Sprzęganie łańcuchów Markowa

Jędrzej Hodor

Piotr Micek

Metody probabilistyczne informatyki

Instytut Informatyki Analitycznej
Uniwersytet Jagielloński

15 stycznia 2024

Sprzęganie łańcuchów Markowa

- każdy skończony, nieredukowalny i nieokresowy łańcuch Markowa ma rozkład stacjonarny
- jak długo należy prowadzić eksperyment aby rozkład aktualnego stanu był bliski rozkładowi stacjonarnemu?

jak szybko łańcuch zbiega do rozkładu stacjonarnego?

przykład

Tasujemy n kart. W każdym kroku:

- ⊛ wybieramy kartę losowo, jednostajnie i niezależnie od poprzednich kroków
- ⊛ kładziemy wybraną kartę na wierzchu talii

przykład

Tasujemy n kart. W każdym kroku:

- ⊛ wybieramy kartę losowo, jednostajnie i niezależnie od poprzednich kroków
- ⊛ kładziemy wybraną kartę na wierzchu talii

X_0 pewna ustalona permutacja początkowa

X_n permutacja kart po n krokach

przykład

Tasujemy n kart. W każdym kroku:

- ⊛ wybieramy kartę losowo, jednostajnie i niezależnie od poprzednich kroków
- ⊛ kładziemy wybraną kartę na wierzchu talii

X_0 pewna ustalona permutacja początkowa

X_n permutacja kart po n krokach

$(X_n)_{n\in\mathbb{N}}$ to łańcuch Markowa (S zbiór stanów)

skończony
nieredukowalny
nieokresowy

przykład

Tasujemy n kart. W każdym kroku:

- ⊛ wybieramy kartę losowo, jednostajnie i niezależnie od poprzednich kroków
- ⊛ kładziemy wybraną kartę na wierzchu talii

X_0 pewna ustalona permutacja początkowa

X_n permutacja kart po n krokach

$(X_n)_{n\in\mathbb{N}}$ to łańcuch Markowa (S zbiór stanów)

$$\left.\begin{array}{l}\text{skończony}\\ \text{nieredukowalny}\\ \text{nieokresowy}\end{array}\right\} (X_n)_{n\in\mathbb{N}} \text{ ma rozkład stacjonarny: } (\pi_x)_{x\in S}$$

przykład

Tasujemy n kart. W każdym kroku:

- ⊛ wybieramy kartę losowo, jednostajnie i niezależnie od poprzednich kroków
- ⊛ kładziemy wybraną kartę na wierzchu talii

X_0 pewna ustalona permutacja początkowa

X_n permutacja kart po n krokach

niech $x \in S$
$N(x)$ – zbiór stanów osiągalnych z x w jednym kroku
$|N(x)| = n$

$$\pi_x = \frac{1}{n} \sum_{y \in N(x)} \pi_y \qquad \forall_{x \in S}$$

przykład

Tasujemy n kart. W każdym kroku:

- ⊛ wybieramy kartę losowo, jednostajnie i niezależnie od poprzednich kroków
- ⊛ kładziemy wybraną kartę na wierzchu talii

X_0 pewna ustalona permutacja początkowa

X_n permutacja kart po n krokach

niech $x \in S$
$N(x)$ – zbiór stanów osiągalnych z x w jednym kroku
$|N(x)| = n$

$$\pi_x = \frac{1}{n} \sum_{y \in N(x)} \pi_y \qquad \forall_{x \in S}$$

rozkład jednostajny spełnia ten układ równań
z unikalności rozkładu stacjonarnego $\Longrightarrow$ $\pi_x = \frac{1}{n!} \quad \forall_{x \in S}$

przykład **a więc to tasowanie ma sens!**

Tasujemy n kart. W każdym kroku:

⊛ wybieramy kartę losowo, jednostajnie i niezależnie od poprzednich kroków

⊛ kładziemy wybraną kartę na wierzchu talii

X_0 pewna ustalona permutacja początkowa

X_n permutacja kart po n krokach

niech $x \in S$
$N(x)$ – zbiór stanów osiągalnych z x w jednym kroku
$|N(x)| = n$

$$\pi_x = \frac{1}{n} \sum_{y \in N(x)} \pi_y \qquad \forall_{x \in S}$$

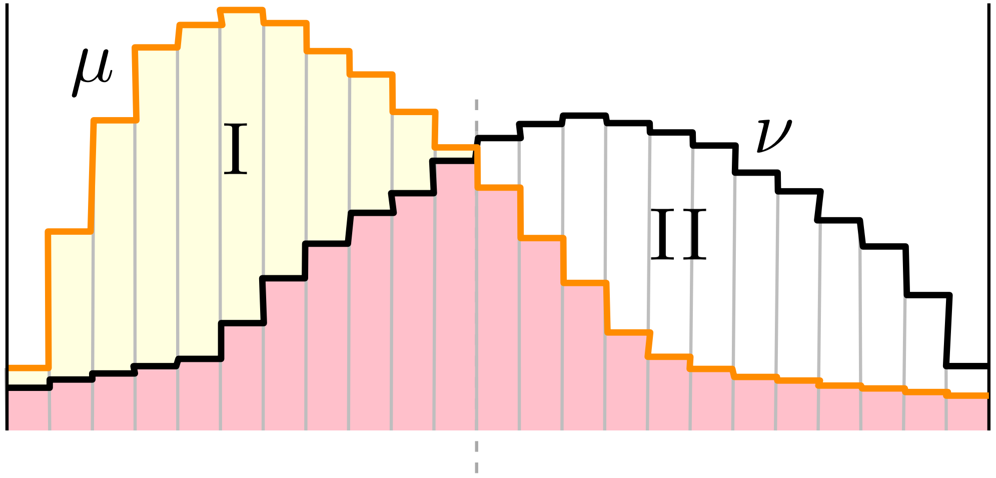
rozkład jednostajny spełnia ten układ równań
z unikalności rozkładu stacjonarnego $\Longrightarrow$ $\pi_x = \frac{1}{n!} \quad \forall_{x \in S}$

norma całkowitego wahania / total variation distance

μ, ν rozkłady prawd. nad skończonym zbiorem S

Normą całkowitego wahania μ i ν nazywamy

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



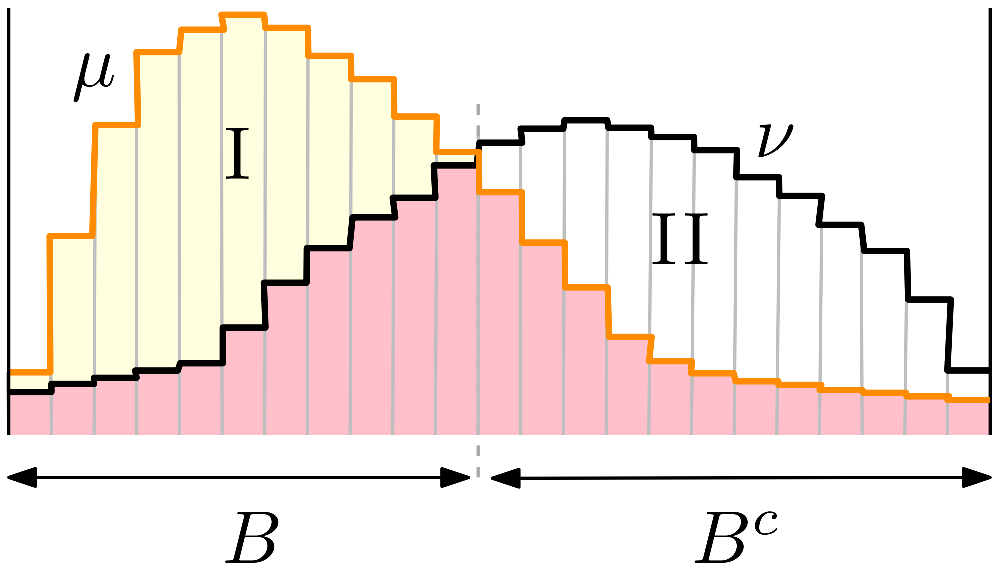
pole pod $\mu = 1$
pole pod $\nu = 1$

norma całkowitego wahania / total variation distance

μ, ν rozkłady prawd. nad skończonym zbiorem S

Normą całkowitego wahania μ i ν nazywamy

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



$$B = \{x \in S \mid \mu(x) \geqslant \nu(x)\}$$

pole pod $\mu = 1$

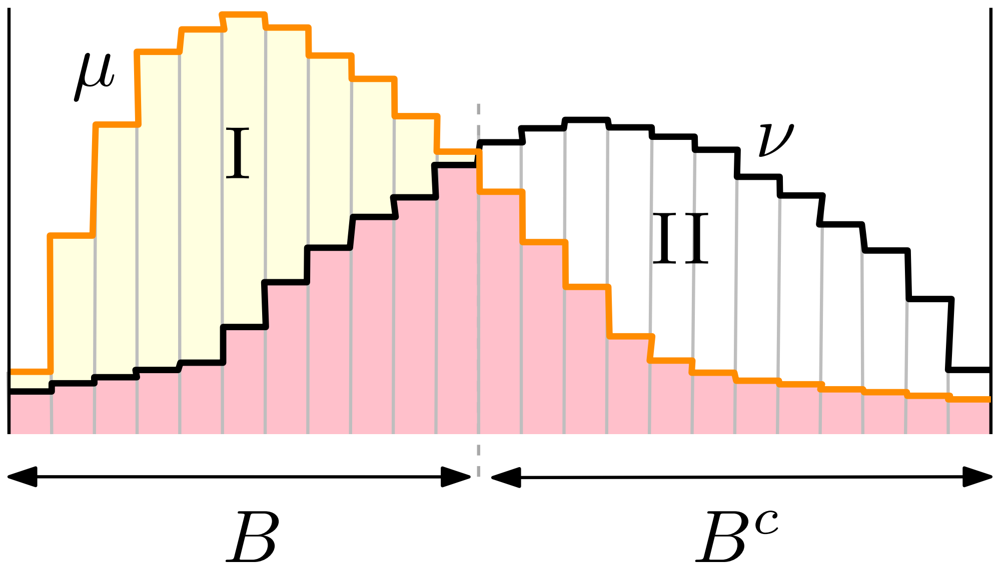
pole pod $\nu = 1$

norma całkowitego wahania / total variation distance

μ, ν rozkłady prawd. nad skończonym zbiorem S

Normą całkowitego wahania μ i ν nazywamy

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



$$B = \{x \in S \mid \mu(x) \geqslant \nu(x)\}$$

pole pod $\mu = 1$
pole pod $\nu = 1$

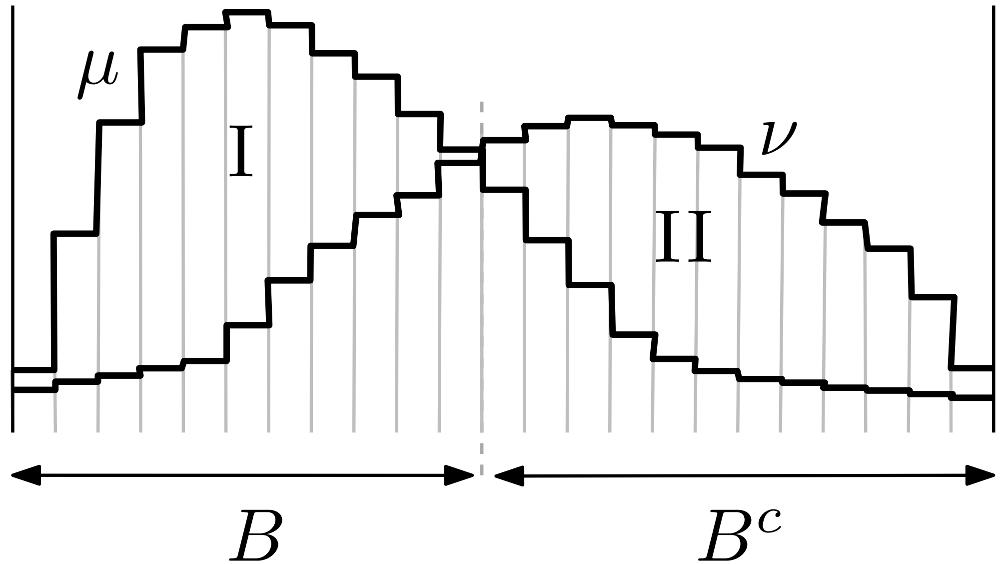
pole I $= \mu(B) - \nu(B)$
pole II $= \nu(B^c) - \mu(B^c)$ $= \|\mu(v) - \nu(v)\|_{\mathsf{TV}}$

Propozycja $$\|\mu - \nu\|_{\mathsf{TV}} = \frac{1}{2} \sum_{x \in S} |\mu(x) - \nu(x)|$$

Propozycja $\quad \|\mu - \nu\|_{\mathsf{TV}} = \dfrac{1}{2} \sum_{x \in S} |\mu(x) - \nu(x)|$

Dowód

$B = \{x \in S \mid \ \mu(x) \geqslant \nu(x)\}$



$$
\begin{aligned}
\|\mu - \nu\|_{\mathsf{TV}} &= \tfrac{1}{2}\left[\mu(B) - \nu(B) + \nu(B^c) - \mu(B^c)\right] \\
&= \tfrac{1}{2}\textstyle\sum_{x \in S} |\mu(x) - \nu(x)|
\end{aligned}
$$

Propozycja μ, ν rozkłady nad skończonym zbiorem S

$$\|\mu - \nu\|_{\mathsf{TV}} =$$

$$= \frac{1}{2} \sup \left\{ \sum_{x \in S} f(x)\mu(x) - \sum_{x \in S} f(x)\nu(x) \ : \ \max_{x \in S} |f(x)| \leqslant 1 \right\}$$

Dowód jeśli $\max_{x \in S} |f(x)| \leqslant 1$ to

$$\frac{1}{2} \left| \sum_{x \in S} f(x)\mu(x) - \sum_{x \in S} f(x)\nu(x) \right| \leqslant \frac{1}{2} \sum_{x \in S} |\mu(x) - \nu(x)|$$

$$= \|\mu - \nu\|_{\mathsf{TV}}$$

co dowodzi „$\geqslant$” w tezie

dla dowodu „$\leqslant$” $\quad B = \{x \in S \mid \ \mu(x) \geqslant \nu(x)\}$

$$f^\star(x) = \begin{cases} 1 & \text{jeśli } x \in B, \\ -1 & \text{jeśli } x \in B^c. \end{cases}$$

$$\frac{1}{2}\left[\sum_{x\in S} f^\star(x)\mu(x) - \sum_{x\in S} f^\star(x)\nu(x)\right] =$$

$$= \frac{1}{2}\left[\sum_{x\in B}(\mu(x)-\nu(x)) + \sum_{x\in B^c}(\nu(x)-\mu(x))\right]$$

$$= \|\mu-\nu\|_{\mathsf{TV}}$$

co dowodzi „$\leqslant$” w tezie $\square$

przykład z tasowaniem kart

$\pi_x = \frac{1}{n!} \quad \forall_x$ (π_x) rozkład stacjonarny

ustalmy $\varepsilon > 0$
tasujemy dopóki obecny stan ma rozkład D taki, że
x ustalona permutacja kart $\|D - \pi\|_{\mathsf{TV}} < \varepsilon$

$$\varepsilon > \|D - \pi\|_{\mathsf{TV}} = \max_{A \subseteq S} |D(A) - \pi(A)|$$
$$\geqslant |D(x) - \tfrac{1}{n!}|$$

$$\frac{1}{n!} - \varepsilon < D(x) < \frac{1}{n!} + \varepsilon$$

dla odpowiednio małego ε to może nam wystarczyć

dla porównania ...

przykład z tasowaniem kart

jeśli ktoś oszukuje przy tasowaniu
i zawsze zostawia A♠ na wierzchu to

D rozkład stacjonarny przy takim tasowaniu
π rozkład jednostajny
B zbiór permutacji z A♠ na wierzchu

$$\begin{aligned} ||D - \pi||_{\mathsf{TV}} &= \max_{A \subseteq S} |D(A) - \pi(A)| \\ &\geqslant |D(B) - \pi(B)| \\ &= |1 - \tfrac{1}{52}| = \tfrac{51}{52} \end{aligned}$$

P	macierz przejścia łańcucha Markowa $(X_t)_{t\in\mathbb{N}}$ (skończony, nieredukowalny, nieokresowy)
S	zbiór stanów
$(\pi_x)_{x\in S}$	rozkład stacjonarny
$P^t(x,\cdot)$	rozkład X_t przy założeniu, że $X_0 = x$

Definiujemy

$$\Delta_x(t) = \|P^t(x,\cdot) - \pi\|_{\mathsf{TV}} \qquad \Delta(t) = \max_{x\in S} \Delta_x(t)$$

$$\tau_x(\varepsilon) = \min\{t \mid \Delta_x(t) \leqslant \varepsilon\} \qquad \tau_{\mathsf{mix}}(\varepsilon) = \max_{x\in S} \tau_x(\varepsilon)$$

$\tau_{\mathsf{mix}}(\cdot)$ czas mieszania łańcucha $(X_t)_{t\in\mathbb{N}}$

$\tau_{\mathsf{mix}} = \tau_{\mathsf{mix}}(1/4)$ ($1/4$ to dość arbitralna stała $< 1/2$)

Jak dowodzić ograniczenia na τ_{mix} zadanego łańcucha Markowa?

μ, ν rozkłady prawd. nad skończonym zbiorem S

Sprzęganiem μ i ν nazywamy dowolną parę zmiennych losowych (X, Y) taką, że X ma rozkład μ
Y ma rozkład ν

Propozycja Niech (X, Y) będzie sprzęganiem μ i ν. Wtedy

$$||\mu - \nu||_{\mathsf{TV}} \leqslant P(X \neq Y)$$

Ponadto, istnieje takie sprzęganie dla którego w powyższej linii zachodzi równość.

Propozycja Niech (X, Y) będzie sprzęganiem μ i ν. Wtedy

$$\|\mu - \nu\|_{\mathsf{TV}} \leqslant P(X \neq Y)$$

Ponadto, istnieje takie sprzęganie dla którego w powyższej linii zachodzi równość.

Propozycja Niech (X, Y) będzie sprzęganiem μ i ν. Wtedy

$$\|\mu - \nu\|_{\mathsf{TV}} \leqslant P(X \neq Y)$$

Ponadto, istnieje takie sprzęganie dla którego w powyższej linii zachodzi równość.

Przykład μ, ν rozkłady nad $\{\text{orzeł}, \text{reszka}\}$

$\mu(\text{orzeł}) = \mu(\text{reszka}) = 0{,}5$

$\nu(\text{orzeł}) = 0{,}6 \quad \nu(\text{reszka}) = 0{,}4$

postarajmy się skonstruować (X, Y) sprzęganie μ i ν takie, że $P(X \neq Y)$ jest jak najmniejsze

	$Y = o$	$Y = r$
$X = o$	$0{,}5$	0
$X = r$	$0{,}1$	$0{,}4$

$$\|\mu - \nu\|_{\mathsf{TV}} = \tfrac{1}{2}(|0{,}5 - 0{,}6| + |0{,}5 - 0{,}4|) = 0{,}1$$

$$P(X \neq Y) = 0{,}1$$

Propozycja Niech (X, Y) będzie sprzęganiem μ i ν. Wtedy

$$\|\mu - \nu\|_{\mathsf{TV}} \leqslant P(X \neq Y)$$

Ponadto, istnieje takie sprzęganie dla którego w powyższej linii zachodzi równość.

Dowód dla dowolnego $A \subseteq S$

$$\begin{aligned} \mu(A) - \nu(A) &= P(X \in A) - P(Y \in A) \\ &\leqslant P(X \in A \cap Y \notin A) \\ &\leqslant P(X \neq Y) \end{aligned}$$

analogicznie

$$\nu(A) - \mu(A) \leqslant P(X \neq Y)$$

zatem

$$\|\mu - \nu\|_{\mathsf{TV}} = \max_{A \subseteq S} |\mu(A) - \nu(A)| \leqslant P(X \neq Y)$$

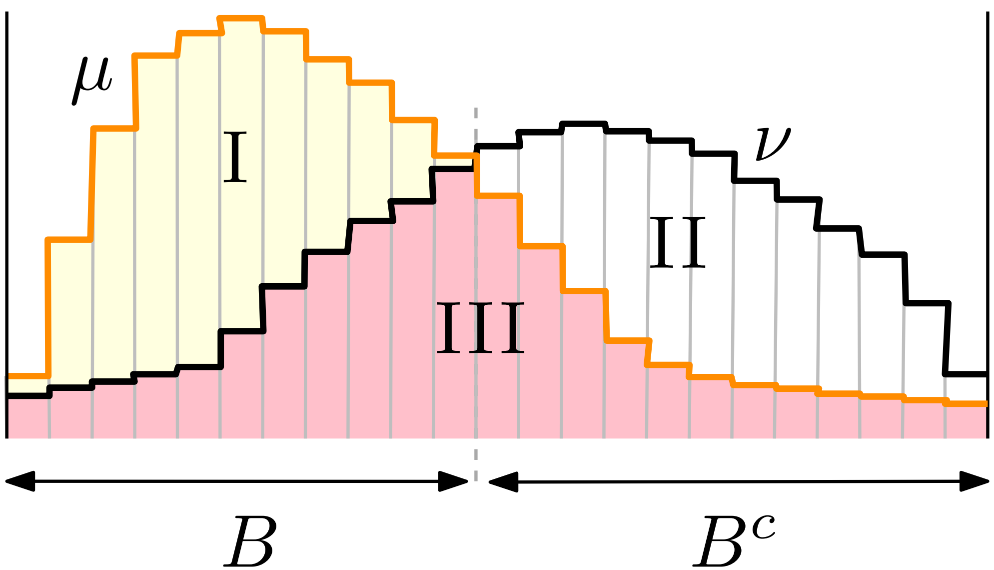
Propozycja Niech (X, Y) będzie sprzęganiem μ i ν. Wtedy

$$\|\mu - \nu\|_{\mathsf{TV}} \leqslant P(X \neq Y)$$

Ponadto, istnieje takie sprzęganie dla którego w powyższej linii zachodzi równość.

Konstrukcja sprzęgania (X, Y) rozkładów μ i ν takiego, że

$$\|\mu - \nu\|_{\mathsf{TV}} = P(X \neq Y)$$



$$p_{\mathrm{I}} + p_{\mathrm{III}} = p_{\mathrm{II}} + p_{\mathrm{III}} = 1$$
$$p_{\mathrm{I}} = p_{\mathrm{II}} = \|\mu - \nu\|_{\mathsf{TV}}$$

$$p_{\mathrm{III}} = 1 - \|\mu - \nu\|_{\mathsf{TV}}$$

$$B = \{x \in S \mid \mu(x) \geqslant \nu(x)\}$$

rzucamy monetą na której orzeł wypada z prawd. p_{III}

jeśli orzeł to

$X = Y = s$ gdzie s wybierany losowo z S z rozkładem

$$\left(\frac{1}{p_{\mathrm{III}}} \min\left(\mu(s), \nu(s)\right) \mid s \in S\right)$$

jeśli reszka to

$X = x$ gdzie x wybierany losowo z S z rozkładem

$$\begin{cases} \frac{\mu(x)-\nu(x)}{1-p_{\mathrm{III}}} & \text{jeśli } \mu(x) \geqslant \nu(x) \\ 0 & \text{wpp} \end{cases}$$

$Y = y$ gdzie y wybierany losowo z S z rozkładem

$$\begin{cases} \frac{\nu(y)-\mu(y)}{1-p_{\mathrm{III}}} & \text{jeśli } \mu(y) < \nu(y) \\ 0 & \text{wpp} \end{cases}$$

$$P(X \neq Y) = 1 - p_{\mathrm{III}} = \|\mu - \nu\|_{\mathsf{TV}}$$

Sprzęganiem łańcuchów Markowa z macierzą przejścia P i zbiorem stanów S jest dowolny łańcuch Markowa $(Z_t = (X_t, Y_t))_{t \in \mathbb{N}}$ na przestrzeni stanów $S \times S$ taki, że

$$\begin{aligned} P(X_{t+1} = x' \mid Z_t = (x, y)) &= P(x, x') \\ P(Y_{t+1} = y' \mid Z_t = (x, y)) &= P(y, y') \end{aligned}$$

dla każdego $t \geqslant 0$, $x, y, x'y' \in S$.

Sprzęganiem łańcuchów Markowa z macierzą przejścia P i zbiorem stanów S jest dowolny łańcuch Markowa $(Z_t = (X_t, Y_t))_{t\in\mathbb{N}}$ na przestrzeni stanów $S \times S$ taki, że

$$\begin{aligned} P(X_{t+1} = x' \mid Z_t = (x,y)) &= P(x,x') \\ P(Y_{t+1} = y' \mid Z_t = (x,y)) &= P(y,y') \end{aligned}$$

dla każdego $t \geqslant 0$, $x, y, x'y' \in S$.

- ⊛ zauważ, że mamy zupełną dowolność w określeniu rozkładu X_0 oraz Y_0
- ⊛ sprzęgane łańcuchy (X_t, Y_t) to dwie równoległe kopie jednego procesu
- ⊛ nie oznacza to, że te dwie kopie mają zawsze te same stany
  nie oznacza to również, że kopie te są niezależne

Będziemy zainteresowani takim sprzęganiem, które

- ⊛ stara się sprowadzić dwie kopie do tego samego stanu niezależnie od pozycji początkowej
- ⊛ jeśli kopie są w tym samym stanie to pozostają już zawsze w tym samym stanie

Lemat o sprzęganiu (skończone, nieredukowalne, nieokresowe)
Niech $((X_t, Y_t))_{t\in\mathbb{N}}$ będzie sprzęganiem łańcuchów z macierzą przejścia P i zbiorem stanów S.
Niech $T \in \mathbb{N}$, $\varepsilon > 0$ będą takie, że dla każdego $x, y \in S$

$$P(X_T \neq Y_T \mid X_0 = x,\ Y_0 = y) \leqslant \varepsilon$$

Wtedy czas mieszania łańcucha z macierzą przejścia P jest ograniczony:

$$\Delta_x(T) \leqslant \varepsilon \quad \forall\, x \in S$$
$$\tau_{\mathsf{mix}}(\varepsilon) \leqslant T$$

Lemat o sprzęganiu (skończone, nieredukowalne, nieokresowe)

Niech $((X_t, Y_t))_{t \in \mathbb{N}}$ będzie sprzęganiem łańcuchów z macierzą przejścia P i zbiorem stanów S.

Niech $T \in \mathbb{N}$, $\varepsilon > 0$ będą takie, że dla każdego $x, y \in S$

$$P(X_T \neq Y_T \mid X_0 = x,\ Y_0 = y) \leqslant \varepsilon$$

Wtedy czas mieszania łańcucha z macierzą przejścia P jest ograniczony:

$$\Delta_x(T) \leqslant \varepsilon \quad \forall\, x \in S$$

$$\tau_{\mathsf{mix}}(\varepsilon) \leqslant T$$

Lemat o sprzęganiu (skończone, nieredukowalne, nieokresowe)
Niech $((X_t, Y_t))_{t \in \mathbb{N}}$ będzie sprzęganiem łańcuchów z macierzą przejścia P i zbiorem stanów S.
Niech $T \in \mathbb{N}$, $\varepsilon > 0$ będą takie, że dla każdego $x, y \in S$

$$P(X_T \neq Y_T \mid X_0 = x,\ Y_0 = y) \leqslant \varepsilon$$

Wtedy czas mieszania łańcucha z macierzą przejścia P jest ograniczony:

$$\Delta_x(T) \leqslant \varepsilon \quad \forall\, x \in S$$
$$\tau_{\mathsf{mix}}(\varepsilon) \leqslant T$$

Dowód

zauważ, że $(X_t, Y_t)_{t \in \mathbb{N}}$ spełnia założenia
niezależnie od tego jak ustalimy rozkład X_0 czy Y_0

ustalmy dowolne $x \in S$
niech $X_0 = x$
niech Y_0 będzie wybrany losowo z rozkładu stacjonarnego π

Lemat o sprzęganiu (skończone, nieredukowalne, nieokresowe)
Niech $((X_t, Y_t))_{t\in\mathbb{N}}$ będzie sprzęganiem łańcuchów z macierzą przejścia P i zbiorem stanów S.
Niech $T \in \mathbb{N}$, $\varepsilon > 0$ będą takie, że dla każdego $x, y \in S$

$$P(X_T \neq Y_T \mid X_0 = x,\ Y_0 = y) \leqslant \varepsilon$$

Wtedy czas mieszania łańcucha z macierzą przejścia P jest ograniczony:

$$\Delta_x(T) \leqslant \varepsilon \quad \forall\, x \in S$$
$$\tau_{\mathsf{mix}}(\varepsilon) \leqslant T$$

Dowód

zauważ, że $(X_t, Y_t)_{t\in\mathbb{N}}$ spełnia założenia niezależnie od tego jak ustalimy rozkład X_0 czy Y_0

ustalmy dowolne $x \in S$
niech $X_0 = x$
niech Y_0 będzie wybrany losowo z rozkładu stacjonarnego π

wtedy $\forall_{t\geqslant 0}\ Y_t$ ma również rozkład π

niech $A \subseteq S$

$$
\begin{aligned}
P(X_T \in A) &\geqslant P(Y_T \in A \cap X_T = Y_T) \\
&= 1 - P(Y_T \notin A \cup X_T \neq Y_T) \\
&\geqslant 1 - P(Y_T \notin A) - P(X_T \neq Y_T) \\
&\geqslant P(Y_T \in A) - \varepsilon \\
&= \pi(A) - \varepsilon
\end{aligned}
$$

niech $A \subseteq S$

$$\begin{aligned} P(X_T \in A) &\geqslant P(Y_T \in A \cap X_T = Y_T) \\ &= 1 - P(Y_T \notin A \cup X_T \neq Y_T) \\ &\geqslant 1 - P(Y_T \notin A) - P(X_T \neq Y_T) \\ &\geqslant P(Y_T \in A) - \varepsilon \\ &= \pi(A) - \varepsilon \end{aligned}$$

analogicznie

$$P(X_T \in S - A) \geqslant \pi(S - A) - \varepsilon$$

czyli

$$1 - P(X_T \in A) \geqslant 1 - \pi(A) - \varepsilon$$

$$P(X_T \in A) \leqslant \pi(A) + \varepsilon$$

niech $A \subseteq S$

$$
\begin{aligned}
P(X_T \in A) &\geqslant P(Y_T \in A \cap X_T = Y_T) \\
&= 1 - P(Y_T \notin A \cup X_T \neq Y_T) \\
&\geqslant 1 - P(Y_T \notin A) - P(X_T \neq Y_T) \\
&\geqslant P(Y_T \in A) - \varepsilon \\
&= \pi(A) - \varepsilon
\end{aligned}
$$

analogicznie

$$P(X_T \in S - A) \geqslant \pi(S - A) - \varepsilon$$

czyli

$$1 - P(X_T \in A) \geqslant 1 - \pi(A) - \varepsilon$$

$$P(X_T \in A) \leqslant \pi(A) + \varepsilon$$

$$\pi(A) - \varepsilon \leqslant P(X_T \in A) \leqslant \pi(A) + \varepsilon$$

$$\forall\, x \in S$$

$$\max_{A \subseteq S} |P^T(x, A) - \pi(A)| = \|P^T(x, \cdot) - \pi\|_{\mathsf{TV}} = \Delta_x(T) \leqslant \varepsilon$$

$$\Delta(T) \leqslant \varepsilon \qquad \Rightarrow \qquad \tau_{\mathsf{mix}}(\varepsilon) \leqslant T \qquad \square$$

przykład z tasowaniem kart

P	macierz przejścia dla tasowania kart
$X_0 = x$	pewna początkowa permutacja kart
(X_t)	oryginalny łańcuch Markowa dla tasowania

$Y_0 = y$	pewna początkowa permutacja kart
(Y_t)	jeśli X_t powstało z X_{t-1} przez przełożenie karty C na szczyt to Y_t powstaje z Y_{t-1} przez przełożenie karty C na szczyt

wtedy $((X_t, Y_t))$ jest sprzęganiem łańcuchów z macierzą P

przykład z tasowaniem kart

P	macierz przejścia dla tasowania kart
$X_0 = x$	pewna początkowa permutacja kart
(X_t)	oryginalny łańcuch Markowa dla tasowania

$Y_0 = y$ pewna początkowa permutacja kart

(Y_t) jeśli X_t powstało z X_{t-1} przez przełożenie karty C na szczyt
to Y_t powstaje z Y_{t-1} przez przełożenie karty C na szczyt

wtedy $((X_t, Y_t))$ jest sprzęganiem łańcuchów z macierzą P

⊛ jeśli karta C została przełożona na szczyt w kroku T
to C będzie w tym samym miejscu w X_t i Y_t $\forall\, t \geqslant T$

⊛ jeśli T to liczba kroków po której łańcuchy się schodzą
to T ma rozkład z problemu kolekcjonera kuponów

przykład z tasowaniem kart

P	macierz przejścia dla tasowania kart
$X_0 = x$	pewna początkowa permutacja kart
(X_t)	oryginalny łańcuch Markowa dla tasowania

$Y_0 = y$ pewna początkowa permutacja kart

(Y_t) jeśli X_t powstało z X_{t-1} przez przełożenie karty C na szczyt
to Y_t powstaje z Y_{t-1} przez przełożenie karty C na szczyt

wtedy $((X_t, Y_t))$ jest sprzęganiem łańcuchów z macierzą P
$t \geqslant n \ln n + cn$

prawd., że karta C nie została przełożona po t krokach

$$\leqslant (1 - \frac{1}{n})^{n \ln n + cn}$$

$$\leqslant e^{-(\ln n + c)} = \frac{e^{-c}}{n}$$

przykład z tasowaniem kart

P	macierz przejścia dla tasowania kart
$X_0 = x$	pewna początkowa permutacja kart
(X_t)	oryginalny łańcuch Markowa dla tasowania

$Y_0 = y$	pewna początkowa permutacja kart

(Y_t) jeśli X_t powstało z X_{t-1} przez przełożenie karty C na szczyt
to Y_t powstaje z Y_{t-1} przez przełożenie karty C na szczyt

wtedy $((X_t, Y_t))$ jest sprzęganiem łańcuchów z macierzą P
$t \geqslant n \ln n + cn$

prawd., że karta C nie została przełożona po t krokach

$$\begin{aligned} &\leqslant (1 - \tfrac{1}{n})^{n \ln n + cn} \\ &\leqslant e^{-(\ln n + c)} = \frac{e^{-c}}{n} \end{aligned} \quad \Rightarrow \quad P(X_t \neq Y_t \mid X_0 = x, Y_0 = y) \leqslant e^{-c}$$

przykład z tasowaniem kart

P	macierz przejścia dla tasowania kart
$X_0 = x$	pewna początkowa permutacja kart
(X_t)	oryginalny łańcuch Markowa dla tasowania

$Y_0 = y$	pewna początkowa permutacja kart

(Y_t) jeśli X_t powstało z X_{t-1} przez przełożenie karty C na szczyt
to Y_t powstaje z Y_{t-1} przez przełożenie karty C na szczyt

wtedy $((X_t, Y_t))$ jest sprzęganiem łańcuchów z macierzą P
$t \geqslant n \ln n + cn$

prawd., że karta C nie została przełożona po t krokach

$$\leqslant (1 - \tfrac{1}{n})^{n \ln n + cn} \qquad \Rightarrow \quad P(X_t \neq Y_t \mid X_0 = x, Y_0 = y) \leqslant e^{-c}$$

$$\leqslant e^{-(\ln n + c)} = \frac{e^{-c}}{n} \qquad c = \ln(4) \qquad \tau_{\mathsf{mix}} \leqslant n \ln(4n)$$

Lemat o monotoniczności

Niech P będzie macierzą przejścia skończonego,nieredukowalnego i nieokresowego łańcucha Markowa ze zbiorem stanów S.
Wtedy dla każdych $t \geqslant 0$, $x \in S$

$$\Delta_x(t+1) \leqslant \Delta_x(t)$$

(czyli $\Delta(t+1) \leqslant \Delta(t)$)

Dowód ustalmy $t \geqslant 0$
ustalmy $x \in S$
niech $y \in S$ wybrany z rozkładu stacjonarnego π

wtedy $$\|P^t(x,\cdot) - P^t(y,\cdot)\|_{\mathsf{TV}} = \|P^t(x,\cdot) - \pi\|_{\mathsf{TV}} = \Delta_x(t)$$

niech (X_t, Y_t) będzie sprzęganiem rozkładów $P^t(x,\cdot)$ i π że

$$P(X_t \neq Y_t) = \Delta_x(t)$$

istnieje z propozycji
o sprzęganiu rozkładów

(X_{t+1}, Y_{t+1})
jeden krok macierzy P z pary stanów (X_t, Y_t) taki, że
jeśli $X_t = Y_t$ to $X_{t+1} = Y_{t+1}$

$$
\begin{aligned}
\Delta_x(t) &= P(X_t \neq Y_t) \\
&\geqslant P(X_{t+1} \neq Y_{t+1}) \\
&\geqslant \|P^{t+1}(x,\cdot) - P^{t+1}(y,\cdot)\|_{\mathsf{TV}} \\
&= \|P^{t+1}(x,\cdot) - \pi\|_{\mathsf{TV}} \\
&= \Delta_x(t+1) \quad \square
\end{aligned}
$$

propozycja
o sprzęganiu rozkładów

Twierdzenie o geometrycznej zbieżności

Niech P będzie macierzą przejścia skończonego, nieredukowalnego i nieokresowego łańcucha Markowa ze zbiorem stanów S i rozkładem stacjonarnym π.

Wtedy istnieje $\alpha \in (0,1)$ i $C > 0$ takie, że $\forall\, n \in \mathbb{N}$

$$\Delta(n) = \max_{x \in S} \|P^n(x,\cdot) - \pi\|_{\mathsf{TV}} \leqslant C\alpha^n$$

Ćwiczenie Istnieje $r \geqslant 1$ takie, że dla każdych $x, y \in S$

$$P^r(x,y) > 0$$

Dowód twierdzenia

ustalmy $r \geqslant 1$ jak w ćwiczeniu

niech $m_y = \min_{x \in S} P^r(x, y)$ dla $y \in S$ zatem $m_y > 0 \ \forall y$

niech $m = \sum_{y \in S} m_y$ $m \in (0, 1]$

skonstruuj sprzęganie $((X_t, Y_t))$ łańcuchów
o macierzy przejścia P^r takie, że

$$\forall_{\substack{t \geqslant 0 \\ y \in S}} \quad P(X_{t+1} = Y_{t+1} = y) \geqslant m_y$$

(i jeśli łańcuchy się zeszły
to już się nie rozchodzą)

wtedy

$$P(X_{t+1} = Y_{t+1}) = \sum_y P(X_{t+1} = Y_{t+1} = y) \geqslant \sum_y m_y = m$$

zatem $P(X_t \neq Y_t) \leqslant (1 - m)^t \qquad \forall\, t \geqslant 1$

niech $n \in \mathbb{N}$

niech $t \geqslant 0$ takie, że

$$n = rt + j \qquad \text{gdzie } j \in [0, r)$$

$$\alpha = (1-m)^{1/r}$$

$$C = \alpha^{-r}$$

$$\begin{aligned}
\Delta_x(n) &\leqslant \Delta_x(rt) && \text{(lemat o monotoniczności)} \\
&= \|P^{rt}(x,\cdot) - \pi\|_{\mathsf{TV}} \\
&\leqslant P(X_t \neq Y_t) && \text{(lemat o sprzęganiu)} \\
&\leqslant (1-m)^t = \alpha^{rt} = \alpha^{-j}\alpha^n \leqslant C\alpha^n
\end{aligned}$$

$\square$

przykład **łańcuch Markowa na zbiorach niezależnych ustalonej wielkości**

G graf

$$n = |V(G)|$$
$$\Delta = \Delta(G)$$

$k \geqslant 1$

M_0 ustalony k-elementowy zbiór niezależny w G

$\forall\, t \geqslant 0 \qquad M_t \to M_{t+1}$:

v wierzchołek wybrany z jednostajnego rozkładu na M_t

w wierzchołek wybrany z jednostajnego rozkładu na $V(G)$

$M_{t+1} = m(v, w, M_t)$ czyli

jeśli $M_t - \{v\} + \{w\}$ jest zbiorem niezależnym to

$$M_{t+1} = M_t - \{v\} + \{w\}$$

wpp $M_{t+1} = M_t$

załóżmy, że $k \leqslant \frac{n}{3(\Delta+1)}$

Ćwiczenie Wykaż, że $(M_t)_{t\in\mathbb{N}}$ jest skończony, nieredukowalny i nieokresowy. Wykaż, że rozkład stacjonarny tego łańcucha jest jednostajny po wszystkich k-elementowych zbiorach niezależnych w G.

chcemy ograniczyć τ_{mix} tego łańcucha

P macierz przejścia $(M_t)_{t\in\mathbb{N}}$

przechodzimy do konstrukcji sprzęgania $((X_t, Y_t))_{t\in\mathbb{N}}$ łańcuchów o macierzy przejścia P

sprzęganie $((X_t, Y_t))_{t \in \mathbb{N}}$

$\forall\, t \geqslant 0$

v wierzchołek wybrany z jednostajnego rozkładu na X_t

w wierzchołek wybrany z jednostajnego rozkładu na $V(G)$

$X_{t+1} = m(v, w, X_t)$

jeśli $v \in Y_t$ to
$Y_{t+1} = m(v, w, Y_t)$
wpp
v' wierzchołek wybrany z jednostajnego rozkładu na $Y_t - X_t$
$Y_{t+1} = m(v', w, Y_t)$

łatwo zauważyć, że $(X_t)_{t \in \mathbb{N}}$ oraz $(Y_t)_{t \in \mathbb{N}}$ mają m. przejścia P

$d_t = |X_t - Y_t|$
oczywiście $d_{t+1} \in \{d_t - 1, d_t, d_t + 1\}$
wykażemy, że jeśli $d_t > 0$ to przejście -1 jest bardziej prawdopodobne od przejścia $+1$

$$P(d_{t+1} = d_t + 1 \mid d_t > 0) \leqslant \frac{k-d_t}{k} \cdot \frac{2d_t(\Delta+1)}{n}$$

$$P(d_{t+1} = d_t - 1 \mid d_t > 0) \geqslant \frac{d_t}{k} \cdot \frac{n-(k+d_t-2)(\Delta+1)}{n}$$

zakładając $d_t > 0$ mamy

$$\begin{aligned}
E(d_{t+1}|d_t) &= d_t - P(d_{t+1} = d_t - 1) + P(d_{t+1} = d_t + 1) \\
&\leqslant d_t - \frac{d_t}{k} \cdot \frac{n-(k+d_t-2)(\Delta+1)}{n} + \frac{k-d_t}{k} \cdot \frac{2d_t(\Delta+1)}{n} \\
&= d_t\left(1 - \frac{n-(k+d_t-2)(\Delta+1)-2(k-d_t)(\Delta+1)}{kn}\right) \\
&= d_t\left(1 - \frac{n-(3k-d_t-2)(\Delta+1)}{kn}\right) \\
&\leqslant d_t\left(1 - \frac{n-3(k-1)(\Delta+1)}{kn}\right)
\end{aligned}$$

zakładając $d_t > 0$ mamy
$E(d_{t+1}|d_t) \leqslant d_t(1 - \frac{n-3(k-1)(\Delta+1)}{kn})$

a jeśli $d_t = 0$ to łańcuchy już nigdy się nie rozejdą
$E(d_{t+1}|d_t = 0) = 0$

zatem $E(d_{t+1}) = E(E(d_{t+1}|d_t)) \leqslant E(d_t)(1 - \frac{n-3(k-1)(\Delta+1)}{kn})$

$$\leqslant d_0(1 - \frac{n-3(k-1)(\Delta+1)}{kn})^t$$

$P(d_t \geqslant 1) \leqslant E(d_t) \leqslant k(1 - \frac{n-3(k-1)(\Delta+1)}{kn})^t$

$$\leqslant ke^{-t(n-3(k-1)(\Delta+1))/(kn)}$$

$\lim_{t\to\infty} ke^{-t(n-3(k-1)(\Delta+1))/(kn)} = 0$

$\Updownarrow$

$n - 3(k-1)(\Delta+1) > 0$

$\Updownarrow$

$k - 1 < \frac{n}{3(\Delta+1)}$

$\Updownarrow$

$k \leqslant \frac{n}{3(\Delta+1)}$

$P(X_t \neq Y_t) \leqslant ke^{-t(n-3(k-1)(\Delta+1))/(kn)}$

$$\leqslant \frac{1}{4}$$

dla $t = t_0 = \ln(4k) \cdot \frac{nk}{n-3(k-1)(\Delta+1)}$

$\tau_{\text{mix}} \leqslant t_0$

przykład łańcuch Markowa na kolorowaniach wierzchołkowych

G graf $\quad n = |V(G)|$, $\Delta = \Delta(G)$ (przypomnienie: $\chi(G) \leqslant \Delta + 1$)

$c \geqslant 1$ $\quad [c] = \{1, \ldots, c\}$ zbiór kolorów

M_0 ustalony kolorowanie wierzchołkowe G kolorami z $[c]$

$\forall\, t \geqslant 0 \quad M_t \to M_{t+1}$:

v wierzchołek wybrany z jednostajnego rozkładu na $V(G)$

ℓ kolor wybrany z jednostajnego rozkładu na $[c]$

M_{t+1} = recolor(v, w, M_t) czyli

jeśli $N(v)$ nie zawiera wierzchołków koloru ℓ w M_t to

M_{t+1} otrzymujemy z M_t zmieniając kolor v na ℓ

wpp $M_{t+1} = M_t$

Ćwiczenie Niech $c \geqslant \Delta + 2$.
Wykaż, że $(M_t)_{t \in \mathbb{N}}$ jest skończony, nieredukowalny i nieokresowy. Wykaż, że rozkład stacjonarny tego łańcucha jest jednostajny po wszystkich poprawnych kolorowaniach wierzchołków G kolorami z $[c]$.

chcemy ograniczyć τ_{mix} tego łańcucha

P macierz przejścia $(M_t)_{t \in \mathbb{N}}$

przechodzimy do konstrukcji sprzęgania $((X_t, Y_t))_{t \in \mathbb{N}}$ łańcuchów o macierzy przejścia P

jeśli $c \geqslant 4\Delta + 1$ to najprostszy pomysł na sprzęganie działa ...

$(X_t, Y_t) \to (X_{t+1}, Y_{t+1})$

- v wierzchołek wybrany z jednostajnego rozkładu na $V(G)$
- ℓ kolor wybrany z jednostajnego rozkładu na $[c]$
- $X_{t+1} = \mathrm{recolor}(v, \ell, X_t)$ $\quad$ $Y_{t+1} = \mathrm{recolor}(v, \ell, Y_t)$

$D_t = \{v \in V(G) \mid X_t(v) \neq Y_t(v)\}$
$d_t = |D_t|$
oczywiście $d_{t+1} \in \{d_t - 1, d_t, d_t + 1\}$
wykażemy, że jeśli $d_t > 0$ to przejście -1 jest bardziej prawdopodobne od przejścia $+1$

$$P(d_{t+1} = d_t - 1 \mid d_t > 0) \geqslant \frac{d_t}{n} \cdot \frac{c-2\Delta}{c}$$

$$\begin{aligned} P(d_{t+1} = d_t + 1 \mid d_t > 0) &\leqslant \textstyle\sum_w P(w \text{ zablokował kolor}) \\ &\leqslant \textstyle\sum_w \frac{\Delta}{n} \cdot \frac{2}{c} \\ &\leqslant \frac{\Delta d_t}{n} \cdot \frac{2}{c} \end{aligned}$$

$$\begin{aligned} E(d_{t+1}|d_t) &= d_t - P(d_{t+1} = d_t - 1) + P(d_{t+1} = d_t + 1) \\ &\leqslant d_t - \frac{d_t}{n} \cdot \frac{c-2\Delta}{c} + \frac{\Delta d_t}{n} \cdot \frac{2}{c} \\ &\leqslant d_t \cdot \left(1 - \frac{c-4\Delta}{nc}\right) \end{aligned}$$

$$E(d_{t+1}|d_t) \leqslant d_t \cdot (1 - \tfrac{c-4\Delta}{nc})$$

$$\begin{aligned} E(d_t) &= E(E(d_{t+1} \mid d_t)) \\ &\leqslant E(d_t) \cdot (1 - \tfrac{c-4\Delta}{nc}) \\ &\leqslant d_0 \cdot (1 - \tfrac{c-4\Delta}{nc})^t \end{aligned}$$

$$\begin{aligned} P(d_t \geqslant 1) &\leqslant E(d_t) \\ &\leqslant c \cdot (1 - \tfrac{c-4\Delta}{nc})^t \\ &\leqslant n e^{-t(c-4\Delta)/(nc)} \end{aligned}$$

$$\lim_{t\to\infty} n e^{-t(c-4\Delta)/(nc)} = 0$$

$$\Updownarrow$$

$$c - 4\Delta > 0$$

$$\Updownarrow$$

$$c \geqslant 4\Delta + 1$$

$$P(X_t \neq Y_t) \leqslant n e^{-t(c-4\Delta)/(nc)} \leqslant \tfrac{1}{4}$$

dla $t = t_0 = \ln(4n) \cdot \frac{nc}{c-4\Delta}$

$\tau_{\mathsf{mix}} \leqslant t_0$

jeśli $c \geqslant 2\Delta + 1$ to możemy sobie poradzić subtelniejszym sprzęganiem ...

$(X_t, Y_t) \to (X_{t+1}, Y_{t+1})$

$D_t = \{v \in V(G) \mid X_t(v) \neq Y_t(v)\}$

$A_t = \{v \in V(G) \mid X_t(v) = Y_t(v)\}$

v wierzchołek wybrany z jednostajnego rozkładu na $V(G)$

ℓ kolor wybrany z jednostajnego rozkładu na $[c]$

jeśli $v \in D_t$ to

$X_{t+1} =$ recolor(v, ℓ, X_t) $Y_{t+1} =$ recolor(v, ℓ, Y_t)

jeśli $v \in A_t$ to

$S_X(v) = X_t(N(v)) - Y_t(N(v))$

$S_Y(v) = Y_t(N(v)) - X_t(N(v))$

$f : [c] \to [c]$ bijekcja

taka, że $f(S_X(v)) \subseteq S_Y(v)$ jeśli $|S_X(v)| \leqslant |S_Y(v)|$

$S_X(v) \supseteq f^{-1}(S_Y(v))$ jeśli $|S_X(v)| > |S_Y(v)|$

$X_{t+1} =$ recolor(v, ℓ, X_t) $X_{t+1} =$ recolor$(v, f(\ell), X_t)$

$$d_t = |D_t|$$

$$d'(v) = \begin{cases} N(v) \cap A_t & \text{jeśli } v \in D_t \\ N(v) \cap D_t & \text{jeśli } v \in A_t \end{cases}$$

$$\sum_{v \in D_t} d'(v) = \sum_{v \in A_t} d'(v) =: m'$$

$$\begin{aligned} P(d_{t+1} = d_t - 1 | d_t > 0) &\geqslant \tfrac{1}{n} \textstyle\sum_{v \in D_t} \frac{c - 2\Delta + d'(v)}{c} \\ &= \tfrac{(c-2\Delta)d_t}{nc} + \tfrac{1}{nc} \textstyle\sum_{v \in D_t} d'(v) \\ &= \tfrac{(c-2\Delta)d_t + m'}{nc} \\ P(d_{t+1} = d_t + 1 | d_t > 0) &\leqslant \tfrac{1}{n} \textstyle\sum_{v \in A_t} \frac{\max(|S_X(v), |S_Y(v)||)}{c} \\ &\leqslant \tfrac{1}{nc} \textstyle\sum_{v \in A_t} d'(v) \\ &\leqslant \tfrac{m'}{nc} \end{aligned}$$

$$P(d_{t+1} = d_t - 1 | d_t > 0) \geqslant \frac{(c-2\Delta)d_t + m'}{nc}$$
$$P(d_{t+1} = d_t + 1 | d_t > 0) \leqslant \frac{m'}{nc}$$

$$\begin{aligned} E(d_{t+1} \mid d_t) &\leqslant d_t - \frac{(c-2\Delta)d_t + m'}{nc} + \frac{m'}{nc} \\ &= d_t\left(1 - \frac{c-2\Delta}{nc}\right) \end{aligned}$$

$$P(d_t \geqslant 1) \leqslant E(d_t) \leqslant n\left(1 - \frac{c-2\Delta}{nc}\right)^t \leqslant n e^{-t(c-2\Delta)/(nc)}$$

$c > 2\Delta$

$$\tau_{\mathsf{mix}} \leqslant \left\lceil \ln(4n) \cdot \frac{nc}{c-2\Delta} \right\rceil$$