MS&E 114/214 Autumn 2024 Practice Problems

These are meant as practice problems. There may be some aspects of this problem set that are open-ended. The actual exam will be quite precise.

1. Revenue sensitivity for production problems.

(a) Write an LP to maximize total revenue from producing chairs and tables using wood, labor hours, and metal parts. Solve the LP using excel and generate the sensitivity report.

Parameters:

- **Revenue per Unit:**
  - Chair: \$50
  - Table: \$70
- **Resource Requirements per Unit of Product:**

Commodity	Chair Requirement	Table Requirement
Wood	2 units	3 units
Labor Hours	4 hours	5 hours
Metal Parts	1 unit	2 units

- **Total Available Resources:**
  - Wood: 100 units
  - Labor Hours: 180 hours
  - Metal Parts: 50 units

(b) Using just the sensitivity report and your optimum solution from part (a), answer as many of the questions as you can. Show your work in the excel sheet clearly. No points for correct answers without a clear explanation. For those parts that you can not answer with just your answer to part (a), explain why.

What would be the maximum revenue if you had:

(i) 10 extra units of wood?
(ii) 10 fewer units of wood?
(iii) 10 extra units of labor?
(iv) 10 fewer units of labor?

2. Using flows to model production

(a) Revisit the discussion problem from 11/22 and make sure you understand how to solve it.

(b) Modify the formulation from the discussion section to handle the constraint that no goods can be stored in the warehouse for more than one time period. So any units manufactured in month i must be used in month i or $i + 1$.

3. Base stock policies

(a) Reproduce the LP for the simple newsvendor problem that we solved in class. State the insight that we derived in class from this LP regarding base stock policies.

(b) Write down an LP you can solve to find the optimum base stock.

(c) *[EC]* Come up with a simpler process for finding the optimum base stock if the holding cost and the lost sales cost are the same (no LPs, just some simple formula).

4. Revenue Management, Extra Credit

You are designing a pricing scheme for a hotel chain for the thanksgiving weekend (the entire weekend sells as one block) which is currently 12 weeks away. The hotel has 200 identical rooms. You can set the price for a room to be $400 or $800, and you are allowed to change the price every week. Here is the estimated number of people who will reserve the room during every week at each price point. Week 1 refers to the week 12 weeks before the thanksgiving weekend, and week 12 is the week just before that weekend.

Week	Estimated Demand at $400	Estimated Demand at $800
1	70	2
2	60	3
3	60	4
4	18	5
5	22	10
6	20	13
7	20	16
8	10	10
9	5	5
10	5	4
11	5	0
12	5	0

(a) The first column of the above table is always at least as large as the second. Why might that be?

(b) Write down the dynamic program that you will use to find the maximum total revenue. *Hint: Write down the state space, the boundary conditions, the formula you will use to calculate the maximum revenue, and the order in which you will fill the table.*

(c) Solve the problem using either Excel or Python and report the maximum revenue overall.

(d) How will you find out what price to set each day? Your dynamic program above will only give you the maximum revenue. You don't have to solve this problem numerically – just let us know in words.

(e) How would you handle the case where the price is not allowed to decrease, because of a price guarantee that has been offered by the hotel?

5. Finance problems

Part 1: Super-Replication of Target Payoff The investor group specifies the following target payoffs in three economic scenarios (*Recession*, *Normal Growth*, *Boom*):

$$y = \begin{bmatrix} 5 \\ 12 \\ 25 \end{bmatrix}.$$

Each of the five asset classes generates different payoffs in these scenarios based on their characteristics:

- Clean Energy Stocks (CE): Payoff increases significantly with economic growth. The payoffs in recession, normal growth, and boom scenarios are 1, 3, and 5 times the investment, respectively.

- Green Bonds (GB): Payoffs are stable across all scenarios, returning 2 times the investment in all states.
- Commodities (CM): High payoffs during recessions and moderate payoffs otherwise. The payoffs are 3, 2, and 2 times the investment in recession, normal growth, and boom scenarios, respectively.
- Real Estate Funds (RE): Payoffs increase with better economic conditions. The payoffs are 3, 4, and 5 times the investment in recession, normal growth, and boom scenarios, respectively.
- Technology ETFs (TE): High payoffs in boom scenarios, with lower returns during recessions. The payoffs are 4, 5, and 6 times the investment in recession, normal growth, and boom scenarios, respectively.

The costs of the asset classes are:

$$c = \begin{bmatrix} 7 \\ 5 \\ 4 \\ 6 \\ 8 \end{bmatrix}.$$

(a) Formulate the linear program to determine the optimal portfolio $x = \begin{bmatrix} x_{\text{CE}} \\ x_{\text{GB}} \\ x_{\text{CM}} \\ x_{\text{RE}} \\ x_{\text{TE}} \end{bmatrix}$ that minimizes the total cost while ensuring the portfolio's payoff meets or exceeds the target payoff y in all scenarios.

Part 2: Arbitrage Detection The analysts have identified the following return matrix for the five asset classes under the same economic scenarios:

$$R = \begin{bmatrix} 1.10 & 0.95 & 1.05 & 1.15 & 1.20 \\ 1.05 & 1.00 & 1.08 & 1.12 & 1.15 \\ 1.20 & 1.10 & 1.15 & 1.25 & 1.30 \end{bmatrix}.$$

(a) Formulate the linear program to determine whether there is an arbitrage opportunity by finding a vector $x = \begin{bmatrix} x_{\text{CE}} \\ x_{\text{GB}} \\ x_{\text{CM}} \\ x_{\text{RE}} \\ x_{\text{TE}} \end{bmatrix}$.

Part 3: Portfolio Optimization with Risk and Return The investor group wants the fund to maximize risk-adjusted returns. The characteristics of the asset classes are:

- Expected returns:

$$\mu = \begin{bmatrix} 0.08 \\ 0.06 \\ 0.04 \\ 0.05 \\ 0.07 \end{bmatrix}.$$

- Variances and covariances (covariance matrix):

$$V = \begin{bmatrix} 0.005 & 0.002 & 0.001 & 0.002 & 0.003 \\ 0.002 & 0.004 & 0.001 & 0.002 & 0.002 \\ 0.001 & 0.001 & 0.003 & 0.001 & 0.002 \\ 0.002 & 0.002 & 0.001 & 0.004 & 0.002 \\ 0.003 & 0.002 & 0.002 & 0.002 & 0.006 \end{bmatrix}.$$

(a) Formulate the optimization problem to minimize portfolio variance while satisfying the constraints:

- Achieve a minimum expected return of 7%,
- Be fully invested,
- Avoid short-selling.

(b) Solve for the optimal weights of the five asset classes using Python

(c) Discuss how the fund would adjust the portfolio weights if the investor group becomes more risk-averse.

6. Machine learning

You are tasked with building a binary classification model to predict whether an email is spam (+1) or not (-1) based on two features: x_1 (length of the email in words) and x_2 (number of links in the email).

The training dataset contains the following labeled examples:

Email	x_1 (Length)	x_2 (Links)	Label (y)
1	50	2	$+1$
2	30	0	-1
3	80	5	$+1$
4	60	1	-1
5	90	3	$+1$

Part 1: Linear Classification

(a) Sketch the training data on a 2D plane with x_1 on the x-axis and x_2 on the y-axis. Clearly label the points and their classes.

(b) Propose a linear decision boundary of the form $w_1x_1 + w_2x_2 = \alpha$. Write the inequality constraints that the weights w_1, w_2 and threshold α must satisfy to separate the data.

(c) Explain why the data may not be perfectly separable. Suggest how a soft-margin support vector machine (SVM) approach can address this issue.

Part 2: Support Vector Machines

(a) Explain why it is essential to evaluate a model on a test set rather than the training set. How does this relate to generalization error?

(b) Discuss why soft-margin SVMs are more practical than hard-margin SVMs for real-world datasets.

(c) If the value of C (penalty parameter) is set very high, what potential problems might arise? Conversely, what issues occur if C is set very low?

7. Basics

Consider the problem of allocating the rights to use GPU clusters among n students. There are m GPU clusters in the department. Cluster j can be assigned to at most c_j students, where each capacity c_j is an integer. The utility of student i from the use of cluster j is u_{ij}. The utilities from the use of multiple clusters are additive. Each student needs to be assigned *at least* two clusters. Fractional assignments are allowed.

(a) Does there exist an integer basic feasible solution? Does this depend on the capacities of the clusters? Does this depend on the utilities?

(b) Formulate the problem of finding a utility-maximizing assignment as an LP. Clearly mention all the parameters, decision variables, the objective function, and constraints. How many decision variables and constraints are there? How many variables will there be in the dual program?

(c) The department is planning to increase the capacity of any one cluster by a *small* amount. Assuming the cost is the same for all clusters, how would you determine which cluster they should invest in? The goal is to maximize the increase in the sum of utilities of all students.

8. Social Choice

This problem is based on a previous HW problem. Consider an election with n voters and m candidates. The goal is to select a committee of $l < m$ winners. Each voter submits a preference ranking on all candidates. We are interested in studying different scoring-based aggregation rules to determine the winning committee in which a score of s_k is given to a candidate ranked at the k-th position, and the l candidates with the highest scores win.

Scoring-based rules: A scoring rule is determined by a sequence of nonnegative real numbers $(s_1, s_2, \ldots, s_m)$ satisfying $1 = s_1 \geq s_2 \geq \ldots \geq s_m = 0$. A candidate who is ranked at the k-th position by a voter i is assigned a score of s_k, and the cumulative score for each candidate is calculated by summing across all voters. Ties may be broken arbitrarily. An example is described in the following for two different scoring vectors.

Example 1 *Consider an election with three candidates A, B and C and three voters. A committee of size $l = 2$ is to be chosen. The first voter ranks $C > A > B$, the second voter ranks $B > A > C$, and the third voter ranks $B > A > C$. We now discuss the winning committee under two different scoring vectors.*

- *Suppose that a scoring vector of $(1, 0.75, 0)$ is used. In this case, the candidate A gets a cumulative score of $0.75 + 0.75 + 0.75 = 2.25$, the candidate B a cumulative score of $0 + 1 + 1 = 2$ and the candidate C a cumulative score of 1 and thus a committee of A and B is the winner.*
- *Suppose that a scoring vector of $(1, 0, 0)$ is used. In this case, the candidate A gets a cumulative score of 0, the candidate B has a cumulative score of 2 and the candidate C has a cumulative score of 1. In this case, a committee of B and C is the winner.*

Thus, different scoring vectors may lead to different winning committees.

(a) Each voter has provided you with a ranking of the candidates. Your goal is to determine whether there is a scoring-based rule that can make a given target committee win the election (ties can be broken in your favour). Model this problem as a Linear Program. Clearly describe what the parameters are, what the decision variables are, what the objective is, and what the constraints are. No credit is given for work that is not legible or where the LP model is not defined crisply and with proper notation.

(b) Now consider the situation where you can re-weight the votes such that each voter's vote is weighted between 1 and 1.2. There is a constraint that the total deviation of the weights from the uniform-weights case (where all weights are 1) is at most $0.05n$. Modify your LP from the previous part for this situation.