DIIO06 – Optimización lineal avanzada – Homework #1

1. **(The general diet problem – 25 pts.)** There are n possible foods and m nutrients. The minimum recommended amount of nutrient j is N_j ($j = 1,\dots,m$). The price per serving of food i is c_i ($i = 1,\dots,n$). The amount of nutrient j that food i provides is a_{ij} ($i = 1,\dots,n$, $j = 1,\dots,m$). In addition, there is a servings-per-day limit of L_i for each food $i = 1,\dots,n$.

   (a) Using variables x_i representing the number of servings of food i to be purchased each day ($i = 1,\dots,n$) formulate a linear program that helps you to decide the daily menu that minimizes the total cost while satisfying the minimum amount of nutrients and the servings-per-day limit. How many variables and constraints does your model have?

   (b) Put the LP obtained in part (a) in the `Simplex Standard form`. What is the meaning of the auxiliary variables $s_1,\dots,s_m$ (associated to the minimum nutrient constraints) and $w_1,\dots,w_n$ (associated to the servings-per-day limit constraints)?

   (c) Create an `XLSX` file containing the data of the *diet problem* that Polly is solving. Recall that the table with the data is written in the corresponding slides of Lecture 1, so you can copy and paste part of the data from there.

   (d) Implement in `Julia`/`JuMP` the model for the *diet problem*. **All** the data of the problem **must be read** from the `XLSX` file from part 1. Solve Polly's *diet problem*. What is the optimal solution? The optimal value?

   (e) Update the cost data in the `XLS` with what you believe are the current prices (in pesos) of the foods Polly is considering. Add to the `XLS`, with all the corresponding data needed to solve the *diet problem*, the following: (i.) The favorite foods of each member of the group. (i.) Two more nutrients. Solve the *diet problem* with the new dataset you obtained from (i.) and (ii.).

   (f) How would your model in part (a) changes if instead of minimizing the total cost you want to minimize the total servings-per-day in the daily menu? Solve it and compare the solution you obtain with the one in part (a).

   (g) How would you add to your model in part (a) the additional constraint that the total number of **foods** purchased should be at most B, with $B < n$? Give an appropriate value to B, solve the model and compare the solution you obtain with the one in part (a).

   (h) Change your model in part (a) to consider the following additional constraint: "you cannot purchase both *Cherry pie* and *Pork with beans*". Solve the model and verify that the condition is satisfied. **Hint:** Use an auxiliary binary variable

   $$z = \begin{cases} 1, & \text{Cherry pie is purchased} \\ 0, & \text{else} \end{cases}$$

   and add two constraints, one relating this new variable to x_5, L_5 and one relating it to x_6, L_6 in a way that if $z = 1$, then *Pork with beans* cannot be purchased and if $z = 0$ *Cherry pie* cannot be purchased.

   (i) The model in part (a.) requires that all m minimum nutrient requirements are satisfied, but perhaps this is asking for too much. Change the model so the requirement is that, *among all m nutrients, at least T minimum nutrient requirements are satisfied* where $T < m$. **Hint:** use the following binary variables for each $j = 1,\dots,m$:

   $$z_j = \begin{cases} 1, & \text{Requirement for nutrient } j \text{ is satisfied} \\ 0, & \text{else} \end{cases}$$

add these variables to the nutrient requirement constraints in a way that the jth nutrient requirement constraint must be satisfied only when $z_j = 1$, and add the constraint $\sum_{j=1}^{m} z_j \geq T$.

2. **(The diet problem revisited – 10 pts.)**

   Recall the diet problem: There are n possible foods and m nutrients. The minimum recommended amount of nutrient j is N_j ($j = 1,\ldots,m$). The price per serving of food i is c_i ($i = 1,\ldots,n$). The amount of nutrient j that food i provides is a_{ij} ($i = 1,\ldots,n$, $j = 1,\ldots,m$). In addition, there is a servings-per-day limit of L_i for each food $i = 1,\ldots,n$.

   (a) Consider the following situation: If you purchase more than $L_i/2$ of food i, then you need to pay a cost $c_i' > c_i$ for any amount in excess of $L_i/2$ (this could be the case if, for instance, the cheaper place only have an inventory of $L_i/2$ for food i and you need to buy the extra food in a more expensive place). Model the diet problem with this consideration in mind as an LP (use continuous variables only!). **Hint:** Use variables y_i that represent the amount of food i that you should purchase at price c_i'.

   (b) Consider the following situation: If you purchase (strictly) more than $L_i/2$ of food i, then there is a bulk discount and the unit price of food i becomes $c_i' < c_i$ (every unit can be purchased at this cheaper price). Model the diet problem with this consideration in mind as a MILP. **Hint:** Use binary variables z_i that tell whether the amount of food i that you purchase is (strictly) more than $L_i/2$ or not.

   (c) Use `Julia`/`JuMP`, Polly's diet problem data and appropriate c_i' values to verify that both your models in parts (a) and (b) are correct. You need to create, implement and solve two models: one with $c_i' > c_i$ values and other with $c_i' < c_i$ values.

3. **(Production planning problem revisited – 25 pts.)**

   A factory operates for T periods. For each period $t = 1,\ldots,T$ there is a demand d_t that has to be satisfied, a cost of production of a unit of p_t, and a setup cost of f_t. Demand d_t does not need to be satisfied by production in period t, units can be produced in previous periods and stored to be used later. The quantity s_0 denotes the initial inventory. The holding cost of storing a unit from period t to period $t+1$ is given by h_t. The goal is to obtain a production plan for the factory that minimizes the total cost.

   In Lecture 3 we modeled this problem by defining the following variables for each $t = 1,\ldots,T$:

   - x_t: production in period t.
   - s_t: end storage in period t.
   - w_t: on/off variable for production in period t (1 if there is a positive production in period t, 0 otherwise).

   The mixed-integer optimization model is:

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} p_t x_t + \sum_{t=1}^{T} h_t s_t + \sum_{t=1}^{T} f_t w_t \\
\text{s.t.} \quad & s_{t-1} + x_t = d_t + s_t, \quad t = 2,\ldots,T \\
& x_1 = d_1 + s_1 \\
& x_t \leq \left(\sum_{s=1}^{T} d_s \right) w_t, \quad t = 1,\ldots,T \\
& x_t, s_t \geq 0, w_t \in \{0,1\}, \quad t = 1,\ldots,T.
\end{aligned}
$$

(a) i. Generate (random) data for $T = 100$ periods (you will use the same data in part c)).
ii. Solve the problem and report the optimal value.
iii. Solve the continuous relaxation of the problem (that is, change the binary variables to be continuous variables with bounds between 0 and 1). Report the optimal value.
iv. What can you say about the values obtained in part ii. and iii.?

(b) The constraints

$$x_t \le \left(\sum_{s=1}^{T} d_s \right) w_t, \quad t = 1, \ldots, T$$

that relate x_t and w_t are so-called *Big M* constraints. There is a constant $M_t = \sum_{s=1}^{T} d_s$ in the right hand side with the purpose of making the constraint redundant if the variable w_t is equal to 1; in this particular case the constant is the same for all these T constraints, but in general it may depend on the constraint. The value of this *big-M* constant can have an impact on performance when optimizing the problem with a MIP solver.

i. Give a better constant M_t for this formulation (the subindex t indicates that this *big-M* constant depends on the index $t = 1, \ldots, T$ of the constraint). This better constant M_t should be smaller than M but in a way that still makes the constraint redundant when $w_t = 1$.
ii. Using `Julia`/JuMP, determine how does the optimal value of the continuous relaxation changes with the choice of the *big-M* constant: compare the value obtained in part a) with the value obtained using the better constants M_t that you found.

(c) In the following questions you will be asked to modify the above LP in order to be able to model **as an LP** some variations of the standard problem (the questions are independent of each other). You may need to change the objective function, add more variables/constraints or change some of the original constraints. **Do not write the complete model again**, you only need to indicate what changes or what should be added.

i. For each period $t = 1, \ldots, T$ the absolute value of the difference between the inventory in that period and the previous period cannot be more than 2 units.
ii. For each period $t = 1, \ldots, T$ the actual values of p_t and h_t are not known with certainty but there are n possible scenarios for their values: p_t^1 and h_t^1 or p_t^2 and h_t^2 or ... or p_t^n and h_t^n. For this reason you want to replace the original objective function by the following *worse case cost* (WCC) nonlinear function:

$$\text{WCC} = \max \left\{ \sum_{t=1}^{T} p_t^i x_t + \sum_{t=1}^{T} h_t^i s_t : i = 1, \ldots, n \right\}.$$

Linearize the new objective function using the technique studied in class.
iii. If the company produces in more than $T/2$ periods, then it has to pay a penalty of \$1000.

(d) i. Using only the following decision variables (for $t = 1, \ldots, T$ and $u \ge t$):
- q_{tu}: production in period t used to satisfy demand in period $u \ge t$.
- w_t: on/off variable for production in period t (1 if there is a positive production in period t, 0 otherwise).

Formulate and solve the production planning problem for the same data you used in part a).
ii. Solve the continuous relaxation of the model in the previous part and report the optimal value. What can you say about the value of the w_t variables?

Instructions.

- All files should be submitted through the link available in Webcursos. You can write your solution in a single `Jupyter` notebook or a combination of a single PDF file and a `Julia` script (so **at most** two files in your submission).
- On every assignment **you must list the names** of all the students you discussed a particular problem and **you must cite all the references** you used in order to achieve your solution (papers, websites, AIs, etc.). Failing to report cooperation or to cite the corresponding sources is considered cheating/plagiarism.
- The homework is due Sunday, April 16 at 11:59pm.