2009

On the Modification of an Eigenvalue Problem that Preserves an Eigenspace

Maxim Maumov

Report Number:
09-004

On the Modification of an Eigenvalue Problem that Preserves an Eigenspace

Maxim Naumov

CSD TR #09-004
April 2009

Computer Science Technical Report, Purdue University -West Lafayette

On the Modification of an Eigenvalue Problem that Preserves an Eigenspace

Maxim Naumov

April 12th, 2009

Abstract Eigenvalue problems arise in many application areas ranging from computational fluid dynamics to information retrieval. In these fields we are often interested in only a few eigenvalues and corresponding eigenvectors of a sparse matrix. In this paper, we comment on the modifications of the eigenvalue problem that can simplify the computation of those eigenpairs. These transformations allow us to avoid difficulties associated with non-Hermitian eigenvalue problems, such as the lack of reliable non-Hermitian eigenvalue solvers, by mapping them into generalized Hermitian eigenvalue problems. Also, they allow us to expose and explore parallelism. They require knowledge of a selected eigenvalue and preserve its eigenspace. The positive definiteness of the Hermitian part is inherited by the matrices in the generalized Hermitian eigenvalue problem. The position of the selected eigenspace in the ordering of the eigenvalues is also preserved under certain conditions. The effect of using approximate eigenvalues in the transformation is analyzed and numerical experiments are presented.

Keywords Modification of the Eigenvalue Problem · non-Hermitian matrix · Hermitian matrix

Mathematics Subject Classification (2000) 65F15 · 65F50 · 65F99

1 Introduction

The solution of eigenvalue problems is of interest in many fields of computational science and engineering. In these fields we are often interested in obtaining only a few of the eigenpairs of a sparse matrix. Several authors have worked on developing methods for this problem, resulting in methods such as subspace iteration [16], Davidson [5], trace minimization [13], implicitly restarted Arnoldi [15], Jacobi-Davidson [14] and LOBPCG [11]. The search for faster algorithms and acceleration of existing ones is the subject of many papers.

In this paper we do not attempt to create a novel scheme, instead we propose a general technique that could be used to modify any eigenvalue problem so that its

Department of Computer Science, Purdue University - West Lafayette,
305 N. University Street, West Lafayette, IN, 47907-2107, United States.

eigenspace of interest is easier to compute. For example, there are several well known modifications of the generalized eigenvalue problem described in Theorem 4.8 [17]. These result from the combination of the matrices A and B already present in the pencil (A, B). Their effect is a simple shift applied to all the eigenvalues that leaves the structure of the spectrum the same.

We consider a modification of the eigenvalue problem that could be considered an extension of the above. We do not insist that the structure of the spectrum remains fixed, but keep invariant the subspace corresponding to a selected eigenvalue. Thus, the proposed technique is applicable to problems, where the eigenvalue, or an approximation of it, is already known, but its eigenspace remains to be computed.

It is well know that eigenvalues converge faster than eigenvectors, hence our approach can be used when the eigenvalue has already been obtained, but the eigenvector has not yet been found with sufficient accuracy. Also, it is ideally suited for applications where the eigenvalue is known beforehand, for example in the computation of PageRank [10,3] in the field of information retrieval. The use of approximate eigenvalues in the proposed transformation is also studied in the next section.

The modification creates a different eigenvalue problem that has an eigenvalue 1 with the corresponding eigenspace being identical to the eigenspace of the selected eigenvalue of the original problem. It results from the observation, that for Hermitian matrices with simple eigenvalues, the eigenvector $\mathbf{u}$ corresponding to the eigenvalue λ in the standard eigenvalue problem

$$A\mathbf{u} = \lambda\mathbf{u} \tag{1}$$

is also the eigenvector corresponding to the eigenvalue 1 of the generalized eigenvalue problem

$$(A + S)\mathbf{u} = \mu(\lambda I + S)\mathbf{u} \tag{2}$$

for some matrix S. Although we used a Hermitian matrix A in (2), such transformation also applies to non-Hermitian matrices.

There are two main contributions of this paper. First, is to show that the modification we propose can be used to transform a non-Hermitian into a generalized Hermitian eigenvalue problem. It is true that if the eigenvalue is known this can also be achieved by formulating the eigenvalue problem as a homogeneous linear system, that can be changed into a Hermitian form by using normal equations, and further reformulated as an eigenvalue problem in which we are interested in the eigenvector corresponding to the eigenvalue zero

$$(A - \lambda I)^T(A - \lambda I)\mathbf{u} = 0\mathbf{u} \tag{3}$$

However, such approach would not only square the condition number of the matrix at hand, require the knowledge of the matrix transpose (not necessarily available), but also would always be working with the original matrix. In our modification we introduce a new matrix S into the equation, which does not square the condition number or require explicit knowledge of the matrix transpose and can be chosen in many different ways, which will be discussed in the next sections. Moreover, we show that if the Hermitian or (premultiplied by i) skew-Hermitian part of the non-Hermitian eigenvalue problem is positive definite, one of the matrices in the generalized Hermitian eigenvalue problem will also be positive definite.

This transformation is very important, because although we could use inverse iteration or Jacobi-Davidson methods to find the smallest eigenpairs of the the non-Hermitian eigenvalue problem directly, both require solution of non-Hermitian systems

of linear equations. If the direct solution of these systems is not feasible, iterative methods such as restarted GMRES or BiCGstab are the remaining alternative. However, even with more advanced preconditioners such as those in ILUPACK [2], these schemes are often very slow to converge, hence preventing us from obtaining the eigenpairs of interest. On the other hand, if we transform the original eigenvalue problem into a generalized Hermitian problem, a variety of other eigenvalue solvers with a well-developed theory of convergence, applicable only to Hermitian problems, also becomes available, e.g. [8, 13, 11].

Second, is to show that our approach can be used to improve the parallelism available in the eigenvalue problem. For instance suppose that we are interested in the smallest eigenvalue and to obtain it apply subspace iteration on

$$A^{-1}\mathbf{u} = \frac{1}{\lambda}\mathbf{u} \tag{4}$$

Notice that we can write matrices

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \text{ and } S = \begin{pmatrix} & -A_{12} \\ -A_{21} & \end{pmatrix} \tag{5}$$

and assume that diagonal blocks A_{11} and A_{22} are nonsingular. Then, using (2) we can write the modified eigenvalue problem as

$$\begin{pmatrix} A_{11} & \\ & A_{22} \end{pmatrix}^{-1} \begin{pmatrix} \lambda I & -A_{12} \\ -A_{21} & \lambda I \end{pmatrix} \mathbf{u} = \frac{1}{\mu}\mathbf{u} \tag{6}$$

If $\frac{1}{\mu}$ is its largest eigenvalue we can also obtain it using subspace iteration. However, instead of making solves with the original matrix, we will only need to make two independent solves with diagonal blocks and a matrix-vector multiplication with off-diagonal blocks. It will be shown that for Hermitian matrices, under some additional conditions, the ordering of the eigenspace in the original and modified eigenvalue problem is the same. Hence, if λ is the smallest eigenvalue of (1), then μ is the smallest eigenvalue of (2) and the above approach can significantly reduce the computational costs.

Thus, the modification we are about to describe may simplify difficulties associated with solving non-Hermitian eigenvalue problems and allow us to expose and explore available parallelism. It can also be used to isolate a Gershgorin disk of the modified problem and enables other interesting scalings.

2 Modification of the Eigenvalue Problem

Let us first restate a very convenient definition of Jordan decomposition, see Theorem 1.22 [17].

Lemma 1 *Let the matrix $A \in \mathbb{C}^{n\times n}$ have k distinct eigenvalues $\lambda_1, \ldots, \lambda_k$ of algebraic multiplicities $m_1, \ldots, m_k$. Then there are unique integers m_{ij} for $i = 1, \ldots, k$ and $j = 1, \ldots, l_i$ satisfying*

$$m_i = \sum_{j=1}^{l_i} m_{ij} \tag{7}$$

and a nonsingular matrix X such that

$$X^{-1}AX = diag(J_1, \ldots, J_k) \tag{8}$$

where $J_i = diag(J_{m_{i1}}, \ldots, J_{m_{il_i}}) \in \mathbb{C}^{m_i \times m_i}$ and $J_{m_{ij}}$ are Jordan blocks of order m_{ij}.

Once again following [17], we partition $X = (X_1, \ldots, X_k)$, $X_i = (X_{i1}, \ldots, X_{il_i})$ and $X_{ij} = (\mathbf{x}_1^{(ij)}, \ldots, \mathbf{x}_{m_{ij}}^{(ij)})$.

Suppose that we are interested in finding an eigenvalue λ_i with multiplicity l_i and corresponding eigenvectors $\mathbf{x}_1^{(ij)}$, in other words, we must solve the standard eigenvalue problem

$$AU_i = \lambda_i U_i \tag{9}$$

where $U_i = (\mathbf{x}_1^{(i1)}, \ldots, \mathbf{x}_1^{(il_i)}) \in \mathbb{C}^{n \times l_i}$.

Let us consider the following generalized eigenvalue problem

$$(A + S)V_i = \mu_i(\alpha I + S)V_i \tag{10}$$

where $V_i \in \mathbb{C}^{n \times l_i}$, $\mu_i, \alpha \in \mathbb{C}$ and $S \in \mathbb{C}^{n \times n}$.

Theorem 1 *If $\alpha = \lambda_i$ then the eigenspace corresponding to the eigenvalue $\mu_i = 1$ of the pencil $(A+S, \alpha I+S)$ is equivalent to the eigenspace corresponding to the eigenvalue λ_i of the matrix A, in other words,*

$$span(V_i) \equiv span(U_i) \tag{11}$$

Proof Let $Y_i \in \mathbb{C}^{n \times l_i}$. If Y_i satisfies (9), then by adding SY_i to both sides of (9) we obtain

$$(A + S)Y_i = (\lambda_i I + S)Y_i \tag{12}$$

thus Y_i is the eigenspace corresponding to the eigenvalue $\mu_i = 1$ of the pencil $(A + S, \alpha I + S)$ in (10), where $\alpha = \lambda_i$. Hence, $\text{span}(U_i) \subseteq \text{span}(V_i)$.

On the other hand, if Y_i satisfies (10) with $\mu_i = 1$ and $\alpha = \lambda_i$, then upon subtracting SY_i and simplifying we obtain

$$AY_i = \alpha Y_i \tag{13}$$

thus (λ_i, Y_i) is an eigenpair of A in (9). Hence, $\text{span}(V_i) \subseteq \text{span}(U_i)$.

Combining both results we obtain that

$$\text{span}(V_i) \equiv \text{span}(U_i) \tag{14}$$

□

Notice that in (10) we are modifying the eigenvalue problem (9) with an arbitrary matrix S, and we are keeping only a selected eigenspace the same, all other eigenvalues and their corresponding eigenspaces can change. Clearly, to use (10) in practice we must already know a particular eigenvalue of (9) and be interested only in its eigenspace.

Although in this paper we are focused on general matrices, we point out the following result concerning the ordering of the eigenvalues after the modification, which applies only to Hermitian matrices.

Theorem 2 *Suppose that A is Hermitian and that $\alpha I + S$ is Hermitian positive definite. Let the eigenvalues of* (9) *in increasing order be denoted by*

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n \tag{15}$$

and those of (10) *by*

$$\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n \tag{16}$$

If we choose $\alpha = \lambda_k$ then $\mu_k = 1$. Hence, the eigenspace corresponding to the k-th eigenvalue of the original eigenvalue problem (9) *is identical to the eigenspace corresponding to the k-th eigenvalue of the modified eigenvalue problem* (10)*.*

Proof Let $\hat{\lambda}_i = \lambda_i - \lambda_k$ and the triplet (ν, ζ, π) denote the i-th eigenvalue and inertia (see Definition 2.4 in [17]), respectively, of the matrix $\hat{A} = A - \lambda_k I$. Consider the eigenvalue problem (10) with $\alpha = \lambda_k$ and eigenvalues shifted by -1 that is given by

$$(A - \lambda_k I)V_i = (\mu_i - 1)(\lambda_k I + S)V_i \tag{17}$$

Let $P = (\lambda_k I + S)^{-1}$ and $\hat{\mu}_i = (\mu_i - 1)$, then we can rewrite (17) as the standard eigenvalue problem

$$P\hat{A}V_i = \hat{\mu}_i V_i \tag{18}$$

Using Theorem 3 in [12] we conclude that

$$\hat{\mu}_i = \theta\hat{\lambda}_i \tag{19}$$

where θ is between the smallest and the largest eigenvalue of P. Since $\theta > 0$, the inertia of $\hat{A}$ and $P\hat{A}$ are identical. Thus, the number of eigenvalues $\mu_i < 1$ is the same as that for which $\lambda_i < \lambda_k$. Similarly for $\mu_i = 1$ & $\lambda_i = \lambda_k$ and $\mu_i > 1$ & $\lambda_i > \lambda_k$. Hence, the position of the selected eigenspace in the ordering of the eigenvalues is preserved. □

Let us now consider what happens when an approximation to the exact eigenvalue is used in the modified problem (10). For simplicity, in the theorem below, we assume that $l_i = 1$, in other words, we work with a simple eigenpair $(\lambda_i, \mathbf{u}_i)$.

Theorem 3 *Suppose that $\lambda_i I + S$ is Hermitian positive definite. Let $\tilde{\lambda}_i = \lambda_i + \epsilon$, with ϵ sufficiently small, be an approximation to a simple eigenvalue λ_i. Then using $\alpha = \tilde{\lambda}_i$, we obtain the approximation $(\tilde{\lambda}_i, \tilde{\boldsymbol{v}}_i)$ to the eigenpair $(\lambda_i, \boldsymbol{v}_i)$ of* (9)*. The residual of this approximate eigenpair satisfies*

$$||\tilde{\boldsymbol{r}}_i||_2 \leq \varphi(\epsilon)||S + \lambda_i I||_2||\boldsymbol{v}_i||_2 \tag{20}$$

where $\varphi(\epsilon)$ is a second degree polynomial in ϵ.

Proof For clarity we will omit subscripts in this proof. Assume that we know an approximate eigenvalue $\tilde{\lambda} = \lambda + \epsilon$ of the original eigenvalue problem (9). Also, suppose that ϵ is sufficiently small so that a simple eigenpair $(\tilde{\mu}, \tilde{\mathbf{v}})$ exists and that we have solved the perturbed modified eigenvalue problem

$$(A + S)\tilde{\mathbf{v}} = \tilde{\mu}(\tilde{\lambda}I + S)\tilde{\mathbf{v}} \tag{21}$$

First, let us treat ϵ as a perturbation and use the first-order perturbation analysis outlined for standard eigenvalue problems in [17] to understand how $(\tilde{\lambda}, \tilde{\mathbf{v}})$ is related

to $(\lambda, \mathbf{v})$. Denote $\bar{A} = A + S$, $\bar{B} = \lambda I + S$ and let $n \times n$ nonsingular matrices $(\mathbf{v}, W)$ and $(\mathbf{y}, Z)^H$ be such that they satisfy

$$\begin{pmatrix} \mathbf{y}^H \\ Z^H \end{pmatrix} \bar{A} \, (\mathbf{v}, W) = \begin{pmatrix} \mu & 0 \\ 0 & M \end{pmatrix} \tag{22}$$

and

$$\begin{pmatrix} \mathbf{y}^H \\ Z^H \end{pmatrix} \bar{B} \, (\mathbf{v}, W) = \begin{pmatrix} 1 & 0 \\ 0 & I \end{pmatrix} \tag{23}$$

Then, let us write $\tilde{\mu} = \mu + \phi$ and $\tilde{\mathbf{v}} = \mathbf{v} + W\mathbf{p}$. Substituting these expressions into (21) , simplifying the equation using the fact that $\bar{A}\mathbf{v} = \mu\bar{B}\mathbf{v}$ and ignoring second- and higher-order terms in ϵ, ϕ, $\mathbf{p}$ or combinations of those we obtain

$$\bar{A}W\mathbf{p} \approx \mu\bar{B}W\mathbf{p} + \mu\epsilon\mathbf{v} + \phi\bar{B}\mathbf{v} \tag{24}$$

Premultiplying by $\mathbf{y}^H$ and using (22) and (23) we obtain

$$\phi \approx -\epsilon\mu\theta \text{ where } \theta = \mathbf{y}^H\mathbf{v} \tag{25}$$

Similarly, premultiplying by Z^H and once again using (22) and (23) we obtain

$$\mathbf{p} \approx \epsilon\mu(M - \mu I)^{-1} Z^H\mathbf{v} \tag{26}$$

Notice that we are only interested in the eigenspace corresponding to eigenvalue $\mu = 1$. Thus, when computing the expression for the residual we will simply omit μ from the formulas (25), (26) and let $\tilde{\mu} = 1 + \phi$ in (21).

Finally, let us obtain the expression for the residual of the original eigenvalue problem (9). Using (21),

$$\tilde{\mathbf{r}} = A\tilde{\mathbf{v}} - \tilde{\lambda}\tilde{\mathbf{v}} = (A + S)\tilde{\mathbf{v}} - (\tilde{\lambda}I + S)\tilde{\mathbf{v}} = \phi((\lambda + \epsilon)I + S)(\mathbf{v} + W\mathbf{p}) \tag{27}$$

Once again, regrouping second-order terms in ϵ, ϕ, $\mathbf{p}$ and their combinations, and using (25), (26) we obtain

$$||\tilde{\mathbf{r}}||_2 \leq |\epsilon||\theta|||S + \lambda I||_2||\mathbf{v}||_2 + O(|\epsilon|^2 + |\epsilon|||\mathbf{p}||_2) \tag{28}$$

Unless $M - \mu I$ or Z in (26) is extremely ill-conditioned, $||\mathbf{p}||_2 \sim O(\epsilon)$. Then,

$$||\tilde{\mathbf{r}}||_2 \leq \varphi(\epsilon)||S + \lambda I||_2||\mathbf{v}||_2 \tag{29}$$

for some second degree polynomial $\varphi(\epsilon)$. □

Hence, the residual of the eigenvalue problem (9) mainly depends on the absolute value of the error ϵ in the approximate eigenvalue, the norm of the eigenvector $\mathbf{v}$, and the norm of the matrix $S + \lambda I$. We should point out that, by using the perturbation theory for the generalized eigenvalue problems in [17], a similar bound on $||\tilde{\mathbf{r}}||_2$ can be obtained even when $\lambda_i I + S$ is not Hermitian positive definite, as long as the pencil $(\bar{A}, \bar{B})$ is regular.

So far we have introduced the modification of an eigenvalue problem that preserves an eigenspace. We have studied some of its properties, such as the ordering of the eigenspace in the modified eigenvalue problem and the effect of using an approximation to the exact eigenvalue in the transformation. Let us now focus on three particular choices for modification matrix S that can be useful in practice.

3 Particular Choices of S

3.1 Diagonal

Let $S = D$, where $D = \text{diag}(d_1, \ldots, d_n)$ is a diagonal matrix. Notice that if $\widetilde{D} = D + \alpha I$ is nonsingular, then from (10) we can write the modified eigenvalue problem as

$$\widetilde{D}^{-1}(A + D)V_i = \mu_i V_i \tag{30}$$

Hence, we are scaling the rows of matrix A and augmenting the resulting diagonal. Such scaling may be useful for very ill-conditioned matrices or can be used to enhance the dominance of certain rows. Notice that this is different from the diagonal similarity transformation D that can be applied directly to (9), yielding

$$D^{-1}AD\widetilde{V}_i = \mu_i \widetilde{V}_i \tag{31}$$

where $\widetilde{V}_i = D^{-1}V_i$. There are situations where one of these might be preferred over the other. For example, if we have a row diagonally dominant matrix and know the particular eigenvalue we seek, the modification (30) can be used to preserve the diagonal dominance and obtain a better conditioned eigenvalue problem. It can also be used to isolate the Gershgorin disk of a particular eigenvalue.

3.2 Rank-k Perturbation

Let $\text{rank}(S) = k$, hence we can express $S = ZY^H$, where $Z, Y \in \mathbb{C}^{n\times k}$ are tall matrices. Then, the modified eigenvalue problem (10) can be rewritten as

$$(A + ZY^H)V_i = \mu_i(\alpha I + ZY^H)V_i \tag{32}$$

Assume for simplicity that $\alpha I + ZY^H$ is nonsingular. Notice that if we know the eigenvalue of interest we can compute its eigenspace by solving a rank-k modification (32) of the original eigenvalue problem.

This approach, for example, could potentially contribute to the computation of PageRank [10,3] in the field of information retrieval. Using the notation of [3], recall that in this problem we are interested in finding the eigenvector corresponding to the eigenvalue 1 of an irreducible stochastic matrix

$$P_c = c\widetilde{P} + (1 - c)\mathbf{e}\mathbf{v}^T \tag{33}$$

where $\widetilde{P}$ is a reducible stochastic matrix, $\mathbf{v}$ is the personalization vector, $\mathbf{e} = (1, \ldots, 1)^T$ and constant $c \in (0, 1)$.

Once again, it is worth noticing that it is completely different from the rank-one perturbation approach described in [7,4,1] where the knowledge of the entire spectrum is used to give an expression for the eigenpairs of the modified eigenvalue problem.

3.3 Hermitian or skew-Hermitian part

Let us rewrite matrix A in terms of its Hermitian $A_H = \frac{1}{2}(A+A^H)$ and skew-Hermitian $A_{SH} = \frac{1}{2}(A - A^H)$ parts, so that

$$A = A_H + A_{SH} \tag{34}$$

Then, choosing $S = -A_H$ in (10) we obtain

$$A_{SH}V_i = \mu_i(\alpha I - A_H)V_i \tag{35}$$

Using the fact that $A_{HH} = iA_{SH}$ is Hermitian, premultiplying both sides by i and letting $\widetilde{\mu}_i = i\mu_i$ we obtain the Hermitian generalized eigenvalue problem

$$A_{HH}V_i = \widetilde{\mu}_i(\alpha I - A_H)V_i \tag{36}$$

It is clear for this particular choice of S that if the (premultiplied by i) skew-Hermitian part of the non-Hermitian eigenvalue problem is positive definite or if the spectrum $\sigma(A_H) \subseteq (-\infty, \alpha)$, one of the matrices in the generalized Hermitian eigenvalue problem inherits positive definiteness. In this case a variety of Hermitian eigenvalue solvers [8, 13,11] can be applied to solve the modified eigenvalue problem (36) and obtain the eigenspace of interest.

4 Numerical Experiments

To illustrate the proposed modifications of the eigenvalue problem let us focus on three particular choices of S discussed earlier: first $S = D$ second $S = ZY^H$ and finally $S = -A_H$. For clarity for the first and second choices we will work with real 5×5 and 7×7 matrices, respectively, while in the third case general sparse nonsymmetric matrices will be used. Matlab is used for numerical computations.

4.1 Experiment 1 - Diagonal Dominance

Consider the matrix

$$A = \begin{pmatrix} 2 & -\frac{3}{2} & & & \\ -\frac{5}{2} & 2 & -\frac{3}{2} & & \\ & -\frac{5}{2} & 2 & -\frac{3}{2} & \\ & & -\frac{5}{2} & 2 & -\frac{3}{2} \\ & & & -\frac{5}{2} & 2 \end{pmatrix} \tag{37}$$

Notice that it is not diagonally dominant and its eigenvalues are

$$-1.35, 0.06, 2.00, 3.94, 5.35 \tag{38}$$

Let us consider the diagonal modification $D = \mathrm{diag}(0, 0, 18, 0, 0)$ with $\alpha = 2.0$. Then the matrix in the modified eigenvalue problem (30) is written as

$$\widetilde{D}^{-1}(A + D) = \begin{pmatrix} 1 & -\frac{3}{4} & & & \\ -\frac{5}{4} & 1 & -\frac{3}{4} & & \\ & -\frac{5}{40} & 1 & -\frac{3}{40} & \\ & & -\frac{5}{4} & 1 & -\frac{3}{4} \\ & & & -\frac{5}{4} & 1 \end{pmatrix} \tag{39}$$

It has a selected diagonally dominant row and its eigenvalues are

$$-0.06, 0.03, 1.00, 1.97, 2.10 \tag{40}$$

with eigenvector corresponding to eigenvalue 1.0 in (39) being the same as eigenvector corresponding to eigenvalue 2.0 in (37).

4.2 Experiment 2 - Gershgorin Disk

Consider the matrix

$$A = \begin{pmatrix} 10 & -3 & & & \\ -6 & 20 & -3 & & \\ & -6 & 30 & -3 & \\ & & -6 & 40 & -3 \\ & & & -6 & 50 \end{pmatrix} \tag{41}$$

It has eigenvalues

$$8.34, 19.87, 30.00, 40.13, 51.66 \tag{42}$$

Once again, let us consider the diagonal modification $D = \text{diag}(0, 0, 270, 0, 0)$ with $\alpha = 30.0$. Then the matrix in the modified eigenvalue problem (30) is

$$\widetilde{D}^{-1}(A + D) = \begin{pmatrix} \frac{1}{3} & -0.1 & & & \\ -0.2 & \frac{2}{3} & -0.1 & & \\ & -0.02 & 1 & -0.01 & \\ & & -0.2 & \frac{4}{3} & -0.1 \\ & & & -0.2 & \frac{5}{3} \end{pmatrix} \tag{43}$$

Notice that the third row in (43) has Gershgorin disk that is isolated from the rest, see Fig. 1. For completeness we mention that the matrix above has eigenvalues

$$0.28, 0.71, 1.00, 1.29, 1.72 \tag{44}$$

and that the eigenvector corresponding to eigenvalue 1.0 of the modified and eigenvalue 30.0 of the original problem match.

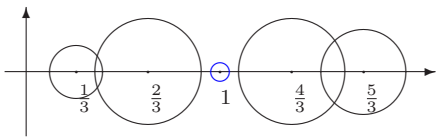

Fig. 1 Plot of Gershgorin disks of the matrix in (43)

4.3 Experiment 3 - Rank-k Update

Consider the tridiagonal matrix resulting from the standard second-order space central discretization of the one dimensional laplace operator with seven points given below

$$A = \left(\begin{array}{ccc|ccc} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ \hline & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{array}\right) \tag{45}$$

Its eigenvalues, rounded up to two digits of accuracy, are

$$0.20, 0.75, 1.55, 2.45, 3.25, 3.80 \tag{46}$$

Suppose we are interested in its smallest eigenvalue $\alpha = 0.20$. Then, the modified eigenvalue problem (32) can be written as

$$\frac{1}{\mu_1}\mathbf{v}_1 = \left(\begin{array}{ccc|ccc} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & & & \\ \hline & & & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{array}\right)^{-1} \left(0.20I + \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}\right) \mathbf{v}_1 \tag{47}$$

Notice that using Theorem 2 we conclude that $\mu_1 \equiv 1$ is the smallest eigenvalue of the modified eigenvalue problem (32) and consequently $\frac{1}{\mu_1}$ is the largest eigenvalue of (47). Hence, we can use a subspace iteration method, which only requires two independent solves, that can be performed in parallel, and a vector scaling to find the eigenvalue of interest. Notice that if we use subspace iteration on the original eigenvalue problem

$$\frac{1}{\lambda_1}\mathbf{u}_1 = A^{-1}\mathbf{u}_1 \tag{48}$$

we would need to make solves with the entire matrix A in (45), which is more computationally expensive. It is well known that eigenvalues converge faster than eigenvectors, so as mentioned earlier this approach can be used when the eigenvalue has already been obtained, but eigenvector has not yet been found with sufficient accuracy.

4.4 Experiment 4 - Non-Hermitian to Hermitian

Finally, consider the two randomly chosen nonsymmetric matrices N_1 and N_2 from The University of Florida Matrix Collection [6], with some of their properties and the pattern of nonzero elements shown in Tab. 1 and Fig. 2-3, respectively.

Table 1 Matrices

	Matrix	Size	Nonzeros	$\|\|.\|\|_2$	Application
N_1:	Graham/graham1	9,035	335,472	8.2e+04	Comp. Fluid Dynamics
N_2:	Hollinger/mark3jac020sc	9,129	52,883	2.3e+06	Economics

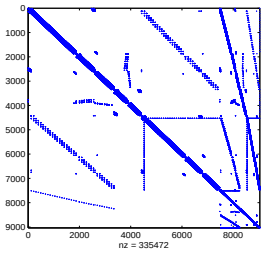

Fig. 2 Matrix N_1

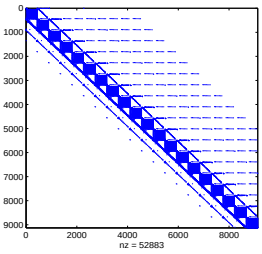

Fig. 3 Matrix N_2

The eigenvalues of N_1 and N_2 are extremely clustered with the five smallest in absolute value shown in arrays $E_1^{(5)}$ and $E_2^{(5)}$ of (49), respectively:

$$E_1^{(5)} = \begin{pmatrix} 6.75 \\ 6.76 \\ 16.7 \\ 16.7 \\ 20.8 \end{pmatrix} \cdot 10^{-10} \text{ and } E_2^{(5)} = \begin{pmatrix} 1.84 \\ 1.86 \\ 1.86 - 0.00469i \\ 1.86 + 0.00469i \\ 3.68 \end{pmatrix} \cdot 10^{-6} \tag{49}$$

Let us suppose that we are interested in the second $\left(6.76 \cdot 10^{-10}, \mathbf{u}_2^{(N_1)}\right)$ and third $\left((1.86 - 0.00469i) \cdot 10^{-6}, \mathbf{u}_3^{(N_2)}\right)$ eigenpairs of the nonsymmetric matrices N_1 and N_2, respectively. Then computing the eigenpairs $\left(i, \mathbf{v}_2^{(N_1)}\right)$ and $\left(i, \mathbf{v}_3^{(N_2)}\right)$ of the generalized Hermitian eigenvalue problem (36) we obtain

$$||\mathbf{u}_2^{(N_1)} - \mathbf{v}_2^{(N_1)}||_2 = 1.65 \cdot 10^{-6} \tag{50}$$

$$||\mathbf{u}_3^{(N_2)} - \mathbf{v}_3^{(N_2)}||_2 = 2.24 \cdot 10^{-3} \tag{51}$$

The accuracy is lower in (51), because clustering of the eigenvalues of the original matrix leads to clustering of the eigenvalues of the modified eigenvalue problem. In fact in the modified eigenvalue problem resulting from N_2 the three eigenvalues closest to i are

$$\begin{gathered} 0.000000000000074 + 0.999999999999974i \\ -0.000000000002871 + 0.999999999996976i \\ -0.000000000010112 + 1.000000000003275i \end{gathered} \tag{52}$$

Suppose that in the process of looking for the eigenvector associated with i, we have computed the eigenvectors $Z = \left(\mathbf{v}_3^{(N_2)}, \mathbf{v}_4^{(N_2)}, \mathbf{v}_5^{(N_2)}\right)$ associated with these three eigenvalues. Then, we may use them to construct a better approximation to $\mathbf{u}_3^{(N_2)}$ by solving the linear least squares problem

$$Z\boldsymbol{\zeta} = \mathbf{u}_3^{(N_2)} \tag{53}$$

with new approximation to the eigenvector $\mathbf{u}_3^{(N_2)}$ being $\bar{\mathbf{v}}_3^{(N_2)} = Z\boldsymbol{\zeta}$. The accuracy of the resulting approximation is

$$||\mathbf{u}_3^{(N_2)} - \bar{\mathbf{v}}_3^{(N_2)}||_2 = 6.14 \cdot 10^{-6} \tag{54}$$

which is sufficient for most of the applications. Notice that in all of the above examples we were indeed able to reduce the solution of a non-Hermitian eigenvalue problem to the solution of a more simple generalized Hermitian eigenvalue problem.

5 Conclusion

In this paper, we explored the modification $(A+S)V_i = \mu_i(\alpha I + S)V_i$ of the eigenvalue problem $AU_i = \lambda_i U_i$. Using $\alpha = \lambda_i$ we noticed that it preserves a selected eigenspace and explored some of the choices for S. The main results described the modification that allows us to transform a non-Hermitian into a generalized Hermitian eigenvalue problem as well as the rank-k update, making possible the application of algorithms available only for Hermitian problems and exposing available parallelism, respectively. We have analyzed the effects of using approximate eigenvalues in the transformation and shown when the ordering of the eigenspace is preserved. Finally, numerical experiments validating the theoretical results were also presented.

References

1. P. Arbenz, W. Gander and G. H. Golub, Restricted Rank Modification of the Symmetric Eigenvalue Problem: Theoretical Considerations, Linear Algebra Appl., 104, pp. 75-95 (1988)
2. Matthias Bollhofer, A Robust ILU Based on Monitoring the Growth of the Inverse Factors, Linear Algebra Appl., 338, pp. 201-218 (2001)
3. C. Brezinski and M. Redivo-Zaglia, The PageRank Vector: Properties, Computation, Approximation, and Acceleration, SIAM J. of Mat. Anal. Appl., 28, pp. 551-575 (2006)
4. J. R. Bunch, C. P. Nielsen and D. Sorensen, Rank-One Modification of the Symmetric Eigenproblem, Numer. Math., 31, pp. 31-48 (1978)
5. E. R. Davidson, The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices, J. Comp. Phys., 17, pp. 87-94 (1975)
6. T. Davis, The University of Florida Sparse Matrix Collection, http://www.cise.ufl.edu/research/sparse/matrices/ (2008)
7. G. Golub, Some Modified Matrix Eigenvalue Problems, SIAM Rev., 15, pp. 318-335 (1973)
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. Nat. Bur. Stand., 47, pp. 45-61 (1951)
9. R. A. Horn and C. R. Johnson, Matrix Analysis, Cambridge University Press, New York, NY, (1999)
10. S. D. Kamvar, T. H. Haveliwala, C. D. Manning and G. H. Golub, Extrapolation Methods for Accelerating PageRank Computations, in 12th WWW, ACM Press, NY, pp. 261-270 (2003)
11. A. V. Knyazev, Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method, SIAM J. Scient. Comp., 23, pp. 517-541 (2001)
12. A. M. Ostrowski, A Quantitative Formulation of Sylvester's Law of Inertia, in N.A.S. USA, pp. 740-744 (1959)
13. A. H. Sameh and J. Wisniewski, A Trace Minimization Algorithm for the Generalized Eigenvalue Problem, SIAM J. Numer. Anal., 19, pp. 1243-1259 (1982)
14. G. L. G. Sleijpen and H. A. Van der Vorst, A Jacobi-Davidson Iteration Method for Linear Eigenvalue Problems, SIAM Rev., 42, pp. 267-293 (2000)
15. D. C. Sorensen, Implicit Application of Polynomial Filters in a k-step Arnoldi Method SIAM J. Matrix Anal. Appl., 13, pp. 357-385 (1992)
16. G. W. Stewart, Accelerating the Orthogonal Iteration for the Eigenvectors of a Hermitian Matrix, Numer. Math., 13, pp. 362-376 (1969)
17. G. W. Stewart, Matrix Algorithms, Volume II: Eigensystems, SIAM, Philadelphia, PA, (2001)