COS 521: Advanced Algorithm Design
Homework 1
Due: Sun, April 7

Collaboration Policy: You may collaborate with other students on these problems. Collaboration is limited to discussion of ideas only, and you should write up the solutions entirely *on your own* and *list your collaborators* as well as *cite any references* (book, paper, etc.) you may have used. Limit your answers for each problem to two pages or less (one page for most problems) – you need to give enough detail to convince the grader.

1. Write a linear program for finding the largest sphere contained inside a given polyhedron $\{x : Ax \leq b\}$ and explain why it is correct. What does it mean if this program is infeasible or unbounded ?

2. Suppose we are given points $(x_1, y_1), \ldots (x_n, y_n)$ in the plane and want to fit a line $y = ax + b$ to them. There are various notions of what a good line is; the most common one (linear regression) seeks to minimize the ℓ_2^2 error:

$$\epsilon_2(a, b) = \sum_i (ax_i - b - y_i)^2$$

and can be solved using linear algebra. Write linear programs which compute lines minimizing the ℓ_1 and ℓ_∞ error, namely

$$\epsilon_1(a, b) = \sum_i |ax_i - b - y_i|,$$
$$\epsilon_\infty(a, b) = \max_i |ax_i - b - y_i|.$$

What is the interpretation of dual feasible solutions for these programs?

3. Suppose we wish to solve the following flow problem: There are n nodes in an undirected graph, and the edges should be thought of as pipes of a certain capacity. For each pair $\{i, j\}$ we wish to send 1 unit of flow between them. All these flows must be routed through the pipes and should not violate any capacity. Let z be the minimum number such that if all edge pipes have capacity z then the flows can be routed in the network.

   (a) Express the problem of finding z as a linear program, and argue that it can be solved in polynomial time.

   (b) Use the multiplicative weights method to design a algorithm that solves the above linear program approximately. How long does your algorithm take to find z correctly up to an additive error $\epsilon > 0$?

4. Consider the following optimization problem with *robust conditions*:

$$\min\{c^T x : x \in \Re^n; Ax \geq b \text{ for any } A \in F\},$$

where $b \in \Re^m$ and F is a set of $m \times n$ matrices:

$$F = \{A; \forall i, j; a_{ij}^{min} \leq a_{ij} \leq a_{ij}^{max}\}.$$

   (a) Considering F as a polytope in $\Re^{m\times n}$, what are the vertices of F ?

   (b) Show that instead of conditions for all $A \in F$, it is enough to consider the vertices of F. Write the resulting linear program. What is its size ? Is this poynomial in the size of the input namely m, n and the sizes of b, c, a_{ij}^{min} and a_{ij}^{max} ?

   (c) Derive a more efficient description of the linear program: Write the conditions on x given by one row of A, for all choices of A. Formulate the condition as a linear program. Use duality and formulate the original problem as a linear program. What is the size of this one ? Is this polynomial in the size of the input ?

5. Consider the LP for the Set Cover problem we discussed in class. It has a variable x_j for every set S_j. Suppose x^* is an optimal solution to this LP. Consider the following randomized algorithm: Pick set S_j with probability x_j^*, where these choices are made independently for each j.

   (a) For any element e_i, compute the probability that e_i is covered by the sets picked by the algorithm. What is the expected cost of the sets picked?

   (b) By repeating this process multiple times, describe how you would obtain a randomized approximation algorithm for Set Cover. (*Hint:* How may iterations do you need to ensure that, with probability $\geq 1/2$, every element is covered?)

6. Given a graph $G(V, E)$, we would like to find a subset of vertices $S \subseteq V$ so as to maximize $\frac{|E(S)|}{|S|}$. Here $E(S)$ is the set of edges in the subgraph induced by S.

   (a) Write a linear programming relaxation for this problem using a variable for every vertex and a variable for every edge.

   (b) Show how the optimum solution to the problem can be obtained from the linear program. (*Hint:* Think about the method we used to obtain a cut from the max-flow LP dual.)