Wstęp do rachunku prawdopodobieństwa
Wykład 12: Łańcuchy Markowa. Twierdzenie ergodyczne. Błądzenie na grafach.

Twierdzenie. *Jeżeli* $p_{ij} = p_{ji}$ *dla wszystkich stanów* i, j *łańcucha Markowa, to łańcuch ten posiada jednostajny rozkład stacjonarny.*

Definicja. *Łańcuch Markowa jest* **nieprzywiedlny**, *gdy dla dowolnych jego stanów* i, j *istnieje takie* $t \geqslant 1$, *że* $p_{ij}^t > 0$.

Definicja. Okresem stanu j *nazywamy liczbę* $d(j) = NWD\{t \geqslant 1 : p_{jj}^t > 0\}$.
Uwaga: okresu nie definiujemy dla stanu j, *gdy* $p_{jj}^t = 0$ *dla wszystkich* $t \geqslant 1$.
Gdy $d(j) > 1$, *wtedy mówimy, że* **stan** j **jest okresowy**, *w przeciwnym wypadku mówimy, że* **stan jest nieokresowy**. *Stany nieposiadające okresu też są nieokresowe.*

Fakt. W nieprzywiedlnym łańcuchu Markowa wszystkie stany mają ten sam okres.

Definicja. *Łańcuch Markowa jest* **nieokresowy**, *gdy wszystkie jego stany są nieokresowe. Łańcuch jest* **ergodyczny**, *gdy jest nieprzywiedlny i nieokresowy.*

Twierdzenie (ergodyczne). *Każdy ergodyczny łańcuch Markowa ma dokładnie jeden rozkład stacjonarny* $\bar{\pi} = (\pi_j)$. *Ponadto dla stanów* i, j *zachodzi* $\lim_{t\to\infty} p_{ij}^t = \pi_j$. *Innymi słowy, niezależnie od początkowego rozkładu* $\bar{\rho}^0$ *zachodzi* $\lim_{t\to\infty} \bar{\rho}^t = \bar{\pi}$.

Błądzeniem klasycznym (spacerem losowym) na grafie G bez wierzchołków izolowanych nazywamy proces, w którym cząstka znajdująca się w jakimś wierzchołku grafu w każdym kroku przemieszcza się do jednego z sąsiadów tego wierzchołka, wybranego z jednakowym prawdopodobieństwem. Procesowi temu odpowiada łańcuch Markowa o zbiorze stanów $V(G)$. Gdy mowa o błądzeniu na grafie, zawsze zakładamy, że graf ten nie ma wierzchołków izolowanych.

Fakt. Błądzenie na grafie G jest łańcuchem nieprzywiedlnym wtedy i tylko wtedy, gdy G jest spójny. Błądzenie na grafie G jest łańcuchem nieokresowym wtedy i tylko wtedy, gdy G nie jest dwudzielny.

Przez $v(G)$ oznaczamy liczbę wierzchołków, a przez $e(G)$ – liczbę krawędzi grafu G.

Twierdzenie. *Wektor* $(deg(w)/(2e(G)))_{w\in V(G)}$ *jest rozkładem stacjonarnym dla błądzenia na grafie* G.

Definicja. Czasem pokrycia *dla błądzenia na grafie* G *nazywamy zmienną losową będącą liczbą kroków, po których każdy wierzchołek grafu* G *został odwiedzony.*

Twierdzenie. *Średni czas pokrycia dla błądzenia po dowolnym spójnym grafie* G, *dla dowolnego stanu początkowego, jest mniejszy niż* $2e(G)v(G)$.

---

1. Zadania na ćwiczenia

Zadanie 1. Przedstaw za pomocą łańcucha Markowa następujące procesy:
(a) Cząstka błądzi klasycznie na ścieżce o 4 wierzchołkach.
(b) Cząstka błądzi (nieklasycznie) na ścieżce o 4 wierzchołkach w następujący sposób: jeśli przybyła do wierzchołka wewnętrznego ścieżki z lewej strony, to w następnym kroku szansa pójścia w lewo jest dwa razy mniejsza od szansy pójścia w prawo. Analogiczna reguła dotyczy przybycia z prawej strony.

Zadanie 2. Komputer generuje tekst złożony z liter A, B, C wg takich reguł: Zaczyna od losowo wybranej litery. Jeżeli ostatnio napisana litera jest samogłoską, następna będzie spółgłoską wybraną z jednakowym prawdopodobieństwem. Jeżeli ostatnio napisana litera jest spółgłoską, następna będzie inna, przy czym samogłoska jest wybierana z prawdopodobieństwem dwa razy większym niż spółgłoska.
(a) Przedstaw ten proces w postaci łańcucha Markowa.
(b) Wyznacz jego rozkład stacjonarny.
(c) Zbadaj, czy łańcuch ten jest nieprzywiedlny i nieokresowy.

Zadanie 3. (rezerwowe) Komputer generuje tekst złożony z liter A, B, C wg takich reguł: Zaczyna od losowo wybranej trójki różnych liter, wybranej spośród tych, które nie zawierają CA ani BAC. Każda następna litera jest wybrana z jednakowym prawdopodobieństwem spośród spełniających dwa warunki:

- litera jest inna od poprzedniej,
- w tekście nie mogą pojawić się wyrazy: BABA, BAC, CA.

(a) Przedstaw ten proces w postaci łańcucha Markowa.
(b) Zbadaj, czy łańcuch jest nieprzywiedlny.
(c) Zbadaj, czy łańcuch jest nieokresowy.

Zadanie 4. (rezerwowe) Ala i Franek rzucają na zmianę standardową monetą (zaczyna Ala). Franek wygrywa, gdy po raz pierwszy w ciągu rzutów pojawi sie ROR, a Ala – gdy pojawi sie OOR. Oblicz prawdopodobieństwo wygranej Ali.

---

2. Zadania domowe

Zadanie 1. Macierz przejścia pewnego łańcucha Markowa dana jest następującym wzorem:

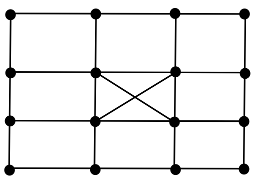
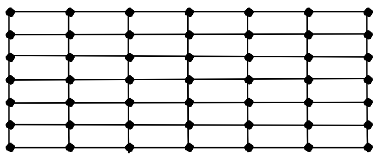
(a) $\begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$ (c) $\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

(a) Sprawdź, czy łańcuch jest nieprzywiedlny.
(b) Sprawdź, czy istnieją w nim stany okresowe.
(c) Wyznacz liczbę rozkładów stacjonarnych.

Zadanie 2. Rozpatrzmy klasyczne błądzenie po cyklu o 5 wierzchołkach w_1, w_2, w_3, w_4, w_5, zaczynające się w wierzchołku w_1.

(a) Czy odpowiadający temu błądzeniu łańcuch Markowa jest nieprzywiedlny? Czy jest okresowy?
(b) Wyznacz rozkład stacjonarny tego łańcucha.
(c) Wyznacz $\lim_{t\to\infty} p^t_{w_1 w_3}$.

Zadanie 3. Uzasadnij, że błądzenie klasyczne na pierwszym grafie jest nieokresowym łańcuchem Markowa, a błądzenie na drugim grafie jest łańcuchem okresowym.

(a)  (b) 

Zadanie 4. Rozważmy bładzenie na pierwszym grafie z poprzedniego zadania. Załóżmy, że startujemy z lewego górnego rogu.

(a) Wyznacz rozkład stacjonarny dla tego błądzenia.
~~(b) Uzasadnij, że wartość oczekiwana czasu potrzebnego do odwiedzenia każdego wierzchołka jest mniejsza niż 2000.~~

Zadanie 5. Rozważamy błądzenie na grafie G o wierzchołkach $w_1, \ldots, w_{10}$, który jest spójny i nie jest dwudzielny. Niech $\bar{\pi} = (\pi_1, \ldots, \pi_{10})$ będzie rozkładem stacjonarnym dla tego błądzenia. Oceń poprawność poniższych zdań. Odpowiedź należy poprzeć albo uzasadnieniem ogólnym, albo przykładem, albo kontrprzykładem.

(a) $p^t_{w_1 w_2} \to \frac{\deg(w_1)}{2e(G)}$ przy $t \to \infty$.
(b) Jeżeli wierzchołek 2 jest stopnia 5, a $p_{1,2}(t) \to \frac{1}{8}$ przy $t \to \infty$, to G ma dokładnie 20 krawędzi.
(c) Jeżeli $\bar{\pi} = (\frac{1}{10}, \ldots, \frac{1}{10})$, to G jest grafem regularnym.
~~(d) Jeśli G jest 4-regularny i startujemy z wierzchołka w_1, to oczekiwana liczba kroków, po których wszystkie wierzchołki będą odwiedzone, jest mniejsza niż 5000.~~