Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Květa Cvrčková

Konvergence markovských řetězců

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Michaela Prokešová, Ph.D.
Studijní program: Matematika, finanční matematika

2009

Na tomto místě bych chtěla poděkovat všem, kteří mi při psaní bakalářské práce pomohli, především pak vedoucí této bakalářské práce RNDr. Michaele Prokešové, Ph.D. za její podporu a cenné připomínky.

Prohlašuji, že jsem svou bakalářskou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 1.8.2009 Květa Cvrčková

Obsah

Název práce: Konvergence markovských řetězců
Autor: Květa Cvrčková
Katedra (ústav): Katedra pravděpodobnosti a matematické statistiky
Vedoucí bakalářské práce: RNDr.Michaela Prokešová,Ph.D.
e-mail vedoucího: prokesov@karlin.mff.cuni.cz

Abstrakt: V předložené práci studujeme vybrané metody používané pro odvození mezí rychlosti konvergence markovského řetězce ke svému stacionárnímu rozdělení. Budeme se zabývat pouze nerozložitelnými aperiodickými markovskými řetězci. Největší pozornost budeme věnovat speciální třídě markovských řetězců, jejichž matice přechodu je v detailní rovnováze se svým stacionárním rozdělením. Také se zaměříme na řetězce definované pomocí grafů. Ukážeme si metodu založenou na průzkumu vlastních čísel matice přechodu, metody založené na studiu vlastností členů matice přechodu a metodu využívající vodivosti vhodného asociovaného grafu. Tyto metody si popíšeme, porovnáme mezi sebou a ukážeme si jejich použití na jednoduchých příkladech.

Klíčová slova: markovský řetězec, rychlost konvergence ke stacionárnímu rozdělení, vodivost grafu, vlastní čísla matice přechodu

Title: Convergence of Markov Chains
Author: Květa Cvrčková
Department: Department of Probability and Mathematical Statistics
Supervisor: RNDr. Michaela Prokešová, Ph.D.
Supervisor's e-mail address: prokesov@karlin.mff.cuni.cz

Abstract: In the present work we study selected methods for deriving bounds for the speed of convergence of Markov chains to their equilibrium distribution. We consider only aperiodic and irreducible Markov chains. We mainly focus on a special case of Markov chains for which the transition matrix and the equilibrium distribution fulfill the detailed balance condition. Moreover we study a class of Markov chains defined by graphs. We will present a method based on investigation of eigenvalues, several methods which use entries of the transition matrix directly and a method which uses the conductance of the asociated graph. We describe these methods and compare obtained results, also by aplications to simple examples.

Keywords: Markov chain, speed of convergence to the equilibrium distribution, graph conductance, eigenvalues of the transition matrix

Použité značení

$\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$	množina přirozených, celých, reálných, komplexních čísel
$\boldsymbol{a}$	sloupcový vektor
$\boldsymbol{a}^\top$	řádkový vektor
$\boldsymbol{A}$	matice
$\boldsymbol{A} > 0$	matice $\boldsymbol{A}$ taková, že $a_{ij} > 0$ pro $\forall i, j \in S$
$\boldsymbol{A}^{-1}$	matice inverzní k matici $\boldsymbol{A}$
$\boldsymbol{A}^k$	k-tá mocnina matice $\boldsymbol{A}$
$\boldsymbol{J}(\lambda, n) = (j_{il})_{i,l=1}^n$	Jordanova buňka, tj. matice taková, že $j_{il} := \begin{cases} \lambda \; i = l \\ 1 \; i = l - 1 \\ 0 \; jinak \end{cases}$
S	stavový prostor markovského řetězce
$\boldsymbol{p}(0) = (p_1, p_2, \ldots)$	počáteční rozdělení markovského řetězce
$\boldsymbol{p}(n) = (p_1, p_2, \ldots)$	rozdělení markovského řetězce v čase n
$\boldsymbol{P} = (p_{ij})_{i,j \in S}$	matice přechodu markovského řetězce
$\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots)$	stacionární rozdělení markovského řetězce
$\boldsymbol{\Pi} = (\pi_{ij})_{i,j \in S}$	matice, jejíž řádky tvoří vektor $\boldsymbol{\pi}^\top$, tj. $\pi_{ij} := \pi_j$
$\boldsymbol{P}^{(k)} = (p_{ij}^{(k)})_{i,j \in S}$	matice přechodu markovského řetězce po k krocích
$\lvert a \rvert$	absolutní hodnota čísla a

Kapitola 1

Úvod

Markovské řetězce jsou jedním z klasických modelů teorie pravděpodobnosti, zkoumaným už v první polovině dvacátého století. My se v naší práci zabýváme nejjednodušší verzí tohoto modelu – markovskými řetězci s diskrétním časem a diskrétní konečnou množinou stavů. Takové markovské řetězce začaly být v osmdesátých letech s nárůstem výpočetní kapacity počítačů velmi důležité i pro aplikace mimo pravděpodobnost a statistiku.

Jako takovou typickou aplikaci v kombinatorice/informatice můžeme uvést přibližné určení počtu nějakých složitých kombinatorických objektů. Obvykle je teoreticky možné určit jejich počet přesně, ale složitost výpočtu roste exponenciálně s velikostí problému a proto se velmi rychle narazí na meze výpočetní kapacity dostupných počítačů a prakticky se přímý výpočet nedá provést. Ale protože často stačí znát jenom aproximaci hledaného počtu, která s dostatečně velkou pravděpodobností je dostatečně přesná, můžeme použít jinou strategii. Místo abychom opravdu počítali objekty, které nás zajímají, zadefinujeme si je jako podmnožinu nějaké množiny S, jejíž velikost známe (grafy s nějakou speciální vlastností na N vrcholech jako podmnožinu všech možných grafů na N vrcholech) a poté generujeme rovnoměrné náhodné vzorky z množiny S a náš hledaný počet aproximujeme velikostí množiny S a pozorovanou proporcí grafů z podmnožiny.

Problém výše uvedeného postupu je ale v tom, že přímé generování z takového rovnoměrného rozdělení (říkejme mu $\boldsymbol{\pi}$) není typicky možné a v této situaci je právě vhodné použít markovské řetězce. Místo přímého generování vzorků z rozdělení $\boldsymbol{\pi}$ totiž zkonstruujeme markovský řetězec na stavovém prostoru S s limitním rozdělením $\boldsymbol{\pi}$ a generujeme-li dostatečně dlouhou realizaci z takového markovského řetězce, pak se rozdělení takto generovaných hodnot bude blížit (limitnímu) rozdělení $\boldsymbol{\pi}$. Obdobná strategie se používá i v mnoha jiných aplikacích jako je analýza obrazu, generování markovských a Gibbsových polí či maximalizace simulovaným žíháním. Podrobnější informace o těchto aplikacích lze najít např. [5] nebo [6], včetně postupů jak najít/zvolit vhodnou matici pravděpodobností přechodu $\boldsymbol{P}$ s limitním rozdělením $\boldsymbol{\pi}$.

My se budeme dále zabývat pouze otázkou, jak určit nebo odhadnout rychlost konvergence marginálního rozdělení markovského řetězce určeného maticí pravděpodobností přechodu $\boldsymbol{P}$ k jeho limitnímu rozdělení $\boldsymbol{\pi}$.

Ukážeme si, že pro aperiodický nerozložitelný řetězec existuje jediné stacionární rozdělení $\boldsymbol{\pi}$ a že právě stacionární rozdělení se shoduje s limitním rozdělením. Proto se v celé práci budeme zabývat pouze těmito řetězci.

V druhé kapitole si pouze připomeneme nejdůležitější věty o konvergenci markovského řetězce. Také si uvedeme několik základních tvrzení o vlastních číslech matice přechodu $\boldsymbol{P}$. Zjistíme, proč právě vlastní čísla matice $\boldsymbol{P}$ budou hrát důležitou roli v naší práci.

V rozsáhlejší třetí kapitole si uvedeme různé metody pro samotný odhad rychlosti konvergence marginálního rozdělení k limitnímu rozdělení. Nejdůležitějším a nejlepším odhadem bude odhad na základě vlastních čísel matice $\boldsymbol{P}$. Bohužel zjistíme, že tento odhad má mnohá omezení.

Ve čtvrté kapitole se pokusíme některá z těchto omezení odstranit. Budeme se zabývat speciálními markovskými řetězci – řetězci definovanými pomocí grafů. Ukážeme si, jak pomocí konstanty přidružené danému řetězci určit rychlost konvergence. Určení této konstanty také nebude jednoduché. Proto si ukážeme tzv. metodu kanonických cest.

Cílem této práce bude odvodit a ukázat si různé metody pro odhad rychlosti konvergence markovského řetězce ke stacionárnímu rozdělení. Ukážeme si jejich aplikaci na jednoduchých příkladech, porovnáme je mezi sebou.

Kapitola 2

Konvergence markovského řetězce

Zde uvádíme seznam vět z literatury, které v práci používáme.

Věta 1 (Prášková, Lachout [4], str.41) *V nerozložitelném řetězci s konečně mnoha stavy jsou všechny stavy trvalé nenulové.*

Věta 2 (Prášková, Lachout [4], str.51) *Nechť je dán nerozložitelný markovský řetězec. Jsou-li všechny jeho stavy trvalé nenulové, stacionární rozdělení existuje a je jediné.*
Jsou-li všechny stavy aperiodické, potom pro stacionární pravděpodobnosti π_j platí

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} > 0 \quad i, j \in S$$

$$\pi_j = \lim_{n\to\infty} p_j(n) > 0 \quad j \in S.$$

Jsou-li všechny stavy periodické, platí

$$\pi_j = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}^{(k)} > 0 \quad i, j \in S$$

$$\pi_j = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_j(k) > 0 \quad j \in S.$$

Věta 3 (Prášková,Lachout [4], str.54) *V nerozložitelném řetězci s konečně mnoha stavy stacionární rozdělení existuje.*

Hlavní roli při studiu metod zjišťování rychlosti konvergence markovského řetězce ke svému stacionárnímu rozdělení hrají vlastní čísla matice $\boldsymbol{P}$. Následující věty napoví, proč tomu tak je.

Věta 4 (Behrends [1], str.71) *Nechť $\boldsymbol{A}$ je stochastická matice. Pak 1 je vlastní číslo a každé vlastní číslo splňuje $\mid \lambda \mid \leq 1$.*

Zřejmě toto platí i pro matici $\boldsymbol{P}$.

Věta 5 (Behrends [1], str.50) *Markovský řetězec je*
(i) nerozložitelný $\Leftrightarrow\ \exists\ k \in \mathbb{N}$ tak, že $(\boldsymbol{P} + \boldsymbol{P}^2 + \ldots + \boldsymbol{P}^k) > 0$
(ii) nerozložitelný aperiodický $\Leftrightarrow\ \exists\ k \in \mathbb{N}$ tak, že $\boldsymbol{P}^k > 0$.

Věta 6 (Behrends [1], str.72) *Nechť $\exists\ k \in \mathbb{N}$ tak, že $\boldsymbol{P}^k > 0$. Pak 1 je jediné vlastní číslo takové, že $\mid \lambda \mid = 1$.*

Věta 7 (Behrends [1], str.75) *Nechť $\exists\ k \in \mathbb{N}$ tak, že $\boldsymbol{P}^k > 0$. Pak existuje invertovatelná matice $\boldsymbol{S}$ taková, že $\boldsymbol{S}^{-1}\boldsymbol{P}\boldsymbol{S}$ má tvar:*

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & \boldsymbol{J}(\lambda_2, n_2) & 0 & \ldots & 0 \\ 0 & 0 & \boldsymbol{J}(\lambda_3, n_3) & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \boldsymbol{J}(\lambda_r, n_r) \end{pmatrix},$$

kde $\boldsymbol{J}(\lambda, n)$ jsou Jordanovy buňky a platí $\mid \lambda_2 \mid, \cdots, \mid \lambda_r \mid < 1$.
Je-li navíc $\boldsymbol{P}$ podobná samoadjungované matici, pak dimenze n_p Jordanových buněk jsou jedna, vlastní čísla matice $\boldsymbol{P}$ jsou reálná a mohou být oindexováva tak, že $1 \geq\mid \lambda_2 \mid\geq\mid \lambda_3 \mid\geq \cdots \geq\mid \lambda_n \mid\geq -1$.

Kapitola 3

Základní metody odhadů

3.1 Úvod

Budeme se zabývat aperiodickými nerozložitelnými markovskými řetězci s konečnou množinou stavů, neboť pro každý z těchto řetězců existuje právě jedno stacionární rozdělení (viz Věta 2).

Bude nás zajímat, jak rychle marginální rozdělení markovského řetězce $\boldsymbol{p}(k)$ konverguje k limitnímu stacionárnímu rozdělení $\boldsymbol{\pi}$, neboli jak rychle $\boldsymbol{P}^k$ konverguje k $\boldsymbol{\Pi}$, matici, jejíž řádky tvoří vektor $\boldsymbol{\pi}^\top$.

Z Věty 7 jsme se dozvěděli, že pro matici přechodu $\boldsymbol{P}$ existuje matice $\boldsymbol{S}$ taková, že $\boldsymbol{S}^{-1}\boldsymbol{P}\boldsymbol{S}$ je sestavena z bloků na diagonále: prvním blokem je číslo 1, ostatními bloky pak Jordanovy buňky $\boldsymbol{J}(\lambda, n)$, kde $\mid \lambda \mid< 1$. Proto $\lim_{k\to\infty}(\boldsymbol{S}^{-1}\boldsymbol{P}\boldsymbol{S})^k = \boldsymbol{S}'$, matici mající 1 v levém horním rohu a ostatní členy nulové. Platí tedy

$$\lim_{k\to\infty} \boldsymbol{P}^k = \boldsymbol{S}\boldsymbol{S}^{-1} \lim_{k\to\infty} \boldsymbol{P}^k \boldsymbol{S}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{S}'\boldsymbol{S}^{-1}, \tag{3.1}$$

což je matice se shodnými řádky, a tudíž z jednoznačnosti stacionárního rozdělení matice $\boldsymbol{\Pi}$. Z rovnice (3.1) také plyne, že $\boldsymbol{P}^k$ nebude k $\boldsymbol{\Pi}$ konvergovat rychleji než $(\boldsymbol{J}(\lambda^*, n))^k$ k 0, kde $\lambda^* := max\{\mid \lambda \mid; \lambda \text{ vlastní číslo}, \lambda \neq 1\}$ značí největší vlastní číslo matice $\boldsymbol{P}$ v absolutní hodnotě. Toto bude základní myšlenkou této kapitoly.

Věta 7 také říká, že pokud $\boldsymbol{P}$ bude podobná samodjungované matici, vyhneme se technickým problémům s úpravou Jordanových bloků a složitému počítání s komplexními čísly. Proto se zaměříme na speciální případy markovských řetězců, kde $\boldsymbol{P}$ a $\boldsymbol{\pi}$ budou splňovat určitou podmínku a splnění této podmínky nám zajistí podobnost matice $\boldsymbol{P}$ samoadjungované matici. Pro tyto speciální případy si uvedeme výpočet mezí na vzdálenost $\boldsymbol{P}$ a $\boldsymbol{\Pi}$ používající vlastních čísel matice $\boldsymbol{P}$.

Ukážeme si, že aplikace tohoto odhadu naráží na mnoho překážek. Hned první překážkou je, že ne vždy umíme najít vlastní čísla matice $\boldsymbol{P}$. Proto se budeme snažit uvést si i odhady, které vlastní čísla nevyužívají. K aplikaci těchto odhadů nebudeme potřebovat nic jiného než znalost matice $\boldsymbol{P}$. Bohužel, v těchto případech dostaneme často mnohem horší výsledky, než v případě odhadu využívacího vlastních čísel.

Řekli jsme, že chceme odhadnout vzdálenost $\boldsymbol{P}^k$ a $\boldsymbol{\Pi}$. Ale jak ji vlastně budeme definovat? Máme mnoho možností. Můžeme ji definovat například jako *absolutní chybu* (tj.$\mid p_{ij}^{(k)} - \pi_j \mid$), *relativní chybu* (tj. $\mid p_{ij}^{(k)} - \pi_j \mid / \pi_j$), chybu $\sum_j \mid p_{ij}^{(k)} - \pi_j \mid$, nebo jakkoli jinak. V naší práci si uvedeme odhady pro různé chyby.

Naším cílem bude odvodit si a úvest věty typu: jestliže $k \geq k_0$, pak vzdálenost mezi $\boldsymbol{P}^k$ a $\boldsymbol{\Pi}$ je menší než ϵ.

Uvedeme si také dostatečné množství příkladů. Ukážeme použití vět na konkrétních jednoduchých příkladech, porovnáme výsledky různých odhadů. Budeme se také zabývat problémy aplikace námi odvozených vět.

3.2 Detailní rovnováha

Jak jsme již řekli, v první části této kapitoly bude naším cílem najít odhad mezí pro rychlost konvergence matice $\boldsymbol{P}$ k matici $\boldsymbol{\Pi}$ pomocí vlastních čísel matice $\boldsymbol{P}$. Věta 7 říká, že bude-li $\boldsymbol{P}$ podobná samoadjungované matici, vyhneme se nepříjemné práci s Jordanovými buňkami a komplexními čísly. Na začátku si proto zavedeme podmínku detailní rovnováhy pro matici $\boldsymbol{P}$ a její stacionární rozdělení $\boldsymbol{\pi}$ – podmínku, která nám zaručí podobnost matice $\boldsymbol{P}$ samoadjungované matici, což si dokážeme v závěru této podkapitoly. Uvedeme si také některé zajímavé vlastnosti takovéto matice $\boldsymbol{P}$ a ukážeme si, že některé speciální matice tuto podmínku vždy splňují.

Zaveďme si tedy nyní podmínku detailní rovnováhy a vysvětleme si její pravděpodobnostní význam.

Definice. Mějme nerozložitelný markovský řetězec. Řekneme, že $\boldsymbol{P}$ *a* $\boldsymbol{\lambda}$ *jsou v detailní rovnováze*, jestliže pro pravděpodobnostní vektor $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_N)^\top$ platí

$$\lambda_i p_{ij} = \lambda_j p_{ji} \quad \forall i, j \in S. \tag{3.2}$$

Potom ale dostáváme

$$\sum_{i \in S} \lambda_i p_{ij} = \sum_{i \in S} \lambda_j p_{ji} = \lambda_j \sum_{i \in S} p_{ji} = \lambda_j \quad \forall j \in S,$$

neboli $\boldsymbol{\lambda}^\top \boldsymbol{P} = \boldsymbol{\lambda}^\top$. Z jednoznačnosti stacionárního rozdělení tudíž plyne, že

$\boldsymbol{\lambda} = \boldsymbol{\pi}$. Pro přehlednost zaměníme (3.2) s podmínkou

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j \in S. \tag{3.3}$$

Představme si, že počátečním rozdělením markovského řetězce je stacionární rozdělení nebo že počáteční rozdělení je libovolné a řetězec „necháme běžet“ dostatečně dlouho, aby $\boldsymbol{p}(n) \approx \boldsymbol{\pi}$. Pak v libovolném následujícím kroku pravděpodobnost nalezení řetězce ve stavu i je π_i. Podmínka detailní rovnováhy pak pro libovolný nerozložitelný aperiodický řetězec znamená, že pro libovolné dva stavy $i, j \in S$ bude součin pravděpodobnosti nalezení řetězce ve stavu i a pravděpodobnosti následného skoku do j (tj. $\pi_i p_{ij}$) shodný se součinem pravděpodobnosti nalezení řetězce ve stavu j a pravděpodobnosti následného skoku do i (tj. $\pi_j p_{ji}$).

Zřejmě pro obecný řetězec tato podmínka splněna nebude, tj. $\pi_i p_{ij}$ bude různé od $\pi_j p_{ji}$.

Pro lepší pochopení podmínky detailní rovnováhy a pro zajímavost se podívejme ještě na jeden zvláštní typ markovských řetězců.

Mějme nerozložitelný řetězec s maticí přechodu $\boldsymbol{P}$ a se stacionárním rozdělením $\boldsymbol{\pi}$ a zkusme najít další stochastickou matici $\tilde{\boldsymbol{P}}$ s tím samým stacionárním rozdělením $\boldsymbol{\pi}$ a s podmínkou

$$\pi_i p_{ij} = \pi_j \tilde{p_{ji}} \quad \forall i, j \in S. \tag{3.4}$$

Překvapivě můžeme jednoduše vzít (3.4) jako definici $\tilde{\boldsymbol{P}}$, tj. definovat

$$\tilde{p_{ji}} := \frac{\pi_i p_{ij}}{\pi_j} \quad (\text{tj.} \quad \tilde{p_{ij}} := \frac{\pi_j p_{ji}}{\pi_i}) \qquad \forall i, j \in S.$$

Ukažme, že tato matice vždy existuje. Víme, že $\pi_i > 0 \quad \forall i \in S$ (viz. Věta 2), proto $\pi_i p_{ji} / \pi_j$ vždy existuje a je kladné. Tudíž je-li $\tilde{\boldsymbol{P}}$ stochastická, bude zároveň nerozložitelná. $\tilde{\boldsymbol{P}}$ opravdu stochastická je, neboť

$$\sum_{j \in S} \tilde{p_{ij}} = \sum_{j \in S} \frac{\pi_j}{\pi_i} p_{ji} = \frac{1}{\pi_i} \sum_{j \in S} \pi_j p_{ji} = \frac{1}{\pi_i} \pi_i = 1 \quad \forall j \in S. \tag{3.5}$$

Ukažme dále, že vektor $\boldsymbol{\pi}$ skutečně je stacionárním rozdělením matice $\tilde{\boldsymbol{P}}$, tj. že

$$\sum_{i \in S} \pi_i \tilde{p_{ij}} = \pi_j \ \forall j \in S.$$

Tedy

$$\sum_{i \in S} \pi_i \tilde{p_{ij}} = \pi_i \frac{\pi_j}{\pi_i} p_{ji} = \pi_j \sum_{i \in S} p_{ji} = \pi_j \quad \forall j \in S.$$

Vidíme, že vlastnosti $\boldsymbol{P}$ a $\boldsymbol{\pi}$ implikují, že $\tilde{\boldsymbol{P}}$ je stochastická, nerozložitelná a splňuje dané podmínky.

Řetězec $\tilde{\boldsymbol{P}}$ přidružený k $\boldsymbol{P}$ budeme nazývat *časově obrácený k původnímu řetězci.*

Všimněme si několika vlastností matic $\boldsymbol{P}$ a $\tilde{\boldsymbol{P}}$.

Definice. Řekneme, že matice $\boldsymbol{A}$ je *dvojně stochastická*, jestliže platí $\sum_j a_{ij} = 1 \ \forall j$ a $\sum_i a_{ij} = 1 \ \forall i$.

Věta 8 *Nechť $\boldsymbol{P}$ je nerozložitelná dvojně stochastická matice. Pak stacionární rozdělení je rovno rovnoměrnému rozdělení.*

Důkaz. Z definice stacionárního rozdělení a z Věty 2 víme, že $\pi_j = \sum_i \pi_i p_{ij}$ a $\lim_{k\to\infty} p_{ij}^{(k)} = \pi_j$. Transponujeme-li matici $\boldsymbol{P}$, dostaneme opět stochastickou nerozložitelnou matici $\boldsymbol{O} := \boldsymbol{P}^\top$, jejíž stacionární rozdělení označme $\boldsymbol{\pi}^o = (\pi_1^o, \pi_2^o, \ldots, \pi_N^o)^\top$. Platí

$$\pi_j^o = \sum_i \pi_i^o o_{ij} = \sum_i \pi_i^o p_{ji} \text{ a } \pi_j^o = \lim_{k\to\infty} o_{ij}^{(k)} = \lim_{k\to\infty} p_{ji}^{(k)}.$$

Ale z jednoznačnosti limity

$$\pi_j = \lim_{k\to\infty} p_{ij}^{(k)} = \lim_{k\to\infty} o_{ji}^{(k)} = \pi_i^o \quad \forall i, j,$$

tj. $\pi_1 = \pi_2 = \ldots = \pi_N = \pi_1^o = \pi_2^o = \ldots = \pi_N^o$. Jelikož $\sum_j \pi_j = 1$, nutně platí $\pi_1 = \pi_2 = \ldots = \pi_N = 1/N$. $\square$

Poznámka. Všimněme si, že symetrická matice je zvláštním případem dvojně stochastické matice.

Věta 9 *Jestliže $\boldsymbol{P}$ je nerozložitelná dvojně stochastická, pak i $\tilde{\boldsymbol{P}}$ je nerozložitelná dvojně stochastická.*

Důkaz. Na straně 12 jsme ukázali, že jestliže bude matice $\boldsymbol{P}$ nerozložitelná, bude rovněž $\tilde{\boldsymbol{P}}$ nerozložitelná. Přímo z definice dvojně stochastické matice a z (3.5) víme, že $\sum_{i=1}^N p_{ij} = 1$, $\sum_{j=1}^N p_{ij} = 1$, $\sum_{j=1}^N \tilde{p}_{ij} = 1$. Předchozí Věta 8 říká, že $\boldsymbol{\pi}$ má rovnoměrné rozdělení. Tvrdíme, že $\sum_{i=1}^N \tilde{p}_{ij} = 1$, což platí, neboť

$$\sum_{i=1}^N \tilde{p}_{ij} = \sum_{i=1}^N \frac{\pi_j}{\pi_i} p_{ji} = \sum_{i=1}^N \frac{\frac{1}{N}}{\frac{1}{N}} p_{ji} = \sum_{i=1}^N p_{ji} = 1 \quad \forall j.$$

$\square$

Věta 10 *Nechť $\boldsymbol{P} > 0$ je dvojně stochastická matice. Pak $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze $\Leftrightarrow$ $\boldsymbol{P}$ je symetrická.*

Důkaz. Věta 5 říká, že $\boldsymbol{P} > 0$ je nerozložitelná aperiodická. Z Věty 8 vyplývá, že stacionární rozdělení je rovno rovnoměrnému. Pak už

$$(\pi_i p_{ij} = \pi_j p_{ji}) \;\Leftrightarrow\; \left(\frac{1}{N} p_{ij} = \frac{1}{N} p_{ji}\right) \;\Leftrightarrow\; (p_{ij} = p_{ji}) \qquad \forall i, j \ .$$

□

Věta 11 *Nechť $\boldsymbol{P}$ je nerozložitelná stochastická matice a nechť všechny její řádky jsou shodné. Pak $\boldsymbol{P}$ je v detailní rovnoráze se stacionárním rozdělením $\boldsymbol{\pi}$.*

Důkaz. Všechny řádky matice $\boldsymbol{P}$ jsou shodné, neboli $p_{1j} = p_{2j} = \ldots = p_{Nj} \quad \forall j$. Proto platí

$$\pi_j = \sum_i \pi_i p_{ij} = \sum_i \pi_i p_{1j} = p_{1j} \sum_i \pi_i = p_{1j} \quad \forall j \ .$$

Podmínka detailní rovnováhy se v našem případě změní na podmínku

$$(\pi_i p_{ij} = \pi_j p_{ji}) \;\Leftrightarrow\; (\pi_i p_{1j} = \pi_j p_{1i}) \;\Leftrightarrow\; (p_{1i} p_{1j} = p_{1j} p_{1i}) \ ,$$

která zřejmě platí pro $\forall i, j$. □

Poznámka. Zamysleme se ještě, proč je vhodné omezit definici časově obráceného řetězce pouze na nerozložitelné řetězce. V případě těchto řetězců je totiž zaručena existence stacinárního rozdělení $\boldsymbol{\pi}^\top = (\pi_1, \pi_2, \ldots, \pi_N)$ (viz Věta 3) a platí $\pi_j > 0 \quad \forall j$(viz Věta 2). Rozložitelné konečné řetězce můžeme rozdělit do dvou skupin:

a) řetězce, které obsahují přechodné stavy. Pak stacionární rozdělení nutně obsahuje nulové členy a my bychom opět dělili nulou. Zároveň by nebylo jasné, jak definovat obrácený přechod ze stavu tvalého do stavu přechodného. Mohli bychom ho například položit roven nule, ale pak by $\tilde{\boldsymbol{P}}$ nebyla stochatická.

b) řetězce se dvěma či více vzájemně nedosažitelnými uzavřenými množinami trvalých stavů. V tomto případě stacionární rozdělení není jediné. Přesto bychom pro každé z těchto stacionárních rozdělení bez nulových prvků dostali tu samou matici $\tilde{\boldsymbol{P}}$. Tedy v tomto případě bychom mohli $\tilde{\boldsymbol{P}}$ definovat i pro tyto řetězce. Uvědomme si, že v tomto případě máme vlastně několik nerozložitelných řetězců spojených v jeden rozložitelný řetězec.

Na závěr této podkapitoly si uvedeme větu shrnující vlastnosti matic $\boldsymbol{P}$, které jsou v detailní rovnováze s $\boldsymbol{\pi}$. Ukážeme si, jak v obecném případě snadno poznat, zda $\boldsymbol{P}$ a $\boldsymbol{\pi}$ splňují podmínku detailní rovnováhy. Jak jsme slíbili v úvodu, dokážeme, že $\boldsymbol{P}$ je podobná samoadjungované matici. Uvedeme si, kolik takovýchto matic existuje.

Věta 12 *Nechť $\boldsymbol{P}$ je nerozložitelná matice se stacionárním rozdělením $\boldsymbol{\pi}$.*
(i) *Označme $\boldsymbol{D} = (d_{ij})$ matici takovou, že $d_{ii} = \sqrt{\pi_i}$ a ostatní členy má nulové.*
Pak $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze $\Leftrightarrow$ $\boldsymbol{DPD}^{-1}$ je symetrická.
(ii) *Označme $\boldsymbol{H}_\pi$ N-dimenzionální reálný vektorový prostor spojený se skalárním součinem*

$$\langle (x_1, \ldots, x_N)^\top, (y_1, \ldots, y_N)^\top \rangle_\pi := \sum \pi_i x_i y_i,$$

a přidružený k $\boldsymbol{P}$ lineárním zobrazením $\boldsymbol{T_P} : \boldsymbol{H}_\pi \to \boldsymbol{H}_\pi,\ (x_i) \to (\sum_j p_{ij} x_j)_i$. Pak $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze $\Leftrightarrow$ $\boldsymbol{T_P}$ splňuje $\langle \boldsymbol{T_P x}, \boldsymbol{y} \rangle_\pi = \langle \boldsymbol{x}, \boldsymbol{T_P y} \rangle_\pi \forall \boldsymbol{x}, \boldsymbol{y}$, tj. $\Leftrightarrow$ $\boldsymbol{T_P}$ je samoadjungovaný operátor na Hilbertově prostoru $(\boldsymbol{H}_\pi, \langle \cdot, \cdot \rangle_\pi)$.
(iii) *Jestliže $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze, pak matice $(\pi_i p_{ij})_{ij}$ je symetrická. Naopak nechť $(a_{ij})_{i,j=1,\ldots,N}$ je symetrická matice s nezápornými členy taková, že $\sum_{ij} a_{ij} = 1$ a $\pi_i := \sum_j a_{ij} > 0 \quad \forall i$. Pak matice $\boldsymbol{P} = (p_{ij})$, kde $p_{ij} := a_{ij}/\pi_i$ je stochastická se stacionárním rozdělením $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_N)^T$, a $\boldsymbol{P}$ a $\boldsymbol{\pi}$ splňují podmínku detailní rovnováhy. Pokud navíc $\boldsymbol{A}^k > 0$, pak také $\boldsymbol{P}^k > 0$. Tedy existuje tolik matic $\boldsymbol{P}$, které jsou v detailní rovnováze se svým stacionárním rozdělením $\boldsymbol{\pi}$, kolik existuje symetrických nezáporných matic.*

Důkaz. (i) Nejdříve si všimněme, jak vypadají matice $\boldsymbol{DP}$ a $\boldsymbol{DPD}^{-1}$: víme, že $\boldsymbol{D}^{-1}$ je matice taková, že $d_{ii}^{-1} = 1/d_{ii} = 1/\sqrt{\pi_i}$ a ostatní členy má nulové. Pak

$$(dp)_{ij} = \sum_{k \in S} d_{ik} p_{kj} = \sqrt{\pi_i} p_{ij},$$

$$(dpd^{-1})_{ij} = \sum_{k \in S} (dp)_{ik} d_{kj}^{-1} = \sum_{k \in S} \sqrt{\pi_i} p_{ik} d_{kj}^{-1} = \sqrt{\pi_i} p_{ij} \frac{1}{\sqrt{\pi_j}} = \sqrt{\frac{\pi_i}{\pi_j}} p_{ij}. \quad (3.6)$$

Nechť tedy platí podmínka detailní rovnováhy pro $\boldsymbol{P}$ a $\boldsymbol{\pi}$. Pak

$$(\pi_i p_{ij} = \pi_j p_{ji}) \Leftrightarrow \left(\sqrt{\pi_i}\sqrt{\pi_i} p_{ij} = \sqrt{\pi_j}\sqrt{\pi_j} p_{ji}\right) \Leftrightarrow \left(\sqrt{\frac{\pi_i}{\pi_j}} p_{ij} = \sqrt{\frac{\pi_j}{\pi_i}} p_{ji}\right) \ \forall i, j.$$

Dostáváme

$$(dpd^{-1})_{ij} = \sqrt{\frac{\pi_i}{\pi_j}} p_{ij} = \sqrt{\frac{\pi_j}{\pi_i}} p_{ji} = (dpd^{-1})_{ji} \quad \forall i, j \quad , \quad (3.7)$$

tedy matice $\boldsymbol{DPD}^{-1}$ je symetrická. Naopak nechť platí (3.7). Vynásobíme-li tuto rovnici $\sqrt{\pi_i\pi_j}$, dostaneme $\pi_i p_{ij} = \pi_j p_{ji}\ \forall i,j$, neboli podmínku detailní rovnováhy pro $\boldsymbol{P}$ a $\boldsymbol{\pi}$.
(ii) Uvědomme si, jak vypadá $\boldsymbol{T_P x}$ a $\langle \boldsymbol{T_P x}, \boldsymbol{y}\rangle_\pi$:

$$\boldsymbol{T_P x} = \boldsymbol{T_P}(x_1, x_2, \dots, x_N)^\top = (\sum_j p_{1j}x_j, \sum_j p_{2j}x_j, \dots, \sum_j p_{Nj}x_j)^\top,$$

$$\begin{aligned}\langle \boldsymbol{T_P x}, \boldsymbol{y}\rangle_\pi &= \langle (\sum_j p_{1j}x_j, \sum_j p_{2j}x_j, \dots, \sum_j p_{Nj}x_j)^\top, (y_1, y_2, \dots, x_N)^\top\rangle_\pi \\ &= \sum_i \pi_i \sum_j p_{ij}x_j y_i = \sum_j \sum_i \pi_i p_{ij} x_j y_i \ .\end{aligned}$$

Podobně

$$\langle \boldsymbol{x}, \boldsymbol{T_p y}\rangle_\pi = \sum_i \sum_j \pi_i p_{ij} x_i y_j \ .$$

Proveďme malý trik-v předchozí rovnici zaměňme i s j. Vůbec nic se tím nezmění, jelikož sčítáme přes stejné množiny. Tak máme:

$$\langle \boldsymbol{x}, \boldsymbol{T_p y}\rangle_\pi = \sum_j \sum_i \pi_j p_{ji} x_j y_i \ .$$

Tedy zřejmě

$$(\pi_i p_{ij} = \pi_j p_{ji}) \quad \Leftrightarrow \quad \left(\sum_j \sum_i \pi_i p_{ij} x_j y_i = \sum_j \sum_i \pi_j p_{ji} x_j y_i \ \forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{H}_\pi \right),$$

neboli $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze $\Leftrightarrow$ $(\langle \boldsymbol{T_P x}, \boldsymbol{y}\rangle_\pi = \langle \boldsymbol{x}, \boldsymbol{T_p y}\rangle_\pi\ \forall \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{H}_\pi)$
(iii) Nechť $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze. Pak

$$(\pi_i p_{ij})_{ij} = \pi_i p_{ij} = \pi_j p_{ji} = (\pi_j p_{ji})_{ji} \ ,$$

tudíž $\boldsymbol{P}$ je skutečně symetrická.

Naopak nechť $(a_{ij})_{i,j=1,\dots,N}$ je matice taková, že

- $a_{ij} = a_{ji} \quad \forall i,j,$
- $a_{ij} \geq 0 \quad \forall i,j,$
- $\sum_{ij} a_{ij} = 1,$
- $\pi_i := \sum_j a_{ij} > 0 \quad \forall i.$

Pak $\boldsymbol{P} = (p_{ij})$, kde $p_{ij} := a_{ij}/\pi_i$, je stochastická, neboť

$$\sum_j p_{ij} = \sum_j \frac{a_{ij}}{\pi_i} = \frac{1}{\pi_i}\sum_j a_{ij} = \frac{1}{\pi_i}\pi_i = 1 \quad \forall i \ ,$$

má stacionární rozdělení $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_N)^T$, protože

$$\sum_i \pi_i p_{ij} = \sum_i \pi_i \frac{a_{ij}}{\pi_i} = \sum_i a_{ij} = \sum_i a_{ji} = \pi_j \quad \forall j \ ,$$

a $\boldsymbol{P}$ a $\boldsymbol{\pi}$ splňují podmínku detailní rovnováhy:

$$\pi_i p_{ij} = \pi_i \frac{a_{ij}}{\pi_i} = a_{ij} = a_{ji} = \pi_j \frac{a_{ji}}{\pi_j} = \pi_j p_{ji} \quad \forall i, j \ .$$

Pokud navíc $\boldsymbol{A}^k > 0$, pak také $\boldsymbol{P}^k > 0$, neboť jestliže $a_{ij} > 0$ a zároveň $\pi_i > 0$, pak nutně $p_{ij} = a_{ij}/\pi_i > 0$. □

3.3 Odhad používající vlastní čísla

Dostáváme se k samotnému odhadu mezí na vzdálenost mezi $\boldsymbol{P}^k$ a $\boldsymbol{\Pi}$. Zabývali jsme se podmínkou detailní rovnováhy a dozvěděli jsme se, že pro matici $\boldsymbol{P}$ v detailní rovnováze s $\boldsymbol{\pi}$ bude existovat matice $\boldsymbol{S}$ taková, že $\boldsymbol{S}^{-1}\boldsymbol{P}\boldsymbol{S}$ bude diagonální matice a na diagonále najdeme vlastní čísla matice $\boldsymbol{P}$ tak, že $\mid \lambda_1 \mid = 1$ a $\mid \lambda_2 \mid, \mid \lambda_3 \mid, \ldots < 1$ (viz Věty 7 a 12). Proto $(\boldsymbol{S}^{-1}\boldsymbol{P}\boldsymbol{S})^k = \boldsymbol{S}^{-1}\boldsymbol{P}^k\boldsymbol{S}$ a

$$\lim_{k\to\infty} \boldsymbol{P}^k = \boldsymbol{S}\boldsymbol{S}^{-1} \lim_{k\to\infty} \boldsymbol{P}^k \boldsymbol{S}\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{S}'\boldsymbol{S}^{-1}.$$

Vidíme, že rychlost konvergence bude nějakým způsobem záviset na λ^* (– druhém největším vlastním čísle v absolutní hodnotě), což nám potvrdí následující věta. Její důkaz si zde uvádět neuvedeme – je dlouhý a technický, v dalším textu jej využívat nebudeme.

Věta 13 *Nechť je dán nerozložitelný aperiodický markovský řetězec takový, že $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze. Definujme* maximální relativní chybu $\delta(k) := max_{ij} \mid p_{ij}^{(k)} - \pi_j \mid /\pi_j$. *Potom platí:*
(i) $\delta(k) \leq (\lambda^*)^k / min_i \pi_i \quad \forall k,$
(ii) $\delta(k) \geq (\lambda^*)^k$ *pro všechna k sudá. Pokud jsou všechna vlastní čísla nezáporná, pak tato rovnice platí pro všechna k.*

Důkaz. Viz Behrends [1], strana □

Rychlost konvergence je tedy přímo úměrná $(\lambda^*)^k$.

Ukažme si odhad mezí na vzdálenost mezi $\boldsymbol{P}^k$ a $\boldsymbol{\Pi}$ podle Věty 13 na příkladě.

Příklad. Mějme matici

$$\boldsymbol{P} = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Odhadněme, jak rychle konverguje $\boldsymbol{P}$ k $\boldsymbol{\Pi}$. Pro jednoduchost si představme, že chceme odhadnout chybu v desátém kroku.
Stacionární rozdělení této matice je $\boldsymbol{\pi} = (\frac{1}{2}, \frac{1}{2})^\top$. Z Věty 12 vyplývá, že $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze, neboť matice

$$\boldsymbol{DPD}^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

je symetrická.

$$\begin{aligned} &\text{Dostáváme } \lambda^* = 1/3 \text{ a } max_{ij} \mid p_{ij}^{(k)} - \pi_j \mid /\pi_j && \leq 2 * (\tfrac{1}{3})^k, \\ & && \geq (\tfrac{1}{3})^k. \\ &\text{Tedy pro k=10 máme } max_{ij} \mid p_{ij}^{(k)} - \pi_j \mid /\pi_j && < 0.0000340, \\ & && > 0.0000168. \end{aligned}$$

Ve skutečnosti je chyba přibližně 0.0000169, tedy blízko dolní hranici. △

Věta 13 dává velice dobré odhady mezí, ale bohužel má také spoustu omezení. Například abychom určili důležitější horní mez potřebujeme znát stacionární rozdělení, nebo alespoň $min_i\ \pi_i$. V praxi proto používáme různé metody pro odhady $min_i\ \pi_i$.

V mnoha aplikacích však známe $\boldsymbol{\pi}$ explicitně, často se dokonce rovná rovnoměrnému rozdělení. Pak Věta 13 dává nejlepší možný odhad $\delta(k) \leq N(\lambda^*)^k$. Všimněme si, že často chceme určit rychlost konvergence pro markovské řetězce s obrovským počtem prvků stavového prostoru. Pak pro malé k bude $N(\lambda^*)^k$ veliké číslo. Uvědomme si ale, že $N(\lambda^*)^k$ velmi rychle klesá, tedy pokud se λ^* nenachází příliš blízko u 1.

Dalším problémem je velmi silný předpoklad podmínky detailní rovnováhy $\boldsymbol{P}$ a $\boldsymbol{\pi}$. V praxi však velmi často modelujeme markovské řetězce tak, aby tuto podmínku splňovaly.

Největším omezením zůstává, že potřebujeme znát λ^* nebo alespoň nějaký jeho rozumný odhad. Máme mnoho možností, jak toho dosáhnout – přímým výpočtem vlastních čísel, různými odhady, nebo například převedením tohoto problému na problém optimalizační. V další kapitole si také ukážeme metodu odhadu mezí pro λ^* pomocí „vodivosti“, čísla přidruženého k řetězci, které někdy můžeme zjistit snadněji než λ^*.

3.4 Odhady používající členy $\boldsymbol{P}$ přímo I

Jak jsme řekli, z různých důvodů v praxi často nelze předchozí větu použít. Není splňena podmínka detailní rovnováhy, neznáme λ^* nebo $\boldsymbol{\pi}$. Ukažme si proto alternativu, kterou lze aplikovat na libovolný nerozložitelný aperiodický markovský řetězec.

Začněme s nerozložitelnou aperiodickou maticí $\boldsymbol{P}$ a uvažujme

$$m_j^{(k)} := min_i \; p_{ij}^{(k)} \;, \quad M_j(k) := max_i \; p_{ij}^{(k)}$$

– minimum a maximum z j-tého sloupce $\boldsymbol{P}^k$. Protože

$$m_j^{(k+1)} = min_i \; p_{ij}^{(k+1)} = min_i \sum_l p_{il} p_{lj}^{(k)} \geq min_i \sum_l p_{il} m_j^{(k)} = m_j^{(k)} \;,$$

$$M_j^{(k+1)} = max_i \; p_{ij}^{(k+1)} = max_i \sum_l p_{il} p_{lj}^{(k)} \leq max_i \sum_l p_{il} M_j^{(k)} = M_j^{(k)} \;,$$

jednoduše dostáváme

$$m_j^{(1)} \leq m_j^{(2)} \leq \ldots M_j^{(2)} \leq M_j^{(1)}. \tag{3.8}$$

Protože $p_{ij}^{(k)} \stackrel{k\to\infty}{\longrightarrow} \pi_j$, z (3.8) plyne, že $p_{ij}^{(k)}, \pi_j \in [m_j^{(k)}, M_j^{(k)}]$, neboli $\mid \pi_j - p_{ij}^{(k)} \mid \leq M_j^{(k)} - m_j^{(k)}$.

Pokračujme ještě dál: označme $\delta := min_{ij} \; p_{ij}$. Pak z $\sum_j p_{ij} = 1$ plyne, že $p_{ij} \leq 1/N$, neboli $\delta \leq 1/N$. Proto $\tau := 1 - N\delta \in [0,1]$. Dokážeme si, že

$$M_j^{(k+1)} - m_j^{(k+1)} \leq \tau(M_j^{(k)} - m_j^{(k)}), \tag{3.9}$$

což zřejmě implikuje

$$M_j^{(k+1)} - m_j^{(k+1)} \leq \tau^k(M_j^{(1)} - m_j^{(1)}) \;.$$

Dostaneme tedy geometricky rychlou konvergenci $p_{ij}^{(k)} \to \pi_j$ pokud $\tau < 1$.

Začněme tedy s důkazem (3.9). Nechť i_0, j_0 jsou pevné. Označme si $\triangle'$(resp.$\triangle''$) množinu (indexů sloupců) $j \in \{1, 2, \ldots, N\}$ takových, že $p_{i_0 j} \geq p_{j_0 j}$ (resp.$p_{i_0 j} < p_{j_0 j}$) a zestručněme $\sum_{j\in\triangle'}$, $\sum_{j\in\triangle''}$ na $\sum'$, $\sum''$.

Všimněme si, že

- $\sum' (p_{i_0 j} - p_{j_0 j}) + \sum'' (p_{i_0 j} - p_{j_0 j}) = \sum_j (p_{i_0 j} - p_{j_0 j})$ $= \sum_j p_{i_0 j} - \sum_j p_{j_0 j} = 0$, neboli $\sum' (p_{i_0 j} - p_{j_0 j}) = -\sum'' (p_{i_0 j} - p_{j_0 j})$,
- součet libovolných N členů matice $\boldsymbol{P}$ bude větší než N-násobek jejího nejmenšího členu. Proto platí i $\sum'' p_{i_0 j} + \sum' p_{j_0 j} \geq N\delta$,
- $\sum' p_{j_0 j} + \sum'' p_{j_0 j} = 1.$

Proto

$$\sum\nolimits^{'}(p_{i_0 j}-p_{j_0 j}) = -\sum\nolimits^{''}(p_{i_0 j}-p_{j_0 j}) = \sum\nolimits^{''} p_{j_0 j} - \sum\nolimits^{''} p_{i_0 j}$$
$$= 1-\sum\nolimits^{'} p_{j_0 j} - \sum\nolimits^{''} p_{i_0 j} \leq 1 - N\delta = \tau$$

Tedy pro libovolné s platí:

$$\begin{aligned} p_{i_0 s}^{(k+1)} - p_{j_0 s}^{(k+1)} &= \sum_j p_{i_0 j} p_{js}^{(k)} - \sum_j p_{j_0 j} p_{js}^{(k)} = \left(\sum_j p_{i_0 j} - \sum_j p_{j_0 j}\right) p_{js}^{(k)} \\ &\leq \sum\nolimits^{'} (p_{i_0 j}-p_{j_0 j}) M_s^{(k)} + \sum\nolimits^{''} (p_{i_0 j}-p_{j_0 j}) m_s^{(k)} \\ &= \sum\nolimits^{'} (p_{i_0 j}-p_{j_0 j}) M_s^{(k)} - \sum\nolimits^{'} (p_{i_0 j}-p_{j_0 j}) m_s^{(k)} \\ &= \left[\sum\nolimits^{'} (p_{i_0 j}-p_{j_0 j})\right] (M_s^{(k)} - m_s^{(k)}) \leq \tau (M_s^{(k)} - m_s^{(k)}) \ . \end{aligned}$$

□

Věta 14 *Nechť je dán nerozložitelný aperiodický markovský řetězec a nechť* $\exists k_0 : \boldsymbol{P}^{k_0} > 0$. *Definujme* $\delta := min_{ij}\ p_{ij}^{k_0}$ *a* $\tau := 1 - N\delta$. *Potom jestliže*
(i) $k_0 = 1$, *pak* $\mid p_{ij}^{(k)} - \pi_j \mid \leq \tau^k$ *pro* $\forall i, j$ *a* $\forall k$,
(ii) $k_0 > 1$, *pak* $\mid p_{ij}^{(k)} - \pi_j \mid \leq \frac{1}{\tau} \tau^{k/k_0}$ *pro* $\forall i, j$ *a* $\forall k$.
Proto se $p_{ij}^{(k)}$ *blíží k* π_j *geometricky rychle pro* $\forall i$.

Důkaz. (i) je přímým důsledkem předchozích výpočtů, pouze si musíme uvědomit, že $M_j^{(1)} - m_j^{(1)} \leq \tau$ (jestliže $M_j^{(1)} = p_{i_0 j}$, pak $1 = M_j^{(1)} + \sum_{s \neq j} p_{i_0 s} \geq M_j^{(1)} + (N-1)\delta$, takže $M_j^{(1)} - m_j^{(1)} \leq 1 - N\delta$).

K důkazu (ii) potřebujeme znát jeden základní fakt o klesajících posloupnostech: Nechť (c_k) je posloupnost nezáporných členů takových, že pro nějaké pevné k_0 je $\tau < 1$ a pro $\forall k^{'}$ platí $c_{k^{'} k_0} \leq \tau^{k^{'}}$. Pak pro libovolné k můžeme psát $k = k^{'} k_0 + s$, kde $s < k$, a proto

$$c_k \leq c_{k^{'} k_0} \leq \tau^{k^{'}} = (\tau^{1/k_0})^k \tau^{-s/k_0} \leq (\tau^{1/k_0})^k / \tau.$$

Zde toto tvrzení aplikujeme na posloupnost $c_k := M_j^{(k)} - m_j^{(k)}$ – monotonii této posloupnosti jsme dokázali již na začátku odstavce 3.6. Tedy dostáváme $M_j^{(k)} - m_j^{(k)} \leq (\tau^{1/k_0})^k / \tau$
Odhad pak plyne z přechodu od $\boldsymbol{P}$ k $\boldsymbol{P}^{k_0}$ a z (i). □

Srovnáme-li Větu 13 a Větu 14, zjistíme, že každá odhaduje jinou chybu – Věta 13 chybu relativní, tj. $max_{ij} \mid p_{ij}^{(k)} - \pi_j \mid / \pi_j$, Věta 14 chybu absolutní, tj. $\mid p_{ij}^{(k)} - \pi_j \mid$. Dalším rozdílem je, že Věta 13 nám dává velice dobrý odhad mezí na vzdálenost $\boldsymbol{P}$ a $\boldsymbol{\pi}$, což o Větě 14 tvrdit nelze. K lepšímu pochopení důsledků Věty 14 si uvedeme 2 příklady.

Příklad. Mějme markovský řetězec s maticí přechodu

$$\boldsymbol{P} := \frac{1}{6}\begin{pmatrix} 3 & 3 \\ 5 & 1 \end{pmatrix}.$$

Můžeme vybrat $k_0 = 1$. Pak $\delta = \frac{1}{6}$, $\tau = 1 - \frac{2}{6} = \frac{2}{3}$. Stacionární rozdělení je $\boldsymbol{\pi}^T = (\frac{5}{8}, \frac{3}{8})$. Věta 14 říká, že

$$\mid p_{i1}^{(k)} - \frac{5}{8} \mid, \mid p_{i2}^{(k)} - \frac{3}{8} \mid \leq \frac{2}{3}^k. \tag{3.10}$$

Pro ilustraci přesnosti si můžeme explicitně spočítat například

$$\boldsymbol{P}^4 := \frac{1}{81}\begin{pmatrix} 51 & 30 \\ 50 & 31 \end{pmatrix}.$$

Pak
$\mid p_{11}^{(4)} - \pi_1 \mid = \mid \frac{51}{81} - \frac{5}{8} \mid = \frac{1}{216}$, $\qquad \mid p_{21}^{(4)} - \pi_1 \mid = \mid \frac{50}{81} - \frac{5}{8} \mid = \frac{5}{648}$,
$\mid p_{12}^{(4)} - \pi_2 \mid = \mid \frac{30}{81} - \frac{3}{8} \mid = \frac{1}{216}$, $\qquad \mid p_{22}^{(4)} - \pi_2 \mid = \mid \frac{31}{81} - \frac{3}{8} \mid = \frac{5}{648}$,
ale z (3.10) víme pouze, že chyba je menší než $\frac{16}{81}$, což je přibližně $\frac{1}{5}$. $\triangle$

Zkusme získat nějaký lepší odhad. Položme $\delta_{k_0} := min_{ij}\, p_{ij}^{(k_0)}$, $\tau_{k_0} := 1 - N\delta_{k_0}$. Pak v k-tém kroku může být chyba ohraničena číslem $(\tau_{k_0})^{(k/k_0)}$ a pro dvě přípustná k_0, k_1 se $(\tau_{k_0})^{(k/k_0)}, (\tau_{k_1})^{(k/k_1)}$ mohou lišit. Uveďme příklad:

Příklad. Mějme markovský řetězec s maticí přechodu

$$\boldsymbol{P} := \frac{1}{10}\begin{pmatrix} 1 & 9 \\ 5 & 5 \end{pmatrix}.$$

Pak

$$\boldsymbol{P}^2 = \frac{1}{50}\begin{pmatrix} 23 & 27 \\ 15 & 35 \end{pmatrix}.$$

Tedy $\delta_1 = \frac{1}{10}$ a $\delta_2 = \frac{15}{50}$,
$\tau_1 = \frac{4}{5}$ a $\tau_2 = \frac{2}{5}$,
a jelikož $\frac{2}{5} < (\frac{4}{5})^2$ je lepší aplikovat Větu 3 s $k_0 = 2$ než $k_0 = 1$.

Dostáváme $\mid p_{ij}^{(k)} - \pi_j \mid \leq \frac{5}{2}(\sqrt{\frac{2}{5}})^k$. Tedy například pro $k = 10$ máme $\mid p_{ij}^{(k)} - \pi_j \mid < 0.0257$. $\triangle$

Jak jsme viděli v předchozím příkladě, pro každé k_0 můžeme dostat jinak dobrý odhad absolutní chyby. Obvykle zkoušíme najít kompromis mezi vybráním k_0 malého, kdy sice samotné provedení výpočtů nepotrvá dlouho, ale nemusíme dostat dostatečně dobrý odhad, a k_0 velkého, kdy výpočet zabere delší dobu, ale odhad dostaneme lepší.

3.5 Odhady používající členy P přímo II

Ukážeme si druhou aproximaci bez vlastních čísel, založenou na Banachově větě o kontrakci. Pro úplnost zopakujme definice úplného metrického prostoru, kontrakce a pevného bodu, dále pak Banachovu větu o kontrakci a větu o pevném bodu zobrazení, jehož mocninou je kontrakce:

Definice. *Metrickým prostorem* nazýváme dvojici (M, ρ), kde M je neprázdná množina, ρ je funkce z $M \times M$ do $[0, \infty)$ splňující následují 3 axiomy:
(i) $\forall x, y \in M \quad \rho(x, y) = 0 \Leftrightarrow x = y$
(ii) $\forall x, y \in M \quad \rho(x, y) = \rho(x, y)$ (symetrie)
(iii)$\forall x, y, z \in M \quad \rho(x, y) \leq \rho(x, z) + \rho(z, y)$ (trojúhelníková nerovnost)

Definice. Řekneme, že posloupnost $(x_k)_{k\in\mathbb{N}}$ je *cauchyovská* v metrickém prostoru (M, ρ), jestliže pro
$\forall \epsilon\ \exists\ k_0 \in \mathbb{N}$: pro $\forall m, n \geq k_0$ platí $\rho(x_m, x_n) < \epsilon$.

Definice. Řekneme, že metrický prostor (M, ρ) je *úplný*, jestliže každá cauchyovská posloupnost $(x_k)_{k\in\mathbb{N}} \subset M$ konverguje v M.

Definice. Nechť (M, ρ) je úplný metrický prostor. Zobrazení $T : M \to M$ nazveme *kontrakcí na* M, jestliže $\exists L \in [0, 1) : \forall x, y \in M : \rho(Tx, Ty) \leq L\rho(x, y)$.
Řekneme, že $x \in M$ je *pevným bodem zobrazení* T, jestliže $Tx = x$.

Věta 15 (Banachova věta o kontrakci) *Nechť (M, ρ) je úplný metrický prostor a nechť $T : M \to M$ je kontrakce. Potom T má na M právě jeden pevný bod, tj. $\exists!\ x \in M : Tx = x$.*

Věta 16 *Nechť (M, ρ) je úplný metrický prostor, $k \in \mathbb{N}$ a nechť $T : M \to M$ má následující vlastnost: jeho k-tá mocnina $T^k = \underbrace{T \circ T \circ T \circ \cdots \circ T}_{k\text{-krát}}$ je kontrakce na M. Potom T má na M právě jeden pevný bod.*

Všimněme si, že pokud je x_0 pevným bodem zobrazení T, tj. pokud $Tx_0 = x_0$, pak pro libovolné $x \in M$ platí:

$$\rho(T^k x, x_0) = \rho(T^k x, T^k x_0) \leq L^k \rho(x, x_0) \quad \forall k\ . \tag{3.11}$$

Proto se $T^k x$ blíží k x_0 geometricky rychle. Toto tvrzení se budeme snažit aplikovat na konvergenci markovského řetězce ke svému stacionárnímu rozdělení.

Představme si, že máme dán markovský řetězec s maticí přechodu $\boldsymbol{P}$. Úplným metrickým prostorem bude v našem případě dvojice $(M, \|\boldsymbol{x}-\boldsymbol{y}\|_1)$, kde M bude značit množinu $M := \{(x_1, x_2, \ldots, x_N) \mid x_i \geq 0,\ \sum_i x_i = 1\}$ a $\|\boldsymbol{x}\|_1$ pak l^1-normu vektoru $\boldsymbol{x}$ ($\|(x_1, x_2, \ldots, x_N)\|_1 := |x_1| + |x_2| + \ldots + |x_N|$). Zobrazení $T : M \to M$ definujeme jako
$T : (x_1, x_2, \ldots, x_N) \mapsto (\sum_j x_j p_{j1}, \sum_j x_j p_{j2}, \ldots, \sum_j x_j p_{jN})$.

Ukažme si některé vlastnosti matice $\boldsymbol{P}$ a zobrazení T.

Věta 17 *Označme $\boldsymbol{r}_i$ i-tý řádek matice $\boldsymbol{P}$ a položme $C_{\boldsymbol{P}} := \frac{1}{2} max_{ij} \parallel \boldsymbol{r}_i - \boldsymbol{r}_j \parallel_1$. Potom platí:*

(i) l^1-průměr konvexního obalu K všech řádků $\boldsymbol{r}_i$ je $2C_{\boldsymbol{P}}$

(ii) $\parallel T\boldsymbol{x} - T\boldsymbol{y} \parallel_1 \leq C_{\boldsymbol{P}} \parallel \boldsymbol{x} - \boldsymbol{y} \parallel_1$ pro $\forall \boldsymbol{x}, \boldsymbol{y} \in M$

(iii) jestliže $\boldsymbol{x_0}$ je pevným bodem zobrazení T, pak $\parallel T^k\boldsymbol{x} - \boldsymbol{x_0} \parallel_1 \leq 2(C_{\boldsymbol{P}})^k$ pro $\forall \boldsymbol{x}, \forall k$.

Důkaz. (i) Označme C' jako průměr K. Dokážeme si, že $2C_{\boldsymbol{P}} \leq C'$ a $C' \leq 2C_{\boldsymbol{P}}$, z čehož plyne $2C_{\boldsymbol{P}} = C'$.

Konvexní obal vektorů $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$ je nejmenší konvexní množina, která obsahuje vektory $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$. Nutně tedy pro $\forall i, j, i \neq j$ platí, že l^1-vzdálenost $\boldsymbol{r}_i$ a $\boldsymbol{r}_j$ je menší nebo rovná l^1-průměru K. Tedy $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| \leq C'$ pro $\forall i, j$, $i \neq j$, tedy $2C_{\boldsymbol{P}} \leq C'$.

Nechť $\boldsymbol{r}_{j_0}$ je pevné. l^1-vzdálenost vektoru $\boldsymbol{r}_{j_0}$ a libovolného jiného vektoru $\boldsymbol{r}_j$ bude menší než $max_{ij} \parallel \boldsymbol{r}_i - \boldsymbol{r}_j \parallel_1$, tedy menší než $2C_{\boldsymbol{P}}$. Tedy všechna $\boldsymbol{r}_j$ leží v konvexní kouli $B(\boldsymbol{r}_{j_0}, 2C_{\boldsymbol{P}})$ se středem v bodě $\boldsymbol{r}_{j_0}$ a poloměrem $2C_{\boldsymbol{P}}$. Tedy $K \subset B(\boldsymbol{r}_{j_0}, 2C_{\boldsymbol{P}})(-$ z definice $K)$, neboli $\|\boldsymbol{x} - \boldsymbol{r}_{j_0}\|_1 \leq 2C_{\boldsymbol{P}}$ pro $\forall \boldsymbol{x} \in K$.

Nyní zafixujme libovolné $\boldsymbol{x}_0 \in K$. Protože $\boldsymbol{r}_j \in B(\boldsymbol{x}_0, 2C_{\boldsymbol{P}})\ \forall j$, opět díky konvexitě dostáváme $K \subset B(\boldsymbol{x}_0, 2C_{\boldsymbol{P}})$, tj. $\|\boldsymbol{x}_0 - \boldsymbol{y}\| \leq 2C_{\boldsymbol{P}}$ pro $\forall \boldsymbol{y} \in K$, neboli $\|\boldsymbol{x} - \boldsymbol{y}\| \leq C_{\boldsymbol{P}}$ pro $\forall \boldsymbol{x}, \boldsymbol{y} \in K$, neboli $C' \leq 2C_{\boldsymbol{P}}$.
(ii) Nechť $\boldsymbol{x}, \boldsymbol{y} \in M$ jsou pevné. Chceme dokázat, že

$$\left\| \sum_j (x_j - y_j)\boldsymbol{r}_j \right\|_1 \leq C_{\boldsymbol{P}} \|\boldsymbol{x} - \boldsymbol{y}\|_1 .$$

Bez újmy na obecnosti předpokládejme, že $\|\boldsymbol{x} - \boldsymbol{y}\|_1 = 2$. Vektory $\boldsymbol{a} := \sum_j x_j \boldsymbol{r}_j$ a $\boldsymbol{b} := \sum_j y_j \boldsymbol{r}_j$ jsou konvexní kombinací vektorů $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$. Z (i) plyne, že $\boldsymbol{a}, \boldsymbol{b} \in K$. Tedy $2C_{\boldsymbol{P}} \geq \|\boldsymbol{a} - \boldsymbol{b}\|_1 = \left\| \sum_j (x_j - y_j)\boldsymbol{r}_j \right\|_1$
(iii) Víme, že $\parallel \boldsymbol{x} - \boldsymbol{x_0} \parallel_1 \leq 2C_{\boldsymbol{P}}$ pro $\forall \boldsymbol{x} \in M$ a že
$\parallel T\boldsymbol{x} - T\boldsymbol{x_0} \parallel_1 \leq C_{\boldsymbol{P}} \parallel \boldsymbol{x} - \boldsymbol{x_0} \parallel_1$ pro $\forall \boldsymbol{x} \in M$. Tedy z (3.11) plyne

$\| T^k\boldsymbol{x} - \boldsymbol{x_0} \|_1 \leq (C_{\boldsymbol{P}})^k \| \boldsymbol{x} - \boldsymbol{x_0} \|_1 \leq 2(C_{\boldsymbol{P}})^k$. □

Nyní jen zbývá si uvědomit, že pevným bodem zobrazení T je vektor $\boldsymbol{\pi}$. Dokážeme si dvě věty dávající odhady chyby $\sum_j \left| p_{ij}^{(k)} - \pi_j \right|$.

Ve větě 18 budeme existenci stacionárního rozdělení předpokládat (viz Věta 2).

Věta 18 *Nechť je dán nerozložitelný aperiodický markovský řetězec s maticí přechodu $\boldsymbol{P}$ a se stacionárním rozdelením $\boldsymbol{\pi}$. Pak*

$$\sum_j \left| p_{ij}^{(k)} - \pi_j \right| \leq 2(C_P)^k \quad \forall i \in \{1, \ldots, N\} \ \forall k \in \mathbb{N} \ .$$

Důkaz. Toto plyne z (iii) předchozí věty. Aplikujeme ji na $\boldsymbol{x_0} = (\pi_1, \pi_2, \ldots, \pi_N)$ a $\boldsymbol{x} = (0, \ldots, 0, 1, 0, \ldots, 0)$ s „1“ na i-té pozici: pak
$T(0, \ldots, 0, 1, 0, \ldots, 0) = (p_{i1}, p_{i2}, \ldots, p_{iN})$,
$T^k(0, \ldots, 0, 1, 0, \ldots, 0) = (p_{i1}^{(k)}, p_{i2}^{(k)}, \ldots, p_{iN}^{(k)})$
a $\left\| T^k\boldsymbol{x} - \boldsymbol{\pi} \right\|_1 = \sum_j \left| p_{ij}^{(k)} - \pi_j \right|$. □

Ve Větě 19 si znovu dokážeme, že pro aperiodické nerozložitelné markovské řetězce stacionární rozdělení existuje a je jediné. Ukážeme si také druhý možný odhad chyby $\sum_j \left| p_{ij}^{(k)} - \pi_j \right|$.

Věta 19 *Nechť je dán nerozložitelný aperiodický markovský řetězec s maticí přechodu $\boldsymbol{P}$. Zvolme k_0 tak, že $\boldsymbol{P}^{k_0} > 0$. Označme $C := C_{\boldsymbol{P}^{k_0}}$ (=0.5$*$ maximální l^1-vzdálenost mezi řádky $\boldsymbol{P}^{k_0}$). Pak*
(i) $C < 1$, proto díky Větě 17(ii) a Banachově větě o pevném bodu, existuje na množině M jediný pevný bod $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)$ zobrazení T^{k_0}, kde $T : (x_1, x_2, \ldots, x_N) \mapsto (x_1, x_2, \ldots, x_N)\boldsymbol{P}$,
(ii) $\boldsymbol{\pi}$ také splňuje $(\pi_1, \pi_2, \ldots, \pi_N) = (\pi_1, \pi_2, \ldots, \pi_N)\boldsymbol{P}$, tudíž $\boldsymbol{\pi}$ je stacionárním rozdělením. Platí

$$\sum_j \left| p_{ij}^{(k)} - \pi_j \right| \leq \frac{2}{C} C^{k/k_0} \quad \forall i \in \{1, \ldots, N\} \ \forall k \in \mathbb{N} \ .$$

Důkaz. (i) Zřejmě pokud dva vektory $\boldsymbol{x}$, $\boldsymbol{y}$ mají kladné členy, pak $\|\boldsymbol{x} - \boldsymbol{y}\|_1 < \|\boldsymbol{x}\|_1 + \|\boldsymbol{y}\|_1$. Toto znamená, že l^1-vzdálenost každých dvou řádků je menší než 2, neboli $C_{\boldsymbol{P}} < 2$, tudíž $C < 1$.
(ii) Z

$$T^{k_0}(T(\boldsymbol{\pi}^\top)) = T^{k_0+1}(\boldsymbol{\pi}^\top) = T(T^{k_0}(\boldsymbol{\pi}^\top)) = T(\boldsymbol{\pi}^\top)$$

plyne, že $\boldsymbol{\pi}^\top$ i $T(\boldsymbol{\pi}^\top)$ jsou pevnými body zobrazení T^{k_0}. Z Banachovy věty o kontrakci víme, že zobrazení T^{k_0} má právě jeden pevný bod, tudíž $\boldsymbol{\pi}^\top = T(\boldsymbol{\pi}^\top)$. Stejně tak platí $\boldsymbol{\pi}^\top = T^k(\boldsymbol{\pi}^\top)$, neboli $\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top\boldsymbol{P}^k$ pro $k \in \mathbb{N}$.

Předpokládejme, že $\boldsymbol{x} = (0, \ldots, 0, 1, 0, \ldots, 0)$ je vektor s „1“ na i-té pozici a že $k \in \mathbb{N}$. k můžeme zapsat jako $k = rk_0 + k^{'}$, kde $0 \leq k^{'} < k_0$ a $r \in \mathbb{N}_0$. Pak

$$\begin{aligned} \sum_j \left| p_{ij}^{(k)} - \pi_j \right| &= \left\| \boldsymbol{x}\boldsymbol{P}^k - \boldsymbol{\pi}^\top \right\|_1 = \left\| \boldsymbol{x}\boldsymbol{P}^k - \boldsymbol{\pi}^\top \boldsymbol{P}^k \right\|_1 \\ &= \left\| (\boldsymbol{x} - \boldsymbol{\pi}^\top)\boldsymbol{P}^{rk_0 + k^{'}} \right\|_1 \leq \left\| (\boldsymbol{x} - \boldsymbol{\pi}^\top)\boldsymbol{P}^{rk_0} \right\|_1 \\ &\leq C^r \left\| \boldsymbol{x} - \boldsymbol{\pi}^\top \right\|_1 \leq 2C^{(k-k^{'})/k_0}. \end{aligned}$$

□

3.6 Příklady

Příklad. Uvažujme nerozložitelnou stochastickou matici

$$\boldsymbol{P} := \frac{1}{10} \begin{pmatrix} 1 & 1 & 8 \\ 3 & 3 & 4 \\ 5 & 5 & 0 \end{pmatrix} .$$

Použijeme výsledky této kapitoly, abychom rozhodli, jak rychle konvergují řádky $\boldsymbol{P}^k$ k $\boldsymbol{\pi}^\top$. Stacionární rozdělení je v tomto případě $\boldsymbol{\pi}^\top = (\frac{5}{16}, \frac{5}{16}, \frac{3}{8})$.

Například $\pi_1 p_{12} \neq \pi_2 p_{21}$ (neboť $\frac{5}{16}\frac{1}{10} \neq \frac{5}{16}\frac{3}{10}$), tedy není splněna podmínka detailní rovnováhy. Spočítejme si také

$$\boldsymbol{P}^2 := \frac{1}{25} \begin{pmatrix} 11 & 11 & 3 \\ 8 & 8 & 9 \\ 5 & 5 & 15 \end{pmatrix} \boldsymbol{P}^3 := \frac{1}{250} \begin{pmatrix} 59 & 59 & 132 \\ 77 & 77 & 96 \\ 95 & 95 & 60 \end{pmatrix}$$

$$\boldsymbol{P}^4 := \frac{1}{625} \begin{pmatrix} 224 & 224 & 177 \\ 197 & 197 & 231 \\ 170 & 170 & 285 \end{pmatrix} \boldsymbol{P}^5 := \frac{1}{6250} \begin{pmatrix} 1781 & 1781 & 2688 \\ 1943 & 1943 & 2364 \\ 2105 & 2105 & 2040 \end{pmatrix}$$

$$\boldsymbol{P}^{10} := \frac{1}{2765625} \begin{pmatrix} 3072671 & 3072671 & 3620283 \\ 3052988 & 3052988 & 3659649 \\ 3033305 & 3033305 & 3699015 \end{pmatrix}$$

Poznámka. Všechny výpočty byly provedeny v programu Mathematica. V následujícím textu uvedeme zaokrouhlená čísla, abychom mohli lépe porovnávat.

Představme si pro jednoduchost, že chceme odhadnou chybu v desátém kroku.

1. Nejprve aplikujme Větu 14 pro

   - $k_0 = 2$. Pak $\delta = 0.12,\ \tau = 0.64$ a $\left|p_{ij}^{(k)} - \pi_j\right| < 1.563 * 0.8^k$.
     Tedy pro $k = 10$ $\left|p_{ij}^{(k)} - \pi_j\right| < 1.168$.
   - $k_0 = 3$. Pak $\delta \doteq 0.236,\ \tau \doteq 0.292$ a $\left|p_{ij}^{(k)} - \pi_j\right| < 3.425 * 0.664^k$.
     Tedy pro $k = 10$ $\left|p_{ij}^{(k)} - \pi_j\right| < 0.057$.
   - $k_0 = 4$. Pak $\delta \doteq 0.272,\ \tau \doteq 0.184$ a $\left|p_{ij}^{(k)} - \pi_j\right| < 5.435 * 0.655^k$.
     Tedy pro $k = 10$ $\left|p_{ij}^{(k)} - \pi_j\right| < 0.079$.
   - $k_0 = 5$. Pak $\delta \doteq 0.285,\ \tau \doteq 0.145$ a $\left|p_{ij}^{(k)} - \pi_j\right| < 6.891 * 0.680^k$.
     Tedy pro $k = 10$ $\left|p_{ij}^{(k)} - \pi_j\right| < 0.146$.

   Vidíme, že nejlepší odhad dává volba $k_0 = 3$. Pro ilustraci ve skutečnosti platí $\left|p_{ij}^{(k)} - \pi_j\right| < 0.005$.

2. Nyní aplikujme Větu 18. Všimněme si, že odhaduje jinou chybu než Věta 14.

   - $C_{\boldsymbol{P}=\frac{4}{5}}$. Proto $\sum_j \left|p_{ij}^{(k)} - \pi_j\right| \leq 2 * 0.8^k$.
     Tedy pro $k = 10$ máme $\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 0.215$.

   Ve skutečnosti $\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 0.009$.

3. Dále zkusme aplikovat Větu 19 opět pro

   - $k_0 = 2$. Pak $C = 0.48$ a$\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 4.167 * 0.693^k$.
     Tedy pro $k = 10$ máme $\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 0.215$.
   - $k_0 = 3$. Pak $C = 0.288$ a$\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 6.945 * 0.661^k$.
     Tedy pro $k = 10$ máme $\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 0.111$.
   - $k_0 = 4$. Pak $C \doteq 0.173$ a$\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 11.575 * 0.645^k$.
     Tedy pro $k = 10$ máme $\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 0.145$.
   - $k_0 = 5$. Pak $C \doteq 0.103$ a$\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 19.291 * 0.636^k$.
     Tedy pro $k = 10$ máme $\sum_j \left|p_{ij}^{(k)} - \pi_j\right| < 0.209$.

Vidíme, že nejlepší odhad pro $k = 10$ dává volba $k_0 = 4$. Ovšem pokud bychom odhadovali vysoké k, vybrali bychom možnost $k_0 = 5$ Jak jsme již zmínili, ve skutečnosti platí $\sum_j \left| p_{ij}^{(k)} - \pi_j \right| < 0.009$.

$\triangle$

Na straně 20 jsme si řekli, že Věty 13 a 14 odhadují jinou chybu. Přesto se pokusíme jejich výsledky srovnat.

Příklad. Mějme jako v příkladě na straně 21 markovský řetězec s maticí přechodu

$$\boldsymbol{P} := \frac{1}{10} \begin{pmatrix} 1 & 9 \\ 5 & 5 \end{pmatrix}.$$

Zkusme pomocí Věty 13 odhadnout absolutní chybu, tj. $\mid p_{ij}^{(k)} - \pi_j \mid$.

Lze jednoduše dokázat, že matice $\boldsymbol{P}$ je v detailní rovnováze se svým stacionárním rozdělením $\boldsymbol{\pi} = \left(\frac{9}{14}, \frac{5}{14}\right)$. Vlastní čísla matice $\boldsymbol{P}$ jsou $\lambda_1 = 1$, $\lambda_2 = \frac{8}{25}$. Věta 13 říká, že $\mid p_{ij}^{(k)} - \pi_j \mid / \pi_j \leq \frac{14}{5} \left(\frac{8}{25}\right)^k$ pro $i, j \in \{1, 2\}$.

Všimněme si, že jestliže pro $i \in \{1, 2\}$ platí $\frac{|p_{i1}^{(k)} - \pi_1|}{\pi_1} = \frac{|p_{i1}^{(k)} - \frac{9}{14}|}{\frac{9}{14}} \leq a$ a $\frac{|p_{i2}^{(k)} - \pi_2|}{\pi_2} = \frac{|p_{i2}^{(k)} - \frac{5}{14}|}{\frac{5}{14}} \leq a$, pak nutně $\mid p_{ij}^{(k)} - \pi_j \mid \leq \frac{9}{14} a$ pro $i, j \in \{1, 2\}$. Tedy v našem příkladě máme $a = \frac{14}{5} \left(\frac{8}{25}\right)^k$, a tudíž $\mid p_{ij}^{(k)} - \pi_j \mid / \pi_j \leq \frac{9}{5} \left(\frac{8}{25}\right)^k$.

Pro $k = 10$ tedy dostáváme $\mid p_{ij}^{(k)} - \pi_j \mid / < 0,0000203$, tedy mnohem lepší odhad než odhad získaný na straně 20 pomocí Věty 14. $\triangle$

Kapitola 4

Vodivost

4.1 Úvod

V předchozí kapitole jsme si odvodili různé odhady rychlosti konvergence $\boldsymbol{P}^k$ k $\boldsymbol{\Pi}$. Nejlepší výsledky nám dávala Věta 13. Jedním z jejích hlavních omezení byla nutnost znalosti vlastních čísel matice $\boldsymbol{P}$, nebo alespoň λ^* – největšího vlastního čísla matice $\boldsymbol{P}$ v absolutní hodnotě.

V této kapitole se zaměříme na speciální metodu odhadu druhého největšího vlastního čísla λ_2. Najdeme horní a spodní mez pro toto číslo.

Začneme s tím, že se budeme se snažit přiřadit danému markovskému řetězci konstantu – tzv. vodivost. Ukážeme si, že právě vodivost souvisí s konvergencí řetězce ke stacionárnímu rozdělení a že pomocí vodivosti můžeme získat meze pro λ_2. Ukážeme si výpočet vodivosti na jednoduchých příkladech. Zjistíme, že je prakticky nemožné určovat vodivost explicitně – pro markovský řetězec, jehož stavový prostor S čítá N členů, bychom museli provést 2^N výpočtů. Proto se omezíme pouze na speciální markovské řetězce – markovské řetězce definované pomocí grafů. Seznámíme s tzv. metodou kanonických cest, která nám pomůže najít velice dobrý odhad pro hranové zvětšení – konstatu, která úzce souvisí s vodivostí.

Na konec této kapitoly si uvedeme metodu kanonických cest pro obecné markovské řetězce a odhad vodivosti pomocí této metody.

Opět si uvedeme mnoho různých příkladů.

4.2 Vodivost

V této podkapitole bude naším prvním cílem si vodivost zadefinovat, vysvětlit její pravděpodobnostní význam a uvést si alternativní definici. Ukážeme si také její výpočet na konkrétních příkladech. Na konci této podkapitoly pak uvedeme jednu metodu odhadu vodivosti a ukážeme vztah mezi vodivostí a druhým největším vlastním číslem matice $\boldsymbol{P}$.

Představme si, že počátečním rozdělením markovského řetězce je stacionární rozdělení nebo že počáteční rozdělení je libovolné a řetězec „necháme běžet“ dostatečně dlouho na to, aby $\boldsymbol{p}(n) \approx \boldsymbol{\pi}$. Pro libovolnou vlastní podmnožinu T množiny S zkoumejme pravděpodobnost, že nachází-li se nyní řetězec v nějakém stavu množiny T, v dalším kroku přejde do stavu mimo tuto množinu, neboli množinu T opustí. Tedy budeme zkoumat podmíněnou pravděpodobnost

$$\begin{aligned} P((X_{n+1} = j \notin T | X_n = i \in T) \mid \boldsymbol{p}(n) = \boldsymbol{\pi}) \\ &= P\left((X_{n+1} = j \notin T | X_n = i \in T) \mid \boldsymbol{p}(0) = \boldsymbol{\pi}\right) \\ &= \frac{\sum_{i \in T, j \notin T} \pi_i p_{ij}}{\sum_{i \in T} \pi_i}. \end{aligned}$$

Budeme dále hledat takovou množinu T, pro kterou je tato pravděpodobnost co nejmenší. A právě takto nalezená pravděpodobnost, nazývejme ji *vodivostí*, úzce souvisí s rychlostí konvergence markovského řetězce.

Uveďme si formální definici vodivosti. K našim předpokladům si přidáme ještě předpoklad detailní rovnováhy $\boldsymbol{P}$ a $\boldsymbol{\pi}$.

Definice. Nechť $\boldsymbol{P}$ je nerozložitelná stochatická matice, která je v detailní rovnováze se svým stacionární rozdělením $\boldsymbol{\pi}$. Pro $T \subset S$ označme

$C_T := \sum_{i \in T} \pi_i$ *kapacitu množiny T,*

$F_T := \sum_{i \in T, j \notin T} \pi_i p_{ij}$ *ergodický tok z T do $S \setminus T$,*

$\Phi_T := \frac{F_T}{C_T}$.

Pak *vodivost* matice $\boldsymbol{P}$ definujeme jako

$$\Phi := min_{T \subset S}\ max \left\{\Phi_T, \Phi_{S \setminus T}\right\}.$$

Malá hodnota Φ_T znamená, že T je pro řetězec něčím jako pastí – pouze s malou pravděpodobností řetězec „vyskočí“ z T. Očekáváme tedy, že vysoká vodivost bude odpovídat rychlé konvergenci k $\boldsymbol{\pi}$.

Všimněme si, že platí:

- $C_{S \setminus T} = \sum_{i \notin T} \pi_i = 1 - \sum_{i \in T} \pi_i = 1 - C_T$
- $F_T = \sum_{i \in T, j \notin T} \pi_i p_{ij} = \sum_{i \in T, j \notin T} \pi_j p_{ji} = F_{S \setminus T}$ (plyne z podmínky detailní rovnováhy).

Proto

$$\left(\Phi_T = \frac{F_T}{C_T} > \frac{F_T}{1 - C_T} = \Phi_{S\setminus T}\right) \;\Leftrightarrow\; \left(C_T < C_{S\setminus T}\right) \;\Leftrightarrow\; \left(C_T < \frac{1}{2}\right).$$

Proto za stejných předpokladů jako v definici na straně 29 můžeme vodivost definovat také jako

$$\Phi := min\left\{\Phi_T | T \subset S, 0 < C_T < \frac{1}{2}\right\}. \tag{4.1}$$

Ukažme si na konkrétních příkladech výpočet vodivosti.

Příklad. Mějme nerozložitelný aperiodický markovský řetězec s maticí $\boldsymbol{P} = \boldsymbol{W}$, maticí, jejíž řádky jsou shodné.
Stacionární rozdělení se shoduje s každým z těchto řádků a je v detailní rovnováze s $\boldsymbol{P}$. Jsou tedy splněny předpoklady pro výpočet vodivosti. Dostáváme

$$\begin{aligned} F_T &= \textstyle\sum_{i\in T, j\notin T} \pi_i p_{ij} = \sum_{i\in T, j\notin T} \pi_i \pi_j \\ &= \textstyle\sum_{i\in T} \pi_i \sum_{j\notin T} \pi_j = C_T C_{S\setminus T}, \\ \Phi_T &= F_T/C_T = C_T C_{S\setminus T}/C_T = C_{S\setminus T}. \end{aligned}$$

Proto lze vodivost hledat následujícím způsobem:

$$\Phi = min\left\{\textstyle\sum_{i\notin \mathrm{T}} \pi_i \mid T \subset S : 0 < \sum_{i\in \mathrm{T}} \pi_i \leq 1/2\right\}.$$

Všimněme si, že nutně platí $\Phi_T \geq 1/2$. △

Příklad. Mějmě markovský řetězec s maticí přechodu

$$\boldsymbol{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}, \; 0 < a, b < 1.$$

Tento markovský řetězec je zřejmě nerozložitelný aperiodický se stacionárním rozdělením $\boldsymbol{\pi} = \left(\frac{b}{(a+b)}, \frac{a}{(a+b)}\right)^\top$ a lze jednoduše ověřit, že $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze. Dostáváme: $\Phi_{\{0\}} = \pi_0 a/\pi_0 = a$, $\Phi_{\{1\}} = \pi_1 b/\pi_1 = b$. Pak

$$\Phi = \left\{ \begin{array}{ll} min\{a, a\} & \Leftrightarrow a > b \\ min\{b, b\} & \Leftrightarrow b > a \end{array} \right\} = max\{a, b\}.$$ △

Příklad. Uvažujme cyklickou náhodnou procházku, kde s pravděpodobností $0 < p < 1/2$ půjdeme o 1 krok doleva, s pravděpodobností p o 1 krok doprava a s $1-2p$ zůstaneme, kde jsme. Matice přechodu pak má tvar

$$\boldsymbol{P} = \begin{pmatrix} 1-2p & p & 0 & \cdots & 0 & p \\ p & 1-2p & p & 0 & \cdots & 0 \\ 0 & p & 1-2p & p & 0 & \cdots \\ 0 & 0 & p & 1-2p & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ p & 0 & \cdots & 0 & p & 1-2p \end{pmatrix}.$$

Zřejmě $\boldsymbol{P}$ je aperiodická nerozložitelná a stacionárním rozdělením je rovnoměrné rozdělení. Všimněme si, že $\boldsymbol{P}$ je dvojně stochastická a symetrická, tudíž vzhledem k Větě 10 jsou $\boldsymbol{P}$ a $\boldsymbol{\pi}$ v detailní rovnováze. Díky rovnoměrnému rozdělení nám při výpočtu Φ stačí uvažovat ty množiny T, jejichž počet členů nepřesahuje $N/2$, matematicky vyjádřeno $r := card(T) \leq N/2$. Všimněme si, že pro dané r ergodický tok závisí pouze na počtu těch stavů sousedících se stavy z T, které neleží v T. Závisí tedy na tom, jak rozptýlené jsou stavy T neboli jak moc jsou stavy z T shluknuty u sebe. Minimální Φ_T bude zřejmě mít množina tvaru $\{k, k+1, \ldots, k+r-1\}$ $mod\ N$, kde se nachází všechny stavy u sebe. V tomto případě $F_T = 2p/N$, $\Phi = \Phi_T = 2p/r^*$, kde $r^* := [N/2]$ značí největší celé číslo menší nebo rovno $N/2$. △

Příklad. Mějme markovský řetězec s maticí přechodu

$$\boldsymbol{P} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}.$$

Zřejmě daný řetězec je nerozložitelný aperiodický. Všimněme si, že $\boldsymbol{P}$ je dvojně stochatická a symetrická tudíž podle Vět 8 a 10 v detailní rovnováze se svým stacionárním rovnoměrným rozdělením $\boldsymbol{\pi}$. Podle (4.1) stačí při výpočtu vodivosti uvažovat minimum z Φ_T přes množiny T:$C_T \leq \frac{1}{2}$, v našem případě tedy přes jednoprvkové podmnožiny S (neboť stacionárním rozdělením je rovnoměrné rozdělení). Tedy $\Phi = min\left\{\Phi_{\{0\}}, \Phi_{\{1\}}, \Phi_{\{2\}}\right\}$. Protože $p_{11} = 0 = p_{22}$, nutně $\Phi_{\{1\}} = \Phi_{\{2\}} = 1$, tedy $\Phi = \Phi_{\{0\}}$. Zbývá jen provést konkrétní výpočet: $\Phi_{\{0\}} = \frac{\pi_0 p_{01} + \pi_0 p_{02}}{\pi_0} = \frac{\frac{1}{3}\frac{1}{3} + \frac{1}{3}\frac{1}{3}}{\frac{1}{3}} = \frac{2}{9}$. △

Ukázali jsme si metody výpočtu vodivosti v některých speciálních případech. Přesvědčili jsme se, že pro obecný markovský řetězec, dokonce i pro případ malého N, bude velice obtížné vypočítat Φ přesně. Proto si ukážeme jednoduchý odhad dolní meze pro vodivost pro obecný markovský řetězec a v dalších kapitolách pak dolní odhady pro vodivost pro určitý speciální typ markovských řetězců.

Věta 20 *Mějme nerozložitelný aperiodický markovský řetězec s maticí přechodu $\boldsymbol{P}$, která je v detailní rovnováze se stacionárním rozdělením $\boldsymbol{\pi}$. Splňuje-li číslo $\gamma > 0$ podmínku $\pi_j \leq \gamma p_{ij}$ pro $\forall i, j$, pak $\Phi \geq \frac{1}{2\gamma}$.*

Důkaz. Pro libovolnou $T \subset S$ takovou, že $C_T \leq \frac{1}{2}$, platí $C_{S \setminus T} \geq \frac{1}{2}$, tudíž

$$C_T \frac{1}{2} \leq C_T C_{S \setminus T} = \sum_{i \in T, j \notin T} \pi_i \pi_j \leq \sum_{i \in T, j \notin T} \gamma \pi_i p_{ij} = \gamma F_T,$$

neboli $\frac{F_T}{C_T} \geq \frac{1}{2\gamma}$ pro $\forall T \subset S : C_T \leq \frac{1}{2}$, neboli $\Phi \geq \frac{1}{2\gamma}$. □

Všimněme si, že tento odhad vodivosti lze použít pro libovolnou aperiodickou nerozložitelnou matici $\boldsymbol{P}$, která je v detailní rovnováze se stacionárním rozdělením $\boldsymbol{\pi}$. Bohužel, v mnoha případech nedává Věta 20 ani uspokojivé výsledky.

Nyní si uvedeme nejdůležitější větu této kapitoly – větu, která ukazuje vztah mezi vodivostí a druhým největším vlastním číslem matice $\boldsymbol{P}$. Proto budeme schopni pomocí Věty 13 a následující Věty 21 odhadnout rychlost konvergence matice $\boldsymbol{P}$ k matici $\boldsymbol{W}$ i bez znalosti vlastních čísel matice $\boldsymbol{P}$.

Věta 21 *Mějme nerozložitelný aperiodický markovský řetězec s maticí přechodu $\boldsymbol{P}$, která je v detailní rovnováze se stacionárním rozdělením $\boldsymbol{\pi}$. Označme $\lambda_1, \lambda_2, \ldots, \lambda_N$ vlastní čísla matice $\boldsymbol{P}$ tak, aby platilo $1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. Potom platí*

i) $\lambda_2 \leq 1 - \Phi^2/2$

ii) $\lambda_2 \geq 1 - 2\Phi$

Důkaz. Důkaz je technický, proto ho neuvádíme. Je uveden v [1] na straně 93. □

4.3 Markovské řetězce definované pomocí grafů

Cílem této podkapitoly bude odvodit odhady vodivosti pro speciální případy markovských řetězců.

Budeme se zabývat markovskými řetězci definovanými pomocí grafů. Každý vrchol grafu bude představovat jeden stav markovského řetězce. Vrcholy budeme označovat příslušnými stavy řetězce a množinu vrcholů označíme V. Jestliže mezi dva vrcholy umístíme hranu, pravděpodobnost přechodu mezi příslušnými stavy bude kladná. Množinu hran označíme E. Tedy

$[(i,j) \in E] \Leftrightarrow [p_{ij} > 0]$. Dále předpokládejme, že E je symetrická:$[(i,j) \in E] \Leftrightarrow [(j,i) \in E]$. Graf bude představován uspořádanou dvojicí $G := (V,E)$. Vzhledem k předpokladu symetrie se tedy budeme zabývat pouze neorientovanými grafy. Graf zřejmě můžeme snadno vizualizovat : vrcholy zakreslíme jako body a pokud $(i,j) \in E$, pak body i a j spojíme úsečkou.

Zopakujme si dva pojmy z teorie grafů, které budeme potřebovat.

Definice. Řekneme, že graf $G = (V,E)$ je *souvislý*, jestliže pro $\forall i,j \in V, i \neq j$ existuje posloupnost $i_1, \ldots, i_s \in V$ taková, že $i = i_1,\ j = i_s$ a $(i_k, i_{k+1}) \in E$ pro $k = 1, \ldots, s-1$.

Definice. Nechť $G = (V,E)$ je souvislý graf. Pro dvojici vrcholů $i,j \in V$ budeme *cestou z i do j* rozumět posloupnost $i_1, \ldots, i_s \in V$ takovou, že $i = i_1,\ j = i_s$ a $(i_k, i_{k+1}) \in E$ pro $k = 1, \ldots, N-1$.

Pro $e \in E$ řekneme, že cesta z i do j *obsahuje hranu* e, pokud existuje k takové, že $(i_k, i_k + 1) = e$.

Zavedeme speciální třídu markovských řetězců, které budou určeny daným grafem a ještě číslem β splňujícím $0 < \beta \leq 1$.

Nechť $G = (V,E)$ je graf. Pro pozdější použití si označme

$d_i := card\,\{j \mid j \neq i, (i,j) \in E\}$	počet hran spojující vrchol i s jinými vrcholy (*stupeň vrcholu*)
$d := max_{i \in V} d_i$	maximální počet hran spojující nějaký vrchol s jinými vrcholy.

V dalším textu budeme předpokládat, že $d > 0$.

Definice. Nechť $G = (V,E)$ je graf. Definujme *markovský řetězec přidružený ke grafu G* jako řetězec, jehož matice přechodu $\boldsymbol{P}$ má tvar

$$p_{ij} := \begin{cases} \beta/d & \text{pokud } i \neq j \text{ a zároveň } (i,j) \in E \\ 0 & \text{pokud } i \neq j \text{ a zároveň } (i,j) \notin E \\ 1 - d_i\beta/d & \text{pokud } i = j. \end{cases}$$

Věta 22 *Markovský řetězec přidružený ke grafu $G = (V,E)$ je symetrický, jeho stacionárním rozdělením je rovnoměrné rozdělení. Jestliže G je souvislý, pak přidružený řetězec je nerozložitelný. Pokud navíc $\beta < 1$, je tento řetězec nerozložitelný aperiodický a v detailní rovnováze se stacionárním rozdělením.*

Důkaz Řetězec je zřejmě symetrický, neboť leží-li hrana mezi vrcholy i a j, $i \neq j$, pak $p_{ij} = \beta/d = p_{ji}$ a neleží-li hrana mezi vrcholy i a j, $i \neq j$, pak $p_{ij} = 0 = p_{ji}$.

Řetězec je nerozložitelný, neboť z definice souvislosti grafu pro každé dva vrcholy i,j existuje posloupnost $(i = i_1, i_2, \ldots, i_s = j)$ taková, že

$(i_k, i_{k+1}) \in E$, neboli $p_{i_k i_{k+1}} > 0$ pro $k = 1, \dots, s-1$. Proto je pravděpodobnost přechodu mezi stavy i a j po s krocích kladná:
$p^{(s)} \geq p_{i_1 i_2} * p_{i_2 i_3} * \cdots * p_{i_{s-1}, i_s} = (\frac{\beta}{d})^s > 0$, tedy stav j je dosažitelný ze stavu i, tedy každý stav je dosažitelný z libovolného jiného stavu.

Jelikož symetrická matice je speciálním případem dvojně stochastické matice, z Věty 8 plyne, že stacionárním rozdělením je rovnoměrné rozdělení. Pokud $\beta < 1$, pak platí $p_{ii} = 1 - d_i\beta/d > 0$ pro $\forall i \in S = V$, tudíž řetězec je aperiodický. Pak už z Věty 10 přímo dostáváme, že $\boldsymbol{P}$ a $\boldsymbol{\pi}$ jsou v detailní rovnováze. $\square$

Všimněme si, že pro řetězec přidružený souvislému grafu G můžeme Φ_T a tedy i vodivost upravit následujícím způsobem. Upravme nejprve Φ_T a podmínku $0 < C_T \leq 1/2$ na:

$$\Phi_T = \frac{\sum\limits_{i \in T,\ j \notin T} \pi_i p_{ij}}{\sum\limits_{i \in T} \pi_i} = \frac{\frac{1}{card(V)} \sum\limits_{i \in T,\ j \notin T} p_{ij}}{\frac{1}{card(V)} \sum\limits_{i \in T} 1} = \frac{\frac{\beta}{d} \sum\limits_{i \in T,\ j \notin T} 1}{\sum\limits_{i \in T} 1} = \frac{\beta}{d} \frac{card(\{i,j\} | i \in T, j \notin T, (i,j) \in E)}{card(T)},$$

$$[0 < C_T \leq 1/2] \Leftrightarrow \left[0 < \sum_{i \in T} \pi_i \leq 1/2)\right] \Leftrightarrow [(0 < card(T) \leq card(V)/2] \,.$$

Pak $\Phi := min\{\Phi_T | T \subset V,\ 0 < card(T) \leq (V/2)\}$, neboli vodivost bude β/d násobkem čísla $\mu = \min\{\mu_T | T \subset V, 0 < card(T) \leq card(V)/2\}$, kde

$$\mu_T = \frac{card(\{i,j\} \,| i \in T, j \notin T, (i,j) \in E)}{card(T)} \tag{4.2}$$

značí počet hran vedoucích pryč z množiny T dělený počtem bodů množiny T.

Definice. Nechť G je souvislý graf. Pak číslo

$$\mu = \min\{\mu_T | T \subset V, 0 < card(T) \leq card(V)/2\} \,,$$

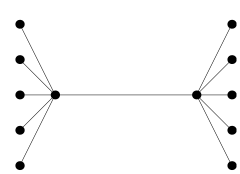
kde

$$\mu_T = \frac{card(\{i,j\} \,| i \in T, j \notin T, (i,j) \in E)}{card(T)},$$

budeme nazývat *hranové zvětšení grafu*.

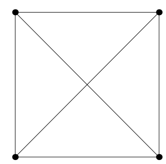
Podívejme se na následující grafy a zkusme zjistit jejich hranové zvětšení.

Příklad.



U tohoto grafu je nejlepší volit T jako množinu všech šesti bodů na levé straně, neboť v tomto případě existuje pouze jedna hrana, krerou se lze dostat z této množiny ven, a ještě toto číslo jedna budeme dělit šesti (počet bodů množiny T), což je největší možné číslo, kterým lze v tomto případě vůbec dělit. Proto ani nemůže existovat žádná jiná množina R s menším μ_R. Tedy $\mu = 1/6$. △

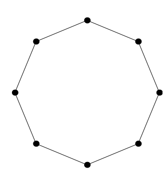
Příklad.



Vzhledem k tvaru tohoto grafu nám stačí porovnat μ_T pro množinu T s jedním či dvěma vrcholy. Dostáváme $\mu_{\{1vrchol\}} = 3/1 = 3$ a $\mu_{\{2vrcholy\}} = 4/2 = 2$. Tedy $\mu = 2$. △

Příklad. V předchozím příkladě jsme spočítali μ pro úplný graf se čtyřmi vrcholy. Podobně bychom postupovali v případě úplného grafu s N vrcholy (grafu, kde každá dvojice vrcholů je spojena hranou). Mezi libovolným vrcholem a všemi ostatními vrcholy leží hrany, tedy nezáleží na tom, které vrcholy přesně do množiny T zařadíme – záleží pouze na jejich počtu. Začněme s množinou T obsahující pouze jeden vrchol, pak $\mu_{\{1 \text{ vrchol}\}} = \frac{N-1}{1}$. Pro T s dvěma vrcholy je $\mu_{\{2 \text{ vrcholy}\}} = \frac{2N-4}{2}$. Všimněme si že pro množinu T s k vrcholy vede z každého vrcholu $k \in T$ $N-k$ vrcholů mimo množinu T. Tedy $\mu_{\{k \text{ vrcholů}\}} = \frac{k*(N-k)}{k} = (N-k)$. Tedy $\mu_{\{k \text{ vrcholů}\}} > \mu_{\{k+1 \text{ vrcholů}\}}$. Proto $\mu = \lceil N/2 \rceil$, kde $\lceil N/2 \rceil$ značí horní celou část čísla $N/2$ △

Příklad.



U tohoto grafu bude zřejmě lepší volit do množiny T vrcholy ležící vedle sebe (pak máme pouze 2 hrany, po kterých lze množinu T opustit). Do množiny T jich budeme chtít zařadit co nejvíc, neboť čím více bodů vedle sebe do T zařadíme, tím bude $\Phi_T = \frac{2}{card(T)}$ menší. Proto $\mu = \frac{2}{[N/2]} = \frac{2}{4} = \frac{1}{2}$.
△

Jak jsme viděli, pro složitější grafy nemusí být ani přímé určení hranového zvětšení grafu snadné,proto si ukážeme, jak je možné hranové zvětšení odhadnout. Ukážeme metodu kanonických cest – metodu odhadování dolní meze pro hranové zvětšení (a tedy i pro vodivost a pomocí Věty i pro λ^*) pro markovské řetězce definované pomocí grafu. Ta je založena na následující větě.

Věta 23 *Nechť G=(V,E) je souvislý graf. Předpokládejme, že pro každé dva stavy s, t máme jednoznačně předepsánu posloupnost ($s = i_1, i_2, \dots, i_{n-1}$, $i_n = t$) – kanonickou cestu z s do t. Dále předpokládejme, že existuje α takové, že pokud $e \in E$, pak existuje nejvýše α dvojic (i, j) různých vrcholů takových, že cesta z i do j obsahuje hranu e. Pak $\mu \geq \frac{card(V)}{2\alpha}$.*

Důkaz. Nejdříve si uvědomme jednu skutečnost o zobrazeních mezi konečnými množinami. Představme si, že máme zobrazení $\Phi : A \to B$. Když budeme vědět, že každý prvek množiny B má maximálně α vzorů (tj. $\Phi^{-1}(\{\mathrm{b}\}) \leq \alpha$), můžeme tvrdit, že počet členů množiny A je maximálně α-násobkem počtu členů množiny B (matematicky zapsáno $\alpha\ card(B) \geq card(A)$). Toto bude stěžejní myšlenkou našeho důkazu.

Vezměme si tedy libovolnou $T \subset V : card(T) \leq card(V)/2$. Označme si $E := \{e_1, e_2, \dots, e_l\}$ množinu hran spojujících body z T s body ležícími mimo T. Dále si položme $A := \{(i, j) | i \in T, j \notin T\}$, $B := \{1, \dots, l\}$.

Definujme zobrazení $\Phi : A \to B$ následovně: mějme prvek (i, j) z množiny A. Pro tento prvek známe kanonickou cestu z i do j. Tato cesta obsahuje minimálně jednu hranu z množiny E a když půjdeme po této cestě, jedna z těchto hran bude poslední. Právě index této hrany přiřadíme prvku (i, j).

Z předpokladů víme, že $\Phi^{-1}(\{\mathrm{b}\}) \leq \alpha$. Počet členů množiny A je $card(T)card(V \setminus T)$. Tedy platí $\alpha l \geq card(T)card(V \setminus T)$. Všimněme si dále, že $card(V \setminus T) \geq card(V)/2$. Proto dostáváme $\frac{l}{card(T)} \geq \frac{card(V)}{2}$ pro každou $T \subset V : card(T) \leq card(V)/2$, tedy i

$$\mu = min\left\{\frac{l}{card(T)} | T \subset V : card(T) \leq card(V)/2\right\} \geq \frac{card(V)}{2}.$$

□

Ukažme si na příkladech použití předchozí věty.

Příklad. Mějme cyklický graf s vrcholy $1, \dots, N$ (viz Příklad na straně 35). Předpokládejme, že kanonické cesty mezi vrcholy budou zadány tímto způsobem: od vrcholu i dojdeme k vrcholu j vždy po směru hodinových ručiček (tedy např. jestliže $N = 8$, pak při přechodu od vrcholu 7 k vrcholu 6 potřebujeme cestu délky 7).

Nyní zafixujme libovolnou hranu-například $e = \{1, 2\}$. Neexistuje žádná cesta z vrcholu 2 do jiného vrcholu, která obsahuje hranu e. Existuje 1 cesta z vrcholu 3 do jiného vrcholu obsahující hranu e (cesta do 2). Z vrcholu 4 máme dvě takovéto cesty. Takto postupujeme dále až do vrcholu 1-z jedničky do jiného vrcholu musíme jít cestou, která vždy obsahuje hranu e, tedy máme $N-1$ těchto cest. Celkem dostáváme $\alpha = \frac{(1+(N-1))}{2}(N-1) = \frac{1}{2}N(N-1)$ cest, obsahujících hranu e. Tedy podle Věty 23 $\mu \geq \frac{1}{N-1}$. Všimněme si, že v Příkladě na straně 35 jsme zjistili, že $\mu = 2/[N/2]$. $\triangle$

Příklad. Mějme stejný graf, jako v předchozím příkladě, ale nyní budou kanonické cesty zadány jiným způsobem: od vrcholu i k vrcholu j půjdeme nejkratší cestou-po směru či proti směru hodinových ručiček, v případě stejně dlouhé cesty oběma směry půjdeme po směru. Pro jednoduchost budeme předpokládat, že N je sudé. Budeme postupovat stejně jako v předchozím případě. Zafixujme e={1,2}. Od 1 k jiným vrcholům máme $N/2$ cest obsahujících hranu e, od N (a od 2) vede $N/2-1$ těchto cest. Takto pokračujeme dál a dostáváme $\alpha = N/2+2(1+2+\cdots+(\frac{N}{2}-1)) = 2*\frac{1}{2}(1+\frac{N}{2}-1)\frac{N}{2} = \frac{N^2}{4}$. Pak $\mu \geq \frac{2}{N}$. Tento odhad je dvakrát lepší, než předchozí. $\triangle$

Příklad. Nechť r je pevné a $V = \{0, 1\}^r$. Mezi dva vrcholy $i = (i_1, \ldots, i_r)$, $j = (j_1, \ldots, j_r)$ umístíme hranu, jestliže se tyto vrcholy budou lišit právě v jedné souřadnici, tj. bude existovat právě jedno i_k, $k \in \{1, \ldots, r\}$ takové, že $i_k \neq j_k$ (proto mezi vrcholy (1111111) a (1011111) leží hrana, mezi (1111111) a (1111100) neleží). Z bodu u do bodu v se dostaneme tak, že souřadnice bodu u budeme postupně z leva do prava měnit na souřadnice bodu v. Tj. například pro $r = 7, u = (0100101), v = (0010111)$ bude kanonická cesta vypadat takto:

$$(0100101) \rightarrow (0000101) \rightarrow (0010101) \rightarrow (0010111)$$

Nyní zafixujme například hranu e ležící mezi
$(i_1, \ldots, i_{s-1}, 0, i_{s+1}, \ldots, i_r)$ a $(i_1, \ldots, i_{s-1}, 1, i_{s+1}, \ldots, i_r)$. Přímo z definice kanonické cesty plyne, že z u do v vede cesta obsahující hranu e právě tehdy, když u je tvaru $(*, \ldots, *, 0, i_{s+1} \ldots, i_r)$, nebo v je tvaru $(i_1, \ldots, i_{s-1}, 1, *, \ldots, *)$, kde $* \in \{0, 1\}$. Tedy máme $\alpha = 2*2^{r-1}$ cest vedoucích přes e (neboť můžeme jít z u do v nebo z v do u). Pak $\mu \geq \frac{2^r}{(2 \cdot 2^{r-1})} = 1$ $\triangle$

Abychom si uvědomili, jak pomocí hranového zvětšení získáme odhad meze na vzdálenost $\boldsymbol{P}$ a $\boldsymbol{\Pi}$ vyřešíme následující příklad.

Příklad. Mějme opět cyklický graf s vrcholy $1, \ldots, N$. Pro jednoduchost předpokládejme, že N je sudé. Pomocí odhadu dolní meze pro hranové zvětšení zjistěme horní mez na vzdálenost $\boldsymbol{P}$ a $\boldsymbol{\Pi}$.

Nejlepší odhad, který jsme získali, je $\mu \geq \frac{2}{N}$. Proto $\phi \geq \frac{2}{N}\frac{\beta}{d} = \frac{2}{N}\frac{\beta}{2} = \frac{\beta}{N}$. $\frac{\beta}{2}$ je pravděpodobnost, že jsme-li v libovolném vrcholu, v dalším kroku půjdeme doleva, nebo pravděpodobnost, že v dalším kroku půjdeme doprava. S pravděpodobností $1-\beta$ zůstaneme, kde jsme.

Pomocí Věty 21 získáme horní mez pro λ_2: $\lambda_2 \leq 1 - (\frac{\beta}{N})^2/2$. Nakonec z Věty 13 plyne, že

$$max_{ij}\frac{\left|p_{ij}^{(k)} - \pi_j\right|}{\pi_j} \leq \frac{\left(1 - \left(\frac{\beta}{N}\right)^2/2\right)^k}{min_i \pi_i} = N\left(1 - \frac{1}{2}\left(\frac{\beta}{N}\right)^2\right)^k.$$

Například pro $N = 8$ a $\beta = \frac{2}{3}$ tedy dostáváme

$$max_{ij}\frac{\left|p_{ij}^{(k)} - \pi_j\right|}{\pi_j} \leq 8\left(1 - \frac{1}{2}\left(\frac{\frac{2}{3}}{8}\right)^2\right)^k = 8\left(\frac{287}{288}\right)^k.$$

$\triangle$

Příklad. Mějme stejný cyklický graf jako v předchozím příkladě. Všimněme si, že v příkladě na straně 35 jsme zjistili, že pro N sudé platí $\mu = \frac{4}{N}$. Pomocí tohoto hranového zvětšení zjistěme meze na vzdálenost $\boldsymbol{P}$ a $\boldsymbol{\Pi}$.

Budeme postupovat podobně jako v předchozím příkladě. Dostáváme $\phi = \frac{4}{N}\frac{\beta}{d} = \frac{4}{N}\frac{\beta}{2} = \frac{2\beta}{N}$.

Pomocí Věty 21 získáme meze pro λ_2:
$$\lambda_2 \leq 1 - \left(\frac{2\beta}{N}\right)^2/2,$$
$$\lambda_2 \geq 1 - 2\frac{2\beta}{N}.$$

Z Věty 13 plyne, že

$$\begin{aligned} max_{ij}\frac{\left|p_{ij}^{(k)} - \pi_j\right|}{\pi_j} &\leq \frac{\left(1 - \left(\frac{2\beta}{N}\right)^2/2\right)^k}{min_i \pi_i} = N\left(1 - 2\left(\frac{\beta}{N}\right)^2\right)^k, \\ max_{ij}\frac{\left|p_{ij}^{(k)} - \pi_j\right|}{\pi_j} &\geq \left(1 - 4\frac{\beta}{N}\right)^k. \end{aligned}$$

Pro $N = 8$ a $\beta = \frac{2}{3}$ tedy dostáváme

$$\begin{aligned} max_{ij}\frac{\left|p_{ij}^{(k)} - \pi_j\right|}{\pi_j} &\leq 8\left(1 - 2\left(\frac{\frac{2}{3}}{8}\right)^2\right)^k = 8\left(\frac{71}{72}\right)^k \\ max_{ij}\frac{\left|p_{ij}^{(k)} - \pi_j\right|}{\pi_j} &\geq \left(1 - 4\frac{\frac{2}{3}}{8}\right)^k = \left(\frac{2}{3}\right)^k. \end{aligned}$$

Všimněme si, že $\frac{71}{72} \doteq 0.986$ a (z předchozího příkladu) $\frac{287}{288} \doteq 0.997$, což se nemusí zdát jako velký rozdíl. Pro $k = 300$ ale dostáváme $\left(\frac{71}{72}\right)^{300} \doteq 0.015$ a $\left(\frac{287}{288}\right)^{300} \doteq 0.352$. $\triangle$

Nalezení dobrého α není jednoduché. Použitím různých kanonických cest při zjišťování odhadu dolní meze pro hranové zvětšení podle Věty 23 můžeme dojít k velmi odlišným výsledkům. Zajímá nás co největší spodní odhad, tedy co největší výsledné μ. I v jednoduchých příkladech si lze přemýšlením nad vhodnou volbou kanonických cest ušetřit mnoho času. Abychom rychle dostali dobrý odhad i ve složitějších případech, museli bychom se blíže věnovat studiu grafů a jejich struktur. Teprve potom bychom dokázali odhadnout kanonické cesty, které povedou alespoň k uspokojivým výsledkům.

4.4 Metoda kanonických cest pro obecné markovské řetězce

V minulé podkapile jsme se seznámili s metodou kanonických cest pro markovské řetězce definované pomocí grafů. Nyní si ukážeme její zobecnění a odhad dolní meze pro vodivost pomocí této zobecněné metody.

Mějme nerozložitelný aperiodický markovský řetězec s maticí přechodu $\boldsymbol{P}$, která je v detailní rovnováze se stacionárním rozdělením $\boldsymbol{\pi}$. Předpokládejme, že pro každé dva stavy s, t máme jednoznačně předepsánu posloupnost $(s = i_1, i_2, \ldots, i_{n-1}, i_n = t)$ – *kanonickou cestu z s do t*. Řekneme, že cesta $(i_1, i_2, \ldots, i_n)$ *obsahuje hranu* (i, j), pokud existuje k takové, že $i_k = i$ a $i_{k+1} = j$.

Předpokládejme dále, že existuje číslo α takové, že pro všechny stavy i, j platí

$$\sum \pi_s \pi_t \leq \alpha \pi_i p_{ij}, \tag{4.3}$$

kde sčítáme přes všechny dvojice (s, t) takové, že cesta z s do t obsahuje hranu (i, j). Dokážeme si, že potom platí

Věta 24 $\Phi \geq \frac{1}{2\alpha}$

Důkaz. Budeme postupovat podobně jako v důkazu Věty 23.

Všimněme si nejdříve, že z (4.3), plyne, že $p_{ij} > 0$ pro $\forall i, j$, neboť pro všechna k platí $\pi_k > 0$ (viz Věta 2).

Vezměme si libovolnou $T \subset V : C_T \leq 1/2$. Budeme mít opět zobrazení $\Phi : A \to B$, kde tentokrát A bude značit množinu $\{\pi_i \pi_j | i \in T, j \notin T\}$, B množinu $\{(i, j) | i \in T, j \notin T\}$. Pro prvek $\pi_a \pi_b$ z množiny A máme opět kanonickou cestu z a do b. Tato cesta obsahuje minimálně jednu hranu z množiny

B a když půjdeme po této cestě, jedna z těchto hran bude poslední. Právě tuto hranu (i, j) přiřadíme prvku $\pi_a\pi_b$.

Uvědomme si, že $\pi_s\pi_t$ si můžeme představit jako určitou váhu dvojice vrcholů s, t. (4.3) říká, že součet vah dvojic vrcholů s, t, takových, že kanonická cesta z s do t obsahuje hranu (i, j) není větší než $\alpha\pi_i p_{ij}$, neboli součet vzorů hrany $(i, j) \in B$ není větší než $\alpha\pi_i p_{ij}$.

Podobnou úvahou jako v důkazu Věty 23 bychom došli k tomu, že sečteme-li dohromady součet vzorů každého z bodů množiny B, dostaneme větší hodnotu, než kdybychom sečetli prvky množiny A. Tedy

$$\sum_{i \in T, j \notin T} \sum_{*} \pi_s\pi_t \geq \sum_{a \in T, b \notin T} \pi_a\pi_b, \tag{4.4}$$

kde $*$ značí sčítání přes všechny dvojice s, t, takové, že cesta z $s \in T$ do $t \notin T$ obsahuje hranu (i, j). Z předpokladů a další úpravou (4.4) dostáváme

$$\alpha F_T = \sum_{i \in T, j \notin T} \alpha\pi_i p_{ij} \geq \sum_{i \in T, j \notin T} \sum_{*} \pi_s\pi_t \geq \sum_{a \in T, b \notin T} \pi_a\pi_b \geq C_T C_{S \setminus T} \geq \frac{1}{2} C_T.$$

Tedy $\frac{F_T}{C_T} \geq \frac{1}{2\alpha}$ pro $\forall T \subset V : C_T \leq \frac{1}{2}$, neboli $\Phi \geq \frac{1}{2\alpha}$. $\square$

Literatura

[1] Behrends E.: *Introduction to Markov Chains with Special Emphasis on Rapid Mixing*, Friedrick Vieweg & Son. Braunschweig, Wiesbaden, 2000.

[2] Bican L.: *Lineární algebra a geometrie*, Academia, Praha, 2002.

[3] Matoušek J., Nešetřil J. E.: *Kapitoly z diskrétní matematiky*, Karolinum, Praha, 2003.

[4] Prášková Z., Lachout P.: *Základy náhodných procesů*, Karolinum, Praha, 2005.

[5] Sinclair A.: *Algorithms for Random Generating and Counting*, Birkhäuser, Boston, 1993.

[6] Winkler G.: *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, Springer-Verlag, Berlin-Heidelberg-New York, 1995.