Characterize distribution of Rayleigh quotients in the numerical range of matrix

Fatemeh Esmaeili Taheri

Munich June 2018

General view

General view

- Definition and basic properties of Numerical range .

General view

- Definition and basic properties of Numerical range .

General view

- Definition and basic properties of Numerical range .
- Characterize distribution of Rayleigh quotients in the numerical range of matrix.

General view

- Definition and basic properties of Numerical range .
- Characterize distribution of Rayleigh quotients in the numerical range of matrix.

General view

- Definition and basic properties of Numerical range .
- Characterize distribution of Rayleigh quotients in the numerical range of matrix.
- Distance from normality.

General view

- Definition and basic properties of Numerical range .
- Characterize distribution of Rayleigh quotients in the numerical range of matrix.
- Distance from normality.

General view

- Definition and basic properties of Numerical range .
- Characterize distribution of Rayleigh quotients in the numerical range of matrix.
- Distance from normality.
- Inverse field of value problem.

General view

- Definition and basic properties of Numerical range .
- Characterize distribution of Rayleigh quotients in the numerical range of matrix.
- Distance from normality.
- Inverse field of value problem.

General view

- Definition and basic properties of Numerical range .
- Characterize distribution of Rayleigh quotients in the numerical range of matrix.
- Distance from normality.
- Inverse field of value problem.
- The possible points of Rayleigh quotients in the numerical range of unknown matrix.

In 1918, Teoplitz introduced the fallowing algebraic form:

In 1918, Teoplitz introduced the fallowing algebraic form:

In 1918, Teoplitz introduced the fallowing algebraic form:

In 1918, Teoplitz introduced the fallowing algebraic form:

In 1918, Teoplitz introduced the fallowing algebraic form:

In 1918, Teoplitz introduced the fallowing algebraic form:

Definition

let A be a $n \times n$ complex matrix, the numerical range of A is a set of x^*Ax such that $x^*x = 1$, denoted by $W(A)$

$$W(A) = \{x^*Ax : x \in \mathbb{C}^n, x^*x = 1\}$$

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ *and* $\alpha, \beta \in C$ *then*

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ *and* $\alpha, \beta \in C$ *then*

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$
4) *If $U, A \in M_n$ and U be a unitary matrix then $w(U^* AU) = w(A)$*

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$
4) *If $U, A \in M_n$ and U be a unitary matrix then* $w(U^* AU) = w(A)$
5) $Re(w(A)) = w(H(A))$

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$
4) *If $U, A \in M_n$ and U be a unitary matrix then* $w(U^* AU) = w(A)$
5) $Re(w(A)) = w(H(A))$
6) $w(A)$ *is a subset of* $\mathbb{C}$.

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$
4) *If $U, A \in M_n$ and U be a unitary matrix then $w(U^* AU) = w(A)$*
5) $Re(w(A)) = w(H(A))$
6) *$w(A)$ is a subset of $\mathbb{C}$.*
7) *$w(A)$ is connected.*

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$
4) *If* $U, A \in M_n$ *and* U *be a unitary matrix then* $w(U^* AU) = w(A)$
5) $Re(w(A)) = w(H(A))$
6) $w(A)$ *is a subset of* $\mathbb{C}$.
7) $w(A)$ *is connected.*
8) $w(A)$ *is convex.*

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$
4) *If $U, A \in M_n$ and U be a unitary matrix then $w(U^* AU) = w(A)$*
5) $Re(w(A)) = w(H(A))$
6) *$w(A)$ is a subset of $\mathbb{C}$.*
7) *$w(A)$ is connected.*
8) *$w(A)$ is convex.*
9) *$w(A)$ is compact.*

Toeplitz showed Numerical range has following properties,

Theorem

If $A, B \in M_n$ and $\alpha, \beta \in C$ then

1) $w(A + \alpha I_n) = w(A) + \alpha, w(\beta A) = \beta w(A)$
2) $\sigma(A) \subseteq w(A)$
3) $w(A + B) \subseteq w(A) + w(B)$
4) *If $U, A \in M_n$ and U be a unitary matrix then $w(U^* AU) = w(A)$*
5) $Re(w(A)) = w(H(A))$
6) *$w(A)$ is a subset of $\mathbb{C}$.*
7) *$w(A)$ is connected.*
8) *$w(A)$ is convex.*
9) *$w(A)$ is compact.*

When we try to locate in the complex plane the eigenvalues of complex matrices, we can use the field of values as a containment region for the spectrum of the matrix.
For an eigenvector $x \in \mathbb{C}^n$ of A with corresponding eigenvalue λ, the equation $Ax = \lambda x$ holds. Taking the inner product with x on both sides, we get

$$(Ax, x) = (\lambda x, x)$$

or

$$\frac{(Ax, x)}{(x, x)} = \lambda.$$

The ratio $\frac{(Ax,x)}{(x,x)}$ is well defined for any nonzero vector $x \in \mathbb{C}^n$ and any matrix $A \in \mathbb{M}_n$, and is called the **Rayleigh quotient** of x with respect to A. Thus, the numerical range comprises all the Rayleigh quotients of the matrix.

Motivation

The field of values of a matrix is the closed convex subset of the complex plane comprising all Rayleigh quotients a set of interest in the

- stability analysis of dynamical system

among other applications.

Motivation

The field of values of a matrix is the closed convex subset of the complex plane comprising all Rayleigh quotients a set of interest in the

- stability analysis of dynamical system
- convergence theory of matrix iteration

among other applications.

FOV points can be generated in several ways. For one, a natural method is to evaluate x^*Ax for a number of random unit vectors and see the given point $p \in \mathbb{C}$ lies inside their convex hull (Numerical range). We produce a pattern of points while the points are randomly located.

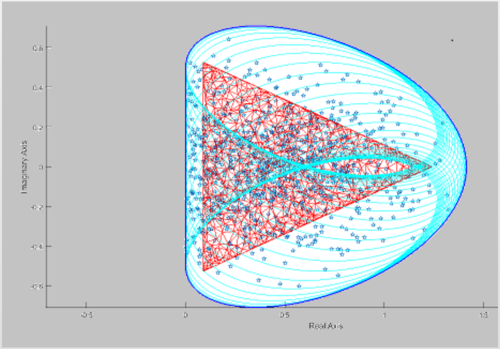
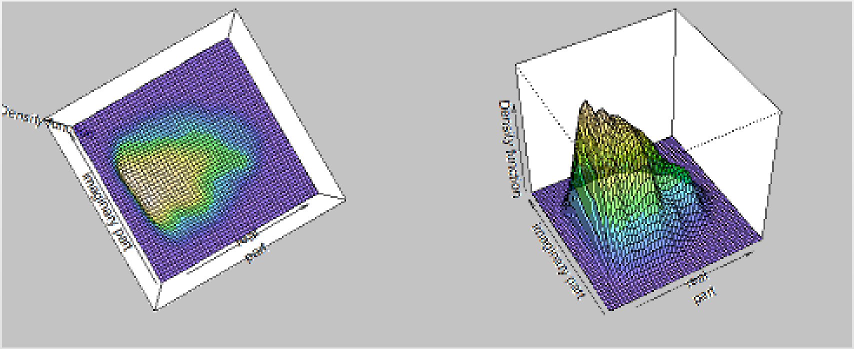
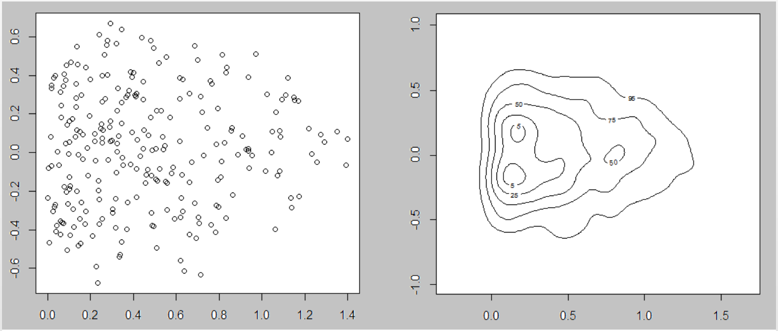


Figure : Blue curve is the boundary of numerical range of matrix $\begin{bmatrix} 0 & \frac{-1}{2} & 0, \frac{1}{2} & 0 & \frac{-1}{2}, 0 & \frac{-1}{2} & -\sqrt{2} \end{bmatrix}$ and red triangle is the covexhull of eigenvalues with 1000 Rayleigh quotients points.

We look at the points with different perspective
(numerical description)



when density is high or low it shows the point concentrating on that region around arbitrary fixed point x is high or low and it shows you can expect that many points around x or not?
In physics it shows density mass they like to know mass density function



Most of the points are probably near the flat portion and more rarely they are located near the border and we try to measure this concentrating and we do not know what the reason is?

these points will gives us much more information than those points which are further away

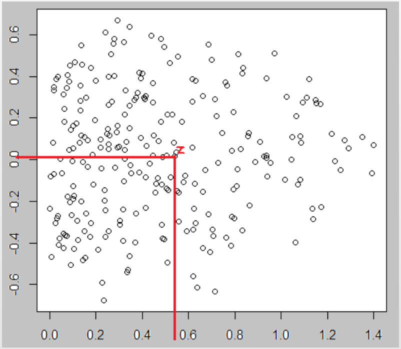
Goal

- our aim is studying the behaviour of the matrix, how Rayleigh quotient distribute themselves throughout the field of values.

density, measure how likely the points is to appear in particular area.

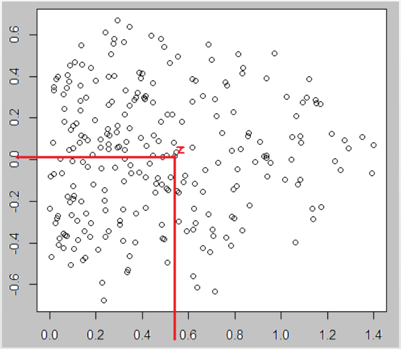
One of the mutivations

Such knowledge would prove useful for analyzing eigenvalue algorithms such as Arnoldi method(for computing eigenvalues) understanding the convergence of such iterative eigensolvers requires insight into how Rayleigh quotients distribute themselves throughout the field of values.



we observe certain pattern of points and trying to guess what density function should be

our density function is not constant so it shows the points do not spread evenly along the region and proportionality is not the same for each points and different proportionality is depend on where the point is located



look at observation window counts number of points in the certain region, count point around x (x is a initial point) and normalized by the size of our window of observation then we could estimate density function integration of these point gives us distribution function.
we are interested on concentration points around x.

A way to define distribution function over the complex plane we just fixed a point on the plane and we try to characterize the probability of complex point, $F : \mathbb{C} \to [0, 1], \quad (z = x_0 + y_0)$

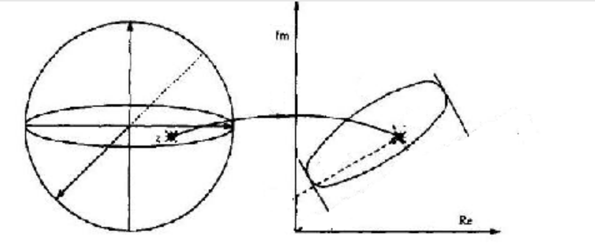
$$F(x_0, y_0) = P(x \leq x_0, y \leq y_0) \approx \frac{1}{n}\Sigma_{i=1}^{n} 1(x_i \leq x_0, y_i \leq y_0)$$

$$1_{\Omega^*}(z) = \begin{cases} 1 & z \in \Omega^* \\ 0 & z \notin \Omega^* \end{cases}$$

we define

$$T : S^n \to \mathbb{C}, \quad T_A(X) = \frac{X^*AX}{X^*X} \in \mathbb{C}$$

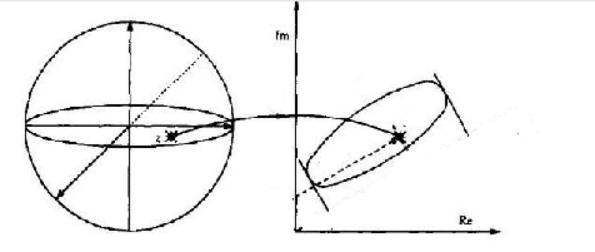
in order to characterize distribution of Rayleigh quotient get the set from the Rayleigh quotient name $\Omega^* \subset W(A) \subset \mathbb{C}$ is a subset of complex number we want to compute the probability of $T_A(X)$ belong to Ω^*, to compute probability of the points of the set, the idea is find the set of the points which are transformed into Ω^*.



To find the distribution of points we should pay attention the distribution of vectors, we have uniform distribution on sphere that is because we try to get back.(we choose point uniformly) analytically means we just comprise the size of $T_A^{-1}(\Omega^*)$ with the size whole sphere we should use lebesgue measure to get the size in two dimensional case compute the area

$$P(T_A(X) \in \Omega^*) = P(X \in T_A^{-1}(\Omega^*)) = \frac{\lambda^{n-1}(T_A^{-1}(\Omega^*))}{\lambda^{n-1}(S^n)} = \int fdx.$$

and f is a density function and describe the concentration of Rayleigh quotient



consider $A \in M_2$ without loss of generality we may assume that the trace of A is zero. A unitary matrix U can be found such that A is transformed into Schur form

$$U^* AU = \widehat{A} = \begin{bmatrix} a & c \\ 0 & -a \end{bmatrix}$$

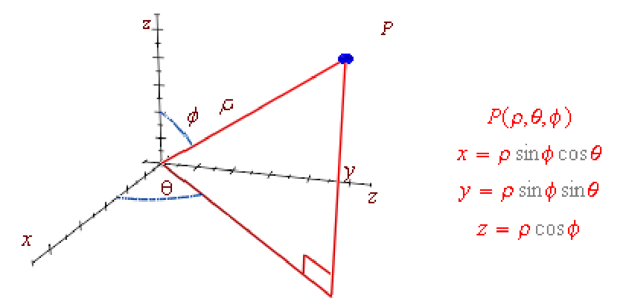
where c is real. Let $z = x + iy \in F(\widehat{A})$.
Without loss of generality, we may consider the unit vector

$$w_z = (\cos u, e^{i\varphi} \sin u)^T,$$

with real nonnegative first coordinate.

Geometrical interpretation of parameters

we choose the vector $w_z = (\cos u, e^{i\varphi}\sin u)^T \in \mathbb{C}^2 \cong \mathbb{R}^4 \cong S^3$. in the sphere,



We have spherical polar coordinates (φ, u) such that

$$x = r\cos u \sin\varphi$$
$$y = r\sin u \sin\varphi$$
$$z = r\cos\varphi$$

we fixed Ω^* in two dimensional plane and compute inverse set, it is the set depending on (u, φ) afterword we compute the area of this set or surface with integrating with respect to u and φ
We find $w_z^* \widehat{A} w_z = a \cos 2u + \frac{c}{2} \sin 2u \cos \varphi + i \frac{c}{2} \sin 2u \sin \varphi$ which easily gives

$$cos2u = \frac{4ax \pm \sqrt{c^4 + 4a^2c^2 4c^2x^2(4c^2 + 16a^2)y^2}}{4a^2 + c^2},$$

this relation determines u, and the relation

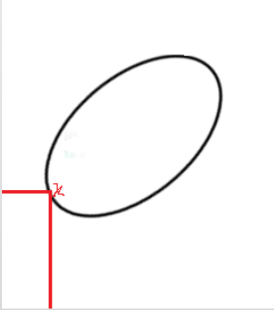
$$sin\varphi = \frac{2y}{csin2u}$$

determines φ.

$$x_1 = \lambda_{min}(H(A)), y_1 = \lambda_{min}(K(A)) \quad s.t \quad A = H + iK,$$

$$T_A^{-1}(\Omega^*) = \{(u, \varphi) : cos2u = \frac{4ax \pm \sqrt{c^4 + 4a^2c^2 4c^2x^2(4c^2 + 16a^2)y^2}}{4a^2 + c^2},$$

$$sin\varphi = \frac{2y}{csin2u}, x_1 < x \leq x_2, y_1 < y \leq y_2\}. \quad 0 \leq u \leq \pi, 0 \leq \varphi \leq 2\pi,$$



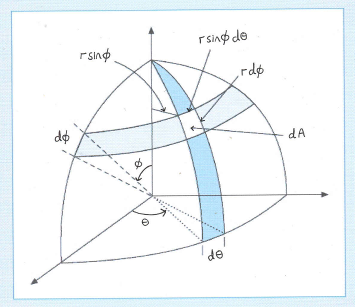
we know the location of extreme point minimum and maximum possible value for x and y

$$u \in (\frac{1}{2} Arc\cos \frac{4ax_1 + \sqrt{c^4 + 4a^2c^2 4c^2 x_1^2 (4c^2 + 16a^2) y_1^2}}{4a^2 + c^2},$$

$$\frac{1}{2} Arc\cos \frac{4ax_2 + \sqrt{c^4 + 4a^2c^2 4c^2 x_2^2 (4c^2 + 16a^2) y_2^2}}{4a^2 + c^2})$$

$$\varphi \in (Arc\sin \frac{2y_1}{c\sin 2u}, Arc\sin \frac{2y_2}{c\sin 2u})$$

we have the subset of sphere on the circumference (we consider normal vector) we want to compute the area set of (u, φ) I will integrating over this subset



distribution function of transformation T_A

$$F_A = P_A(x_1 < x \leq x_2, y_1 < y \leq y_2) = \frac{1}{4\pi} \int_{u_1}^{u_2} \int_{\varphi_1}^{\varphi_2} \sin \varphi d\varphi du$$

$$F_A = \frac{-1}{4\pi}\int_{u_1}^{u_2} \cos\varphi_2 - \cos\varphi_1 du$$

$$F_A = \frac{-1}{4\pi}\int_{u_1}^{u_2} [\cos(Arc\sin(\frac{2y_2}{c}\csc 2u_2)) - \cos(Arc\sin(\frac{2y_1}{c}\csc 2u_1))]du$$

in order to find density function we should differentiate of F_A respect to x and y so density function is $f_A = \frac{\partial F_A}{\partial x \partial y}$

It explains for a given matrix how likely to find the Rayleigh quotient over here or there in each region inside the numerical range.

this characterization might provide some information can be useful how fast the convergence depending on how Rayleigh quotient behave in numerical range then we can say about convergence

Distance from the normality

Things I would ask is How can I measure the distance between two matrices with respect to the difference between density functions they have.
to distinguish between one matrix and the normal matrix try to get an idea of distance between their distribution function
$F : \mathbb{C} \to [0, 1], \quad F(z) = P(x \leq x_0, y \leq y_0) \quad (z = x_0 + y_0)$
there are several ways to define the distance between two distributions

- K-S method $Sup_z|F(z) - G(z)|$
- Anderson-Darling distance $\int_{\mathbb{C}}(F(z) - G(z))^2 w(z)dz$,
- Hellinger distance and Wasserstein distance also they are most popular methods

try to characterize measure distance It should be a one to one correspondence, you give a numerical range matrix you get a one density function and if you have density function you get a matrix, it should be one to one.

Inverse Problem (sort of inverse problem by statistical motivation)

try to test this observation (Rayleigh quotients) produce with the matrix is compatible with the matrix which is produced by inverse problem or not

Likelihood Method

$$Sup\Sigma_{i=1}^{n} log(f(a, c, x_i, y_i))$$

where f is density function and x_i, y_i are real part and imaginary part of random field of value points, with optimization we get a and c two entries of a matrix.

So in general, if we have computational numerical points which are generated by unknown matrix on the plane we can identify the unknown two by two matrix.

The possible points of Field of value points of matrix.

Second thing I would ask is how can I realize that the pattern of points can be the field of values points of matrix we have a bunch of points,

The possible points of Field of value points of matrix.

Second thing I would ask is how can I realize that the pattern of points can be the field of values points of matrix we have a bunch of points,

- we try to find the estimate distribution function
  $G : \mathbb{C} \to [0,1], \quad G(z) = P(x \leq x_0, y \leq y_0) \quad (z = x_0 + y_0)$

The possible points of Field of value points of matrix.

Second thing I would ask is how can I realize that the pattern of points can be the field of values points of matrix we have a bunch of points,

- we try to find the estimate distribution function $G : \mathbb{C} \to [0,1], \quad G(z) = P(x \leq x_0, y \leq y_0) \quad (z = x_0 + y_0)$

The possible points of Field of value points of matrix.

Second thing I would ask is how can I realize that the pattern of points can be the field of values points of matrix we have a bunch of points,

- we try to find the estimate distribution function
  $G : \mathbb{C} \to [0,1], \quad G(z) = P(x \leq x_0, y \leq y_0) \quad (z = x_0 + y_0)$
- we find the distance between this estimated function and exact distribution function
  $Sup_z |F(z) - G(z)|$

if this value less than epsilon we get the result.

Compression to the two-by-two case

For A an n-square complex matrix, a two dimensional real orthogonal compression of A is a two-by-two square matrix

$$A_{\gamma\tau} = \begin{bmatrix} (A\gamma, \gamma) & (A\tau, \gamma) \\ (A\gamma, \tau) & (A\tau, \tau) \end{bmatrix}, \tag{1}$$

with γ, τ real orthonormal column n-tuples.

Theorem

(Marcus-Pesce) Let A be an n-square complex matrix. Then

$$W(A) = \bigcup_{\gamma,\tau} W(A_{\gamma\tau}),$$

where $A_{\gamma\tau}$ is the matrix, and γ, τ run over all pairs of real orthonormal vectors.

Bibliography I

[4]D. Williams, *A Course in Probability and Statistics,* Cambridge University Press, Cambridge (2001).

[5] I.F santos, W.G Manteiga, J. Mateu, *Consistent smooth bootstrap kernel intensity estimation for inhomogeneous spatial Poisson point processes,* Scandinavian Journal of Statistics.. 252 (2016) 416435.

[3] N. Bebiano, J. da Providencia, A. Nata and J. P.da Providencia *Revising the inerse fied of values probem,*

[6] R. Cordan, *A simple algorithm for the inverse field of values problem,* Inverse Problems, 25 (2009), 115019, (9 pages). Czechoslovsk Mathematical Journal. 42 (2014) 1-12.

[4]R. Horn, C. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, Cambridge (1991).

Bibliography II

[12] J. William Helton and I.M. Spitkovsky, *The possible shapes of numerical ranges,* Operators and Matrices, 6 (2012) 607-611.

[14] David W. Kribs , Aron Pasieka , Martin Laforest , Colm Ryan and Marcus P. da Silva, *Research problems on numerical ranges in quantum computing,* Linear and Multilinear Algebra, 57 (2009) 491-502.

" Thanks "