EIGEN-PAIRS OF PERTURBED MATRICES

M.THAMBAN NAIR

Department of Mathematics
Indian Institute of Technology Madras
Chennai-600 036, INDIA
E-Mail: mtnair@iitm.ac.in

Abstract

Given a simple eigenvalue and a corresponding eigenvector of a matrix $A$ we obtain an estimate for the norm of the perturbation matrix $E$ so that the perturbed matrix $A + E$ has a simple eigenvalue and an eigenvector which is normalized using a left eigenvector of $A$. Our procedure of establishing the existence of the eigen-pair is by finding a fixed point of certain equation, which suggests iterative refinements for yielding the required eigen-pair, and establishes that the perturbed eigen-pair is simple as well. The method adopted here is in variant with a procedure of Stewart (1987) and also simpler than the the analysis carried out recently by Varadharaj (2003).

Mathematics Subject Classification: 47A10, 47A75, 65F15, 65J10

Key words: Eigenvalues, eigenvectors, eigen-pairs, algebraic multiplicity, geometric multiplicity, index, contraction map, iterative refinements.

1 Introduction

For positive integers $m, n$ we denote by $\mathbb{C}^{m \times n}$ the set of all $m \times n$ matrices with its entries in the complex field $\mathbb{C}$. If $A_1 \in \mathbb{C}^{n \times k}$ and $A_2 \in \mathbb{C}^{n \times r}$, then $(A_1, A_2)$ is the matrix
in $\mathbb{C}^{n \times (k+r)}$ with its first columns are those in $A_1$, and the remaining $r$ columns are those in $A_2$. We shall denote the set $\mathbb{C}^{n \times 1}$ by $\mathbb{C}^n$. Thus elements from $\mathbb{C}^n$ are sometimes called *vectors* and are usually denoted by small case letters such as $x, y, w, p, q$ etc. For $A \in \mathbb{C}^{m \times n}$, its conjugate transpose will be denoted by $A^*$, and

$$N(A) := \{x \in \mathbb{C}^n : Ax = 0\}, \quad R(A) = \{Ax : x \in \mathbb{C}^n\}.$$  

Clearly the set $R(A)$ is the space spanned by the columns of $A$ which is also denoted by $\text{span}(A)$. We note that $u \in \text{span}(A)$ if and only if there exists a vector $v \in \mathbb{C}^k$ such that $u = Av$.

For $x \in \mathbb{C}^n$, its norm $\|x\|$ is defined as the positive square root of $x^*x$, and the norm of a matrix $A \in \mathbb{C}^{m \times n}$, denoted by $\|A\|$, is defined by

$$\|A\| := \sup\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\}.$$  

For $x, y \in \mathbb{C}^k$, the quantity $y^*x$ is denoted by $\langle x, y \rangle$. We say that vectors in a set $S \subseteq \mathbb{C}^n$ of $\mathbb{C}^n$ are said to be *orthogonal* if $\langle x, y \rangle = 0$ for every distinct $x, y$ in $S$; and the vectors in $S$ are called *orthonormal* if they are orthogonal and $\|x\| = 1$ for every $x \in S$. It is seen that for $A \in \mathbb{C}^{m \times n}$

$$\langle Ax, y \rangle = \langle x, A^*y \rangle \quad \forall x \in \mathbb{C}^n, y \in \mathbb{C}^n.$$

Suppose $A \in \mathbb{C}^{n \times n}$. Recall that $\lambda_0 \in \mathbb{C}$ is an *eigenvalue* of $A$ if there exists a nonzero vector $x_0 \in \mathbb{C}^n$ such that

$$Ax_0 = \lambda_0 x_0,$$

and in that case $x_0$ is called an *eigenvector* corresponding to the eigenvalue $\lambda_0$. It is a well-known fact that $\lambda_0 \in \mathbb{C}$ is an eigenvalue of $A$ if and only if the $\lambda_0$ is a zero of the characteristic polynomial

$$p_A(z) := \det(A - zI), \quad z \in \mathbb{C}.$$  

Here, $I$ denotes the identity matrix in $\mathbb{C}^{n \times n}$. An eigenvalue $\lambda_0$ of $A$ is called a *simple eigenvalue* if $\lambda_0$ is a simple zero of $p_A(z)$.

Suppose $\lambda_0$ is an eigenvalue of $A \in \mathbb{C}^{n \times n}$ with a corresponding eigenvector $x_0$ of norm 1. It is desirable to know the effect on $\lambda_0$ and $x_0$ when $A$ is perturbed by another
matrix $E \in \mathbb{C}^{n \times n}$. It is a well known fact that, if $\|E\|$ is small enough, and if $\lambda_0$ is an eigenvalue of $A$ with a corresponding eigenvector $x_0$, then $A + E$ has an eigenvalue $\lambda$ which is close to $\lambda_0$; but, there need not exist an eigenvector $x$ for $A + E$ which is ‘close to’ $x_0$. Here, closeness of vectors may be measured in terms of the angle between them or by the norm of their difference.

To illustrate this, consider the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{pmatrix}.$$  

It is easily seen that $\lambda_0 = 1$ is an eigenvalue of $A$ with two linearly independent eigenvectors $u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and $\lambda_1 = 1 + \varepsilon$, $\lambda_2 = 1 - \varepsilon$ are eigenvalues of $A + E$ with corresponding eigenvectors $v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ respectively. Clearly the acute angle between $u_1, v_1$ (respectively, $u_2, v_2$) is $\pi/4$, no matter what the value of $\varepsilon$ is, although $\|E\|$ is small for small enough $\varepsilon$.

Such a bad situation does not prevail if the eigenvalue $\lambda_0$ under consideration is simple. This issue has been discussed in the literature in great detail in the context of operators (cf. Kato [1], Stewart [6, 7], Nair [4]) as well as in the context of matrices (cf. Stewart [8]). Another article of Stewart which is quite educative in the context of matrix eigenvalue is [9] which also directed towards obtaining approximations to invariant subspaces as in Nair [3]. The method adopted here is in slight variant with the procedure of Stewart [9], and also simpler than the analysis carried out recently by Varadharaj [10]. A main feature of the result in this paper is that a quantity $\varepsilon_0$ is specified such that for all $E$ with $\|E\| < \varepsilon_0$, the errors $|\lambda - \lambda_0|$ and $\|x - x_0\|$ are expressible in terms of $\|E\|$ which also ensure that the perturbed eigenpair is simple.

2 Preliminary Notions

Suppose $A \in \mathbb{C}^{n \times n}$. Recall that $\lambda \in \mathbb{C}$ is an eigenvalue of $A$ if and only if there exists nonzero $x \in \mathbb{C}^n$ such that

$$Ax = \lambda x.$$
The following theorem can be verified easily. For a detailed proof see Nair [5].

**THEOREM 2.1** Suppose $\lambda$ is an eigenvalue of $A$. Then we have following:

(a) $N((A - \lambda I)^k) \subseteq N((A - \lambda I)^{k+1})$ for all $k \in \mathbb{N}$.

(b) There exists $k \in \mathbb{N}$ such that $N((A - \lambda I)^k) = N((A - \lambda I)^{k+1})$.

(c) If $k \in \mathbb{N}$ is such that $N((A - \lambda I)^k) = N((A - \lambda I)^{k+1})$, then $N((A - \lambda I)^k) = N((A - \lambda I)^{k+j})$ for all $j \in \mathbb{N}$.

Suppose $\lambda$ is an eigenvalue of $A$. Then the number

$$\ell := \min\{k \in \mathbb{N} : N((A - \lambda I)^k) = N((A - \lambda I)^{k+1})\}$$

is called the *index* or *ascent* of $\lambda$. The following result is proved in Nair [5].

**THEOREM 2.2** Suppose $\lambda$ is an eigenvalue of $A$ and $\ell$ is the index of $\lambda$. Then $N((A - \lambda I)^{\ell})$ and $R((A - \lambda I)^{\ell})$ are invariant under $A$, and

$$\mathbb{C}^n = N((A - \lambda I)^{\ell}) + R((A - \lambda I)^{\ell}), \quad N((A - \lambda I)^{\ell}) \cap R((A - \lambda I)^{\ell}) = \emptyset.$$ 

Suppose $\lambda$ is an eigenvalue of $A$ and $\ell$ is the index of $\lambda$. Let $m = \dim N((A - \lambda I)^{\ell})$. Then, by Theorem 2.2, $\dim R((A - \lambda I)^{\ell}) = n - m$ and there exists a basis $\mathcal{B}$ of $\mathbb{C}^n$ such that with respect to $\mathcal{B}$, $A$ can be represented as

$$[A]_{\mathcal{B}} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

where $A_1 \in \mathbb{C}^{m \times m}$ and $A_2 \in \mathbb{C}^{(n-m) \times (n-m)}$. Moreover $\lambda$ is the only eigenvalue of $A_1$, and $\lambda$ is not an eigenvalue of $A_2$. It also follows from the above representation that $\det(A - zI) = \det(A_1 - zI_1) \det(A_2 - zI_2)$ where $I_1$ and $I_2$ are the identity matrices on $\mathbb{C}^{m \times m}$ and $\mathbb{C}^{(n-m) \times (n-m)}$ respectively. Thus, it follows that $m$ is the multiplicity of the $\lambda$ as the root of the characteristic polynomial of $A$. This number $m$ is called the *algebraic multiplicity* of $\lambda$, and the space $N((A - \lambda I)^{\ell})$ is called the generalized eigenspace of $A$ associated with the eigenvalue $\lambda$. The number $g := \dim N(A - \lambda I)$, the dimension of the eigenspace $N(A - \lambda I)$, is called the *geometric multiplicity* of $\lambda$. It is shown in Nair [5] (see also Limaye and Nair [2]) that

$$\ell + g - 1 \leq m \leq \ell g.$$
It can be seen that an eigenvalue is simple iff its algebraic multiplicity is 1. If $\lambda$ is a simple eigenvalue of $A$ with a corresponding eigenvector $x$, then we may say the pair $(\lambda, x)$ as simple eigen-pair.

In our analysis for obtaining an eigen-pair for a perturbed matrix $A + E$, we not only make use of an eigen-pair $(\lambda_0, x_0)$ of $A$ but also an eigenvector of $A^*$ associated with the eigenvalue $\bar{\lambda}_0$. It is easily seen that $\lambda_0$ is a simple eigenvalue of $A$ if and only if $\bar{\lambda}_0$ is a simple eigenvalue of $A^*$. We observe the following.

**THEOREM 2.3** Suppose $\lambda_0$ is a simple eigenvalue of $A$, and $x_0, y_0 \in \mathbb{C}^n$ are nonzero vectors such that $Ax_0 = \lambda_0 x_0$ and $A^* y_0 = \bar{\lambda}_0 y_0$. Then $\langle x_0, y_0 \rangle \neq 0$.

**Proof.** Since $\lambda_0$ is a simple eigenvalue of $A$, the spaces $N(A - \lambda_0 I)$ and $N(A^* - \bar{\lambda}_0 I)$ are one-dimensional. If $\langle x_0, y_0 \rangle = 0$, then $x_0 \in [N(A^* - \bar{\lambda}_0 I)]^\perp$. But $[N(A^* - \bar{\lambda}_0 I)]^\perp = R(A - \lambda_0 I)$. Thus $\langle x_0, y_0 \rangle = 0$ implies $x_0 \in N(A - \lambda_0 I) \cap R(A - \lambda_0 I)$. By Theorem 2.2, $N(A - \lambda_0 I) \cap R(A - \lambda_0 I) = \{0\}$. Thus we arrive at a contradiction to the assumption that $x_0 \neq 0$. $\blacksquare$

### 3 Eigen-Pairs of a Perturbed Matrix

Let $\lambda_0$ be a simple eigenvalue of $A$ and $x_0$ be a corresponding eigenvector such that $\|x_0\| = 1$. Suppose $y_0^* \in \mathbb{C}^n$ is an eigenvector of $A^*$ corresponding to the eigenvalue $\bar{\lambda}_0$. Then, by Theorem 2.3, $y_0^* x_0 \neq 0$. Without loss of generality we can assume that $y_0^* x_0 = 1$.

Now we consider the question of finding an eigen-pair $(\lambda, x)$ of a perturbed matrix

$$\tilde{A} := A + E,$$

which is close to $(\lambda_0, x_0)$ when $\|E\|$ is small. We shall, in fact, look for an eigenvector $x$ which also satisfies the normalization condition

$$y_0^* x = 1.$$

The following theorem characterizes such eigenvectors
THEOREM 3.1 There exists an eigenvector $x$ for $A + E$ such that $y_0^* x = 1$ if and only if there exists $u \in \{y_0\}^\perp$ such that

$$(A - \lambda_0 I) u = (\mu I - E)(x_0 + u),$$

with $\mu = y_0^* E x_0 + y_0^* Eu$, and in that case $x = x_0 + u$.

Proof. Suppose $u \in \mathbb{C}^n$ and $x := x_0 + u$. Then $\langle x, y_0 \rangle = 1$ if and only if $u \in \{y_0\}^\perp$.

Also, for $u \in \mathbb{C}^n$ and $\mu \in \mathbb{C}$, $x = x_0 + u$, and $\lambda = \lambda_0 + \mu$,

$$(A + E)x = \lambda x$$

if and only if

$$(A - \lambda_0 I) u = (\mu I - E)(x_0 + u).$$

Since $y_0^* (A - \lambda_0 I) = 0$, $y_0^* x_0 = 1$, and $y_0^* u = 0$ for $u \in \{y_0\}^\perp$, it follows that the relation $(A - \lambda_0 I) u = (\mu I - E)(x_0 + u)$ for $u \in \{y_0\}^\perp$ implies $\mu = y_0^* E x_0 + y_0^* Eu$. This completes the theorem. ■

In order to discuss the solvability of (3.1) for a $u \in \{y_0\}^\perp$, we would like to represent (3.1) in an alternate manner. For this purpose, let $Y_0 \in \mathbb{C}^{n \times (n-1)}$ be such that its columns are orthonormal and they span $\{y_0\}^\perp$. Then clearly, $Y_0^* Y_0 = I_0$, the identity matrix in $\mathbb{C}^{(n-1) \times (n-1)}$. Moreover,

$$\{y_0\}^\perp = [N(A^* - \bar{\lambda}_0)]^\perp = R(A - \lambda_0 I)$$

and for $u \in \{y_0\}^\perp$, there exists $v \in \mathbb{C}^{n-1}$ such that $u = Y_0 v$. Also for every $u \in \{y_0\}^\perp$,

$$Y_0 Y_0^* u = u.$$

Indeed, if $u \in \{y_0\}^\perp$, then there exists $v \in \mathbb{C}^{n-1}$ such that $u = Y_0 v$ so that $Y_0^* u = Y_0^* Y_0 v = v$ and $Y_0 Y_0^* u = Y_0 v = u$.

We write

$$L_0 := Y_0^* A Y_0.$$

THEOREM 3.2 The matrix $L_0 - \lambda_0 I_0$ is invertible, and for $u, v$ in $\{y_0\}^\perp$,

$$(A - \lambda_0 I) u = v \iff u = Y_0 (L_0 - \lambda_0 I_0)^{-1} Y_0^* v.$$
Proof. Note \( L_0 - \lambda_0 I_0 = Y_0^*(A - \lambda_0 I)Y_0 \). Now, suppose \( w \in \mathbb{C}^{n-1} \) is such that 
\((Y_0^*AY_0 - \lambda_0 I_0)w = 0\). Then we have \( Y_0^*(A - \lambda_0 I)Y_0 w = 0 \). Since 
\[ \{y_0\}^\perp = [N(A^* - \lambda_0)]^\perp = R(A - \lambda_0 I) \]
is invariant under \( A - \lambda_0 I \), it follows that \( z := (A - \lambda_0 I)Y_0 w \in \{y_0\}^\perp \), so that \( Y_0^*z = z \). Hence 
\[ z = Y_0Y_0^*z = Y_0^*(A - \lambda_0 I)Y_0 w = 0. \]
Since \( A - \lambda_0 I \) is injective on \( \{y_0\}^\perp \), \( z = (A - \lambda_0 I)Y_0 w \) implies \( Y_0 w = 0 \), and consequently, \( w = Y_0^*Y_0 w = 0 \). Thus \( L_0 - \lambda_0 I_0 \) is invertible.

Now for \( u, v \) in \( \{y_0\}^\perp \)
\[ (A - \lambda_0 I)u = v \iff (A - \lambda_0 I)Y_0Y_0^*u = v \]
\[ \iff Y_0^*(A - \lambda_0 I)Y_0Y_0^*u = v \]
\[ \iff Y_0(Y_0^*AY_0 - \lambda_0 I_0)Y_0^*u = v \]
\[ \iff (Y_0^*AY_0 - \lambda_0 I_0)Y_0^*u = Y_0^*v \]
\[ \iff Y_0^*u = (Y_0^*AY_0 - \lambda_0 I_0)^{-1}Y_0^*v \]
\[ \iff u = Y_0(Y_0^*AY_0 - \lambda_0 I_0)^{-1}Y_0^*v \]
This completes the proof of the theorem. ■

Let \( S := Y_0(L_0 - \lambda_0 I_0)^{-1}Y_0^* \). For \( u \in \{y_0\}^\perp \), let 
\[ F(u) := S[x_0y_0^*Ex_0 + x_0y_0^*Eu - Ex_0 - Eu + uy_0^*Ex_0 + uy_0^*Eu]. \]
(3.2)
Note that \( F(u) \in \{y_0\}^\perp \), as columns of \( Y_0 \) span the space \( \{y_0\}^\perp \) and range of \( S \) is contained in the range of \( Y_0 \).

**THEOREM 3.3** For \( u \in \{y_0\}^\perp \), let \( F(u) \) be defined as in (3.2). Then \( A + E \) has an eigenvector \( x \) satisfying \( y_0^*x = 1 \) if and only if \( F \) has a fixed point in \( \{y_0\}^\perp \).

**Proof.** For \( u \in \{y_0\}^\perp \), let \( \mu = y_0^*Ex_0 + y_0^*Eu \). Then 
\[ S(\mu I - E)(x_0 + u) = S[x_0y_0^*Ex_0 + x_0y_0^*Eu - Ex_0 - Eu + uy_0^*Ex_0 + uy_0^*Eu]. \]
Hence, taking \( v = (\mu I - E)(x_0 + u) \) in Theorem 3.2, it follows that equation (3.1) holds for \( u \in \{y_0\}^\perp \) if and only if \( u = F(u) \) holds. ■

7
By the above theorem, the problem of finding an eigenvector \( x \) of \( A + E \) satisfying \( y_0^* x = 1 \) is reduced that of finding a fixed point for the function \( F \) in some subset of \( \{y_0\}^\perp \). Our next task is to identify such a subset. For this purpose the following technical lemma would be of use. First some notations.

For positive real numbers \( a, b, c, \delta \), let
\[
\alpha := ab\delta, \quad \beta := a\delta(1 + 2b), \quad \gamma := a\delta(1 + b), \quad \varepsilon_0 := \frac{1}{a[1 + 2b + 2\sqrt{b(b + 1)}]}.
\]
(3.3)

We shall also make use of the function
\[
g(t) := \begin{cases} 
1, & \text{if } t = 0, \\
\frac{1 - \sqrt{1 - 4t}}{2t}, & \text{if } 0 < t < 1/4.
\end{cases}
\]

It is seen that \( 1 \leq g(t) \leq 2 \) for all \( t \in [0, 1/4] \). We observe that
\[
\delta < \varepsilon_0 \iff \beta < 1 \quad \text{and} \quad (1 - \beta) > 2\sqrt{\alpha\gamma}.
\]
Hence, the proof of the following lemma is obvious.

**LEMMA 3.4** If \( \delta < \varepsilon_0 \), then \( \beta < 1 \) and \( \omega := \frac{\alpha\gamma}{(1 - \beta)^2} < \frac{1}{4} \), and in that case,
\[
\alpha\rho^2 + \beta\rho + \gamma = \rho,
\]
where \( \rho := \frac{\gamma}{1 - \beta} g(\omega) \).

Hereafter, we take
\[
a := \| (L_0 - \lambda_0 I_0)^{-1} \|, \quad b := \| y_0^* \|, \quad \delta := \| E \|,
\]
(3.4)
and \( \alpha, \beta, \gamma, \delta \) and \( \varepsilon_0 \) are as in (3.3). We note that \( \| Y_0^* \|^2 = \| Y_0 \|^2 = \| Y_0^* Y_0 \| = \| I_0 \| = 1 \) so that
\[
\| S \| = \| Y_0 (L_0 - \lambda_0 I_0)^{-1} Y_0^* \| \leq a.
\]
For \( r > 0 \), let
\[
D_r := \{ u \in \{y_0\}^\perp : \| u \| \leq r \}.
\]
(3.5)
THEOREM 3.5 Let $\delta < \varepsilon_0$, and $\rho := \frac{\gamma}{1-\beta} g(\omega)$. Let $D_r$ and $F$ be as in (3.5) and (3.2) respectively. Then

(i) $F(u) \in D_{\rho}$ for all $u \in D_{\rho}$,

(ii) $F : D_{\rho} \rightarrow D_{\rho}$ is a contraction map.

Proof. Since $\delta < \varepsilon_0$, we know by Lemma 3.4 that $\beta < 1$, so that $\rho$ is well defined. Let $u \in D_{\rho}$. Then, since $\|S\| \leq a$ and

$$
\|F(u)\| \leq \|S[x_0y_0^*Ex_0 + x_0y_0^*Eu - Ex_0 - Eu + uy_0^*Ex_0 + uy_0^*Eu]\|
\leq a[b\delta + b\delta\rho + \delta + \delta\rho + \rho b\delta + \rho b\delta\rho]
\leq a\delta[b\rho^2 + (1 + 2b)\rho + (1 + b)]
= \alpha\rho^2 + \beta\rho + \gamma.
$$

Recall that $\rho$ is a zero of the equation $\alpha\rho^2 + \beta\rho + \gamma = \rho$. We already know that $F(u) \in \{y_0\}^\perp$ for all $u \in \{y_0\}^\perp$. Thus, $F(u) \in D_{\rho}$ for all $u \in D_{\rho}$, proving (i).

To prove (ii), let $u_1, u_2 \in D_{\rho}$. Then it is seen that

$$
F(u_1) - F(u_2) = S[x_0y_0^*E(u_1 - u_2) + (u_1 - u_2)y_0^*Ex_0 - E(u_1 - u_2)
+ (u_1 - u_2)y_0^*Eu + u_2y_0^*E(u_1 - u_2)].
$$

so that we have

$$
\|F(u_1) - F(u_2)\| \leq a\delta[b + b + 1 + b\rho + b\rho]\|u_1 - u_2\| = \kappa\|u_1 - u_2\|,
$$

(3.6)

where

$$
\kappa := a\delta[(1 + 2b) + 2b\rho] = \beta + 2\alpha\rho.
$$

But, $2\alpha\rho = (1 - \beta) - \sqrt{(1 - \beta)^2 - 4\alpha\gamma}$. Thus $\kappa = \beta + 2\alpha\rho < 1$ and hence $F : D_{\rho} \rightarrow D_{\rho}$ is a contraction map. ■

By Theorem 3.5, we can conclude that $F$ has a fixed point in $D_{\rho}$ whenever $\delta < \varepsilon_0$.

Now we are in a position to state the main theorem of this section.

THEOREM 3.6 For $u \in \{y_0\}^\perp$, let $F(u)$ be defined as in (3.2). If $\delta := \|E\| < \varepsilon_0$, then $A + E$ has an eigenvector $x$ satisfying $y_0^*x = 1$. Moreover,

$$
\|x - x_0\| \leq \frac{\gamma}{1-\beta} g(\omega) \leq \eta\|E\|,
$$

9
\[ |\lambda - \lambda_0| \leq b\|E\|(1 + \eta\|E\|), \]
\[ |\lambda - \lambda_1| \leq b\eta\|E\|^2, \]
where
\[ \eta = a\sqrt{\frac{b + 1}{b}}[1 + 2b + 2\sqrt{b(b + 1)}], \quad \lambda_1 = y_0^*(A + E)x_0. \]

**Proof.** Suppose \( \delta := \|E\| < \varepsilon_0 \). Then, by Theorems 3.3 and 3.5, we can assert that \( A + E \) has an eigenvector \( x \) satisfying \( y_0^*x = 1 \). Moreover,
\[ \|x - x_0\| \leq \frac{\gamma}{1 - \beta}g(\omega) \leq \frac{2\gamma}{1 - \beta}. \]
Now we observe that, since \( \delta < \varepsilon_0 \), we have \( a\delta < \frac{1}{1 + 2b + 2\sqrt{b(b + 1)}} \) so that
\[ 1 - \beta = 1 - (1 + 2b)a\delta > 1 - \frac{1 + 2b}{(1 + 2b) + 2\sqrt{b(b + 1)}} = \frac{2\sqrt{b(b + 1)}}{(1 + 2b) + 2\sqrt{b(b + 1)}}. \]
Hence, it follows that
\[ \|x - x_0\| \leq \frac{2\gamma}{1 - \beta} \leq \eta\delta = \eta\|E\|, \quad \text{(3.7)} \]
where \( \eta = a\sqrt{\frac{b + 1}{b}}[1 + 2b + 2\sqrt{b(b + 1)}] \).

We also know that the corresponding eigenvalue \( \lambda \) is given by
\[ \lambda = y_0^*(A + E)x = \lambda_0 + \mu, \quad \mu := y_0^*Ex = y_0^*Ex_0 + y_0^*E(x - x_0). \]
Hence,
\[ |\lambda - \lambda_0| \leq |y_0^*Ex_0| + |y_0^*E(x - x_0)| \leq b\|E\|(1 + \|x - x_0\|). \]
Using the estimate for \( \|x - x_0\| \) given in (3.7),
\[ |\lambda - \lambda_0| \leq b\|E\|(1 + \eta\|E\|). \]
Since \( y_0^*A(x - x_0) = y_0^*Au = \lambda_0y_0^*u = 0 \), we have
\[ \lambda = y_0^*(A + E)x = y_0^*(A + E)x_0 + y_0^*E(x - x_0). \]
Hence, taking
\[ \lambda_1 := y_0^*(A + E)x_0, \]
we have
\[ |\lambda - \lambda_1| = |y_0^*E(x - x_0)| \leq b\eta\|E\|^2. \]
This completes the proof. \( \blacksquare \)
4 Simplicity of a Perturbed Eigen-Pair

Now we establish that the eigenvalue $\lambda$ of $A + E$ associated with the eigenvector $x$ which satisfies $y_0^*x = 1$ is simple. The idea is to find matrices $P$ and $Q$ such that $PQ = I$ and

$$P(A + E)Q = \begin{pmatrix} \lambda & w^* \\ 0 & L \end{pmatrix}$$ (4.1)

for some $w \in \mathbb{C}^{n-1}$ and $L \in \mathbb{C}^{(n-1)\times(n-1)}$. Once this is done, it follows that $\lambda$ is a simple eigenvalue of $A + E$ if and only if $\lambda$ is not an eigenvalue of $L$. Thus, our task is to look for $P, Q, L$ satisfying the equation (3.4), and then show that $\lambda$ is not an eigenvalue of $L$. For this requirement, we take

$$P = \begin{pmatrix} y_0^* \\ Y_0^* - Y_0^*xY_0^* \end{pmatrix}, \quad Q = \begin{pmatrix} x & Y_0 \end{pmatrix}.$$

Then it can be verified that $PQ = I$ and

$$P(A + E)Q = \begin{pmatrix} \lambda & w^* \\ 0 & L \end{pmatrix},$$

with $w^* = y_0^*(A + E)Y_0$ and

$$L = L_0 + Y_0^*EY_0 - Y_0^*x_0y_0^*EY_0 + Y_0^*w_0^*EY_0.$$

Now $\lambda$ is a simple eigenvalue of $A + E$ if and only if $\lambda$ is not an eigenvalue of $L$, i.e. if and only if $L - \lambda I_0$ is injective. For establishing the injectivity of $L - \lambda I_0$ we shall make use of the following lemma.

**Lemma 4.1** If $C \in \mathbb{C}^{n\times n}$ is such that $\|C\| < 1$, then $I - C$ is injective.

**Proof.** We observe that for $v \in \mathbb{C}^n$,

$$\|(I - C)v\| \geq \|v\| - \|Cv\| \geq (1 - \|C\|)\|v\|.$$  

From this the result follows.  

**Theorem 4.2** Suppose $\|E\| < \varepsilon_0$. Then the eigenvalue $\lambda$ of $A + E$ corresponding to the eigenvector $x$ as in Theorem 3.6 is a simple eigenvalue.
Proof. We have to prove that $L - \lambda I_0$ is injective. Since
\[ L - \lambda I_0 = (L_0 - \lambda_0 I_0) - (\mu I_0 - M) = (L_0 - \lambda_0 I_0)(I_0 - (L_0 - \lambda_0 I_0)^{-1}(\mu I_0 - M)) \]
where
\[ M = Y_0^*EY_0 - Y_0^*x_0y_0^*EY_0 + Y_0^*uy_0^*EY_0, \quad \mu = y_0^*Ex_0 + y_0^*Eu \]
by Lemma 4.1 it is enough to prove that
\[ \|(L_0 - \lambda_0 I_0)^{-1}(\mu I_0 - M)\| < 1. \]
Since $\|u\| \leq \rho$, we see from (3.4) that
\[ \|(L_0 - \lambda_0 I_0)^{-1}(\mu I_0 - M)\| \leq a\delta[(1 + 2b) + 2b\rho] = \beta + 2\alpha\rho = \kappa < 1. \]
This completes the proof. ■

5 Iterative Refinements

Theorem 3.6 shows how much close the eigen-pair $(\lambda_0, x_0)$ of the unperturbed matrix $A$ to an eigen-pair $(\lambda, x)$ of the perturbed matrix $A + E$. Now, a natural query would be whether we can obtain better approximations to $(\lambda, x)$ than $(\lambda_0, x_0)$ without having to solve another eigenvalue problem. We show that it is possible by making use of the function $F$ defined in (3.2). For this, we define $u_1, u_2, \ldots$ iteratively as
\[ u_n := F(u_{n-1}), \quad n = 1, 2, \ldots \]
with $u_0 = 0$. Then let
\[ x_n := x_0 + u_n, \quad \lambda_n = \lambda_0 + y_0^*E(x_0 + u_{n-1}), \quad n = 1, 2, \ldots \]

The following theorem gives satisfactory answer to our query.

THEOREM 5.1 Let $\delta := \|E\| < \varepsilon_0$. Then the eigenvector $x := x_0 + u$ and the corresponding eigenvalue $\lambda$ satisfy the following relations:

(i) $\|x - x_n\| \leq \kappa^n \rho = O(\delta^{n+1})$.

(ii) $|\lambda - \lambda_n| \leq b\delta \kappa^{n-1} \rho \leq \kappa^n \rho = O(\delta^{n+1})$

In particular, $\lambda_n \to \lambda$ and $x_n \to x$ as $n \to \infty$. 

12
Proof. Recall from Theorems 3.5 and 3.6 that if $\delta < \varepsilon_0$, then there exists a unique $u \in \{y_0\}^\perp$ such that $u = F(u)$ where $F$ is as in (3.2), and $x := x_0 + u$ is an eigenvector of $A + E$ satisfying $y_0^*x = 1$. Hence, for $n = 0, 1, 2, \ldots$, it can be seen that
\[
\|x - x_{n+1}\| = \|u - u_{n+1}\| = \|F(u) - F(u_n)\| \leq \kappa \|u - u_n\|,
\]
where
\[
\kappa := \beta + 2\alpha \rho = \beta + (1 - \beta)(1 - \sqrt{1 - 4\omega}) < 1.
\]
Thus, for $n = 1, 2, \ldots$
\[
\|x - x_n\| \leq \kappa \|u - u_{n-1}\| \leq \kappa^n \|u\| \leq \kappa^n \rho.
\]
Now, consider
\[
|\lambda - \lambda_n| = |\lambda_0 + y_0^*E(x_0 + u) - \lambda_0 - y_0^*E(x_0 + u_{n-1})|
= |y_0^*E(u - u_{n-1})|
\leq \|y_0^*\| \|E\| \|u - u_{n-1}\|
\leq b\delta \kappa^{n-1} \rho.
\]
Clearly, since $\kappa < 1$, we have $\|x - x_n\| \rightarrow 0$ and $|\lambda - \lambda_n| \rightarrow 0$ as $n \rightarrow \infty$. ■

References


