UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Matematika – 1. stopnja

Aljaž Ostrež

Matrične potence

Delo diplomskega seminarja

Mentorica: izr. prof. dr. Marjeta Kramar Fijavž
Somentor: doc. dr. Pavle Boškoski

Ljubljana, 2021

Kazalo

Matrične potence

Povzetek

V delu diplomskega seminarja obravnavamo matrične potence in zaporedja matričnih potenc. Sprva navedemo motivacijska zgleda in naredimo pregled pojmov iz linearne algebre, potrebnih za razumevanje besedila. Nato pogledamo matrične polinome in funkcije ter navedemo nekaj osnovnih pojmov iz spektralne teorije. V osrednjem delu si ogledamo koordinatna zaporedja, asimptotsko obnašanje matričnih zaporedij in Cesàrovo konvergenco. Na koncu znanje, pridobljeno tekom diplomskega dela, uporabimo za analizo markovskih verig.

Powers of Matrices

Abstract

In the thesis we discuss matrix powers and sequences of matrix powers. We first give motivational examples and make an overview of the concepts from linear algebra needed to understand the text. We look at matrix polynomials and functions and list some basic concepts from spectral theory. In the central part, we look at coordinate sequences, asymptotic behavior of matrix sequences, and Cesàro summability. Finally, the knowledge gained during the thesis is used to analyze Markov chains.

Math. Subj. Class. (2020): 15A16, 15A18, 60J10

Ključne besede: matrika, matrične potence, markovska veriga, lastna vrednost, spektralni projektor, matrično zaporedje, Cesàrova konvergenca

Keywords: matrix, powers of matrices, Markov chain, eigenvalue, spectal projection, matrix sequence, Cesàro summability

1. Uvod

Matrike so matematični objekti, ki so zelo uporabni v različnih vejah matematike in ostalih znanostih. V teoriji grafov matrika sosednosti enolično določa graf, ki je lahko utežen ali pa neutežen. V linearni algebri jih uporabljamo na primer za zapis linearnih preslikav med vektorskimi prostori in za proučevanje koeficientov sistemov linearnih enačb. V nadaljevanju bodo vse matrike kompleksne. To pomeni, da so koordinate matrike $A = [a_{ij}]_{1\le i,j\le n}$ kompleksna števila, tj. $a_{ij} \in \mathbb{C}$ za $i, j = 1, \ldots, n$.

V nekaterih primerih pa nam poleg matrik pridejo prav tudi matrične funkcije. Eden od primerov matričnih funkcij je funkcija

$$f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$$

s predpisom

$$f(A) = A^k$$

za $k \in \mathbb{N}$, kar je ravno k-ta potenca matrike $A \in \mathbb{C}^{n\times n}$. S pomočjo te funkcije lahko sestavimo zaporedje oblike $\left(A^k\right)_{k\in\mathbb{N}}$, ki je zelo uporabno v slučajnih procesih (stran 6).

Želeli bomo opisati asimptotsko obnašanje matričnih zaporedij s splošnim členom $A_n = A^n$, kjer je $A \in \mathbb{R}^{m\times m}$. To obnašanje se lahko razbere že iz spektra matrike, včasih pa celo le iz spektralnega radija.

Večino predstavljenih rezultatov povzemamo iz knjige [1, 2. poglavje, 3. poglavje, dod. A].

Z naslednjimi zgledi bomo še dodatno motivirali uporabo matričnih potenc.

1.1. **Potence in Jordanova forma.**

Najprej si poglejmo uporabo matričnih potenc na primeru Fibonaccijevega zaporedja.

Zgled 1.1 (Fibonaccijevo zaporedje)**.** Poznamo rekurzivno formulo za Fibonaccijevo zaporedje:

$$\begin{aligned} &f_0 = 0, f_1 = 1 \\ &f_{n+1} = f_n + f_{n-1}, n \geq 1. \end{aligned}$$

Z uvedbo zaporedja $g_n = f_{n-1}$ dobimo sistem:

$$\begin{aligned} f_{n+1} &= f_n + g_n \\ g_{n+1} &= f_n \end{aligned}$$

ob začetnih pogojih $f_1 = 1$ in $g_1 = 0$. Ta sistem lahko zapišemo v matrični obliki:

$$\begin{bmatrix} f_{n+1} \\ g_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_n \\ g_n \end{bmatrix}$$

za vsak $n \in \mathbb{N}$. Rekurzivno uporabljamo zgornji predpis, da dobimo

$$\begin{bmatrix} f_{n+1} \\ g_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_n \\ g_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^2 \begin{bmatrix} f_{n-1} \\ g_{n-1} \end{bmatrix} = \cdots = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} f_1 \\ g_1 \end{bmatrix}$$

oziroma

$$\begin{bmatrix} f_{n+1} \\ g_{n+1} \end{bmatrix} = \begin{bmatrix} f_{n+1} \\ f_n \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Z uporabo matričnih potenc torej lahko dobimo ekplicitno formulo za splošni člen v Fibonaccijevemu zaporedju

$$f_n = a_{21},$$

kjer je a_{21} prvi element v drugi vrstici matrike:

$$A^n = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n.$$

Tekom študija smo že spoznali en način za računanje matričnih potenc – potence matrik lahko računamo s prevedbo na Jordanovo formo.

Naj bo $A \in \mathbb{R}^{m \times m}$ matrika. Potem velja:

$$A^n = PJ^nP^{-1},$$

kjer je J Jordanova forma matrike A, P pa pripadajoča prehodna matrika. Vemo, da je

$$J^n = \begin{bmatrix} J_1^n & & \\ & \ddots & \\ & & J_k^n \end{bmatrix},$$

kjer so $J_i, i = 1, \ldots, k$, Jordanovi bloki. Za Jordanove bloke velja

$$J_i^n = \begin{bmatrix} \lambda_i^n & \binom{n}{1}\lambda_i^{n-1} & \cdots & \binom{n}{m_i}\lambda_i^{n-m_i} \\ & \lambda_i^n & \cdots & \binom{n}{m_i-1}\lambda_i^{n-(m_i-1)} \\ & & \ddots & \vdots \\ & & & \lambda_i^n \end{bmatrix},$$

pri čemer je λ_i lastna vrednost matrike A, $m_i \times m_i$ pa velikost Jordanovega bloka. V posebnem primeru, ko je J diagonalna matrika, velja:

$$J^n = \begin{bmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_m^n \end{bmatrix}.$$

Z izračunom lastnih vrednosti in lastnih vektorjev bi za matriko:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

dobili Jordanovo formo (oziroma v tem primeru kar diagonalizacijo) matrike A in prehodno matriko:

$$J = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}, \qquad P = \begin{bmatrix} \frac{1-\sqrt{5}}{2} & \frac{1+\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}.$$

Naš sistem enačb lahko zapišemo kot:

$$\begin{bmatrix} f_{n+1} \\ f_n \end{bmatrix} = A^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} = PJ^nP^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Izračunamo, da je:

$$f_n = \frac{\sqrt{5}}{5} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right], \quad n \in \mathbb{N}_0.$$

S pomočjo matričnih potenc smo iz rekurzivnega predpisa dobili eksplicitno formulo zaporedja. $\diamondsuit$

Zakaj potrebujemo nove metode za računanje matričnih potenc, če jih že znamo računati s pomočjo Jordanove forme? Izkaže se, da je izračun Jordanove forme numerično zahteven, poleg tega pa nam Jordanova forma omogoča potenciranje matrik le v končnih dimenzijah. Velikokrat nam tudi ni potrebno izračunati celotne matrike A^n, ampak želimo le poznati nekatere njene lastnosti. Namesto Jordanove forme bomo uporabili spektralni razcep, ki ni odvisen od izbire baze matrike A.

1.2. **Markovske verige.** Uporabo matričnih potenc bomo prikazali na primeru markovskih verig. Navedimo definicijo markovske verige.

Definicija 1.2. Naj bo S števna množica, ki jo poimenujemo *množica stanj*. Njene elemente $s \in S$ imenujemo *stanja*. *Slučajni proces* (z diskretnim časom) je vsako zaporedje diskretnih slučajnih spremenljivk $X_0, X_1, \ldots, X_n, \ldots$, katerih zaloga vrednosti leži v S. To zaporedje imenujemo *markovska veriga*, če ima markovsko lastnost:

$$P\left(X_n = s_n \middle| \, X_0 = s_0, X_1 = s_1, \ldots, X_{n-1} = s_{n-1}\right) = P\left(X_n = s_n \middle| \, X_{n-1} = s_{n-1}\right),$$

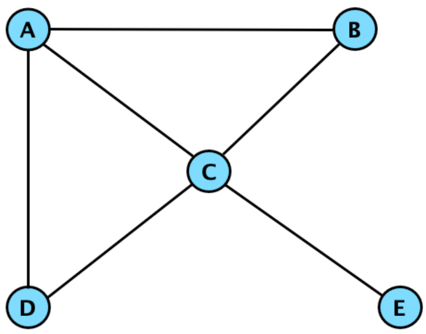
tj. verjetnost stanja na n-tem koraku je odvisna le od stanja na $(n-1)$-tem koraku. Vpeljemo še pojma *prehodne verjetnosti na* n*-tem koraku*

$$p_{ij} = P\left(X_n = s_j \middle| \, X_{n-1} = s_i\right)$$

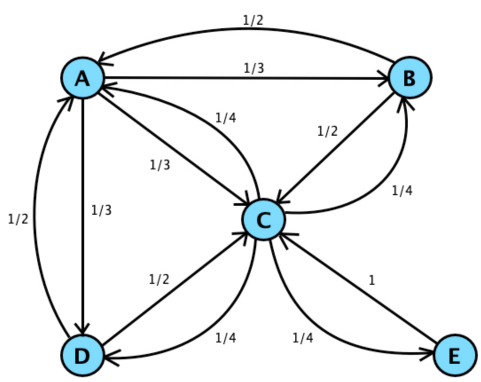
in *prehodne matrike* $P = [p_{ij}]_{1 \leq i,j \leq n}$.

Poglejmo si zgled markovske verige, ki bo hkrati tudi motiviral računanje limite zaporedja $A_n = A^n$, če ta obstaja.

Zgled 1.3. Izberimo si vozlišče v poljubnem neusmerjenem grafu. Izberemo si naključnega soseda izbranega vozlišča in se premaknemo v njega. Postopek ponavljamo. Kolikšen je delež obiskov določenega vozlišča v grafu po dolgem času? Ker soseda izbiramo naključno, začetnemu grafu priredimo utežen usmerjen graf, kjer so uteži verjetnosti, da se premaknemo v povezanega soseda. Tu privzamemo, da so izbire sosedov enako verjetne, za vozlišče stopnje k so vse verjetnosti enake $\frac{1}{k}$.



Slika 1. Začetni neusmerjen graf.



Slika 2. Prirejen utežen usmerjen graf.

Naša množica stanj je $S = \{A, B, C, D, E\}$, prehodna matrika pa bo kar matrika sosednosti prirejenega uteženega usmerjenega grafa

$$P = \begin{array}{c} \\ {\scriptstyle A} \\ {\scriptstyle B} \\ {\scriptstyle C} \\ {\scriptstyle D} \\ {\scriptstyle E} \end{array} \begin{array}{c} \begin{array}{ccccc} {\scriptstyle A} & {\scriptstyle B} & {\scriptstyle C} & {\scriptstyle D} & {\scriptstyle E} \end{array} \\ \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}.$$

Ker je izbira naslednjega vozlišča v zaporedju odvisna le od trenutnega vozlišča, ima naše zaporedje markovsko lastnost, torej bo naša pot v grafu markovska veriga. Kaj nam pa pove prehodna matrika?

- Stolpci matrike P nam povejo, kakšna je verjetnost, da bomo prišli v določeno vozlišče v naslednjem koraku, če smo trenutno v vozlišču, ki ga predstavlja vrstica.
- Stolpci matrike P^2 nam povejo, kakšna je verjetnost, da bomo prišli v določeno vozlišče čez 2 koraka.
- ...
- Stolpci matrike P^n nam povejo, kakšna je verjetnost, da bomo prišli v določeno vozlišče čez n korakov.

Za naš primer se izkaže celo, da obstaja limita:

$$\lim_{n\to\infty} P^n = \begin{bmatrix} 0,25 & 0,17 & 0,33 & 0,17 & 0,08 \\ 0,25 & 0,17 & 0,33 & 0,17 & 0,08 \\ 0,25 & 0,17 & 0,33 & 0,17 & 0,08 \\ 0,25 & 0,17 & 0,33 & 0,17 & 0,08 \\ 0,25 & 0,17 & 0,33 & 0,17 & 0,08 \end{bmatrix}.$$

Ta limita nam pove delež obiskov za vsako vozlišče (če bi zelo dolgo potovali). Limita je tudi neodvisna od izbire začetnega vozlišča, če je le graf krepko povezan [2, Theorem 5.1]. $\diamondsuit$

1.3. **Ponovitev iz linearne algebre.** Naj bo $A \in \mathbb{C}^{n\times n}$ matrika dimenzije $n \times n$. Definiramo *množico polinomov matrike* A s predpisom

$$\mathcal{P}_A := \Big\{ \sum_{i=0}^{m} \alpha_i A^i \mid \alpha_i \in \mathbb{C}, m \in \mathbb{N}_0 \Big\} \subset \mathbb{C}^{n\times n}.$$

Če je $p(x) = \sum_{i=0}^{m} \alpha_i x^i$ polinom, bomo pisali

$$p(A) := \sum_{i=0}^{m} \alpha_i A^i.$$

Definiramo preslikavo

$$\begin{aligned} \Phi_A : \mathbb{C}[x] &\longrightarrow \mathcal{P}_A \\ p &\longmapsto p(A), \end{aligned} \tag{1}$$

kjer je $\mathbb{C}[x]$ množica polinomov s koeficienti iz $\mathbb{C}$. Jedro preslikave Φ_A je

$$\ker(\Phi_A) = \Big\{ p \in \mathbb{C}[x] \mid p(A) = 0 \Big\},$$

torej, $\ker(\Phi_A)$ je množica vseh polinomov, za katere je $p(A) = 0$.

Definicija 1.4. *Minimalni polinom* $m_A \in \mathbb{C}[x]$ je neničelen monični polinom najmanjše stopnje iz $\ker(\Phi_A)$, tj. monični polinom p najnižje stopnje, za katerega je $p(A) = 0$.

Opomba 1.5. *Monični polinom* je polinom z vodilnim koeficientom enakim 1.

Spomnimo se, da je minimalni polinom m_A enolično določen z matriko A. Potrebovali bomo še pojem *lastne vrednosti* in *lastnega vektorja* matrike A, osvežimo pa še definicijo *karakterističnega polinoma* matrike A. Poleg tega navedimo še definicijo *lastnega podprostora.*

Definicija 1.6. Naj bo $A \in \mathbb{C}^{n\times n}$.

(1) Število $\lambda \in \mathbb{C}$ imenujemo *lastna vrednost* matrike A, če obstaja tak vektor $x \in \mathbb{C}^n$, $x \neq 0$, da velja
$$Ax = \lambda x.$$
Vektorji x, za katere velja zgornja enačba, so pripadajoči *lastni vektorji.*

(2) *Lastni podprostor* matrike A, ki pripada lastni vrednosti λ, je definiran kot $\ker(A - \lambda I)$.

(3) Polinom
$$\Delta_A(\lambda) = \det(A - \lambda I) \in \mathcal{P}_A$$
v spremenljivki λ imenujemo *karakteristični polinom* matrike A.

(4) *Algebraična večkratnost* lastne vrednosti λ je večkratnost λ kot ničle polinoma $\Delta_A(\lambda)$. *Geometrična večkratnost* lastne vrednosti λ pa je dimenzija pripadajočega lastnega podprostora $\ker(A - \lambda I)$.

Matrika $A - \lambda I$ je *nilpotentna* reda ν, kjer je ν geometrična večkratnost lastne vrednosti λ. To pomeni, da velja $(A - \lambda I)^k = 0$ za $k \geq \nu$.

Brez dokaza se spomnimo naslednje trditve, s pomočjo katere med drugim tudi na roke iščemo lastne vrednosti matrike A.

Trditev 1.7. *Naj bo $A \in \mathbb{C}^{n\times n}$, m_A pripadajoči minimalni polinom, Δ_A pripadajoči karakteristični polinom in $\lambda \in \mathbb{C}$. Naslednje trditve so ekvivalentne:*

(1) *λ je lastna vrednost matrike A,*
(2) *matrika $A - \lambda I$ ni obrnljiva,*
(3) *λ je ničla polinoma m_A,*
(4) *λ je ničla polinoma Δ_A.*

Iz zgornje trditve sledi, da lahko minimalni polinom m_A matrike A zapišemo kot
$$m_A(z) = (z - \lambda_1)^{\nu_1} (z - \lambda_2)^{\nu_2} \cdots (z - \lambda_m)^{\nu_m},$$
kjer $\lambda_1, \ldots, \lambda_m$ lastne vrednosti matrike A, $\nu_1, \ldots, \nu_m$ pa pripadajoče geometrične večkratnosti.

V nadaljevanju bomo omenjali tudi *matrične* in *operatorske norme*, zato se spomnimo lastnosti norme.

Definicija 1.8. *Matrična norma* na $\mathbb{C}^{n\times n}$ je takšna preslikava $\|\cdot\| : \mathbb{C}^{n\times n} \to \mathbb{R}$, da za vsaki matriki $A, B \in \mathbb{C}^{n\times n}$ in vsak skalar $\alpha \in \mathbb{C}$ velja:

(1) $\|A\| \geq 0$ in $\|A\| = 0$ natanko tedaj, ko je $A = 0$ (pozitivnost),
(2) $\|\alpha A\| = |\alpha|\|A\|$ (homogenost),
(3) $\|A + B\| \leq \|A\| + \|B\|$ (trikotniška neenakost),
(4) $\|AB\| \leq \|A\|\|B\|$ (submultiplikativnost).

Matrične norme, ki so porojene iz vektorskih norm s predpisom

$$\|A\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} \quad \text{za } x \in \mathbb{C}^n,$$

imenujemo *operatorske norme.*

Zgled 1.9. Matrične norme, ki so porojene iz vektorskih p-norm, imenujemo *p-norme.* To so operatorske norme in so definirane kot

$$\|A\|_p = \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p},$$

kjer je $A \in \mathbb{C}^{n\times n}$. Dva predstavnika p-norm, ki imata še posebej poenostavljen predpis, sta:

- *1-norma*, ki je definirana kot

  $$\|A\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}|,$$

  tj. največja od vsot absolutnih vrednosti elementov v isti vrstici,
- ∞*-norma*, ki je definirana kot

  $$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|,$$

  tj. največja od vsot absolutnih vrednosti elementov v istem stolpcu.

$\diamondsuit$

Operatorska norma je torej podrazdred matričnih norm in ima nekaj zanimivih lastnosti.

Trditev 1.10. *Naj bo $A \in \mathbb{C}^{n\times n}$, $x \in \mathbb{C}^{n\times n}$ in $\|\cdot\|$ operatorska norma na $\mathbb{C}^{n\times n}$. Potem velja*

$$\|Ax\| \leq \|A\| \cdot \|x\|.$$

Dokaz. Za $x = 0$ trditev očitno velja. Za $x \neq 0$ pa velja

$$\frac{\|Ax\|}{\|x\|} \leq \max_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \|A\|,$$

torej je $\|Ax\| \leq \|A\| \cdot \|x\|$. $\square$

Posledica 1.11. *Naj bo $A \in \mathbb{C}^{n\times n}$ in $\|\cdot\|$ neka operatorska norma na $\mathbb{C}^{n\times n}$. Potem velja neenakost*

$$\|A\| \geq \max\{|\lambda| \mid \lambda \textit{ je lastna vrednost matrike } A\}.$$

Dokaz. Po definiciji lastne vrednosti velja $Ax = \lambda x$ za neki lastni vektor $x \in \mathbb{C}^{n\times n}$, $x \neq 0$. Zato velja tudi

$$\|Ax\| = \|\lambda x\| = |\lambda| \cdot \|x\|.$$

Ker je $\|\cdot\|$ operatorska norma, po trditvi 1.10 velja

$$\|Ax\| \leq \|A\| \cdot \|x\|.$$

Sledi, da je

$$|\lambda| \cdot \|x\| \leq \|A\| \cdot \|x\|,$$

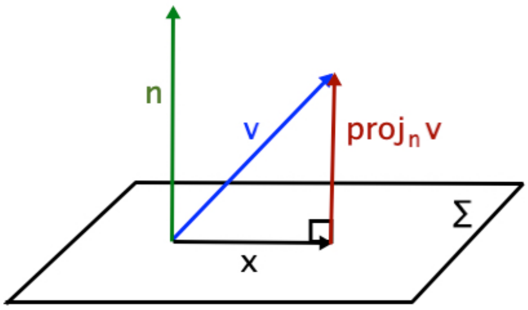
torej je $|\lambda| \leq \|A\|$. To velja za vsako lastno vrednost λ. Posledica sledi. $\square$

Definicija 1.12. Matrika $P \in \mathbb{C}^{n\times n}$ je *projektor*, če velja $P^2 = P$.

Zgled 1.13. Poglejmo si primer pravokotnega projektorja v prostoru $\mathbb{R}^3$. Za pravokotni projektor velja, da je $\ker P \perp \operatorname{Im} P$. Projicirajmo vektor $v = (1,1,0)$ pravokotno na ravnino Σ, ki je podana z enačbo $x+y+z=0$. Naj bo $n=(1,1,1)$ normalni vektor te ravnine, ki ni normiran. Vemo, da s formulo

$$\operatorname{proj}_n(v) = \frac{n\cdot v}{n\cdot n}\cdot n \tag{2}$$

projiciramo vektor v na vektor n. S tem lahko izračunamo tudi projekcijo vektorja v na ravnino Σ.



SLIKA 3. Projekcija vektorja v na ravnino Σ.

Če vstavimo v in n v formulo (2), dobimo

$$\operatorname{proj}_n(v) = \frac{2}{3}\cdot(1,1,1).$$

Ker velja $x = v - \operatorname{proj}_n(v)$, je torej

$$x = \operatorname{proj}_\Sigma(v) = \frac{1}{3}\cdot(1,1,-2).$$

Bolj zanimiv kot le izračun $\operatorname{proj}_\Sigma(v)$ pa je izračun projektorja, ki vektor v projicira na vektor n. Če pišemo skalarni produkt kot množenje matrik, tj. $a\cdot b = a^T\cdot b$, iz asociativnosti matričnega množenja sledi

$$\operatorname{proj}_n(v) = \frac{n^Tv}{n^Tn}\cdot n = \frac{\left(n^Tv\right)n}{n^Tn} = \frac{n\left(n^Tv\right)}{n^Tn} = \frac{nn^T}{n^Tn}\cdot v.$$

S P označimo matriko $\frac{nn^T}{n^Tn}$. V našem primeru izračunamo, da je

$$P = \frac{1}{3}\cdot\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix},$$

velja pa tudi $P^2 = P$, kar pomeni, da je P res projektor. Enačba $\operatorname{proj}_n(v) = Pv$ pa sledi direktno iz izpeljave. $\diamondsuit$

Definicija 1.14. Naj bosta X in Y vektorska prostora nad $\mathbb{C}$, za katera velja $X\cap Y=\emptyset$. Tedaj vsoti

$$X+Y = \{x+y \mid x\in X, y\in Y\}$$

pravimo *direktna vsota* in jo označimo z $X\oplus Y$.

Trditev 1.15. *Matrika $P\in\mathbb{C}^{n\times n}$ je projektor natanko tedaj, ko je*

$$\mathbb{C}^n = \ker P\oplus \operatorname{Im} P \quad \textit{in} \quad P|_{\operatorname{Im} P} = I.$$

Dokaz. ($\Leftarrow$) $P|_{\operatorname{Im} P} = I$ implicira $P\,(Px) = Px$ za vsak $x \in \mathbb{C}^n$, tj. $P^2 = P$, kar je ravno definicija projektorja.

($\Rightarrow$)Naj bo $x \in \mathbb{C}^n$ in $y \in \operatorname{Im} P$, tako da velja $y = Px$. Definirajmo $z := x - y$. Potem je $x = z + y$ in

$$Pz = Px - Py = Px - P^2x = Px - Px = 0,$$

torej je $z \in \ker P$. Lahko sklepamo tudi, da iz $y \in \operatorname{Im} P$ sledi $Py = P^2y = Px = y$.

Dokažimo še, da je razcep na direktno vsoto enoličen. Naj bosta $y' \in \operatorname{Im} P$ in $z' \in \ker P$ takšna, da je $x = y' + z'$. Iz računa

$$y = Px = Py' + Pz' = y'$$

sledi, da je $y = y'$ in $z = z'$. □

2. Spektralni razcep in matrične funkcije

2.1. **Matrični polinomi.** V podpoglavju 1.3 smo že definirali množico matričnih polinomov $\mathcal{P}_A$ in minimalni polinom m_A matrike $A \in \mathbb{C}^{n\times n}$, pa tudi preslikavo Φ_A iz množice polinomov $\mathbb{C}[x]$ v $\mathcal{P}_A$.

Trditev 2.1. *Naj bo m_A minimalni polinom. Potem je*

$$p \in \ker(\Phi_A) \Leftrightarrow p = m_A \cdot q \quad za \quad q \in \mathbb{C}[x],$$

tj. $p(A) = 0$ natanko tedaj, ko je $p(A)$ večkratnik minimalnega polinoma m_A.

Dokaz. Najprej dokažimo implikacijo ($\Leftarrow$). Če je $p = m_A \cdot q$ za nek $q \in \mathbb{C}[x]$, potem je

$$p(A) = m_A(A) \cdot q(A) = 0 \cdot q(A) = 0,$$

torej je $p \in \ker(\Phi_A)$.

Dokažimo še ($\Rightarrow$). Naj bo $p \in \mathbb{C}[x]$, za katerega velja $p(A) = 0$. Če delimo p z m_A, dobimo:

$$p = m_A \cdot q + r$$

za $q, r \in \mathbb{C}[x]$, kjer je $\operatorname{st}(r) < \operatorname{st}(m_A)$. Velja

$$0 = p(A) = \underbrace{m_A(A)}_{=0} \cdot q(A) + r(A) = r(A),$$

torej je $r \in \ker(\Phi_A)$. Ker ima m_A najmanjšo stopnjo v $\ker(\Phi_A)$, sklepamo, da je $r \equiv 0$. □

Za polinome $p, q, r \in \mathbb{C}[x]$ bomo uporabljali zapis

$$p \equiv q \mod r \iff p - q = s \cdot r \;\text{ za nek } s \in \mathbb{C}[x].$$

Poglejmo si posledico trditve 2.1

Posledica 2.2. *Za $p, q \in \mathbb{C}[x]$ velja*

$$p(A) = q(A) \iff p \equiv q \mod m_A.$$

V posebnem primeru velja, da je $p(A)$ projektor natanko tedaj, ko je

$$p^2 \equiv p \mod m_A.$$

Opazimo torej, da je lahko $p(A) = q(A)$ tudi, če polinoma $p, q \in \mathbb{C}[x]$ nista enaka. Naslednja trditev pa nam pove, da lahko enakost po $\mod m_A$ preverimo le s pomočjo lastnih vrednosti matrike A in njihovih večkratnosti.

Trditev 2.3. *Naj bosta $p, q \in \mathbb{C}[x]$ in naj bo λ_i ničla polinomov p in q z večkratnostjo ν_i. Potem velja*

$$p \equiv q \mod m_A \iff p^{(\nu)}(\lambda_i) = q^{(\nu)}(\lambda_i) \text{ za vsak } i = 1, \ldots, m \text{ in } \nu = 0, \ldots, \nu_i - 1.$$

Dokaz. Trditev sledi iz dejstva, da za nek $n \in \mathbb{N}$ in neničeln $s \in \mathbb{C}[x]$ ter fiksen $\lambda_i \in \mathbb{C}$ velja

$$(p - q)(x) = (x - \lambda_i)^n s(x) \iff (p - q)^{(\nu)}(\lambda_i) = 0 \text{ za } i = 0, 1, \ldots, n - 1,$$

kar smo že dokazali pri študiju algebre (knjiga [5, poglavje 7.5]). □

2.2. **Gladke matrične funkcije.** Naj bo $A \in \mathbb{C}^{n \times n}$ matrika z minimalnim polinomom m_A ter $\lambda_1, \ldots, \lambda_m$ ničle m_A z večkratnostimi $\nu_1, \ldots, \nu_m$. Množico funkcij, ki so definirane in neskončnokrat odvedljive na neki okolici $\{\lambda_1, \ldots, \lambda_m\}$, označimo kot

$$C_A^\infty := \{f : D(f) \to \mathbb{C} \mid \exists U \subset D(f) \text{ odprta}, \{\lambda_1, \ldots, \lambda_m\} \subset U, f|_U \in C^\infty\},$$

kjer je $D(f) \subset \mathbb{C}$ definicijsko območje funkcije f.

Pri uporabi funkcije $f \in \mathbb{C}_A^\infty$ na matriki A si bomo pomagali z interpolacijskimi polinomi. *Interpolacijski polinom* je polinom stopnje največ n, ki se v $n + 1$ točkah $(x_0, f(x_0)), (x_1, f(x_1)), \ldots, (x_n, f(x_n))$ ujema s funkcijo f. Če zahtevamo, da se za $i = 0, \ldots, n$ v točkah $(x_i, f(x_i)$ ujemajo tudi vsi odvodi do k_i-tega, pa je stopnja interpolacijskega polinoma navzgor omejena z $\sum_{i=0}^n (k_i + 1) - 1$.

Definicija 2.4. Za $f \in \mathbb{C}_A^\infty$ definiramo matriko $f(A)$ s pomočjo interpolacijskega polinoma:

$$f(A) := \Phi_A(p_f) = p_f(A),$$

kjer je p_f polinomska interpolacija funkcije f, ki zadošča pogoju

$$f^{(\nu)}(\lambda_i) = p_f^{(\nu)}(\lambda_i)$$

za vsak $i = 1, \ldots, m$ in $\nu = 0, \ldots, \nu_i - 1$.

S to definicijo lahko matrične polinome razširimo na množico gladkih matričnih funkcij:

$$\begin{aligned} \widetilde{\Phi} : C_A^\infty &\longrightarrow \mathcal{P}_A \\ \widetilde{\Phi}_A(f) &= \Phi_A(p_f) = p_f(A) \end{aligned} \tag{3}$$

Iz definicije 2.4 in trditev 2.1, 2.3 sledi:

Trditev 2.5. *Naj bodo oznake kot zgoraj. Potem veljajo naslednje trditve:*

(1) *Definicija $\widetilde{\Phi}_A(f)$ ni odvisna od izbire metode za polinomsko interpolacijo p_f.*
(2) *Preslikava $\widetilde{\Phi}_A$ je razširitev funkcije Φ_A.*
(3) *$\widetilde{\Phi}$ je homomorfizem algeber, kar pomeni, da velja:*

$$\widetilde{\Phi}_A(\lambda f + \mu g) = \lambda \widetilde{\Phi}_A(f) + \mu \widetilde{\Phi}_A(g),$$
$$\widetilde{\Phi}_A(f \cdot g) = \widetilde{\Phi}_A(f) \cdot \widetilde{\Phi}_A(g)$$

za $\lambda, \mu \in \mathbb{C}$ in $f, g \in C_A^\infty$.

Za odprto množico $U \subset \mathbb{C}$ definirajmo *karakteristično funkcijo* χ_U s predpisom

$$\chi_U(\lambda) = \begin{cases} 1, \lambda \in U, \\ 0, \lambda \notin U \end{cases} .$$

Funkcija χ_U je idempotent, kar pomeni, da velja $\chi_U \cdot \chi_U = \chi_U$. Po posledici 2.2 sklepamo, da je $\chi_U(A) \in \mathcal{P}_A$ projektor, ki komutira z A. Vzemimo posebno množico takšnih projektorjev.

Definicija 2.6. Naj bodo $U_1, \ldots, U_m \subset \mathbb{C}$ odprte množice, ki zadostujejo pogojema:

(1) $\lambda_i \in U_i$ za $i = 1, \ldots, m$ in
(2) $U_i \cap U_j = \emptyset$ za $i \neq j$.

S χ_i označimo karakteristično funkcijo množice U_i. Matrike

$$P_i := \chi_i(A) \in \mathcal{P}_A \text{ za } i = 1, \ldots, m, \tag{4}$$

so vse projektorji, ki jih imenujemo *spektralni projektorji*, njihovo zalogo vrednosti pa označimo kot

$$X_i := \operatorname{Im} P_i = P_i \mathbb{C}^n. \tag{5}$$

Opomba 2.7.

(1) Omenimo, da je P_i neodvisna od izbire U_i, poleg tega pa velja še $P_i \neq 0$ za vsak i.
(2) Poznamo formulo za izračun spektralnih projektorjev diagonalizabilnih matrik (knjiga [6, 7. poglavje, str. 529, enačba 7.3.11]). Če je A diagonalizabilna matrika z lastnimi vrednostmi $\lambda_1, \ldots, \lambda_m$, potem spektralni projektor za lastno vrednost λ_i izračunamo kot:

$$P_i = \prod_{\substack{j=1 \\ j \neq i}}^{m} \frac{A - \lambda_j I}{\lambda_i - \lambda_j}, \quad i = 1, \ldots, m. \tag{6}$$

Za naslednji izrek moramo razumeti pojem invariantnosti.

Definicija 2.8. Podprostor Y prostora X je *invarianten za A* oz. *A-invarianten*, kadar velja $Ay \in Y$ za vsak $y \in Y$, kar lahko zapišemo kot $AY \subseteq Y$.

Projektorji P_i komutirajo z A, zato so podprostori X_i invariantni za A, in velja naslednje.

Izrek 2.9. *Naj $A \in \mathbb{C}^{n \times n}$ matrika z lastnimi vrednostmi $\lambda_1, \ldots, \lambda_m$ in z geometrijskimi večkratnostmi $\nu_1, \ldots, \nu_m$. Če vzamemo P_i in X_i kot v* (4) *in* (5), *potem je*

$$\mathbb{C}^n = X_1 \oplus \cdots \oplus X_m \tag{7}$$

direktni razcep na A-invariantne podprostore z lastnostjo, da je matrika $A - \lambda_i I$ na X_i nilpotentna reda ν_i, $i = 1, \ldots, m$.

Dokaz. Ker so množice U_i iz definicije 2.6 disjunktne, velja $\chi_i\left(\sum_{i \neq j} \chi_j\right) = 0$, torej tudi
$P_i\left(\sum_{i \neq j} P_j\right) = 0$. Iz tega sledi, da je

$$X_i \cap \left(\bigoplus_{i \neq j} X_j\right) = \{0\}.$$

Ker je $\{\lambda_1, \ldots, \lambda_m\} \subset \bigcup_{i=1}^m U_i =: U$, je

$$\sum_{i=1}^m P_i = \widetilde{\Phi}_A\left(\sum_{i=1}^m \chi_i\right) = \widetilde{\Phi}_A(\chi_U) = I,$$

torej $X_1 \oplus \cdots \oplus X_m = X$.

Spomnimo, da je matrika A nilpotentna reda k, če velja $A^k = 0$ in $A^{k-1} \neq 0$. Za vsak fiksen i je $g_i(\lambda) = (\lambda - \lambda_i)_i^\nu \chi_i(\lambda)$ funkcija v domeni preslikave $\widetilde{\Phi}_A$, ki se v točkah $\lambda_1, \ldots, \lambda_m$ ujema z ničelno funkcijo $\mathbb{O} \in C_A^\infty$, vključno z vsemi odvodi. Zaradi lastnosti $\widetilde{\Phi}_A$ iz trditve 2.5 in definicije 2.4 mora veljati

$$(A - \lambda I)^{\nu_i} P_i = \left((\lambda - \lambda_i)^{\nu_i}\, \chi_i\right)(A) = \mathbb{O}(A) = 0$$

za $i = 1, \ldots, m$ in $\lambda \in \mathbb{C}$. Definirajmo funkcijo

$$f_i(\lambda) = (\lambda - \lambda_i)^{\nu_i - 1}\chi_i(\lambda),\ \lambda \in \mathbb{C}.$$

Ker je

$$f_i(A) = (A - \lambda_i I)^{\nu_i - 1} P_i,$$

velja $(A - \lambda_i I)^{\nu_i - 1} P_i \neq 0$, torej je matrika $A - \lambda_i I$ nilpotentna reda ν_i na X_i. □

Definicija 2.10. Razcepu $\mathbb{C}^n$ na invariante podprostore iz formule (7) rečemo *spektralni razcep* prostora $\mathbb{C}^n$, ki pripada matriki A.

Spektralni razcep nam pove, da je matrika A podobna bločno diagonalni matriki,

$$A \sim \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{bmatrix},$$

kjer ima A_i eno samo lastno vrednost, in sicer λ_i. Iz tega dejstva bi lahko slutili, da bi za podprostore X_i vzeli kar korenske podprostore $\ker(A - \lambda_i I)^{\nu_i}$ za lastne vrednosti λ_i.

Trditev 2.11. *X_i so korenski podprostori matirke A za lastne vrednosti λ_i:*

$$X_i = \ker(A - \lambda_i I)^{\nu_i},$$

kjer je ν_i geometrijska večkratnost lastne vrednosti λ_i.

Dokaz. Trditev 2.9 nam pove, da je $A - \lambda_i I$ nilpotentna reda ν_i na X_i, zato je

$$X_i \subset \ker(A - \lambda_i I)^{\nu_i}. \tag{8}$$

Recimo, da je vsebovanost stroga, tj. obstaja neničeln vektor $x \in \ker(A - \lambda_i I)^{\nu_i} \setminus X_i$. Za neki $i \neq j$ potem obstaja $y \neq 0$, $y := P_j x \in X_j \cap \ker(A - \lambda_i I)^{\nu_i}$. Vzamemo največje število $p \in \mathbb{N}$, za katerega velja

$$z := (A - \lambda_i I)^p y \neq 0.$$

Ker je X_j A-invarianten, je $z \in X_j$. Iz

$$(A - \lambda_i I) z = (A - \lambda_i I)^{p+1} y = 0$$

sledi, da je z lastni vektor za lastno vrednost λ_i, tj. $Az = \lambda_i z$, iz česar sledi, da

$$(A - \lambda_j I)^{\nu_j} z = \left(A - \lambda_i I + \lambda_i I - \lambda_j I\right)^{\nu_j} z = (\lambda_i - \lambda_j)^{\nu_j} z \neq 0,$$

kar je protislovje z lastnostjo, da je $A - \lambda_j I$ na X_j nilpotentna reda ν_j. □

S pomočjo spektralnega razcepa v izreku 2.9 lahko izračunamo $f(A)$ brez iskanja interpolacijskih polinomov.

Izrek 2.12. *Naj bo $A \in \mathbb{C}^{n\times n}$ z lastnimi vrednostmi $\lambda_1, \lambda_2, \ldots, \lambda_m$, ki imajo po vrsti večkratnosti $\nu_1, \nu_2, \ldots, \nu_m$. Definirajmo projektorje P_i kot v (4). Tedaj za vsako funkcijo $f \in C_A^\infty$ velja*

$$f(A) = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i-1} \frac{f^{(\nu)}(\lambda_i)}{\nu!}(A-\lambda_i)^\nu P_i.$$

Dokaz. Funkcija

$$g(\lambda) = \sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i-1} \frac{f^{(\nu)}(\lambda_i)}{\nu!}(\lambda-\lambda_i)^\nu \chi_i(\lambda), \quad \lambda \in \mathbb{C},$$

sovpada z f, vključno z vsemi pomembnimi odvodi, na vseh točkah $\lambda_1, \ldots, \lambda_m$. Po trditvah 2.3, 2.5 in definiciji 2.4 sledi, da je $f(A) = g(A)$. □

Posledica 2.13. *Naj bodo predpostavke enake kot v trditvi 2.12. Za $k \in \mathbb{N}$ velja*

$$A^k = \sum_{i=1}^{m} \sum_{\nu=0}^{\min\{\nu_i-1,k\}} \binom{k}{\nu} \lambda_i^{k-\nu}(A-\lambda_i)^\nu P_i. \tag{9}$$

Dokaz. V trditvi 2.12 za $k \in \mathbb{N}$ uporabimo funkcijo $f(\lambda) = \lambda^k$. □

2.3. **Spektralna teorija.** Tu zberimo še nekaj rezultatov o spektru matrik, ki jih bomo kasneje potrebovali.

Definicija 2.14. Naj bo $A \in \mathbb{C}^{n\times n}$. Množico

$$\sigma(A) := \{\lambda_1, \ldots, \lambda_m\}$$

imenujemo *spekter* matrike A. *Spektralni radij* definiramo kot

$$r(A) := \max\{|\lambda| \mid \lambda \in \sigma(A)\},$$

tj. po absolutni vrednosti največja lastna vrednost matrike A.

Naslednja trditev nam pove, kako se preslika spekter matrike za neko preslikavo iz C_A^∞.

Izrek 2.15 (Izrek o preslikavi spektra)**.** *Naj bo $A \in \mathbb{C}^{n\times n}$ in $f \in C_A^\infty$. Potem velja*

$$\sigma\big(f(A)\big) = f\big(\sigma(A)\big) := \{f(\lambda) \mid \lambda \in \sigma(A)\}.$$

Dokaz. Naj bo λ poljubna lastna vrednost matrike A in $\mu \notin f(\sigma(A))$. Definirajmo $u(\lambda) := \frac{1}{f(\lambda)-\mu} \in C_A^\infty$, torej je $u(\lambda)\cdot(f(\lambda)-\mu) = 1$ na okolici $\sigma(A)$. Če v to enačbo vstavimo matriko A, dobimo

$$u(A)\cdot(f(A)-\mu I) = (f(A)-\mu I)\cdot u(A) = I,$$

torej je matrika $f(A)-\mu I$ obrnljiva, zato $\mu \notin \sigma(f(A))$.

Obratno, naj bo x_i lastni vektor matrike A, ki pripada lastni vrednosti λ_i za $i = 1, \ldots, m$. Po izreku 2.9 za vse $i \neq j$ velja $P_j x_i = 0$ in $P_i x_i = x_i$. Po trditvi 2.12 sledi

$$f(A)x_i = f(\lambda_i)x_i,$$

torej je $f(\lambda_i) \in \sigma(f(A))$. □

Lema 2.16. *Naj bo* $\|\cdot\|$ *norma na* $\mathbb{C}^{n\times n}$*. Za* $A \in \mathbb{C}^{n\times n}$ *in* $\mu > r(A)$ *obstajata konstanti* $N > 0$ *in* $M > 1$*, da za vse* $k \in \mathbb{N}$ *velja*

$$N \cdot r(A)^k \leq \left\|A^k\right\| \leq M \cdot \mu^k.$$

Če je $\|\cdot\|$ *operatorska norma, lahko izberemo* $N = 1$*.*

Dokaz zgornje leme bo sledil v nadaljevanju (na strani 18). Iz te leme sledi zanimiva formula, s pomočjo katere lahko spektralni radij matrike izračunamo s pomočjo matričnih potenc.

Trditev 2.17 (Gelfandova formula)**.** *Za matriko* $A \in \mathbb{C}^{n\times n}$ *velja naslednje.*

(1) $r(A) = \lim_{k\to\infty}\left\|A^k\right\|^{1/k}$ *za vsako matrično normo* $\|\cdot\|$ *na* $\mathbb{C}^{n\times n}$*.*
(2) *Če je* $\|\cdot\|$ *operatorska norma na* $\mathbb{C}^{n\times n}$*, potem je* $r(A) = \inf_{k>0}\left\|A^k\right\|^{1/k}$*.*

Dokaz. Vzemimo k-ti koren enačbe (2.16):

$$N^{1/k} r(A) \leq \left\|A^k\right\|^{1/k} \leq M^{1/k}\mu.$$

Če je $\|\cdot\|$ operatorska norma, za dopustno izbiro $N = 1$ dobimo

$$r(A) = \inf_{k>0}\left\|A^k\right\|^{1/k}. \qquad \square$$

3. Zaporedja matričnih potenc

V tem poglavju bomo opazovali asimptotsko obnašanje matričnih zaporedij oblike $(A^k)_{k\in\mathbb{N}}$. Izkazalo se bo, da nekatere lastnosti lahko karakteriziramo s spektralnim radijem matrike A. Da pridemo do tega rezultata, moramo najprej povedati nekaj o koordinatnih zaporedjih.

3.1. **Koordinatna zaporedja.**

Lema 3.1. *Naj bo* $A \in \mathbb{C}^{n\times n}$ *z lastnimi vrednostmi* $\lambda_1, \ldots, \lambda_m$ *z večkratnostmi* $\nu_1, \ldots, \nu_m$*.*

(1) *Za* $i = 1, \ldots, m$ *in* $0 \neq z \in X_i$ *je množica*

$$\{(A - \lambda_i I)^\nu z \mid \nu = 0, \ldots, \nu_i - 1\}\backslash\{0\}$$

linearno neodvisna v X_i*.*

(2) *Množica*

$$B_A = \{(A - \lambda_i I)^\nu P_i \mid i = 1, \ldots, m;\ \nu = 0, \ldots, \nu_i - 1\}$$

je linearno neodvisna v $\mathbb{C}^n$*.*

Dokaz.

(1) Naj bo $i \in \{1, \ldots, m\}$ in $0 \neq z \in \mathbb{C}^n$. Preveriti moramo, da so vektorji $(A - \lambda_i I)^\nu z$, $\nu = 0, \ldots, \nu_i - 1$ linearno neodvisni. Ker je matrika $A - \lambda_i I$ nilpotentna reda ν_i, je

$$(A - \lambda_i I)^\nu = 0 \text{ za } \nu \geq \nu_i. \tag{10}$$

Po definiciji linearne neodvisnosti mora iz enačbe

$$\sum_{\nu=0}^{\nu_i - 1} \alpha_\nu (A - \lambda_i I)^\nu z = 0 \tag{11}$$

slediti, da je $\alpha_\nu = 0$ za $\nu = 0, \dots, \nu_i - 1$. Če enačbo (11) pomnožimo z $(A - \lambda_i I)^{\nu_i - 1}$ z leve, z upoštevanjem (10) dobimo

$$0 = \alpha_0 (A - \lambda_i I)^{\nu_i - 1} z.$$

Ker $(A - \lambda_i I)^{\nu_i - 1} z \neq 0$, je $\alpha_0 = 0$. Če bi enačbo (11) pomnožili z $(A - \lambda_i I)^{\nu_i - 2}$ in upoštevali, da je $\alpha_0 = 0$, bi dobili $\alpha_1 = 0$, itd. Sledi, da so

$$\alpha_0 = \cdots = \alpha_{\nu_i - 1} = 0$$

in posledično vektorji $(A - \lambda_i I)^\nu z$, $\nu = 0, \dots, \nu_i - 1$ linearno neodvisni.

(2) Podobno kot v dokazu točke (1) sledi, da je množica

$$B_A^{(i)} := \{(A - \lambda_i I)^\nu P_i \mid \nu = 0, \dots, \nu_i - 1\}$$

linearno neodvisna v $\mathbb{C}^{n \times n}$ za vsak $i = 1, \dots, m$. Dokazati moramo, da so matrike $(A - \lambda_i I)^\nu P_i$, $i = 1, \dots, m$, $\nu = 0, \dots, \nu_i - 1$ linearno neodvisne. Denimo, da je:

$$\sum_{i=1}^{m} \sum_{\nu=0}^{\nu_i - 1} \alpha_{i,\nu} (A - \lambda_i I)^\nu P_i = 0.$$

Če to enačbo za fiksen i pomnožimo s P_i in upoštevamo, da je $P_i \cdot P_j = 0$ za $i \neq j$ in $P_i^2 = P_i$, dobimo

$$\sum_{\nu=0}^{\nu_i - 1} \alpha_{i,\nu} (A - \lambda_i I)^\nu P_i = 0.$$

Iz linearne neodvisnosti množice $B_A^{(i)}$ sledi, da je $\alpha_{i,\nu} = 0$ za $i = 1, \dots, m$ in $\nu = 0, \dots, \nu_i - 1$.

□

Opomba 3.2. Množica matrik $\{A_1, \dots, A_k\}$ je linearno neodvisna, če iz

$$\sum_{i=1}^{k} \alpha_i A_i = 0$$

sledi, da so $\alpha_1, \dots, \alpha_k = 0$.

Zanima nas asimptotsko obnašanje zaporedij oblike $(A^k)_{k \in \mathbb{N}}$. Za razumevanje asimptotskega obnašanja bomo uporabili neodvisnost množice B_A iz leme 3.1. Če to množico razširimo do baze $\mathbb{B}_A$ prostora $\mathbb{C}^{n \times n}$, potem iz formule (9) sledi, da so neničelne kooordinate zaporedja $(A^k)_{k \in \mathbb{N}}$ v bazi $\mathbb{B}_A$ enake

$$\left\{ \binom{k}{\nu} \lambda_i^{k-\nu} \mid i = 1, \dots, m;\ \nu = 0, \dots, \nu_i - 1 \right\}$$

(ker nas zanima $k \to \infty$, lahko poenostavimo zgornjo mejo z $\min\{\nu_i - 1, k\}$ na $\nu_i - 1$).

Opomba 3.3. *Baza* $\mathbb{B}$ vektorskega prostora matrik $\mathbb{C}^{n \times n}$ je takšna množica velikosti n^2 baznih matrik $B_i \in \mathbb{C}^{n \times n}$, da lahko vsako matriko $A \in \mathbb{C}^{n \times n}$ zapišemo kot linearno kombinacijo baznih matrik, tj. $A = \alpha_1 B_1 + \cdots + \alpha_{n^2} B_{n^2}$ za $\alpha_1, \dots, \alpha_{n^2} \in \mathbb{C}$.

Podobne koordinate dobimo, če gledamo zaporedje $(A^k x)_{k \in \mathbb{N}}$, torej če gledamo množico

$$\{(A - \lambda_i I)^\nu P_i x \mid i = 1, \dots, m;\ \nu = 0, \dots, \nu_i - 1\} \backslash \{0\}.$$

Ker je konvergenca v končno dimenzionalnih vektorskih prostorih ekvivalentna konvergenci po koordinatah, se (neodvisno od izbire baze) obnašanje zaporedja $(A^k)_{k\in\mathbb{N}}$, ko gre $k \to \infty$, odraža z obnašanjem *koordinatnih zaporedij*

$$z_{\lambda,\nu}(k) := \binom{k}{\nu} \lambda^{k-\nu}$$

za $\lambda \in \sigma(A), \nu = 0, \ldots, n-1$ (tukaj upoštevamo, da je $\nu_i \leq n$ za vsak i). V primeru, da vsa koordinatna zaporedja konvergirajo, konvergira tudi zaporedje $(A^k)_{k\in\mathbb{N}}$ oziroma $(A^k x)_{k\in\mathbb{N}}$. Koordinate limite zaporedja $(A^k)_{k\in\mathbb{N}}$, ko gre $k \to \infty$, se izražajo s pripadajočimi limitami koordinatnih zaporedij.

Konvergenca koordinatnih zaporedij se lahko razbere iz velikosti lastne vrednosti λ in pripadajoče večkratnosti ν.

- Če je $|\lambda| < 1$, potem gre $z_{\lambda,\nu}(k) \to 0$, ko gre $k \to \infty$, za vse ν, ker je $\lim_{k\to\infty} k^\nu \lambda^k = 0$.
- Če je $|\lambda| > 1$, potem gre $|z_{\lambda,\nu}(k)| \to \infty$, ko gre $k \to \infty$, za vse ν.
- Če je $|\lambda| = 1$ in $\nu = 0$, potem je $z_{\lambda,0}(k) = \lambda^k$.
- Če je $|\lambda| = 1$ in $\nu \geq 1$, potem gre $|z_{\lambda,\nu}(k)| \to \infty$, ko gre $k \to \infty$.

Z zgornjimi dejstvi lahko opišemo asimptotsko obnašanje zaporedja $(A^k x)_{k\in\mathbb{N}}$ za $x \in X_i = \operatorname{Im} P_i$ za lastno vrednost λ_i z večkratnostjo ν_i.

Z uvedbo koordinatnih zaporedij lahko dokažemo lemo 2.16, iz katere sledi Gelfandova formula 2.17.

Dokaz. (Lema 2.16) Po izreku o preslikavi spektra 2.15 velja

$$\big(r(A)\big)^k = r(A^k),$$

torej izbira $N = 1$ za operatorske norme sledi iz posledice 1.11. Ocena

$$N \cdot r(A)^k \leq \left\| A^k \right\|$$

za neki $N > 0$ sledi iz dejstva, da so vse norme na končno dimenzionalnih prostorih ekvivalentne.

Dokazati moramo še oceno navzgor. Koordinate zaporedja $(A^k)_{k\in\mathbb{N}}$ glede na bazo $\mathbb{C}^{n\times n}$, ki vsebuje množico

$$B_A := \{(A - \lambda_i I)^\nu P_i \mid i = 1, \ldots, m;\ \nu = 0, \ldots, \nu_i - 1\},$$

so enake

$$z_{\lambda,\nu}(k) := \binom{k}{\nu} \lambda^{k-\nu},$$

kjer je $i = 1, \ldots, m$ in $\nu = 0, \ldots, \nu_i - 1$. Ker velja ocena

$$\binom{k}{\nu} \leq k^\nu,$$

in ker za $|\lambda| < 1$ velja $k^\nu \lambda^{k-\nu} \to \infty$, ko gre $k \to \infty$, za vse $\nu \in \mathbb{N}$, je zaporedje $\frac{1}{\mu^k}(A^k)_{k\in\mathbb{N}}$ omejeno, ko gre $k \to \infty$. Torej je

$$\frac{1}{\mu^k}\left\| A^k \right\|_\infty \leq M \text{ oziroma } \left\| A^k \right\|_\infty \leq M\mu^k$$

za neko konstanto $M \in \mathbb{N}$ in vsak $n \in \mathbb{N}$. Ocena

$$\left\| A^k \right\| \leq M\mu^k$$

spet sledi iz dejstva, da so vse norme ekvivalentne. □

Opomba 3.4. To, da so vse matrične norme ekvivalentne, pomeni, da za poljubni normi $\|\cdot\|_a$ in $\|\cdot\|_b$ obstajata konstanti $C_1, C_2 > 0$, da za vsako matriko $A \in \mathbb{C}^{n\times n}$ velja

$$C_1\|A\|_a \leq \|A\|_b \leq C_2\|A\|_a .$$

3.2. **Asimptotsko obnašanje.**

Definicija 3.5. Naj bo $A \in \mathbb{C}^{n\times n}$ in $\|\cdot\|$ neka matrična norma. Zaporedje $(A^k)_{k\in\mathbb{N}}$ je:

- *omejeno*, če je $\sup_{k\in\mathbb{N}}\left\|A^k\right\| < \infty$,
- *stabilno*, če je $\lim_{k\to\infty}\left\|A^k\right\| = 0$,
- *konvergentno*, če je $\lim_{k\to\infty} A^k = P$ za neko matriko $P \in \mathbb{C}^{n\times n}$,
- *periodično* s periodo p, če je $A^p = I$,
- *Cesàrovo konvergentno*, če obstaja limita $\lim_{k\to\infty}\left(\frac{1}{k}\sum_{l=0}^{k-1} A^l\right)$.

Členom $A^{(k)} = \frac{1}{k}\sum_{l=0}^{k-1} A^l$ pravimo *Cesàrova povprečja.*

Pokazali bomo, da se asimptotsko obnašanje $(A^k)_{k\in\mathbb{N}}$ odraža z velikostjo $r(A)$ v primerjavi z 1. Pred tem se dogovorimo za nekaj oznak.

Koreni enote v $\mathbb{C}$ so rešitve enačbe $z^q = 1$, torej

$$\Gamma_q := \{e^{\frac{2k\pi i}{q}} \mid k = 0, \ldots, q-1\}.$$

Z Γ bomo označili enotsko krožnico v $\mathbb{C}$.

V izreku bomo uporabili tudi naslednjo definicijo.

Definicija 3.6. Lastna vrednost $\lambda_0 > 0$ matrike $A \in \mathbb{C}^{n\times n}$ je *radialno dominantna lastna vrednost*, če je $\lambda_0 \in \sigma(A)$ in velja $|\lambda| < \lambda_0$ za vse $\lambda \in \sigma(A)\backslash\{\lambda_0\}$.

Izrek 3.7. *Naj bo $A \in \mathbb{C}^{n\times n}$. Za matrično zaporedje $(A^k)_{k\in\mathbb{N}}$ veljajo naslednje trditve.*

(1) *Zaporedje $(A^k)_{k\in\mathbb{N}}$ je stabilno natanko tedaj, ko je $r(A) < 1$.*
(2) *Zaporedje $(A^k)_{k\in\mathbb{N}}$ je omejeno natanko tedaj, ko je $r(A) \leq 1$ in imajo vse lastne vrednosti λ_i, za katere velja $|\lambda_i| = 1$, večkratnost 1.*
(3) *Zaporedje $(A^k)_{k\in\mathbb{N}}$ je periodično s periodo p natanko tedaj, ko je omejeno in je $\sigma(A) \subseteq \Gamma_p$.*
(4) *$\lambda_1 = 1 \in \sigma(A)$ in $\lim_{k\to\infty} A^k = P_1$ (P_1 je spektralni projektor A za lastno vrednost λ_1) natanko tedaj, ko je λ_1 radialno dominantna lastna vrednost z večkratnostjo 1.*

Dokaz.

(1) Sledi direktno iz leme 2.16.
(2) ($\Rightarrow$) Recimo, da je zaporedje $(A^k)_{k\in\mathbb{N}}$ omejeno. Iz leme 2.16 sledi, da mora biti v tem primeru $r(A) \leq 1$. Če ima lastna vrednost $\lambda \in \sigma(A)$, za katero velja $|\lambda| = 1$, večkratnost $\nu > 1$, potem koordinatno zaporedje $z_{\lambda,\nu-1}(k)$ zaporedja $(A^k)_{k\in\mathbb{N}}$ glede na bazo $\mathbb{B}_A$ ni omejeno, zato tudi matrično zaporedje ni omejeno. Prišli smo do protislovja.
($\Leftarrow$) Če je $r(A) \leq 1$ in imajo vse lastne vrednosti λ, za katere velja $|\lambda| = 1$, večkratnost 1, potem so koordinatna zaporedja zaporedja $(A^k)_{k\in\mathbb{N}}$ glede na bazo $\mathbb{B}_A$ omejena.

(3) ($\Leftarrow$) Naj bo $(A^k)_{k\in\mathbb{N}}$ omejeno in $\sigma(A) \subseteq \Gamma_p$. Iz (2) sledi, da imajo lastne vrednosti večkratnost 1, zato se formula iz (9) poenostavi v

$$A^k = \sum_{i=1}^{m} \lambda_i^k P_i \text{ za } k \in \mathbb{N}$$

Ker je $\lambda_i^p = 1$ za vse $i = 1, \ldots, m$, sledi $A^p = I$.
($\Rightarrow$) Če je $(A^k)_{k\in\mathbb{N}}$ periodično, tj. $A^p = I$ za neki $p \in \mathbb{N}$, je zagotovo omejeno in za vsak $\lambda \in \sigma(A)$ velja $\lambda^p = 1$, iz česar sledi, da je $\sigma(A) \subseteq \Gamma_p$.

(4) ($\Rightarrow$) Naj bo $\lim_{k\to\infty} A^k = P_1$. Ker limita ni enaka 0, nam (1) pove, da $r(A) \geq 1$. Ker limita obstaja, pomeni, da je zaporedje omejeno, torej iz (2) sledi, da je $r(A) \leq 1$. Sklepamo, da je $r(A) = 1$.

Če obstaja lastna vrednost $\lambda \neq 1$, da velja $|\lambda| = 1$, potem obstaja koordinatno zaporedje zaporedja $(A^k)_{k\in\mathbb{N}}$ glede na bazo $\mathbb{B}_A$ s predpisom $z_{\lambda,0}(k) = \lambda^k$, ki divergira. Lastna vrednost 1 je torej edini kandidat za radialno dominantno lastno vrednost. Če ima 1 večkratnost več kot 1, potem ima $(A^k)_{k\in\mathbb{N}}$ koordinato $z_{\lambda,1}(k) = k$, ki spet divergira. Lastna vrednost 1 ima torej večkratnost 1.
($\Leftarrow$) Če je 1 radialno dominantna lastna vrednost z večkratnostjo 1, potem koordinatno zaporedje $z_{\lambda,0}(k)$ glede na bazo $\mathbb{B}_A$ konvergira. Iz (1) sledi, da koordinatna zaporedja za lastne vrednosti $|\lambda| < 1$ konvergirajo k 0, zato je $\lim_{k\to\infty} A^k = P_1$.

$\square$

Limita zaporedja $(A^k)_{k\in\mathbb{N}}$ torej obstaja samo v dveh primerih:

(1) Ko je $r(A) < 1$ oziroma je zaporedje stabilno. V tem primeru je

$$\lim_{k\to\infty} A^k = 0.$$

(2) Ko je $r(A) = 1 = \lambda_1$ in je λ_1 radialno dominantna lastna vrednost z večkratnostjo 1. Tedaj je

$$\lim_{k\to\infty} A^k = P_1.$$

Poglejmo si uporabo izreka na nekaj zgledih.

Zgled 3.8.

(1) Naj bo

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

S pomočjo karakterističnega polinoma poiščimo lastne vrednosti matrike A.

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2,$$

torej ima A lastno vrednost $\lambda = 1$. $r(A) = 1$, lastna vrednost $\lambda = 1$ pa ima večkratnost 2, zato zaporedje po točki (2) iz izreka ni omejeno. Uporabimo formulo (9) in izračunamo, da lahko splošni člen zaporedja zapišemo kot

$$A^k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}.$$

(2) Naj bo

$$B = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Poiščimo lastne vrednosti matrike B.

$$\det(B-\lambda I) = \begin{vmatrix} \frac{1}{2}-\lambda & 0 \\ \frac{1}{2} & \frac{1}{2}-\lambda \end{vmatrix} = \left(\frac{1}{2}-\lambda\right)^2.$$

B ima lastno vrednost $\lambda = 1/2$ z večkratnostjo 2. Ker je $r(A) = 1/2 < 1$, je zaporedje $(B^k)_{k\in\mathbb{N}}$ stabilno, torej je $\lim_{k\to\infty} B^k = 0$. Splošni člen zaporedja je po formuli (9) enak

$$B^k = \begin{bmatrix} \frac{1}{2^k} & 0 \\ \frac{k}{2^k} & \frac{1}{2^k} \end{bmatrix}.$$

(3) Naj bo

$$C = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Poiščimo lastne vrednosti matrike C.

$$\det(C-\lambda I) = \begin{vmatrix} \frac{1}{2}-\lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}-\lambda \end{vmatrix} = \left(\frac{1}{2}-\lambda\right)^2 - \frac{1}{4} = \lambda^2 - \lambda = \lambda(\lambda-1).$$

Lastni vrednosti C sta $\lambda_1 = 0$ in $\lambda_2 = 1$. Ker je $\lambda_2 = 1$ radialno dominantna lastna vrednost, je $\lim_{k\to\infty} C^k = P_2$, kjer je P_2 spektralni projektor za lastno vrednost λ_2. Ker je matrika C dimenzije 2 in ima dve različni lastni vrednosti, je C diagonalizabilna, torej lahko P_2 izračunamo s pomočjo formule za spektralne projektorje diagonalizabilnih matrik (6) in dobimo $P_2 = C$.

(4) Naj bo

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

Poiščimo lastne vrednosti matrike D.

$$\begin{aligned}\det(D-\lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & 0 & -\lambda \end{vmatrix} &= (-\lambda)^3 + 1 = \\ &= -(\lambda-1)(\lambda^2+\lambda+1) = \\ &= -(\lambda-1)\left(\lambda + \frac{1}{2} - i\frac{\sqrt{3}}{2}\right)\left(\lambda + \frac{1}{2} + i\frac{\sqrt{3}}{2}\right),\end{aligned}$$

torej ima D lastne vrednosti

$$\lambda_1 = 1,\ \lambda_2 = -\frac{1}{2} + i\frac{\sqrt{3}}{2},\ \lambda_3 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

Absolutne vrednosti vseh lastnih vrednosti so 1, zato je $r(D) = 1$, velja pa še $\sigma(D) \subseteq \Gamma_3$. Po točki (3) iz izreka sledi, da je zaporedje $(D^k)_{k\in\mathbb{N}}$ periodično s periodo 3. Členi zaporedja so:

$$D^1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, D^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, D^3 = I, D^4 = D^1, \ldots$$

To zaporedje je očitno tudi omejeno, saj zavzame le 3 vrednosti.

$\diamondsuit$

3.3. **Cesârova konvergenca.** V tem podpoglavju bomo povedali nekaj več o Cesàrovi konvergenci, ki smo jo definirali v definiciji 3.5 na strani 19.

Dokaz naslednje leme bomo izpustili, bralec ga lahko najde v [1, str. 38, lema 3.9].

Lema 3.9 (Kronecker). *Naj bodo $\lambda_1, \ldots, \lambda_r \in \Gamma$, kjer je Γ enotska kompleksna krožnica. Tedaj obstaja zaporedje $(s_k)_{k\in\mathbb{N}} \subset \mathbb{N}$, da velja:*

$$\lim_{k\to\infty} \lambda_i^{s_k} = 1$$

za vse $i = 1, \ldots, r$.

Iz Kroneckerjeve leme sledi naslednja trditev, ki pove nekaj o konvergenci podzaporedja zaporedja $(A^k)_{k\in\mathbb{N}}$.

Izrek 3.10. *Naj bo $A \in \mathbb{R}^{n\times n}$. Tedaj sta naslednji trditvi ekvivalentni.*

(1) *Obstaja konvergentno podzaporedje zaporedja $(A^k)_{k\in\mathbb{N}}$, ki konvergira k limiti $P \neq 0$.*
(2) *$r(A) = 1$ in vse lastne vrednosti λ, za katere velja $|\lambda| = 1$, imajo večkratnost 1.*

V tem primeru je limita oblike

$$P = \sum_{|\lambda_i|=1} P_i,$$

kjer so P_i spektralni projektorji, ki pripadajo lastnim vrednostim λ_i, za katere velja $|\lambda_i| = 1$.

Dokaz. (1) $\Rightarrow$ (2): Naj obstaja podzaporedje zaporedja $(A^k)_{k\in\mathbb{N}}$, ki konvergira k limiti $P \neq 0$. Označimo z I indeksno množico tega zaporedja. Implikacija sledi iz točke (2) v izreku 3.7.
(2) $\Rightarrow$ (1): Ker je $r(A) = 1$ in imajo vse lastne vrednosti, ki ležijo na kompleksni enotski krožnici, večkratnost 1, po točki (2) iz izreka 3.7 sledi, da je zaporedje $(A^k)_{k\in\mathbb{N}}$ omejeno. Iz analize vemo, da ima vsako omejeno zaporedje neko podzaporedje, ki je konvergentno.

Dokazati moramo še, da je limita podzaporedja enaka $P = \sum_{|\lambda_i|=1} P_i$. Naj bodo $\lambda_1, \ldots, \lambda_r$ lastne vrednosti, za katere velja $|\lambda_i| = 1$ za vsak $i = 1, \ldots, r$. Iz Kroneckerjeve leme sledi, da obstaja zaporedje $(s_k)_{k\in\mathbb{N}} \subset \mathbb{N}$, da velja

$$\lim_{k\to\infty} \lambda_i^{s_k} = 1, \quad i = 1, \ldots, r.$$

Za podzaporedje $(A^{s_k})_{k\in\mathbb{N}}$ uporabimo formulo (9) in upoštevamo, da imajo lastne vrednosti λ_i večkratnost 1:

$$\lim_{k\to\infty} A^{s_k} = \sum_{i=1}^{r} \lambda_i^{s_k} P_i = \sum_{i=1}^{r} P_i.$$

S tem smo dokazali ustrezno obliko limite. $\square$

Definirajmo izraz za zaporedje, za katerega veljajo lastnosti iz prejšnjega izreka.

Definicija 3.11. Matriko $A \in \mathbb{R}^{n\times n}$ imenujemo *spektralna skrčitev*, če ustreza ekvivalentnim lastnostim iz izreka 3.10.

Zgled 3.12. Primer spektralne skrčitve je matrika

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

za katero smo v zgledu 3.8 izračunali, da ima 3 različne lastne vrednosti na kompleksni enotski krožnici:

$$\lambda_1 = 1, \ \lambda_2 = -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \ \lambda_3 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}. \qquad \diamondsuit$$

Sedaj si poglejmo še Cesàrovo konvergenco, torej konvergenco zaporedja s členi $A^{(k)} = \frac{1}{k}\sum_{\nu=0}^{k-1} A^\nu$. Videli bomo, da je zaporedje $(A^k)_{k\in\mathbb{N}}$ Cesàrovo konvergentno natanko tedaj, ko je omejeno (torej, ko velja $r(A) < 1$ ali pa je A spektralna skrčitev).

Izrek 3.13. *Naj bo $A \in \mathbb{R}^{n\times n}$. Naslednje izjave so ekvivalentne.*

(1) *Zaporedje $(A^k)_{k\in\mathbb{N}}$ je Cesàrovo konvergentno.*
(2) $\lim_{k\to\infty}\left(k^{-1}A^k\right) = 0$.
(3) *Zaporedje $(A^k)_{k\in\mathbb{N}}$ je omejeno.*
(4) $r(A) \le 1$ *in imajo vse lastne vrednosti λ, za katere velja $|\lambda| = 1$, večkratnost* 1.

Če je zaporedje $(A^k)_{k\in\mathbb{N}}$ Cesàrovo konvergentno, tj. velja katerikoli od zgornjih pogojev, potem zaporedje $(A^{(k)})_{k\in\mathbb{N}}$ konvergira proti 0, če $1 \notin \sigma(A)$, oziroma proti spektralni projekciji P_1, ki pripada lastni vrednosti $1 \in \sigma(A)$.

Dokaz. Implikacija (1) ⇒ (2) sledi, ker lahko zapišemo

$$A^k = kA^{(k)} - (k-1)A^{(k-1)}.$$

(2) ⇒ (3): Če zaporedje $(A^k)_{k\in\mathbb{N}}$ ni omejeno, limita $\lim_{k\to\infty}(k^{-1}A^k)$ ne obstaja.
(3) ⇒ (4): Sledi direktno iz točke (2) v izreku 3.7.
(4) ⇒ (1): Poglejmo si koordinatno zaporedje zaporedja $(A^{(k)})_{k\in\mathbb{N}}$ glede na bazo $\mathbb{B}_T$:

$$z_{\lambda,\nu}^{(k)} := \frac{1}{k}\sum_{l=0}^{k-1} z_{\lambda,\nu}(l) = \frac{1}{k}\sum_{l=0}^{k-1}\binom{l}{\nu}\lambda^{l-\nu}.$$

To zaporedje očitno divergira, če $|\lambda| > 1$ in konvergira, če $|\lambda| < 1$. Za $|\lambda| = 1$ pa je to zaporedje omejeno le, če je $\nu = 0$, torej kadar ima λ večkratnost 0.
Pokazati moramo še, kam konvergira zaporedje $(A^{(k)})_{k\in\mathbb{N}}$ v odvisnosti od spektralnega razcepa. Če $\lambda \neq 1$, potem je

$$z_{\lambda,0}^{(k)} = \frac{1}{k}\sum_{l=0}^{k-1}\binom{l}{0}\lambda^l = \frac{1}{k}\frac{\lambda^k - 1}{\lambda - 1}.$$

To zaporedje torej konvergira proti 0, ko $k \to \infty$, če $|\lambda| \le 1$, $\lambda \neq 1$. Če pa je $\lambda = 1$, potem je $z_{1,0}^{(k)} = 1$ za vsak $k \in \mathbb{N}$, zato v tem primeru zaporedje $(A^{(k)})_{k\in\mathbb{N}}$ konvergira proti P_1, ko gre $k \to \infty$.

□

Poglejmo si zgled matrike, katere zaporedje potenc je Cesàrovo konvergentno.

Zgled 3.14. V zgledu 3.8 smo srečali matriko

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

ki je imela 3 različne lastne vrednosti na kompleksni enotski krožnici, ki so imele večkratnost 1. Po izreku 3.13 sledi, da je zaporedje $(D^k)_{k\in\mathbb{N}}$ Cesàrovo konvergentno.

Ugotovili smo že, da je zaporedje periodično s periodo 3, torej zagotovo ni konvergentno. Poglejmo pa si zaporedje

$$D^{(k)} = \left(\frac{1}{k} \sum_{l=0}^{k-1} D^l \right)_{k\in\mathbb{N}}$$

in poskusimo izračunati limito tega zaporedja. Za predstavo si oglejmo nekaj začetnih členov zgornjega zaporedja:

$$D^{(1)} = I, \qquad D^{(2)} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \qquad D^{(3)} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

$$D^{(4)} = \frac{1}{4}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad D^{(5)} = \frac{1}{5}\begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \end{bmatrix}, \quad D^{(6)} = \frac{1}{6}\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} = D^{(3)}.$$

Izrek 3.13 pove, da je limita zaporedja $\left(D^{(k)}\right)_{k\in\mathbb{N}}$ v primeru, ko je $\sigma(D) = 1$, enaka spektralnemu projektorju P_1 za lastno vrednost 1 matrike D. Izračunamo, da je

$$\lim_{k\to\infty} D^{(k)} = P_1 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \qquad \diamondsuit$$

4. Primer uporabe

Pogledali si bomo primer uporabe matričnih potenc na markovskih verigah. Prvi del poglavja bomo namenili aplikaciji teorije na markovskih verigah, nato pa si bomo pogledali še zgled markovske verige.

Najprej navedimo definicijo pojma *stohastičnost*.

Definicija 4.1. Kvadratna matrika $A = [a_{ij}]_{1\le i,j\le n}$ je *stohastična matrika*, če za njene člene velja $0 \le a_{ij} \le 1$ za vsaka $i, j = 1, \ldots, n$ in je $\sum_{j=1}^{n} a_{ij} = 1$ za vsak $i = 1, \ldots, n$ (tj. vsota vseh elementov v isti vrstici je enaka 1).

Podobno definiramo *stohastični vektor* $v = (v_1, \ldots, v_n)^T$, za katerega velja, da je $0 \le v_i \le 1$ za vsak $i = 1, \ldots, n$ in velja $\sum_{i=1}^{n} v_i = 1$ (tj. vsota komponent je enaka 1).

V posebnem primeru, ko so vse komponente matrike oz. vektorja strogo večje od 0, matriki oz. vektorju pravimo *pozitivna stohastična matrika* oz. *pozitivni stohastični vektor*.

Opomba 4.2.

(1) Ponekod v literaturi, ki opisuje markovske verige, vektorje pišejo v vrstici. Tedaj je stohastična matrika tista matrika, za katero velja, da je vsota elementov v stolpcih (in ne v vrsticah) enaka 1. V tem primeru vektorje z matriko preslikamo tako, da matriko množimo iz desne.

(2) Za stohastične matrike velja, da vektor $(1, 1, \dots, 1)^T$ preslikajo samega vase.

Naj bo $P = [p_{ij}]_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$ prehodna matrika (diskretne) markovske verige z množico stanj $S = \{s_1, \dots, s_n\}$. P je stohastična matrika. Stohastični vektor $p(k) = \big(p_1(k), p_2(k), \dots, p_n(k)\big)^T$, kjer je $p_i(k)$ verjetnost, da je markovski proces v stanju s_i po k-tem koraku, imenujemo *porazdelitev verjetnosti na k-tem koraku.* Iz lastnosti $p(k) = P^T p(k-1)$ sledi, da lahko vektor $p(k)$ na k-tem koraku dobimo s pomočjo začetne porazdelitve in prehodne matrike s formulo

$$p(k) = (P^k)^T p(0). \tag{12}$$

Vidimo, da je limitni porazdelitveni vektor $\lim_{k\to\infty} p(k)$ odvisen od limite zaporedja $(P^k)_{k\in\mathbb{N}}$. Z znanjem, ki ga imamo iz prejšnjih poglavij, lahko obnašanje konvergence predstavimo s spektralnimi lastnostmi matrike P. Poglejmo si nekaj spektralnih lastnosti stohastičnih matrik.

Trditev 4.3. *Za prehodno matriko P velja:*

(1) *$r(P) = 1$ je lastna vrednost matrike P s pripadajočim lastnim vektorjem*
$$\mathbb{1} = (1, 1, \dots, 1)^T.$$
(2) *Vse lastne vrednosti λ matrike P na kompleksni enotski krožnici imajo večkratnost* 1.

Dokaz.

(1) Ker je P stohastična, po definiciji sledi, da je $P^k \mathbb{1} = \mathbb{1}$ za vse $k \in \mathbb{N}$. Po Gelfandovi formuli (trditev 2.17) sledi, da je $r(P) = 1$ in da je 1 lastna vrednost s pripadajočim lastnim vektorjem $\mathbb{1}$.
(2) Ker je $\left\|P^k\right\|_\infty = 1$ za vse $k \in \mathbb{N}$, je zaporedje $(P^k)_{k\in\mathbb{N}}$ omejeno. Iz točke (4) izreka 3.13 sledi, da imajo vse lastne vrednosti na enotski krožnici večkratnost 1.

□

Iz izreka 3.13 in dejstva, da je zaporedje $(P^k)_{k\in\mathbb{N}}$ omejeno, sledi, da je zaporedje $(P^k)_{k\in\mathbb{N}}$ Cesàrovo konvergentno. Iz istega izreka sledi, da zaporedje Cesàrovih povprečij $\left(\frac{1}{k}\sum_{l=0}^{k-1} P^l\right)_{k\in\mathbb{N}}$ konvergira k spektralnemu projektorju P_1, ki pripada lastni vrednosti 1 matrike P. Vendar pa zaporedje $(P^k)_{k\in\mathbb{N}}$ ne konvergira, razen če je 1 radialno dominantna lastna vrednost matrike P (izrek 3.7, trditev (4)).

Poglejmo si, kako si lahko predstavljamo Cesàrova povprečja v markovskih verigah. Fiksiramo stanje s_j in definirajmo zaporedje slučajnih spremenljivk $(X_i)_{i\in\mathbb{N}_0}$ s členi

$$X_i = \begin{cases} 1, & \text{če je na } i\text{-tem koraku veriga v stanju } s_j, \\ 0, & \text{sicer.} \end{cases}$$

Potem izraz $\frac{1}{k}\sum_{i=0}^{k-1} X_i$ predstavlja delež časa, ko je veriga v stanju s_j, po $k-1$ korakih. Ker je pričakovana vrednost slučajne spremenljivke X_i enaka

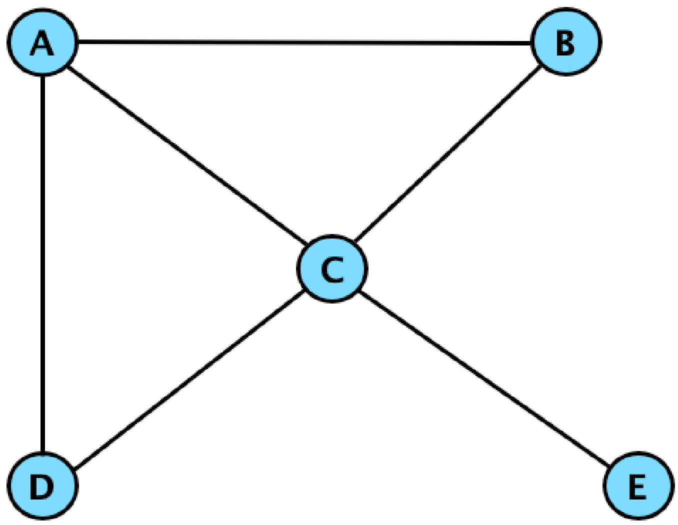
$$E(X_i) = \big(p(i)\big)_j,$$

iz linearnosti pričakovane vrednosti velja

$$E\left(\frac{1}{k}\sum_{i=0}^{k-1} X_i\right) = \frac{1}{k}\sum_{i=0}^{k-1} E(X_i) = \frac{1}{k}\sum_{i=0}^{k-1} \big(p(i)\big)_j.$$

To pomeni, da j-ta komponenta limite Cesàrovih povprečij predstavlja delež časa, ko je veriga v stanju s_j, po zelo dolgem času.

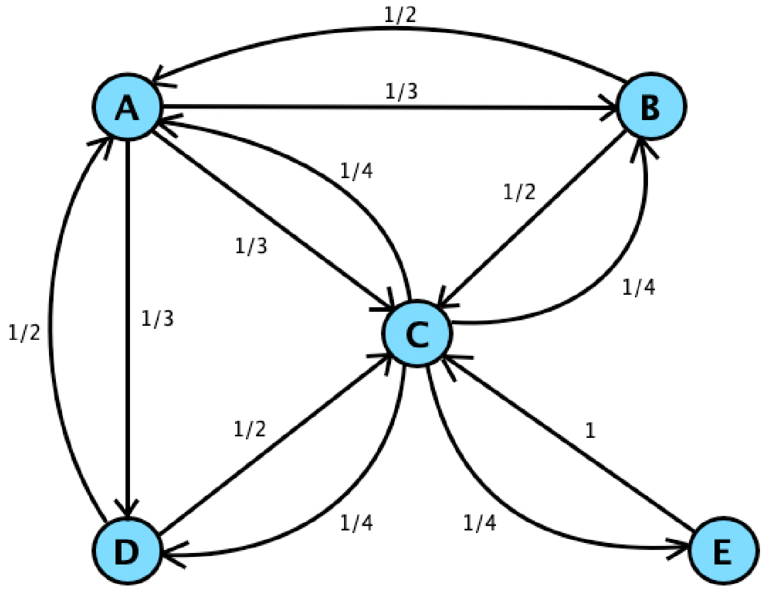
4.1. **Primer markovske verige.** Za konec predstavimo še realni problem markovske verige. Podjetje ima tovarne v petih mestih, naj bodo to mesta A, B, C, D, E. Nadzornik proizvodnje mora na vsake toliko časa priti v mesto A, B, C, D, E na nadzor proizvodnje. Katero bo naslednje mesto, nadzornik izbere naključno, ampak letalske poti ne povezujejo vseh mest med sabo. Kako so z dvosmernimi letalskimi potmi povezana mesta, prikazuje naslednji graf.



SLIKA 4. Dvosmerne letalske povezave med mesti A, B, C, D, E.

Če je nadzornik na začetku v mestu A, kakšne so verjetnosti, da se v k-tem koraku nahaja v mestu A, B, C, D, E? Kolikšen delež svojega časa nadzornik preživi v mestih A, B, C, D, E po dolgem času?

Seveda smo že tekom opisa problema zaslutili, da gre za problem iz podpoglavja 1.2, ki pa si ga bomo sedaj nekoliko bolj podrobno pogledali. Že v omenjenem podpoglavju smo neusmerjenemu grafu priredili usmerjenega, ki odraža verjetnosti izbire mesta naslednjega postanka.



SLIKA 5. Usmerjeni graf, ki prikazuje verjetnosti, da bo nadzornik izbral določeno mesto kot naslednji postanek.

Problem lahko interpetiramo kot markovsko verigo z množico stanj

$$S = \{A, B, C, D, E\}$$

in prehodno matriko

$$P = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1/3 & 1/3 & 1/3 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 0 & 1/4 & 1/4 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array},$$

ki je hkrati tudi matrika sosednosti usmerjenega grafa na sliki 5. Z nekoliko truda bi za to matriko izračunali lastne vrednosti

$$\begin{aligned} \lambda_1 &= 1, \\ \lambda_2 &= -\frac{1}{2}, \\ \lambda_3 &= 0, \\ \lambda_4 &= \frac{1}{12}\left(-3 + \sqrt{33}\right), \\ \lambda_5 &= \frac{1}{12}\left(-3 - \sqrt{33}\right). \end{aligned}$$

Vse lastne vrednosti razen λ_1 so po absolutni vrednosti strogo manjši od 1, torej po izreku 3.7 sledi, da je zaporedje $(P^k)_{k\in\mathbb{N}}$ konvergentno z limito P_1, torej spektralnim projektorjem za lastno vrednost λ_1.

Recimo, da je nadzornik na začetku v mestu A. Porazdelitev verjetnosti stanj na 0-tem koraku bo torej enaka $p(0) = (1, 0, 0, 0, 0)^T$. Po formuli (12) lahko izračunamo porazdelitev verjetnosti stanj na k-tem koraku

$$p(k) = \left(P^k\right)^T p(0)$$

za vsak $k \in \mathbb{N}$.

Zanima nas še, kakšna je porazdelitev verjetnosti stanj po zelo dolgem času, tj. $p(k)$, ko $k \to \infty$. Vemo, da je

$$\lim_{k\to\infty} P^k = P_1,$$

kjer je P_1 spektralni projektor za lastno vrednost $\lambda_1 = 1$. Ker ima matrika $P \in \mathbb{R}^{5\times 5}$ 5 različnih lastnih vrednosti, lahko P_1 izračunamo po formuli (6) in dobimo

$$P_1 = \frac{1}{2}\begin{bmatrix} 1/2 & 1/3 & 2/3 & 1/3 & 1/6 \\ 1/2 & 1/3 & 2/3 & 1/3 & 1/6 \\ 1/2 & 1/3 & 2/3 & 1/3 & 1/6 \\ 1/2 & 1/3 & 2/3 & 1/3 & 1/6 \\ 1/2 & 1/3 & 2/3 & 1/3 & 1/6 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/6 & 1/3 & 1/6 & 1/12 \\ 1/4 & 1/6 & 1/3 & 1/6 & 1/12 \\ 1/4 & 1/6 & 1/3 & 1/6 & 1/12 \\ 1/4 & 1/6 & 1/3 & 1/6 & 1/12 \\ 1/4 & 1/6 & 1/3 & 1/6 & 1/12 \end{bmatrix}.$$

Sedaj lahko izračunamo limito $p(k)$, ko gre $k \to \infty$:

$$\begin{aligned} p := \lim_{k\to\infty} p(k) = \lim_{k\to\infty} \left(P^k\right)^T p(0) = P_1^T\, p(0) = \\ = \left(\frac{1}{4}, \frac{1}{6}, \frac{1}{3}, \frac{1}{6}, \frac{1}{12}\right)^T. \end{aligned}$$

Prva komponenta vektorja p nam pove, kolikšen delež časa bo nadzornik preživel v mestu A. Druga komponenta nam pove, kolikšen delež časa bo nadzornik preživel v mestu B. Enako sklepamo za mesta C, D, E.

Slovar strokovnih izrazov

A-invarianten podprostor A-invariant subspace
Cesàrova konvergenca Cesàro summability
koordinatno zaporedje coordinate sequence
lastna vrednost eigenvalue
lastni vektor eigenvector
lastni podprostor eigenspace
markovska veriga Markov chain
minimalni polinom minimal polynomial
množica stanj state space
prehodna matrika transition matrix
radialno dominantna lastna vrednost radially dominant eigenvalue
spekter spectrum
spektralna skrčitev spectral contraction
spektralni projektor spectral projection
spektralni radij spectral radius
spektralni razcep spectral decomposition
stohastičnost stochasticity

Literatura

[1] A. Bátkai, M. Kramar Fijavž in A. Rhandi, *Positive operator semigroups from finite to infinite dimensions*, Birkhäuser/Springer, 2017.

[2] V. Guruswami in R. Kannan, *Computer science theory for the information age, Spring 2012*, verzija 2. 2012, [ogled 17. 3. 2021], dostopno na `https://www.cs.cmu.edu/~venkatg/teaching/CStheory-infoage/hopcroft-kannan-feb2012.pdf`.

[3] T. Košir, *Lastne vrednosti in lastni vektorji (8. poglavje iz predavanj)*, [ogled 17. 3. 2021], dostopno na `https://www.fmf.uni-lj.si/~kosir/poucevanje/skripta/lastne.pdf`.

[4] B. Lavrič, *Algebra* 1 *(kratek pregled rezultatov in definicij)*, [ogled 17. 3. 2021], dostopno na `https://www.fmf.uni-lj.si/~lavric/algebra%201%20-%20pregled.pdf`.

[5] M. Brešar, *Uvod v algebro*, DMFA – založništvo, Ljubljana, 2018, dostopno tudi na `https://www.fmf.uni-lj.si/~bresar/documents/UvodVAlgebro2018.pdf`

[6] C. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[7] HarvardX (28. 2. 2020), Introducing Markov Chains [Video], Youtube, [ogled 10. 10. 2020], dostopno na `https://www.youtube.com/watch?v=JHwyHIz6a8A`