**CSCI 5654-Fall13**: Assignment #1 (75 points).
**Assigned date:** Friday 1/17/2020,
**Due date:** Tuesday, 01/28/2020 (before class)
**Instructions:** Please upload your HW as a PDF file along with code that you wrote for running the LP pro

---

**P1. (5 points)** Convert the following problem to the standard form:

$$\begin{array}{rll} \textbf{minimize} & 3x_1 - 5x_2 & \\ \textbf{s.t.} & 4x_1 + x_2 & \geq \; -4 \\ & 2x_1 - x_2 & \geq \; -8 \\ & x_1 + 2x_2 & \leq \; 4 \\ & x_1 & \geq \; 0 \end{array}$$

**Solution.**
First the easy transformations convert $\geq$ to $\leq$, min to max.

$$\begin{array}{rll} \textbf{maximize} & -3x_1 + 5x_2 & \\ \textbf{s.t.} & -4x_1 - x_2 & \leq \; 4 \\ & -2x_1 + x_2 & \leq \; 8 \\ & x_1 + 2x_2 & \leq \; 4 \\ & x_1 & \geq \; 0 \end{array}$$

We are missing $x_2 \geq 0$ constraint. Therefore, let us write $x_2 = y_2 - z_2$.

$$\begin{array}{rll} \textbf{maximize} & -3x_1 + 5y_2 - 5z_2 & \\ \textbf{s.t.} & -4x_1 - y_2 + z_2 & \leq \; 4 \\ & -2x_1 + y_2 - z_2 & \leq \; 8 \\ & x_1 + 2y_2 - 2z_2 & \leq \; 4 \\ & x_1, y_2, z_2 & \geq \; 0 \end{array}$$

**P2. (10 points)** (a) Add slack variables to the standard form that you obtained in (P1) above and write down the **initial dictionary**. (b) Perform one step of simplex on the initial dictionary. (c) Write down clearly which variable is entering and which variable is leaving. (d) Show the resulting dictionary after pivoting.

**Solution.** Let us write $x_2 := y_2 - z_2$.

$$\begin{array}{rll} \textbf{maximize} & -3x_1 + 5y_2 - 5z_2 & \\ \textbf{s.t.} & -4x_1 - y_2 + z_2 & \leq \; 4 \\ & -2x_1 + y_2 - z_2 & \leq \; 8 \\ & x_1 + 2y_2 - 2z_2 & \leq \; 4 \\ & x_1, y_2, z_2 & \geq \; 0 \end{array}$$

The slack form is

$$
\begin{array}{rlcl}
\textbf{maximize} & -3x_1 + 5y_2 - 5z_2 & & \\
\textbf{s.t.} & -4x_1 - y_2 + z_2 + \mathbf{x_3} & = & 4 \\
& -2x_1 + y_2 - z_2 + \mathbf{x_4} & = & 8 \\
& x_1 + 2y_2 - 2z_2 + \mathbf{x_5} & = & 4 \\
& x_1, y_2, z_2, x_3, x_4, x_5 & \geq & 0
\end{array}
$$

and the initial dictionary is

$$
\begin{array}{r|cccc}
x_3 & 4 & +4x_1 & +y_2 & -z_2 \\
x_4 & 8 & +2x_1 & -y_2 & +z_2 \\
x_5 & 4 & -x_1 & -2y_2 & +2z_2 \\
\hline
z & 0 & -3x_1 & +5y_2 & -5z_2
\end{array}
$$

$y_2$ enters and $x_5$ leaves.

$$
\begin{array}{r|cccc}
x_3 & 6 & +\frac{7}{2}x_1 & -\frac{1}{2}x_5 & -2z_2 \\
x_4 & 6 & +\frac{5}{2}x_1 & +\frac{1}{2}x_5 & \\
y_2 & 2 & -\frac{1}{2}x_1 & -\frac{1}{2}x_5 & +z_2 \\
\hline
z & 10 & -\frac{11}{2}x_1 & -\frac{5}{2}x_5 & +0z_2
\end{array}
$$

**P3.(20 points)** An investment adviser wishes to recommend an ideal investment for her client. After lots of market research, she has compiled a table of investment as expressed in the spreadsheet that has been distributed with this assignment. The total investment is \$10,000.

The data in Fig. 1 shows the investment options and categorizes them based on risk $A - D$, market segment and whether the investment is in an eco-friendly business. The client has specified minimum and maximum investment limits for each category. This data is also given as a CSV file `investmentData.csv` for your convenience.

Setup and solve a linear program that maximizes the expected profit while respecting the maximum and minimum investment percentages for each category.

Is the problem feasible? Is it bounded? What is the optimal solution? **Note:** Upload the code you used to solve the LP either as a GLPK model, matlab files or Python code. Also turn in the outputs that the solver gave you as a `problem3.txt` file.

**Solution.**

I have posted files: `invest.model` online as part of the solution and the output from GLPK in `invest.output`. They work with GLPK.

The investment data can be thought of as:

- Number of investments $n$.
- A vector of expected profits $\mathbf{pr}$.
- A vector of costs $\mathbf{c}$.
- A set of categories for risk, sector, eco friendliness and so on. Each category is represented by a vector of $0, 1$.

  For instance, think of the category Risk A. We can represent that by a vector:

$$r: \ [1;0;1;0;0;1;1;0;0;0;1\cdots]$$

where $r_j = 1$ means that the investment is in fact a risk A investment. $r_j = 0$ means it is'nt.

- Let $m$ be the number of categories and let $\mathbf{r_j}$ for $j \in [1, m]$ represent the 0-1 vector that identifies category $j$.
- Each category has an upper and lower bound on total investment allowed in that category.
- Let $\ell$, $\mathbf{u}$ be the upper and lower bound vectors as fractions of the total investment $T$.
- All investments sum up to the total $T$.

The LP can be formulated as:

$$\begin{array}{llcccl} \max & & & \mathbf{pr}^{\intercal}\mathbf{x} & & \\ \mathsf{s.t.} & \ell_j T & \leq & \mathbf{c}^{\intercal}\,\mathsf{diag}(r_j)\mathbf{x} & \leq & u_j T, \quad 1 \leq j \leq mT \\ & & & \sum_{j=1}^{n} x_j & = & T \\ & & & \mathbf{x} & \geq & 0 \end{array}$$

Note that $\mathsf{diag}(\mathbf{v})$ for a $n \times 1$ vector creates a $n \times n$ matrix with all off diagonal entries as 0 and the diagonal entries corresponding to $\mathbf{v}$.

```
bash-3.2$ glpsol --math invest.model
GLPSOL: GLPK LP/MIP Solver, v4.60
Parameter(s) specified in the command line:
 --math invest.model
Reading model section from invest.model...
Reading data section from invest.model...
155 lines were read
Generating obj...
Generating totInvest...
Generating rBnds...
Generating sBnds...
Generating eBnds...
Model has been successfully generated
GLPK Simplex Optimizer, v4.60
11 rows, 15 columns, 70 non-zeros
Preprocessing...
10 rows, 15 columns, 55 non-zeros
Scaling...
 A: min|aij| =  1.961e+00  max|aij| =  1.214e+01  ratio =  6.189e+00
GM: min|aij| =  1.000e+00  max|aij| =  1.000e+00  ratio =  1.000e+00
EQ: min|aij| =  1.000e+00  max|aij| =  1.000e+00  ratio =  1.000e+00
Constructing initial basis...
Size of triangular part is 10
      0: obj =   5.353052649e+03 inf =   2.150e+04 (6)
      6: obj =   5.938527931e+03 inf =   0.000e+00 (0)
*    11: obj =   8.513474798e+03 inf =   0.000e+00 (0)
OPTIMAL LP SOLUTION FOUND
Time used:   0.0 secs
Memory used: 0.1 Mb (137848 bytes)
Optimal Value: 8513.474798
Display statement at line 30
x[1].val = 0
x[2].val = 0
x[3].val = 467.144191840548
```

```
x[4].val = 0
x[5].val = 0
x[6].val = 0
x[7].val = 0
x[8].val = 0
x[9].val = 0
x[10].val = 147.637795275591
x[11].val = 0
x[12].val = 53.763440860215
x[13].val = 0
x[14].val = 459.685575066655
x[15].val = 0
-----
Optimizer says to invest 467.144192 units of investment 3
Optimizer says to invest 147.637795 units of investment 10
Optimizer says to invest 53.763441 units of investment 12
Optimizer says to invest 459.685575 units of investment 14
Bon appetit!
 -----
Model has been successfully processed
```

**P4 (10 points)** A set $S$ is *convex* if any two points $\mathbf{x}_1, \mathbf{x}_2 \in S$, the point $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in S$ for all $\lambda \in [0, 1]$.

Consider the linear program:

$$\max\ \mathbf{c}^t\mathbf{x} \text{ s.t. } A\mathbf{x} \leq \mathbf{b}\,.$$

(A) Prove that the set of feasible solutions $F : \{\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}\}$ is a convex set. **Note:** The proof is just 5 lines at most.

(B) Let $\mathbf{x}_1 \neq \mathbf{x}_2$ be two distinct optimal solutions for the problem both achieving the optimal objective value $z^* = \mathbf{c}^t\mathbf{x_1} = \mathbf{c}^t\mathbf{x_2}$. Show that there are infinitely many optimal solutions for the linear program. **Note:** The proof requires 3-5 sentences at most.

**Solution.**
**(A)** Let $\mathbf{x}_1, \mathbf{x}_2 \in F$. We have that $A\mathbf{x}_1 \leq \mathbf{b}_1$ and $A\mathbf{x}_2 \leq \mathbf{b}_2$. We wish to prove that $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in F$.

$$A(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) = \lambda A\mathbf{x}_1 + (1-\lambda)A\mathbf{x}_2$$

Since $\lambda \geq 0$, we have $\lambda A\mathbf{x}_1 \leq \lambda\mathbf{b}_1$. Likewise, since $(1-\lambda) \geq 0$, we have $(1-\lambda)A\mathbf{x}_2 \leq (1-\lambda)\mathbf{b}_2$. Combining,

$$\begin{aligned} A(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) &= \lambda A\mathbf{x}_1 + (1-\lambda)A\mathbf{x}_2 \\ &\leq \lambda\mathbf{b}_1 + (1-\lambda)\mathbf{b}_2 \\ &\leq \mathbf{b} \end{aligned}$$

This immediately shows that $\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in F$.

(B) Let $\mathbf{x}_1, \mathbf{x}_2$ be such that $\mathbf{c}^t\mathbf{x}_1 = \mathbf{c}^t\mathbf{x}_2 = z^*$. For any choice of $\lambda \in [0,1]$, we obtain that $\mathbf{x}(\lambda) = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ satisfies the constraints $A\mathbf{x} \leq \mathbf{b}$ and furthermore,

$$\begin{aligned}\mathbf{c}^t\mathbf{x}(\lambda) &= \mathbf{c}^t(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2)\\ &= \lambda\mathbf{c}^t\mathbf{x}_1 + (1-\lambda)\mathbf{c}^t\mathbf{x}_2\\ &= \lambda z^* + (1-\lambda)z^*\\ &= z^*\end{aligned}$$

Thus, if $\mathbf{x}_1 \neq \mathbf{x}_2$ then we conclude the existence of infinitely many solutions.

**P5 (15 points)** This problem explores how to formulate constraints in linear programming. Let $\mathbf{x} : (x_1, \ldots, x_n)$ be $n$ decision variables such that $x_1, \ldots, x_n \geq 0$. We wish to formulate the following constraints. Which of them can be expressed as constraints in a linear programming problem? If yes, explain how you would formulate them. If no, give a short proof. Note that linear programming problems allow constraints of the form $\sum a_j x_j \leq b$ for real numbers $a_j$, for $j = 1, \ldots, n$ and $b$.

(a) (2 points) $x_5$ must be between $x_1$ and $x_3$.

(b) (3 points) $x_1$ must be between 40% and 60% of the sum of all $x_1, \ldots, x_n$.

(c) (10 points) At least one of the variables $x_1, \ldots, x_n$ must be 0. (**Hint:** Is this a "convex" constraint?).

**Solution.**
(a) $x_1 \leq x_5$ and $x_5 \leq x_3$.
(b) $x_1 \leq 0.6\sum_{j=1}^n x_j$ and $x_1 \geq 0.4\sum_{j=1}^n x_j$.
(c) This is not possible to express as a system of linear inequalities since it is a non-convex set. For instance take $n = 3$ and consider $(1,0,1)$ and $(0,1,1)$ both of which satisfy the constraint. However, their mid point $(0.5, 0.5, 1)$ does not.

**P6 (15 points, HARDER)** Let $x_1, x_2$ be two real-valued variables which satisfy the constraints $0 \leq x_1 \leq 10$ and $0 \leq x_2 \leq 15$. Using indicator variables $w_1, w_2 \in \{0,1\}$ write down a system of linear inequalities involving $(x_1, x_2, w_1, w_2)$ that express the nonlinear constraint $x_1x_2 = 0$. (**Hint:** Suppose we want to make sure that $w_1 = 1$ indicates that $x_1 = 0$ but $w_1 = 0$ places no constraints on $x_1$. How do we achieve this? Just write down the constraints involving $x_1, x_2, w_1, w_2$ with at most 4 lines of explanations).

**Solution.**

$$0 \leq x_1 \leq 10w_1,\ 0 \leq x_2 \leq 15w_2,\ w_1 + w_2 \geq 1,\ w_1, w_2 \in \{0,1\}\,.$$

If $w_1 = 1$ then note that the first constraint reduces to $x_1 = 0$, and likewise for $w_2$ and $x_2$. The constraint $w_1 + w_2 \geq 1$ now expresses that at least one of $x_1$ or $x_2$ must be 0.

| ID | Expected Profit/ Unit | Price / Unit | Risk Category | Investment Market | Eco Friendly ? |
|---|---|---|---|---|---|
| | | | | | |
| 1 | 1.451 | 2.563 | A | Tech | Y |
| 2 | 2.683 | 4.307 | B | Finance | Y |
| 3 | 5.898 | 6.422 | C | Finance | Y |
| 4 | 2.102 | 3.488 | A | PetroChem | N |
| 5 | 5.709 | 6.581 | B | Finance | N |
| 6 | 4.519 | 8.993 | D | Finance | Y |
| 7 | 7.176 | 11.481 | C | Finance | Y |
| 8 | 6.075 | 11.730 | B | Tech | Y |
| 9 | 5.718 | 9.270 | B | PetroChem | Y |
| 10 | 7.442 | 10.160 | A | Automobile | Y |
| 11 | 1.234 | 1.961 | C | Tech | Y |
| 12 | 4.680 | 9.300 | D | Automobile | N |
| 13 | 7.229 | 11.672 | A | PetroChem | Y |
| 14 | 9.589 | 10.877 | B | Automobile | N |
| 15 | 6.497 | 12.137 | C | Finance | N |
| | | | | | |

| Risk Categories | Min | Max |
|---|---|---|
| A | 1500 | 3500 |
| B | 4500 | 6500 |
| C | 1000 | 3000 |
| D | 500 | 2500 |

| Investment Market | Min | Max |
|---|---|---|
| Tech | 0 | 3000 |
| Finance | 0 | 4000 |
| PetroChem | 0 | 5000 |
| Automobile | 0 | 7000 |

| EcoFriendly | Min | Max |
|---|---|---|
| Y | 2000 | 10000 |

Figure 1: Investment Data.