Exercise Set V

These exercises are for your own benefit. Feel free to collaborate and share your answers with other students. Solve as many problems as you can and ask for help if you get stuck for too long. Problems marked * are more difficult but also more fun :).

These problems are taken from various sources at EPFL and on the Internet, too numerous to cite individually.

1 *(homework problem from previous year)* **Randomized rounding.** Consider the standard linear programming relaxation of Set Cover that we saw in class. We gave a randomized rounding algorithm for the Set Cover problem. Use similar techniques to give an algorithm that, with probability at least a positive constant, returns a collection of sets that cover at least 90% of the elements and has cost at most a constant factor larger than the LP solution.

2 (*) Consider the LP-rounding algorithm for Set Cover that works as follows:

1. Solve the LP relaxation to obtain an optimal solution x^*.
2. Return the solution $\{S : x^*_S > 0\}$, i.e., containing all sets with a positive value in the fractional solution.

Use the complementarity slackness conditions to prove that the algorithm is an f-approximation algorithm, where f is the frequency (i.e., the maximum number of sets that any element belongs to).

3 Consider the following grumpy commuters problem. There are n persons. Each person $i = 1, \ldots, n$ has to choose a path from his/her home to work from a set $\mathcal{P}_i$ of potential paths (in an underlying graph with n vertices that models the road network). In addition, person i has a personal preference of paths modeled with a cost function $c_i : \mathcal{P}_i \to \mathbb{R}_+$. The goal is to select paths $p_1 \in \mathcal{P}_1, p_2 \in \mathcal{P}_2, \ldots, p_n \in \mathcal{P}_n$ (one for each person) so that

1. The paths $p_1, \ldots, p_n$ are edge-disjoint (the commuters are grumpy and do not want to share roads).
2. The total cost $c_1(p_1) + c_2(p_2) + \cdots + c_n(p_n)$ is minimized.

3a Write an exact integer linear program for the grumpy commuters problem.

3b *(*) Give* a randomized rounding algorithm that returns a set of paths $p_1 \in \mathcal{P}_1, p_2 \in \mathcal{P}_2, \ldots, p_n \in \mathcal{P}_n$ satisfying:

1. The total cost $c_1(p_1) + c_2(p_2) + \cdots + c_n(p_n)$ is at most a constant factor times the cost of the optimal value of the LP-relaxation.
2. Each edge is used by at most $100 \cdot \log n / \log \log n$ paths.

You may assume that the underlying graph has n vertices.

Hint: for the analysis (of the second condition) the following specialized Chernoff bound[1] *will be useful: Let $X_1, \ldots, X_n$ be n* independent *random variables taking values in $\{0, 1\}$. If $\mathbb{E}[X_1 + \cdots + X_n] \le 1$, then*

$$\Pr[X_1 + \cdots + X_n > 100 \cdot \log n / \log \log n] < \frac{1}{n^3}.$$

4 Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Consider the following linear program with n variables:

$$\begin{aligned} \textbf{maximize} \quad & c^T x \\ \textbf{subject to} \quad & Ax = b \\ & x \ge 0 \end{aligned}$$

Show that any extreme point x^* has at most m non-zero entries, i.e., $|\{i : x^*_i > 0\}| \le m$.

Hint: what happens if the columns corresponding to non-zero entries in x^* are linearly dependent?

(If you are in a good mood you can prove the following stronger statement: x^* is an extreme point if and only if the columns of A corresponding to non-zero entries of x^* are linearly independent.)

5 Consider the following quadratic programming relaxation of the Max Cut problem on $G = (V, E)$:

$$\begin{aligned} \textbf{maximize} \quad & \sum_{\{i,j\} \in E} (1 - x_i) x_j + x_i (1 - x_j) \\ \textbf{subject to} \quad & x_i \in [0, 1] \quad \forall i \in V \end{aligned}$$

Show that the optimal value of the quadratic relaxation actually equals the value of an optimal cut. (Unfortunately, this does not give an exact algorithm for Max Cut as the above quadratic program is NP-hard to solve (so is Max Cut).)

Hint: analyze basic randomized rounding.

[1] Chernoff bounds are an extremely useful tool for proving concentration of sums of independent random variables taking values in $\{0, 1\}$. We will revisit them later in the course.