Applied Mathematics and Computation 244 (2014) 132–141

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

On the geometry of the envelope of a matrix

Panayiotis J. Psarrakos [a,*], Michael J. Tsatsomeros [b]

[a] Department of Mathematics, National Technical University of Athens, Zografou Campus, 15780 Athens, Greece
[b] Mathematics Department, Washington State University, Pullman, WA 99164-3113, USA

ARTICLE INFO

Keywords:
Eigenvalue bounds
Numerical range
Cubic curve
Envelope

ABSTRACT

The envelope, $\mathcal{E}(A)$, of a complex square matrix A is a region in the complex plane that contains the spectrum of A and is contained in the numerical range of A. The envelope is compact but not necessarily convex or connected. The connected components of $\mathcal{E}(A)$ have the potential of isolating the eigenvalues of A, leading us to study its geometry, boundary, and number of components. We also examine the envelope of normal matrices and similarities. In the process, we observe that $\mathcal{E}(A)$ contains the 2-rank numerical range of A.

1. Introduction

The *envelope* of a complex square matrix A, denoted by $\mathcal{E}(A)$, is an eigenvalue containment region that was introduced in [14]. Evidently, the envelope represents a theoretically, computationally and visually attractive way to localize the spectrum of A by isolating the eigenvalues in its connected components.

The concept and definition of the envelope are based on an inequality proven in [1] that the (real and imaginary parts of the) eigenvalues of A must satisfy. This inequality allows one to replace the half-plane to the left of the largest eigenvalue of the hermitian part of A by a smaller region that contains the spectrum of A. Thus, upon rotating a matrix A through all angles in $[0, 2\pi)$, the envelope arises as a region that contains the eigenvalues and is contained in the numerical range, $F(A)$. The precise definition and illustrations of $\mathcal{E}(A)$ can be found in Section 3.

The rendering of $\mathcal{E}(A)$ is akin to the process for $F(A)$, essentially requiring knowledge of the first but also the second largest eigenvalues of the hermitian part of $e^{i\theta}A$ for a range of angles in $[0, 2\pi)$.

The envelope has properties similar to $F(A)$, e.g., it is compact, invariant under unitary similarities and homogeneous; it is not, however, necessarily convex or connected. The aim of this paper is to further understand the properties and features of $\mathcal{E}(A)$ as they pertain to its geometry, boundary, number of components, and containment of eigenvalues. In particular, we study the case of normal matrices and eigenvalues, and make comparisons to the numerical range. In the process, we discover that the envelope contains the 2-rank numerical range of A introduced in [2].

This paper is organized as follows. In Section 2, we describe the notions relevant to the definition and study of the envelope. In Section 3, the envelope is defined formally, its basic properties are reviewed, and its relation to the 2-rank numerical range is established. Section 4 contains results on extremal eigenvalues, normal matrices (Section 4.1) and similarities (Section 4.2), and the effects of such assumptions on the geometry of the envelope are examined. Finally, a result on the eigenvectors of the right-most eigenvalues is given in Section 5, and some conclusions are presented in Section 6.

* Corresponding author.
E-mail addresses: ppsarr@math.ntua.gr (P.J. Psarrakos), tsat@math.wsu.edu (M.J. Tsatsomeros).

http://dx.doi.org/10.1016/j.amc.2014.06.077

2. Definitions and preliminaries

Let $A \in \mathbb{C}^{n\times n}$ ($n \geqslant 2$) be an $n \times n$ complex matrix with spectrum $\sigma(A)$. Consider the *hermitian* and *skew-hermitian parts* of A, $H(A) = (A + A^*)/2$ and $S(A) = (A - A^*)/2$, respectively, and let $\delta_1(A) \geqslant \delta_2(A) \geqslant \cdots \geqslant \delta_n(A)$ denote the eigenvalues of $H(A)$ in a nonincreasing order. Let also $y_1 \in \mathbb{C}^n$ be a unit (with respect to the Euclidean vector norm) eigenvector of $H(A) correspondingto$\color{black}{\delta}_{1}(A).

2.1. The standard numerical range

The *numerical range* (also known as the *field of values*) of A is defined as

$$F(A) = \{v^*Av \in \mathbb{C} : v \in \mathbb{C}^n \text{ with } v^*v = 1\}.$$

It is a *compact* and *convex* subset of $\mathbb{C}$ that contains $\sigma(A)$ and is a useful concept in understanding matrices and operators; see [6, Chapter 1] and the references therein.

For an angle $\theta \in [0, 2\pi)$, we consider the largest eigenvalue $\delta_1(e^{i\theta}A)$ and an associated unit eigenvector $y_1(\theta)$ of the hermitian matrix $H(e^{i\theta}A)$. Then, the point $z_\theta = y_1(\theta)^*Ay_1(\theta)$ lies on the boundary of $F(A)$, denoted by $\partial F(A)$, and the line $\mathcal{L}_\theta = \{e^{-i\theta}(\delta_1(e^{i\theta}A) + it) : t \in \mathbb{R}\}$ is tangential to $\partial F(A)$ at z_θ [6,7]. Furthermore, $\mathcal{L}_\theta$ defines the closed half-plane $\mathcal{H}_{in}(A,\theta) = \{e^{-i\theta}(s + it) : s, t \in \mathbb{R} \text{ with } s \leqslant \delta_1(e^{i\theta}A)\}$, which contains $F(A)$. Hence, $F(A)$ can be written as an infinite intersection of closed half-planes [6, Theorem 1.5.12], namely,

$$F(A) = \bigcap_{\theta\in[0,2\pi)} \{e^{-i\theta}(s + it) : s, t \in \mathbb{R} \text{ with } s \leqslant \delta_1(e^{i\theta}A)\} = \bigcap_{\theta\in[0,2\pi)} \mathcal{H}_{in}(A,\theta). \tag{1}$$

2.2. The k-rank numerical range

For $1 \leqslant k \leqslant n-1$, the *k-rank numerical range* of matrix $A \in \mathbb{C}^{n\times n}$ is defined as

$$\begin{aligned}\Lambda_k(A) &= \{\mu \in \mathbb{C} : PAP = \mu P \text{ for some rank-}k\text{ orthogonal projection } P \in \mathbb{C}^{n\times n}\}\\ &= \left\{\mu \in \mathbb{C} : X^*AX = \mu I_k \text{ for some } X \in \mathbb{C}^{n\times k} \text{ such that } X^*X = I_k\right\},\end{aligned}$$

and is a natural generalization of the standard numerical range, in the sense that $\Lambda_1(A)$ coincides with $F(A)$. This set was introduced in [2] and has attracted attention because of its role in quantum information theory; specifically, it is closely connected to the construction of quantum error correction codes for noisy quantum channels (see [2,3,8] and the references therein). The range $\Lambda_k(A)$ is a *compact* and *convex* subset of the complex plane [16] and is given by the explicit formula [11, Theorem 2.2]

$$\Lambda_k(A) = \bigcap_{\theta\in[0,2\pi)} \{e^{-i\theta}(s + it) : s, t \in \mathbb{R} \text{ with } s \leqslant \delta_k(e^{i\theta}A)\}. \tag{2}$$

Moreover, $\Lambda_k(A)$ is invariant under unitary similarity and satisfies $\Lambda_{n-1}(A) \subseteq \Lambda_{n-2}(A) \subseteq \cdots \subseteq \Lambda_2(A) \subseteq \Lambda_1(A) = F(A)$. For $k \geqslant 2$, $\Lambda_k(A)$ does not necessarily contain all of the eigenvalues of A and, in fact, may be empty [10].

If the matrix $A \in \mathbb{C}^{n\times n}$ is normal with (not necessarily distinct) eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then (2) implies that (see Corollary 2.4 of [11])

$$\Lambda_k(A) = \bigcap_{1\leqslant j_1<j_2<\cdots<j_{n-k+1}\leqslant n} \operatorname{conv}\{\lambda_{j_1}, \lambda_{j_2}, \ldots, \lambda_{j_{n-k+1}}\}, \tag{3}$$

where $\operatorname{conv}\{\cdot\}$ denotes the convex hull. Efficient techniques to generate $\Lambda_k(A)$ for normal A, using half-planes determined by the eigenvalues instead of formula (3), are proposed in [4].

2.3. The cubic curve $\Gamma(A)$

For matrix $A \in \mathbb{C}^{n\times n}$, define the nonnegative quantities $\mathrm{v}(A) = \|S(A)y_1\|_2^2$ and $\mathrm{u}(A) = \operatorname{Im}(y_1^*S(A)y_1) \leqslant \|S(A)y_1\|_2 = \sqrt{\mathrm{v}(A)}$, where $\|\cdot\|_2$ denotes the *spectral matrix norm* (i.e., the norm subordinate to the Euclidean vector norm). Adam and Tsatsomeros [1], extending a methodology of [12], derived the following theorem.

Theorem 2.1. [1, Theorem 3.1] *Let $A \in \mathbb{C}^{n\times n}$. Then, for every eigenvalue $\lambda \in \sigma(A)$,*

$$(\operatorname{Re}\lambda - \delta_2(A))(\operatorname{Im}\lambda - \mathrm{u}(A))^2 \leqslant (\delta_1(A) - \operatorname{Re}\lambda)[\mathrm{v}(A) - \mathrm{u}(A)^2 + (\operatorname{Re}\lambda - \delta_2(A))(\operatorname{Re}\lambda - \delta_1(A))].$$

Motivated by the above result, the authors of [1] introduced and studied the algebraic curve

$$\Gamma(A) = \left\{s + it : s, t \in \mathbb{R},\ (\delta_2(A) - s)[(\delta_1(A) - s)^2 + (\mathrm{u}(A) - t)^2] + (\delta_1(A) - s)(\mathrm{v}(A) - \mathrm{u}(A)^2) = 0\right\}.$$