

Spektralne metode za računanje meta-stabilnih stanja Markovljevih lanaca

Blažeka, Ivana

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:095482>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-08-07**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



Spektralne metode za računanje meta-stabilnih stanja Markovljevih lanaca

Blažeka, Ivana

Master's thesis / Diplomski rad

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:095482>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-06-20**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU
PRIRODOSLOVNO–MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Ivana Blažeka

**SPEKTRALNE METODE ZA
RAČUNANJE META-STABILNIH
STANJA MARKOVLJEVIH LANACA**

Diplomski rad

Voditelj rada:
prof. dr. sc. Zlatko Drmač

Zagreb, travanj, 2021.

Ovaj diplomski rad obranjen je dana _____ pred ispitnim povjerenstvom u sastavu:

1. _____, predsjednik
2. _____, član
3. _____, član

Povjerenstvo je rad ocijenilo ocjenom _____.

Potpisi članova povjerenstva:

1. _____
2. _____
3. _____

Ovaj rad posvećujem svojoj obitelji.

Zahvaljujem roditeljima na svojoj ljubavi i podršci koju su pružali tijekom cijelog studija i pomogli da postanem ovo što jesam unatoč raznim životnim preprekama.

Zahvaljujem bratu i kolegi Mateju na beskrajnoj vjeri u mene i savjetima bez kojih bi sve bilo puno drugačije.

Zahvaljujem sestri Mihaeli i bratu Damjanu na podršci i razumijevanju.

Zahvaljujem svim svojim dragim prijateljima.

Zahvaljujem mentoru, prof. dr. sc. Zlatku Drmaču, na pomoći u izradi ovog rada i susretljivosti.

Sadržaj

Sadržaj	iv
Uvod	2
1 Osnovne definicije i teoremi	3
1.1 Markovljevi lanci	3
1.2 Teorija stohastičkih matrica	8
1.3 Dekompozicije matrice	15
2 SVD pristup	23
2.1 Potrebni teorijski rezultati	24
2.2 Algoritam	30
3 GenPCCA	33
3.1 Klasteriranje prostora stanja	33
3.2 Slučaj reverzibilnih Markovljevih lanaca	34
3.3 Slučaj nereverzibilnih Markovljevih lanaca, GenPCCA	39
3.4 Transformacijska matrica A	41
3.5 Algoritam	47
4 Particioniranje grafa	51
4.1 Višeklasni kriteriji podjele	51
4.2 Relaksirano rješenje problema	53
4.3 Optimalno diskretno rješenje	55
4.4 Algoritam	59
5 Primjeri obrađenih metoda	61
5.1 Konstruirani primjer malih dimenzija	61
5.2 Primjer molekule n-pentana Ph300	64
5.3 Primjer molekule n-pentana Ph500	68

Bibliografija	73
A MATLAB kodovi	81
A.1 Analiza meta-stabilnih stanja reverzibilnih Markovljevih lanaca - SVD pristup	81
A.2 Analiza meta-stabilnih stanja nereverzibilnih Markovljevih lanaca - GenP-CCA	82
A.3 Analiza meta-stabilnih stanja reverzibilnih Markovljevih lanaca - particioniranje grafa	88

Uvod

Markovljevim lancima opisujemo pravilnosti prijelaza raznih sustava iz nekog stanja u drugo. Promatraju se sustavi iz raznih područja znanosti - od prirodnih i medicinskih, preko tehničkih i poljoprivrednih, do društvenih, umjetničkih i humanističkih. Svaki Markovljev lanac koji opisuje neki sustav ima određen skup stanja koja taj sustav može poprimiti. Što je skup stanja u kojem Markovljev lanac poprima svoje vrijednosti veći, to je teže dalje analizirati pravilnosti prijelaza iz jednog stanja u drugo. Problem je moguće riješiti grupiranjem stanja u klastere.

Postoji velika klasa prirodnih i umjetnih sustava koji su u stanju „nesigurne stabilnosti“, odnosno doimaju se stabilnima u jednom stanju kroz dulje vremensko razdoblje, a onda se pod određenim utjecajima prebace u novo naizgled stabilno stanje. Pošto će takvi sustavi sigurno napustiti određeno stanje nakon dovoljno dugog vremenskog perioda, ne možemo ih klasificirati pod „stabilne“. Također, nema ih smisla nazivati ni „nestabilnima“. Takvi sustavi su metastabilni, a klasteriranjem je moguće smanjiti broj mogućih stanja koje pripadni Markovljev lanac može poprimiti, a da on i dalje dobro opisuje stohastički proces.

Motivacija za nastanak metoda klasteriranja opisanih u ovom radu bila je proučavanje konformacija molekula. Konformacija je prostorni oblik molekule, odnosno stvarni prostorni razmještaj atoma u molekuli u nekom trenutku. Molekule su trodimenzionalni objekti, a zbog rotacija atoma oko jednostrukih veza dolaze u više mogućih konformacija. Svako konformaciji pripada određena vrijednost potencijalne energije zbog djelovanja privlačnih i odbojnih sila. Kada neka konformacija ima potencijalnu energiju nižu od energija svih bliskih joj konformacija, ona je stabilnija. Konformacije koje odgovaraju lokalnim minimumima na energetskej površini posebno se nazivaju konformeri. Konformacije koje odgovaraju lokalnim maksimumima na energetskej površini su prijelazna stanja između lokalnih minimalnih konformera. Rotacije oko jednostrukih veza uključuju prevladavanje rotacijske energetske barijere, kako bi se jedan konformer pretvorio u drugi. Mnogi konformeri neke molekule imaju sličnu strukturu i kemijska svojstva te ih nema smisla sve zasebno promatrati. Zato ih je poželjno klasterirati na način da vjerojatnost prelaska molekule iz jednog klastera konformacija u drugi bude jako mala.

Dinamika konformacija molekule može se uspješno opisati pomoću reverzibilnog Markovljevog lanca s konačnim skupom stanja. Neke od metoda klasteriranja u ovom radu su

poopćene i na nereverzibilne Markovljeve lance kako bi se mogle koristiti u svrhu klasiranja stanja nekih drugih sustava. Osnova svih danih metoda bit će analiza spektra prijelazne matrice danih Markovljevih lanaca, odnosno analiza Perronovog klastera.

U prvom poglavlju bit će dani osnovni pojmovi i rezultati koji su bitni za sve tri proučavane metode i koriste se u cijelom radu. Sljedeća tri poglavlja obrađuju svako po jednu metodu. Zatim slijedi primjena metoda na nekim jednostanijim primjerima i na primjeru molekule pentana (C_5H_{12}). U dodatku na kraju rada, nakon literature, sažetka i životopisa, bit će priloženi MATLAB kodovi koji realiziraju obrađene metode.

Poglavlje 1

Osnovne definicije i teoremi

1.1 Markovljevi lanci

U ovom radu bit će predstavljene metode klasteriranja konačnog skupa stanja Markovljevog lanca. Markovljev lanac ćemo označiti s X , njegov skup stanja pomoću S , a prijelaznu matricu s P . U nekim metodama koristit će se stacionarna distribucija koja se tradicionalno označava s π . Slijede neke definicije i rezultati koji se odnose općenito na Markovljeve lance, a koristit će se u ovom radu. Kao izvor je poslužio [15], gdje se mogu naći i dodatne informacije o Markovljevim lancima.

Započnimo s osnovnom definicijom Markovljevog lanca i pojmova koji se direktno vežu iz njega.

Definicija 1.1.1. *Neka je S prebrojiv skup. Slučajni proces $X = (X_n : n \geq 0)$ definiran na vjerojatnostnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ s vrijednostima u skupu S je Markovljev lanac ako vrijedi*

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i) \quad (1.1)$$

za svaki $n \geq 0$ i za sve $i_0, \dots, i_{n-1}, i, j \in S$ za koje su obje uvjetne vjerojatnosti dobro definirane. Svojstvo u relaciji naziva se Markovljevim svojstvom.

Definicija 1.1.2. *Matrica $P = (p_{ij} : i, j \in S)$ naziva se stohastičkom matricom po retcima ako je $p_{ij} \geq 0$ za sve $i, j \in S$, te*

$$\sum_{j \in S} p_{ij} = 1, \text{ za sve } i \in S.$$

Definicija 1.1.3. *Neka je $\eta = (\eta_i : i \in S)$ vjerojatnosna distribucija na S i neka je $P = (p_{ij} : i, j \in S)$ stohastička matrica. Markovljev lanac $X = (X_n : n \geq 0)$ je homogen s početnom distribucijom η i prijelaznom matricom P ako vrijedi*

1. $\mathbb{P}(X_0 = i) = \eta_i$ za sve $i \in S$
2. $\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = p_{ij}$, za svaki $n \geq 0$ i $i_0, \dots, i_{n-1}, i, j \in S$.

Napomena 1.1.4. Neka je P prijelazna matrica Markovljevog lanca X . Neka je n -ta potencija matrice P dana s $P^n = (p_{ij}^{(n)} : i, j \in S)$. Tada za svaki $n \geq 0$ vrijedi $\mathbb{P}_i(X_n = j) = p_{ij}^{(n)}$, odnosno ako Markovljev lanac kreće iz stanja i , tada je vjerojatnost da nakon n koraka bude u stanju j jednaka (i, j) -tom elementu n -te potencije prijelazne matrice P .

Slijede rezultati vezani za reverzibilnost i stacionarnost Markovljevog lanca.

Definicija 1.1.5. Slučajni proces $X = (X_n : n \geq 0)$ definiran na vjerojatnosnom prostoru $(\Omega, \mathcal{F}, \mathbb{P})$ zove se stacionaran ako za sve $k \geq 0$ i sve $n \geq 0$, slučajni vektori (X_0, X_1, \dots, X_k) i $(X_n, X_{n+1}, \dots, X_{n+k})$ imaju istu distribuciju (u odnosu na vjerojatnost \mathbb{P}).

Definicija 1.1.6. Neka je $X = (X_n : n \geq 0)$ Markovljev lanac s prebrojivim skupom stanja S i prijelaznom matricom P . Vjerojatnosna distribucija $\pi = (\pi_i : i \in S)$ na S je stacionarna distribucija (odnosno invarijantna distribucija) Markovljevog lanca X (odnosno prijelazne matrice P) ako vrijedi

$$\pi^T = \pi^T P, \quad (1.2)$$

odnosno po komponentama

$$\pi_j = \sum_{k \in S} \pi_k p_{kj}, \text{ za sve } j \in S. \quad (1.3)$$

Napomena 1.1.7. Ako je u definiciji 1.1.6 Markovljevom lancu X distribucija π inicijalna, tada je on stacionaran. Preciznije, X je stacionaran uz vjerojatnost

$$\mathbb{P}_\pi = \sum_{i \in S} \pi_i \mathbb{P}_i.$$

Definicija 1.1.8. Za stanja $i, j \in S$ kažemo da je j dostižno iz i , u oznaci $i \rightarrow j$, ako vrijedi $\mathbb{P}_i(T_j < \infty) > 0$, pri čemu je $T_j = \min\{n \geq 0 : X_n = j\}$ prvo vrijeme dolaska u stanje $j \in S$.

Definicija 1.1.9. Markovljev lanac je ireducibilan ako se prostor stanja S sastoji od samo jedne klase komuniciranja, odnosno za sve $i, j \in S$ vrijedi $i \rightarrow j$ i $j \rightarrow i$.

Primjer 1.1.10. Promotrimo sljedeće dvije prijelazne matrice:

$$P_1 = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.9 & 0.1 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0.7 & 0.3 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 0.8 & 0 & 0.2 \\ 0.7 & 0 & 0.1 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.1 & 0.9 & 0 \\ 0.9 & 0 & 0 & 0 & 0.1 & 0 \end{bmatrix}.$$

Markovljev lanac s prijelaznom matricom P_1 je reducibilan. Lanac poprima ili stanja $\{1, 2\}$, ili $\{3, 4\}$. Nemoguće je preći iz jednog podskupa stanja u drugi.

Za lanac s prijelaznom matricom P_2 se lako vidi da je ireducibilan. U kojem god stanju se lanac trenutno nalazio, još uvijek može doći do bilo kojeg drugog stanja.

Definicija 1.1.11. Stanje $i \in S$ je povratno ako vrijedi $\mathbb{P}_i(T_i < \infty) = 1$. Stanje $i \in S$ je pozitivno povratno ako je $\mathbb{E}_i(T_i) < \infty$. Pritom je $T_i = \min\{n > 0 : X_n = i\}$ prvo vrijeme povratka u stanje $i \in S$.

Teorem 1.1.12. Neka je X ireducibilan Markovljev lanac s prijelaznom matricom P . Sljedeće tvrdnje su ekvivalentne:

1. Svako stanje je pozitivno povratno.
2. Postoji pozitivno povratno stanje $i \in S$.
3. X ima stacionarnu distribuciju π .

Nadalje, ako vrijedi 3. tvrdnja, tada je $\pi_j = \frac{1}{\mathbb{E}_j(T_j)}$ pri čemu je $T_j = \min\{n > 0 : X_n = j\}$, za svaki $j \in S$.

Napomena 1.1.13. Iz teorema 1.1.12 slijedi da ireducibilan i povratan Markovljev lanac ima jedinstvenu stacionarnu distribuciju π , odnosno do na množenje pozitivnom konstantom jedinstvenu invarijantnu mjeru.

Primjer 1.1.14. Neka je $X = (X_n : n \geq 0)$ Markovljev lanac na dvočlanom skupu stanja $S = \{1, 2\}$, s prijelaznom matricom

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

Lanac je ireducibilan i povratan. Stoga postoji jedinstvena stacionarna distribucija $\pi = (\frac{1}{2}, \frac{1}{2})$.

Napomena 1.1.15. Potpunu simetriju u vremenu, odnosno simetrične prošlost i budućnost, moguće je postići ukoliko je početna distribucija lanca stacionarna.

Dalje definiran uvjet detaljne ravnoteže će nam biti bitan u slučaju reverzibilnih Markovljevih lanaca u poglavljima 3 i 4.

Neka je $X = (X_n : n \geq 0)$ Markovljev lanac s prostorom stanja S , prijelaznom matricom P i stacionarnom distribucijom π takvom da je $\pi_i > 0$, za svaki $i \in S$. Definiramo matricu $\hat{P} = (\hat{p}_{ij}, i, j \in S)$ s

$$\pi_i \hat{p}_{ij} = \pi_j p_{ji}.$$

Lako se vidi da je \hat{P} stohastička matrica i da je π stacionarna distribucija i za \hat{P} :

$$\sum_{j \in S} \hat{p}_{ij} = \frac{1}{\pi_i} \sum_{j \in S} \pi_j p_{ji} = 1,$$

$$\sum_{j \in S} \pi_j \hat{p}_{ji} = \sum_{j \in S} \pi_i p_{ij} = \pi_i \sum_{j \in S} p_{ij} = \pi_i.$$

Neka je sada $X = (X_n : n \geq 0)$ (π, P) -ireducibilan Markovljev lanac, π stacionarna distribucija, i neka je za $N \in \mathbb{N}$ i $0 \leq n \leq N$ definirano $Y_n = X_{N-n}$. Tada iz teorema 10.1 u [15] vidimo da je $Y = (Y_n : 0 \leq n \leq N)$ (π, \hat{P}) -ireducibilan Markovljev lanac kojem je π stacionarna distribucija.

Definicija 1.1.16. Za stohastičku matricu P i mjeru λ kažemo da su u detaljnoj ravnoteži ako vrijedi

$$\lambda_i p_{ij} = \lambda_j p_{ji}, \text{ za sve } i, j \in S.$$

Lema 1.1.17. Ako su P i λ u detaljnoj ravnoteži, tada je λ invarijantna mjera za P .

Dokaz.

$$\sum_{i \in S} \lambda_i p_{ij} = \sum_{i \in S} \lambda_j p_{ji} = \lambda_j$$

zbog definicije detaljne ravnoteže i činjenice da je P stohastička matrica. \square

Definicija 1.1.18. Neka je λ distribucija na S . Za ireducibilan (λ, P) -Markovljev lanac $X = (X_n : n \geq 0)$ kažemo da je reverzibilan, ako je za sve $N \geq 1$, $(X_{N-n} : 0 \leq n \leq N)$ ponovno (λ, P) -Markovljev lanac.

Propozicija 1.1.19. Neka je X ireducibilan (λ, P) -Markovljev lanac. Ekvivalentno je:

1. X je reverzibilan.
2. P i λ su u detaljnoj ravnoteži.

Dokaz.

Prvo neka je X reverzibilan. Tada je za $N = 1$ proces $Y = (Y_0, Y_1) = (X_1, X_0)$ (λ, P) -Markovljev. To znači da je za $i, j \in S$

$$\mathbb{P}(X_1 = j, X_0 = i) = \mathbb{P}(Y_0 = j, Y_1 = i) = \lambda_j p_{ji}.$$

S druge strane, je X (λ, P) -Markovljev i imamo $\mathbb{P}(X_0 = i, X_1 = j) = \lambda_i p_{ij}$. Dakle λ i P su u detaljnoj ravnoteži.

Obratno, neka su P i λ u detaljnoj ravnoteži. Tada je

$$\hat{p}_{ij} = \frac{p_{ji}\lambda_j}{\lambda_i} = p_{ij},$$

odnosno $\hat{P} = P$. Iz leme 1.1.17, λ je stacionarna distribucija, a za $N \geq 1$ je $(X_{N-n} : 0 \leq n \leq N)$ (λ, P) -Markovljev lanac. \square

U sljedećem primjeru ćemo definirati ireducibilan reverzibilan Markovljev lanac koji predstavlja slučajnu šetnju na grafu. U poglavlju 4 bit će dana metoda klasteriranja skupa stanja takvog Markovljevog lanca.

Primjer 1.1.20. *Primjer reverzibilnog Markovljevog lanca je slučajna šetnja na grafu.*

Graf je uređen par $G = (V, E)$, gdje je V skup vrhova, a E skup bridova grafa. Vrijedi da je $E \subseteq V \times V$ pri čemu $(i, j) \in E \iff (j, i) \in E$ jer brid koji povezuje vrhove i, j , povezuje također j, i . Graf je povezan ako za svaka dva vrha i, j postoji konačan niz bridova $(i, i_1), (i_1, i_2), \dots, (i_n, j)$ koji ih povezuje. Neka je graf konačan.

Neka je W simetrična matrica s težinama bridova, odnosno svakom bridu $e \in E$, $e = (i, j) = (j, i)$ je pridružena njegova težina $w_{ij} = w_{ji}$. Ukoliko $e = (i, j) = (j, i) \notin E$, tada je $w_{ij} = w_{ji} = 0$. Stupanj vrha $i \in V$ definiramo kao $w_i = \sum_{j \in V} w_{ij}$.

Slušajna šetnja po grafu G je Markovljev lanac $X = (X_n : n \geq 0)$ sa skupom stanja V i prijelaznom matricom $P = (p_{ij} : i, j \in V)$ gdje je

$$p_{ij} = \frac{w_{ij}}{w_i}.$$

Iz pretpostavke o povezanosti grafa, slučajna šetnja X je ireducibilan Markovljev lanac. Neka je $w = \sum_{i \in V} w_i$. Definirajmo $\pi = (\pi_i : i \in V)$ s

$$\pi_i = \frac{w_i}{w}.$$

Vrijedi

$$\pi_i p_{ij} = \frac{w_i w_{ij}}{w w_i} = \frac{w_{ij}}{w} = \frac{w_{ji}}{w} = \frac{w_j w_{ji}}{w w_j} = \pi_j p_{ji},$$

odnosno P i π su u detaljnoj ravnoteži. Slijedi da je π stacionarna distribucija od X . X je reverzibilan.

Definicija 1.1.21. *Neka je $X = (X_n : n \geq 0)$ Markovljev lanac s prostorom stanja S i prijelaznom matricom P . S π_t označimo distribuciju od X_t , odnosno Markovljevog lanca u trenutku t :*

$$(\pi_t)_i = P(X_t = i), \quad \forall i \in S.$$

Dalje, vrijedi

$$\pi_{t+1}^T = \pi_t^T P. \quad (1.4)$$

Napomena 1.1.22. Iz jednakosti (1.4), uz činjenicu $P^{n+m} = P^n P^m$, imamo da propagacija distribucije za n koraka glasi

$$\pi_{t+n}^T = \pi_t^T P^n. \quad (1.5)$$

1.2 Teorija stohastičkih matrica

Meta-stabilna stanja Markovljevog lanca bit će analizirana na temelju pripadne prijelazne matrice. Prema definiciji, prijelazna matrica je stohastička po retcima. U metodama će biti primijenjeni razni rezultati vezani uz stohastičke matrice, od kojih je najznačajniji Perron-Frobeniusov teorem. Korišteni su izvori [8] i [12].

Započinjemo s osnovnim definicijama.

Definicija 1.2.1. Vektor $v \in \mathbb{R}^n$ nazivamo pozitivnim (nenegativnim) i pišemo $v > 0$ ($v \geq 0$) ako je svaki njegov element pozitivan (nenegativan). Slično, matricu $T \in \mathbb{R}^{n \times n}$ nazivamo pozitivnom (nenegativnom) i pišemo $T > 0$ ($T \geq 0$) ako je svaki njen element pozitivan (nenegativan).

Definicija 1.2.2. Matricu T nazivamo reducibilnom ako postoji permutacijska matrica $P \in \mathbb{R}^{n \times n}$ tako da

$$PTP^T = \begin{bmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{bmatrix}, \quad (1.6)$$

gdje su T_{11} i T_{22} kvadratne podmatrice. U suprotnom, T nazivamo ireducibilnom matricom.

Definicija 1.2.3. Nenegativnu matricu $T \in \mathbb{R}^{n \times n}$ nazivamo ireducibilnom ako za svaki $i, j \in \{1, \dots, n\}$ postoji $k \geq 1$ takav da je $T_{ij}^k > 0$.

Definicija 1.2.4. Matricu $T \in \mathbb{R}^{n \times n}$ nazivamo dijagonalno dominantnom ako za sve $i = 1, \dots, n$ vrijedi

$$|t_{ii}| > \sum_{j=1, j \neq i}^n |t_{ij}|.$$

Napomena 1.2.5. Vektor jedinica $(1, \dots, 1)$ označavamo s $\mathbf{1}$.

Definicija 1.2.6. Skalar $\lambda \in \mathbb{R}$ nazivamo svojstvena vrijednost matrice $T \in \mathbb{R}^{n \times n}$ ako postoji vektor $v \in \mathbb{R}^n, v \neq 0$, takav da $Tv = \lambda v$. Vektor v naziva se desnim svojstvenim vektorom. Analogno, vektor $w \in \mathbb{R}^n, w \neq 0$ za koji vrijedi $w^T T = \lambda w^T$ zovemo lijevim svojstvenim vektorom.

Definicija 1.2.7. Neka su $\lambda_i, i = 1, \dots, n$ svojstvene vrijednosti matrice $T \in \mathbb{R}^{n \times n}$. Vrijednost $\varphi(T) = \max_{1 \leq i \leq n} |\lambda_i|$ nazivamo spektralnim radijusom od T .

Definicija 1.2.8. Neka je $T \in \mathbb{R}^{n \times n}$ stohastička matrica. Njenu svojstvenu vrijednost $\lambda_1 = 1$ nazivamo Perronovim korijenom. Pripadni desni svojstveni vektor je vektor jedinica $\mathbf{1}$, $T\mathbf{1} = \mathbf{1}$. Ukoliko postoji pripadni lijevi nenegativni svojstveni vektor čija suma elemenata je 1, $\pi T = \pi$, tada je to stacionarna distribucija Markovljevog lanca čija je T prijelazna matrica.

Definicija 1.2.9. Matricu T nazivamo blok-stohastičkom ako je T blok-dijagonalna, $T = \text{diag}(T_1, \dots, T_m)$, a matrice $T_i \in \mathbb{R}^{n_i \times n_i}$, $i = 1, \dots, m$, su stohastičke.

Napomena 1.2.10. Neka je za svaki blok T_i matrice T definiran skup S_i od n_i indeksa koji odgovaraju bloku T_i . Vrijedi

$$\bigcup_{i=1}^m S_i = \{1, \dots, n\}, \quad S_i \cap S_j = \emptyset \text{ za } i \neq j.$$

Definiramo blok $T(S_i, S_j)$ matrice T koji sadrži vrijednosti T_{ij} , gdje $i \in S_i, j \in S_j$.

Slijede rezultati vezani uz stohastičke matrice koji će biti korišteni.

Teorem 1.2.11 (Perron-Frobenius). Neka je $T \in \mathbb{R}^{n \times n}$ ireducibilna i nenegativna matrica. Tada vrijedi:

1. T ima pozitivnu svojstvenu vrijednost jednaku spektralnom radijusu $\varphi(T)$.
2. Svojstvenoj vrijednosti $\varphi(T)$ pripada pozitivan desni svojstveni vektor x .
3. $\varphi(T)$ je svojstvena vrijednost matrice T algebarske i geometrijske kratnosti 1.

Svojstvenu vrijednost $\varphi(T)$ nazivamo Perronovim korijenom, a pripadni svojstveni vektor Perronovim svojstvenim vektorom.

Dokaz.

1. (*Dokaz tvrdnji 1 i 2*). Definirajmo skup $X = \{x \in \mathbb{R}^n : x \geq 0, x \neq 0\}$. Definirajmo funkciju

$$r(x) = \min \left\{ \frac{(Tx)_i}{x_i} : x_i \neq 0 \right\}, \quad x \in X.$$

Jer je T linearni operator, r je neprekidna funkcija za svaki $x > 0 \in X$, osim eventualno na rubu od X (pogledati primjer 1.2.15).

Promotrimo što je zapravo funkcija r . Neka je $x \in X$ proizvoljan. Usporedimo vektore Tx i x . Tamo gdje x ima pozitivne elemente, ima ih i Tx . To znači da za svaki pozitivan element x_i postoji pozitivan faktor skaliranja α_i takav da $(Tx)_i = \alpha_i \cdot x_i$. Slijedi da je $r(x)$ najmanji takav α_i , pozitivan realni broj. Iz toga je očito da je $r(x)x \leq Tx$, štoviše $r(x) = \max \{\varphi : \varphi x \leq Tx\}$.

- a) Želimo naći supremum funkcije r na skupu X . Dokazat ćemo egzistenciju vektora z takvog da vrijedi $r(z) = \sup \{r(x) : x \in X\}$.

Weierstrassov teorem nam kaže da neprekidna funkcija na nepraznom kompaktnom skupu u \mathbb{R} postiže svoj maksimum (i minimum). Bolzano-Weierstrassov teorem nam kaže da je skup u \mathbb{R}^n kompaktan akko je ograničen i zatvoren. Skup X nije niti ograničen, niti zatvoren.

Vrijedi $r(x) = r(\alpha x)$ za svaki $\alpha > 0$. Stoga vrijedi $\sup \{r(x) \mid x \in X\} = \sup \{r(x) \mid x \in X^{(1)}\}$, pri čemu je

$$X^{(1)} = \{x \in X \mid \|x\| = 1\}.$$

Skup $X^{(1)}$ je ograničen i zatvoren, odnosno kompaktan. Ali $X^{(1)}$ sadrži i vektore koji nisu pozitivni, te u kojima r ima prekid.

Stoga definirajmo skup

$$Y = \{y \mid y = (T + I)^{n-1}x, x \in X^{(1)}\}.$$

Skup Y je kompaktan kao slika linearnog operatora nad kompaktnim skupom. Prema lemi 1.2.12, $y > 0, \forall y \in Y$ i r je neprekidna na Y . Primjenom Weierstrassovog teorema, postoji vektor $z \in Y$ takav da $r(z) = \max \{r(y) \mid y \in Y\}$.

Neka je $x \in X$ proizvoljan i neka je $y = (T + I)^{n-1}x$. Neka je φ takav da vrijedi $\varphi x \leq Tx$. S obje strane nejednakosti primjenimo pozitivnu matricu $(T + I)^{n-1}$ koja komutira s matricom T . Dobivamo $\varphi y \leq Ty$. Iz navedenog slijedi $r(x) \leq r(y)$.

Imamo:

$$\begin{aligned} \sup \{r(x) \mid x \in X\} &= \sup \{r(x) \mid x \in X^{(1)}\} \\ &\leq \sup \{r(y) \mid y \in Y\} = \max \{r(y) \mid y \in Y\} = r(z). \end{aligned}$$

- b) Neka je $\hat{\varphi} = r(z)$. Neka je $u \in X$ proizvoljan vektor koji zadovoljava $r(u) = \hat{\varphi}$.
- i. $z > 0 \Rightarrow Tz > 0 \Rightarrow r(z) > 0 \Rightarrow \hat{\varphi} > 0$.
 - ii. Neka je $y_u = (T + I)^{n-1}u$. Imamo od prije $\hat{\varphi}u \leq Tu$ i želimo pokazati da vrijedi jednakost. Prestpostavimo suprotno. Tada prema lemi 1.2.12:

$$\begin{aligned} (T + I)^{n-1}(Tu - \hat{\varphi}u) > 0 &\Rightarrow \hat{\varphi}y_u < Ty_u \Rightarrow \exists \epsilon > 0 \quad \text{t.d.} \\ (\hat{\varphi} + \epsilon)y_u < Ty_u &\Rightarrow \hat{\varphi} + \epsilon < r(y_u) \Rightarrow \text{kontradikcija.} \end{aligned}$$

Dakle vrijedi $\hat{\varphi}u = Tu$, odnosno $\hat{\varphi}$ je **svojstvena vrijednost matrice \mathbf{T}** , a vektor u je pripadni **svojstveni vektor**.

- iii. $u > 0$ jer je $y_u = (T + I)^{n-1}u = (\hat{\varphi} + 1)^{n-1}u$ i $y_u > 0$ po lemi 1.2.12.

- iv. Neka je α (kompleksna) svojstvena vrijednost matrice T s pripadnim svojstvenim vektorom v . Iz jednakosti $Tv = \alpha v$ slijedi $(Tv)^+ = (\alpha v)^+$. Tada za svaki $j = 1, \dots, n$ vrijedi

$$\begin{aligned} ((\alpha v)^+)_j &= |\alpha v_j| = |\alpha| |v_j| = (|\alpha| v^+)_j, \\ ((Tv)^+)_j &= |T_{j1}v_1 + \dots + T_{jn}v_n| \leq T_{j1}|v_1| + \dots + T_{jn}|v_n| = (Tv^+)_j. \end{aligned} \quad (1.7)$$

Slijedi $|\alpha| v^+ \leq Tv^+$. Jer je $|\alpha| \leq r(v^+) \leq \hat{\varphi}$. Slijedi da je $\hat{\varphi}$ **spektralni radijus matrice T** .

- v. Primjenimo argument u točki 1(b)iv za $\alpha = \hat{\varphi}$. Znamo da vrijedi $\hat{\varphi} v^+ \leq Tv^+$ i $r(v^+) = \hat{\varphi}$. To znači da je, prema dosadašnjim zaključcima, v^+ pozitivan svojstveni vektor matrice T koji pripada svojstvenoj vrijednosti $\hat{\varphi}$, odnosno spektralnom radijusu matrice T . Vrijedi $Tv^+ = \hat{\varphi} v^+$. To znači da u nejednakosti u (1.7) u točki 1(b)iv vrijedi jednakost. Slijedi da kompleksni brojevi v_1, \dots, v_n imaju isti argument ϕ , odnosno $v = e^{i\phi} v^+$. Zaključujemo da su svi svojstveni vektori koji odgovaraju spektralnom radijusu $\hat{\varphi}$ kolinearni s pozitivnim svojstvenim vektorom. Dodatno, $\hat{\varphi}$ ne može imati dva linearno nezavisna pozitivna svojstvena (ta dva pozitivna svojstvena vektora bi mogla dati novi svojstveni vektor kao nenegativnu linearnu kombinaciju što je nemoguće prema točki 1(b)iii). Slijedi da **spektralni radijus ima geometrijsku kratnost 1**.

2. Iz činjenice da je geometrijska kratnost jednaka 1 slijedi da u Jordanovoj formi J matrice T ima jedinstvenu ćeliju koja sadrži vrijednost $\varphi(T)$. Za kraj je potrebno dokazati da je dimenzija te ćelije 1×1 iz čega slijedi da $\varphi(T)$ ima algebarsku kratnost jednaku 1.

Promotrimo matricu T^T . Matrica T^T je također nenegativna i ireducibilna, te dijeli karakteristični polinom s matricom T . Također postoji pozitivan vektor y takav da $T^T y = \varphi(T)y$.

Neka je T matrica linearnog operatora \mathcal{T} u nekoj bazi. Ortogonalni komplement $y^\perp = \{x \mid y^T x = 0\}$ od y je invarijantan na \mathcal{T} :

$$y^T x = 0 \Rightarrow y^T (Tx) = \varphi(T)y^T x = 0.$$

Sa z označimo svojstveni vektor matrice T koji pripada svojstvenoj vrijednosti $\varphi(T)$. Uočimo da $z \notin y^\perp$ jer su y i z pozitivni.

Slijedi da je \mathbb{R}^n suma dva invarijantna potprostora: y^\perp i svojstvenog potprostora $\langle z \rangle$. Neka je $y^\perp = \langle y^{(2)}, \dots, y^{(n)} \rangle$. Matrica operatora \mathcal{T} zapisana u bazi $(z, y^{(2)}, \dots, y^{(n)})$ ima

blok-dijagonalnu formu $T' = \begin{bmatrix} \varphi(T) & 0 \\ 0 & Y \end{bmatrix}$, pri čemu je Y neka matrica dimenzije

$n - 1$. Dovođenjem Y u Jordanovu formu, matrica T' također postiže Jordanovu formu. T i T' imaju istu Jordanovu formu jer su ekvivalentne. Kao prvi blok u gornjem lijevom kutu ostaje 1×1 blok s vrijednosti φ . Slijedi: **algebarska kratnost svojstvene vrijednosti $\varphi(T)$ je 1.**

□

Lema 1.2.12. *Neka je $T \in \mathbb{R}^{n \times n}$, $T \geq 0$, ireducibilna matrica. Tada je $(T + I)^{n-1} > 0$.*

Dokaz. Matrica B je pozitivna akko je vektor $Bx > 0$ za svaki $x \geq 0$, $x \neq 0$. Usporedimo broj nenegativnih koordinata vektora $x \geq 0$, $x \neq 0$ i vektora $y = (T + I)x$: ako je $x_j > 0$, tada je $(Tx)_j + x_j > 0$. Slijedi da je skup koordinata vektora y jednakih 0 podskup skupa koordinata vektora x jednakih 0.

Pretpostavimo da su ta dva skupa jednaka. Permutirajmo vektore x, y pomoću permutacije π , tako da $x = [u \ 0]^T$ i $y = [v \ 0]^T$, $u, v > 0$ jednake veličine. Permutirana matrica $\pi(T)$ je reprezentirana matricom $\begin{bmatrix} B & C \\ D & F \end{bmatrix}$, pri čemu su B, F kvadratne matrice, B jednake veličine kao u .

Iz jednakosti

$$\begin{bmatrix} B & C \\ D & F \end{bmatrix} \begin{bmatrix} u \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ 0 \end{bmatrix}$$

implicira $D = 0$ jer $u > 0$. Slijedi da je matrica T reducibilna što je kontradikcija s pretpostavkom leme. Stoga, vektor y ima strogo manje koordinata jednakih 0 nego vektor x .

Uzmimo u obzir da je broj nul-koordinata vektora x najviše $n - 1$ i primijenimo prethodni zaključak na vektore $x, (T + I)x, \dots, (T + I)^{n-2}x$. Slijedi $(T + I)^{n-1} > 0$. □

Napomena 1.2.13. *Neka je $v \in \mathbb{R}^n$ proizvoljan vektor. Tada je $v^+ = [|v_1|, \dots, |v_n|]$.*

Napomena 1.2.14. *Jordanova matrica ima formu*

$$\begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_s \end{bmatrix}$$

pri čemu svaki blok J_k ima formu

$$\begin{bmatrix} \lambda_k & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_k \end{bmatrix}$$

i naziva se Jordanovom ćelijom. Jordanova ćelija J_k dimenzije 1×1 sadrži samo vrijednost λ_k .

Jordanov teorem kaže da svaka matrica $T \in \mathbb{C}^{n \times n}$ ima, do na raspored blokova, jedinstvenu ekvivalentnu Jordanovu matricu. Ta Jordanova matrica se zove Jordanova normalna forma od T .

Pošto ekvivalentne matrice imaju isti karakteristični polinom, brojevi λ_k u Jordanovoj normalnoj formi od T su svojstvene vrijednosti od T . Svaka se ponavlja onoliko puta kolika joj je algebarska kratnost.

Ako je j -ti stupac matrice J prvi stupac neke ćelije J_k , tada je on svojstveni vektor koji odgovara λ_k .

Primjer 1.2.15. *Slijedi primjer točke prekida funkcije r iz skupa $X^{(1)}$ u dokazu teorema 1.2.11 u dijelu 1a). Neka je $T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $x^{(k)} = \begin{bmatrix} 1/k \\ 1 \end{bmatrix}$, $x = \lim_{k \rightarrow \infty} x^{(k)}$. Tada vrijedi $r(x^{(k)}) = k$ i $r(x) = 0$, odnosno $\lim_{k \rightarrow \infty} r(x^{(k)}) \neq r(x)$, pa funkcija r nije neprekidna u točki x .*

Nakon Perron-Frobeniusovog teorema slijede još neke činjenice koje će se koristiti pri analizi prijelazne matrice u metodama koje ovaj rad obrađuje:

Teorem 1.2.16. *Neka je $T \in \mathbb{R}^{n \times n}$ reverzibilna i ireducibilna stohastička matrica sa stacionarnim vektorom π . Tada imamo skup desnih svojstvenih vektora r_i , $i = 1, \dots, n$ matrice T koji odgovaraju svojstvenim vrijednostima $\lambda_1, \dots, \lambda_n$ i čine ortonormalnu bazu u \mathbb{R}^n s obzirom na težinski skalarni produkt:*

$$\langle v, w \rangle_\pi = \sum_{i=1}^n v_i w_i \pi_i, \quad (1.8)$$

odnosno vrijedi $\langle r_i, r_j \rangle_\pi = \delta_{ij}$.

Desni svojstveni vektori matrice T mogu biti preslikani u lijeve svojstvene vektore na sljedeći način:

$$l_i = D_\pi r_i, \forall i = 1, \dots, n, \quad (1.9)$$

pri čemu je D_π dijagonalna matrica sa stacionarnim vektorom π na dijagonali.

Dokaz. Definirajmo matricu $S = D_\pi^{-\frac{1}{2}} T D_\pi^{\frac{1}{2}}$. Tada je $S_{ij} = \pi_i^{\frac{1}{2}} \pi_j^{-\frac{1}{2}} T_{ij}$. Ova matrica je simetrična zbog uvjeta detaljne ravnoteže (definicija 1.1.16) i slična matrici T pa imaju iste svojstvene vrijednosti. Zbog simetrije (propozicija 1.2.18), možemo naći n ortonormalnih svojstvenih vektora v_1, \dots, v_n matrice S koji pripadaju svojstvenim vrijednostima $\lambda_1, \dots, \lambda_n$. Slijedi da je $r_i = D_\pi^{-\frac{1}{2}} v_i$ svojstveni vektor od T s istom svojstvenom vrijednošću:

$$T r_i = T D_\pi^{-\frac{1}{2}} v_i = D_\pi^{-\frac{1}{2}} (D_\pi^{\frac{1}{2}} T D_\pi^{-\frac{1}{2}}) v_i = \lambda_i D_\pi^{-\frac{1}{2}} v_i = \lambda_i r_i.$$

Slijedi:

$$\langle r_i, r_j \rangle_\pi = \langle D_\pi^{-\frac{1}{2}} v_i, D_\pi^{-\frac{1}{2}} v_j \rangle_\pi = \langle v_i, D_\pi^{-1} v_j \rangle_\pi = \sum_{k=1}^n (v_i)_k (v_j)_k \frac{\pi_k}{\pi_k} = \delta_{ij}.$$

Dalje, neka je $l_i = D_\pi r_i$ za proizvoljan $i = 1, \dots, n$:

$$\begin{aligned} l_i^T T &= (D_\pi r_i)^T T = r_i^T D_\pi T = r_i^T D_\pi^{\frac{1}{2}} D_\pi^{\frac{1}{2}} T D_\pi^{-\frac{1}{2}} D_\pi^{\frac{1}{2}} = r_i^T D_\pi^{\frac{1}{2}} S D_\pi^{\frac{1}{2}} = v_i^T S^T D_\pi^{\frac{1}{2}} \\ &= (S v_i)^T D_\pi^{\frac{1}{2}} = \lambda_i v_i^T D_\pi^{\frac{1}{2}} = \lambda_i r_i^T D_\pi^{\frac{1}{2}} D_\pi^{\frac{1}{2}} = \lambda_i (D_\pi r_i)^T = \lambda_i l_i^T. \end{aligned}$$

□

Propozicija 1.2.17. *Svojtvene vrijednosti realne simetrične matrice su realne.*

Dokaz. Neka je x svojstveni vektor simetrične matrice A pridružen svojstvenoj vrijednosti λ . Tada je:

$$\lambda(x|x) = (\lambda x|x) = (Ax|x) = (x|Ax) = (x|\lambda x) = \bar{\lambda}(x|x)$$

pa je

$$(\lambda - \bar{\lambda})(x|x) = 0.$$

Jer je $x \neq 0$, vrijedi $(x|x) > 0$ pa slijedi $\lambda = \bar{\lambda}$. Znači da je $\lambda \in \mathbb{R}$.

□

Propozicija 1.2.18. *Svojstveni vektori pridruženi različitim svojstvenim vrijednostima realne simetrične matrice su ortogonalni.*

Dokaz. Neka su x, y svojstveni vektori matrice A pridruženi svojstvenim vrijednostima λ, μ , $\lambda \neq \mu$. Vrijedi $Ax = \lambda x$, $Ay = \mu y$, $x, y \neq 0$. Tada je

$$\lambda(x|y) = (\lambda x|y) = (Ax|y) = (x|Ay) = (x|\mu y) = \mu(x|y)$$

iz čega slijedi

$$(\lambda - \mu)(x|y) = 0.$$

Kako je $\lambda \neq \mu$, slijedi $(x|y) = 0$ pa su x i y ortogonalni.

□

Napomena 1.2.19. *Formula za skalarni umnožak vektora u \mathbb{C}^n je $(x|y) = \sum_{i=1}^n x_i \bar{y}_i$.*

1.3 Dekompozicije matrice

Temelj za analizu meta-stabilnih stanja Markovljevih lanaca je spektar pripadne prijelazne matrice. U tu svrhu, u poglavlju 2 ovog rada će se koristiti dekompozicija na singularne vrijednosti, a u poglavlju 3. Schurova dekompozicija. Ovdje su dane neke činjenice i osnovni teoremi vezani uz te dekompozicije, a izvori su [9] i [11].

Dekompozicija matrice na singularne vrijednosti

Definicija 1.3.1. *Neka je $A \in \mathbb{C}^{m \times n}$ za $m, n \in \mathbb{N}$. Rastav matrice*

$$A = U\Sigma V^*$$

nazivamo dekompozicija na singularne vrijednosti matrice A , ako su $U \in \mathbb{C}^{m \times m}$ i $V \in \mathbb{C}^{n \times n}$ unitarne, a $\Sigma \in \mathbb{C}^{m \times n}$ dijagonalna

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_{\min(m,n)}),$$

pri čemu vrijedi $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$ i te brojeve nazivamo singularnim vrijednostima matrice A . Stupce matrice U nazivamo lijevi, a stupce matrice V nazivamo desni singularni vektori matrice A .

Napomena 1.3.2. *Dekompozicija na singularne vrijednosti postoji za svaku matricu.*

Teorem 1.3.3 (SVD). *Ako je $A \in \mathbb{C}^{m \times n}$, tada postoje unitarne matrice $U \in \mathbb{C}^{m \times m}$ i $V \in \mathbb{C}^{n \times n}$ takve da je*

$$A = U\Sigma V^*, \quad \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_{\min(m,n)}), \quad (1.10)$$

pri čemu vrijedi $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)} \geq 0$.

Dokaz. Jedinična sfera u \mathbb{C}^n je ograničen i zatvoren skup, što znači da je kompaktan pa svaka neprekidna funkcija na njemu dostiže minimum i maksimum. Funkcija $f(x) = \|Ax\|_2$ je neprekidna pa postoji jedinični vektor $v \in \mathbb{C}^n$ takav da je

$$\|Av\|_2 = \max\{\|Ax\|_2 : \|x\|_2 = 1, x \in \mathbb{X}^m\}. \quad (1.11)$$

Ako je $\|Av\|_2 = 0$, onda je $A = 0$ i faktorizacija u iskazu teorema je trivijalna uz $\Sigma = 0$ i s proizvoljnim unitarnim matricama U i V reda m , odnosno n .

Ako je $\|Av\|_2 \geq 0$, stavimo $\sigma_1 = \|Av\|_2$ i formiramo jedinični vektor

$$u_1 = \frac{Av}{\sigma_1} \in \mathbb{C}^m.$$

Nadopunom u_1 s $m - 1$ vektora do baze u \mathbb{C}^m i onda primjenom Gram-Schmidtovog postupka ortogonalizacije, dobijemo ortonormiranu bazu u_1, u_2, \dots, u_m . Time smo dobili unitarnu matricu $U_1 = [u_1, u_2, \dots, u_m]$. Slično, za $v_1 = v$ postoji $n - 1$ ortonormiranih vektora $v_2, \dots, v_n \in \mathbb{C}^n$, takvih da je matrica $V_1 = [v_1, v_2, \dots, v_n]$ unitarna. Tada je

$$\begin{aligned} A_1 = U^*AV_1 &= \begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_m^* \end{bmatrix} \begin{bmatrix} Av_1 & Av_2 & \cdots & Av_n \end{bmatrix} \\ &= \begin{bmatrix} u_1^* \\ u_2^* \\ \vdots \\ u_m^* \end{bmatrix} \begin{bmatrix} \sigma_1 v_1 & Av_2 & \cdots & Av_n \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 & u_1^*Av_2 & \cdots & u_1^*Av_n \\ 0 & u_2^*Av_2 & \cdots & u_2^*Av_n \\ \vdots & \vdots & & \vdots \\ 0 & u_m^*Av_2 & \cdots & u_m^*Av_n \end{bmatrix} = \begin{bmatrix} \sigma_1 & w^* \\ 0 & A_2 \end{bmatrix} \end{aligned} \quad (1.12)$$

gdje je $w \in \mathbb{C}^{m-1}$, $A_2 \in \mathbb{C}^{(m-1) \times (n-1)}$. Za jedinični vektor koji odgovara prvom retku matrice iz jednakosti (1.12),

$$y = \frac{1}{\sqrt{\sigma_1^2 + w^*w}} \begin{bmatrix} \sigma_1 \\ w \end{bmatrix},$$

zbog unitarne invarijantnosti euklidske norme, vrijedi

$$\begin{aligned} \|A(V_1y)\|_2^2 &= \|(U_1A_1V_1^*)V_1y\|_2^2 = \|A_1y\|_2^2 \\ &= \left\| \begin{bmatrix} \sigma_1 & w^* \\ 0 & A_2 \end{bmatrix} \frac{1}{\sqrt{\sigma_1^2 + w^*w}} \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2^2 \\ &= \left(\frac{1}{\sqrt{\sigma_1^2 + w^*w}} \right)^2 \left\| \begin{bmatrix} \sigma_1 & w^* \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} \sigma_1 \\ w \end{bmatrix} \right\|_2^2 \\ &= \frac{1}{\sigma_1^2 + w^*w} \left\| \begin{bmatrix} \sigma_1^2 + w^*w \\ A_2w \end{bmatrix} \right\|_2^2 \\ &= \frac{(\sigma_1^2 + w^*w)^2 + \|A_2w\|_2^2}{\sigma_1^2 + w^*w} \geq \sigma_1^2 + w^*w. \end{aligned} \quad (1.13)$$

Ako je $w \neq 0$, onda je gornji izraz strogo veći od σ_1^2 . Iz (1.11) i $\sigma_1 = \|Av\|_2$ znamo da je σ_1 maksimalan pa dolazi do kontradikcije te zaključujemo da je $w = 0$.

Stoga je

$$A_1 = U_1^* A V_1 = \begin{bmatrix} \sigma_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

Ponavljanjem istog postupka za matricu $A_2 \in \mathbb{C}^{(m-1) \times (n-1)}$, dobivamo unitarne matrice U i V , kao produkt unitarnih matrica dobivenih nakon svakog koraka.

Ako je $m \geq n$, taj postupak vodi do dijagonalne matrice Σ .

Ako je $m \leq n$, primijenimo postupak opisan u dokazu teorema na matricu A^* . Nakon dobivene dekompozicije

$$A^* = \tilde{U} \tilde{\Sigma} \tilde{V}^*$$

i kompleksnog transponiranja obje matrice u toj jednakosti, dobijemo

$$A = \tilde{V} \tilde{\Sigma} \tilde{U}^*,$$

što je tražena dekompozicija na singularne vrijednosti od A , pri čemu moramo još preimenovali \tilde{V} u U , \tilde{U} u V i $\tilde{\Sigma}$ u Σ . \square

Lema 1.3.4. *Neka je A blok-stohastička matrica kao u definiciji 1.2.9 i neka je $B = PAP^T$, gdje je P bilo koja permutacija matrice. Neka je*

$$B = U \Sigma V^T$$

SVD od B kao u formuli (1.10). Tada je

$$A = \tilde{U} \tilde{\Sigma} \tilde{V}^T, \tag{1.14}$$

gdje je $\tilde{U} = P^T U$ i $\tilde{V}^T = V^T P$, SVD od A . Matrice A i B imaju jednake singularne vrijednosti.

Dokaz. Za svaku permutacijsku matricu P vrijedi $PP^T = I$. Iz toga slijedi da su matrice $\tilde{U} = P^T U$ i $\tilde{V}^T = V^T P$ ortogonalne. Tvrdnja slijedi iz definicije dekompozicije matrice na singularne vrijednosti. \square

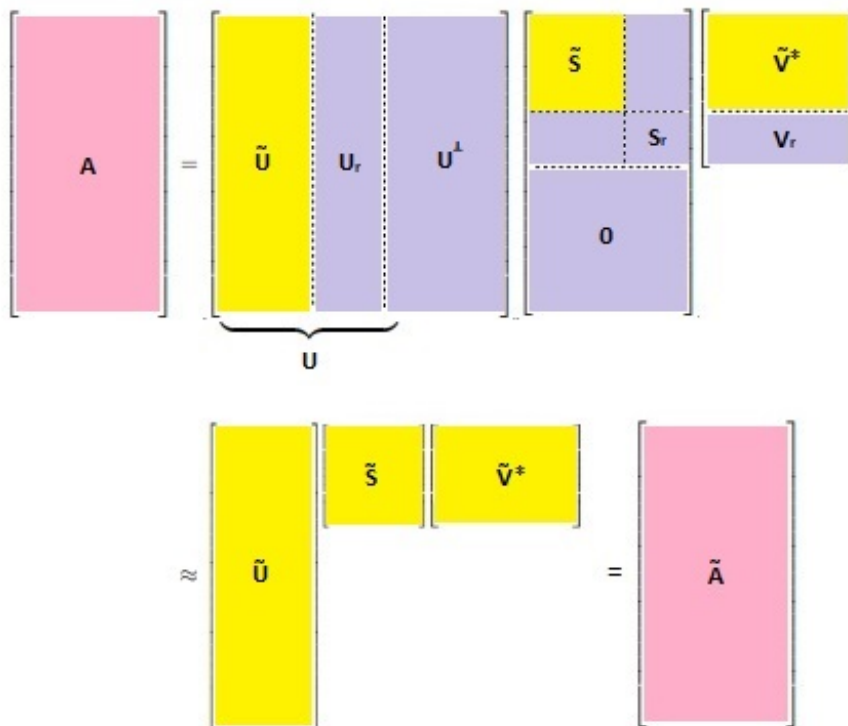
Jedna od najkorisnijih upotreba SVD-a je optimalna aproksimacija matrice matricom nižeg ranga. Zapravo, SVD nam pruža cijeli niz aproksimacija različite optimalnosti pošto se zasniva na odabiru broja vodećih singularnih vrijednosti početne matrice.

Teorem 1.3.5 (Eckart-Young-Minsky). *Neka je dana matrica $A \in \mathbb{C}^{m \times n}$ i $A = U \Sigma V^*$ njena SVD kao u teoremu 1.3.3. Njena optimalna aproksimacija ranga $r < \text{rang}(A)$ dana je s*

$$\underset{\tilde{A}}{\text{argmin}} \|A - \tilde{A}\|_2 = \tilde{U} \tilde{\Sigma} \tilde{V}^*, \tag{1.15}$$

pri čemu su \tilde{U} , \tilde{V} prvih r stupaca matrica U , V , a $\tilde{\Sigma}$ prvi blok dimenzije $r \times r$ matrice Σ .

Dokaz teorema može se naći u [3].



Slika 1.1: Shematski prikaz krnjeg SVD iz teorema 1.3.5, uz oznaku $S = \Sigma$

Primjer 1.3.6. Izravna korist teorema 1.3.5 se najslikovitije može prikazati pomoću kompresije slika. Siva slika je zapravo matrica piksela, a slika u boji ima i treću dimenziju zbog crvene, zelene i plave boje.

Slika 1.2 je originalna verzija fotografije mojih ljubimaca s matricom ranga 405. Upotrebom teorema 1.3.5 možemo značajno smanjiti broj podataka koje moramo pamti.

Primjenom teorema 1.3.5 smanjivan je rang matrice fotografije. Dok je fotografija na slici 1.3 s matricom ranga 100 još uvijek bez vidljivih promjena, do značajnijeg kvarenja kvalitete je došlo tek oko ranga 25 što možemo vidjeti na slici 1.4, iako je još uvijek očito što fotografija prikazuje.

Daljnijim smanjenjem ranga, kvaliteta fotografije se rapidno pogoršava. Već na slici 1.5 bi možda bilo upitno o kojim se životinjama radi kad nam ne bi bio predložen original.



Slika 1.2: Originalna slika ranga 405



Slika 1.3: Slika ranga 100



Slika 1.4: Slika ranga 25



Slika 1.5: Slika ranga 10

Primjer 1.3.7. U programskom jeziku MATLAB, SVD dekompozicija matrice A može se dobiti naredbom $[U,S,V] = \text{svd}(A)$:

$$A = \begin{bmatrix} 10 & 10 \\ 3 & 9 \\ 20 & 1 \\ 14 & 14 \end{bmatrix}, \quad U = \begin{bmatrix} -0.4544 & -0.2454 & -0.6251 & -0.5853 \\ -0.2418 & -0.4729 & 0.7429 & -0.4074 \\ -0.5747 & 0.7734 & 0.2346 & -0.1287 \\ -0.6362 & -0.3436 & -0.0478 & 0.6891 \end{bmatrix},$$

$$S = \begin{bmatrix} 30.3558 & 0 \\ 0 & 12.7092 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad V = \begin{bmatrix} -0.8457 & 0.5337 \\ -0.5337 & -0.8457 \end{bmatrix}.$$

Ako želimo samo prvih k singularnih vrijednosti i vektora matrice A , za to postoji naredba $[U,S,V] = \text{svds}(A,k)$.

Schurova dekompozicija

Teorem 1.3.8 (Schurova dekompozicija). Neka je $A \in \mathbb{C}^{n \times n}$ i neka su $\lambda_1, \dots, \lambda_n$ svojstvene vrijednosti od A u proizvoljnom poretku. Tada postoje unitarna matrica U i gornjetrokutasta matrica $T = [t_{ij}]$ takve da je

$$A = UTU^*, \quad t_{ii} = \lambda_i, i \in \{1, \dots, n\}. \quad (1.16)$$

Ako je $A \in \mathbb{R}^{n \times n}$ i ako su sve svojstvene vrijednosti od A realne, onda je T također realna i U se može odabrati da bude realna ortogonalna.

Definicija 1.3.9. Dekompoziciju $A = UTU^*$ zovemo Schurova dekompozicija od A , a matrica T se zove Schurova forma od A .

Dokaz. Dokaz se provodi matematičkom indukcijom po n .

Baza Uzmimo $n = 1$ i imamo $A = 1 \cdot A \cdot 1^*$. Skalar A nazivamo degeneriranom 1×1 gornjetrokutastom matricom, a 1 je unitarna 1×1 matrica.

Korak Neka je $A \in \mathbb{C}^{n \times n}$ i pretpostavimo da tvrdnja vrijedi za svaku $(n-1) \times (n-1)$ matricu.

- Promatrajmo svojstvenu vrijednost λ_1 i pripadni svojstveni vektor u_1 tako da je

$$Au_1 = \lambda_1 u_1, \quad \|u_1\| = 1.$$

- Skup $\{u_1\}$ nadopunimo s $\{u_2, \dots, u_n\}$ do ortonormirane baze u \mathbb{C}^n .

- Definirajmo ortonormiranu matricu $V_2 = [u_2, \dots, u_n] \in \mathbb{C}^{n \times (n-1)}$. Tada je matrica $U_1 = [u_1 \quad V] \in \mathbb{C}^{n \times n}$ unitarna matrica za koju vrijedi

$$U_1^* A U_1 = \begin{bmatrix} u_1^* \\ V_2^* \end{bmatrix} [A u_1 \quad A v_2] = \begin{bmatrix} u_1^* \\ V_2^* \end{bmatrix} [\lambda_1 u_1 \quad A V_2] = \begin{bmatrix} \lambda_1 & u_1^* A V_2 \\ 0 & A_2 \end{bmatrix},$$

pri čemu je $A_2 = V_2^* A V_2 \in \mathbb{C}^{(n-1) \times (n-1)}$.

- Iz činjenice da je

$$\det(A - \lambda I_n) = (\lambda_1 - \lambda) \cdot \det(A_2 - I_{n-1}),$$

slijedi da su $\lambda_2, \dots, \lambda_n$ svojstvene vrijednosti od A_2 .

- Po pretpostavci indukcije postoji unitarna matrica $U_2 \in \mathbb{C}^{(n-1) \times (n-1)}$ i gornjetrokutasta matrica $T_2 \in \mathbb{C}^{(n-1) \times (n-1)}$ s $\lambda_2, \dots, \lambda_n$ na dijagonali, takve da

$$A_2 = U_2 T U_2^*.$$

- Definirajmo sada

$$U = U_1 = \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} \in \mathbb{C}^{n \times n},$$

za koju vrijedi sljedeće:

$$U^* U = \begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} U_1^* U_1 \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & U_2^* U_2 \end{bmatrix} = I_n,$$

pa je U unitarna matrica sa svojstvom

$$\begin{aligned} U^* A U &= \begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} U_1^* A U_1 \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & U_2^* \end{bmatrix} \begin{bmatrix} \lambda_1 & u_1^* A V_2 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & U_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & u_1^* A V_2 U_2 \\ 0 & U_2^* A_2 U_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & u_1^* A V_2 U_2 \\ 0 & T_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & & & & \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ 0 & 0 & 0 & \ddots & * \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} = T. \end{aligned}$$

□

Napomena 1.3.10. Schurova forma nije jedinstvena jer se svojstvene vrijednosti na dijagonali mogu pojaviti u bilo kojem poretku.

Napomena 1.3.11. Numeričko računanje Schurove dekompozicije svodi se na beskonačan niz transformacija sličnosti koje sustavno reduciraju elemente ispod glavne dijagonale i osiguravaju trokutastu formu tek u limesu.

Primjer 1.3.12. U programskom jeziku MATLAB, Schurova dekompozicija matrice A može se dobiti naredbom $[U,T] = \text{schur}(A)$:

$$A = \begin{bmatrix} 16 & 18 & 4 \\ 15 & 7 & 0 \\ 18 & 14 & 15 \end{bmatrix},$$

$$U = \begin{bmatrix} 0.5287 & 0.6086 & -0.5917 \\ 0.3104 & -0.7874 & -0.5325 \\ 0.7900 & -0.0979 & 0.6052 \end{bmatrix}, \quad T = \begin{bmatrix} 32.5464 & -2.6439 & -19.8378 \\ 0 & -5.6357 & 0.6855 \\ 0 & 0 & 11.0893 \end{bmatrix}$$

Poglavlje 2

Analiza meta-stabilnih stanja Markovljevih lanaca pomoću strukture predznaka singularnih vektora

Sljedeći pristup traženja meta-stabilnih stanja Markovljevih lanaca je poopćenje pristupa danom u [5] koji je pogodan samo za reverzibilne Markovljeve lance i koristi svojstvene vektore umjesto singularnih. Općenito, ideja je pronaći permutaciju dane stohastičke matrice Markovljevog lanca tako da rezultirajuća matrica bude blok-dijagonalno dominantna. Tada dominantni dijagonalni blokovi predstavljaju različita meta-stabilna stanja lanca.

Prednost ove metode nad onom danom u [5] je, osim što se može primijeniti i na nereverzibilne Markovljeve lance, i ta što ona ne koristi stacionarnu distribuciju Markovljevog lanca čije je računanje često skupo i loše uvjetovano. Prednost nad nekim drugim pristupima koji koriste spektralne metode je također ušteda, pošto su u svakom koraku algoritma potrebna samo prva dva singularna vektora. Također, nije potrebno unaprijed znati broj meta-stabilnih stanja, nego se on računa paralelno s izvođenjem algoritma. Iz posljednjega slijedi da nam stoga nije nužna niti velika razlika unutar spektra tranzicijske matrice.

Osnovna ideja u ovom algoritmu je izračunati singularni vektor koji pripada drugoj najvećoj singularnoj vrijednosti, uzlazno sortirati njegove elemente i tu permutaciju primijeniti na matricu prijelaza. Postoji više drugih načina kako iskoristiti dekompoziciju na singularne vrijednosti u svrhu klasteriranja, kao npr. u [1], [7] i [17].

Za početak će biti obrađeni potrebni teorijski rezultati za Markovljeve lance s blok-dijagonalnom prijelaznom matricom, a zatim na one sa skoro blok-dijagonalnom prijelaznom matricom, tzv razdvojene i skoro razdvojene Markovljeve lance. Naposljetku će biti predstavljen algoritam.

2.1 Teorijska pozadina

Razdvojeni Markovljevi lanci

Neka je skup $S = \{1, 2, \dots, N\} = S_1 \oplus \dots \oplus S_k$ dekompozicija konačnog skupa stanja Markovljevog lanca u k disjunktih podskupova od kojih svaki predstavlja prostor stanja u kojem se Markovljev lanac nalazi "beskonačno mnogo vremena", odnosno stabilno stanje. Nakon odgovarajuće permutacije, prijelazna matrica B Markovljevog lanca bit će strogo blok-stohastička s podmatricama $\{A_{11}, \dots, A_{kk}\}$ kao u definiciji 1.2.9.

$$PBP^T = A = \begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_{mm} \end{bmatrix}. \quad (2.1)$$

Teorem 2.1.1. *Neka je A blok-stohastička matrica kao u formuli (2.1) s m ireducibilnih blokova A_1, \dots, A_m na dijagonali. Neka je*

$$A = \tilde{U}\Sigma\tilde{V}^T$$

SVD od A (definicija 1.3.1) i neka su $\tilde{u}_1, \dots, \tilde{u}_m$ m lijevih singularnih vektora tako da \tilde{u}_i odgovara najvećoj singularnoj vrijednosti bloka A_i za svaki $i = 1, \dots, m$. Neka vrijedi da je produkt $A_i A_i^T$ ireducibilna matrica za svaki $i = 1, \dots, m$.

Svakom stanju s_i pridružimo njegovu strukturu predznaka:

$$\text{sign}(s_i) := [\text{sgn}(\tilde{u}_1)_i, \dots, \text{sgn}(\tilde{u}_m)_i], \quad (2.2)$$

gdje

$$\text{sgn} : \mathbb{R} \rightarrow \{-1, 0, 1\}$$

$$x \rightarrow \begin{cases} 1, & x > 0 \\ 0, & x = 0. \\ -1, & x < 0 \end{cases}$$

Tada,

1. stanja koja pripadaju istom bloku unutar A posjeduju jednaku strukturu predznaka, odnosno za proizvoljan blok A_j i svaki $k, l \in S_j$ imamo $\text{sign}(s_k) = \text{sign}(s_l)$;
2. stanja koja pripadaju različitim blokovima unutar A posjeduju različitu strukturu predznaka, odnosno za svaka dva bloka A_i, A_j gdje $i \neq j$ i sve $k \in S_i, l \in S_j$ imamo $\text{sign}(s_k) \neq \text{sign}(s_l)$.

Dokaz. 1. Iz jednakosti (1.10) u teoremu 1.3.3 slijedi $AA^T = \tilde{U}\Sigma^2\tilde{U}^T$. Stoga su lijevi singularni vektori matrice A jednaki svojstvenim vektorima matrice AA^T , a singularne vrijednosti matrice A korijenima svojstvenih vrijednosti matrice AA^T .

Neka je $i \in \{1, \dots, m\}$ proizvoljan. Vrijedi $A_i A_i^T \geq 0$. Također, pretpostavili smo da je produkt matrica $A_i A_i^T$ ireducibilna matrica (definicija 1.6).

Imamo $A_i A_i^T \geq 0$ i $A_i A_i^T$ ireducibilna, pa možemo primijeniti Perron-Frobeniusov teorem 1.2.11. Postoji jednostavna svojstvena vrijednost $\varphi(A_i A_i^T)$ i pripadajući strogo pozitivan svojstveni vektor \hat{u}_i .

Stoga je vektor

$$\tilde{u}_i = [0, \dots, 0, \hat{u}_i^T, 0, \dots, 0]^T \quad (2.3)$$

svojstveni vektor od AA^T koji pripada najvećoj svojstvenoj vrijednosti bloka $A_i A_i^T$, odnosno lijevi singularni vektor koji pripada najvećoj singularnoj vrijednosti bloka A_i .

Iz ovoga slijedi da stanja koja pripadaju istom bloku imaju istu strukturu predznaka.

2. Zato što sva stanja istog bloka imaju jednaku strukturu predznaka, bez smanjenja općenitosti možemo pretpostaviti da svakom bloku A_i pripada samo jedno stanje kako bismo dobili kvadratnu matricu. Jer je $\tilde{U} = [\tilde{u}_1, \dots, \tilde{u}_m] \in \mathbb{R}^{m \times m}$ ortogonalna, tada je i \tilde{U}^T ortogonalna, odnosno ortogonalni su redovi matrice \tilde{U} . Stoga je nemoguće da dva stanja imaju jednaku strukturu predznaka.

□

Primjer 2.1.2. *Primjer strukture predznaka za Markovljev lanac sa sedam mogućih stanja i prijelaznom matricom koja se sastoji od tri bloka veličina $2 \times 2, 3 \times 3, 2 \times 2$.*

$$\begin{array}{c} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \end{array} \begin{array}{c} \left[\begin{array}{c} + \\ + \\ + \\ 0 \\ 0 \\ 0 \\ 0 \end{array} \right] \left[\begin{array}{c} \tilde{u}_1 \\ \tilde{u}_2 \\ \tilde{u}_3 \end{array} \right] \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ + \\ + \\ 0 \\ 0 \end{array} \right] \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ + \\ + \end{array} \right] \end{array}$$

Stanja s_1, s_2 koja pripadaju prvom bloku imaju strukturu predznaka $(+, 0, 0)$. Stanja s_3, s_4, s_5 koja pripadaju drugom bloku imaju strukturu predznaka $(0, +, 0)$. Stanja s_6, s_7 koja pripadaju trećem bloku imaju strukturu predznaka $(0, 0, +)$.

Neka druga ortogonalna kombinacija singularnih vektora $\tilde{u}_1, \tilde{u}_2, \tilde{u}_3$ mogla bi na primjer imati sljedeću strukturu predznaka:

$$\begin{array}{c}
 s_1 \\
 s_2 \\
 s_3 \\
 s_4 \\
 s_5 \\
 s_6 \\
 s_7
 \end{array}
 \begin{array}{c}
 v_1 \\
 v_2 \\
 v_3
 \end{array}
 \begin{array}{c}
 + \\
 + \\
 + \\
 + \\
 + \\
 + \\
 +
 \end{array}
 \begin{array}{c}
 - \\
 - \\
 + \\
 + \\
 - \\
 - \\
 -
 \end{array}
 \begin{array}{c}
 - \\
 - \\
 - \\
 - \\
 - \\
 + \\
 +
 \end{array}$$

Ponovno vidimo da stanja koja pripadaju istom bloku imaju istu strukturu predznaka.

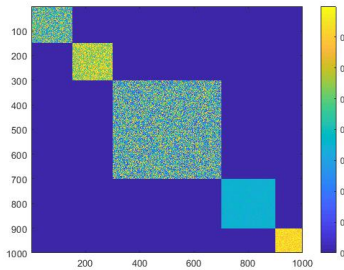
Napomena 2.1.3. Među pretpostavkama teorema 2.1.1 nalazi se da je produkt $A_i A_i^T$ ireducibilna matrica za svaki $i = 1, \dots, m$. U praksi je ta pretpostavka gotovo uvijek zadovoljena jer su svi elementi matrice pozitivni. Ukoliko ta pretpostavka nije zadovoljena, nule se mogu zamijeniti vrlo malim pozitivnim elementima koji neće utjecati na konačni rezultat. Što se dešava bez takve korekcije opisano je u članku [14].

Skoro razdvojeni Markovljevi lanci

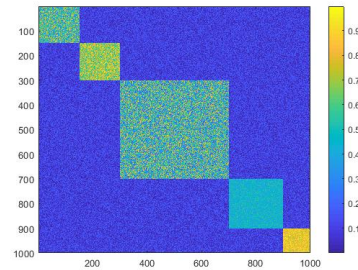
U primjenama iz stvarnog života, Markovljevi lanci su rijetko razdvojeni. Najčešće dobivamo skoro razdvojene Markovljeve lance, odnosno lance kod kojih se prijelazna matrica B može pomoću permutacijske matrice transformirati u

$$PBP^T = A + E = \begin{bmatrix} A_{11} & E_{12} & \cdots & E_{1m} \\ E_{21} & A_{22} & \cdots & E_{2m} \\ \vdots & \vdots & & \vdots \\ E_{m1} & E_{m2} & \cdots & A_{mm} \end{bmatrix}, \quad (2.4)$$

pri čemu su elementi matrice E jako mali. U ovom slučaju se Markovljev lanac u svakom skupu indeksa S_i nalazi "jako dugo", ali ne i "beskonačno" kao u razdvojenom slučaju.



Slika 2.1: Razdvojen Markovljev lanac



Slika 2.2: Skoro razdvojen Markovljev lanac

Slike 2.1 i 2.2 prikazuju prijelazne matrice Markovljevih lanaca nakon odgovarajućih permutacija kao u jednakostima (2.1) i (2.4).

Da bismo našli te permutacije, potrebno je definirati mjere kojima ćemo određivati koliko su dijagonalni blokovi veliki, odnosno vandijagonalni blokovi mali.

Napomena 2.1.4. *Neka su $S_k, S_l \subseteq \{1, \dots, n\}$ skupovi indeksa. B_{kl} ćemo označavati $B(S_k, S_l)$ kao u napomeni 1.2.10, odnosno B_{kl} će biti blok unutar matrice B koji sadrži elemente b_{ij} gdje $i \in S_k, j \in S_l$. Radi jednostavnosti, blokove B_{kk} kraće označavamo s B_k za svaki $k \in \{1, \dots, n\}$.*

U svrhu definicije prigodne norme za blok $B_{k,l}$ za svaki $k, l \in S$, poslužiti ćemo se uvjetnom vjerojatnošću prelaska iz bloka B_k u blok B_l . Definirajmo je na sljedeći način:

Definicija 2.1.5 (Uvjetna vjerojatnost prijelaza). *Neka je $B \in \mathbb{R}^{n \times n}$ stohastička matrica. Neka je $v \in \mathbb{R}^n$ inicijalna distribucija Markovljevog lanca. Neka su S_k i S_l disjunktni skupovi indeksa i B_k, B_l odgovarajući blokovi matrice B . Uvjetna vjerojatnost prijelaza iz B_k u B_l je dana s*

$$\omega_v(B_k, B_l) = \frac{\sum_{i \in S_k, j \in S_l} v_i b_{ij}}{\sum_{i \in S_k} v_i}. \quad (2.5)$$

Specijalno, u slučaju $l = k$, uvjetna vjerojatnost prijelaza $\omega_\pi(B_k, B_k)$ je vjerojatnost ostanka u bloku B_k .

Napomena 2.1.6. *Vektor v u definiciji 2.1.5 predstavlja inicijalnu distribuciju Markovljevog lanca. U svrhu pojednostavljivanja algoritma, pretpostavit ćemo da umjesto v može doći bilo koji pozitivan vektor v takav da $\sum_{i=1}^n v_i = 1$. Na temelju te pretpostavke, slijedi definicija v -norme. Sve o uvjetnoj vjerojatnosti može se naći u [4].*

Definicija 2.1.7 (v -norma). *Za svaki vektor $v > 0$ definiramo v -normu nenegativnog bloka B_{kl} matrice B s*

$$\|B_{kl}\|_v := \omega_v(B_k, B_l). \quad (2.6)$$

Slijedi definicija matrice veza koja sadrži informaciju o v -normama blokova B_{kl} , odnosno "vjerojatnosti prelaska" iz bloka B_k u blok B_l za svaki $k, l \in S$ ili ostanka u trenutnom bloku.

Definicija 2.1.8 (Matrica veza). *Neka su $S_1, \dots, S_m \subseteq \{1, \dots, n\}$ skupovi indeksa, $S_i \cap S_j = \emptyset, i \neq j$ i $\cup_{i=1}^m S_i = \{1, \dots, n\}$. Matrica veza dekompozicije je dana pomoću stohastičke matrice W_v definirane s*

$$(W_v)_{kl} = \omega_v(B_k, B_l), k, l = 1, \dots, m. \quad (2.7)$$

Napomena 2.1.9 (π -norma). *U [5] se kao vektor v uzima stacionarna distribucija π (definicija 1.1.6). Tada normu u definiciji 2.1.7 nazivamo π -norma.*

Napomena 2.1.10 (1-norma). v -norma u kojoj za vektor v uzmemo jedinični vektor nazivamo **1-norma**, a formula poprima sljedeći oblik

$$\|B_k\|_1 = \frac{1}{n_k} \sum_{i \in S_k, j \in S_l} b_{ij}. \quad (2.8)$$

Zapravo se radi o prosječnoj sumi redaka matrice.

Lema 2.1.11. Neka je B stohastička matrica. Neka je $S_k \subseteq \{1, \dots, n\}$ skup n_k indeksa i B_k pripadni blok unutar B . Nadalje, neka je v pozitivan vektor i v_{\min}, v_{\max} njegov najmanji i najveći element od svih elemenata s indeksima u S_k . Vrijedi

$$\|B_k\|_v \leq \frac{v_{\max}}{v_{\min}} \|B_k\|_1 \leq \frac{v_{\max}^2}{v_{\min}^2} \|B_k\|_v. \quad (2.9)$$

Dokaz. Iz formula (2.5) i (2.8) dobivamo

$$\|B_k\|_v = \frac{\sum_{i,j \in S_k} v_i b_{ij}}{\sum_{i \in S_k} v_i}$$

$$\|B_k\|_1 = \frac{1}{n_k} \sum_{i,j \in S_k} b_{ij}.$$

Jer je $v_{\min} \leq v_i \leq v_{\max}$ za svaki $i \in S_k$, imamo

$$\|B_k\|_v \leq \frac{\sum_{i,j \in S_k} v_{\max} b_{ij}}{\sum_{i \in S_k} v_{\min}} = \frac{v_{\max}}{v_{\min}} \frac{1}{n_k} \sum_{i,j \in S_k} b_{ij} = \frac{v_{\max}}{v_{\min}} \|B_k\|_1.$$

Na sličan način imamo i

$$\|B_k\|_v \geq \frac{\sum_{i,j \in S_k} v_{\min} b_{ij}}{\sum_{i \in S_k} v_{\max}} = \frac{v_{\min}}{v_{\max}} \|B_k\|_1.$$

Slijedi tvrdnja. □

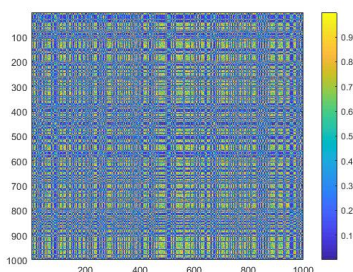
Napomena 2.1.12. Neka su η i v dvije norme na vektorskom prostoru V . Tada ih nazivamo ekvivalentnima ako postoje $a, b > 0$ takvi da za svaki vektor $v \in V$ vrijedi

$$a \cdot \eta(v) \leq v(v) \leq b \cdot \eta(v).$$

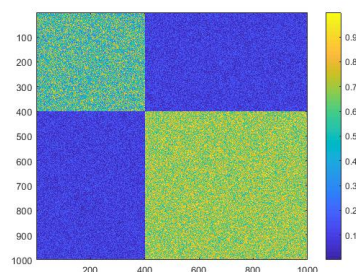
One tada induciraju istu topologiju na V .

Napomena 2.1.13. U algoritmu će biti korištena **1-norma** kako bi se izbjeglo nepotrebno i skupo računanje stacionarne distribucije Markoljevog lanca. Lema 2.1.11 nam daje potrebnu ekvivalenciju **1-norme** i proizvoljne v -norme. Korištenjem svake norme, algoritam daje iste blokove prijelazne matrice, razlika je jedino u matrici veza.

Definicija 2.1.14 (Skora razdvojenost). Za dva bloka matrice B kažemo da su skoro razdvojena ako je **1-norma** svakog bloka veća od dane vrijednosti $r = 1 - \delta$ pri čemu je $\delta > 0$ jako mali. Za matricu B kažemo da je skoro razdvojena ako se sastoji od m dijagonalnih blokova i pripadna matrica veza (definicija 2.1.8) je dijagonalno dominantna.



Slika 2.3: Promatrani blok u matrici B



Slika 2.4: Permutirani blok u matrici B

Slika 2.5: Primjenom odgovarajuće permutacije na blok sa slike 2.3 dobili smo dva istaknuta podbloka na slici 2.4. Računanjem njihove **1-norme** zaključujemo da li se radi o skoro razdvojenim blokovima.

2.2 Algoritam

U ovom odjeljku će biti predstavljen algoritam čiji je cilj naći permutaciju stohastičke matrice tako da permutirana matrica bude blok dijagonalno-dominantna oblika kao u formuli (2.4). Dijagonalno dominantni blokovi bit će pronalazeni rekursivno.

Algoritam za dva bloka

Neka je $PBP^T = A + E$ matrica oblika kao u jednakosti (2.4) i neka ima dva bloka ($m = 2$). Matrica P nije poznata.

Bez smanjena općenitosti, prema napomeni 2.1.3, možemo pretpostaviti da je $BB^T \geq 0$ ireducibilna i da možemo primijeniti Perron-Frobeniusom teorem 1.2.11. Slijedi da B ima jednostavnu singularnu vrijednost σ_1 i pripadni pozitivan lijevi singularni vektor u_1 . Promotrimo i drugu najveću singularnu vrijednost σ_2 i pripadni lijevi singularni vektor u_2 .

Iz ortogonalnosti singularnih vektora slijedi da u_1 i u_2 imaju promjenu u strukturi predznaka, a to znači da postoje stanja i, j koja pripadaju različitim blokovima.

Sortiramo uzlazno drugi singularni vektor i tu istu permutaciju primijenimo na matricu B . Dobivenu matricu $\tilde{B} = \tilde{P}B\tilde{P}^T$ podijelimo na dva bloka tako da je prvi blok veličine jednake broju negativnih vrijednosti u u_2 , a drugi blok veličine jednake broju pozitivnih vrijednosti u u_2 .

Prema definiciji 2.1.14 provjerimo radi li se zaista o skoro razdvojenim blokovima.

Algoritam za više blokova

Rekursivnom primjenom procedure permutacije velikog bloka koji sadrži dva manja bloka proširujemo algoritam na matrice s više blokova.

Potrebno je izračunati drugi lijevi singularni vektor, sortirati ga i primijeniti permutaciju na čitavu matricu B . Zatim slijedi provjera je li 1-norma potencijalnih blokova veća od zadanog praga.

Ako je, tada su otkrivena dva bloka te nastavljamo s procedurom na svakom pojedinom bloku od ta dva.

Ako nije, tada je nemoguće dalje razdvojiti blok i algoritam završava.

Algoritam

Input: matrica B , prag $r = 1 - \delta$

Output: broj m pronađenih blokova u B i veličine $n_i, i = 1, \dots, m$, permutacijska matrica P takva da $PTP^T = A + E$

1. Izračunaj drugi lijevi singularni vektor u_2 od B .
2. Sortiraj u_2 i dobivenu permutaciju P primijeni na matricu B .
3. Koristeći strukturu predznaka vektora u_2 odredi blokove B_1, B_2 .
4. Veličina B_1 je broj negativnih vrijednosti u u_2 . Veličina B_2 je broj pozitivnih vrijednosti u u_2 .
5. **Ako** je norma dijagonalnih blokova veća od r **onda**
 - Pronađena su dva bloka. Razdvoji ih.
 - Rekurzivno primijeni postupak od 1. koraka na oba bloka.
6. **Inače**
 - Trenutni blok se ne može dalje reducirati.
 - Povećaj brojač blokova za jedan. STOP.

Poglavlje 3

Analiza meta-stabilnih stanja nereverzibilnih Markovljevih lanaca - GenPCCA

Pristup predstavljen u prethodnom poglavlju je jedno od poopćenja metode dane u [5]. Iz te metode je dalje razvijen robusniji PCCA algoritam, a iz njega još robusniji PCCA+ algoritam. Oba su predstavljena u članku [6]. Metoda klasteriranja Perron Cluster Analysis (PCCA+) je primjenjiva za reverzibilne Markovljeve lance. Generalizirana PCCA metoda (GenPCCA) uključuje i nereverzibilne Markovljeve lance. Pristup je isti kao u PCCA+, samo se problem svojstvenih vrijednosti mijenja realnom Schurovom dekompozicijom.

U prošlom poglavlju smo stanja Markovljevog lanca klasterirali mijenjajući poredak stupaca i redaka prijelazne matrice P , tako da ona postane blok dijagonalno-dominantna. Takav algoritam svako stanje na jedinstven način smješta u neki klaster. Naziv za taj princip klasteriranja je ”*crisp clustering*”.

Sada ćemo uzeti u obzir da postoje ”rubna stanja” koja ne možemo sa sigurnošću smjestiti u jedinstveni klaster. Primjenit ćemo ”*fuzzy clustering*”.

3.1 Klasteriranje prostora stanja

Klasteriranje može biti obavljeno na diskretan način (*crisp clustering*), kao što je to u PCCA algoritmu, a predstavljeno sljedećim vrijednostima:

$$\mathcal{X}_i \in \{0, 1\}^n, \quad \mathcal{X}_{i,k} = \begin{cases} 1, & \text{stanje } i \text{ pripada klasteru } k, \\ 0, & \text{inače} \end{cases} . \quad (3.1)$$

U PCCA+ proučavamo robusniji princip klasteriranja. Definiramo matricu članstva \mathcal{X} dimenzije $N \times n$ kao

$$\mathcal{X}_i \in [0, 1]^n, \quad \sum_{j=1}^n \mathcal{X}_{ij} = 1, \quad i = 1, \dots, N. \quad (3.2)$$

Matrica \mathcal{X} opisuje članstvo svakog stanja u svakom klasteru, ali ne daje nam direktnu informaciju o tome kojem klasteru pripada pojedino stanje kao što je to u jednakosti (3.1).

3.2 Slučaj reverzibilnih Markovljevih lanaca

Pretpostavimo da je Markovljev lanac X ireducibilan i reverzibilan. Tada on ima jedinstvenu stacionarnu distribuciju $\pi = (\pi_1, \dots, \pi_N)$ (teorem 1.1.12), odnosno vrijedi

$$\pi^T P = \pi^T.$$

Neka je X stacionaran (definicija 1.1.5, napomena 1.1.7). Vrijedi da su π i P u detaljnoj ravnoteži (teorem 1.1.19):

$$\pi_i P_{ij} = \pi_j P_{ji}, \quad \forall i, j \in S.$$

Slijedi da prijelazna matrica P može biti transformirana u takozvanu simetričnu matricu relativnih tranzicijskih frekvencija \bar{P} s realnim svojstvenim vrijednostima

$$\bar{P} = D_\pi P, \quad (3.3)$$

pri čemu je D_π dijagonalna matrica s distribucijom π na dijagonali (dokaz teorema 1.2.16).

Prijelaznu matricu P dimenzije $N \times N$ Markovljevog lanca mijenjamo $n \times n$ prijelaznom matricom P_c koja nam opisuje prijelaze lanca između n klastera. Traži se projekcija $G(P) = P_c$ matrice P tako da ona ima smisla, odnosno da vrijedi

$$(G(P))^k = G(P^k), \quad (3.4)$$

odnosno da je tranzicijska matrica k -tog koraka projekcije jednaka projekciji tranzicijske matrice k -tog koraka.

Napomena 3.2.1. Iz definicije 1.1.21 i napomene 1.1.22 vidimo da Markovljev lanac možemo promatrati kao slijed distribucija umjesto kao slijed stanja iz skupa S . Ukoliko je $\eta \in \mathbb{R}^n$ vjerojatnosna distribucija u danom trenutku lanca, tada je $\hat{\eta} = P^T \eta$ vjerojatnosna distribucija u idućem koraku.

Projekcija prijelazne matrice P

Promatramo princip klasteriranja definiran u jednakosti (3.2). Svaki stupac matrice \mathcal{X} predstavlja jedan klaster. Zanima nas vjerojatnost prelaska iz jednog klastera u drugi, odnosno iz klastera \mathcal{X}_k u klaster \mathcal{X}_l , $k, l = 1, \dots, n$.

Prema [10], za diskretan princip klasteriranja predstavljen jednakošću (3.1), projicirana matrica frekvencija $\bar{P}_C \in \mathbb{R}^{n \times n}$ definirana je s

$$\bar{P}_C(k, l) = \sum_{i, j: \mathcal{X}_{ik}=1, \mathcal{X}_{jl}=1} \bar{P}(i, j),$$

odnosno prikazano pomoću matrice članstva \mathcal{X} :

$$\bar{P}_C = \mathcal{X}^T \bar{P} \mathcal{X} = \mathcal{X}^T D_\pi P \mathcal{X}$$

Stohastička matrica P_C novog Markovljevog lanca dobije se skaliranjem redova matrice \bar{P}_C :

$$P_C = (\mathcal{X}^T D_\pi \mathcal{X})^{-1} \mathcal{X}^T D_\pi P \mathcal{X} = \tilde{D}^{-1} \mathcal{X}^T D_\pi P \mathcal{X} \quad (3.5)$$

Propozicija 3.2.2. *Stacionarna distribucija novog Markovljevog lanca je*

$$\pi_C = \mathcal{X}^T \pi = \text{diag}(\tilde{D}_2). \quad (3.6)$$

Dokaz.

$$\pi_C^T P_C = \text{diag}(\tilde{D}_2)^T \tilde{D}_2^{-1} \mathcal{X}^T D_\pi P \mathcal{X} = \mathbf{1}_n^T \mathcal{X}^T D_\pi P \mathcal{X} = \mathbf{1}_N D_\pi P \mathcal{X} = \pi^T P \mathcal{X} = \pi^T \mathcal{X} = \pi_C^T.$$

□

Napomena 3.2.3. *Funkcija $\text{diag}(\cdot)$ vraća dijagonalu matrice u argumentu.*

Napomena 3.2.4. *Ako je matrica \mathcal{X} definirana na diskretan način kao u jednakosti (3.1), tada je $\tilde{D} = \tilde{D}_2$.*

Jer proizlaze samo iz činjenice da su sume redaka matrice \mathcal{X} i P jednake 1, rezultati (3.5) i (3.6) se mogu koristiti i za princip klasteriranja dan u jednakosti (3.2).

Definicija 3.2.5. *Neka je matrica P_C kao u jednakosti (3.5). Metastabilnost klastera $\mathcal{X}_1, \dots, \mathcal{X}_n$ definira se kao trag matrice P_C .*

Matrica članstva \mathcal{X}

Teorem 3.2.6. *Ako je Markovljev lanac reverzibilan, matrica članstva \mathcal{X} može se dobiti kao*

$$\mathcal{X} = YA, \quad (3.7)$$

pri čemu je $Y \in \mathbb{R}^{N \times n}$ matrica n svojstvenih vektora prijelazne matrice P koji odgovaraju prvim n najvećim svojstvenim vrijednostima, a matrica $A \in \mathbb{R}^{n \times n}$ nesingularna transformacijska matrica. Broj klastera n je broj svojstvenih vrijednosti blizu Perronovog korijena 1.

Dokaz. 1. Za početak opravdajmo odabir matrice Y kao matrice svojstvenih vektora prijelazne matrice. Prvo promotrimo idealan slučaj razdvojenog Markovljevog lanca s prijelaznom matricom \tilde{P} permutiranom u blok-matricu kao u definiciji 1.2.9. Svaki blok predstavlja jedan od klastera S_1, \dots, S_n . Suma svakog retka matrice \tilde{P} je 1, što se može izraziti pomoću

$$\tilde{P}\mathbf{1}_N = \mathbf{1}_N.$$

Neka je matrica $\tilde{\mathcal{X}}$ matrica članstva definirana kao u jednakosti (3.1), odnosno $\tilde{\mathcal{X}}_{i,j} = 1$ za svaki $i \in S_j$, a 0 inače. To znači da vrijedi

$$\tilde{P}\tilde{\mathcal{X}} = \tilde{\mathcal{X}}.$$

Ovo je jednadžba svojstvenih vektora.

Za svaku regularnu transformacijsku matricu $\tilde{A} \in \mathbb{R}^{n \times n}$, stupci matrice $\tilde{Y} \in \mathbb{R}^{N \times n}$, $\tilde{Y} = \tilde{\mathcal{X}}\tilde{A}^{-1}$ su svojstveni vektori matrice \tilde{P} koji odgovaraju svojstvenoj vrijednosti $\lambda = 1$. Za reducibilnu matricu \tilde{P} , n klastera se može prepoznati po jednakim strukturama redova matrice $\tilde{\mathcal{X}}$. Obratno, ako je \tilde{Y} matrica prvih n svojstvenih vektora matrice \tilde{P} , tada postoji regularna transformacijska matrica \tilde{A} takva da je

$$\tilde{\mathcal{X}} = \tilde{Y}\tilde{A}$$

tražena matrica članstva s nulama i jedinicama.

2. Sada pomotrimo naš manje idealan slučaj, odnosno matricu P . Intuitivno, što je P po strukturi sličnija \tilde{P} , to je matrica \mathcal{X} sličnija matrici $\tilde{\mathcal{X}}$. Prema [6, 5.str, (1.9)], za $\epsilon \geq 0$ vrijedi

$$Y = \tilde{Y} + \epsilon Y^{(1)} + \mathcal{O}(\epsilon^2). \quad (3.8)$$

Nadalje, prema [6, Lemma 1.1] vrijedi sljedeća tvrdnja. Postoji matrica $B \in \mathbb{R}^{n \times n}$ takva da se $Y^{(1)}$ u formuli (3.8) može zapisati kao

$$Y^{(1)} = \tilde{\mathcal{X}}B,$$

pri čemu je $\tilde{\mathcal{X}}$ kao i ranije definirana matrica po dijelovima konstantnih svojstvenih vektora matrice \tilde{P} . Iz toga slijedi da \tilde{Y} i $\tilde{Y} + \epsilon Y^{(1)}$ razapinju isti potprostor. Iz toga i iz jednakosti (3.8) možemo zaključiti

$$\mathcal{X} = \tilde{\mathcal{X}} + O(\epsilon^2).$$

3. Za kraj, opravdajmo odabir n od N svojstvenih vektora koji pripadaju svojstvenim vrijednostima iz Perronovog klastera. Koristeći teorem 1.2.16, napišimo matricu P_C definiranu u jednakosti (3.5) u drugačijoj formulaciji:

$$\begin{aligned} P_C &= (\mathcal{X}^T D_\pi \mathcal{X})^{-1} \mathcal{X}^T D_\pi P \mathcal{X} = (\mathcal{X}^T D_\pi \mathcal{X})^{-1} \mathcal{X}^T D_\pi P Y A \\ &= (\mathcal{X}^T D_\pi \mathcal{X})^{-1} \mathcal{X}^T D_\pi Y \Delta A = (\mathcal{X}^T D_\pi \mathcal{X})^{-1} \mathcal{X}^T D_\pi \mathcal{X} A^{-1} \Delta A \\ &= S A^{-1} \Delta A, \end{aligned}$$

pri čemu su $\Delta = \text{diag}(\lambda_1, \dots, \lambda_n)$, $S = \left(\frac{\langle X_i, X_j \rangle_\pi}{\langle X_i \rangle_\pi} \right)_{i,j=1,\dots,n}$.

Pretpostavimo da vrijedi

$$\|A^{-1} \Delta A - I_n\|_1 < 1.$$

Tada imamo

$$\text{trag}(\Delta) - O(\epsilon^2) \leq \text{trag}(P_C) \leq \text{trag}(\Delta).$$

Kompletan dokaz zadnje činjenice može se naći u [6]. Zato uzimamo svojstvene vektore prvih n najvećih svojstvenih vrijednosti. Time dobivamo maksimalnu metastabilnost uz optimalan broj klastera. □

Markovljevošć matrice P_C

Matrica Y prvih n svojstvenih vektora zadovoljava sljedeće uvjete:

- *uvjet invarijantnog prostora:*

$$P Y = Y \Delta \tag{3.9}$$

za $\Delta \in \mathbb{R}^{n \times n}$.

- *relacija ortogonalnosti:*

$$Y^T D_\pi Y = I_{n \times n}, \tag{3.10}$$

Sukladno formuli (3.6), ako je η distribucija lanca u trenutku t kao u definiciji 1.1.21, tada će

$$\eta_C = \mathcal{X}^T \eta \tag{3.11}$$

biti njena projekcija na klastere, odnosno distribucija novog lanca u trenutku t .

Pošto je Markovljev lanac reverzibilan, kao što je dokazano u teoremu 1.2.16, lijeve svojstvene vektore matrice P možemo dobiti s

$$Z = D_\pi Y. \quad (3.12)$$

Uočimo da vrijedi $Z^T Y = Y^T D_\pi Y = I$, gdje je $I \in \mathbb{R}^{n \times n}$ jedinična matrica.

Lanac je stacionaran pa je distribucija η linearna kombinacija lijevih svojstvenih vektora prijelazne matrice P

$$\eta = Z\alpha, \quad (3.13)$$

gdje je $\alpha \in \mathbb{R}^n$ vektor koeficijenata.

Iz formule (3.11) čitamo da je projekcijska matrica dana s

$$\Pi^f = \mathcal{X}^T = A^T Y^T. \quad (3.14)$$

Iz prethodnih jednakosti dobivamo

$$A^T \alpha = A^T I \alpha = A^T Y^T D_\pi Y \alpha = \{(3.14), (3.12)\} = \mathcal{X}^T Z \alpha = \{(3.13)\} = \mathcal{X}^T \eta = \{(3.11)\} = \eta_C.$$

Definirajmo projekciju unazad

$$\Pi^b = D_\pi Y A^{-T}. \quad (3.15)$$

Vrijedi

$$\begin{aligned} \Pi^b \eta_C &= (D_\pi Y A^{-T}) \eta_C = (D_\pi Y A^{-T}) (A^T Y^T D_\pi Y \alpha) \\ &= D_\pi Y (Y^T D_\pi) Y \alpha \\ &= D_\pi Y Z^T Y \alpha \\ &= D_\pi Y \alpha \\ &= Z \alpha \\ &= \eta. \end{aligned}$$

Stavimo i $\alpha = Y^T \eta$. Promatramo li Markovljev lanac kao slijed distribucija, sljedeća lema nam govori o tome da su za distribucije propagirane pomoću P^T , pripadne projicirane distribucije propagirane pomoću P_C^T bez greške.

Lema 3.2.7. *Propagacija projiciranje distribucije komutira s projekcijom propagirane distribucije, odnosno*

$$\Pi^b (P_C^T)^k \Pi^f \eta = (P^T)^k \eta.$$

Dokaz. Neka je $k \in \mathbb{N}$ broj koraka za koji se pomičemo. Vrijedi

$$\begin{aligned} \Pi^b (P_C^T)^k \Pi^f \eta &= D_\pi Y A^{-T} [(\mathcal{X}^T P^T D_\pi \mathcal{X})(\mathcal{X}^T D_\pi \mathcal{X})^{-1}]^k A^T Y^T \eta \\ &= D_\pi Y A^{-T} [(A^T Y^T P^T D_\pi Y A)(A^T Y^T D_\pi Y A)^{-1}]^k A^T Y^T \eta \\ &= D_\pi Y A^{-T} [A^T \Delta A^{-T}]^k A^T Y^T \eta = D_\pi Y \Delta^k Y^T \eta = D_\pi Y \Delta^k \alpha = (P^T)^k D_\pi \mathcal{X} \alpha \\ &= (P^T)^k \eta \end{aligned}$$

pri čemu je $\Delta \in \mathbb{R}^{n \times n}$ matrica dominantnih realnih svojstvenih vrijednosti. □

3.3 Slučaj nereverzibilnih Markovljevih lanaca, GenPCCA

Sada ćemo promatrati općenitiji slučaj, odnosno nereverzibilne Markovljeve lance čija prijelazna matrica u pravilu ima kompleksni spektar i prilagoditi dosadašnje rezultate za PCCA+ tako da umjesto svojstvenih vrijednosti koristimo Schurovu dekompoziciju.

Inicijalnu distribuciju Markovljevog lanca označimo s $\eta = (\eta_1, \dots, \eta_N)$.

Matrica članstva X

Označimo s D_η dijagonalnu matricu s vektorom η na dijagonali. Ideja je pronaći matricu Y i Δ koje zadovoljavaju iste uvjete kao i za reverzibilne Markovljeve lance, odnosno uvjet invarijantnog prostora (3.9) i relaciju ortogonalnosti (3.10).

Promotrimo matricu

$$\tilde{P} = D_\eta^{0.5} P D_\eta^{-0.5}.$$

Napomena 3.3.1. Neka je $D \in \mathbb{R}^{n \times n}$ proizvoljna dijagonalna matrica. U ovom radu, oznaka $D^{0.5}$ označava dijagonalnu matricu čiji dijagonalni elementi su korijeni dijagonalnih elemenata matrice D u istom poretku.

Napomena 3.3.2. Prema teoremu 1.3.8, Schurova dekompozicija matrice \tilde{P} je oblika $\tilde{P} = Q \Delta Q^{-1}$ pri čemu je Q ortonormalna matrica čiji stupci se nazivaju Schurovim vektorima, a matrica Δ zvana Schurova forma je gornje-trokutasta matrica s 1-1 i 2-2 blokovima na dijagonali, ovisno o tome je li svojstvena vrijednost matrice \tilde{P} realna ili kompleksna.

Napomena 3.3.3. Analogno odabiru prvih n svojstvenih vektora u slučaju reverzibilnih Markovljevih lanaca, sada biramo prvih n Schurovih vektora matrice \tilde{P} . Sredimo matricu Δ tako da gornji $n \times n$ blok sadrži n najvećih svojstvenih vrijednosti i ponovno izračunamo Schurove vektore koji odgovaraju danim svojstvenim vrijednostima. Za izračun prvih n Schurovih vektora koristit će se algoritam SRSchur opisan u [2].

Neka je \tilde{Y} matrica n vodećih Schurovih vektora od \tilde{P} . Vrijedi

$$\begin{aligned}\tilde{P}\tilde{Y} = \tilde{Y}\Delta &\iff D_\eta^{0.5}PD_\eta^{-0.5}\eta Y = \eta Y\Delta \\ &\iff PD_\eta^{-0.5}\tilde{Y} = D_\eta^{-0.5}\tilde{Y}\Delta \\ &\iff PY = Y\Delta, \quad Y = D_\eta^{-0.5}\tilde{Y}.\end{aligned}\tag{3.16}$$

Matrica Y očito zadovoljava uvjet invarijantnog potprostora (3.9), ali i relaciju ortogonalnosti (3.10) zbog ortonormalnosti matrice \tilde{Y} :

$$Y^T D_\eta Y = \tilde{Y}^T D_\eta^{-0.5} D_\eta D_\eta^{-0.5} \tilde{Y} = \tilde{Y}^T \tilde{Y} = I_{n \times n}.\tag{3.17}$$

Analogno matrici P_C definiranoj formulom (3.5), projekcija matrice P neka je sada dana s

$$G(P) = (\mathcal{X}^T D_\eta \mathcal{X})^{-1} (\mathcal{X}^T D_\eta P \mathcal{X}) = \tilde{D}^{-1} (\mathcal{X}^T D_\eta P \mathcal{X}).\tag{3.18}$$

Teorem 3.3.4. Neka je $G(P) = (\mathcal{X}^T D_\eta \mathcal{X})^{-1} (\mathcal{X}^T D_\eta P \mathcal{X})$, gdje je $\mathcal{X} = YA$, Y matrica n vodećih Schurovih vektora i D_η dijagonalna matrica s inicijalnom distribucijom danog Markovljevog lanca na dijagonali. Vrijedi

$$(G(P))^k = G(P^k).$$

Dokaz.

$$\begin{aligned}G(P) &= (\mathcal{X}^T D_\eta \mathcal{X})^{-1} (\mathcal{X}^T D_\eta P \mathcal{X}) = (A^T Y^T D_\eta Y A)^{-1} (A^T Y^T D_\eta P Y A) \\ &= (A^T Y^T D_\eta Y A)^{-1} (A^T Y^T D_\eta Y \Delta A) = (A^T A)^{-1} (A^T \Delta A) = A^{-1} \Delta A.\end{aligned}$$

Dalje vrijedi

$$(G(P))^k = (A^{-1} \Delta A)^k = A^{-1} \Delta^k A = G(P^k).$$

□

3.4 Transformacijska matrica A

Rješenje problema $X = YA$ nije jedinstveno. Broj transformacijskih matrica A tako da $X = YA$ bude stohastička matrica je beskonačan. Na primjer, uvijek možemo naći matricu A tako da svaki element matrice X bude jednak $1/n$. Problemu traženja matrice A pristupamo kao problemu optimizacije.

Dopustivi skup

Tražimo dopustivi skup matrica $F_A \subset \mathbb{R}^{n \times n}$ za matricu A .

Jer je P stohastička matrica, vektor $e = (1, \dots, 1)^T$ je svojstveni vektor koji odgovara njenoj najvećoj svojstvenoj vrijednosti $\lambda_1 = 1$. Pošto su stupci matrice Y uređeni tako da odgovaraju Schurovoj formi u silazno uređenom obliku, imamo $Y_{i,1} = 1, i = 1, \dots, N$. Koristeći tu činjenicu i $X = YA$, preformulirajmo sljedeće uvjete koje zadovoljava matrica X :

- $\sum_{j=1}^n X_{ij} = 1$ (particija jedinice)
- $X_{ij} \geq 0$ za svaki $i = 1, \dots, n$ (pozitivnost)

u uvjete zadane pomoću matrica Y i A :

$$A_{1,j} \geq - \sum_{k=2}^n Y_{i,k} A_{k,j}, i = 1, \dots, N, j = 1, \dots, n \quad \text{pozitivnost} \quad (3.19)$$

$$A_{i,1} = \delta_{i,1} - \sum_{j=2}^n A_{i,j}, i = 1, \dots, n \quad \text{particija jedinice} \quad (3.20)$$

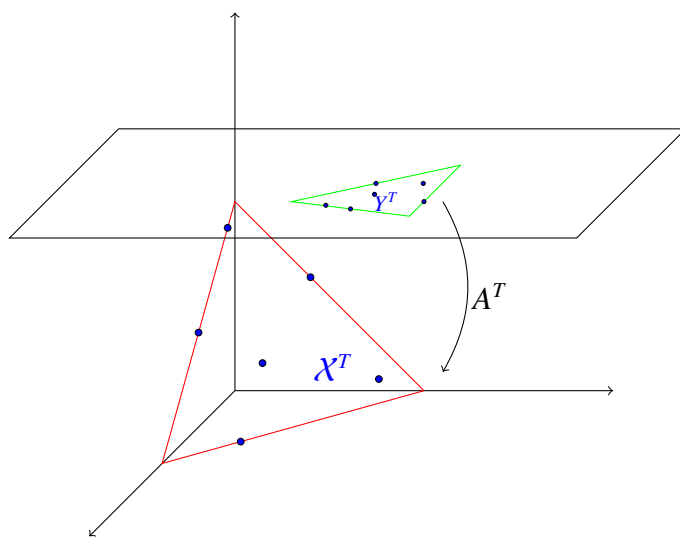
Dani uvjeti za F_A su linearni po A iz čega zaključujemo da je dopustivi skup F_A konveksni politop (definicija 3.4.1).

Definicija 3.4.1. Neka je $S = \{x_1, \dots, x_m\} \subset \mathbb{R}^n$ konačan skup. Konveksnu ljusku skupa S , odnosno skup svih konveksnih kombinacija točaka iz S

$$x = \sum_{i=1}^m \alpha_i x_i, \quad \alpha_i \geq 0, i = 1, \dots, m, \quad \sum_{i=1}^m \alpha_i = 1,$$

nazivamo konveksnim politopom s generatorima x_1, \dots, x_m .

Definicija 3.4.2. K – simpleks je k -dimenzionalni politop koji je konveksna ljuska svojih $k + 1$ vrhova. Standardni $(n - 1)$ – simpleks je konveksna ljuska vektora kanonske baze $\{e_1, \dots, e_n\}$ koji su, osim linearno, i afino nezavisni.



Slika 3.1: Shematski prikaz linearne transformacije redaka matrice Y na standardni 2-simpleks.

Napomena 3.4.3. Za točke $x_1, \dots, x_m \in \mathbb{R}^n$ kažemo da su afino nezavisne ako su jednakosti

$$\sum_{i=1}^m \alpha_i x_i = 0 \quad i \quad \sum_{i=1}^m \alpha_i = 0$$

moгуće jedino za $\alpha_1 = \alpha_2 = \dots = \alpha_m = 0$.

Shvatimo N redaka matrice X kao točke u \mathbb{R}^n . Iz uvjeta particije i pozitivnosti za matricu X slijedi ta te točke leže na standardnom $(n-1)$ -simpleksu Δ . Vrijedi $X^T = A^T Y^T$ što znači da matrica A linearno povezuje N redaka matrice Y sa simpleksom Δ . Iz $Y_{i,1} = 1, i = 1, \dots, N$ znamo da svi retci matrice Y leže na hiperravnini s prvom komponentom 1. Shematski prikaz nalazi se na slici 3.1.

Teorem 3.4.4 (Teorem o vrhu). *Ako linearni program ima optimalno rješenje (maksimum ili minimum), ono se mora nalaziti u vrhu izvedbenog područja.*

Napomena 3.4.5. *Zbog teorema 3.4.4, skup vrhova $v(F_A) \subset F_A$ igra važnu ulogu. Vrijedi $F_A \subset \mathbb{R}^{n \times n}$ pa je vrh $A \in v(F_A)$ određen s n^2 linearnih ograničenja od ograničenja danih u (3.19) i (3.20). Dalje, moguće je karakterizirati podskup od $2n$ aktivnih ograničenja.*

Lema 3.4.6. *Neka je $A \in F_A$ regularna i $X = YA$. Ako postoji indeks $i \in \{1, \dots, n\}$ tako da za svaki $l = 1, \dots, N$ vrijedi*

$$0 < X(l, i),$$

tada $A \notin v(F_A)$.

Dokaz. Bez smanjena općenitosti, neka je $i = 1$. Definirajmo

$$A^* = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}$$

i

$$0 < \delta := \min_{j=1, \dots, N} \mathcal{X}(j, 1) < 1, \quad B := \frac{A - \delta A^*}{1 - \delta}.$$

Provjerom uvjeta slijedi $A^*, B \in F_A$. Zato što vrijedi $A = (1 - \delta)B + \delta A$, A nije vrh konveksnog skupa F_A . \square

Iz leme 3.4.6 slijedi da ako je A vrh od F_A , za svaki $i = 1, \dots, n$ postoji $l \in \{1, \dots, N\}$ takav da $X_{l,i} = 0$. Zato je za svaki $i = 1, \dots, n$ aktivna barem jedna od jednakosti u (3.19).

Stoga možemo definirati skup $F'_A \subset F_A$, $v(F_A) \subset F'_A$ sljedećim jednakostima:

$$A_{j,1} = \delta_{j,1} - \sum_{i=2}^n A_{j,i}, \quad j = 1, \dots, n, \quad (3.21)$$

$$A_{1,i} = - \min_{l=1, \dots, N} \sum_{j=2}^n A_{j,i} Y_{l,j}, \quad i = 1, \dots, n. \quad (3.22)$$

Skup F'_A sadrži optimalan vrh A iz F_A , a dimenzija problema je reducirana.

Napomena 3.4.7. Osim u jednakosti (3.21) za $j = 1$, svaka jednakost koja definira skup F'_A je invarijantna s obzirom na $A \rightarrow \gamma A, \gamma > 0$. Ta činjenica se koristi dalje u algoritmu za konstrukciju dopustivih matrica A i u transformaciji problema u problem neograničene optimizacije.

Primjer 3.4.8. Lako se vidi da F_A nije prazan skup jer sadrži točku $A(j, i) = \delta_{j,1} n_C^{-1}$.

$$A = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Uvjet pozitivnosti (3.19):

$$A_{1,j} = \frac{1}{n} \geq - \sum_{k=2}^n Y_{i,k} A_{k,j} = - \sum_{k=2}^n X_{i,k} \cdot 0 = 0, \quad j = 1, \dots, n.$$

Uvjet particije jedinice (3.20):

$$A_{1,1} = \delta_{1,1} - \sum_{j=2}^n A_{i,j} = 1 - (n-1) \cdot \frac{1}{n} = \frac{1}{n}$$

$$A_{i,1} = \delta_{i,1} - \sum_{j=2}^n A_{i,j} = 0 - 0 = 0, \quad i = 2, \dots, n.$$

Funkcija cilja

Slijede dvije prirodne funkcije cilja koje bi trebale biti maksimizirane da se dobije optimalna transformacijska matrica A :

- *Uvjet maksimalnog skaliranja.* Idealno, za svaki klaster \mathcal{X}_j bi trebalo postojati stanje $i \in \{1, \dots, N\}$ koje mu sigurno pripada, odnosno

$$\max_{i=1, \dots, N} \mathcal{X}_{i,j} = 1, \quad j = 1, \dots, n.$$

Pošto ne promatramo idealan slučaj, dovoljno je zahtijevati da vrijedi

$$\max_{i=1, \dots, N} \mathcal{X}_{i,j} \approx 1, \quad j = 1, \dots, n.$$

Slijedi da je cilj je maksimizirati sljedeću funkciju:

$$I_1(A) := \sum_{j=1}^n \max_{i=1, \dots, N} \mathcal{X}_{i,j} \leq n.$$

- *Uvjet maksimalne metastabilnosti.* Optimizacijskom problemu se može pristupiti i kao traženju maksimalne metastabilnosti (definicija 3.2.5), odnosno minimizaciji vjerojatnosti prelaženja sustava iz jednog klastera stanja u drugi.

$$\begin{aligned} P_C &= \tilde{D}^{-1} \mathcal{X}^T D_\eta P \mathcal{X} \\ &= \tilde{D}^{-1} A^T Y^T D_\eta P Y A \\ &= (3.16) = \tilde{D}^{-1} A^T Y^T D_\eta Y \Delta A \\ &= (3.17) = \tilde{D}^{-1} A^T \Delta A \end{aligned}$$

$$I_2 := \text{trag}(P_C) = \text{trag}(\tilde{D}^{-1} A^T \Delta A) \leq \sum_{i=1}^n \lambda_i.$$

Lema 3.4.9. [16, lema3.6] Dijagonalni elementi d_i , $i = 1, \dots, n$, matrice \tilde{D}^{-1} mogu se dobiti iz transformacijske matrice A:

$$d_i = A_{1,i}, \quad i = 1, \dots, n.$$

Stoga funkcija cilja I_2 ima oblik

$$I_2(A) := \text{trag}(W) = \sum_{i=1}^n \lambda_i \sum_{j=1}^n \frac{A_{i,j}^2}{A_{1,j}}.$$

Lema 3.4.10. Dopustiv skup F_A za matricu A definiran ograničenjima (3.20) i (3.19) je konveksni politop. Funkcije cilja I_1 i I_2 su konveksne na F_A . Za I_1 i I_2 postoje optimalni vrhovi skupa politopa F_A koji ih maksimiziraju, odnosno postoji rješenje $A \in v(F_A) \subset F'_A$ optimizacijskog problema, pri čemu je F'_A definiran pomoću jednakosti (3.21) i (3.22).

Dokaz. Ograničenja kojima je definiran F_A su linearna. Prema definiciji konveksnog politopa (definicija 3.4.1), F_A je politop.

Konveksnost funkcije cilja I_1 : $g(A) := \sum_{i=1}^n Y_{l,i} A_{i,j}$ za $l \in \{1, \dots, N\}$ je linearna funkcija i time konveksna. Maksimum linearnih funkcija je konveksna funkcija. Suma konveksnih funkcija je konveksna funkcija. Kompozicija ovih konveksnih funkcija daje konveksnu funkciju.

Konveksnost funkcije cilja I_2 : Definirajmo funkciju $f(a, b) := a^2 b^{-1}$. Promatramo f za $b > 0$ jer pozitivnost $b = A_{1,j}$ slijedi iz leme 3.4.9. Hessianova matrica

$$H_f(a, b) = \begin{bmatrix} \frac{2}{b} & \frac{-2a}{b^2} \\ \frac{-2a}{b^2} & \frac{2a^2}{b^3} \end{bmatrix}$$

sa svojstvenim vrijednostima $\lambda_1 = 0$, $\lambda_2 = \frac{a^2+b^2}{b^3}$ je pozitivno semi-definitna. Time je funkcija I_2 konveksna.

Funkcija I_1 je odozgo ograničena s n , a funkcija I_2 s $\sum_{i=1}^n \lambda_i$. Supremum konveksne neprekidne funkcije nad konveksnim politopom postiže se u vrhovima politopa prema lemi 3.4.6.

□

Optimizacija bez ograničenja

Dobiveni optimizacijski problem je zbog brojnih uvjeta za dopustivi skup vremenski vrlo zahtjevan. Stoga primjenjujemo optimizaciju bez ograničenja.

Kao što je prikazano na slici 3.1, redovi matrice X leže na standardnom simpleksu Δ . Vrijedi:

$$x \in \Delta \iff \exists \gamma_1, \dots, \gamma_n \in [0, 1], x = \sum_{j=1}^n \gamma_j e_j, \sum_{j=1}^n \gamma_j = 1,$$

pri čemu su $e_1, \dots, e_n \in \mathbb{R}^n$ jedinični vektori. Simpleks Δ linearnom transformacijom A^{-1} preslikavamo u drugi simpleks kojeg ćemo označiti s Δ_2 . On sadrži redove matrice Y dobivene preslikavanjem redova matrice X .

$$\Delta_2 = \Delta A^{-1}.$$

Vrhovi simpleksa Δ_2 dani su s

$$v_j^T = e_j^T A^{-1} = A_{j,\cdot}^{-1}$$

(3.23)

za $j = 1, \dots, n$. Zaključujemo da su redovi matrice A^{-1} jednaki vrhovima simpleksa Δ_2 .

Neka je zadovoljen uvjet maksimalnog skaliranja

$$\max_{i=1, \dots, N} X_{i,j} = 1, \quad \forall j = 1, \dots, n.$$

Drugim riječima, neka za svaki $j = 1, \dots, n$ postoji red i matrice X takav da je $X_{i,\cdot} = e_j^T$, odnosno vrh simpleksa Δ . Iz toga slijedi da je $Y_{i,\cdot} = v_j^T$ vrh simpleksa Δ_2 . Iz jednakosti (3.4) slijedi da je

$$A_{j,\cdot}^{-1} = Y_{i,\cdot}. \quad (3.24)$$

To znači da transformacijsku matricu A možemo dobiti adekvatnim odabirom n redaka matrice Y .

Iz [16, Lema 3.13, Teorem 3.14] imamo ekvivalenciju sljedećih tvrdnji:

- Konveksna ljuska matrice Y je simpleks.
- Ako je A matrica dobivena algoritmom 3.26 koji implementira teoriju ove podsekcije, tada je ona dopustiva, odnosno $A \in F_A$.
- $A \in v(F_A)$ i stoga A maksimizira funkciju cilja I_1 iz čega slijedi da je zadovoljen uvjet maksimalnog skaliranja.

Tada je rješenje A jedinstveno do na permutaciju stupaca. To slijedi iz činjenice da je A^{-1} jedinstvena do na permutaciju redaka koji predstavljaju vrhove simpleksa čiji redosljed odabira nije bitan.

Za X općenito vrijedi da je uvjet maksimalnog skaliranja skoro zadovoljen. Tada je konveksna ljuska matrice Y blizu simpleksu, a algoritam 3.26 daje matricu koja je blizu rješenju.

Napomena 3.4.11. Neka je $I = \{ind_1, \dots, ind_n\}$ skup indeksa redova matrice Y koji su aproksimacija vrhova simpleksa Δ_2 . Tada

$$Y_{ind_j, \cdot} \approx v_j \xrightarrow{A} e_j$$

za svaki $j = 1, \dots, n$. Fiksirajmo $j \in \{1, \dots, n\}$. Vrijedi $\max_{i=1, \dots, N} X_{i,j} = X_{ind_j, i}$ pa funkciju I_1 možemo napisati u obliku

$$I_1(A) = \sum_{j=1}^n \sum_{i=1}^n Y_{ind_j, i} A_{i,j}.$$

3.5 Algoritam

Algoritam 3.25 (Algoritam preslikavanja indeksa).

Među N redaka matrice Y tražimo n vrhova simpleksa Δ_2 . Za početni vrh ima smisla odabrati onaj najudaljeniji od ishodišta. Zatim biramo točku najudaljeniju od odabranog vrha i nju uzimamo za drugi vrh. Konstruirana je hiperravnina minimalne dimenezije. Idući vrh je točka maksimalno udaljena od dobivene hiperravnine, a tražimo ga pomoću ortogonalizacije. Konstruira se nova hiperravnina određena i tim dodanim vrhom. Postupak se ponavlja dok ne dobijemo svih n vrhova.

Input: matrica $Y \in \mathbb{R}^{N \times n}$ definirana u jednakostima (3.16)

Output: vektor indeksa ind duljine n

1. Pronađi $k \in \{1, \dots, N\}$ za koji je $\|Y(k, :)\|_2$ maksimalna.
2. Postavi $ind(1) = k, v_1 = Y(ind(1), :)$.
3. Za $i = 1, \dots, N$ postavi $Y(i, :) \leftarrow Y(i, :) - v_1$.
4. Za $j = 2, \dots, n$ pronadži daljnje vrhove:
 - 4.1. Pronađi $k \in \{1, \dots, N\}$ za koji je $\|Y(k, :)\|_2$ maksimalna.
 - 4.2. Postavi $ind(j) = k$.

4.3. Postavi $Y \leftarrow Y/\|Y(ind(j), :)\|_2$.

4.4. Postavi $v_j = Y(ind(j), :)$.

4.5. Za $i = 1, \dots, N$ postavi $Y(i, :) \leftarrow Y(i, :) - v_j Y(i, :)^T v_j$.

Algoritam 3.26 (Inicijalna točka za optimizaciju).

Za iterativni postupak optimizacije bez ograničenja potrebno je odabrati inicijalnu matricu A_0 . Prema formuli (3.24), to možemo napraviti na sljedeći način.

Input: vektor indeksa ind dobiven algoritmom 3.25, matrica $Y \in \mathbb{R}^{N \times n}$

Output: inicijalna matrica $A_0 \in \mathbb{R}^{n \times n}$

1. Izračunaj matricu A_0^{-1} pomoću jednakosti (3.24):

$$A_0^{-1} = \begin{bmatrix} Y(ind(1), 1) & \dots & Y(ind(1), n) \\ \vdots & \ddots & \vdots \\ Y(ind(n), 1) & \dots & Y(ind(n), n) \end{bmatrix}.$$

2. Izračunaj A_0 .

Algoritam 3.27 (Transformacija u dopustivu točku).

Neka je $M \in \mathbb{R}^{n \times n}$ proizvoljna matrica. Cilj ovog algoritma je dobiti projekciju matrice M na dopustivi skup $F'_A \subset F_A$. Koraci 1 i 2 prilagođavaju matricu M uvjetima (3.21) i (3.22), a korak 3 osigurava da suma prvog retka matrice bude 1 zbog uvjeta (3.21) za $i = 1$.

Input: proizvoljna matrica $M \in \mathbb{R}^{n \times n}$

Output: dopustiva matrica $A \in \mathbb{R}^{n \times n}$

1. $M(j, 1) := -\sum_{i=2}^n M(j, i)$ za $j = 2, \dots, n$.

2. $M(1, i) := -\min_{l=1, \dots, N} \sum_{j=2}^n M(j, i) Y(l, j)$ za $i = 1, \dots, n$.

3. $\gamma := \sum_{i=1}^n M(1, i)$

$A(i, j) := M(i, j)/\gamma$ za $i, j = 1, \dots, n$.

Algoritam 3.28 (PCCA+).

PCCA+ algoritam implementira sve prethodno navedene algoritme i na temelju matrice Y definirane u jednakostima (3.16) računa transformacijsku matricu A . Primjenom algoritma 3.27 u koraku 2 korištena je nediferencijabilna rutina. Zato Deuffhard i Weber za lokalnu optimizaciju preporučuju korištenje metode bez derivacija.

Input: matrica $Y \in \mathbb{R}^{N \times n}$

Output: dopustiva transformacijska matrica A

1. Pomoću algoritma 3.25 i 3.26 izračunaj početnu iteraciju za lokalnu optimizaciju A_0 .
2. Izvedi iterativnu lokalnu optimizaciju s funkcijom I_1 ili I_2 počevši s A_0 . U svakom koraku lokalne optimizacije $A_k \rightarrow A_{k+1}$ zamijeni samo elemente $A_k(i, j)$, $i, j \neq 1$, a za preostale elemente matrice A_{k+1} primjeni algoritam 3.27.
3. Ako je zadovoljen kriterij konvergencije, zaustavi algoritam.

Algoritam 3.29 (GenPCCA).

Ovaj algoritam računa matricu Y prema formuli (3.16) i primjenjuje PCCA+ algoritam koji računa traženu transformacijsku matricu A . Algoritam vraća dobivene matrice Y i A te traženu projekciju $G(P)$ prijelazne matrice P koja se lako dobije pomoću formule (3.18).

Input: prijelazna matrica P , inicijalna distribucija η , broj klastera n

Output: matrica $Y \in \mathbb{R}^{N \times n}$, transformacijska matrica A , projekcija $G(P)$

1. Izračunaj dijagonalnu matricu $D_\eta^{0.5}$.
2. Izračunaj realnu Schurovu dekompoziciju matrice $\tilde{P} = D_\eta^{0.5} P D_\eta^{-0.5}$.
3. Pomoću metode SRSchur, sortiraj Schurovu matricu tako da n Schurovih vrijednosti blizu Perronovog korijena 1 budu u gornjem lijevom dijelu Schurove matrice. Izdvoji vodeće Schurove vektore \tilde{Y} iz Schurove matrice.
4. Izračunaj $Y = D_\eta^{-0.5} \tilde{Y}$.
5. Primjeni PCCA+ algoritam na Y i izračunaj X .

Poglavlje 4

Analiza meta-stabilnih stanja reverzibilnih Markovljevih lanaca - particioniranje grafa

Neka je $G = (V, E, W)$ povezan težinski graf, gdje je V skup vrhova, E skup bridova grafa, a W nenegativna simetrična matrica s težinama bridova koja nam daje informaciju koliko vjerojatno dva vrha pripadaju istom klasteru. Iz primjera 1.1.20 vidimo da graf predstavlja ireducibilan reverzibilan Markovljev lanac sa skupom stanja $V = \{1, \dots, N\}$. Prijelazna matrica P definirana je s $p_{ij} = \frac{w_{ij}}{w_i}$, $i, j = 1, \dots, N$, a stacionarna distribucija $\pi = (\pi_1, \dots, \pi_N)$ s $\pi_i = \frac{w_i}{w}$, $i = 1, \dots, N$, pri čemu je $w_i = \sum_{j \in V} w_{ij}$ i $w = \sum_{i \in V} w_i$.

Skup stanja V želimo podijeliti u n klastera $\Gamma_V^n = \{V_1, \dots, V_n\}$, $V = \cup_{l=1}^n V_l$, $V_k \cap V_l = \emptyset$, $\forall k \neq l$, tako da bude maksimizirana metastabilnost klastera. To radi sljedeća metoda s izvorom iz [13].

4.1 Višeklasni kriteriji podjele

Neka su $A, B \subset V$. Definirat ćemo sljedeće pojmove:

Definicija 4.1.1. *Ukupna težinska veza između skupova A i B dana je s*

$$\text{links}(A, B) = \sum_{i \in A, j \in B} w_{ij}. \quad (4.1)$$

Stupanj skupa A definiramo kao ukupnu težinsku vezu skupa A sa svim vrhovima u V :

$$\text{degree}(A) = \text{links}(A, V). \quad (4.2)$$

Napomena 4.1.2. Normalizirana ukupna težinska veza između A i B dana je s

$$\text{linkratio}(A, B) = \frac{\text{links}(A, B)}{\text{degree}(A)}. \quad (4.3)$$

Od posebnog interesa su nam vrijednosti $\text{linkratio}(A, A)$, koji mjeri povezanost skupa A sa samim sobom, i $\text{linkratio}(A, V \setminus A)$ koji mjeri nepovezanost A sa samim sobom.

Dobro klasteriranje podrazumijeva dobru povezanost unutar svakog pojedinog klastera i slabu povezanost s vanjskim vrhovima. Ta dva svojstva mjere se pomoću:

$$\text{knassoc}(\Gamma_V^n) = \frac{1}{n} \sum_{l=1}^n \text{linkratio}(V_l, V_l), \quad (4.4)$$

$$\text{kncuts}(\Gamma_V^n) = \frac{1}{n} \sum_{l=1}^n \text{linkratio}(V_l, V \setminus V_l). \quad (4.5)$$

Nazivi funkcija dolaze od eng. naziva *K-way normalized associations* i *normalized cuts criteria*.

Napomena 4.1.3. Funkcije knassoc i kncuts su direktno povezane s

$$\text{knassoc}(\Gamma_V^n) + \text{kncuts}(\Gamma_V^n) = 1.$$

Što je vrijednost funkcije knassoc veća, time je vrijednost funkcije kncuts manja i obratno.

Neka je matrica $X = [X_1, \dots, X_n] \in \{0, 1\}^{n \times n}$ definirana kao u formuli (3.1).

Definicija 4.1.4. Matrica stupnjeva za simetričnu matricu težina W dana je kao dijagonalna matrica sa sumama redova matrice W na dijagonali, preciznije

$$D = \text{Diag}(W\mathbf{1}_N), \quad (4.6)$$

pri čemu funkcija Diag vraća dijagonalnu matricu s vektorom u argumentu na dijagonali.

Napomena 4.1.5. U primjeru 1.1.20 vidimo da se radi o invarijantnoj mjeri Markovljevog lanca na dijagonali matrice D .

Kako je već i navedeno, želimo maksimizirati funkciju

$$\epsilon(\Gamma_V^n) := \text{knassoc}(\Gamma_V^n) = \frac{1}{n} \sum_{l=1}^n \frac{\text{links}(V_l, V_l)}{\text{degree}(V_l)}. \quad (4.7)$$

Vrijednosti u brojniku i nazivniku moguće je zapisati na sljedeći način:

$$\text{links}(V_l, V_l) = X_l^T W X_l,$$

$$\text{degree}(V_l) = X_l^T DX_l.$$

Stoga dobivamo sljedeći optimizacijski problem zvan PCNX:

$$\begin{cases} \epsilon(X) = \frac{1}{n} \sum_{l=1}^n \frac{X_l^T W X_l}{X_l^T D X_l} \rightarrow \max \\ X \in \{0, 1\}^{n \times n} \\ X \mathbf{1}_n = \mathbf{1}_n \end{cases}, \quad (4.8)$$

odnosno želimo naći takav X da metastabilnost klastera (definicija 3.2.5) bude maksimalna moguća.

4.2 Relaksirano rješenje problema

Definicija 4.2.1. *Definirajmo skaliranu matricu particije $Z = X(X^T DX)^{\frac{1}{2}}$. Dio $X^T DX$ je dijagonalna matrica, što znači da su stupci matrice Z stupci matrice X skalirani inverzom korijena stupnjeva particija.*

Pojednostavimo formulu (4.8) pomoću matrice Z :

$$\epsilon(X) = \frac{1}{n} \text{trag}(Z^T W Z), \quad (4.9)$$

pri čemu Z prirodno ograničavamo s

$$Z^T Z = (X^T DX)^{-\frac{1}{2}} X^T DX (X^T DX)^{\frac{1}{2}} = I_n.$$

Dakle, konačno imamo novi optimizacijski problem zvan PCNZ:

$$\begin{cases} \epsilon(Z) = \frac{1}{n} \text{trag}(Z^T W Z) \\ Z^T D Z = I_n \end{cases}. \quad (4.10)$$

Relaksacijom Z u neprekidnu domenu, diskretan problem prebacujemo u neprekidan optimizacijski problem.

Propozicija 4.2.2. *Neka je $R \in \mathbb{R}^{n \times n}$. Ako je Z dopustivo rješenje PCNZ programa, tada je i $\{ZR : R^T R = I_n\}$ skup rješenja. Dalje, vrijedi*

$$\epsilon(ZR) = \epsilon(Z).$$

Dokaz. Neka je Z dopustivo rješenje. Tada vrijedi $Z^T D Z = I_n$.

Za ZR slijedi:

$$(ZR)^T D (ZR) = R^T Z^T D Z R = R^T R = I_n.$$

Dalje, uz korištenje činjenice $\text{trag}(AB) = \text{trag}(BA)$,

$$\begin{aligned}\epsilon(ZR) &= \frac{1}{n} \text{trag}((ZR)^T W(ZR)) = \frac{1}{n} \text{trag}(R^T Z^T WZR) = \frac{1}{n} \text{trag}([R^T][Z^T WZR]) \\ &= \frac{1}{n} \text{trag}(Z^T WZRR^T) = \frac{1}{n} \text{trag}(Z^T WZ(R^T R)^T) = \frac{1}{n} \text{trag}(Z^T WZ) \\ &= \epsilon(Z).\end{aligned}$$

□

Napomena 4.2.3. Iz prethodne propozicije slijedi da dopustivo rješenje ostaje dopustivo pri rotaciji, odnosno refleksiji.

Iz primjera 1.1.20 vidimo da je $P = D^{-1}W$ prijelazna matrica Markovljevog lanca. Najveća svojstvena vrijednost stohastičke matrice je 1 s pripadnim jediničnim svojstvenim vektorom $\mathbf{1}_N$ (definicija 1.2.8).

Propozicija 4.2.4. Neka su $V, S \in \mathbb{R}^{n \times n}$, $S = \text{Diag}(s)$ matrice takve da vrijedi

$$PV = VS.$$

pri čemu su svojstvene vrijednosti $s = (s_1, \dots, s_N)$ poredane u nerastući poredak.

Problem svojstvenih vrijednosti možemo prebaciti u simetričnu formu:

$$\begin{aligned}PV = VS &\Leftrightarrow (D^{-\frac{1}{2}} W D^{-\frac{1}{2}}) D^{\frac{1}{2}} V = D^{\frac{1}{2}} VS \\ &\Leftrightarrow (D^{-\frac{1}{2}} W D^{-\frac{1}{2}}) \bar{V} = \bar{V} S, \\ \bar{V}^T \bar{V} &= I_N.\end{aligned}$$

Posljednja jednakost slijedi iz korolara 1.2.18. Vidimo da su V i S realni (korolar 1.2.17) i da bilo kojih n različitih svojstvenih vektora čini kandidata za lokalno optimalno rješenje PCNZ programa:

$$\epsilon([V_{\pi_1}, \dots, V_{\pi_n}]) = \frac{1}{n} \sum_{l=1}^n s_{\pi_l}, \quad (4.11)$$

pri čemu je $\{\pi_1, \dots, \pi_n\} \subseteq \{1, \dots, N\}$.

Iz (4.11) slijedi da se lokalno optimalno rješenje očito postiže za $(\pi_1, \dots, \pi_n) = (1, \dots, n)$:

$$\begin{aligned}Z^* &= [V_1, \dots, V_n] \\ \Lambda^* &= \text{Diag}([s_1, \dots, s_n]).\end{aligned} \quad (4.12)$$

Time je

$$\epsilon(Z^*) = \frac{1}{n} \text{trag}(\Lambda^*) = \max_{Z^T DZ = I_n} .$$

Prema propoziciji prop1, globalno optimalno rješenje PCNZ programa nije jedinstveno, već je to cijeli potprostor razapet s prvih n najvećih svojstvenih vektora od P :

$$\{Z^*R : R^T R = I_n, PZ^* = Z^* \Lambda^*\}. \quad (4.13)$$

Napomena 4.2.5. Z^*R nisu svojstvene vrijednosti od P (osim ako $s_1 = \dots = s_n$).

Imamo optimalno rješenje Z i želimo ga transformirati natrag u prostor particijskih matrica tipa (3.1).

Imamo:

$$\begin{aligned} Z &= f(X) = X(X^T D X)^{-\frac{1}{2}} \Rightarrow \\ X &= f^{-1}(Z) = \text{Diag}(\text{diag}^{-\frac{1}{2}}(ZZ^T))Z. \end{aligned} \quad (4.14)$$

Napomena 4.2.6. Funkcija diag vraća dijagonalu matrice kao vektor stupac.

Promotrimo retke matrice Z kao koordinate N točaka u \mathbb{R}^n . Funkcija f^{-1} normalizira njihove duljine tako da leže na jediničnoj hipersferi.

Koristeći činjenicu $RR^T = I_n$, vrijedi:

$$\begin{aligned} f^{-1}(Z^*R) &= \text{Diag}(\text{diag}^{-\frac{1}{2}}(Z^*R(Z^*R^T)))Z^*R = \\ &= \text{Diag}(\text{diag}^{-\frac{1}{2}}(Z^*Z^{*T}))Z^*R = f^{-1}(Z^*)R = \tilde{X}^*R. \end{aligned}$$

Zato su optimalna rješenja elementi prostora

$$\{\tilde{X}^*R : \tilde{X}^* = f^{-1}(Z^*), R^T R = I_n\}. \quad (4.15)$$

Napomena 4.2.7. Iako $Z_1^* = (\mathbf{1}_N^T D \mathbf{1}_N)^{-\frac{1}{2}}$ ima svojstvo da $Z_1^* = c \cdot \mathbf{1}_N$, $c \in \mathbb{R}$, no to ne treba vrijediti za X_1^* kada je $n > 1$.

4.3 Optimalno diskretno rješenje

Napomena 4.3.1. Općenito, optimalno neprekidno rješenje iz $\{\tilde{X}^*R : \tilde{X}^* = f^{-1}(Z^*), R^T R = I_n\}$ nije u dopustivom skupu PCNX programa, ali možemo naći "blisko" diskretno rješenje koje zadovoljava uvjete originalnog programa PCNX.

Napomena 4.3.2. Frobeniusova norma realne matrice A dana je s

$$\|A\|_F = \sqrt{\text{tr}(AA^T)}.$$

Teorem 4.3.3. *Neka je $\tilde{X}^* = f^{-1}(Z^*)$. Optimalna diskretna particija X^* je ona koja zadovoljava sljedeći program zvan POD:*

$$\begin{cases} \phi(X, R) = \|X - \tilde{X}^* R\|_F^2 \rightarrow \min \\ X \in \{0, 1\}^{n \times n}, \quad X \mathbf{1}_n = \mathbf{1}_N \\ R^T R = I_n \end{cases} . \quad (4.16)$$

Lokalno optimalno rješenje (X^, R^*) ovog bilinearnog programa može se naći iterativno:*

- *Pretpostavimo da imamo R^* . POD je reduciran na program PODX u X :*

$$\begin{cases} \phi(X) = \|X - \tilde{X}^* R^*\|_F^2 \rightarrow \min \\ X \in \{0, 1\}^{n \times n}, \quad X \mathbf{1}_n = \mathbf{1}_N \end{cases} .$$

Označimo $\tilde{X} = \tilde{X}^ R^*$. Optimalno rješenje je dano s:*

$$X^*(i, l) = \langle l = \operatorname{argmax}_{k=1, \dots, n} \tilde{X}(i, k) \rangle, \quad i \in V, \quad (4.17)$$

pri čemu je $\langle \cdot \rangle$ jednako 1 ukoliko je argument istinit, a 0 u suprotnom.

1. *Sada pretpostavimo da imamo X^* . POD je reduciran na PODR u R :*

$$\begin{cases} \phi(X) = \|X^* - \tilde{X}^* R\|_F^2 \rightarrow \min \\ R^T R = I_n \end{cases} .$$

Optimalno rješenje je dano s:

$$R^* = \tilde{U} U^T, \quad (4.18)$$

pri čemu su \tilde{U} i U singularni vektori:

$$X^{*T} \tilde{X}^* = U \Omega U^T, \quad \Omega = \operatorname{Diag}(\omega),$$

*pri čemu je (U, Ω, \tilde{U}) SVD od $X^{*T} \tilde{X}^*$ i vrijedi $U^T U = I_n$, $\tilde{U}^T \tilde{U} = I_n$ i $\omega_1 \geq \dots \geq \omega_n$.*

Dokaz. Vrijedi:

$$\begin{aligned} \phi(X, R) &= \|X - \tilde{X}^* R\|_F^2 = \operatorname{trag}((X - \tilde{X}^* R)(X - \tilde{X}^* R)^T) = \operatorname{trag}((X - \tilde{X}^* R)(X^T - R^T \tilde{X}^{*T})) \\ &= \operatorname{trag}(XX^T - XR^T \tilde{X}^{*T} - \tilde{X}^* R X^T + \tilde{X}^* R R^T \tilde{X}^{*T}) = \operatorname{trag}(XX^T) + \operatorname{trag}(\tilde{X}^* \tilde{X}^{*T}) \\ &= \|X\|_F^2 + \|\tilde{X}^*\|_F^2 - \operatorname{trag}(XR^T \tilde{X}^{*T} - \tilde{X}^* R X^T) = 2N - 2\operatorname{trag}(XR^T \tilde{X}^{*T}). \end{aligned}$$

*Slijedi da je za minimizaciju funkcije (X, R) dovoljno maksimizirati $\operatorname{trag}(XR^T \tilde{X}^{*T})$.*

PODX Neka je $R = R^*$. Maksimiziramo $\text{trag}(XR^{*T}\tilde{X}^{*T})$. Svaki element u $\text{diag}(XR^T\tilde{X}^{*T})$ može se samostalno optimizirati. Kao rezultat dobivamo (4.17).

PODR Neka je $X = X^*$. Definirajmo Lagrangeovu funkciju sa simetričnim multiplikatorom Λ :

$$L(R, \Lambda) = \text{trag}(X^*R^T\tilde{X}^{*T}) - \text{trag}(\Lambda^T(R^T R - I_n)).$$

Optimalno rješenje mora zadovoljavati:

$$\begin{aligned} L_R(R, \Lambda) &= \tilde{X}^{*T}X^* - R\Lambda = 0 \\ \Leftrightarrow \tilde{X}^{*T}X^* &= R\Lambda \Leftrightarrow \Lambda^* = R^{*T}\tilde{X}^{*T}X^* \\ \text{i } L_\Lambda(R, \Lambda) &= R^T R - I_n = 0. \end{aligned} \quad (4.19)$$

Imamo:

$$\begin{aligned} \Lambda^{*T}\Lambda^* &= X^{*T}\tilde{X}^*R^*R^{*T}\tilde{X}^{*T}X^* \\ &= X^{*T}\tilde{X}^*(X^{*T}\tilde{X}^{*T})^T \\ &= U\Omega\tilde{U}\tilde{U}^T\Omega^T U^T \\ &= U\Omega^2 U^T \end{aligned}$$

pri čemu je $U\Omega\tilde{U}$ SVD od $X^{*T}\tilde{X}^*$. Jer vrijedi $\Lambda = \Lambda^T$, imamo $\Lambda^* = U\Omega U^T$.

Iz jednakosti (4.19) slijedi:

$$\begin{aligned} R^*\Lambda^* &= \tilde{X}^{*T}X^* \\ R^* &= \tilde{X}^{*T}X^*\Lambda^{*-1} = \tilde{U}^T\Omega U^T(U\Omega U^T)^{-1} = \tilde{U}^T\Omega U^T U\Omega^{-1}U^T = \tilde{U}^T U^T \end{aligned}$$

Vrijedi $R^{*T}R = U\tilde{U}\tilde{U}^T U^T = I_n$.

Imamo:

$$\begin{aligned} \text{trag}(X^*R^{*T}\tilde{X}^{*T}) &= \text{trag}(R^*U\Omega\tilde{U}) = \text{trag}(\tilde{U}U^T U\Omega\tilde{U}) \\ &= \text{trag}(\tilde{U}\Omega\tilde{U}) = \text{trag}((\tilde{U}^T)^T(\Omega\tilde{U})) = \text{trag}(\tilde{U}^T\tilde{U}^T\Omega) = \text{trag}(\Omega). \end{aligned}$$

Iz toga slijedi:

$$\phi(R^*) = 2N - 2\text{trag}(\Omega).$$

Vidimo da što je veći $\text{trag}(\Omega)$, to je X^* bliže \tilde{X}^*R^* .

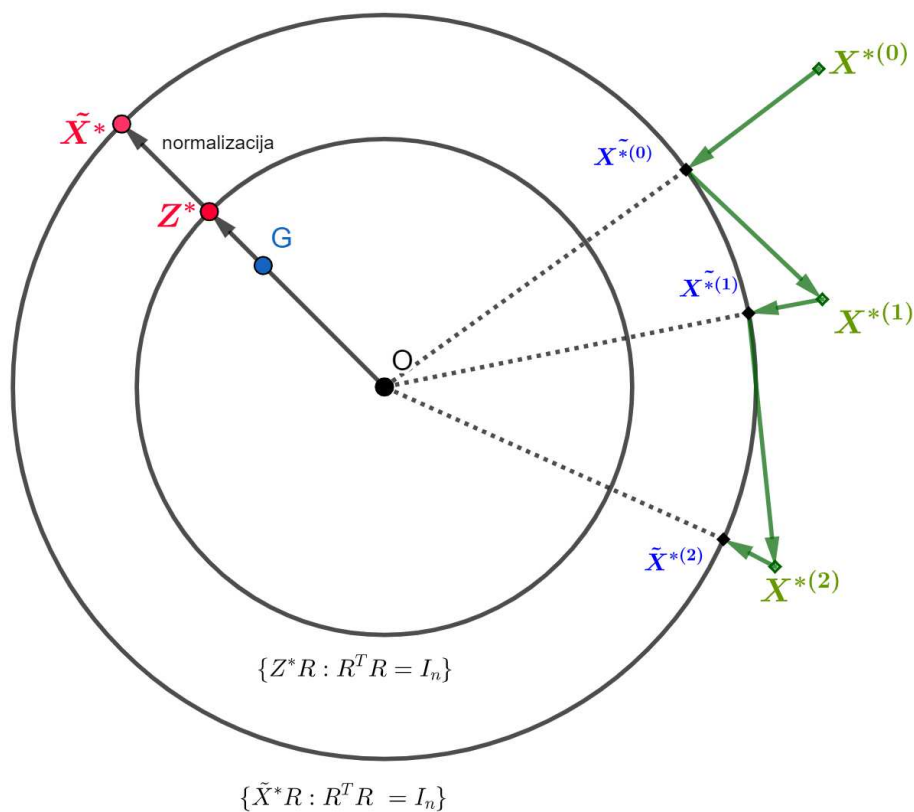
□

Svaki element prostora $\{\tilde{X}^*R^* : \tilde{X}^* = f^{-1}(Z^*), R^T R = I_n\}$ je optimalno rješenje PCNZ programa, neovisno o odabiru R . Stoga, cilj je naći diskretno optimalno rješenje programa PCNX dovoljno blizu bilo kojem elementu tog prostora. Za početak biramo proizvoljni

početni X^* u $\{0, 1\}^{n \times n}$ i zatim rješavanjem PODR i PODX programa postupno doći do onog najbližem danom prostoru neprekidnih rješenja.

Jedna od opcija je inicijalizirati "gotovo ortogonalni" R^* odabirom n "što ortogonalnijih" redova matrice \tilde{X}^* . Matricu X^* dobijemo zatim pomoću formule (4.17) što je zapravo obično klasteriranje točaka (redova matrice $\tilde{X}^* R^*$) s n centara jedinične duljine.

Dalje, riješimo PODR program tako da nađemo optimum $\tilde{X}^* R^*$ najbliži danom početnom X^* . Zatim dobivenom optimalnom $\tilde{X}^* R^*$ tražimo najbliže diskretno rješenje PCNX programa. Postupak je grafički prikazan na slici 4.1.



Slika 4.1: Grafički opis algoritma

4.4 Algoritam

Input: simetrična matrica težina W , broj klasa n

Output: particijska matrica X^*

1. Izračunaj matricu stupnjeva $D = \text{Diag}(W\mathbf{1}_N)$.
2. Nađi Z^* kao rješenje problema svojstvenih vrijednosti:

$$PZ^*(:, 1:n) = Z^*(:, 1:n)S(:, 1:n)$$

3. Normaliziraj Z^*

$$\tilde{X}^* = \text{Diag}(\text{diag}^{-\frac{1}{2}}(Z^*Z^{*T}))Z^*$$

4. Izračunaj "skoro ortogonalni" R^* :

Odaberi proizvoljan $i \in \{1, \dots, N\}$

$$\Rightarrow R^*(:, 1) = [\tilde{X}^*(i, 1), \dots, \tilde{X}^*(i, n)]^T$$

Za $k = 2, \dots, n$:

$$c = c + \text{abs}(\tilde{X}^*R_{k-1}^*)$$

$$i = \text{argmin}(c)$$

$$R^*(:, k) = [\tilde{X}^*(i, 1), \dots, \tilde{X}^*(i, n)]^T$$

5. Nađi optimalan X^* :

$$\tilde{X} = \tilde{X}^*R^*$$

$$X^*(i, l) = \left\langle l = \text{argmax}_{k \in [n]} \tilde{X}(i, k) \right\rangle, \quad i \in V, \quad l \in \{1, \dots, n\}$$

6. Nađi optimalan R^* :

$$(U, \Omega, \tilde{U}) = \text{SVD}(X^{*T}\tilde{X}^*)$$

$$r_2 = \text{tr}(\Omega)$$

Ako $|r_2 - r| < \epsilon$:

vрати X^*

Inače:

$$r = r_2$$

$$R^* = \tilde{U}U^T$$

7. Vрати se na 5. korak

Poglavlje 5

Primjeri obrađenih metoda

Obrađeni algoritmi bit će testirani na umjetno konstruiranoj matrici malih dimenzija i dvjema matricama koje pripadaju molekuli n-pentana na različitim temperaturama.

Kod primjera malih dimenzija bit će moguće detaljnije prikazati rezultate i proučiti rad funkcija. Kod preostala dva primjera bit će dani samo glavni rezultati.

5.1 Konstruirani primjer malih dimenzija

Za jednostavniji primjer uzet ćemo blok-dijagonalno dominantnu matricu A i pretpostaviti da je to prijelazna matrica nekog Markovljevog lanca sa skupom stanja S koje sadrži sedam stanja. Dominantni blokovi su istaknuti različitim bojama i predstavljaju klustere stanja - dva klastera sadrže dva stanja, a treći preostala tri stanja.

$$A = \begin{bmatrix} 0.2018 & 0.7752 & 0.0003 & 0.0093 & 0.0048 & 0.0064 & 0.0022 \\ 0.3950 & 0.5888 & 0.0027 & 0.0003 & 0.0043 & 0.0016 & 0.0073 \\ 0.0093 & 0.0072 & 0.2911 & 0.2949 & 0.3938 & 0.0012 & 0.0025 \\ 0.0064 & 0.0038 & 0.1947 & 0.1947 & 0.5580 & 0.0048 & 0.0049 \\ 0.0003 & 0.0063 & 0.1034 & 0.1028 & 0.7713 & 0.0092 & 0.0067 \\ 0.0082 & 0.0016 & 0.0067 & 0.0076 & 0.0027 & 0.6762 & 0.2970 \\ 0.0089 & 0.0068 & 0.0030 & 0.0018 & 0.0065 & 0.1014 & 0.8715 \end{bmatrix}$$

Na matricu A primjenimo odabranu permutaciju P i dobivamo matricu $B = PBP^T$ na kojoj će biti testiran program. Očekujemo da, primjenom algoritma, iz matrice B ponovno

izvučemo blokove istaknute u matrici A .

$$B = \begin{bmatrix} 0.2911 & 0.0093 & 0.0072 & 0.0012 & 0.3938 & 0.0025 & 0.2949 \\ 0.0003 & 0.2018 & 0.7752 & 0.0064 & 0.0048 & 0.0022 & 0.0093 \\ 0.0027 & 0.3950 & 0.5888 & 0.0016 & 0.0043 & 0.0073 & 0.0003 \\ 0.0067 & 0.0082 & 0.0016 & 0.6762 & 0.0027 & 0.2970 & 0.0076 \\ 0.1034 & 0.0003 & 0.0063 & 0.0092 & 0.7713 & 0.0067 & 0.1028 \\ 0.0030 & 0.0089 & 0.0068 & 0.1014 & 0.0065 & 0.8715 & 0.0018 \\ 0.1947 & 0.0064 & 0.0038 & 0.0048 & 0.5880 & 0.0049 & 0.1974 \end{bmatrix}$$

Pomoću MATLAB funkcije `eigs` analiziramo spektar matrice B i promatramo koliko svojstvenih vrijednosti se nalazi u Perronovom klasteru. Nije potrebno računati sve svojstvene vrijednosti jer je očekivano da će tek manji broj njih pripadati Perronovom klasteru. To u ovom primjeru nije značajno, ali je ušteda kod matrica većih dimenzija.

Prve četiri svojstvene vrijednosti su 1.0000, 0.9707, 0.9621, 0.5746 iz čega slijedi da je Perronov klaster dimenzije 3. Zaključak odgovara početnoj matrici A . Zato pretpostavljamo da će algoritmi dijeliti stanja ovog lanca u tri klastera.

Primjena - SVD pristup

Koristimo funkciju danu u dodatku A.1. Ulazni parametri su prijelazna matrica B , dimenzija prijelazne matrice $nB = 7$, prag blizu vrijednosti jedan $r = 0.97$, početne vrijednosti za broj blokova i dimenzije blokova $m = 0$ i $n = 0$, pomoćne vrijednosti funkcije $a = 1$ i $b = 7$, te početna jedinična permutacijska matrica $P = I_7$. Funkcija vraća broj blokova m , dimenzije blokova n i permutacijsku matricu čijom primjenom na matricu B dobivamo traženu blok-dijagonalno dominantnu matricu.

S obzirom na to kako je konstruirana matrica B , očekivano je da funkcija rezultira s 3 bloka na dijagonali, veličina 2×2 , 3×3 , 2×2 , ne nužno tim redoslijedom.

Dobivene izlazne vrijednosti funkcije su broj blokova $m = 3$, dimenzije blokova $n = [2, 2, 3]$ i permutacijska matrica

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

Rezultirajuća matrica odgovara očekivanome. Blokovi su isti kao u matrici A , samo u drugačijem redosljedu i iznutra permutiranih elemenata.

$$P * B * P' = \begin{bmatrix} 0.8715 & 0.1014 & 0.0068 & 0.0089 & 0.0065 & 0.0018 & 0.0030 \\ 0.2970 & 0.6762 & 0.0016 & 0.0082 & 0.0027 & 0.0076 & 0.0067 \\ 0.0073 & 0.0016 & 0.5888 & 0.3950 & 0.0043 & 0.0003 & 0.0027 \\ 0.0022 & 0.0064 & 0.7752 & 0.2018 & 0.0048 & 0.0093 & 0.0003 \\ 0.0067 & 0.0092 & 0.0063 & 0.0003 & 0.7713 & 0.1028 & 0.1034 \\ 0.0049 & 0.0048 & 0.0038 & 0.0064 & 0.5880 & 0.1974 & 0.1947 \\ 0.0025 & 0.0012 & 0.0072 & 0.0093 & 0.3938 & 0.2949 & 0.2911 \end{bmatrix}.$$

Ako su stanja Markovljevog lanca promatramo u poretku kao u matrici B , stanja $\{1, 2, 3, 4, 5, 6, 7\}$ su klasterirana na sljedeći način: $\{6, 4\}, \{3, 2\}, \{1, 5, 7\}$.

Stohastička matrica veza W_1 je očekivano dijagonalno dominantna:

$$W_1 = \begin{bmatrix} 0.9731 & 0.0128 & 0.0142 \\ 0.0087 & 0.9804 & 0.0109 \\ 0.0097 & 0.0111 & 0.9792 \end{bmatrix}.$$

Primjena - GenPCCA

Koristimo funkciju danu u dodatku A.2. Cilj prethodne metode je bio vratiti permutiranu ulaznu matricu tako da ona bude blok-dijagonalno dominantna, te stohastičku matricu veza između klastera, odnosno vjerojatnosti prelaska sustava iz jednog klastera stanja u drugi. Metoda GenPCCA vraća matricu članstva \mathcal{X} i prijelaznu matricu novog Markovljevog lanca $G(B)$ čije vrijednosti možemo usporediti s matricom veza iz prethodne metode.

Ulazni parametri su prijelazna matrica B , početna distribucija y za koju ćemo uzeti stacionarnu distribuciju lanca $y = [0.0672, 0.1215, 0.2332, 0.0626, 0.2958, 0.1556, 0.0641]$, te broj klastera $n = 3$.

Rezultati funkcije su

$$\mathcal{X} = \begin{bmatrix} 0.0069 & 0.0018 & 0.9912 \\ 0.0148 & 0.9837 & 0.0016 \\ 0.0069 & 0.9832 & 0.0098 \\ 0.9883 & 0.0018 & 0.0098 \\ 0.0000 & 0.0154 & 0.9846 \\ 0.9772 & 0.0228 & -0.0000 \\ 0.0116 & -0.0000 & 0.9884 \end{bmatrix},$$

$$G(B) = \begin{bmatrix} 0.9752 & 0.0149 & 0.0099 \\ 0.0056 & 0.9772 & 0.0172 \\ 0.0081 & 0.0115 & 0.9804 \end{bmatrix}.$$

Slijedi, kao i kod prethodne metode, da su stanja najvjerojatnije klasterirana na sljedeći način: {6, 4}, {3, 2}, {1, 5, 7}. Prijelazna matrica $G(B)$ je dijagonalno dominantna i ne odstupa previše od matrice veza W_1 .

Primjena - particioniranje grafa

Koristimo funkciju danu u dodatku A.3. Ova metoda vraća samo matricu članstva X i to baziranu na diskretnom principu klasteriranja s nulama i jedinicama. Ovu metodu možemo svejedno usporediti s prethodnom, odnosno matricu X s matricom \mathcal{X} .

Ulazne vrijednosti funkcije su prijelazna matrica B i broj klastera $n = 3$.

Rezultat funkcije je:

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

Redoslijed klastera u matricama \mathcal{X} i X je različit, ali dobivamo isti rezultat - metoda je, kao i u oba prethodna slučaja, stanja klasterirala na sljedeći način: {6, 4}, {3, 2}, {1, 5, 7}.

5.2 Primjer molekule n-pentana Ph300

Analizirat ćemo prijelaznu matricu stanja molekule n-pentana Ph300. Matrica je dimenzije 255×255 . Originalna matrica nam ne odaje blok-dijagonalno dominantnu strukturu.

Izračunom određenog broja najvećih svojstvenih vrijednosti, uočavamo da je dimenzija Perronovog klastera sedam s vrijednostima 1.0000, 0.9859, 0.9840, 0.9820, 0.9751, 0.9414 i 0.9380. Sljedeća najveća svojstvena vrijednost bi bila 0.5986 i ona očito ne pripada Perronovom klasteru.

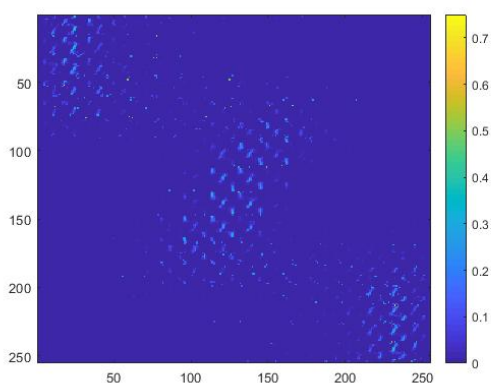
Slijedi da je skup stanja pripadnog Markovljevog lanca potrebno klasterirati u sedam klastera.

Matrice dobivene u ovom primjeru su, za razliku od prethodnog primjera, prevelike za prikaz i koristit ćemo slikovni prikaz pomoću boja.

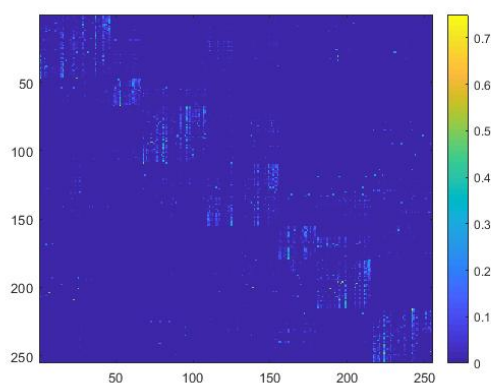
Primjena - SVD pristup

Koristimo funkciju danu u dodatku A.1. Ulazni parametri su prijelazna matrica B , dimenzija prijelazne matrice $nB = 255$, prag blizu vrijednosti jedan $r = 0.5$, početne vrijednosti za broj blokova i dimenzije blokova $m = 0$ i $n = 0$, pomoćne vrijednosti funkcije $a = 1$ i $b = 255$, te početna jedinična permutacijska matrica $P = I_{255}$. Funkcija vraća broj blokova m , dimenzije blokova n i permutacijsku matricu čijom primjenom na matricu prijelaznu matricu dobivamo traženu blok-dijagonalno dominantnu matricu.

Dobivene izlazne vrijednosti funkcije su broj blokova $m = 7$, dimenzije blokova $n = [47, 20, 42, 46, 24, 36, 40]$ i permutacijska matrica. Na slici 5.13 jasno se vidi sedam dijagonalno-dominantnih blokova, odnosno klastera stanja.



Slika 5.1: Ph300 - originalna matrica



Slika 5.2: Ph300 - permutirana matrica

Pripada dijagonalno dominantna matrica veza glasi:

$$W_1 = \begin{bmatrix} 0.8027 & 0.0376 & 0.0023 & 0.0984 & 0.0016 & 0.0361 & 0 \\ 0.0665 & 0.8178 & 0.1110 & 0.0047 & 0 & 0 & 0 \\ 0.0020 & 0.0160 & 0.8478 & 0.0883 & 0 & 0.0016 & 0.0443 \\ 0.0528 & 0.0008 & 0.0473 & 0.6691 & 0.0145 & 0.1208 & 0.0948 \\ 0.0208 & 0 & 0 & 0.0117 & 0.6590 & 0.2315 & 0.0770 \\ 0.0671 & 0 & 0.0011 & 0.0635 & 0.0539 & 0.8141 & 0.0004 \\ 0.0108 & 0 & 0.0303 & 0.0402 & 0.0664 & 0.0076 & 0.8448 \end{bmatrix}$$

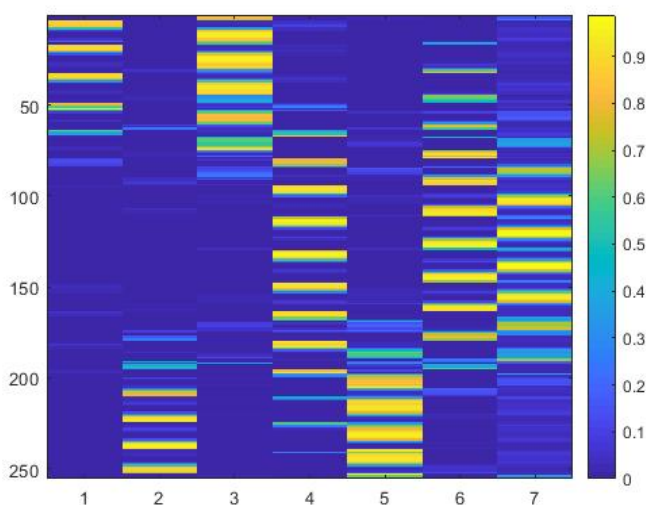
Primjena - GenPCCA

Koristimo funkciju danu u dodatku A.2.

Ulazni parametri su prijelazna matrica P , početna distribucija y za koju ćemo uzeti stacionarnu distribuciju lanca, te broj klastera $n = 7$.

Metoda GenPCCA vraća matricu članstva X i prijelaznu matricu novog Markovljevog lanca $G(P)$.

Rezultati funkcije su:



Slika 5.3: Ph300 - matrica članstva X

$$G(P) = \begin{bmatrix} 0.9449 & 0.0001 & 0.0281 & 0.0263 & -0.0000 & 0.0001 & 0.0006 \\ 0.0001 & 0.9418 & 0.0001 & -0.0000 & 0.0296 & 0.0268 & 0.0015 \\ 0.0027 & 0.0001 & 0.9820 & 0.0001 & 0.0012 & 0.0003 & 0.0135 \\ 0.0024 & 0.0001 & 0.0001 & 0.9839 & 0.0002 & 0.0009 & 0.0123 \\ 0.0001 & 0.0033 & 0.0011 & 0.0001 & 0.9820 & -0.0000 & 0.0135 \\ 0.0001 & 0.0027 & 0.0002 & 0.0008 & 0.0001 & 0.9848 & 0.0114 \\ 0.0001 & 0.0001 & 0.0032 & 0.0030 & 0.0035 & 0.0031 & 0.9869 \end{bmatrix}$$

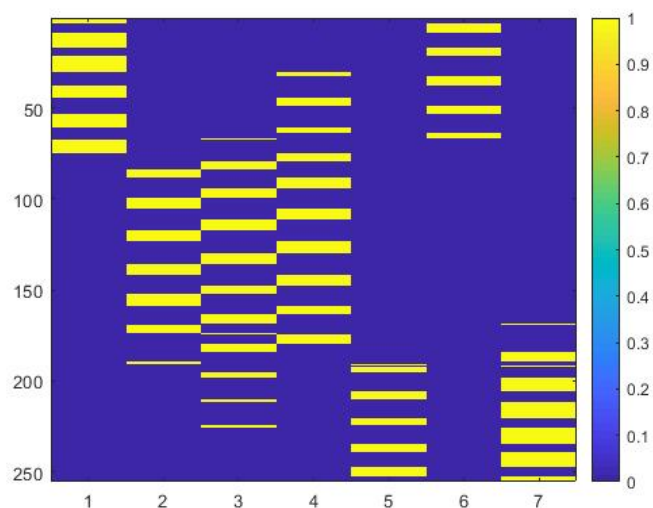
Na prikazu matrice članstva jasno se vidi kako su otprilike klasterirana stanja danog Markovljevog lanca. Nova matrica prijelaza je očekivano dijagonalno dominantna.

Primjena - particioniranje grafa

Koristimo funkciju danu u dodatku A.3.

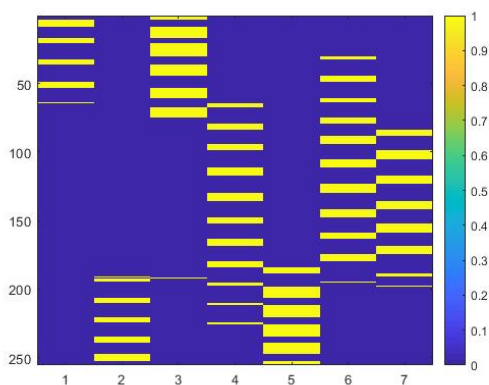
Ulazni parametri su prijelazna matrica P i broj klastera $n = 7$. Metoda vraća samo matricu članstva X , baziranu na diskretnom principu klasteriranja s nulama i jedinicama.

Rezultat funkcije glasi:

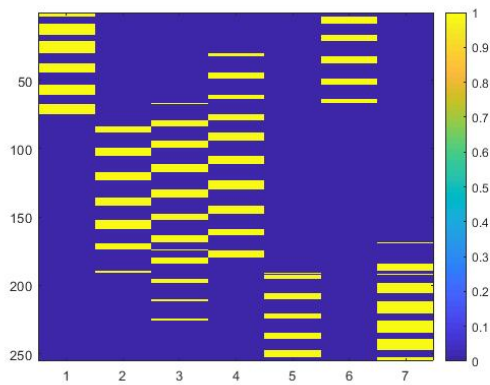


Slika 5.4: Ph300 - matrica članstva X

Prikaz matrice članstva jasno nam odaje podjelu stanja po klasterima. Usporedimo li s matricom članstva dobivene prethodnom metodom i uzmemo li u obzir drugačiji poredak klastera, vidimo da su obje metode dale isti rezultat - dakako, ova metoda je stanja klasterirala na diskretan način za razliku od prethode. Zbog toga je i slika matrice X ispala mnogo jasnija. Za jasniju usporedbu, matricu članstva X iz metode GenPCCA smo diskretizirali.



Slika 5.5: Metoda GenPCCA - diskretizirana



Slika 5.6: Metoda graf-pristup

Slika 5.7: Ph300 - matrica članstva X

5.3 Primjer molekule n-pentana Ph500

Prijelazna matrica u ovom primjeru pripada molekuli n-pentana Ph500. Matrica je dimenzije 307×307 . Kao i u prethodnom primjeru, originalna matrica ne odaje blok-dijagonalno dominantnu strukturu.

Prvih šest najvećih svojstvenih vrijednosti glasi 1.0000, 0.9287, 0.9106, 0.9082, 0.8754, 0.8100. Zatim slijede 0.8021 i -0.4413 . Svojstvena vrijednost -0.4413 očito ne pripada Perronovom klasteru. Vrijednost 0.8100 je dosta udaljena od preostalih pet svojstvenih vrijednosti pa za nju možemo odrediti da ne pripada Perronovom klasteru i da je njegova dimenzija pet, odnosno da skup stanja dijelimo na pet klastera.

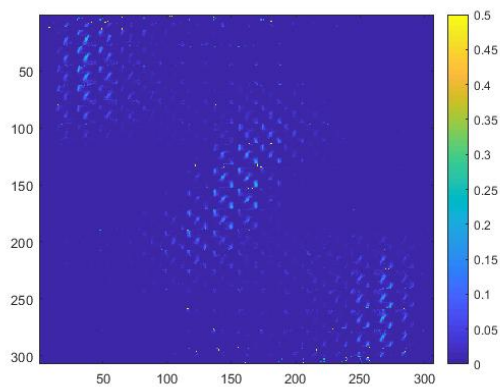
Slijedi da je skup stanja pripadnog Markovljevog lanca potrebno podijeliti u pet klastera.

Primjena - SVD pristup

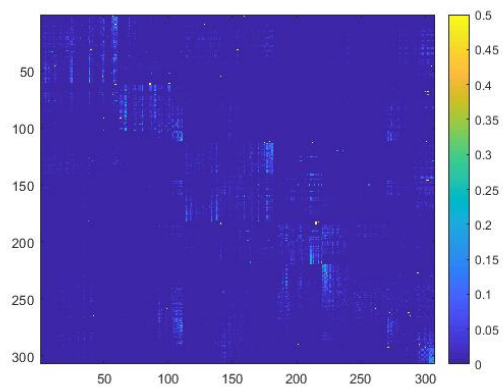
Ponovno koristimo funkciju danu u dodatku A.1. Sada su ulazni parametri prijelazna matrica B , dimenzija prijelazne matrice $nB = 307$, prag blizu vrijednosti jedan $r = 0.5$, početne vrijednosti za broj blokova i dimenzije blokova $m = 0$ i $n = 0$, pomoćne vrijednosti funkcije $a = 1$ i $b = 307$, te početna jedinična permutacijska matrica $P = I_{307}$.

Dobivene izlazne vrijednosti funkcije su broj blokova $m = 5$, dimenzije blokova $n = [60, 51, 71, 37, 88]$ i permutacijska matrica. Na slici 5.9 jasno se vidi sedam dijagonalno-

dominantnih blokova, odnosno klastera stanja.



Slika 5.8: Ph500 - originalna matrica



Slika 5.9: Ph500 - permutirana matrica

Pripada dijagnoalno dominantna matrica veza glasi:

$$W_1 = \begin{bmatrix} 0.6549 & 0.0716 & 0.1411 & 0.0013 & 0.1311 \\ 0.1440 & 0.6725 & 0.0348 & 0.0034 & 0.1452 \\ 0.0718 & 0.0288 & 0.6150 & 0.1016 & 0.1829 \\ 0.0042 & 0.0228 & 0.1336 & 0.5523 & 0.2871 \\ 0.0471 & 0.1606 & 0.0713 & 0.0960 & 0.6250 \end{bmatrix}$$

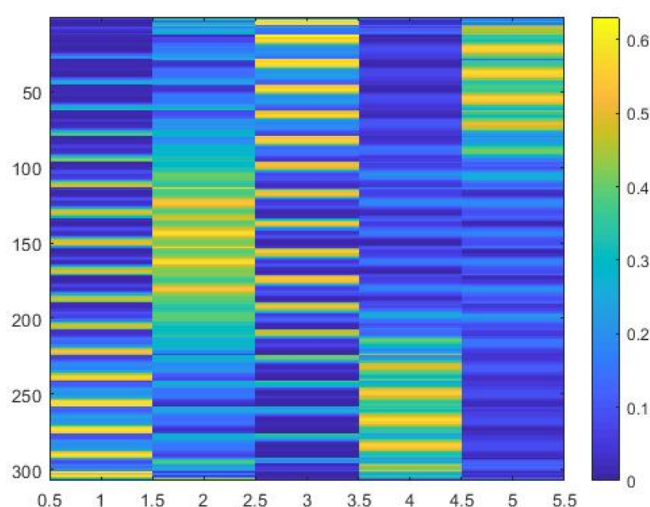
Primjena - GenPCCA

Koristimo funkciju danu u dodatku A.2.

Ulazni parametri su prijelazna matrica Ph500, početna distribucija y za koju ćemo uzeti stacionarnu distribuciju lanca, te broj klastera $n = 5$.

Metoda GenPCCA vraća matricu članstva X i prijelaznu matricu novog Markovljevog lanca $G(P)$.

Rezultati funkcije su:



Slika 5.10: Ph500 - matrica članstva X

$$G(P) = \begin{bmatrix} 0.9190 & 0.0505 & -0.0020 & 0.0202 & 0.0123 \\ 0.0208 & 0.9250 & 0.0212 & 0.0164 & 0.0166 \\ -0.0029 & 0.0499 & 0.9220 & 0.0102 & 0.0208 \\ 0.0157 & 0.0333 & 0.0075 & 0.9284 & 0.0151 \\ 0.0065 & 0.0347 & 0.0164 & 0.0140 & 0.9284 \end{bmatrix}$$

Na prikazu matrice članstva vidi se kako su otprilike klasterirana stanja danog Markovljevog lanca. U prethodnom primjeru s prijelaznom matricom Ph300, metoda je dala jasnije rezultate. Ali promotrimo li prethodnu metodu u ovom primjeru, vidimo da je i ona rezultirala "slabijim" rezultatima, odnosno klasteriranje je obavljeno, ali matrica veza W_1 na dijagonali ima vrijednosti tek nešto veće od 0.5. Metastabilnost stanja ove molekule

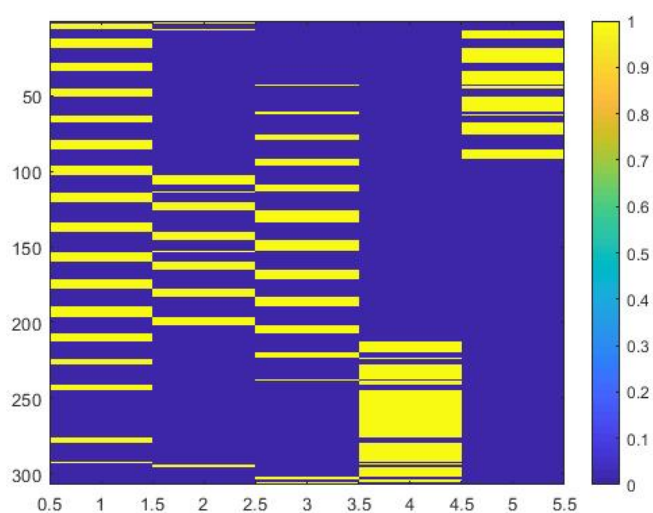
postoji, ali je očito manja nego kod molekule Ph300. Nova matrica prijelaza je očekivano dijagonalno dominantna.

Primjena - particioniranje grafa

Koristimo funkciju danu u dodatku A.3.

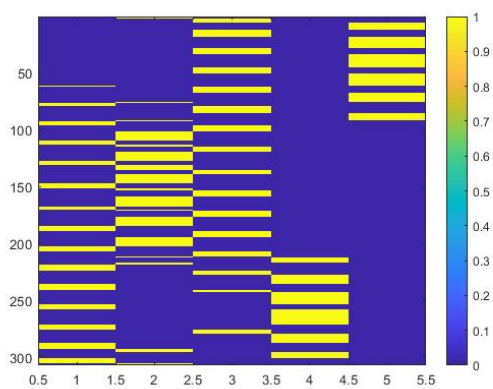
Ulazni parametri su prijelazna matrica Ph500 i broj klastera $n = 5$. Metoda vraća samo matricu članstva X , baziranu na diskretnom principu klasteriranja s nulama i jedinicama.

Rezultat funkcije glasi:

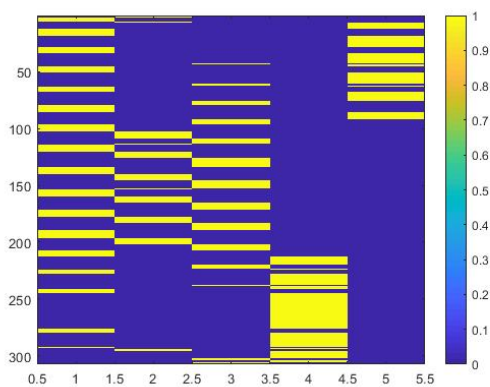


Slika 5.11: Ph500 - matrica članstva X

Prikaz matrice članstva ponovno nam jasno odaje podjelu stanja po klasterima. Sada je malo teže usporediti rezultate ove i prethodne metode, ali vidimo da se klasteri, uzmemo li u obzir drugačiji poredak, uglavnom podudaraju. Za jasniju usporedbu, matricu članstva X iz metode GenPCCA smo diskretizirali.



Slika 5.12: Metoda GenPCCA - diskretizirana



Slika 5.13: Metoda graf-pristup

Slika 5.14: Ph500 - matrica članstva X

Bibliografija

- [1] Jonathan Atkins, Erik Gunnar Boman i Bruce Hendrickson, *A spectral algorithm for seriation and the consecutive ones problem*, SIAM Journal on Computing **28** (1998), br. 1, 297–310.
- [2] Jan Brandts, *Matlab code for sorting real Schur forms*, Numerical linear algebra with applications **9** (2002), br. 3, 249–261.
- [3] Steve Brunton i José Nathan Kutz, *Singular value decomposition*, (2015).
- [4] Kai Lai Chung i Kailai Zhong, *A course in probability theory*, Academic press, 2001.
- [5] Peter Deuffhard, Wilhelm Huisinga, Alexander Fischer i Christof Schütte, *Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains*, Linear Algebra and its Applications **315** (2000), br. 1-3, 39–59.
- [6] Peter Deuffhard i Marcus Weber, *Robust Perron cluster analysis in conformation dynamics*, Linear algebra and its applications **398** (2005), 161–184.
- [7] Petros Drineas, Alan Frieze, Ravi Kannan, Santosh Vempala i Deodhar Vinay, *Clustering in Large Graphs and Matrices.*, SODA, sv. 99, Citeseer, 1999, str. 291–299.
- [8] David Fritzsche, Volker Mehrmann, Daniel Szyld i Elena Virnik, *An SVD approach to identifying meta-stable states of Markov chains*, (2007).
- [9] Mihaela Hrnjkaš, *SVD dekompozicija i primjene*, Disertacija, Josip Juraj Strossmayer University of Osijek. Department of Mathematics, 2018.
- [10] Susanna Kube i Marcus Weber, *A coarse graining method for the identification of transition rates between molecular conformations*, The Journal of chemical physics **126** (2007), br. 2, 024103.
- [11] Ana Mrkojević, *Schurova dekompozicija matrice*, Disertacija, Josip Juraj Strossmayer University of Osijek. Department of Mathematics, 2019.

- [12] Arseny M. Shur, *Detailed Proof of The PerronFrobenius Theorem*, (2016).
- [13] X Yu Stella i Jianbo Shi, *Multiclass spectral clustering*, 2003.
- [14] Ryan Tifenbach, *Issues concerning “An SVD approach to identifying metastable states of Markov chains”*, (2009).
- [15] Zoran Vondraček, *Markovljevi lanci*, (2008).
- [16] Marcus Weber, *Meshless Methods in Confirmation Dynamics*, (2006).
- [17] Dimitrios Zeimpekis i Efstratios Gallopoulos, *TMG: A MATLAB toolbox for generating term-document matrices from text collections*, Grouping multidimensional data, Springer, 2006, str. 187–210.

Sažetak

Ovaj rad obrađuje tri metode za klasteriranje konačnog skupa stanja Markovljevog lanca. Sve tri metode bazirane su na spektralnoj analizi, odnosno Perron-Frobeniusovom klasteru svojstvenih vrijednosti matrice prijelaza. Metode su prvo teorijski obrađene, a zatim testirane na odabranim primjerima.

Kao najbolju metodu izdvojila bih metodu GenPCCA. Za početak, ona daje najviše korisnih informacija - prijelaznu matricu koja pripada novodobivenom Markovljevom lancu i matricu članstva iz koje se jasno iščitava kojem klasteru pripada koje stanje uz vjerojatnosti drugih, manje izglednih mogućnosti. Dalje, čini se i najpreciznijom. Uz analizu spektra i Schurovih vektora, slijedi i postupak optimizacije koji rješenje dodatno približava onom idealnom.

Metoda koja pri analizi pristupa Markovljevom lancu kao grafu je također vrlo precizna. Koristi metode optimizacije uz samu analizu spektra. Nedostatak je da daje i najmanje informacija. Dobivena matrica članstva jednoznačno svakom stanju daje jedinstveni klaster bez informacije o drugim mogućim varijantama, a nemamo novu matricu prijelaza ili nešto slično tome. Rezultati su se uglavnom poklapali s onima dobivenim metodom GenPCCA.

Metoda bazirana na analizi singularnih vrijednosti je u sva tri primjera vrlo lijepo klasterirala stanja i rezultati su se poklapali s onima dobivenim metodom GenPCCA. Također nam daje i matricu veza koja općenito nema Markovljevo svojstvo, ali je korisna informacija. Metoda je vjerojatno najmanje precizna s obzirom da koristi samo spektralnu analizu, bez dodatnih optimizacija.

Summary

This thesis deals with three methods of clustering a finite set of states of the Markov chain. All three methods are based on spectral analysis, i.e., on Perron-Frobenius cluster of eigenvalues of transition matrix. Each method is first theoretically analyzed and then tested on selected examples.

As the best method, I would single out the GenPCCA method. First, it provides more useful information than other methods - the transition matrix of newly acquired Markov chain, the membership matrix which gives us clear information about how likely each state is to belong to a particular cluster. Furthermore, it seems to be the most accurate. In addition to the analysis of the spectrum and Schur vectors, the optimization procedure follows, which brings the solution even closer to the ideal one.

The method that approaches to the Markov chain as a graph is also very precise in the analysis. It uses optimization methods in addition to the spectral analysis itself. However, the disadvantage is that it gives even the least information. The resulting membership matrix uniquely gives each state a unique cluster without information about other possible variants, and we do not get a new transition matrix or anything similar to that. The results mostly coincided with those obtained by the GenPCCA method.

The method based on the analysis of singular values clustered the states very nicely in all three examples and the results coincided with those obtained by the GenPCCA method. It also gives us a coupling matrix that generally does not have a Markov property, but is a useful information. The method is probably intended to be the least precise as it uses only spectral analysis, without additional optimizations.

Životopis

Ivana Blažeka rođena je 30. srpnja 1996. godine u Zagrebu, Hrvatska. Nakon završene V. gimnazije u Zagrebu 2015. godine, upisala je Preddiplomski studij Matematike na Prirodoslovno-matematičkom fakultetu Sveučilišta u Zagrebu. Nakon stečene titule prvostupnice matematike (*univ. bacc. math*), upisuje Diplomski sveučilišni studij Matematička statistika.

Dodatak A

MATLAB kodovi

A.1 Analiza meta-stabilnih stanja reverzibilnih Markovljevih lanaca - SVD pristup

Algoritam

```
1 function [m, n, P] = SVD_metastable(B, nB, a, b, r, m, n, P)
2     %izracun prva dva svojstvena vektora
3     [U,S,V] = svds(B,2);
4     %sortiranje drugog singularnog vektora i kreiranje P
5     [u2, I] = sort(U(:,2));
6     Pp = eye(size(u2,1));
7     Pp = Pp(I,:);
8     %primjena permutacije na matricu B
9     B2 = Pp*B*Pp';
10    %identifikacija dva potencijalna bloka
11    ind = find(diff(sign(u2)));
12
13    D = eye(nB);
14    D(a:b,a:b)=Pp;
15    P = D*P;
16    %izdvojimo dobivene potencijalne blokove
17    B_prvi = B2(1:ind,1:ind);
18    B_drugi= B2(ind+1:size(B,1), ind+1:size(B,1));
19    %racunanje 1-norme blokova
20    normaB1 = sum(sum(B_prvi))/size(B_prvi,1);
21    normaB2 = sum(sum(B_drugi))/size(B_drugi,1);
22    if normaB1 > r & normaB2 > r
23        [m, n, P] = SVD_metastable(B_prvi, nB, a, a+ind-1, r, m, n, P);
24        [m, n, P] = SVD_metastable(B_drugi, nB, ind+a, b, r, m, n, P);
25    return
```



```

26     else
27         m = m + 1;
28         n(m) = size(B,1);
29         return;
30     end
31 end

1 function[W] = coupling_matrix(Q, n, m)
2     k = 0;
3     for i = 1:m
4         l = 0;
5         for j = 1:m
6             W(i,j) = sum(sum(Q(k+1 : k+n(i), l+1 : l+n(j))))/n(i);
7             l = l + n(j);
8         end
9         k = k + n(i);
10    end
11
12 end

```

A.2 Analiza meta-stabilnih stanja nereverzibilnih Markovljevih lanaca - GenPCCA

GenPCCA

```

1 function[Y, GP] = GenPcca(P, y, n)
2
3     %1. korak: izra un dijagonalne matrice i matrice P2
4     y = sqrt(y);
5     D = diag(y);
6     P2 = D*P*(diag(1./y));
7     %2. korak: realna Schurova dekompozicija matrice P2
8     [Q,R] = schur(P2, 'real');
9     %3. korak: sortiranje Shurove dekompozicije i stvaranje matrice X2
10    %prvih n Schurovih vektora
11    [Q,R] = SRSchur(Q,R,1,n);
12    X2(:,1:n) = Q(:,1:n);
13    %4. korak: Izra un X (P*X=X*Del)
14    X = (diag(1./y))*X2;
15    X(:,1) = ones(size(P,1),1);
16    %5. korak: primjena PCCA na X
17    [A] = PCCAplus(X);
18    Y = X*A;
19    GP = inv(Y'*D*Y)*(Y'*D*P*Y);
20 end

```

PCCA+

```
1 function[A] = PCCAplus(X)
2     n = size(X,2);
3     ind = indeksi(X);
4     A0 = pocetna_iteracija(ind, X);
5     A0 = transformacija(A0, X);
6
7     for i = 1:n
8         X2(i,:) = X(ind(i),:);
9     end
10
11     fun = @(z)f(X,X2,z);
12     options = optimset('MaxFunEvals',1);
13     for i=1:500
14         A0 = fminsearch(fun,A0,options);
15         A0 = transformacija(A0,X);
16     end
17     A = transformacija(A0,X);
18
19 end
```

Transformacija u dopustivu točku

```
1 function[A] = transformacija(M, X)
2     [N,n] = size(X);
3     for i = 2:n
4         M(i,1) = -1*sum(M(i,2:n));
5     end
6
7     for j = 1:n
8         G = X(:,2:n); H = M(2:n,j);
9         c = G*H;
10        M(1,j) = -min(c);
11    end
12    gamma = sum(M(1,:));
13    A = M/gamma;
14 end
```

Preslikavanje indeksa

```
1 function[ind] = indeksi(Y)
2     [N,n] = size(Y);
3     for i=1:N
4         c(i) = norm(Y(i,:),2);
5     end
6     [m,k] = max(c);
```

```

7   ind(1) = k;
8   v = Y(k,:);
9   Y = Y - v;
10
11  for j=2:n
12      for i=1:N
13          c(i) = norm(Y(i,:),2);
14      end
15      [m,k] = max(c);
16      ind(j) = k;
17      Y = Y/norm(Y(k,:),2);
18      v = Y(k,:);
19      for i=1:N
20          Y(i,:) = Y(i,:) - v*Y(i,:)'*v;
21      end
22  end
23 end

```

Inicijalna točka za optimizaciju

```

1 function[A0] = pocetna_iteracija(ind, X)
2     n = size(X,2);
3     for i = 1:n
4         for j = 1:n
5             A00(i,j) = X(ind(i),j);
6         end
7     end
8     A0 = inv(A00);
9 end

```

Funkcija cilja

```

1 function[s] = f(X,X2,A)
2     A = transformacija(A,X);
3     s = 0;
4     n = size(A,1);
5     for i = 1:n
6         for j = 1:n
7             s = s + X2(j,i)*A(i,j);
8         end
9     end
10    s=-s;
11 end

```

Sortirani Schurovi vektori

Funkcija SRSchur.m preuzeta je od autora Jan Brands sa sveučilišta u Amsterdamu i javno je dostupna.

```
1 function [Q,R,ap] = SRSchur(Q,R,z,b);
2
3 r = find(abs(diag(R,-1)) > 100*eps);
4 s = 1:size(R,1)+1;
5 s(r+1) = [];
6
7 for k=1:length(s)-1;
8     sk = s(k);
9     if s(k+1)-sk == 2
10        [Q,R] = normalize(Q,R,sk:s(k+1)-1);
11        p(k) = R(sk,sk)+sqrt(R(sk+1,sk)*R(sk,sk+1));
12    else
13        p(k) = R(s(k),s(k));
14    end
15 end
16
17 for k = swaplist(p,s,z,b);
18     v = s(k):s(k+1)-1;
19     w = s(k+1):s(k+2)-1;
20     nrA = norm(R([v,w],[v,w]),inf);
21     [Q,R] = swap(Q,R,v,w);
22     s(k+1) = s(k)+s(k+2)-s(k+1);
23     v = s(k):s(k+1)-1;
24     w = s(k+1):s(k+2)-1;
25     if length(v)==2
26        [Q,R] = normalize(Q,R,v);
27     end
28     if length(w)==2
29        [Q,R] = normalize(Q,R,w);
30     end
31     ap(k) = norm(R(w,v),inf)/(10*eps*nrA);
32 end
33
34 R = R - tril(R,-2);
35 for j=2:length(s)-1; R(s(j),s(j)-1)=0; end
36
37
38 % -----%
39
40 function [U,S] = normalize(U,S,v);
41 n = size(S,1);
42 Q = rot(S(v,v));
43 S(:,v) = S(:,v)*Q;
```

```

44 S(v,:) = Q'*S(v,:);
45 U(:,v) = U(:,v)*Q;
46
47 % -----%
48
49 function Q = rot(X);
50 c = 1; s = 0;
51 if X(1,1)~=X(2,2);
52     tau = (X(1,2)+X(2,1))/(X(1,1)-X(2,2));
53     off = sqrt(tau^2+1);
54     v = [tau - off, tau + off];
55     [d,w] = min(abs(v));
56     c = 1/sqrt(1+v(w)^2);
57     s = v(w)*c;
58 end
59 Q = [c -s;s c];
60
61 % -----%
62
63 function v = swaplist(p,s,z,b);
64 n = length(p);
65 k = 0; v = [];
66 srted = 0;
67 q = diff(s);
68 fini = 0;
69 while ~fini
70     k = k+1;
71     [dum,j] = select(p(k:n),z);
72     p(k:n+1) = [p(j+k-1) p(k:n)];
73     p(j+k) = [];
74     q(k:n+1) = [q(j+k-1) q(k:n)];
75     q(j+k) = [];
76     v = [v,j+k-2:-1:k];
77     srted = srted + q(k);
78     fini = (k==n-1)|(k==b)|(srted==b)|((srted==b+1)&(b~=0));
79 end
80
81 % -----%
82
83 function [val,pos] = select(p,z);
84 y = real(z)+abs(imag(z))*i;
85 [val pos] = min(abs(p-y));
86
87 % -----%
88
89 function [U,S] = swap(U,S,v,w);
90 [p,q] = size(S(v,w)); Ip = eye(p); Iq = eye(q);

```

```

91 r = [];
92 for j=1:q
93     r = [r;S(v,w(j))];
94 end
95 K = kron(Iq,S(v,v))-kron(S(w,w)',Ip);
96 [L,H,P,Q] = lu_complpiv(K);
97 e = min(abs(diag(H)));
98 sigp = 1:p*q;
99 for k = 1:p*q-1;
100     sigp([k,P(k)]) = sigp([P(k),k]);
101 end
102 r = e*r(sigp);
103 x = (H\ (L\r));
104 sigq = 1:p*q;
105 for k = 1:p*q-1;
106     sigq([k,Q(k)]) = sigq([Q(k),k]);
107 end
108 x(sigq) = x;
109 X = [];
110 for j=1:q
111     X = [X,x((j-1)*p+1:j*p)];
112 end
113 [Q,R] = qr([-X;e*Iq]);
114 S(:, [v,w]) = S(:, [v,w])*Q;
115 S([v,w], :) = Q'*S([v,w], :);
116 U(:, [v,w]) = U(:, [v,w])*Q;
117
118 % -----%
119
120 function [L,U,P,Q] = lu_complpiv(A);
121 P = []; Q = []; n = size(A,1);
122 for k=1:n-1;
123     [a,r] = max(abs(A(k:n,k:n)));
124     [dummy,c] = max(abs(a));
125     cl = c+k-1;
126     rw = r(c)+k-1;
127     A([k,rw], :) = A([rw,k], :);
128     A(:, [k,cl]) = A(:, [cl,k]);
129     P(k) = rw; Q(k) = cl;
130     if A(k,k) ~= 0;
131         rs = k+1:n;
132         A(rs,k) = A(rs,k)/A(k,k);
133         A(rs,rs) = A(rs,rs)-A(rs,k)*A(k,rs);
134     end
135 end
136 U = tril(A')'; L = tril(A,-1) + eye(n);

```

A.3 Analiza meta-stabilnih stanja reverzibilnih Markovljevih lanaca - particioniranje grafa

```
1 function[Xk] = graf_pristup(P, n)
2
3     N = size(P,1);
4     [Z,D] = eigs(P,n);
5
6     Q = sqrt(diag(Z*Z'));
7     for i = 1:size(Q,1)
8         Q(i) = 1/Q(i);
9     end
10    Q = diag(Q);
11    X = Q*Z;
12
13    R(:,1) = X(1,:)' ;
14    c = zeros(N,1);
15    for k = 2:n
16        c = c + abs(X*R(:,k-1));
17        [y,i] = min(c);
18        R(:,k) = X(i,:)' ;
19    end
20
21    r = 0;
22    r2 = 100;
23
24    while abs(r2-r) > eps
25        X2 = X*R;
26        Xk = zeros(N,n);
27
28        for i = 1:N
29            [m,l] = max(X2(i,:));
30            Xk(i,l) = 1;
31        end
32
33        [U,S,V] = svd(Xk'*X2);
34        r2 = trace(S);
35        r = r2;
36        R = V*U';
37    end
38 end
```