Assignment 3

COMP 3801- Fall 2024

1 Instructions

Each question is worth 10 marks. The assignment is due in Brighspace before the due date. Please write clearly and answer questions precisely. As a thumb rule, the answer should be limited to ≤ 2 written pages, with ample spacing between lines and in margins, for each Question. Always start a new question on a new page, starting with Question 1, followed by Question 2, ..., Question n. Please cite all the references you have used/consulted as the source of information for each question. If a question asks you to Prove or Show, please clearly spell out the proof - the technique used - and show each step of the proof. Don't expect (partial) marks if the main idea isn't clear to us.

2 Problems

1. Assume that the total population of the Province of Ontario is 16 million. Each resident of Ontario is either a fan of the Ottawa Senators or Maple Leafs Hockey Club. Each year, 5% fans of Senators switch their aligns to Leafs, and the 10% of Leafs fans switch their aligns to Senators. Suppose that this trend continues year after year, and the population of Ontario doesn't change, what will be the steady state of the number of fans of each of the Hockey Clubs? Is that affected by the initial number of fans each team has?

2. (a) Let A and B be two $n \times n$ Markov matrices. Is the product $C = AB$ always a Markov matrix? Justify.

   (b) Construct a directed graph $G = (V, E)$ with $|V| = 5$ and $|E| \geq 6$ such that between each pair of vertices $u, v \in V$ there is a directed path from u to v. Find the page rank of each vertex with and without the teleportation probability of 0.2. (Use some math software to compute eigenvalues/eigenvectors).

3. Suppose a compulsory textbook for a course costs \$200 to purchase or \$20/week to rent. You love to take this course, but you are unsure whether you will stay until the end of this course. Assume that the whole course runs for 16 weeks, and you can drop it at any point. Suggest your strategy: should you rent the book or buy it? Justify, preferably using competitive analysis. (Suppose you rent the book for three weeks and then decide to buy. It will cost you a total of 3*\$20 +\$200=\$260, whereas after three weeks, if you drop the course, it costs you only \$60). Note that once you buy the book, you can't return it.

4. Consider the balance algorithm, where we have two advertisers with the same starting budget and the bid values are identical. In this case, it was shown that the Balance algorithm always

achieves revenue of at least 3/4 of the revenue obtained by an optimal off-line algorithm. How can the analysis be modified if the advertises have unequal budgets to begin with?

5. Suppose we have three advertisers, A_1, A_2, and A_3, where each with a budget of B dollars (B is a positive integer). Advertiser A_1 bids only for the item of type a, A_2 bids for items of types a and b, and A_3 bids for items of types a, b, and c. Assume that the click-through rate is \$1 for each of the advertisers. Assume that the online sequence of $3B$ queries consists of a random permutation of B queries for a, B queries for b, and B queries for c. Is it true that the competitive ratio of the BALANCE algorithm is at least 2/3? (Note that an optimal offline algorithm earns a revenue of $3B$). For example, if $B = 4$, a possible query sequence may be baaccacbbbca

6. Show that for a real symmetric matrix $S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, its eigenvectors are $x_1 = \begin{bmatrix} b \\ \lambda_1 - a \end{bmatrix}$ and $x_2 = \begin{bmatrix} \lambda_2 - d \\ c \end{bmatrix}$ where λ_1 and λ_2 are the eigenvalues, respectively. Will x_1 and x_2 be the eigenvectors even if S isn't symmetric?

7. Show that for any real matrix A of dimensions $m \times n$, $A^T A$ is a symmetric positive semi definite matrix, i.e. all its eigenvalues are non-negative.

8. Let S be a $n \times n$ real symmetric matrix and let $v_1, \ldots, v_n$ be its n orthonormal eigenvectors corresponding to its n eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$, respectively. Let $x \in \Re^n$ be any vector, and we can express it as a linear combination of vectors $v_1, \ldots, v_n$, i.e., $x = \alpha_1 v_1 + \ldots + \alpha_n v_n$, for some constants $\alpha_1, \ldots, \alpha_n$. Answer the following:

   (a) Show that $x^T x = \alpha_1^2 + \ldots + \alpha_n^2$.

   (b) Show that $x^T S x = \lambda_1 \alpha_1^2 + \ldots + \lambda_n \alpha_n^2$.

   (c) For a vector $x \in \Re^n$, consider the expression $f(x) = \frac{x^T S x}{x^T x}$. Show that for any eigenvector v_i of S, $f(v_i) = \lambda_i$.

   (d) Show that $\max\limits_{x \in \Re^n} \frac{x^T S x}{x^T x} = \lambda_1$.

9. Let $G = (V, E)$ be a simple undirected graph. A subset $S \subseteq V$ is said to *hit* E, if for each edge $e = (u, v) \in E$, $u \in S$ or $v \in S$. A set $S^* \subseteq V$ is said to be an *optimal hitting set* if it has the smallest cardinality among all hitting sets of G.

   The graph G is presented in an online manner as follows. We know all its vertices $V = \{1, \ldots, n\}$ in advance, and the edge set is empty. At each time instance $t = 1, \ldots, |E|$, an edge e_t between a pair of vertices is added to G. Thus, at each time instance t, we have the graph $G_t = (V, E_t = \bigcup_{i=1}^{t} e_i)$.

   We need to devise an online algorithm that maintains a hitting set $S_t \subseteq V$ of G_t, for $t = 1, \ldots, |E|$, whose cardinality is at most two times the cardinality of an optimal (offline) hitting set. I.e., if S_t^* is an optimal (offline) hitting set of G_t, $|S_t| \leq 2|S_t^*|$, for $t = 1, \ldots, |E|$. Explain your online algorithm and analyze its competitive ratio.

   Note: This is an online algorithm. Therefore, whichever vertices are added to S_t, at time t, will remain in the hitting set till the end of the algorithm. I.e. $S_1 \subseteq S_2 \subseteq \cdots \subseteq S_{|E|}$.

10. You are standing in the middle of a very long horizontal street at the location M, where houses are on the North side of the road, and the street is oriented East-West. Houses are

equally spaced, say every 100 meters. Each home has a unique number, but the numbering is chaotic and doesn't follow a sequential order. In fact, we can't make any assumptions about the distribution of the house numbers. We want to visit house #3801, but we don't have any way of knowing whether it is on the East or the West side of M. We employ the following strategy to find the house #3801.

(a) Initialize $x := 100$ meters.

(b) Repeat the following steps till success:

   i. From M, travel distance of x towards East. If we find #3801 on the way, we stop the search and terminate.

   ii. Return to M. Set $x := 2x$.

   iii. From M, travel distance of x towards West. We stop and terminate the search if we find #3801 on the way.

   iv. Return to M. Set $x := 2x$.

Let D be the distance from M to the house # 3801. Assume that $D \geq 100$. Show that the total distance that we travel using the above method is $\leq 9D$. Note that the algorithm doesn't know the value of D.