Pseudospektar matrica

Ivica Nakić Mihael Petrić

Sažetak

Pojam svojstvenih vrijednosti i spektra matrice odavno se istražuje i koristi u matematici. Svojstvene vrijednosti matrica u mnogim slučajevima daju izvrstan uvid u svojstva samih matrica, no katkada ne daju dovoljno informacija za rješavanje problema na koje se može naići. Takvi se slučajevi pojavljuju u raznim granama matematike kao npr. teorije operatora i teorije Markovljevih lanaca i ostalih znanosti, od populacijske ekologije, preko laserske tehnologije, kvantne mehanike i hidrodinamike.

Katkada se preciznije informacije o matrici mogu dobiti korištenjem pseudospektra te je cilj ovog članka dati čitatelju osnovne informacije o ovom zanimljivom poopćenju pojma spektra.

U ovom će članku biti navedeni osnovni pojmovi vezani uz pseudospektar, ekvivalentne definicije pseudospektra, odnos prema običnom spektru matrica, kao i neka osnovna svojstva.

1 Definicije pseudospektra

Promatrano u okvirima primijenjene matematike pitanje „Je li $\mathbf{A}$ singularna matrica?" često nema puno smisla. Naime, proizvoljno mala perturbacija matrice može promijeniti odgovor na to pitanje iz pozitivnog u negativan. Budući da se pitanje „Je li λ svojstvena vrijednost matrice $\mathbf{A}$?" može ekvivalentno postaviti u obliku „Je li $\lambda - \mathbf{A}$ singularna matrica?", i ovdje nailazimo na isti problem.

Stoga je potrebno preformulirati ovo pitanje tako da uzmemo u obzir osjetljivost na perturbacije. Do pogodne formulacije dolazimo na osnovi sljedeće činjenice: što je matrica $\mathbf{A}$ bliža singularnoj matrici, to je matrica $\mathbf{A}^{-1}$ veća, u smislu da je norma te matrice veći broj.

Dakle, možemo postaviti pitanje „Je li $\|\mathbf{A}^{-1}\|$ velika?".

Ako je odgovor da, onda je matrica jako blizu nekoj singularnoj matrici, te je za praktične svrhe možemo smatrati singularnom. Obrnuto, ako je odgovor ne, matrica će i uz male smetnje ostati regularnom.

Ako ovo rezoniranje primijenimo na problem svojstvenih vrijednosti, dolazimo do zaključka da bi od interesa mogli biti oni brojevi z za koje $\|(z-\mathbf{A})^{-1}\|$ ima veliku vrijednost.

Ovim slijedom razmišljanja dolazimo do prve definicije pseudospektra.

Definicija 1.1 (1. definicija pseudospektra). *Neka je $\mathbf{A} \in \mathbb{C}^{N\times N}$ i $\varepsilon > 0$ proizvoljan.* ε-pseudospektar $\sigma_\varepsilon(\mathbf{A})$ *matrice $\mathbf{A}$ u normi $\|\cdot\|$ definiran je s*

$$\sigma_\varepsilon = \{z \in \mathbb{C} : \|(z-\mathbf{A})^{-1}\| > \varepsilon^{-1}\}. \tag{1}$$

Važno je uočiti da ova definicija ovisi o izboru norme $\|\cdot\|$, te se stoga katkada govoti o 2–pseudospektru, 1–pseudospektru i ∞–pseudospektru, ako je pripadna norma 2–norma, 1–norma, odnosno ∞–norma.

Mi ćemo se najčešće koristiti spektralnom 2-normom u oznaci $\|\cdot\|_2$.

Matricu $(z-\mathbf{A})^{-1}$ zovemo *rezolventom matrice* $\mathbf{A}$.

U (1), kao i u ostatku rada, podrazumijevamo da vrijedi:

$$\|(z-\mathbf{A})^{-1}\| = \infty \quad \text{za} \quad z \in \sigma(\mathbf{A}), \tag{2}$$

gdje $\sigma(\mathbf{A})$ predstavlja spektar matrice $\mathbf{A}$.

Odavde slijedi da je ε-pseudospektar podskup kompleksne ravnine koji uvijek sadržava spektar pripadne matrice, i to za svaki $\varepsilon > 0$. Ako definiramo funkciju $f(z) := \|(z-\mathbf{A})^{-1}\|$ i iskoristimo činjenicu da je $\|\cdot\|$ neprekidna funkcija, lako se zaključuje da je ε-pseudospektar otvoren skup kao praslika otvorenog skupa po neprekidnoj funkciji:

$$\sigma_\varepsilon(\mathbf{A}) = f^{-1}(\langle\varepsilon^{-1}, \infty]).$$

Drugim riječima, *ε-pseudospektar matrice otvoren je podskup kompleksne ravnine omeđen s ε^{-1} nivo krivuljom norme rezolvente matrice.*

Intuitivno možemo pretpostaviti da je $\|(z-\mathbf{A})^{-1}\|$ velika upravo onda kada je točka z veoma blizu svojstvenoj vrijednosti matrice $\mathbf{A}$. No, kao što ćemo poslije i demonstrirati, točnost naše intuicije ovisi o izboru matrične norme i normalnosti same matrice. Kod normalnih matrica, kada je $\|\cdot\| = \|\cdot\|_2$, $\|(z-\mathbf{A})^{-1}\|$ velika točno onda kada je točka z blizu svojstvenoj vrijednosti matrice $\mathbf{A}$ (vidi Sliku 2). Važnost pseudospektra dolazi do izražaja kod matrica koje nemaju svojstvo normalnosti, a za koje norma $\|(z-\mathbf{A})^{-1}\|$ može biti velika čak i kada je točka z daleko od spektra matrice (vidi Sliku 2).

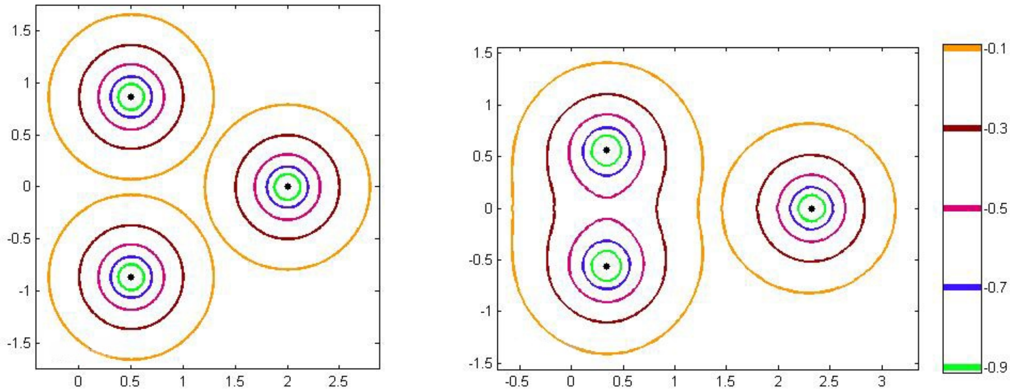
Neka je $\mathbf{A}$ matrica s potpunim skupom svojstvenih vektora $\{\mathbf{v}_j\}$, $\mathbf{V}$ neka je $N\times N$ matrica čiji je j-ti stupac vektor $\mathbf{v}_j$ i Λ dijagonalna $N \times \mathbb{N}$ matrica s j-tom svojstvenom vrijednošću λ_j na j-tom mjestu dijagonale. Znamo da tada matricu $\mathbf{A}$ možemo zapisati kao

$$\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^{-1}.$$

ε-pseudospektar će biti od važnosti kod matrica za koje je

$$\|\mathbf{V}\|\|\mathbf{V}^{-1}\| \gg 1, \tag{3}$$

uz odabranu normu $\|\cdot\|$.



Slika 1: ε-pseudospektar matrice u 2-normi. Različito obojene linije predstavljaju granice pseudospektra za različite vrijednosti ε. Lijeva slika prikazuje pseudospektar normalne matrice, a desna pseudospektar matrice koja nema svojstvo normalnosti.

Primijenimo li na matricu proizvoljno malu perturbaciju i tada gledamo spektar novodobivene matrice, dolazimo do alternativne definicije pseudospektra.

Definicija 1.2 (2. definicija pseudospektra). *$\sigma_\varepsilon(\mathbf{A})$ u normi $\|\cdot\|$ je skup svih $z \in \mathbb{C}$ takvih da vrijedi*

$$z \in \sigma(\mathbf{A} + \mathbf{E}) \tag{4}$$

za neku matricu $\mathbf{E} \in \mathbb{C}^{N\times N}$ pri čemu je $\|\mathbf{E}\| < \varepsilon$.

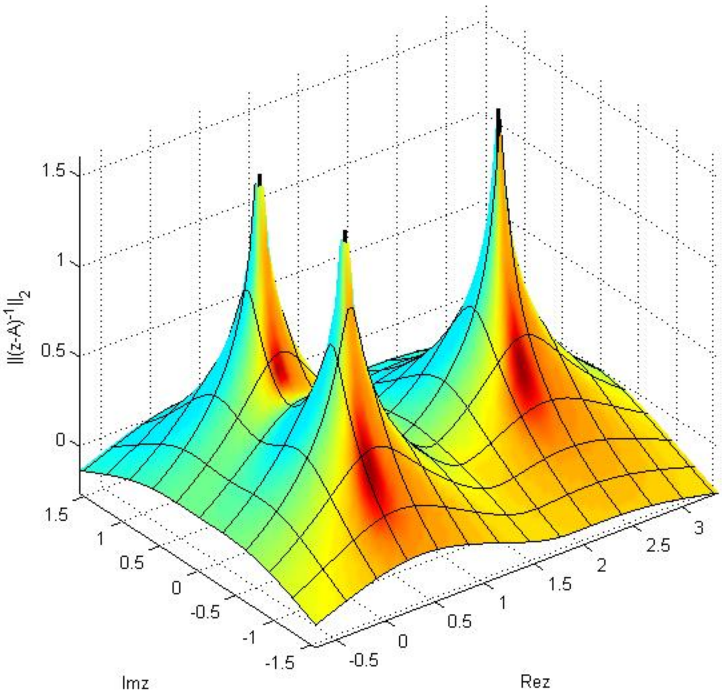
Drugim riječima, *ε–pseudospektar je skup brojeva koji su svojstvene vrijednosti neke perturbirane matrice $\mathbf{A} + \mathbf{E}$ gdje je $\|\mathbf{E}\| < \varepsilon$.*

Iz ove definicije očito slijedi da za pseudospektre vezane uz raličite ε vrijedi

$$\sigma_{\varepsilon_1} \subseteq \sigma_{\varepsilon_2}, \quad 0 < \varepsilon_1 \leq \varepsilon_2, \tag{5}$$

te također da je presjek svih mogućih pseudospektara matrice upravo spektar matrice $\mathbf{A}$,

$$\bigcap_{\varepsilon>0} \sigma_\varepsilon(\mathbf{A}) = \sigma(\mathbf{A}) \tag{6}$$



Slika 2: 3D prikaz pseudospektra matrice. Na z-osi nalazi se norma rezolvente matrice kao funkcija od $z \in \mathbb{C}$. Vidimo da je ona velika za brojeve blizu svojstvenim vrijednostima matrice.

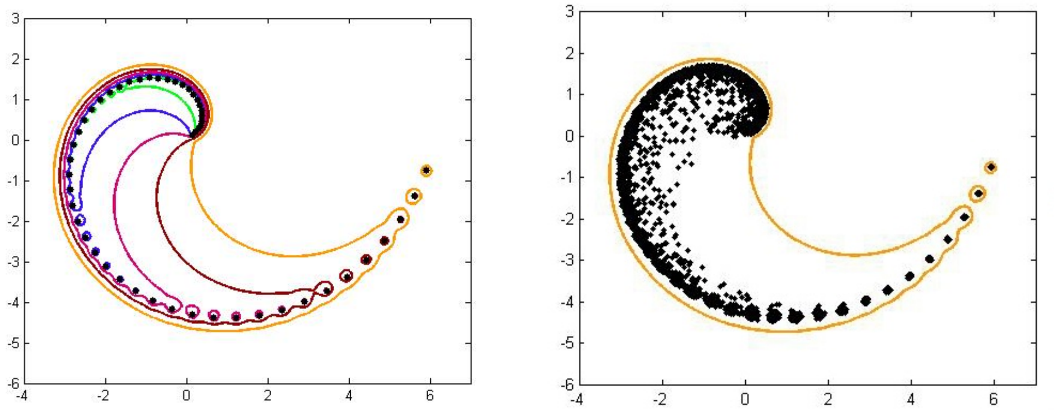
Definicija 1.3 (3. definicija pseudospektra). *$\sigma_\varepsilon(\mathbf{A})$ u normi $\|\cdot\|$ je skup svih $z \in \mathbb{C}$ t.d.*

$$\|(z - \mathbf{A})\mathbf{v}\| < \varepsilon \tag{7}$$

za neki $\mathbf{v} \in \mathbb{C}^N$ pri čemu je $\|\mathbf{v}\| = 1$.

Broj z u (7) zovemo ε-pseudo svojstvena vrijednost matrice $\mathbf{A}$, a $\mathbf{v}$ zovemo odgovarajući ε-pseudo svojstveni vektor. Drugim riječima, *ε-pseudospektar je skup svih ε-pseudo svojstvenih vrijednosti.*

Slika 3 prikazuje 50×50 Basor-Morrisonovu matricu i ilustrira ekvivalenciju prve i treće definicije pseudospektra za matricu koja nije normalna. Na lijevoj slici vidimo granice ε-pseudospektra za različite ε, a na desnoj svojstvene vrijednosti 100 slučajno perturbiranih matrica. Očito je da su svojstvene vrijednosti matrice iznimno osjetljive na perturbacije, a ovaj primjer također lijepo prikazuje moguću geometrijsku strukturu matrica u kompleksnoj ravnini koja isprva nije vidljiva iz samog spektra matrice.



Slika 3: Pseudospektar 50×50 Basor-Morrisonove matrice. Lijeva slika prikazuje svojstvene vrijednosti matrice (točke) i granice 2-norme ε-pseudospektra za $\varepsilon = 10^{-1}, 10^{-1.5}, 10^{-2}, 10^{-2.5}, 10^{-3}$. Desna slika prikazuje pozicije svojstvenih vrijednosti 100 slučajno perturbiranih matrica $\mathbf{A} + \mathbf{E}$, gdje svaka matrica $\mathbf{E}$ ima nezavisne normalno distribuirane kompleksne vrijednosti i vrijedi $\|\mathbf{E}\| = 10^{-1}$. Kada bi u obzir bile uzete sve moguće perturbacije uz $\|\mathbf{E}\| < \varepsilon^{-1}$, točke bi ispunile cijelo područje omeđeno najvećom granicom 2-norme s lijeve slike.

Slijedi teorem koji dokazuje da su ove tri definicije pseudospektra zaista ekvivalentne.

Teorem 1.4 (Ekvivalencija definicija pseudospektra). *Tri gore navedene definicije pseudospektra su ekvivalentne za proizvoljnu matricu $\mathbf{A} \in \mathbb{C}^{N \times N}$.*

Dokaz. Ako je $z \in \sigma(\mathbf{A})$ ekvivalencija je očita, pa ćemo pretpostaviti da $z \notin \sigma(\mathbf{A})$, što povlači da $(z - \mathbf{A})^{-1}$ zaista postoji.

Da bismo dokazali (4)⇒(7), pretpostavimo da vrijedi $(\mathbf{A} + \mathbf{E})\mathbf{v} = z\mathbf{v}$ za neki $\mathbf{E} \in \mathbb{C}^{N \times N}$ s $\|\mathbf{E}\| < \varepsilon$ i $\mathbf{v} \neq 0$, $\mathbf{v} \in \mathbb{C}^N$ (možemo pretpostaviti da je vektor $\mathbf{v}$ normaliziran, $\|\mathbf{v}\| = 1$). Tada vrijedi

$$\|(z - \mathbf{A})\mathbf{v}\| = \|\mathbf{E}\mathbf{v}\| < \varepsilon,$$

što je i trebalo dokazati.

Da bismo dokazali (7)⇒(1), pretpostavimo da vrijedi $(z - \mathbf{A})\mathbf{v} = s\mathbf{u}$ za neke $\mathbf{v}, \mathbf{u} \in \mathbb{C}^N$ takve da $\|\mathbf{v}\| = \|\mathbf{u}\| = 1$ i $0 < s < \varepsilon$. Tada je $(z - \mathbf{A})^{-1}\mathbf{u} = s^{-1}\mathbf{v}$ te s jedne strane vrijedi

$$\|(z-\mathbf{A})^{-1}\mathbf{u}\| \leq \|(z-\mathbf{A})^{-1}\|\,\|\mathbf{u}\| = \|(z-\mathbf{A})^{-1}\|,$$

a s druge strane imamo

$$\|(z-\mathbf{A})^{-1}\mathbf{u}\| = \|s^{-1}\mathbf{v}\| = s^{-1}. \tag{8}$$

Stoga

$$\|(z-\mathbf{A})^{-1}\| \geq s^{-1} > \varepsilon^{-1}.$$

Da bismo dokazali (1)⇒(4), pretpostavimo da je $\|(z-\mathbf{A})^{-1}\| > \varepsilon^{-1}$. Tada postoje $\mathbf{u}, \mathbf{v} \in \mathbb{C}^N$, $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ i $0 < s < \varepsilon$ takvi da vrijedi

$$(z-\mathbf{A})^{-1}\mathbf{u} = s^{-1}\mathbf{v},$$

iz čega slijedi

$$s\mathbf{u} = (z-\mathbf{A})\mathbf{v} = z\mathbf{v} - \mathbf{A}\mathbf{v}.$$

Želimo pokazati da je $z \in \sigma(\mathbf{A}+\mathbf{E})$ za neku matricu $\|\mathbf{E}\| \leq \varepsilon$. Za to je dovoljno pokazati da postoji matrica $\mathbf{E} \in \mathbb{C}^{N\times N}$ takva da vrijedi $\|\mathbf{E}\| \leq s$ i $\mathbf{E}\mathbf{v} = s\mathbf{u}$. Naime, ako takva matrica postoji, tada je $\mathbf{v}$ svojstveni vektor matrice $\mathbf{A}+\mathbf{E}$ s pripadnom svojstvenom vrijednošću z:

$$(\mathbf{A}+\mathbf{E})\mathbf{v} = z\mathbf{v}, \quad \|\mathbf{E}\| \leq \varepsilon.$$

Pokazat ćemo da za $\mathbf{E}$ možemo uzeti matricu ranga 1 oblika $\mathbf{E} = s\mathbf{u}\mathbf{w}^*$ za neki $\mathbf{w} \in \mathbb{C}^N$ t.d. $\mathbf{w}^*\mathbf{v} = 1$.

Ako je $\|\cdot\| = \|\cdot\|_2$ tvrdnja je očita uzimanjem $\mathbf{w} = \mathbf{v}$.

U slučaju neke druge norme $\|\cdot\|$, možemo se poslužiti korolarom Hahn–Banachova teorema (vidi npr. [5, str. 60.]) koji garantira postojanje linearnog funkcionala $\mathbf{L}$ na $\mathbb{C}^N$ za koji vrijedi $\|\mathbf{L}(\mathbf{v})\| = 1$ i $\|\mathbf{L}\| = 1$. Budući da su svi linearni funkcionali na $\mathbb{C}^N$ oblika $\mathbf{F}(\mathbf{x}) = \mathbf{y}^*\mathbf{x}$ za neki $\mathbf{y}$, znači da postoji vektor $\mathbf{w}$ takav da vrijedi $\mathbf{L}(\mathbf{x}) = \mathbf{w}^*\mathbf{x}$. Sada iz

$$\|\mathbf{E}\mathbf{x}\| = s\|\mathbf{u}\mathbf{w}^*\mathbf{x}\| = s\|\mathbf{u}\||\mathbf{L}(\mathbf{x})| \leq s\|\mathbf{x}\| \text{ za svaki } \mathbf{x}$$

i

$$\mathbf{E}\mathbf{v} = s\mathbf{u}\mathbf{w}^*\mathbf{v} = s\mathbf{u}\mathbf{L}(\mathbf{v}) = s\mathbf{u}$$

slijedi da $\mathbf{w}$ ima tražena svojstva. □

Prisjetimo se da su singularne vrijednosti matrice $\mathbf{A}$ svojstvene vrijednosti matrice $\sqrt{\mathbf{A}^*\mathbf{A}}$. Lako se vidi da u slučaju 2–norme vrijedi

$$\|(z-\mathbf{A})^{-1}\|_2 = [s_{\min}(z-\mathbf{A})]^{-1} \quad \text{i} \quad \|(z-\mathbf{A})\|_2 = [s_{\max}(z-\mathbf{A})], \tag{9}$$

gdje $s_{\min}(z - \mathbf{A})$ predstavlja najmanju, a $s_{\max}(z - \mathbf{A})$ najveću singularnu vrijednost matrice $z - \mathbf{A}$. Ovime dolazimo do četvrte i posljednje definicije pseudospektra u ovom radu.

Definicija 1.5 (4. definicija pseudospektra). *Za $\|\cdot\| = \|\cdot\|_2$, $\sigma_\varepsilon(\mathbf{A})$ je skup svih $z \in \mathbb{C}$ takvih da vrijedi*

$$s_{\min}(z - \mathbf{A}) < \varepsilon. \tag{10}$$

Iz (9) je očito da je (10) ekvivalentno s (1), a samim time i s ostalim definicijama pseudospektra prema teoremu 1.4. U dokazu teorema 1.4 imali smo matricu $\mathbf{E} = s\mathbf{u}\mathbf{v}^*$ ranga 1. Sada s možemo smatrati najmanjom singularnom vrijednošću, a $\mathbf{u}$ i $\mathbf{v}$ pripadnim lijevim, odnosno desnim singularnim vektorom matrice $z - \mathbf{A}$.

2 Pseudospektar normalnih matrica

Za početak primijetimo da ako je $\mathbf{U}$ unitarna matrica ($\mathbf{U}^* = \mathbf{U}^{-1}$), tada vrijedi

$$(z - \mathbf{U}\mathbf{A}\mathbf{U}^*)^{-1} = [\mathbf{U}(z - \mathbf{A})\mathbf{U}^*]^{-1} = \mathbf{U}(z - \mathbf{A})^{-1}\mathbf{U}^* \tag{11}$$

te zbog toga vrijedi i

$$\|(z - \mathbf{U}\mathbf{A}\mathbf{U}^*)^{-1}\|_2 = \|(z - \mathbf{A})^{-1}\|_2, \quad \forall z \in \mathbb{C}.$$

To povlači da je norma rezolvente invarijantna na unitarno slične transformacije, te da isto pravilo vrijedi i za pseudospektar:

$$\sigma_\varepsilon(\mathbf{A}) = \sigma_\varepsilon(\mathbf{U}\mathbf{A}\mathbf{U}^*), \quad \forall \varepsilon \geq 0. \tag{12}$$

Definicija 2.1. *Matrica $\mathbf{A} \in \mathbb{C}^{N \times N}$ je* normalna *ako ima potpun skup ortogonalnih svojstvenih vektora tj. ako je unitarno dijagonalizabilna:*

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^*, \tag{13}$$

gdje je $\mathbf{U}$ unitarna matrica, a Λ dijagonalna matrica sa svojstvenim vrijednostima na dijagonali.

Napomena 2.2. *Ekvivalentna definicija kaže da je matrica $\mathbf{A} \in \mathbb{C}^{N \times N}$* normalna *ako vrijedi $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$.*

Dakle, normalna matrica je matrica koja ima posebno svojstvo da postoji unitarna transformacija koja je transformira u dijagonalnu matricu. Za normalnu matricu, ε-pseudospektar je zapravo unija otvorenih ε-kugli oko točaka spektra matrice (vidi sliku 1). Tada norma rezolvente zadovoljava:

$$\|(z-\mathbf{A})^{-1}\|_2 = \frac{1}{\operatorname{dist}(z,\sigma(\mathbf{A}))}, \tag{14}$$

gdje $\operatorname{dist}(z,\sigma(\mathbf{A}))$ predstavlja uobičajenu udaljenost točke z od skupa $\sigma(\mathbf{A})$ u kompleksnoj ravnini.

Prije nego što iskažemo idući teorem, objasnit ćemo notaciju kojom ćemo se koristiti. Otvorenu ε-kuglu označavat ćemo s

$$\Delta_\varepsilon = \{z \in \mathbb{C} : |z| < \varepsilon\}. \tag{15}$$

Suma dvaju skupova imat će uobičajeno značenje te će vrijediti:

$$\sigma(\mathbf{A}) + \Delta_\varepsilon = \{z : z_1 + z_2, z_1 \in \sigma(\mathbf{A}), z_2 \in \Delta_\varepsilon\} = \{z : \operatorname{dist}(z,\sigma(\mathbf{A})) < \varepsilon\}.$$

Teorem 2.3 (Pseudospektar normalnih matrica). *Za bilo koju matricu* $\mathbf{A} \in \mathbb{C}^{N\times N}$ *vrijedi*

$$\sigma_\varepsilon(\mathbf{A}) \supseteq \sigma(\mathbf{A}) + \Delta_\varepsilon, \quad \forall\varepsilon > 0, \tag{16}$$

a ako je $\mathbf{A}$ *normalna i* $\|\cdot\| = \|\cdot\|_2$*, tada je*

$$\sigma_\varepsilon(\mathbf{A}) = \sigma(\mathbf{A}) + \Delta_\varepsilon, \quad \forall\varepsilon > 0. \tag{17}$$

Obratno, ako je $\|\cdot\| = \|\cdot\|_2$*, tada* (17) *povlači da je* $\mathbf{A}$ *normalna matrica.*

Dokaz. Ako je z svojstvena vrijednost matrice $\mathbf{A}$, tada je i $z+\delta$ svojstvena vrijednost od $\mathbf{A}+\delta\mathbf{I}$ za $\forall\delta \in \mathbb{C}$. Budući da je $\|\delta\mathbf{I}\| = |\delta|$, (16) vrijedi. Za dokazivanje (17) primijetimo da ako je $\mathbf{A}$ normalna, možemo bez smanjenja općenitosti pretpostaviti da je i dijagonalna. Pretpostavka neće imati nikakvog utjecaja na norme ako je $\|\cdot\| = \|\cdot\|_2$. Dijagonalni elementi matrice $\mathbf{A}$ tada su jednaki svojstvenim vrijednostima λ_j. Tada je i rezolventa dijagonalna matrica te zbog toga vrijedi (14). Definicija 1.1 pseudospektra tada povlači (17).

Za dokaz obrata definirajmo najprije skup $\tau_\varepsilon(\mathbf{A}) = \{z \in \mathbb{C} : \|(\mathbf{A}-z)^{-1}\|_2^{-1} < \varepsilon\}$. Tada znamo da vrijedi

$$\tau_\varepsilon(\mathbf{A}) = \sigma(\mathbf{A}) + \Delta_\varepsilon = \{z \in \mathbb{C} : \operatorname{dist}(z,\sigma(\mathbf{A})) < \varepsilon\}, \quad \forall\varepsilon > 0. \tag{18}$$

Također vrijedi

$$\bigcap_{n\in\mathbb{N}} \tau_{\varepsilon+\frac{1}{n}} = \bigcap_{n\in\mathbb{N}} \sigma(\mathbf{A}) + \Delta_{\varepsilon+\frac{1}{n}}.$$

Lako se vidi da vrijedi

$$\bigcap_{n\in\mathbb{N}} \tau_{\varepsilon+\frac{1}{n}} = \{z \in \mathbb{C} : \|(\mathbf{A}-z)^{-1}\|_2^{-1} \leq \varepsilon\}$$

te

$$\bigcap_{n\in\mathbb{N}} \sigma(\mathbf{A}) + \Delta_{\varepsilon+\frac{1}{n}} = \{z \in \mathbb{C} : \text{dist}(z, \sigma(\mathbf{A})) \leq \varepsilon\}.$$

Stoga vrijedi i

$$\tau_\varepsilon^c \bigcap \left(\bigcap_{n\in\mathbb{N}} \tau_{\varepsilon+\frac{1}{n}}\right) = (\sigma(\mathbf{A}) + \Delta_\varepsilon)^c \bigcap \left(\bigcap_{n\in\mathbb{N}} \sigma(\mathbf{A}) + \Delta_{\varepsilon+\frac{1}{n}}\right),$$

što se može zapisati i kao

$$\{z \in \mathbb{C} : \|(\mathbf{A}-z)^{-1}\|_2^{-1} = \varepsilon\} = \{z \in \mathbb{C} : \text{dist}(z, \sigma(\mathbf{A})) = \varepsilon\} \quad \forall\varepsilon > 0. \tag{19}$$

Neka je $\mathrm{A} = \mathbf{ULU}^*$ Schurova dekompozicija matrice $\mathbf{A}$, gdje je $\mathbf{L}$ donjetrokutasta matrica ([4]). Tada su dijagonalni elementi matrice $\mathbf{L}$ svojstvene vrijednosti od $\mathbf{A}$, te vrijedi $(\mathbf{A}-z)^{-1} = (\mathbf{ULU}^* - z)^{-1} = (\mathbf{U}(\mathbf{L}-z)\mathbf{U}^*)^{-1} = \mathbf{U}(\mathbf{L}-z)^{-1}\mathbf{U}^*$ te stoga vrijedi i

$$\|(\mathbf{A}-z)^{-1}\|_2 = \|(\mathbf{L}-z)^{-1}\|_2. \tag{20}$$

Ako pokažemo da je $\mathbf{L} = [\varphi_{ij}]$ dijagonalna matrica, onda smo gotovi jer je tada očito $\mathbf{A}$ normalna matrica.

Odaberemo proizvoljan $i_0 \in \{1, \ldots, N\}$ te neke $z \in \mathbb{C}$ i $\varepsilon > 0$ takve da $\min_i |\varphi_{ii} - z| = |\varphi_{i_0 i_0} - z| = \varepsilon$. To znači da vrijedi $\text{dist}(z, \sigma(\mathbf{A})) = \varepsilon$, te stoga po pretpostavci vrijedi $\|(\mathbf{A}-z)^{-1}\|_2^{-1} = \varepsilon$, tj.

$$\|(\mathbf{L}-z)^{-1}\|_2 = \frac{1}{\varepsilon}. \tag{21}$$

Izračunajmo sada $\|(\mathbf{L}-z)^{-1}\|_2$. Znamo da vrijedi $\|(\mathbf{L}-z)^{-1}\|_2 = \lambda_{\text{max}}[(\mathbf{L}-z)^{-1}(\mathbf{L}-z)^{-*}]$, a kako je

$$\mathbf{L} - z = \begin{bmatrix} \varphi_{11} - z & 0 & \ldots & 0 \\ \varphi_{12} & \varphi_{22} - z & \ddots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \varphi_{1n} & \varphi_{2n} & \cdots & \varphi_{nn} - z \end{bmatrix},$$

lako je vidjeti da je inverz matrice $\mathbf{L} - z$ ponovno donjetrokutasta matrica oblika

$$(\mathbf{L}-z)^{-1} = \begin{bmatrix} (\varphi_{11}-z)^{-1} & 0 & \cdots & 0 \\ * & (\varphi_{22}-z)^{-1} & \ddots & 0 \\ \vdots & \vdots & \ddots & 0 \\ * & * & \cdots & (\varphi_{nn}-z)^{-1} \end{bmatrix}. \tag{22}$$

Nadalje, $(\mathbf{L}-z)^{-1}(\mathbf{L}-z)^{-*}$ je hermitska matrica, što povlači $\sigma(\mathbf{A}) \geq \operatorname{diag}((\mathbf{L}-z)^{-1}(\mathbf{L}-z)^{-*})$ ([4]). Stoga

$$\begin{aligned} \|(\mathbf{L}-z)^{-1}\|_2^2 &\geq \max_i \left((\mathbf{L}-z)^{-1}(\mathbf{L}-z)^{-*} \right)_{ii} \\ &= \max_i \left(\sum_{j=1}^{i-1} \left| \left((\mathbf{L}-z)^{-1} \right)_{ij} \right|^2 + |\varphi_{ii}-z|^{-2} \right) \\ &\geq |\varphi_{i_0 i_0}|^{-2} + \sum_{j=1}^{i_0-1} \left| \left((\mathbf{L}-z)^{-1} \right)_{i_0 j} \right|^2 \\ &= \frac{1}{\varepsilon^2} + \sum_{j=1}^{i_0-1} \left| \left((\mathbf{L}-z)^{-1} \right)_{i_0 j} \right|^2. \end{aligned}$$

Budući da je $\|(\mathbf{L}-z)^{-1}\|_2^2 = \frac{1}{\varepsilon^2}$, slijedi da je $((\mathbf{L}-z)^{-1})_{i_0 j} = 0$ za $j = 1, \ldots, i_0 - 1$. Budući da je i_0 bio proizvoljno odabran, slijedi da je $(\mathbf{L}-z)^{-1}$ dijagonalna matrica, što povlači da je $\mathbf{L}$ dijagonalna matrica te smo dokazali tvrdnju.

□

3 Svojstva pseudospektra

Pretpostavimo da je matrica $\mathbf{A}$ dijagonalizabilna, ali ne nužno i normalna. Neka $\mathbf{V} \in \mathbb{C}^{N \times N}$ predstavlja matricu svojstvenih vrijednosti matrice $\mathbf{A}$, pri čemu vrijedi:

$$\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^{-1},$$

gdje je Λ dijagonalna $N \times N$ matrica sa svojstvenim vrijednostima λ_j na dijagonali. Ako je $\|\cdot\| = \|\cdot\|_2$, onda je kondicijski broj ove baze svojstvenih vektora dan s

$$\kappa(\mathbf{V}) \equiv \|\mathbf{V}\|_2 \|\mathbf{V}^{-1}\|_2 = \frac{s_{\max}(\mathbf{V})}{s_{\min}(\mathbf{V})}, \tag{23}$$

pri čemu su $s_{\max}(\mathbf{V})$ i $s_{\min}(\mathbf{V})$ najveća i najmanja singularna vrijednost matrice $\mathbf{V}$. Budući da matrica $\mathbf{V}$ nije jedinstvena, $\kappa(\mathbf{V})$ nije jedinstveno definiran za danu matricu $\mathbf{A}$. No, ako su sve svojstvene vrijednosti matrice $\mathbf{A}$ različite, $\kappa(\mathbf{V})$ postaje jedinstven ako normaliziramo svojstvene vektore ($\|\mathbf{v}_j\| = 1$).

Općenito, za $\kappa(\mathbf{V})$ vrijedi $1 \leq \kappa(\mathbf{V}) < \infty$, a $\kappa(\mathbf{V}) = 1$ ako i samo ako je matrica $\mathbf{A}$ normalna. Kondicija matrice $\mathbf{V}$ daje nam gornju granicu za kondicije pojedinačnih svojstvenih vrijednosti matrice $\mathbf{A}$. Ovime dolazimo do Bauer-Fikeova teorema.

Teorem 3.1 (Bauer-Fikeov teorem). *Neka je $\mathbf{A} \in \mathbb{C}^{N\times N}$ dijagonalizabilna matrica, $\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^{-1}$. Tada, uz $\|\cdot\| = \|\cdot\|_2$, $\forall\varepsilon > 0$ vrijedi*

$$\sigma(\mathbf{A}) + \Delta_\varepsilon \subseteq \sigma_\varepsilon(\mathbf{A}) \subseteq \sigma(\mathbf{A}) + \Delta_{\varepsilon\kappa(\mathbf{V})}. \tag{24}$$

Dokaz. Prva inkluzija je dokazana u (16). Za dokazivanje druge inkluzije u tvrdnji teorema računamo:

$$(z - \mathbf{A})^{-1} = (z - \mathbf{V}\Lambda\mathbf{V}^{-1})^{-1} = [\mathbf{V}(z - \Lambda)\mathbf{V}^{-1}]^{-1} = \mathbf{V}(z - \Lambda)^{-1}\mathbf{V}^{-1},$$

što povlači

$$\|(z - \mathbf{A})^{-1}\|_2 \leq \kappa(\mathbf{V})\|(z - \Lambda)^{-1}\|_2 = \frac{\kappa(\mathbf{V})}{\text{dist}(z, \sigma(\mathbf{A}))}.$$

Sada dokaz slijedi iz definicije 1.1.

□

Sljedeći teorem navodi neka od osnovnih svojstava pseudospektra.

Teorem 3.2 (Svojstva pseudospektra). *Neka je $\mathbf{A} \in \mathbb{C}^{N\times N}$ i $\varepsilon > 0$ proizvoljan.*

1. *$\sigma_\varepsilon(\mathbf{A})$ je neprazan, otvoren i ograničen skup, s najviše N komponenti povezanosti, od kojih svaka komponenta sadržava jednu ili više svojstvenih vrijednosti.*

2. *Ako je $\|\cdot\| = \|\cdot\|_2$, tada $\sigma_\varepsilon(\mathbf{A}^*) = \overline{\sigma_\varepsilon(\mathbf{A})}$.*

3. *Ako je $\|\cdot\| = \|\cdot\|_2$, tada $\sigma_\varepsilon(\mathbf{A}_1 \oplus \mathbf{A}_2) = \sigma_\varepsilon(\mathbf{A}_1) \cup \sigma_\varepsilon(\mathbf{A}_2)$.*

4. *Za proizvoljan $c \in \mathbb{C}$ vrijedi $\sigma_\varepsilon(\mathbf{A} + c) = c + \sigma_\varepsilon(\mathbf{A})$.*

5. *Za proizvoljan $c \in \mathbb{C}, c \neq 0$ vrijedi $\sigma_{|c|\varepsilon}(c\mathbf{A}) = c\sigma_\varepsilon(\mathbf{A})$.*

U dijelu (iii), $\mathbf{A}_1 \oplus \mathbf{A}_2$ predstavlja direktnu sumu dviju kvadratnih matrica. Pri tome matrice ne moraju biti istih dimenzija, a njihova direktna suma je blok dijagonalna matrica

$$\mathbf{A}_1 \oplus \mathbf{A}_2 = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}.$$

Prije nego što započnemo s dokazom teorema, navest ćemo važan rezultat o subharmoničnosti rezolvente kojima ćemo se koristiti u samom dokazu.

Definicija 3.3 ([1]). *Neka je U otvoren podskup od $\mathbb{C}$ i $f : U \to \mathbb{R}$ neprekidna funkcija. Kažemo da je f* subharmonička funkcija na *U ako za svaku zatvorenu kuglu $\overline{K}(a, r) \subset U$ sa središtem u a i radijusa r vrijedi*

$$f(a) \leq \frac{1}{2\pi} \int_0^{2\pi} f(a + re^{i\theta}) d\theta.$$

Teorem 3.4 ([1], princip maksimuma). *Ako je S ograničen podskup od $\mathbb{C}$ i $f : \overline{S} \to \mathbb{R}$ neprekidna funkcija koja je subharmonička na S, tada je*

$$\sup f(S) = \sup f(\partial S). \tag{25}$$

Teorem 3.5 ([3]). *Ako je f holomorfna funkcija na S, tada je $\|f(\cdot)\|$ subharmonička funkcija na S. Slijedi da $\|f(\cdot)\|$ može imati maksimum na skupu S samo ako je konstantne vrijednosti na cijelom skupu S.*

Korolar 3.6 ([1]). *Norma rezolvente $\|(z - \mathbf{A})^{-1}\|$ je subharmonička funkcija za $z \notin \sigma(\mathbf{A})$, što povlači da zadovoljava princip maksimuma. Također vrijedi*

$$\|(z - \mathbf{A})^{-1}\| \geq \frac{1}{\operatorname{dist}(z, \sigma(\mathbf{A}))}. \tag{26}$$

Dokaz teorema 3.2. Dokazi tvrdnji (ii), (iii) i (iv) su jednostavni pa ih nećemo navoditi.

(i)

Nepraznost, otvorenost i ograničenost skupa $\sigma_\varepsilon(\mathbf{A})$ već smo dokazali. Preostaje nam pokazati da se $\sigma_\varepsilon(\mathbf{A})$ sastoji od najviše N komponenti povezanosti, od kojih svaka sadržava jednu ili više svojstvenih vrijednosti. Koristit ćemo se gore navedenim rezultatom.

Pretpostavimo da unutar neke komponente povezanosti nema svojstvenih vrijednosti. Tada je rezolventa holomorfna na tom skupu te je norma rezolvente subharmonička funkcija. Po principu maksimuma, supremum te funkcije dostiže se na rubu te komponente povezanosti. No u našem je slučaju rub podskup skupa

$\{z : \|(\mathbf{A}-z)^{-1}\| = \frac{1}{\varepsilon}\}$ te stoga dolazimo do kontradikcije s činjenicom da je komponenta povezanosti podskup skupa $\{z : \|(\mathbf{A}-z)^{-1}\| > \frac{1}{\varepsilon}\}$.

Ovime smo tvrdnju dokazali za ograničene komponente povezanosti. Neograničenih komponenti povezanosti uopće ni nema, budući da $\|(\mathbf{A}-z)^{-1}\| \to 0$ kad $|z| \to \infty$, te stoga zaključujemo da nema komponenti povezanosti koje ne sadržavaju barem jednu svojstvenu vrijednost matrice $\mathbf{A}$. Budući da $\mathbf{A}$ ima najviše N različitih svojstvenih vrijednosti, komponenti povezanosti također može biti najviše N.

(v)

Pretpostavimo da je $z \in \sigma_{|c|\varepsilon}(c\mathbf{A})$. Tvrdnja je očigledna ako je $c = 0$ ili $\varepsilon = 0$. Stoga, uz $c \neq 0$ i $\varepsilon \neq 0$, definicija 1.1 pseudospektra povlači da je

$$(|c|\varepsilon)^{-1} < \|(z - c\mathbf{A})^{-1}\| = |c|^{-1}\|(\frac{z}{c} - \mathbf{A})^{-1}\|.$$

Sada je jasno da je $\frac{z}{c} \in \sigma_\varepsilon(\mathbf{A})$, odnosno $z \in c\sigma_\varepsilon(\mathbf{A})$, što je i trebalo pokazati. $\square$

4 Primjeri

Promatrat ćemo tridijagonalnu Toeplitzovu matricu

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & & & \\ \frac{1}{4} & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{4} & 0 & 1 \\ & & & \frac{1}{4} & 0 \end{bmatrix} \in \mathbb{C}^{N\times N}. \tag{27}$$

Iako je ova matrica nesimetrična, možemo je svesti na simetričan oblik sljedećom transformacijom

$$\mathbf{DAD}^{-1} = \mathbf{S}, \tag{28}$$

gdje je $\mathbf{D} = \mathrm{diag}(2, 4, ..., 2^N)$ i

$$\mathbf{S} = \begin{bmatrix} 0 & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{2} & 0 & \frac{1}{2} \\ & & & \frac{1}{2} & 0 \end{bmatrix} \in \mathbb{C}^{N\times N}. \tag{29}$$

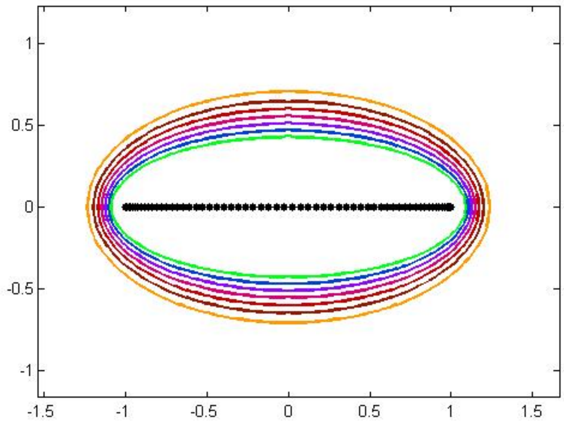
Matrice **A** i **S** imaju iste svojstvene vrijednosti te se može izračunati da se spektar matrice **A** sastoji od N različitih svojstvenih vrijednosti iz intervala $\langle -1, 1\rangle$

$$\lambda_k(\mathbf{A}) = \lambda_k(\mathbf{S}) = \cos\frac{k\pi}{N+1}, \quad 1 \leq k \leq N. \tag{30}$$

S druge strane, pseudospektar matrice **A** znatno odudara od realne osi za različite vrijednosti ε. Slika 4 prikazuje granice $\sigma_\varepsilon(\mathbf{A})$ za $\varepsilon = 10^{-2}, 10^{-3}, ..., 10^{-8}$ uz $\|\cdot\| = \|\cdot\|_2$ i otkriva široka ovalna područja pseudospektra u kompleksnoj ravnini. ε-pseudospektar zapravo je približno jednak području koje omeđuje elipsa koja je slika kružnice $|z| = \varepsilon^{1/N}$ pri preslikavanju

$$f(z) = z^{-1} + \frac{1}{4}z. \tag{31}$$

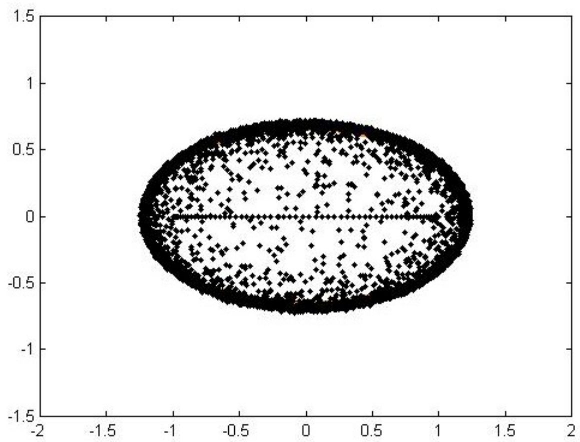
Slika 5 prikazuje svojstvene vrijednosti 100 slučajno perturbiranih matrica $\mathbf{A}+\mathbf{E}$, pri čemu vrijedi $\|\mathbf{E}\| = 10^{-3}$ uz $\|\cdot\| = \|\cdot\|_2$. Veza s elipsama ponovno je jasno vidljiva.



Slika 4: Granice ε-pseudospektra $\sigma_\varepsilon(\mathbf{A})$ u 2-normi za $\varepsilon = 10^{-2}, 10^{-3}, ..., 10^{-8}$ matrice (27) dimenzije $N = 64$. Svojstvene vrijednosti matrice prikazane su crnim točkama.

5 Povijesni pregled i literatura

Slučajevi u kojima svojstvene vrijednosti i spektar matrice ne daju zadovoljavajuće odgovore na postavljena pitanja usko su vezani uz matrice koje nemaju svojstvo



Slika 5: Pozicije svojstvenih vrijednosti 100 matrica $\mathbf{A} + \mathbf{E}$, gdje je $\mathbf{A}$ tridijagonalna Toeplitzova matrica (27) dimenzije $N = 64$, a svaka $\mathbf{E}$ je slučajna matrica uz $\|\mathbf{E}\| = 10^{-3}$. Svojstvene vrijednosti matrice $\mathbf{A}$ su realne, ali perturbacije inducirane matricom $\mathbf{E}$ ih pomiču u kompleksnu ravninu, blizu elipse definirane s (31) uz $|z| = 0.001^{1/64}$. Ponovno je $\|\cdot\| = \|\cdot\|_2$.

normalnosti. Potrebu za istraživanjem pseudospektra prepoznao je von Neumann još u 1930–ima, no zbog teškoća s velikim brojem složenih računskih operacija i nedovoljnog stupnja razvijenosti računala koja bi ih mogla izvesti, pojam pseudospektra nije zaživio sve do druge polovine 20. stoljeća. Tako na potrebu za istraživanjem pseudospektra ponovno upućuju Varah 1967. i 1979. ([12] i [11]), Landau 1975. ([6]) i Godunov 1982. ([2]). Prvi rad potpuno posvećen pseudospektru pod nazivom „Pseudospectra of matrices" napisao je Trefethen 1992. godine ([9]). U tom je radu iznio osnovnu teoriju pseudospektra popraćenu s 13 primjera. Trefethen nastavlja istraživanje pseudospektra te nekoliko godina poslije izdaje i rad „Pseudospectra of Linear Operators" ([8]) kojim produbljuje teoriju iza ideje pseudospektra i proširuje područje njegove primjene. Tim radovima pseudospektar postaje poznat široj javnosti te se javlja znatno veći interes među znanstvenicima za razradom ovog relativno novog pojma u matematici.

Ovaj članak temelji se na diplomskom radu drugog autora, napisanom pod vodstvom prvog autora [7], koji se pak temelji na knjizi [10].

[1] Burkcel, R. B.: *An introduction to classical complex analysis*, Academic Press, New York, 1979

[2] Godunov, S. K.: *Guaranteed accuracy in numerical linear algebra*, Kluwer Academic Publishers, 1993

[3] Hille, E, i Phillips, R. S.: *Functional analysis and semi-groups*, AMS, Providence, 1957

[4] Horn, R. A. i Johnson C. R.: *Matrix analysis*, Cambridge University Press, 1985

[5] Kurepa, S.: *Funkcionalna analiza*, Školska knjiga, Zagreb, 1990.

[6] Landau, H. J.: On Szegõ's eigenvalue distribution theorem and non-Hermitian kernels, *J. d'Analyse Math.*, no. 28, str. 335.–357., 1975

[7] Petrić, M.: *Pseudospektar*, Diplmoski rad, PMF–MO, Sveučilište u Zagrebu, 2010.

[8] Trefethen, L. N.: Pseudospectra of Linear Operators, *SIAM review*, vol. 39, no. 3, str. 383.–406., 1997

[9] Trefethen, L. N.: Pseudospectra of matrices, u Griffiths, D. F. i Watson, G. A. (ur.):*Numerical Analysis 1991*, str. 234.–266., 1992

[10] Trefethen, L. N. i Embree, M.: *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*, Princeton University Press, Princeton, 2005

[11] Varah, J. M.: On the Separation of Two Matrices, *SIAM Journal on Numerical Analysis*, vol. 16, no. 2, str. 216.–222., 1979

[12] Varah, J. M.: The computation of bounds for the invariant subspaces of a general matrix operator, Technical report CS 66, Computer Science Department, Stanford University, 1967