Probabilistic techniques - tutorials

Problem set #5 - Markov chains.

Release: **December 7, 2023**. Hints: **January 4, 2024**. Deadline: **January 11, 2024**. Send solutions to dbulavka+pt23@kam.mff.cuni.cz.

1. By homogeneous discrete time Markov chain we mean a Markov chain $X_0, X_1, \dots$ such that for every n we have $\Pr[X_{n+1} = x | X_n = y] = \Pr[X_n = x | X_{n-1} = y]$.

   (a) Let $Y_0, Y_1, \dots$ be integers chosen independently uniformly from $\{-1, 0, 1\}$ and let $X_n = \max\{Y_0, \dots, Y_n\}$. Decide and justify whether $X_0, X_1, \dots$ is a homogeneous discrete time Markov chain. **[1]**

   Hint: Use the following equality of events, $\{X_n = x\} \cap \{X_{n-1} = y\} = \{\max\{Y_n, y\} = x\}$.

   (b) Let $Y_0, Y_1, \dots$ be integers chosen independently uniformly from $\{-1, 0, 1\}$ and let $X_n = \max\{Y_n, Y_{n-1}\}$. Decide and justify whether $X_0, X_1, \dots$ is a homogeneous discrete time Markov chain. **[1]**

   Hint: Compare the probability $\Pr[X_2 = -1 | X_1 = 0, X_0 = -1]$ and $\Pr[X_2 = -1 | X_1 = 0]$.

2. Construct satisfiable 2-SAT formulae with n variables (for arbitrary large n) such that the randomized 2-SAT algorithm from the lecture takes $\Omega(n^2)$ steps (in expectation) to find a satisfying assignment. (Recall that in each step the algorithm picks an arbitrary unsatisfied clause and flips one of its variables) **[2]**

   Hint: Consider the 2-SAT formula $\bigwedge_{i \neq j \in [n]} (x_i \vee \neg x_j)$ and follow the same procedure as in the lecture. Notice that this one has two possible solutions, all 1's or all 0's

3. We say that a graph G is triangle 2-colorable if there is a 2-coloring of the vertices such that no triangle in G is monochromatic. (For example, a 3-colorable graph is triangle 2-colorable.) Let G be a 3-colorable graph. Start with $c\colon V(G) \to \{0, 1\}$ being constant 0, and then do the following: while there is a monochromatic triangle in G (with respect to c), pick one of its vertices at random and flip its color. Show that the expected number of steps for this process to find a suitable triangle 2-coloring is polynomial in $|V(G)|$. **[3]**

   Hint: Take l be a 3-coloring of G, c' to be a triangle 2-coloring. Let V_i denote the vertices of color i by l. Set X_i to denote the random variable representing the number of vertices from $V_1 \cup V_2$ on which c and c' coincide on step i of the algorithm. Proceed as in the lecture when analyzing the 2-SAT algorithm.

4. Show that in one communicating class of a Markov chain, there cannot be a state which is recurrent and another state which is transient. **[2]**

   Hint: Use that a state i is tanscient iff $\sum_{n \geq 0} P^n_{i,i} < \infty$ and $P^{n+a+b}_{i,i} \geq P^a_{i,j} P^n_{j,j} P^b_{j,i}$.

5. In a finite Markov chain show that the following holds:

   (a) There is at least one recurrent state. **[2]**

   Hint: Use the characterization of states showed in the tutorial as well as the identity $\sum_i P^n_{j,i} = 0 = 1$ for all n, where the sum is over all states.

(b) All recurrent states are positive recurrent (i. e., the expected time until the process returns to this state is less than infinity). **[2]**

Hint: Say we want to show that state i is positive recurrent. Given a particular execution of the Markov chain, if the return time is k, then it must have spent k_j time in state j, for all j different from i, before returning to i. Set $N_{i,j}$ to be the number of visits to j before returning to i, startin from i. Upper bound the expectation of $N_{i,j}$.

6. Let X be a Poisson random variable with mean $\lambda \in \mathbb{N}$. Prove that $\Pr[X \geq \lambda] \geq 1/2$. **[2]**

   Hint: Show that $\Pr[X = \lambda + h] \geq \Pr[X = \lambda - h - 1]$ for every $0 \leq h \leq \lambda - 1$.

7. Throw m balls uniformly independently into n bins, where the bins are numbered from 0 to $n-1$. We say that there is a k-gap starting at j if bins $j, j+1, \ldots, j+k-1$ are all empty.

   (a) Compute the expected number of k-gaps. **[1]**

   Hint: Write the random variable denoting the number of k-gaps as a sum of indicator random variables for each k-gap and use linearity of expectation.

   (b) Prove a Chernoff-like bound for the number of k-gaps X, that is, the inequality

   $$\Pr[X \geq (1+\delta)\,\mathbb{E}[X]] \leq e^{-c\,\mathbb{E}[X]\delta^2}$$

   where $0 < \delta < 1$ and c is a positive constant. **[4]**

   Hint: If you let $X_i = 1$ when there is a k-gap starting at bin i, then there are dependencies between X_i and X_{i+l}. To avoid these dependencies, you might consider X_i and X_{i+k}