15.072: The Analytics Edge **Fall 2021**
Problem Set 3: Large-scale linear optimization

Out: November 08; Due: November 23, 11:59 pm.

Guidelines:

- *When asked to "formulate an optimization model", no computation is required. It involves writing the model mathematically. This is a required step toward implementing the model and getting a solution.*
- *For each optimization problem, please select a relative termination tolerance of 0.0001, i.e., any algorithm should not terminate until it finds a solution within 0.01% of the true optimal solution. When building your model, you can set the tolerance as follows, in JuMP v.0.21.5.*

```
mod = JuMP.Model(JuMP.optimizer_with_attributes(() -> Gurobi.Optimizer(),
    "MIPGap" => 0.0001))
```

Optimizing Waste Management

You are in charge of a public authority in a large region, where municipal waste management has grown to be a significant issue. In the past, the authority has focused on waste collection by acquiring new trucks and optimizing collection routes. However, municipal solid waste is still disposed in open-air dumps, which burn waste from time to time. This is widely recognized to be the worst possible disposal technique!

Part A. Building landfills

Your task is to build sanitary landfills for waste disposal. Waste is generated in 50 metropolitan centers, and you want to locate the landfills strategically to minimize the costs of transporting waste from the centers to the landfills. You have a budget equivalent to 5 landfills in the region.

Your region can be represented as a square of 100 miles over 100 miles. You are given the following data:

- **`centers.csv`**: indicates the location of the 50 centers. The first column represents the coordinate along the x axis (in miles, between 0 and 100), and the second column represents the coordinate along the y axis (in miles, between 0 and 100).
- **`landfills.csv`**: indicates 15 sites for potential sanitary landfills. Again, the first column represents the coordinate along the x axis (in miles, between 0 and 100), and the second column represents the coordinate along the y axis (in miles, between 0 and 100).
- **`q.csv`**: indicates the amount of waste (in tons) to collect every day from each of the 50 centers.

You define the following optimization problem:

Sets

$$\begin{aligned} i &:= 1, \cdots, 50: && \text{centers} \\ j &= 1, \cdots, 15: && \text{potential landfills} \end{aligned}$$

Parameters

$$\begin{aligned} q_i &: && \text{amount of waste generated at center } i \\ d_{ij} &: && \text{(Euclidean) distance between center } i \text{ and potential landfill } j \end{aligned}$$

Hint: Recall that the Euclidean distance is given as follows:

$$d_{ij} = \sqrt{(\text{x-axis of center } i \; - \; \text{x-axis of landfill } j)^2 + (\text{y-axis of center } i \; - \; \text{y-axis of landfill } j)^2}.$$

Decision variables

$$x_{ij}: \quad \text{amount of waste transported from center } i \text{ to landfill } j$$

$$z_j = \begin{cases} 1 & \text{if landfill } j \text{ is build} \\ 0 & \text{otherwise} \end{cases}$$

Formulation

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{50}\sum_{j=1}^{15} d_{ij}x_{ij} && (1)\\
\text{subject to} \quad & \sum_{j=1}^{15} x_{ij} = q_i, \quad \forall i = 1,\cdots,50 && (2)\\
& x_{ij} \leq 1{,}000{,}000 \times z_j, \quad \forall i = 1,\cdots,50,\ \forall j = 1,\cdots,15 && (3)\\
& \sum_{j=1}^{15} z_j \leq 5 && (4)\\
& x_{ij} \geq 0, \quad \forall i = 1,\cdots,50,\ \forall j = 1,\cdots,15 && (5)\\
& z_j \in \{0,1\}, \quad \forall j = 1,\cdots,15 && (6)
\end{aligned}$$

a. Justify the objective function and each constraint, and highlight any assumption. **[10 pts]**

b. Implement the model in Julia. Where do you build landfills (please show it graphically), and what is the total distance traveled by the waste? **[10 pts]**

Part B. Building landfills and transfer stations

You are provided extra budget, but given environmental concerns, you cannot build more than 5 landfills. Instead, you can build transfer stations for waste compaction. Compaction essentially reduces the size of solid waste, making it less expensive to transport. The transport cost of uncompacted waste is estimated at \$1 per mile per ton, and that of compacted waste is estimated at \$0.5 per mile per ton. The cost of building and operating each transfer station is equivalent to \$10,000 per day (this is essentially a fixed cost; there is no variable cost of waste compacting). Each transfer station can treat up to 2,000 tons of waste per day.

You want to determine where to build landfills for waste disposal and where to build transfer stations for waste compaction. Collected waste can be transported directly to a landfill as uncompacted waste; alternatively, it can first go to a transfer station as uncompacted waste and then to a landfill as compacted waste. You are given the following data (in addition to those you already have):

- `stations.csv`: indicates 50 sites for potential transfer stations. Again, the first column represents the coordinate along the x axis (in miles, between 0 and 100), and the second column represents the coordinate along the y axis (in miles, between 0 and 100).

c. Write a linear optimization model that determines the 5 landfills and the transfer stations to build. Make sure to define your sets, parameters and decision variables carefully. Your objective is to minimize total costs. You should still ensure that all the generated waste is collected and disposed at a landfill, and to capture the various elements outlined above. **[20 pts]**

d. Implement the model in Julia. How many transfer stations do you build? What is the fraction of waste that gets compacted? Compare the total distance traveled, the transportation costs, and the total daily costs to your earlier solution. **[20 pts]**

Part C. A second region

Your success did not go unnoticed. You have now been asked to apply your latest model (Question e.) to a neighboring region, also of 100 miles per 100 miles. It is located on the east side of your original region (with an x axis spanning 100 to 200 miles, and a y axis spanning 0 to 100 miles), and has 40 centers. It is also planning to build 5 landfills and it is considering 15 landfill sites and 50 transfer station sites. You obtain the equivalent data in the following files: **`centers2.csv`**, **`stations2.csv`**, **`landfills2.csv`**, and **`q2.csv`**.

For simplicity, let us refer to the first region as "Region A" and to this new region as "Region B".

e. Implement your model in Region B. Where do you build landfills (please show it graphically)? What is the total daily cost? **[10 pts]**

When your leadership (of Region A) hears about this interest, you feel a sense of excitement: "What if we planned the municipal waste management system together across the two regions? There has been a lot of talk to achieve greater collaboration between us, and we would love to showcase this as an example of our partnership. Plus, we might be able to find efficiencies if we worked together. What do you think?"

f. Implement the model for the two regions combined. Which landfills do you build, and what is the total daily cost? **[10 pts]**

g. Write a short memo in response to your leadership, providing a clear recommendation, quantitative evidence to support it, the intuition behind it, and practical concerns. **[20 pts]**