Week 11: Abstract Formulation

Session 22: Creating an Abstract Formulation from Scratch

Tips on Creating an Abstract Formulation

1. First create a concrete formulation or at least fragments of a concrete formulation. Write the abstract formulation by generalizing this, and afterward, manually expand the abstract formulation using made-up data and ensure that you get back the concrete formulation you started with.
2. Be familiar with **commonly used patterns**: capacity constraints, demand constraints, flow conservation constraints, conflicts, pre-requisites, conditional right hand sides, big-M, ...
3. Watch for **red flags**:
  i) undefined indices or variables;
  ii) clashing definitions of indices or variables;
  iii) hard-coded numbers that are not inherent in the logic;
  iv) summation signs with nothing under it.
  v) the objective or one side of a constraint is not a linear expression of decision variables.

In-Class Exercise: Checking for red flags

Read through the following examples and verify that there are no red flags.

Examples from Last Session (for your ease of reference)

Example 1. Abstract Formulation for Assortment Planning

Data:

- B: the set of books.
- G: the set of genres.
- B_g: the set of books of genre g.
- r_g: the number of books required for genre g.

Decision Variables:

- x_b: whether to carry book b. (Binary)

Objective and constraints:

$$\begin{aligned}
\text{Minimize:} && \sum_{b \in B} x_b \\
\text{subject to:} \\
\text{(Enough books in genre)} && \sum_{b \in B_g} x_b & \ge r_g && \text{ for each genre $g \in G$.}
\end{aligned}$$

Corresponding Latex

```
$$\begin{aligned}
\text{Minimize:} && \sum_{b \in B} x_b \\
\text{subject to:} \\
\text{(Enough books in genre)} && \sum_{b \in B_g} x_b & \ge r_g && \text{ for each genre $g \in G$.}
\end{aligned}$$
```

Example 2. Abstract Formulation for Food Production

Data:

Flow conservation constraints

change in inventory = inflow − outflow

- n: number of months.
- T: set of months. $T = \{1, 2, 3, \cdots, n\}$.
- p_t: price of oil in month t.
- d_t: demand in month t.
- s: amount of oil that can be stored at any time.

Decision Variables:

- x_t: amount of oil to buy in month t. (Continuous)
- y_t: amount of oil stored at the end of month t. (Continuous)

Objective and Constraints:

$$\begin{aligned}
\text{Minimize: } && \sum_{t \in T} p_tx_t \\
\text{s.t. } && \\
&& y_1 &= x_1 - d_1 \\
&& y_t &= x_t + y_{t-1} - d_t && \text{for each month $t \in \{2,3,\cdots,n\}$.} \\
&& y_t &\le s && \text{for each month $t \in T$.} \\
&& x_t, y_t & \ge 0
\end{aligned}$$

Corresponding Latex

```
$$\begin{aligned}
\text{Minimize: } && \sum_{t \in T} p_tx_t \\
\text{s.t. } && \\
&& y_1 &= x_1 - d_1 \\
&& y_t &= x_t + y_{t-1} - d_t && \text{for each month $t \in \{2,3,\cdots,n\}$.} \\
&& y_t &\le s && \text{for each month $t \in T$.}  \\
&& x_t, y_t & \ge 0
\end{aligned}$$
```

Example 3. Abstract Formulation for Project Selection

Data:

- P: set of projects
- C: set of conflicts. Each $(p_1, p_2) \in C$ is a pair of projects that conflicts with one another.
- R: set of prerequisite pairs. Each $(p_1, p_2) \in R$ is a pair such that project p_1 is a prerequisite to project p_2.

Decision Variables: x_p: whether to pursue project p. (Binary)
Objective and Constraints:

$$\begin{aligned}
\text{Maximize} && \sum_{p \in P} x_p \\
\text{s.t.} && \\
&& x_{p_1} + x_{p_2} \le 1 && \text{For each conflicting pairs } (p_1, p_2) \in C. \\
&& x_{p_1} \ge x_{p_2} && \text{For each pair } (p_1, p_2) \text{ such that } p_1 \text{ is a prereq to } p_2.
\end{aligned}$$

conflict constraints

pre-requisite constraints.

Corresponding Latex

```
$$\begin{aligned}
\text{Maximize} && \sum_{p \in P} x_p \\
\text{s.t.} && \\
&& x_{p_1}+x_{p_2} &\le 1 && \text{For each conflicting pairs $(p_1,p_2) \in C$.} \\
&& x_{p_2} & \le x_{p_1} && \text{For each pair $(p_1,p_2)$ such that $p_1$ is a prereq to $p_2$.}
\end{aligned}$$
```

Example 4. Abstract Formulation for Supply Chain Planning

Data:

- F: set of FCs (fulfillment centers).
- R: set of regions.
- c_{fr}: unit cost of transporting from FC f to region r.
- q_f: capacity of FC f.
- d_r: demand of region r.

Decision Variables:

- x_{fr}: Amount to transport from FC f to region r. (Continuous)

Objective:

$$\text{Minimize: } \sum_{f \in F, r \in R} c_{fr}x_{fr}$$

Constraints:

$$\begin{aligned}
\text{(Capacity)} && \sum_{r \in R} x_{fr} &\le q_f && \text{for each FC $f \in F$.} \\
\text{(Demand)} && \sum_{f \in F} x_{fr} &\ge d_r && \text{for each region $r \in R$.} \\
&& x_{fr} &\ge 0 && \text{for each $f \in F$ and $r \in R$.}
\end{aligned}$$

Corresponding Latex:

```
$$\text{Minimize: } \sum_{f \in F, r \in R} c_{fr}x_{fr}$$
$$\begin{aligned}
\text{(Capacity)} && \sum_{r \in R} x_{fr} &\le q_f && \text{for each FC $f \in F$.} \\
\text{(Demand)} && \sum_{f \in F} x_{fr} &\ge d_r && \text{for each region $r \in R$.} \\
 && x_{fr} &\ge 0 && \text{for each $f \in F$ and $r \in R$.}
\end{aligned}$$
```

Example 5. Abstract Formulation for Box Selection

Data:

- n: the number of item types, which is also equal to the number of box types.
- $I = \{1, 2, \cdots, n\}$: the set of box types, with the labels ordered in increasing box sizes.
- s_i: the size of box type i. This is also the variable cost of producing boxes of this type.
- d_i: the demand for item i.
- $M = \sum_{i=1}^{n} d_i$: the total demand of all items.
- F: the fixed cost of making each type of box. (In the concrete formulation, $F = 1000$.)

Decision Variables:

- Y_i: how many boxes to make of box type i. (Integer)

- Z_i: whether to make the mold for box type i. (Binary)

Objective and Constraints:

$$\begin{aligned}
\text{Minimize:} && \sum_{i=1}^n s_iY_i + F\sum_{i=1}^n Z_i \\
\text{s.t.} && \\
\text{(Demand)} && \sum_{j=i}^n Y_j & \ge \sum_{j=i}^n d_j && \text{for each $i \in I$.} \\
\text{(On/Off)} && Y_i & \le MZ_i && \text{for each $i \in I$.} \\
&& Y_i & \ge 0 && \text{for each $i \in I$.}
\end{aligned}$$

If $Z_i = 1$
$0 \le Y_i \le 99999$

If $Z_i = 0$
$0 \le Y_i \le 0$

Corresponding Latex:

```
$$\begin{aligned}
\text{Minimize:} &&  \sum_{i=1}^n s_iY_i + F\sum_{i=1}^n Z_i \\
\text{s.t.} && \\
\text{(Demand)} && \sum_{j=i}^n Y_j & \ge \sum_{j=i}^n d_j && \text{for each $i \in I$.} \\
\text{(On/Off)} && Y_i & \le MZ_i && \text{for each $i \in I$.} \\
&& Y_i & \ge 0 && \text{for each $i \in I$.}
\end{aligned}$$
```

Exercise 11.3: Assigning Consultants to Projects

Download the Jupyter notebook attached to the Blackboard link for this exercise and submit it there. The notebook contains the following question.

Trojan Consulting would like to assign consultants to projects in a way that minimizes total travel costs while satisfying the skill requirements of each project and avoiding assigning the same consultant to two projects with overlapping time frames.

In the following example, there are four consultants, each of whom may possess one or more of two possible skills. (A checkmark indicates whether the person has the skill.) Each project requires at least a certain number of consultants of each skill. If a consultant has both skills, he/she can count toward the number required for both skills, and the travel cost may potentially be less as one less person would be needed. Projects 1 and 2 have conflicting timelines, so the same consultant cannot be assigned to both. Similarly, projects 2 and 3 are also in conflict. But the same consultant may be assigned to projects 1 and 3.

Consultant	Accounting	Operations
Alice	√	
Bob	√	√
Charlie		√
Daphne	√	√

Project	Accounting	Operations
P1	2	1
P2	1	1
P3	0	2

Costs	P1	P2	P3
Alice	10	0	5
Bob	8	15	13
Charlie	0	5	10
Daphne	10	3	0

Write an abstract formulation of a linear optimization model to find a cost-minimizing assignment that satisfies all of the above constraints. Your formulation needs to be general enough to handle arbitrarily many consultants, projects, skills, as well as arbitrary information on skills of consultants, requirements of projects, conflicts between projects, and travel costs.

Data:

- C : set of consultants
- P : set of projects.
- S : set of skills.
- r_{ps} : no. of consultants required for project p of skill s..
- K : set of pair of projects that conflict

Decision Variables: whether to assign to each consultant each project

x_{A1}, x_{A2}, x_{A3} x_{ij} : whether to assign i to project j. (Binary)

$x_{B1}, x_{B2}, \ldots$

Objective and Constraints:

Minimize cost $10x_{A1} + 0x_{A2} + \ldots$

Constraints: P1-Acc $x_{A1} + x_{B1} + x_{D1} \geq 2$

Demand: P2-Acc $x_{A1} + x_{B1} + x_{D1} \geq 1$

Conflict : Alice can't do 1 & 2 : $x_{A1} + x_{A2} \leq 1$

" " " 2 & 3 : $x_{A2} + x_{A3} \leq 1$

Bob " " " : . . . . .

Abstract: $\min \sum_{i \in I, j \in J} c_{ij} x_{ij}$

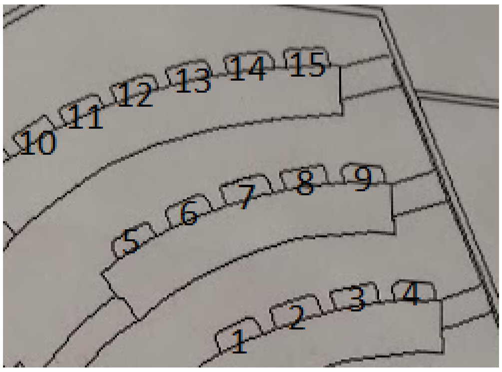
$\sum_{i \in I} x_{ij} \geq r_{jk}$ for each project j & skill k.

$x_{ip_1} + x_{ip_2} \leq 1$ for each consultant i & pair of projects that conflict. $(p_1, p_2) \in K$

Exercise 11.4 Classroom Seating under Social Distancing

Download the Jupyter notebook attached to the Blackboard link for this exercise and submit it there. The notebook contains the following question.

A committee at USC Marshall is exploring the viability of in-person instruction while observing social distancing guidelines. One challenge is that certain classrooms have tables and seats bolted to the floor, and the seats cannot be moved unless the rooms undergo significant remodeling. As an illustration, the following image is a portion of the floor plan for JKP204, and the numbers in the image correspond to the individual seats. As you can see, the distance between adjacent seats can be quite close, and the room would be overly dense if every seat is used. Since the seats cannot be moved, only a subset of them can be used to seat students.



Your task is to formulate an optimization problem to maximize the number of students that can be safely seated in the current classrooms without remodeling. The committee has access to the detailed floor plans of every classroom, and they have labelled every seat as above and measured its exactly position in terms of x-y coordinates, so they can easily compute the distance between any two seats. (For simplicity, the distance between two seated students is defined to be the straight-line distance between the center of the two seats.) Based on discussions with public health officials, the committee has summarized the requirements for safely seating students as follows:

1. The minimum distance between any two seated students is at least 6 feet.
2. For every seated student, the number of other students seated within a 9 feet radius is at most 3, so the density of the room is kept low. (In other words, if we draw a circle centered at a seated student with a radius of 9 feet, then there are at most 4 students seated strictly inside this circle, including the first student.)

Write an abstract formulation of a LP/MIP to solve the above problem, by listing all the data variables, decision variables, objective, and constraints. You may define any data variables that can be straightforwardly calculated based on the information the commitee has access to, but your definition must be completely clear and without ambiguities. Your formulation must be flexible enough to handle an arbitrary floor plan, not only the one shown above, and your objective and constraints must all be linear.

Data:

Decision Variables:

X_i: whether to seat student on seat i

Objective and Constraints:

Maximize # of seats used max $X_1 + \dots\dots + X_n$

Constraints:

min distance (conflict): $X_{i_1} + X_{i_2} \le 1$ for every pair of i_1 & i_2 within 6 ft of each other.

density:

If $X_1 = 0$

$X_2 + X \dots \le n$

If $X_1 = 1$

$\underbrace{X_2 + \dots + }_{\text{set of seats within 9 ft of each other.}} \le 3$