Računanje izotropnih vektorjev

Avtor: Mirjam Pergar
Mentor: izred. prof. dr. Bor Plestenjak

24. november 2015

Vsebina

Vsebina

Vsebina

Vsebina

Problem

Za dano nesingularno, kvadratno $n \times n$ matriko A z realnimi ali kompleksnimi elementi, nas zanima izračun enotskega vektorja b z realnimi ali kompleksnimi elementi, tako da velja:

$$b^* Ab = 0. \tag{1}$$

Vektor b, za katerega velja (1) in $b^* b = 1$ imenujemo **izotropni vektor**. Bolj splošen je inverzni problem zaloge vrednosti, kjer iščemo enotski vektor b, za katerega velja:

$$b^* Ab = \mu, \tag{2}$$

kjer je μ dano kompleksno število.

- Problem (2) je možno prevesti na problem (1) za drugo matriko, saj je (2) enako

$$b^*(A - \mu I)b = 0.$$

- Če μ lastna vrednost matrike A, potem je rešitev pripadajoč lastni vektor matrike A.
- Če μ ni lastna vrednost matrike A, je $A - \mu I$ nesingularna in je potreben izračun izotropnega vektorja te matrike.
- Od sedaj naprej bomo vse vrednosti enačili z 0.

Zaloga vrednosti

Definicija

Zaloga vrednosti *matrike $A \in \mathbb{C}^{n\times n}$ je podmnožica kompleksne ravnine, definirana kot*

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

Zaloga vrednosti

Definicija

Zaloga vrednosti *matrike $A \in \mathbb{C}^{n\times n}$ je podmnožica kompleksne ravnine, definirana kot*

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

- Očitno je $W(A)$ množica vseh Rayleighovih kvocientov matrike A.
- Izhodišče mora biti v $W(A)$, če hočemo, da ima (1) vsaj eno rešitev.
- Označimo s $\sigma(A)$ množico vseh lastnih vrednosti matrike A, ki jo imenujemo spekter.

Lastnosti zaloge vrednosti

1. $W(A)$ je konveksna, zaprta in omejena.
2. $\sigma(A) \subseteq W(A)$.
3. Za vsako unitarno matriko U je $W(U^*AU) = W(A)$.
4. $W(A + zI) = W(A) + z$ in $W(zA) = zW(A)$ za vsako kompleksno število z.
5. Rob zaloge vrednosti $W(A), \partial W(A)$ je kosoma algebrajska krivulja, in vsaka točka v kateri $\partial W(A)$ ni diferenciabilna je lastna vrednost matrike A.
6. Če je A normalna, potem $W(A) = \mathrm{Co}(\sigma(A))$, kjer s Co označimo zaprto konveksno ogrinjačo množice.
7. $W(A)$ je daljica na realni osi, če in samo če je A hermitska.

Izrek

Naj imata A in b realne ali kompleksne elemente. Potem veljajo enakosti:

$$b^* Ab = 0 \Leftrightarrow b^*(A + A^*)b = 0 \quad \mathit{in} \quad b^*(A - A^*)b = 0.$$

Izrek

Naj imata A in b realne ali kompleksne elemente. Potem veljajo enakosti:

$$b^* A b = 0 \Leftrightarrow b^*(A + A^*)b = 0 \quad in \quad b^*(A - A^*)b = 0.$$

Dokaz.

($\Rightarrow$) Če velja $b^* A b = 0$, je tudi $(b^* A b)^* = b^* A^* b = 0$.
Preoblikujemo prvo enačbo na desni v $b^* A b + b^* A^* b$ dobimo 0.
Drugo enačbo dokažemo na podoben način.
($\Leftarrow$) S seštevkom enačb na desni dobimo enačbo na levi:

$$\begin{aligned} b^*(A + A^*)b + b^*(A - A^*)b &= 0 \\ b^*(2A)b &= 0 \\ b^* A b &= 0 \end{aligned}$$

$\square$

- Če velja le $b^*(A + A^*)b = 0 \Rightarrow \Re(b^*Ab) = 0$.
- Če velja le $b^*(A - A^*)b = 0 \Rightarrow \Im(b^*Ab) = 0$.
- Označimo z $A_{sim} = A^T$ simetrični del matrike A in z $A_{psim} = -A^T$ poševno-simetrični del matrike A.

- Če velja le $b^*(A + A^*)b = 0 \Rightarrow \Re(b^*Ab) = 0$.
- Če velja le $b^*(A - A^*)b = 0 \Rightarrow \Im(b^*Ab) = 0$.
- Označimo z $A_{sim} = A^T$ simetrični del matrike A in z $A_{psim} = -A^T$ poševno-simetrični del matrike A.

Lema

Izotropni vektorji matrike A so identični izotropnim vektorjem njenega simetričnega dela.

- Če velja le $b^*(A+A^*)b = 0 \Rightarrow \Re(b^*Ab) = 0$.
- Če velja le $b^*(A-A^*)b = 0 \Rightarrow \Im(b^*Ab) = 0$.
- Označimo z $A_{sim} = A^T$ simetrični del matrike A in z $A_{psim} = -A^T$ poševno-simetrični del matrike A.

Lema

Izotropni vektorji matrike A so identični izotropnim vektorjem njenega simetričnega dela.

Dokaz.

To sledi iz $b^TAb = b^TA_{sim}b + b^TA_{psim}b = b^TA_{sim}b$, kjer je z A_{sim} označen simetrični del matrike A in z A_{psim} poševno-simetrični del matrike A. □

Velja enakost:

$$b^T Ab = 0 \Leftrightarrow b^T(A + A^T)b = 0.$$

Velja enakost:

$$b^T A b = 0 \Leftrightarrow b^T (A + A^T) b = 0.$$

- Hermitski del matrike A bomo označili s $H = (A + A^*)/2$.
- Poševno-hermitski del matrike A bomo označili z $\tilde{K} = (A - A^*)/2 = \imath K$.

Uporaba

- Preučevanje konvergence nekaterih iterativnih metod za reševanje linearnih sistemov, npr. GMRES.
- Aplikacije v numerični analizi, diferencialnih enačbah, teoriji sistemov itd.

Realne matrike

Ko je A realna matrika, nas zanima kako se izračuna rešitev enačbe:

$$b^* H b = 0, \tag{3}$$

kjer je H realna in simetrična matrika (t.j. $H = H^T$).

Realne matrike

Ko je A realna matrika, nas zanima kako se izračuna rešitev enačbe:

$$b^* H b = 0, \tag{3}$$

kjer je H realna in simetrična matrika (t.j. $H = H^T$).

Vemo:

- $W(A)$ simetrična glede na realno os.
- $0 \in W(A)$, če in samo če $\lambda_n \leq 0 \leq \lambda_1$, kjer sta λ_n in λ_1 najmanjša in največja lastna vrednost matrike H.

Naj bosta x_1 in x_n realna lastna vektorja, pripadajoča λ_1 in λ_n. Potem sta:

- $x_1^T A x_1 = x_1^T H x_1 = \lambda_1$.
- $x_n^T A x_n = x_n^T H x_n = \lambda_n$.

realni točki na skrajni levi in skrajni desni zaloge vrednosti $W(A)$ na realni osi.

Naj bosta x_1 in x_n realna lastna vektorja, pripadajoča λ_1 in λ_n. Potem sta:

- $x_1^T A x_1 = x_1^T H x_1 = \lambda_1$.
- $x_n^T A x_n = x_n^T H x_n = \lambda_n$.

realni točki na skrajni levi in skrajni desni zaloge vrednosti $W(A)$ na realni osi.

- Realne rešitve (3) izračunamo z uporabo lastnih vektorjev matrike H.
- Predpostavimo, da iščemo vektorje b z normo 1.
- H zapišemo kot
$$H = X \Lambda X^T,$$
kjer je Λ matrika, ki ima na diagonali lastne vrednosti λ_i, ki so realna števila in X je ortogonalna matrika lastnih vektorjev, tako da $X^T X = I$.

- Uporabimo ta spektralni razcep v (3):

$$b^* Hb = b^* X \Lambda X^T b = 0.$$

- Označimo s $c = X^T b$ vektor projekcije b na lastne vektorje matrike H.

- Uporabimo ta spektralni razcep v (3):

$$b^* H b = b^* X \Lambda X^T b = 0.$$

- Označimo s $c = X^T b$ vektor projekcije b na lastne vektorje matrike H.

Izrek

Naj bo b rešitev problema (3)*. Potem vektor $c = X^T b$ s komponentami c_i zadošča naslednjima enačbama:*

$$\sum_{i=1}^{n} \lambda_i |c_i|^2 = 0, \tag{4}$$

$$\sum_{i=1}^{n} |c_i|^2 = 1. \tag{5}$$

Dokaz.

Enačbo (4) dokažemo tako, da $c = X^T b$ oz. $c^* = b^* X$ vstavimo v (3) in dobimo

$$b^* H b = b^* X \Lambda X^T b = c^* \Lambda c = 0.$$

Ker je Λ diagonalna matrika, lahko $c^* \Lambda c$ zapišemo kot vsoto komponent $\bar{c}_i \lambda_i c_i = \lambda_i |c_i|^2$, ko $i = 1, 2, ...n$. Za enačbo (5) vemo, da je $\|b\|_2 = 1$. Če normo zapišemo s c dobimo

$$\|b\|_2 = \|Xc\|_2 = \|c\|_2 = 1,$$

saj je X ortogonalna matrika. $\square$

Iskanje izotropnih vektorjev

Če predpostavimo, da nimajo vse lastne vrednosti H enakega predznaka, potem mora za najmanjšo lastno vrednost λ_n veljati $\lambda_n < 0$. Naj bo $k < n$ tak, da je $\lambda_k > 0$ in $0 < t < 1, t \in \mathbb{R}$. Označimo $|c_n|^2 = t, |c_k|^2 = 1 - t$ in $c_i = 0, i \neq n, k$, kar velja zaradi enačbe (5), $t + (1 - t) = 1$. Iz (4) mora veljati enačba:

$$\lambda_n t + \lambda_k (1 - t) = 0,$$

katere rešitev je:

$$t_s = \frac{\lambda_k}{\lambda_k - \lambda_n}. \tag{6}$$

- Absolutna vrednost c_n (oz. c_k) je kvadratni koren od t_s (oz. $1 - t_s$).
- Ker je $b = Xc$, sta dve realni rešitvi:

$$b_1 = \sqrt{t_s} x_n + \sqrt{1 - t_s} x_k, \quad b_2 = -\sqrt{t_s} x_n + \sqrt{1 - t_s} x_k,$$

kjer sta x_n in x_k lastna vektorja pripadajoča λ_n in λ_k.

- Absolutna vrednost c_n (oz. c_k) je kvadratni koren od t_s (oz. $1 - t_s$).
- Ker je $b = Xc$, sta dve realni rešitvi:

$$b_1 = \sqrt{t_s}x_n + \sqrt{1 - t_s}x_k, \quad b_2 = -\sqrt{t_s}x_n + \sqrt{1 - t_s}x_k,$$

kjer sta x_n in x_k lastna vektorja pripadajoča λ_n in λ_k.

- Ker imata izraza v rešitvah enaka imenovalca, lahko rešitvi zapišemo kot:

$$b_1 = \sqrt{\lambda_k}x_n + \sqrt{|\lambda_n|}x_k, \quad b_2 = -\sqrt{\lambda_k}x_n + \sqrt{|\lambda_n|}x_k$$

.

- Absolutna vrednost c_n (oz. c_k) je kvadratni koren od t_s (oz. $1 - t_s$).
- Ker je $b = Xc$, sta dve realni rešitvi:

$$b_1 = \sqrt{t_s} x_n + \sqrt{1 - t_s} x_k, \quad b_2 = -\sqrt{t_s} x_n + \sqrt{1 - t_s} x_k,$$

kjer sta x_n in x_k lastna vektorja pripadajoča λ_n in λ_k.
- Ker imata izraza v rešitvah enaka imenovalca, lahko rešitvi zapišemo kot:

$$b_1 = \sqrt{\lambda_k} x_n + \sqrt{|\lambda_n|} x_k, \quad b_2 = -\sqrt{\lambda_k} x_n + \sqrt{|\lambda_n|} x_k$$

.
- Vektor mora biti normiran:

$$b_1 = \sqrt{\frac{\lambda_k}{\lambda_k + |\lambda_n|}} x_n + \sqrt{\frac{|\lambda_n|}{\lambda_k + |\lambda_n|}} x_k,$$

$$b_2 = -\sqrt{\frac{\lambda_k}{\lambda_k + |\lambda_n|}} x_n + \sqrt{\frac{|\lambda_n|}{\lambda_k + |\lambda_n|}} x_k.$$

Število rešitev

Uporabimo lahko vsak par pozitivnih in negativnih lastnih vrednosti. Ta postopek lahko vrne toliko rešitev kot je dvakratno število parov lastnih vrednosti matrike H z nasprotnimi predznaki, če so vse lastne vrednosti različne. Predpostavimo, da je b realen.

Število rešitev

Uporabimo lahko vsak par pozitivnih in negativnih lastnih vrednosti. Ta postopek lahko vrne toliko rešitev kot je dvakratno število parov lastnih vrednosti matrike H z nasprotnimi predznaki, če so vse lastne vrednosti različne. Predpostavimo, da je b realen.

Posledica

Dobljena izotropna vektorja sta ortogonalna ($b_1^T b_2 = 0$), če in samo če $\lambda_k = -\lambda_n$.

Število rešitev

Uporabimo lahko vsak par pozitivnih in negativnih lastnih vrednosti. Ta postopek lahko vrne toliko rešitev kot je dvakratno število parov lastnih vrednosti matrike H z nasprotnimi predznaki, če so vse lastne vrednosti različne. Predpostavimo, da je b realen.

Posledica

Dobljena izotropna vektorja sta ortogonalna ($b_1^T b_2 = 0$), če in samo če $\lambda_k = -\lambda_n$.

Dokaz.

$$b_1^T b_2 = (\sqrt{\lambda_k} x_n + \sqrt{|\lambda_n|} x_k)^T (-\sqrt{\lambda_k} x_1 + \sqrt{|\lambda_n|} x_k) = -(\lambda_n + \lambda_k).$$

□

Neskončno rešitev

Ko sta A in b realna smo dokazali naslednji izrek:

Izrek

Če je A realna in nedefinitna (t.j. ni pozitivno in negativno definitna), potem obstajata najmanj dva neodvisna realna izotropna vektorja.

Neskončno rešitev

Ko sta A in b realna smo dokazali naslednji izrek:

Izrek

Če je A realna in nedefinitna (t.j. ni pozitivno in negativno definitna), potem obstajata najmanj dva neodvisna realna izotropna vektorja.

- Pokazati želimo, da imamo neskončno število realnih rešitev in jih izračunati.
- Potrebujemo vsaj 3 različne lastne vrednosti z različnimi predznaki.

- Predpostavimo, da $\lambda_1 < 0 < \lambda_2 < \lambda_3$.

- Predpostavimo, da $\lambda_1 < 0 < \lambda_2 < \lambda_3$.
- Naj bo $t_1 = |c_1|^2$, $t_2 = |c_2|^2$.

- Predpostavimo, da $\lambda_1 < 0 < \lambda_2 < \lambda_3$.
- Naj bo $t_1 = |c_1|^2$, $t_2 = |c_2|^2$.

Veljati mora enačba (5)

$$\lambda_1 t_1 + \lambda_2 t_2 + \lambda_3 (1 - t_1 - t_2) = 0$$

oz.

$$(\lambda_1 - \lambda_3) t_1 + (\lambda_2 - \lambda_3) t_2 + \lambda_3 = 0,$$

s pogoji: $t_i \geq 0, i = 1, 2$ in $t_1 + t_2 \leq 1$.

Neskončno rešitev

- Dobimo premico definirano v (t_1, t_2) ravnini

$$t_2 = \frac{\lambda_3}{\lambda_3 - \lambda_2} - \frac{\lambda_3 - \lambda_1}{\lambda_3 - \lambda_2} t_1 .$$

Neskončno rešitev

- Dobimo premico definirano v (t_1, t_2) ravnini

$$t_2 = \frac{\lambda_3}{\lambda_3 - \lambda_2} - \frac{\lambda_3 - \lambda_1}{\lambda_3 - \lambda_2} t_1 .$$

- Pogoji za t_1, t_2 definirajo trikotnik.
- Potrebno je preveriti, če premica seka trikotnik.

Neskončno rešitev

- Dobimo premico definirano v (t_1, t_2) ravnini

$$t_2 = \frac{\lambda_3}{\lambda_3 - \lambda_2} - \frac{\lambda_3 - \lambda_1}{\lambda_3 - \lambda_2} t_1 .$$

- Pogoji za t_1, t_2 definirajo trikotnik.
- Potrebno je preveriti, če premica seka trikotnik.
- t_1-os seka pri $\lambda_3/(\lambda_3 - \lambda_1) > 1$.
- t_2-os seka pri $\lambda_3/(\lambda_3 - \lambda_2) < 1$.

Neskončno rešitev

- Dobimo premico definirano v (t_1, t_2) ravnini

$$t_2 = \frac{\lambda_3}{\lambda_3 - \lambda_2} - \frac{\lambda_3 - \lambda_1}{\lambda_3 - \lambda_2} t_1 .$$

- Pogoji za t_1, t_2 definirajo trikotnik.
- Potrebno je preveriti, če premica seka trikotnik.
- t_1-os seka pri $\lambda_3/(\lambda_3 - \lambda_1) > 1$.
- t_2-os seka pri $\lambda_3/(\lambda_3 - \lambda_2) < 1$.
- Vse dopustne vrednosti za t_1 in t_2 so dane z daljico v trikotniku.
- Zato obstaja neskončno število možnih pozitivnih parov (t_1, t_2).

Zgled

Poglejmo si enostaven zgled za matriko $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

Očitno so lastne vrednosti $\lambda_1 = -1, \lambda_2 = 1$ in $\lambda_3 = 2$. Drugače izračunamo lastne vrednosti in vektorje v Matlabu s pomočjo ukaza `[X,D] = eig(A)`. Kjer je X matrika lastnih vektorjev in D matrika lastnih vrednosti.

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

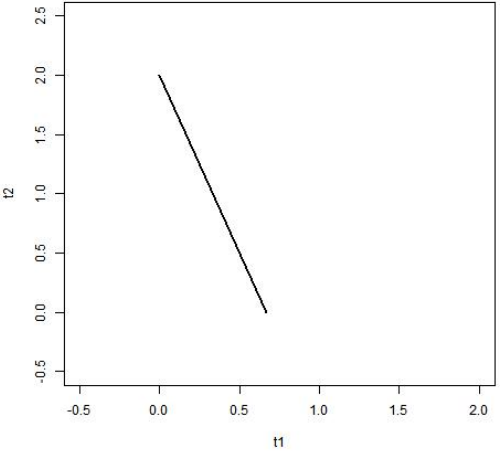
$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Zgled

Enačba premice je tedaj $t_2 = 2 - 3t_1$.

Zgled

Enačba premice je tedaj $t_2 = 2 - 3t_1$.

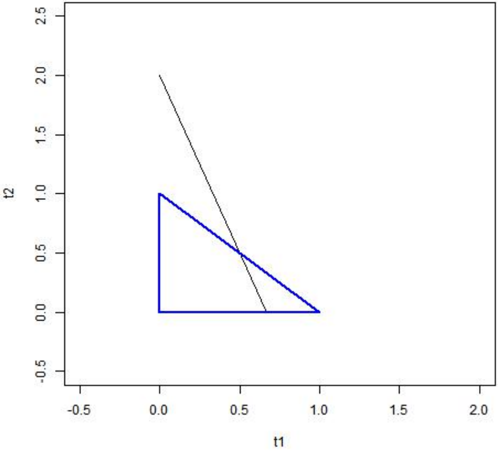


Zgled

Narišemo omejitve: $t_1, t_2 \geq 0$ in $t_1 + t_2 \leq 1$.

Zgled

Narišemo omejitve: $t_1, t_2 \geq 0$ in $t_1 + t_2 \leq 1$.

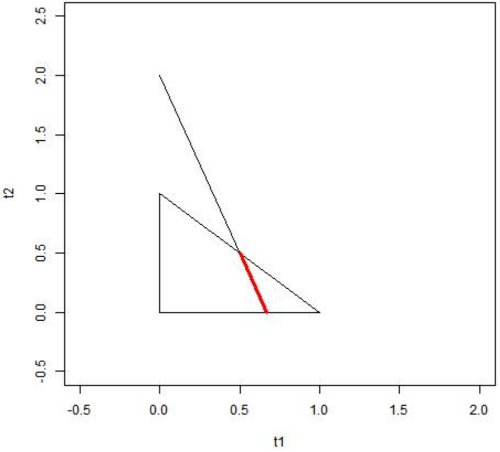


Zgled

Dopustne rešitve za t_1 in t_2 so dane z daljico v trikotniku.

Zgled

Dopustne rešitve za t_1 in t_2 so dane z daljico v trikotniku.



Zgled

Iz daljice lahko izberemo katerikoli par točk (t_1, t_2), npr. $(0.5, 0.5)$. Potem vemo kako izgleda vektor $c = \begin{bmatrix} \sqrt{0.5} \\ \sqrt{0.5} \\ 0 \end{bmatrix}$. Iz enačbe $c = X^T b$, dobimo

$$b = Xc = c = \begin{bmatrix} \sqrt{0.5} \\ \sqrt{0.5} \\ 0 \end{bmatrix}.$$

Seveda je rešitev tudi $b = -c$. Tako dobimo neskončno izotropnih vektorjev b.

Neskončno rešitev

Ta problem za iskanje koeficientov je v treh dimenzijah in možne rešitve t-ja so v eni dimenziji. Takšna konstrukcija pripelje do naslednjega izreka:

Izrek

Če je $n > 2$ in A je realna in nedefinitna, ima matrika H vsaj tri različne lastne vrednosti z različnimi predznaki. Potem obstaja neskončno število realnih izotropnih vektorjev.

Neskončno rešitev

Ta problem za iskanje koeficientov je v treh dimenzijah in možne rešitve t-ja so v eni dimenziji. Takšna konstrukcija pripelje do naslednjega izreka:

Izrek

Če je $n > 2$ in A je realna in nedefinitna, ima matrika H vsaj tri različne lastne vrednosti z različnimi predznaki. Potem obstaja neskončno število realnih izotropnih vektorjev.

Splošen zapis

$$\sum_{i=1}^{k-1}(\lambda_i - \lambda_k)t_i + \lambda_k = 0, \quad t_i \geq 0, i = 1, \dots, k-1, \sum_{i=1}^{k-1} t_i \leq 1.$$

Kompleksne matrike

Predstavljeni bodo algoritmi naslednjih avtorjev:

1. *Meurant*.
2. *Carden*.
3. *Chorianopoulos, Psarrakos in Uhlig*.

Algoritme bomo numerično primerjali, saj bi radi, da algoritem vrne izotropni vektor s čim manj računanja.

Meurant 1.

- V nekaterih primerih lahko izračunamo rešitve s samo enim računanjem lastnih vrednosti in vektorjev matrike K.
- Uporabimo lastne vektorje matrike H.
- Z uporabo treh lastnih vektorjev H, obstaja neskončno rešitev dobljenih na daljici v trikotniku omejitev.
- Če sta imaginarna dela, ki ustrezata robnima točkama daljice, različnih predznakov, potem iz izreka o povprečni vrednosti sledi, da obstaja točka na daljici, ki ima ničeln imaginarni del.

Meurant 2.

- Uporabimo lastne vrednosti in lastne vektorje matrike $K = (A - A^*)/(2i)$, ki je hermitska.
- S kombiniranjem lastnih vektorjev matrike K pripadajočim k pozitivnim in negativnim lastnim vrednostim, lahko (v nekaterih primerih) izračunamo dva vektorja b_1 in b_2, taka da
$$\alpha_1 = \Re(b_1^* A b_1) < 0$$
in
$$\alpha_2 = \Re(b_2^* A b_2) > 0.$$

Literatura

1. G. Meurant, *The computation of isotropic vectors*, Numer. Alg. **60** (2012) 193–204.
2. R. Carden, *A simple algorithm for the inverse field of values problem*, Inverse Probl. **25** (2009) 1–9
3. C. Chorianopoulos, P. Psarrakos in F. Uhlig, *A method for the inverse numerical range problem*, Electron. J. Linear Algebra **20** (2010) 198–206
4. N. Ciblak, H. Lipkin, *Orthonormal isotropic vector bases*, In: Proceedings of DETC'98, 1998 ASME Design Engineering Technical Conferences (1998).
5. Johnson, C. R., *Numerical determination of the field of values of a general complex matrix*, SIAM J. Numer. Anal. **15** (1978) 595–602.