MATH 11247: Optimization under Uncertainty
Assignment #2

Due Date: February 28, 4pm

Instructions

- This is an individual assignment.
- You **must** cite **any** references (texts, papers, websites, etc.) you have used to help you solve these problems.
- Submit your written/typed solutions via Gradescope **(pdf file only)**.
- For the implementation questions, namely Question 1 part (c) and Question 3, please add a screenshot of your code's output (i.e., what is printed to the screen) to your written answer, and additionally submit your codes as separate files.
- For any clarifications on the questions, feel free to post in Piazza.

---

1 ScotRail has bought a new train for its Edinburgh-Glasgow route and is trying to design its interior by deciding on how to partition the available space into economy, business, and first-class sections. The maximum capacity of the train is 200 passengers if all seats are allocated to the economy class. A first-class seat requires twice as much space as an economy seat, while a business-class seat requires 1.5 times the space of an economy seat. The profit from business-class and first-class tickets is, respectively, twice and thrice the revenue of an economy-class ticket. The sections will be fixed after partitioning and cannot be changed.

ScotRail can face (only) the four demand scenarios provided in Table 1. The strict policy of the company is not to have overselling in any section.

Scenario	Probability	First Class	Business Class	Economy Class
1	0.4	25	60	200
2	0.3	20	40	180
3	0.2	10	25	175
4	0.1	5	10	150

Table 1: Demand scenarios for ScotRail

(a) [15pts] Using the scenarios in Table 1, create a stochastic program that will help ScotRail in partitioning the train in order to maximize expected profit.

(b) [10pts] Explain whether your model has complete recourse and/or relatively complete recourse.

(c) [15pts] Solve your model using the modeling environment of your choice and provide an optimal solution. Determine the value of the stochastic solution (VSS) and the expected value of the perfect information (EVPI).

2 A company is planning to assign workers to tasks in an optimal way. The assignment must be made such that each worker is assigned exactly one task, and each task is assigned exactly one worker. There are a total of n workers and n tasks. The company models this problem as a **minimum-cost perfect matching** problem on a bipartite graph.

The problem is represented as an undirected weighted bipartite graph $G = (V, E)$. The node set V consists of two partitions: V_1 (workers) and V_2 (tasks), with $|V_1| = |V_2| = n$. The edge set E contains possible worker-task assignments; note that this is usually skill-based, thus it is usually the case that $E \subset \{\{v_1, v_2\} : v_1 \in V_1, v_2 \in V_2\}$. Each edge $e = \{i, j\} \in E$ represents assigning worker $i \in V_1$ to task $j \in V_2$ and has an associated cost c_e, which represents the cost of assigning worker i to task j. The goal is to find a perfect matching in the bipartite graph that minimizes the total assignment cost. An example is given in the below figure.

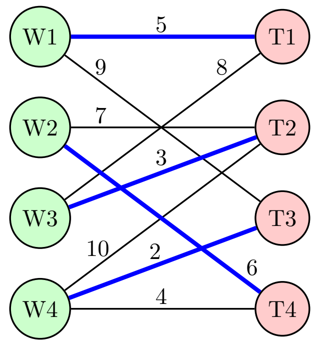

Figure 1: An example weigthed bipartite graph and its minimum weight perfect matching (highlighted by blue edges)

Defining the binary decision variables

$$x_e = \begin{cases} 1, & \text{if edge } e \text{ is included in the matching,} \\ 0, & \text{otherwise,} \end{cases}$$

we can formulate this problem as the following integer program:

$$\begin{aligned} \min \quad & \sum_{e \in E} c_e x_e \\ \text{s.t.} \quad & \sum_{e \in \delta(v)} x_e = 1 \quad \forall v \in V \quad \text{(each worker and task must be matched exactly once)} \\ & x_e \in \{0, 1\} \qquad \forall e \in E \quad \text{(binary matching decisions)} \end{aligned}$$

where $\delta(v)$ denotes the set of edges incident to node v.

Now, consider the following extension to a two-stage decision-making framework:

- In the first stage, we must choose a **matching** with only, exactly, B edges (where $B < n$), incurring the associated costs of c_e for $e \in E$ for the selected edges.
- In the second stage, we must complete the first-stage matching into a **perfect matching**, where additionally selected edges incur cost of d_e^s for $e \in E$ under scenario s. The uncertain second-stage costs are modeled using a set of scenarios $s \in S$, each occurring with probability p_s.

- The objective is to minimize the sum of the total cost of the edges selected in the first-stage and the expected total cost of the edges selected in the second stage.

(a) [8pts] Modify the given deterministic integer program to formulate a two-stage stochastic programming (2SP) model.

(b) [10pts] Show that the obtained 2SP model does not have relatively complete recourse for the instance provided in Figure 1 when $B = 2$.

(c) [7pts] Provide some cases (i.e., families of instances) for which the obtained 2SP model will always have relatively complete recourse.

3 [15pts] Consider the two-stage stochastic programming model (obtained from a supply chain application):

$$\begin{aligned}
\min \quad & \sum_{n \in N} \theta_n x_n + \sum_{k \in K} p_k Q_k(x) \\
\text{s.t.} \quad & \sum_{n \in N} x_n \le I \\
& x_n \ge 0 \quad \forall n \in N
\end{aligned}$$

where, for any $k \in K$,

$$\begin{aligned}
Q_k(x) = \min \quad & \sum_{n \in N} \theta'_n (u_n^k + v_n^k) + h z_n^k + g s_n^k \\
\text{s.t.} \quad & I + \sum_{n \in N} u_n^k \ge \sum_{n \in N} v_n^k + \sum_{n \in N} x_n && \forall k \in K \\
& Y_n + x_n + v_n^k + s_n^k = d_n^k + z_n^k + u_n^k && \forall n \in N, k \in K \\
& u_n^k, v_n^k, z_n^k, s_n^k \ge 0 && \forall n \in N, k \in K
\end{aligned}$$

A data file "data.py" and a data reader file "ReadData.py" are provided. Solve the stochastic programming model for the provided instance using the modeling environment of your choice and provide the optimal value, an optimal first-stage solution, and the solution time.

4 [20pts] **(Quick research)** Song and Dinh [1] used stochastic programming to minimize the misinformation spread in social media networks. Clearly explain the model formulation, given as model (1) in their paper. It would be helpful to create a simple numerical example (i.e., specify an example graph, etc.) to explain the model.

References

[1] Yongjia Song and Thang N Dinh. Optimal containment of misinformation in social media: A scenario-based approach. In *International Conference on Combinatorial Optimization and Applications*, pages 547–556. Springer, 2014.