

Univerzita Komenského v Bratislave
Fakulta Matematiky, Fyziky a Informatiky

PageRank algoritmus

Bakalárska práca

Študijný program: Informatika

Študijný odbor: 9.2.1 Informatika

Školiace pracovisko: Katedra Informatiky

Školiteľ: RNDr. Martin Sleziak, PhD

Ondrej Mikuláš

Bratislava, 2010

Podakovanie

Týmto sa chcem poďakovať môjmu vedúcemu RNDr. Martinovi Sleziakovi, PhD za výber témy, cenné rady a podporu.

Čestné prehlásenie

Čestne prehlasujem, že som túto bakalársku prácu vypracoval samostatne s použitím citovaných zdrojov.

.....

Abstrakt

Autor: Ondrej Mikuláš
Názov bakalárskej práce: PageRank algoritmus
Škola: Univerzita Komenského v Bratislave
Fakulta: Fakulta matematiky, fyziky a informatiky
Katedra: Katedra informatiky
Vedúci bakalárskej práce: RNDr. Martin Sleziak, PhD
Rozsah práce: 49 strán
Bratislava, jún 2010

V práci sa zaoberáme algoritmom PageRank, ktorý slúži na zoradovanie webových stránok podľa popularity. Odvodíme potrebné poznatky z lineárnej algebry a pomocou nich popíšeme PageRank algoritmus a overíme jeho správnosť a konvergenciu.

Ďalej vysvetlíme praktický význam týchto teoretických poznatkov pre fungovanie algoritmu. Na záver otestujeme správanie algoritmu na niektorých špeciálnych sieťach s použitím vlastnej implementácie.

KĹÚČOVÉ SLOVÁ: PageRank algoritmus, lineárna algebra, vlastný vektor, konvergencia

Abstract

Author: Ondrej Mikuláš
Thesis title: PageRank algoritmus
University: Univerzita Komenského v Bratislave
Faculty: Fakulta matematiky, fyziky a informatiky
Department: Katedra informatiky
Advisor: RNDr. Martin Sleziak, PhD
Thesis length: 49 strán
Bratislava, june 2010

The main topic of this thesis is PageRank algorithm, which ranks the webpages according to their popularity. Our knowledge of PageRank algorithm is based on the chapters from linear algebra.

Then we verify its correctness and properties of convergence. We also explain the effects of this knowledge on running of the algorithm. At the end we test behaviour of the algorithm on special webs with our implementation.

KEY WORDS: PageRank algorithm, linear algebra, eigenvector,

Predhovor

Google je celosvetový vyhľadávač informácií na internete. Pre niektorých laikov je Google synonymom samotného internetu, čo pri istej miere zjednodušenia je skoro pravda, pretože vo svojich datacentrách má fakticky celý dostupný internet zaindexovaný.

Google robí jedným z najpoužívanejších vyhľadávačov najmä fakt, že pri hľadaní stránok k danému kľúčovému výrazu zobrazí *dobré výsledky* na prvých miestach. Pri iných vyhľadávačoch väčšinou nastane problém, že na prvých miestach sa umiestnia nerelevantné stránky, ktoré iba obsahujú daný výraz. Za úspechom Googlu stojí najmä jeho PageRank algoritmus, ktorý ohodnocuje dôležitosť každej webovej stránky. Následne používateľovi zobrazí najviac dôležité a väčšinou aj najviac relevantné stránky ako prvé.

PageRank algoritmus priraduje dôležitosť stránkam na základe vlastného vektora upravenej matice liniek. Pre analýzu algoritmu sú preto potrebné základné znalosti lineárnej algebry. Prácu sa snažíme sprístupniť čo najviac záujemcom o danú problematiku, preto podrobne vysvetľujeme potrebné teoretické predpoklady. V práci je tiež veľa prehľadných obrázkov, na ktorých ukazujeme priebeh algoritmu na konkrétnych weboch.

Pri príprave práce sme ako literatúru používali hlavne [LM] a [BL], v ktorých autori popisujú základné aspekty algoritmu.

Obsah

1	Úvod	1
2	Základný vzťah algoritmu PageRank	3
2.1	Štruktúra webu	3
2.2	Dôležitosť stránky	4
2.3	Rovnica $\mathbf{x} = \mathbf{x}\mathbf{A}$	6
2.3.1	Náhodný surfer	7
2.4	Matica liniek	8
2.4.1	Nejednoznačné riešenie	10
2.5	Úprava stochastickej matice	12
3	Podmienky konverencie	14
3.1	Ohraničenie vlastných čísel	14
3.2	Konvergenca pre diagonalizovateľné matice	15
3.3	Vlastné čísla matice \mathbf{G}	18
3.4	Konvergenca pre všeobecnú maticu	21
3.4.1	Jordanov normálny tvar	21
3.4.2	Rozklad matice \mathbf{G}	23
3.4.3	Konvergenca členov rozkladu	24
3.4.4	Závislosť výsledného ohodnotenia od počiatočného vektora \mathbf{x}_0	26
3.5	Rýchlosť konverencie	27
3.6	Parameter α	30
4	Implementácia a správanie algoritmu	31
4.1	Správanie algoritmu na skutočnom webe	31
4.2	Implementácia algoritmu	33
4.3	Použitie algoritmu	35
4.3.1	Iteračný proces s maticou liniek \mathbf{A}	35
4.3.2	Iteračný proces s upravenou maticou liniek \mathbf{S}	36
4.3.3	Iteračný proces s \mathbf{G}	39

<i>OBSAH</i>	viii
4.4 Grafy s cyklom z ktorého nevychádza žiadna linka	43
5 Využitie PageRank algoritmu	46
6 Záver	47

Kapitola 1

Úvod

Veľa z dnešných vyhľadávačov používa dvojkrokový proces pri hľadaní stránok ku danému výrazu. V prvom kroku sa pozbierajú všetky dokumenty, ktoré obsahujú hľadaný výraz, alebo výraz podobný v sémantickom zmysle. Tento proces môže viesť k tisícom relevantných stránok ku danému výrazu. Aby bol tento zoznam použiteľný treba zoradiť stránky v ňom na základe nejakého kritéria. Jednou z možností je zoradiť stránky na základe informácii plynúcich zo štruktúry webu. Presnejšie zo štruktúry hypertextových odkazov medzi jednotlivými stránkami. Jeden z úspešných algoritmov využívajúci štruktúru liniek je PageRank použitý vo vyhľadávači Google.

Vyhľadávač Google používa kombinované skóre, pričom hodnotí stránky aj na základe ich obsahu. Celkové poradie stránok získa tak, že stránky prislúchajúce hľadanému výrazu skombinuje s výsledkami PageRank algoritmu. V práci sa budem zaoberať iba PageRank algoritmom a štruktúrou hypertextových odkazov medzi jednotlivými stránkami. Pre ďalšie informácie o tom ako skombinovať skóre odporúčam pozrieť napríklad [LM, Kapitola 2.3].

Samotný PageRank algoritmus je pomenovanie procesu, ktorého vstupnými dátami je štruktúra webu a jeho výstupom je usporiadaný zoznam stránok v tomto webe podľa popularity. Musíme si teda vysvetliť, že kedy bude nejaká stránka populárnejšia, dôležitejšia ako iná.

V kapitole 2 vysvetlíme čo budeme rozumieť pod pojmom dôležitosť stránky. Ukážeme že dôležitosť jednotlivých stránok webu získame ako riešenie sústavy rovníc. Túto sústavu, budeme riešiť iteratívne. PageRank algoritmom budeme odkazovať práve na tento iteratívny proces pomocou ktorého vyriešime sústavu rovníc. Ďalej určíme podmienky, ktoré musia byť splnené aby existovalo jednoznačné riešenie. Na záver odvodíme vzťah, ktorý má potenciál splniť všetky stanovené podmienky a s ktorým budeme ďalej pracovať.

Iteračné metódy sú užitočné pri riešení veľkých sústav rovníc postupným pribli-

žovaním sa k správneému výsledku. Nie každá iteračná metóda je vhodná na výpočet ľubovoľnej sústavy rovníc, niektoré metódy sa hodia iba pre určité typy rovníc. Aby bola metóda funkčná, musí konvergovať, približovať sa k správneému výsledku. V opačnom prípade, ak sa metóda odchyľuje od správneho riešenia, hovoríme o divergencii. My sme zvolili jednu z najjednoduchších iteračných metód. Jej popis sa tiež dozvieme v kapitole 2.

V kapitole 3 si povieme niečo viac o iteračnom procese použitom k určeniu výsledného ohodnotenia. Stanovíme podmienky ktoré musia byť splnené, aby sme iteračným procesom dostali riešenie. Ukážeme, že ak iteračná metóda skonverguje k nejakému výsledku, tak ten výsledok budeme považovať za výsledné ohodnotenie. Ukážeme, že podmienky konvergence závisia od veľkosti vlastných čísel špeciálnej matice, ktorú dostaneme na základe štruktúry webu. V tejto kapitole sa tiež venujeme rýchlosti konvergence.

Spomínanú sústavu rovníc možno riešiť aj inými iteračnými metódami. V kapitole 4 vysvetlíme, prečo je metóda s ktorou pracujeme vhodná. Následne vytvoríme vlastný algoritmus, ktorým budeme realizovať túto iteračnú metódu. Ukážeme výpočet ohodnotenia na špeciálnych weboch (sieťach). Budeme sa snažiť aplikovať teoretické predpoklady, ktoré sme odvodili v kapitolách 2 a 3, aby sme vysvetlili vlastnosti výsledného ohodnotenia, prípadne počtu krokov potrebného k jeho dosiahnutiu.

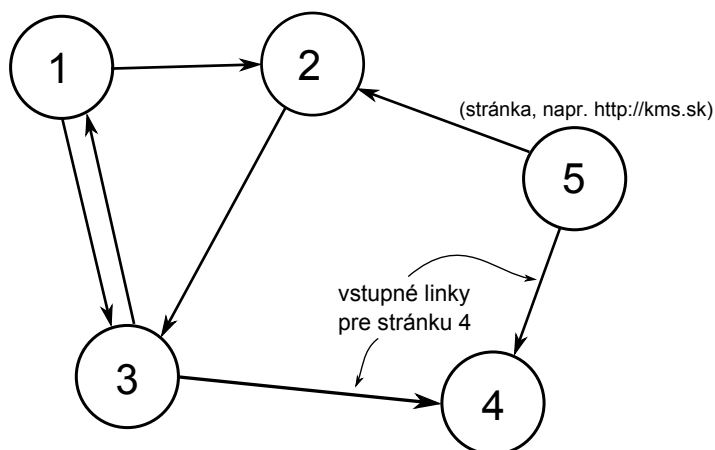
V poslednej kapitole 5 uvádzame pre zaujímavosť niekoľko ďalších využití PageRank algoritmu.

Kapitola 2

Základný vzťah algoritmu PageRank

2.1 Štruktúra webu

Vstupnými dátami pre PageRank algoritmus je hypertextová štruktúra liniek medzi jednotlivými stránkami webu. Predpokladajme, že web obsahuje n stránok/dokumentov. Každú stránku označíme celým číslom k , $1 \leq k \leq n$. Príklad webu je na obrázku 2.1.



Obr. 2.1: Príklad webu

Šípka zo stránky A do stránky B symbolizuje, že stránka A obsahuje hypertextový odkaz na stránku B . Web je teda príkladom orientovaného grafu.¹ Linky ktoré smerujú na stránku nazveme *vstupné* linky.

2.2 Dôležitosť stránky

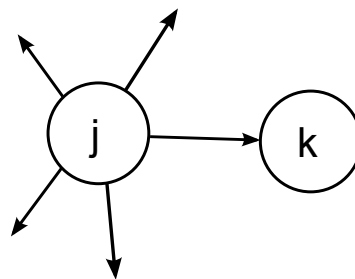
Základná myšlienka algoritmu PageRank je priradiť dôležitosť stránke na základe liniek, ktoré smerujú na túto stránku. Pod spojením dvoch stránok môžeme rozumieť odporúčanie. Linka z mojej stránky na tvoju je moje odporúčanie pre tvoju stránku. Stránka s väčším počtom odporúčaní (ktoré je určené počtom *vstupných* liniek) bude dôležitejšia ako stránka s menším.

Rovnako ako pri iných systémoch využívajúcich odporúčania (citácie článkov, osobné odporúčanie) je významné ako dôležitý je odporúčateľ. Napríklad odporúčenie od Billa Gatesa určite zaváži viac ako odporúčenie od mojich 10 neznámych kolegov alebo učiteľov.

Na druhej strane ak je odporúčateľ veľmi štedrý a za svoj život napísal vyše 20 000 odporúčaní, tak jeho odporúčenie stratí na váhe. Preto je potrebné brať ohľad na počet liniek čo vychádza z jednotlivých stránok.

Dôležitosť stránky preto získame tak, že sčítame dôležitosti stránok, ktoré ukazujú na našu stránku, a zároveň ich vydělíme počtom liniek čo z nich vychádzajú.

Symbolom x_k označíme dôležitosť stránky k . Dôležitosť stránky bude nezáporné reálne číslo. Pričom $x_j > x_k$ indikuje, že stránka j je dôležitejšia ako stránka k . Ak $x_k = 0$ tak stránka k má najmenšie možné ohodnotenie. Pozrime sa na situáciu na obrázku



Obr. 2.2: Stránka j navýši skóre stránky k o x_j/n_j , kde n_j je počet liniek vychádzajúcich zo stránky j

2.2. Stránka j ukazuje na stránku k (odporúča ju) a zároveň zo stránky j vychádza

¹Graf obsahuje množinu vrcholov (v našom prípade webových stránok) a množinu hrán. Každá hrana spája dvojicu vrcholov. Graf je *neorientovaný* ak hrany nemajú smer. Graf je *orientovaný*, ak každá hrana (v prípade webu linka) má smer a vieme povedať ktorý je začiatkový a ktorý koncový vrchol.

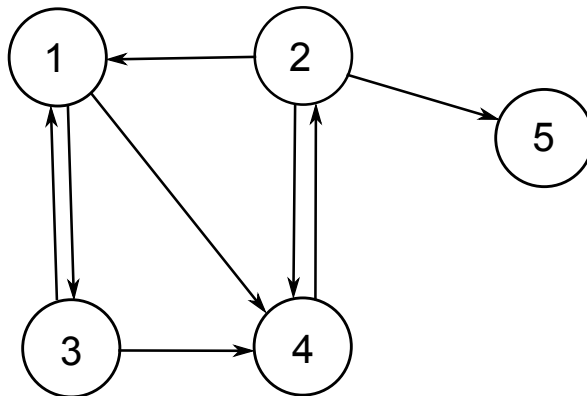
n_j liniek. Potom stránka j navýši skóre stránky k o x_j/n_j . Stránka j rozdelí svoju dôležitosť (odporúčania) rovnomerne na stránky na ktoré ukazuje. Poďme aplikovať túto myšlienku na web s n stránkami. Nech $L_k \subset \{1, 2, \dots, n\}$ značí množinu stránok, ktoré ukazujú na stránku k . Pre každú stránku k spočítame dôležitosť nasledovne:

$$x_k = \sum_{j \in L_k} \frac{x_j}{n_j} \quad (2.1)$$

pričom predpokladáme, že n_j je kladné číslo (stránka j ukazuje minimálne na stránku k , keďže $j \in L_k$).

Aplikujme vzťah (2.1) na web z obrázka 2.3 ktorý obsahuje 5 stránok. Pre stránku 1 spočítame dôležitosť ako $x_1 = x_2/3 + x_3/2$. Stránka 2 ukazuje okrem stránky 1 na stránky 3 a 5, rozdelí svoje odporúčania na tri časti. Stránka 3 ukazuje okrem stránky 1 na stránku 4, preto stránke 1 predá polovičné odporúčanie. Rovnako môžeme zapísať rovnice aj pre ostatné stránky

$$\begin{aligned} x_2 &= x_4 \\ x_3 &= x_1/2 \\ x_4 &= x_1/2 + x_2/3 + x_3/2 \\ x_5 &= x_2/3 \end{aligned}$$



Obr. 2.3: Web s 5 stránkami

Spolu s rovnicou pre x_1 nám vznikla sústava 5 lineárnych rovníc, ktorú môžeme zapísať $\mathbf{Ax}^T = \mathbf{x}^T$, kde $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$ a

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Po transponovaní dostávame rovnicu $\mathbf{x} = \mathbf{x}\mathbf{A}$. Pre každý web vieme zostrojiť takúto maticu \mathbf{A} . Túto maticu budeme nazývať *matica liniek*. O matici liniek si povieme neskôr v časti 2.4. Náš cieľ teda bude nájsť taký vektor \mathbf{x} ohodnotení stránok pre ktorý je splnená rovnica

$$\mathbf{x} = \mathbf{x}\mathbf{A} \quad (2.2)$$

V ďalšej časti sa budeme zaoberať riešením tejto rovnice.

2.3 Rovnica $\mathbf{x} = \mathbf{x}\mathbf{A}$

Problém rovnice (2.2) je ten, že nepoznáme vektor ohodnotení na pravej strane. Inak povedané nevieme akú dôležitosť majú stránky ukazujúce na ľubovoľnú stránku. Na vyriešenie tohoto problému použili Brin a Page iteratívny proces. Predpokladajme, že všetky stránky majú na začiatku rovnakú dôležitosť. Pre web s n stránkami to bude:

$$\mathbf{x}_0 = (1/n, 1/n, \dots, 1/n)$$

Keď vynásobíme ohodnotenie \mathbf{x}_0 maticou liniek

$$\mathbf{x}_1 = \mathbf{x}_0\mathbf{A}$$

dostaneme nové ohodnotenie dôležitosti \mathbf{x}_1 , v ktorom sa prejaví štruktúra liniek a odlišia sa menej dôležité stránky od dôležitejších. Ak chceme, aby hlas dôležitejších zavážil viac, zoberieme predchádzajúci výsledok (ohodnotenie \mathbf{x}_1) a vynásobíme ho maticou \mathbf{A} . V každom kroku počítame nové ohodnotenie \mathbf{x}_{k+1} na základe predošlého podľa vzťahu

$$\mathbf{x}_{k+1} = \mathbf{x}_k\mathbf{A} \quad (2.3)$$

Postupne počítame:

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{x}_0\mathbf{A} \\ \mathbf{x}_2 &= \mathbf{x}_1\mathbf{A} = \mathbf{x}_0\mathbf{A}^2 \\ &\vdots \\ \mathbf{x}_n &= \mathbf{x}_{n-1}\mathbf{A} = \mathbf{x}_0\mathbf{A}^n \end{aligned}$$

V prípade, že vektor dôležitosti jednotlivých stránok \mathbf{x}_n bude konvergovať k nejakej hodnote, tak je prirodzené, že túto hodnotu budeme považovať za výsledné ohodnotenie. V tomto prípade existuje limita:

$$\mathbf{x} = \lim_{n \rightarrow \infty} \mathbf{x}_0\mathbf{A}^n$$

potom platí aj

$$\mathbf{x}\mathbf{A} = \lim_{n \rightarrow \infty} \mathbf{x}_0 \mathbf{A}^{n+1} = \mathbf{x}. \quad (2.4)$$

Tu sme využili fakt, že v priestore \mathbb{R}^n je každé lineárne zobrazenie spojité. V našom prípade je zobrazenie realizované násobením maticou liniek \mathbf{A} . Teraz nasleduje stručné zdôvodnenie tohoto faktu tak, ako je uvedené v [S, kapitola 3.6]

Ak označíme $\|\mathbf{A}\|_{max} = \max_{i,j} |a_{ij}|$, tak očividne pre každý vektor \mathbf{z} taký, že $\max |z_i| < \delta$ platí, že všetky súradnice vektora $\mathbf{z}\mathbf{A}$ nepresahujú v absolútnej hodnote $\delta n \|\mathbf{A}\|_{max}$. Takže ak máme dané $\epsilon > 0$ a dvojicu vektorov takú, že pre všetky ich súradnice platí

$$|x_i - y_i| < \frac{\epsilon}{n \|\mathbf{A}\|_{max}},$$

tak potom

$$\mathbf{x}\mathbf{A} - \mathbf{y}\mathbf{A} = (\mathbf{x} - \mathbf{y})\mathbf{A} < n \|\mathbf{A}\|_{max} \frac{\epsilon}{n \|\mathbf{A}\|_{max}} = \epsilon.$$

Pri skúmaní riešení rovnice (2.2) iteračným spôsobom sa objavuje niekoľko otázok na ktoré je treba zodpovedať.

- Bude iteratívny proces (2.3) pokračovať do nekonečna, alebo skonverguje ku nejakej hodnote?
- Aké vlastnosti musí mať matica liniek, aby proces skonvergoval?
- Skonverguje ku niečomu čo súvisí s hľadaním ohodnotení pre stránky webu?
- Zmení sa vektor ku ktorému skonverguje proces, pri zmene štartovacieho vektora \mathbf{x}_0 ?
- V prípade že proces skonverguje, koľkokrát musíme vykonať násobenie aby sme dostali výsledok s dostatočnou presnosťou?

Existencii vektora, ku ktorému proces skonverguje, sa budeme venovať v ďalších častiach tejto kapitoly. O konvergencii si povieme neskôr v kapitole 3. Najprv ešte uvedieme zaujímavé prepojenie problému hľadania dôležitosti pre stránky webu s teóriou pravdepodobnosti.

2.3.1 Náhodný surfer

Ďalšia možnosť ako interpretovať násobenie maticou liniek je sledovanie pohybu *náhodného surfera* po webe. Predstavme si surfera, ktorý sleduje hyperlinkovú štruktúru webu a na každej stránke si vyberie linku, ktorou bude pokračovať ďalej s rovnakou pravdepodobnosťou. Problém vznikne, keď príde surfer na stránku z ktorej nevychádza žiadna linka. Vtedy si náhodne vyberie ľubovoľnú stránku a pokračuje v browsovaní

ďalej. Predpokladajme, že vektor \mathbf{x}_n konverguje ku vektoru \mathbf{y} . Súradnicu y_i vektora \mathbf{y} môžeme považovať za hodnotu pravdepodobnosti s ktorou sa surfer nachádza na stránke i v prípade, že ho necháme surfovať neobmedzene dlho.

Vektor \mathbf{x}_0 symbolizuje počiatočnú pravdepodobnosť s ktorou sa surfer na začiatku nachádza v jednotlivých stránkach. Vo všeobecnosti môže mať hodnotu inú ako $\mathbf{x}_0 = (1/n, 1/n, \dots, 1/n)$. Pre tento vektor musí platiť, že hodnoty jeho súradníc sú nezáporné a ich súčet je 1, keďže hovoríme o pravdepodobnosti.

Toto pozorovanie môžeme interpretovať maticou liniek tak, že nenulové prvky a_{ij} v nejakom riadku i (pre stránku i) symbolizujú pravdepodobnosť s ktorou bude surfer pokračovať na danej stránke. Nulové riadky v matici liniek nahradíme vektorom $1/n \mathbf{e}$, kde $\mathbf{e} \in \mathbb{R}^n$ je vektor, ktorý má na všetkých súradniciach hodnotu $1/n$. Znamená to, že na ľubovoľnú stránku sa presunie s rovnakou pravdepodobnosťou $1/n$. Upravená matica liniek potom vyzerá nasledovne:

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

Práve tento pohľad dáva do súvisu PageRank algoritmus s markovovskými reťazcami. Tie sa študujú v teórii stochastických procesov, čo je matematická oblasť patriaca do teórie pravdepodobnosti.

V matici liniek sme nahradili nulové riadky vektorom $1/n \mathbf{e}$. V časti 2.4 ukážeme, že táto úprava nám zabezpečí aby takáto upravená matica mala nové vlastnosti ktoré budeme môcť využiť.

2.4 Matica liniek

Matica liniek má zaujímavé vlastnosti. Každý stránke webu prislúcha jeden riadok v tejto matici. Nejaký prvok a_{ij} matice liniek \mathbf{A} je nenulový práve vtedy keď stránka i ukazuje na stránku j . Súčet v každom riadku je buď jedna, alebo nula. Niektoré riadky v matici liniek obsahujú samé nuly. Ku týmto riadkom prislúchajú presne tie stránky z ktorých nevychádza žiadna linka. Tieto vrcholy nazveme *slepé*. Ukážeme, že slepé vrcholy spôsobujú problém a pre maticu liniek webu so slepými vrcholmi nemusí existovať riešenie rovnice.

Pripomenieme teraz definíciu vlastného čísla, vlastného vektora a charakteristického polynómu, ktoré budeme v ďalšom texte často používať.

Definícia 2.4.1. Nech \mathbf{A} je štvorcová matica nad poľom F . Prvok $c \in F$ nazveme *vlastným číslom* matice \mathbf{A} , ak existuje nenulový vektor $\mathbf{x} \in F^n$ taký, že $\mathbf{x}\mathbf{A} = c\mathbf{x}$.

Nenulový vektor $\mathbf{x} \in F^n$ nazývame *vlastným vektorom* matice \mathbf{A} , ak existuje $c \in F$ (c môže byť aj nula) také, že $\mathbf{x}\mathbf{A} = c\mathbf{x}$.

Definícia 2.4.2. Hodnotu determinantu $|\mathbf{A} - \mathbf{xI}|$ nazývame *charakteristický polynóm* matice \mathbf{A} . Nájst' vlastné čísla matice znamená nájsť korene charakteristického polynómu.

Riešiť rovnicu (2.2) znamená nájsť vlastné vektory ku vlastnému číslu 1 matice \mathbf{A} . Ku každému webu vieme zostrojiť takúto maticu. No nie je zaručené že táto matica bude mať vlastné číslo 1. Musíme teda upraviť maticu liniek, aby vždy mala vlastnú hodnotu 1.

Definícia 2.4.3. Štvorcová matica A je *riadkovo stochastická* ak sú všetky jej prvky nezáporné a v každom riadku je súčet prvkov 1.

Tvrdenie 2.4.1. Každá riadkovo stochastická matica \mathbf{A} má vlastné číslo 1.

Dôkaz. Nech má matica \mathbf{A} rozmery $n \times n$. Označme \mathbf{e} n -rozmerný riadkový vektor, ktorý má na všetkých súradniciach číslo 1. Keďže matica \mathbf{A} je riadkovo stochastická, tak platí $\mathbf{A}\mathbf{e}^T = \mathbf{e}^T$. Vektor $\mathbf{A}\mathbf{e}^T$ má na k -tej súradnici presne súčet prvkov v k -tom riadku.

Keď obidve strany rovnosti transponujeme dostaneme rovnosť $\mathbf{e}\mathbf{A}^T = \mathbf{e}$. Teda 1 je vlastné číslo matice \mathbf{A}^T . Matice \mathbf{A} a \mathbf{A}^T majú rovnaký charakteristický polynóm, preto majú rovnaké vlastné čísla. Kvôli tomu má matica \mathbf{A} vlastné číslo 1. \square

Z prechádzajúceho tvrdenia plynie, že keby sme upravili maticu liniek tak, aby v nej neboli nulové riadky, mali by sme zaručenú existenciu vlastného vektora k vlastnej hodnote 1. Tým pádom by sme mali kandidáta na výsledné ohodnotenie. Nahradíme teraz nulové riadky v matici liniek vektorom $1/n \mathbf{e}$ rovnako ako v časti, kde sme hovorili o náhodnom surferovi. Túto úpravu môžeme zapísať rovnicou

$$\mathbf{S} = \mathbf{A} + \mathbf{a}^T(1/n \mathbf{e}),$$

kde súradnica a_i vektora \mathbf{a} je 1 ak je vrchol i slepý, v opačnom prípade je 0. V ďalšom texte budem symbolom \mathbf{S} označovať maticu, ktorú dostanem nahradením nulových riadkov vektorom $1/n \mathbf{e}$. Túto maticu budeme v ďalšom texte nazývať *upravená matica liniek*. Vždy keď budeme hovoriť o matici \mathbf{S} pre web, tak budeme mať na mysli práve túto maticu, ktorú sme dostali nahradením nulových riadkov v matici liniek \mathbf{A} .

Pre maticu \mathbf{S} , ktorú sme dostali predchádzajúcou úpravou platí, že existuje vektor \mathbf{x} , taký že $\mathbf{x}\mathbf{S} = \mathbf{x}$ (výsledné ohodnotenie). Nemáme však zaručené, že existuje jediný

taký vektor. Symbolom $V_1(\mathbf{S})$ budeme označovať vektorový priestor vlastných vektorov prislúchajúcich ku vlastnému číslu 1.

2.4.1 Nejednoznačné riešenie

Pre naše potreby je vhodné aby dimenzia priestoru $V_1(\mathbf{S})$ bola rovná jednej. V tom prípade by existoval jediný vektor \mathbf{y} (až na skalárny násobok), ktorý by generoval priestor $V_1(\mathbf{S})$. Nasledujúce tvrdenia nám určia ďalšie podmienky, ktoré zabezpečia, aby dimenzia priestoru $V_1(\mathbf{S})$ bola jedna.

Tvrdenie 2.4.2. *Nech matica \mathbf{S} je riadkovo stochastická a $s_{ij} > 0$ pre $i, j = 1, \dots, n$. Potom každý vektor $\mathbf{y} \in V_1(\mathbf{S})$ má všetky súradnice rovnakého znamienka (všetky nezáporné, alebo všetky nekladné).*

Dôkaz. Pri dôkaze tohoto tvrdenia využijeme nerovnosť $|\sum_{i=1}^n y_i| \leq \sum_{i=1}^n |y_i|$. Táto nerovnosť je ostrá pre každý vektor, ktorý obsahuje prvky so zmiešanými znamienkami.

Z predpokladu je vektor \mathbf{y} vlastný vektor ku vlastnému číslu 1, preto $\mathbf{y} = \mathbf{yS}$. Predpokladajme teraz sporom, že vektor \mathbf{y} má na niektorých súradniciach rôzne znamienka. Platí potom ostrá nerovnosť:

$$|y_i| = \left| \sum_{j=1}^n s_{ji} y_j \right| < \sum_{j=1}^n s_{ji} |y_j|.$$

Pre $i = 1, \dots, n$ sčítame vyjadrenia y_i , dostávame tak

$$\sum_{i=1}^n |y_i| < \sum_{i=1}^n \sum_{j=1}^n s_{ji} |y_j| \stackrel{(1)}{=} \sum_{j=1}^n |y_j| \underbrace{\sum_{i=1}^n s_{ji}}_1 = \sum_{j=1}^n |y_j|$$

Pri úprave (1) sme vymenili poradie sumácie. Dostali sme nerovnosť

$$\sum_{i=1}^n |y_i| < \sum_{j=1}^n |y_j|$$

čo je spor. Platí teda, že vektor \mathbf{y} nemôže mať na niektorých súradniciach rôzne znamienka. \square

Tvrdenie 2.4.3. *Nech \mathbf{a} a \mathbf{b} sú lineárne nezávislé vektory z priestoru \mathbb{R}^n , $n \geq 2$. Potom existujú koeficienty $c, d \in \mathbb{R}$ také že vektor $c\mathbf{a} + d\mathbf{b}$ obsahuje súradnice so zmiešanými znamienkami.*

Dôkaz. Lineárna nezávislosť implikuje, že \mathbf{a} a \mathbf{b} sú nenulové vektory. Ak pre vektor $\mathbf{a} = (a_1, \dots, a_n)$ platí $\sum_{i=1}^n a_i = 0$, tak určite obsahuje súradnice so smiešanými znamienkami. Stačí zvoliť $c = 1$ a $d = 0$. Naopak, ak súčet súradníc vektora \mathbf{b} je nula, zvolíme $c = 0$ a $d = 1$.

V prípade, že žiadny z vektorov nedáva nulový súčet súradníc, zvolíme c a d tak, aby $c \sum_{i=1}^n a_i + d \sum_{i=1}^n b_i = 0$. Kvôli lineárnej nezávislosti vektorov \mathbf{a} a \mathbf{b} bude vektor $c\mathbf{a} + d\mathbf{b}$ nenulový. Dostávame, že znamienka jeho súradníc nemôžu byť rovnaké. \square

Vďaka tvrdeniam 2.4.2 a 2.4.3 dostávame, že ak je matica \mathbf{S} kladná ($s_{ij} > 0, i, j \in 1, \dots, n$) a riadkovo stochastická, tak neexistujú lineárne nezávislé vektory vo $V_1(\mathbf{S})$. Preto je dimenzia priestoru $V_1(\mathbf{S})$ rovná 1. Preto existuje práve jeden vektor $\mathbf{x} \in V_1(\mathbf{S})$ taký, že $\sum_{i=1}^n x_i = 1$.

Ukázali sme, že existuje jediný kandidát na výsledné ohodnotenie a to vlastný vektor prislúchajúci ku vlastnému číslu 1. V časti 2.3 sme zase ukázali, že ak iteratívny proces konverguje ku nejakému vektoru, tak tento vektor spĺňa rovnicu (2.4). Je teda vlastným vektorom ku vlastnému číslu 1. Iteratívny proces preto skonverguje práve k vlastnému vektoru matice \mathbf{S} a teda ku výslednému ohodnoteniu dôležitosti stránok.

Teraz ukážeme, že lineárne zobrazenie prislúchajúce riadkovo stochastickej matici \mathbf{S} nemení súčet súradníc vektora. V praxi to znamená, že ak začneme iteratívny proces so štartovacím vektorom $\mathbf{x}_0 = (1/n, 1/n, \dots, 1/n)$, ktorý má súčet súradníc 1, tak aj výsledný vektor bude mať súčet súradníc 1.

Tvrdenie 2.4.4. *Nech \mathbf{S} je riadkovo stochastická matica, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ a $\mathbf{y} = (y_1, y_2, \dots, y_n) = \mathbf{xS}$. Potom platí $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$*

Dôkaz. Pre i -tu súradnicu vektora \mathbf{y} platí

$$y_i = \sum_{j=1}^n s_{ji} x_j$$

Sčítaním týchto rovníc dostaneme

$$\sum_{i=1}^n y_i = \sum_{i=1}^n \sum_{j=1}^n s_{ji} x_j = \sum_{j=1}^n x_j \underbrace{\sum_{i=1}^n s_{ji}}_1 = \sum_{j=1}^n x_j$$

\square

Predchádzajúce tvrdenia nám lepšie objasňuje pohyb náhodného surfera po webe o ktorom sme hovorili v podkapitole 2.3.1 Počiatočnú pravdepodobnosť, že sa surfer nachádza v jednotlivých stránkach sme označovali \mathbf{x}_0 . Tento vektor má súčet súradníc 1. Pravdepodobnosť s ktorou bude surfer v jednotlivých vrcholoch po jednom kroku dostaneme tak, že vynásobíme vektor \mathbf{x}_0 upravenou maticou liniek \mathbf{S} . Pre nové rozloženie bude tiež platiť, že súčet pravdepodobností s ktorou sa nachádza v jednotlivých vrcholoch je jedna.

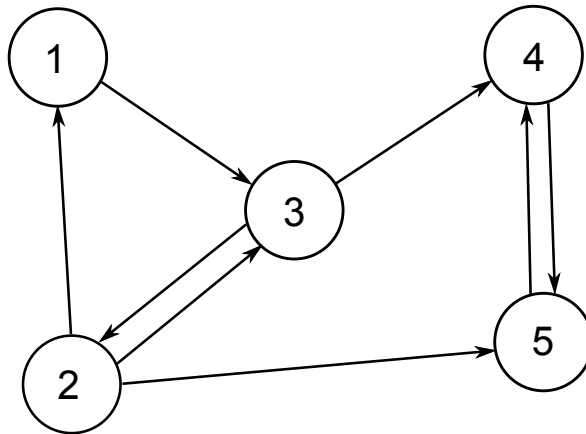
V prípade, že matica liniek \mathbf{A} má nulový riadok, tak nebude platiť, že súčet ohodnotení v každom kroku je 1. Preto v iteračnom procese s maticou liniek \mathbf{A} nebude

vektor ohodnotenia v k -tom kroku zodpovedať pravdepodobnosti výskytu v jednotlivých stránkach. Súčet súradníc vektora ohodnotenia v k -tom kroku môže dokonca klesať. Tento problém ukazujeme v časti 4.3.1.

V tvrdení 2.4.3 je daná ďalšia podmienka na riadkovo stochastickú maticu \mathbf{S} . Aby sme zaručili existenciu jednoznačného riešenia musíme upraviť maticu \mathbf{S} prislúchajúcu matici liniek tak, aby všetky jej prvky boli kladné.

2.5 Úprava stochastickej matice

Problém slepých vrcholov sme vyriešili tak, že sme nulové riadky, ktoré im prislúchajú v matici liniek nahradili vektorom $1/n\mathbf{e}$. Ďalší problém ktorý treba vyriešiť je problém cyklov. Pozrime sa na web z obrázka 2.4. Keď spustíme iteračný pro-



Obr. 2.4: Graf s cyklom

ces s počiatočným vektorom $\mathbf{x}_0 = (1/5, 1/5, 1/5, 1/5, 1/5)$ a stochastickou maticou prislúchajúcou k tomuto webu, tak po 15 iteráciách dostaneme ohodnotenie $\mathbf{x} = (0.0002, 0.0004, 0.0005, 0.4662, 0.5328)$. Dôležitosť sa naakumulovala v stránkach 4 a 5. Keď sa raz dostal náhodný surfer do týchto stránok, tak z nich už ďalej nevyšiel. Totižto do zvyšku webu nevychádza žiadna linka. Celkové skóre sa tým pádom rozdelí medzi stránky, ktoré sú súčasťou nejakého cyklu. Stránky, ktoré nie sú súčasťou žiadneho cyklu skončia s ohodnotením veľmi blízkym nule. Tomuto potrebujeme zabrániť, pretože nechceme aby výsledné ohodnotenie malo túto vlastnosť. V časti 4.3.2 ukazujeme vývoj ohodnotenia pre jednotlivé kroky iteračného procesu pre tento graf.

Označme \mathbf{E} štvorcovú maticu veľkosti $n \times n$, ktorá má všetky prvky rovné $1/n$. Nahradme maticu \mathbf{S} maticou

$$\mathbf{G} = \alpha\mathbf{S} + (1 - \alpha)\mathbf{E} = \alpha\mathbf{S} + (1 - \alpha)\frac{1}{n}\mathbf{e}^T\mathbf{e},$$

kde $0 \leq \alpha \leq 1$. Matica \mathbf{G} zodpovedá tomu, že pohybujúci sa surfer v istom percente prípadov (určenom parametrom $1 - \alpha$) rozhodne nesledovať niektorú z vychádzajúcich liniek zo stránky, ale novú stránku si vyberie úplne náhodne. Môžeme povedať, že sa „teleportuje“ na inú stránku. Preto maticu \mathbf{E} budeme nazývať *teleportačná* matica. Podme teraz vyšetriť aké vlastnosti má táto nová matica.

Matica \mathbf{G} je riadkovo stochastická pre ľubovoľné $\alpha \in \langle 0, 1 \rangle$, lebo súčet parametrom α a $(1 - \alpha)$ je 1. Vyššie sme ukázali, že ak $\alpha \in \langle 0, 1 \rangle$, tak potom dimenzia priestoru $V_1(\mathbf{G})$ bude 1, lebo všetky prvky matice \mathbf{G} sú kladné. Bude preto existovať jednoznačné ohodnotenie stránok.

Pohybujeme sa medzi dvoma extrémnymi hodnotami. Ak je parameter α rovný nule, tak $\mathbf{G} = \mathbf{E}$. V tomto prípade nebude vôbec záležať na matici liniek. Všetky stránky budú ohodnotené rovnako $x_i = 1/n$. Na druhej strane ak $\alpha = 1$, tak riešime pôvodnú úlohu pri ktorej sú prítomné problémy s cyklami. Predpokladajme teda, že pre parameter α platí $\alpha \in (0, 1)$, aby nenastala žiadna z týchto situácií.

V nasledujúcej kapitole 3 budeme hovoriť o účinku tohoto parametra podrobnejšie.

V tejto kapitole sme odvodili vzťah pre maticu \mathbf{G} , s ktorou budeme násobiť počiatočný vektor ohodnotenia \mathbf{x}_0 , ktorý má tvar

$$\mathbf{G} = \alpha \mathbf{S} + (1 - \alpha) \mathbf{E}.$$

Maticou \mathbf{G} budeme odteraz stále označovať maticu z predošlého vzťahu. Odkazujeme ňou teda na dvojnásobne upravenú maticu liniek \mathbf{A} (najprv sme prešli od \mathbf{A} ku upravenej matici liniek \mathbf{S} a potom od matice \mathbf{S} ku matici \mathbf{G}). S maticou \mathbf{G} budeme viac pracovať v nasledujúcej kapitole, v ktorej budeme riešiť problém konvergencie iteračného procesu.

Kapitola 3

Podmienky konvergenzie

V minulej kapitole sme ukázali, že existuje jediný kandidát na ohodnotenie a to vlastný vektor matice \mathbf{G} prislúchajúci ku vlastnému číslu 1. Tento vektor môžeme získať pomocou iteračného procesu, ktorý vieme zapísať rovnicou $\mathbf{x}_{k+1} = \mathbf{x}_k \mathbf{G}$. Pričom začneme s nejakým počiatočným vektorom ohodnotení \mathbf{x}_0 . Ak proces skonverguje, tak pre vektor

$$\lim_{n \rightarrow \infty} \mathbf{x}_0 \mathbf{G}^n$$

bude platiť

$$\mathbf{x} \mathbf{G} = \mathbf{x}.$$

V tejto kapitole ukážeme, že iteratívny proces s maticou \mathbf{G} konverguje ku riešeniu. Najprv dokážeme konvergenciu pre špeciálny typ matíc, ktoré sú diagonalizovateľné matice, potom pre všeobecné. Ďalej sa budeme venovať parametru α , ktorý sa vyskytuje v matici \mathbf{G} . Ukážeme, že tento parameter výrazne ovplyvňuje rýchlosť konvergenzie.

3.1 Ohraničenie vlastných čísel

V závere kapitoly 2 sme odvodili vzťah pre maticu \mathbf{G} v ktorom sa vyskytuje parameter α . O tejto matici nevieme povedať veľa, vďaka jej nejednoznačnosti. Sústreďme sa preto najprv na upravenú maticu liniek \mathbf{S} , ktorá sa vyskytuje vo vyjadrení matice \mathbf{G} . Konvergenciu k riešeniu najprv dokážeme pre špeciálny typ matíc \mathbf{G} . Použitú metódu budeme môcť s menšími úpravami aplikovať aj na všeobecnú maticu \mathbf{G} . Teraz ukážeme, že vlastnosti konvergenzie iteračného procesu závisia od vlastných čísel upravenej matice liniek \mathbf{S} . Nasledujúce tvrdenie hovorí o tom, že vlastné čísla riadkovo stochastickej matice sú ohraničené.

Tvrdenie 3.1.1. *Nech matica \mathbf{S} je riadkovo stochastická a $s_{ij} \geq 0$ pre $i, j = 1, \dots, n$. Ak λ je jej vlastné číslo, tak $|\lambda| \leq 1$.*

Dôkaz. Pre vlastné číslo λ a príslušný vlastný vektor $\mathbf{x} = (x_1, x_2, \dots, x_n)$ máme:

$$\lambda x_i = \sum_{j=1}^n s_{ji} x_j$$

$$|\lambda| |x_i| = \left| \sum_{j=1}^n s_{ji} x_j \right|$$

Sčítaním týchto nerovností dostaneme

$$|\lambda| \sum_{i=1}^n |x_i| \leq \sum_{i=1}^n \sum_{j=1}^n s_{ji} |x_j| = \sum_{j=1}^n |x_j| \underbrace{\sum_{i=1}^n s_{ji}}_1 = \sum_{j=1}^n |x_j|$$

Vektor \mathbf{x} je nenulový, preto po vydelení oboch strán nerovnosti kladným číslom

$$\sum_{j=1}^n |x_j|$$

dostaneme $|\lambda| \leq 1$. □

3.2 Konvergencia pre diagonalizovateľné matice

V tejto časti budeme hovoriť o konvergencii iteračného procesu pre špeciálny typ matíc, ktoré sú *diagonalizovateľné*.

Definícia 3.2.1. Nech \mathbf{A} a \mathbf{B} sú štvorcové matice nad poľom \mathbb{F} . Hovoríme, že matice \mathbf{A} a \mathbf{B} sú podobné, ak existuje regulárna matica \mathbf{P} taká, že $\mathbf{B} = \mathbf{PAP}^{-1}$.

Definícia 3.2.2. Matica \mathbf{A} je *diagonalizovateľná*, ak existuje taká diagonálna matica \mathbf{D} , že \mathbf{A} a \mathbf{D} sú podobné. Diagonálna matica je taká, ktorá má nenulové prvky iba na diagonále.

Ak je matica \mathbf{S} podobná s diagonálnou maticou \mathbf{D} , ktorá má na diagonále prvky $\lambda_1, \dots, \lambda_n$, tak platia rovnosti

$$\mathbf{PSP}^{-1} = \mathbf{D}$$

$$\mathbf{PS} = \mathbf{DP}$$

Ak označíme riadky matice \mathbf{P} ako $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$, potom z predchádzajúcej rovnosti

dostávame

$$\begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \mathbf{S} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{a}_1 \mathbf{S} \\ \vdots \\ \mathbf{a}_n \mathbf{S} \end{pmatrix} = \begin{pmatrix} \lambda_1 \mathbf{a}_1 \\ \vdots \\ \lambda_n \mathbf{a}_n \end{pmatrix}$$

Ak porovnáme riadky v poslednej rovnosti dostaneme pre každé i vzťah

$$\mathbf{a}_i \mathbf{S} = \lambda_i \mathbf{a}_i.$$

Preto sú $\lambda_1, \dots, \lambda_n$ vlastné čísla matice \mathbf{S} a $\mathbf{a}_1, \dots, \mathbf{a}_n$ ku nim prislúchajúce vlastné vektory. Ak je matica \mathbf{S} podobná s diagonálnou, tak táto diagonálna matica má na uhlopriečke vlastné čísla matice \mathbf{S} .

Nasledujúce tvrdenie bude kľúčové ku dôkazu, že iteračný proces skonverguje. V odbornej literatúre ho možno nájsť ako vetu o *spektrálnom rozklade*.

Tvrdenie 3.2.1. *Ak je štvorcová matica \mathbf{S} podobná s diagonálnou maticou \mathbf{D} , čo má na uhlopriečke hodnoty $\lambda_1, \dots, \lambda_n$, tak existujú matice $\mathbf{G}_1, \dots, \mathbf{G}_n$, také, že platí:*

$$\mathbf{S} = \lambda_1 \mathbf{G}_1 + \lambda_2 \mathbf{G}_2 + \cdots + \lambda_n \mathbf{G}_n \quad (3.1)$$

a súčasne

(i) Platí rovnosť $\mathbf{G}_1 + \mathbf{G}_2 + \cdots + \mathbf{G}_n = \mathbf{I}$.

(ii) Pre každé $i = 1, \dots, n$ platí $\mathbf{G}_i^2 = \mathbf{G}_i$.

(iii) Pre $i \neq j$; $i, j = 1, \dots, n$ platí $\mathbf{G}_i \mathbf{G}_j = \|\mathbf{0}\|$.

Dôkaz. Podľa predpokladu existuje regulárna matica \mathbf{P} taká, že $\mathbf{PSP}^{-1} = \mathbf{D}$. Platí teda $\mathbf{S} = \mathbf{P}^{-1} \mathbf{D} \mathbf{P}$. Označme stĺpce matice \mathbf{P}^{-1} ako $\mathbf{a}_1^T, \dots, \mathbf{a}_n^T$ a riadky matice \mathbf{P} ako $\mathbf{b}_1, \dots, \mathbf{b}_n$.

$$\mathbf{P}^{-1} = (\mathbf{a}_1^T, \dots, \mathbf{a}_n^T) \quad \mathbf{P} = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix}$$

Dosadíme teraz do rovnosti $\mathbf{S} = \mathbf{P}^{-1} \mathbf{D} \mathbf{P}$ vyjadrenie \mathbf{P}^{-1} a \mathbf{P} . Dostávame:

$$\mathbf{S} = (\mathbf{a}_1^T, \dots, \mathbf{a}_n^T) \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} = \lambda_1 \mathbf{a}_1^T \mathbf{b}_1 + \cdots + \lambda_n \mathbf{a}_n^T \mathbf{b}_n.$$

Označme $\mathbf{G}_i = \mathbf{a}_i^T \mathbf{b}_i$. Pre takto definované matice $\mathbf{G}_1, \dots, \mathbf{G}_n$ platí, že sú veľkosti $n \times n$. Ďalej pre ne dostávame

$$\mathbf{S} = \lambda_1 \mathbf{G}_1 + \lambda_2 \mathbf{G}_2 + \dots + \lambda_n \mathbf{G}_n$$

Teraz potrebujeme overiť tvrdenia (i) – (iii).

Rovnosť $\mathbf{I} = \mathbf{P}^{-1} \mathbf{P}$ môžeme prepísať ako

$$\mathbf{I} = (\mathbf{a}_1^T, \dots, \mathbf{a}_n^T) \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} = \mathbf{a}_1^T \mathbf{b}_1 + \dots + \mathbf{a}_n^T \mathbf{b}_n = \mathbf{G}_1 + \dots + \mathbf{G}_n.$$

Čo je dôkaz tvrdenia (i). Ak vynásobíme $\mathbf{P} \mathbf{P}^{-1}$, tak dostaneme

$$\mathbf{I} = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} (\mathbf{a}_1^T, \dots, \mathbf{a}_n^T) = \begin{pmatrix} \mathbf{b}_1 \mathbf{a}_1^T & \dots & \mathbf{b}_1 \mathbf{a}_n^T \\ \vdots & \ddots & \vdots \\ \mathbf{b}_n \mathbf{a}_1^T & \dots & \mathbf{b}_n \mathbf{a}_n^T \end{pmatrix}$$

Porovnaním matíc na ľavej a pravej strane dostaneme $\mathbf{b}_i \mathbf{a}_i^T = 1$ a $\mathbf{b}_i \mathbf{a}_j^T = 0$ pre $i \neq j$. Môžeme teraz dosadiť do vyjadrení \mathbf{G}_i^2 a $\mathbf{G}_i \mathbf{G}_j$, pričom sa držíme pôvodného označenia $\mathbf{G}_i = \mathbf{a}_i^T \mathbf{b}_i$. Dostávame:

$$\mathbf{G}_i^2 = \mathbf{a}_i^T \underbrace{\mathbf{b}_i \mathbf{a}_i^T}_1 \mathbf{b}_i = \mathbf{a}_i^T \mathbf{b}_i = \mathbf{G}_i$$

a

$$\mathbf{G}_i \mathbf{G}_j = \mathbf{a}_i^T \underbrace{\mathbf{b}_i \mathbf{a}_j^T}_0 \mathbf{b}_j = \|\mathbf{0}\|.$$

Čím sme dokázali platnosť tvrdení (ii) a (iii). □

Pre ohodnotenie stránok v k -tom kroku iteračného procesu platí

$$\mathbf{x}_k = \mathbf{x}_0 \mathbf{S}^k.$$

Na základe prechádzajúcej vety môžeme prepísať výraz \mathbf{S}^k do nasledujúceho tvaru

$$\mathbf{S}^k = \lambda_1^k \mathbf{G}_1 + \lambda_2^k \mathbf{G}_2 + \dots + \lambda_n^k \mathbf{G}_n \quad (3.2)$$

Vo vyjadrení pre k -ty krok iteračného procesu sa vyskytujú vlastné čísla matice \mathbf{S} . Riadkovo stochastická matica \mathbf{S} má vlastné číslo 1. Bez ujmy na všeobecnosti môžeme predpokladať, že $1 = \lambda_1 \geq |\lambda_2| \geq \dots \geq |\lambda_n| \geq 0$.¹ V opačnom prípade môžeme vlastné čísla preznačiť. Dostávame teda

$$\mathbf{x}_k = \mathbf{x}_0 \mathbf{G}_1 + \lambda_2^k \mathbf{x}_0 \mathbf{G}_2 + \dots + \lambda_n^k \mathbf{x}_0 \mathbf{G}_n$$

¹Vlastné čísla môžu byť aj komplexné, môžeme ich však usporiadať podľa absolútnej hodnoty.

V úvode kapitoly sme vyslovili tvrdenie 3.1.1, podľa ktorého sú vlastné čísla matice \mathbf{S} ohraňované. Pre tie vlastné čísla λ_i , ktoré sú v absolútnej hodnote ostro menšie ako 1, platí

$$\lambda_i^k \longrightarrow 0.$$

Medzi vlastnými číslami matice \mathbf{S} sa môže vyskytovať niekoľko vlastných čísel, ktoré majú absolútnu hodnotu rovnú jednej (všetky komplexné čísla s veľkosťou 1). Ukázali sme že určite je medzi nimi číslo 1. Môže byť medzi nimi tiež vlastné číslo -1 . Práve prítomnosť iných vlastných čísel s absolútnou hodnotou 1 naruší konvergenciu. Potom vo vyjadrení (3.2) získaného zo spektrálneho rozkladu bude príslušná matica \mathbf{G}_i meniť hodnoty na základe mocniny λ_k . Absolútna hodnota prvkov v príslušnej matici bude rovnaká. V prípade vlastného čísla -1 sa budú hodnoty meniť z kladných na záporné. Preto je prítomnosť vlastných čísel s absolútnou hodnotou 1 iných ako vlastné číslo 1 nežiaduca. V nasledujúcej časti ukážeme, že matica \mathbf{G} nebude mať iné vlastné čísla s absolútnou hodnotou 1 ako vlastné číslo 1, ak bude $\alpha \in (0, 1)$.

3.3 Vlastné čísla matice \mathbf{G}

Doteraz sme pracovali s maticou \mathbf{S} . Vyskytli sa problémy s tým, že iteračný proces nemusí skonvergovať. Tieto problémy s vlastným číslom -1 vyrieši prechod ku matici \mathbf{G} . V nasledujúcej časti ukážeme, že vlastné čísla matice \mathbf{G} vieme vypočítavať na základe vlastných čísel matice \mathbf{S} .

Teraz si uvedieme niekoľko tvrdení, ktoré nám pomôžu pri analýze vlastných hodnôt matice \mathbf{G} . V nasledujúcom tvrdení ukážeme, charakteristický polynóm dvoch podobných matic je rovnaký.

Tvrdenie 3.3.1. *Nech \mathbf{A} a \mathbf{B} sú štvorcové matice typu $n \times n$ nad poľom \mathbb{F} . Ak \mathbf{A} a \mathbf{B} sú podobné, tak $ch_{\mathbf{A}}(x) = ch_{\mathbf{B}}(x)$.*

Dôkaz. Keďže \mathbf{A} a \mathbf{B} sú podobné, tak existuje regulárna matica \mathbf{P} taká že $\mathbf{A} = \mathbf{PBP}^{-1}$. Na základe definície 2.4.2 môžeme písať

$$\begin{aligned} ch_{\mathbf{A}}(x) &= |\mathbf{A} - x\mathbf{I}| = |\mathbf{PBP}^{-1} - x\mathbf{I}| = |\mathbf{PBP}^{-1} - x\mathbf{PIP}^{-1}| = \\ &= |\mathbf{P}(\mathbf{B} - x\mathbf{I})\mathbf{P}^{-1}| = |\mathbf{P}||\mathbf{B} - x\mathbf{I}||\mathbf{P}^{-1}| = |\mathbf{PP}^{-1}||\mathbf{B} - x\mathbf{I}| = \\ &= |\mathbf{I}||\mathbf{B} - x\mathbf{I}| = |\mathbf{B} - x\mathbf{I}| = ch_{\mathbf{B}}(x) \end{aligned}$$

□

Keďže korene charakteristického polynómu sú vlastné čísla matice, tak z predchádzajúceho tvrdenia plynie, že podobné matice majú rovnaké vlastné čísla.

Nasledujúce tvrdenie uvediem bez dôkazu. Jeho dôkaz je náročnejší a vyžaduje hlbšie poznatky z lineárnej algebry. Dôkaz možno nájsť napr. v [M, kapitola 7.7].

Tvrdenie 3.3.2. *Každá štvorcová matica \mathbf{G} je podobná s hornou trojuholníkovou maticou, ktorá má na uhlopriečke vlastné hodnoty matice \mathbf{G} .*

Toto tvrdenie presnejšie sformulujeme v časti 3.4, kde sa budeme zaoberať konvergenciou pre všeobecné matice.

Tvrdenie 3.3.3. *Nech \mathbf{S} je riadkovo stochastická matica a jej vlastné čísla sú $1, \lambda_2, \dots, \lambda_n$. Potom matica*

$$\mathbf{G} = \alpha \mathbf{S} + (1 - \alpha) \frac{1}{n} \mathbf{e}^T \mathbf{e}$$

má vlastné čísla $1, \alpha \lambda_2, \dots, \alpha \lambda_n$, kde $\mathbf{e} = (1, 1, \dots, 1)$.

Dôkaz. Matica \mathbf{S} je riadkovo stochastická, platí preto

$$\mathbf{S} \mathbf{e}^T = \mathbf{e}^T.$$

Nech \mathbf{Q} je ľubovoľná regulárna matica, ktorej prvý stĺpec je \mathbf{e} .

$$\mathbf{S} = \begin{pmatrix} \mathbf{e}^T & \mathbf{X} \end{pmatrix}$$

Ak označíme prvý riadok inverznej matice \mathbf{Q}^{-1} ako \mathbf{y} , tak platí:

$$\mathbf{Q}^{-1} \mathbf{Q} = \begin{pmatrix} \mathbf{y} \\ \mathbf{Y} \end{pmatrix} \begin{pmatrix} \mathbf{e}^T & \mathbf{X} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \mathbf{e}^T & \mathbf{y} \mathbf{X} \\ \mathbf{Y} \mathbf{e}^T & \mathbf{Y} \mathbf{X} \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0}^T & \mathbf{I} \end{pmatrix}. \quad (3.3)$$

Ak vynásobíme maticu \mathbf{S} zľava maticou \mathbf{Q} a sprava maticou \mathbf{Q}^{-1} dostávame:

$$\mathbf{Q}^{-1} \mathbf{S} \mathbf{Q} = \begin{pmatrix} \mathbf{y} \\ \mathbf{Y} \end{pmatrix} \mathbf{S} \begin{pmatrix} \mathbf{e}^T & \mathbf{X} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \mathbf{S} \\ \mathbf{Y} \mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{e}^T & \mathbf{X} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \mathbf{S} \mathbf{e}^T & \mathbf{y} \mathbf{S} \mathbf{X} \\ \mathbf{Y} \mathbf{S} \mathbf{e}^T & \mathbf{Y} \mathbf{S} \mathbf{X} \end{pmatrix}$$

Na základe (3.3) dostávame

$$\begin{aligned} \mathbf{y} \mathbf{S} \mathbf{e}^T &= \mathbf{y} \mathbf{e}^T = 1 \\ \mathbf{Y} \mathbf{S} \mathbf{e}^T &= \mathbf{Y} \mathbf{e}^T = \mathbf{0}^T \end{aligned}$$

Dosadením do vzťahu pre $\mathbf{Q}^{-1} \mathbf{S} \mathbf{Q}$ dostaneme

$$\mathbf{Q}^{-1} \mathbf{S} \mathbf{Q} = \begin{pmatrix} 1 & \mathbf{y} \mathbf{S} \mathbf{X} \\ \mathbf{0}^T & \mathbf{Y} \mathbf{S} \mathbf{X} \end{pmatrix} = \mathbf{B}$$

Z poslednej rovnice dostávame, že matica \mathbf{S} je podobná s maticou \mathbf{B} . Podľa tvrdenia 3.3.1 je charakteristický polynóm matice \mathbf{S} rovnaký ako charakteristický polynóm matice \mathbf{B} . Pre char. polynóm matice \mathbf{B} platí

$$ch_{\mathbf{B}}(x) = |\mathbf{B} - x \mathbf{I}| = \begin{vmatrix} 1 - x & \mathbf{y} \mathbf{S} \mathbf{X} \\ \mathbf{0}^T & \mathbf{Y} \mathbf{S} \mathbf{X} - \mathbf{I}' \end{vmatrix},$$

kde \mathbf{I}' je diagonálna matica veľkosti $(n-1) \times (n-1)$, ktorá má na diagonále samé jednotky. Keď rozvinieme determinant matice $\mathbf{B} - x\mathbf{I}$ podľa prvého stĺpca, dostávame

$$ch_{\mathbf{B}}(x) = |\mathbf{B} - x\mathbf{I}| = (1-x)|\mathbf{YSX} - \mathbf{I}'|$$

Keďže $ch_{\mathbf{S}}(x) = ch_{\mathbf{B}}(x)$ a $ch_{\mathbf{S}}(x) = (1-x)(\lambda_2 - x) \cdots (\lambda_n - x)$, tak korene charakteristického polynómu pre maticu \mathbf{YSX} sú $\lambda_2, \dots, \lambda_n$. Inak povedané matica \mathbf{YSX} má vlastné hodnoty $\lambda_2, \dots, \lambda_n$.

Vynásobme teraz maticu $\mathbf{e}^T \mathbf{e}$ zľava maticou \mathbf{Q}^{-1} a sprava maticou, dostaneme

$$\begin{aligned} \mathbf{Q}^{-1} \mathbf{e}^T \mathbf{e} \mathbf{Q} &= \begin{pmatrix} \mathbf{y} \\ \mathbf{Y} \end{pmatrix} \mathbf{e}^T \mathbf{e} \begin{pmatrix} \mathbf{e}^T & \mathbf{X} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \mathbf{e}^T \\ \mathbf{Y} \mathbf{e}^T \end{pmatrix} \begin{pmatrix} \mathbf{e} \mathbf{e}^T & \mathbf{e} \mathbf{X} \end{pmatrix} = \\ &= \begin{pmatrix} 1 \\ \mathbf{0}^T \end{pmatrix} \begin{pmatrix} n & \mathbf{e} \mathbf{X} \end{pmatrix} = \begin{pmatrix} n & \mathbf{e} \mathbf{X} \\ \mathbf{0}^T & \|\mathbf{0}\| \end{pmatrix} \end{aligned}$$

Skombinujme teraz vyjadrenia pre $\mathbf{Q}^{-1} \mathbf{S} \mathbf{Q}$ a pre $\mathbf{Q}^{-1} \mathbf{e}^T \mathbf{e} \mathbf{Q}$, aby sme dostali vyjadrenie pre maticu \mathbf{G} .

$$\begin{aligned} \mathbf{Q}^{-1} \mathbf{G} \mathbf{Q} &= \alpha \mathbf{Q}^{-1} \mathbf{S} \mathbf{Q} + (1-\alpha) \frac{1}{n} \mathbf{e}^T \mathbf{e} \mathbf{Q} = \\ &= \alpha \begin{pmatrix} 1 & \mathbf{y} \mathbf{S} \mathbf{X} \\ \mathbf{0}^T & \mathbf{Y} \mathbf{S} \mathbf{X} \end{pmatrix} + (1-\alpha) \frac{1}{n} \begin{pmatrix} n & \mathbf{e} \mathbf{X} \\ \mathbf{0}^T & \|\mathbf{0}\| \end{pmatrix} = \\ &= \begin{pmatrix} 1 & \alpha \mathbf{y} \mathbf{S} \mathbf{X} + (1-\alpha) \frac{1}{n} \mathbf{e} \mathbf{X} \\ \mathbf{0}^T & \alpha \mathbf{Y} \mathbf{S} \mathbf{X} \end{pmatrix} = \mathbf{C} \end{aligned}$$

Vidíme, že matica \mathbf{G} je podobná s maticou \mathbf{C} , ktorá má v prvom stĺpci číslo 1 a potom samé nuly. Stačí nám zistiť aké vlastné čísla má matica \mathbf{C} , z podobnosti bude mať rovnaké matica \mathbf{G} .

Pre charakteristický polynóm matice \mathbf{C} platí

$$ch_{\mathbf{C}}(x) = |\mathbf{C} - x\mathbf{I}| = (1-x)|\alpha \mathbf{Y} \mathbf{S} \mathbf{X} - \mathbf{I}'| = ch_{\mathbf{G}}(x)$$

Takže matica \mathbf{G} má vlastné číslo 1 a všetky vlastné čísla matice $\alpha \mathbf{Y} \mathbf{S} \mathbf{X}$. Pre maticu $\mathbf{Y} \mathbf{S} \mathbf{X}$ platí na základe tvrdenia 3.3.2, že je podobná s hornou trojuholíkovou maticou, ktorá má na diagonále čísla $\lambda_2, \dots, \lambda_n$. Potom matica $\alpha \mathbf{Y} \mathbf{S} \mathbf{X}$ je podobná s hornou trojuholíkovou, čo má na diagonále čísla $\alpha \lambda_2, \dots, \alpha \lambda_n$. Teda matica $\alpha \mathbf{Y} \mathbf{S} \mathbf{X}$ má vlastné čísla $\alpha \lambda_2, \dots, \alpha \lambda_n$.

Dostávame, že matica \mathbf{G} má vlastné čísla $1, \alpha \lambda_2, \dots, \alpha \lambda_n$. \square

Na maticu \mathbf{G} kladieme podmienku, aby bola podobná s diagonálnou maticou. Na základe vzťahu (3.2) môžeme napísať vyjadrenie pre jej k -tu mocninu

$$\mathbf{G}^k = \lambda_1^k \mathbf{G}_1 + \lambda_2^k \mathbf{G}_2 + \cdots + \lambda_n^k \mathbf{G}_n \quad \lambda_1 = 1. \quad (3.4)$$

Odteraz budeme symbolmi $\lambda_1, \dots, \lambda_n$ označovať vlastné čísla matice \mathbf{G} . Z predchádzajúceho tvrdenia plynie, že absolútna hodnota vlastných čísel matice \mathbf{G} okrem čísla 1 leží v intervale $\langle \alpha, 0 \rangle$. Preto pre všetky $2 \leq i \leq n$ platí, že $\lambda_i^k \rightarrow 0$, pre rastúce k . Iteračný proces $\mathbf{x}_0 \mathbf{G}^k$ teda skonverguje ku hľadanému ohodnoteniu.

Z vyjadrenia (3.4) plynie, že rýchlosť konvergenencie závisí od druhého najväčšieho vlastného čísla. Pre vlastné čísla matice \mathbf{G} platí $\alpha \geq |\lambda_2| \geq \dots \geq |\lambda_n|$. Parametrom α teda vieme kontrolovať rýchlosť konvergenencie iteračného procesu (vieme ohraničiť druhé najväčšie vlastné číslo). Viac o vplyve parametra α povieme v časti 3.6. Teraz nasleduje dôkaz konvergenencie pre všeobecnú maticu \mathbf{G} .

3.4 Konvergenca pre všeobecnú maticu

V tejto časti dokážeme konvergenziu pre všeobecnú maticu \mathbf{G} . Budeme postupovať podobne ako v dôkaze pre diagonalizovateľné matice. Nájdeme vyjadrenie pre \mathbf{G}^k . V predošlej časti sme našli jednoduchý tvar pre k -tu mocninu diagonalizovateľnej matice na základe tvrdenia o spektrálnom rozklade. Mierne zložitejší rozklad existuje aj pre všeobecnú maticu.

Teraz presnejšie sformulujeme tvrdenie 3.3.2, na základe ktorého budeme vedieť nájsť vyjadrenie pre k -tu mocninu matice \mathbf{G} .

3.4.1 Jordanov normálny tvar

Definícia 3.4.1. *Jordanov blok* veľkosti k prislúchajúci číslu λ je matica veľkosti $k \times k$ tvaru

$$J_*(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda & 1 & 0 \\ 0 & \dots & 0 & 0 & \lambda & 1 \\ 0 & \dots & 0 & 0 & 0 & \lambda \end{pmatrix}$$

Tvrdenie 3.4.1. *Pre každú štvorcovú maticu \mathbf{A} veľkosti $n \times n$ nad \mathbb{C} s vlastnými číslami $\lambda_1, \dots, \lambda_s$ existuje taká regulárna matica \mathbf{P} , že*

$$\mathbf{PAP}^{-1} = \mathbf{J} = \begin{pmatrix} \mathbf{J}(\lambda_1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(\lambda_2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}(\lambda_s) \end{pmatrix}. \quad (3.5)$$

- matica \mathbf{J} sa nazýva *Jordanov normálny tvar matice \mathbf{J}* .

- \mathbf{J} má jeden Jordanov segment $\mathbf{J}(\lambda_j)$ pre každé vlastné číslo λ_j , $j = 1, \dots, s$.
- Každý segment $\mathbf{J}(\lambda_j)$ sa skladá z t_j Jordanových blokov $\mathbf{J}_*(\lambda_j)$, kde t_j je počet lineárne nezávislých vektorov prislúchajúcich ku vlastnému číslu λ_j .

$$\mathbf{J}(\lambda_j) = \begin{pmatrix} \mathbf{J}_1(\lambda_j) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2(\lambda_j) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}_{t_j}(\lambda_j) \end{pmatrix}$$

- Platí, že matica \mathbf{J} je jednoznačne určená až na poradie Jordanových blokov na diagonále.
- Ďalej platí, že dve matice \mathbf{A} a \mathbf{B} sú podobné práve vtedy, keď majú rovnaký Jordanov tvar (až na poradie Jordanových blokov).

Dôkaz. Opäť pripomínáme, že toto tvrdenie je bez dôkazu. Jeden možný dôkaz, ktorý využíva teóriu modulov možno nájsť v [Z, kapitola 20] \square

Predpokladajme, že vlastné čísla matice \mathbf{G} sú usporiadané podľa veľkosti t.j. $\lambda_1 = 1 > \alpha \geq |\lambda_2| \geq \dots \geq |\lambda_s| \geq 0$. Môžeme tiež predpokladať, že Jordanove segmenty v matici \mathbf{J} pre vlastné čísla $\lambda_1, \dots, \lambda_s$ sú na diagonále v tomto poradí. Preto

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}(1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(\lambda_2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}(\lambda_s) \end{pmatrix}.$$

Rovnako ako v prípade diagonalizovateľnej matice vieme získať mocninu matice \mathbf{G} prechodom k mocnine inej matice. Napíšme teraz vyjadrenie pre k -tu mocninu matice \mathbf{G}

$$\mathbf{G}^k = \mathbf{P}^{-1} \mathbf{J}^k \mathbf{P},$$

$$\mathbf{J}^k = \begin{pmatrix} \mathbf{J}(1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(\lambda_2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}(\lambda_s) \end{pmatrix}^k$$

Tu si treba uvedomiť, že matica \mathbf{J} je blokovo diagonálna a každý jej Jordanov segment je štvorcová matica. Zároveň každý Jordanov segment je tiež blokovo diagonálna matica s Jordanovými blokmi na diagonále. Preto keď umocňujeme maticu \mathbf{J} , tak okrem

štvorcových blokov na diagonále budú všade nuly. Schematicky môžeme k -tu mocninu môžeme zapísať nasledovne:

$$\mathbf{J}^k = \begin{pmatrix} \mathbf{J}^k(1) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}^k(\lambda_2) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}^k(\lambda_s) \end{pmatrix}, \text{ pre každý segment}$$

$$\mathbf{J}^k(\lambda_j) = \begin{pmatrix} \mathbf{J}_1^k(\lambda_j) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2^k(\lambda_j) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}_{t_j}^k(\lambda_j) \end{pmatrix}$$

Mocnina Jordanovho tvaru je teda blokovo diagonálna matica s mocninami jednotlivých blokov na diagonále. Zapišme teraz pomocou nej k -tu mocninu matice \mathbf{G} . Označme $\mathbf{Q}_l(\lambda_j)$, $l = 1, \dots, t_j$ skupinu stĺpcových vektorov matice \mathbf{P}^{-1} , ktorá prislúcha Jordanovmu bloku $\mathbf{J}_l(\lambda_j)$. Podobne $\mathbf{P}_l(\lambda_j)$, $l = 1, \dots, t_j$ skupinu riadkových matice \mathbf{P} , ktorá prislúcha Jordanovmu bloku $\mathbf{J}_l(\lambda_j)$. Dostávame

$$\mathbf{P}^{-1} = \left(\mathbf{Q}_1(\lambda_1) \mid \dots \mid \mathbf{Q}_l(\lambda_j) \mid \dots \mid \mathbf{Q}_{t_s}(\lambda_s) \right) \quad \mathbf{P} = \begin{pmatrix} \mathbf{P}_1(\lambda_1) \\ \vdots \\ \mathbf{P}_l(\lambda_j) \\ \vdots \\ \mathbf{P}_{t_s}(\lambda_s) \end{pmatrix}$$

$$\begin{aligned} \mathbf{G}^k &= \mathbf{P}^{-1} \mathbf{J}^k \mathbf{P} = \\ &= \left(\mathbf{Q}_1(\lambda_1) \mid \dots \mid \mathbf{Q}_{t_s}(\lambda_s) \right) \begin{pmatrix} \mathbf{J}_1^k(\lambda_1) & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{J}_{t_j}^k(\lambda_s) \end{pmatrix} \begin{pmatrix} \mathbf{P}_1(\lambda_1) \\ \vdots \\ \mathbf{P}_{t_s}(\lambda_s) \end{pmatrix} \end{aligned}$$

3.4.2 Rozklad matice \mathbf{G}

Maticu \mathbf{G}^k vieme rozložiť na súčet matíc podobne ako v 3.2. Matice v rozklade budú mať tvar $\mathbf{Q}_l(\lambda_j)' \mathbf{J}_l^k(\lambda_j)' \mathbf{P}_l(\lambda_j)'$. Na základe týchto matíc budeme vedieť analyzovať konvergenciu pre mocninu matice \mathbf{G} .

$$\begin{aligned} \mathbf{G}^k &= \mathbf{Q}_1(\lambda_1)' \mathbf{J}_1^k(\lambda_1)' \mathbf{P}_1(\lambda_1)' + \dots + \mathbf{Q}_l(\lambda_j)' \mathbf{J}_l^k(\lambda_j)' \mathbf{P}_l(\lambda_j)' + \dots + \\ &\quad + \dots + \mathbf{Q}_{t_s}(\lambda_s)' \mathbf{J}_{t_s}^k(\lambda_s)' \mathbf{P}_{t_s}(\lambda_s)' \end{aligned} \quad (3.6)$$

Tieto matice vzniknú z matíc \mathbf{P}^{-1} , \mathbf{J}^k , \mathbf{P} tak, že z matice \mathbf{J}^k vyberieme jeden blok $\mathbf{J}_l^k(\lambda_j)$ a v maticiach \mathbf{P}^{-1} , resp. \mathbf{P} necháme len prislúchajúce stĺpce, resp. riadky. Pre

jeden blok dostávame

$$\mathbf{Q}_l(\lambda_j)' \mathbf{J}_1^k(\lambda_j)' \mathbf{P}_l(\lambda_j)' = \left(\mathbf{0} \mid \mathbf{Q}_l(\lambda_j) \mid \mathbf{0} \right) \begin{pmatrix} \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \mathbf{J}_1^k(\lambda_j) & \vdots \\ \mathbf{0} & \dots & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{P}_l(\lambda_j) \\ \mathbf{0} \end{pmatrix}$$

Dostali sme podobný rozklad ako (3.2), s tým rozdielom, že v (3.6) môžu mať bloky veľkosť väčšiu ako 1. Predtým sme pracovali s diagonálnou maticou. Diagonálna matica je špeciálny prípad Jordanovho normálneho tvaru, v ktorom všetky Jordanove bloky majú veľkosť 1.

Teraz popíšeme mocninu jedného bloku, ktorý sa vyskytuje v rozklade matice \mathbf{G} . Predpokladajme, že blok má veľkosť $m \times m$. Matematickou indukciou možno dokázať, že

$$\mathbf{J}_*^k(\lambda_i) = \begin{pmatrix} \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \binom{k}{2}\lambda_i^{k-2} & \dots & \binom{k}{m-2}\lambda_i^{k-m+2} & \binom{k}{m-1}\lambda_i^{k-m+1} \\ 0 & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} & \dots & \binom{k}{m-3}\lambda_i^{k-m+3} & \binom{k}{m-2}\lambda_i^{k-m+2} \\ 0 & 0 & \lambda_i^k & \dots & \binom{k}{m-4}\lambda_i^{k-m+4} & \binom{k}{m-3}\lambda_i^{k-m+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_i^k & \binom{k}{1}\lambda_i^{k-1} \\ 0 & 0 & 0 & \dots & 0 & \lambda_i^k \end{pmatrix}$$

Mocnina Jordanovho bloku je horná trojuholníková matica, v ktorej sa vyskytujú výrazy tvaru $\binom{k}{j}\lambda_i^{k-j}$, pričom $0 \leq j \leq m-1$. Ak $k < j$, tak pre binomický koeficient $\binom{k}{j}$ platí, že je rovný nule.

Binomický koeficient $\binom{k}{j}$ je polynóm stupňa najviac $m-1$ s premennou k . Ďalej λ_i^{k-j} je exponenciálna funkcia, ktorá pre vlastné číslo 1 nadobúda hodnotu 1. Pre ostatné vlastné čísla, ktorých absolútna hodnota je menšia ako α je táto funkcia klesajúca. Keď uvažujeme iteračný proces, tak $k \rightarrow \infty$. Exponenciálna funkcia klesá asymptoticky rýchlejšie ako rastie polynomiálna, preto

$$\binom{k}{j}\lambda_i^{k-j} \rightarrow 0.$$

3.4.3 Konvergencia členov rozkladu

Pre všetky vlastné čísla λ_j , kde $j = 2, \dots, s$ platí $|\lambda_j| \in \langle 0, \alpha \rangle$. Preto pre všetky Jordanove bloky prislúchajúce vlastným číslam rôznym od 1 platí, že konvergujú k nulovej matici. Pre každý prvok J_{ij} matice $\mathbf{J}_1^k(\lambda_j)'$, ktorá má rozmer $n \times n$ preto platí, že je nulový, alebo

$$\lim_{k \rightarrow \infty} J_{ij} = 0$$

Teraz ukážeme, že každá matica $\mathbf{Q}_l(\lambda_j)' \mathbf{J}_1^k(\lambda_j)' \mathbf{P}_l(\lambda_j)'$, $j \geq 2$ zo súčiny (3.6) konverguje ku nulovej matici. Označme q_{ij} prvky matice $\mathbf{Q}_l(\lambda_j)'$ a p_{ij} prvky matice $\mathbf{P}_l(\lambda_j)'$.

Dostávame

$$\mathbf{Q}_l(\lambda_j)' \mathbf{J}_1^k(\lambda_j)' \mathbf{P}_l(\lambda_j)' = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{pmatrix} \begin{pmatrix} J_{11} & J_{12} & \cdots & J_{1n} \\ J_{21} & J_{22} & \cdots & J_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ J_{n1} & J_{n2} & \cdots & J_{nn} \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

Pre prvok z_{rs} v r -tom riadku a s -tom stĺpci $r, s = 1, \dots, n$ matice

$\mathbf{Q}_l(\lambda_j)' \mathbf{J}_1^k(\lambda_j)' \mathbf{P}_l(\lambda_j)'$ platí

$$z_{rs} = \sum_{i=1}^n \sum_{j=1}^n q_{ri} J_{ij} p_{js}.$$

Vo vyjadrení pre ľubovoľný prvok z_{rs} sčítame výrazy tvaru $M \cdot J_{ij}$, kde M je reálne číslo. Vyššie sme ukázali, že pre každé z čísel J_{ij} je nula alebo má limitu rovnú nule.

Preto pre limitu platí

$$\lim_{k \rightarrow \infty} z_{rs} = \lim_{k \rightarrow \infty} \sum_{i=1}^n \sum_{j=1}^n q_{ri} J_{ij} p_{js} = 0.$$

Každý prvok matice $\mathbf{Q}_l(\lambda_j)' \mathbf{J}_1^k(\lambda_j)' \mathbf{P}_l(\lambda_j)'$ konverguje k nule, preto aj celá matica konverguje k nule.

Pre mocniny blokov prislúchajúcich ku vlastnému číslu 1 nemáme zaručenú konvergenciu. Dokonca sa môže stať, že nejaká hodnota v bloku môže neobmedzene rásť. Keby mal blok $\mathbf{J}_1(1)$ veľkosť väčšiu ako 1, tak by sa v mocnине tohoto bloku vyskytovalo aspoň jedno kombinačné číslo $\binom{k}{j}$. Potom číslo $\binom{k}{j} 1^{k-j}$ bude rásť s rastúcim k . Musíme preto overiť akú štruktúru majú Jordanove bloky prislúchajúci vlastnému číslu 1. Jordanov segment $\mathbf{J}(1)$ sa skladá iba z jedného bloku, pretože vlastnému číslu 1 prislúcha vektorový priestor vlastných vektorov s dimenziou 1. Platí preto $\mathbf{J}(1) = \mathbf{J}_1(1)$. V nasledujúcom tvrdení ukážeme, že vlastnému číslu 1 prislúcha blok veľkosti 1.

Tvrdenie 3.4.2. *Nech \mathbf{G} je riadkovo stochastická matica, potom v jej Jordanovom normálnom tvare prislúcha vlastnému číslu 1 jediný blok, ktorý má veľkosť 1.*

Dôkaz. Vyššie sme ukázali, že vlastnému číslu 1 prislúcha jediný blok. Predpokladajme teraz sporom, že Jordanov blok prislúchajúci vlastnému číslu jedna má veľkosť $m \geq 2$. Nech je tento blok prvý na diagonále. Na základe tvrdenia 3.4.1 existuje regulárna matica \mathbf{P} taká, že

$$\mathbf{P}\mathbf{G} = \mathbf{J}\mathbf{P} = \begin{pmatrix} \mathbf{J}_1(\lambda_1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}(\lambda_2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{J}(\lambda_s) \end{pmatrix}.$$

Ak označíme riadky matice \mathbf{P} ako $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$, tak pre prvých m vektorov platí

$$\begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{pmatrix} \mathbf{G} = \begin{pmatrix} 1 & 1 & \dots & 0 & 0 \\ 0 & 1 & \ddots & 0 & 0 \\ & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & & 1 & 1 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_m \end{pmatrix}$$

Pre vektory $\mathbf{a}_{m-1}, \mathbf{a}_m$ platí

$$\begin{aligned} \mathbf{a}_{m-1} \mathbf{G} &= \mathbf{a}_m + \mathbf{a}_{m-1} \\ \mathbf{a}_m \mathbf{G} &= \mathbf{a}_m \end{aligned}$$

Potom $\mathbf{a}_{m-1}(\mathbf{G} - \mathbf{I}) = \mathbf{a}_m$. Poďme teraz zistiť aký je súčet súradníc vektora $\mathbf{a}_{m-1}(\mathbf{G} - \mathbf{I})$. Na základe tvrdenia 2.4.4 násobenie riadkovo stochastickou maticou nemení súčet súradníc. Ak označíme S súčet súradníc vektora \mathbf{a}_{m-1} , tak vektor

$$\mathbf{a}_{m-1}(\mathbf{G} - \mathbf{I}) = \mathbf{a}_{m-1} \mathbf{G} - \mathbf{a}_{m-1}$$

má súčet súradníc rovný $S - S = 0$. Teda vektor \mathbf{a}_m má súčet súradníc rovný nule. Zároveň platí, že $\mathbf{a}_m \in V_1(\mathbf{G})$ je vlastný vektor prislúchajúci číslu 1. Keďže vlastný vektor je podľa definície nenulový, dostávame, že \mathbf{a}_m má súradnice so zmiešanými znamienkami. Podľa tvrdenia 2.4.2 však dostávame spor, pretože pre každý vlastný vektor prislúchajúci číslu 1 platí, že jeho súradnice majú rovnaké znamienka. \square

Ukázali sme, že iteračný proces s maticou \mathbf{G} skonverguje k vektoru \mathbf{x} , ktorý je zároveň vlastným vektorom matice \mathbf{G} pre vlastné číslo 1, teda je ohodnotením dôležitosti pre stránky.

3.4.4 Závislosť výsledného ohodnotenia od počiatočného vektora \mathbf{x}_0

Doteraz sme odpovedali na väčšinu otázok, čo sú uvedené na strane 7. Potrebujeme ešte ukázať, že výsledný vektor je v istom zmysle jednoznačný. Iteračný proces začíname s ohodnotením \mathbf{x}_0 , ktoré má súčet súradníc jedna.

Tvrdenie 2.4.4 hovorí, že násobenie riadkovo stochastickou maticou nemení súčet zložiek vektora. Ďalej sme ukázali, že iteračná metóda pre maticu \mathbf{G} skonverguje k riešeniu. Preto pre výsledné ohodnotenie \mathbf{x} , ktoré získame ako limitu

$$\mathbf{x} = \lim_{n \rightarrow \infty} \mathbf{x}_0 \mathbf{G}^n$$

platí, že má súčet súradníc 1. Vektor \mathbf{x} je zároveň vlastný vektor prislúchajúci vlastnému číslu 1. Dimenzia priestoru $V_1(G)$ vlastných vektorov prislúchajúcich ku číslu 1 je 1. Preto existuje jediný vektor $\mathbf{x}' \in V_1(G)$ ktorý má súčet súradníc 1. Dostávame, že $\mathbf{x} = \mathbf{x}'$.

Inak povedané iteračný proces skonverguje pre počiatočný vektor, ktorý má súčet súradníc 1 vždy k tomu istému ohodnoteniu.

3.5 Rýchlosť konvergenzie

Zatiaľ sme ukázali, že iteračný proces skonverguje, no nikde sme sa nevenovali rýchlosti konvergenzie. V tejto časti ukážeme, že rýchlosť konvergenzie asymptoticky závisí od veľkosti druhého najväčšieho vlastného čísla λ_2 . V [QSS, kapitola 5.3.1] je zdôvodnenie pre diagonalizovateľné matice. My toto tvrdenie rozšírime a ukážeme jeho platnosť aj pre všeobecné matice.

Ideme hľadať vyjadrenie k -teho kroku iteračného procesu $\mathbf{x}_0 \mathbf{G}^k$ so štartovacím vektorom \mathbf{x}_0 . Na základe tvrdenia 3.4.1 existuje regulárna matica \mathbf{P} taká, že

$$\begin{aligned}\mathbf{P}\mathbf{G} &= \mathbf{J}\mathbf{P} \\ \mathbf{G} &= \mathbf{P}^{-1}\mathbf{J}\mathbf{P},\end{aligned}$$

kde \mathbf{J} je Jordanov normálny tvar matice \mathbf{G} . Riadky matice \mathbf{P} označme $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$. Tieto vektory sú lineárne nezávislé a tvoria bázu priestoru \mathbb{C}^n , nazývame ich aj zo-všeobecnené vlastné vektory. Vyjadrime teraz štartovací vektor \mathbf{x}_0 v báze $\mathbf{p}_1, \dots, \mathbf{p}_n$. Dostávame

$$\mathbf{x}_0 = \alpha_1 \mathbf{p}_1 + \dots + \alpha_n \mathbf{p}_n, \quad \alpha_i \in \mathbb{C}, \quad i = 1, \dots, n,$$

čo môžeme zapísať v maticovom tvare

$$\mathbf{x}_0 = \mathbf{a}\mathbf{P}, \quad \mathbf{a} = (\alpha_1, \dots, \alpha_n).$$

Potom platí

$$\mathbf{x}_k = \mathbf{x}_0 \mathbf{G}^k = \mathbf{a}\mathbf{P}\mathbf{P}^{-1}\mathbf{J}^k\mathbf{P} = \mathbf{a}\mathbf{J}^k\mathbf{P}. \quad (3.7)$$

Jordanov blok prislúchajúci vlastnému číslu 1 je jediný a má rozmer 1×1 . Prislúchajúci vektor \mathbf{p}_1 je zároveň vlastným vektorom pre vlastné číslo 1. Na základe rovnice (3.7)

môžeme napísať vyjadrenie pre k -ty krok iteračného procesu.

$$\begin{aligned} \mathbf{x}_k &= \mathbf{a} \mathbf{J}^k \mathbf{P} = \\ &= \mathbf{a} \begin{pmatrix} 1 & & & & \\ & \mathbf{J}_1^k(\lambda_2) & & & \\ & & \ddots & & \\ & & & \mathbf{J}_1^k(\lambda_j) & \\ & & & & \ddots \\ & & & & & \mathbf{J}_{t_s}^k(\lambda_s) \end{pmatrix} \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{P}_1(\lambda_2) \\ \vdots \\ \mathbf{P}_1(\lambda_j) \\ \vdots \\ \mathbf{P}_{t_s}(\lambda_s) \end{pmatrix} \end{aligned} \quad (3.8)$$

Pripomeniem, že $\mathbf{J}_1^k(\lambda_j)$ je l -tý Jordanov blok pre vlastné číslo λ_j a $\mathbf{P}_1(\lambda_j)$ ku nemu prislúchajúcu skupinu riadkových vektorov \mathbf{p}_i matice \mathbf{P} .

Výsledkom súčiny (3.8) bude vyjadrenie vektora \mathbf{x}_k v báze $\mathbf{p}_1, \dots, \mathbf{p}_n$. Hľadáme teda koeficienty pred vektormi $\mathbf{p}_1, \dots, \mathbf{p}_n$. Pred \mathbf{p}_1 bude koeficient α_1 , kvôli tomu, že Jordanov blok prislúchajúci číslu 1 je samotná jednotka v ľavom hornom rohu matice \mathbf{J}^k .

Teraz spočítame koeficienty pred ostatnými vektormi \mathbf{p}_i . Pozrime sa na jeden Jordanov blok $\mathbf{J}_*^k(\lambda_i)$. Nech má tento blok rozmery $m \times m$. Označme $\mathbf{p}_1^*, \dots, \mathbf{p}_m^*$ príslušné riadkové vektory matice \mathbf{P} . V súčine $\mathbf{a} \mathbf{J}^k \mathbf{P}$ tomuto bloku zodpovedá výraz

$$(\alpha_1^*, \dots, \alpha_m^*) \mathbf{J}_*^k(\lambda_i) (\mathbf{p}_1^*, \dots, \mathbf{p}_m^*)^T, \quad (3.9)$$

kde $\alpha_1^*, \dots, \alpha_m^*$ sú súradnice vektora \mathbf{a} prislúchajúce bloku $\mathbf{J}_*^k(\lambda_i)$ a

Označme $p_1^*, \dots, p_m^*, p_i^* \in \mathbb{C}$ koeficienty pred vektormi $\mathbf{p}_1^*, \dots, \mathbf{p}_m^*$ v súčine (3.9). Potom môžeme zapísať rovnicu

$$\begin{aligned} p_1^* \mathbf{p}_1^* + \dots + p_m^* \mathbf{p}_m^* &= (\alpha_1^*, \dots, \alpha_m^*) \mathbf{J}_*^k(\lambda_i) (\mathbf{p}_1^*, \dots, \mathbf{p}_m^*)^T \\ \mathbf{J}_*^k &= \begin{pmatrix} \lambda_i^k & \binom{k}{1} \lambda_i^{k-1} & & \binom{k}{m-2} \lambda_i^{k-m+2} & \binom{k}{m-1} \lambda_i^{k-m+1} \\ 0 & \lambda_i^k & \dots & \binom{k}{m-3} \lambda_i^{k-m+3} & \binom{k}{m-2} \lambda_i^{k-m+2} \\ & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & \lambda_i^k & \binom{k}{1} \lambda_i^{k-1} \\ 0 & 0 & \dots & 0 & \lambda_i^k \end{pmatrix} \end{aligned}$$

Pre koeficienty p_1^*, \dots, p_m^* dostávame vzťahy

$$\begin{aligned} p_1^* &= \alpha_1^* \lambda_i^k \\ p_2^* &= \alpha_1^* \binom{k}{1} \lambda_i^{k-1} + \alpha_2^* \lambda_i^k \\ p_3^* &= \alpha_1^* \binom{k}{2} \lambda_i^{k-2} + \alpha_2^* \binom{k}{1} \lambda_i^{k-1} + \alpha_3^* \lambda_i^k \\ &\vdots \\ p_m^* &= \sum_{j=1}^m \alpha_j^* \binom{k}{m-j} \lambda_i^{k-m+j} \end{aligned}$$

Vyjadrili sme súradnice pred vektormi \mathbf{p}_i prislúchajúcim jednému bloku. Môžeme preto zapísať vyjadrenie vektora \mathbf{x}_k v báze $\mathbf{p}_1, \dots, \mathbf{p}_n$. Každú zo súradníc p_i , kde $i = 2, \dots, n$ vieme napísať v tvare $\lambda_i^k R_i(k)$. Polynóm $R_i(k)$ s premennou k je stupňa najviac n (sú v ňom obsiahnuté výrazy s kombinačnými číslami a príslušné koeficienty α_j spolu s mocninou vlastného čísla na diagonále prislúchajemu danému Jordanovmu bloku). Pričom λ_i^k je k -ta mocnina vlastného čísla, ktoré sa nachádza v matici \mathbf{J}^k na diagonále v i -tom riadku. Vektor ohodnotenia v k -tom kroku procesu preto bude vyzeráť nasledovne

$$\begin{aligned} \mathbf{x}_k = \mathbf{aJ}^n \mathbf{P} &= \alpha_1 \mathbf{p}_1 + R_2(k) \lambda_{t_2+1}^k \mathbf{p}_2 + \dots + R_{t_2+1}(k) \lambda_2^k \mathbf{p}_{t_2+1} + \dots + \\ &+ R_n(k) \lambda_s^k \mathbf{p}_n. \end{aligned} \quad (3.10)$$

Vo vyjadrení vektora \mathbf{x}_k v báze $\mathbf{p}_1, \dots, \mathbf{p}_n$ je pred vektorom \mathbf{p}_1 koeficient α_1 . Koeficienty pred ostatnými vektormi $\mathbf{p}_2, \dots, \mathbf{p}_n$ obsahujú mocniny vlastných čísel $\lambda_2, \dots, \lambda_s$. Pre tieto čísla platí, $1 > \alpha \geq |\lambda_2| \geq \dots \geq |\lambda_s|$. Preto ak $k \rightarrow \infty$, tak potom bude platiť $\lambda_i \rightarrow 0$, kde $i = 2 \dots s$. Keďže exponenciálna funkcia klesá asymptoticky rýchlejšie ako rastie polynomiálna, tak potom platí pre každý z koeficientov

$$\lim_{k \rightarrow \infty} R_l(k) \lambda_i^k = 0.$$

Preto existuje limita

$$\lim_{k \rightarrow \infty} \mathbf{x}_k = \alpha_1 \mathbf{p}_1.$$

Vektor \mathbf{p}_1 je zároveň vlastným vektorom k vlastnému číslu 1. Táto limita presne zodpovedá tomu že iteračný proces konverguje ku vlastnému vektoru pre vlastné číslo 1.

Z koeficientov pre vektormi $\mathbf{p}_2, \dots, \mathbf{p}_n$ klesajú najpomalšie tie, ktoré obsahujú vlastné číslo s najväčšou možnou absolútnou hodnotou $|\lambda_2|$. Preto je rýchlosť konvergenie závislá od veľkosti druhého najväčšieho vlastného čísla.

Ukázali, že iteračný proces konverguje ku vlastnému vektoru matice \mathbf{G} , ktorý považujeme za výsledné ohodnotenie dôležitosti jednotlivých stránok. Teraz budeme hovoriť o parametre α , ktorým vieme okrem iného meniť veľkosť vlastných čísel matice \mathbf{G} .

3.6 Parameter α

Pôvodnú maticu liniek \mathbf{S} sme upravili na maticu \mathbf{G} , ktorá má oproti \mathbf{S} nové vlastnosti. Vo vyjadrení matice \mathbf{G}

$$\mathbf{G} = \alpha\mathbf{S} + (1 - \alpha)\frac{1}{n}\mathbf{e}^T\mathbf{e}$$

sa vyskytuje parameter $\alpha \in \langle 0, 1 \rangle$. Týmto parametrom vieme kontrolovať percento času, ktorý strávi surfer sledovaním hyperlinkovej štruktúry webu oproti teleportovaní na ľubovoľnú stránku. Predpokladajme, že $\alpha = 0.7$. Potom sa surfer v 70% prípadoch rozhodne sledovať štruktúru webu a vo zvyšných 30% prípadoch sa teleportuje na ľubovoľnú stránku.

Pre diagonalizovateľnú resp. všeobecnú maticu \mathbf{G} sme odvodili vzťah (3.1) resp. (3.6). V oboch sa vyskytujú vlastné hodnoty. Parametrom α v predpise matice \mathbf{G} preto vieme regulovať rýchlosť konvergenzie. Pre vlastné čísla matice \mathbf{G} platí tvrdenie 3.3.3

Zakladatelia Googlu Brin a Page vo svojich článkoch navrhli veľkosť nastaviť $\alpha = 0.85$ [BPMW]. Môžeme sa opýtať, prečo nie 0.9, alebo 0.6. Čím viac sa α približuje ku číslu 1, tým viac sa matica \mathbf{G} podobá na maticu \mathbf{S} . Väčšiu váhu prikladáme štruktúre internetu. Na druhej strane sa spomaľuje rýchlosť konvergenzie iteračného procesu.

Naopak ak bude parameter α blízky nule, tak bude veľmi málo záležať na štruktúre webu (vo väčšine prípadov si surfer vyberie náhodnú stránku).

Ukazuje sa, že zvolenie $\alpha = 0.85$ je vhodný kompromis medzi rýchlosťou konvergenzie a relevantnosťou výsledného ohodnotenia. V časti 4.3.3 budeme analyzovať aký vplyv má zmena parametru na rýchlosť konvergenzie.

Kapitola 4

Implementácia a správanie algoritmu

V predchádzajúcich kapitolách sme overili, že iteračný proces, ktorý je základom PageRank algoritmu naozaj skonverguje ku vektoru, ktorého zložky považujeme za dôležitosť jednotlivých stránok webu. V tejto kapitole povieme niečo o problémoch, ktoré sa vyskytujú pri použití PageRank algoritmu na skutočnom webe. Zároveň vysvetlíme praktický význam dôsledkov, ktoré sme odvodili. Navrhujeme vlastný algoritmus, ktorým budeme realizovať iteračný proces. Ďalej ukážeme správanie sa PageRank algoritmu na špeciálnych sieťach. Praktickými ukázkami zároveň umožníme čitateľovi lepšie porozumieť teórii, ktorej sme sa doteraz venovali.

4.1 Správanie algoritmu na skutočnom webe

V súčasnom webe sa nachádza približne 8 000 000 000 stránok s verejným obsahom ku ktorým má Google vyhľadávač prístup. Hovoríme preto o veľmi vysokej hodnote n . Matica liniek \mathbf{A} veľkosti $n \times n$ je zostrojená na základe webu. V praxi sa však jedná o veľmi riedku maticu. Z každej stránky vychádza priemerne 10 liniek. Väčšina jej prvkov sú teda nuly.

Hľadáme vlastný vektor k vlastnému číslu 1 matice \mathbf{G} . Vlastný vektor môžeme spočítať viacerými spôsobmi. Okrem iteračného procesu s ktorým sme sa doteraz zaoberali, existuje aj klasická metóda z lineárnej algebry, kde riešime homogénnu sústavu lineárnych rovníc

$$\begin{aligned}\mathbf{x}(\mathbf{I} - \mathbf{G}) &= (0, \dots, 0) \\ \mathbf{x}(1, \dots, 1)^T &= 1.\end{aligned}$$

No pre veľké n s ktorým pracujeme je neprípustné riešiť túto sústavu priamou metódou

(napr. Gaussovou elimináciou), pretože by sme potrebovali prejsť aspoň raz každý prvok matice \mathbf{G} čo je $O(n^2)$ operácií. Priamu metódu môžeme použiť pre veľmi malé hodnoty n , no čelíme veľkej nepresnosti.

Iteračná metóda¹ o ktorej sme doteraz hovorili je jedna z najstarších a najjednoduchších metód na hľadanie vlastných čísel a vektorov. My sme ju použili na výpočet vlastného vektora ku vlastnému číslu 1. Táto metóda je najpomalšia z podobných použiteľných metód (Gauss-Seidelova, Jacobiho, GMRES, BICGSTAB, atď. viac o týchto metódach možno nájsť v [BBC⁺]). Brin a Page sa rozhodli napriek tomu použiť túto metódu. Neskôr spomenieme prečo tak urobili.

Iteračný proces je realizovaný násobením $\mathbf{x}\mathbf{A}$, kde $\mathbf{x} \in \mathbb{R}^n$. Jeden krok násobenia preto zaberie približne $O(n^2)$ operácií násobenia a sčítovania. Pre $n = 8\,000\,000\,000$ je tento odhad zložitosti neprípustný už pre jeden krok procesu. Vzhľadom na riedkosť matice \mathbf{A} môžeme tento odhad upraviť. Matica \mathbf{A} obsahuje približne $10n$ nenulových prvkov. Na uloženie jej prvkov môžeme preto použiť inú dátovú štruktúru ako pole, napríklad spájaný zoznam susedov. O ďalších možnostiach uloženia hodnôt, resp. problémoch s tým spojených sa možno dočítať v [LM, kapitola 8.1]. Násobenie vektora \mathbf{x} maticou \mathbf{A} potom zaberie $O(n)$ operácií.

Matica \mathbf{A} vo všeobecnosti nemá vlastnosti potrebné ku konvergencii. Preto sme najprv nahradili nenulové riadky vektorom $(1/n, 1/n, \dots, 1/n)$ a získali sme maticu \mathbf{S} . Následne maticu \mathbf{G} . Úpravou na maticu \mathbf{G} stratíme pôvodnú riedkosť matice \mathbf{A} . Matica \mathbf{G} totižto neobsahuje nulový prvok.² Krok iteračného procesu $\mathbf{x}_k = \mathbf{x}_{k-1}\mathbf{G}$ preto nevieme priamo realizovať. Maticu \mathbf{G} však môžeme rozpísať na tvar

$$\begin{aligned}\mathbf{G} &= \alpha\mathbf{S} + (1 - \alpha)\frac{1}{n}\mathbf{E} = \\ &= \alpha(\mathbf{A} + \mathbf{a}^T(1/n\mathbf{e})) + (1 - \alpha)\frac{1}{n}\mathbf{e}\mathbf{e}^T.\end{aligned}$$

Pripomenieme, že vektor \mathbf{a} má číslo jedna na súradniciach, ktoré prislúchajú slepým vrcholom, ostatné súradnice má nulové.

Jeden krok iterácie preto vyzerá nasledovne

$$\begin{aligned}\mathbf{x}_k &= \mathbf{x}_{k-1}\mathbf{G} = \\ &= \mathbf{x}_{k-1}(\alpha(\mathbf{A} + \mathbf{a}^T(1/n\mathbf{e}))) + (1 - \alpha)\frac{1}{n}\mathbf{x}_{k-1}\mathbf{e}\mathbf{e}^T = \\ &= \alpha\mathbf{x}_{k-1}\mathbf{A} + (\alpha\mathbf{x}_{k-1}\mathbf{a}^T + 1 - \alpha)\frac{1}{n}\mathbf{e}\end{aligned}\tag{4.1}$$

Vyššie sme ukázali, že na násobenie $\mathbf{x}_{k-1}\mathbf{A}$ potrebujeme $O(n)$ operácií. Rovnako $O(n)$ operácii stačí na výpočet vektora $(\alpha\mathbf{x}_{k-1}\mathbf{a}^T + 1 - \alpha)(1/n)\mathbf{e}$.

¹V angličtine známa pod pojmom *power method*.

²Samozrejme uvažujeme, že $\alpha \in (0, 1)$.

Vidíme, že iteračný proces vieme realizovať aj pre maticu \mathbf{G} . Z tvaru (4.1) vidno, že nemáme veľké nároky na uskladnenie hodnôt. Hodnoty matice \mathbf{A} , môžeme uložiť do spájaného zoznamu, kde pre každý vrchol grafu budeme mať uložený zoznam liniek na ktorý ukazuje. Okrem neho stačí uložiť vektor \mathbf{a}^T . Toto je asi hlavný dôvod, pre ktorý použili Brin a Page túto metódu v roku 1998. Teraz spomenieme niekoľko výhod tejto metódy, oproti ostatným možným.

Iteračná metóda, ktorú používame nevyžaduje úpravy matice \mathbf{G} , ktorou násobíme. Veľa iných iteračných metód mení tvar matice počas každého kroku. Úpravy matice \mathbf{G} počas iteračného procesu sú kvôli jej veľkosti a spôsobu uloženia nemožné. Ďalší dôvod prečo je práve táto metóda najvhodnejšia je, že nemá veľké pamäťové nároky. Okrem matice \mathbf{A} a vektora \mathbf{a} potrebujeme uložiť aktuálny vektor ohodnotenia \mathbf{x}_k . Iné metódy ako GMRES, alebo BICGSTAB sú síce rýchlejšie, no vyžadujú výpočet 10 ďalších vektorov dĺžky n , pri každom kroku. Pri hodnote $n = 8\,000\,000\,000$ je každý ďalší vektor, ktorý treba uložiť veľká záťaž na pamäť a celkovú rýchlosť.

Zrejme najzávažnejší dôvod prečo sa používa práve táto metóda je malý počet iterácií, ktoré sú potrebné. Brin a Page zverejnili vo svojej práci z roku 1998 [BPMW], že stačí iba 50 – 100 iterácií, aby sme dostali výsledné skóre s dostatočnou presnosťou. Ukázali sme, že každá iterácia vyžaduje $O(n)$ operácií, pretože matica \mathbf{A} je riedka. Je veľmi ťažké nájsť metódu, ktorá prekoná $O(n)$ krokov. Vektor ohodnotenia netreba rátať s veľkou presnosťou, pretože ku získaniu celkového poradia stránky v prehliadači sa kombinuje zo skóre získaným na základe obsahu stránok. Viac o počte krokov potrebných k dosiahnutiu postačujúcej presnosti sa možno dočítať v [LM, kapitola 8.3].

4.2 Implementácia algoritmu

PageRank algoritmus vypočíta ohodnotenie dôležitosti stránok ako výsledok iteračného procesu (4.1). Podľa tohoto vyjadrenia vieme zostrojiť jednoduchý algoritmus ktorý bude realizovať iteračný proces. Samotný algoritmus som implementoval v jazyku MATLAB³, ktorý umožňuje jednoduchú prácu s vektormi a maticami.

Základ algoritmu tvorí cyklus, v ktorom postupne počítame vektory \mathbf{x}_k .

```
x_0 = (1/n)*ones(1,n);
x_old = x_0;
while (terminating_condition)
    x_new = step(x_old);
    x_old = x_new;
end
```

³MATLAB[®] je registrovaná značka spoločnosti The MathWorks, Inc.

```
return x_new;
```

Podmienkou `terminating_condition` vieme určiť, kedy má iteračný proces skončiť. Funkciou `step` realizujeme násobenie maticou \mathbf{G} . Proces môžeme ukončiť buď po niekoľkých krokoch, alebo po dosiahnutí určitej presnosti vektora \mathbf{x}_{new} (vektora \mathbf{x}_k). Presnosť môžeme merať spočítaním *jednotkovej normy* rozdielu dvoch po sebe idúcich ohodnotení $\mathbf{x}_k - \mathbf{x}_{k-1}$. Hodnotu tejto normy budeme volať *relatívna chyba*.

Definícia 4.2.1. *Jednotková norma* vektora \mathbf{v} je $\|\mathbf{v}\|_1 = \sum_{i=1}^n |v_i|$, kde v_1, \dots, v_n sú súradnice vektora \mathbf{v} .

Cyklus potom vyzerá nasledovne

```
while (terminating_condition)
    x_new = step(x_old);
    terminating_condition = norm(x_new-x_old,2) >= epsilon;
    x_old = x_new;
end
```

Ďalšia možnosť je spočítať ohodnotenie pre rádovo veľký počet krokov, ktoré bude mať veľkú presnosť. Následne počítajú sa normu rozdielu $\mathbf{x}_k - \mathbf{x}$, kde \mathbf{x} je ohodnotenie s veľkou presnosťou. Hodnotu tejto normy budeme volať *absolútna chyba*.

```
terminating_condition = norm(x_new-rank,2) >= epsilon;
```

V tomto vyjadrení je `rank` ohodnotenie dôležitosti s veľkou presnosťou.

Parametrom `epsilon` vieme nastaviť požadovanú presnosť. Funkciu `step` vieme rozpísať na základe (4.1) nasledovne

```
function [x_new]=step(x_old)
    x_new = alpha*x_old*A +
            (alpha*(x_old*a)+1-alpha)*((1/n)*ones(1,n));
end
```

Na výpočet zatiaľ používame neoptimálne násobenie maticou \mathbf{A} , preto funkciu `step` mierne zmeníme. Maticu liniek \mathbf{A} uložíme ako zoznam susedov, do poľa `list`. V poli `outlinks` máme pre každý vrchol uložený počet vychádzajúcich liniek. Zmenená funkcia `step` bude vyzeráť nasledovne

```
function [x_new]=step(x_old)
    xx_new = zeros(1,n);
    for i = 1:n
        for j = 1:outlinks(i)
```

```

        xx_new(list{i}(1,j)) = (xx_new(list{i}(1,j)) +
            alpha*x_old(i)*1/outlinks(i));
    end
end
x_new = xx_new +
    (alpha*(x_old*a)+1-alpha)*((1/n)*ones(1,n));
end

```

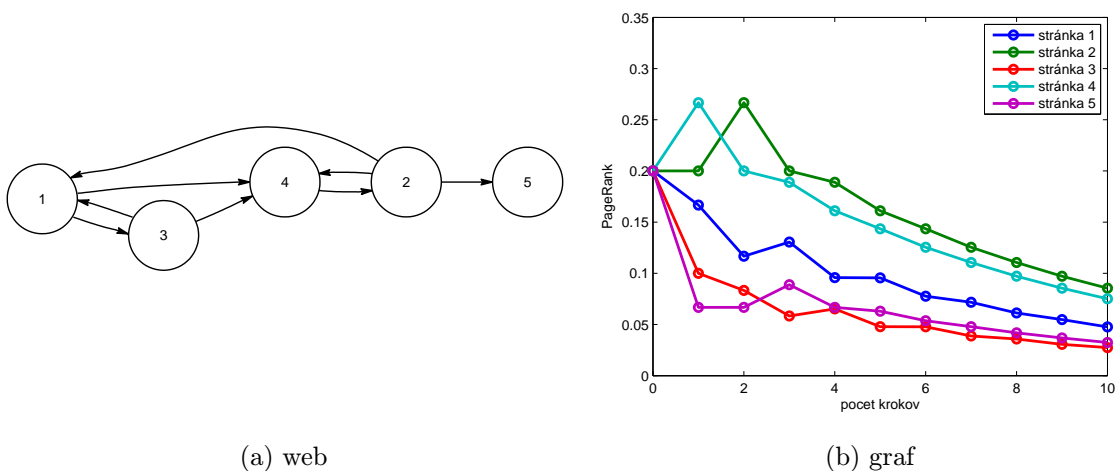
S navrhnutým algoritmom budeme teraz budeme teraz realizovať výpočet ohodnotenia pre konkrétne siete.

4.3 Použitie algoritmu

Algoritmus budeme stále spúšťať s počiatočným ohodnotením $\mathbf{x}_0 = (1/n, 1/n, \dots, 1/n)$. V jednotlivých častiach postupne preberieme problémy, kvôli ktorým sme museli meniť maticu \mathbf{A} , resp. \mathbf{S} . Následne budeme meniť parameter α , ktorý sa vyskytuje v matici \mathbf{G} a ukážeme aký vplyv má jeho zmena na počet krokov algoritmu a na výsledné ohodnotenie.

4.3.1 Iteračný proces s maticou liniek \mathbf{A}

V podkapitole 2.4 sme prešli od matice liniek \mathbf{A} ku matici \mathbf{S} nahradením nulových riadkov. V tomto príklade ukážeme iteračný proces na webe, ktorý obsahuje slepý vrchol. Web na obrázku 4.1 obsahuje slepý vrchol; stránku 5. Keď na ňom spustíme iteračný



Obr. 4.1: Iteračný proces s webom so slepým vrcholom

proces, tak po každom kroku sa zmenší súčet súradníc vektora \mathbf{x}_k o hodnotu $\mathbf{x}_{k-1}(5)$

z predchádzajúceho kroku iterácie.⁴ Matica \mathbf{A} ktorou násobíme vektor ohodnotenia nie je riadkovo stochastická. V matici \mathbf{A} prislúcha stránke 5 nulový riadok. Ohodnotenie stránok v k -tom kroku vieme dostať na základe ohodnotenia v $k - 1$ -vom kroku nasledovne

$$\begin{aligned} \mathbf{x}_k(1) &= \mathbf{x}_{k-1}(2)/3 + \mathbf{x}_{k-1}(3)/2 \\ \mathbf{x}_k(2) &= \mathbf{x}_{k-1}(4) \\ \mathbf{x}_k(3) &= \mathbf{x}_{k-1}(1)/2 \\ \mathbf{x}_k(4) &= \mathbf{x}_{k-1}(1)/2 + \mathbf{x}_{k-1}(2)/3 + \mathbf{x}_{k-1}(3)/2 \\ \mathbf{x}_k(5) &= \mathbf{x}_{k-1}(2)/3 \end{aligned}$$

Dôležitosť stránok $1, \dots, 5$ v k -tom kroku počítame iba na základe ohodnotenia stránok $1, \dots, 4$. Preto v k -tom kroku sa zmenší súčet súradníc práve o ohodnotenie stránky 5 v $k - 1$ -vom kroku.

V príslušnom grafe máme znázornené ohodnotenie stránok pre prvých 10 krokov. Vidíme, že súčet ohodnotení jednotlivých stránok sa v každom kroku znižuje.

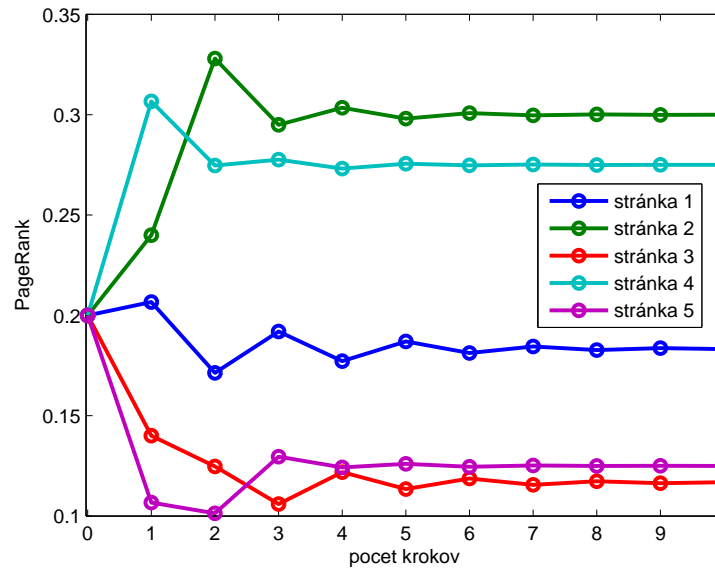
4.3.2 Iteračný proces s upravenou maticou liniek \mathbf{S}

Nahradíme teraz riadok v matici liniek \mathbf{A} z webu na obrázku 4.1, ktorý prislúcha slepému vrcholu vektorom $(1/5, 1/5, \dots, 1/5)$. Dostávame tak upravenú maticu liniek

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

Pre maticu \mathbf{S} platí, že je riadkovo stochastická, preto vektor nemení súčet súradníc po vynásobení touto maticou. Iteračný proces je znázornený na obrázku 4.2

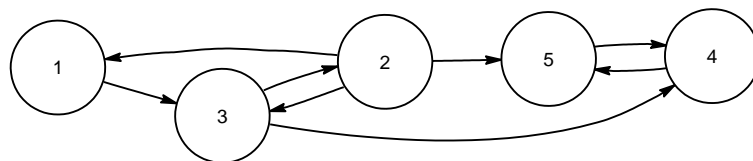
⁴Výrazom $\mathbf{x}_k(i)$ značíme hodnotu i -tej súradnice vektora ohodnotenia v k -tom kroku procesu



Obr. 4.2: Iteračný proces s webom so slepým vrcholom

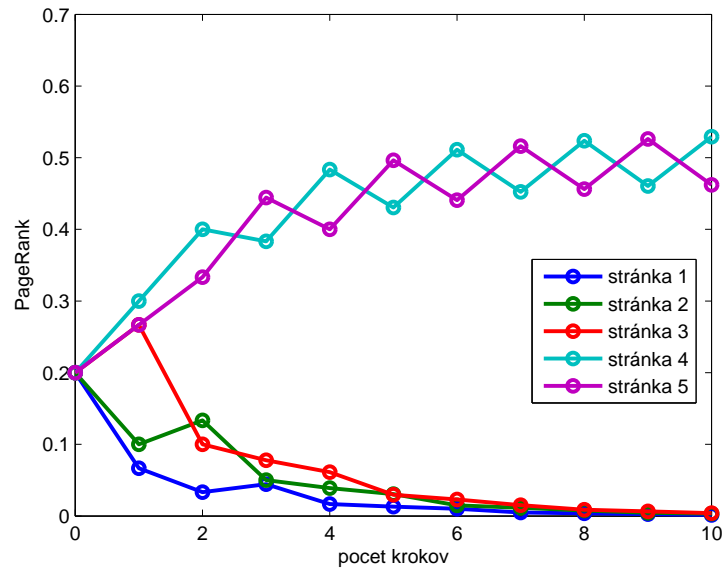
Vidíme, že stránky majú po desiatich krokoch rovnaké poradie ako v prípade iterácie s neupravenou maticou liniek. Rozdiel je v tom, že po každom kroku je súčet ohodnotení stránok rovný jednej.

Web s ktorým sme doteraz pracovali neobsahoval cyklus, z ktorého nevychádza linka do zvyšku grafu. Odstránenie problémov s takýmito cyklami bol jeden z dôvodov prechodu ku matici \mathbf{G} s parametrom α . Na obrázku 4.3 je jednoduchý graf, ktorý obsahuje cyklus z ktorého nevychádza žiadna linka. Iteračný proces pre tento graf je znázornený na obrázku 4.4.



Obr. 4.3: Web, ktorý obsahuje cyklus s vrcholmi 4 a 5

Ohodnotenie má v tomto prípade tendenciu naakumulovať sa vo vrcholoch cyklu. Zodpovedá to presne predstave náhodného surfera, ktorý po návšteve stránky 4 alebo 5 zostane v cykle tvorenom týmito stránkami. Nebude môcť navštíviť iné stránky. Prechodom ku matici \mathbf{G} problém s cyklom vyriešime, pretože umožníme surferovi v istom percente prípadov určenom parametrom α teleportovať sa na inú stránku. Iteračnému procesu pre maticu \mathbf{G} sa budeme venovať v časti 4.3.3.



Obr. 4.4: Iteračný proces s webom, ktorý obsahuje cyklus

Ostaňme ešte pri matici \mathbf{S} . Ľahko môžeme nájsť graf pre ktorý iteračný proces neskonverguje a to aj v prípade, že graf neobsahuje cyklus. Príklad takéhoto grafu je na obrázku 4.5a. Graf je tvorený z troch vrcholov. Rovnice pomocou ktorých vieme vypočítať ohodnotenie vyzerajú nasledovne

$$\mathbf{x}_k(1) = \mathbf{x}_{k-1}(2) + \mathbf{x}_{k-1}(3)$$

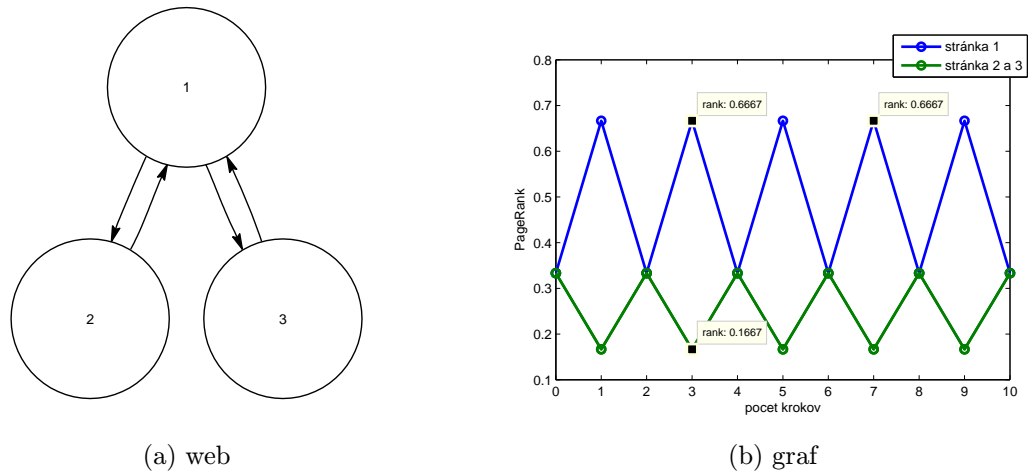
$$\mathbf{x}_k(2) = \mathbf{x}_{k-1}(1)/2$$

$$\mathbf{x}_k(3) = \mathbf{x}_{k-1}(1)/2$$

Keď začneme iteračný proces s ohodnotením $\mathbf{x}_0 = (1/3, 1/3, 1/3)$, tak ohodnotenie v ďalších krokoch bude $\mathbf{x}_1 = (2/3, 1/6, 1/6)$, $\mathbf{x}_2 = (1/3, 1/3, 1/3)$. Ohodnotenie sa bude v ďalších krokoch periodicky opakovať s periódou 2. Prvých 10 krokov procesu je znázornených na obrázku 4.5b. Pri dôkaze tvrdenia 3.2.1 sme spomenuli, že pri diagonalizovateľných maticiach robí problém prítomnosť vlastných čísel s absolútnou hodnotou rovnou 1 iných ako vlastné číslo 1. Vlastné čísla upravenej matice liniek \mathbf{S} , ktorá prislúcha webu 4.5a má vlastné hodnoty 1, 0 a -1 . Jordanov tvar tejto matice (až na poradie blokov) je

$$J_{\mathbf{S}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Matica \mathbf{S} je teda diagonalizovateľná (každý Jordanov blok má veľkosť 1). Prítomnosť vlastného čísla -1 teda zabraňuje konvergencii.

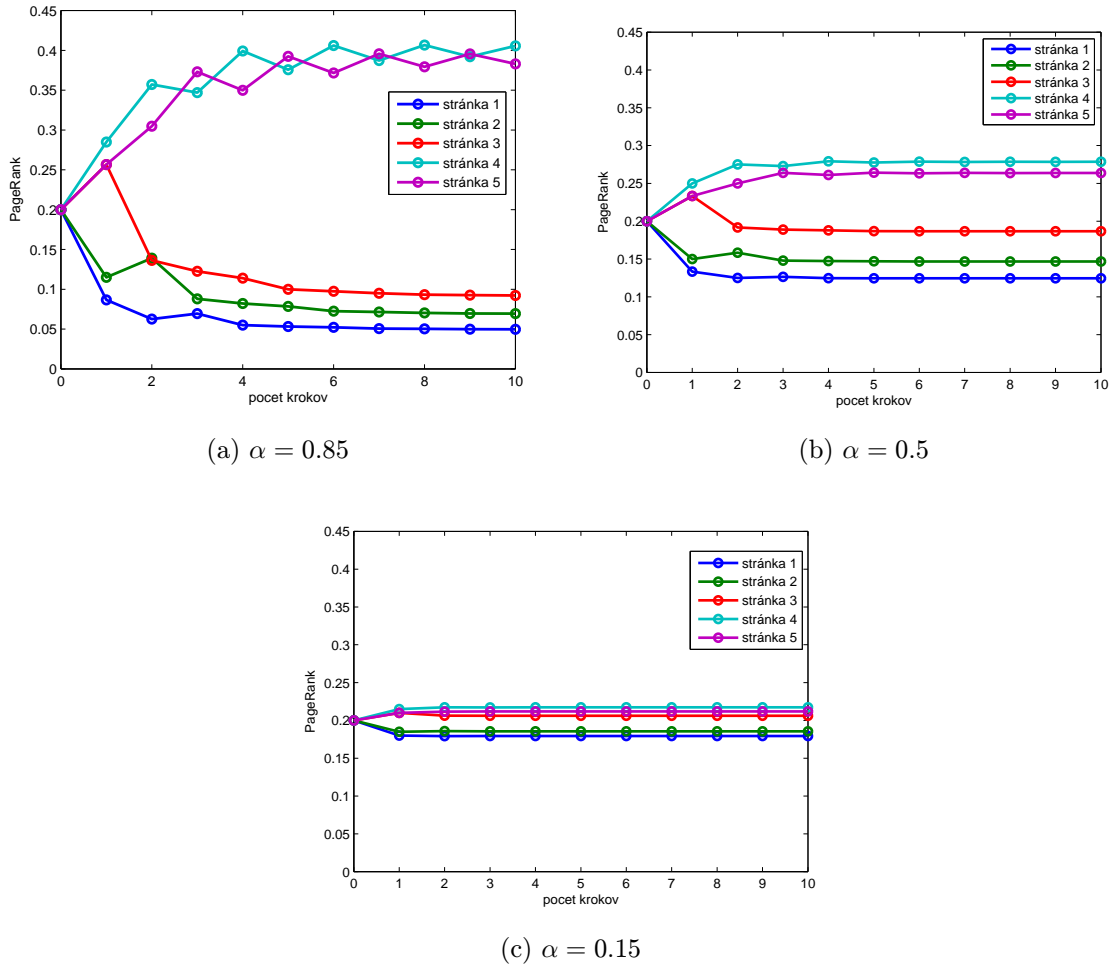
Obr. 4.5: Iteračný proces s webom, ktorý má vlastné číslo -1 .

4.3.3 Iteračný proces s \mathbf{G}

Prechodom ku matici \mathbf{G} vyriešime problém cyklov. Zároveň budeme mať zaručenú konvergenciu. V tejto časti sa budeme venovať hlavne zmene ohodnotenia pri zmene parametru α v matici \mathbf{G} . Pripomeňme, matica \mathbf{G} má tvar

$$\mathbf{G} = \alpha\mathbf{S} + (1 - \alpha)\mathbf{E} \quad \alpha \in (0, 1).$$

Pri webe 4.4 sme mali problém s cyklom na vrcholoch 4 a 5. Spustíme iteračný proces s maticou \mathbf{G} a nastavme parameter α postupne 0.85, 0.5, 0.15. Vidíme, že pre hodnotu $\alpha = 0.85$ dostaneme nenulové ohodnotenie aj pre stránky 1, 2 a 3. Keď zmenšujeme hodnotu α , tak ohodnotenie každej stránky sa blíži ku hodnote $1/5$. Môžeme to zdôvodniť tým, že pri menšom α prikladáme väčšiu váhu teleportačnej matici \mathbf{E} .

Obr. 4.6: Iteračný proces s maticou \mathbf{G} , pre rôzne hodnoty parametru α .

Iteračný proces pre web z obrázka 4.5a neskonvergoval ku riešeniu, ale ohodnotenie sa periodicky menilo kvôli prítomnosti vlastného čísla -1 v matici \mathbf{S} . Matica \mathbf{G} má potom pri voľbe koeficientu $\alpha = 0.85$ vlastné čísla 1 , 0 a -0.85 . Iteračný proces s maticou \mathbf{G} pri voľbe parametra $\alpha = 0.85$ je na obrázku 4.7. Výsledné ohodnotenie stránok je vektor $(0.4864, 0.2568, 0.2568)$.

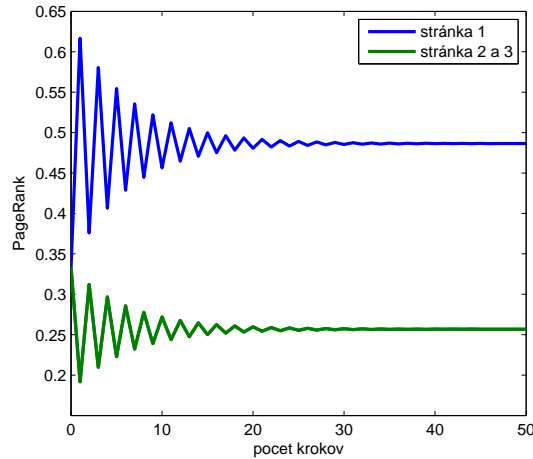
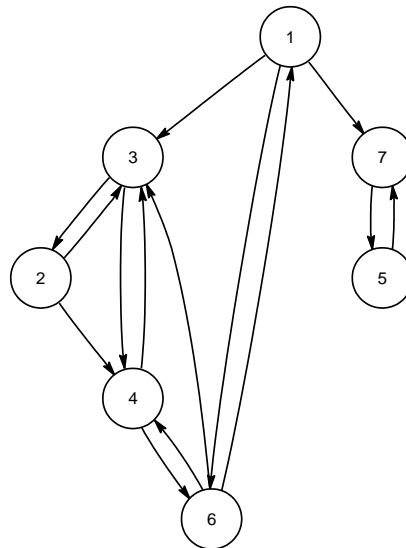
V nasledujúcom príklade je web, ktorého upravená matica liniek \mathbf{S} má okrem vlastného čísla 1 , vlastné číslo -1 s absolútnu hodnotu rovnou číslu 1 . Matica \mathbf{S} má navyše vlastné číslo 0.9805 , ktoré je blízke číslu 1 .

1.0000	0.9805	0.0980	0.0000	-0.5000	-0.5784	-1.0000
--------	--------	--------	--------	---------	---------	---------

Tabuľka 4.1: vlastné čísla upravenej matice \mathbf{S} webu 4.8

Podľa rovnice (3.10) a po nej nasledujúcich úvah plynie, že pri takejto voľbe vlastného čísla by mala byť konvergencia ku riešeniu najpomalšia.

Podľa tvrdenia 3.3.3 môžeme povedať, že matica \mathbf{G} má druhé najväčšie vlastné

Obr. 4.7: Iteračný proces pre web 4.5a, ktorý pre $\alpha = 1$ nekonvergovalObr. 4.8: Web, ktorého matica \mathbf{S} má druhé najväčšie vlastné číslo blízke 1

číslo rovné $\alpha\lambda_2$. V tomto príklade dokonca platí $|\lambda_2| = 1$. Ak pre absolútnu hodnotu druhého najväčšie číslo λ_2 platí

$$|\lambda_2| \sim 1,$$

tak potom

$$(\alpha\lambda_2)^k \sim \alpha^k$$

Budeme sa preto snažiť ukázať na konkrétnom príklade že rýchlosť konvergencie je približne taká ako ide mocnina α^k k nule. Všetky uvedené odhady sú veľmi hrubé. Predpokladajme, že chceme spočítať ohodnotenie s presnosťou na d desatinných miest. Musí potom platiť

$$\alpha^k < 10^{-d}.$$

Iteračný proces potom teda musí mať aspoň $-d/\log_{10} \alpha$ krokov, aby sme dosiahli presnosť d desatinných miest. Môžeme teraz urobiť hrubý odhad na počet krokov potrebných ku skonvergovaní.

Predpokladajme, že ohodnotenie chceme spočítať s presnosťou na 5 desatinných miest. Pre rôzne parametre α potom dostávame

α	počet krokov
0.5	17
0.75	56
0.85	70
0.95	224

Tabuľka 4.2: odhadnutý počet krokov potrebný pre výpočet ohodnotenia s presnosťou na 5 desatinných miest

Pozrime sa teraz na iteračný proces pre web 4.8. Iteračný procese pre web 4.8 sme zastavili, keď bola *absolútna chyba* menšia ako $\epsilon = 10^{-5}$. Pričom v k -tom kroku sme porovnávali absolútnu chybu počítali na základe vektora ohodnotenia získaného po 1000 krokoch. Iteračný proces teda zastavil, keď $\|\mathbf{x}_k - \mathbf{x}_{1000}\|_1 < 10^{-5}$.

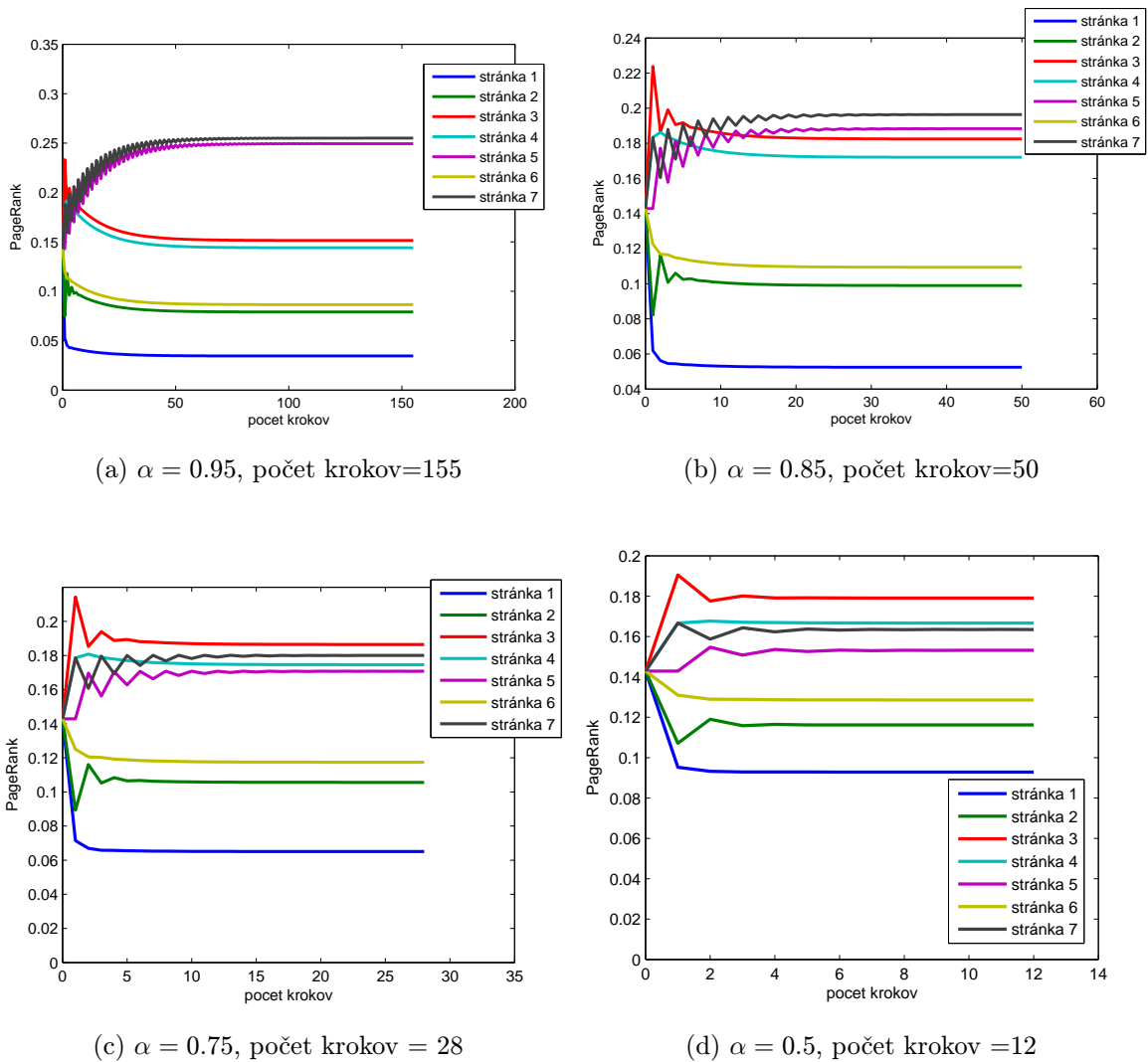
Pri zastavení iteračného na základe *relatívnej chyby* $\|\mathbf{x}_k - \mathbf{x}_{k-1}\|_1 < 10^{-5}$ boli počty krokov podobné.

α	počet krokov
0.5	13
0.75	30
0.85	54
0.95	168

Tabuľka 4.3: počty krokov pri zastavení iteračného procesu na základe relatívnej chyby

Počet krokov potrebný ku dosiahnutiu výsledku s presnosťou $\epsilon = 10^{-5}$ pre rôzne parametre α je veľmi podobný ako odhad v tabuľke 4.1, hoci tento odhad bol veľmi hrubý.

Za povšimnutie stojí, že web 4.8 obsahuje cyklus na vrcholoch 5 a 7 z ktorého nevychádza linka do zvyšku grafu. Vidíme, že pre zväčšujúce sa hodnotu α majú stránky tohoto cyklu vysoké ohodnotenie. V podkapitole 4.4 ukážeme, že každý graf s cyklom párnej dĺžky s ktorého nevychádza linka do zvyšku webu má vlastné číslo -1 .

Obr. 4.9: Počty krokov potrebných k dosiahnutiu presnosti 5 des. miest pre rôzne α .

4.4 Grafy s cyklom z ktorého nevychádza žiadna linka

Motivácia pre túto podkapitolu je ponúknuť jednoduché algebraické vysvetlenie faktu, že pre grafy s cyklom (z ktorých nevychádza žiadna linka) iteračná metóda s maticou \mathbf{S} neskonverguje. V teoretickej časti sme pomocou príkladu náhodného surfera ukázali, že máme problém s maticami, ktoré obsahujú takýto cyklus. Ďalej sme ukázali, že konvergencii (minimálne pre diagonalizovateľné matice \mathbf{S}) metóda neskonverguje kvôli prítomnosti iných vlastných čísel s absolútnou hodnotou 1 ako vlastné číslo 1. V tejto časti ukážeme, že grafy s cyklom obsahujú tieto vlastné hodnoty s absolútnou hodnotou 1.

Ukážeme, charakteristický polynóm matice \mathbf{S} pre tieto grafy obsahuje faktor $x^k - 1$, kde k je počet vrcholov v cykle.

Rovnica $x^k - 1 = 0$ má v komplexných číslach k riešení, preto charakteristický polynóm $ch_{\mathbf{S}}(x)$ má aspoň k koreňov s absolútnou hodnotou 1, teda, že matica \mathbf{S} má aspoň k vlastných čísel s veľkosťou 1. Okrem iného to znamená, že ak má cyklus párnou dĺžku $2k$, tak obsahuje vlastné číslo -1 , lebo $(-1)^{2k} - 1 = 0$.

Uvažujme jednoduchý graf, ktorého matica \mathbf{S} vyzerá nasledovne

$$\mathbf{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Charakteristický polynóm tejto matice je presne $ch_{\mathbf{S}}(x) = x^2 - 1$. Vlastné čísla matice \mathbf{S} sú teda 1 a -1 . Pre vektor $(1, -1)$ platí

$$(1, -1) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (-1)(1, -1)$$

Vektor $(1, -1)$ je teda vlastným vektorom pre vlastné číslo -1 . Matica cyklu vyzerá nasledovne

$$\mathbf{C}_k = \begin{pmatrix} 0 & 1 & 0 & & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \vdots & \\ & \vdots & & \ddots & \ddots & \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & & 0 & 0 \end{pmatrix}$$

Keď máme graf s n vrcholmi, ktorý obsahuje cyklus s k vrcholmi, tak existuje k riadkov a k stĺpcov v matici liniek, ktorých prieniky tvoria podmaticu \mathbf{C}' rozmerov $k \times k$. Riadky tejto matice môžeme preusporiadať tak, aby sme dostali maticu \mathbf{C} . Preznačme vrcholy tohoto grafu tak, aby vrcholy ktoré tvoria cyklus boli očíslované $1, \dots, k$ a aby platil cyklus $1 \rightarrow 2 \rightarrow \dots \rightarrow k \rightarrow 1$. Matica liniek \mathbf{S} potom bude matica, ktorá bude mať podmaticu \mathbf{C}_k v ľavom hornom rohu.

$$\mathbf{S} = \begin{pmatrix} \mathbf{C}_k & \mathbf{0} \\ \mathbf{X} & \mathbf{Y} \end{pmatrix} \quad (4.2)$$

Preznačenie znázorníme symbolicky s maticou \mathbf{C}_2 , ktorú presunieme do ľavého horného rohu.

$$\begin{pmatrix} * & 1 & * & 0 & * \\ \vdots & & \vdots & & \vdots \\ * & \dots & * & \dots & * \\ \vdots & & \vdots & & \vdots \\ * & 0 & * & 1 & * \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ * & * & * & * & * \\ \vdots & \vdots & * & * & * \\ * & * & * & * & * \end{pmatrix} \quad (4.3)$$

Pre maticu \mathbf{S} čo sme dostali preznačením platí, že má vlastné číslo -1 .

$$(1, -1, 0, \dots, 0) \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ * & * & * & * & * \\ \vdots & \vdots & * & * & * \\ * & * & * & * & * \end{pmatrix} = (-1)(1, -1, 0, \dots, 0)$$

Podobným spôsobom ide ukázať, že pre v prípade, že cyklus má párnú dĺžku $k = 2l$, tak vieme nájsť vlastný vektor $(\underbrace{1, -1, \dots, 1, -1}_{2l}, 0, \dots, 0)$ Pre blokovú maticu platí nasledovné tvrdenie

Tvrdenie 4.4.1. *Nech $1 \leq k < n$ a $\mathbf{A} \in \mathbb{F}^{n \times n}$ je bloková matica tvaru*

$$\mathbf{S} = \begin{pmatrix} \mathbf{C}_k & \mathbf{0} \\ \mathbf{X} & \mathbf{Y} \end{pmatrix},$$

kde $\mathbf{C}_k \in \mathbb{F}^{k \times k}$, $\mathbf{X} \in \mathbb{F}^{(n-k) \times k}$ a $\mathbf{Y} \in \mathbb{F}^{(n-k) \times (n-k)}$. Potom pre determinanty platí

$$|\mathbf{S}| = |\mathbf{C}_k| |\mathbf{Y}|.$$

Dôkaz. Dôkaz tohoto tvrdenia možno nájsť v [Z, tvrdenie 10.2.2]. □

Charakteristický polynóm matice \mathbf{S} teda môžeme rozložiť na súčin

$$ch_{\mathbf{S}}(x) = |\mathbf{S} - x\mathbf{I}| = |\mathbf{C}_k - x\mathbf{I}| |\mathbf{Y} - x\mathbf{I}|.$$

Zároveň pre determinant matice cyklu platí $|\mathbf{C}_k - x\mathbf{I}| = x^k - 1$ [BC, strana 166]. Preto pre charakteristický polynóm matice \mathbf{S} platí

$$ch_{\mathbf{S}}(x) = (x^k - 1) |\mathbf{Y} - x\mathbf{I}|.$$

Ako sme povedali vyššie rovnica $x^k - 1 = 0$ má k rôznych komplexných koreňov s veľkosťou 1. V matici s cyklom sa teda vyskytnú vlastné čísla s absolútnou hodnotou 1 iné ako číslo 1, ktoré kazia konvergenciu.

Kapitola 5

Využitie PageRank algoritmu

Algoritmus PageRank sa primárne využíva na usporiadanie stránok podľa dôležitosti, no možno ho využiť aj na iné účely. V tejto kapitole spomeniem niekoľko ďalších možných využití algoritmu PageRank

V článku [GM] ukazujú autori ako aplikovať PageRank algoritmus na ohodnotenie tímov v National Football League (NFL). Pod tímami si predstavili webové stránky a následne použili doterajšie výsledky v sezóne na vytvorenie liniek medzi nimi. Podarilo sa im zoradiť tímy NFL podľa sily a určiť ktorý tím má najväčšiu šancu dominovať pre zvyšok sezóny.

V článku [BRdS] autori predpovedajú pohyb človeka v meste s použitím algoritmu PageRank. Mesto reprezentujú ako graf, kde ulice sú linky a križovatky vrcholy.

Na webe [WIK] možno nájsť ďalšie informácie o tom, že PageRank bol použitý aj na hodnotenie článkov na základe citácií.

Kapitola 6

Záver

Naším cieľom bolo vysvetliť fungovanie algoritmu PageRank, s ktorým sa nepriamo stretávame vždy, keď zadávame nejaký kľúčový výraz do vyhľadávača Google a čakáme, že nám vyhodí relevantné stránky. Vysvetlili sme, že ohodnotenie stránok získame ako riešenie špeciálnej sústavy rovníc, ktorú zostrojíme na základe webu.

PageRank algoritmus hľadá riešenie tejto sústavy iteratívnou metódou. Ukázali sme, že tento spôsob je vhodný a že po určitom počte iterácii dosiahne ohodnotenie s dostatočnou presnosťou. Ďalej sme sa venovali parametru v tomto algoritme, ktorým vieme ovplyvniť rýchlosť konvergenencie.

Prácu sme sa snažili koncipovať tak, aby bola prístupná čo najširšiemu publiku. K jej porozumeniu je potrebné poznať základné pojmy z algebry ako je vektor, vektorový priestor, matica a determinant matice. V práci sa nachádza niekoľko tvrdení, ktoré sú podrobne dokázané, prípadne je uvedený odkaz na literatúru, kde je možné nájsť ich dôkaz.

Myšlienky vysvetliť algoritmus sme sa držali najmä v kapitole o praktických ukážkach behu algoritmu. Na niekoľkých konkrétnych sieťach sme ukázali beh algoritmu. Prehľadnými obrázkami sme demonštrovali beh algoritmu a ukázali ako sa mení ohodnotenie v priebehu niekoľkých krokoch iterácie.

Problematika okolo algoritmu PageRank je veľmi široká a ponúka veľký priestor pre ďalšiu prácu. Môžeme napríklad skúmať aký vplyv má zmena parametru α v matici \mathbf{G} na celkové poradie stránok, prípadne aký vplyv má malá zmena štruktúry webu na poradie stránok. Ďalšie smery ktorými sa možno uberať môžeme nájsť v použitej literatúre.

Táto práca môže slúžiť ako odrazový mostík pre všetkých, ktorých táto téma zaujala a ktorý by chceli skúmať ďalšie aspekty usporiadania stránok podľa popularity.

Literatúra

- [BBC⁺] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. SIAM, Philadelphia, 1994. 2nd edition.
- [BC] R. A. Brualdi and D. Cvetković. *A combinatorial approach to matrix theory and its applications*. Taylor & Francis Group, LLC, 6000 Broken Sound Parkway NW, Suite 300, 2009.
- [BL] Kurt Bryan and Tanya Leise. The \$25,000,000,000 eigenvector – the linear algebra behind Google. *SIAM Review*, 48(3):569–581, 2006. <http://www.rose-hulman.edu/~bryan/googleFinalVersionFixed.pdf>.
- [BPMW] S. Brin, L. Page, R. Motwami, and T. Winograd. *The PageRank citation ranking: Bringing order to the Web*. Computer Science Department, Stanford University, 1999.
- [BRdS] Johan Bollen, Marko A. Rodriguez, and Herbert Van de Sompel. Journal status. *CoRR*, abs/cs/0601030, 2006.
- [GM] A. Y. Govan and C. D. Meyer. Ranking national football league teams using Google’s PageRank. <http://www.ncsu.edu/crsc/reports/ftp/pdf/crsc-tr06-19.pdf>.
- [LM] A. N. Langville and C. D. Meyer. *Google’s PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, Princeton, 2006.
- [M] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, 2000.
- [QSS] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*. Springer, New York, 2000. Texts in Applied Mathematics 37.

- [S] M. Sleziak. Algebra 3. Poznámky k prednáške,
<http://thales.doa.fmph.uniba.sk/sleziak/vyuka/>.
- [WIK] Wikipedia. <http://en.wikipedia.org/wiki/PageRank>.
- [Z] P. Zlatoš. Lineárna algebra a geometria. Poznámky k prednáške,
<http://thales.doa.fmph.uniba.sk/zlatos/>.