COMPUTATION OF ANTIEIGENVALUES

MORTEZA SEDDIGHIN

We will use a theorem by M. G. Kreĭn and properties of the numerical range of an operator to develop a method of computation for the first antieigenvalue of a strictly accretive operator.

1. Introduction

An operator on a Hilbert space is called accretive if $\mathrm{Re}(Tf, f) \geq 0$ and strictly accretive if $\mathrm{Re}(Tf, f) > 0$, for every vector f. For an accretive operator or matrix T on a Hilbert space, the first antieigenvalue of T, denoted by $\mu_1(T)$, is defined by Gustafson to be

$$\mu_1(T) = \inf_{Tf \neq 0} \frac{\mathrm{Re}(Tf, f)}{\|Tf\| \|f\|} \tag{1.1}$$

(see [3, 4, 5]). The quantity $\mu_1(T)$ is also denoted by $\cos T$ and is called the cosine of T. Definition (1.1) is equivalent to

$$\mu_1(T) = \inf_{\substack{Tf \neq 0 \\ \|f\|=1}} \frac{\mathrm{Re}(Tf, f)}{\|Tf\|}. \tag{1.2}$$

$\mu_1(T)$ measures the maximum turning capability of T. A vector f for which the inf in (1.1) is attained is called an antieigenvector of T. Cosine of T has important applications in pure operator theory as well as numerical analysis (see [3, 4, 6]). In [7, 8, 11], we have computed the first antieigenvalue of normal accretive operators in terms of eigenvalues of these operators. In [12], we computed the first antieigenvalue for certain classes of matrices on the real field. In this note, we will focus on computation and approximation of the first antieigenvalue of general accretive operators. Our work in this note is a generalization of our work in [7, 8, 11] to certain classes of nonnormal operators. We will start with the case of a general two-by-two matrix.

2. Computing $\mu_1(T)$ for an accertive two-by-two matrix

THEOREM 2.1. *Let T be a strictly accretive two-by-two matrix and let $S = \operatorname{Re} T + iT^*T$. There exist the following two cases. If S has two distinct eigenvalues $\alpha = \alpha_1 + i\alpha_2$, $\beta = \beta_1 + i\beta_2$ and the corresponding eigenvectors, so normalized that $\|f\| = \|g\| = 1$, then $\mu_1(T) = \sqrt{x_0^2/y_0}$, where (x_0, y_0) is the point where a member of the family of convex functions $y = x^2/k$ touches just one point of the ellipse whose equation is*

$$
\left(\gamma^2 \frac{(\beta_2 - \alpha_2)^2}{|\alpha - \beta|^2} + \frac{(\beta_1 - \alpha_1)^2}{|\alpha - \beta|^2} \right) \left(x - \frac{1}{2}\alpha_1 - \frac{1}{2}\beta_1 \right)^2
$$
$$
+ 2(\gamma^2 - 1) \frac{\beta_1 - \alpha_1}{|\alpha - \beta|^2} (\beta_2 - \alpha_2) \left(x - \frac{1}{2}\alpha_1 - \frac{1}{2}\beta_1 \right) \left(y - \frac{1}{2}\alpha_2 - \frac{1}{2}\beta_2 \right) \tag{2.1}
$$
$$
+ \left(\gamma^2 \frac{(\beta_1 - \alpha_1)^2}{|\alpha - \beta|^2} + \frac{(\beta_2 - \alpha_2)^2}{|\alpha - \beta|^2} \right) \left(y - \frac{1}{2}\alpha_2 - \frac{1}{2}\beta_2 \right)^2 - \frac{1}{4}\gamma^2 \frac{|\alpha - \beta|^2}{\delta^2} = 0,
$$

where $\gamma = \|(f,g)\|$ and $\delta = \sqrt{1 - \gamma^2}$. If S has only one eigenvalue $\alpha = \alpha_1 + \alpha_2$, then $\mu_1(T) = \sqrt{x_0^2/y_0}$, where (x_0, y_0) is the point where a member of the family of convex functions $y = x^2/k$ touches just one point of the circle whose equation is

$$
(x - \alpha_1)^2 + (y - \alpha_2)^2 = \frac{\|S - \alpha\|^2}{4}. \tag{2.2}
$$

Proof. Let $\Omega(S)$ be the numerical range of S. By Kreĭn [10], we have

$$
\mu_1^2(T) = \inf \left\{ \frac{x^2}{y} : x + iy \in \Omega(S) \right\}. \tag{2.3}
$$

On the other hand, by Halmos [9], if α and β are distinct, then $\Omega(S)$ is the closed elliptical disk with foci at α and β whose major axis is $|\alpha - \beta|/\delta$ and minor axis is $\gamma(|\alpha - \beta|/\delta)$. If S has only one eigenvalue α, then $\Omega(S)$ is the circular disk with center at α and radius $\|S - \alpha\|/2$. Using the convexity property of the curves $f(x, y) = x^2/y = k$, we conclude that the minimum of the function $f(x, y) = x^2/y$ on $\Omega(S)$ is exactly at the point where a member of the family of convex functions $y = x^2/k$ touches one point of $\Omega(S)$. This point of contact must be on the boundary of $\Omega(S)$, which is an ellipse or a circle. We denote the boundary of $\Omega(S)$ by $\partial\Omega(S)$. First, assume that S has two distinct eigenvalues $\alpha = \alpha_1 + i\alpha_2$ and $\beta = \beta_1 + \beta_2$. Let $d = |\alpha - \beta|$. If we select the system of coordinate axis x'-y' so that x'-axis coincides with the major axis of $\Omega(S)$, then the equation of $\partial\Omega(S)$ with respect to x'-y' system is

$$
\frac{(x')^2}{(d/2\delta)^2} + \frac{(y')^2}{(\gamma d/2\delta)^2} = 1. \tag{2.4}
$$

To obtain the equation of this ellipse with respect to x-y axis, we need to perform a rotation and a translation of axis. This will give us (2.1) which represents the equation of $\partial\Omega(S)$ with respect to x-y axis. Next, assume that $\alpha = \alpha_1 + \alpha_2$ is the only eigenvalue of S, then the equation of $\partial\Omega(S)$ is (2.2). □

Example 2.2. Let $T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, then $S = \begin{bmatrix} 1+2i & 0.5+i \\ 0.5+i & 1+i \end{bmatrix}$. The matrix S has two distinct eigenvalues. Numerical approximation of these eigenvalues and their corresponding normalized eigenvectors gives $\alpha = 1.4551 + 2.5987i$ and $\beta = 0.54491 + 0.40132i$ with corresponding normalized eigenvectors

$$f = (0.22718 + 0.79806i, 0.2496 + 0.4992i),$$
$$g = (0.39816 - 0.39106i, -0.68941 + 0.46169i),$$

(2.5)

respectively. Using the notations of the previous theorem, we have $d = |\alpha - \beta| = 2.3784$, $\gamma = \|(f,g)\| = 0.17156$, and $\delta = \sqrt{1 - \gamma^2} = 0.98517$. Hence $d/2\delta = 2.3784/(2(0.98517)) = 1.2071$, $(d/2\delta)^2 = 1.4571$, $\gamma d/2\delta = 0.20709$, and $(\gamma d/2\delta)^2 = 4.2886 \times 10^{-2}$. Therefore the equation of $\partial\Omega(S)$ with respect to x'-y' coordinate system is $(x')^2/1.4571 + (y')^2/(4.2886 \times 10^{-2}) - 1 = 0$. Note that the center of this ellipse (with respect to x-y coordinate system) is at $((\alpha_1 + \beta_1)/2, (\alpha_2 + \beta_2)/2) = (1, 1.5)$. Also note that if θ is the angle between x' and x axis, then $\cos\theta = (\alpha_1 - \beta_1)/d = 0.38269$ and $\sin\theta = (\alpha_2 - \beta_2)/d = 0.92389$. Therefore, the equation of $\partial\Omega(S)$ with respect to x-y axis is

$$\frac{(0.38269(x-1) - 0.92389(y-1.5))^2}{1.4571} + \frac{(0.92389(x-1) + 0.38269(y-1.5))^2}{4.2886 \times 10^{-2}} - 1 = 0$$

(2.6)

which further simplifies to

$$1.25x^2 - 4.0001x + 1.0xy + 3.2501 - 1.75y + 0.25y^2 = 0.$$

(2.7)

Now to find infimum of the function $f(x,y) = x^2/y$ on $\partial\Omega(S)$, we use the Lagrange multiplier method. The resulting equations are

$$1.25x^2 - 4.0001x + 1.0xy + 3.2501 - 1.75y + 0.25y^2 = 0,$$
$$\frac{2x}{y} = \lambda \frac{\partial(1.25x^2 - 4.0001x + 1.0xy + 3.2501 - 1.75y + 0.25y^2)}{\partial x},$$
$$\frac{-x^2}{y^2} = \lambda \frac{\partial(1.25x^2 - 4.0001x + 1.0xy + 3.2501 - 1.75y + 0.25y^2)}{\partial y}.$$

(2.8)

The system (2.8) is solved numerically and the solution is $\lambda = -1.9019$, $x = 0.5019$, and $y = 2.5374$. Therefore $\mu_1(T) = \sqrt{0.5019^2/2.5374} = 0.31508$.

In [7, 8, 11], we used (1.2) to first compute antieigenvectors, and then from there compute antieigenvalues for normal matrices and normal operators. Also in [12], we used (1.2) to first compute antieigenvectors, and then from there compute antieigenvalues for arbitrary matrices on the field of real numbers. However, it is very difficult to use (1.2) to compute antieigenvectors in the case of an arbitrary matrix or operator on the field of complex numbers. To see this, consider the accertive matrix $T = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and let $f = \frac{z_1}{z_2}$ represent a unit vector. Assume $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$. Direct computations show

that the problem of finding $\inf_{\|Tf\|\neq 0}(\mathrm{Re}(Tf,f)/\|Tf\|)$, subject to the constraint $\|f\| = 1$, is equivalent to the problem of finding the infimum of the function

$$f(x_1,y_1,x_2,y_2) = \frac{x_2 x_1 + x_2^2 + y_2 y_1 + y_2^2 + x_1^2 + y_1^2}{\sqrt{(2x_1^2 + 2y_1^2 + 2x_2 x_1 + x_2^2 + 2y_2 y_1 + y_2^2)}} \tag{2.9}$$

subject to the constraint

$$x_1^2 + y_1^2 + x_2^2 + y_2^2 = 1. \tag{2.10}$$

If one wishes to solve this optimization problem using Lagrange multiplier method, then the following nonlinear system consisting of five equations and five unknowns will result

$$(x_2 + 2x_1)\frac{x_1^2 + y_1^2 + x_2 x_1 + y_2 y_1}{\left(\sqrt{(2x_1^2 + 2y_1^2 + 2x_2 x_1 + x_2^2 + 2y_2 y_1 + y_2^2)}\right)^3} = 2\lambda x_1,$$

$$(y_2 + 2y_1)\frac{x_1^2 + y_1^2 + x_2 x_1 + y_2 y_1}{\left(\sqrt{(2x_1^2 + 2y_1^2 + 2x_2 x_1 + x_2^2 + 2y_2 y_1 + y_2^2)}\right)^3} = 2\lambda y_1,$$

$$\frac{x_1^3 + y_1^2 x_1 + 4x_2 x_1^2 + 3x_2^2 x_1 + y_2 y_1 x_1 + 3y_1^2 x_2 + x_2^3 + 3y_2 y_1 x_2 + y_2^2 x_2}{\left(\sqrt{(2x_1^2 + 2y_1^2 + 2x_2 x_1 + x_2^2 + 2y_2 y_1 + y_2^2)}\right)^3} = 2\lambda x_2, \tag{2.11}$$

$$\frac{x_1^2 y_1 + y_1^3 + x_2 x_1 y_1 + 4y_2 y_1^2 + 3y_2^2 y_1 + 3x_1^2 y_2 + 3x_2 x_1 y_2 + x_2^2 y_2 + y_2^3}{\left(\sqrt{(2x_1^2 + 2y_1^2 + 2x_2 x_1 + x_2^2 + 2y_2 y_1 + y_2^2)}\right)^3} = 2\lambda y_2,$$

$$x_1^2 + y_1^2 + x_2^2 + y_2^2 = 1.$$

Notice that three equations and three unknowns resulted when Lagrange multiplier method was applied in Example 2.2 above to compute the antieigenvalue for the same matrix.

In computing $\mu_1(T)$ for a two-by-two matrix, we took advantage of the fact that the minimum of the convex function $f(x,y) = x^2/y$ on the convex set $\Omega(S)$ is attained on the boundary of $\Omega(S)$, where $\Omega(S)$ is the numerical range of the operator $S = \mathrm{Re}\,T + iT^*T$. The same argument can also be made in the case of a general matrix or operator. However, If T is not a two-by-two matrix, then the boundary of $\Omega(S)$ is no longer an ellipse. Indeed, the boundary of the numerical range of a finite-dimensional matrix can be very complicated. We can generalize the technique of computing $\mu_1(T)$ for two-dimensional matrices to matrices which are direct sums of two-by-two matrices. For instance, suppose that $T = T_1 \oplus T_2$, where T_1 and T_2 are accretive two-by-two matrices. If as before we define $S = \mathrm{Re}\,T + iT^*T$, then $S = S_1 \oplus S_2$ where $S_1 = \mathrm{Re}\,T_1 + iT_1^*T_1$ and $S_2 = \mathrm{Re}\,T_2 + iT_2^*T_2$. Hence, by Halmos [9], we have $\Omega(S) = \mathrm{co}(\Omega(S_1),\Omega(S_2))$, where $\mathrm{co}(\Omega(S_1),\Omega(S_2))$ denotes the convex hull of the numerical ranges of S_1 and S_2. Hence, the boundary of $\Omega(S)$ consists of a finite number of possible elliptic, circular, and flat (straight) arcs. The minimum of $f(x,y) = x^2/y$ on each arc can be obtained using Lagrange multiplier method as demonstrated in Example 2.2 above. Therefore, $\mu_1^2(T)$ is the smallest of minimum values of $f(x,y) = x^2/y$ on those arcs. This can be further generalized to the case when T is a direct sum of a finite number of two-by-two accretive matrices $T_1, T_2, T_3, \ldots, T_n$.

3. The general case

To apply these techniques to a general operator T, we need to understand the boundary of the numerical range of the operator S where $S = \operatorname{Re} T + iT^*T$. If $\Omega(S)$ is closed, then its boundary $\partial\Omega(S)$ consists of a countable set of differentiable arcs plus a countable set of points at which the boundary is not differentiable. Furthermore, the points at which the boundary is not differentiable are eigenvalues of S (see [2]). This leads us to the following theorem.

Theorem 3.1. *Let T be a strictly accretive operator and $S = \operatorname{Re} T + iT^*T$. Assume that $\Omega(S)$ is closed and $\partial\Omega(S) = M \cup N$, where*

$$M = \{(\alpha_i, \beta_i) : \alpha_i + i\beta_i \in \partial\Omega(S),\ \partial\Omega(S) \text{ is not differentiable at } (\alpha_i, \beta_i)\} \qquad (3.1)$$

and $N = \cup_i \gamma_i$, where each γ_i is a differentiable arc represented by the differentiable function $g_i(x, y) = 0$. Let N be the set of all points (u_i, v_i) such that (u_i, v_i) is the solution of the Lagrange multiplier equation

$$g_i(x, y) = 0,$$
$$\frac{2x}{y} = \lambda \frac{\partial g_i(x, y)}{\partial x}, \qquad (3.2)$$
$$\frac{-x^2}{y^2} = \lambda \frac{\partial g_i(x, y)}{\partial y},$$

at which the function $f(x, y) = x^2/y$ attains its minimum on γ_i. If

$$E = \left\{ \sqrt{\frac{\alpha_i^2}{\beta_i}} : (\alpha_i, \beta_i) \in M \right\},$$
$$F = \left\{ \sqrt{\frac{u_i^2}{v_i}} : (u_i, v_i) \in N \right\}, \qquad (3.3)$$

then $\mu_1(T) = \inf(E \cup F)$.

Under certain conditions, the differentiable arcs γ_i are simple arcs such as conic sections and straight lines. In those situations, equations (3.2) can be solved explicitly. One of those special cases is discussed in the next theorem. Recall that for two operators U and V, we have $U \geq V$ if $(Uf, f) \geq (Vf, f)$ for every vector f. An operator A is hyponormal if $A^*A \geq AA^*$. An operator A is seminormal if either A or A^* is hyponormal. In the following theorem, $\sigma(T)$ denotes the spectrum of an operator T, and $\sigma_p(T)$ denotes the point spectrum of T (i.e., $\sigma_p(T)$ denotes the set of all eigenvalues of T).

Theorem 3.2. *Let T be a bounded strictly accretive operator such that the numerical range of $T^*T\operatorname{Re} T$ is a subset of upper half-plane or lower half-plane, that is, $\operatorname{Im}(T^*T\operatorname{Re} Tf, f) \geq 0$ for all vectors f or $\operatorname{Im}(T^*T\operatorname{Re} Tf, f) \leq 0$ for all vectors f. Let $S = \operatorname{Re} T + iT^*T$, and $C = \sigma(S) \cap \partial\operatorname{co}(\sigma(S))$, where $\partial\operatorname{co}(\sigma(S))$ denotes the boundary of the convex hull of $\sigma(S)$. Assume that $\sigma_p(S) \cap C$ is dense in C. Let $\mathbb{R}^+$ denote the set of all positive numbers and define*

the set M by

$$M = \{\lambda_i = \alpha_i + \beta_i : \lambda_i \in \sigma_p(S), \ \partial\Omega(S) \text{ is not differentiable at } \lambda_i\}. \tag{3.4}$$

If $E = \{\sqrt{\alpha_i^2/\beta_i} : (\alpha_i, \beta_i) \in M\}$ *and*

$$F = \left\{ \frac{2\sqrt{(\alpha_i\beta_j - \alpha_{j\beta_i})(\alpha_j - \alpha_i)}}{\beta_j - \beta_i} : \alpha_i + \beta_i \in M, \ \alpha_j + \beta_j \in M \right\} \cap \mathbb{R}^+, \tag{3.5}$$

then $\mu_1(T) = \inf(E \cup F)$.

Proof. Straightforward computations show that $(S^*Sf, f) = \|T^*Tf\|^2 + \|\operatorname{Re} Tf\|^2 + 2\operatorname{Im}(T^*T\operatorname{Re} Tf, f)$ and $(SS^*f, f) = \|T^*Tf\|^2 + \|\operatorname{Re} Tf\|^2 - 2\operatorname{Im}(T^*T\operatorname{Re} Tf, f)$. Therefore, $(S^*Sf, f) - (SS^*f, f) = 4\operatorname{Im}(T^*T\operatorname{Re} Tf, f)$. Hence if $\operatorname{Im}(T^*T\operatorname{Re} Tf, f) \geq 0$ for all vectors f or $\operatorname{Im}(T^*T\operatorname{Re} Tf, f) \leq 0$ for all vectors f, then S is hyponormal or S^* is hyponormal. This implies that S is a seminormal operator. By [1], the denseness of $\sigma_p(S) \cap C$ in C implies that $\Omega(S)$ is closed and $\Omega(S) = \operatorname{co}(\sigma_p(S))$. Therefore, in this case the differentiable curves γ_i are simple line segments whose endpoints are pairs of points from set M. First assume that the infimum of the function $f(x, y) = x^2/y$ is attained at an interior point of one of these line segments. We denote the endpoints of this line segment by (α_i, β_i) and (α_j, β_j). The slope of this line segment is

$$m = \frac{\beta_j - \beta_i}{\alpha_j - \alpha_i} \tag{3.6}$$

and its equation is

$$y - \beta_i - m(x - \alpha_i) = 0. \tag{3.7}$$

To find the infimum of the function $f(x, y) = x^2/y$ on this line segment, we use the Lagrange multiplier method. This will give

$$\frac{2x}{y} = -\lambda m,$$
$$\frac{-x^2}{y^2} = \lambda. \tag{3.8}$$

Eliminating λ from (3.8) yields

$$y = \frac{mx}{2}. \tag{3.9}$$

If we now substitute y from (3.9) and m from (3.6) into (3.7) and solve the resulting equation for x, we obtain

$$x = 2\frac{\alpha_i\beta_j - \alpha_j\beta_i}{\beta_2 - \beta_1}. \tag{3.10}$$

Substituting x from (3.10) into (3.9) yields

$$y = \frac{\alpha_i\beta_j - \alpha_{j\beta_i}}{\alpha_j - \alpha_i}. \tag{3.11}$$

If we evaluate $f(x,y) = x^2/y$ at x and y given by (3.10) and (3.11), we obtain

$$\frac{4(\alpha_i\beta_j - \alpha_{j\beta_i})(\alpha_j - \alpha_i)}{(\beta_j - \beta_i)^2}. \tag{3.12}$$

Hence from the fact that $\mu_1^2(T) = \inf\{x^2/y : x + iy \in \Omega(S)\}$, we conclude that $\mu_1(T) = 2\sqrt{(\alpha_i\beta_j - \alpha_{j\beta_i})(\alpha_j - \alpha_i)}/(\beta_j - \beta_i)$. Next, assume that the infimum of the function $f(x,y) = x^2/y$ on $\partial\Omega(S)$ occurs at a nondifferentiable point of $\partial\Omega(S)$. Since a nondifferential point of $\partial\Omega(S)$ is an eigenvalue of S, then $\mu_1(T)$ is equal to an element from set E.   □

Note that every finite-dimensional hyponormal matrix is normal (see [9]). The author and Karl Gustafson have treated the cases of normal matrices and normal operators in [7, 8, 11].

Acknowledgment

References

[1]   G. De Barra, *The boundary of the numerical range*, Glasg. Math. J. **22** (1981), no. 1, 69–72.
[2]   W. F. Donoghue Jr., *On the numerical range of a bounded operator*, Michigan Math. J. **4** (1957), 261–263.
[3]   K. Gustafson, *The angle of an operator and positive operator products*, Bull. Amer. Math. Soc. **74** (1968), 488–492.
[4]   ———, *Positive (noncommuting) operator products and semi-groups*, Math. Z. **105** (1968), 160–172.
[5]   ———, *Antieigenvalue inequalities in operator theory*, Inequalities III (O. Shisha, ed.), Academic Press, New York, 1972, pp. 115–119.
[6]   K. Gustafson and D. Rao, *Numerical Range*, Universitext, Springer, New York, 1997.
[7]   K. Gustafson and M. Seddighin, *Antieigenvalue bounds*, J. Math. Anal. Appl. **143** (1989), no. 2, 327–340.
[8]   ———, *A note on total antieigenvectors*, J. Math. Anal. Appl. **178** (1993), no. 2, 603–611.
[9]   P. R. Halmos, *A Hilbert Space Problem Book*, 2nd ed., Graduate Texts in Mathematics, vol. 19, Springer, New York, 1982.
[10] M. G. Kreĭn, *Angular localization of the spectrum of a multiplicative integral in a Hilbert space*, Funct. Anal. Appl. **3** (1969), no. 1, 73–74.
[11] M. Seddighin, *Antieigenvalues and total antieigenvalues of normal operators*, J. Math. Anal. Appl. **274** (2002), no. 1, 239–254.
[12] ———, *Optimally rotated vectors*, Int. J. Math. Math. Sci. **2003** (2003), no. 63, 4015–4023.

Morteza Seddighin: Mathematics Department, Division of Natural Science and Mathematics, Indiana University East, Richmond, IN 47374, USA

E-mail address: mseddigh@indiana.edu