

Politechnika Warszawska

WYDZIAŁ MATEMATYKI
I NAUK INFORMACYJNYCH



Praca dyplomowa licencjacka

na kierunku Matematyka i Analiza Danych

Metody Monte Carlo oparte na łańcuchach Markowa związane z pewnym
problemem z ekologii

Julia Girtler

Numer albumu 320555

promotor

dr inż. Joanna Matysiak

WARSZAWA 2024

Streszczenie

Metody Monte Carlo oparte na łańcuchach Markowa związane z pewnym problemem z ekologii

Niniejsza praca skupia się na rozwiązaniu problemu z zakresu ekologii związanego z współwystępowaniem gatunków ptaków na poszczególnych wyspach. Wprowadza miary współzależności, które opisują poziom współzawodnictwa gatunków. Wyjściowy problem ekologiczny daje się sformułować w języku macierzowym i dotyczy określenia prawdopodobieństwa wylosowania macierzy składającej się z zer i jedynek (o ustalonych sumach w wierszach i kolumnach), dla której miary współzależności będą miały takie same lub większe wartości niż badane. W pracy przedstawione zostaną sposoby generowania macierzy zgodnie z rozkładem jednostajnym przy użyciu algorytmu Metropolisa-Hastingsa. Szczegółowo zostanie zaprezentowana technika sklejania łańcuchów Markowa, co okaże się użyteczne w obrębie problematyki pracy.

Słowa kluczowe: łańcuch Markowa, algorytm Metropolisa-Hastingsa, spacer losowy po grafie, szybkość zbieżności, MCMC

Abstract

Some Markov Chain Monte Carlo methods related to a problem of ecology

This paper focuses on solving an ecological problem related to the co-occurrence of bird species on individual islands. It introduces measures of interdependence that describe the level of competition among species. The initial ecological problem can be formulated in terms of matrices and concerns determining the probability of drawing a matrix consisting of zeros and ones (with fixed sums in rows and columns), for which the interdependence measures will have the same or greater values than those examined. The paper presents methods for generating matrices according to a uniform probability distribution using the Metropolis-Hastings algorithm. Additionally, it details the technique of lumping Markov chains, which proves useful within the scope of the problem addressed in the paper.

Keywords: Markov chain, random walk, Metropolis-Hastings algorithm, rate of convergence, MCMC

Spis treści

1. Wprowadzenie	11
2. Losowe generowanie macierzy	15
2.1. Macierze zamienne	15
2.2. Algorytm zamian 2×2	16
2.3. Spacer po grafie	17
2.4. Algorytm Metropolis-Hastingsa	21
3. Szybkość zbieżności	26
3.1. Wartości własne macierzy stochastycznej	26
3.2. Ograniczenie normy	28
3.3. Wnioski	30
4. Sklejanie łańcuchów Markowa	31
4.1. Idea sklejanego łańcucha Markowa	31
4.2. Macierzowa reprezentacja sklejanego łańcucha Markowa	33
4.3. Wartości własne macierzy przejścia po sklejaniu	35
5. Podsumowanie	37
5.1. Analiza przykładu	37
5.2. Wnioski	41

1. Wprowadzenie

Otoczający nas świat przyrody jest obserwowany przez ludzkość od samego początku. Wraz z postępem technologicznym i naukowym jesteśmy w stanie odpowiedzieć na coraz więcej pytań pojawiających się przy analizowaniu życia zwierząt.

Przyglądając się życiu ptaków, możemy dostrzec, że niektóre gatunki występują na danych wyspach, a inne nigdy się na nich nie pojawiają. Tworzą one między sobą społeczności, w których niektóre z gatunków unikają siebie, nie bytując wspólnie na jednej wyspie, inne natomiast dzielą ze sobą swoje siedliska. Problem ten opisany jest w [3]. Czy jesteśmy w stanie zaobserwować pewne wzorce dotyczące rozmieszczenia gatunków ptaków? Jeśli tak, to na ile prawdopodobne jest, że prawidłowości te są wynikiem przypadku? W niniejszej pracy wykorzystujemy metodę Monte Carlo Markov Chain (MCMC) do odpowiedzenia na zadane pytania.

W pierwszym rozdziale przyjrzymy się tabeli występowania gatunków ptaków na wyspach Galapagos, wprowadzając tym samym definicje niezbędne do jej dalszego studiowania. Przedstawiony zostanie problem ekologiczny, który po odpowiednim przeformułowaniu, będzie miał matematyczne rozwiązanie. Naszym głównym zadaniem będzie znalezienie sposobu na generowanie macierzy zgodnie z jednostajnym rozkładem prawdopodobieństwa. W drugim rozdziale przedstawimy algorytm zamian 2×2 , za pomocą którego będziemy losować macierze zer i jedynek. Zostanie on także zinterpretowany jako spacer losowy po grafie, dzięki czemu będziemy mogli wyrazić go za pomocą odpowiednio zdefiniowanego łańcucha Markowa oraz zastosować algorytm Metropolisa-Hastingsa. W trzecim rozdziale zbadamy potrzebną ilość kroków, aby każda macierz występowania ptaków była generowana z jednakowym prawdopodobieństwem. Wykażemy związek między szybkością zbieżności, a drugą największą wartością własną macierzy przejścia odpowiadającej łańcuchowi Markowa związanego ze spacerem po grafie. W czwartym rozdziale przedstawimy strategię sklejanego łańcucha Markowa. Na samym końcu zaprezentujemy przykład, na którym podsumujemy dotychczas zebrane wyniki.

Tabela 1.1: Macierz występowania gatunków ptaków na wyspach Galapagos

Gatunek	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	suma
A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	17
B	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	1	1	14
C	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	17
D	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	13
E	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	12
F	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	11
G	1	1	1	1	1	0	1	1	0	0	0	0	0	1	0	1	1	10
H	1	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	10
I	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	10
J	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	6
K	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2
L	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
M	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
suma	11	10	10	10	10	9	9	9	8	8	7	4	4	4	3	3	3	122

Powyższa tabela przedstawia występowanie poszczególnych gatunków ptaków na różnych wyspach. Wartość 1 informuje o tym, że wyspa jest zamieszkiwana przez dany gatunek, a 0, że nie jest ona siedliskiem tego gatunku.

Na podstawie danych w tabeli tj. macierzy składającej się z jedynek i zer o wymiarach $\# \text{gatunki} \times \# \text{wyspy}$ możemy określić stopień rywalizacji dla danego zbioru gatunków ptaków, czyli na ile jedne gatunki wykluczają występowanie innych. Poniżej wprowadzimy kilka miar współzależności.

Definicja 1.1. *Szachownicą* nazwiemy parę gatunków, taką, że iloczyn skalarny wektorów powstałych z wierszy odpowiadającym tym gatunkom wynosi 0. Intuicyjnie jest to taka para gatunków, która nie występuje razem na żadnej wyspie.

Przykład 1.2. Dla naszych danych szachownicą są gatunki M i J.

$$J = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T$$

$$M = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T$$

Otrzymujemy, że $J^T \bullet M = 0$, czyli gatunki J i M tworzą szachownicę.

Definicja 1.3. *Jednostkową szachownicą* nazywamy parę gatunków i parę wysp, taką, że każdy z gatunków mieszka na dokładnie jednej wyspie. Pojedyncze szachownice mogą mieć dwie postacie:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (1.1)$$

Przykład 1.4. Jednostkową szachownicę tworzą gatunki J i M z wyspami 4 i 5. Ma ona postać:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Dla naszych danych istnieje 10 szachownic spośród $\binom{13}{2} = 78$ możliwych par. Jednostkowych szachownic jest natomiast 333 spośród $\binom{13}{2}\binom{17}{2} = 10608$.

Będziemy próbowali odpowiedzieć na następujące pytanie.

Problem ekologiczny

Mając daną macierz występowania gatunków i wartości miar współzależności, jakie jest prawdopodobieństwo, że gdy wylosujemy macierz, której sumy wierszy i kolumn są takie same co wyjściowej macierzy miary współzależności będą miały takie same lub większe wartości.

W teorii rozwiązanie powyższego problemu jest bardzo proste. Wystarczy podzielić liczbę macierzy spełniających dany warunek przez liczbę wszystkich możliwych macierzy o konkretnej sumie wierszy i kolumn. W praktyce jednak jest to niewykonalne. Dla naszego przykładu macierzy o takiej samej sumie kolumn i wierszy jest ponad 6.71×10^{16} . Jest to na tyle duża liczba, że wygenerowanie tylu macierzy, a następnie sprawdzenie, czy spełniają one warunki jest nieosiągalne.

Intuicyjne rozwiązanie

$$p = \frac{\text{liczba wszystkich macierzy z wymaganym poziomem konkurencyjności}}{\text{liczba wszystkich macierzy}}$$

Jesteśmy jednak w stanie odpowiedzieć na to pytanie, ponieważ powołując się na prawo wielkich liczb, powyższy wynik z dużym prawdopodobieństwem będzie wynosił:

$$p = \frac{\text{liczba losowych macierzy z wymaganym poziomem konkurencyjności}}{\text{liczba losowych macierzy}}$$

gdzie przez losowe macierze rozumiemy macierze generowane zgodnie z rozkładem jednostajnym.

Pozostaje nam zatem wygenerować losowo, zgodnie z rozkładem jednostajnym, macierze 0 i 1 z konkretnymi sumami wierszowymi i kolumnowymi. Jest to zadanie, które da się rozwiązać stosując matematyczne narzędzia.

Definicja 1.5. Niech $n, m \in \mathbb{N}$ oraz $w \in \mathbb{R}^n$, $k \in \mathbb{R}^m$. Przez $\mathcal{A}(w, k)$ oznaczać będziemy zbiór macierzy z 1 i 0, w których wektor sum wierszowych jest równy w , a wektor sum kolumnowych jest równy k .

Przykład 1.6. Niech $w = (1, 1, 1)$, $k = (2, 1)$. Wówczas wszystkie macierze należące do rodziny $\mathcal{A}(w, k)$ to:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Matematyczne wyzwanie

W jaki sposób generować macierze należące do zbioru $\mathcal{A}(w, k)$, aby każda z nich była wybierana z jednakowym prawdopodobieństwem?

Celem niniejszej pracy jest przedstawienie sposobu rozwiązywania powyższego problemu.

2. Losowe generowanie macierzy

W poniższym rozdziale skupimy się na rozwiązaniu problemu generowania macierzy należących do zbioru $\mathcal{A}(w, k)$ zgodnie z rozkładem jednostajnym na $\mathcal{A}(w, k)$. W tym celu zostanie wprowadzony algorytm zamian, a następnie zmodyfikowany poprzez zastosowanie MCMC (Monte Carlo Markov Chain). Pokazane zostanie również, że algorytmy zamian mogą być przedstawione za pomocą grafów, co w dalszych rozdziałach pomoże nam je ulepszać.

2.1. Macierze zamienne

Definicja 2.1. *Zamianą* nazwiemy modyfikację macierzy polegającą na wyborze dwóch kolumn, bądź wierszy, a następnie zamianie ich miejscami (założmy, że wybraliśmy i -ty i j -ty wierszy macierzy, wówczas w zmodyfikowanej macierzy i -ty wiersz jest teraz j -tym wierszem, a j -ty - i -tym).

Definicja 2.2. O dwóch macierzach powiemy, że są *zamiennie* jeśli mają takie same sumy wierszowe i kolumnowe.

Jeśli rozpatrzymy macierze 2×2 to jedynymi zamiennymi macierzami są pojedyncze szachownice:

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Jeśli w większej macierzy wybierzemy podmacierz A_1 lub A_2 i zastąpimy ją jej zamiennym odpowiednikiem to sumy wierszowe i kolumnowe pozostaną takie same.

Twierdzenie 2.3. Dowolne dwie macierze $A, A^* \in \mathcal{A}(w, k)$ o wymiarach $m \times n$ są połączone ze sobą skończoną liczbą zamian podmacierzy 2×2 .

Dowód twierdzenia znajdują się w książce [9].

2.2. Algorytmy zamian 2×2

Bazując na powyższym twierdzeniu, ekolodzy stosują następujący algorytm do losowego generowania macierzy.

Algorytm zamian 2×2

1. Losujemy bez zwracania z jednakowym prawdopodobieństwem dwa rzędy oraz dwie kolumny
2. Czy wybrana podmacierz 2×2 jest zamienna? Jeśli tak to dokonujemy tej zamiany, jeśli nie to wracamy do kroku pierwszego

Powtarzając powyższy algorytm dużą liczbę razy otrzymamy macierz losową, jednak okazuje się, że nie wszystkie macierze uzyskane za pomocą poniższego algorytmu otrzymuje się z jednakowym prawdopodobieństwem.

Powyższy algorytm pokażemy na prostym przykładzie.

Przykład 2.4. Rozpatrzmy wszystkie macierze należące do $\mathcal{A}(w, k)$ gdzie $w = (1, 1, 1)$, $k = (1, 1, 1)$. Mają one postać:

$$a = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad c = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$d = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad e = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad f = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Wybierając drugi i trzeci wiersz oraz pierwszą i drugą kolumnę macierzy a otrzymujemy macierz 2×2 : $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ Jest ona zamienna, więc możemy zamienić ją na macierz $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ W ten sposób otrzymujemy macierz b . Jest to jeden krok algorytmu 2×2 .

2.3. Spacer po grafie

Na powyższym przykładzie mogliśmy zaobserwować, że pewne macierze są ze sobą połączone poprzez pojedyncze zamiany. Ten fakt pozwala nam na interpretację graficzną algorytmu zamian 2×2 . Możemy przyjąć, że macierze należące do zbioru $\mathcal{A}(w, k)$ są wierzchołkami grafu, które są połączone krawędzią wtedy i tylko wtedy, gdy z jednego wierzchołka do drugiego da się przejść stosując pojedynczą zamianę 2×2 .

Na potrzeby dalszych rozważań przypomnimy podstawowe twierdzenia dotyczące teorii grafów.

Definicja 2.5. Niech $G(V, E)$ będzie grafem z $|V|$ wierzchołkami i $|E|$ krawędziami.

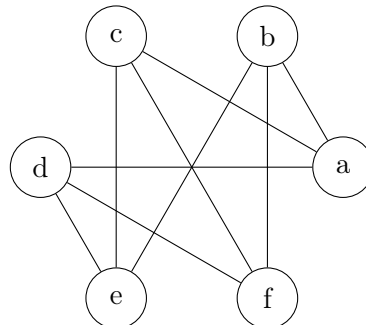
Spacerem losowym $\{X_t\}$ po grafie G nazywamy proces, w którym z każdego wierzchołka $v \in V$ grafu G przechodzimy do innego wierzchołka (jeśli te są ze sobą połączone krawędzią) z prawdopodobieństwem $\frac{1}{\deg(v)}$, gdzie $\deg(v)$ to stopień wierzchołka v .

Losowy spacer po grafie $G(V, E)$ jest łańcuchem Markowa $\{X_t\}$ o przestrzeni stanów V , gdzie z macierz przejścia $P = (p_{ij})_{i,j \in V}$ wyraża się w następujący sposób:

$$p_{ij} = \begin{cases} \frac{1}{\deg(i)}, & \text{jeśli } ij \in E \\ 0, & \text{wpp} \end{cases} \quad (2.1)$$

Przez $p_{ij}^{(t)}$ oznaczmy prawdopodobieństwo, że po t krokach łańcuch ze stanu i przejdzie do stanu j .

Wróćmy do przykładu 2.4. Niech wierzchołkami naszego grafu G będą macierze tj. $V = \{a, b, c, d, e, f\}$ oraz niech wierzchołki będą połączone, jeśli z jednej macierzy do drugiej można przejść w jednym kroku algorytmu zamiany 2×2 . Wówczas nasz graf ma następującą postać:



Łańcuch Markowa związany z tym grafem ma macierz przejścia $P =$

$$\begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \end{pmatrix}$$

Definicja 2.6. *Rozkładem stacjonarnym* π łańcucha Markowa $(X_n)_n$ o macierzy przejścia P nazwiemy wektor π taki, że $\pi P = \pi$.

Twierdzenie 2.7 (Twierdzenie ergodyczne). Niech $(X_n)_n$ będzie nieprzywiedlnym i nieokresowym łańcuchem Markowa o przestrzeni stanów S z rozkładem stacjonarnym π .

Wówczas rozkład ten jest jedyny oraz dla każdego $i, j \in V$ $\lim_{t \rightarrow \infty} p_{ij}^{(t)} = \pi_j$.

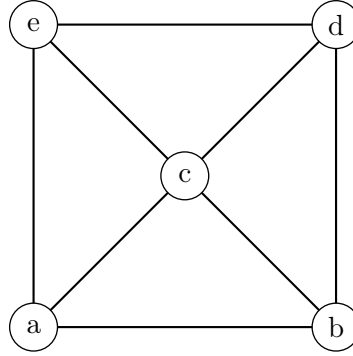
Wróćmy do naszego problemu. Zamierzamy losować macierze z jednakowym prawdopodobieństwem. Przekładając to na spacer losowy po grafie, oznacza to, że niezależnie od wyboru punktu startowego po dużej liczbie t kroków znajdziemy się w każdym z wierzchołków z tym samym prawdopodobieństwem. W naszym przykładzie oznaczałoby to, że jest $\frac{1}{6}$ szans, że znajdziemy się w wierzchołku a , $\frac{1}{6}$ szans, że znajdziemy się w wierzchołku b i tak dalej. Możemy zatem obliczyć rozkład stacjonarny łańcucha, i jeśli spełnione są warunki Twierdzenia ergodycznego to przyjąć go jako rozkład graniczny. Interpretacją rozkładu granicznego jest interesujące nas prawdopodobieństwo tj. jaka jest szansa, że po t krokach łańcuch znajdzie się w i -tym stanie. Zatem aby losować macierze z jednakowym prawdopodobieństwem rozkład graniczny musi mieć postać rozkładu jednostajnego.

Rozkład stacjonarny dla naszego przykładu ma postać $\pi = (\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$. Oznacza to, że algorytm zamiany 2×2 dla wybranego przykładu działa poprawnie i z prawdopodobieństwem jednakowym dla każdej macierzy spełniającej warunki generuje losową macierz.

Wybrany przykład jest jednak podany dla niewielkiej liczby gatunków i wysp. Ponadto, w każdym wierszu i kolumnie suma jest identyczna, tj. równa 1. W rzeczywistości macierze są większych rozmiarów oraz o bardziej różnorodnych wektorach sum i wierszy. Czy algorytm zmian 2×2 działa dla większej liczby zmiennych?

Przykład 2.8. Niech dany będzie zbiór macierzy $\mathcal{A}(w, k)$ gdzie $w = (2, 1, 1)$, $k = (1, 2, 1)$. Wówczas wszystkie możliwe macierze to:

$$a = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad c = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad d = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad e = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



Odpowiadająca macierz przejścia jest postaci:

$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

Zatem rozkład stacjonarny takiego łańcucha ma postać $\pi = (\frac{3}{16}, \frac{3}{16}, \frac{4}{16}, \frac{3}{16}, \frac{3}{16})$.

Możemy zauważyć zatem, że w stanie c po t krokach znajdziemy się z większym prawdopodobieństwem niż w stanie a, b, d, e . Oznacza to, że macierz:

$$c = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

będzie generowana z większym prawdopodobieństwem.

Można zauważyć, że rozkład graniczny ma bezpośredni związek ze stopniem wierzchołków.

W ogólności zachodzi następujące twierdzenie:

Twierdzenie 2.9. Niech dany będzie łańcuch Markowa o przestrzeni stanów V i odpowiadający mu spacer losowy po grafie $G(V, E)$, przy czym $ij \in E$ jeśli $p_{ij} \neq 0$. Wówczas rozkład stacjonarny i przy spełnieniu odpowiednich warunków rozkład graniczny ma postać:

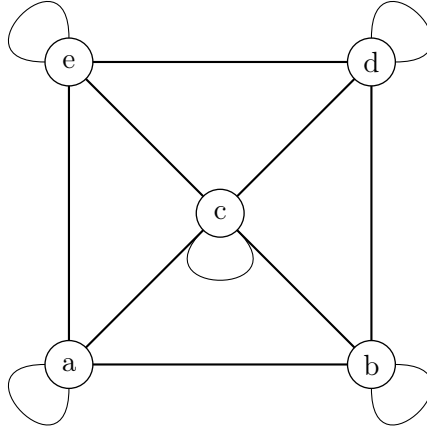
$$\pi_i = \frac{\deg(i)}{2|E|} \quad (2.2)$$

Dowód. Przez $P_{\cdot i}$ oznaczmy i -tą kolumnę macierzy P . Wiemy również, że $p_{ji} = \frac{1}{\deg(j)}$. Zauważmy, że zachodzi równość:

$$\pi P_{\cdot i} = \sum_{j=1}^{|V|} \pi_j p_{ji} = \sum_{ij \in E} \frac{\deg(j)}{2|E|} \frac{1}{\deg(j)} = \sum_{ij \in E} \frac{1}{2|E|} = \frac{\deg(i)}{2|E|} = \pi_i \quad (2.3)$$

Jeśli spełnione są warunki twierdzenia ergodycznego, to rozkład stacjonarny jest jedyny, więc równy $\pi_i = \frac{\deg(i)}{2|E|}$ oraz jest on rozkładem granicznym. \square

Zatem, aby rozkład graniczny był rozkładem jednostajnym każdy wierzchołek grafu musi mieć ten sam stopień. Nasuwa się więc, poprawiony algorytm zamiany 2×2 . Spróbujmy liczyć każdy krok, niezależnie czy wylosowana podmacierz 2×2 jest zamienna. Wówczas w przykładzie 2.8 macierz przejścia i odpowiadający temu łańcuchowi graf miałyby taką postać:



$$P = \begin{pmatrix} \frac{2}{3} & \frac{1}{9} & \frac{1}{9} & 0 & \frac{1}{9} \\ \frac{1}{9} & \frac{2}{3} & \frac{1}{9} & \frac{1}{9} & 0 \\ \frac{1}{9} & \frac{1}{9} & \frac{5}{9} & \frac{1}{9} & \frac{1}{9} \\ 0 & \frac{1}{9} & \frac{1}{9} & \frac{2}{3} & \frac{1}{9} \\ \frac{1}{9} & 0 & \frac{1}{9} & \frac{1}{9} & \frac{2}{3} \end{pmatrix}, \pi = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$$

Istnieje $\binom{3}{2}\binom{3}{2} = 9$ możliwości wyboru podmacierzy 2×2 , a więc każdy wierzchołek musi być dziewiętego stopnia. W powyższym przykładzie pętle przy wierzchołkach a, b, d, e są sześciokrotne, ponieważ za każdym razem gdy podmacierz 2×2 nie jest zamienna dodajemy jedną krawędź, analogicznie przy wierzchołku c pięciokrotne.

2.4. ALGORYTM METROPOLISA-HASTINGSA

Dla przykładów z macierzami o dużej liczbie kolumn i wierszy liczba pętli staje się na tyle duża, że poprawiony algorytm zamiany 2×2 jest mało wydajny. Wprowadźmy zatem algorytm, który będzie w stanie rozwiązać nasz problem.

2.4. Algorytm Metropolisa-Hastingsa

W tym podrozdziale przedstawimy bardziej efektywny sposób rozwiązywania powyższego problemu, który pozwoli nam na pozbycie się komplikacji związanymi z wielokrotnymi pętlami przy wierzchołkach.

Algorytm Metropolisa-Hastingsa jest przykładem próbkowania Monte Carlo łańcuchami Markowa. MCMC używane jest w sytuacjach, gdy chcemy losować pewne zmienne X_i z zadaniem przez nas prawdopodobieństwem p_i . W tym celu tworzymy łańcuch Markowa, którego rozkład stacjonarny π będzie mieć współrzędne p_i .

Algorytm Metropolisa-Hastingsa

Chcemy skonstruować łańcuch Markowa $(X_n)_n$ taki, że rozkład stacjonarny jest rozkładem jednostajnym. Niech $Q = (q_{ij})_{i,j \in S}$ będzie macierzą przejścia dowolnego nieprzywiedlnego łańcucha na S , gdzie S to przestrzeń stanów. Załóżmy, że umiemy próbować z wierszy macierzy Q .

Niech $X_n = i$, wówczas:

1. wybieramy *stan propozycji* - jest to element z przestrzeni stanów S wybrany zgodnie z Q tzn, dla każdego $j \in S$ prawdopodobieństwo jego wyboru wynosi q_{ij}
2. obliczamy funkcję akceptacji, która ma postać: $a(i, j) = \frac{\pi_j q_{ji}}{\pi_i q_{ij}}$
3. jeśli $a(i, j) \geq 1$ wtedy $X_{n+1} = j$
4. jeśli $a(i, j) < 1$ wtedy $X_{n+1} = \begin{cases} j & \text{z prawdopodobieństwem } a(i, j) \\ i & \text{z prawdopodobieństwem } 1 - a(i, j) \end{cases}$

$$\text{Zatem } X_{n+1} = \begin{cases} j & \text{z prawdopodobieństwem } \min(1, a(i, j)) \\ i & \text{z prawdopodobieństwem } 1 - a(i, j) \end{cases}$$

Wprowadzimy teraz definicje i twierdzenia związane z czasową odwracalnością łańcucha, w celu pokazania, że rozkład π używany przy konstrukcji łańcucha Markowa za pomocą algorytmu Metropolisa-Hastingsa jest w naszym przypadku rozkładem granicznym.

Definicja 2.10. Łańcuch Markowa $\{X_t\}$ nazwiemy *czasowo odwracalnym* względem rozkładu stacjonarengo π , jeśli dla wszystkich $i, j \in \{1, 2, \dots, n\}$ zachodzi równość $\pi(i)P(i, j) = \pi(j)P(j, i)$.

Twierdzenie 2.11. Jeśli łańcuch Markowa $(X_n)_n$ jest czasowo odwracalny względem rozkładu π to π jest rozkładem stacjonarnym tego łańcucha.

Dowód. Jeśli łańcuch Markowa o macierzy przejścia P jest czasowo odwracalny względem π to dla wszystkich i, j należących do przestrzeni stanów zachodzi:

$$\pi_i p_{ij} = \pi_j p_{ji} \quad (2.4)$$

$$\sum_i \pi_i p_{ij} = \sum_i \pi_j p_{ji} \quad (2.5)$$

$$\sum_i \pi_i p_{ij} = \pi_j \sum_i p_{ji} \quad (2.6)$$

$$\sum_i \pi_i p_{ij} = \pi_j \quad (2.7)$$

□

Twierdzenie 2.12. Łańcuch powstały za pomocą algorytmu Metropolisa-Hastingsa jest czasowo odwracalny względem π .

Dowód. Należy pokazać, że $\pi_i p_{ij} = \pi_j p_{ji}$. Rozważymy dwa przypadki.

Niech funkcja akceptacji $a(i, j) < 1$. Wówczas $a(j, i) \geq 1$, więc $p_{ji} = q_{ji}$ oraz $p_{ij} = a(i, j)q_{ij}$

$$\pi_j p_{ji} = \pi_j q_{ji} = \pi_j \frac{\pi_i a(i, j) q_{ij}}{\pi_j} = \pi_i a(i, j) q_{ij} = \pi_i p_{ij} \quad (2.8)$$

Dla przypadku, gdy funkcja akceptacji $a(i, j) \geq 1$ mamy $p_{ij} = a(i, j)q_{ij}$

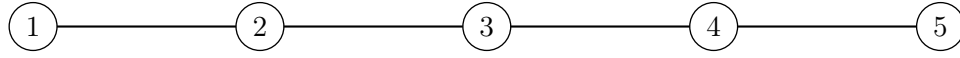
$$\pi_i p_{ij} = \pi_j q_{ij} = \pi_i \frac{\pi_j a(j, i) q_{ji}}{\pi_i} = \pi_j a(j, i) q_{ji} = \pi_j p_{ji} \quad (2.9)$$

Zatem łańcuch wygenerowany za pomocą algorytmu Metropolis-Hastingsa jest czasowo odwracalny względem π . □

Mamy zatem zapewnioną czasową odwracalność względem π dla łańcucha Markowa konstruowanego za pomocą algorytmu Metropolisa-Hastingsa. Na mocy Twierdzenia 2.12 jest on rozkładem stacjonarnym, a tym samym granicznym, ponieważ są spełnione założenie twierdzenia ergodycznego.

Zaprezentujemy działanie algorytmu na prostym przykładzie.

Przykład 2.13. Opisany zostanie algorytm Metropolisa-Hastingsa dla rozkładu jednostajnego na $[5] := \{1, 2, 3, 4, 5\}$. Rolę propozycji gra symetryczny spacer losowy po $[5]$ z ekranami odbijającymi w 1 i 5.



$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Macierz Q jest macierzą propozycji. Wybierzmy najpierw stan propozycji, niech $X_n = 1$. Wówczas dla:

$j = 2$ mamy $a_{1j} = \frac{\frac{1}{5} \frac{1}{2}}{\frac{1}{5} 1} = \frac{1}{2}$, a pozostałych $a_{1j} = 0$.

Mamy zatem:

$$a_{1j} = \begin{cases} \frac{1}{2} & \text{dla } j = 1, 2 \\ 0 & \text{dla } j = 3, 4, 5 \end{cases} \quad \text{oraz}$$

$$X_{n+1} = \begin{cases} 1, 2 & \text{z prawdopodobieństwem } \frac{1}{2} \\ 3, 4, 5 & \text{z prawdopodobieństwem } 0 \end{cases}.$$

Sytuacja jest analogiczna, gdy wybierzemy stan propozycji $X_n = 5$.

Wybierzmy teraz stan propozycji $X_n = 2$. Wówczas otrzymujemy:

$$X_{n+1} = \begin{cases} 1, 3 & \text{z prawdopodobieństwem } \frac{1}{2} \\ 2, 4, 5 & \text{z prawdopodobieństwem } 0 \end{cases}.$$

Podobnie jest, gdy stan propozycji będzie 3 lub 4.

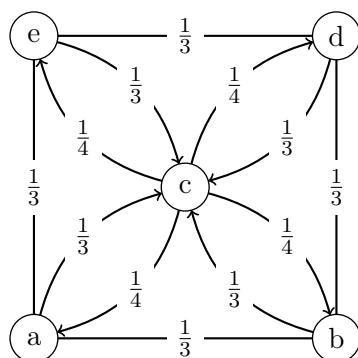
Powstały łańcuch Markowa $(X_n)_{n \in S}$ ma zatem macierz przejścia:

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

i rozkład stacjonarny, tym samym graniczny: $\pi = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$

Przykład 2.14. Wróćmy do Przykładu 2.8, w którym za pomocą algorytmu zamian 2×2 generowaliśmy macierz według rozkładu, który nie był rozkładem jednostajnym. Pokażemy, że za pomocą algorytmu zamian 2×2 z Metropolisa-Hastingsa, jesteśmy w stanie generować macierze z jednakowym prawdopodobieństwem.

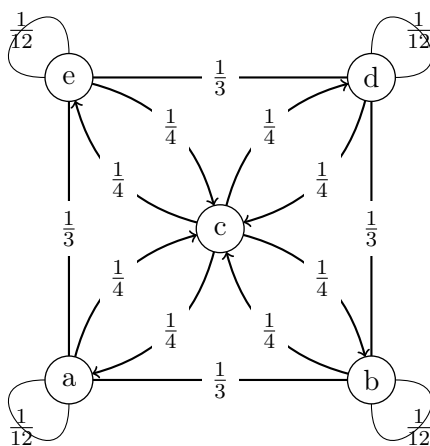
Rozkład stacjonarny i tym samym graniczny w algorytmie zamian 2×2 ma na trzeciej współrzędnej wartość inną od pozostałych. $\pi = (\frac{3}{16}, \frac{3}{16}, \frac{4}{16}, \frac{3}{16}, \frac{3}{16})$



$$P = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}$$

Niech macierz P będzie naszą macierzą propozycji w algorytmie Metropolisa-Hastingsa. Żądany przez nas rozkład ma postać: $\tilde{\pi} = (\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$

Wówczas otrzymany przez nas łańcuch Markowa, ma graf i macierz przejścia następującej postaci:



2.4. ALGORYTM METROPOLISA-HASTINGSA

$$\tilde{P} = \begin{pmatrix} \frac{1}{12} & \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{12} & \frac{1}{4} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{12} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{4} & \frac{1}{3} & \frac{1}{12} \end{pmatrix}$$

Macierz \tilde{P} odpowiada nowemu łańcuchowi Markowa skonstruowanemu za pomocą algorytmu Metropolisa-Hastingsa. Rozkład stacjonarny, a tym samym graniczny macierzy \tilde{P} jest rozkładem jednostajnym. W ulepszonym algorytmie zamiany 2×2 każda macierz po wykonaniu odpowiedniej liczby kroków jest generowana z jednakowym prawdopodobieństwem.

3. Szybkość zbieżności

W tym rozdziale skupimy się na zbadaniu liczby zamian 2×2 , które należy wykonać, aby wygenerowana przez nas macierz mogła być uznana za wylosowaną zgodnie z rozkładem jednostajnym.

Z definicji rozkładu granicznego, który jest w naszym przypadku jednocześnie rozkładem stacjonarym wiemy, że po podniesieniu macierzy przejścia P do pewnej potęgi t jej wiersze będą tworzyły wektory zbliżone do rozkładu granicznego.

Na ile zatem duża powinna być liczba kroków t , aby uzyskane macierze losowane były zgodnie z rozkładem jednostajnym?

Dla danych takich jak macierz z gatunkami ptaków kilka kroków nie jest w stanie oddalić nas od oryginalnej macierzy. Liczba wszystkich macierzy, z równymi sumami kolumn i wierszy dla naszego przykładu wynosi około 10^{17} . Czy to oznacza, że musimy wykonać tyle kroków aby otrzymać losową macierz?

Okazuje się, że dla odwracalnych łańcuchów Markowa (więc w szczególności dla losowych spacerów po grafie) odpowiedź na to pytanie jest zawarta w wartościach własnych macierzy przejścia P . Dokładną informację podaje nam druga co do wielkości wartość własna, co omówimy w poniższym rozdziale. Pomocne przy rozważanym problemie okażą się pozycje [6],[7] i [8].

3.1. Wartości własne macierzy stochastycznej

Poniższe wyniki związane są z szybkością zbieżności do rozkładu stacjonarnego. Przedstawimy zagadnienia związane z analizą spektralną.

Dana jest macierz stochastyczna P o wymiarach $n \times n$ oraz jej rozkład stacjonarny π .

Definicja 3.1 (Wartość własna). *Wartością własną* macierzy P nazywamy liczbę $\lambda \in \mathbb{C}$, taką, która spełnia równość $Pv = \lambda v$ dla pewnego niezerowego wektora $v \in \mathbb{C}^n$.

Niech $\lambda_1, \lambda_2, \dots, \lambda_n$ będą wartościami własnymi macierzy P .

Twierdzenie 3.2. Macierz stochastyczna P taka, że dla każdego $i, j \in \{1, 2, \dots, n\}$ $P(i, j) > 0$ ma dokładnie jedną wartość własną równą 1. Pozostałe wartości własne są co do modułu mniejsze od 1.

Dowód. Wprost z definicji rozkładu stacjonarnego mamy, że $P\pi = \pi$. Zatem 1 jest wartością własną macierzy P . Bez utraty ogólności założmy, że $\lambda_1 = 1$ oraz niech λ_2 będzie drugą co do modułu największą wartością własną. Wybierzmy indeks i taki, że $|v_i| \geq |v_j|$ dla każdego $j \in \{1, \dots, n\}$, gdzie v_i jest wartością na i -tej współrzędnej wektora v .

Wówczas mamy:

$$|\lambda v_i| = |(Pv)_i| = \left| \sum_{j=1}^n P(i, j)v_j \right| \leq \sum_{j=1}^n |v_j|P(i, j) \leq \sum_{j=1}^n |v_i| = |v_i|. \quad (3.1)$$

Otrzymaliśmy zatem, że $|\lambda_*| \leq 1$. Równość zachodzi tylko wtedy, gdy wektor v jest stały, czyli gdy wszystkie jego elementy są sobie równe, tzn. $v_1 = v_2 = \dots = v_n$. Z tego wynika, że jeśli macierz P jest macierzą diagonalizowalną, to wartość własna $\lambda_1 = 1$ jest jedyna.

Zajmijmy się przypadkiem, gdy macierz P nie jest diagonalizowalna. Wówczas istnieje możliwość, że wartość własna 1 jest częścią większej klatki Jordana. Wtedy dla pewnego wektora v mielibyśmy $Pv = v + u$ gdzie $u = (1, 1, \dots, 1)$. Wybierając indeks i taki, że $\Re v_i \geq \Re v_j$ dla każdego $j \in \{1, 2, \dots, n\}$ otrzymujemy:

$$1 + \Re v_i = \Re(Pv)_i = \Re \sum_{j=1}^n (P(i, j)v_j) \leq \sum_{j=1}^n P(i, j)v_j \leq \Re \sum_{j=1}^n P(i, j)v_i = \Re v_i. \quad (3.2)$$

Otrzymaliśmy sprzeczność, zatem $|\lambda_2| < 1$.

□

Twierdzenie 3.3. Niech A będzie macierzą symetryczną rzeczywistą o wymiarach $m \times m$. Wtedy wszystkie wartości własne macierzy A są rzeczywiste i istnieje baza ortonormalna złożona z rzeczywistych wektorów własnych.

$$A = \sum_{j=1}^m \lambda_j u_j u_j^T, \quad (3.3)$$

gdzie:

$u = (u_j)_{j=1}^m$ - baza złożona z wektorów własnych, λ_j - wartość własna odpowiadająca wektorowi własnemu u_j , $j = 1, \dots, m$.

Dowód powyższego twierdzenia możemy znaleźć w pozycji [6].

Powyższe warunki dotyczące symetrycznej macierzy A zastosujemy do macierzy przejścia P łańcucha Markowa, który jest nieprzywiedlny i nieokresowy.

Zachodzą następujące własności:

1. λ_1 jest prostą wartością własną odpowiadającą wektorowi własnemu $\mathbf{1} = (1, \dots, 1)^T$
2. $\lambda_j, j = 2, \dots, m: |\lambda_j| < 1$
3. $P^n = \sum_{j=1}^m \lambda_j^n u_j u_j^T, u_1 = \mathbf{1} \frac{1}{m}$

3.2. Ograniczenie normy

W celu zbadania szybkości zbieżności macierzy P^n do rozkładu stacjonarnego π , rozważmy różnicę:

$$P^n(x, y) - \pi(y) = P^n(x, y) - \frac{1}{m}, \quad (3.4)$$

która jest (x,y)-elementem macierzy

$$A^n = P^n - u_1 u_1^T = \sum_{j=2}^m \lambda_j^n u_j u_j^T \quad (3.5)$$

dla rozkładu $\pi = (\frac{1}{m}, \dots, \frac{1}{m})$.

W celu przedstawienia wyniku związanego z szybkością zbieżności łańcucha Markowa (o symetrycznej macierzy przejścia P), w zależności od wartości własnych, wprowadzimy pojęcie odległości dwóch miar.

Definicja 3.4. *Odległością dwóch miar μ, ν określonych na przestrzeni mierzalnej (X, \mathcal{B}) nazywamy*

$$\|\mu - \nu\|_{TV} := \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)|, \quad (3.6)$$

gdzie $\|\cdot\|_{TV}$ - norma całkowitej wariacji

Jeśli X jest co najwyżej przeliczalna, to:

$$\|\mu - \nu\|_{TV} := \frac{1}{2} \sum_{x \in X} |\mu(x) - \nu(x)| \quad (3.7)$$

Twierdzenie 3.5. Niech $(X_n)_n$ będzie nieprzywiedlnym i nieokresowym łańcuchem Markowa określonym na skończonej przestrzeni stanów S ($\#S = m$), o symetrycznej macierzy przejścia P . Wtedy dla każdego $i \in S$:

$$\|P_{i,\cdot} - \pi\|_{TV} \leq \frac{\sqrt{m}}{2} |\lambda_2|^n, \quad (3.8)$$

gdzie $\sigma(P) = \{1, \lambda_2, \dots, \lambda_m\}$ takie że $1 > |\lambda_2| \geq \dots \geq |\lambda_m|$

3.2. OGRANICZENIE NORMY

Dowód. niech $\mathbf{1} = (1, \dots, 1)^T$

wówczas mamy $P\mathbf{1} = \mathbf{1}$, co prowadzi do:

$$\mathbf{1}^T P^t = \mathbf{1}^T \quad (3.9)$$

z faktu, że $P^T = P$ mamy, że $\mathbf{1}^T P = \mathbf{1}^T$, więc

$$\pi := \frac{1}{m} \mathbf{1}^T \quad (3.10)$$

jest rozkładem stacjonarnym dla łańcucha $(X_n)_n$, oraz ponieważ jest jedyny, to jest on rozkładem granicznym, zatem:

$$\pi = \left(\frac{1}{m}, \dots, \frac{1}{m}\right) \quad (3.11)$$

ustalmy $i \in S$

$$\|P_{i,\cdot}^n - \pi\|_{TV} := \frac{1}{2} \sum_{j \in S} |p_n(i, j) - \frac{1}{m}| \quad (3.12)$$

Z nierówności Cauchy'ego-Schwarza:

$$\sum |x_j y_j| \leq \sqrt{\sum x_j^2} \sqrt{\sum y_j^2} \quad (3.13)$$

dla:

$$\begin{cases} x_j = \left| p_n(i, j) - \frac{1}{m} \right| \sqrt{m} \\ y_j = \frac{1}{\sqrt{m}} \end{cases} \quad (3.14)$$

$$\begin{aligned} \frac{1}{2} \left(\sum_{j \in S} \left| p_n(i, j) - \frac{1}{m} \right| \right)^2 &\leq \frac{1}{2} \left(\sum_{j \in S} m \left(p_n(i, j) - \frac{1}{m} \right)^2 \right) \left(\sum_{j \in S} \frac{1}{m} \right) \\ &= \frac{1}{2} \sum_{j \in S} m \left(p_n^2(i, j) - \frac{2}{m} p_n(i, j) + \frac{1}{m^2} \right) \stackrel{*}{=} \frac{1}{2} \left(\sum_{j \in S} m (p_n(i, j) p_n(j, i) - 1) \right) \\ &\stackrel{**}{=} \frac{1}{2} (m p_{2n}(i, i) - 1) = \frac{1}{2} m \left(p_{2n}(i, i) - \frac{1}{m} \right) \end{aligned} \quad (3.15)$$

gdzie (*) wynika z faktu, że P jest macierzą symetryczną, więc P^n jest także symetryczna, a (**) to zastosowanie Równania Chapmana-Kołmogorowa. Załóżmy, że:

$$P = \sum_{j=1}^m \lambda_j u_j u_j^T \quad (3.16)$$

gdzie u_j to kolumnowe prawostronne ortonormalne wektory własne, $u_1 = \frac{1}{\sqrt{m}} \mathbf{1}$

$$P = \frac{1}{m} \mathbf{1} \mathbf{1}^T + \sum_{j=2}^m \lambda_j u_j u_j^T \quad (3.17)$$

Niech $A := P - \frac{1}{m} \mathbf{1} \mathbf{1}^T$, możemy zauważyć, że

$$A = \sum_{j=2}^m \lambda_j u_j u_j^T \quad (3.18)$$

zatem A jest symetryczna, $\sigma(A) : 0, |\lambda_2| \geq \dots \geq |\lambda_m|$ oraz

$$A^n = P^n - \frac{1}{m} \mathbf{1}\mathbf{1}^T \quad (3.19)$$

zatem otrzymujemy:

$$\|P_{i,\cdot}^n - \pi\|_{TV} \leq \frac{1}{2} \sqrt{m(p_{2n}(i, i) - \frac{1}{m})} = \frac{1}{2} \sqrt{m(A^{2n})_{i,i}} \quad (3.20)$$

ale

$$\begin{aligned} (A^{2n})_{i,i} &= e_i^T A^{2n} e_i = e_i^T A^n A^n e_i \\ &= (A^n e_i)^T (A^n e_i) = \|A^n e_i\|_2^2 \leq \|A^n\|_2^2 \|e_i\|_2^2 \leq \left(\max_{\lambda \in \sigma(A^n)} |\lambda| \right)^2 = \lambda_2^{2n} \end{aligned} \quad (3.21)$$

gdzie $e_i = (0, 0, \dots, 1, \dots, 0)^T$ z jedynką na i -tym miejscu,

$\|A\|_2 = \max_{\lambda \in \sigma(A)} |\lambda|$ i $\sigma(A^n) : 0, |\lambda_2|^n \geq \dots \geq |\lambda_m|^n$. Zatem końcowo:

$$\|P_{i,\cdot}^n - \pi\|_{TV} \leq \frac{1}{2} \sqrt{m \lambda_2^{2n}} = \frac{\sqrt{m}}{2} |\lambda_2|^n \quad (3.22)$$

□

Co jeśli macierz P jest niesymetryczna? Wtedy jest ona podobna do macierzy symetrycznej z rzeczywistymi wartościami własnymi.

3.3. Wnioski

Z powyższych rozważań możemy wywnioskować, że im mniejsza wartość λ_2 , tym mniej kroków jest potrzebnych, aby uzyskać $\|P_{i,\cdot}^n - \pi\|_{TV} > \epsilon$ dla dowolnego $\epsilon \geq 0$. W praktyce może być jednak trudno bezpośrednio korzystać z tej teorii. Po pierwsze, niektóre rodzaje kroków wykonuje się dłużej niż inne, więc mniejsza liczba kroków nie jest automatycznie lepsza.

Po drugie, znaczenie ma również ilość stanów. Dla dużych przestrzeni obliczenie macierzy przejścia nie jest praktyczne, w takim przypadku implementacja zależy od znalezienia alternatywy dla bezpośredniego działania na macierzy P , na przykład poprzez zamiany 2×2 . Bez znajomości P nie można zazwyczaj znaleźć jego wartości własnych, co uniemożliwia nam użycie $\frac{\sqrt{m}}{2} |\lambda_2|$ do określenia liczby kroków potrzebnych do zajścia nierówności $\|P_{i,\cdot}^n - \pi\|_{TV} \leq \epsilon$.

Niemniej jednak granica $\frac{\sqrt{m}}{2} |\lambda_2|$ jest użyteczna, ponieważ sugeruje, że możemy przyspieszyć zbieżność, stosując strategię, które zmniejszają wartość drugiej największej wartości własnej.

4. Sklejanie łańcuchów Markowa

W tym rozdziale przyjrzymy się bliżej modyfikacjom pierwotnego łańcucha Markowa. Omówiona przez nas strategia - sklejanie, będzie miała na celu zmniejszenie liczby stanów danego łańcucha, co przyczyni się do zwiększania przestrzeni spektralnej. Zmniejszenie drugiej największej wartości własnej prowadzi do przyspieszenia zbieżności. Niniejszy rozdział pisany jest w oparciu o pozycję [4]. W celu przybliżenia tematu przedstawimy sklejanie na kilku przykładach. Rozpatrywać będziemy tylko przypadki, gdy łańcuch Markowa jest jednorodny.

4.1. Idea sklejania łańcucha Markowa

Dany jest łańcuch Markowa z n stanami o macierzy przejścia P i z rozkładem początkowym π . Niech $L = \{L_1, L_2, \dots, L_r\}$ będzie podziałem zbioru stanów. Definiujemy nowy proces, o r stanach, w którym prawdopodobieństwa przejścia wyrażają się w następujący sposób:

$$\mathbb{P}_\pi(X_0 \in L_i) \tag{4.1}$$

$$\mathbb{P}_\pi(X_n \in L_r | X_{n-1} \in L_s, \dots, X_0 \in L_i) \tag{4.2}$$

Powyższy proces nazywamy procesem sklejonym. Nie zawsze jednak proces jest w stanie być sklejony.

Definicja 4.1. Łańcuch Markowa $(X_n)_n$ jest *sklejalny względem partycji* jeżeli dla dowolnego wektora początkowego π proces zdefiniowany za pomocą (4.1), (4.2) jest łańcuchem Markowa o prawdopodobieństwach przejścia niezależnych od wyboru rozkładu początkowego π .

Zanim wprowadzimy twierdzenie, które przedstawi nam warunek konieczny i dostateczny na to, aby proces był sklejalny, przedstawimy intuicyjnie sposób w jaki sklejamy łańcuchy Markowa.

Przykład 4.2. Dany jest łańcuch Markowa ze stanami $\{1, 2, 3, 4\}$ z podziałem

$L = \{\{1, 2\}, \{3, 4\}\}$ i jego macierz przejścia $P = \begin{pmatrix} \frac{1}{8} & \frac{3}{8} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{16} & \frac{7}{16} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{9} & \frac{2}{9} \\ \frac{1}{6} & \frac{3}{6} & \frac{1}{9} & \frac{2}{9} \end{pmatrix}$. Sklejony łańcuch Markowa

ma dwa stany: $\{1, 2\}, \{3, 4\}$, a jego nowa macierz przejścia wygląda w następujący sposób.

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

Treść poniżej wprowadzonego twierdzenie pochodzi z pozycji [4].

Twierdzenie 4.3. Warunkiem koniecznym i dostatecznym na to, aby łańcuch Markowa (X_n) był sklejalny względem partycji $L = \{L_1, L_2, \dots, L_r\}$ jest to, aby dla dowolnych zbiorów $L_k, L_l \in L$ dla każdego n $\mathbb{P}(X_n \in L_l | X_{n-1} = s_k)$ miało tą samą wartość dla każdego $s_k \in L_k$ niezależnie od n .

Dowód. \Rightarrow (dostateczność) Zaprezentowany zostanie szkic dowodu. Pokażemy prawdziwość implikacji w prawą stronę dla przypadku, gdy $n = 2$.

$$\begin{aligned} \mathbb{P}(X_2 \in L_{x_2} | X_1 \in L_{x_1}, X_0 \in L_{x_0}) &= \frac{\mathbb{P}(X_2 \in L_{x_2} \cap X_1 \in L_{x_1} \cap X_0 \in L_{x_0})}{\mathbb{P}(X_1 \in L_{x_1}, X_0 \in L_{x_0})} \\ &= \frac{\sum_{x_2 \in L_{x_2}} \sum_{x_1 \in L_{x_1}} \sum_{x_0 \in L_{x_0}} \mathbb{P}(X_2 = x_2, X_1 = x_1, X_0 = x_0)}{\sum_{x_1 \in L_{x_1}} \sum_{x_0 \in L_{x_0}} \mathbb{P}(X_1 = x_1, X_0 = x_0)} \\ &= \frac{\sum_{x_2 \in L_{x_2}} \sum_{x_1 \in L_{x_1}} \sum_{x_0 \in L_{x_0}} \mathbb{P}(X_2 = x_2 | X_1 = x_1) \mathbb{P}(X_1 = x_1 | X_0 = x_0) \mathbb{P}(X_0 = x_0)}{\sum_{x_1 \in L_{x_1}} \sum_{x_0 \in L_{x_0}} \mathbb{P}(X_1 = x_1 | X_0 = x_0) \mathbb{P}(X_0 = x_0)} = \star \end{aligned} \quad (4.3)$$

Ponieważ z założenia $\sum_{x_2 \in L_{x_2}} \mathbb{P}(X_2 = x_2 | X_1 = x_1)$ ma taką samą wartość dla każdego $x_1 \in L_{x_1}$ to nie zależy od wyboru x_1 , więc możemy ją wyłączyć przed sumy. Oznaczmy powyższą sumę przez $p_{1,2}$. Tak samo zrobimy z $p_{0,1}$. Mamy zatem:

$$\begin{aligned} \star &= \frac{p_{1,2} \sum_{x_0 \in L_{x_0}} \sum_{x_1 \in L_{x_1}} \mathbb{P}(X_1 = x_1 | X_0 = x_0) \mathbb{P}(X_0 = x_0)}{\sum_{x_1 \in L_{x_1}} \sum_{x_0 \in L_{x_0}} \mathbb{P}(X_1 = x_1 | X_0 = x_0) \mathbb{P}(X_0 = x_0)} \\ &= \frac{p_{1,2} p_{0,1} \sum_{x_0 \in L_{x_0}} \mathbb{P}(X_0 = x_0)}{p_{0,1} \sum_{x_0 \in L_{x_0}} \mathbb{P}(X_0 = x_0)} = p_{1,2} = \mathbb{P}(X_2 \in L_{x_2} | X_1 \in L_{x_1}) \end{aligned} \quad (4.4)$$

Pokazaliśmy zatem, że jest to łańcuch Markowa.

\Leftarrow (konieczność) implikacja w drugą stronę wynika wprost z definicji łańcucha sklejalnego. \square

Przykład 4.4. Zaprezentujemy przykład łańcucha Markowa, który nie jest sklejalny w odniesieniu do danej partycji.

Niech dany będzie łańcuch Markowa o przestrzeni stanów $S = \{1, 2, 3, 4, 5\}$ i podział $L = \{\{1, 2\}, \{3, 4, 5\}\}$ o macierzy przejścia:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

Warunkiem koniecznym, aby łańcuch był sklejalny w odniesieniu do danej partycji jest to, aby dla każdej pary L_k, L_l i dla każdego stanu i_1, i_2 należącego do L_k

$$\sum_{j \in L_l} p(i_1, j) = \sum_{j \in L_l} p(i_2, j).$$

Weźmy $L_k = \{1, 2\}$ i $L_l = \{3, 4, 5\}$. Wówczas sprawdzimy czy zachowana jest równość $p(1, 3) + p(1, 4) + p(1, 5) = p(2, 3) + p(2, 4) + p(2, 5)$. Podstawiając wartości odczytane z macierzy otrzymujemy sprzeczność: $\frac{1}{4} = \frac{1}{3}$. Oznacza to, że łańcuch ten nie jest sklejalny w odniesieniu do partycji L .

4.2. Macierzowa reprezentacja sklejonego łańcucha Markowa

Założmy, że nasz łańcuch Markowa o n stanach jest sklejalny w odniesieniu do partycji L . Przez macierz U oznaczmy taką macierz $r \times n$, w której i -ty wiersz jest wektorem prawdopodobieństwa mającym równe wartości na współrzędnych u_{ij} dla $j \in L_i$, na reszcie miejsc są zera. Macierz V o wymiarach $n \times r$ niech będzie taką macierzą, w której j -ta kolumna to wektor zer z jedynkami na i -tym miejscu jeśli $i \in L_j$.

Twierdzenie 4.5. Łańcuch Markowa o macierzy przejścia P jest sklejalny w odniesieniu do podziału L wtedy i tylko wtedy, gdy $VUPV = PV$.

Dowód. Najpierw założmy, że łańcuch jest lumpable w odniesieniu do podziału L .

$$VU = \left[\begin{array}{c|c|c|c} \mathbf{W}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{W}_2 & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}_r \end{array} \right]$$

W_1, W_2, \dots, W_r są macierzami stochastycznymi. Warunek, który chcemy udowodnić oznacza, że kolumny macierzy PV są wektorami własnymi macierzy VU odpowiadającymi wartości własnej 1. Z tego, że łańcuch jest sklejalny wiemy, że prawdopodobieństwo przejścia z każdego

stanu z L_l do zbioru L_k musi być takie samo, więc wyrazy w kolumnie macierzy PV odpowiadające L_k są takie same (każda z macierzy W_1, \dots, W_r jest macierzą o stałych elementach). Oznacza to, że kolumny PV muszą być wektorami własnymi odpowiadającymi wartości własnej 1, a więc zachodzi $VUPV = PV$.

Teraz założmy, że zachodzi $VUPV = PV$. Wówczas mamy, że kolumny PV są wektorami własnymi VU odpowiadającymi wartości własnej 1. Ponieważ każda z macierzy W_1, \dots, W_r odpowiada pewnemu łańcuchowi Markowa (macierze prawdopodobieństw), stąd wszystkie współczynniki PV odpowiadające jednemu ze zbiorów z L muszą być takie same, co dowodzi temu, że łańcuch jest sklejalny.

□

Przykład 4.6. Dla macierzy P z przykładu 4.3 mamy

$$V = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, U = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Macierz przejścia dla sklejonego łańcucha wyraża się wzorem $\hat{P} = UPV$

$$\hat{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{1}{8} & \frac{3}{8} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{16} & \frac{7}{16} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{9} & \frac{2}{9} \\ \frac{1}{6} & \frac{3}{6} & \frac{1}{9} & \frac{2}{9} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Własności macierzy V i VU

1. Kolumny macierzy V są ortogonalne (iloczyn skalarny dowolnej pary jest równy 0). Możemy to interpretować w ten sposób, że jeden stan należy tylko do jednego zbioru z L .
2. $VU = V(V^T V)^{-1} V^T$ jest macierzą symetryczną. Macierz ta jest macierzą przekształcenia rzutu na przestrzeń stworzoną przez kolumny V . Zatem mamy warunek sklejalności: $PV = VUPV$

Twierdzenie 4.7. Niech macierz P będzie macierzą przejścia łańcucha Markowa, a macierz \hat{P} macierzą przejścia sklejonego łańcucha Markowa ze względu na podział L . Wówczas zachodzi równość:

$$\hat{P}^k = UP^kV \quad (4.6)$$

Dowód. Dowód zostanie przeprowadzony przez indukcję.

Dla $k = 2$ mamy $\hat{P}^2 = UPVUPV$. Z warunku na lumpability mamy, że $PV = VUPV$ zatem $\hat{P}^2 = UPVUPV = UPPV = UP^2V$.

Założmy, że twierdzenie jest prawdziwe dla $k - 1$. Pokażemy, że zachodzi ono dla k .

$$\hat{P}^k = \hat{P}^{k-1}\hat{P} = UP^{k-1}VUPV = UP^kV. \quad (4.7)$$

□

4.3. Wartości własne macierzy przejścia po sklejaniu

Przykład 4.8. Wróćmy do przykładu 2.14. Nasza macierz ma postać

$$\tilde{P} = \begin{pmatrix} \frac{1}{12} & \frac{1}{3} & \frac{1}{4} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{12} & \frac{1}{4} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{12} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{4} & \frac{1}{12} & \frac{1}{3} \end{pmatrix}, V = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, U = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\text{Wówczas } \hat{P} = U\tilde{P}V = \begin{pmatrix} \frac{5}{12} & \frac{1}{4} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{4} & \frac{5}{12} \end{pmatrix}$$

Przyjrzyjmy się teraz wartościom własnym tych macierzy. Zbiorem wartości macierzy \tilde{P} jest $\{1, -\frac{7}{12}, -\frac{1}{4}, \frac{1}{12}, \frac{1}{12}\}$, natomiast macierzy \hat{P} : $\{1, -\frac{1}{4}, \frac{1}{12}\}$.

Drugą co do modułu wartością własną macierzy \tilde{P} jest $-\frac{7}{12}$, a macierzy \hat{P} : $-\frac{1}{4}$. Proces sklejania zapewnił nam zatem zwiększenie przerwy spektralnej $(1 - \lambda_*)$.

Możemy zauważyć, że zbiór wartości własnych macierzy \hat{P} jest podzbiorem wartości własnych macierzy \tilde{P} .

Twierdzenie 4.9. Niech macierz \hat{P} będzie macierzą przejścia sklejonego łańcucha Markowa o macierzy przejścia P o wymiarach $n \times n$ ze względu na podział $L = \{L_1, L_2, \dots, L_r\}$. Wówczas zbiór wartości własnych macierzy \hat{P} jest podzbiorem wartości własnych macierzy P .

Dowód. Niech $W_P(\lambda) = 0$ będzie równaniem charakterystycznym macierzy P . $W_P(\lambda)$ jest wielomianem n -tego stopnia, zatem możemy go zapisać w następującej postaci: $W_P(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0$, gdzie $a_n, \dots, a_0 \in \mathbb{R}$

Przypomnijmy twierdzenie Cayleya-Hamiltona. Każda macierz kwadratowa jest rozwiązaniem jej równania charakterystycznego.

Mamy zatem $W_P(P) = 0$. Wykorzystując to oraz fakt, że $\hat{P}^n = UP^nV$ otrzymujemy:

$$W_P(P) = a_n P^n + a_{n-1} P^{n-1} + \dots + a_1 P + a_0 I = 0 \quad (4.8)$$

$$UW_P(P)V = a_n UP^nV + a_{n-1}UP^{n-1}V + \dots + a_1UPV + a_0I = 0 \quad (4.9)$$

$$Uw_P(P)V = a_n\hat{P}^nV + a_{n-1}\hat{P}^{n-1}V + \dots + a_1U\hat{P}V + a_0I = 0 \quad (4.10)$$

$$W_P(\hat{P}) = 0 \quad (4.11)$$

Otrzymaliśmy zatem, że macierz \hat{P} jest rozwiązaniem równania charakterystycznego macierzy P . Oznacza to, że wartości własne macierzy \hat{P} są również wartościami własnymi macierzy P . \square

Powyższe twierdzenie gwarantuje nam, że sklejając łańcuch Markowa przerwa spektralna zwiększy się lub pozostanie bez zmian. Działa to korzystnie na przyspieszanie zbieżności, co wykazaliśmy w rozdziale trzecim.

5. Podsumowanie

W poniższym rozdziale podsumujemy zgromadzone dotychczas informacje na przykładzie. Przedstawione zostanie dwukrotne sklejenie łańcucha Markowa o początkowej liczbie stanów wynoszącej 12, która końcowo zredukuje się do 4. Następnie zostanie zastosowany algorytm Metropolisa-Hastingsa w celu stworzenia łańcucha Markowa z jednostajnym rozkładem stacjonarnym (w oparciu o macierz propozycji sklejonego pierwotnie łańcucha).

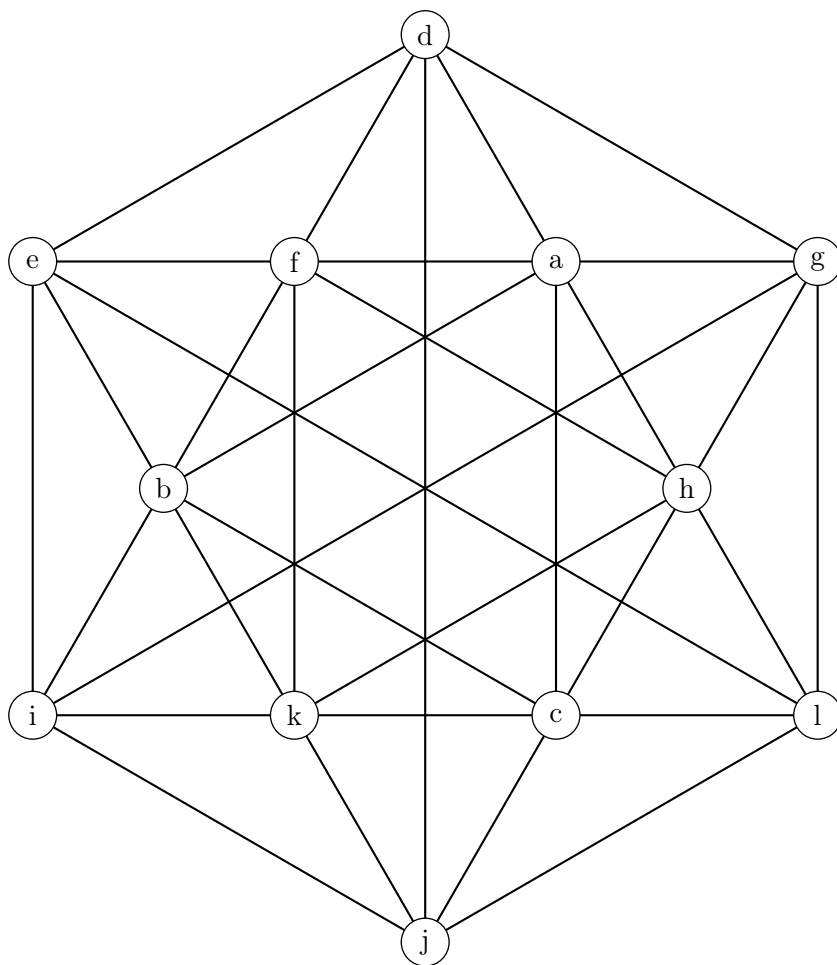
5.1. Analiza przykładu

Przykład 5.1. Niech dany będzie zbiór macierzy $\mathcal{A}(w, k)$ gdzie $w = (2, 1, 2)$, a $k = (2, 1, 1, 1)$.

Wówczas wszystkie możliwe macierze to:

$$\begin{aligned} a &= \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}, c = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, d = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}, \\ e &= \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}, f = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, g = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}, h = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \\ i &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, j = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, k = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, l = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Powyższy przykład możemy zareprezentować na grafie, gdzie wierzchołki są macierzami i są one połączone krawędziami, gdy z jednej macierzy do drugiej możemy przejść za pomocą pojedynczej zamiany 2×2 .

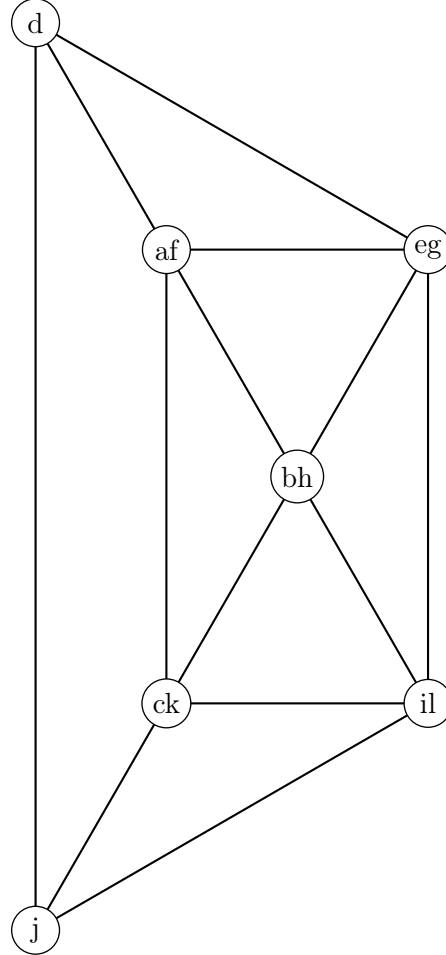


Macierz przejścia łańcucha Markowa odpowiadającego powyższemu grafowi wyraża się w następujący sposób:

$$P = \begin{array}{c|cccccccccccc} & a & f & b & h & c & k & e & g & l & i & j & d \\ \hline a & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & \frac{1}{6} & 0 & 0 & 0 & \frac{1}{6} \\ f & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{6} \\ b & \frac{1}{6} & \frac{1}{6} & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & \frac{1}{6} & 0 & 0 \\ h & \frac{1}{6} & \frac{1}{6} & 0 & 0 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 \\ c & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & 0 & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 \\ k & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & 0 \\ e & 0 & \frac{1}{5} & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ g & \frac{1}{5} & 0 & 0 & \frac{1}{5} & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ l & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{1}{5} & 0 \\ i & 0 & 0 & \frac{1}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{1}{5} & 0 \\ j & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ d & \frac{1}{5} & \frac{1}{5} & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 & 0 & \frac{1}{5} & 0 \end{array}$$

5.1. ANALIZA PRZYKŁADU

Stosując udowodnione w rozdziale czwartym twierdzenie możemy zauważyć, że łańcuch ten jest składalny w odniesieniu do partycji $L = \{\{a, f\}, \{c, k\}, \{e, g\}, \{i, l\}, \{b, h\}, \{d, j\}\}$. Oznacza to, że przy składaniu odpowiednie macierze będą tworzyły jeden wierzchołek. Graf odpowiadający sklejonymu łańcuchowi ma następującą postać:

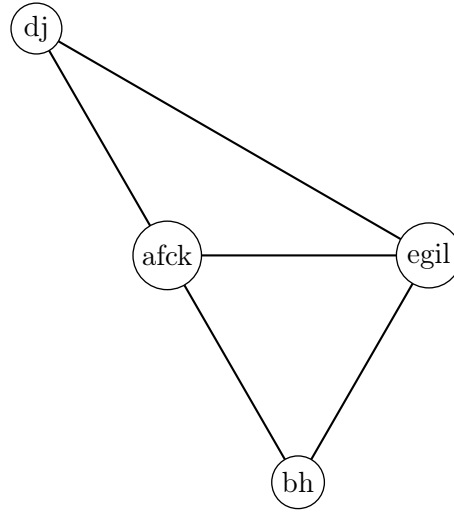


Macierz przejścia łańcucha Markowa odpowiadającego powyższemu grafowi wyraża się w następujący sposób:

$$P = \begin{array}{c|ccccccc} & af & ck & eg & il & bh & d & j \\ \hline af & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{3} & 0 & \frac{1}{6} \\ kc & \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\ ge & \frac{1}{5} & 0 & 0 & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ il & 0 & \frac{1}{5} & \frac{2}{5} & 0 & \frac{1}{5} & \frac{1}{5} & 0 \\ bh & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 \\ j & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & 0 & \frac{1}{6} \\ d & \frac{1}{6} & 0 & \frac{1}{6} & 0 & 0 & \frac{1}{6} & 0 \end{array}$$

Złożony przez nas łańcuch da się jednak jeszcze złożyć w odniesieniu do partycji

$$L = \{\{af, ck\}, \{eg, il\}, \{j, d\}, bh\}$$



Macierz przejścia łańcucha Markowa odpowiadającego powyższemu grafowi wyraża się w następujący sposób:

$$P_3 = \begin{array}{c|cccc} & afck & egil & bh & dj \\ \hline afck & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ egil & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & \frac{1}{5} \\ bh & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ dj & \frac{2}{5} & \frac{2}{5} & 0 & \frac{1}{5} \end{array}$$

Na sklejonym łańcuchu Markowa o macierzy przejścia P_3 zastosujemy algorytm Metropolisa-Hastingsa.

Niech macierzą propozycji będzie macierz P_3 , a żądanym przez nas rozkładem:

$$\pi = \left(\frac{4}{12}, \frac{4}{12}, \frac{2}{12}, \frac{2}{12}\right).$$

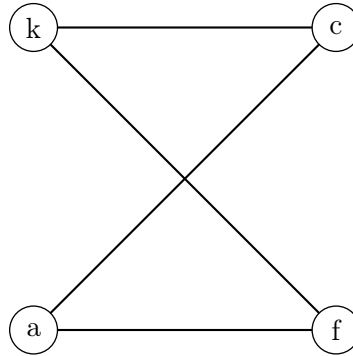
Powyższa postać rozkładu stacjonarnego wynika z tego, że chcemy aby każda z macierzy (a, b, c, \dots, l) była generowana z jednakowym prawdopodobieństwem. Musimy zatem dostosować prawdopodobieństwo wygenerowania grupy macierzy do jej rozmiaru. Przykładowo w grupie $afck$ są 4 macierze, wszystkich łącznie jest 12, więc współrzędna w rozkładzie stacjonarnym będzie wynosiła $\frac{4}{12}$.

5.2. WNIOSKI

Nasza nowa macierz powstała za pomocą algorytmu Metropolisa-Hastingsa ma postać:

$$\tilde{P}_3 = \begin{array}{c|cccc} & afck & egil & bh & dj \\ \hline afck & \frac{1}{3} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ egil & \frac{1}{6} & \frac{7}{15} & \frac{1}{6} & \frac{1}{5} \\ bh & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ dj & \frac{1}{3} & \frac{2}{5} & 0 & \frac{4}{15} \end{array}$$

Po wylosowaniu grupy macierzy zostaje nam zastosować zwykły algorytm zamian 2×2 . Możemy wylosowaną grupę macierzy potraktować jako podgraf. Ponieważ w każdej grupie macierze są tego samego rodzaju, każdy wierzchołek podgrafu ma ten sam stopień. Oznacza to, że macierz przejścia odpowiadająca spacerowi losowemu po tym podgrafie będzie miała rozkład stacjonarny, który będzie rozkładem jednostajnym. Pokażemy powyższe rozumowanie na grupie macierzy *afck*. Graf odpowiadający spacerowi losowemu po grupie *afck* ma następującą postać:



a macierz przejścia związana z odpowiadającym łańcuchem Markowa:

$$\tilde{P}_3 = \begin{array}{c|cccc} & a & f & c & k \\ \hline a & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ f & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ c & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ k & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}$$

Rozkład stacjonarny tego łańcucha jest rozkładem jednostajnym, $\pi = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.

Dzięki sklepaniu zapewniliśmy sobie przyspieszenie zbieżności, tj. wykonanie mniejszej ilości zamian 2×2 .

5.2. Wnioski

Niniejsza praca dostarcza nam sposobów na generowanie macierzy zgodnie z jednostajnym rozkładem prawdopodobieństwa. Na potrzeby ekologicznego problemu wprowadzone zostały algorytmy, które mogliśmy ulepszać korzystając z teorii spektralnej oraz teorii łańcuchów Markowa.

Metody Monte Carlo oparte na teorii łańcuchów Markowa (MCMC) - algorytm Metropolis-Hastingsa okazał się niezwykle przydatny w postawionym przez nas matematycznym zadaniu. Omówione metody i wyniki mogą być wykorzystane do analizy bardziej złożonych tabeli wysp i ptaków o większych rozmiarach. Matematyka odgrywa kluczową rolę w ekologii, dostarczając narzędzi do analizy złożonych problemów.

Bibliografia

- [1] D.R Barr, M.U. Thomas An Eigenvector Condition for Markov Chain Lumpability, 1977
- [2] R. A. Brualdi, Matrices of zeros and ones with fixed row and column sum vectors, *Linear Algebra and its Applications*, Volume 33, October 1980, Pages 159-231
- [3] Cobb, George W.; Chen, Yung-Pin, An application of Markov chain Monte Carlo to community, *Amer. Math. Monthly* 110 (2003), no. 4, 265–288.
- [4] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*, D. Van Nostrand, Princeton, 1960, rozdział 6
- [5] Steven P. Lalley, *Convergence rates of markov chains*, rozdział 3
- [6] David A. Levin, Yuval Peres, *Markov Chains and Mixing Times*, Chapter 12, 2017
- [7] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer-Verlag, New York, 2001.
- [8] Jeffrey S. Rosenthal, *Convergence Rates for Markov Chains*, *SIAM Review*, Sep., 1995, Vol. 37, No. 3 (Sep., 1995), 387-405
- [9] H. J. Ryser, Combinatorial properties of matrices of zeros and ones, *Canad. J. Math.* 9 (1957) 371-377.

Spis tabel

1.1	Macierz występowania gatunków ptaków na wyspach Galapagos	12
-----	---	----