A Theorist's Toolkit **CMU 15-751*, Fall 2016**

Problem Set 6

Due: Noon, Monday Nov. 7, email the pdf to toolkit2016homework@gmail.com

---

Homework policy: Exactly the same as last time.

---

1. Consider the system of inequalities

$$\begin{aligned} a^{(1)} \cdot x &\le b_1 \\ a^{(2)} \cdot x &\le b_2 \\ &\cdots \\ a^{(m)} \cdot x &\le b_m, \end{aligned}$$

where $x = (x_1, \dots, x_n)$ is a vector of n real variables, and where $a^{(i)} \in \mathbb{R}^n$, $b_i \in \mathbb{R}$ are constants. In class we discussed how you could attempt to prove that this system is infeasible (i.e., has no real solution) by finding numbers $\lambda_1, \dots, \lambda_m \ge 0$ such that, if you multiply the ith inequality above by λ_i and add up the results, you get the inequality $0 \le -1$. By analyzing "Fourier–Motzkin Elimination", we showed that such λ_i's exist if and only if the system is infeasible. This fact is called the *Farkas Lemma*.

Assume the above system is indeed feasible. Suppose we now further consider the *optimization* problem of maximizing $c \cdot x$ over all x satisfying the above inequalities, where $c \in \mathbb{R}^n$. Let M be the maximum possible value, and assume M is finite. As mentioned in class, one way you could try to prove an upper bound on M is to find numbers $y_1, \dots, y_m \ge 0$ such that if you multiply the ith inequality above by y_i and sum up the results, you get an inequality of the form $c \cdot x \le \beta$. Explain why any β attainable this way is an upper bound on M, explain why finding $y_1, \dots, y_m$ achieving the minimum possible value of β is an instance of Linear Programming, and finally prove that the minimum possible value of β is exactly M. For the last of these, your proof should use the Farkas Lemma as a black box.

2. Let $A \in \mathbb{R}^{m \times n}$ be an $m \times n$ matrix, let $b \in \mathbb{R}^m$ be a column vector, and let $c \in \mathbb{R}^n$ be a column vector. Let $K = \{x \in \mathbb{R}^n : Ax \le b\}$, where we think of x as a column vector, and where "$\le$" is element-wise. Assume $K \ne \emptyset$. Let $M = \max\{c^\top x : x \in K\}$ and assume $M < \infty$. Prove that the following system of inequalities in variables $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ is feasible:

$$Ax \le b, \quad A^\top y = c, \quad y \ge 0, \quad b^\top y \le c^\top x.$$

Furthermore, show that if (x, y) is any solution to the above system, then x is an optimal solution to the LP $\max\{c^\top x : x \in K\}$. You may appeal to Problem 1. (Remark: This shows that one can efficiently reduce the problem of optimizing a linear polynomial over linear inequalities to the problem of finding any solution to a system of linear inequalities.)

3. For the following problem you are given n "data points": each is a pair (X_i, y_i), where $X_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$. (For intuition, you might like to think of the special case $d = 1$.)

   (a) Suppose you wish to find the best-fitting hyperplane in "absolute error loss": this means you want to find the hyperplane $a \cdot x + b = 0$ (where $a \in \mathbb{R}^d$ and $b \in \mathbb{R}$) which minimizes $\sum_{i=1}^{n} |a \cdot X_i + b - y_i|$. Formulate the task as a linear program.

(b) Suppose that each y_i is either $+1$ or -1. We say $a \cdot x + b = 0$ is a *weakly separating hyperplane* for the data points if $a \cdot X_i + b \geq 0$ whenever $y_i = +1$ and $a \cdot X_i + b \leq 0$ whenever $y_i = -1$. Formulate an LP which is feasible if and only if there is a weakly separating hyperplane.

(c) Continuing the above, say that $a \cdot x + b = 0$ is a *separating hyperplane* if $a \cdot X_i + b > 0$ whenever $y_i = +1$ and $a \cdot X_i + b < 0$ whenever $y_i = -1$. Show that linear programming can be used to decide if there is a separating hyperplane.

(d) Continuing the above, suppose there is no separating hyperplane. Show that the hyperplane which minimizes "hinge loss" can be found using linear programming. Here the hinge loss of $a \cdot x + b$ on a point $(X_i, +1)$ is defined to be 0 if $a \cdot X_i + b \geq 1$ and $1 - (a \cdot X_i + b)$ otherwise; the hinge loss on a point $(X_i - 1)$ is defined to be 0 if $a \cdot X_i + b \leq -1$ and $(a \cdot X_i + b) - (-1)$ otherwise.

(e) For all of the above problems, what if instead of hyperplanes we are looking for "quadratic surfaces"? This means equations of the form

$$\sum_{1 \leq k \leq \ell \leq d} c_{k\ell} x_k x_\ell + \sum_{1 \leq k \leq d} a_k x_k + b = 0.$$

Explain how we can still solve all of the problems using polynomial-sized LPs.

4. Recall the LP relaxation for Minimum Vertex-Cover:

$$\begin{aligned} \min \quad & \sum_{v \in V} c_v x_v \\ \text{s.t.} \quad & 0 \leq x_v \leq 1 \quad \text{for all } v \in V, \\ & x_u + x_v \geq 1 \quad \text{for all } (u, v) \in E. \end{aligned}$$

(a) Let $\widetilde{x}$ be any feasible solution for the LP. Define another solution x^+ by

$$x_v^+ = \begin{cases} \widetilde{x}_v + \varepsilon & \text{if } \frac{1}{2} < \widetilde{x}_v < 1, \\ \widetilde{x}_v - \varepsilon & \text{if } 0 < \widetilde{x}_v < \frac{1}{2}, \\ \widetilde{x}_v & \text{if } \widetilde{x}_v \in \{0, \frac{1}{2}, 1\}. \end{cases}$$

Similarly define the solution x^-, replacing ε with $-\varepsilon$. Prove that one can find $\varepsilon > 0$ such that both x^+ and x^- are feasible for the LP. (Hint: there are at least four cases.)

(b) Show that every extreme point x^* of the LP is *half-integral*, i.e. $x_v^* \in \{0, \frac{1}{2}, 1\}$ for all $v \in V$.

5. Recall that a matrix $A \in \mathbb{R}^{m \times n}$ is said to be *totally unimodular* if every square submatrix of A has determinant either -1, 0, or $+1$. (In particular, every entry of A must be either -1, 0, or $+1$.)

(a) Let $b \in \mathbb{Z}^m$ be an integer vector and let $K = \{x \in \mathbb{R}^n : Ax \leq b\}$. Show that K is an "integral polytope", meaning that every "extreme point" of K has integer coordinates. (Recall that an extreme point is a point in K that cannot be written as the average of two distinct points in K. Hint: Cramer's rule.)

(b) Let G be a simple undirected graph with n vertices and m edges. Let $A \in \{0, 1\}^{m \times n}$ be its *incidence matrix*, in which $A_{i,j} = 1$ if and only if the ith edge of G touches the jth vertex. It is a fact that G is bipartite if and only if A is totally unimodular. Prove either the "if" or the "only if" part of this statement (your choice).

6. Consider the following algorithmic problem. The input consists of n "variables" $v_1, \ldots, v_n$ (to be assigned values 0 or 1) as well as a list of "constraints". Each constraint is either of the form "$v_i = 0$", "$v_i = 1$", or "$v_i \leq v_j$". It is easy to decide if there is an assignment to the variables satisfying all the constraints, but if the answer is "no", we'd still like to find an assignment that falsifies as few constraints as possible.

   Given an instance, let OPT denote the number of constraints falsified by the best assignment. Find an analyze a poly(n) time algorithm that outputs an assignment falsifying at most $C \cdot \mathrm{OPT}$ constraints, where C is a universal constant of your choice. Actually, your algorithm may be *randomized* and merely have the property that on any input, the *expected* number of constraints falsified by the algorithm's output assignment is at most $C \cdot \mathrm{OPT}$.

   (Hint: First formulate an ILP that exactly captures the task; you may need to use more than n variables. Next, relax to an LP. Finally find and analyze a "randomized rounding algorithm".)