(x^{(k)})$ form a monotonously decreasing sequence which is bounded from below by λ_1. Therefore, the difference of consecutive Rayleigh quotients, i.e., $\lambda(x^{(k)}) - \lambda((A-\sigma I)^{-1}x^{(k)})$, converges to 0. In the next lemma the latter difference is shown to be an upper bound for a certain norm of the residual vectors which proves convergence of $x^{(k)}$ to an eigenvector.

Lemma 2.4. *For* $y \in \mathbb{R}^s$, $y \neq 0$ *let the residual vector be given by* $r(y) = Ay - \lambda(y)y$ *and let* $\|y\|_{A-\sigma I}^2 = (y, (A - \sigma I)y)$*. On the assumptions of Theorem 2.1 for any* $x \in \mathbb{R}^s$ *with* $\|x\|^2 = (x, x) = 1$ *it holds that*

$$\|r((A - \sigma I)^{-1}x)\|_{A-\sigma I}^2 \leq \lambda(x) - \lambda((A - \sigma I)^{-1}x). \tag{19}$$

Proof. For any nonzero $z \in \mathbb{R}^s$ it holds $0 < \sigma < \lambda_1 \leq \lambda(z)$. Multiplication of the last inequality with the (due to the Cauchy inequality non-negative) factor $(A^2z, z) - \lambda(z)(z, Az)$ leads to

$$\sigma \left((A^2z, z) - \lambda(z)(z, Az)\right) \leq \lambda(z) \left((A^2z, z) - \lambda(z)(z, Az)\right).$$

The latter inequality is equivalent to

$$\frac{\|r(z)\|_{A-\sigma I}^2}{\|(A - \sigma I)z\|^2} \leq \lambda((A - \sigma I)z) - \lambda(z) \tag{20}$$

which can be verified by writing the norms and Rayleigh quotients in (20) in terms of $(z, A^k z)$, $k = 0, 1, 2, 3$. Inequality (20) proves (19) using the substitution $x = (A - \sigma I)z / \|(A - \sigma I)z\|$. □

By Lemma 2.4 the residual vectors $r(x^{(k)})$ of inverse iteration (1) converge to the null vector. Thus the iterates $x^{(k)}$ converge to an eigenvector of A.

Remark 2.5. *Theorem 2.1 is restricted to symmetric positive definite matrices. To give an example of an indefinite matrix, let* $A = \mathrm{diag}(-3, 1)$ *and* $x = (1, 2)^T$. *Then*

$$\lambda(A^{-1}x) = \frac{33}{37} > \lambda(x) = \frac{1}{5},$$

since the component corresponding to the eigenvalue -3 *is damped out most rapidly. Hence inverse iteration for indefinite symmetric matrices is no longer a descent scheme for the Rayleigh quotient.*

REFERENCES

1. H. Wielandt. *Mathematische Werke/Mathematical works. Vol. 2.* Walter de Gruyter & Co., Berlin, 1996. Linear algebra and analysis, With essays on some of Wielandt's works by G. P. Barker, E. R. Barnes, J. von Below et al, Edited and with a preface by Bertram Huppert and Hans Schneider.
2. I. Ipsen. Computing an eigenvector with inverse iteration. *SIAM Review*, 39:254–291, 1997.
3. B.N. Parlett. *The symmetric eigenvalue problem.* Prentice Hall, Englewood Cliffs New Jersey, 1980.
4. F. Chatelin. *Eigenvalues of matrices.* Wiley, Chichester, 1993.
5. Gene H. Golub and Charles F. Van Loan. *Matrix computations.* Johns Hopkins University Press, Baltimore, MD, third edition, 1996.
6. I. Ipsen. *A history of inverse iteration*, volume in Helmut Wielandt, Mathematische Werke, Mathematical Works, Vol. 2: Linear Algebra and Analysis, pages 464–472. Walter de Gruyter, Berlin, 1996.
7. G.H. Golub and Q. Ye. Inexact inverse iteration for generalized eigenvalue problems. *BIT*, 40(4):671–684, 2000.
8. A.V. Knyazev. Computation of eigenvalues and eigenvectors for mesh problems: algorithms and error estimates. (In Russian), Dept. Numerical Math., USSR Academy of Sciences, Moscow, 1986.
9. A.V. Knyazev. Convergence rate estimates for iterative methods for a mesh symmetric eigenvalue problem. *Russian Journal of Numerical Analysis and Mathematical Modelling*, 2:371–396, 1987.
10. 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 and its Applications*, 358:95–114, 2003.
11. Y. Lai, K. Lin, and W. Lin. An inexact inverse iteration for large sparse eigenvalue problems. *Numer. Linear Algebra and its Applications*, 4:425–437, 1997.
12. K. Neymeyr. A geometric theory for preconditioned inverse iteration. I: Extrema of the Rayleigh quotient. *Linear Algebra and its Applications*, 322:61–85, 2001.
13. Y. Notay. Convergence analysis of inexact Rayleigh quotient iterations. *SIAM Journal on Matrix Analysis and Applications*, 24:627–644, 2003.
14. P. Smit and M. Paardekooper. The effects of inexact solvers in algorithms for symmetric eigenvalue problems. *Linear Algebra and its Applications*, 287:337–357, 1999.