Lesson Plan

- Group Exercise 1. Airline scheduling
- Group Exercise 2. Firehouse location
- Solutions
- Formulation Strength
Exercise 1. Airline scheduling

- Need to operate a number of flights
  – e.g., LAX to ORD, 9am

- Routes are feasible sequences of flights for a single plane considering factors such as turnaround time.

- Find an assignment of plane equipment to a set of possible routes (e.g. LAX-ORD, ORD-JFK, JFK-LAX).

- Each route has an associated cost (e.g., based on distance.)

- **Goal**: select a set of routes so that each flight is operated **exactly** once.

### Example set of possible routes

<table>
<thead>
<tr>
<th>Route</th>
<th>SFO-LAX</th>
<th>SFO-DEN</th>
<th>SFO-SEA</th>
<th>LAX-ORD</th>
<th>LAX-SFO</th>
<th>ORD-DEN</th>
<th>ORD-SEA</th>
<th>DEN-SFO</th>
<th>DEN-ORD</th>
<th>SEA-SFO</th>
<th>SEA-LAX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**COST ($1000's)**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

e.g., route 6 does SFO-SEA then SEA-LAX then LAX-SFO.
Exercise 1: Two variations

• **Equipment scheduling:**
  – Equipment routes $j \in J$ and flights $i \in I$. Let $c_j$ denote the cost of using route $j$.
  – **Goal:** Minimize total cost, subject to each flight covered by exactly one plane.

• **Crew scheduling:**
  – Once planes scheduled, can now schedule crews to “crew routes.” Set of crew routes may be different from equipment routes because turnaround time between flights different. Also: it is ok to have more than one crew on a flight.
  – Crew routes $k \in K$, flights $i \in I$. Cost $c_k$.
  – **Goal:** Minimize total cost, subject to each flight covered by one or more crew.
Exercise 2: Firehouse location

• J potential sites (m total), I districts (n total).
• Each site can serve every district, but distance $d_{ij} \geq 0$ from site j to district i.
• Each district $i \in I$ has population $p_i$.
• Cost $f_j(s_j) = K_j + c_js_j$ to use firehouse at j to service $s_j$ people in total. $K_j$ fixed cost, $c_j$ variable cost.
• Budget B.

• Goal: determine firehouse sites and assignment of districts to these sites to minimize max distance to a firehouse. (Each district needs to be assigned to exactly one site.) Don’t exceed budget.
Correct formulations

• **Definition.** A formulation is **correct** if the set of feasible values for decision variables corresponds to the set of feasible decisions in the problem to be modeled.

• Correctness => optimal solution will solve intended problem for any choice of objective.

• A “true model” of the problem.

Alternate Formulations
Formulation (P₁):
• If m=2 sites, and n=4 districts:
  • \( x_{11} + x_{21} + x_{31} + x_{41} \leq 4 \ y_1 \)
  • \( x_{12} + x_{22} + x_{32} + x_{42} \leq 4 \ y_2 \)
  • m constraints
Alternate Formulations

Formulation (P₁):  
• If m=2 sites, and n=4 districts:
  • $x_{11} + x_{21} + x_{31} + x_{41} \leq 4y_1$
  • $x_{12} + x_{22} + x_{32} + x_{42} \leq 4y_2$
  • m constraints

Formulation (P₂):
• $x_{11} \leq y_1$, $x_{21} \leq y_1$, $x_{31} \leq y_1$, $x_{41} \leq y_1$
• $x_{12} \leq y_2$, $x_{22} \leq y_2$, $x_{32} \leq y_2$, $x_{42} \leq y_2$
• mn constraints

• Which is easier to solve? Why?

LP Relaxations

• A **polyhedron** is a set that can be described in the form $P = \{x \in \mathbb{R}^n : Ax \leq b\}$
• Fact: the feasible set of any LP can be described as a polyhedron.
LP Relaxations

• A **polyhedron** is a set that can be described in the form \( P = \{ x \in \mathbb{R}^n : Ax \leq b \} \).

• **Fact**: the feasible set of any LP can be described as a polyhedron.

• **Definition**: The **LP relaxation** of an integer program replaces all integer variables with continuous variables.

• **Definition**: The **polyhedron** (\( P \)) of an IP is the feasible region for the LP relaxation.

**Illustrating correct formulations**

Suppose \( P_0 \) is a correct formulation. Why are \( P_1 \) and \( P_2 \) also correct? Which is “ideal”??
Formulation Strength

**Definition.** Given two correct IP formulations, formulation 2 is **stronger** than formulation 1 if $P_2 \subset P_1$.

Equivalently: formulation 2 is “tighter”.

A stronger formulation is preferred because it eliminates some fractional solutions. Tends to makes the IP easier to solve.

**Example: Facility Location**

- $P_1$: $\sum_i x_{ij} \leq ny_j$, for each $j$
- $P_2$: $x_{ij} \leq y_j$, for each $i, j$
Example: Facility Location

- $P_1$: $\sum_i x_{ij} \leq n y_j$, for each $j$
- $P_2$: $x_{ij} \leq y_j$, for each $i, j$
- We have $P_2 \subseteq P_1$ because $P_2$ implies $P_1$. By summing over districts $i$, then $\sum_i x_{ij} \leq n y_j$.
- But is $P_2$ stronger?

\[ \begin{array}{c}
\text{sites} \\
1 \\
2 \\
\text{districts} \\
1 \\
2 \\
3 \\
4 \\
\end{array} \]

\begin{align*}
&x_{11} = 1 \\
&y_1 = 1/2 \\
&x_{21} = 1 \\
&y_2 = 1/2 \\
&x_{32} = 1 \\
&x_{42} = 1
\end{align*}

\begin{align*}
&x_{11} + x_{21} + x_{31} + x_{41} \leq 4y_1 \\
&x_{21} + x_{22} + x_{32} + x_{42} \leq 4y_2
\end{align*}
Example: Facility Location

- $P_1$: $\sum_i x_{ij} \leq n y_j$, for each $j$
- $P_2$: $x_{ij} \leq y_j$, for each $i, j$
- We have $P_2 \subseteq P_1$ because $P_2$ implies $P_1$. By summing over districts $i$, then $\sum_i x_{ij} \leq n y_j$.
- But is $P_2$ stronger?

![Diagram of Facility Location](image)

Fractional solution in $P_1$ but not $P_2$. Conclude $P_2 \subset P_1$.

On Good IP Formulations

Correctness is most important!

But: to make IPs faster to solve:

- Prefer formulations that are “strong” and exclude fractional solutions. This may involve using more constraints.
- Generally try to minimize the number of integer variables.
- If using “big Ms” make them as small as possible.
Summary: Modeling via IPs

• IPs are very expressive. Able to model lots of real-world problems.
• Modeling tricks often use indicator variables and “big M” constants.
• There can be multiple correct formulations, and they may differ in speed to solve.
• Prefer stronger formulations, i.e. those with tighter LP relaxations.