On the construction of nearest defective matrices to a normal matrix

Rafikul Alam*
Department of Mathematics
Indian Institute of Technology Guwahati
Guwahati - 781 039, INDIA

Abstract. Let A be an n-by-n normal matrix with n distinct eigenvalues. We describe a simple procedure to construct a defective matrix whose distance from A, among all the defective matrices, is the shortest. Our construction requires only an appropriate pair of eigenvalues of A and their corresponding eigenvectors. We also show that the nearest defective matrices influence the evolution of spectral projections of A.

Keywords: Defective matrix, multiple eigenvalue, pseudospectra, eigenvector, spectral projection.

1 Introduction

Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix with n distinct eigenvalues. Let $d(A)$ be the distance of A from the set of defective matrices, that is,

$$d(A) := \inf\{\|A - A'\| : A' \text{ is defective}\}, \tag{1.1}$$

where $\|\cdot\|$ is either the 2-norm or the Frobenius norm. For a non-normal B, the determination of $d(B)$ and a defective matrix B' such that $\|B - B'\| = d(B)$ is a nontrivial task and is widely known as Wilkinson's problem (see, for example, [8], [9], [10], [6], [2], [5], [4], [1]). By contrast, since A is normal, it is easy to see that for the 2-norm

$$d(A) = \frac{1}{2}\min_{i\neq j} |\lambda_i - \lambda_j| \tag{1.2}$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. However, it appears that the construction of a defective matrix A' such that $\|A - A'\| = d(A)$ is nontrivial. Note that the existence of such an A' would imply that the infimum in (1.1) is actually the minimum. Since the set of defective matrices is not a closed set, it is not obvious that the infimum in (1.1) is, in fact, the minimum.

As an illustration, consider $A := \operatorname{diag}(1+\epsilon,\ 1-\epsilon)$, where $\epsilon > 0$. Then by (1.2) we have $d(A) = \epsilon$. The matrix $A' := \begin{bmatrix} 1+\epsilon/2 & -\epsilon/2 \\ \epsilon/2 & 1-\epsilon/2 \end{bmatrix}$ is defective and $\|A - A'\| = \epsilon = d(A)$. Note that the construction of A' is not obvious.

The main purpose of this note is to describe a construction of a defective matrix A' such that $\|A - A'\| = d(A)$. Our construction requires only a pair of eigenvalues λ_i and λ_j such

*Current address: Department of Mathematics, University of Manchester, Manchester M13 9PL, U.K., email: ralam@maths.man.ac.uk, rafikul@yahoo.com

that $|\lambda_i - \lambda_j| = 2d(A)$ and their corresponding eigenvectors. The constructed A' is a rank one perturbation of A. Consequently, the distance of A from the set of defective matrices is the same for the 2-norm and the Frobenius norm. We also derive a result that provides an understanding of structures of the nearest defective matrices to A. Finally, we briefly show that the nearest defective matrices influence the evolution of spectral projections of A.

Notation. For a complex n-by-n matrix A, we denote the spectrum of A by $\Lambda(A)$. The 2-norm of A, denoted by $\|A\|_2$, is defined by $\|A\|_2 := \max_{\|x\|=1} \|Ax\|_2$, where $\|x\|_2 := (\sum_{j=1}^n |x_j|^2)^{1/2}$. The Frobenius norm of A, denoted by $\|A\|_F$, is defined by $\|A\|_2 := \sqrt{\mathrm{trace}(A^*A)}$.

2 Construction of nearest defective matrices

Recall that $A \in \mathbb{C}^{n\times n}$ is a normal matrix with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Let $\mathcal{D}$ denote the set of matrices having a multiple eigenvalue. Then $\mathcal{D}$ is a closed set and the set of defective matrices is dense in $\mathcal{D}$. Hence it follows that

$$d(A) := \inf\{\|A - A'\| : A' \text{ is defective}\} = \min\{\|A - B\| : B \in \mathcal{D}\}. \tag{2.1}$$

The ϵ-pseudospectra [7] of matrices provide a convenient setting in which (1.1) can be solved. For $\epsilon > 0$, we consider the closed ball $\mathbf{A}(\epsilon) := \{A' : \|A - A'\| \leq \epsilon\}$ and the ϵ-pseudospectrum

$$\Lambda_\epsilon(A) := \bigcup_{A' \in \mathbf{A}(\epsilon)} \Lambda(A'). \tag{2.2}$$

It is easy to see that $\Lambda_\epsilon(A)$ is the same for the 2-norm and the Frobenius norm. We mention that for the Frobenius norm, (2.2) is not equivalent to the more widely used definition

$$\Lambda_\epsilon(A) := \{z \in \mathbb{C} : \|(A - zI)^{-1}\|_F \geq \epsilon^{-1}\},$$

where it is assumed that $\|(A - zI)^{-1}\|_F = \infty$ whenever $z \in \Lambda(A)$ [7].

From (2.1) it follows that $d(A)$ is the radius of the largest open ball centred at A whose interior consists of simple matrices but the boundary contains a matrix having a multiple eigenvalue. Hence $d(A) = \epsilon$ if and only if the interior of $\mathbf{A}(\epsilon)$ consists of simple matrices but $\partial\mathbf{A}(\epsilon) \cap \mathcal{D} \neq \emptyset$, where $\partial\mathbf{A}(\epsilon)$ is the boundary of $\mathbf{A}(\epsilon)$. Thus if $\epsilon := d(A)$ then at least two eigenvalues of A must move and coalesce when A varies in $\mathbf{A}(\epsilon)$. Since A is normal, we have

$$\Lambda_\epsilon(A) = \bigcup_{j=1}^n B[\lambda_j, \epsilon] \tag{2.3}$$

where $B[\lambda_j, \epsilon] := \{z : |\lambda_j - z| \leq \epsilon\}$. Hence if two eigenvalues of A coalesce when A varies in $\mathbf{A}(\epsilon)$ then at least two disks in (2.3) must coalesce, that is, have a common tangent. The smallest value of ϵ for which at least two disks coalesce is given by $\epsilon = \min_{i\neq j} |\lambda_i - \lambda_j|/2$. Define the **gap** in the spectrum $\Lambda(A)$ by

$$\mathrm{gap}(A) := \min_{i\neq j} |\lambda_i - \lambda_j|.$$

Note that $\Lambda_\epsilon(A)$ consists of n disjoint disks if and only if $\epsilon < \mathrm{gap}(A)/2$. Further, if λ_i and λ_j are such that $|\lambda_i - \lambda_j| = \mathrm{gap}(A)$ then the disks $B[\lambda_i, \epsilon]$ and $B[\lambda_j, \epsilon]$ coalesce at $(\lambda_i + \lambda_j)/2$ for $\epsilon = \mathrm{gap}(A)/2$. This immediately shows that $d(A) \geq \mathrm{gap}(A)/2$.

Note that $d(A)$ is invariant under unitary similarity transformations of A. Hence we may assume that A is diagonal because A is normal. However, this assumption neither reduces the complexity of the description of our construction process nor simplifies its proof. Hence we present the construction in a most general form. Let $d_2(A)$ and $d_F(A)$ denote $d(A)$ with respect to the 2-norm and the Frobenius norm, respectively. The following result identifies two nearest defective matrices to A.

Theorem 2.1 *Let A be normal with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ and the corresponding unit eigenvectors $x_1, \ldots, x_n$, respectively. Let λ_i and λ_j be such that $|\lambda_i - \lambda_j| = \text{gap}(A)$. Define*

$$A' := A - \frac{(\lambda_i - \lambda_j)}{4}(x_i - x_j)(x_i + x_j)^* \text{ and } A'' := A - \frac{(\lambda_i - \lambda_j)}{4}(x_i + x_j)(x_i - x_j)^*.$$

Then A' and A'' are defective, and $\|A - A'\| = \|A - A''\| = \text{gap}(A)/2 = d_2(A) = d_F(A)$, where the norm is either the 2-norm or the Frobenius norm.

Proof: First, note that $\|A - A'\|_2 = \|A - A'\|_F = \text{gap}(A)/2$. The same is true for A''. Recall that $d(A) \geq \text{gap}(A)/2$. Hence to prove the result, it is enough to show that both A' and A'' are defective. Let $z_0 := (\lambda_i + \lambda_j)/2$. We show that z_0 is a defective eigenvalue of A' and A''.

Set $x := (x_i + x_j)/\sqrt{2}$ and $y := (x_i - x_j)/\sqrt{2}$. Then it follows that y and x are unit left and right eigenvectors of A' corresponding to z_0. Indeed,

$$A'x = \frac{1}{\sqrt{2}}\left((\lambda_i x_i + \lambda_j x_j) - \frac{(\lambda_i - \lambda_j)}{2}(x_i - x_j)\right) = z_0 x.$$

Similarly, $y^* A' = z_0 y^*$. It is well known that an eigenvalue is multiple if and only if it has a pair of left and right eigenvectors which are orthogonal ([10], pp.11). Since $y^* x = 0$, it follows that z_0 is a multiple eigenvalue of A'. Hence $d_2(A) = d_F(A) = \|A - A'\| = \text{gap}(A)/2$.

Next, we show that z_0 is defective and has algebraic multiplicity 2. Note that the unit eigenvector x_k of A corresponding to λ_k is orthogonal to both x and y for all $k \neq i, j$. Hence it follows that $A' x_k = \lambda_k x_k$ for all $k \neq i, j$. In other words, the $n-2$ simple eigenpairs (λ_k, x_k), for all $k \neq i, j$, of A are also simple eigenpairs of A'. This shows that the algebraic multiplicity of z_0 is 2. Next, we have

$$A'y = \frac{(\lambda_i x_i - \lambda_j x_j)}{\sqrt{2}} = \frac{(\lambda_i - \lambda_j)}{2} x + z_0 y.$$

Consider the unitary matrix $U := [x, y, X]$, where the columns of X are the eigenvectors x_k for all $k \neq i, j$. Then it follows that $U^* A' U = \text{diag}(A_1, D)$, where $A_1 := \begin{bmatrix} z_0 & (\lambda_i - \lambda_j)/2 \\ 0 & z_0 \end{bmatrix}$ and $D := \text{diag}(\lambda_k)$ for all $k \neq i, j$. This shows that z_0 is defective.

Similarly, A'' is defective and $U^* A'' U = \text{diag}(A_2, D)$, where $A_2 := \begin{bmatrix} z_0 & 0 \\ (\lambda_i - \lambda_j)/2 & z_0 \end{bmatrix}$. This completes the proof. ■

It follows from the above proof that A' is unitarily similar to A''. Observe that for each eigenvalue pair (λ_i, λ_j) such that $|\lambda_i - \lambda_j| = \text{gap}(A)$, there are at least two defective matrices at which the infimum in (1.1) is attained. Therefore, if A has m such pairs then there are at least $2m$ defective matrices at which the infimum in (1.1) is attained. Further, the proof shows that each of these nearest defective matrices has exactly one elementary quadratic divisor, that is, its Jordan canonical form contains only one 2-by-2 Jordan block. For example, if $A := \mathrm{diag}(1, -1, i, -i)$ then $d(A) = 0.7$ and there are 4 such eigenvalue pairs. Hence there are at least 8 defective matrices nearest to A. Finally, note that λ_i and λ_j, and their corresponding eigenvectors completely determine the smallest perturbations that deform A to defective matrices. It is therefore not surprising that under such a perturbation the pair (λ_i, λ_j) is deformed to a defective eigenvalue and the rest of the eigenstructure of A remains intact.

Next, we show that the nearest defective matrices are necessarily of the form as described in Theorem 2.1. To that end, we rewrite these matrices as follows. Set

$$v_i := \frac{x_i + x_j}{\sqrt{2}}\mathrm{sign}(\lambda_i - \lambda_j),\ \ v_j := \frac{x_i - x_j}{\sqrt{2}}\mathrm{sign}(\lambda_i - \lambda_j),\ \ u_i := \frac{x_i - x_j}{\sqrt{2}},\ \ u_j := \frac{x_i + x_j}{\sqrt{2}},$$

where $\mathrm{sign}(t) = \bar{t}/|t|$. Then the defective matrices A' and A'' can be rewritten as

$$A' = A - d(A)u_i v_i^* \text{ and } A'' = A - d(A)u_j v_j^*.$$

It is easy to see that u_k and v_k are unit left and right singular vectors of $A - \frac{1}{2}(\lambda_i + \lambda_j)I$ corresponding to the smallest singular value $d(A)$ and that $v_k^* u_k = 0$, for $k = i, j$. The following result which is, in a sense, the converse of Theorem 2.1 provides an understanding of structures of the nearest defective matrices to A.

Theorem 2.2 *Suppose that* $\|A - A'\|_2 = \|A - A'\|_F = d(A)$. *If* A' *has a multiple eigenvalue* z_0 *then the following holds.*

(i) z_0 *is a defective eigenvalue of algebraic multiplicity* 2 *and there is a pair of eigenvalues* λ_i *and* λ_j *of* A *such that* $z_0 := (\lambda_i + \lambda_j)/2$.

(ii) A' *is of the form* $A' = A - d(A)uv^*$, *where* u *and* v *are unit left and right singular vectors of* $A - z_0 I$ *corresponding to the smallest singular value* $d(A)$ *such that* $u^*v = 0$. *Also,* u *and* v *are unit left and right eigenvectors of* A' *corresponding to* z_0.

Proof: Set $\epsilon := d(A)$. Then in view of (2.3), $\Lambda_\epsilon(A)$ consists of n disks $B[\lambda_j, \epsilon]$, $j = 1, 2, \ldots, n$, of which at least two have a common tangent. Since z_0 is a multiple eigenvalue of A', z_0 is the common boundary point of two disks, say, $B[\lambda_i, \epsilon]$ and $B[\lambda_j, \epsilon]$. This shows that $|\lambda_i - \lambda_j| = \mathrm{gap}(A) = 2d(A)$ and that $z_0 = (\lambda_i + \lambda_j)/2$. It is also clear that the algebraic multiplicity of z_0 is 2.

Evidently, $A - A'$ is a rank one matrix. Hence $A' = A - \epsilon uv^*$ for some unit vectors u and v. We show that u and v are left and right eigenvectors of A'. Let y and x be unit left and right eigenvectors of A' corresponding to z_0. Since $(A' - z_0 I)x = (A - z_0 I)x - \epsilon uv^*x = 0$, taking norm we have $\epsilon \leq \|(A - z_0 I)x\|_2 = \epsilon|v^*x|$. This shows that $|v^*x| = 1$, that is, $x = \alpha v$ for some $|\alpha| = 1$. Similarly, $y^*(A' - z_0 I) = y^*(A - z_0 I) - \epsilon(y^*u)v^* = 0$, shows that $y = \beta u$ for some $|\beta| = 1$. Hence u and v are unit left and right eigenvectors of A' corresponding to z_0, and that the geometric multiplicity of z_0 is one. Clearly u and v are left and right singular vector of $A - z_0 I$ corresponding to ϵ. Since z_0 is nonderogatory, the left and right eigenvectors of A' corresponding to z_0 are orthogonal. Hence $v^*u = 0$. This completes the proof. ■

Note that the singular vectors u and v in Theorem 2.2 can be taken, for example, to be $u = u_i$ and $v = v_i$, where u_i and v_i are given above. A few remarks are as follows.

1. It is clear that the smallest perturbation required to coalesce three eigenvalues of A is strictly bigger than $d(A)$. In other words, the distance of A from the nearest matrix having a triple eigenvalue is strictly bigger than $d(A)$. This follows from the fact that three disks cannot coalesce, that is, have a common tangent, at a single point.

2. Set
$$\delta_i := \mathrm{dist}(\lambda_i, \Lambda(A) \setminus \{\lambda_i\}).$$
Let λ_i and λ_j be such that $|\lambda_i - \lambda_j| = \delta_i$. Then, in the same way, it can be shown that $\delta_i/2$ is the magnitude of smallest perturbation of A required to coalesce λ_i with λ_j at $(\lambda_i + \lambda_j)/2$. In other words, the radius of the largest open ball centred at A on which the spectral projection P_i associated with A and λ_i is continuous is given by $\delta_i/2$. The nearest defective matrices to A where P_i is discontinuous may be constructed as outlined in Theorem 2.1.

We have seen that A is close to a defective matrix if and only if it has a pair of close eigenvalues. Thus having close eigenvalues is a manifestation of close to being defective and vice-versa. How do the nearest defective matrices affect the spectral evolution of A? As we shall see now, they strongly influence the evolution of the spectral projections but their influence on the spectral variation of A seems to be quite subtle. In a sense, the nearest defective matrices represent singularity for the eigenproblem $Au = \lambda u$.

Observe that $\mathrm{gap}(A) = \min_i \delta_i$. As mentioned above, $\delta_i/2$ is the radius of the largest open ball centred at A on which P_i is continuous. In fact, more is true. For each $\|E\|_2 = 1$, $A + tE$ has a simple eigenvalue $\lambda_i(t)$ near λ_i for all $|t| < \delta_i/2$. Further, the spectral projection $P_i(t)$ associated with $A + tE$ and $\lambda_i(t)$ is analytic and

$$P_i(t) = P_i + (-P_i E S_i - S_i E P_i)t + \mathcal{O}(t^2)$$

for all $|t| < \delta_i/2$, where S_i is the reduced resolvent of $A - \lambda_i I$ (see, [3], pp.75-77). Since A is normal, we have $\|S\|_2 = 1/\delta_i$ consequently

$$\|P_i(t) - P_i\|_2 \leq \frac{|t|}{\delta_i/2} + \mathcal{O}(|t|^2).$$

Now, if λ_j is such that $|\lambda_i - \lambda_j| = \delta_i$ then it is easy to see that

$$A' := A - \frac{(\lambda_i - \lambda_j)}{4}(x_i - x_j)(x_i + x_j)^*$$

is defective and $\|A - A'\|_2 = \|A - A'\|_F = \delta_i/2$, where x_i and x_j are unit eigenvectors of A corresponding to λ_i and λ_j, respectively. Taking $E := -\frac{1}{2}(x_i - x_j)(x_i + x_j)^*$, it follows that $\|E\|_2 = \|E\|_F = 1$ and, that $P(t)$ is discontinuous at $t := (\lambda_i - \lambda_j)/2$ and $|t| = \delta_i/2$. This shows that $A' \in \partial \mathbf{A}(\epsilon)$ for $\epsilon := \delta_i/2$, and P_i is discontinuous at A'; the rank of P_i increases from one to two at A'. Hence we have the following result.

Theorem 2.3 *The spectral projection P_i associated with A and λ_i is continuous on $\mathbf{A}(\epsilon)$ if and only if $\epsilon < \delta_i/2$. Let A' be a matrix such that $\|A - A'\|_2 < \delta_i/2$. Then A' has a simple eigenvalue λ_i' near λ_i. Let P_i' denote the spectral projection associated with A' and λ_i'. Then*

$$\|P_i - P_i'\|_2 \leq \frac{\|A - A'\|_2}{\delta_i/2} + \mathcal{O}(\|A - A'\|_2^2).$$

Since $d(A) = \min_i \delta_i/2$, the following result is immediate.

Corollary 2.1 *Each spectral projection P_i is continuous on $\mathbf{A}(\epsilon)$ if and only if $\epsilon < d(A)$. If $\|A - A'\|_2 < d(A)$ then*

$$\|P_i - P_i'\|_2 \leq \frac{\|A - A'\|_2}{d(A)} + \mathcal{O}(\|A - A'\|_2^2)$$

for all $i = 1, 2, \ldots, n$.

In view of the above results, we conclude that $1/d(A)$ is a measure of ill-conditioning of the eigenvalue problem $Au = \lambda u$.

Acknowledgement. The author thanks Professor Nick Higham for reading the manuscript and making many suggestions which have significantly improved the presentation.

References

[1] R. Alam and S. Bora, On sensitivity of eigenvalues and eigendecompositions of matrices, submitted to *Linear Algebra & Appl.*

[2] J. W. Demmel, On condition numbers and the distance to the nearest ill-posed problem, *Numer. Math.*, 51(1987), pp.251-289.

[3] T. Kato, *Perturbation Theory for Linear Operators*, Springer, New York, 1976.

[4] R. A. Lippert and A. Edelman, The computation and sensitivity of double eigenvalues, *Advances in Computational Mathematics: Proc. Gaungzhou International Symposium,* Z. Chen, Y. Li, C. A. Micchelli and Y. Xu, eds., pp.353-393, Dekker, New York, 1999.

[5] A. N. Malyshev, A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues, *Numer. Math.,* 83(1999), pp.443-454.

[6] A. Ruhe, Properties of a matrix with a very ill-conditioned eigenproblem, *Numer. Math.,* 15(1970), pp.57-60.

[7] L. N. Trefethen, Pseudospectra of matrices, in D. F. Griffiths and G. A. Watson, Numerical Analysis 1991, Longman, (1992), pp.234-266.

[8] J. H. Wilkinson, Notes on matrices with a very ill-conditioned eigenproblem, *Numer. Math.,* 19(1972), pp.175-178.

[9] J. H. Wilkinson, On neighbouring matrices with quadratic elementary divisors, *Numer. Math.,* 44(1984), pp.1-21.

[10] J. H. Wilkinson, Sensitivity of eigenvalues, *Utilitas Mathematica,* 25(1984), pp.5-76.