



**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

BAKALÁŘSKÁ PRÁCE

Jakub Ditrich

Rychlost konvergence Markovových řetězců – dolní meze pro mixing

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Michaela Prokešová, Ph.D.

Studijní program: Matematika

Studijní obor: Obecná matematika

Praha 2022

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Chtěl bych poděkovat RNDr. Michaele Prokešové, Ph.D. za veškeré rady, pomoc a opravy, které mi vlídně poskytla. Nebýt jejího vedení, tato práce neskončí ve zdárném cíli. Také děkuji všem svým blízkým, kteří si vyslechli mé lamentování když jsem nevěděl, kam se hnout, a motivovali mne pokračovat.

Název práce: Rychlost konvergence Markovových řetězců – dolní meze pro mixing

Autor: Jakub Ditrich

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Michaela Prokešová, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: V práci se zabýváme rychlostí konvergence ireducibilních a aperiodických homogenních markovských řetězců s konečnou diskrétní množinou stavů. Přesněji dolními odhady času potřebného k dostatečnému přiblížení marginálního rozdělení řetězce ke stacionárnímu rozdělení, takzvaným časem mixingů. Ukážeme různé metody odvození odhadů, patřičně je namotivujeme, zformulujeme a dokážeme. Nakonec ukážeme jejich použití na vhodných příkladech. Výsledkem je přehledný soupis tří metod, které se za účelem získání dolní meze dají použít.

Klíčová slova: markovský řetězec, rychlost konvergence, dolní mez pro čas mixingů

Title: Mixing of Markov chains – lower bounds for mixing

Author: Jakub Ditrich

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Michaela Prokešová, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The focus of the thesis is the convergence of irreducible aperiodic homogeneous Markov chains with a finite and discrete set of states. Specifically, lower bounds on the time needed for the chain's marginal probability distribution to be sufficiently close to the stationary distribution, so called mixing time. Multiple methods are introduced, properly motivated and proven. Finally, each method is demonstrated on a suitable example. The result is an overview of three methods that can be used to derive lower bounds for mixing time.

Keywords: Markov chain, speed of convergence, lower bound for mixing time

Obsah

Značení	2
Úvod	3
1 Základní definice	4
2 Konvergence ke stacionárnímu rozdělení	6
3 Dolní meze pro mixing	9
3.1 Počítací mez	10
3.2 Míra zúžení	13
3.3 Rozlišující statistiky	16
Seznam použité literatury	24

Značení

t_{mix}	Čas mixingu
S^c	Doplňěk množiny S
$Q(A,B)$	Hranová míra mezi množinami A a B
$\mathbb{1}$	Indikátor jevu
$\mathbf{1}$	Jednotkový vektor
Cov_ν	Kovariance počítaná vzhledem k rozdělení ν
\mathbf{Cov}_ν	Kovariance za podmínky, že řetězec má počáteční rozdělení ν
P	Matice přechodu markovského řetězce
$\Phi(S)$	Míra zúžení množiny S
Φ_\star	Míra zúžení markovského řetězce
\mathbb{Z}^+	Množina kladných celých čísel
\mathcal{X}	Množina stavů markovského řetězce
(Ω, \mathcal{A})	Měřitelný prostor
X	Náhodná veličina
ν_f	Obraz míry ν při zobrazení f
$\mathbf{P}(A B)$	Podmíněná pravděpodobnost jevu A za podmínky B
\mathbf{P}	Pravděpodobnost
\mathbf{P}_ν	Pravděpodobnost pokud řetězec má počáteční rozdělení ν
\mathbf{P}_x	Pravděpodobnost pokud řetězec začínal ve stavu x
$P(x,y)$	Pravděpodobnost přechodu ze stavu x do stavu y
$P^t(x,y)$	Pravděpodobnost přechodu ze stavu x do stavu y v t krocích
$P^t(x,\cdot)$	Pravděpodobnostní rozdělení na \mathcal{X} po t krocích ze stavu x
e_x	Řádkový vektor délky $ \mathcal{X} $ s 1 na pozici odpovídající x a 0 jinde
Var	Rozptyl
Var_ν	Rozptyl počítaný vzhledem k rozdělení ν
\mathbf{Var}_ν	Rozptyl za podmínky, že řetězec má počáteční rozdělení ν
π	Stacionární rozdělení markovského řetězce
Ω	Stavový prostor
\mathbf{E}	Střední hodnota
\mathbf{E}_ν	Střední hodnota počítaná vzhledem k rozdělení ν
\mathbf{E}_ν	Střední hodnota za podmínky, že řetězec má počáteční rozdělení ν
$ A $	Velikost množiny A
$\ \mu - \nu\ _{TV}$	Vzdálenost měř v totální variaci
$d(t)$	Vzdálenost marginálního rozdělení od stacionárního v čase t

Úvod

Markovské řetězce jsou speciálním případem náhodných procesů, které, díky svým restriktivním, však ne-nereálným předpokladům, nacházejí mnohá uplatnění v každodenních aplikacích. Jednou z nejčastějších otázek, jež tyto aplikace pokládají, je: zda-li a jak rychle konvergují k jistému rovnovážnému stavu.

V této práci nejprve určíme za jakých podmínek bude řetězec konvergovat, a poté se budeme zabývat dolními odhady na rychlost konvergence, tedy kolik nejméně kroků je třeba uskutečnit, aby celkové pravděpodobnostní rozdělení na stavovém prostoru bylo „dostatečně“ blízko ke stacionárnímu. Tento stav budeme nazývat *mixing* a budeme tedy hovořit o *času mixing*.

V prvních kapitolách si připomeneme základní značení a vlastnosti homogenních markovských řetězců s diskrétním stavovým prostorem. Poté připomeneme co je to stacionární rozdělení na množině stavů, ukážeme za jakých podmínek k němu řetězec konverguje a určíme rychlost této konvergence.

Ve zbytku práce se soustředíme na různé způsoby určení dolních odhadů pro čas mixing a doplníme je o názorné příklady ilustrující jejich využití. Jmenovitě se podíváme na jednoduchou počítací metodu odhadu, poté odhad pomocí takzvaného *bottleneck ratio* (budeme překládat jako *hranová míra*), a finálně se podíváme na metodu využívající rozlišujících statistik. Každou metodu následuje příklad, kde se ukáže její použití.

Celá práce se opírá o poznatky z knihy od Levina a Perese (Levin a Peres, 2017), přesněji o kapitoly 1 a 3 pro úvod do markovských řetězců, a poté o kapitolu 7, která se zabývá právě dolními odhady pro čas mixing. Hlavním přínosem autora je doplnění jednotlivých kroků v důkazech a vysvětleních, samostatná aplikace představených metod na vhodně zvolené příklady a celkové zpřístupnění textu pro potenciální zájemce o tuto problematiku.

1. Základní definice

Úvodem si uvedeme základní definice a věty, které budeme k další práci potřebovat. Obsah této kapitoly by měl být čtenáři relativně známý, jejím účelem je hlavně zavést používané značení a stanovit předpokládané čtenářovy znalosti, ze kterých budeme vycházet.

V této práci se zabýváme pouze *náhodnými procesy s diskrétním časem* a to odráží i zde použitá definice.

Definice 1 (Náhodný proces). *Nechť $(\Omega, \mathcal{A}, \mathbf{P})$ je pravděpodobnostní prostor a pro každé $t \in \mathbb{N}$ je dána náhodná veličina $X_t : (\Omega, \mathcal{A}) \rightarrow (H, \mathcal{H})$, kde (H, \mathcal{H}) je měřitelný prostor. Pak $\{X_t; t \in \mathbb{N}\}$ nazveme náhodným procesem.*

Místo celého stavového prostoru se budeme většinou zabývat pouze jeho malou podmnožinou, a to nejmenší měřitelnou množinou $\mathcal{X} \subseteq H$ takovou, že $\mathbf{P}(X_t \in \mathcal{X}) = 1$ pro všechna $t \in \mathbb{N}$, nazveme ji *množina stavů*. Zde budeme pro potřeby naší práce nadále uvažovat \mathcal{X} pouze konečnou.

Tento úvod nás plynule vede ke stěžejní definici této práce a to definici markovského řetězce:

Definice 2 (Markovský řetězec). *Náhodný proces $\{X_t; t \in \mathbb{N}\}$ s množinou stavů \mathcal{X} nazveme markovským řetězcem, pokud $\forall x, y \in \mathcal{X}, t \geq 1$ a pro každý jev $M_{t-1} = \bigcap_{s=0}^{t-1} \{X_s = i_s\}, i_0, \dots, i_{t-1} \in \mathcal{X}$ takový, že $\mathbf{P}(M_{t-1} \cap \{X_t = x\}) \neq 0$ platí:*

$$\mathbf{P}(X_{t+1} = y | M_{t-1} \cap \{X_t = x\}) = \mathbf{P}(X_{t+1} = y | X_t = x)$$

V práci se budeme omezovat pouze na řetězce, kde $\mathbf{P}(X_{t+1} = y | X_t = x)$ nezávisí na hodnotě t , takzvané *homogenní markovské řetězce*. Pro zjednodušení tedy budeme značit $P(x, y) = \mathbf{P}(X_{t+1} = y | X_t = x)$ a budeme tento koncept nazývat *pravděpodobnost přechodu ze stavu x do stavu y* .

Vzhledem k tomu, že uvažujeme \mathcal{X} konečné, můžeme tyto pravděpodobnosti přechodu uspořádat do čtvercové matice rozměru $|\mathcal{X}| \times |\mathcal{X}|$, kde prvek na pozici (x, y) odpovídá pravděpodobnosti přechodu z x do y . Matici budeme značit P a nazveme ji *matice přechodu*. Její základní vlastnosti uvedeme v následujícím lemmatu bez důkazu.

Lemma 1.1. *Matice přechodu markovského řetězce má nezáporné prvky a její řádkové součty jsou rovny jedné, tedy:*

$$\sum_{y \in \mathcal{X}} P(x, y) = 1, \forall x \in \mathcal{X}.$$

Maticím s touto vlastností budeme říkat stochastické. Dále platí:

$$\mathbf{P}(X_{t+2} = y | X_t = x) = \sum_{z \in \mathcal{X}} P(x, z)P(z, y) = e_x P P e_y^T, \quad y, x \in \mathcal{X}, t \geq 1,$$

kde e_x značí řádkový vektor s jednotkou na pozici odpovídající x a nulami jinde

Z lemmatu máme, že časový posun v řetězci o více kroků můžeme reprezentovat adekvátní maticovou mocninou matice přechodu. Řádky těchto matic nám

přirozeně definují pravděpodobnostní rozdělení $P(x, \cdot)$ na \mathcal{X} . V případě více kroků budeme psát $P^t(x, \cdot)$.

Pravděpodobnostní rozdělení vektoru X_0 budeme nazývat *počáteční rozdělení* a budeme jej (stejně jako všechna další pravděpodobnostní rozdělení na \mathcal{X}) reprezentovat řádkovým vektorem o délce $|\mathcal{X}|$. Pravděpodobnost nějakého jevu, je-li μ počátečním rozdělením, budeme značit $\mathbf{P}_\mu(\cdot)$. V případě, že $\mu(x) = 1$ pro nějaký stav $x \in \mathcal{X}$, a my tedy víme, v jakém stavu řetězec začal, píšeme jednoduše $\mathbf{P}_x(\cdot)$.

Odteď budeme uvažovat pouze *aperiodické a ireducibilní¹ homogenní markovské řetězce s diskrétním časem a konečnou množinou stavů* a budeme je nazývat markovskými řetězci.

Pokračujeme dalším klíčovým pojmem a důležitým tvrzením:

Definice 3 (Stacionární rozdělení markovského řetězce). *Pravděpodobnostní rozdělení π na množině stavů \mathcal{X} markovského řetězce s maticí přechodu P nazveme stacionárním, pokud splňuje:*

$$\pi P = \pi$$

Tvrzení 1.2 (Existence a jednoznačnost stacionárního rozdělení). *Každý aperiodický a ireducibilní homogenní markovský řetězec s diskrétním časem a konečnou množinou stavů má stacionární rozdělení a to je jednoznačně určené.*

Důkaz. Levin a Peres (2017, kapitola 1.5, rozloženo mezi více tvrzení) □

Dále zde zavedeme značení, které použijeme později v práci.

Značení. *Nechť ν je pravděpodobnostní rozdělení na \mathcal{X} mějme měřitelnou funkci $f : \mathcal{X} \rightarrow \Theta$. Potom obraz míry ν budeme značit následovně:*

$$\nu_f(A) := \nu(f^{-1}(A)), \quad A \in \mathcal{A},$$

kde \mathcal{A} je σ -algebra na Θ .

Dále, je-li $\Theta \subseteq \mathbb{R}$, E_ν bude značit střední hodnotu počítanou vzhledem k rozdělení ν , tedy:

$$E_\nu f(X) := \sum_{x \in \mathcal{X}} f(x) \nu(x).$$

Oproti tomu \mathbf{E}_x bude značit střední hodnotu za podmínky, že řetězec začínal s rozdělením ν , tedy specificky pro $x \in \mathcal{X}$:

$$\mathbf{E}_x f(X_t) = \mathbf{E}[f(X_t) | X_0 = x].$$

Obdobně i Var_ν a Cov_ν značí rozptyl a kovarianci počítané vzhledem ke rozdělení ν a \mathbf{Var}_x a \mathbf{Cov}_x značí rozptyl a kovarianci za podmínky, že řetězec začal s rozdělením ν .

Kapitolu zakončíme formulací obecně známé Čebyševovy nerovnosti.

Tvrzení 1.3 (Čebyševova nerovnost). *Nechť má reálná náhodná veličina X konečný druhý moment. Potom pro $a > 0$ platí:*

$$\mathbf{P}(|X - \mathbf{E} X| \geq a) \leq \frac{\text{Var}(X)}{a^2}.$$

¹Definice těchto pojmů lze nalézt v libovolné základní učebnici o markovských řetězcích například v Prášková a Lachout (1998, kapitola 2).

2. Konvergence ke stacionárnímu rozdělení

Teď již máme základní vlastnosti námi studovaných markovských řetězců a víme, že každý má své stacionární rozdělení. V této kapitole se budeme zabývat otázkou konvergence markovských řetězců k jejich stacionárnímu rozdělení.

Ke konvergenci ale nejprve potřebujeme nějakou metriku, podle které ji budeme určovat. V našem případě půjde o metriku mezi pravděpodobnostními rozděleními na společné množině stavů.

Definice 4 (Vzdálenost v totální variaci). *Vzdálenost v totální variaci dvou pravděpodobnostních rozdělení μ a ν na konečné množině stavů \mathcal{X} definujeme následovně:*

$$\|\mu - \nu\|_{TV} = \max_{A \subseteq \mathcal{X}} |\mu(A) - \nu(A)|$$

Vzdálenost dvou rozdělení tedy určíme jako největší rozdíl v pravděpodobnostech, které jednotlivá rozdělení přiřadí stejnému jevu. Důkaz, že jde o metriku vynecháme.

V následujícím tvrzení ukážeme alternativní vyjádření vzdálenosti v totální variaci, které se nám bude hodit při důkazu konvergence.

Tvrzení 2.1. *Nechť μ a ν jsou pravděpodobnostní rozdělení na \mathcal{X} . Potom platí:*

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mu(x) - \nu(x)|$$

Důkaz. Levin a Peres (2017, strana 48, Proposition 4.2). □

Dále budeme potřebovat následující lemma, jehož důkaz je přímočarý výpočet, využívající lemma 1.1.

Lemma 2.2. *Nechť π je pravděpodobnostní rozdělení a Π je stochastická matice s řádky rovnými π . Potom pro každou stochastickou matici S platí: $S\Pi = \Pi$ a pro každou matici M takovou, že $\pi M = \pi$ platí: $\Pi M = \Pi$.*

Ještě bez důkazu zformulujeme tvrzení teorie čísel, které také budeme potřebovat.

Tvrzení 2.3. *Je-li $S \subset \mathbb{Z}^+$ uzavřená vzhledem ke sčítání a její největší společný dělitel je 1, pak existuje m_S takové, že všechna $m \geq m_S$ náležejí do S .*

Důkaz. Levin a Peres (2017, strana 19, lemma 1.30) jako důkaz obecnějšího tvrzení, jehož toto je přímočará aplikace. □

Tvrzení 2.3 hned využijeme v následujícím pomocném tvrzení, které pro přehlednost vyjmeme z důkazu věty o konvergenci.

Tvrzení 2.4. *Pro aperiodický a ireducibilní markovský řetězec s konečnou množinou stavů \mathcal{X} existuje $k \in \mathbb{N}$ takové, že matice přechodu P^k má všechny prvky kladné.*

Důkaz. Označme $\Gamma(x)$ množinu všech t pro které je $P^t(x,x) > 0$. Potom pro $s, t \in \Gamma(x)$ platí:

$$P^{t+s}(x,x) \geq P^t(x,x)P^s(x,x) > 0.$$

Tedy $t + s \in \Gamma(x)$ a vidíme, že $\Gamma(x)$ je uzavřená vzhledem ke sčítání. Z aperiodicity řetězce je největší společný dělitel $\Gamma(x)$ roven 1. Dle tvrzení 2.3 tedy pro každé $x \in \mathcal{X}$ existuje $t(x)$ takové, že pro $t \geq t(x)$ je $P^t(x,x) > 0$.

Zvolme teď $x \in \mathcal{X}$ pevné. Z nerozložitelnosti existuje pro každé $y \in \mathcal{X}$ nějaké $r(x,y)$ takové, že $P^{r(x,y)}(x,y) > 0$. Označme dále $r(x) = \max_{y \in \mathcal{X} \setminus \{x\}} r(x,y)$, potom pro $t \geq t(x) + r(x)$ platí:

$$P^t(x,y) \geq P^{t-r(x,y)}(x,x)P^{r(x,y)}(x,y) > 0, \quad \forall y \in \mathcal{X}. \quad (2.1)$$

Takto máme $t(x)$ a $r(x)$ pro každé $x \in \mathcal{X}$.

Označme $k = \max_{x \in \mathcal{X}} (t(x) + r(x))$, potom, nahrazením k za t ve (2.1) pro libovolné $x \in \mathcal{X}$, máme $P^k(x,y) > 0$ pro všechna $x, y \in \mathcal{X}$. □

Teď jsme připraveni zformulovat a dokázat stěžejní větu, na které je tato práce založena.

Věta 2.5 (O konvergenci). *Nechť P je matice přechodu aperiodického a ireducibilního markovského řetězce s konečnou množinou stavů \mathcal{X} a se stacionárním rozdělením π , pak existují konstanty $\alpha \in (0,1)$ a $C > 0$ takové, že:*

$$\max_{x \in \mathcal{X}} \|P^t(x, \cdot) - \pi\|_{TV} \leq C\alpha^t.$$

Důkaz. Dle tvrzení 2.4 nalezneme $k \in \mathbb{N}$ tak, aby matice P^k měla pouze kladné prvky. Dále označme Π matici s řádky rovnými π a nalezněme $\delta > 0$ takové, že:

$$\delta\pi(y) < P^k(x,y); \quad \forall x, y \in \mathcal{X}.$$

Poté lze nalézt stochastickou matici Q takovou, že:

$$P^k = \delta\Pi + (1 - \delta)Q. \quad (2.2)$$

Pro zjednodušení zápisu budeme matici P^k značit R . Dále položme $\theta = (1 - \delta)$.

Teď indukci ukážeme, že pro $n \in \mathbb{N}$ platí:

$$R^n = P^{kn} = (1 - \theta^n)\Pi + \theta^n Q^n. \quad (2.3)$$

Pro $n = 1$ jsme rovnost ukázali výše, předpokládejme tedy, že platí pro n a počítejme:

$$\begin{aligned} R^{n+1} &= R^n R = ((1 - \theta^n)\Pi + \theta^n Q^n) R \\ &= (1 - \theta^n)\Pi R + \theta^n Q^n ((1 - \theta)\Pi + \theta Q) \\ &= (1 - \theta^n)\Pi R + \theta^n (1 - \theta) Q^n \Pi + \theta^{n+1} Q^{n+1}, \end{aligned}$$

kde ve druhé rovnosti využíváme indukčního předpokladu a ve druhém členu ve třetí rovnosti vyžíváme (2.2).

Dále využijeme lemma 2.2, neb $\pi R = \pi P^k = \pi$, a proto $\Pi R = \Pi$ a také platí: $Q^n \Pi = Q^{n-1} \Pi = \dots = \Pi$ protože Q je stochastickou maticí. Máme tak:

$$\begin{aligned} R^{n+1} &= (1 - \theta^n) \Pi + \theta^n (1 - \theta) \Pi + \theta^{n+1} Q^{n+1} \\ &= \Pi - \theta^n \Pi + \theta^n \Pi - \theta^{n+1} \Pi + \theta^{n+1} Q^{n+1} \\ &= (1 - \theta^{n+1}) \Pi + \theta^{n+1} Q^{n+1}. \end{aligned}$$

Ukázali jsme tak kýženou rovnost (2.3).

Dále pro $r \in \{1, \dots, k-1\}$ platí:

$$P^{nk+r} = ((1 - \theta^n) \Pi + \theta^n Q^n) P^r = (1 - \theta^n) \Pi + \theta^n Q^n P^r = \Pi - \theta^n \Pi + \theta^n Q^n P^r.$$

Odečtením Π tak máme:

$$P^{nk+r} - \Pi = \theta^n (Q^n P^r - \Pi).$$

Podívejme se na řádkové součty na obou stranách. Na levé straně je v každém řádku, dle tvrzení 2.1, dvojnásobek vzdálenosti v totální variaci pro jednotlivá počáteční $x \in \mathcal{X}$. Dvojnásobek řádkových součtů uvnitř závorčky na pravé straně lze shora odhadnout maximální vzdáleností v totální variaci, což je 1.

Dostáváme tak pro každé $x \in \mathcal{X}$:

$$\|P^{nk+r}(x, \cdot) - \pi\|_{TV} \leq \theta^n = \theta^{n+1-1} = \theta^{\frac{nk+k}{k}-1} \leq \theta^{\frac{nk+r}{k}-1},$$

kde poslední nerovnost platí, neboť je $r < k$ a $\theta < 1$.

Pro zakončení důkazu nám stačí označit $nk + r = t$, $\theta^{1/k} = \alpha$, $\theta^{-1} = C$ a vzít na levé straně maximum přes všechna $x \in \mathcal{X}$. □

Máme teď horní odhad na vzdálenost marginálního rozdělení X_t od stacionárního rozdělení, který exponenciálně klesá s přibývajícím posuny v řetězci. Tedy jsme ukázali, že každý markovský řetězec s libovolným počátečním rozdělením bude s rostoucím časem t konvergovat ke svému stacionárnímu rozdělení.

Po zbytek práce se budeme zabývat rychlostí této konvergence, je tedy vhodné pro $t \in \mathbb{N}$ definovat funkci:

$$d(t) := \max_{x \in \mathcal{X}} \|P^t(x, \cdot) - \pi\|_{TV},$$

měřící vzdálenost marginálního rozdělení řetězce začínajícího v libovolném stavu od stacionárního rozdělení v závislosti na čase t . Díky tomu tak můžeme zadefinovat poslední důležitý pojem.

Definice 5 (Čas mixingů). *Časem mixingů pro dané $\varepsilon \in [0,1]$ budeme nazývat hodnotu:*

$$t_{mix}(\varepsilon) := \min\{t \in \mathbb{N} : d(t) \leq \varepsilon\}.$$

Tedy čas mixingů závisí na námi určené dostatečně malé vzdálenosti od stacionárního rozdělení. V zájmu dodržování konvence, stanovíme v této práci dostatečně malou vzdálenost na $1/4$. Tedy budeme značit:

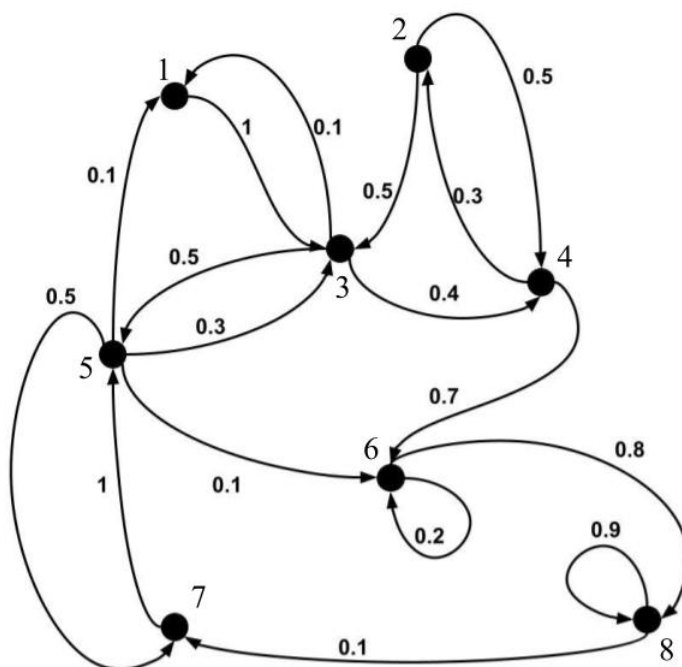
$$t_{mix} := t_{mix}\left(\frac{1}{4}\right).$$

3. Dolní meze pro mixing

V této kapitole uvedeme postupně tři metody, kterými lze určit dolní mez pro čas mixing. Všechny metody důkladně odvodíme a přidáme k nim názorný příklad, na kterém je použijeme.

Všechny řetězce s konečnou množinou stavů lze reprezentovat jako náhodnou procházku po orientovaném grafu s ohodnocenými hranami. V této reprezentaci vrcholy odpovídají jednotlivým stavům, které jsou spojené hranou, pokud lze z jednoho stavu přejít do druhého v jednom kroku s nenulovou pravděpodobností. Tyto pravděpodobnosti přechodů jsou pak hodnotami hran. Máme tedy jednoznačnou souvislost mezi maticí přechodu a orientovaným grafem. Příklad takové reprezentace, i s odpovídající maticí přechodu, lze vidět na obrázku 3.1.

$$\begin{array}{c}
 \begin{array}{cccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
 \begin{array}{c}
 1 \\
 2 \\
 3 \\
 4 \\
 5 \\
 6 \\
 7 \\
 8
 \end{array}
 & \left(\begin{array}{cccccccc}
 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\
 0.1 & 0 & 0 & 0.4 & 0.5 & 0 & 0 & 0 \\
 0 & 0.3 & 0 & 0 & 0 & 0.7 & 0 & 0 \\
 0.1 & 0 & 0.3 & 0 & 0 & 0.1 & 0.5 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0.2 & 0 & 0.8 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0.1 & 0.9
 \end{array} \right)
 \end{array}
 \end{array}$$



Obrázek 3.1: Ukázka reprezentace markovského řetězce

Na tuto reprezentaci se budeme dále odkazovat jako na *grafovou reprezentaci*.

3.1 Počítací mez

Pro řetězce, jejichž stacionární rozdělení je rovno rovnoměrnému, je jasné, že pokud po t krocích není dosažitelná dostatečně velká část \mathcal{X} , nemůže být pravděpodobnostní rozdělení X_t blízko stacionárnímu.

Na této úvaze bude stavět náš první dolní odhad, který vychází z grafové reprezentace. Nalezneme stav $x \in \mathcal{X}$ takový, že v reprezentaci má jemu odpovídající vrchol nejvyšší stupeň Δ , tedy lze z něj přejít do nejvíce jiných stavů. Potom, začne-li řetězec ve stavu x_0 , a označíme-li $\mathcal{X}_t^{x_0}$ množinu stavů dosažitelných z x_0 v čase t , platí: $\Delta^t \geq |\mathcal{X}_t^{x_0}|$. A tedy:

$$\|P^t(x_0, \cdot) - \pi\|_{TV} \geq P^t(x_0, \mathcal{X}_t^{x_0}) - \pi(\mathcal{X}_t^{x_0}) = 1 - \frac{|\mathcal{X}_t^{x_0}|}{|\mathcal{X}|} \geq 1 - \frac{\Delta^t}{|\mathcal{X}|},$$

kde první nerovnost vychází z definice vzdálenosti v totální variaci.

Jestliže pak pro dané $\epsilon > 0$ platí: $\Delta^t < (1 - \epsilon)|\mathcal{X}|$, jednoduchou úpravou pak máme, že:

$$1 - \frac{\Delta^t}{|\mathcal{X}|} \geq \epsilon.$$

Z toho lze, za použití jednoduché práce s logaritmy, odvodit náš první dolní odhad.

Tvrzení 3.1. *Pro aperiodický ireducibilní markovský řetězec s konečnou množinou stavů a stacionárním rozdělením rovným rovnoměrnému platí:*

$$t_{mix}(\epsilon) \geq \frac{\log(|\mathcal{X}|(1 - \epsilon))}{\log(\Delta)}, \quad \epsilon \in (0,1).$$

Důkaz. Proveden odvozením přímo nad tvrzením. □

Příklad. Položme $\mathcal{X}_n = GL_n(\mathbb{F}_2)$, tedy množinu všech invertibilních matic o rozměru $n \times n$ nad dvojprvkovým tělesem \mathbb{F}_2 . Na této množině budeme uvažovat markovský řetězec, který v jednom kroku zvolí libovolnou uspořádanou dvojici (i, j) ; $i, j \in \{1, \dots, n\}$, $i \neq j$ a přičte i -tý řádek k j -tému. Z formulace příkladu je zřejmé, že budeme uvažovat $n \geq 2$.

Nejprve ukážeme, že existuje $\gamma > 0$ takové, že:

$$\frac{|\mathcal{X}_n|}{2^{n^2}} \rightarrow \gamma, \quad \text{pro } n \rightarrow \infty. \quad (3.1)$$

Zřejmé je 2^{n^2} počet všech $n \times n$ matic nad \mathbb{F}_2 , ukazujeme tedy, že poměr invertibilních matic ke všem je limitně konstantní.

Podívejme se na velikost \mathcal{X}_n kombinatoricky. Libovolná matice nad \mathbb{F}_2 je invertibilní právě tehdy, když jsou její řádky lineárně nezávislé. To je nad \mathbb{F}_2 ekvivalentní tomu, že žádný řádek nesmí být součtem libovolné podmnožiny ostatních řádků nebo nesmí být nulový.

Začneme-li podle tohoto požadavku budovat matici z $GL_n(\mathbb{F}_2)$ řádek po řádku, dostaneme toto:

1. První řádek může být libovolný vektor nad \mathbb{F}_2 délky n až na nulový. Existuje tak $2^n - 1$ možností jeho volby.
2. Druhý řádek může být také libovolný, až na nulový a řádek nad ním. Tedy je $2^n - 1 - 1$ možností pro jeho volbu.
3. Volba třetího řádku je omezena nulovým, dvěma předchozími a jejich součtem. Tedy je $2^n - 1 - 2 - 1$ možností.
4. Obecně k -tý řádek má 2^n možností volby bez nulového vektoru, všech jednotlivých řádků před ním, všech součtů dvou řádků, všech součtů trojic řádků, a všech dalších součtů j -tic předcházejících řádků pro j od 1 do $k - 1$. Tedy možností pro jeho volbu tak je:

$$2^n - \binom{k-1}{0} - \binom{k-1}{1} - \dots - \binom{k-1}{k-1} = 2^n - \sum_{j=1}^{k-1} \binom{k-1}{j-1}.$$

Součty jednotlivých j -tic jsou různé, neb jsou řádky lineárně nezávislé a tedy v sumě není žádný započítán vícekrát.

Pro poslední sumu existuje vzorec pro její hodnotu, vycházející z vlastností binomických koeficientů. Počet možností pro volbu k -tého řádku lze tedy zapsat následovně:

$$2^n - \sum_{j=1}^{k-1} \binom{k-1}{j-1} = 2^n - 2^{k-1}.$$

Abychom získali velikost \mathcal{X}_n , stačí počet možností pro jednotlivé řádky mezi sebou vynásobit. Máme tak:

$$|\mathcal{X}_n| = \prod_{k=1}^n (2^n - 2^{k-1}) = \prod_{k=1}^n 2^n (1 - 2^{k-1-n}) = 2^{n^2} \prod_{k=1}^n (1 - 2^{-k}), \quad (3.2)$$

kde poslední rovnost jsme získali přerovněním součinu.

Chceme-li tedy prověřit limitní chování podílu (3.1), stačí ukázat, že:

$$\prod_{k=1}^n (1 - 2^{-k}) \rightarrow \gamma; \text{ pro } n \rightarrow \infty. \quad (3.3)$$

K tomu využijeme Eulerova vzorce pro vyjádření sinu jako nekonečného součinu:

$$\sin x = x \prod_{k=1}^{\infty} \left(1 - \frac{x^2}{k^2 \pi^2}\right), \quad (3.4)$$

jehož důkaz provedl například Eberlein (1977) ve svém článku s využitím minima znalostí z komplexní analýzy.

Zřejmě existuje k_0 takové, že pro každé $k \geq k_0$: $2^k \geq 4k^2$, potom, pro n dost velké, platí:

$$\prod_{k=1}^n (1 - 2^{-k}) = \prod_{k=1}^{k_0} (1 - 2^{-k}) \prod_{k=k_0}^n \left(1 - \frac{1}{2^k}\right) \geq \prod_{k=1}^{k_0} (1 - 2^{-k}) \prod_{k=k_0}^n \left(1 - \frac{1}{4k^2}\right)$$

Položíme-li tedy ve (3.4) $x = \pi/2$ dostaneme:

$$\begin{aligned}
\prod_{k=1}^n (1 - 2^{-k}) &\geq \prod_{k=1}^{k_0} (1 - 2^{-k}) \prod_{k=k_0}^n \left(1 - \frac{1}{4k^2}\right) \\
&= \prod_{k=1}^{k_0} (1 - 2^{-k}) \prod_{k=1}^{k_0} \left(1 - \frac{1}{4k^2}\right)^{-1} \prod_{k=1}^n \left(1 - \frac{\left(\frac{\pi}{2}\right)^2}{k^2\pi^2}\right) \\
&\xrightarrow{n \rightarrow \infty} \frac{2}{\pi} \sin \frac{\pi}{2} \prod_{k=1}^{k_0} (1 - 2^{-k}) \prod_{k=1}^{k_0} \left(1 - \frac{1}{4k^2}\right)^{-1} \\
&= \frac{2}{\pi} \prod_{k=1}^{k_0} (1 - 2^{-k}) \prod_{k=1}^{k_0} \left(1 - \frac{1}{4k^2}\right)^{-1} > 0
\end{aligned}$$

Tedy víme, že hodnota (3.3) je zdola omezena nezápornou konstantou pro každé $n \in \mathbb{N}$. Všimneme si, že jde o monotónně klesající posloupnost, neb s rostoucím n násobíme více činiteli menšími, než jedna. A proto jistě existuje její limita $\gamma > 0$.

Teď se zaměříme na samotnou konvergenci uvažovaného řetězce, kde využijeme tuto ukázanou vlastnost. Ukážeme, že pro daný markovský řetězec platí:

$$t_{mix} \geq c \frac{n^2}{\log n},$$

kde $c > 0$ je konstanta.

Víme, že invertibilní matice lze sérií přičítání násobků řádků k jiným (takzvaných elementárních řádkových úprav) transformovat na matici jednotkovou. Takto jsou všechny stavy navzájem dosažitelné přes jednotkovou matici a řetězec je ireducibilní. Aperiodicitu lze snadno ověřit nalezením cyklů stavů o liché délce, neb cykly délky 2 zřejmě existují. Zároveň je právě $n(n-1)$ uspořádaných dvojic indexů řádků, a tedy lze z každého stavu v jednom kroku přejít do $n(n-1)$ různých stavů.

Celkem z toho plyne, že v grafové reprezentaci lze řetězec reprezentovat náhodnou procházkou po $n(n-1)$ -regulárním grafu. Stacionární rozdělení je tedy rovno rovnoměrnému na \mathcal{X}_n , což lze snadno ověřit výpočtem.

Řetězec splňuje předpoklady tvrzení 3.1 a platí pro něj:

$$t_{mix} \geq \frac{\log\left(\frac{3}{4} |\mathcal{X}_n|\right)}{\log n(n-1)}.$$

Tento odhad budeme dále upravovat. Ukážeme, že posloupnost:

$$\frac{\log\left(\frac{3}{4} |\mathcal{X}_n|\right)}{\log 2^{n^2}}$$

má limitu v intervalu $(0, \infty)$ a tedy ji lze zdola odhadnout konstantou.

Využijeme vyjádření (3.2) a počítáme:

$$\begin{aligned}
\frac{\log\left(\frac{3}{4} |\mathcal{X}_n|\right)}{\log 2^{n^2}} &= \frac{\log\left(\frac{3}{4} 2^{n^2} \prod_{k=1}^n (1 - 2^{-k})\right)}{\log 2^{n^2}} = \frac{\log 2^{n^2} + \log\left(\frac{3}{4} \prod_{k=1}^n (1 - 2^{-k})\right)}{\log 2^{n^2}} \\
&= 1 + \frac{\log\left(\frac{3}{4} \prod_{k=1}^n (1 - 2^{-k})\right)}{\log 2^{n^2}} \xrightarrow{n \rightarrow \infty} 1.
\end{aligned}$$

Limita zde plyne z toho, že logaritmus v čitateli jde limitně k $\log \frac{3\gamma}{4}$, jak jsme ukázali výše, zatímco logaritmus ve jmenovateli roste nade všechny meze. Z vlastností logaritmu tedy pro všechna $n \geq 2$ platí:

$$\frac{\log\left(\frac{3}{4}|\mathcal{X}_n|\right)}{\log 2^{n^2}} \in \left(\frac{\log \frac{9}{2}}{\log 2^4}, 1\right),$$

neb $|\mathcal{X}_2| = 6$.

Můžeme tedy upravovat:

$$\frac{\log\left(\frac{3}{4}|\mathcal{X}_n|\right)}{\log n(n-1)} = \frac{\log\left(\frac{3}{4}|\mathcal{X}_n|\right)}{\log 2^{n^2}} \frac{\log 2^{n^2}}{\log(n^2-n)} \geq \frac{\log \frac{9}{2} n^2 \log 2}{\log 2^4 \log n^2} = \frac{\log \frac{9}{2} n^2}{8 \log n}.$$

Celkem jsme takto ukázali, že:

$$t_{mix} \geq \frac{\log \frac{9}{2} n^2}{8 \log n}.$$

A máme tak dolní odhad na rychlost růstu t_{mix} s rostoucím n , který jsme chtěli ukázat. ▲

3.2 Míra zúžení

Další odhad, který ukážeme, využívá geometrických vlastností grafové reprezentace. Jsou-li některé části stavového prostoru těžko dosažitelné z jiných částí (tedy pravděpodobnost přechodu mezi nimi je nízká), potom rychlost konvergence bude také patřičně pomalejší.

Abychom tuto intuici mohli formalizovat zavedeme nový pojem:

Definice 6 (Hranová míra). *Pro stavy $x, y \in \mathcal{X}$ markovského řetězce s maticí přechodu P definujeme:*

$$Q(x, y) = \pi(x)P(x, y),$$

kde π je stacionární rozdělení řetězce. Pro množiny $A, B \subset \mathcal{X}$ definujeme dále:

$$Q(A, B) = \sum_{x \in A, y \in B} Q(x, y).$$

Hodnotu $Q(A, B)$ nazveme hranová míra mezi množinami A a B .

Hodnota $Q(A, B)$ tedy udává pravděpodobnost, že se v jednom kroku řetězce přešlo z množiny A do množiny B , je-li počáteční rozdělení řetězce rovno rovnoměrnému. Na základě předchozí definice můžeme dále zavést *míru zúžení*:

Definice 7 (Míra zúžení). *Míra zúžení neprázdné množiny $S \subset \mathcal{X}$ je dána jako:*

$$\Phi(S) := \frac{Q(S, S^c)}{\pi(S)},$$

kde S^c značí doplněk množiny S . Dále míru zúžení celého řetězce definujeme takto:

$$\Phi_* := \min_{S \subset \mathcal{X}: \pi(S) \leq \frac{1}{2}} \Phi(S).$$

Vybaveni těmito pojmy můžeme ukázat další dolní mez pro čas mixingů.

Věta 3.2. *Pro aperiodický a ireducibilní markovský řetězec s konečnou množinou stavů platí:*

$$t_{mix} \geq \frac{1}{4\Phi_\star}.$$

Důkaz. Uvažujme řetězec, jehož počáteční rozdělení je rovno stacionárnímu rozdělení π , takto bude platit $\mathbf{P}_\pi(X_t = x) = \pi(x)$ pro všechna $t \in \mathbb{N}$ a $x \in \mathcal{X}$. Pak pro neprázdnou množinu $A \subset \mathcal{X}$ a $t \in \mathbb{N}$ platí:

$$\begin{aligned} \mathbf{P}_\pi(X_0 \in A, X_t \in A^c) &\leq \sum_{r=1}^t \mathbf{P}_\pi(X_{r-1} \in A, X_r \in A^c) \\ &= \sum_{r=1}^t \mathbf{P}_\pi(X_{r-1} \in A) \mathbf{P}_\pi(X_r \in A^c | X_{r-1} \in A) \\ &= t \mathbf{P}_\pi(X_0 \in A) \mathbf{P}_\pi(X_1 \in A^c | X_0 \in A) \\ &= t \mathbf{P}_\pi(X_0 \in A, X_1 \in A^c) = tQ(A, A^c), \end{aligned}$$

kde ve druhé rovnosti využíváme nezávislosti distribuce X_t na čase a homogenity řetězce a v poslední rovnosti používáme definici Q .

Vydělením obou stran $\pi(A)$ získáme:

$$\begin{aligned} \mathbf{P}_\pi(X_t \in A^c | X_0 \in A) &\leq t\Phi(A), \\ 1 - \mathbf{P}_\pi(X_t \in A^c | X_0 \in A) &\geq 1 - t\Phi(A), \\ \mathbf{P}_\pi(X_t \in A | X_0 \in A) &\geq 1 - t\Phi(A). \end{aligned} \tag{3.5}$$

Upravíme levou stranu nerovnosti, kde využijeme, že $\pi(A) = \sum_{x \in A} \pi(x)$:

$$\begin{aligned} \mathbf{P}_\pi(X_t \in A | X_0 \in A) &= \frac{\mathbf{P}_\pi(X_t \in A, X_0 \in A)}{\pi(A)} = \frac{\sum_{x \in A} \mathbf{P}_\pi(X_t \in A, X_0 = x)}{\sum_{x \in A} \pi(x)} \\ &= \frac{\sum_{x \in A} \pi(x) P^t(x, A)}{\sum_{x \in A} \pi(x)}. \end{aligned}$$

Podíl, který jsme získali, je vážený průměr hodnot $P^t(x, A)$ pro $x \in A$. Tedy jistě existuje $x' \in A$ takové, že s využitím nerovnosti (3.5) platí:

$$P^t(x', A) \geq \frac{\sum_{x \in A} \pi(x) P^t(x, A)}{\sum_{x \in A} \pi(x)} = \mathbf{P}_\pi(X_t \in A | X_0 \in A) \geq 1 - t\Phi(A).$$

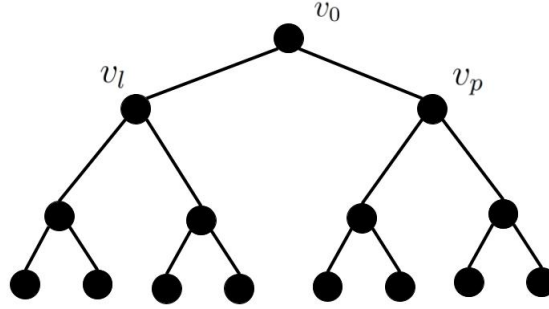
Od obou stran odečteme $\pi(A)$ a z definice vzdálenosti v totální variaci dostaneme:

$$d(t) \geq \left\| P^t(x', \cdot) - \pi \right\|_{TV} \geq 1 - t\Phi(A) - \pi(A). \tag{3.6}$$

Nalezneme A' takové, že $\pi(A') \leq 1/2$ a $\Phi(A') = \Phi_\star$, pak:

$$d(t) \geq 1 - t\Phi_\star - \pi(A') \geq \frac{1}{2} - t\Phi_\star,$$

Finálně, je-li $t < 1/(4\Phi_\star)$, pak je $d(t) > 1/4$. A tedy $t_{mix} \geq 1/(4\Phi_\star)$. □



Obrázek 3.2: Binární zakořeněný strom pro $k = 3$.

Příklad. Uvažujme náhodnou procházku po grafu zakořeněného binárního stromu hloubky k . Takto nazýváme graf s jedním vrcholem v_0 (= kořenem), který je spojen hranou se dvěma dalšími vrcholy v_p, v_l a každý z nich má další dva sousedy. Graf se takto dále větví do k pater. Ilustraci lze vidět na obrázku 3.2.

Označíme-li $|v|$ délku nejkratší cesty z vrcholu v do v_0 a celkový počet vrcholů grafu: $n = 2^{k+1} - 1$, pak má tento řetězec stacionární rozdělení tvaru:

$$\pi(v) = \begin{cases} \frac{2}{2n-2} & \text{pro } v = v_0, \\ \frac{3}{2n-2} & \text{pro } 0 < |v| < k, \\ \frac{1}{2n-2} & \text{pro } |v| = k. \end{cases}$$

Což lze snadno ověřit přímým výpočtem. Označme S množinu všech vrcholů, které jsou dosažitelné z v_p , aniž bychom prošli v_0 . Tedy S je celá pravá větev stromu. Potom máme:

$$\begin{aligned} \pi(S) &= \sum_{v \in S} \pi(v) = (2^{k-1} - 1) \frac{3}{2n-2} + 2^{k-1} \frac{1}{2n-2} \\ &= \frac{3 \cdot 2^{k-1} - 3 + 2^{k-1}}{2n-2} = \frac{2 \cdot 2^{k-1} + 2^{k-1} + 2^{k-1} - 3}{2n-2} \\ &= \frac{2^{k+1} - 1 - 2}{2n-2} = \frac{n-2}{2n-2}. \end{aligned}$$

Dále, neb S lze opustit pouze z vrcholu v_p , platí:

$$Q(S, S^c) = \pi(v_r) P(v_r, v_0) = \frac{3}{2n-2} \frac{1}{3} = \frac{1}{2n-2}.$$

A tedy dle definice Φ :

$$\Phi(S) = \frac{1}{2n-2} \left(\frac{n-2}{2n-2} \right)^{-1} = \frac{1}{n-2}.$$

Celkem dle tvrzení 3.2 dostáváme:

$$t_{mix} \geq \frac{1}{4\Phi(S)} = \frac{n-2}{4}.$$

▲

3.3 Rozlišující statistiky

Poslední dolní odhad, který v této práci budeme probírat, stojí na volbě vhodné rozlišující statistiky. Pojmeme *statistika* zde myslíme měřitelnou funkci $f : \mathcal{X} \rightarrow \mathbb{R}$. Tuto statistiku budeme volit tak, aby se rozdíl středních hodnot $\mathbf{E}_x f(X_t)$ a $\mathbf{E}_\pi f(X)$, dal dobře odhadnout zdola. V následující větě ukážeme, jak tento odhad využít k odhadnutí vzdálenosti samotných rozdělání.

Věta 3.3. *Pro funkci $f : \mathcal{X} \rightarrow \mathbb{R}$ položme $\sigma_\star^2 := \max\{\text{Var}_\mu(f(X)), \text{Var}_\nu(f(X))\}$. Pokud existuje $r > 0$ takové, že:*

$$|\mathbf{E}_\nu f(X) - \mathbf{E}_\mu f(X)| \geq r\sigma_\star, \quad (3.7)$$

pak platí:

$$\|\mu - \nu\|_{TV} \geq 1 - \frac{8}{r^2}.$$

Důsledek. *Speciálně, pokud pro markovský řetězec s maticí přechodu P a stacionárním rozdělením π funkce f splňuje:*

$$|\mathbf{E}_x f(X_t) - \mathbf{E}_\pi f(X)| \geq r\sigma_\star,$$

pro $\sigma_\star^2 = \max\{\text{Var}_\pi(f(X)), \mathbf{Var}_x(f(X_t))\}$, $x \in \mathcal{X}$ a $r > 0$. Potom jest:

$$\|P^t(x, \cdot) - \pi\|_{TV} \geq 1 - \frac{8}{r^2}.$$

Důkaz věty 3.3. Bez újmy na obecnosti předpokládejme, že $\mathbf{E}_\mu f(X) \leq \mathbf{E}_\nu f(X)$. Dále položme $A = (\mathbf{E}_\mu f(X) + r\sigma_\star \frac{1}{2}, \infty)$. Potom pro náhodnou veličinou X s hodnotami v \mathcal{X} takovou, že má rozdělení rovné μ platí:

$$\begin{aligned} \mu_f(A) &= \mathbf{P}(f(X) \in A) = \mathbf{P}\left(f(X) \geq \mathbf{E}_\mu f(X) + \frac{r\sigma_\star}{2}\right) \\ &= \mathbf{P}\left(f(X) - \mathbf{E}_\mu f(X) \geq \frac{r\sigma_\star}{2}\right) \leq \mathbf{P}\left(|f(X) - \mathbf{E}_\mu f(X)| \geq \frac{r\sigma_\star}{2}\right) \\ &\leq \frac{4 \text{Var}_\mu(f(X))}{r^2 \sigma_\star^2} \leq \frac{4}{r^2}, \end{aligned}$$

kde v předposlední nerovnosti využíváme Čebyševovy nerovnosti (tvrzení 1.3).

Naopak pro náhodnou veličinou Y s hodnotami v \mathcal{X} a rozdělení rovným ν je:

$$\begin{aligned} \nu_f(A) &= \mathbf{P}(f(Y) \in A) = \mathbf{P}\left(f(Y) \geq \mathbf{E}_\mu f(Y) + \frac{r\sigma_\star}{2}\right) \\ &= \mathbf{P}\left(f(Y) \geq \mathbf{E}_\nu f(Y) - (\mathbf{E}_\nu f(Y) - \mathbf{E}_\mu f(Y)) + \frac{r\sigma_\star}{2}\right) \\ &\geq \mathbf{P}\left(f(Y) \geq \mathbf{E}_\nu f(Y) - \frac{r\sigma_\star}{2}\right) = \mathbf{P}\left(\mathbf{E}_\nu f(Y) - f(Y) \leq \frac{r\sigma_\star}{2}\right) \\ &= 1 - \mathbf{P}\left(\mathbf{E}_\nu f(Y) - f(Y) > \frac{r\sigma_\star}{2}\right) \geq 1 - \mathbf{P}\left(|f(Y) - \mathbf{E}_\nu f(Y)| \geq \frac{r\sigma_\star}{2}\right) \\ &\geq 1 - \frac{4 \text{Var}(f(Y))}{r^2 \sigma_\star^2} \geq 1 - \frac{4}{r^2}, \end{aligned}$$

kde v první nerovnosti využíváme (3.7) z předpokladu a v předposlední opět Čebyševovu nerovnost.

Z předchozích nerovností pak dostaneme:

$$\|\mu_f - \nu_f\|_{TV} \geq \nu_f(A) - \mu_f(A) \geq 1 - \frac{4}{r^2} - \frac{4}{r^2} = 1 - \frac{8}{r^2}.$$

Finálně, díky konečnosti \mathcal{X} , platí:

$$\begin{aligned} \|\mu_f - \nu_f\|_{TV} &= \max_{A \subset \Theta} |\mu_f(A) - \nu_f(A)| = \max_{A \subset \Theta} |\mu(f^{-1}(A)) - \nu(f^{-1}(A))| \\ &\leq \max_{B \subset \mathcal{X}} |\mu(B) - \nu(B)| = \|\mu - \nu\|_{TV} \end{aligned}$$

□

Příklad. Uvažujme obarvení prázdného grafu (V, \emptyset) pomocí q barev, tedy přiřazení jednoho prvku z množiny $Q = \{1, \dots, q\}$ každému $v \in V$, za podmínky že sousední vrcholy nemají stejnou barvu. Tato podmínka v našem případě vypadá, neboť v prázdném grafu nemá žádný vrchol sousedy.

Náš markovský řetězec bude náhodná procházka po všech obarveních tohoto grafu. Tedy množina stavů \mathcal{X} jsou všechna možná obarvení. Krok řetězce sestává ze zvolení vrcholu dle rovnoměrného rozdělení na V , a jeho přebarvení na náhodně zvolenou přípustnou barvu (v našem případě libovolnou), také dle rovnoměrného rozdělení na Q . Jinými slovy je množina stavů řetězce $\mathcal{X} = Q^n$, kde $n = |V|$ je počet vrcholů v grafu. Jednotlivé stavy řetězce pak značíme vektorem $\mathbf{X}_t = (x_t^{(1)}, \dots, x_t^{(n)})$, kde v jednom kroku se zvolí jedna pozice a té se přiřadí nová hodnota z množiny Q , která může být stejná, jako ta dosavadní.

Tento řetězec má zřejmě konečnou množinu stavů, je nerozložitelný a aperioidický. Výpočtem lze dále ukázat, že stacionárním rozdělením tohoto řetězce je rozdělení rovnoměrné.

Ukážeme, že existuje konstanta c_q pro libovolnou dvojici n a q taková, že poté již platí:

$$t_{mix} \geq \frac{1}{2}n \log n - c_q n.$$

Odtěď předpokládejme, že $q \geq 2$, neboť jinak by byla $|\mathcal{X}| = 1$ a $t_{mix} = 0$. Stejně tak předpokládáme $n \geq 2$, neb v případě $n = 1$ by pro oba možné počáteční stavy platilo: $t_{mix} = 1$.

Jako vhodnou rozlišující statistiku zvolíme $W(\mathbf{X}_t) = \sum_{i=1}^n x_t^{(i)}$. Potom lze snadno spočítat střední hodnotu a rozptyl vzhledem ke stacionárnímu rozdělení π , neboť pro každý vrchol $x^{(i)}$ platí: $\mathbf{P}(x^{(i)} = k) = 1/q$, je-li $k \in Q$. Tyto pravděpodobnosti jsou pro různé vrcholy na sobě navzájem nezávislé. Takže platí:

$$\mathbf{E}_\pi W(\mathbf{X}) = \mathbf{E}_\pi \sum_{i=1}^n x^{(i)} = \sum_{i=1}^n \mathbf{E}_\pi x^{(i)} = n \sum_{k=1}^q \frac{k}{q} = n \frac{q(q+1)}{2q} = \frac{n(q+1)}{2} \quad (3.8)$$

$$\begin{aligned}
\text{Var}_\pi W(\mathbf{X}) &= \sum_{i=1}^n \text{Var}_\pi x^{(i)} = n \text{Var}_\pi x^{(1)} = n \left(\sum_{k=1}^q \frac{k^2}{q} - \frac{(q+1)^2}{4} \right) \\
&= n \left(\frac{q(q+1)(2q+1)}{6q} - \frac{(q+1)^2}{4} \right) \\
&= n \left(\frac{2q^2 + 3q + 1}{6} - \frac{q^2 + 2q + 1}{4} \right) = n \left(\frac{q^2 - 1}{12} \right),
\end{aligned} \tag{3.9}$$

kde v první rovnosti jsme využili nezávislosti $x^{(i)}$.

Dále, pro použití věty 3.3, potřebujeme střední hodnotu a rozptyl $W(\mathbf{X}_t)$ pro nějaké počáteční rozdělení. Zvolíme rozdělení, které s pravděpodobností 1 začíná ve stavu, kdy mají všechny vrcholy barvu 1. Tedy $\mathbf{X}_0 = \mathbf{1}$, kde $\mathbf{1}$ značí jednotkový vektor.

Pro výpočet obou momentů použijeme ještě pomocnou statistiku R_t , kterou definujeme pomocí indikátorů následovně:

$$\begin{aligned}
Y_t^{(i)} &= \mathbb{1}_{\{\text{Vrchol } x^{(i)} \text{ ještě nebyl zvolen k přebarvení v čase } t\}}, \\
R_t &= \sum_{i=1}^n Y_t^{(i)}.
\end{aligned}$$

Za pomoci této statistiky již snadno spočteme podmíněnou střední hodnotu $W(\mathbf{X}_t)$, neb víme, že nepřebarvené vrcholy mají barvu 1, a ty zvolené k přebarvení mohou mít barvu libovolnou:

$$\begin{aligned}
\mathbf{E}_1[W(\mathbf{X}_t)|R_t] &= 1 \cdot R_t + (n - R_t) \sum_{k=1}^q \frac{k}{q} = R_t + (n - R_t) \frac{q(q+1)}{2q} \\
&= R_t + (n - R_t) \frac{q+1}{2} = n \frac{q+1}{2} + \frac{1-q}{2} R_t.
\end{aligned} \tag{3.10}$$

Dále spočteme střední hodnotu R_t :

$$\mathbf{E}_1 R_t = \sum_{i=1}^n \mathbf{E}_1 Y_t^{(i)} = n \mathbf{P}_1(Y_t^{(1)} = 1) = n \left(1 - \frac{1}{n}\right)^t, \tag{3.11}$$

kde poslední rovnost plyne z faktu, že $Y_t^{(i)} = 1$ pouze, pokud se t -krát zvolil některý z n jiných vrcholů.

Dáme-li oba výsledky dohromady, dostaneme:

$$\begin{aligned}
\mathbf{E}_1 W(\mathbf{X}_t) &= \mathbf{E}_1 (\mathbf{E}_1[W(\mathbf{X}_t)|R_t]) = \mathbf{E}_1 \left(n \frac{q+1}{2} + \frac{q-1}{2} R_t \right) \\
&= n \frac{q+1}{2} + \frac{1-q}{2} n \left(1 - \frac{1}{n}\right)^t.
\end{aligned} \tag{3.12}$$

Teď se podíváme na podmíněný rozptyl. Budeme nejprve podmiňovat znalostí celého vektoru $\mathbf{Y}_t = (Y_t^{(1)}, \dots, Y_t^{(n)})$, poté ukážeme, že to je ekvivalentní rozptylu podmíněnému pouze hodnotou R_t .

Nejprve si povšimněme následující rovnosti:

$$\begin{aligned}
\mathbf{E}_1[W(\mathbf{X}_t)|\mathbf{Y}_t] &= \mathbf{E}_1 \left[\sum_{i=1}^n \left(x_t^{(i)}(1 - Y_t^{(i)}) + Y_t^{(i)} \right) \middle| \mathbf{Y}_t \right] \\
&= \mathbf{E}_1 \left[\sum_{i=1}^n x_t^{(i)}(1 - Y_t^{(i)}) \middle| \mathbf{Y}_t \right] + \mathbf{E}_1 \left[\sum_{i=1}^n Y_t^{(i)} \middle| \mathbf{Y}_t \right] \\
&= \sum_{i=1}^n (1 - Y_t^{(i)}) \mathbf{E}_1 \left[x_t^{(i)} \middle| Y_t^{(i)} \right] + \sum_{i=1}^n Y_t^{(i)} \\
&= \sum_{i=1}^{n-R_t} \mathbf{E}_1 \left[x_t^{(i)} \middle| Y_t^{(i)} = 1 \right] + R_t = (n - R_t) \sum_{k=1}^q \frac{k}{q} + R_t,
\end{aligned} \tag{3.13}$$

kde jsme v předposlední rovnosti použili definici R_t jako součtu $Y_t^{(i)}$ a bez újmy na obecnosti předpokládali, že $Y_t^{(i)} = 0$ pro $i \in \{1, \dots, n - R_t\}$. Tedy vidíme, že se podmíněné střední hodnoty v (3.10) a (3.13) rovnají skoro jistě.

Teď tedy počítejme podmíněný rozptyl. Využijeme toho, že víme-li, které vrcholy již byly, či nebyly, zvoleny k přebarvení, jsou jejich barvy vzájemně nezávislé:

$$\begin{aligned}
\mathbf{Var}_1[W(\mathbf{X}_t)|\mathbf{Y}_t] &= \mathbf{Var}_1 \left[\sum_{i=1}^n \left(x_t^{(i)}(1 - Y_t^{(i)}) + Y_t^{(i)} \right) \middle| \mathbf{Y}_t \right] \\
&= \mathbf{Var}_1 \left[\sum_{i=1}^n x_t^{(i)}(1 - Y_t^{(i)}) + \sum_{i=1}^n Y_t^{(i)} \middle| \mathbf{Y}_t \right] \\
&= \sum_{i=1}^n (1 - Y_t^{(i)})^2 \mathbf{Var}_1 \left[x_t^{(i)} \middle| Y_t^{(i)} \right] \\
&= \sum_{i=1}^{n-R_t} \mathbf{Var}_1 \left[x_t^{(i)} \middle| Y_t^{(i)} = 0 \right] \\
&= (n - R_t) \mathbf{Var}_1 \left[x_t^{(1)} \middle| Y_t^{(1)} = 0 \right] \\
&= (n - R_t) \left(\sum_{k=1}^q \frac{k^2}{q} - \frac{(q+1)^2}{4} \right) = (n - R_t) \frac{q^2 - 1}{12}.
\end{aligned} \tag{3.14}$$

Opět jsme ve čtvrté rovnosti bez újmy na obecnosti předpokládali, že $Y_t^{(i)} = 0$ pro $i \in \{1, \dots, n - R_t\}$. Také jsme využili toho, že hodnotu závorky jsme již spočetli v (3.9).

Zároveň z definice podmíněného rozptylu a rovnosti (3.10) a (3.13) platí:

$$\begin{aligned}
\mathbf{Var}_1[W(\mathbf{X}_t)|\mathbf{Y}_t] &= \mathbf{E}_1 \left[(W(\mathbf{X}_t) - \mathbf{E}_1[W(\mathbf{X}_t)|\mathbf{Y}_t])^2 \middle| \mathbf{Y}_t \right] \\
&= \mathbf{E}_1 \left[(W(\mathbf{X}_t) - \mathbf{E}_1[W(\mathbf{X}_t)|R_t])^2 \middle| \mathbf{Y}_t \right].
\end{aligned} \tag{3.15}$$

Tedy zajímá-li nás $\mathbf{Var}_1[W(\mathbf{X}_t)|R_t]$, můžeme využít jednoduché vlastnosti podmíněné střední hodnoty a s pomocí výsledků (3.14) a (3.15) dostaneme:

$$\begin{aligned}
\mathbf{Var}_1[W(\mathbf{X}_t)|R_t] &= \mathbf{E}_1 \left[(W(\mathbf{X}_t) - \mathbf{E}_1[W(\mathbf{X}_t)|R_t])^2 \middle| R_t \right] \\
&= \mathbf{E}_1 \left[\mathbf{E}_1 \left[(W(\mathbf{X}_t) - \mathbf{E}_1[W(\mathbf{X}_t)|R_t])^2 \middle| \mathbf{Y}_t \right] \middle| R_t \right] \\
&= \mathbf{E}_1 \left[\mathbf{Var}_1[W(\mathbf{X}_t)|\mathbf{Y}_t] \middle| R_t \right] = \mathbf{E}_1 \left[(n - R_t) \frac{q^2 - 1}{12} \middle| R_t \right] \\
&= (n - R_t) \frac{q^2 - 1}{12}.
\end{aligned} \tag{3.16}$$

Druhá rovnost platí, neb σ -algebra generovaná \mathbf{Y}_t je nadmnožinou σ -algebry generované R_t .

Teď se posuneme k samotnému vyjádření $\mathbf{Var}_1 W(\mathbf{X}_t)$. Využijeme zde známé rovnosti pro rozptyl, abychom mohli použít výsledky (3.10), (3.11) a (3.16):

$$\begin{aligned} \mathbf{Var}_1 W(\mathbf{X}_t) &= \mathbf{Var}_1 (\mathbf{E}_1[W(\mathbf{X}_t)|R_t]) + \mathbf{E}_1 (\mathbf{Var}_1[W(\mathbf{X}_t)|R_t]) \\ &= \mathbf{Var}_1 \left(n \frac{q+1}{2} + \frac{1-q}{2} R_t \right) + \mathbf{E}_1 (n - R_t) \frac{q^2 - 1}{12} \\ &= \frac{(1-q)^2}{4} \mathbf{Var}_1 R_t + n \frac{q^2 - 1}{12} - \frac{q^2 - 1}{12} n \left(1 - \frac{1}{n} \right)^t. \end{aligned} \quad (3.17)$$

Teď zafixujme:

$$\begin{aligned} t_n &:= \frac{1}{2}(n-1) \log n - (c_q - 1)n \\ &\geq \frac{1}{2}(n-1) \log n - (c_q - 1)n + \left(\frac{\log n}{2} - n \right) = \frac{1}{2}n \log n - c_q n, \end{aligned}$$

což platí, neb $\log n^{1/2} < n$. Ukážeme-li tedy, že pro takto zvolené t_n již platí: $\|P^{t_n}(\mathbf{1}, \cdot) - \pi\|_{TV} > 1/4$ nezávisle na hodnotě n , je důkaz hotov.

Pro další postup potřebujeme znát hodnotu σ_* . Na to potřebujeme zjistit, zda $\mathbf{Var}_1 W(\mathbf{X}_{t_n}) \leq \text{Var}_\pi W(X)$, či naopak. Platnost této nerovnosti závisí na specifické volbě n a q , proto ji zde nebudeme ověřovat. Místo toho nalezneme odhad pro obě varianty a sloučíme je.

Nejprve tedy předpokládejme, že nerovnost platí. Máme tak:

$$\sigma_* = \sqrt{n \frac{q^2 - 1}{12}}.$$

Spočítejme si teď levou stranu nerovnosti (3.7) na základě výsledků (3.8) a (3.12):

$$\begin{aligned} |\mathbf{E}_1 W(\mathbf{X}_{t_n}) - \mathbf{E}_\pi W(\mathbf{X})| &= \left| n \frac{q+1}{2} + n \frac{1-q}{2} \left(1 - \frac{1}{n} \right)^{t_n} - n \frac{q+1}{2} \right| \\ &= n \frac{q-1}{2} \left(1 - \frac{1}{n} \right)^{t_n}. \end{aligned}$$

A k ní nalezneme vhodné r :

$$\begin{aligned} n \frac{q-1}{2} \left(1 - \frac{1}{n} \right)^{t_n} &\geq r \sqrt{n \frac{q^2 - 1}{12}}, \\ n^2 \frac{(q-1)^2}{4} \left(1 - \frac{1}{n} \right)^{2t_n} &\geq r^2 n \frac{q^2 - 1}{12}, \\ 3n \frac{q-1}{q+1} \left(1 - \frac{1}{n} \right)^{2t_n} &\geq r^2. \end{aligned}$$

Hodnotu r^2 zvolíme tak, aby splňovala rovnost. Potom dle věty 3.3, po dosa-

zení za t_n , platí:

$$\begin{aligned} \left\| P^{t_n}(\mathbf{1}, \cdot) - \pi \right\|_{TV} &\geq 1 - \frac{8}{3n} \frac{q+1}{q-1} \left(1 - \frac{1}{n}\right)^{-2t_n} \\ &= 1 - \frac{8}{3n} \frac{q+1}{q-1} \left(1 - \frac{1}{n}\right)^{-(n-1)\log n + 2(c_q-1)n} \\ &= 1 - \frac{8}{3n} \frac{q+1}{q-1} \left(\left(1 - \frac{1}{n}\right)^{(n-1)} \right)^{-\log n} \left(\left(1 - \frac{1}{n}\right)^n \right)^{2(c_q-1)}. \end{aligned}$$

Využijme teď snadno ověřitelného faktu¹, že $\left(1 - \frac{1}{n}\right)^{(n-1)} > e^{-1} > \left(1 - \frac{1}{n}\right)^n$ a dále odhadujeme zdola:

$$\begin{aligned} \left\| P^{t_n}(\mathbf{1}, \cdot) - \pi \right\|_{TV} &\geq 1 - \frac{8}{3n} \frac{q+1}{q-1} e^{-(-\log n)} e^{-2(c_q-1)} \\ &= 1 - \frac{8}{3} \frac{q+1}{q-1} e^{-2(c_q-1)}. \end{aligned}$$

A chceme, aby celý tento výraz byl větší než 1/4:

$$\begin{aligned} 1 - \frac{8}{3} \frac{q+1}{q-1} e^{-2(c_q-1)} &> \frac{1}{4}, \\ \frac{3}{4} \frac{3q-1}{8q+1} &> e^{-2(c_q-1)}, \\ \log \left(\frac{9}{32} \frac{q-1}{q+1} \right) &> -2(c_q-1), \\ \log \sqrt{\frac{32q+1}{9q-1}} &< c_q - 1. \end{aligned} \tag{3.18}$$

Teď naopak předpokládejme, že $\mathbf{Var}_1 W(\mathbf{X}_{t_n}) \geq \mathbf{Var}_\pi W(X)$.

Hodnotu $\mathbf{Var}_1 W(\mathbf{X}_{t_n})$ máme již z (3.17) vyjádřenou v závislosti na $\mathbf{Var}_1 R_{t_n}$, vyjádřeme si tedy i to detailněji:

$$\begin{aligned} \mathbf{Var}_1 R_{t_n} &= \mathbf{Var}_1 \left(\sum_{i=1}^n Y_{t_n}^{(i)} \right) = \sum_{i=1}^n \mathbf{Var}_1 Y_{t_n}^{(i)} + \frac{1}{2} \sum_{i \neq j} \mathbf{Cov}_1(Y_{t_n}^{(i)}, Y_{t_n}^{(j)}) \\ &= n \mathbf{Var}_1 Y_{t_n}^{(1)} + \frac{n(n-1)}{2} \mathbf{Cov}_1(Y_{t_n}^{(1)}, Y_{t_n}^{(2)}) \\ &\leq n \mathbf{Var}_1 Y_{t_n}^{(1)} = n \left(1 - \frac{1}{n}\right)^{t_n} \left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right), \end{aligned}$$

kde nerovnost plyne z toho, že korelace jednotlivých indikátorů je záporná, což lze jednoduše ověřit přímým výpočtem.

Dosazením právě určeného horního odhadu $\mathbf{Var}_1 R_{t_n}$ do (3.17) máme teď novou volbu σ_* :

$$\sigma_* = \sqrt{\frac{(q-1)^2}{4} \mathbf{Var}_1 Y_{t_n}^{(1)} + n \frac{q^2-1}{12} - n \frac{q^2-1}{12} \left(1 - \frac{1}{n}\right)^{t_n}}.$$

Zvětšením σ_* nenarušíme předpoklady věty 3.3, pouze zhoršíme získaný odhad.

¹Jde o monotónní posloupnosti s limitou e^{-1} , kdy první je klesající a druhá rostoucí

V nadcházejícím výpočtu budeme ještě potřebovat následující nerovnost vycházející z našeho aktuálního předpokladu o nerovnosti obou rozptylů:

$$\begin{aligned}
\frac{(1-q)^2}{4} n \mathbf{Var}_1 Y_{t_n}^{(1)} + n \frac{q^2-1}{12} - \frac{q^2-1}{12} n \left(1 - \frac{1}{n}\right)^{t_n} &\geq n \frac{q^2-1}{12}, \\
\frac{(1-q)^2}{4} n \left(1 - \frac{1}{n}\right)^{t_n} \left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right) &\geq \frac{q^2-1}{12} n \left(1 - \frac{1}{n}\right)^{t_n}, \\
\left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right) &\geq \frac{q+1}{3(q-1)}, \\
3 \frac{q-1}{q+1} \left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right) &\geq 1. \tag{3.19}
\end{aligned}$$

Opět teď budeme hledat r dle (3.7):

$$\begin{aligned}
\frac{n}{2}(q-1) \left(1 - \frac{1}{n}\right)^{t_n} &\geq r \sqrt{\frac{(q-1)^2}{4} n \mathbf{Var}_1 Y_{t_n}^{(1)} + n \frac{q^2-1}{12} - n \frac{q^2-1}{12} \left(1 - \frac{1}{n}\right)^{t_n}}, \\
\frac{n^2(q-1)^2}{4r^2} \left(1 - \frac{1}{n}\right)^{2t_n} &\geq n \frac{(q-1)^2}{4} \mathbf{Var}_1 Y_{t_n}^{(1)} + n \frac{q^2-1}{12} - n \frac{q^2-1}{12} \left(1 - \frac{1}{n}\right)^{t_n}, \\
\frac{n^2(q-1)^2}{4r^2} \left(1 - \frac{1}{n}\right)^{2t_n} &\geq n \frac{q^2-1}{12} \left(3 \frac{q-1}{q+1} \mathbf{Var}_1 Y_{t_n}^{(1)} + 1 - \left(1 - \frac{1}{n}\right)^{t_n}\right), \\
\frac{3n}{r^2} \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{2t_n} &\geq 3 \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{t_n} \left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right) + 1 - \left(1 - \frac{1}{n}\right)^{t_n}, \\
\frac{3n}{r^2} \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{2t_n} &\geq \left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right) \left(3 \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{t_n} + 1\right), \\
3n \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{2t_n} &\geq r^2 \left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right) \left(3 \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{t_n} + 1\right).
\end{aligned}$$

Teď zvětšíme pravou stranu, čímž nenarušíme předpoklady věty 3.3. Vynecháme první závorku a ve druhé zvětšíme jednotkový člen dle (3.19):

$$\begin{aligned}
3n \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{2t_n} &\geq r^2 \left(3 \frac{q-1}{q+1} \left(1 - \frac{1}{n}\right)^{t_n} + 3 \frac{q-1}{q+1} \left(1 - \left(1 - \frac{1}{n}\right)^{t_n}\right)\right), \\
n \left(1 - \frac{1}{n}\right)^{2t_n} &\geq r^2 \left(\left(1 - \frac{1}{n}\right)^{t_n} + 1 - \left(1 - \frac{1}{n}\right)^{t_n}\right), \\
n \left(1 - \frac{1}{n}\right)^{2t_n} &\geq r^2.
\end{aligned}$$

I zde r^2 zvolíme tak, aby splňovala rovnost. A opět tedy dle věty 3.3 platí:

$$\begin{aligned}
\|P^{t_n}(x, \cdot) - \pi\|_{TV} &\geq 1 - \frac{8}{n} \left(1 - \frac{1}{n}\right)^{-2t_n} \geq 1 - \frac{8}{n} \left(1 - \frac{1}{n}\right)^{-(n-1) \log n + 2(c_q-1)n} \\
&\geq 1 - \frac{8}{n} \left(\left(1 - \frac{1}{n}\right)^{(n-1)}\right)^{\log n} \left(\left(1 - \frac{1}{n}\right)^n\right)^{2(c_q-1)} \\
&\geq 1 - \frac{8}{n} n e^{-2(c_q-1)} = 1 - 8e^{-2(c_q-1)}.
\end{aligned}$$

Zde jsme opět využili nerovností s e^{-1} .

Finálně srovnáme dolní mez s $1/4$:

$$\begin{aligned} 1 - 8e^{-2(c_q-1)} &> \frac{1}{4}, \\ \frac{3}{32} &> e^{-2(c_q-1)}, \\ \log \sqrt{\frac{32}{3}} &< c_q - 1. \end{aligned}$$

Porovnáme tuto dolní mez s (3.18) a dostaneme, že pro $q \geq 2$ je:

$$\frac{32}{3} \geq \frac{32}{9} \frac{q+1}{q-1}$$

Tedy pro volbu $c > \log \sqrt{32/3} + 1$ jest $t_{mix} \geq 1/2n \log n - cn$. ▲

Tento příklad byl cvičením z Levin a Peres (2017, Strana 98, Excercise 7.3), kde bylo cílem ukázat, že výsledná hranice bude záviset na počtu barev q . Bohužel dolní meze získané metodou rozlišujících statistik silně závisí na volbě funkce f . Ta se volí tak, aby dobře rozlišovala mezi stacionárním rozdělením a marginálním rozdělením řetězce začínajícího ve specifickém stavu, a aby se její momenty daly dobře odhadnout. Námi zvolená funkce $W(X)$ bohužel není nejlepší možná volba a za to tedy platíme zhoršením odhadu.

Lepší volbou by například byla:

$$Z(X) = \sum_{i=1}^n \mathbb{1}_{\{\text{vrchol } x^{(i)} \text{ má barvu } 1\}}.$$

Potom by se daly získat následující dílčí výsledky:

$$\begin{aligned} \mathbf{E}_\pi Z(X) &= \frac{n}{q}, & \text{Var}_\pi Z(X) &\leq \frac{n}{4}, \\ \mathbf{E}_1 Z(X_t) &= \frac{n}{q} + \frac{n(q-1)}{q} \left(1 - \frac{1}{n}\right)^t, & \sigma_\star^2 &= \max\{\text{Var}_\pi Z(X), \mathbf{Var}_1 Z(X_t)\} \leq \frac{n}{4}. \end{aligned}$$

Bez nutnosti dělení postupu na dvě větve bychom tak získali odhad, kde už by hodnota c_q závisela na hodnotě q a byla by zdatelně nižší, než hodnota c z předchozího příkladu. Tímto postupem bychom tedy získali odhad t_{mix} , který je lepší (= vyšší), než ten, který jsme získali s naší volbou rozlišujících statistiky.

Seznam použité literatury

- EBERLEIN, W. (1977). On Euler's infinite product for the sine. *Journal of Mathematical Analysis and Applications*, **58**(1), 147–151. ISSN 0022-247X.
- LEVIN, D. A. a PERES, Y. (2017). *Markov chains and mixing times*. American Mathematical Society, Providence, Rhode Island. ISBN 978-1-4704-2962-1.
- PRÁŠKOVÁ, Z. a LACHOUT, P. (1998). *Základy náhodných procesů*. Karolinum, Praha. ISBN 8071846880.