15.093: Optimization **Fall 2023**

Assignment 2: Linear optimization and network optimization

Assigned: September 21; Due: October 02.

Exercise 3.20 from *Introduction to Linear Optimization*, Bertsimas & Tsisiklis [30 pts]

Exercise 7.23 from *Introduction to Linear Optimization*, Bertsimas & Tsisiklis [30 pts]

Cloud computing [40 pts]

A cloud computing provider needs to process 1,000 jobs, indexed by $i = 1, \cdots, 1,000$ with 20 machines, indexed by $j = 1, \cdots, 20$. We denote by c_{ij} the energy consumption resulting from performing job i on machine j. The objective of the problem is to find an assignment that minimizes the total energy consumption, while ensuring that each job gets fully processed.

a. Assume that each machine j has finite capacity, and can process up to C_j jobs. Formulate the problem as a linear optimization model. Is it a network flow? If so, propose a network representation, including the demand and supply at each node, the cost of each arc, and the capacity on each arc. If not, explain why not. Is the solution guaranteed to be integral?

b. You have access to the following data files:
   - energy.csv: A matrix of size 1,000×20 that indicates the energy consumption c_{ij} for each job i (rows) and each machine j (columns)
   - capacity.csv: A vector of size 20 that indicates the capacity C_j of machine j.

   Implement the model computationally, and report (i) the optimal energy consumption; (ii) the energy consumption that would be achieved if each job could be assigned to any machine without any consideration for machine capacities; and (iii) the number of jobs that are not (fully) assigned to the machine with the lowest energy consumption.

c. Actually, all jobs are not created equal. Accordingly, we replace the machine capacity constraint by a machine utilization constraint. Let r_{ij} be the utilization of machine j when performing job i, and let U_j be the maximal utilization of machine j. Formulate the problem as a linear optimization model. Is it a network flow? If so, propose a network representation, including the demand and supply at each node, the cost of each arc, and the capacity on each arc. If not, explain why not. Is the solution guaranteed to be integral?

d. You have access to the following data files:
   - utilization.csv: A matrix of size 1,000×20 that indicates the utilization r_{ij} for each job i (rows) and each machine j (columns)
   - maxutil.csv: A vector of size 20 that indicates the maximal utilization U_j of machine j.

   Implement the model computationally, and report (i) the optimal energy consumption; (ii) the energy consumption that would be achieved if each job could be assigned to any machine without any consideration for machine utilization; and (iii) the number of jobs that are not (fully) assigned to the machine with the lowest energy consumption.