

Co stała Poincaré mówi o luce spektralnej łańcucha Markowa

Wojciech Pachowiak

Wydział Matematyki i Informatyki, Uniwersytet im. Adama Mickiewicza w Poznaniu

1. TL;DR

Dany jest skończony, *jednorodny*, *nierozkładalny* i *nieokresowy* łańcuch Markowa o macierzy przejścia P . Interesuje nas, jak szybko potęgi macierzy P zbiegają do *rozkładu stacjonarnego*. Jednym ze sposobów oszacowania tej prędkości jest znalezienie ograniczenia na *lukę spektralną* macierzy P . Dla *leniwych* łańcuchów *odwracalnych* możemy to osiągnąć poprzez wyznaczenie wielkości zwanej *stałą Poincaré*. Mówi nam ona o „wąskich gardłach” występujących w podległym grafie rozważanego łańcucha Markowa (z tego powodu jest często porównywana do *stałej Cheegera*, znanej z teorii grafów) i dzięki temu dostarcza dolnego ograniczenia na lukę spektralną.

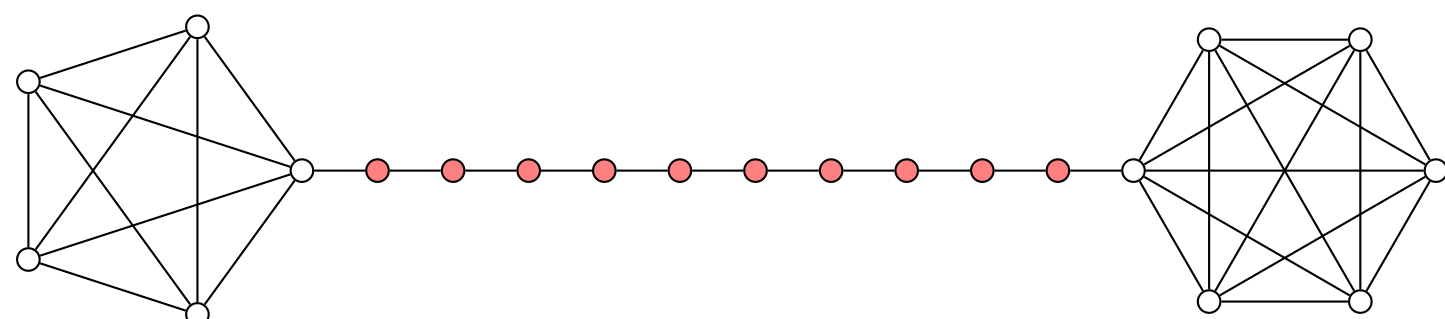


Fig. 1: Każda ścieżka między grafami K_5 i K_6 musi przejść przez „wąskie gardło” (kolor czerwony).

2. Łańcuchy Markowa

Skończony zbiór $\mathcal{X} = \{1, 2, \dots, n\}$ nazwijmy **przestrzenią stanów**. Łańcuch Markowa na $\mathcal{X}$ możemy przedstawiać na dwa sposoby:

- jako *macierz przejścia* P rozmiaru $n \times n$, która jest **kwadratowa**, **nieujemna** oraz **stochastyczna** (wyrazy w każdym wierszu sumują się do 1),
- jako skierowany graf z wagami, gdzie wagi krawędzi wychodzących z dowolnego wierzchołka sumują się do 1.

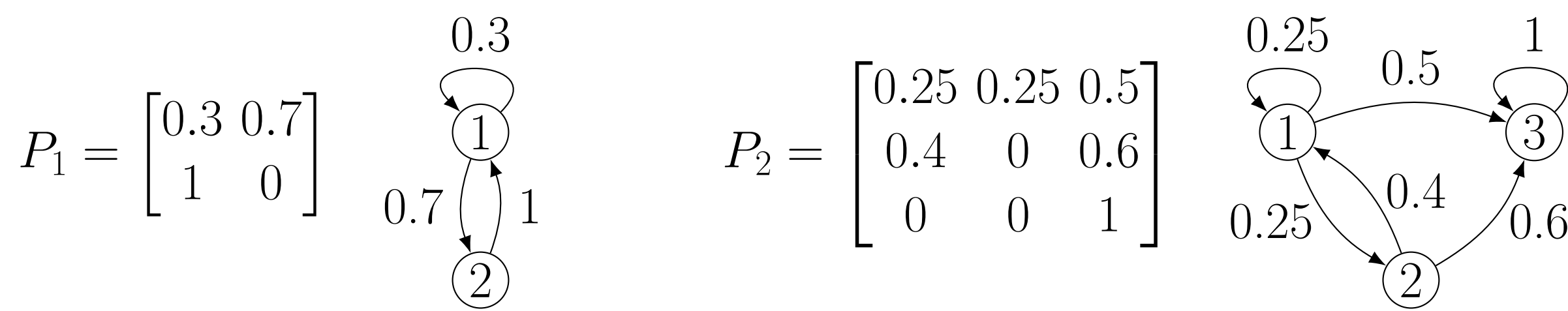


Fig. 2: Dwa przykłady skończonych jednorodnych łańcuchów Markowa wraz z ich macierzami przejścia.

Będziemy rozważać wyłącznie łańcuchy **jednorodne**, w których prawdopodobieństwa przejścia nie zmieniają się w czasie. Ponadto, każdy łańcuch spełnia **własność Markowa**, która mówi, że przyszłość zależy tylko od stanu teraźniejszego.

Fakt. Wyraz $P^t(x, y)$ mówi nam o *prawdopodobieństwie przejścia ze stanu x do stanu y w t krokach*.

W reszcie postera będziemy zakładać, że nasza macierz przejścia P jest: **nierozkładalna**, **nieokresowa** i **odwracalna** względem wektora **rozkładu stacjonarnego**, który oznaczmy przez π .

Twierdzenie. Wiersze kolejnych potęg macierzy P zbiegają do rozkładu π , czyli $P^t(x, y) \xrightarrow{t \rightarrow \infty} \pi(y)$. Ponadto, wektor π jest ściśle dodatni.

3. Luka spektralna i diagonalizacja

Twierdzenie. Wartości własne macierzy P są rzeczywiste i mają postać:

$$1 = \lambda_1 > \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_{n-1} \geq \lambda_n > -1.$$

Definicja. Zdefiniujmy liczbę:

$$\lambda_* = \max \{|\lambda_i| : 2 \leq i \leq n\} = \max \{\lambda_2, -\lambda_n\}.$$

Wtedy $1 - \lambda_*$ nazywamy **luką spektralną** macierzy P .

Twierdzenie. Macierz P jest diagonalizowalna, tzn.

$$P^t(x, y) = \sum_{j=1}^n f_j(x) u_j(y) \lambda_j^t = \pi(y) + \sum_{j=2}^n f_j(x) u_j(y) \lambda_j^t = \pi(y) + O(\lambda_*^t),$$

gdzie f_i i u_i^T są, odpowiednio, prawym i lewym wektorem własnym odpowiadającym wartości własnej λ_i .

Wniosek. Ponieważ $\lambda_*^t \xrightarrow{t \rightarrow \infty} 0$, więc luka spektralna mówi nam o szybkości zbiegania do rozkładu stacjonarnego (im bliższa jest zeru, tym wolniejsze zbieganie).

4. Spektrum leniwego łańcucha

Definicja. Łańcuch o macierzy przejścia $L = (I + P)/2$ nazywamy **leniwą wersją** łańcucha P . Innymi słowy,

$$L(x, y) = \begin{cases} \frac{1}{2} + P(x, x), & x = y \\ \frac{1}{2} P(x, y), & x \neq y. \end{cases}$$

Równoważna interpretacja: przed każdym krokiem rzucamy monetą i albo się nie ruszamy, albo ruszamy się zgodnie z P .

Fakt. Jeśli π jest rozkładem stacjonarnym łańcucha Markowa P , to jest również rozkładem stacjonarnym łańcucha L .

Fakt. Niech β_i będzie i -tą największą wartością własną łańcucha L . Wtedy $\beta_i = (1 + \lambda_i)/2$. Zatem, wartości własne L są nieujemne i należą do przedziału $[1, 0)$ (w przypadku P był to przedział $[1, -1)$).

Fakt. Luka spektralna łańcucha L wynosi $1 - \beta_*$, gdzie $\beta_* = \beta_2 = (1 + \lambda_2)/2$.

Wniosek. Aby oszacować lukę spektralną leniwego łańcucha, potrzebujemy wyłącznie β_2 bez konieczności znajdowania β_n .

5. Stała Poincaré

Przez $G = (\mathcal{X}, E)$ oznaczmy podległy graf macierzy przejścia L .

Definicja. Dla każdej uporządkowanej pary różnych stanów $i, j \in \mathcal{X}$ wybierzmy arbitralnie jedną skierowaną ścieżkę prostą γ_{ij} (wierzchołki mogą się powtarzać, ale krawędzie już nie; kierunek krawędzi musi być zgodny). Przez Γ oznaczmy zbiór wszystkich wybranych ścieżek; moc tego zbioru to $|\Gamma| = |\mathcal{X}|(|\mathcal{X}| - 1)$. Wtedy **stałą Poincaré** nazywamy liczbę

$$\kappa = \kappa(\Gamma) = \max_{(i,j) \in E} \left[\sum_{\gamma_{ij} \ni (i,j)} |\gamma_{xy}|_Q \pi(x) \pi(y) \right],$$

gdzie $|\gamma_{xy}|_Q = \sum_{(a,b) \in \gamma_{xy}} \frac{1}{\pi(a)L(a,b)}$.

Obserwacja. Liczba κ mierzy „wąskie gardła” (patrz: Fig. 1) w grafie G . Jest duża, gdy 1) wiele ścieżek z Γ przechodzi przez tę samą krawędź, 2) wielkość $|\gamma_{xy}|_Q$ jest duża, czyli ścieżka γ_{xy} jest długa i składa się z krawędzi o niskich prawdopodobieństwach.

Twierdzenie. Mamy następujące dolne oszacowanie na lukę spektralną macierzy L :

$$\frac{1}{\kappa} \leq 1 - \beta_*.$$

6. Prosty spacer losowy na grafie

Przyjmijmy, że S jest leniwą wersją prostego spaceru losowego na grafie G . Wtedy

$$S(x, y) = \begin{cases} \frac{1}{2 \deg(x)}, & x \neq y \\ \frac{1}{2}, & x = y \end{cases} \quad \text{oraz} \quad \pi(x) = \frac{\deg(x)}{2|E|}.$$

Lukę spektralną łańcucha S oznaczmy przez $1 - \beta_*$.

Fakt. Dla dowolnego zbioru ścieżek Γ mamy:

$$\kappa \leq \frac{L_\Gamma b_\Gamma \Delta^2}{|E|},$$

gdzie Δ jest maksymalnym stopniem wierzchołka w grafie G , $L_\Gamma := \max \{|\gamma| : \gamma \in \Gamma\}$ jest długością najdłuższej ścieżki z Γ , natomiast $b_\Gamma := \max_{e \in E} |\{\gamma \in \Gamma : e \in \gamma\}|$ jest liczbą ścieżek z Γ przechodzących przez „najpopularniejszą” krawędź.

Wniosek. Korzystając z Twierdzenia z poprzedniej sekcji, otrzymujemy:

$$\frac{|E|}{L_\Gamma b_\Gamma \Delta^2} \leq 1 - \beta_*.$$