Wrocław, 22 marca 2020

Lista zadań nr 1

Symulacje i algorytmiczne zastosowanie łańcuchów Markowa

1 Rachunek prawdopodobieństwa

Zadanie 1. Niech T będzie zmienną losową przyjmującą wartości całkowite nieujemne. Pokaż, że $ET = \sum_{k=0}^{\infty} P(T > k)$.

Zadanie 2. Udowodnij następujący lemat

Lemat 1 (Nierówność Markowa)**.** *Niech X będzie nieujemną zmienną losową. Wtedy $\forall r > 0$*

$$P(X \geqslant r) \leqslant \frac{EX}{r}$$

Zadanie 3. Udowodnij następujący lemat

Lemat 2 (Nierówność Czebyszewa)**.** *Niech X będzie zmienną losową o średniej $\mu < \infty$ i wariancji $\sigma^2 < \infty$. Wtedy $\forall r > 0$*

$$P(|X - \mu| \geqslant r) \leqslant \frac{\sigma^2}{r^2}$$

Zadanie 4. Pokaż, że dla każdego x rzeczywistego mamy $1 + x \leqslant e^x$

Zadanie 5. Udowodnij następujący lemat

Lemat 3 (Nierówność Chernoffa)**.** *Niech X_i będą niezależnymi zmiennymi losowymi o rozkładzie $P(X_i = 1) = p_i = 1 - P(X_i = 0)$. Niech $S = \sum_{i=1}^{n} X_i$ oraz $\mu = ES = \sum_{i=1}^{n} EX_i = \sum_{i=1}^{n} p_i$. Wtedy dla każdego $\varepsilon > 0$ zachodzi*

$$P(S > (1+\varepsilon)\mu) \leqslant \left(\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right)^{\mu}$$

oraz

$$P(S < (1-\varepsilon)\mu) \leqslant \left(\frac{e^{-\varepsilon}}{(1-\varepsilon)^{1-\varepsilon}}\right)^{\mu}$$

Zadanie 6. Pokaż, iż poprzednie zadanie implikuje

$$P(S < (1-\varepsilon)\mu) \leqslant e^{\frac{-\mu\varepsilon^2}{2}}$$

2 Łańcuchy Markowa

Zadanie 7. *D*la Przykładu 1. z wykładu, tj. dla prostego symetryczneg błądzenia po 4 stanach: $E = \{1, 2, 3, 4\}$ z macierzą przejść

$$\mathbf{P} = \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{bmatrix}$$

z rozkładem początkowym $\mu = (1, 0, 0, 0)$

a) Oblicz μ^3

b) Oblicz μ^n ("zgadując" rozwiązanie i udowadniając indukcyjnie, że jest ono poprawne).

Zadanie 8. *M*odel pogody Gothenburga (podobny do tego co było na wykładzie). Rozważmy uproszczony model pogody: rozważamy tylko dwa stany: "deszcz" i "słońce". Po dniu deszczowym następuje dzień deszczowy z prawd. 0.75, natomiast po dniu słonecznym następuje dzień słoneczny z prawd. również 0.75. Oznaczmy przestrzeń stanów $E = \{\mathbf{e}_1, \mathbf{e}_2\}$, gdzie $\mathbf{e}_1$ oznacza dzień deszczowy, a $\mathbf{e}_2$ dzień słoneczny. Wtedy macierz przejść wygląda następująco:

$$\mathbf{P} = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}$$

Załóżmy, że $\mu = (0, 1)$.

a) Oblicz μ^2.

b) Oblicz μ^n stosując diagonalizację macierzy $\mathbf{P}$.

Zadanie 9. *R*ozważmy inny model pogody: rozważamy trzy stany: "ulewa", "lekki deszcz" i "słońce". Oznaczmy przestrzeń stanów $E = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, gdzie $\mathbf{e}_1$ oznacza dzień ulewny, $\mathbf{e}_2$ oznacza dzień z lekkim deszczem, a $\mathbf{e}_3$ dzień słoneczny. Wtedy macierz przejść wygląda następująco:

$$\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/6 & 1/3 & 1/2 \\ 0 & 1/3 & 2/3 \end{bmatrix}$$

Oblicz średni czas powrotu do dnia słonecznego (tzn. ile średnio dni mija między dwoma dniami słonecznymi).

Zadanie 10. *W* modelu pogody "L.A.", tj. modelu podobnym do poprzedniego zadania, ale z macierzą przejść

$$\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 \\ 1/10 & 9/10 \end{bmatrix}$$

(przykład rozważany na wykładzie). Podaj wzór na μ^k dla

a) $\mu = (1, 0)$, b) $\mu = (1/2, 1/2)$, c) $\mu = (1/6, 5/6)$

Czy w którymś z przypadków można zaobserwować coś ciekawego? Co to może oznaczać? Do czego zbiega μ^k, w każdym z przypadków, gdy $k \to \infty$?

Zadanie 11. Niech $(X_0, X_1, \ldots)$ będzie jednorodnym łańcuchem Markowa z macierzą przejść $\mathbf{P}$ i przestrzenią stanów $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_M\}$, tzn. $P(X_{n+1} = \mathbf{e}_j | X_n = \mathbf{e}_i) = \mathbf{P}(\mathbf{e}_i, \mathbf{e}_j)$. Pokaż, że

$$P(X_{n+m} = \mathbf{e}_j | X_n = \mathbf{e}_i) = \mathbf{P}^m(\mathbf{e}_i, \mathbf{e}_j)$$

Zadanie 12. Dla tasowania kart Top-To-Random dla $n = 3$ kart narysuj graf przejść oraz podaj macierz przejść $\mathbf{P}$. Wylicz również $\mathbf{P}^2$.

Zadanie 13. Dla dwóch miar probabilistycznych μ, ν na $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_M\}$ definiujemy *normę całkowitego wahania*: $d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{\mathbf{e} \in E} |\mu(\mathbf{e}) - \nu(\mathbf{e})|$.

Oblicz normę całkowitego wahania między rozkładami:
$\mu \sim Bin(4, 1/2), \nu \sim Bin(4, 1/4)$

Zadanie 14. Na przestrzeni $E = \{1, 2\}$ zdefiniujmy: $\mu^k = \left[\frac{1}{6} + \frac{1}{3}\left(\frac{2}{5}\right)^k, \frac{5}{6} - \frac{1}{3}\left(\frac{2}{5}\right)^k\right]$, oraz $\nu = \left[\frac{1}{6}, \frac{5}{6}\right]$.

a) Oblicz $d_{TV}(\mu^k, \nu)$.

b) Dla jakiego (całkowitego) k norma całkowitego wahania między μ^k i ν jest mniejsza od 0.01?

Zadanie 15. Rozważmy n kart. Niech μ będzie rozkładem jednostajnym na wszystkich permutacjach n kart. Natomiast niech ν będzie rozkładem jednostajnym na wszystkich takich permutacjach, w których pierwsza karta jest ustalona. Oblicz $d_{TV}(\mu, \nu)$.