Problem Set 7

Due: Wednesday, October 26, 2016.

Collaboration policy: collaboration is *strongly encouraged.* However, remember that

1. You must write up your own solutions, independently.
2. You must record the name of every collaborator.
3. You must actually participate in solving all the problems. This is difficult in very large groups, so keep your collaboration groups limited to 3 people in a given week.
4. **No bibles. This includes solutions posted to problems in previous years.**

Problem 1. In class, we went from an *infeasibility witness* for linear systems to a *lower bound* witness for linear programs. There is a nice intermediary, an *infeasiblity witness* for linear programs. Farkas' lemma states that *exactly one* of the following two systems has a solution:

1. $Ax = b$ and $x \geq 0$
2. $yA \geq 0$ but $yb < 0$

We will prove Farkas' lemma.

(a) Give a direct proof that the two systems cannot both be feasible ("weak duality").

(b) Consider the linear program $\min yb \mid yA \geq 0$. Show that it can only have two possible answers. (An "answer" to either be any real number, "unbounded," or "infeasible.")

(c) Prove Farkas' lemma by considering the associated linear program and applying strong duality. (Typically, Farkas' lemma is proved first and then used in the proof of LP duality.)

Problem 2. Suppose you are given two polyhedra $P = \{x \mid Ax \leq b\}$ and $Q = \{x \mid Dx \leq e\}$.

(a) Using duality, (or Farkas' lemma), prove that if polyhedra P and Q have empty intersection (i.e., no point is in both), then there are $y, z \geq 0$ such that $yA+zD = 0$ but $yb + ze < 0$.

(b) Conclude that if polyhedra P and Q have empty intersection (i.e., no point is in both), then there is a separating hyperplane for P and Q (i.e., a vector c such that $c \cdot x < c \cdot w$ for all $x \in P$ and $w \in Q$). **Hint:** consider $c = yA$ from the previous part.

(c) Conclude that given the two polyhedra, there is an easily verifiable answer as to whether or not the two polyhedra have a point in common.

NONCOLLABORATIVE Problem 3. Another way to formulate the maximum-flow problem as a linear program is via flow decomposition. Suppose we consider all (exponentially many) s-t paths P in the network G, and let f_P be the amount of flow on path P. Then maximum flow says to find

$$z = \max \sum f_P$$

subject to

$$\sum_{P \ni e} f_P \leq u_e \qquad \text{for all edges } e,$$

$$f_P \geq 0 \qquad \text{for all paths } P.$$

(The first constraint says that the total flow on all paths through e must be less than u_e.) Take the dual of this linear program and give an English explanation of the objective and constraints.

Problem 4. Consider a directed graph in which edges have costs (possibly negative, representing profits). Suppose you want to find a *minimum mean cycle* in this graph: one with the minimum ratio of cost to length (number of edges). Going around such a cycle repeatedly (assuming it is negative) provides you with the maximum possible profit per unit length/time, so it is the fastest way to earn money if you are, for example, a delivery service. Finding a minimum mean cycle is also an essential step in certain strongly polynomial min cost flow algorithms. Consider the following linear program, with a variable f_{ij} for each edge (i,j):

$$\begin{aligned} w &= \min \sum c_{ij} f_{ij} \\ \sum_j f_{ij} - f_{ji} &= 0 \quad (\forall i) \\ \sum_{i,j} f_{ij} &= 1 \\ f_{ij} &\geq 0 \end{aligned}$$

(a) Explain why this captures the minimum mean cycle problem. (**Hint:** $\{f_{ij}\}$ is a circulation, so it can be decomposed into cycles.)

(b) Give the dual of this linear program—it will involve maximizing a certain variable λ.

(c) Give an explanation (in terms of min-cost-flow reduced costs) for why this dual formulation also captures minimum mean cycles. (**Hint:** how much is added to the cost of a k-edge cycle?)

(d) Let's assume the costs c_{ij} are integers. Suggest a combinatorial algorithm (not based on linear programming) that uses binary search to find the right λ to solve the dual problem. Can you use this to find a minimum mean cycle? **Note:** to know when you can terminate the search, you will need to lower bound the difference between the smallest and next smallest mean cost of a cycle.

Problem 5. Although the dual can tell you a lot about the structure of a problem, knowing an optimal dual solution does not in general help you solve the primal problem. Suppose we had an algorithm that could optimize an LP with an $m \times n$ constraint matrix in $O((m+n)^k)$ time given an optimal solution to the dual LP, where $k > 2$ is some positive constant. The algorithm is not guaranteed to work if given an incorrect solution to the dual.

(a) Argue that any LP optimization problem can be transformed into the following form:

$$\begin{array}{ll} \text{minimize} & 0 \cdot x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

(This LP has optimum value 0 if it is feasible, and ∞ if it is infeasible.) **Hint:** Recall our argument that optimizing an LP can be reduced to checking feasibility of an LP.

(b) What is the dual of this linear program?

(c) Argue that, if the primal is feasible, then the dual has an obvious optimum solution.

(d) Deduce that, given the hypothetical algorithm above, you can build an LP algorithm that will solve any LP *without* knowing beforehand a dual solution, in the same asymptotic time bounds as the algorithm above.

OPTIONAL Problem 6. Markov chains. An $n \times n$ matrix P is **stochastic** if all entries are nonnegative and every row sums to 1, that is $\sum_j p_{ij} = 1$ (so each row can be thought of as taking a convex combination). Stochastic matrices are used to represent the **transition matrices** of **Markov chains**—random walks through a series of states. The term p_{ij} represents the probability, if you are in a current state i, that your next state will be j (thus the sum-to-one rule). If you have a probability distribution π over your current state, where π_i denotes the probability you are in state i, then after a transition with probability defined by P, your new probability distribution is πP.

Use duality to prove that for any stochastic matrix P, there is a **nonzero** vector $\pi \geq 0$ such that $\pi P = \pi$. The vector π can be normalized to 1, in which case it represents a probability distribution that is **stationary** under the action of the transition matrix—that is, if π is the probability distribution on what state you are in before a transition, it is also the probability distribution after the transition. This proves that every Markov chain has a stationary probability distribution.

Hint: you must somehow express the constraints $\pi \geq 0, \pi \neq 0$, i.e. some entry in π is positive. Consider the constraint $\sum \pi_i = 1$. (Note that $\pi > 0$ means that each entry of π is at least 0, and at least one entry is greater than 0.)

Problem 7. How long did you spend on this problem set? Please answer this question using the Google form that is located on the course website. This problem is mandatory, and thus counts towards your final grade. It is due by the Monday 2:30pm after the pset due date.