TD7: Separation Oracles and LP Rounding

Exercise 1. Prove the following theorem.

(Separating Hyperplane Theorem) Let K be a non-empty closed convex subset of $\mathbb{R}^n$ and let $\mathbf{x}^* \in \mathbb{R}^\mathbf{n}$ be a vector that does not belong to K. Then, there exists some vector $\mathbf{c} \in \mathbb{R}^\mathbf{n}$ such that $\mathbf{c} \cdot \mathbf{x}^* < \mathbf{c} \cdot \mathbf{x}$ for all $\mathbf{x} \in \mathbf{K}$.

Exercise 2. The *maximum acyclic subgraph* problem is defined as follows. Given a directed graph $G = (V, A)$, the goal is to find a subset of the edges $S \subset A$ such that S is acyclic (contains no directed cycles) and $|S|$ is maximized. Consider the following linear programming relaxation:

$$\begin{aligned} &\max \sum_{ij \in A} x_{ij} \\ &\sum_{ij \in C} x_{ij} \leq |C| - 1, \quad \text{for all directed cycles } C \in A, \\ &0 \leq x_{ij} \leq 1, \quad \text{for all edges } ij \text{ in } A. \end{aligned} \tag{$P_{acyclic}$}$$

Give a polynomial-time separation oracle for $(P_{acyclic})$.

Exercise 3. In class, we gave a $\frac{3}{4}$-approximation algorithm for the maximum satisfiability problem. Give a tight example for this algorithm. In other words, give an instance for which the expected value of the solution returned by the algorithm is $\frac{3}{4}OPT$.

Exercise 4. We introduced the *set cover* problem in class. Given a universe $\mathcal{U} = \{e_1, e_2, \ldots, e_n\}$ and a family of sets $\mathcal{T} = \{T_1, T_2, \ldots, T_m\}$ where $T_j \subseteq U$, find a subset of $\mathcal{T}$ (i.e. a subset of the sets in $\mathcal{T}$) such that each element in $\mathcal{U}$ is contained in at least one chosen set.

$$\begin{aligned} &\min \sum_{S \in \mathcal{T}} x_S \\ &\sum_{S : e \in S} x_S \geq 1, \quad \text{for all } e \in \mathcal{U}, \\ &x_S \geq 0, \quad \text{for all } S \in \mathcal{T}. \end{aligned}$$

Construct an instance with an integrality gap greater than 1. What is the largest integrality gap you can find?

Exercise 5. Vertex Cover

Given a graph $G = (V, E)$, recall that the *vertex cover* problem is to find a minimum cardinality subset $S \subset V$ such that for each edge $ij \in E$, at least one endpoint belongs to S. Assume that we have found an extreme point $\mathbf{x}^* \in \mathbb{R}^{|\mathbf{V}|}$ such that $\mathbf{x}^*$ is an optimal solution to the linear programming relaxation for vertex cover. We saw in class that $\mathbf{x}^*$ is half-integral.

a. Let V_0, V_1 and $V_{1/2}$ be the set of vertices valued respectively 0,1 and 1/2. Let $OPT_{1/2}$ be the minimum value of a vertex cover in the graph induced by $V_{1/2}$. Let OPT be the minimum value of a vertex cover in G. Show that $OPT \leq OPT_{1/2} + |V_1|$.

b. Let V_{OPT} be a set achieving OPT. Show that $|V_{\text{OPT}} \cap V_0| \geq |V_1 \setminus V_{\text{OPT}}|$.

c. Deduce conversely that $OPT \geq OPT_{1/2} + |V_1|$.

d. Given a graph G and an integer k, we want to output TRUE if G has a vertex cover of size at most k. Show that this problem admits a polytime reduction to an equivalent instance (G', k') in which G' has at most $2k$ vertices.

Exercise 6. Let $\mathcal{U}$ be a universe of elements and let $T_1, \ldots, T_m$ be a collection of subsets of $\mathcal{U}$. Recall that the (unweighted) *hitting set* problem is to find a minimum cardinality subset $S \subseteq \mathcal{U}$ such that for each $j = 1, 2, \ldots, m$, the intersection $T_j \cap S \neq \emptyset$. Let $y^* \in \mathbb{R}^{|\mathcal{U}|}$ be an optimal solution for the linear programming relaxation of the hitting set problem, i.e. y^* is a vector of values between 0 and 1 indexed by elements in $\mathcal{U}$.

a. Is it true that if all T_i have size 3, then y^* is third integral, i.e. only has coordinates 0, 1/3, 2/3, and 1?

b. Propose a construction in which the minimum non-zero value of the coordinates in y^* can be arbitrarily small.

Exercise 7. Relaxations for Multicommodity Flows

Given a graph $G = (V, E)$ and k pairs (s_i, t_i) (where $s_i, t_i \in V$ for all $k = 1, \ldots, k$), our goal is to find a path from s_i to t_i for $i = 1, \ldots, k$ so that the maximum *edge congestion* is minimized. Let P_i denote the set of all paths from s_i to t_i. Consider the following linear programming relaxation.

$$\begin{aligned} &\min C \\ &\sum_{p \in P_i} x_p = 1, \quad \text{for all } i = 1, \ldots, k, \\ &\sum_{p : e \in p} x_p \leq C, \quad \text{for all } e \in E, \\ &x_p \geq 0. \end{aligned} \tag{P^1_{flow}}$$

As noted in class, this linear program may have an exponential number of variables. Consider another linear programming relaxation with a polynomial number of variables. In this relaxation, x_{ie} represents the number of paths using edge e.

$$\begin{aligned} &\min C \\ &\sum_{e \in \delta^+(v)} x_{ie} = \sum_{e \in \delta^-(v)} x_{ie}, \quad \text{for all } i = 1, \ldots, k \text{ and } v \neq s_i, t_i, \\ &\sum_{e \in \delta^-(s_i)} x_{ie} = \sum_{e \in \delta^+(t_i)} x_{ie} = 1, \quad \text{for all } i = 1, \ldots, k, \\ &\sum_{i=1}^{k} x_{ie} \leq C, \quad \text{for all } e \in E \\ &x_{ie} \geq 0. \end{aligned} \tag{P^2_{flow}}$$

Show that relaxations (P^1_{flow}) and (P^2_{flow}) are equivalent in the sense that an optimal solution for (P^1_{flow}) can be converted to an optimal solution for (P^2_{flow}) and vice versa.

Exercise 8. Chernoff Bounds

In the U.S. State of Massachusetts, there was a primary in March 2016. Approximately 1.2 million votes were cast with 49.9% for Hillary Clinton and 48.5% for Bernie Sanders (i.e. the "official" results). An Edison Research exit poll of 1,297 voters found that 45.7% and 52.3% of those surveyed voted for Clinton and Sanders, respectively. Assuming that the official election results are correct, what is the probability that an exit poll of the given sample size finds that at most 45.7% respondents voted for Hillary Clinton? (Assume that each of the 1,297 samples in the exit poll is drawn independently and uniformly at random from the total population of 1.2 million voters.)