

## Introduction to Optimization

MS&E 111/ENGR 62, Autumn 2007-2008, Stanford University

Instructor: Ashish Goel

Handout 6: Homework 3. Given 10/15/07. Due 10/24/07 in class.

Collaboration policy: There are two Excel problems: 1 and 5. You must do problem 1 on your own. You can solve Problem 5 with a partner. If you choose to do so, both of you should turn in a copy of your Answer Reports and clearly indicate who you worked with. Additionally, on any problem you can discuss general strategies with other students in this class but cannot collaborate on the actual final answer. You cannot discuss the HW with anyone not in the class.

**Problem 1** Consider the problem of separation of private and public schools discussed in class.

a) Using the notation from class (e.g.  $U, V, \alpha, x, \delta^+, \delta^-$ ), set up an LP that will minimize the maximum violation (i.e. the max of the positive and negative violation) subject to the constraint that the sum of all the positive and negative violations is less than equal to some constant  $B$ .

b) Using the following data (same as from class):

Private Schools		
College	cost of living	tuition
MIT	116.3	34986
Yale	123.1	34530
Harvard	116.3	31456
Princeton	130	34202
Stanford	138.9	34800
Valparaiso	92.9	23200

Public Schools		
College	cost of living	tuition
Wisconsin Madison	94.4	21438
Washington-Seattle	104.6	22131
Berkeley	138.9	28004
Minnesota Twin cities	99.1	19580
Texas Austin	89.6	27468

Solve this LP in Excel for both  $B = 4$  and  $B = 6$  and attach your Answer Reports for both problems.

c) What trivial solution do you obtain if you solve this LP without the constraint that the sum of all errors must be less than  $B$ ? (Hint: you should not need to actually solve the LP.)

d) What problem would you solve to determine the minimum  $B$  value for which this problem has a solution?

e) For what range of values of  $B$  could we be confident that we would always get a solution, and the solution would not be the one you discussed in part c)?

**Problem 2** Consider a knapsack problem where a thief has a knapsack with a maximum volume  $U$  and weight limit  $W$ . The thief is in a warehouse with  $N$  goods. Each good  $i$  has a weight  $w_i$ , a volume  $u_i$ , and a value  $v_i$ . Additionally, we assume each good to be infinitely divisible, implying not only that we can take fractions of goods, but that we can pack the goods into our knapsack such that the entire space is filled.

The thief wants to maximize the value in the knapsack subject to the two capacity constraints.

- a) Set up an LP that will solve this problem. As always, be sure to clearly identify your decision variables.
- b) Assuming there is exactly one optimal solution, what is the maximum number of goods that the thief would take fractionally? Concisely explain why you know this to be true.
- c) Assume now that some subset  $P$  of the  $N$  goods are precious metals. In addition to the other constraints already mentioned, the thief would like to cap the value of precious metals in the knapsack at  $M$ . Again assuming there is exactly one optimal solution, what is the maximum number of goods that the thief would take fractionally now? Why?
- d) Assume now that the thief has decided to enlist some friends, so that there are now  $k$  thieves with identical knapsacks, each limited to taking volume up to  $U$ , weight up to  $W$ , and metals valued up to  $M$ . The thieves wish to maximize the aggregate value gathered by the entire set of thieves. Again assuming there is exactly one optimal solution, what is the maximum number of goods that will be taken fractionally by our set of thieves in aggregate (i.e. after you add up the totals taken by all thieves)?

**Problem 3** ARC Consultants needs to know how long it will take to complete a particular job. The job consists of  $n$  different tasks. Each task  $k$  takes  $t_k$  hours to complete (regardless of how many people work on it). Additionally, certain tasks cannot be performed until other tasks have been completed. This information is represented in a  $n \times n$  matrix  $P$  where the  $(i, j)$ th element of the matrix is a 1 if task  $j$  must be performed before task  $i$  and a 0 otherwise. (Note that all elements on the diagonal must be zero as a task cannot be performed before itself.) There is no limit on the number of tasks that can be performed at once.

Set up a linear program that determines how soon ARC can complete the job. As always, clearly specify your decision variables, objective function and constraints.

**Problem 4** a) (VRM 3.9.4) Is  $\mathbb{R}^N$  convex? Show that it is, or explain why not.

b) Give an example of a non-linear function and linear constraints such that the feasible region is bounded but there is no basic feasible solution that is optimum.

c) Prove that the region  $|x| \geq 1$  is not convex.

**Problem 5** Consider the following two sets of sample data,  $U$  and  $V$  where each row represents a point in  $\mathbb{R}^2$ :

$$U = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & .5 \end{bmatrix}.$$

$$V = \begin{bmatrix} 5 & .5 \\ 3 & 3 \\ 0 & .5 \\ 2 & 0 \end{bmatrix}.$$

a) Using the technique from class, in Excel find a linear separator in  $\mathbb{R}^2$  that minimizes the sum of the error terms. Attach your Answer Report.

b) Plot the points and the linear separator. You can do this by hand or with the program of your choice. Would you say that this is a good separator?

c) Create 4x4 matrices  $U'$  and  $V'$  where each row of  $U'$  is of the form  $U'_{i*} = [U_{i1}, U_{i2}, U_{i1}^2, U_{i2}^2]$ , and likewise for  $V'$ .

d) Using Excel find a linear separator that minimizes the sum of error terms for this data. Attach your Answer Report.

e) Can you determine from your answer to d) if this is a good separator? If so, how?

f) Recall that a non-rotated ellipse in  $Re^2$  is of the form  $a(x - x_0)^2 + b(y - y_0)^2 = c$ . Does your solution from d) correspond to an ellipse in the original 2 dimensional space of the problem? If so, write the equation that specifies the ellipse.