

UNIVERZA V LJUBLJANI  
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Manca Cör

**Izpeljava metod za iskanje ničel polinomov z uporabo  
optimizacije**

Delo diplomskega seminarja

Mentor: prof. dr. Bor Plestenjak

Ljubljana, 2019

## KAZALO

1. Uvod	4
2. Predpostavke	4
3. Izpeljava osnovnih metod	5
3.1. Newtonova metoda	6
3.2. Metoda Ostrovskega	7
3.3. Laguerrova metoda	7
4. Izpeljava naprednejših metod	8
4.1. Izboljšana Newtonova metoda	8
4.2. Diskretna Laguerrova metoda	10
5. Teoretični rezultati	10
6. Računanje lastnih vrednosti Toeplitzovih matrik	13
7. Primerjava metod za problem iskanja najmanjše lastne vrednosti simetrične tridiagonalne matrike	16
8. Zaključek	25
Slovar strokovnih izrazov	26
Literatura	26
Slike	27
Tabele	27

## Izpeljava metod za iskanje ničel polinomov z uporabo optimizacije

### POVZETEK

V tem diplomskem seminarju obravnavamo naslednje metode za iskanje ničel polinomov: Newtonovo metodo, metodo Ostrovskega in Laguerrovo metodo. Izpeljali jih bomo s pomočjo vezanega optimizacijskega problema, potem pa na podoben način izpeljali še dve naprednejši metodi: izboljšano Newtonovo metodo in diskretno Laguerrovo metodo. Dokazali bomo nekaj izrekov, ki nam povedo, za koliko lahko povečan korak posamezne metode preseže najmanjšo ničlo. Metode bomo še numerično testirali in jih med seboj primerjali na problemu iskanja najmanjše lastne vrednosti simetrične tridiagonalne matrike. Primerjali bomo število potrebnih korakov za dovolj dober približek, njihovo časovno zahtevnost in red konvergence.

## Derivation of polynomial zerofinders via optimization problems

### ABSTRACT

In this diploma seminar we study the following polynomial zerofinders: Newton's method, Ostrowski's method and Laguerre's method. We will derive them via a constrained optimization problem and will also derive some new methods using the same approach: the improved Newton's method and the discrete Laguerre's method. We will prove some theorems that give us a bound on how far a magnified step for a given method can overshoot the smallest zero of a polynomial. We will test the methods numerically and compare them to one another. We will do that for the problem of finding the smallest eigenvalue of a symmetric tridiagonal matrix. We will compare the number of steps we need for a good approximation of the smallest eigenvalue, the time complexity of the methods and rate of convergence.

**Math. Subj. Class. (2010):** 65H04 , 65H05, 65F15

**Ključne besede:** Newtonova metoda, metoda Ostrovskega, Laguerrova metoda, izboljšana Newtonova metoda, diskretna Laguerrova metoda

**Keywords:** Newton's method, Ostrowski's method, Laguerre's method, improved Newton's method, discrete Laguerre's method

## 1. UVOD

Newtonova metoda je dobro znana iterativna metoda za reševanje enačb oblike  $f(x) = 0$ . Rekurzivno je dana kot

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

za primerno izbiro začetnega približka  $x_0$ . Poznamo dva standardna načina izpeljave: analitičnega in geometrijskega.

Geometrijska ideja Newtonove oz. tangentne metode je, da funkcijo  $f$  v točki  $(x_k, f(x_k))$  aproksimiramo s tangento in potem za približek ničle funkcije  $f$  vzamemo ničlo tangente, torej presečišče tangente z osjo  $x$ . Enačba tangente skozi  $(x_k, f(x_k))$  je  $y = f(x_k) + f'(x_k)(x - x_k)$ . Dobimo, da je presečišče tangente z osjo  $x$  v točki  $x_k - \frac{f(x_k)}{f'(x_k)}$ . To točko vzamemo za naslednji približek  $x_{k+1}$  in metodo ponavljamo za  $k = 0, 1, 2, \dots$

Analitično Newtonovo metodo izpeljemo s pomočjo razvoja  $f(x_k + h)$  v Taylorjevo vrsto, kjer je  $x_k$  približek za ničlo funkcije  $f$ :

$$f(x_k + h) = f(x_k) + f'(x_k)h + \frac{f''(x_k)h^2}{2!} + \dots$$

Če upoštevamo le prva dva člena razvoja, dobimo  $f(x_k + h) \approx f(x_k) + f'(x_k)h$ . Če predpostavimo  $f'(x_k) \neq 0$ , iz enačbe  $f(x_k + h) = 0$  dobimo  $h = -\frac{f(x_k)}{f'(x_k)}$ .

Ta načina izpeljave sta ekvivalentna, če je  $f$  realna funkcija in iščemo njene realne ničle, sicer pa je analitična izpeljava splošnejša.

V tem diplomskem seminarju bom pokazala, kako se Newtonovo metodo lahko izpelje s pomočjo vezanega optimizacijskega problema. Reševali bomo enačbo  $p(x) = 0$  za najmanjšo ničlo, kjer je  $p$  polinom s samimi realnimi ničlami. Kasneje bom isto naredila še za nekaj drugih metod za iskanje ničel in predstavila nekaj novih metod, ki jih lahko izpeljemo s podobnim postopkom. Nekatere izmed dobljenih metod lahko uporabljamo le za iskanje ničel polinomov, medtem ko so nekatere (na primer Newtonova metoda) uporabne za splošne nelinearne enačbe.

## 2. PREDPOSTAVKE

Naj bo  $p$  polinom stopnje  $n$ , definiran kot

$$p(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = (x - \xi_1)(x - \xi_2) \dots (x - \xi_n),$$

kjer so vse ničle  $\xi_j$  realne in urejene kot  $\xi_1 \leq \xi_2 \leq \dots \leq \xi_n$ . Če je  $p$  karakteristični polinom matrike  $A$ , potem njegove ničle predstavljajo lastne vrednosti matrike  $A$ . Lahko preverimo, da je:

•

$$\sum_{j=1}^n (x - \xi_j) = nx + a_{n-1},$$

kjer pri prvem enačaju upoštevamo, da je

$$\begin{aligned} (x - \xi_1)(x - \xi_2) \dots (x - \xi_n) &= x^n - (\xi_1 + \xi_2 + \dots + \xi_n)x^{n-1} + \dots \\ &= x^n + a_{n-1}x^{n-1} + \dots, \end{aligned}$$

torej  $a_{n-1} = -\sum_{j=1}^n \xi_j$ .

•

$$\sum_{j=1}^n \frac{1}{x - \xi_j} = \frac{p'(x)}{p(x)}.$$

To smo dobili tako, da smo izraz  $p(x) = (x - \xi_1)(x - \xi_2) \dots (x - \xi_n)$  logaritmirali in tako dobili:

$$\begin{aligned} \log(p(x)) &= \log((x - \xi_1)(x - \xi_2) \dots (x - \xi_n)) \\ &= \sum_{j=1}^n \log(x - \xi_j) \end{aligned}$$

Sedaj obe strani odvajamo in dobimo:

$$(1) \quad \frac{1}{p(x)} \cdot p'(x) = \sum_{j=1}^n \frac{1}{x - \xi_j},$$

kar pa je ravno to, kar smo želeli pokazati.

•

$$\sum_{j=1}^n \frac{1}{(x - \xi_j)^2} = \left( \frac{p'(x)}{p(x)} \right)^2 - \frac{p''(x)}{p(x)}.$$

To smo dobili tako, da smo odvajali enačbo (1):

$$\begin{aligned} \left( \frac{p'(x)}{p(x)} \right)' &= \left( \sum_{j=1}^n \frac{1}{x - \xi_j} \right)', \\ \frac{p''(x)p(x) - p'(x)p'(x)}{p^2(x)} &= \sum_{j=1}^n \frac{-1}{(x - \xi_j)^2}. \end{aligned}$$

Če zgornjo enakost pomnožimo z  $-1$  in jo nekoliko preuredimo, dobimo ravno:

$$\left( \frac{p'(x)}{p(x)} \right)^2 - \frac{p''(x)}{p(x)} = \sum_{j=1}^n \frac{1}{(x - \xi_j)^2}.$$

V prihodnje bomo zaradi lažjega zapisa uporabljali naslednje oznake

$$(2) \quad \alpha_x = \sum_{j=1}^n (x - \xi_j), \quad \beta_x = \sum_{j=1}^n \frac{1}{x - \xi_j}, \quad \gamma_x = \sum_{j=1}^n \frac{1}{(x - \xi_j)^2}.$$

### 3. IZPELJAVA OSNOVNIH METOD

Kot sem omenila že v uvodu, lahko z uporabo optimizacije izpeljemo tudi nekaj novih metod za iskanje ničel, ne le klasičnih. Glavno idejo izpeljave z uporabo optimizacije bom vseeno najprej prikazala pri izpeljavi Newtonove metode, metode Ostrovskega in Laguerrove metode. Da pa bomo vedeli, kaj želimo izpeljati, si najprej pogledjmo osnovne korake vseh treh omenjenih metod:

• **Newtonova metoda**

Osnovni korak:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

- **Metoda Ostrovskega**

Osnovni korak:

$$x_{k+1} = x_k - K(x_k),$$

kjer je

$$K(x) = \frac{f(x)/f'(x)}{\sqrt{1 - f(x)f''(x)/f'(x)^2}}.$$

- **Laguerreova metoda**

Osnovni korak:  $x_{k+1} = x_k - a$ , kjer je  $a$  definiran kot

$$a = \frac{n}{\beta_x \pm \sqrt{(n-1)(n\gamma_x - \beta_x^2)}},$$

in sta  $\beta_x$  in  $\gamma_x$  definirana kot v enačbi (2).

**3.1. Newtonova metoda.** Naj bo polinom  $p$  stopnje  $n$  in naj ustreza predpostavkam. Izračunamo lahko spodnjo mejo za najmanjšo ničlo, če rešimo optimizacijski problem za  $\bar{x}$  z intervala  $(-\infty, \xi_1)$ :

$$\min \left\{ \zeta_1 : \sum_{j=1}^n \frac{1}{\bar{x} - \zeta_j} = \beta_{\bar{x}}; \zeta_j > \bar{x}, j = 1, 2, \dots, n \right\}.$$

Iščemo torej najmanjšo možno ničlo vseh polinomov z danim premaknjenim momentom njihovih ničel. Rešitev tega problema je spodnja meja najmanjše ničle  $p$ . Če vpeljemo novo spremenljivko  $\omega_j = 1/(\bar{x} - \zeta_j)$ , lahko naš optimizacijski problem zapišemo takole:

$$\min \left\{ \omega_1 : \sum_{j=1}^n \omega_j = \beta_{\bar{x}}; \omega_j \leq 0, j = 1, 2, \dots, n \right\}.$$

Omejitve pri tem optimizacijskem problemu nam dajo del hiperravnine v  $\mathbb{R}^n$ . Optimalna rešitev je v tem primeru  $\omega_1 = \beta_{\bar{x}}$  in  $\omega_j = 0$  za  $j \neq 1$ , kar lahko zapišemo tudi kot  $\zeta_1 = \bar{x} - 1/\beta_{\bar{x}}$ . To mejo lahko iterativno izboljšamo, če  $\bar{x}$  zamenjamo z izračunano omejitvijo in postopek ponovimo. Ker je  $\beta_{\bar{x}}$  definiran kot  $\beta_{\bar{x}} = \frac{p'(\bar{x})}{p(\bar{x})}$ , je torej  $\zeta_1 = \bar{x} - \frac{p(\bar{x})}{p'(\bar{x})}$ . Opazimo, da dobimo ravno korak Newtonove metode. Če najmanjša ničla ni enostavna, dobimo sorodne rezultate. Ugotovimo, da je v tem primeru rezultat zgornjega optimizacijskega problema enak rezultatu Newtonove metode, le da je razmerje  $\frac{p(\bar{x})}{p'(\bar{x})}$  pomnoženo z večkratnostjo ničle.

Do rahle izboljšave Newtonove metode pri tej izpeljavi pride, če je znana zgornja meja  $b$  za največjo ničlo polinoma. Med pogoje v optimizacijskem problemu tedaj za vsak  $\zeta_j$  dodamo pogoj  $\bar{x} < \zeta_j \leq b$  in dobimo rešitev:

$$\zeta_1 = \bar{x} - \left( \beta_{\bar{x}} - \frac{n-1}{\bar{x}-b} \right)^{-1}, \quad \zeta_j = b \quad (j \neq 1).$$

Iz tega rezultata dobimo iterativno metodo, ki jo lahko zapišemo kot

$$x_{k+1} = x_k - \frac{1}{1 - \frac{n-1}{x_k-b} \cdot \frac{p(x)}{p'(x)}} \cdot \frac{p(x)}{p'(x)},$$

kjer smo upoštevali, da je  $\beta_{\bar{x}}$  enak  $\frac{p'(\bar{x})}{p(\bar{x})}$ .

**3.2. Metoda Ostrovskega.** Tudi to metodo lahko izpeljemo s pomočjo vezanega optimizacijskega problema za  $\bar{x}$  na  $(-\infty, \xi_1)$ :

$$\min \left\{ \zeta_1 : \sum_{j=1}^n \frac{1}{(\bar{x} - \zeta_j)^2} = \gamma_{\bar{x}}; \zeta_j > \bar{x}, j = 1, 2, \dots, n \right\}.$$

Če spet uporabimo isto spremenljivko kot pri izpeljavi Newtonove metode, torej  $\omega_j = 1/(\bar{x} - \zeta_j)$ , dobimo problem:

$$\min \left\{ \omega_1 : \sum_{j=1}^n \omega_j^2 = \gamma_{\bar{x}}; \omega_j \leq 0, j = 1, 2, \dots, n \right\}.$$

Omejitve pri tem problemu nam dajo del sfere v  $\mathbb{R}^n$ . Optimalna rešitev je v tem problemu dana z  $\omega_1 = \sqrt{\gamma_{\bar{x}}}$  in  $\omega_j = 0$  za  $j \neq 1$ , kar nam potem da  $\zeta_1 = \bar{x} - 1/\sqrt{\gamma_{\bar{x}}}$ .  $1/\sqrt{\gamma_{\bar{x}}}$  lahko zapišemo kot:

$$\frac{1}{\sqrt{\gamma_{\bar{x}}}} = -\frac{1}{\sqrt{1 - \frac{p''(\bar{x})p(\bar{x})}{p'^2(\bar{x})}}} \cdot \frac{p(\bar{x})}{p'(\bar{x})},$$

v čemer prepoznamo  $K(\bar{x})$  iz koraka metode Ostrovskega.

**3.3. Laguerrova metoda.** Do sedaj smo obravnavali le optimizacijske probleme z eno samo omejitvijo, vendar pa lahko v vezan optimizacijski problem dodamo več omejitev. Naslednji problem vsebuje dve:

$$\min \left\{ \zeta_1 : \sum_{j=1}^n \frac{1}{\bar{x} - \zeta_j} = \beta_{\bar{x}}; \sum_{j=1}^n \frac{1}{(\bar{x} - \zeta_j)^2} = \gamma_{\bar{x}}; \zeta_j > \bar{x}, j = 1, 2, \dots, n \right\}.$$

Tudi v tem primeru uvedemo novo spremenljivko  $\omega_j = 1/(\bar{x} - \zeta_j)$ , ki nam pretvori problem v:

$$\min \left\{ \omega_1 : \sum_{j=1}^n \omega_j = \beta_{\bar{x}}; \sum_{j=1}^n \omega_j^2 = \gamma_{\bar{x}}; \omega_j \leq 0, j = 1, 2, \dots, n \right\}.$$

V tem primeru nam omejitve dajo presek hiperravnine in sfere v  $\mathbb{R}^n$ . Če predpostavimo, da  $\omega_j \neq 0$  za vse  $j$ , so pogoji za optimalno rešitev dani z

$$1 + \mu + 2\nu\omega_1 = 0, \quad \mu + 2\nu\omega_j = 0 \quad (j \neq 1),$$

kjer sta  $\mu$  in  $\nu$  Lagrangeeva multiplikatorja. Pogoji nam povedo, da so vsi  $\omega_j$  ( $j = 2, \dots, n$ ) enaki. Če vzamemo  $y = \omega_j$  za  $j \neq 1$  in to vstavimo v omejitve, dobimo:

$$\omega_1 + (n-1)y = \beta_{\bar{x}}, \quad \omega_1^2 + (n-1)y^2 = \gamma_{\bar{x}}.$$

Če rešujemo enačbe za  $\omega_1$  in iščemo minimum danega problema, dobimo:

$$\omega_1 = \frac{\beta_{\bar{x}} - \sqrt{(n-1)(n\gamma_{\bar{x}} - \beta_{\bar{x}}^2)}}{n}.$$

Edine druge rešitve, ki bi zadostile pogojem, so tiste, kjer je en ali več  $\omega_j$  za  $j \geq 2$  enakih nič, vendar lahko pokažemo, da bi te rešitve dale večjo vrednost  $\omega_1$ , zato ne bi dale globalnega minimuma. Za  $\zeta_1$  tako dobimo:

$$\zeta_1 = \bar{x} + \frac{n}{-\beta_{\bar{x}} + \sqrt{(n-1)(n\gamma_{\bar{x}} - \beta_{\bar{x}}^2)}}.$$

Ta rešitev predstavlja en korak Laguerreove metode.

Iz minimizacijskega problema lahko direktno sklepamo, da bi Laguerreova metoda morala, vsaj za izračun najnižje ničle, konvergirati hitreje od Newtonove metode in metode Ostrovskega. To se zgodi, ker je množica dopustnih rešitev manjša in nam zato da večjo minimalno vrednost. Vse tri metode konvergirajo monotono proti  $\xi_1$ , če je začetni približek vzet iz intervala  $(-\infty, \xi_1)$ .

#### 4. IZPELJAVA NAPREDNEJŠIH METOD

V tem poglavju bom prikazala še izpeljavo dveh naprednejših metod za iskanje ničel polinomov. Prva temelji na Newtonovi, druga pa na Laguerrovi metodi.

**4.1. Izboljšana Newtonova metoda.** Izboljšano Newtonovo metodo dobimo, če rešimo spodnji optimizacijski problem, pri katerem so rešitve končne in strogo večje od  $\bar{x}$ :

$$\min \left\{ \zeta_1 : \sum_{j=1}^n (\bar{x} - \zeta_j) = \alpha_{\bar{x}}; \sum_{j=1}^n \frac{1}{\bar{x} - \zeta_j} = \beta_{\bar{x}}; \zeta_j > \bar{x}, j = 1, 2, \dots, n \right\}.$$

Z metodo Lagrangeevih multiplikatorjev, kjer je prva omejitev dobila multiplikator  $\mu$ , druga omejitev pa multiplikator  $\nu$ , za ta optimizacijski problem dobimo sistem enačb:

$$1 - \mu + \frac{\nu}{(\bar{x} - \zeta_1)^2} = 0, \quad -\mu + \frac{\nu}{\bar{x} - \zeta_j} = 0 \quad \text{za } j \neq 1.$$

Tudi tu iz pogojev vidimo, da je  $n - 1$  rešitev  $\bar{x} - \zeta_j$  za  $j = 2, \dots, n$  enakih. Če njihovo vrednost označimo z  $y$ , dobimo naslednji enačbi:

$$(3) \quad (\bar{x} - \zeta_1) + (n - 1)y = \alpha_{\bar{x}},$$

$$(4) \quad \frac{1}{\bar{x} - \zeta_1} + \frac{n - 1}{y} = \beta_{\bar{x}}.$$

Iz (4) dobimo:

$$y = \frac{(n - 1)(\bar{x} - \zeta_1)}{\beta_{\bar{x}}(\bar{x} - \zeta_1) - 1},$$

pri čemer je  $\bar{x} - \zeta_j \neq 1/\beta_{\bar{x}}$ . Če to vstavimo v enačbo (3), dobimo kvadratno enačbo za  $\bar{x} - \zeta_1$ :

$$\beta_{\bar{x}}(\bar{x} - \zeta_1)^2 - (\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2)(\bar{x} - \zeta_1) + \alpha_{\bar{x}} = 0.$$

Ker minimiziramo  $\zeta_1$ , nas zanima največja vrednost  $\bar{x} - \zeta_1$ , torej

$$\bar{x} - \zeta_1 = \frac{(\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2) - \sqrt{(\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2)^2 - 4\alpha_{\bar{x}}\beta_{\bar{x}}}}{2\beta_{\bar{x}}}.$$

Preden analiziramo te ničle, opazujmo še neenakost, ki velja za harmonično<sup>1</sup> in aritmetično sredino števil  $a_1, a_2, \dots, a_n$ :

$$\frac{a_1 + a_2 + \dots + a_n}{n} \geq \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_n}}.$$

V našem primeru za  $a_j$  vzamemo  $(\bar{x} - \zeta_j)$ . Tako dobimo:

<sup>1</sup>Harmonična sredina števil  $a_1, a_2, \dots, a_n$  je definirana kot  $H(a_1, \dots, a_n) = \frac{n}{\sum_{i=1}^n \frac{1}{a_i}}$ .



$$\begin{aligned}
\frac{(\bar{x} - \zeta_1) + (\bar{x} - \zeta_2) + \cdots + (\bar{x} - \zeta_n)}{n} &\geq \frac{n}{\frac{1}{\bar{x} - \zeta_1} + \frac{1}{\bar{x} - \zeta_2} + \cdots + \frac{1}{\bar{x} - \zeta_n}}, \\
\frac{\alpha_{\bar{x}}}{n} &\geq \frac{n}{\beta_{\bar{x}}}, \\
(5) \quad \alpha_{\bar{x}}\beta_{\bar{x}} &\geq n^2,
\end{aligned}$$

kjer smo uporabili definiciji  $\alpha_{\bar{x}}$  in  $\beta_{\bar{x}}$  iz (2). Z nekaj preurejanja lahko diskriminanto napišemo tudi tako:

$$(\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2)^2 - 4\alpha_{\bar{x}}\beta_{\bar{x}} = (\alpha_{\bar{x}}\beta_{\bar{x}} - n^2)(\alpha_{\bar{x}}\beta_{\bar{x}} - (n - 2)^2).$$

Oba faktorja na desni strani zgornje enačbe sta pozitivna zaradi (5), kar nam pove, da je tudi diskriminanta pozitivna.

Nobena druga rešitev ne zadostuje omejitvam iz optimizacijskega problema. Če rešitev za  $\bar{x} - \zeta_1$  razširimo z

$$(\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2) + \sqrt{(\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2)^2 - 4\alpha_{\bar{x}}\beta_{\bar{x}}},$$

dobimo:

$$\begin{aligned}
\bar{x} - \zeta_1 &= \frac{2\alpha_{\bar{x}}\beta_{\bar{x}}}{(\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2) + \sqrt{(\alpha_{\bar{x}}\beta_{\bar{x}} + 1 - (n - 1)^2)^2 - 4\alpha_{\bar{x}}\beta_{\bar{x}}}} \cdot \frac{1}{\beta_{\bar{x}}} \\
&= \frac{2\alpha_{\bar{x}}\beta_{\bar{x}}}{(\alpha_{\bar{x}}\beta_{\bar{x}} - n(n - 2)) + \sqrt{(\alpha_{\bar{x}}\beta_{\bar{x}} - n^2)(\alpha_{\bar{x}}\beta_{\bar{x}} - (n - 2)^2)}} \cdot \frac{1}{\beta_{\bar{x}}}.
\end{aligned}$$

Od tu lahko nato izrazimo  $\zeta_1$ :

$$\zeta_1 = \bar{x} - \frac{2\alpha_{\bar{x}}\beta_{\bar{x}}}{(\alpha_{\bar{x}}\beta_{\bar{x}} - n(n - 2)) + \sqrt{(\alpha_{\bar{x}}\beta_{\bar{x}} - n^2)(\alpha_{\bar{x}}\beta_{\bar{x}} - (n - 2)^2)}} \cdot \frac{1}{\beta_{\bar{x}}}.$$

Če namesto  $\alpha_{\bar{x}}\beta_{\bar{x}}$  pišemo  $s$ , lahko definiramo  $\phi(s)$  kot

$$(6) \quad \phi(s) = \frac{2s}{(s - n(n - 2)) + \sqrt{(s - n^2)(s - (n - 2)^2)}}.$$

Rešitev našega minimizacijskega problema je v tem primeru  $\zeta_1 = \bar{x} - \phi(s)/\beta_{\bar{x}}$ , pri čemer je  $s \geq n^2$ . Funkcija  $\phi$  ima na svojem definicijskem območju naslednje lastnosti:

- $\phi(s)' < 0$ ,
- $\phi(n^2) = n$ ,
- $\lim_{s \rightarrow +\infty} \phi(s) = 1$ .

Vidimo, da je dobljena rešitev optimizacijskega problema  $\zeta_1 = \bar{x} - \phi(s)/\beta_{\bar{x}}$  podobna koraku Newtonove metode, le da je  $-1/\beta_{\bar{x}}$  pomnožen še s  $\phi(s)$ . Iz tega lahko sklepamo, da ta metoda ni nikoli slabša od Newtonove metode in da je lahko precej hitrejša.

Ko se spremenljivke, na katerih izvajamo iteracijo, bližajo najmanjši ničli, linearni pogoj prispeva vse manj in manj, drugi pogoj pa postaja pomembnejši. Ko gre število korakov proti  $\infty$ , ta metoda postane enaka Newtonovi metodi. Njena glavna prednost je v tem, da z njo dobimo boljše začetne približke, ko so vrednosti še daleč stran od ničle. Dodajanje linearne omejitve v optimizacijski problem bi podobno izboljšalo tudi Laguerrovo metodo in metodo Ostrovskega.

**4.2. Diskretna Laguerrova metoda.** Včasih je težko izračunati drugi odvod neke funkcije, zato se takrat uporablja diskretna verzija Laguerrove metode. V tem primeru imamo v optimizacijskem problemu oba pogoja ista kot v Newtonovi metodi, vendar vsakega računamo v drugi točki. Če sta ti dve točki  $x = x_0$  in  $x = x_1$ , kjer je  $x_0 < x_1 < \xi_1$ , potem ustrezeni optimizacijski problem izgleda takole:

$$\min \left\{ \zeta_1 : \sum_{j=1}^n \frac{1}{x_0 - \zeta_j} = \beta_{x_0}; \sum_{j=1}^n \frac{1}{x_1 - \zeta_j} = \beta_{x_1}; \zeta_j > x_1, j = 1, 2, \dots, n \right\}.$$

Rešitev tega optimizacijskega problema je

$$\zeta_1 = \frac{x_0 + x_1}{2} + \frac{n - \frac{1}{4} \left( \frac{\Delta\beta}{\Delta x} + S \right) (\Delta x)^2}{-\frac{1}{2}(\beta_{x_0} + \beta_{x_1}) + \sqrt{S \left( 1 - n + S \frac{(\Delta x)^2}{4} \right)}},$$

kjer je  $\Delta\beta = \beta_{x_1} - \beta_{x_0}$ ,  $\Delta x = x_1 - x_0$  in  $S = \beta_{x_0}\beta_{x_1} + n \frac{\Delta\beta}{\Delta x}$ . Iterirane spremenljivke konvergirajo monotonno proti najmanjši ničli  $\zeta_1$ , če je naš začetni približek z intervala levo od  $\xi_1$ , torej z intervala  $(-\infty, \xi_1)$ . Ta metoda ima podobne lastnosti kot navadna Laguerrova metoda.

Kljub temu da smo za vse zgoraj predstavljene metode predpostavili zelo stroge pogoje, ki veljajo za najmanjšo ničlo, jih lahko bolj ali manj uspešno uporabimo tudi za iskanje drugih ničel. Uporabimo jih lahko tudi za iskanje ničel funkcij, ki niso nujno polinomi.

Predstavljene metode k najmanjši ničli konvergirajo monotonno, vendar pa se njihove lastnosti za izračun ostalih ničel razlikujejo od metode do metode, tega pa na tem mestu ne bomo obravnavali.

## 5. TEORETIČNI REZULTATI

Če malo spremenimo začetni optimizacijski problem, lahko dokažemo sledeče izreke, ki nam dajo mejo, do koder lahko povečan korak za predstavljene metode preseže najmanjšo ničlo polinoma.

**Izrek 5.1. (Newtonova metoda).** *Imamo polinom  $p$   $n$ -te stopnje, ki ima  $s$  različnih realnih ničel, urejenih kot  $\xi_1 < \xi_2 < \xi_3 < \dots < \xi_s$ , kjer ima ničla  $\xi_j$  večkratnost  $q_j$ . Če je  $\bar{x} < \xi_1$  in  $0 < \theta \leq \frac{1}{2}$ , potem drži:*

$$\bar{x} - \left( (1 - \theta)q_1 + \theta q_2 + 2\sqrt{q_1 q_2 \theta(1 - \theta)} \right) \frac{1}{\beta_{\bar{x}}} \leq (1 - \theta)\xi_1 + \theta\xi_2.$$

*Dokaz:*  $\theta$  izberemo z intervala  $(0, \frac{1}{2}]$  in nato rešujemo optimizacijski problem:

$$\min \left\{ (1 - \theta)\zeta_1 + \theta\zeta_2 : \sum_{j=1}^n \frac{1}{\bar{x} - \zeta_j} = \beta_{\bar{x}}; \zeta_j > \bar{x}, j = 1, 2, \dots, n \right\}.$$

Opazimo, da lahko omejitve zapišemo v obliki:

$$(7) \quad \frac{q_1}{\bar{x} - \zeta_1} + \frac{q_2}{\bar{x} - \zeta_2} + \sum_{j=3}^s \frac{q_j}{\bar{x} - \zeta_j} = \beta_{\bar{x}}.$$

S pomočjo Lagrangeevih multiplikatorjev dobimo enačbe:

$$(1 - \theta) + \frac{\mu q_1}{(\bar{x} - \zeta_1)^2} = 0, \quad \theta + \frac{\mu q_2}{(\bar{x} - \zeta_2)^2} = 0, \quad \frac{\mu q_j}{(\bar{x} - \zeta_j)^2} = 0 \quad (j \neq 1, 2).$$

Iz teh pogojev izrazimo  $\mu$  in dobimo:

$$\mu = \frac{-(1-\theta)(\bar{x} - \zeta_1)^2}{q_1}, \quad \mu = \frac{-\theta(\bar{x} - \zeta_2)^2}{q_2}.$$

Izenačimo obe enačbi za  $\mu$  in dobimo:

$$\begin{aligned} \frac{-(1-\theta)(\bar{x} - \zeta_1)^2}{q_1} &= \frac{-\theta(\bar{x} - \zeta_2)^2}{q_2}, \\ \frac{(\bar{x} - \zeta_2)^2}{(\bar{x} - \zeta_1)^2} &= \frac{q_2(1-\theta)}{q_1\theta}, \\ \frac{\bar{x} - \zeta_2}{\bar{x} - \zeta_1} &= \left( \frac{q_2(1-\theta)}{q_1\theta} \right)^{\frac{1}{2}}. \end{aligned}$$

Od tu dobimo rešitve:

$$\bar{x} - \zeta_2 = \left( \frac{q_2(1-\theta)}{q_1\theta} \right)^{\frac{1}{2}} (\bar{x} - \zeta_1), \quad \zeta_j = +\infty \quad (j \neq 1, 2).$$

To lahko vstavimo v enačbo (7) in dobimo

$$\bar{x} - \zeta_1 = \left( q_1 + \left( \frac{q_1 q_2 \theta}{1-\theta} \right)^{\frac{1}{2}} \right) \frac{1}{\beta_{\bar{x}}}, \quad \bar{x} - \zeta_2 = \left( q_2 + \left( \frac{q_1 q_2 (1-\theta)}{\theta} \right)^{\frac{1}{2}} \right) \frac{1}{\beta_{\bar{x}}}.$$

Ti dve enačbi nam vrneti minimalno vrednost, ker morajo rešitve biti večje od  $\bar{x}$  in ker nobena druga rešitev ne zadošča pogojem za optimalnost. Iz tega lahko izračunamo najmanjšo vrednost:

$$(1-\theta)\zeta_1 + \theta\zeta_2 = \bar{x} - \left( (1-\theta)q_1 + \theta q_2 + 2\sqrt{q_1 q_2 (1-\theta)} \right) \frac{1}{\beta_{\bar{x}}}.$$

V oklepaju na desni strani enačbe prepoznamo popolni kvadrat izraza

$$\sqrt{(1-\theta)q_1} + \sqrt{\theta q_2},$$

zato lahko zgornjo enačbo napišemo kot:

$$(1-\theta)\zeta_1 + \theta\zeta_2 = \bar{x} - \left( \sqrt{(1-\theta)q_1} + \sqrt{\theta q_2} \right)^2 \frac{1}{\beta_{\bar{x}}}.$$

Sedaj opazimo, da za vsaka dva para točk  $(a, b)$  in  $(c, d)$ , kjer velja  $0 \leq a \leq b$  in  $0 \leq c \leq d$ , velja  $ad + bc \leq ac + bd$ . To lahko preprosto preverimo z obravnavanjem neenačbe  $(b-a)(d-c) \geq 0$ . Če upoštevamo to, smo za minimalno vrednost dobili:

$$\xi_1 \max\{\theta, q - \theta\} + \xi_2 \min\{\theta, 1 - \theta\},$$

ker  $\xi_1$  in  $\xi_2$  zadoščata več omejitvam kot tem, ki so predstavljene tu. Ker je  $0 < \theta \leq \frac{1}{2}$ , je ta dokaz končan. ■

Če imamo enostavne ničle ( $q_1 = q_2 = 1$ ), nam izrek 5.1 pove, da pri minimumu velja:

$$(1-\theta)\zeta_1 + \theta\zeta_2 = \bar{x} - \left( 1 + 2\sqrt{\theta(1-\theta)} \right) \frac{1}{\beta_{\bar{x}}}.$$

Ko je  $\theta = \frac{1}{2}$ , to pomeni, da dvojni korak Newtonove metode nikoli ne preseže aritmetične sredine dveh najmanjših ničel, če le začnemo v točki z intervala  $(-\infty, \xi_1)$ . Če najmanjše ničle niso enostavne, za večje korake Newtonove metode veljajo iste lastnosti. Za minimizacijo lahko seveda uporabljamo tudi druge podobne funkcije, kot je npr. harmonična sredina.

Zelo podoben rezultat dobimo tudi pri metodi Ostrovskega.

**Izrek 5.2. (Metoda Ostrovskega).** Če ima polinom  $p$   $n$ -te stopnje same realne ničle, ki so urejene kot  $\xi_1 < \xi_2 < \xi_3 < \dots < \xi_s$ , in ima  $\xi_j$  stopnjo  $q_j$ , potem v primeru, da je  $\bar{x} < \xi_1$  in  $0 < \theta \leq \frac{1}{2}$ , velja:

$$\bar{x} + \left( \sqrt{(1-\theta)^2 q_1 + (\theta(1-\theta)^2 q_1 \sqrt{q_2})^{\frac{2}{3}}} + \sqrt{\theta^2 q_2 + ((1-\theta)\theta^2 q_2 \sqrt{q_1})^{\frac{2}{3}}} \right) \frac{1}{\sqrt{\gamma_{\bar{x}}}} \leq (1-\theta)\xi_1 + \theta\xi_2.$$

Tega izreka ne bom dokazala, ker je dokaz zelo podoben kot pri izreku 5.1.

V primeru enostavnih ničel, torej  $q_1 = q_2 = 1$ , nam izrek 5.2 da enačbo:

$$(1-\theta)\zeta_1 + \theta\zeta_2 = \bar{x} + \left( \sqrt{\theta^2 + (\theta^2(q-\theta))^{\frac{2}{3}}} + \sqrt{(1-\theta)^2 + (\theta(1-\theta)^2)^{\frac{2}{3}}} \right) \frac{1}{\sqrt{\gamma_{\bar{x}}}}.$$

Ko je  $\theta = \frac{1}{2}$  to pomeni, da če korak metode Ostrovskega pomnožimo s  $\sqrt{2}$  in je naš začetni približek z intervala  $(-\infty, \xi_1)$ , ne bomo nikoli prekoračili aritmetične sredine najmanjših dveh ničel.

Podoben izrek lahko dokažemo tudi za izboljšano Newtonovo metodo ter za Laguerrovo in diskretno Laguerrovo metodo, a je to v splošnem zelo težko. Zato bomo pri naslednjem izreku predpostavili, da so vse ničle enostavne in  $\theta = \frac{1}{2}$ .

**Izrek 5.3. (Izboljšana Newtonova metoda).** Če ima polinom  $n$ -te stopnje  $p$  same enostavne realne ničle  $\xi_j$ , urejene kot  $\xi_1 < \xi_2 < \dots < \xi_n$ , in če je  $\bar{x} < \xi_1$ , velja:

$$\bar{x} - \frac{2\alpha_{\bar{x}}\beta_{\bar{x}}}{(\alpha_{\bar{x}}\beta_{\bar{x}} - n(n-4)) + \sqrt{(\alpha_{\bar{x}}\beta_{\bar{x}} - n^2)(\alpha_{\bar{x}}\beta_{\bar{x}} - (n-4)^2)}} \cdot \frac{2}{\beta_{\bar{x}}} \leq \frac{1}{2}(\xi_1 + \xi_2).$$

*Dokaz:* Rešujemo optimizacijski problem:

$$\min \left\{ \zeta_1 + \zeta_2 : \sum_{j=1}^n (\bar{x} - \zeta_j) = \alpha_{\bar{x}}; \sum_{j=1}^n \frac{1}{\bar{x} - \zeta_j} = \beta_{\bar{x}}; \zeta_j > 0, j = 1, 2, \dots, n \right\}.$$

Rešitve so končne in različne od  $\bar{x}$ . Optimalna rešitev mora zadoščati naslednjim pogojem, ki jih dobimo s pomočjo Lagrangeevih multiplikatorjev:

$$\begin{aligned} 1 - \mu + \frac{\nu}{(\bar{x} - \zeta_1)^2} &= 0, \\ 1 - \mu + \frac{\nu}{(\bar{x} - \zeta_2)^2} &= 0, \\ -\mu + \frac{\nu}{(\bar{x} - \zeta_j)^2} &= 0, \quad (j \neq 1, 2). \end{aligned}$$

Ker v funkciji, ki jo želimo minimizirati,  $\zeta_1$  in  $\zeta_2$  nastopata simetrično in imata enaka pogoja za optimalnost, sklepamo, da velja  $\zeta_1 = \zeta_2$ . Iz zgornjih pogojev lahko sklepamo tudi, da so vse  $\zeta_j$  enake, ko je  $j \neq 1, 2$ . Če to vstavimo v omejitve in uvedemo spremenljivko  $y = \bar{x} - \zeta_j$  ( $j \neq 1, 2$ ), dobimo:

$$(8) \quad 2(\bar{x} - \zeta_1) + (n-2)y = \alpha_{\bar{x}}, \quad \frac{2}{\bar{x} - \zeta_1} + \frac{n-2}{y} = \beta_{\bar{x}}.$$

Opazimo, da dobimo enake enačbe kot v (4), le da  $\alpha_{\bar{x}}$  zamenjamo z  $\frac{\alpha_{\bar{x}}}{2}$ ,  $\beta_{\bar{x}}$  z  $\frac{\beta_{\bar{x}}}{2}$ ,  $n$  pa z  $\frac{n}{2}$ . Če rešimo enačbi iz (8), dobimo:

$$\zeta_1 = \bar{x} - \frac{2\alpha_{\bar{x}}\beta_{\bar{x}}}{(\alpha_{\bar{x}}\beta_{\bar{x}} - n(n-4)) + \sqrt{(\alpha_{\bar{x}}\beta_{\bar{x}} - n^2)(\alpha_{\bar{x}}\beta_{\bar{x}} - (n-4)^2)}} \cdot \frac{2}{\beta_{\bar{x}}}.$$

Nobena druga rešitev ne zadošča potrebnim pogojem, zato smo dobili minimalni  $\zeta_1$ . Ker je  $\zeta_1 = \zeta_2$ , za minimum aritmetične sredine  $\zeta_1$  in  $\zeta_2$  dobimo kar isto vrednost kot za  $\zeta_1$ . ■

Analogni rezultati obstajajo tudi za Laguerrovo in diskretno Laguerrovo metodo. Izreka za ti dve metodi ne bom dokazala, ker je dokaz zelo podoben kot pri izreku 5.3.

**Izrek 5.4. (Laguerrova in diskretna Laguerrova metoda).** Če ima polinom  $n$ -te stopnje  $p$  same enostavne realne ničle  $\xi_j$ , urejene kot  $\xi_1 < \xi_2 < \dots < \xi_n$ , in če je  $\bar{x} < \bar{y} < \xi_1$ , velja:

$$(9) \quad \bar{x} + \frac{\sqrt{2n}}{-\sqrt{2}\beta_{\bar{x}} + \sqrt{(n-2)(n\gamma_{\bar{x}} - \beta_{\bar{x}}^2)}} \leq \frac{1}{2}(\xi_1 + \xi_2),$$

$$(10) \quad \frac{\bar{x} + \bar{y}}{2} + \frac{n - \frac{1}{4}\left(\frac{\Delta\beta}{\Delta x} + \frac{S}{2}\right)(\Delta x)^2}{-\frac{1}{2}(\beta_{\bar{x}} + \beta_{\bar{y}}) + \frac{1}{\sqrt{2}}\sqrt{S\left(2 - n + S\frac{(\Delta x)^2}{2}\right)}} \leq \frac{1}{2}(\xi_1 + \xi_2),$$

kjer je  $\Delta\beta = \beta_{\bar{y}} - \beta_{\bar{x}}$ ,  $\Delta x = \bar{y} - \bar{x}$  in  $S = \beta_{\bar{x}}\beta_{\bar{y}} + n\frac{\Delta\beta}{\Delta x}$ .

**Opomba 5.5.** Rezultati, ki smo jih dobili pri izrekih, bi ostali enaki za običajne korake, če bi bila najmanjša ničla dvojna.

## 6. RAČUNANJE LASTNIH VREDNOSTI TOEPLITZOVIH MATRIK

Kot ponazoritev si najprej pogledjmo nekaj numeričnih primerjav. V članku Melmana in Gragga [1] je podanih nekaj primerov, kjer je izračunana najmanjša lastna vrednost Toeplitzove matrike – matrike, ki ima po diagonalah konstantne vrednosti. Najmanjša lastna vrednost je bila izračunana pri več kot 500 matrikah. Te so bile treh tipov:

- **CVL matrike**

To so matrike, ki so definirane kot  $T = \mu \sum_{k=1}^n \xi_k T_{2\pi\theta_k}$ , kjer je  $n$  dimenzija matrike  $T$ ,  $\theta_k$  in  $\xi_k$  pa sta enakomerno zvezno porazdeljeni slučajni spremenljivki na intervalu  $(0,1)$ . Parameter  $\mu$  je izbran tako, da je  $T_{kk} = 1$  za vsak  $k = 1, \dots, n$  in  $(T_{\theta})_{ij} = \cos(\theta(i-j))$ . Takšne matrike so pozitivno semidefinitne.

- **KMS matrike**

To so matrike, ki so definirane z

$$T_{ij} = \nu^{|i-j|} \quad (0 < \nu < 1; i, j = 1, \dots, n),$$

kjer je  $n$  dimenzija matrike. Takšne matrike so pozitivno definitne.

- **DDM matrike**

To so diagonalno dominantne simetrične Toeplitzove matrike. Kvadratna matrika je diagonalno dominantna, če zanjo velja:

$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$  za vse  $i$ , kjer je  $a_{ij}$  element  $i$ -te vrstice in  $j$ -tega stolpca matrike.

V obravnavanih matrikah je prva vrstica podana z vektorjem, katerega komponente so porazdeljene enakomerno zvezno na intervalu  $(0,1)$ , vrednost na položaju  $(1,1)$  pa je  $2n$ , kjer je  $n$  dimenzija matrike. Če je diagonalno dominantna matrika simetrična (kar v tem primeru je), vemo, da je pozitivno semidefinitna.

Iskanje najmanjše lastne vrednosti je potekalo s pomočjo iskanja najmanjše ničle karakterističnega polinoma. Poglejmo si, kako je bil ta izračunan. Najprej definiramo matriki  $T_i$  in  $E_i$ , ki sta obe velikosti  $i \times i$ .

$$T_i = \begin{bmatrix} t_0 & t_1 & \dots & t_{i-1} \\ t_1 & t_0 & \dots & t_{i-2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{i-1} & t_{i-2} & \dots & t_0 \end{bmatrix}, \quad E_i = \begin{bmatrix} 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 1 & 0 & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ 1 & \dots & 0 & 0 & 0 \end{bmatrix}.$$

Definiramo tudi vektor  $t^{(i)} = [t_1, t_2, \dots, t_{i-1}, t_i]^T$ , ki je ravno prva vrstica matrike  $T_{i+1}$ , iz katere smo izpustili  $t_0$ ,  $I_i$  pa naj bo identiteta velikosti  $i \times i$ . Karakteristični polinom matrike  $T_i$  označimo s  $p_i(\lambda) = \det(T_i - \lambda I_i)$ .

**Izrek 6.1.** Če so  $\lambda_i(T_i) \leq \lambda_{i-1}(T_i) \leq \dots \leq \lambda_1(T_i)$  lastne vrednosti matrike  $T_i$ , veljajo naslednje lastnosti:

$$(11) \quad p_i(\lambda) > 0, \quad p_i'(\lambda) < 0, \quad p_i''(\lambda) > 0, \quad 1 \leq i \leq n, \quad \lambda < \lambda_n(T_n),$$

$$(12) \quad p_{i+1}(\lambda) = p_i(\lambda)(t_0 - \lambda - t^{(i)T} E_i (T_i - \lambda I_i)^{-1} E_i t^{(i)}).$$

Da bomo lahko dokazali izrek 6.1, potrebujemo še Cauchyjev izrek o prepletanju lastnih vrednosti.

**Izrek 6.2. (Cauchyjev izrek o prepletanju lastnih vrednosti).** Naj bo  $A \in \mathbb{R}^{n \times n}$  simetrična matrika in naj bo  $B \in \mathbb{R}^{m \times m}$ , kjer je  $m < n$ , neka podmatrika matrike  $A$ , ki jo dobimo tako, da iz matrike  $A$  odstranimo  $i$ -to vrstico in  $i$ -ti stolpec za nekatere vrednosti  $i$ . Recimo, da so  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  lastne vrednosti matrike  $A$ ,  $\beta_1 \leq \beta_2 \leq \dots \leq \beta_m$  pa lastne vrednosti matrike  $B$ . Potem velja:

$$\lambda_k \leq \beta_k \leq \lambda_{k+n-m}, \quad \text{za } k = 1, \dots, m.$$

Če je  $m = n - 1$ , potem velja:

$$\lambda_1 \leq \beta_1 \leq \lambda_2 \leq \beta_2 \leq \dots \leq \beta_{n-1} \leq \lambda_n.$$

*Dokaz:* Brez škode za splošnost predpostavimo, da je  $A = \begin{bmatrix} B & X^T \\ X & Z \end{bmatrix}$ . Naj bodo  $\{x_1, \dots, x_n\}$  lastni vektorji matrike  $A$  in  $\{y_1, \dots, y_m\}$  lastni vektorji matrike  $B$ . Definiramo naslednje vektorske prostore:

$$V = \text{Lin}(x_k, \dots, x_n), \quad W = \text{Lin}(y_1, \dots, y_m), \quad \widetilde{W} = \left\{ \begin{bmatrix} w \\ 0 \end{bmatrix} \in \mathbb{R}^n, w \in W \right\}.$$

Ker je  $\dim(V) = n - k + 1$  in  $\dim(\widetilde{W}) = \dim(W) = k$ , obstaja vektor  $\tilde{w} \in V \cap \widetilde{W}$ ,  $\tilde{w} = \begin{bmatrix} w \\ 0 \end{bmatrix}$  za nek neničelni  $w \in W$ . Potem velja:

$$\begin{bmatrix} w^T & 0 \end{bmatrix} \begin{bmatrix} B & X^T \\ X & Z \end{bmatrix} \begin{bmatrix} w \\ 0 \end{bmatrix} = w^T B w.$$

Vemo, da je  $\lambda_k = \min_{x \in V} \frac{x^T A x}{x^T x}$  in  $\beta_k = \max_{x \in W} \frac{x^T B x}{x^T x}$ . Od tod sledi

$$\lambda_k \leq \frac{\tilde{w}^T A \tilde{w}}{\tilde{w}^T \tilde{w}} = \frac{w^T B w}{w^T w} \leq \beta_k.$$

Drugo neenakost dokažemo podobno. Definiramo vektorske prostore

$$V = \text{Lin}(x_1, \dots, x_{k+n-m}), \quad W = \text{Lin}(y_k, \dots, y_m), \quad \tilde{W} = \left\{ \begin{bmatrix} w \\ 0 \end{bmatrix} \in \mathbb{R}^n, w \in W \right\}.$$

Ker je  $\dim(V) = k+n-m$ ,  $\dim(\tilde{W}) = \dim(W) = m-k+1$ , spet obstaja vektor  $\tilde{w}$  v preseku  $V \cap \tilde{W}$ , kjer je  $\tilde{w} = \begin{bmatrix} w \\ 0 \end{bmatrix}$  za nek  $w \in W$ . Spet velja enakost  $\tilde{w}^T A \tilde{w} = w^T B w$ .

Sledi:

$$\lambda_{k+n-m} = \max_{x \in V} \frac{x^T A x}{x^T x} \geq \frac{\tilde{w}^T A \tilde{w}}{\tilde{w}^T \tilde{w}} = \frac{w^T B w}{w^T w} \geq \min_{x \in W} \frac{x^T B x}{x^T x} = \beta_k,$$

s čimer je izrek dokazan. ■

*Dokaz izreka 6.1:* Karakteristični polinom matrice  $T_i$  je oblike

$$(13) \quad p_i(\lambda) = \prod_{j=1}^i (\lambda_j(T_i) - \lambda).$$

- Vemo, da je  $\lambda < \lambda_n(T_n)$ . Po izreku 6.2 velja, da je  $\lambda_n(T_n) \leq \lambda_j(T_i)$  za vsak  $i = 1, \dots, n$ ,  $j = 1, \dots, i$ . Ker je matrica  $T_i$  pozitivno semidefinitna, so  $\lambda_1(T_i) \geq \dots \geq \lambda_i(T_i) \geq 0$  in zato  $(\lambda_j(T_i) - \lambda) > 0$  za vsaka  $1 \leq i \leq n$ ,  $1 \leq j \leq i$ . Od tod sledi, da je  $p_i(\lambda) > 0$  za  $i = 1, \dots, n$ .
- Če odvajamo polinom, podan v (13), dobimo:

$$p_i'(\lambda) = - \sum_{k=1}^i \frac{p_i(\lambda)}{\lambda_k(T_i) - \lambda}.$$

Vemo, da je  $\frac{p_i(\lambda)}{\lambda_k(T_i) - \lambda} > 0$  za vsak  $k = 1, \dots, i$ , zato je  $p_i'(\lambda) < 0$ .

- Drugi odvod polinoma, podanega v (13), izgleda takole:

$$p_i''(\lambda) = \sum_{k=1}^i \sum_{j \neq k} \frac{p_i(\lambda)}{(\lambda_k(T_i) - \lambda)(\lambda_j(T_i) - \lambda)}.$$

Podobno kot prej lahko vidimo, da bo  $\frac{p_i(\lambda)}{(\lambda_k(T_i) - \lambda)(\lambda_j(T_i) - \lambda)} > 0$  za vsaka  $k, j = 1, \dots, i$ ,  $j \neq k$  in zato je  $p_i''(\lambda) > 0$ .

S tem smo dokazali lastnosti v (11). Želimo dokazati še rekurzivno formulo v (12). Oglejmo si, kako izgleda matrica  $T_{i+1} - \lambda I_{i+1}$ :

$$T_{i+1} - \lambda I_{i+1} = \begin{bmatrix} T_i - \lambda I_i & E_i t^{(i)} \\ t^{(i)T} E_i & t_0 - \lambda \end{bmatrix}.$$

Ker je  $T_{i+1} - \lambda I_{i+1}$  bločna matrica, lahko njeno determinanto zapišemo kot

$$(14) \quad \det(T_{i+1} - \lambda I_{i+1}) = \det(T_i - \lambda I_i) \cdot \det(t_0 - \lambda - t^{(i)T} E_i (T_i - \lambda I_i)^{-1} E_i t^{(i)}),$$

kjer smo upoštevali, da lahko determinanto bločne matrice  $L = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ , če je  $A$  obrnljiva, izračunamo s formulo:

$$\det(L) = \det(A) \cdot \det(D - CA^{-1}B).$$

Opazimo, da je matrika  $t_0 - \lambda - t^{(i)T} E_i (T_i - \lambda I_i)^{-1} E_i t^{(i)}$  velikosti  $1 \times 1$ , kar pomeni, da je determinanta enaka kar njej sami. Ker je  $\det(T_i - \lambda I_i) = p_i(\lambda)$ , lahko enakost iz (14) zapišemo kot

$$p_{i+1}(\lambda) = p_i(\lambda) \left( t_0 - \lambda - t^{(i)T} E_i (T_i - \lambda I_i)^{-1} E_i t^{(i)} \right),$$

kar pa je ravno to, kar smo želeli dokazati. ■

Karakteristični polinom matrice torej lahko izračunamo rekurzivno po formuli (12). Avtorja članka [1] sta lahko uporabila metode, ki smo jih izpeljali prej. Za primerjavo sta izbrala navadno in izboljšano Newtonovo metodo ter diskretno Laguerrovo metodo, pri kateri je bil prvi korak narejen z izboljšano Newtonovo metodo. Ker so vse matrice pozitivno semidefinitne, KMS matrice pa so celo pozitivno definitne, vemo, da imajo lastne vrednosti večje ali enake 0. Vemo torej, da je 0 najmanjša možna lastna vrednost, kar potem pripomore k ustrezni izbiri začetnega približka. Za matrice velikosti  $100 \times 100$  sta avtorja dobila rezultate v tabeli 1. V njej vidimo povprečno število korakov, ki jih je posamezna metoda potrebovala za približek z natančnostjo  $10^{-8}$ .

	CVL	KMS	DDM
Newtonova metoda	10.46	125.39	231.13
izboljšana Newtonova metoda	9.98	97.61	20.40
diskretna Laguerrova metoda	7.20	25.29	14.91

TABELA 1. Primerjava metod za  $n = 100$

Vidimo, da je diskretna Laguerrova metoda potrebovala občutno najmanj korakov za dovolj dober približek. Opazimo tudi, da ima izboljšana Newtonova metoda res veliko hitrejšo konvergenco od osnovne verzije.

## 7. PRIMERJAVA METOD ZA PROBLEM ISKANJA NAJMANJŠE LASTNE VREDNOSTI SIMETRIČNE TRIDIAGONALNE MATRIKE

Oglejmo si primer, ko imamo podano matriko velikosti  $n \times n$ , za katero želimo poiskati najmanjšo lastno vrednost. Obravnavamo samo simetrične matrice, ker vemo, da so vse lastne vrednosti simetričnih matrik realne, kar zahtevamo v naših predpostavkah v poglavju (2). Če vzamemo pozitivno definitne ali pozitivno semidefinitne matrice, vemo, da nobena izmed lastnih vrednosti ni manjša od 0, zato lahko 0 vzamemo za začetni približek. V nasprotnem primeru lahko za ocenjevanje začetnega približka uporabimo naslednji izrek:

**Izrek 7.1. Geršgorinov izrek** *Naj bo  $A = [a_{ij}]$  kompleksna matrika dimenzije  $n \times n$  in naj bo  $R_i = \sum_{i \neq j} |a_{ij}|$ . Velja, da vsaka lastna vrednost matrice  $A$  leži v enem izmed krogov  $K_i$  v kompleksni ravnini s središčem v  $a_{ii}$  in radijem  $R_i$ , za  $i=1, \dots, n$ .*



*Dokaz:* Naj bo  $\lambda$  lastna vrednost matrike  $A$  in naj bo  $x = (x_j)$  pripadajoč lastni vektor. Za  $x$  naj velja, da je  $x_i = 1$  za neko komponento  $x_i$  in  $|x_j| \leq 1$  za  $j \neq i$ . Tak vektor lahko dobimo tako, da vse komponente lastnega vektorja za  $\lambda$  delimo s komponento, ki ima največjo absolutno vrednost. Ker velja  $Ax = \lambda x$ , velja tudi:  $\sum_j a_{ij}x_j = \lambda x_i = \lambda$ , ker je  $x_i$  enak 1. Vsoto zapišemo malce drugače:

$$\sum_{j \neq i} a_{ij}x_j + a_{ii}x_i = \sum_{j \neq i} a_{ij}x_j + a_{ii} = \lambda.$$

Sedaj na obeh straneh te enačbe odštejemo  $a_{ii}$  in uporabimo trikotniško neenakost.

$$|\lambda - a_{ii}| = \left| \sum_{j \neq i} a_{ij}x_j \right| \leq \sum_{j \neq i} |a_{ij}| |x_j| \leq \sum_{j \neq i} |a_{ij}| = R_i. \quad \blacksquare$$

Ker imamo v simetričnih matrikah samo realne lastne vrednosti, vemo, da se vse lastne vrednosti nahajajo v uniji intervalov

$$\bigcup_i [a_{ii} - R_i, a_{ii} + R_i], \quad i = 1, \dots, n.$$

Za začetni približek torej lahko vzamemo  $\min_i \{a_{ii} - R_i\}$ .

Recimo, da imamo podano tridiagonalno simetrično matriko  $T$ :

$$(15) \quad T = \begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix}, \quad b_i \neq 0 \text{ za vsak } i.$$

Dokažemo lahko, da so lastne vrednosti take matrike vedno enostavne.

**Trditev 7.2.** *Nerazcepna simetrična tridiagonalna matrika ima same enostavne lastne vrednosti.*

*Dokaz:* Naj bo  $T$  simetrična tridiagonalna matrika, kot je definirana v (15). Tako definirana matrika je nerazcepna, saj je  $b_i \neq 0$  za vsak  $i = 1, \dots, n - 1$ . Za vsak  $\lambda \in \mathbb{R}$  je  $T - \lambda I$  ranga vsaj  $n - 1$ , saj je prvih  $n - 1$  stolpcev očitno linearno neodvisnih. Ker je  $T$  simetrična, vemo, da se algebraična in geometrijska večkratnost vsake lastne vrednosti ujemata. To pomeni, da se število, kolikokrat se lastna vrednost  $\lambda$  pojavlja kot ničla karakterističnega polinoma, ujema z dimenzijo  $\ker(T - \lambda I)$ . Vemo tudi, da velja enakost  $\text{rang}(T - \lambda I) + \dim(\ker(T - \lambda I)) = n$ . Ker je v našem primeru  $\text{rang}(T - \lambda I) = n - 1$ , sledi, da je  $\dim(\ker(T - \lambda I)) = 1$  za vsak  $\lambda$ , torej so vse lastne vrednosti enostavne.  $\blacksquare$

V knjigi [3] vidimo, da lahko karakteristični polinom  $f_r(\lambda) := \det(T_r - \lambda I)$  vodilne  $r \times r$  podmatrike  $T_r$  izračunamo rekurzivno:

$$(16) \quad f_r(\lambda) = (a_r - \lambda)f_{r-1}(\lambda) - b_{r-1}^2 f_{r-2}(\lambda)$$

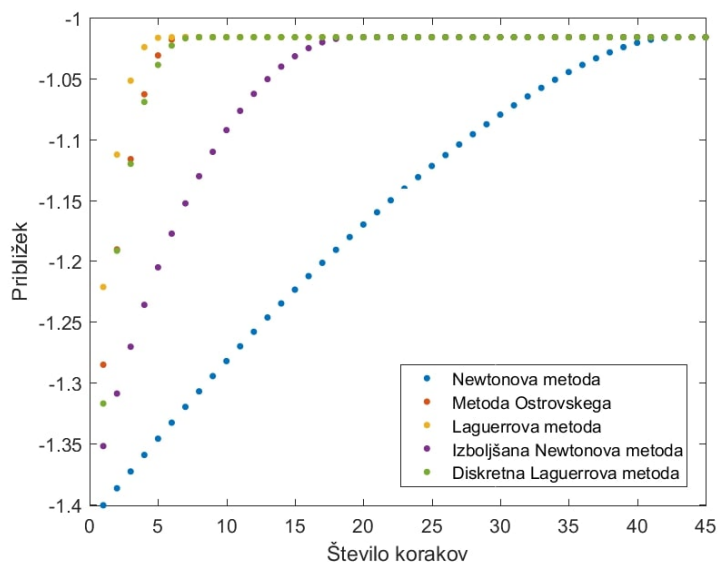
za  $r = 2, \dots, n$ , kjer velja  $f_0(\lambda) = 1$  in  $f_1(\lambda) = a_1 - \lambda$ . V naših metodah za iskanje ničel potrebujemo tudi prvi in drugi odvod. Dobimo ju tako, da poiščemo prvi in drugi odvod rekurzivne formule (16):

$$(17) \quad \begin{aligned} f'_r(\lambda) &= -f_{r-1}(\lambda) + (a_r - \lambda)f'_{r-1} - b_{r-1}^2 f'_{r-2}(\lambda), \\ f'_0(\lambda) &= 0, \quad f'_1(\lambda) = -1, \end{aligned}$$

$$(18) \quad \begin{aligned} f_r''(\lambda) &= -2f_{r-1}' + f_{r-1}''(\lambda)(a_r - \lambda) - b_{r-1}^2 f_{r-2}''(\lambda), \\ f_0''(\lambda) &= 0, \quad f_1''(\lambda) = 0. \end{aligned}$$

Za uporabo naših metod je dovolj, če znamo izračunati le vrednosti karakterističnega polinoma v točkah naših približkov in vrednosti njegovih prvih in drugih odvodov. Ker za karakteristični polinom tridiagonalne matrike to lahko naredimo s formulami (16), (17) in (18), s tem prihranimo pri številu operacij, ki bi jih sicer potrebovali za izračun karakterističnega polinoma.

Za primerjavo smo vzeli naključno simetrično tridiagonalno matriko velikosti  $100 \times 100$ , katere elementi so naključne vrednosti z enakomerno porazdelitvijo na intervalu  $[0, 1]$ . Začetni približek smo izračunali s pomočjo izreka 7.1, naša zahtevana natančnost pa je bila  $10^{-15}$ . Na sliki 1 vidimo, kako so konvergirale naše metode.



SLIKA 1. Primerjava metod pri tridiagonalni matriki velikosti  $100 \times 100$

Opazimo, da je najhitreje konvergirala Laguerrova metoda, ki je za dovolj natančen približek potrebovala le 8 korakov, najpočasneje pa je konvergirala Newtonova metoda, ki je potrebovala 45 korakov.

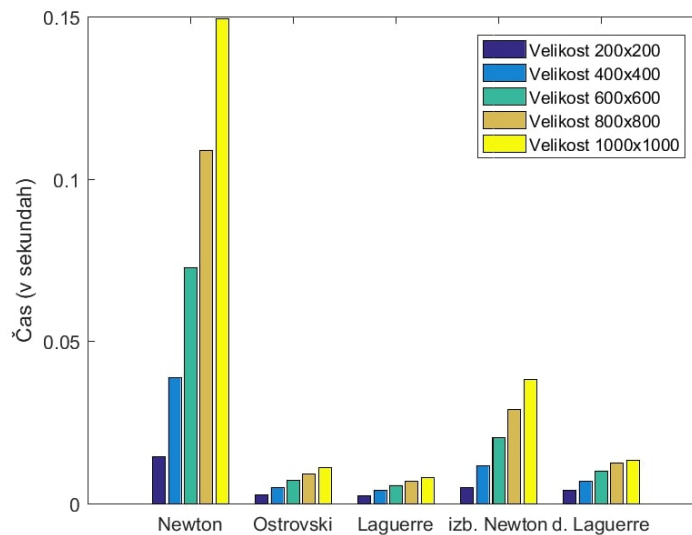
Oglejmo si sedaj še tabelo 2, ki prikazuje povprečno število korakov, ki jih posamezna metoda potrebuje za izračun dovolj dobrega približka za najmanjšo lastno vrednost podane matrike. Rezultati so povprečje za 200 naključno generiranih tridiagonalnih simetričnih matrik, ki imajo vrednosti naključno vzete z intervala  $[0, 0.5]$  v skladu z enakomerno zvezno porazdelitvijo. Povprečno število korakov smo izračunali za več različnih velikosti matrik:  $200 \times 200$ ,  $400 \times 400$ ,  $600 \times 600$ ,  $800 \times 800$  in  $1000 \times 1000$ .

Laguerrova metoda je pri vseh različnih velikostih matrik potrebovala najmanj korakov, največ pa Newtonova metoda. Vidimo, da izboljšana Newtonova metoda res bistveno izboljša hitrost konvergence in za dovolj dober približek porabi med četrtino in tretjino korakov, ki jih za enako dober približek potrebuje klasična Newtonova metoda.

Metoda	Povprečno št. korakov				
	$200 \times 200$	$400 \times 400$	$600 \times 600$	$800 \times 800$	$1000 \times 1000$
Newtonova metoda	83.215	150.033	221.070	286.845	341.63
Metoda Ostrovskega	9.960	12.140	14.115	15.505	16.200
Laguerrova metoda	7.440	8.670	9.775	10.555	10.710
Izbolj. Newtonova metoda	28.080	44.465	61.570	76.980	87.215
Diskr. Laguerrova metoda	11.705	13.455	15.150	16.245	16.875

TABELA 2. Število potrebnih iteracij za dovolj dober približek

Število korakov, ki jih za dovolj dober približek potrebuje neka metoda, ni nujno najboljši pokazatelj učinkovitosti metode. Če za izračun enega koraka na primer porabi veliko več operacij kot ostale metode, se lahko zgodi, da bomo z ostalimi metodami hitreje naredili več potrebnih korakov za dovolj dober približek, kot s časovno zahtevno metodo manj potrebnih korakov. Preveriti moramo torej, kako se obnesejo naše metode. Vemo, da je Laguerrova metoda najučinkovitejša kar se tiče števila potrebnih korakov za dovolj dober približek, vemo pa tudi, da je za en korak Laguerrove metode potrebnih več operacij, ker moramo izračunati tudi drugi odvod.



SLIKA 2. Čas, ki ga metode potrebujejo za dovolj dober približek

Če pogledamo sliko 2 in jo primerjamo s tabelo 2, vidimo, da več časa za dovolj dober približek potrebujejo metode, ki potrebujejo tudi večje število korakov. Izkaže se torej, da se v našem primeru splača investirati več operacij na vsakem koraku za Laguerrovo metodo, saj je skupni čas računanja vseeno najnižji.

Definirajmo sedaj tridiagonalno matriko s konstantnimi vrednostmi na diagonalni in na obdiagonalah:

$$(19) \quad A_n = \begin{bmatrix} a & b & & & \\ c & a & b & & \\ & c & a & b & \\ & & \ddots & \ddots & \ddots \\ & & & c & a & b \\ & & & & c & a \end{bmatrix}, \quad bc \neq 0, \text{ kjer je } A_n \text{ velikosti } n \times n.$$

V članku [9] je izpeljana formula, s katero lahko direktno izračunamo lastne vrednosti matrike  $A_n$ :

$$(20) \quad \lambda_k = a + 2\sqrt{bc} \cdot \cos \frac{k\pi}{n+1}, \quad k \in \{1, 2, \dots, n\}.$$

Ker se v našem primeru ukvarjamo s simetričnimi tridiagonalnimi matrikami, bomo obravnavali primer, ko velja  $b = c$ . Formula (20) se torej poenostavi v:

$$(21) \quad \lambda_k = a + 2b \cdot \cos \frac{k\pi}{n+1}, \quad k \in \{1, 2, \dots, n\}.$$

Ker sedaj poznamo natančno vrednost najmanjše lastne vrednosti, lahko vidimo, kakšen je red konvergence naših metod. Primerjamo lahko namreč približke pridobljene z našimi metodami in točno vrednost lastne vrednosti, ki jo izračunamo s formulo (21).

Najprej si oglejmo, kaj iz teorije vemo o redu konvergence naših metod. Vemo, da je red konvergence zaporedja približkov  $(x_k)$  enak  $p$ , če obstaja takšna konstanta  $C > 0$ , da za ničlo  $\xi$  velja

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - \xi|}{|x_k - \xi|^p} = C.$$

Definirajmo sedaj napako po  $k$  korakih iteracije:  $e_k = x_k - \xi$ . Vidimo, da ko gre  $k \rightarrow \infty$ , velja:  $|e_{k+1}| \approx C \cdot |e_k|^p$  in  $|e_k| \approx C \cdot |e_{k-1}|^p$ . Iz tega sledi:

$$\frac{|e_{k+1}|}{|e_k|} \approx \frac{C \cdot |e_k|^p}{C \cdot |e_{k-1}|^p} \approx \left| \frac{e_k}{e_{k-1}} \right|^p.$$

Če levo in desno stran logaritmiramo, lahko iz enačbe izrazimo  $p$ .

$$(22) \quad \begin{aligned} \log \left| \frac{e_{k+1}}{e_k} \right| &\approx p \cdot \log \left| \frac{e_k}{e_{k-1}} \right|, \\ p &\approx \frac{\log \left| \frac{e_{k+1}}{e_k} \right|}{\log \left| \frac{e_k}{e_{k-1}} \right|}, \\ p &\approx \frac{\log |(x_{k+1} - \xi)/(x_k - \xi)|}{\log |(x_k - \xi)/(x_{k-1} - \xi)|}. \end{aligned}$$

Za uporabo formule (22) moramo poznati vrednosti razmerja napak na dveh zaporednih korakih. V našem primeru to lahko izračunamo, saj znamo izračunati tudi

natančno vrednost lastne vrednosti. Za izračun reda konvergence lahko torej uporabimo to formulo. Če točne vrednosti lastne vrednosti ne poznamo, lahko približek za  $p$  izračunamo tudi s pomočjo naslednje formule:

$$(23) \quad p \approx \frac{\log |(x_{k+1} - x_k)/(x_k - x_{k-1})|}{\log |(x_k - x_{k-1})/(x_{k-1} - x_{k-2})|}.$$

Celotno izpeljavo formule (23) lahko najdemo v [12].

V teoriji ob naših predpostavkah velja, da ima Newtonova metoda kvadratičen red konvergence, če je  $p''(\xi) \neq 0$ , saj zaradi trditve 7.2 vemo, da so vse ničle karakterističnega polinoma enostavne. Če imamo večkratno ničlo, pa v splošnem velja, da je konvergenca linearna. To lahko vidimo v knjigi [3].

Glede na to da se izboljšana Newtonova metoda od klasične razlikuje le v prvih nekaj približkih, na dolgi rok pa sta metodi enaki, tudi za izboljšano Newtonovo metodo pričakujemo kvadratično konvergenco.

Iz [2] vemo, da za zaporedje približkov  $(x_k)$  za enostavno ničlo, pridobljenih z metodo Ostrovskega, velja:

$$\lim_{k \rightarrow \infty} \frac{\xi - x_{k+1}}{(\xi - x_k)^3} = \frac{3p''(\xi)^2 - 4p'(\xi)p'''(\xi)}{24p'(\xi)^2}.$$

Ker je desna stran neodvisna od  $k$ , je to konstanta, ki jo iščemo, in je red konvergence metode Ostrovskega enak 3 za enostavne ničle. Če imamo večkratne ničle, pa velja za  $r > 0$ , kjer je  $r$  kratnost ničle  $\xi$ :

$$\lim_{k \rightarrow \infty} \frac{\xi - x_{k+1}}{\xi - x_k} = 1 - \frac{1}{\sqrt[r]{r}}.$$

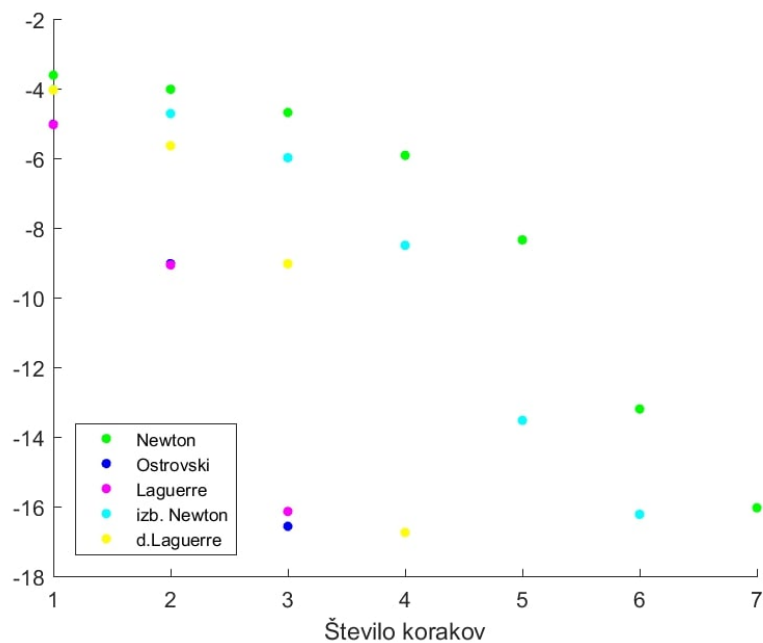
Konvergenca je v tem primeru torej linearna. Iz [5] vemo tudi, da je konvergenca Laguerrove metode kubična, če imamo enostavne ničle, v bližini večkratnih ničel pa je linearna. Laguerrova metoda sicer spada med metode za iskanje ničel, ki konvergirajo ne glede na izbiro začetnega približka.

Poglejmo torej graf na sliki 3, ki prikazuje odvisnost  $\log_{10}(|x_k - \lambda|)$  od števila korakov, kjer je  $x_k$   $k$ -ti približek za najmanjšo lastno vrednost,  $\lambda$  pa točna vrednost lastne vrednosti, izračunana s formulo (21). Za izračun smo vzeli naključno generirano simetrično tridiagonalno matriko velikosti  $100 \times 100$ , ki ima po diagonalah konstantne vrednosti, ki so naključno vzete z intervala  $[-10, 10]$  v skladu z enakomerno zvezno porazdelitvijo.

V tabeli 3 so zbrane vrednosti, ki so vrisane na grafu 3. Po formuli (22) smo za vsako metodo posebej iz vsakih treh zaporednih vrednosti izračunali še oceno za red konvergence  $p_i$ . Te vrednosti so zbrane v tabeli 4

št. koraka	1	2	3	4	5	6	7
Newton	-3.612	-4.019	-4.685	-5.915	-8.3442	-13.201	-16.039
Ostrovski	-5.023	-9.026	-16.571				
Laguerre	-5.033	-9.063	-16.144				
Izbolj. Newton	-4.029	-4.714	-5.986	-8.500	-13.527	-16.225	
Diskr. Laguerre	-4.035	-5.641	-9.029	-16.747			

TABELA 3.  $\log_{10} |x_k - \lambda|$  za različne  $k$



SLIKA 3. Odvisnost  $\log_{10}(|x_k - \lambda|)$  od števila korakov

	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$
Newton	1.6370	1.8479	1.9749	1.9990	0.5843
Ostrovski	1.8848				
Laguerre	1.7567				
Izbolj. Newton	1.8555	1.9776	1.9993	0.5366	
Diskr. Laguerre	2.1097	2.2781			

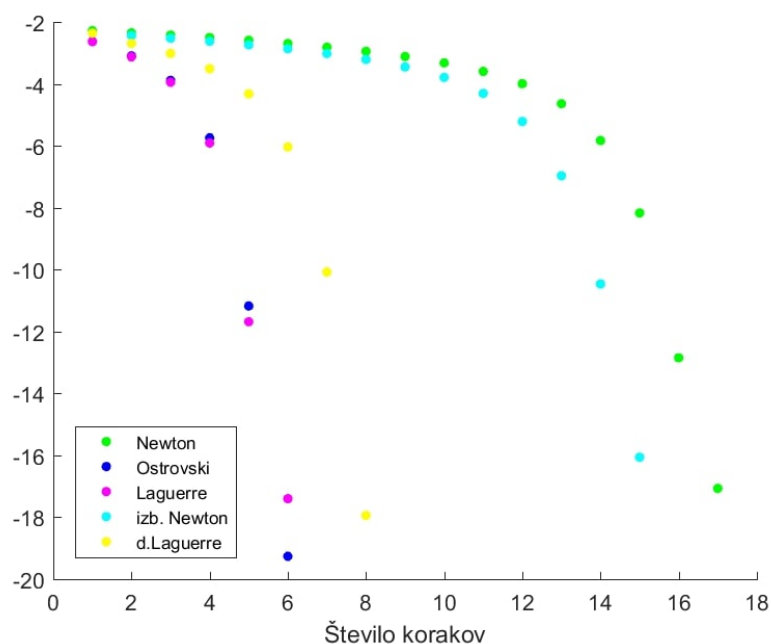
TABELA 4. Izračunan  $p$  za različne trojice zaporednih približkov

Vidimo, da je vrednost na zadnjem koraku smiselno odstraniti iz obravnave, saj nam zaradi numeričnih napak pri zadnjem koraku vrne napačno vrednost. Pri metodi Ostrovskega in Laguerrovi metodi zadnjih približkov ne moremo kar tako izpustiti, saj imamo potem premalo približkov za izračun reda konvergence. Zato bomo začetni približek zmotili za  $5 \cdot 10^{-3}$ . V tem primeru pri metodi Ostrovskega in Laguerrovi metodi pridobimo še 3 nove približke. Na sliki 4 vidimo, kako so konvergirale ostale metode, red konvergence pa bomo ponovno izračunali le za omenjeni dve metodi. Rezultati so zbrani v tabeli 5. V tabeli smo že izpustili zadnjo, napačno, oceno reda konvergence.

	$p_1$	$p_2$	$p_3$
Ostrovski	1.6690	2.34330	2.9301
Laguerre	1.6851	2.3834	2.9417

TABELA 5. Izračunan  $p$  za metodo Ostrovskega in Laguerrovo metodo

Vidimo, da se red konvergence vseh metod, izračunan iz predzadnjih treh približkov, ujema s teorijo. Sklepamo lahko tudi, da je red konvergence diskretne



SLIKA 4. Odvisnost  $\log_{10}(|x_k - \lambda|)$  od števila korakov pri novem začetnem približku

Laguerrove metode sicer slabši od kubičnega, a je vseeno precej boljši od kvadratičnega. Če opazujemo tabelo 2 in sliko 2, vidimo, da je v primeru naših matrik diskretna Laguerrova metoda po hitrosti konvergence in časovni zahtevnosti precej bližje kubičnim kot kvadratičnim metodam.

Izračunajmo  $p$  še za simetrično tridiagonalno matriko, ki po diagonalah nima konstantnih vrednosti. Ker za takšno matriko ne znamo natančno izračunati vrednosti njene najmanjše lastne vrednosti, v tem primeru uporabimo formulo (23). Red konvergence izračunamo na primeru simetrične tridiagonalne matrike velikosti  $500 \times 500$ , ki ima vrednosti naključno izbrane z intervala  $[0,1]$  v skladu z enakomerno zvezno porazdelitvijo. Za izračun reda konvergence uporabimo zadnje 4 približke. Dobimo rezultate, zbrane v tabeli 6.

Po trditvi 7.2 vemo, da imajo matrike, ki jih obravnavamo, same enostavne lastne vrednosti, zato reda konvergence ob večkratnih lastnih vrednostih ne moremo preveriti. Vseeno pa lahko najdemo primer matrike, ki ima vse lastne vrednosti skoraj dvojne, torej lastne vrednosti nastopajo v parih, kjer sta lastni vrednosti v paru zelo blizu ena drugi. Tak primer je matrika  $M_\epsilon$  velikosti  $2n$ , ki izgleda takole:





Iz

$$\|(\Lambda - \lambda(\epsilon)I)^{-1}\| = \frac{1}{\min_{i=1,\dots,n} |\lambda_i - \lambda(\epsilon)|}$$

sledi

$$1 \leq \frac{1}{\min_{i=1,\dots,n} |\lambda_i - \lambda(\epsilon)|} \cdot \epsilon \|X^{-1}\| \cdot \|X\| \cdot \|E\|,$$

zato je

$$\min_{i=1,\dots,n} |\lambda_i - \lambda(\epsilon)| \leq \epsilon \kappa(X) \|E\|.$$

■

**Izrek 7.5. (Weylov izrek).** Če sta  $A$  in  $E$  simetrični matriki in so  $\lambda_n \leq \dots \leq \lambda_1$  lastne vrednosti matrike  $A$ ,  $\tilde{\lambda}_n \leq \dots \leq \tilde{\lambda}_1$  pa lastne vrednosti matrike  $A + E$ , potem za  $i = 1, \dots, n$  velja

$$|\lambda_i - \tilde{\lambda}_i| \leq \|E\|.$$

*Dokaz* To je neposredna posledica izreka 7.4: če sta  $A$  in  $E$  simetrični matriki, potem je  $X$  iz diagonalizacije ortogonalna matrika, torej je  $\kappa(X) = 1$ . ■

Vidimo, da z uporabo izreka 7.5 lahko ocenimo, za koliko se bodo med sabo razlikovale lastne vrednosti matrike  $M_\epsilon$  in lastne vrednosti matrike  $M_0$ . Izračunati moramo le  $\|E\|_2$ , kar pa je v našem primeru enostavno. Ker je  $E$  simetrična matrika, vemo, da je  $\|E\|_2 = \max(|\lambda_1(E)|, |\lambda_{2n}(E)|)$ , kjer je  $\lambda_1(E)$  največja,  $\lambda_{2n}(E)$  pa najmanjša lastna vrednost matrike  $E$ . V našem primeru je tako  $\|E\|_2 = \epsilon$ . Velja torej, da se lastne vrednosti matrike  $M_\epsilon$  od lastnih vrednosti matrike  $M_0$  ne razlikujejo za več kot  $\epsilon$ . Velja tudi, da manjši kot bo  $\epsilon$ , bližje skupaj bodo ležali pari lastnih vrednosti v  $M_\epsilon$ .

Za naš izračun približka za red konvergence smo vzeli matriko, kjer je  $n = 250$  in  $\epsilon = 10^{-8}$ ,  $a$  je slučajni vektor dolžine  $n$ ,  $b$  pa slučajni vektor dolžine  $n - 1$ . Vrednosti obeh vektorjev smo v skladu z enakomerno zvezno porazdelitvijo vzeli z intervala  $(0, 1)$ . Ocene za red konvergence  $p$  so zbrane v tabeli 6. Teorija pravi, da lahko pričakujemo  $p \approx 1$  za vse naše metode. Vidimo, da se naši rezultati lepo skladajo s teorijo, saj smo pri vseh metodah dobili vrednosti blizu 1.

metoda	$p$ za enojne ničle	$p$ za skoraj dvojne ničle
Newtonova metoda	2.00086	0.93424
metoda Ostrovskega	2.95055	0.91553
Laguerrova metoda	2.88682	0.83399
izboljšana Newtonova metoda	2.00024	0.94609
diskretna Laguerrova metoda	2.40360	0.95491

TABELA 6. Red konvergence različnih metod za enojne in skoraj dvojne ničle

## 8. ZAKLJUČEK

V tem diplomskem seminarju smo s pomočjo vezanih optimizacijskih problemov izpeljali pet metod za iskanje ničel polinomov. Te metode smo med sabo primerjali na problemu iskanja najmanjše lastne vrednosti tridiagonalne simetrične matrike. Želeli smo ugotoviti, katero metodo se nam najbolj splača uporabiti, ko rešujemo

tak problem. Že s pomočjo slike 1 in tabele 2 smo ugotovili, da Laguerrova metoda potrebuje najmanj korakov, da pride do približka za lastno vrednost, v katerem je vrednost karakterističnega polinoma dovolj blizu 0.

Ker za korak Laguerrove metode in metode Ostrovskega potrebujemo izračun drugega odvoda karakterističnega polinoma, porabimo za en korak teh dveh metod več operacij kot za ostale metode. Zanimalo nas je torej, ali se nam skupno splača vložiti večje število operacij v posamezni korak in uporabiti Laguerrovo metodo ali bomo s kakšno drugo metodo hitreje opravili več potrebnih korakov za dovolj dober približek. Slika 2 nam pove, da je za naš obravnavani problem Laguerrova metoda vseeno najboljša izbira, saj je skupni čas računanja pri njej najkrajši. Ugotovili smo tudi, da je izboljšana Newtonova metoda v praksi bistveno boljša kot Newtonova metoda, čeprav sta v teoriji metodi na dolgi rok enaki.

Zanimal nas je tudi red konvergence posameznih metod. Pri vseh metodah, razen pri diskretni Laguerrovi metodi, smo teoretično znali izračunati red konvergence, zanimalo pa nas je še, kakšen red konvergence imajo metode v našem primeru. Rezultati, ki smo jih dobili v tabelah 4, 5 in 6, nam povedo, da se eksperimentalni rezultati skladajo s teorijo, kar velja tako za konvergenco v bližini enostavnih kot tudi v bližini dvojnih ničel.

#### SLOVAR STROKOVNIH IZRAZOV

**diagonally dominant matrix** diagonalno dominantna matrika  
**magnified step** povečan korak  
**shifted moment** premaknjen moment  
**harmonic mean** harmonična sredina  
**zerofinders** metode za iskanje ničel

#### LITERATURA

- [1] B. Gragg in A. Melman, *An optimization framework for polynomial zerofinders*, Am. Math. Mon. **113**, 794 - 804 (2006).
- [2] A. M. Ostrowski, *Solution of Equations and Systems of Equations*, Academic Press, New York, London, 1966.
- [3] B. Plestenjak, *Razširjen uvod v numerične metode*, Matematika - fizika št. 52, DMFA - založništvo, Ljubljana, 2015.
- [4] K. A. Redish, *On Laguerre's method*, Int. J. Sci. Math. Educ. **5**, 91 - 102 (1974).
- [5] A. Ralston, *A First Course in Numerical Analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill Kōgakusha, New York, Tokio, 1965.
- [6] X. Zou, *Analysis of the quasi-Laguerre method*, Numer. Math. **82**, 491 - 519 (1999).
- [7] N. Mastronardi in D. Boley, *Computing the smallest eigenpair of a symmetric positive definite Toeplitz matrix*, SIAM J. Sci. Comput. **20**, 1921 - 1927 (1999).
- [8] L. V. Foster, *Generalizations of Laguerre's method: higher order methods*, SIAM J. Numer. Anal. **18**, 1004 - 1018 (1981).
- [9] Wen-Chyuan Yueh, *Eigenvalues of several tridiagonal matrices*, Appl. Math. E-Notes **5**, 66 - 74 (2005).
- [10] J. R. Silvester, *Determinants of block matrices*, Math. Gaz. **84**, 460 - 467 (2000).
- [11] H. Möller, *Convergence and visualization of Laguerre's rootfinding algorithm*, [ogled 6. 9. 2018], dostopno na <https://arxiv.org/pdf/1501.02168.pdf>.
- [12] J. R. Senning, *Computing and estimating the rate of convergence*, [ogled 14. 5. 2019], dostopno na <http://www.math-cs.gordon.edu/courses/ma342/handouts/rate.pdf>.
- [13] D. P. Williamson, *Spectral Graph Theory, Lecture 4*, [ogled 4. 6. 2019], dostopno na <https://people.orie.cornell.edu/dpw/orie6334/lecture4.pdf>.

- [14] P. D. Powell, *Calculating Determinants of Block Matrices*, [ogled 5. 6. 2019], dostopno na [https://www.researchgate.net/publication/51966305\\_Calculating\\_Determinants\\_of\\_Block\\_Matrices](https://www.researchgate.net/publication/51966305_Calculating_Determinants_of_Block_Matrices).
- [15] *Diagonally dominant matrix* - *Wikipedia*, [ogled 4. 9. 2018], dostopno na [https://en.wikipedia.org/wiki/Diagonally\\_dominant\\_matrix](https://en.wikipedia.org/wiki/Diagonally_dominant_matrix).
- [16] *Harmonic mean* - *Wikipedia*, [ogled 6. 9. 2018], dostopno na [https://en.wikipedia.org/wiki/Harmonic\\_mean](https://en.wikipedia.org/wiki/Harmonic_mean).
- [17] *Gershgorin circle theorem* - *Wikipedia*, [ogled 11. 9. 2018], dostopno na [https://en.wikipedia.org/wiki/Gershgorin\\_circle\\_theorem](https://en.wikipedia.org/wiki/Gershgorin_circle_theorem).
- [18] *Rank - nullity theorem* - *Wikipedia*, [ogled 11. 4. 2019], dostopno na [https://en.wikipedia.org/wiki/Rank-nullity\\_theorem](https://en.wikipedia.org/wiki/Rank-nullity_theorem).
- [19] *Matrix Manual: Matrix Property Proofs*, [ogled 5. 6. 2019], dostopno na <http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/proof003.html>.

#### SLIKE

1	Primerjava metod pri tridiagonalni matriki velikosti $100 \times 100$	18
2	Čas, ki ga metode potrebujejo za dovolj dober približek	19
3	Odvisnost $\log_{10}( x_k - \lambda )$ od števila korakov	22
4	Odvisnost $\log_{10}( x_k - \lambda )$ od števila korakov pri novem začetnem približku	23

#### TABELE

1	Primerjava metod za $n = 100$	16
2	Število potrebnih iteracij za dovolj dober približek	19
3	$\log_{10}  x_k - \lambda $ za različne $k$	21
4	Izračunan $p$ za različne trojice zaporednih približkov	22
5	Izračunan $p$ za metodo Ostrovskega in Laguerrovo metodo	22
6	Red konvergence različnih metod za enojne in skoraj dvojne ničle	25