

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
Fakulta jaderná a fyzikálně inženýrská

Laplaceova matice a její aplikace

Laplacian matrix and its applications

Bakalářská práce

Autor: **Daniel Khol**

Vedoucí práce: **Doc. RNDr. Jan Vybíral Ph.D.**

Akademický rok: 2021/2022

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student:	Daniel Khol
Studijní program:	Aplikace přírodních věd
Studijní obor:	Matematické inženýrství
Studijní zaměření:	Aplikované matematicko-stochastické metody
Název práce (česky):	Laplaceova matice a její aplikace
Název práce (anglicky):	Laplacian matrix and its applications

Pokyny pro vypracování:

- 1) Student se seznámí s pojmem Laplaceovy matice a jeho vlastnostmi a podá o nich ucelený přehled.
- 2) Student sestaví Laplaceovu matici pro vybrané třídy grafů (kompletní graf, strom, orientovaný graf) a najde jejich spektra.
- 3) Student podá přehled o některých aplikacích Laplaceovy matice a jejího využití v problémech teorie grafů a teorie Markovovských řetězců včetně numerických simulací na vybraných grafech.

Doporučená literatura:

- 1) A. E. Brouwer, W. H. Haemers, Spectra of graphs. Universitext, Springer, New York, 2012.
- 2) F. Chung, Laplacians and the Cheeger Inequality for Directed Graphs. Annals of Combinatorics 9, 2005, 1-19.
- 3) M. Fiedler, Algebraic connectivity of graphs. Czech. Math. J. 23 (98), 1973, 298-305.
- 4) B. Mohar, Some applications of Laplace eigenvalues of graphs. In 'Graph symmetry', Springer, Dordrecht, 1997, 225-275.

Jméno a pracoviště vedoucího bakalářské práce:

doc. RNDr. Jan Vybíral, Ph.D.

Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze, Trojanova 13, 120 00 Praha 2


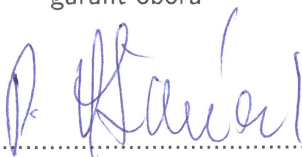
Jméno a pracoviště konzultanta:

Datum zadání bakalářské práce: 31.10.2020

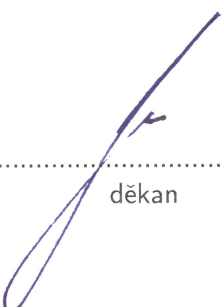
Datum odevzdání bakalářské práce: 7.7.2021

Doba platnosti zadání je dva roky od data zadání.

V Praze dne 30.10.2020


.....
garant oboru

.....
vedoucí katedry




.....
děkan

Poděkování:

Chtěl bych zde poděkovat především svému školiteli docentu Janu Vybiralovi za pečlivost, ochotu, vstřícnost, odborné i lidské zázemí a svatou trpělivost při vedení mé bakalářské práce.

Čestné prohlášení:

Prohlašuji, že jsem tuto práci vypracoval samostatně a uvedl jsem všechnu použitou literaturu.

V Praze dne 1. července 2022

Daniel Khol

Název práce:

Laplaceova matice a její aplikace

Autor: Daniel Khol

Obor: Matematické inženýrství

Zaměření: Aplikované matematicko-stochastické metody

Druh práce: Bakalářská práce

Vedoucí práce: Doc. RNDr. Jan Vybíral Ph.D., Katedra matematiky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze

Abstrakt: Grafy a Markovské řetězce je možné reprezentovat pomocí matic. Jednou z nejběžnějších forem reprezentace pro grafy je Laplaceova matice. Tato práce o ní dává ucelený přehled a nachází Laplaceova spektra několika základních typů grafů. Dále shrnuje základní obecné spektrální vlastnosti Laplaceovy matice a normalizované Laplaceovy matice pro neorientované grafy, jejich vzájemný vztah i vztah s maticí přechodu z teorie Markovských řetězců. Dále se práce věnuje aplikaci daných matic, primárně ve spectral clusteringu. Vedle toho definujeme novou Laplaceovu matici pro orientované grafy a ověříme její vlastnosti. Vše je na konci ukázáno formou numerické simulace na třech grafech.

Klíčová slova: algebraická konektivita, Laplaceova matice, Laplaceovo spektrum, Markovské řetězce, náhodné procházky, normalizovaná Laplaceova matice, orientované grafy, spectral clustering, spektrální analýza, vlastní čísla

Title:

Laplace matrix and its applications

Author: Daniel Khol

Abstract: Graphs and Markov chains can be represented by matrixes. One of the most common representations is the Laplacian matrix. This thesis summaries knowledge about the Laplacian matrix and prove Laplacians of some particular graphs. Then it summarises the basic generic properties of the Laplacian matrix and the Normalized Laplacian matrix of unoriented graphs, their relations and their relation with the Transition matrix from Markov chain theory. Then we consider the application of these matrixes, especially in spectral clustering. In the thesis, we also analyze the new Laplacian matrix for oriented graphs and proves its properties. Everything concludes in the form of numerical simulation on three particular graphs.

Key words: Algebraic connectivity, Eigenvalues, Laplace spectrum, Laplacian matrix, Markov chains, Normalized Laplacian matrix, Oriented graphs, Random walk, Spectral analysis, Spectral clustering

Obsah

Úvod	9
1 Laplaceova matice	11
1.1 Základy teorie grafů	11
1.2 Definice a základní vlastnosti	13
1.3 Spektrální analýza matice sousednosti	16
1.4 Spektrální analýza Laplaceovy matice	17
1.5 Spektra vybraných grafů	24
1.5.1 Úplný graf	24
1.5.2 Kružnice	26
1.5.3 Cesta	28
1.5.4 Hvězda	29
1.5.5 Strom	30
1.6 Algebraická konektivita	32
2 Laplaceova matice pro orientované grafy	35
2.1 Různé definice Laplaceových matic pro orientované grafy	35
2.2 Abstraktní Laplaceova matice	36
2.3 Spektrální analýza abstraktní Laplaceovy matice	38
3 Náhodné procházky	43
3.1 Laplaceovy matice pro Markovské řetězce	43
3.2 Spektrální analýza normalizované Laplaceovy matice	45
3.3 Konvergence ke stacionární distribuci	48
3.4 Clustering Markovských řetězců	49
4 Spectral clustering	51
4.1 Neexaktní úvod	51
4.2 Popis metody	52
4.3 Řešené problémy	53
4.3.1 Příklad 1	54
4.3.2 Příklad 2	60
4.3.3 Příklad 3	66
Závěr	69

Úvod

Laplaceova matice je společně s maticí sousednosti základním způsobem, jak popsat graf formou lineární algebry. Laplaceova matice nese v podstatě stejnou informaci jako matice sousednosti a graf je oběma maticemi jednoznačně popsán. Pokusíme se zjistit jaké informace nám může přinést zkoumání spektra Laplaceovy matice. Také budeme analyzovat výhody Laplaceovy matice oproti matici sousednosti.

Řada reálných problémů je popsána teorií grafů a lineární algebra nabízí jednoduchý způsob jak je reprezentovat v počítači. V klasickém matematickém modelování může být příkladem modelování difuze na grafu, popis molekul v organické chemii, či analýza elektrických obvodů. V informačních technologiích pak patří grafy a náhodné procházky na nich k základům celého fungování webu. Jako příklad můžeme uvést webové vyhledávače, jako Google, které využívají náhodné procházky na grafech k řazení výsledků vyhledávání. Ovšem najdeme je i v řadě dalších aplikací, například vyhledávání nejrychlejších cest za pomoci genetických algoritmů, bez kterých by nefungovaly moderní navigace. Jako poslední významnou aplikaci grafů, která vyžaduje jejich reprezentaci v počítači, můžeme uvést strojové učení a práci s "big data". Často zde zkoumáme dělení velkých datových souborů na clustery, například v sociálních sítích. Další využití je pak například strojové učení v evoluční teorii her.

Bakalářská práce se skládá z několika částí. V první, spíše kompilační, části zadefinujeme klasickou Laplaceovu matici na neorientovaných grafech. Budeme zkoumat její vlastnosti a provedeme její základní spektrální analýzu. Dále pak nalezneme a dokážeme Laplaceova spektra pro konkrétní grafy: úplný graf, kružnice, cesta a hvězda. Pro stromy udáme přehled o jejich vlastnostech a na závěr podkapitoly uvedeme Kirchhoffovu větu. Poté budeme provádět přípravné práce pro spectral clustering, zde zadefinujeme algebraickou konektivitu a Fiedlerův vektor.

V druhé části, tvořené spíše vlastním výzkumem, se pokusíme o vhodné zadefinování Laplaceovy matice pro orientované grafy. Poté ověříme, jestli má stejně výhodné spektrální vlastnosti jako klasická Laplaceova matice pro neorientované grafy. Tyto vlastnosti porovnáme s jinými maticemi popisující orientované grafy. Budeme se snažit zjistit, jestli tato definice řeší problém clusteringu lépe než, v současné době běžný, postup využívající symetrizaci orientovaného grafu. Ten má problém v tom, že nerozlišuje, jestli mezi vrcholy existuje orientovaná hrana jedním, nebo oba směry. To může vést k nepřesnostem, které si ve čtvrté kapitole ukážeme.

Ve třetí části pak budeme zkoumat normalizovanou Laplaceovu matici. Zaměříme se na její vztah k Markovským řetězcům, Laplaceově matici pro náhodné procházky a matici přechodu. Vyslovíme a dokážeme věty, které propojují spektrum normalizované Laplaceovy matice s maticí přechodu. Jednu podkapitolu věnujeme vztahu mezi spektrem normalizované Laplaceovy matice a rychlostí konvergence ke stacionární distribuci příslušného Markovského řetězce.

Ve čtvrté části se zaměříme na jednu konkrétní aplikaci spektra Laplaceovy matice, konkrétně spectral clustering. Zde se pokusíme vysvětlit její propojení s algebraickou konektivitou a Fiedlerovým vektorem. Poté vysvětlíme blízkost tohoto problému pro normalizovanou Laplaceovu matici a klasickou Laplaceovu matici. Na závěr aplikujeme spectral clustering na tři grafy. Na prvním názorně ukážeme

fungování metod. U druhého příkladu porovnáme výsledky spectral clusteringu při použití Laplaceovy matice a normalizované Laplaceovy matice. Dále spočítáme i abstraktní Laplaceovu matici a srovnáme její výsledky s výsledky Laplaceovy matice a normalizované Laplaceovy matice. Poté spočítáme abstraktní Laplaceovu matici pro graf upravený do slabě souvislého, nebo rovnou nesouvislého tvaru. Na poslední orientovaný graf pak aplikujeme novou abstraktní Laplaceovu matici, na symetrizaci pak užitíme klasickou Laplaceovu matici a provedeme standardní spectral clustering pro dva clustery. Následně budeme diskutovat rozdíly výsledků.

Kapitola 1

Laplaceova matice

1.1 Základy teorie grafů

Na začátku si vyjasníme určité pojmy z teorie grafů [14]. V první části práce budeme uvažovat neorientované grafy G , jako uspořádané dvojice (V, E) s tím, že V je neprázdná množina vrcholů a E je množina hran jako dvoubodových podmnožin V . Čtenář si neorientovaný graf může představit jako body (vrcholy) připojené čarami (hranami). Není těžké rozmyslet, že je to poměrně dobrá aproximace řady systémů z reálného světa. Například silniční síť si můžeme představit jako neorientovaný graf, kde silnice budou hrany a křižovatky vrcholy. Samozřejmě i z tohoto příkladu by nás hned mohlo napadnout, že by bylo dobré do grafu nějakým způsobem zakomponovat i další informace. Například jestli je silnice jednosměrná, či obousměrná, nebo jak je dlouhá, což budeme řešit v dalších odstavcích. Mnoho čtenářů může znát (hlavně z anglofonní literatury) obecnější definice pojmu neorientovaný graf. Pak bude naše definice přibližně odpovídat tomu co se při obecnějších definicích označuje jako prostý neorientovaný graf bez smyček. Jako v budeme označovat prvky množiny V a $e = \{x, y\}$ budeme označovat prvky množiny E . Velmi často nám přijde vhod říci, jestli mezi dvěma konkrétními vrcholy v_i, v_j existuje hrana, tedy pokud pro

$$v_i, v_j \in V : \{v_i, v_j\} \in E,$$

označíme takové vrcholy jako připojené. Velká část teorie Laplaceovy matice se dá vybudovat s jistými obtížemi i pro nekonečné grafy. My ovšem budeme v této práci uvažovat jen grafy konečné. Tak označujeme neorientované grafy s konečnou kardinalitou množiny V .

Nyní se dostaneme k definování určitých pojmů, které budeme v práci užívat. Stupněm vrcholu neorientovaného grafu označíme počet hran grafu obsahujících vrchol v , tedy

$$\deg(v) = |\{e \in E : v \in e\}|.$$

Jak již víme z úvodu, tato práce se bude týkat i náhodných procházek na grafech. Kdybychom použili předchozí ilustrační příklad se silniční sítí, je možné si náhodnou procházku představit jako kdybychom se chtěli jen tak projet a budeme náhodně jezdit po okolí. Samozřejmě se nabízí otázka jak vůbec popsat trasu, kterou jsme jeli a kde jsme skončili. Na neorientovaných grafech k tomu použijeme pojem sled, což je posloupnost vrcholů v_1, \dots, v_j , pro kterou platí, že mezi každými dvěma po sobě jdoucími vrcholy je hrana, tedy

$$\forall i \in \{2, \dots, j\} : \{v_{i-1}, v_i\} \in E.$$

Důležitá vlastnost, kterou budeme v rámci Laplaceovy matice zkoumat, je souvislost. Neorientovaný graf je souvislý právě tehdy, když pro každé dva vrcholy z V platí, že mezi nimi existuje sled. Pro řadu

definic či důkazů budeme taktéž potřebovat pojem podgraf. Tak označujeme neorientovaný graf, který z původního neorientovaného grafu vznikl odebráním některých hran nebo vrcholů. Stejně tak můžeme zadefinovat poměrně logický pojem propojenost, kdy dva vrcholy budou propojené právě tehdy, když mezi nimi existuje sled. Graf je taktéž možné rozdělit na více grafů za podmínky, že je beze zbytku rozdělíme na více disjunktních podmnožin, které stále splňují definici neorientovaného grafu. Jinými slovy, graf můžeme rozdělit jen v případě, že nám nezůstane nějaká hrana bez vrcholu. Je tedy zřejmé, že to bude pouze v případě, že pro všechny vrcholy každé disjunktní podmnožiny bude platit, že nejsou propojené s jakýmkoliv vrcholem z jiné podmnožiny.

V druhé části textu budeme uvažovat orientované grafy. Tedy opět dvojice (V, E) , kde E je množina orientovaných hran jako podmnožina $V \times V$. Prvky množiny E budeme označovat $e = (x, y)$ s tím, že e vychází z x a přichází do y . Když se vrátíme k příkladu se silniční sítí, tak jsme tímto dokázali do pojmu graf zakomponovat orientaci. Tedy pro silniční síť už umíme zavést jak obousměrné silnice, tak jednosměrné silnice. Pro každý orientovaný graf $G = (V, E)$ je možné najít neorientovaný graf $G = (V, E')$ jako jeho symetrizaci s tím, že najdeme symetrickou množinu hran jako

$$E' = \{(x, y); (x, y) \in E \vee (y, x) \in E\}.$$

Symetrizace v podstatě znamená, že z grafu vymažeme informaci o orientaci. Jestliže je mezi dvěma vrcholy hrana jedním či oběma směry řekneme, že je mezi nimi neorientovaná hrana. Z toho by nás mohlo napadnout, že v určitém smyslu jsou neorientované grafy jen speciálním případem grafů a zároveň může neorientovaný graf pro určité případy substituovat graf orientovaný. Jejich vztahu se často využívá právě v Laplaceově matici, protože je obtížné přijít s užitečnou definicí Laplaceovy matice pro orientované grafy. Pojmy definované pro neorientované grafy je možné zobecnit i do grafů orientovaných. První významnou změnu uděláme v definici souvislosti, kdy námi definovanou souvislost na neorientovaných grafech nazveme silnou souvislostí na grafech orientovaných. Jako slabou souvislost nazveme situaci kdy symetrizace daného grafu je souvislá podle definice pro neorientované grafy. Dále pak definujeme stupeň orientovaného grafu jako dvouprvkový vektor obsahující $\deg_{out}(v)$ a $\deg_{in}(v)$ s tím, že

$$\deg_{out}(v) = |\{e \in E : \exists v_p \in V : e = (v, v_p)\}|$$

je počet vycházejících hran z v a

$$\deg_{in}(v) = |\{e \in E : \exists v_p \in V : e = (v_p, v)\}|$$

je počet přicházejících hran do v .

V části věnované normalizované Laplaceově matici a Markovským řetězcům budeme potřebovat obohatit naši definici $G = (V, E)$ tak, abychom mohli přiřadit různým hranám různé pravděpodobnosti. To uděláme tím způsobem, že každé hraně $e \in E$ přisoudíme váhu $\omega(e) \in \mathbb{R}^+$. To se běžně formalizuje formou váhové funkce $\omega : E \rightarrow \mathbb{R}^+$. Takovýto graf pak nazýváme vážený neorientovaný graf, či vážený orientovaný graf. V této práci však pouze zmíníme vážené neorientované grafy. Kdybychom si to opět měli představit jako silniční síť. Nyní jsme obohatili naši definici například o délku jednotlivých silnic mezi křižovatkami. Mnozí čtenáři se mohli setkat s odlišnou definicí, kdy již na začátku je definován graf, který připouští rovnoběžné hrany. Tedy, že mezi dvěma vrcholy může existovat více než jedna hrana. Hrany poté fungují jako váhy. Výsledný efekt je ovšem, při naší aplikaci, naprosto stejný. Hlavně ve spojitosti s aperiodičností Markovských řetězců budeme potřebovat, aby graf byl aperiodický. Na to první potřebojeme znát pojem kružnice, tak graf nazvu pokud pro jeho množiny V a E platí, že

$$V = \{1, 2, \dots, k\} \wedge E = \{(i, i+1); i = 1, \dots, k-1\} \cup \{(1, k)\}.$$

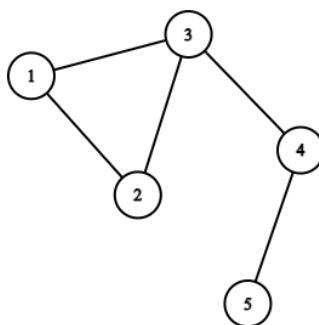
V případě, že v grafu najdeme všechny podgrafy, které splňují definici kružnice a zjistíme, že největší společný dělitel délky těchto kružnic je jedna, nazveme graf aperiodickým. Délkou kružnice myslíme počet jejích hran.

Úmluva: V situacích, kdy je z kontextu jasné o jakém grafu mluvíme, budeme v důkazech či vysvětlivkách používat pouze pojem "graf". Nebude-li řečeno jinak, či nebude-li z kontextu zcela zřejmý jiný účel daného písmene (např e v exponenciále), tak písmenem G budeme obecně označovat různé grafy, v jejich vrcholy, V příslušnou množinu vrcholů, e jejich orientované či neorientované hrany a E příslušnou množinu hran. Písmenem k pak budeme zpravidla označovat počet vrcholů daného grafu.

1.2 Definice a základní vlastnosti

Nyní budeme definovat Laplaceovu matici za pomoci matice stupňovitosti a matice sousednosti. Pro lepší pochopení budeme názorně ukazovat tvorbu jednotlivých matic pro graf H z obrázku 1.1.

Obrázek 1.1: Neorientovaný graf H



Naším cílem bude sestavit matici, která jasně a jednoznačně popisuje konečný a neorientovaný graf. Začneme definicí matice stupňovitosti, která nám udá stupně vrcholů.

Definice 1.2.1. *Necht' G je neorientovaný graf s k vrcholy. Maticí stupňovitosti $D \in \mathbb{R}^{k,k}$ grafu G definujeme*

$$D_{i,j} = \begin{cases} \deg(v_i), & \text{když } i = j \\ 0, & \text{jinak.} \end{cases}$$

Pro náš graf H tedy máme následující matici stupňovitosti.

$$D(H) = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Je nutné si uvědomit důvody a dopady požadavku neorientovanosti. Stupeň orientovaného grafu je totiž dvouprvkový vektor, tedy tato definice jde velmi těžko zobecnit do orientovaných grafů, což později uvidíme. Vzhledem k tomu, že aplikace ve kterých bychom potřebovali využít Laplaceovu matici jsou často orientované grafy je toto nejslabší místo celé definice, které se pokusíme řešit.

Budeme pokračovat maticí sousednosti, která nám popíše, jak konkrétně jsou dané vrcholy připojené.

Definice 1.2.2. *Necht' G je neorientovaný graf s k vrcholy. Maticí sousednosti $A \in \mathbb{R}^{k,k}$ grafu G definujeme*

$$A_{i,j} = \begin{cases} 1, & \text{když } \{v_i, v_j\} \in E \\ 0, & \text{jinak.} \end{cases}$$

Nyní se můžeme podívat jak vypadá matice sousednosti pro náš graf H .

$$A(H) = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Samozřejmě si můžeme všimnout, že by nám pro jednoznačný popis bohatě stačila matice sousednosti. Jak ovšem později zjistíme, tak z řady důvodů nám v definici matice, kterou popisujeme graf, přijde vhod i matice stupňovitosti.

Až budeme matici sousednosti zobecňovat do orientovaných grafů, tak dvouprvkové množiny jednoduše nahradíme uspořádanými dvojicemi. Kdybychom potřebovali v normalizované Laplaceově matici přidat váhy jednotlivým hranám, nahradíme jedničky v matici sousednosti váhami ω příslušící jednotlivým hranám. Z toho také vidíme, proč je matice sousednosti tak oblíbená pro reprezentaci grafů v počítačích. Mimo případů, kdy bychom se v definicích opravdu výrazně vzdalovali od původních myšlenek teorie grafů (např. multihrany, které v této práci ani nedefinujeme), se nám vždy podaří jednoduše zobecnit matici sousednosti tak, aby graf jednoznačně popsala.

Úmluva: Kvůli velmi častému používání těchto pojmů budeme, nebude-li řečeno jinak, písmenem D po zbytek práce označovat matici stupňovitosti. Písmenem A budeme pak označovat matici sousednosti pro příslušný graf.

Nyní se konečně dostáváme k definici Laplaceovy matice. Jak jsme již zmínili, k její definici využijeme matici stupňovitosti a matici sousednosti. V případě, že bychom k matici stupňovitosti přičetli matici sousednosti, získali bychom bezznaménkovou Laplaceovu matici. Ta se ovšem pro naše aplikace nehodí. Naopak klasickou Laplaceovu matici získáme tím, že od matice stupňovitosti matici sousednosti odečteme.

Definice 1.2.3. *Necht' G s k vrcholy je konečný neorientovaný graf. Necht' D je jeho matice stupňovitosti a A matice sousednosti. Matici $L = D - A$ nazveme Laplaceovou maticí grafu G .*

Tedy pro náš graf H máme následující Laplaceovu matici.

$$L(H) = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

Laplaceova matice má tedy na diagonále počet vrcholů, ke kterým je daný vrchol připojen a mimo diagonálu -1 na místech připojených vrcholů. Nabízí se otázka, proč je matice právě Laplaceova. Důvodem je spojitost s Laplaceovým operátorem. Jednou z možností, jak lépe porozumět tomuto vztahu, je matice incidence, kterou se budeme zabývat později. Další možností je diskrétní Laplaceův operátor na grafech [5].

Věta 1.2.1. *Necht' G je neorientovaný graf a i, j jsou indexy jeho vrcholů. Necht' $g : V \rightarrow \mathbb{R}$ a L je Laplaceova matice G . Pak pro všechna i platí, že*

$$(Lg)_i = \sum_{j \sim i}^k (g_i - g_j), \quad (1.1)$$

kde symbolem $j \sim i$ myslíme součet přes všechny vrcholy mezi kterými existuje hrana.

$(Lg)_i$ nazveme Diskrétním Laplaceovým operátorem. Že se jedná o ekvivalentní definici Laplaceovy matice je zřejmé. Na nediagonálních připojených prvcích vznikne -1 . Na diagonále pak počet hran příslušících danému vrcholu.

Nyní se dostáváme k samotnému propojení s klasickým Laplaceovým operátorem na funkci g

$$\nabla^2 g = \left(\frac{\partial^2 g}{\partial x_1^2} + \frac{\partial^2 g}{\partial x_2^2} + \dots + \frac{\partial^2 g}{\partial x_n^2} \right).$$

Budeme vycházet z Laplaceovy rovnice $\nabla^2 g = 0$, její řešení je dle [15] ekvivalentní variačnímu problému Dirichletovy energie

$$H(g) = \frac{1}{2} \int_{\Omega} \|\nabla g\|^2,$$

kde Ω je otevřená podmnožina \mathbb{R}^n , tu pro diskrétní verzi nahradíme množinou hran. Stejně tak gradient můžeme nahradit numerickou derivací, což vede na diskrétní verzi

$$\begin{aligned} H(g) &= \frac{1}{2} \sum_{e \in E} (g_x - g_y)^2 = \frac{1}{2} \sum_{j \sim i}^k (g_i^2 - 2g_i g_j + g_j^2) = \frac{1}{2} \sum_{i=1}^k g_i^2 - \sum_{j \sim i}^k g_i g_j + \frac{1}{2} \sum_{j=1}^k g_j^2 = \\ &= \sum_{i=1}^k g_i^2 - \sum_{j \sim i}^k g_i g_j = g^T Dg - g^T Ag = g^T Lg, \end{aligned}$$

kde $e = \{x, y\}$. Tedy máme dvě, pro další důkazy důležité, rovnice

$$g^T Lg = \frac{1}{2} \sum_{e \in E} (g_x - g_y)^2 \quad (1.2)$$

$$H(g) = g^T Lg, \quad (1.3)$$

druhou rovnici můžeme přepsat na

$$Lg = H(g)g,$$

což je rovnice pro vlastní čísla Laplaceovy matice L .

Definice 1.2.4. *Necht' B je libovolná čtvercová matice, pokud*

- a) $\forall i : |B_{i,i}| > \sum_{i \neq j} |B_{i,j}| \wedge \forall j : |B_{j,j}| > \sum_{i \neq j} |B_{i,j}|$, pak B nazvu *silně diagonálně dominantní maticí*,
- b) $\forall i : |B_{i,i}| \geq \sum_{i \neq j} |B_{i,j}| \wedge \forall j : |B_{j,j}| \geq \sum_{i \neq j} |B_{i,j}|$, pak B nazvu *slabě diagonálně dominantní maticí*.

Věta 1.2.2 (Základní vlastnosti Laplaceovy matice). *Necht' G je konečný neorientovaný graf a L je jeho Laplaceova matice, pak*

- a) L je symetrická,
- b) součet každého řádku a sloupce matice L je 0,
- c) L je singulární,
- d) L je slabě diagonálně dominantní matice.

Důkaz. a) Graf je neorientovaný, tedy pokud existuje hrana mezi v_i a v_j , tak existuje i hrana mezi v_j a v_i . Matice L je tedy symetrická.

b) Graf je neorientovaný, tedy $\deg(v_i)$ se musí rovnat počtu napojených hran. Tedy z definice matice D a A je zřejmé, že se v každém sloupci a řádku odečtou na nulu.

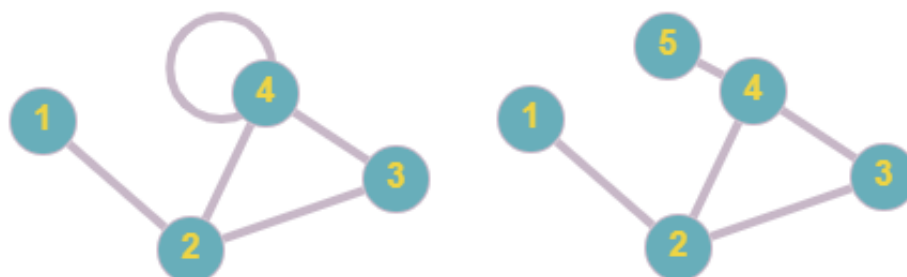
c) Z předchozího bodu vyplývá, že součet všech řádků je 0, tedy matice musí být singulární z Frobeniovovy věty.

d) Z bodu b) vyplývá, že součet všech řádků a sloupců je 0. Z definice L plyne, že přesně polovina součtu prvků v absolutní hodnotě bude v diagonálním prvku. To je z definice slabě diagonální matice. \square

1.3 Spektrální analýza matice sousednosti

Nyní krátce shrneme základní spektrální vlastnosti matice sousednosti. Můžeme si všimnout, že matice sousednosti má jisté výhody oproti Laplaceově matici, například může reprezentovat grafy se smyčkami. Smyčka je rozšíření naší definice grafu, kdy u vrcholů v_i rozšíříme množinu dvouprvkových množin E i o jednoprvkové množiny. Ty budeme reprezentovat jako $\{v_i\}$. Důvod, proč nedává smysl zavádět do Laplaceovy matice smyčky je, že bychom tím ztratili jednu z nejdůležitějších vlastností Laplaceovy matice. A to, že součet všech řádků i sloupců je 0. Tato absence se dá ovšem pro řadu výsledných aplikací poměrně jednoduše nahradit. Například tím, že vytvoříme nový vrchol, který bude připojen jen k vrcholu, který měl původně smyčku, čímž v podstatě smyčku vytvoříme. To je zobrazeno na obrázku 1.2.

Obrázek 1.2: Řešení problémů se smyčkou



Definice 1.3.1. Necht' G je konečný neorientovaný graf a A je příslušná matice sousednosti. Adjenčním spektrem grafu G nazveme spektrum matice A .

Úmluva: Běžně se vlastní čísla matice sousednosti označují jako vlastní čísla grafu. Stejně se ovšem často označují i vlastní čísla Laplaceovy matice daného grafu. Vzhledem k tématu této práce tedy termíny "vlastní čísla grafu", "vlastní vektory grafu" a "spektrum grafu" vyhradíme pouze pro Laplaceovu matici.

Definice 1.3.2. Necht' G je konečný neorientovaný graf a A je příslušná matice sousednosti. Adjenčním charakteristickým polynomem grafu G nazveme

$$p_a(t) = \det(It - A),$$

kde I je matice identity.

Z definice matice sousednosti plyne několik zjevných vlastností. Vzhledem k tomu že A je reálná a symetrická, tak adjenční vlastní čísla jsou reálná. Ze stejného důvodu je i algebraická a geometrická násobnost adjenčních vlastních čísel stejná. Z toho taktéž plyne i to, že A je diagonalizovatelná.

Mohli bychom nalézt i další vlastnosti spektra matice sousednosti a je možné nalézt literaturu tím se zabývající. V této podkapitole jsme vlastnosti spektra matice sousednosti jen krátce zmínili za účelem, abychom později viděli výhody matice Laplaceovy. Všimněme si, že například matice A zjevně není obecně pozitivně definitní, či pozitivně semi-definitní. To by se nám pro spektrální analýzu velmi hodilo.

1.4 Spektrální analýza Laplaceovy matice

V této podkapitole první zadefinujeme Laplaceovo spektrum. Poté budeme dokazovat jeho základní spektrální vlastnosti. Už víme, že Laplaceova matice je symetrická a singulární, a že součet jejích řádků a sloupců je nula. Začneme samotnou definicí spektra Laplaceovy matice.

Definice 1.4.1. Necht' G je konečný neorientovaný graf a L je příslušná Laplaceova matice. Spektrem grafu G nazveme spektrum matice L .

Definice 1.4.2. Necht' G je konečný neorientovaný graf a L je příslušná Laplaceova matice. Charakteristickým polynomem grafu G nazvu

$$p_l(t) = \det(It - L),$$

kde I je matice identity.

Stejně, jako v případě matice sousednosti, máme několik zjevných vlastností. Vzhledem k tomu že L je reálná a symetrická, vlastní čísla jsou reálná, algebraická a geometrická násobnost vlastních čísel je stejná a L je diagonalizovatelná. Z vlastností reálných a symetrických matic taktéž plyne, že vlastní vektory L budou ortogonální. Nyní se dozvíme jeden z důvodů proč jsme se v definici tolik snažili, aby součet všech řádků a sloupců byl roven nule. Plyne z toho totiž, že hned víme nejméně jedno vlastní číslo a vlastní vektor Laplaceovy matice.

Věta 1.4.1. *Necht' G je konečný neorientovaný graf s k vrcholy. Necht' L je příslušná Laplaceova matice, λ_i její vlastní čísla a X_i vlastní vektory L s tím, že $i \in \{1, \dots, k\}$. Pak $\lambda_1 = 0$ a příslušný vlastní vektor $X_1 = (1, \dots, 1)^T$.*

Důkaz. L je z věty 1.2.2. singulární. To ze základů lineární algebry znamená, že alespoň jedno vlastní číslo musí být nulové. Z toho, že $\lambda = 0$ plyne, že $LX = 0X$. Dle předchozí věty 1.2.2. je součet všech řádků roven nule. Tedy $(1, \dots, 1)^T$ je vlastním vektorem. □

Důvod proč jsme nyní označili 0 jako λ_1 se dozvíme později. Zatím si to odůvodníme tím, že se jedná o nejmenší vlastní číslo v absolutní hodnotě.

Nyní se dostáváme k již zmíněné matici incidence. Výraznou změnou oproti předchozím definicím, které vyžadovaly pouze očíslování vrcholů bude, že nyní je třeba očíslovat i hrany. Vzhledem k důvodům, které se dozvíme až v podkapitole o orientovaných grafech zavádíme matici incidence, která se liší od toho, s čím se mohli čtenáři zatím setkat jako s maticí incidence a orientovanou maticí incidence.

Definice 1.4.3. *Necht' $G = (V, E)$ je konečný neorientovaný graf. Necht' $V = \{v_1, \dots, v_k\}$ a $E = \{e_1, \dots, e_l\}$. Zvolme pevné $e_j = \{v_m, v_n\}$, kde $j \in \{1, \dots, l\}$ s tím, že $m, n \in \{1, \dots, k\}$ a zároveň $m < n$. Maticí Incidence $M \in \mathbb{R}^{k, 2l}$ nazveme matici, kde*

$$\begin{aligned} M_{m,j} &= +1 \\ M_{m,(j+l)} &= -1 \\ M_{n,j} &= -1 \\ M_{n,(j+l)} &= +1 \\ M_{i,(j)} &= 0 \text{ pro } i \notin \{m, n\} \\ M_{i,(j+l)} &= 0 \text{ pro } i \notin \{m, n\} \end{aligned}$$

To pro náš graf H znamená následující matici incidence.

$$M(H) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

U této definice se na nějakou dobu zastavíme. Jako první si uvědomíme, že naše definice není úplně jednoznačná, z důvodu, že $m < n$ není dobře definována. Zavedení orientace je totiž ponecháno zcela na

naší libovůli. Matice incidence každého grafu tak není nezávislá na zvoleném číslování m a n . Ve skutečnosti jsme tedy zavedli 2^l matic. Hned v další větě zjistíme, že vzhledem k aplikaci matice incidence to není důležité.

Jako druhé je opět nutné upozornit na to, že toto není standardní definice matice incidence pro neorientované grafy, ani žádná její verze. Ve standardní definici se totiž zpravidla počítá jen přes jednu orientaci. My naproti tomu počítáme přes obě orientace. To umožňuje snadné zobecnění na orientované grafy. Pro neorientované grafy má tak naše definice dvojnásobný počet sloupců oproti standardní definici. Pro výslednou aplikaci v Laplaceově matici se bude od standardní teorie lišit vznikem $\frac{1}{2}$.

Jako třetí je dobré se opět vrátit ke vztahu Laplaceovy matice a Laplaceova operátoru. Matice incidence by se totiž dala interpretovat jako "derivace" grafu. Poté nebude příliš překvapivé, čemu se rovná MM^T

Věta 1.4.2. *Nechť G je konečný neorientovaný graf, L je příslušná Laplaceova matice a M příslušná matice incidence, pak*

$$L = \frac{1}{2}MM^T. \quad (1.4)$$

Důkaz. První krátce zanalyzujeme definici 1.4.3., není těžké poznat a můžeme to vidět i na matici $M(H)$, že pro neorientované grafy bude mít matice M následující tvar

$$M = \begin{bmatrix} -M_o & M_o \end{bmatrix}, \quad (1.5)$$

což odpovídá dvěma orientacím, které jsme u každé neorientované hrany mohli zvolit.

Začneme prvky, když $i = j$. Nechť n je počet hran G pak

$$(MM^T)_{i,i} = \sum_{l=1}^n M_{i,l}M_{l,i}^T = \sum_{l=1}^n M_{i,l}M_{i,l} = \sum_{l=1}^n M_{i,l}^2 = \sum_{e \in V} M_{e,v_i}^2 = \sum_{e \in v_i} 1 = 2 \deg(v_i). \quad (1.6)$$

Zamysleme se, proč se nám v posledním rovná se rovnice (1.6) objevila dvojka. Je to způsobeno tím, že počítáme přes obě orientace, tedy každá hrana tam bude reprezentována dvakrát.

Nyní budeme uvažovat $i \neq j$, kde ovšem platí, že mezi i a j je hrana.

$$(MM^T)_{i,j} = \sum_{l=1}^n M_{i,l}M_{l,j}^T = \sum_{l=1}^n M_{i,l}M_{j,l} = \sum_{e \in V} M_{e,v_i}M_{e,v_j} = -2. \quad (1.7)$$

Opět se zamysleme nad posledním rovná se rovnice (1.7). Musíme si uvědomit, že opět počítáme přes obě orientace. Tedy připojení obou vrcholů se propíše dvakrát.

Je zřejmé, že když $i \neq j$ a zároveň mezi i a j není hrana, tak je matice na těchto místech nulová. Tedy po složení s předchozími výsledky máme $MM^T = 2L$.

□

Z vlastnosti násobení matic tedy vidíme, že je irelevantní, jakou orientaci jednotlivým hranám neorientovaného grafu přiřadíme. Rovnost (1.4) tedy nemusí splňovat právě jedna matice $k \times 2l$ a mimo triviálních příkladů tomu tak ani pro žádný graf nebude. Nám ovšem pro důkaz silných spektrálních vlastností stačí znát jen jednu takovou matici.

Věta 1.4.3. *Nechť G je konečný neorientovaný graf a L je příslušná Laplaceova matice, pak L je pozitivně semi-definitní.*

Důkaz. Pokud známe rovnici (1.4), mohli bychom tuto vlastnost označit za zjevnou. Pro názornost si ji ovšem ukážeme. $\frac{1}{2}$ můžeme z rovnice (1.4) samozřejmě zanedbat. Použijeme definici vlastních čísel

$$\lambda_i X_i = L X_i$$

pro $i \in \{1, \dots, k\}$. To přepíšeme na

$$\lambda_i = X_i^T L X_i = X_i^T M M^T X_i = (M^T X_i)^T M^T X_i = \|M^T X_i\|^2 \geq 0,$$

Vzhledem k tomu, že z věty 1.4.1. víme, že $\lambda_1 = 0$, tak L je pozitivně semi-definitní a ne pozitivně definitní. □

Touto větou jsme se dostali k jednomu z hlavních důvodů zavedení celé teorie Laplaceovy matice. Pro pozitivně semi-definitní matice totiž máme efektivní numerické algoritmy, schopné najít vlastní čísla dané matice. Nezáporná vlastní čísla nám také otevírají řadu aplikací. Například později rozebíraný spectral clustering. Dále to samozřejmě ospravedlňuje, proč jsme již dříve začali označovat $\lambda_1 = 0$, protože se vždy bude jednat o nejmenší vlastní číslo Laplaceovy matice. Nyní se ovšem budeme věnovat následku další důležité vlastnosti Laplaceovy matice a to, že součet všech řádků a sloupců je roven nule. Spolu s ortogonalitou vlastních vektorů Laplaceovy matice nám to umožňuje získat informace i o prvcích ostatních vlastních vektorů L .

Věta 1.4.4. *Necht' G je konečný neorientovaný graf a L je příslušná Laplaceova matice. Necht' λ_i jsou její vlastní čísla a $X_i = (x_{i,1}, \dots, x_{i,k})$ příslušné vlastní vektory. Pak pro každé X_i , které nepřísluší $\lambda = 0$, platí*

$$\sum_{j=1}^k x_{i,j} = 0. \tag{1.8}$$

Důkaz. Již jsme dokázali, že $(1, \dots, 1)^T$ je vlastním vektorem. Dále víme že X_i jsou ortogonální. Tedy všechny vlastní vektory jsou kolmé na vektor $(1, \dots, 1)^T$. Tedy skalární součin prvků vlastních vektorů (x_1, \dots, x_k) a $(1, \dots, 1)^T$ je roven 0. Z čehož nám plyne tvrzení věty, tedy $\sum_j x_j = 0$. □

Tedy mimo vektoru $(1, \dots, 1)^T$, nebo obecně vektorů, které přísluší $\lambda = 0$ jsou součty všech jejich prvků rovny nule. To lze často v důkazech využívat a je to další výhoda oproti používání pouze matice sousednosti. V další části podkapitoly se budeme zabývat vztahy mezi souvislostí neorientovaných grafů a jejich spektry. Začneme tím, jak vlastně bude vypadat Laplaceova matice nespojitelného grafu.

Lemma 1.4.5. *Necht' G je konečný neorientovaný graf. Necht' L je příslušná Laplaceova matice a G_i , $i \in \{1, \dots, j\}$ jsou nespojitelné souvislé grafy, ze kterých se skládá G . Pak je možné vrcholy G přeznačit tak, aby Laplaceova matice G byla blokově diagonální s tím, že jednotlivé bloky původní matice L jsou Laplaceovy matice jednotlivých grafů G_i .*

Důkaz. Vrcholy grafu můžeme libovolně přeznačit. Necht' k_i je počet vrcholů pro každé G_i . Samozřejmě

$$\sum_{i=1}^j k_i = k,$$

kde j je počet nepropojených souvislých grafů a k počet vrcholů G . Nyní přeznačíme vrcholy G tak, že vrcholy G_1 označíme $1, \dots, k_1$, vrcholy G_2 označíme $k_1 + 1, \dots, k_1 + k_2$ a tak budeme pokračovat až k G_j . Z předchozí rovnice jasně plyne, že tím postihneme všechny vrcholy dané matice.

Po našem přeznačení vrcholů vezměme prvních k_1 sloupců a k_1 řádků matice L . Vzhledem k tomu, že vrcholy grafu G_1 nemají napojení s žádným dalším vrcholem mimo G_1 je zřejmé následující. Pro každý řádek i sloupec bude platit, že mimo prvních k_1 hodnot daného sloupce nebo řádku budou všude nuly. Toto můžeme udělat pro každé G_i , tedy L je bloková matice tvaru

$$L = \begin{bmatrix} L_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & L_2 & \ddots & \ddots & & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & L_{j-1} & 0 \\ 0 & 0 & \cdots & 0 & 0 & L_j \end{bmatrix}, \quad (1.9)$$

kde L_i jsou submatice odpovídající vrcholům příslušných G_i . Takovéto submatice mají na diagonále stupeň vrcholu původní matice a prvky mimo diagonálu značí sousednost v rámci G_i . Jedná se tedy o Laplaceovy matice daných souvislých grafů G_i . □

Vlastnosti v tomto lemmatu by někteří označili až za zjevné, ovšem jsou pro další důkazy velmi důležité. Ale teď se již vydáme vstříc spektrální analýze s tím, že začneme zkoumat spektrum pro souvislé grafy. Ukážeme si totiž, že pro souvislé grafy bude mít spektrum pouze jedno nulové vlastní číslo.

Věta 1.4.6. *Necht' G je konečný neorientovaný graf a L je příslušná Laplaceova matice, pak G je souvislý právě tehdy, když algebraická násobnost $\lambda = 0$ je rovna jedné.*

Důkaz. Začneme důkazem, že pokud je G je souvislý, pak je algebraická násobnost $\lambda = 0$ rovna jedné. Začneme tím, že vezmeme druhé nejmenší vlastní číslo λ_2 . Dále si uvědomíme, že příslušný vlastní vektor X_2 musí být kolmý na vektor $(1, \dots, 1)^T$. Poté využijeme diskrétní Laplaceův operátor (1.1)

$$(Lx)_i = \lambda_2 x_{2,i} = \sum_{i \sim j}^k (x_{2,i} - x_{2,j})^2. \quad (1.10)$$

Zjevně vidíme, že

$$\sum_{i \sim j}^k (x_{2,i} - x_{2,j})^2 \geq 0. \quad (1.11)$$

Jak již víme, $(1, \dots, 1)^T$ nemůže být vlastním vektorem pro λ_2 . Tedy musí existovat nějaké $\{v_i, v_j\} \in E$ pro které platí, že $x_{2,i}$ se nerovná $x_{2,j}$. Samozřejmě se nabízí otázka co se stane, když mezi v_i a v_j není hrana. Poté mezi nimi musí existovat alespoň posloupnost vrcholů, které jsou spojeny hranou. Když tuto

posloupnost nazveme sled, tak vidíme, že tato podmínka je splněna z definice souvislosti. To implikuje, že

$$\sum_{i \sim j}^k (x_{2,i} - x_{2,j})^2 > 0.$$

A tedy všechna vlastní čísla L mimo toho příslušícího vektoru $(1, \dots, 1)^T$ musí být nenulové, pokud je graf souvislý.

Pokračujeme tím, že pokud je algebraická násobnost $\lambda = 0$ jedna, pak je graf souvislý. Důkaz ukažme sporem. Budeme tedy předpokládat, že náš graf G se skládá z dvou nepropojených grafů G_1 a G_2 . To by ovšem znamenalo, že dle lemma 1.4.5. můžeme graf rozdělit na dva grafy s tím, že Laplaceova matice grafu G vypadá následovně

$$L = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix}, \quad (1.12)$$

kde L_1 a L_2 jsou Laplaceovy matice G_1 a G_2 . Budeme předpokládat, že G_1 náleží vrcholy v_1, \dots, v_i a G_2 náleží vrcholy $v_{i+1} \dots v_k$.

Dle věty 1.4.1 platí, že každá Laplaceova matice má alespoň jedno vlastní číslo rovno nule. Můžeme tedy jednoduše nalézt dva vektory X a X' , které budou nezávislými vlastními vektory L ,

$$X = (x_1, \dots, x_i, x_{i+1}, \dots, x_k)^T = (1, \dots, 1, 0, \dots, 0)^T, \quad (1.13)$$

$$X' = (x'_1, \dots, x'_i, x'_{i+1}, \dots, x'_k)^T = (0, \dots, 0, 1, \dots, 1)^T, \quad (1.14)$$

čímž jsme ukázali spor, protože algebraická násobnost $\lambda = 0$ není jedna. □

V rovnici (1.10) a (1.11) můžeme opět vidět důkaz pozitivní semi-definitnosti Laplaceovy matice. Tedy vidíme, že kdybychom se pokusili zadefinovat jinou Laplaceovu matici a chtěli zachovat důležitý předpoklad pozitivní semi-definitnosti, je jednou z možností zajistit, aby se rovnala nějaké formě diskrétního Laplaceova operátoru. Nyní si ukážeme, že algebraická násobnost nuly přímo odpovídá počtu nepropojených grafů.

Věta 1.4.7. *Necht' G je konečný neorientovaný graf a L je příslušná Laplaceova matice. Necht' se G skládá z nepropojených souvislých grafů G_i , kde $i \in \{1, \dots, j\}$. Pak algebraická násobnost $\lambda = 0$ je rovna j .*

Důkaz. Začneme důkazem, že algebraická násobnost je větší nebo rovna j . Vyjdeme z lemmatu 1.4.5. s tím, že každému G_i přísluší Laplaceova matice L_i . Z toho lze stejným způsobem jako v důkazu věty 1.4.6. zjistit, že máme alespoň j lineárně nezávislých vlastních vektorů X_i , které splňují, že $LX_i = 0$ pro všechna i . Tím je dokázána první nerovnost.

Máme tedy alespoň j vlastních vektorů, kterým přísluší $\lambda = 0$, nyní dokážeme, že vektory X_i jsou jediné vlastní vektory náležící $\lambda = 0$, opomeneme-li lineární kombinace X_i . Nyní přeznačíme vrcholy G stejně jako v lemmatu 1.4.5. Tedy vrcholy G_1 označíme $1, \dots, k_1$, vrcholy G_2 označíme $k_1 + 1, \dots, k_1 + k_2$ a tak až k G_j . Vlastní vektory X_i odpovídající jednotlivým G_i budou vypadat stejně jako v důkazu věty 1.4.6. Pro větší přehlednost je tentokrát zapíšeme v maticovém tvaru, pro matici vlastních vektorů $X = (X_1^T, \dots, X_j^T)$

$$X = \begin{bmatrix} X_1^T \\ \vdots \\ X_j^T \end{bmatrix} = \begin{bmatrix} 1, \dots, 1 & 0, \dots, 0 & 0, \dots, 0 & \dots & 0, \dots, 0 & 0, \dots, 0 \\ 0, \dots, 0 & 1, \dots, 1 & 0, \dots, 0 & \ddots & 0, \dots, 0 & 0, \dots, 0 \\ 0, \dots, 0 & 0, \dots, 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0, \dots, 0 \\ 0, \dots, 0 & 0, \dots, 0 & \ddots & \ddots & 1, \dots, 1 & 0, \dots, 0 \\ 0, \dots, 0 & 0, \dots, 0 & \dots & 0, \dots, 0 & 0, \dots, 0 & 1, \dots, 1 \end{bmatrix} \quad (1.15)$$

Rovnice (1.12) tedy říká, že pro každé G_i bude mít vektor X_i jedničky pro sloupce L , které odpovídají vrcholům daného grafu G_i a všude jinde nuly.

Nyní vyjdeme z rovnice (1.9), kterou přenásobíme libovolným vlastním vektorem

$$X_m = (x_{m,1}, \dots, x_{m,k_1}, \dots, x_{m,k_2}, \dots, \dots, x_{m,k_{j-1}}, \dots, x_{m,k}) = (I_{m,1}, \dots, I_{m,j}),$$

to nám dává následující rovnici

$$LX_m = \begin{bmatrix} L_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & L_2 & \ddots & \ddots & & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & L_{j-1} & 0 \\ 0 & 0 & \dots & 0 & 0 & L_j \end{bmatrix} \begin{bmatrix} I_{m,1} \\ I_{m,2} \\ \vdots \\ \vdots \\ I_{m,j-1} \\ I_{m,j} \end{bmatrix} = \begin{bmatrix} L_1 I_{m,1} \\ L_2 I_{m,2} \\ \vdots \\ \vdots \\ L_{j-1} I_{m,j-1} \\ L_j I_{m,j} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix}. \quad (1.16)$$

Tím se nám ovšem hledání příslušných vlastních vektorů X_m rozpadá na j rovnic tvaru

$$L_i I_{m,i} = 0.$$

To jsou ovšem naše rovnice pro X_i . Tím jsme dokázali, že X_i jsou všechny lineárně nezávislé vlastní vektory L , které přísluší $\lambda = 0$. Tím jsme dokázali, že algebraická násobnost $\lambda = 0$ je právě j . \square

Zamysleme se ještě jednou nad tím co jsme získali. Máme Laplaceovu matici, kterou není těžké získat a její vlastní čísla, které není obtížné numericky spočítat, protože Laplaceova matice je pozitivně semi-definitní. Dále víme, že algebraická násobnost nuly odpovídá počtu nepropojených grafů. Tedy problém ověření souvislosti grafů je fakticky vyřešen.

1.5 Spektra vybraných grafů

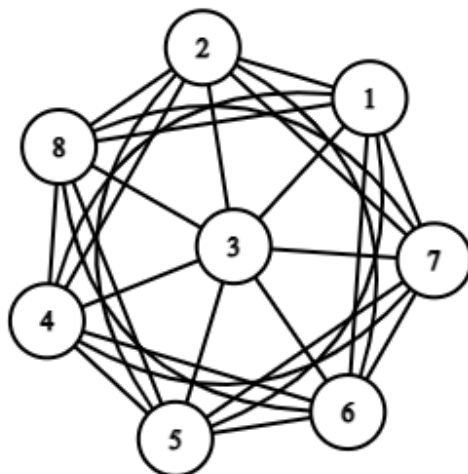
Nyní si zvolíme několik konkrétních tříd grafů, konkrétně úplný graf, kružnici, cestu, hvězdu a strom. U prvních čtyřech ukážeme Laplaceovu matici a nalezneme spektra. U stromů pak zmíníme některé vlastnosti a některé dokážeme.

1.5.1 Úplný graf

Definice 1.5.1. Neorientovaný graf K_k nazveme úplný, právě tehdy když

$$V = \{1, 2, \dots, k\} \wedge E = \binom{V}{2}.$$

Obrázek 1.3: Úplný graf



Věta 1.5.1. Necht' K_k je úplný graf s k vrcholy, jeho Laplaceova matice L je pak:

$$L(K_k) = \begin{bmatrix} k-1 & -1 & -1 & & -1 & -1 \\ -1 & k-1 & \ddots & \ddots & & -1 \\ -1 & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & -1 \\ -1 & & \ddots & \ddots & k-1 & -1 \\ -1 & -1 & & -1 & -1 & k-1 \end{bmatrix}$$

a vlastní čísla K_k jsou $\lambda_1 = 0$ a $\lambda_2, \dots, \lambda_k = k$ s algebraickou násobností $k-1$.

Důkaz. Úplný graf je souvislý, z čehož dle věty 1.4.1. plyne, že existuje právě jedno $\lambda_1 = 0$. Vyjdeme z definice vlastních čísel

$$\lambda_i X_i = LX_i.$$

V úplném grafu z každého vrcholu vede právě $k - 1$ hran, z předchozí matice $L(K_k)$ lze vidět, že

$$\lambda_i x_j^i = (k - 1)x_j^i - \sum_{\substack{l=1 \\ l \neq j}}^k x_l^i = kx_j^i - x_j^i - \sum_{\substack{l=1 \\ l \neq j}}^k x_l^i$$

Dále x_j^i doplní v $\sum_l x_l^i$ členy pro které $j \neq l$, z čehož vznikne

$$\lambda_i x_j^i = kx_j^i - \sum_{l=1}^k x_l^i.$$

Z věty 1.4.4. ovšem již víme, že $\sum_j x_j = 0$, což znamená, že

$$\lambda_i x_j^i = kx_j^i,$$

z čehož plyne, že $\lambda_i = k$ a po dosazení do definice $kX_i = LX_i$ pro $i \in \{2, \dots, k\}$

□

Z tvaru vlastních čísel úplného grafu je také vidět jedna zajímavá vlastnost algebraické konektivity. Velmi často totiž roste s velikostí grafů. To ovšem neplatí u normalizované Laplaceovy matice.

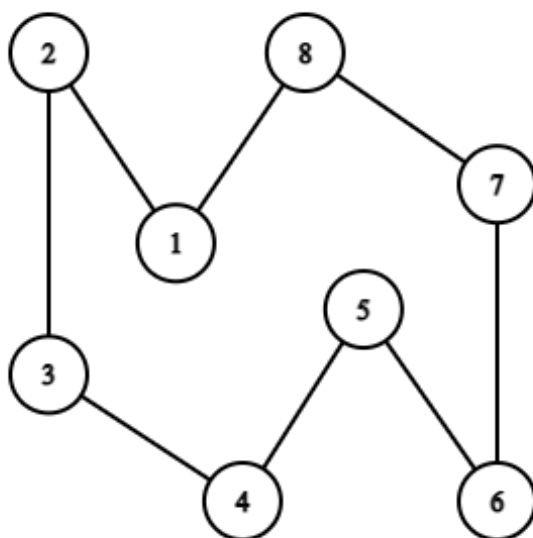
1.5.2 Kružnice

Kružnici jsme již jednou definovali, ale pro připomenutí.

Definice 1.5.2. Neorientovaný graf C_k nazveme kružnicí právě tehdy, když

$$V = \{1, 2, \dots, k\} \wedge E = \{\{i, i + 1\}; i = 1, \dots, k - 1\} \cup \{\{1, k\}\}.$$

Obrázek 1.4: Kružnice



Věta 1.5.2. Necht' C_k je kružnice s k vrcholy. Její Laplaceova matice L je pak:

$$L(C_k) = \begin{bmatrix} 2 & -1 & 0 & \dots & 0 & -1 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & 2 & -1 \\ -1 & 0 & \dots & 0 & -1 & 2 \end{bmatrix}$$

a vlastní čísla C_k jsou:

$$\lambda_m = 2 - 2 \cos\left(\frac{2\pi(m-1)}{k}\right)$$

pro $m \in \{1, \dots, k\}$ s algebraickou násobností pro všechna vlastní čísla, mimo $\lambda_m = 2$, rovnou dvěma.

Důkaz. Kružnice je souvislá, z čehož dle věty 1.4.1. plyne, že existuje právě jedno $\lambda_1 = 0$. Vyjdeme z článku [13]. Odhadneme řešení v následujícím tvaru pro $m > 1$:

$$z_k(j) = \exp\left(\frac{2\pi i(m-1)j}{k}\right),$$

kde $j \in \{1, \dots, k\}$. Na toto řešení pak aplikujeme Laplaceovu matici a algebraickými úpravami se ji pokusíme vyjádřit.

$$\begin{aligned} (Lz_m)(j) &= 2 \exp\left(\frac{2\pi i(m-1)j}{k}\right) - \exp\left(\frac{2\pi i(m-1)(j-1)}{k}\right) - \exp\left(\frac{2\pi i(m-1)(j+1)}{k}\right) \\ &= \exp\left(\frac{2\pi i(m-1)j}{k}\right) \left(2 - \exp\left(\frac{2\pi i(m-1)}{k}\right) - \exp\left(\frac{-2\pi i(m-1)}{k}\right)\right) \\ &= z_m(j) \left(2 - 2 \cos\left(\frac{2\pi(m-1)}{k}\right)\right). \end{aligned}$$

Nyní už jen využijeme definici vlastních čísel

$$\lambda_m z_m = Lz_m.$$

Z čehož plyne, že

$$\lambda_m = 2 - 2 \cos\left(\frac{2\pi(m-1)}{k}\right)$$

pro $m > 1$. Když si uvědomíme, že pro $m = 1$ máme

$$\lambda_m = 2 - 2 \cos\left(\frac{2\pi 0}{k}\right) = 2 - 2 = 0,$$

což odpovídá λ_1 , tak můžeme napsat, že $m = \{1, \dots, k\}$.

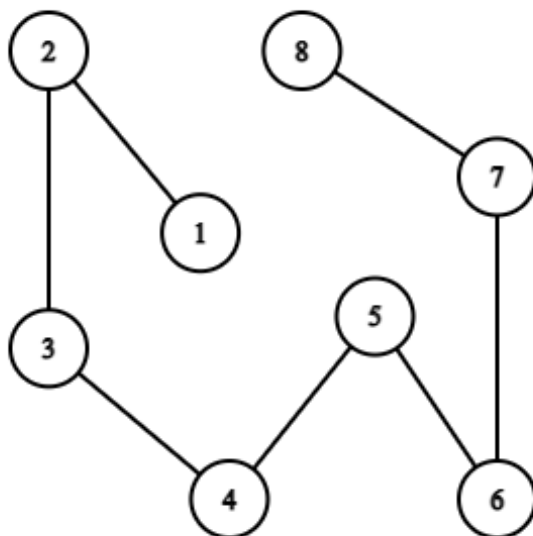
□

1.5.3 Cesta

Definice 1.5.3. Neorientovaný graf P_k nazveme cestou, právě tehdy když

$$V = \{1, 2, \dots, k\} \wedge E = \{\{i-1, i\}; i = 1, \dots, k\}.$$

Obrázek 1.5: Cesta



Věta 1.5.3. Necht' P_k je cesta s k vrcholy. Její Laplaceova matice L je pak:

$$L(P_k) = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & 2 & -1 \\ 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix}$$

a vlastní čísla P_k jsou:

$$\lambda_m = 2 - 2 \cos\left(\frac{\pi(m-1)}{k}\right)$$

pro $m = \{1, \dots, k\}$ s algebraickou násobností 1.

Důkaz. Pro důkaz si stačí uvědomit, že cestu získáme odebráním jedné hrany z kružnice. Tedy důkaz bude naprosto stejný jako ve větě 1.5.2. Jen tentokrát použijeme řešení ve tvaru

$$z_k(j) = \exp\left(\frac{\pi i(m-1)j}{k}\right),$$

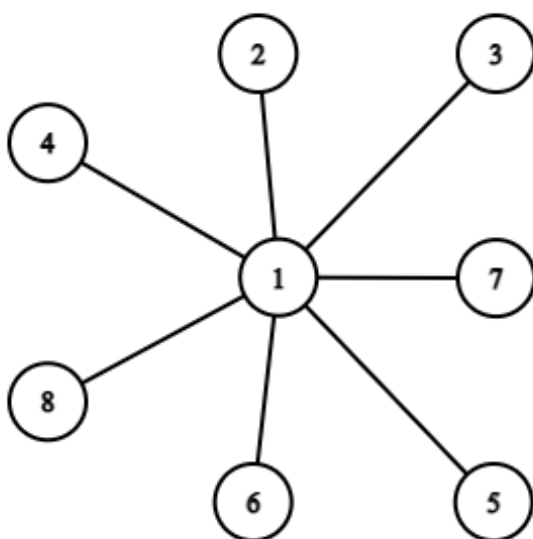
□

1.5.4 Hvězda

Definice 1.5.4. Neorientovaný graf S_k nazveme hvězdou právě tehdy, když

$$V = \{1, 2, \dots, k\} \wedge E = \{\{1, i\}; i = 2, \dots, k\}.$$

Obrázek 1.6: Hvězda



Věta 1.5.4. Necht' S_k je Hvězda s k vrcholy. Její Laplaceova matice L je pak:

$$L(S_k) = \begin{bmatrix} k-1 & -1 & -1 & \dots & -1 & -1 \\ -1 & 1 & 0 & \dots & 0 & 0 \\ -1 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ -1 & 0 & \ddots & \ddots & 1 & 0 \\ -1 & 0 & \dots & 0 & 0 & 1 \end{bmatrix}$$

a vlastní čísla S_k jsou:

$\lambda_1 = 0, \lambda_k = k$ a $\lambda_2, \dots, \lambda_{k-1} = 1$ s algebraickou násobností $k-2$.

Důkaz. Hvězda je souvislá, z čehož dle věty 1.4.1. plyne, že existuje právě jedno $\lambda_1 = 0$. Dále vyslovíme očividnou hypotézu, že $\lambda_k = k$. Té náleží vektor

$$X_k = (k-1, -1, \dots, -1).$$

Nyní vynásobíme vektor X_1 s maticí L . Výsledkem je

$$(k^2 - k, -k, \dots, -k) = kX_k.$$

Tím jsme ověřili, že $LX_k = kX_k$. Ukázali jsme tedy, že k je vlastní číslo. Dále budeme uvažovat odčítání sloupců matice po dvou:

$$L(e_i - e_{i+1}) = e_i - e_{i+1},$$

z čehož plyne, že $e_i - e_{i+1}$ je vlastní vektor s vlastním číslem 1. Tím jsme našli dalších minimálně $k - 2$ lineárně nezávislých vlastních vektorů. Známe tedy všech k vlastních čísel, a tím jsme našli spektrum L .

□

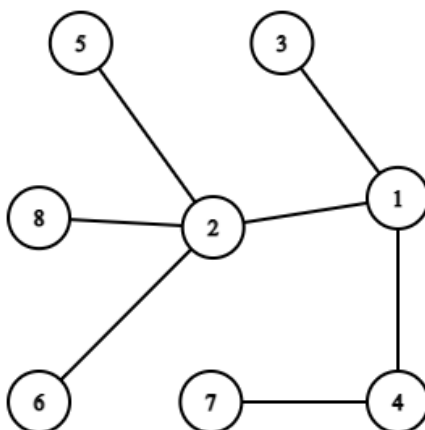
1.5.5 Strom

Jako poslední rozebereme spektrální vlastnosti stromů. Na rozdíl od předchozích typů grafů zde nezáiskáme obecně spektrum, ani obecný tvar Laplaceovy matice. Strom je totiž výrazně obecnější definice, než předchozí grafy. Je tedy znám pouze obecný tvar spektra extrémálních případů stromu, tedy hvězdy a cesty.

Laplaceova matice má ovšem výrazné aplikace v oblasti stromů, dokonce se s nimi pojí i jedna z jejich prvních aplikací v rámci elektrických obvodů. Proto se pokusíme alespoň ukázat nějaké vlastnosti spekter stromů. Začneme tím co vlastně strom je.

Definice 1.5.5. *Nechť G je konečný a souvislý neorientovaný graf. G pak nazvu stromem právě tehdy, když pro každý podgraf platí, že není kružnicí.*

Obrázek 1.7: Strom



Na obrázku 1.7 můžeme vidět graf stromu. Ať vybereme jakýkoliv vrchol a přejdeme z něj do jiného vrcholu cesta se nám rozvětví do dalších možností kam jít. Jestliže nepoužijeme dvakrát jednu hranu, tak se do výchozího vrcholu nevrátíme. Z tohoto důvodu jsou stromy velmi populární v teorii Markovských řetězců, ale například i kombinatorické teorii her, nebo Feynmanových diagramech. Abychom si uvědomili, jak vypadá Laplaceova matice stromu, je dobré pochopit, jak fungují stupně vrcholů stromu [22].

Věta 1.5.5. *Necht' G je strom s k vrcholy, pak pro aritmetický průměr stupňů vrcholů \deg_a , kde*

$$\deg_a = \sum_{i=1}^k \frac{\deg(v_i)}{k}$$

platí, že

$$\deg_a < 2. \quad (1.17)$$

Důkaz. Stačí si uvědomit jak vypadá strom, pak máme stupeň vrcholu

$$\deg_a = \frac{2(k-1)}{k} = 2 - \frac{2}{k},$$

z čehož je zřejmé, že průměr \deg_a musí být vždy menší než dva. □

Tato vlastnost samozřejmě není příliš překvapivá, když si uvědomíme jak vypadá graf stromu. Pro pochopení následující vlastnosti je nutné zadefinovat pojem Laplaceova energie.

Definice 1.5.6. *Necht' G je konečný neorientovaný graf s k vrcholy. Necht' λ_i jsou jeho vlastní čísla a \deg_a průměrný stupeň vrcholu, pak Laplaceovou energií E_L nazvu*

$$E_L = \sum_{i=1}^k |\lambda_i - \deg_a|$$

Nyní tedy k samotné vlastnosti Laplaceovy matice stromu.

Věta 1.5.6. *Necht' G je strom a \deg_a průměrný stupeň jeho vrcholů, pak pro nejméně polovinu jeho vlastních čísel λ_i platí, že*

$$\lambda_i < \deg_a < 2$$

Důkaz. Tuto nerovnost rigorózně dokazovat nebudeme. Samotný rigorózní důkaz je možné najít [8]. Důkaz je ovšem velmi technický a dlouhý.

Přesto se, pro pochopení, pokusíme alespoň heuristicky danou větu vysvětlit ve slabší verzi

$$\lambda_i < 2.$$

První se zamyslíme, co nám vlastně říká věta 1.5.5, o Laplaceově matici. Samozřejmě říká, že průměr prvků na diagonále bude menší než dva. Pokud víme jak vypadají běžné grafy stromů, může toho být dosaženo tak, že bude na diagonále velké množství jedniček. Mimo případů, kdy vezmeme cestu, či strom takového typu, že velká část podgrafů bude cesta. Ovšem i vrcholy takovýchto posloupností budou mít stupeň příslušných vrcholů menší než dva.

Z kapitoly o algebraické konektivě vyplývá spojitost mezi vlastními čísly a "propojeností" grafu. Jednou ze snah o formalizaci pojmu "propojenost" je právě Laplaceova energie E_L . Zkusme se zamyslet, jaké dva stromy by mohli být "nejpropojenější". Dva očividní kandidáti jsou extrémní případy, tedy cesta a hvězda. Ovšem když vyjdeme z věty 1.5.4., tak u hvězdy je zřejmé, že větu 1.5.6. splňuje. Opět i z věty 1.5.3. a znalosti chování funkce kosinus vyplývá, že nejméně polovina vlastních čísel je menší než dva. To je jedna ze základních myšlenek celého důkazu. Bylo by nutné dokázat a využít, že právě cesta je strom s nejmenší Laplaceovou energií. □

Na základě této věty jde dovodit [8], že u stromů se kterými se běžně setkáváme v aplikacích, existuje jen několik velkých vlastních čísel a zbylá jsou velmi malá. Jako extrémní případ můžeme ukázat hvězdu, její spektrum také použijeme v další větě.

Věta 1.5.7. *Necht' G je strom s k vrcholy a λ_i jsou jeho vlastní čísla, pak*

$$\max_{i \in \{1, \dots, k\}} \lambda_i \leq k.$$

Důkaz. Dle věty z [7] víme, že hvězda má větší největší vlastní číslo než jakýkoliv jiný strom se stejným počtem vrcholů. Ve spojení s větou 1.5.4. z toho triviálně plyne tvrzení věty. □

Laplaceova matice má výraznou aplikaci v otázce koster grafu. Vyvrcholením její aplikace je Kirchhoffova věta, jež řeší jeden z nejzákladnějších matematických problémů v této teorii. Tím je počet koster která přísluší grafu [11].

Definice 1.5.7. *Kostrou grafu $G = (V, E)$ nazvu jeho libovolný souvislý podgraf $G' = (V, E')$, který je zároveň stromem.*

Věta 1.5.8 (Kirchhoffova věta). *Necht' G je konečný a souvislý neorientovaný graf a n_G je počet koster grafu G . Necht' L je jeho Laplaceova matice, $\lambda_1 \dots \lambda_k$ jsou Laplaceova vlastní čísla a $\lambda_1 = 0$, pak*

$$n_G = \prod_{i=2}^k \lambda_i.$$

Důkaz. Věta se běžně dokazuje na základním kurzu teorie grafů (včetně FJFI), proto jí nebudeme znovu dokazovat. Přidáme odkaz na originální důkaz [12]. □

1.6 Algebraická konektivita

Nyní v krátkosti teoreticky zadefinujeme algebraickou konektivitu a Fiedlerův vektor [6]. Víc se tomuto tématu budeme věnovat v podkapitole o clusteringu Markovských řetězců a kapitole o spectral clusteringu, kde se oba pojmy hojně využívají.

Definice 1.6.1. *Necht' G je konečný neorientovaný graf a L je příslušná Laplaceova matice. Algebraickou konektivitou λ_2 nazvu druhé nejmenší vlastní číslo matice L a příslušný vlastní vektor X_2 nazvu Fiedlerovým vektorem.*

Jak si později ukážeme, při vhodně zdefinovaném L je možné pro některé aplikace podmínku neorientovanosti vypustit. Stejně tak si ukážeme, že algebraická konektivita funguje i pro normalizovanou Laplaceovu matici.

Věta 1.6.1. *Necht' G je konečný neorientovaný graf a L je příslušná Laplaceova matice, pak G je souvislý právě tehdy, když $\lambda_2 \neq 0$.*

Důkaz. V podstatě jen jiná formulace věty 1.4.6.

□

Nyní si uvědomíme, co vlastně počítáme, když hledáme algebraickou konektivitu.

Věta 1.6.2. *Nechť G je konečný a souvislý neorientovaný graf. Poté výpočet algebraické konektivity λ_2 je ekvivalentní problému:*

$$\lambda_2 = \min_{\sum_i x_i = 0 \wedge x \neq 0} \sum_{(i,j) \in E} (x_i - x_j)^2.$$

Důkaz. Vyjdeme z Courant-Fischerovy věty [10]:

$$\lambda_2 = \min_x \frac{x^T L x}{x^T x}.$$

Dále si uvědomíme, že kvůli kolmosti můžeme jmenovatele vypustit, protože $x^T x = 1$. Poté provedeme několik úprav

$$x^T L x = \sum_{(i,j)=1} x_i L_{i,j} x_j = \sum_{(i,j)=1} x_i L_{i,j} x_j = \sum_{(i,j) \in E} (x_i^2 + x_j^2 - 2x_i x_j) = \sum_{(i,j) \in E} (x_i - x_j)^2$$

Poté si musíme uvědomit, že sčítáme přes kolmé X . Tedy využijeme podmínku $\sum_i x_i = 0$.

□

Kapitola 2

Laplaceova matice pro orientované grafy

2.1 Různé definice Laplaceových matic pro orientované grafy

Ačkoliv existuje poměrně dobrá teorie k Laplaceovým maticím neorientovaných grafů, o orientovaných grafech toho bylo napsáno velmi málo. Laplaceova matice nachází své aplikace z velké části v oblastech, kde se využívají orientované grafy. Většinou v praxi bývá s orientovaným grafem provedena symetrizace, kterou je změněn na graf neorientovaný. Toto má ovšem významný problém, protože celý postup zcela ignoruje jestli je mezi vrcholy orientovaná hrana v jednom či oběma směry. To může vést k chybám jako například v Příkladu 3 v pozdějších numerických simulacích. Bývá to uváděno i jako problém v aplikacích [3].

V této části se pokusíme zkonstruovat Laplaceovu matici L pro orientovaný graf tak, aby měla nějaké využitelné vlastnosti. Jako kritérium zadáme, aby měla podobné vlastnosti jako klasická Laplaceova matice na neorientovaných grafech.

Jako první si vysvětlíme problémy, kterým budeme čelit při zobecnění Laplaceovy matice pro neorientované grafy.

Matice sousednosti A pro orientovaný graf

Zobecnit naši definici matice sousednosti do orientovaných grafů je triviální, stačí nahradit $\{v_i, v_j\}$ uspořádanou dvojicí (v_i, v_j) . Ovšem zde již vidíme některé problémy. Matice A nebude v obecných případech symetrická a tedy ani L nemůže být symetrická. Z čehož plyne, že vlastní čísla nemusí být reálná, ani kladná, algebraická a geometrická násobnost se nemusí rovnat atd.

Matice stupňovitosti D pro orientovaný graf

Vzhledem k tomu, že stupeň vrcholu je u orientovaných grafů dvouprvkový vektor. Matice D vůbec nebude moci být maticí, protože má na diagonále dvouprvkový vektor. Bude se tedy jednat o tenzor 3. řádu o rozměrech $k \times k \times 2$, kde k je počet vrcholů.

Definice využívající jednoduché zobecnění je zjevně nesmyslná, kvůli rozměru daných objektů. To je ovšem možné řešit součtem vycházejících a přicházejících hran pro každý vrchol [16]. To se formalizuje součtem D_{in} a D_{out} , které jsou definované stejně jako matice stupňovitosti, jen mají na diagonále $\deg_{in}(v_i)$ respektive $\deg_{out}(v_i)$.

Definice 2.1.1. *Necht' G s k vrcholy je konečný orientovaný graf. Necht' D_{in} je matice stupňovitosti pro příchozí hrany do vrcholu, D_{out} je matice stupňovitosti pro hrany vycházející z vrcholu. Necht' A je matice sousednosti grafu G . Laplaceovu matici pro orientované grafy $L \in \mathbb{R}^{k,k}$, pak nazvu*

$$L = D_{in} + D_{out} - A.$$

Laplaceova matice pro orientované grafy se používá v některých aplikacích, leč pro naše potřeby nemá v podstatě žádné využitelné vlastnosti. Zjevně nemusí být symetrická, což dosti komplikuje jakoukoliv spektrální analýzu. Jediné spektrálně zajímavé vlastnosti, které jsme získali, se týkaly silně souvislých grafů.

Věta 2.1.1. *Necht' G je silně souvislý orientovaný graf. Pak je její Laplaceova matice pro orientované grafy L regulární.*

Důkaz. G je silně souvislý, tedy do každého vrcholu musí přicházet alespoň jedna hrana a zároveň alespoň jedna odcházet. To znamená, že nikdy nemůže dojít k tomu, že by prvek na diagonále byl ve svém sloupci nebo řádku jediný. Z toho plyne, že součet je vždy ostře větší než ostatní prvky v řádku nebo sloupci. Z toho vyplývá, že L je silně diagonálně dominantní a taková matice je vždy regulární. □

Povedlo se nám dokázat tuto vlastnost pro silně souvislé orientované grafy. Vyslovme ovšem hypotézu, že tato věta bude platit pro libovolný orientovaný graf bez osamocené vrcholu. Osamoceným vrcholem myslíme vrchol, který nenáleží žádné hraně. Tuto hypotézu zatím nebylo možné dokázat a je ponechána k pozdějšímu výzkumu.

Věta 2.1.2. *Necht' G je konečný a silně souvislý orientovaný graf. Necht' L je její Laplaceova matice pro orientované grafy a λ její vlastní čísla, pak*

$$Re(\lambda_i) \geq 0.$$

Důkaz. V důkaze věty 2.1.1. již bylo ukázáno, že L je silně diagonálně dominantní matice a pro ty platí, že $Re(\lambda_i) \geq 0$. □

S výjimkou těchto dvou vlastností se nepodařilo nalézt žádné další významné spektrální vlastnosti této matice. To je ve shodě s běžnou literaturou, která se spektrem této matice nezajímá.

Existuje několik dalších způsobů jak definovat Laplaceovu matici pro orientované grafy. Většinou toho ovšem o jejím spektru obecně umíme říci velmi málo. Proto je v praxi většinou užívána symetrizace původního orientovaného grafu.

2.2 Abstraktní Laplaceova matice

Pokusíme se o novou definici za pomoci matice incidence. Na to potřebujeme zobecnit matici incidence do orientovaných grafů. To bude paradoxně jednodušší než pro neorientované grafy.

Definice 2.2.1. Necht' $G = (V, E)$ je konečný orientovaný graf. Necht' $V = \{v_1, \dots, v_k\}$ a $E = \{e_1, \dots, e_l\}$. Maticí Incidence $M \in \mathbb{R}^{k,l}$ nazveme matici, kde

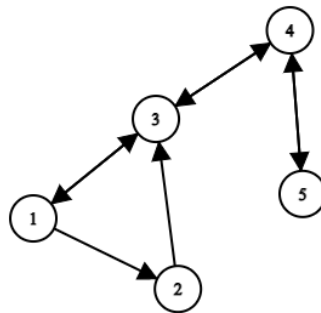
$$\begin{aligned} M_{i,j} &= +1, \text{ když } e_j \text{ je hrana vycházející z } v_i \\ M_{i,j} &= -1, \text{ když } e_j \text{ je hrana přicházející do } v_i \\ M_{i,k} &= 0 \text{ jinak} \end{aligned}$$

Zkusme si rozmyslet vztah k definici původní matice incidence 1.4.3. Novou maticí incidence bychom mohli získat tak, že orientovaný graf symetrizujeme a najdeme klasickou maticí incidence jeho symetrizace. Poté pouze odebereme sloupce, které náležejí chybějícím orientovaným hranám.

Z toho nám taktéž vychází zajímavá vlastnost. Pokud pro všechny hrany orientovaného grafu e_i , kde $e_i = (x, y)$, existuje $e_j = (y, x)$, bude matice incidence takového orientovaného grafu stejná jako matice incidence symetrizace daného grafu. Uvažujme, že sloupce můžeme libovolně zaměnit. To je důležitá vlastnost, nýbrž nám umožňuje brát v tomto případě grafy neorientované jen jako speciální případ grafů orientovaných.

Kvůli významnosti tohoto pojmu si jej ukážeme na konkrétním grafu G

Obrázek 2.1: Orientovaný graf G



Tomu náležejí následující matice incidence M .

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix}.$$

Jasně z toho vidíme, že matice se bude lišit podle toho, jak zvolíme číslování hran. Nyní přikročíme k samotné definici abstraktní Laplaceovy matice.

Definice 2.2.2. Necht' G s k vrcholy je konečný orientovaný graf. Necht' M je jeho matice incidence. Abstraktní Laplaceovu matici $L^a \in \mathbb{R}^{k,k}$ pak definují jako

$$L^a = \frac{1}{2} M^T M. \quad (2.1)$$

Tato definice vychází ze specifického zadefinování matice incidence, viz definice 1.4.3. a jejího přenesení do orientovaných grafů. Ačkoliv je Laplaceova matice pro neorientované grafy běžně definována přes matici incidence, tak tuto úpravu definice pro orientované grafy se nepodařilo nalézt kdekoli v dostupné literatuře.

Věta 2.2.1. *Necht' G je konečný orientovaný graf, pro který platí, že pro libovolné $e_i \in E$ kde $e_i = (x, y)$ existuje $e_j \in E$ kde $e_j = (y, x)$ a necht' G' je symetrizace G . Pak Laplaceova matice G' a abstraktní Laplaceova matice G je stejná.*

Důkaz. Důkaz je zřejmý ze způsobu definice abstraktní Laplaceovy matice a věty 1.4.2. □

Tato věta je naprosto triviální a zřejmá, ovšem velice důležitá. Jak již jsme psali v úvodu, neorientované grafy jsou v určitém smyslu jen speciálním případem grafů orientovaných. Je proto velmi praktické, aby definice objektů pro neorientované grafy odpovídaly stejně se chovajícím grafům orientovaným. Kdybychom se vrátili k příkladu z úvodu, bylo by zvláštní, kdybychom měli graf dálniční sítě, kde by všechny silnice byly obousměrné a na naší Laplaceově matici by se něco změnilo, kdybychom přidali informaci o orientaci.

2.3 Spektrální analýza abstraktní Laplaceovy matice

Nyní se pokusíme zjistit, jestli má abstraktní Laplaceova matice stejně silné spektrální vlastnosti jako klasická Laplaceova matice.

Věta 2.3.1. *Necht' G je konečný orientovaný graf. Necht' L^a je abstraktní Laplaceova matice G , λ_i vlastní čísla L^a a X_i příslušné vlastní vektory, pak*

- a) L^a je reálná,
- b) L^a je symetrická,
- c) λ_i jsou reálná,
- d) λ_i jsou nezáporná,
- e) geometrická a algebraická násobnost λ_i je stejná pro všechna i ,
- f) L^a je diagonalizovatelná,
- g) X_i jsou ortogonální.

Důkaz. Důkazy jsou víceméně zjevné.

- a) M je reálná matice a součin dvou reálných matic je opět reálná matice.
- b) Vychází z definice L^a jako MM^T
- c) Již jsme ve ukázali, že L^a je reálná a symetrická, což ze základů lineární algebry znamená, že vlastní čísla jsou reálná.
- d) Vlastnost je zjevná, protože $L^a = \frac{1}{2}MM^T$. Důkaz je stejný jako ve větě 1.4.3.
- e) Protože se nám definicí abstraktní Laplaceovy matice povedlo zachovat reálnost a symetrii i pro orientované grafy, tak tato vlastnost bude platit stejně jako u klasické Laplaceovy matice.
- f) Matice má z předchozího bodu stejnou algebraickou a geometrickou násobnost a z toho opět plyne diagonalizovatelnost.
- g) Opět vychází ze silných vlastností reálných symetrických matic. □

Všechny tyto vlastnosti vychází jen z toho, že jsme abstraktní Laplaceovu matici definovali jako součin reálné matice M a její matice transponované M^T . Z důvodu výrazně jiného způsobu definice abstraktní Laplaceovy matice a Laplaceovy matice zvolíme trochu odlišný způsob stavby teorie. Singularita byla u Laplaceovy matice zjevná, zde to bude poněkud složitější.

Věta 2.3.2. *Nechť G je konečný a orientovaný graf. Nechť je dána abstraktní Laplaceova matice L^a , pak L^a je singulární.*

Důkaz. Vyjdeme z definice 2.2.1. Když $i = j$. Nechť n je počet orientovaných hran G , pak

$$(MM^T)_{i,i} = \sum_{l=1}^n M_{i,l}M_{l,i}^T = \sum_{l=1}^n M_{i,l}M_{i,l} = \sum_{l=1}^n M_{i,l}^2 = \sum_{e \in V} M_{e,v_i}^2 = \sum_{e \in v_i} 1 = \deg_{in}(v_i) + \deg_{out}(v_i). \quad (2.2)$$

Využili jsme ten samý postup, který jsme použili například v důkazu věty 1.4.2. Dokázali jsme, že součin MM^T vytvoří na diagonále součet všech vycházejících a přicházejících hran.

Pro $i \neq j$, v případě že mezi vrcholy v_i a v_j existuje hrana s libovolnou orientací, budeme pokračovat stejným způsobem.

$$(MM^T)_{i,j} = \sum_{l=1}^n M_{i,l}M_{l,j}^T = \sum_{l=1}^n M_{i,l}M_{j,l} = \sum_{e \in V} M_{e,v_i}M_{e,v_j} \quad (2.3)$$

Vidíme, že protože matice M již nemá obecně tvar (1.5), odpověď není tak jednoznačná. Jestliže mezi nimi existuje právě jedna ze hran (v_i, v_j) , (v_j, v_i) pak,

$$\sum_{e \in V} M_{e,v_i}M_{e,v_j} = -1 \quad (2.4)$$

Pokud existují obě hrany mezi (v_i, v_j) , (v_j, v_i) , pak

$$\sum_{e \in V} M_{e,v_i}M_{e,v_j} = -2 \quad (2.5)$$

Pokud $i \neq j$ a mezi v_i a v_j neexistuje hrana libovolné orientace, pak z násobení matic vyplývá, že

$$MM^T_{i,j} = 0.$$

Tímto jsme v podstatě ukázali tvar abstraktní Laplaceovy matice. Stačí náš výsledek vydělit dvěma. Na diagonále je polovina součtu přicházejících a vycházejících hran $\frac{\deg_{in}(v_i) + \deg_{out}(v_i)}{2}$. Mimo jsou pak na místech připojených vrcholů buď -0.5 nebo -1 podle toho, jestli je mezi v_i a v_j hrana jen jedné orientace nebo obou.

Z toho ovšem vyplývá, že na každém řádku je mimo diagonálu -0.5 za každou hranu přicházející do vrcholu, či hranu vycházející z daného vrcholu. Nebo -1 pokud existují obě. Ovšem na diagonále máme $\frac{\deg_{in}(v_i) + \deg_{out}(v_i)}{2}$. Tedy součet každého řádku abstraktní Laplaceovy matice je roven nule. Z čehož nám plyne singularita. □

V podstatě jsme tedy dokázali, že stejně jako u klasické Laplaceovy matice i u abstraktní je součet všech řádků roven nule. Z toho plyne, že L^a je slabě diagonálně dominantní. Nyní využijeme několik dalších důsledků této vlastnosti.

Věta 2.3.3. *Nechť G je konečný orientovaný graf. Nechť je dána jeho abstraktní Laplaceova matice L^a , pak alespoň jedno z jejích vlastních čísel $\lambda_1 = 0$ a příslušný vlastní vektor $X_1 = (1, \dots, 1)^T$.*

Důkaz. Důkaz věty je zřejmý z předchozí věty 2.3.2. □

Věta 2.3.4. *Necht' G je konečný orientovaný graf. Necht' je dána abstraktní Laplaceova matice L^a , pak L^a je pozitivně semi-definitní.*

Důkaz. Vychází ze způsobu definice abstraktní Laplaceovy matice jako MM^T dle 2.3.2. □

Vzhledem k tomu, co jsem již o spektru abstraktní Laplaceovy matice dokázali, nás příliš nepřekvapí, že i vlastnost prvků vlastních vektorů X_i zůstane stejná.

Věta 2.3.5. *Necht' G je konečný orientovaný graf a L^a je příslušná abstraktní Laplaceova matice. Necht' λ_i jsou její vlastní čísla a $X_i = (x_{i,1}, \dots, x_{i,k})$ příslušné vlastní vektory. Pak pro každé X_i , které nepřísluší $\lambda = 0$, platí, že*

$$\sum_{j=1}^k x_{i,j} = 0. \quad (2.6)$$

Důkaz. Vidíme, že jsme ve větě 2.3.1. již dokázali, že vektory X_i jsou ortogonální. Z věty 2.3.3. dále známe tvar $X_1 = (1, \dots, 1)^T$, což znamená, že důkaz je stejný jako u věty 1.4.4. o klasické Laplaceově matici. □

Nyní krátce probereme i věty o souvislosti grafů a jejich propojení se spektrem abstraktní Laplaceovy matice. Požadavek souvislosti pouze nahradíme požadavkem slabé souvislosti. Důkazy jsou buď úplně stejné, nebo extrémně podobné, jako u klasické Laplaceovy matice. To vychází z toho, že jsme pro abstraktní Laplaceovu matici již ukázali stejné vlastnosti jako jsme měli pro klasickou. Druhou věcí je, že pro všechny tyto vlastnosti je zjevně důležité jestli existuje nějaká hrana, její orientace je podružná.

Lemma 2.3.6. *Necht' G je konečný orientovaný graf a L^a je příslušná abstraktní Laplaceova matice. Necht' G_i , $i \in \{1, \dots, j\}$ jsou nepropojené slabě souvislé grafy, ze kterých se skládá G . Pak je možné vrcholy G přeznačit tak, aby Laplaceova matice G byla blokově diagonální s tím, že jednotlivé bloky původní matice L^a jsou abstraktní Laplaceovy matice jednotlivých grafů G_i .*

Důkaz. Stačí si uvědomit, že připojení vrcholů se při naší definici zachovává i pro slabou souvislost a tak můžeme použít důkaz lemma 1.4.5. □

Důležité je ještě zdůraznit, že silně souvislé grafy jsou podmnožinou grafů souvislých slabě.

Věta 2.3.7. *Necht' G je konečný orientovaný graf a L^a je příslušná abstraktní Laplaceova matice, pak G je slabě souvislý právě tehdy, když algebraická násobnost $\lambda = 0$ je rovna jedné.*

Důkaz. Stejně jako v případě minulé věty si stačí uvědomit, že pro důkaz věty 1.4.6. je nepodstatné, zda se prvky mimo diagonálu rovnají $-0, 5$, nebo -1 . Důležité je, že na místech připojených vrcholů je nenulové číslo. □

Věta 2.3.8. *Necht' G je konečný a orientovaný graf a L^a je příslušná abstraktní Laplaceova matice. Necht' se G skládá z nepropojených slabě souvislých grafů G_i , kde $i \in \{1, \dots, j\}$. Pak algebraická násobnost $\lambda = 0$ je rovna j .*

Důkaz. Využijeme větu 2.3.7. a postupujeme přesně podle věty 1.4.7. Důkaz není podobný, ale naprosto stejný.

□

Jak vidíme, definice abstraktní Laplaceovy matice nám garantuje stejné základní spektrální vlastnosti jako klasická Laplaceova matice, což by mělo znamenat, že v následných aplikacích by měla fungovat obdobně jako klasická verze.

Kapitola 3

Náhodné procházky

3.1 Laplaceovy matice pro Markovské řetězce

V této podkapitole se budeme zabývat propojením teorie Markovských řetězců a Laplaceových matic [20]. Budeme uvažovat diskrétní Markovovské řetězce. Pro ty potřebujeme zavést množinu stavů $S = \{1, 2, \dots, s\}$. Ten bude u nás odpovídat množině vrcholů. Dále pak potřebujeme zavést diskrétní čas $T = \{1, 2, \dots\}$. Markovský řetězec je zadán dvěma strukturami. Maticí přechodu a vektorem pravděpodobností. Matice přechodu je obecně definována jako $P = [p_{ij}]$ kde $i, j \in S$ a p_{ij} značí pravděpodobnost přechodu ze stavu i do stavu j . Vektor pravděpodobností $p(t) = [p_1(t), p_2(t), \dots, p_s(t)]$ kde $t \in T$ nám pak udává, s jakou pravděpodobností se nacházíme v každém stavu z S , což může záviset na počátečních podmínkách. Markovův řetězec je pak popsán rovnicí $p(t+1) = p(t) * P$. Prvky vektoru $p(t)$ i matice P jsou pravděpodobnosti, tedy by měly být z intervalu $\langle 0, 1 \rangle$. Protože se jedná o úplný rozklad jevu (V každém kroku je nutné změnit stav.), tak P musí být stochastická matice. Stochastická matice je reálná čtvercová matice, kde součet každého řádku musí být jedna. Zmíňme ještě, že je možné provést obecnější definici, kde i P bude závislá na t .

V případě, že se čtenář ještě nesetkal s teorií Markovských řetězců, bude pro pochopení jednodušší přejít přímo k náhodným procházkám na grafech. Budeme tedy uvažovat náhodnou procházku na neorientovaném grafu [19]. Tu zadáme tak, že v každém kroku je možné přejít z libovolného vrcholu $v_i \in V$ do každého vrcholu v_j , pokud existuje $e = \{v_i, v_j\}$ s tím, že do každého vrcholu v_j přejdeme s pravděpodobností $\frac{1}{\deg(v_j)}$. V případě že bychom měli jeden samostatný vrchol a řetězec by začal v něm, prohlásíme Markovský řetězec za ukončený. Mohli bychom problém zavést obecněji a nerozložit pravděpodobnosti rovnoměrně. To by bylo velmi jednoduché, stačilo by hranám přiřadit váhy ω . Ovšem toto zavedení nám umožňuje elegantně zavést matici přechodu, využívající původní matici sousednosti [2].

Definice 3.1.1. *Necht' G je konečný a souvislý neorientovaný graf s k vrcholy. Necht' A je matice sousednosti a D matice stupňovitosti grafu G . Pak matici přechodu $P = \mathbb{R}^{k,k}$ definujeme jako*

$$P = D^{-1}A \tag{3.1}$$

Nyní zavedeme normalizovanou Laplaceovu matici. Její souvislost s náhodnými procházkami nemusí být vidět na první pohled. Ovšem ukážeme souvislost mezi spektry normalizované Laplaceovy matice a matice přechodu [9].

Definice 3.1.2. *Necht' G s k vrcholy je konečný a souvislý neorientovaný graf. Necht' L je jeho Laplaceova matice a D jeho matice stupňovitosti. Pak normalizovanou Laplaceovu matici $L^n = \mathbb{R}^{k,k}$ definujeme jako*

$$L^n = D^{-1/2}LD^{-1/2} \quad (3.2)$$

Matice zjevně zachovává symetrii, protože A je symetrická a $D^{-1/2}$ je diagonální matice. Jako mezistupeň si zadefinujeme Laplaceovu matici pro náhodné procházky. Jak brzy ukážeme, tato matice má velmi blízko k matici přechodu.

Definice 3.1.3. *Necht' G s k vrcholy je konečný a souvislý neorientovaný graf. Necht' L je jeho Laplaceova matice a D jeho matice stupňovitosti. Pak Laplaceovu matici pro náhodné procházky $L' = \mathbb{R}^{k,k}$ definujeme jako*

$$L' = D^{-1}L \quad (3.3)$$

Právě jsme si zavedli definice různých matic, které se používají pro popisy náhodných procházek. Nyní se je pokusíme propojit. Laplaceovu matici pro náhodné procházky můžeme přepsat jako

$$L' = D^{-1}L = D^{-1}(D - A) = I - D^{-1}A = I - P. \quad (3.4)$$

Z tohoto tvaru jsou dobře vidět základní vlastnosti Laplaceovy matice pro náhodné procházky. P je ze znalosti Markovských řetězců vždy stochastická matice. Je tedy zřejmé, že součet řádků L' bude nula. Z definice P vidíme, že P nemusí být symetrická. Tedy ani L' nemusí být symetrická. Ovšem pro potřeby spektrální analýzy by se nám hodila vlastnost symetrie. Převédeme tedy problém hledání jejích vlastních čísel na normalizovanou Laplaceovu matici.

Věta 3.1.1. *Necht' G je konečný a souvislý neorientovaný graf. Pak L' a L^n grafu G jsou podobné matice.*

Důkaz. Začneme definicí normalizované Laplaceovy matice

$$L^n = D^{-1/2}LD^{-1/2} = D^{-1/2}(D - A)D^{-1/2} = I - D^{-1/2}AD^{-1/2}. \quad (3.5)$$

Z toho vyjádříme, že

$$A = D^{1/2}(I - L^n)D^{1/2} \quad (3.6)$$

To dosadíme do (3.3)

$$L' = D^{-1}L = D^{-1}(D - A) = I - D^{-1}A = I - D^{-1/2}(I - L^n)D^{1/2} = D^{-1/2}L^nD^{1/2} \quad (3.7)$$

Což znamená, že jsou L' a L^n podobné z definice podobnosti matic.

□

Nyní jsme dokončili přípravné práce a budeme pokračovat k hlavním výsledkům teorie.

Věta 3.1.2. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L^n je normalizovaná Laplaceova matice G , X_i jsou její vlastní vektory a λ_i její vlastní čísla. Necht' P je matice přechodu náhodné procházky na G , Y_i jsou její vlastní vektory a μ_i její vlastní čísla. Necht' D je matice stupňovitosti G , pak*

$$X_i = D^{1/2}Y_i \wedge \lambda_i = 1 - \mu_i.$$

Důkaz. Začneme přechodem spektra normalizované Laplaceovy matice L^n na spektrum Laplaceovy matice pro náhodné procházky L' . Z věty 3.1.1. víme, že L^n a L' jsou podobné matice, tedy λ_i jsou vlastními čísly matice L' . Dále z rovnice (3.7) a znalosti vlastností podobných matic víme, že vlastní vektory L' jsou $D^{-1/2}X_i$. Tím máme první část důkazu hotovou.

Nyní přejdeme ze spektra L' ke spektru P . Na důkaz stačí dosadit do definice vlastních čísel Laplaceovu matici pro náhodné procházky,

$$L'D^{-1/2}X_i = \lambda_i D^{-1/2}X_i,$$

poté dosadíme z rovnice (3.4),

$$(I - P)D^{-1/2}X_i = \lambda_i D^{-1/2}X_i,$$

což nakonec upravíme na

$$PD^{-1/2}X_i = (1 - \lambda_i)D^{-1/2}X_i.$$

Z toho je jasné vidět že vlastní vektory se zachovaly a vlastní čísla se jen posunula o jedna. Což bylo očekávané, protože L' je k P v podstatě jen posunutí. Tedy $D^{-1/2}X_i = Y_i$ a $\mu = 1 - \lambda$. □

Tímto jsme tedy našli způsob, jak se vyhnout komplikovanému hledání spekter matice přechodu a Laplaceovy matice pro náhodné procházky a nahradit je normalizovanou Laplaceovou maticí. Jejím spektrálním vlastnostem se budeme věnovat v další podkapitole.

3.2 Spektrální analýza normalizované Laplaceovy matice

Nyní přistoupíme ke zkoumání spektra normalizované Laplaceovy matice. Na rozdíl od klasické Laplaceovy matice a abstraktní Laplaceovy matice bude postrádat některé vlastnosti běžně asociované s Laplaceovou maticí. Součet řádků se nebude muset rovnat nule a matice nemusí být diagonálně dominantní.

Věta 3.2.1. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L^n je jeho normalizovaná Laplaceova matice. λ_i jsou vlastní čísla a X_i vlastní vektory L^n . Pak*

- a) L^n je reálná,
- b) L^n je symetrická,
- c) algebraická a geometrická násobnost λ_i je stejná,
- d) L^n je diagonalizovatelná,
- e) X_i jsou ortogonální.

Důkaz. a) Z rovnice máme (3.4), že

$$L^n = I - D^{-1/2}AD^{-1/2}.$$

Jednotková matice I je bezpochyby reálná. D je diagonální a kladná, tedy její odmocnina je reálná, což znamená, že odčítáme a násobíme reálné matice, tedy L^n je taktéž reálná.

b) Opět použijeme rovnici (3.4). Jednotková matice I je symetrická. A je pro neorientované grafy taktéž symetrická. $D^{-1/2}$ je diagonální matice, tedy i $D^{-1/2}AD^{-1/2}$ je symetrická. Tedy odečítáme dvě symetrické matice a z toho plyne, že L^n je symetrická.

- c) Pouze důsledek reálné symetrické matice.
- d) Vychází z předchozího bodu.
- e) Opět jen vlastnost reálných symetrických matic.

□

Věta 3.2.2. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L^n je jeho normalizovaná Laplaceova matice a D jeho matice stupňovitosti. Necht' λ_i jsou její vlastní čísla a X_i jsou její vlastní vektory, pak nejmenší vlastní číslo $\lambda_1 = 0$ a příslušný vlastní vektor je $X_1 = D^{1/2}(1, \dots, 1)^T$.*

Důkaz. Dosadíme X_1 a $\lambda_1 = 0$ do definice vlastních čísel pro L^n .

$$L^n D^{1/2}(1, \dots, 1)^T = \lambda D^{1/2}(1, \dots, 1)^T.$$

Poté rozepíšeme normalizovanou Laplaceovu matici

$$D^{-1/2}LD^{-1/2}D^{1/2}(1, \dots, 1)^T = \lambda D^{1/2}(1, \dots, 1)^T.$$

Z čehož plyne

$$D^{1/2}D^{-1/2}LD^{-1/2}D^{1/2}(1, \dots, 1)^T = \lambda(1, \dots, 1)^T.$$

Což po vykrácení znamená

$$L(1, \dots, 1)^T = \lambda(1, \dots, 1)^T$$

Tím nám vznikla klasická Laplaceova matice. Z věty 1.4.1. víme, že vektor $(1, \dots, 1)^T$ je jejím vlastním vektorem pro $\lambda = 0$. Tím jsme ověřili tvrzení věty.

□

Definice 3.2.1. *Necht' G je konečný a souvislý neorientovaný graf s k vrcholy a l hranami. Necht' M je jeho matice incidence a D jeho matice stupňovitosti. Normalizovanou maticí incidence $T \in \mathbb{R}^{k,2l}$ pak nazvu*

$$T = D^{-1/2}M.$$

Věta 3.2.3. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L^n je jeho normalizovaná Laplaceova matice a necht' T je její normalizovaná matice incidence, pak*

$$L = \frac{1}{2}TT^T.$$

Důkaz. Vyjdeme z definice normalizované Laplaceovy matice (3.2).

$$L^n = D^{-1/2}LD^{-1/2} = D^{-1/2}\frac{1}{2}MM^TD^{-1/2} = \frac{1}{2}(D^{-1/2}M)(M^TD^{-1/2}) = \frac{1}{2}(D^{-1/2}M)(D^{-1/2}M)^T \quad (3.8)$$

Tedy jsme našli naši matici $T = D^{-1/2}M$.

□

Díky tomu jsme opět získali velmi silnou spektrální vlastnost pozitivní semi-definitnosti.

Věta 3.2.4. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L^n je jeho normalizovaná Laplaceova matice, pak L^n je pozitivně semi-definitní.*

Důkaz. Jedná se o součin matice a její transpozice. Opět budeme postupovat stejně jako u věty 1.4.3. \square

Nyní si konečně ukážeme proč se normalizovaná Laplaceova matice nazývá normalizovaná [21].

Věta 3.2.5. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L^n je jeho normalizovaná Laplaceova matice a λ_i její vlastní čísla. Pak pro všechna i platí*

$$\lambda_i \in (0, 2).$$

Důkaz. Již víme, že $\lambda_i \geq 0$ pro všechna i z věty 3.2.4.

Pro druhou hranici budeme potřebovat Rayleighův koeficient, a to konkrétně vlastnost pro symetrické matice B . Konkrétně, že pro největší vlastní vektor je $\frac{X^T B X}{X^T X}$ rovno spektrálnímu poloměru B . Začneme s $X^T L^n X$, poté použijeme rovnici (3.4)

$$X^T (I + D^{-1/2} A D^{-1/2}) X = X^T X + X^T D^{-1/2} A D^{-1/2} X$$

Poté rovnici přepíšeme do sumačního zápisu

$$X^T X + X^T D^{-1/2} A D^{-1/2} X = \sum_{i \in V} \left(\frac{X_i}{\sqrt{\deg(i)}} \right)^2 - \sum_{i, j \in E} \frac{X_i X_j}{\sqrt{\deg(i) \deg(j)}}.$$

Z čehož plyne

$$X^T X + X^T D^{-1/2} A D^{-1/2} X = \sum_{i, j \in E} \left(\frac{X_i}{\sqrt{\deg_i}} - \frac{X_j}{\sqrt{\deg_j}} \right)^2 \geq 0.$$

Z této nerovnosti můžeme podmínku přeformulovat na

$$X^T X \geq -X^T D^{-1/2} A D^{-1/2} X.$$

Poté na obě strany vhodně rozšíříme,

$$2X^T X \geq X^T X - X^T D^{-1/2} A D^{-1/2} X = X^T D^{-1/2} L D^{-1/2} X = X^T L^n X.$$

Z čehož plyne, že

$$\frac{X^T L^n X}{X^T X} \leq 2.$$

čímž jsme dokázali druhé omezení. \square

3.3 Konvergence ke stacionární distribuci

Jedním z hlavních výsledků teorie Markovských procesů je, že konečné, nerozložitelné a aperiodické řetězce mají jednoznačný stacionární stav [20]. Přímé definice z teorie Markovských řetězců nebudeme uvádět, toto tvrzení si raději přeformulujeme pro naše potřeby. To, že Markovský řetězec je konečný, je ekvivalentní tomu, že náhodná procházka je na konečném grafu. Nerozložitelný Markovský řetězec pak znamená, že náhodná procházka se odehrává na souvislém grafu. To že je Markovský řetězec aperiodický je ekvivalentní tomu, že náhodná procházka je na aperiodickém grafu.

Tvrdíme tedy, že náhodná procházka na konečném, souvislém a aperiodickém neorientovaném grafu, má právě jeden stacionární stav. Tím rozumíme, že neohledně na volbu vektoru pravděpodobností p platí pro matici přechodu P

$$\lim_{n \rightarrow \infty} pP^n = \pi,$$

kde π je stacionární stav.

Ovšem nabízí se otázka, jak rychle konverguje. Budeme tedy hledat, jak daleko bude v určitý čas Markovský řetězec od svého stacionárního stavu. První, co musíme udělat, abychom toho dosáhli, je zavést metriku, kterou budeme měřit vzdálenost mezi stavy. Celou dobu budeme vycházet ze článků [4] a [18].

Definice 3.3.1. *Necht' je dán konečný, souvislý a aperiodický neorientovaný graf G . Necht' je dána náhodná procházka na něm. Necht' V je množina vrcholů a necht' f, h jsou dvě distribuce na V , pak vzdálenost distribucí f a h definuji jako:*

$$\|f - h\|_d = \max_{\mu \in V} |f(\mu) - h(\mu)| \quad (3.9)$$

Definice 3.3.2. *Necht' je dán konečný, souvislý a aperiodický neorientovaný graf G . Necht' je dána náhodná procházka na něm. Necht' V je množina vrcholů a necht' $n \in \mathbb{N}$. Necht' P je Matice přechodu a p vektor počátečních pravděpodobností. Pak vzdáleností od stacionárního stavu $d(n, p)$ nazvu*

$$d(n, p) = \max_{v \in V} \|pP^n - \pi\|_d.$$

Definice 3.3.3. *Necht' je dán konečný, souvislý a aperiodický neorientovaný graf G . Necht' je dána náhodná procházka na něm a vzdálenost od stacionárního stavu $d(n)$. Pak definujeme rychlost konvergence jako*

$$t(\epsilon, p) = \min\{n : d(n, p) \leq \epsilon\}.$$

Definice 3.3.4. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L^n je jeho normalizovaná Laplaceova matice. Necht' λ_2 je druhé nejmenší a λ_{k-1} druhé největší vlastní číslo dané matice. Pak spektrální mezerou nazvu*

$$\alpha = \min(\lambda_2, 2 - \lambda_{k-1}).$$

Nyní se po přípravných definicích dostáváme ke klíčové větě.

Věta 3.3.1. *Necht' je dán konečný, souvislý a aperiodický neorientovaný graf G . Necht' je dána náhodná procházka na něm. V je množina vrcholů G , π stacionární stav náhodné procházky a α jeho spektrální mezera. Pak pro rychlost konvergence $t(\epsilon)$ platí*

$$\left(\frac{1}{\alpha} - 1\right) \log\left(\frac{1}{2\epsilon}\right) \leq t(\epsilon) < 1 + \frac{1}{\alpha} \log\left(\frac{1}{\beta\epsilon}\right), \quad (3.10)$$

kde $\beta := \min\{\pi(v) : v \in V\}$.

Důkaz. Důkaz vychází z článku [4], poté si stačí z věty 3.1.2. uvědomit vztah mezi vlastními čísly normalizované Laplaceovy matice a matice přechodu. □

Provádět důkaz věty 3.3.1. nemá z pohledu práce smysl, protože je veden čistě v rámci obecné teorie Markovských řetězců. Nabízí se tedy otázka, proč to vůbec zahrnujeme do práce o Laplaceově matici. Důvodem je, že je jednodušší získat spektrum normalizované Laplaceovy matice, než matice přechodu. Toto se dá provést u řady dalších problémů, kde problematickou spektrální analýzu matice přechodu nahradíme jednodušší spektrální analýzou symetrické normalizované Laplaceovy matice.

3.4 Clustering Markovských řetězců

K čemu je clustering Markovských řetězců ukážeme na následující motivační úloze.

Motivace:

Uvažujme náhodnou procházku na grafu sítě veřejné dopravy v Evropě. Tedy na libovolné přestupní stanici v Evropě se ztratí teenager, ten si vybere náhodný spoj, a při cestě pak náhodně vystoupí na nějaké další přestupní stanici. Tam postup opakuje. Aby byl graf souvislý, budeme brát pouze stanice-vrcholy, pro které existuje sled z a do Hlavního nádraží v Praze.

Zjištění pravděpodobnostní distribuce po n krocích je možné umocňováním matice přechodu. V případě, že tato náhodná procházka bude trvat nekonečně dlouho, tak se nutností zkonstruovat takovou matici pro celou Evropu skutečně nevyhneme. Chování v nekonečnu bezpochyby matematicky zajímavé, ale běžně se nám ovšem v aplikacích, například ve strojovém učení, hodí sice velký, avšak konečný počet kroků.

Představme si tedy, že počáteční stav bude například ve stanici metra v Liverpoolu. Ačkoliv by daný teenager mohl být teoreticky už 6 hodiny poté v Paříži, veřejnou dopravu Velké Británie a kontinentální Evropy spojuje jediná hrana. Tedy taková situace je velmi nepravděpodobná při zadaných podmínkách. Můžeme tedy předpokládat, že pravděpodobnost cesty z Muzea na Florenc nebude mít ve střednědobém časovém horizontu významný vliv na ztraceného teenagera v Liverpoolu a že Britské policii bude velmi pravděpodobně stačit používat matici přechodu sítě veřejné dopravy Velké Británie.

Z toho vychází potřeba rozdělit grafy a jim příslušné náhodné procházky na podgrafy. Poté buď zkoumáme jen vztahy mezi těmito podgrafy, nebo zkoumáme samotné podgrafy a nebudeme řešit jejich napojení. Popřípadě například definovat tunel pod La Manche jako absorbující stav.

Clustering se typicky provádí na neorientovaných nebo symetrizovaných grafech a používá se k němu standardní Laplaceova matice. Ovšem normalizovaná Laplaceova matice by měla dávat podobné výsledky, jak uvidíme.

Definice 3.4.1. *Necht' C a B jsou symetrické matice. Necht' λ_1 je řešení klasického problému vlastních čísel matice C :*

$$CX_1 = \lambda_1 X_1.$$

Pak λ_2 je řešení zobecněného problému vlastních čísel matice C :

$$CX_2 = \lambda_2 BX_2.$$

Věta 3.4.1. *Necht' G je konečný a souvislý neorientovaný graf. Necht' L je jeho Laplaceova matice a L^n je jeho normalizovaná Laplaceova matice, pak vlastní čísla L^n jsou řešením zobecněného problému vlastních čísel matice L .*

Důkaz. Začneme dosazením do zobecněného problému vlastních čísel. Budeme předpokládat $B = D$

$$LX = \lambda DX$$

$$D^{-1/2}LX = \lambda D^{1/2}X$$

$$L^n D^{1/2}X = \lambda D^{1/2}X$$

dále označíme $X^n = D^{1/2}X$

$$L^n X^n = \lambda X^n$$

Z čehož máme standardní problém vlastních čísel pro L^n .

□

Tato věta zdůvodňuje zajímavý vztah mezi Laplaceovou maticí a normalizovanou Laplaceovou maticí, který si v praxi ukážeme později. A to, že při použití spectral clusteringu jsou výsledky pro obě matice podobné. To vede k tomu, že se často v aplikacích neorientovaný graf Markovského řetězce rozdělí klasickou Laplaceovou maticí. Tento postup platí i v případech, že náhodnou procházku na grafu zobecníme. I když při kroku nebudou mít všechny připojené vrcholy stejnou pravděpodobnost, klasická Laplaceova matice bude stále fungovat pro dostatečně velké clustery [3]. Výjimkou jsou extrémní případy, kdy například některé hrany budou mít pravděpodobnost 90%, nebo když v grafu bude nějaký obecný trend. Obecně lze říct, že čím bude obecná náhodná procházka bližší námi zadané náhodné procházce, tím bude vztah mezi spektrem Laplaceovy matice a spektrem normalizované Laplaceovy matice silnější.

Kapitola 4

Spectral clustering

V této kapitole se budeme zabývat aplikací Spectral clusteringu pro tři grafy, nalezneme jejich clustery skrze Laplaceovu matici a normovanou Laplaceovu matici a porovnáme je. Poté využijeme novou definici abstraktní Laplaceovy matice a uvidíme, jestli dává dobré výsledky. Dále grafy upravíme do stavu nesouvislého a slabě souvislého tvaru. Taktéž vyzkoušíme jak se změní abstraktní Laplaceova matice přidáním absorpčního stavu. Na konci porovnáme účinnost klasického postupu symetrizací u orientovaných grafů s abstraktní Laplaceovou maticí. V celé kapitole se vyskytují informace z [3], stejně tak z online kurzů náležících přednášce Mining of Massive Datasets, které je kniha [3] učebnicí a z článku [17].

4.1 Neexaktní úvod

Než se začneme zabývat spectral clusteringem, je potřeba udělat velmi matematicky neexaktní úvod do clusteringu jako takového. Problémem je, že nyní budeme muset opustit oblast exaktní matematiky, protože se zatím nepodařilo úspěšně nalézt definici pojmu cluster.

Clustering grafů je název pro operaci, kdy se velký graf snažíme rozdělit na menší podgrafy nějakých intuitivních celků. Příkladem může být snaha rozdělit uživatele sociálních sítí do sociálních bublin. Pro tento účel existuje několik velmi rychlých a dobře fungujících algoritmů, z nichž zdaleka nejpoužívanější je k-mean algoritmus. Fungování k-mean algoritmu by se jednoduše dalo vysvětlit následujícím způsobem:

- 1) Zvolíme náhodně k stejně velkých clusterů.
- 2) Najdeme centroid clusteru (Může se vzít geometrický střed, těžiště, liší se podle konkrétní aplikace algoritmu.).
- 3) Rozdělíme graf podle blízkosti centroidu.
- 4) Body 2) a 3) opakujeme.

Problémem tohoto algoritmu je, že neexistuje žádná matematická garance konvergence pro různé počáteční volby. Tedy můžeme mít různé výsledky a algoritmus na konci vyžaduje kontrolu.

Druhým významným problémem je, že algoritmus předpokládá přibližnou sférickou symetrii vrcholů v clusteru. Použijme opět původní modelový příklad se silniční sítí. Představme si podlouhlý stát, například Chile. Chile je dlouhá kolem 4300 kilometrů a místy i méně než 100 kilometrů široká. V rámci Chile existuje hustá silniční síť a jen několik silnic ji spojuje se sousední Argentinou. Bylo by tedy logické

rozdělit graf silniční sítě jihu Jižní ameriky na Argentinu a Chile (Opomeňme nyní Ohňovou zemi.). Ovšem k-mean algoritmus by s tímto měl problém, protože Chile není ani přibližně sféricky symetrická. S dělením Česka a Polska by problém nebyl.

Naproti tomu spectral clustering nemá žádnou počáteční volbu a zajímá ho výhradně propojenost daných vrcholů, tedy nevyžaduje sférickou symetrii.

Spectral clustering vychází z práce českého matematika Prof. RNDr. Miroslava Fiedlera, DrSc. (1926-2015) o algebraické konektivitě a Fiedlerových vektorech. Je nutné zdůraznit, že proč přesně metoda funguje, není exaktně matematicky dokázáno. Jakýkoliv důkaz by narazil již u definice problému, protože clustering jak grafů, tak Markovských řetězců není dobře definován. Snažíme se tedy graf intuitivně rozdělit na celky které vzájemně nejsou dobře propojeny.

Zatím jsme druhé nejmenší vlastní číslo λ_2 v otázce souvislosti brali diskrétně z věty 1.4.6. Tedy to že λ_2 je nenulové, implikuje, že graf je souvislý. To, že λ_2 je nulové, implikuje, že graf je nesouvislý. Pojďme si ovšem tento princip zespojit. Tedy čím větší je algebraická konektivita, tím vyšší je souvislost daného grafu. Stejně tak, čím nižší je algebraická konektivita, tím nižší je souvislost grafu až k nesouvislému grafu, který má algebraickou konektivitu rovnou nule. To můžeme velmi hezky vidět na námi dokázaných speciálních případech grafů. Úplný graf o kterém bychom intuitivně řekli, že je velmi dobře propojený, má dle věty 1.5.1. velmi vysoké λ_2 . Naopak cesta a kružnice, které bychom označili za velmi málo propojené, mají dle 1.5.2. a 1.5.3. velmi malé λ_2 .

Pro pochopení Fiedlerova vektoru X_2 využijeme věty 1.4.6. a 1.4.7. Pro připomenutí si ukážeme tvar Laplaceovy matice nesouvislého grafu skládajícího se ze dvou souvislých grafů dle rovnice (1.12)

$$L = \begin{bmatrix} L_1 & 0 \\ 0 & L_2 \end{bmatrix}.$$

Dle rovnice (1.13) a (1.14) mají jejich vlastní vektory X_1 a X_2 náležící λ_1 a λ_2 hodnoty patřící oběma skupinám vrcholů stejné. Samozřejmě bychom mohli například za vektor X_1 vzít vektor samých jedniček, pak by ovšem stejné hodnoty pro obě skupiny vrcholů musely být v X_2 . To se opět pokusíme přenést do grafů souvislých. Podmínku, že prvky X_2 náležící jednomu z dvou souvislých grafů jsou stejné, nahradíme tím, že jejich prvky vektoru X_2 mají podobnou hodnotu.

V podstatě bychom pro souvislý graf, který má dva clusterly, měli matici podobnou té z rovnice (1.12). Jen místo nulových matic v I. a III. kvadrantu by zde byli velmi řídké matice. To nás také vede k tomu proč by abstraktní Laplaceova matice měla být lepší pro dělení orientovaných grafů. Pokud bude hrana jen orientovaná, bude jí dávat menší význam (hodnotu), než když mezi vrcholy budou existovat hrany obě. Tím se liší od postupu se symetrizací orientovaného grafu, který mezi tím rozdíl nedělá.

4.2 Popis metody

Běžně je definována a používána pro konečné a neorientované grafy G . My si ovšem ukážeme, že metoda dobře funguje i pro jiné grafy, při správné volbě matice.

Metoda pro dva clusterly

- 1) Nalezneme Laplaceovu matici L grafu G .
- 2) Nalezneme algebraickou konektivitu λ_2 matice L .
- 3) V případě, že $\lambda_2 = 0$, nalezneme první souvislé grafy G_i . Pro každý pokračujeme bodem 1).
- 4) V případě, že $\lambda_2 \neq 0$, nalezneme Fiedlerův vektor.

5) Rozdělíme vrcholy dle hodnoty Fiedlerova vektoru na kladné a záporné a z nich pak složíme dva clustery.

Toto je klasická metoda, kterou budeme používat i při následných numerických simulacích. Vždy není nutné dělit vrcholy na ty s hodnotou Fiedlerova vektoru na větší a menší než nula. Také se jako středový bod používá například medián, aritmetický průměr hodnot Fiedlerova vektoru, či se různě iteruje. To už jsou spíše optimalizační úlohy.

Metoda pro více clusterů - čistě spektrální

- 1) Nalezneme Laplaceovu matici G .
- 2) Nalezneme algebraickou konektivitu G .
- 3) V případě, že $\lambda_2 = 0$, nalezneme souvislé grafy G_i a pro každý pokračujeme bodem 1).
- 4) V případě, že $\lambda_2 \neq 0$, nalezneme Fiedlerův vektor.
- 5) Rozdělíme vrcholy dle hodnoty Fiedlerova vektoru na kladné a záporné a z nich pak složíme dva clustery.
- 6) Z těchto dvou clusterů uděláme nepropojené souvislé grafy G_1 a G_2 a pokračujeme bodem 1) zvlášť na G_1 a G_2 .
- 7) Provádíme body 1-6) dokud rozdíl hodnot ve vlastních vektorech neklesne pod určitou hranici, nebo dokud nemáme určený počet clusterů.

Problém této metody je, že musíme poté výsledky zkontrolovat. K dělení nemuselo docházet po dvou a tak se nám mohl jeden hledaný cluster roztrhnout na dva, nebo dva menší zůstat jako jeden. Proto se tato metoda v praxi příliš nepoužívá.

Metoda pro více clusterů - smíšená

- 1) Nalezneme Laplaceovu matici G
- 2) Nalezneme algebraickou konektivitu G
- 3) V případě, že $\lambda_2 = 0$, nalezneme souvislé grafy G_i a pro každý pokračujeme bodem 1).
- 4) V případě, že $\lambda_2 \neq 0$ vypočteme pro ně jedno z nejmenších vlastních čísel λ_j a příslušný vlastní vektor X_i^j . (Nemusíme tedy použít nutně Fiedlerův vektor X_i^2 .)
- 5) Na hodnoty vektoru X_i^j aplikujeme algoritmus k-mean, který nám daný vektor rozdělí na vrcholy s podobnou hodnotou X_i^j .

Tato metoda dává v praxi velmi dobré výsledky, protože k-mean pracuje na úrovni čísel mnohem lépe než na úrovni grafů. Tím je schopen vykompenzovat různé problémy, které má samostatné dělení na poloviny čistým spektrálním clusteringem.

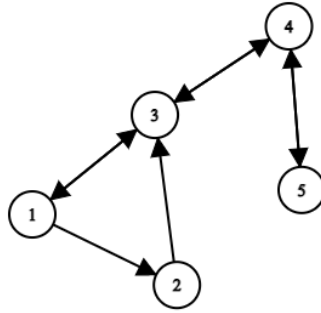
4.3 Řešené problémy

Pro numerickou simulaci jsme zvolili tři grafy. Na prvním si ukážeme princip metody. Na druhém si ukážeme, že všechny metody dávají podobné, ovšem ne stejné, výsledky. Ve třetím příkladě pak ukážeme výhody abstraktní Laplaceovy matice. V incidenčních maticích jsou úmyslně ponechány nulové sloupce matic, aby bylo možno lépe pochopit vliv úprav těchto grafů na dané matice.

4.3.1 Příklad 1

Máme následující orientovaný graf G_1 z dřívější ukázky.

Obrázek 4.1: Orientovaný graf G_1



Symetrizovaný graf

Začneme tím, že graf na obrázku 4.1 symetrizujeme na graf G_1^s . Poté nalezneme matici incidence M_1^s .

$$M_1^s = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Tu sice pro zjištění Laplaceovy matice pro neorientované grafy nepotřebujeme, ale je to výhodné pro pochopení konstrukce abstraktní Laplaceovy matice. Dále pak najdeme z definice Laplaceovu matici L_1^s .

$$L_1^s = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2 = 0.5188$$

$$X_2 = \begin{bmatrix} 0.4193 \\ 0.4193 \\ 0.2018 \\ -0.3380 \\ -0.7024 \end{bmatrix}$$

To by nám rozdělilo graf na dva clustery, první složen z vrcholů {1, 2, 3}, druhý {4, 5}, což by bylo logické rozdělení daného grafu.

Symetrizovaný graf - normalizovaná Laplaceova matice

Najdeme normalizovanou Laplaceovu matici grafu G_1^s . Již známe matici L_1^s .

$$L_1^s = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

Nyní ji z rovnice (3.2) znormalizujeme

$$L_1^n = D^{-1/2} L_1^s D^{-1/2}.$$

Z čehož plyne, že

$$L_1^n = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{\sqrt{6}}{6} & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{\sqrt{6}}{6} & 0 & 0 \\ -\frac{\sqrt{6}}{6} & -\frac{\sqrt{6}}{6} & 1 & -\frac{\sqrt{6}}{6} & 0 \\ 0 & 0 & -\frac{\sqrt{6}}{6} & 1 & -\frac{\sqrt{2}}{2} \\ 0 & 0 & 0 & -\frac{\sqrt{2}}{2} & 1 \end{pmatrix}.$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2 = 0.3459$$

$$X_2 = \begin{bmatrix} -0.3964 \\ -0.3946 \\ -0.1496 \\ 0.5531 \\ 0.5979 \end{bmatrix}$$

To by nám graf rozdělilo na stejné clustery jakožto běžná Laplaceova matice, což jsme z teorie očekávali.

Orientovaný graf

Přistoupíme k výpočtu samotné Laplaceovy matice L_1 grafu G_1 . Matici M_1 tohoto grafu už známe z dřívějších.

$$M_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Využijeme definici abstraktní Laplaceovy matice (2.1):

$$L_1 = \frac{1}{2}M_1(M_1)^T$$

a máme

$$L_1 = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} & -1 & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ -1 & -\frac{1}{2} & \frac{5}{2} & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

Sice se nejedná o klasickou Laplaceovu matici, ale můžeme si všimnout, že splňuje většinu běžných vlastností jako symetrie atd., což jsme již dokázali.

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

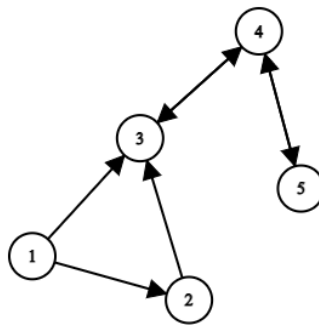
$$\lambda_2 = 0.4791$$

$$X_2 = \begin{bmatrix} 0.3838 \\ 0.5031 \\ 0.1403 \\ -0.3518 \\ -0.6754 \end{bmatrix}$$

To by nám rozdělilo graf na stejné clusterly jako klasická metoda.

Slabě souvislý graf

Nyní zkusíme graf předefinovat tak, aby byl slabě souvislý, ale zároveň nebyl silně souvislý.

Obrázek 4.2: Slabě souvislý graf G_1^a 

Nalezneme matici incidence M_1^a grafu G_1^a .

$$M_1^a = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Opět využijeme definici abstraktní Laplaceovy matice (2.1) k získání L_1^a .

$$L_1^a = \begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

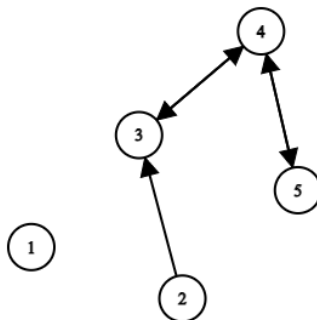
Dále nalezneme algebraickou konektivitu.

$$\lambda_2 = 0.4293$$

Slabá souvislost tedy nezajišťuje algebraickou konektivitu rovnou nule, na to potřebujeme plnou nesouvislost. Opět se jedná o již dokázaný fakt.

Nesouvislý graf

Nyní si graf upravíme na nesouvislý graf G_1^n .

Obrázek 4.3: Nesouvislý graf G_1^n 

Nalezneme matici incidence M_1^n .

$$M_1^n = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Z rovnice (2.1) nalezneme L_1^n .

$$L_1^n = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{3}{2} & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

Zde vidíme, že samostatný vrchol se chová stejně jako v případě klasické Laplaceovy matice. Vytvořil nám v matici nulový řádek a nulový sloupec.

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2 = 0$$

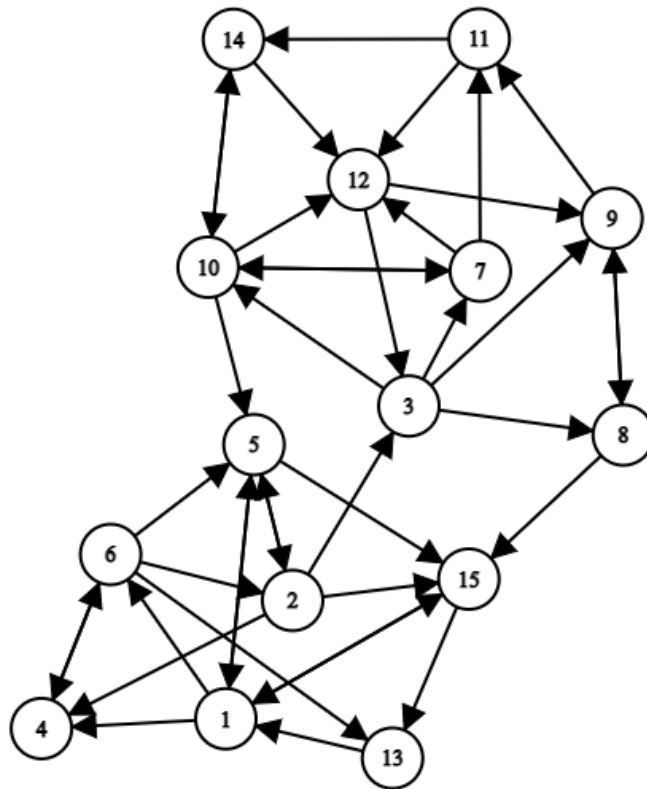
$$X_2 = \begin{bmatrix} 0.0 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix}$$

Graf je nesouvislý, tedy i druhé nejmenší vlastní číslo musí být rovno nule, což odpovídá teorii.

4.3.2 Příklad 2

Nyní přikročíme k trochu většímu orientovanému grafu G_2 .

Obrázek 4.4: Velký orientovaný graf G_2



Symetrizovaný graf

Stejně jako poprvé graf symetrizujeme na graf G_2^s a nalezneme matici L_2^s z definice.

$$L_2^s = \begin{pmatrix} 5 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 5 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & -1 & 6 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & 0 & -1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 3 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 5 & -1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ -1 & -1 & 0 & -1 & -1 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 4 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 3 & -1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 4 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & -1 & 0 & 0 & 5 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & -1 & -1 & -1 & 6 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 3 & 0 \\ -1 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 5 \end{pmatrix}$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2 = 0.5436$$

$$X_2 = \begin{pmatrix} -0.3048 \\ -0.1945 \\ 0.1580 \\ -0.3264 \\ -0.1847 \\ -0.3017 \\ 0.2736 \\ 0.0765 \\ 0.2409 \\ 0.1894 \\ 0.3217 \\ 0.2758 \\ -0.3332 \\ 0.3207 \\ -0.2112 \end{pmatrix}$$

To nám graf dělí na clusterly z vrcholů $\{1, 2, 4, 5, 6, 13, 15\}$ a $\{3, 7, 8, 9, 10, 11, 12, 14\}$, což vypadá jako logický způsob, jak tento graf rozdělit.

Symetrizovaný graf - normalizovaná Laplaceova matice

Budeme počítat normalizovanou Laplaceovu matici grafu G_2^s . Matici L_2^s známe a z rovnice (3.2) nalezneme L_2^n

$$\begin{pmatrix} 1 & 0 & 0 & -\frac{\sqrt{15}}{15} & -\frac{1}{5} & -\frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{15}}{15} & 0 & -\frac{1}{5} \\ 0 & 1 & -\frac{\sqrt{30}}{30} & -\frac{\sqrt{15}}{15} & -\frac{1}{5} & -\frac{1}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{5} \\ 0 & -\frac{\sqrt{30}}{30} & 1 & 0 & 0 & 0 & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{2}}{6} & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{30}}{30} & 0 & -\frac{1}{6} & 0 & 0 & 0 \\ -\frac{\sqrt{15}}{15} & -\frac{\sqrt{15}}{15} & 0 & 1 & 0 & -\frac{\sqrt{15}}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{5} & -\frac{1}{5} & 0 & 0 & 1 & -\frac{1}{5} & 0 & 0 & 0 & -\frac{1}{5} & 0 & 0 & 0 & 0 & -\frac{1}{5} \\ -\frac{1}{5} & -\frac{1}{5} & 0 & -\frac{\sqrt{15}}{15} & -\frac{1}{5} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{15}}{15} & 0 & 0 \\ 0 & 0 & -\frac{\sqrt{6}}{12} & 0 & 0 & 0 & 1 & 0 & 0 & -\frac{\sqrt{5}}{10} & -\frac{1}{4} & -\frac{\sqrt{6}}{12} & 0 & 0 & 0 \\ 0 & 0 & -\frac{\sqrt{2}}{6} & 0 & 0 & 0 & 0 & 1 & -\frac{\sqrt{3}}{6} & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{15}}{15} \\ 0 & 0 & -\frac{\sqrt{6}}{12} & 0 & 0 & 0 & 0 & -\frac{\sqrt{3}}{6} & 1 & 0 & -\frac{1}{4} & -\frac{\sqrt{6}}{12} & 0 & 0 & 0 \\ 0 & 0 & -\frac{\sqrt{30}}{30} & 0 & -\frac{1}{5} & 0 & -\frac{\sqrt{5}}{10} & 0 & 0 & 1 & 0 & -\frac{\sqrt{30}}{30} & 0 & -\frac{\sqrt{15}}{15} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{4} & 0 & -\frac{1}{4} & 0 & 1 & -\frac{\sqrt{6}}{12} & 0 & -\frac{\sqrt{3}}{6} & 0 \\ 0 & 0 & -\frac{1}{6} & 0 & 0 & 0 & -\frac{\sqrt{6}}{12} & 0 & -\frac{\sqrt{6}}{12} & -\frac{\sqrt{30}}{30} & -\frac{\sqrt{6}}{12} & 1 & 0 & -\frac{\sqrt{2}}{6} & 0 \\ -\frac{\sqrt{15}}{15} & 0 & 0 & 0 & 0 & -\frac{\sqrt{15}}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -\frac{\sqrt{15}}{15} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{\sqrt{15}}{15} & -\frac{\sqrt{3}}{6} & -\frac{\sqrt{2}}{6} & 0 & 1 & 0 \\ -\frac{1}{5} & -\frac{1}{5} & 0 & 0 & -\frac{1}{5} & 0 & 0 & -\frac{\sqrt{15}}{15} & 0 & 0 & 0 & 0 & -\frac{\sqrt{15}}{15} & 0 & 1 \end{pmatrix}$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2 = 0.1287$$

$$X_2 = \begin{bmatrix} 0.3383 \\ 0.2174 \\ -0.1953 \\ 0.2638 \\ 0.2116 \\ 0.3344 \\ -0.2660 \\ -0.0591 \\ -0.2325 \\ -0.2068 \\ -0.3065 \\ -0.3381 \\ 0.2701 \\ -0.2543 \\ 0.2387 \end{bmatrix}$$

To nám graf dělí na stejné clusterly jako výpočet klasické Laplaceovy matice. Můžeme si všimnout, že díky normalizaci je číslo algebraické konektivity výrazně menší.

Orientovaný graf

Pokračujeme hledáním L_2 z rovnice (2.1). Matici M_2 zde psát nebudeme z důvodu rozměrů matice. Napíšeme tedy rovnou L_2 .

$$L_2 = \begin{pmatrix} \frac{7}{2} & 0 & 0 & -\frac{1}{2} & -1 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & -1 \\ 0 & 3 & -\frac{1}{2} & -\frac{1}{2} & -1 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 3 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & 2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & \frac{7}{2} & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 0 & -1 & -\frac{1}{2} & 3 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & \frac{5}{2} & 0 & 0 & -1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -1 & \frac{5}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & -1 & 0 & 0 & \frac{7}{2} & 0 & -\frac{1}{2} & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 2 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 3 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 2 & 0 \\ -1 & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 3 \end{pmatrix}'$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

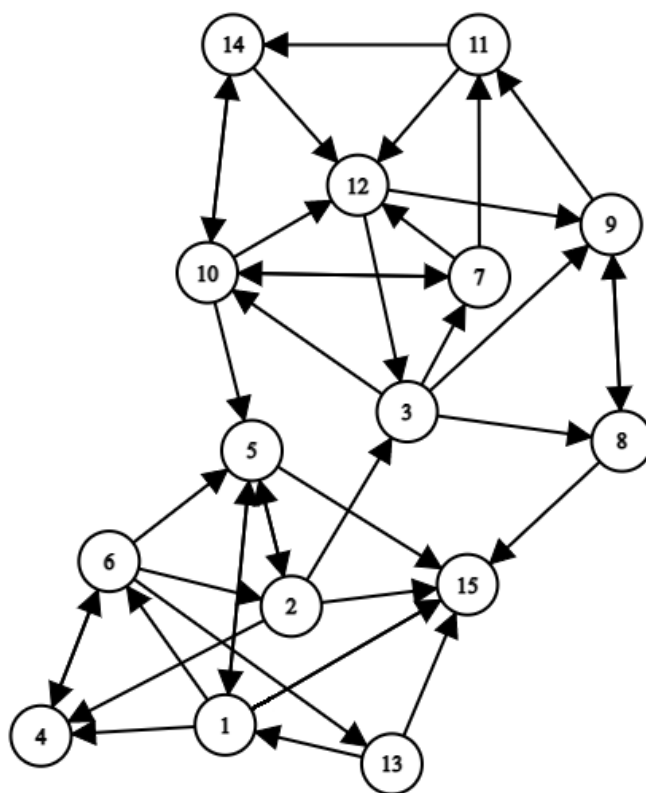
$$\lambda_2 = 0.2980$$

$$X_2 = \begin{bmatrix} 0.2835 \\ 0.2017 \\ -0.1662 \\ 0.3245 \\ 0.2007 \\ 0.3097 \\ -0.2693 \\ -0.1156 \\ -0.2240 \\ -0.2153 \\ -0.3139 \\ -0.2754 \\ 0.3385 \\ -0.2996 \\ 0.2206 \end{bmatrix}$$

To nám dělí graf G_2 na stejné clusterly jako u neorientovaného grafu. Tedy vidíme, že abstraktní Laplaceova matice dává přibližně stejné výsledky jako klasická Laplaceova matice.

Graf s absorpčním stavem

Drobnou úpravou změním graf na následující graf G_2^a . Tento graf je zvláštní tím, že vrchol 15 má pouze hrany, přicházející do vrcholu. V případě, že bychom rozšířili naši teorii Markovských řetězců o orientované grafy, nazýval by se tento vrchol absorpčním stavem.

Obrázek 4.5: Graf s absorpčním stavem G_2^a 

Pokračujeme hledáním L_2^a z rovnice (2.1).

$$L_2^a = \begin{pmatrix} 3 & 0 & 0 & -\frac{1}{2} & -1 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 3 & -\frac{1}{2} & -\frac{1}{2} & -1 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 3 & 0 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & 2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & \frac{7}{2} & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 0 & -1 & -\frac{1}{2} & 3 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & \frac{5}{2} & 0 & 0 & -1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -1 & \frac{5}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & -1 & 0 & 0 & \frac{7}{2} & 0 & -\frac{1}{2} & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 2 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & 3 & 0 & -\frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -\frac{1}{2} & -\frac{1}{2} & 0 & 2 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & \frac{5}{2} \end{pmatrix}$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2^a = 0.2953$$

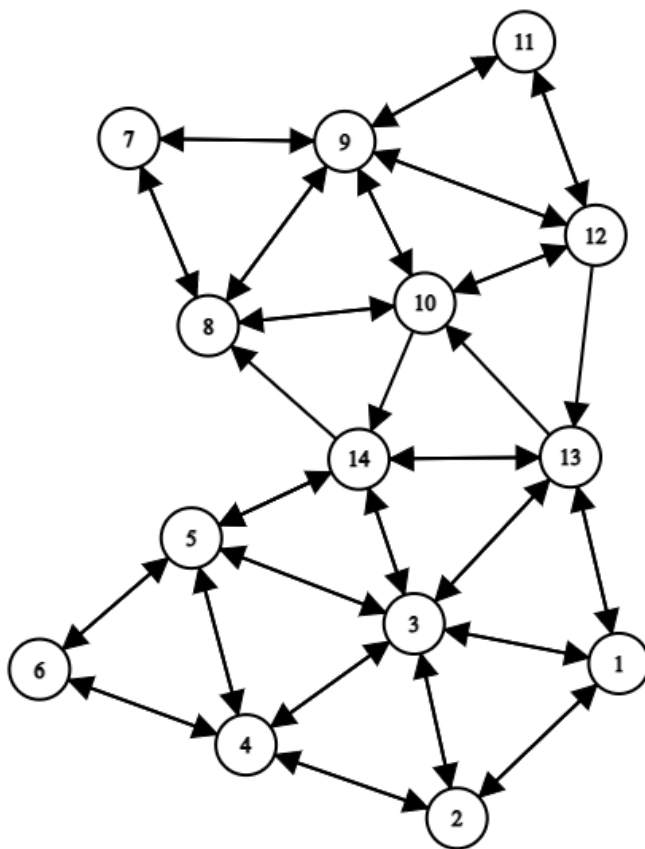
$$X_2 = \begin{pmatrix} 0.2933 \\ 0.2002 \\ -0.1667 \\ 0.3278 \\ 0.2016 \\ 0.3120 \\ -0.2678 \\ -0.1206 \\ -0.2256 \\ -0.2138 \\ -0.3124 \\ -0.2743 \\ 0.3371 \\ -0.2975 \\ 0.2067 \end{pmatrix} 3$$

To se příliš neliší od silně souvislé verze, což pro nás není zcela ideální, protože to ukazuje, že naše matice není s takovými eventualitami schopna pracovat.

4.3.3 Příklad 3

Nyní jsme ukázali, že jak klasický postup se symetrizací grafu, tak naše abstraktní Laplaceova matice dávají dobré výsledky. Nabízí se otázka jestli má námi definovaná matice nějaké výhody. To si ukážeme na následujícím příkladu.

Obrázek 4.6: Orientovaný graf G_3



Symetrizovaný graf

Opět graf G_3 symetrizujeme na graf G_3^s a najdeme matici L_3^s z definice.

$$L_3^s = \begin{pmatrix} 3 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 6 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & -1 & -1 & 4 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 4 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 4 & -1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 5 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 5 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & 4 & -1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 5 & -1 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & -1 & 5 \end{pmatrix}$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2^a = 0.5002$$

$$X_2 = \begin{bmatrix} 0.1935 \\ 0.2881 \\ 0.1978 \\ 0.3290 \\ 0.2677 \\ 0.3978 \\ -0.3619 \\ -0.2366 \\ -0.3061 \\ -0.1722 \\ -0.3653 \\ -0.2417 \\ -0.0023 \\ 0.0121 \end{bmatrix}$$

To nám graf dělí poněkud zvláště na clustery z vrcholů $\{1, 2, 3, 4, 5, 6, 14\}$ a $\{7, 8, 9, 10, 11, 12, 13\}$, což ovšem nevypadá příliš logicky. Vrchol 13 by neměl patřit k hornímu clusteru. Existuje jen jedna možnost, jak se do něj dostat z horního clusteru. Stejně tak existuje jen jedna možnost, jak se z něj dostat do horního clusteru. Naproti tomu existují tři možnosti, jak se dostat z i do dolního clusteru z vrcholu 13. Ještě k tomu oba vrcholy 13 a 14 mají hodnoty X_2 velmi blízké nule.

Orientovaný graf

Nyní naproti tomu nalezneme abstraktní Laplaceovu matici L_3 grafu G_3 . Opět nalezneme matici incidence M_3 , kterou z nebudeme psát z důvodu rozměrů. Opět použijeme rovnici (2.1) k nalezení L_3 .

$$L_3 = \begin{pmatrix} 3 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 6 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & -1 & -1 & 4 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 4 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & \frac{7}{2} & -1 & -1 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 5 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 4 & 0 & -1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 & \frac{7}{2} & -\frac{1}{2} & 0 \\ -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 4 & -1 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & -1 & 4 \end{pmatrix}$$

Dále nalezneme algebraickou konektivitu a Fiedlerův vektor.

$$\lambda_2 = 0.3617$$

$$X_2 = \begin{bmatrix} -0.2159 \\ -0.2778 \\ -0.2158 \\ -0.30120 \\ -0.2598 \\ -0.3424 \\ 0.3553 \\ 0.2673 \\ 0.3148 \\ 0.2117 \\ 0.3564 \\ 0.2692 \\ -0.0761 \\ -0.0858 \end{bmatrix}$$

To nám graf dělí na dva clustery z vrcholů $\{1, 2, 3, 4, 5, 6, 13, 14\}$ a $\{7, 8, 9, 10, 11, 12\}$, což je mnohem lepší dělení, než nám dal původní postup. Za prvé, vrcholy 13 a 14 patří k dolnímu clusteru. Za druhé, hodnoty X_2 náležící vrcholům 13 a 14 jsou o téměř řád vyšší, než v postupu se symetrizací. To znamená jednoznačnější rozdělení.

Závěr

V práci je definována Laplaceova matice a je podán ucelený přehled o jejích vlastnostech a spektru. Stejně tak jsou nalezena spektra Laplaceovy matice pro úplný graf, kružnici, cestu a hvězdu. Dále pak byly rozebrány spektrální vlastnosti stromů. Poté byla klasickým způsobem zdefinována algebraická konektivita a Fiedlerův vektor. V rámci kapitoly jsou i krátce rozebrány vlastnosti matice sousednosti.

V další části práce jsme zdefinovali novou abstraktní Laplaceovu matici pro orientované grafy. Dokázali jsme, že pro ni platí stejné spektrální vlastnosti jako pro běžnou Laplaceovu matici. Z toho jsme vyvodil, že pro ni taktéž budou fungovat aplikace algebraické konektivity a Fiedlerova vektoru ve spectral clusteringu.

Ve třetí části jsme zdefinovali matice běžně asociované s náhodnou procházkou na neorientovaném grafu a dokázali vztahy jejich spekter. Následně jsme ukázali, že spektrální vlastnosti normalizované Laplaceovy matice jsou podobné vlastnostem klasické Laplaceovy matice. Poté jsme rozebrali vztah spektra normalizované Laplaceovy matice a rychlosti konvergence ke stacionární distribuci. Také jsme vysvětlili clustering Markovských řetězců a vztah Markovských řetězců k náhodným procházkám na grafech. Z toho jsme vyvodili, že clustering Markovských řetězců můžeme nahradit clusteringem grafů.

Nakonec jsme přistoupili k samotnému spectral clusteringu grafů. Vysvětlili jsme si dané metody a aplikovali je na několik grafů. Naše zjištění byla ve shodě s teorií a naším očekáváním. Ukázalo se, že pro námi zdefinovanou náhodnou procházku, dávají normalizované Laplaceova matice a klasická Laplaceova matice stejné výsledky. Naopak jsme zjistili, že abstraktní Laplaceova matice dává pro některé orientované grafy lepší výsledky než klasický postup se symetrizací. To je možné přičítat tomu, že existence jen jedné orientované hrany má v rámci abstraktní Laplaceovy matice menší váhu než existence orientovaných hran oběma směry.

Seznam obrázků

1.1	Neorientovaný graf H	13
1.2	Řešení problémů se smyčkou	17
1.3	Úplný graf	24
1.4	Kružnice	26
1.5	Cesta	28
1.6	Hvězda	29
1.7	Strom	30
2.1	Orientovaný graf G	37
4.1	Orientovaný graf G_1	54
4.2	Slabě souvislý graf G_1^a	58
4.3	Nesouvislý graf G_1^n	59
4.4	Velký orientovaný graf G_2	60
4.5	Graf s absorpčním stavem G_2^a	64
4.6	Orientovaný graf G_3	66

Literatura

- [1] A. E. Brouwer, W. H. Haemers: *Spectra of Graphs*. Springer,2011.
- [2] M. Juvan *Some applications of Laplace eigenvalues of graphs*. In 'Graph symmetry' * B. Mohar. Springer,Dordrecht,2011,225-275.
- [3] J. Leskovec, A. Rajaraman, J. D. Ullman: *Mining of Massive Datasets*. Cambridge University Press,2011.
- [4] D. A. Levin, Y. Peres, and E. L. Wilmer: *Markov Chains and Mixing Times*. American Mathematical Society,2009.
- [5] M. Wardetzky: *Discrete Laplace Operators*. . An Excursion Through Discrete Differential Geometry: AMS Short Course, Discrete Differential Geometry,2020.
- [6] M. Fiedler: *Algebraic connectivity of graphs*. Czechoslovak Mathematical Journal 23(98), 1973,298-305.
- [7] I. Gutman: *The star is the tree with greatest greatest laplacian eigenvalue*. Kragujevac Journal of Mathematics 24,2002.
- [8] R. O.Braga, V. M. Rodrigues, V. Trevisan: *On the distribution of Laplacian eigenvalues of trees*. Discrete Mathematics 313(21), 2010,2382-2389.
- [9] M. S. Cavers,: *The normalized laplacian matrix and general Randic Index of graphs*,Ph.D. thesis,University of Regina,Regina, Saskatchewan,2010.
- [10] J. H. Gallier: *Notes on Elementary Spectral Graph Theory. Applications to Graph Clustering Using Normalized Cuts*,ArXiv 1311(2492), 2013.
- [11] Xiao-Dong Zhang: *The Laplacian eigenvalues of graphs: a survey*. arXiv 1111(2897),2011.
- [12] Gustav R. Kirchhoff: *Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird*. Annalen der Physik 148,1847.
- [13] A. J. Ganesh,: *Introduction to spectral graph theory*. University of Bristol,2015.
- [14] J. Matoušek, J. Nešetřil: *Kapitoly z diskrétní matematiky*. Univerzita Karlova v Praze, Nakladatelství Karolinum,2009.
- [15] O. Tegby: *The Laplacian in its different guises*. Examensarbete i matematik, Uppsala University, Uppsala,2019.

- [16] J. Dong, L. Lin: *Laplacian matrices of general complex weighted directed graphs*. Linear Algebra and its Applications 510, 2016, 1-9.
- [17] A. I. Ng, M. I. Jordan, Y. Weiss: *On spectral clustering: Analysis and an algorithm*. Advances in Neural Information Processing Systems 14, 2002, 849-856.
- [18] S. Kale: *Eigenvalues and Mixing Time*. University of Dartmouth, 2013.
- [19] F. Chung: *Laplacians and the Cheeger Inequality for Directed Graphs*. Annals of Combinatorics 9, 2005, 1-19.
- [20] R. Douc, E. Moulines, P. Priouret, P. Soulier: *Markov Chains*. Springer, 2018.
- [21] D. P. Williamson: *Normalized Adjacency and Laplacian Matrices*. Spectral Graph Theory, Lecture 7, Cornell University, September 13, 2016.
- [22] László Lovász: *Combinatorial Problems and Exercises*. American Mathematical Society, 1993.