New CS 473: Algorithms, Spring 2015
Homework 10

Due Tuesday, April 28, 2015 at 5pm

This is the last graded homework.

1. Given points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ in the plane, the ***linear regression problem*** asks for real numbers a and b such that the line $y = ax + b$ fits the points as closely as possible, according to some criterion. The most common fit criterion is the ***L_2 error***, defined as follows:

$$\varepsilon_2(a, b) = \sum_{i=1}^{n} (y_i - ax_i - b)^2.$$

(This is the error metric *(ordinary/linear) least squares*.)

But there are several other ways of measuring how well a line fits a set of points, some of which can be optimized via linear programming.

(a) The ***L_1 error*** (or *total absolute deviation*) of the line $y = ax + b$ is the sum of the vertical distances from the given points to the line:

$$\varepsilon_1(a, b) = \sum_{i=1}^{n} \left| y_i - ax_i - b \right|.$$

Describe a linear program whose solution (a, b) describes the line with minimum L_1 error.

(b) The ***L_∞ error*** (or *maximum absolute deviation*) of the line $y = ax + b$ is the maximum vertical distance from any given point to the line::

$$\varepsilon_\infty(a, b) = \max_{i=1}^{n} \left| y_i - ax_i - b \right|.$$

Describe a linear program whose solution (a, b) describes the line with minimum L_∞ error.

2. (a) Give a linear-programming formulation of the maximum-cardinality bipartite matching problem. The input is a bipartite graph $G = (L \cup R, E)$, where every edge connects a vertex in L ("on the left") with a vertex in R ("on the right"). The output is the largest matching in G. Your linear program should have one variable for each edge.

(b) Now dualize the linear program from part (a). What do the dual variables represent? What does the objective function represent? What problem is this!?

3. An ***integer program*** is a linear program with the additional constraint that the variables must take only integer values. Prove that deciding whether a given integer program has a feasible solution is NP-hard. *[Hint: Any NP-complete decision problem can be formulated as an integer program. Choose your favorite!]*

New CS 473 Spring 2015 — Homework 10 Problem 1

Name:	NetID:
Name:	NetID:
Name:	NetID:

(a) Describe a linear program whose solution describes the line with minimum L_1 error.

(b) Describe a linear program whose solution describes the line with minimum L_∞ error.

New CS 473 Spring 2015 — Homework 10 Problem 2

Name:	NetID:
Name:	NetID:
Name:	NetID:

(a) Give a linear-programming formulation of the maximum-cardinality bipartite matching problem.

(b) Now dualize the linear program from part (a). What do the dual variables represent? What does the objective function represent? What problem is this!?

New CS 473 Spring 2015 — Homework 10 Problem 3

Name:	NetID:
Name:	NetID:
Name:	NetID:

Prove that deciding whether a given integer program has a feasible solution is NP-hard.
