TPI – 11

Łańcuchy Markowa – generowanie struktur dyskretnych, czasy pokrycia

Zadania domowe, przygotowujące do kartkówki

ZOT jest nieobowiązkowy; zainteresowani oddają na ćwiczeniach jego pisemne rozwiązanie.

A1. Korzystając z twierdzenia z wykładu, oszacuj z góry średni czas pokrycia dla klasycznego błądzenia po ścieżce o 3 wierzchołkach.

A2. Rozważmy łancuch Markowa, ktorego stanami są wszystkie 4 podzbiory zbioru $X = \{1, 2\}$. Zasady przejścia dla tego łańcucha są nastepujące:

- Mając zbiór A, rzuć monetą.
- Jeśli wypadł orzeł, to wylosuj jeden z dwóch elementów zbioru X, każdy z jednakowym prawdopodobieństwem. Zmień A na $A \cup \{x\}$.
- Jeśli wypadła reszka, to wylosuj jeden z dwóch elementów zbioru X, każdy z jednakowym prawdopodobieństwem. Zmień A na $A \setminus \{x\}$.

(a) Wyznacz macierz przejścia tego łańcucha.
(b) Wyznacz jego rozkład stacjonarny.
(c) Czy łańcuch jest ergodyczny?
(d) Czy łańcuch jest odwracalny?

A3. Rozważmy dwa łancuchy Markowa, L_1 i L_2, ktorych stanami jest pięć zbiorów niezależnych ścieżki na trzech wierzchołkach (jeden zbiór pusty, trzy jednoelementowe i jeden dwuelementowy).

- Zasady przejścia dla łańcucha L_1 są nastepujące:
  - Mając zbiór niezależny I, rzuć monetą.
  - Jeśli wypadł orzeł, to wylosuj jeden z trzech wierzcholkow v, każdy z jednakowym prawdopodobieństwem. Gdy $v \notin I$ i $I \cup \{v\}$ jest niezależny, wtedy zamień I na $I \cup \{v\}$; w przeciwnym przypadku zostań w I.
  - Jeśli wypadła reszka, to wylosuj jeden z trzech wierzcholkow v, każdy z jednakowym prawdopodobieństwem. Gdy $v \in I$, wtedy zamień I na $I \setminus \{v\}$; w przeciwnym przypadku zostań w I.
- Zasady przejścia dla łańcucha L_2 są takie:
  - Mając zbiór niezależny I, rzuć monetą.
  - Jeśli wypadł orzeł i istnieje przynajmniej jeden taki zbiór niezależny J, że $J \supset I$ i $|J \setminus I| = 1$, wybierz losowo jeden z takich zbiorów (każdy z jednakowym prawdopodobieństwem) i zamień I na J.
  - Jeśli wypadła reszka i istnieje przynajmniej jeden taki zbiór niezależny J, że $J \subset I$ i $|I \setminus J| = 1$, wybierz losowo jeden z takich zbiorów (każdy z jednakowym prawdopodobieństwem) i zamień I na J.
  - W pozostalych przypadkach zostań w I.

Wyznacz macierze przejścia i rozkłady stacjonarne dla obu łańcuchów.

A4. Opisz zwięźle ideę algorytmu, generującego losowy zbiór niezależny w grafie (w tym zbiór pusty) w taki sposób, że

(a) łańcuch Markowa dla tego procesu posiada rozkład stacjonarny, w którym wszystkie podzbiory są jednakowo prawdopodobne.
(b) po wielu krokach otrzymamy każdy podzbiór niezależny z prawdopodobieństwem w przybliżeniu jednakowym.
(c) łańcuch Markowa dla tego procesu posiada rozkład stacjonarny, w którym każdy ze zbiorów niezależnych A występuje z prawdopobieństwem $\frac{2^{|A|}}{\sum_Z 2^{|Z|}}$, gdzie suma w mianowniku przebiega wszystkie zbiory niezależne grafu (w tym zbiór pusty).

(d) łańcuch Markowa dla tego procesu posiada rozkład stacjonarny, w którym każdy ze zbiorów niezależnych A występuje z prawdopobieństwem $\frac{(1/2)^{|A|}}{\sum_Z (1/2)^{|Z|}}$, gdzie suma w mianowniku przebiega wszystkie zbiory niezależne grafu (w tym zbiór pusty).

Dokładnie uzasadnij poprawność algorytmu.

A5. Generujemy lasy rozpięte grafu K_3 następująco. Zaczynamy od dowolnego lasu rozpiętego, a w każdym kroku, mając wygenerowany las H, wybieramy z jednakowym prawdopodobieństwem jakąś krawędź z $E(K_3)$, a następnie rzucamy symetryczną monetą.

- Jeżeli wypadł orzeł i wybrana krawędź należy do lasu H, to ją usuwamy.
- Jeżeli wypadła reszka i wybrana krawędź nie należy do H, to dodajemy ją do H (o ile nie domyka cyklu).
- W przeciwnym wypadku nic nie robimy.

Budujemy łańcuch Markowa odpowiadający powyższemu algorytmowi, w którym stanami są wszystkie lasy rozpięte w K_3.

(a) Ile stanów ma ten łańcuch?
(b) Proszę wybrać w K_3 drzewo rozpięte i podać wiersz macierzy przejścia łańcucha, odpowiadający temu stanowi.
(c) Czy łańcuch jest ergodyczny?
(d) Czy łańcuch jest odwracalny?
(e) Jaki jest jego rozkład stacjonarny?
(f) Czy możemy stwierdzić, że po wielu krokach każdy las rozpięty będzie w przybliżeniu jednakowo prawdopodobny?
(g) Czy możemy stwierdzić, że po wielu krokach każde drzewo rozpięte będzie w przybliżeniu jednakowo prawdopodobne?

A6. Generujemy losowo drzewo rozpięte grafu K_3 poprzez wyróżnienie pewnych krawędzi wg następujących reguł. Zaczynamy od wierzchołka nr 1 i pustego zbioru krawędzi wyróżnionych. W każdym kroku przemieszczamy się wzdłuż kawędzi do wierzchołka sąsiedniego, wybranego z jednakowym prawdopodobieństwem. Przebytą krawędź dołączamy do wyróżnionych wtedy i tylko wtedy, gdy nie domyka cyklu zbudowanego z krawędzi wyróżnionych.

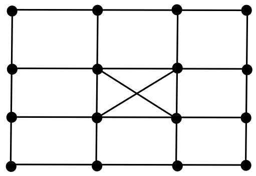
(a) Uzasadnij, że oczekiwana liczba kroków, po których wyróżnione krawędzie utworzą drzewo rozpięte, jest mniejsza niż 20.
(b) Dla każdego drzewa rozpiętego wyznacz prawdopodbieństwo jego wygenerowania po 4 krokach, to znaczy prawdopodobieństwo, że po 4 krokach wszystkie krawędzie tego drzewa są wyróżnione.
(c) Czy możemy stwierdzić, że po wielu krokach każde drzewo rozpięte będzie w przybliżeniu jednakowo prawdopodobne?

ZOT 7. Mysz i kot błądzą klasycznie (i niezależnie) na tym samym spójnym grafie 8-regularnym o 20 wierzchołkach, ktie jest dwudzielny. Na początku znajdują sią w pewnych niesąsiednich wierzchołkach. Gdy znajdą się w tym samym wierzchołku, kot zjada mysz. Uzasadnij, że oczekiwana liczba kroków, po których mysz zostanie zjedzona, jest mniejsza niż $625 \cdot 2^{15}$.

Zadania na ćwiczenia

Zad.1. Oblicz dokładnie średni czas pokrycia dla klasycznego błądzenia po ścieżce o 3 wierzchołkach.

Zad.2. Rozważmy błądzenie klasyczne na poniższym grafie.



Załóżmy, że startujemy z lewego górnego rogu. Oszacuj średni czas pokrycia z góry i z dołu.

Zad.3. Oto znany problem plecakowy. Mamy n przedmiotów, o zadanych objętościach. Mamy też plecak o znanej pojemności. Plecak możemy zapakować wybierając dowolnie zestaw przedmiotów (w tym zestaw pusty), o ile suma ich objętości nie przekracza pojemności plecaka. Chcemy wygenerować losowo różne zapakowania plecaka (czyli zestaw przedmiotów, który się w plecaku mieści). Zaproponuj algorytm oparty na łańcuchu Markowa o zbiorze stanów złożonym z wszystkich możliwych zestawów w plecaku, o poniższej własności.

(a) Łańcuch Markowa odpowiadający temu algorytmowi posiada rozkład stacjonarny, który każdemu zapakowaniu plecaka przypisuje jednakowe prawdopodobieństwo.
(b) Po wielu krokach otrzymamy każde zapakowanie z prawdopodobieństwem w przybliżeniu jednakowym.
(c) Po wielu krokach przybliżone prawdopodobieństwo danego zapakowania będzie funkcją malejącą względem liczby przedmiotów w plecaku, a dokładniej – będzie wprost proporcjonalne do $1/4^m$, gdzie m jest liczbą przedmiotów w plecaku.

Zakładamy, że stworzenie listy wszystkich możliwych do zapakowania zestawów jest nieekonomiczne.

Zad.4. (rezerwowe) Generujemy lasy rozpięte spójnego grafu G następująco. Zaczynamy od dowolnego lasu rozpiętego, a w każdym kroku, mając wygenerowany las H, wybieramy jednostajnie krawędź z $E(G)$ i z prawdopodobieństwem $\frac{1}{2}$ usuwamy ją z H (jeśli nie jest ona krawędzią H, to nic nie robimy), a z prawdopodobieństwem $\frac{1}{2}$ dodajemy ją do H (o ile nie domyka cyklu).

(a) Czy możemy stwierdzić, że po wielu krokach każdy las rozpięty będzie w przybliżeniu jednakowo prawdopodobny?
(b) Czy możemy stwierdzić, że po wielu krokach każde drzewo rozpięte będzie w przybliżeniu jednakowo prawdopodobne?

Zadania do samodzielnego rozwiązania później

B1. Pchła błądzi klasycznie po cyklu o 5 wierzchołkach. Oszacuj z góry i z dołu średni czas pokrycia dla tego łańcucha Markowa.

B2. Tasujemy talię n ($n \geqslant 2$) kart w taki sposób: zaczynamy od dowolnego stosu, wybieramy losowo (z jednakowym prawdopodobieństwem) jedną z n kart i kładziemy ją na wierzch.

(a) Sprawdź, że rozkład jednostajny jest rozkładem stacjonarnym tego łańcucha.
(b) Uzasadnij, że łańcuch jest nieokresowy i nieprzywiedlny.
(c) Czy po wielu krokach uzuskamy każde potasowanie z prawdopodobieństwem w przybliżeniu jednakowym?

Zad.5. Lisek Chytrusek zaproponował poniższy algorytm (oparty na łańcuchu Markowa) do dobrego potasowania 52 kart. Celem liska było stworzyć algorytm, który po wielu krokach (np. 100tys) wygeneruje każde potasowanie z prawdopodobieństwem w przybliżeniu jednakowym.

Algorytm liska: *Zaczynam od dowolnego potasowania. W każdym kroku wybieram losowo kolejno dwie, niekoniecznie różne, karty (każdą losuję jednostajnie) i zamieniam miejscami.*

Sprawdź poprawność algorytmu, tzn. czy rzyczywiście po wielu krokach każde potasowanie będzie w przybliżeniu tak samo prawdopodobne.

B3. Zaproponuj algorytm oparty na łańcuchu Markowa o 2^k stanach, który po wielu krokach wygeneruje każdy wierzchołek k-kostki z prawdopodobieństwem w przybliżeniu jednakowym. Zakładamy, że k jest duże i stworzenie listy wszystkich 2^k wierzchołków k-kostki jest nieekonomiczne.

B4. Opisz zwięźle ideę algorytmu, generującego losowe skojarzenia w grafie (w tym zbiór pusty), w taki sposób, by każde ze skojarzeń M wylosowane zostało z prawdopobieństwem

(a) prawie jednakowym.

(b) bliskim
$$\frac{5^{|M|}}{\sum_S 5^{|S|}},$$
gdzie $|M|$ jest liczbą krawędzi w skojarzeniu M, a suma w mianowniku przebiega wszystkie skojarzenia grafu (i zbiór pusty).

(c) bliskim
$$\frac{1}{3^{|M|} \cdot \sum_S (1/3)^{|S|}},$$
gdzie $|M|$ jest liczbą krawędzi w skojarzeniu M, a suma w mianowniku przebiega wszystkie skojarzenia grafu (i zbiór pusty).

Najważniejszą częścią tego zadania jest uzasadnienie poprawności algorytmu.

B5. Mamy 3 kolory i graf $G = C_4$, czyli będący cyklem o 4 wierzchołkach. Oto reguły przejścia dla algorytmu opartego na łańcuchu Markowa, którego zbiorem stanów są wszystkie właściwe kolorowania grafu G:

- Mając kolorowanie właściwe, wylosuj jeden z wierzcholkow $v \in V(G)$, każdy z jednakowym prawdopodobieństwem, następnie wylosuj jeden z kolorów i, każdy z jednakowym prawdopodobieństwem.
- Jeśli po przemalowaniu wierzchołka v na kolor i kolorowanie pozostaje właściwe, przemaluj v na kolor i.
- W przeciwnym przypadku pozostaw kolorowanie bez zmian.

(a) Ile stanów ma łańcuch Markowa odpowiadający podanemu algorytmowi?

(b) Wyznacz pierwszy wiersz macierzy przejścia tego łańcucha (stan numer jeden wybierz samodzielnie).

(c) Czy po bardzo dużej liczbie kroków każde właściwe kolorowanie będzie w przybliżeniu jednakowo prawdopodobne, niezależnie jaki był stan początkowy?

(d) Jeśli w poprzednim podpunkcie zapomniałeś sprawdzić, czy łańcuch jest ergodyczny, zrób to teraz.

B6. Mamy n przedmiotów, o zadanych objętościach. Mamy też plecak o znanej pojemności. Plecak możemy zapakować, wybierając dowolnie zestaw przedmiotów (w tym zestaw pusty), o ile suma ich objętości nie przekracza pojemności plecaka. Zaproponuj algorytm oparty na łańcuchu Markowa, który po wielu krokach wybierze zestaw przedmiotów możliwy do zapakowania w taki sposób, by przybliżone prawdopodobieństwo wylosowania zestawu było funkcją rosnącą liczby przedmiotów w zestawie. Najważniejszą częścią zadania jest uzasadnienie poprawności algorytmu.

B7. Oceń poprawność poniższych zdań. Odpowiedź należy poprzeć albo uzasadnieniem ogólnym, albo przykładem, albo kontrprzykładem.

(a) Jeśli błądzimy klasycznie na grafie G, który jest 4-regularny, to oczekiwana liczba kroków, po których wszystkie wierzchołki będą odwiedzone, jest mniejsza niż 5000.

(b) Jeśli błądzimy klasycznie na cyklu i oczekiwana liczba kroków, po których wszystkie wierzchołki będą odwiedzone, jest większa niż 1000, to cykl ten ma więcej niż 20 wierzchołków.