Modeling and Solving Discrete Optimization Problems in Practice

Juan Pablo Vielma
Massachusetts Institute of Technology

18.095 - Mathematics Lecture Series.
Cambridge, MA, IAP 2018.
Combinatorial Example: Assignment Problem

- Assign $n$ workers to $m$ tasks to complete all tasks
- At most one task per worker
- Worker $i$ takes $t_{i,j}$ hours to complete task $j$
- Minimize total time worked

Graph:
- Worker and task nodes
- Arcs between worker and task nodes

```
Workers
n = 3

Tasks
m = 2

$\begin{align*}
\text{Workers} \\
1 & t_{1,j} = 4 \\
2 & t_{2,j} = 4 \\
3 & t_{3,j} = 7
\end{align*}$

\begin{align*}
\text{Tasks} \\
1 \\
2
\end{align*}
```
Combinatorial Example: Assignment Problem

- Assign \( n \) workers to \( m \) tasks to complete all tasks
- At most one task per worker
- Worker \( i \) takes \( t_{i,j} \) hours to complete task \( j \)
- Minimize total time worked
- Variables: \( x_{i,j} = 1 \) if worker \( i \) is assigned to task \( j \) and 0 o.w.

\[
\begin{align*}
\min & \quad \sum_{i=1}^{n} \sum_{j=1}^{m} t_{i,j} x_{i,j} \\
\text{s.t.} & \quad \sum_{j=1}^{m} x_{i,j} \leq 1 \quad \forall i \in \{1, \ldots, n\} \quad \text{Worker constraints} \\
& \quad \sum_{i=1}^{n} x_{i,j} \geq 1 \quad \forall j \in \{1, \ldots, m\} \quad \text{Task constraints} \\
& \quad x_{i,j} \in \{0, 1\} \quad \forall i \in \{1, \ldots, n\}, \ j \in \{1, \ldots, m\}
\end{align*}
\]
Traveling Salesman Problem : Visit all Cities Once
Formulation for Traveling Salesman Problem

$$[n] := \{1, \ldots, n\}$$

$$\min \sum_{i,j=1}^{n} d(i, j) x_{i,j}$$

s.t.

$$\sum_{j=1}^{n} x_{i,j} = 1 \quad \forall i \in [n]$$

$$\sum_{i=1}^{n} x_{i,j} = 1 \quad \forall j \in [n]$$

$$x_{i,i} = 0 \quad \forall i \in [n]$$

$$x_{i,j} \in \{0, 1\} \quad \forall i, j \in [n]$$

Homework Question 1: Add missing constraints
Hint: You will need around $2^n$ inequalities
Mixed Integer Programming (MIP)

- Discrete and continuous variables or combinatorial constraints on continuous variables.
- Example: Find minimum volume ellipsoid that contains 90% of data points
Dear Mr. Hunter, Dr. Vielma, & Dr. Zaman,

On behalf of The Greater Boston Food Bank (GBFB), I want to thank you for your recent gifts. Your $16,709.98 contribution will help our neighbors who struggle to have enough to eat and will promote healthy lives and communities in eastern Massachusetts. Generous and dedicated individuals like you have enabled GBFB to progress in our mission to End Hunger Here and work toward your generous donations of $1,500.00 and $15,209.98, received on Wednesday, November 25, 2015, and Tuesday, May 3, 2016, respectively, will provide more than 50,000 nutritious meals to those at risk of hunger throughout eastern Massachusetts.

As one of the largest food banks in the country, GBFB is proud to be a leader in the conversation around food insecurity, using our voice to highlight the importance of nutrition. We aim to improve community health by distributing safe, nutritious food and we are dedicated to increasing the volume of fresh produce we provide. Last year, GBFB distributed a record-breaking 54 million pounds of food, with 25% being fresh produce, a number that will grow to 35% over the next few years. We will continue to ensure that at least 80% of our food is of the highest nutritional quality.

For single-parent families, veterans, low-income seniors, and community college students, the healthiest foods are often those too expensive to afford. GBFB works with our network of donors, food industry and corporate partners, and member agencies, to acquire and distribute high-quality and nutrient-rich items sourced from local and national manufacturers, retailers and growers. By ensuring access to nutrient-dense food, GBFB helps to provide the building blocks for healthy bodies and minds for those in need.

GBFB is committed to the efficient use of resources. We are proud that 91 cents of every dollar donated goes directly to hunger-relief efforts, helping us to earn a 4-star rating from Charity Navigator. With your help, we employ innovative warehouse technology and food industry expertise to maximize our efficiency to have the greatest impact.

You have my earnest gratitude for your partnership in helping to End Hunger Here.

Sincerely,

Suzanne J. Battit
Vice President of External Affairs and Advancement

We gratefully acknowledge your gift and confirm that no goods or services were provided in consideration of this charitable support.

Please retain this letter for your tax records. Our Tax ID# is 04-2717782.

Download Code from Github:

https://github.com/dscotthunter/Fantasy-Hockey-IP-Code

How hard is MIP: Traveling Salesman Problem?

Paradoxes, Contradictions, and the Limits of Science

Many research results define boundaries of what cannot be known, predicted, or described. Classifying these limitations shows us the structure of science and reason.

Noson S. Yanofsky

“A computer would have to check all these possible routes to find the shortest one.”
MIP = Avoid (Complete) Enumeration

- Number of tours for 49 cities $= \frac{48!}{2} \approx 10^{60}$
- Fastest supercomputer $\approx 10^{17}$ flops
- Assuming one floating point operation per tour:
  $> 10^{35}$ years $\approx 10^{25}$ times the age of the universe!
- How long does it take on an iphone?
  - $< 1$ sec ! Dantzig, Fulkerson and Johnson $\leftarrow$ in 54’
  - Even theoretically hard MIPs “can” be solved:
    - Open-source solvers: GLPK, CBC, etc.
    - Commercial: Gurobi, CPLEX, etc.
    - Modeling Language: JuMP
Easy MIP through \texttt{julia} & \texttt{JuMP}

- \texttt{julia} : general purpose programming language
  - download \url{https://julialang.org/downloads/}
    then click or run from command line

- \texttt{JuMP} : modeling language for optimization

- \texttt{GLPK} : Open-source MIP solver
  - \texttt{julia} > Pkg.add("JuMP"); Pkg.add("GLPKMathProgInterface")

- Can also try JuliaBox on web
  - \url{https://www.juliabox.com/}
Assignment problem:

\[
\text{min } \sum_{i=1}^{n} \sum_{j=1}^{m} t_{i,j} x_{i,j} \\
\text{s.t. } \sum_{j=1}^{m} x_{i,j} \leq 1 \quad \forall i \in \{1, \ldots, n\} \\
\sum_{i=1}^{n} x_{i,j} \geq 1 \quad \forall j \in \{1, \ldots, m\} \\
x_{i,j} \in \{0,1\} \quad \forall i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}
\]

```julia
model = Model(solver=GLPKSolverMIP());
@variable(model, x[1:n, 1:m], Bin);
@objective(model, Min, sum(t[i,j]*x[i,j] for i in 1:n, j in 1:m));
@constraint(model, [i=1:n], sum(x[i,j] for j in 1:m) <= 1);
@constraint(model, [j=1:m], sum(x[i,j] for i in 1:n) >= 1);
```

Homework Question 2: Solve problem with random cost Complete file in website.
Solving MIPs: Step 1 = Linear Programming

\[
\begin{align*}
\text{max} & \quad x_2 \\
\text{s.t.} & \quad x_1 + x_2 \leq 1 \\
& \quad -x_1 - x_2 \leq 1 \\
& \quad +x_1 - x_2 \leq 1 \\
& \quad -x_1 + x_2 \leq 1 \\
x_1, x_2 & \in \mathbb{Z}
\end{align*}
\]

Linear Programming (LP) Relaxation

- Solving LPs is easy in theory and practice.
- One reason = LP duality
  - Suppose I guess optimum \( x_1 = 0 \) and \( x_2 = 1 \).
  - How do I prove that for all solutions of LP \( x_2 \leq 1 \)?

\[
\left( \frac{1}{2} \right) \times \left( x_1 + x_2 \leq 1 \right) + \left( \frac{1}{2} \right) \times \left( -x_1 + x_2 \leq 1 \right) \\
\]

\( x_2 \leq 1 \)
Solving MIPs: Step 1 = Linear Programming

Linear Programming (LP) Relaxation

- LP relaxation always gives a (upper) bound on the MIP:
  - If solution of LP is “integer” then you solved the MIP
  - LP solvers return “corner” solution, which fixes “multiple optima” (e.g. $\max x_1 + x_2$)
  - Homework Question 3: Solve LP relaxation of assignment problem with JuMP. Is solution integer?
Solving MIPs: Step 2 = Branch-and-Bound

\[
\begin{align*}
\text{max } z := & \quad x_2 \\
& 3x_1 + 2x_2 \leq 6 \\
& -2x_1 + x_2 \leq 0 \\
& x_1, x_2 \geq 0 \\
& x_1, x_2 \in \mathbb{Z}
\end{align*}
\]

Linear Programming (LP) Relaxation

Homework Question 4:
Prove \( x_2 \leq \frac{12}{7} \) for LP Relaxation.
Modern MIP Solvers = B&B++

- Really branch-and-cut:
  - Use cuts to improve LP relaxation.
- Elaborate heuristics: Rounding +++
- Preprocessing: fixing variables by logical implications.
- Advanced management of B&B tree.
- Extensive tuning of parameters and techniques.

CPLEX

Gurobi 7.0 Performance Benchmarks

SCIP

GLPK

COIN-OR CBC

GNU
Cutting Plane Example: Chátil-Gomory Cuts

\[ P := \left\{ x \in \mathbb{R}^2 : \begin{align*} x_1 + x_2 &\leq 3, \\ 5x_1 - 3x_2 &\leq 3 \end{align*} \right\} \]

\[ H := \left\{ x \in \mathbb{R}^2 : \begin{align*} 4x_1 + 3x_2 &\leq 10.5 \\ \in \mathbb{Z} \end{align*} \right\} \]

if \( x \in \mathbb{Z}^2 \)

\[ 4x_1 + 3x_2 \leq [10.5] \]

Valid for \( H \cap \mathbb{Z}^2 \)

Valid for \( P \cap \mathbb{Z}^2 \)

\[ \frac{27}{8} ( x_1 + x_2 \leq 3 ) + \frac{1}{8} ( 5x_1 - 3x_2 \leq 3 ) \Rightarrow 4x_1 + 3x_2 \leq 10.5 \]
Branch-and-Bound and Cuts (Branch-and-Cut)

\[ \max z := x_2 \]
\[ 3x_1 + 2x_2 \leq 6 \quad x_2 \leq \lfloor 1.71 \rfloor = 1 \]
\[ -2x_1 + x_2 \leq 0 \]
\[ x_1, x_2 \geq 0 \]
\[ x_1, x_2 \in \mathbb{Z} \]

(1) FRAC
\[ z = 1.71 \]

(2) FRAC
\[ z = 1 \]

(3) INT
\[ z = 0 \]

(4) INT
\[ z = 1 \]

\[ x_1 \leq 0 \quad x_1 \geq 1 \]
Branch-and-Bound and Cuts (Branch-and-Cut)

Homework Question 5:
Add two more Chátal-Gomory cuts so the LP relaxation with all cuts solves the MIP.

Hint:

\[
\begin{align*}
\frac{1}{3} (3x_1 + 2x_2 &\leq 6) \\
+ \quad &\frac{1}{3} (x_2 \leq 1) \\
\hline
\end{align*}
\]

and

\[
\begin{align*}
\frac{1}{2} (-2x_1 + x_2 &\leq 0) \\
+ \quad &\frac{1}{2} (x_2 \leq 1) \\
\hline
\end{align*}
\]

\[
\begin{align*}
\text{max } z := x_2 \\
3x_1 + 2x_2 &\leq 6 \\
x_2 &\leq \lfloor 1.71 \rfloor = 1 \\
-2x_1 + x_2 &\leq 0 \\
x_1, x_2 &\geq 0 \\
x_1, x_2 &\in \mathbb{Z}
\end{align*}
\]
No Enumeration = Keep Adding Cuts

- Number of tours for 49 cities $= \frac{48!}{2} \approx 10^{60}$
- Fastest supercomputer $\approx 10^{17}$ flops
- Assuming one floating point operation per tour:
  $> 10^{35}$ years $\approx 10^{25}$ times the age of the universe!
- How long does it take on an iphone?
  - $< 1$ sec ! Dantzig, Fulkerson and Johnson $\Rightarrow$ in 54’
  - This is how DFJ solved the problem by hand in 54’
  - In practice Branch-and-Cut is better.
- More details in Concord TSP App
  - Cutting plane tutorial for TSP
  - http://www.math.uwaterloo.ca/tsp/iphone/
Easy Problems : LP Relaxation Always Integral

Consequence of LP duality: Kőnig's theorem

• **Largest Matching**
  – Pick edges, at most one edge per node

• **Smallest Node Cover**
  – Pick nodes that touch all edges
Classes and Links

• Julia, JuMP and Optimization
  – http://www.juliaopt.org

• 15.053 Optimization Methods in Business Analytics
  – Modeling and computation
  – Instructor: James B. Orlin

• 18.453 Combinatorial Optimization
  – Theory and algorithms
  – Instructor: Michel Goemans
MIP & Daily Fantasy Sports
# Example Entry

**LINEUP**

<table>
<thead>
<tr>
<th>POS</th>
<th>PLAYER</th>
<th>OPP</th>
<th>FPPG</th>
<th>SALARY</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Jussi Jokinen</td>
<td>Fla@Anh</td>
<td>3.1</td>
<td>$5,300</td>
</tr>
<tr>
<td>C</td>
<td>Brandon Sutter</td>
<td>Pit@Van</td>
<td>3.0</td>
<td>$4,400</td>
</tr>
<tr>
<td>W</td>
<td>Nikolaj Ehlers</td>
<td>Wpg@Tor</td>
<td>3.9</td>
<td>$4,800</td>
</tr>
<tr>
<td>W</td>
<td>Daniel Sedin</td>
<td>Pit@Van</td>
<td>3.8</td>
<td>$6,400</td>
</tr>
<tr>
<td>W</td>
<td>Radim Vrbata</td>
<td>Pit@Van</td>
<td>3.4</td>
<td>$5,800</td>
</tr>
<tr>
<td>D</td>
<td>Brian Campbell</td>
<td>Fla@Anh</td>
<td>2.6</td>
<td>$4,100</td>
</tr>
<tr>
<td>D</td>
<td>Morgan Rielly</td>
<td>Wpg@Tor</td>
<td>3.5</td>
<td>$4,200</td>
</tr>
<tr>
<td>G</td>
<td>Corey Crawford P</td>
<td>StL@Chi</td>
<td>6.3</td>
<td>$7,800</td>
</tr>
<tr>
<td>UTIL</td>
<td>Blake Wheeler</td>
<td>Wpg@Tor</td>
<td>4.8</td>
<td>$7,200</td>
</tr>
</tbody>
</table>

Avg. Rem. / Player: $0  
Rem. Salary: $0
100% of the money in the top 20% lineups
26% of the money in the top 10 lineups (0.04%)
Building a Lineup
MIP Formulation

• L lineups : indexed by $l$

• 9 players per lineup: indexed by $p$

• Decision variables

\[ x_{pl} = \begin{cases} 
1, & \text{if player } p \text{ in lineup } l \\ 
0, & \text{otherwise} 
\end{cases} \]
Basic Feasibility

• Basic constraints:
  § 9 different players
  § Salary less than $50,000

\[
\sum_{p=1}^{N} c_p x_{pl} \leq $50,000, \quad \text{(budget constraint)}
\]

\[
\sum_{p=1}^{N} x_{pl} = 9, \quad \text{(lineup size constraint)}
\]

\[x_{pl} \in \{0, 1\}, \quad 1 \leq p \leq N.\]
Position Feasibility

• Between 2 and 3 centers
• Between 3 and 4 wingers
• Between 2 and 3 defensemen
• 1 goalie

Position constraints

\[ 2 \leq \sum_{p \in C} x_{pl} \leq 3, \quad \text{(center constraint)} \]

\[ 3 \leq \sum_{u \in W} x_{pl} \leq 4, \quad \text{(winger constraint)} \]

\[ 2 \leq \sum_{u \in D} x_{pl} \leq 3, \quad \text{(defencemen constraint)} \]

\[ \sum_{u \in G} x_{pl} = 1 \quad \text{(goalie constraint)} \]
Team Feasibility

- At least 3 different NHL teams

Team constraints

\[ t_i \leq \sum_{p \in T_i} x_{pl}, \quad \forall \ i \in \{1, \ldots, N_T\} \]

\[ \sum_{i=1}^{N_T} t_i \geq 3, \]

\[ t_i \in \{0, 1\}, \quad \forall \ i \in \{1, \ldots, N_T\}. \]
Maximize Points

- Forecasted points for player p: $f_p$

### Table 1: Points system for NHL contests in DraftKings.

<table>
<thead>
<tr>
<th>Score type</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal</td>
<td>3</td>
</tr>
<tr>
<td>Assist</td>
<td>2</td>
</tr>
<tr>
<td>Shot on Goal</td>
<td>0.5</td>
</tr>
<tr>
<td>Blocked Shot</td>
<td>0.5</td>
</tr>
<tr>
<td>Short Handed Point Bonus (Goal/Assist)</td>
<td>1</td>
</tr>
<tr>
<td>Shootout Goal</td>
<td>0.2</td>
</tr>
<tr>
<td>Hat Trick Bonus</td>
<td>1.5</td>
</tr>
<tr>
<td>Win (goalie only)</td>
<td>3</td>
</tr>
<tr>
<td>Save (goalie only)</td>
<td>0.2</td>
</tr>
<tr>
<td>Goal allowed (goalie only)</td>
<td>-1</td>
</tr>
<tr>
<td>Shutout Bonus (goalie only)</td>
<td>2</td>
</tr>
</tbody>
</table>

Points Objective Function

$$\sum_{p=1}^{N} f_p x_{pl}$$
## Lineup

<table>
<thead>
<tr>
<th>Projections</th>
<th>$9500</th>
<th>$2700</th>
<th>$4600</th>
<th>$3800</th>
<th>$4600</th>
<th>$6400</th>
<th>$5200</th>
<th>$5100</th>
<th>$8000</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>UTIL</td>
<td>D</td>
<td>D</td>
<td>C</td>
<td>C</td>
<td>W</td>
<td>W</td>
<td>G</td>
<td></td>
</tr>
</tbody>
</table>

23 points on average
Need > 38 points for a chance to win
Increase variance to have a chance
Structural Correlations : Teams
Structural Correlations: Lines

- Goal = 3 pt, assist = 2 pt
Structural Correlations : Lines = Stacking

• At least 1 complete line (3 players per line)
• At least 2 partial lines (at least 2 players per line)

1 complete line constraint

$$3v_i \leq \sum_{p \in L_i} x_{pl}, \quad \forall i \in \{1, \ldots , N_L\}$$

$$\sum_{i=1}^{N_L} v_i \geq 1$$

$$v_i \in \{0, 1\}, \quad \forall i \in \{1, \ldots , N_L\}.$$ 

2 partial lines constraint

$$2w_i \leq \sum_{p \in L_i} x_{pl}, \quad \forall i \in \{1, \ldots , N_L\}$$

$$\sum_{i=1}^{N_L} w_i \geq 2$$

$$w_i \in \{0, 1\}, \quad \forall i \in \{1, \ldots , N_L\}.$$
Structural Correlations:
Goalie Against Opposing Players
Structural Correlations:
Goalie Against Opposing Players

• No skater against goalie

No skater against goalie constraint

\[ 6x_{pl} + \sum_{q \in \text{Opponents}_p} x_{ql} \leq 6, \quad \forall p \in G \]
Good, but not great chance
Play many diverse Lineups

• Make sure lineup l has no more than $\gamma$ players in common with lineups 1 to l-1

Diversity constraint

$$\sum_{p=1}^{N} x_{p_k}^* x_{pl} \leq \gamma, \ k = 1, \ldots, l - 1$$
Were we able to do it?


200 lineups
Policy Change

200 lineups -> 100 lineups
Dear Mr. Hunter, Dr. Vielma, & Dr. Zaman,

On behalf of The Greater Boston Food Bank (GBFB), I want to thank you for your recent gifts. Your $16,709.98 contribution will help our neighbors who struggle to have enough to eat and will promote healthy lives and communities in eastern Massachusetts. Generous and dedicated individuals like you have enabled GBFB to progress in our mission to End Hunger Here and work toward

Your generous donations of $1,500.00 and $15,209.98, received on Wednesday, November 25, 2015, and Tuesday, May 3, 2016, respectively, will provide more than 50,000 nutritious meals to those at risk of hunger throughout eastern Massachusetts.

As one of the largest food banks in the country, GBFB is proud to be a leader in the conversation around food insecurity, using our voice to highlight the importance of nutrition. We aim to improve community health by distributing safe, nutritious food and we are dedicated to increasing the volume of fresh produce we provide. Last year, GBFB distributed a record-breaking 54 million pounds of food, with 25% being fresh produce, a number that will grow to 35% over the next few years. We will continue to ensure that at least 80% of our food is of the highest nutritional quality.

For single-parent families, veterans, low-income seniors, and community college students, the healthiest foods are often those too expensive to afford. GBFB works with our network of donors, food industry and corporate partners, and member agencies, to acquire and distribute high-quality and nutrient-rich items sourced from local and national manufacturers, retailers and growers. By ensuring access to nutrient-dense food, GBFB helps to provide the building blocks for healthy bodies and minds for those in need.

GBFB is committed to the efficient use of resources. We are proud that 91 cents of every dollar donated goes directly to hunger-relief efforts, helping us to earn a 4-star rating from Charity Navigator. With your help, we employ innovative warehouse technology and food industry expertise to maximize our efficiency to have the greatest impact.

You have my earnest gratitude for your partnership in helping to End Hunger Here.

Sincerely,

Suzanne J. Battit
Vice President of External Affairs and Advancement

We gratefully acknowledge your gift and confirm that no goods or services were provided in consideration of this charitable support. Please retain this letter for your tax records. Our Tax ID# is 04-2717782.

December 12, 2015

100 lineups
How can you do it?

Download Code from Github: https://github.com/dscotthunter/Fantasy-Hockey-IP-Code

Performance Time < 30 Minutes