

Session 19: Abstract Formulation I (with Solutions)

Example (Concrete formulation from last session's Q1)

Amazon.com is expanding its business by launching a physical store in West LA. As the manager, you need to select which bestsellers to carry at the store's grand opening. The following table provides the list of Top 10 Bestsellers in Literature & Fiction, along with their genres. Note that some bestsellers belong to more than one genre.

Rank \ Genre	Literary	Sci-Fi	Romance	Thriller
1	✓			
2		✓		✓
3			✓	✓
4	✓		✓	
5	✓			
6			✓	
7		✓		
8				✓
9	✓	✓		
10			✓	

You wish to carry the minimum number of bestsellers, while ensuring that there are at least two bestsellers in each genre. Formulate this as an optimization problem.

Decision variables: Let x_i denote whether to carry book i , where $i \in \{1, 2, \dots, 10\}$. (Binary)

Objective and Constraints:

Minimize: $x_1 + x_2 + \dots + x_{10}$

subject to:

(Literary) $x_1 + x_4 + x_5 + x_9 \geq 2$

(Sci-Fi) $x_2 + x_7 + x_9 \geq 2$

(Romance) $x_3 + x_4 + x_6 + x_{10} \geq 2$

(Thriller) $x_2 + x_3 + x_8 \geq 2$

Step 3a. Use variables to encode all input data

Data:

- I : the set of books.
- J : the set of genres.
- a_{ij} : a binary data variable (not decision variable) denoting whether book i is of genre j . (These corresponds to the checkmarks in the original question.)
- q_j : how many books do we need of genre j .

Step 3b. Formulate the LP/MIP in terms of only data and decision variables

Decision Variables: For each book $i \in I$, let x_i denote whether to carry the book. (Binary)

Objective and constraints:

$$\begin{aligned} \text{Minimize: } & \sum_{i \in I} x_i \\ \text{subject to: } & \\ \text{(Enough books in genre)} & \sum_{i \in I} a_{ij} x_i \geq q_j \quad \text{for each genre } j \in J. \end{aligned}$$

Alternative encoding of data

Data:

- I : the set of books.
- J : the set of genres.
- I_j : the set of books of genre j .
- q_j : how many books we need of genre j .

Decision Variables: For each book $i \in I$, let x_i be a decision variable denoting whether to carry the book. (Binary)

Objective and constraints:

$$\begin{aligned} \text{Minimize: } & \sum_{i \in I} x_i \\ \text{subject to: } & \\ \text{(Enough books in genre)} & \sum_{i \in I_j} x_i \geq q_j \quad \text{for each genre } j \in J. \end{aligned}$$

Q1 (Emergency Vehicle Location)

The city of Metropolis is divided into nine districts and is considering seven possible sites to place emergency vehicles. The table below shows the time (minutes) it takes for an emergency vehicle to travel from each district to each site. (The column labels are sites and row labels are districts.)

District \ Row	1	2	3	4	5	6	7
1	5	3	4	3	8	9	0
2	3	6	5	4	8	0	3
3	4	3	6	8	10	3	2
4	6	0	2	7	3	2	5
5	2	8	2	5	0	6	8
6	2	6	4	0	7	3	5
7	0	12	5	5	5	7	2
8	10	9	0	2	3	5	7
9	2	4	5	7	3	4	5

Formulate an optimization problem to find the minimum number of sites so that all districts are within three minutes of an emergency vehicle, then generalize this to be able to handle arbitrary data of a similar format.

Step 1. Identify the decision, objective, and constraints in English

Decision: For each of the seven sites, decide whether or not to use it to place emergency vehicles.

Objective: Minimize the number of sites used.

Constraints: For each of the 9 districts, we need there to be at least one site that is within three minutes at which we placed emergency vehicles. In other words, for each district,

$$(\text{\# of sites within three minutes with emergency vehicle}) \geq 1$$

Step 2. Formulate the optimization as linear expressions of decision variables

Decision variables: Let x_i be a binary variable denoting whether we place an emergency vehicle at site i , where $i \in \{1, 2, \dots, 7\}$.

Objective and constraints:

$$\begin{aligned} & \text{Minimize} && x_1 + x_2 + \dots + x_7 \\ & \text{subject to:} && \\ & (\text{District 1}) && x_2 + x_4 + x_7 \geq 1 \\ & (\text{District 2}) && x_1 + x_6 + x_7 \geq 1 \\ & (\text{District 3}) && x_2 + x_6 + x_7 \geq 1 \\ & (\text{District 4}) && x_2 + x_3 + x_5 + x_6 \geq 1 \\ & (\text{District 5}) && x_1 + x_3 + x_5 \geq 1 \\ & (\text{District 6}) && x_1 + x_4 + x_6 \geq 1 \\ & (\text{District 7}) && x_1 + x_7 \geq 1 \\ & (\text{District 8}) && x_3 + x_4 + x_5 \geq 1 \\ & (\text{District 9}) && x_1 + x_5 \geq 1 \end{aligned}$$

Step 3a. Use variables to encode all input data

Data:

- I : the set of sites.
- J : the set of districts.
- a_{ij} : a binary variable for whether site i is sufficiently close to district j . (This data can be obtained from the distance table by checking whether each entry is no more than the given threshold.)

Step 3b. Formulate the LP/MIP in terms of only data and decision variables

Decision variables: For each site $i \in I$, let x_i be a binary variable denoting whether we place an emergency vehicle at this site.

Objective and constraints:

$$\begin{aligned} & \text{Minimize} && \sum_{i \in I} x_i \\ & \text{subject to:} && \\ & (\text{Close by site exists}) && \sum_{i \in I} a_{ij} x_i \geq 1 \quad \text{for each district } j \in J. \end{aligned}$$

Alternative encoding of data

The following formulation is also correct. Note that we encode the data using sets instead of using a binary matrix. Moreover, the constraint $x_i \in \{0, 1\}$ is unnecessary if we specify clearly that the variable is binary.

Data:

- I : the set of sites.
- J : the set of districts.
- I_j : the set of sites sufficiently close to district j .

Decision variables: For each site i , let x_i denote whether we place an emergency vehicle at this site. (binary)

$$\begin{aligned} & \text{Minimize} && \sum_{i \in I} x_i \\ & \text{subject to:} \\ & \text{(Close by site exists)} && \sum_{i \in I_j} x_i \geq 1 \quad \text{for each district } j. \end{aligned}$$

Q2 (Investment Planning Revisited)

Recall the investment planning problem from last session, with the following concrete formulation:

Decision Variables:

- Let x_A, x_B, x_C, x_D, x_E denote how much to invest in each of the five options. (continuous)
- Let y_0, y_1, y_2, y_3 denote the cash at hand at the end of each year. (Continuous)

Objective and constraints:

$$\begin{aligned} & \text{Maximize:} && y_3 \\ & \text{subject to:} \\ & \text{(Limit on investment)} && x_A, x_B, x_C, x_D, x_E \leq 75000 \\ & \text{(Cash flow in year 0)} && 10000 - x_A - x_C - x_D = y_0 \\ & \text{(Cash flow in year 1)} && y_0 + 0.5x_A + 1.2x_C - x_B = y_1 \\ & \text{(Cash flow in year 2)} && y_1 + x_A + 0.5x_B - x_E = y_2 \\ & \text{(Cash flow in year 3)} && y_2 + x_B + 1.9x_D + 1.5x_E = y_3 \\ & \text{(Non-negative investments)} && x_A, x_B, x_C, x_D, x_E \geq 0 \\ & \text{(Non-negative cash flow)} && y_0, y_1, y_2, y_3 \geq 0 \end{aligned}$$

Use data variables to represent all numbers and re-formulate the above into an abstract formulation, which can handle arbitrarily many investment options and arbitrarily many years. The parameters 75000 and 10000 should also be represented using parameter variables L and C . The objective should be to maximize the cash on hand at the end of year m .

Step 3a. Use variables to encode all input data

Data:

- I : the set of investment options.
- m : the number of years to optimize.
- L : the maximum amount to invest in a single option.
- C : the cash on hand at the beginning of year 0.
- a_{it} : equal to -1 if investment option i should be invested in year t , and otherwise equal to the percentage return of investment i in year t .

Step 3b. Formulate the LP/MIP in terms of only data and decision variables

Decision variables:

- x_i : the amount of cash to invest in option $i \in I$. (Continuous)
- y_t : the amount of cash at hand at the end of year $t \in \{0, 1, \dots, m\}$. (Continuous)

Objective and constraints:

$$\begin{array}{ll} \text{Maximize} & y_m \\ \text{subject to:} & \\ \text{(Limit on investment)} & x_i \leq L \quad \text{for all } i \in I. \\ \text{(Balance in Year 0)} & C + \sum_{i \in I} a_{i0} x_{i0} = y_0 \\ \text{(Balance in Other Years)} & y_{t-1} + \sum_{i \in I} a_{it} x_{it} = y_t \quad \text{for all } t \in \{1, 2, \dots, m\}. \\ \text{(Non-negative investments)} & x_i \geq 0 \quad \text{for all } i \in I. \\ \text{(Non-negative cash flow)} & y_t \geq 0 \quad \text{for all } 0 \leq t \leq m. \end{array}$$