ORIE 5380, CS 5727: Optimization Methods
Homework Assignment 5

Due October 23, 12:00 pm

Please submit a single PDF document formatted to print and show all your work clearly. Feel free to scan and submit handwritten work. Do not spend too much time on wordprocessing your answers.

1. Consider the linear program

$$\begin{aligned} \max \quad & -2x_1 + 2x_2 - x_3 + 3x_4 \\ \text{st} \quad & -3x_1 + x_2 + 4x_3 + x_4 \leq 0 \\ & 3x_1 - x_2 - 3x_3 - 2x_4 \leq 3 \\ & x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

Use the simplex method to check whether this linear program is unbounded or has alternative optimal solutions. If the problem is unbounded then give a ray that proves the problem is unbounded. If the problem has alternative optimal solutions, then state two possible optimal solutions.

2. Consider the linear program

$$\begin{aligned} \max \quad & 5x_1 + 7x_2 - 12x_3 - 10x_4 \\ \text{st} \quad & 2x_1 - 2x_2 - 3x_3 - 2x_4 \leq 6 \\ & 2x_1 + 5x_2 - 4x_3 - 4x_4 \leq 3 \\ & x_1, x_2, x_3, x_4 \geq 0. \end{aligned}$$

Use the simplex method to check whether this linear program is unbounded or has alternative optimal solutions. If the problem is unbounded then give a ray that proves the problem is unbounded. If the problem has alternative optimal solutions, then state two possible optimal solutions.

3. In regression problems we have a set of data $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, and the goal is to fit a line of the form $y = ax + b$ for coefficients a and b that we must choose. (Here we assume that the x and y values are real-valued.) In least-squares regression, we choose the coefficients a and b so as to minimize the sum of squared errors $\sum_{i=1}^{n}(y_i - ax_i - b)^2$. Least-squares regression can give poor results when there are *outliers* in the data, where some of the x or y values are corrupted. A more robust approach to such corruption is to use L_1 regression. In L_1 regression, we instead minimize $\sum_{i=1}^{n}|y_i - ax_i - b|$. This objective function doesn't square the errors, so any large errors have a smaller effect on the objective than with least-squares regression.

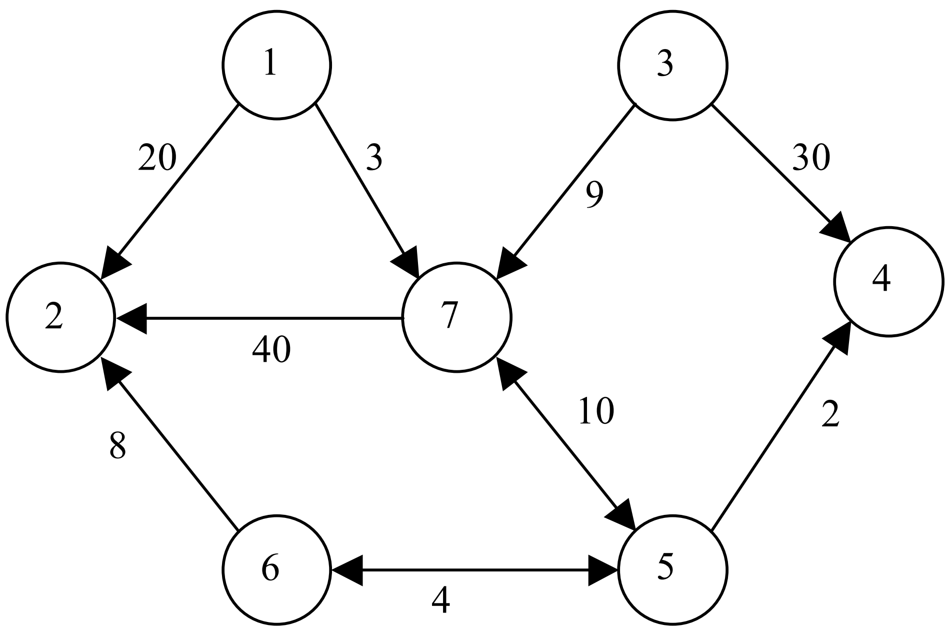
   (a) Formulate the L_1-regression problem as a linear program.

   (b) Use your L_1-regression formulation with the data in the attached Excel file (see `Modules/Homework Attachments` on Canvas) to identify the best choice of the coefficients a and b. This (made up) data gives the travel time y in minutes for trips of a given distance x in kilometers in a city. Report the estimated values of a and b.

4. The figure below represents an oil pipeline network. The different nodes represent pumping and/or receiving stations. The lengths in miles of the different segments of the network are shown on the respective arcs. The bi-directional segments allow flows in both directions. The supplies at stations 1 and 3 are respectively 80 and 70 barrels per day. The demands at the stations 2, 4 and 5 are respectively 80, 30 and 40 barrels per day. Assume that the transportation cost is proportional to the distance and we are interested in minimizing the total transportation cost.

   (a) Formulate the problem as a min-cost network flow problem (i.e. write the problem that consists of an objective function and constraints).

   (b) Find the minimum-cost transportation schedule by using optimization software. Report the optimal solution in an easily understood fashion, e.g., by drawing the flows on a copy of the network.



1. Consider the linear program

$$\begin{array}{ll} \max & -2x_1 + 2x_2 - x_3 + 3x_4 \\ \text{st} & -3x_1 + x_2 + 4x_3 + x_4 \le 0 \\ & 3x_1 - x_2 - 3x_3 - 2x_4 \le 3 \\ & x_1, x_2, x_3, x_4 \ge 0. \end{array}$$

Use the simplex method to check whether this linear program is unbounded or has alternative optimal solutions. If the problem is unbounded then give a ray that proves the problem is unbounded. If the problem has alternative optimal solutions, then state two possible optimal solutions.

$$\begin{array}{llllllll} \max & -2x_1 & +2x_2 & -x_3 & +3x_4 & & & = z \\ \text{s.t.} & -3x_1 & +x_2 & +4x_3 & +x_4 & +w_1 & & = 0 \\ & 3x_1 & -x_2 & -3x_3 & -2x_4 & & +w_2 & = 3 \end{array}$$

basis = $\{w_1, w_2\}$

non basis = $\{x_1, x_2, x_3, x_4\}$.

basic feasible solution = $(0,0,0,0,0,3)$

Entering Variable: x_4

leaving Variable: w_1

$$\begin{array}{llllllll} \max & 7x_1 & -x_2 & -13x_3 & & -3w_1 & & = z \\ \text{s.t.} & -3x_1 & +x_2 & +4x_3 & +x_4 & +w_1 & & = 0 \\ & -3x_1 & +x_2 & +5x_3 & & +2w_1 & +w_2 & = 3 \end{array}$$

basis = $\{x_4, w_2\}$

non basis = $\{x_1, x_2, x_3, w_1\}$.

basic feasible solution = $(0,0,0,0,0,3)$ $\quad z = 12$

Entering Variable: x_1

leaving Variable: None.

Consider decision variable x_1 with a negative coefficient in both 1st & 2nd constrains, those constrans doesn't impose any restrictions on how much we can increase the value of x_3. Thus the linear programming is unbounded.

$$-3x_1 + x_4 = 0$$

$$-3x_1 + w_2 = 3$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 0 \\ 3 \\ 0 \\ 3 \end{bmatrix}$$

$$-3(0+1t) \cancel{+(0+0t)} \cancel{+4(0+0t)} + (0+3t) \le 0$$

$$3(0+1t) \cancel{-(0+0t)} \cancel{-3(0+0t)} - 2(0+3t) \le 3$$

$$\begin{cases} -3t+3t \le 0 \\ 3t-6t \le 3 \end{cases} \Rightarrow t \ge -1$$

$$\begin{cases} x_1 = t \ge 0 \\ x_2 = 0 \\ x_3 = 0 \\ x_4 = 3t \ge 0 \end{cases} \Rightarrow t \ge 0$$

Therefor, solution is feasible for all value of $t \ge 0$.

objective $= -2t + 3*3t = 7t$. $\Rightarrow$ Extreme Ray.

3. In regression problems we have a set of data $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, and the goal is to fit a line of the form $y = ax + b$ for coefficients a and b that we must choose. (Here we assume that the x and y values are real-valued.) In least-squares regression, we choose the coefficients a and b so as to minimize the sum of squared errors $\sum_{i=1}^{n}(y_i - ax_i - b)^2$. Least-squares regression can give poor results when there are *outliers* in the data, where some of the x or y values are corrupted. A more robust approach to such corruption is to use L_1 regression. In L_1 regression, we instead minimize $\sum_{i=1}^{n}|y_i - ax_i - b|$. This objective function doesn't square the errors, so any large errors have a smaller effect on the objective than with least-squares regression.

   (a) Formulate the L_1-regression problem as a linear program.

   (b) Use your L_1-regression formulation with the data in the attached Excel file (see `Modules/Homework Attachments` on Canvas) to identify the best choice of the coefficients a and b. This (made up) data gives the travel time y in minutes for trips of a given distance x in kilometers in a city. Report the estimated values of a and b.

a). Let $z_i = |y_i - ax_i - b|$

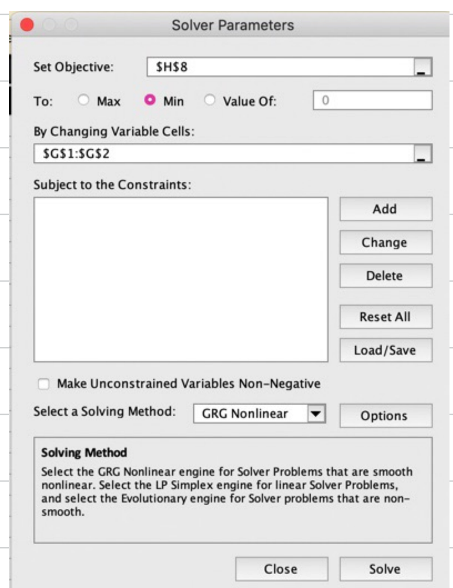
$$\text{minimize} \quad \sum_{i=1}^{n} z_i$$

$$\text{s.t.} \quad y_i - ax_i - b \le z_i$$

$$-y_i + ax_i + b \le z_i$$

b)

	A	B	C	D	E	F	G
1	Distance (km)	Time (min)	ax+b	zi		a=	2.174796116926
2	4.1	9.6	=$G$1*A2+$G$2	=ABS(B2-C2)		b=	0.647886096842388
3	4.9	14.5	=$G$1*A3+$G$2	=ABS(B3-C3)		sum_z =	=SUM(D2:D112)
4	4.7	12.9	=$G$1*A4+$G$2	=ABS(B4-C4)			
5	2.2	2.2	=$G$1*A5+$G$2	=ABS(B5-C5)			
6	5.5	10.4	=$G$1*A6+$G$2	=ABS(B6-C6)			
7	3.5	9	=$G$1*A7+$G$2	=ABS(B7-C7)			
8	4.2	7	=$G$1*A8+$G$2	=ABS(B8-C8)			
9	3.2	4.6	=$G$1*A9+$G$2	=ABS(B9-C9)			
10	5.4	12	=$G$1*A10+$G$2	=ABS(B10-C10)			
11	1.6	4.2	=$G$1*A11+$G$2	=ABS(B11-C11)			
12	2.5	4.3	=$G$1*A12+$G$2	=ABS(B12-C12)			
13	1.1	2.2	=$G$1*A13+$G$2	=ABS(B13-C13)			
14	4.1	12.7	=$G$1*A14+$G$2	=ABS(B14-C14)			
15	1.8	2.7	=$G$1*A15+$G$2	=ABS(B15-C15)			
16	3.1	9.9	=$G$1*A16+$G$2	=ABS(B16-C16)			
17	1.8	7.7	=$G$1*A17+$G$2	=ABS(B17-C17)			



We subjent to $z_i = |y_i - ax_i - b|$ and get the $\begin{cases} a = 2.175 \\ b = 0.648 \end{cases}$ as coeffient for the regression. $\Rightarrow y = 2.175x + 0.648$

& the minimul $\sum_{i=1}^{n} |y_i - ax_i - b| = 226.961$.