Mixed Integer Programming (MIP) for Daily Fantasy Sports, Statistics and Marketing

Juan Pablo Vielma

Massachusetts Institute of Technology

AM/ES 121, SEAS, Harvard.
Boston, MA, November, 2016.
MIP & Daily Fantasy Sports
### Example Entry

![Draft Kings Logo](image)

![Hockey Sticks](image)

**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</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
Integer Programming Formulation

• We will make a bunch of lineups consisting of 9 players each

• Use an integer programming approach to find these lineups

Decision variables

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

• 9 different players
• Salary less than $50,000

Basic constraints

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

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

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

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

Position constraints

\[
\begin{align*}
2 & \leq \sum_{p \in C} x_{pl} \leq 3, \quad \text{(center constraint)} \\
3 & \leq \sum_{u \in W} x_{pl} \leq 4, \quad \text{(winger constraint)} \\
2 & \leq \sum_{u \in D} x_{pl} \leq 3, \quad \text{(defensemen constraint)} \\
\sum_{u \in G} x_{pl} &= 1 \quad \text{(goalie constraint)}
\end{align*}
\]
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

$$\sum_{p=1}^{N} f_p x_p l$$
Lineup

Projections:  5.4  2.5  3.4  3.0  3.2  4.2  3.5  3.4  5.7
$9500 $2700 $4600 $3800 $4600 $6400 $5200 $5100 $8000

W  UTIL  D  D  C  C  W  W  G

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 Skaters

- 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

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


200 lineups
Policy Change

200 lineups -> 100 lineups
Were we able to continue it?

December 12, 2015

The Greater Boston FOOD BANK

> $15K

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.
How can you do it?

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

Performance Time
< 30 Minutes

Runtime (minutes)

CBC
GLPK
Gurobi

Solver
MIP and Statistics: Inference for the Chilean Earthquake
The 2010 Chilean Earthquake

Case study: impact of an earthquake on educational outcomes
6th Strongest in Recorded History (8.8)
Impact on Educational Achievement? PSU = SAT
Earthquake Intensity + Great Demographic Info

Case study: impact of an earthquake on educational outcomes

José R. Zubizarreta (Columbia)

Matching using Integer Programming

11/14/2016 10 / 30
Randomized experiment

• Treatment / control have similar characteristics (covariates).
Covariate Balance Important for Inference

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Dose 1</th>
<th>Dose 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>462</td>
<td>462</td>
</tr>
<tr>
<td>Female</td>
<td>538</td>
<td>538</td>
</tr>
<tr>
<td>School SES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Mid-low</td>
<td>327</td>
<td>327</td>
</tr>
<tr>
<td>Medium</td>
<td>294</td>
<td>294</td>
</tr>
<tr>
<td>Mid-high</td>
<td>189</td>
<td>189</td>
</tr>
<tr>
<td>High</td>
<td>115</td>
<td>115</td>
</tr>
<tr>
<td>Mother’s education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>335</td>
<td>335</td>
</tr>
<tr>
<td>Secondary</td>
<td>426</td>
<td>426</td>
</tr>
<tr>
<td>Technical</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>College</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>Missing</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

Case study: impact of an earthquake on educational outcomes

Intensity of the earthquake in peak ground acceleration (PGA)
Observational Study: e.g. After Earthquake

- Treatment / control can have different characteristics.

Solution = Matching?
Matching

Treated Units: $\mathcal{T} = \{t_1, \ldots, t_T\}$

Control Units: $\mathcal{C} = \{c_1, \ldots, c_C\}$

Observed Covariates: $\mathcal{P} = \{p_1, \ldots, p_P\}$

Covariate Values: $x^t = (x^t_p)_{p \in \mathcal{P}}, \quad t \in \mathcal{T}$

$x^c = (x^c_p)_{p \in \mathcal{P}}, \quad c \in \mathcal{C}$
Nearest Neighbor Matching

\[
\begin{align*}
\text{minimize } & \sum_{t \in \mathcal{T}} \sum_{c \in \mathcal{C}} \delta_{t,c} m_{t,c} \\
\text{subject to } & \sum_{c \in \mathcal{C}} m_{t,c} = 1, \ t \in \mathcal{T} \\
& \sum_{t \in \mathcal{T}} m_{t,c} \leq 1, \ c \in \mathcal{C} \\
& 0 \leq m_{t,c} \leq 1, \ m_{t,c} \in \{0, 1\}, \ t \in \mathcal{T}, \ c \in \mathcal{C}
\end{align*}
\]

- e.g. \( \delta_{t,c} = \|x^t - x^c\|_2 \)
- Easy to solve
Maximum Cardinality Matching

\[
\begin{align*}
\text{max} & \quad \sum_{t \in T} \sum_{c \in C} m_{t,c} \\
\text{s.t.} & \quad \sum_{t \in T} m_{t,c} \leq 1, \quad \forall c \in C \\
& \quad \sum_{c \in C} m_{t,c} \leq 1, \quad \forall t \in T \\
& \quad \sum_{t \in T_{p,k}} \sum_{c \notin C_{p,k}} m_{t,c} = \sum_{t \notin T_{p,k}} \sum_{c \in C_{p,k}} m_{t,c} \\
& \quad \forall p \in P, k \in \mathcal{K}(p) \\
& \quad m_{t,c} \in \{0, 1\} \\
& \quad \forall t \in T, \; c \in C.
\end{align*}
\]

- Very hard to solve ( and very hard to understand! )
Advanced Maximum Cardinality Matching

\[
\text{max } \sum_{t \in \mathcal{T}} x_t \\
\text{s.t.} \\
\sum_{t \in \mathcal{T}} x_t = \sum_{c \in \mathcal{C}} y_c, \\
\sum_{t \in \mathcal{T}_{p,k}} x_t = \sum_{c \in \mathcal{C}_{p,k}} y_c, \quad \forall p \in \mathcal{P}, \quad k \in \mathcal{K}(p) \\
x_t \in \{0, 1\} \quad \forall t \in \mathcal{T} \\
y_c \in \{0, 1\} \quad \forall c \in \mathcal{C}.
\]

- Matching without matching variables
- Easy to solve: Small, but inherits matching properties

\[
\mathcal{K}(p) = \{x^c_p\}_{c \in \mathcal{P}} \cup \{x^t_p\}_{t \in \mathcal{T}} \\
\mathcal{C}_{p,k} = \{c \in \mathcal{C} : x^c_p = k\} \\
\mathcal{T}_{p,k} = \{t \in \mathcal{T} : x^t_p = k\}
\]
<table>
<thead>
<tr>
<th>Feature</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIMCE school (decile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIMCE student (decile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPA ranking (decile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attendance (decile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catholic school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High SES school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid–High SES school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid SES school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid–Low SES school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public School</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Voucher School</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

José R. Zubizarreta (Columbia)
Matching using Integer Programming
11/14/2016 18 / 30
Can Also do Multiple Doses

- Dose
  1. No quake
  2. Medium quake
  3. Bad quake

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Dose 1</th>
<th>Dose 2</th>
<th>Dose 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>462</td>
<td>462</td>
<td>462</td>
</tr>
<tr>
<td>Female</td>
<td>538</td>
<td>538</td>
<td>538</td>
</tr>
<tr>
<td>School SES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>75</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>Mid-low</td>
<td>327</td>
<td>327</td>
<td>327</td>
</tr>
<tr>
<td>Medium</td>
<td>294</td>
<td>294</td>
<td>294</td>
</tr>
<tr>
<td>Mid-high</td>
<td>189</td>
<td>189</td>
<td>189</td>
</tr>
<tr>
<td>High</td>
<td>115</td>
<td>115</td>
<td>115</td>
</tr>
<tr>
<td>Mother’s education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>335</td>
<td>335</td>
<td>335</td>
</tr>
<tr>
<td>Secondary</td>
<td>426</td>
<td>426</td>
<td>426</td>
</tr>
<tr>
<td>Technical</td>
<td>114</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>College</td>
<td>114</td>
<td>114</td>
<td>114</td>
</tr>
<tr>
<td>Missing</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>
Relative (To no Quake) Attendance Impact

Case study results

Pair differences in outcomes (w.r.t. dose 1): attendance

José R. Zubizarreta (Columbia)
Relative (To no Quake) PSU Score Impact

![Boxplot comparing PSU scores for medium and bad quakes across different dose levels.]

- **Medium quake**
  - Doses: -300, -200, -100, 0, 100, 200, 300
  - PSU Scores: 2, 3, 4, 5

- **Bad quake**
  - Doses: -300, -200, -100, 0, 100, 200, 300
  - PSU Scores: 2, 3, 4, 5, 6, 7, 8, 9, 10

José R. Zubizarreta (Columbia)
Matching using Integer Programming
11/14/2016 29/30
MIP and Marketing: Chewbacca or BB-8?
## Adaptive Preference Questionnaires

<table>
<thead>
<tr>
<th>Feature</th>
<th>SX530</th>
<th>RX100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zoom</td>
<td>50x</td>
<td>3.6x</td>
</tr>
<tr>
<td>Prize</td>
<td>$249.99</td>
<td>$399.99</td>
</tr>
<tr>
<td>Weight</td>
<td>15.68 ounces</td>
<td>7.5 ounces</td>
</tr>
<tr>
<td>Prefer</td>
<td>✓</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>TG-4</th>
<th>Galaxy 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterproof</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Prize</td>
<td>$249.99</td>
<td>$399.99</td>
</tr>
<tr>
<td>Weight</td>
<td>7.36 lb</td>
<td>7.5 lb</td>
</tr>
<tr>
<td>Prefer</td>
<td>□</td>
<td>✓</td>
</tr>
</tbody>
</table>

---

We recommend:
Choice-based Conjoint Analysis (CBCA)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Chewbacca</th>
<th>BB-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wookiee</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Droid</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Blaster</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>I would buy toy</td>
<td>☑</td>
<td>☐</td>
</tr>
</tbody>
</table>

Product Profile: $x^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = x^2$
Preference Model and Geometric Interpretation

• Utilities for 2 products, d features, logit model

\[ U_1 = \beta \cdot x^1 + \epsilon_1 = \sum_{i=1}^{d} \beta_i x_i^1 + \epsilon_1 \]
\[ U_2 = \beta \cdot x^2 + \epsilon_2 = \sum_{i=1}^{d} \beta_i x_i^2 + \epsilon_2 \]

• Utility maximizing customer
  – Geometric interpretation of preference for product 1 without error

\[ x^1 \geq x^2 \iff U_1 \geq U_2 \]
Next Question = Minimize (Expected) Volume

Good estimator for $\beta$?  Center of ellipsoid $\beta_0$

- [✓] 1. Choice balance
- [✗] 2. Postchoice symmetry
With Error = Volume of Ellipsoid \( f(x^1, x^2) \)
Rules of Thumb Still Good For Ellipsoid Volume

\[(\beta - \mu)' \cdot \Sigma^{-1} \cdot (\beta - \mu) \leq r\]

- **Choice balance:**
  - Minimize distance to center
    \[\mu \cdot (x^1 - x^2)\]

- **Postchoice symmetry:**
  - Maximize variance of question
    \[(x^1 - x^2)' \cdot \Sigma \cdot (x^1 - x^2)\]
“Simple” Formula for Expected Volume

- Expected Volume = Non-convex function $f(d, v)$ of distance:
  \[ d := \mu \cdot (x^1 - x^2) \]
- Variance:
  \[ v := (x^1 - x^2)' \cdot \sum (x^1 - x^2) \]

Can evaluate $f(d, v)$ with 1-dim integral 😊
Optimization Model

\[
\begin{align*}
\text{min} & \quad f(d, v) \quad \times \\
\text{s.t.} & \quad \mu \cdot (x^1 - x^2) = d \quad \checkmark \\
& \quad (x^1 - x^2)' \cdot \sum \cdot (x^1 - x^2) = v \quad \xmark \\
& \quad A^1 x^1 + A^2 x^2 \leq b \quad \checkmark \\
\text{Formulation trick:} & \quad \text{linearize } x^k_i \cdot x^l_j \quad x^1 \neq x^2 \quad \xmark \\
& \quad x^1, x^2 \in \{0, 1\}^n
\end{align*}
\]
Technique 2: Piecewise Linear Functions

- **D-efficiency** = Non-convex function $f(d, v)$
  
  **distance**: $d := \mu \cdot (x^1 - x^2)$
  
  **variance**: $v := (x^1 - x^2)' \cdot \sum \cdot (x^1 - x^2)$

Can evaluate $f(d, v)$ with 1-dim integral 😊

**Piecewise Linear Interpolation**

**MIP formulation**
Computational Performance

- Advanced formulations provide an computational advantage
- Advantage is significantly more important for free solvers
- State of the art commercial solvers can be significantly better that free solvers
- Still, free is free!
Summary and Main Messages

• Always choose Chewbacca!

• How to YOU use MIP / Optimization / OR / Analytics?
  – Study for the 2\textsuperscript{nd} midterm!
  – How about grad school down the river?
    • Masters of Business Analytics / OR
    • Ph.D. in Operations Research

https://orc.mit.edu
How Hard is MIP?
How hard is MIP: Traveling Salesman Problem?

“A computer would have to check all these possible routes to find the shortest one.”

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
MIP = Avoid 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?
  - Less than a second!
  - 4 iterations of cutting plane method!
  - Dantzig, Fulkerson and Johnson 1954 did it by hand!
  - For more info see tutorial in ConcordeTSP app
  - Cutting planes are the key for effectively solving (even NP-hard) MIP problems in practice.
50+ Years of MIP = Significant Solver Speedups

• Algorithmic Improvements (Machine Independent):
  – Commercial, but free for academic use

• (Reasonably) effective free / open source solvers:
  – GLPK, **COIN-OR (CBC)** and SCIP (only for non-commercial)

• Easy to use, fast and versatile modeling languages
  – Julia based JuMP modelling language
    – [http://julialang.org](http://julialang.org)
    – [http://www.juliaopt.org](http://www.juliaopt.org)
Technique 1: Binary Quadratic $x^1, x^2 \in \{0, 1\}^n$

$$(x^1 - x^2)' \cdot \sum \cdot (x^1 - x^2) = \nu$$

$X^l_{i,j} = x^l_i \cdot x^l_j \quad (l \in \{1, 2\}, \quad i, j \in \{1, \ldots, n\})$

$X^l_{i,j} \leq x^l_i, \quad X^l_{i,j} \leq x^l_j, \quad X^l_{i,j} \geq x^l_i + x^l_j - 1, \quad X^l_{i,j} \geq 0$

$W_{i,j} = x^1_i \cdot x^2_j$

$W_{i,j} \leq x^1_i, \quad W_{i,j} \leq x^2_j, \quad W_{i,j} \geq x^1_i + x^2_j - 1, \quad W_{i,j} \geq 0$

$$\sum_{i,j=1}^{n} (X^1_{i,j} + X^2_{i,j} - W_{i,j} - W_{j,i}) \sum_{i,j} = \nu$$
Technique 1: Binary Quadratic \( x^1, x^2 \in \{0, 1\}^n \)

\[
x^1 \neq x^2 \iff \|x^1 - x^2\|_2^2 \geq 1
\]

\[
X^l_{i,j} = x^l_i \cdot x^l_j \quad (l \in \{1, 2\}, \quad i, j \in \{1, \ldots, n\}) : \\
X^l_{i,j} \leq x^l_i, \quad X^l_{i,j} \leq x^l_j, \quad X^l_{i,j} \geq x^l_i + x^l_j - 1, \quad X^l_{i,j} \geq 0
\]

\[
W_{i,j} = x^1_i \cdot x^2_j : \\
W_{i,j} \leq x^1_i, \quad W_{i,j} \leq x^2_j, \quad W_{i,j} \geq x^1_i + x^2_j - 1, \quad W_{i,j} \geq 0
\]

\[
\sum_{i,j=1}^{n} (X^1_{i,j} + X^2_{i,j} - W_{i,j} - W_{j,i}) \geq 1
\]
2. Modeling Piecewise Linear Functions

An appropriate way of modeling a piecewise linear function $f : \mathcal{D} \rightarrow \mathbb{R}^n$ is to model its epigraph given by $\text{epi}(f) = \{(x, z) \in \mathcal{D} \times \mathbb{R} : f(x) \leq z\}$. For example, the epigraph of the function $f$ depicted in Figure 2(a) is shown in Figure 2(b).

For simplicity, we assume that the function domain $\mathcal{D}$ is bounded and $f$ is only used in a constraint of the form $f(x) \leq 0$ or as an objective function that is being minimized. We then need a model of $\text{epi}(f)$ since $f(x) \leq 0$ can be modeled as $(x, z) \in \text{epi}(f), z \leq 0$ and the minimization of $f$ can be achieved by minimizing $z$ subject to $(x, z) \in \text{epi}(f)$. For continuous functions we can also work with its graph, but modeling the epigraph will allow us to extend most of the results to some discontinuous functions and will simplify the analysis of formulation properties.

### Simple Formulation for Univariate Functions

\[
z = f(x)
\]

\[
\begin{pmatrix} x \\ z \end{pmatrix} = \sum_{j=1}^{5} \begin{pmatrix} d_j \\ f(d_j) \end{pmatrix} \lambda_j
\]

\[
1 = \sum_{j=1}^{5} \lambda_j, \quad \lambda_j \geq 0
\]

\[
y \in \{0, 1\}^4, \quad \sum_{i=1}^{4} y_i = 1
\]

\[
0 \leq \lambda_1 \leq y_1
\]

\[
0 \leq \lambda_2 \leq y_1 + y_2
\]

\[
0 \leq \lambda_3 \leq y_2 + y_3
\]

\[
0 \leq \lambda_4 \leq y_3 + y_4
\]

\[
0 \leq \lambda_5 \leq y_4
\]

Size = $O$ (# of segments)

Non-Ideal: Fractional Extreme Points
2. Modeling Piecewise Linear Functions

An appropriate way of modeling a piecewise linear function $f: \mathbb{D} \rightarrow \mathbb{R}^n$ is to model its epigraph given by $\text{epi}(f) := \{(x, z) \in \mathbb{D} \times \mathbb{R}: f(x) \leq z\}$. For example, the epigraph of the function in Figure 2(a) is depicted in Figure 2(b).

For simplicity, we assume that the function domain $\mathbb{D}$ is bounded and $f$ is only used in a constraint of the form $f(x) \leq 0$ or as an objective function that is being minimized. We then need a model of $\text{epi}(f)$ since $f(x) \leq 0$ can be modeled as $(x, z) \in \text{epi}(f), z \leq 0$ and the minimization of $f$ can be achieved by minimizing $z$ subject to $(x, z) \in \text{epi}(f)$.

For continuous functions we can also work with its graph, but modeling the epigraph will allow us to extend most of the results to some discontinuous functions and will simplify the analysis of formulation properties.

Advanced Formulation for Univariate Functions

$$z = f(x)$$

$$\begin{pmatrix} x \\ z \end{pmatrix} = \sum_{j=1}^{5} \begin{pmatrix} d_j \\ f(d_j) \end{pmatrix} \lambda_j$$

$$1 = \sum_{j=1}^{5} \lambda_j, \quad \lambda_j \geq 0$$

$$y \in \{0, 1\}^2$$

$$0 \leq \lambda_1 + \lambda_5 \leq 1 - y_1$$

$$0 \leq \lambda_3 \leq y_1$$

$$0 \leq \lambda_4 + \lambda_5 \leq 1 - y_2$$

$$0 \leq \lambda_1 + \lambda_2 \leq y_2$$

Size $= O(\log_2 \# \text{ of segments})$

Ideal: Integral Extreme Points