Plan for today

- Integer Programming
  - Power of integer variables in formulations

  - Fun with Formulations
    - Assignment Problem
    - Facility Location Problem

  - How to Solve?
    - Fixing Variables

$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$

INTEGER PROGRAMMING

… where we see the power of integer variables and algorithms to solve optimization problems involving integer variables

Integer Programming

An **integer programming (IP)** model is similar to a linear programming model, except all variables have to take integer values

- if all variables have to be either 0 or 1, the model is sometimes referred to **Binary IP (BIP)**
- if some variables have to be integers while others can take real values, the model is referred to **Mixed IP (MIP)**
- when both objective and constraints in an IP are linear, the IP model is also referred to as a **Integer Linear Program (ILP) or Mixed ILP (MILP)**
- If objectives or constraints are nonlinear then it is known as **non-linear IP**
- Non-linear IP is usually very difficult to solve (ILP is hard enough)

Linear Programming (LP)

$$\begin{gathered} \max\ Z = 3x_1 + 5x_2 \\ 3x_1 + 4x_2 \le 25 \\ x_1 + x_2 \le 20 \\ x_1 \ge 0, x_2 \ge 0 \end{gathered}$$

Integer Programming (IP)

$$\begin{gathered} \max\ Z = 3x_1 + 5x_2 \\ 3x_1 + 4x_2 \le 25 \\ x_1 + x_2 \le 20 \\ x_1, x_2 \text{ integers} \end{gathered}$$

Binary Integer Programming (BIP)

$$\begin{gathered} \max\ Z = 3x_1 + 5x_2 \\ 3x_1 + 4x_2 \le 25 \\ x_1 + x_2 \le 20 \\ x_1, x_2 \in \{0,1\} \end{gathered}$$

Mixed Integer Programming (MIP)

$$\begin{gathered} \max\ Z = 3x_1 + 5x_2 \\ 3x_1 + 4x_2 \le 25 \\ x_1 + x_2 \le 20 \\ x_1 \ge 0, x_2 \text{ integer} \end{gathered}$$

Recall:
- A function is **linear** if it satisfies additivity and proportionality
- Constraint $f(x) \le b_i$ is linear if f is linear

$$\begin{gathered} \max\ Z = 3x_1^2 + 5x_1x_2 \\ 3x_1 + 4x_2 \le 25 \\ x_1 + x_2 \le 20 \\ x_1 \ge 0, x_2 \text{ integer} \end{gathered}$$

Non-linear IP

POWER OF INTEGER VARIABLES

IP Formulations

… where we see the power of integer variables in mathematical modeling

Logical relationships, e.g., exclusion/inclusion

Examples:

1. Cake or Ice cream?

u_1: satisfaction of eating a cake, say $u_1 = 3$
u_2: satisfaction of eating ice cream, say $u_2 = 5$
Can eat at most one of the two
Goal: Maximize satisfaction

$$\max u_1 x_1 + u_2 x_2$$
$$x_1 + x_2 \leq 1$$
$$x_1, x_2 \in \{0,1\}$$

BILP

Logical relationship:
by saying yes to ice cream, you're saying no to cake (exclusion)

2. Course selection

Can select from *n* courses
First *m* courses are required core courses
Have to take at least three core courses
Cannot spend more than *h* hours per week
g_i: credit from course *i*
t_i: time to be spent per week on course *i*

Goal: maximize total credits

$$\max g_1 x_1 + g_2 x_2 + \cdots + g_n x_n$$
$$x_1 + x_2 + \cdots + x_m \geq 3$$
$$t_1 x_1 + t_2 x_2 + \cdots + t_n x_n \leq h$$
$$x_1, \ldots, x_n \in \{0,1\}$$

BILP

Logical relationship:
certain number of elements (3 core courses) must be a part of the solution (inclusion)

Fixed Investment Problem

B-Mobile is considering using exactly one of the two technologies to build a network

- LTE1 has a fixed cost $100k and
  can serve up to 1000 customers at a cost of $70 per customer
- LTE2 has a fixed cost $80k and
  can serve up to 2000 customers at a cost of $90 per customer
- each customer generates a revenue of $150

Question: how many customers to serve and by which technology

$y_i = 1$ if technology i is selected, $i = 1,2$
$x_i =$ number of customers served by technology i, $i = 1,2$

(obj. in $1000s)

$$\max 0.08\, x_1 + 0.06 x_2 - 100 y_1 - 80 y_2$$
$$x_1 \leq 1000 y_1$$
$$x_2 \leq 2000 y_2$$
$$y_1 + y_2 = 1$$
$$x_1, x_2 \geq 0$$
$$y_1, y_2 \in \{0,1\}$$
x_1, x_2 integers

MILP

Sequential decisions

Example:

Suppose that you have \$20m to invest in two projects. The return from each is as follows:

Project 1: 1st installment (max \$7m): -3%, 2nd installment (max \$8m): 10%, 3rd installment (max \$5m): 5%

Project 2: 1st installment (max \$10m): 8%, 2nd installment (max \$10m): 4%

Fine-print: Investment will be counted towards next installment only if the previous installment's complete max is invested

Question: how much should you invest in each project in each installment?

What if you try to formulate this as a LP?

project 1
x_1: 1st installment investment
x_2: 2nd installment investment
x_3: 3rd installment investment
project 2
x_4: 1st installment investment
x_5: 2nd installment investment

$$\max Z = -0.03x_1 + 0.1x_2 + 0.05x_3 + 0.08x_4 + 0.04x_5$$
$$x_1 \le 7$$
$$x_2 \le 8$$
$$x_3 \le 5$$
$$x_4 \le 10$$
$$x_5 \le 10$$
$$x_1 + x_2 + x_3 + x_4 + x_5 \le 20$$
$$x_1, x_2, x_3, x_4, x_5 \ge 0$$

Optimum will be $x_1^* = 0, x_2^* = 8, x_3^* = 2, x_4^* = 10, x_5^* = 0$

Wait a min…

Sequential decisions

Example:

Suppose that you have \$20m to invest in two projects. The return from each is as follows:

Project 1: 1st installment (max \$7m): -3%, 2nd installment (max \$8m): 10%, 3rd installment (max \$5m): 5%

Project 2: 1st installment (max \$10m): 8%, 2nd installment (max \$10m): 4%

Fine-print: Investment will be counted towards next installment only if the previous installment's complete max is invested

Question: how much should you invest in each project in each installment?

MILP formulation

$$\max -0.03x_1 + 0.1x_2 + 0.05x_3 + 0.08x_4 + 0.04x_5$$

$$x_1 \le 7$$

$$y_1 \le \frac{x_1}{7}, \quad x_2 \le 8y_1$$

if $x_1 < 7$, then $y_1 = 0$, and thus $x_2 = 0$ (so the second installment cannot start until after the max \$7m is invested in the first installment)

$$y_2 \le \frac{x_2}{8}, \quad x_3 \le 5y_2$$

similarly, if $x_2 < 8$, then $y_2 = 0$, and thus $x_3 = 0$

$$x_4 \le 10$$

$$y_3 \le \frac{x_4}{10}, \quad x_5 \le 10y_3$$

if $x_4 < 10$, then $y_3 = 0$, and thus $x_5 = 0$

$$x_1 + x_2 + x_3 + x_4 + x_5 \le 20$$

$$x_1, x_2, x_3, x_4, x_5 \ge 0$$

$$y_1, y_2, y_3 \in \{0,1\}$$

project 1
x_1: 1st installment investment
x_2: 2nd installment investment
x_3: 3rd installment investment
project 2
x_4: 1st installment investment
x_5: 2nd installment investment

Indicator variables:
project 1
y_1: indicates if 1st installment max is invested
y_2: indicates if 2nd installment max is invested
project 2
y_3: indicates if 1st installment max is invested

IP formulation: choice of constraints

Suppose that you are given a set of *m* constraints,

$$
\begin{aligned}
f_1(x_1, x_2, \dots, x_n) &\le d_1 \\
f_2(x_1, x_2, \dots, x_n) &\le d_2 \\
&\vdots \\
f_m(x_1, x_2, \dots, x_n) &\le d_m
\end{aligned}
$$

and your solution needs to satisfy at least *k* of them

Indicator variables:
y_i: indicates if ith constraint is ineffective

$$
\begin{aligned}
&f_1(x_1, x_2, \dots, x_n) \le d_1 + My_1 \\
&f_2(x_1, x_2, \dots, x_n) \le d_2 + My_2 \\
&\vdots \\
&f_m(x_1, x_2, \dots, x_n) \le d_m + My_m \\
&y_1 + y_2 + \cdots + y_m \le m - k \\
&y_1, \dots, y_m \in \{0,1\}
\end{aligned}
$$

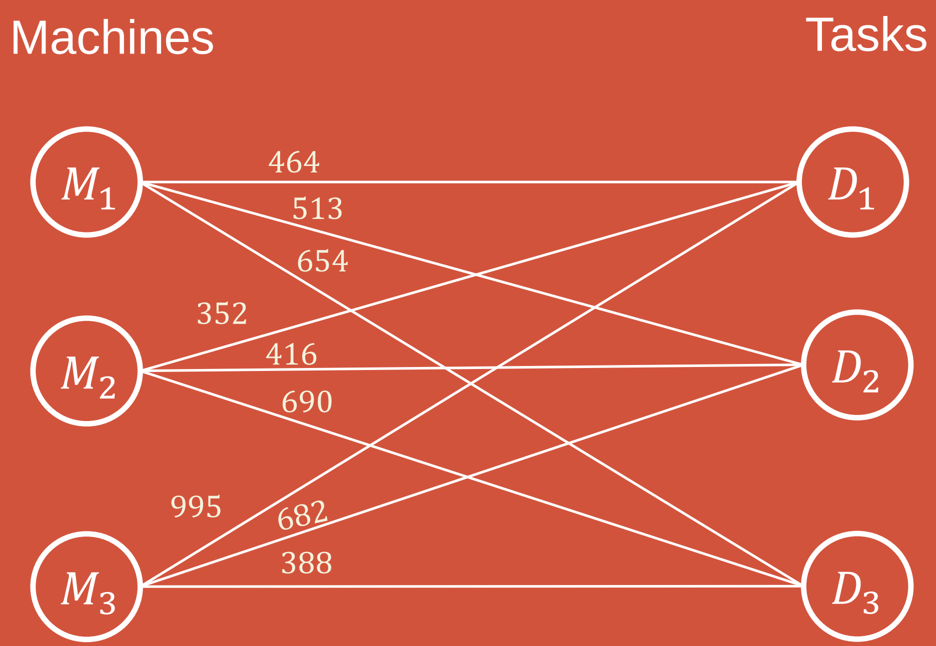
BIP

Interpretation: (Recall M is very large) If $y_i = 1$, the right-hand side of constraint i becomes very large, so the constraint is ineffective (as if it does not exist)

However you can make at most $m - k$ constraints ineffective

FUN WITH FORMULATIONS

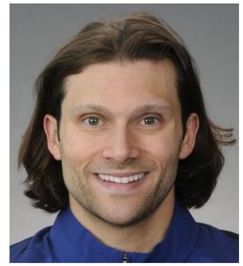
Flaunt (or test) your formulation skills!



ASSIGNMENT PROBLEM

Swimming: 4x100 relay medley

Operations … what??

	Adrian	Miller	Phelps	Murphy
Backstroke	$51.77s$	$51.99s$	$52.33s$	$51.85s$
Breaststroke	$58.86s$	$58.87s$	$58.91s$	$58.95s$
Butterfly	$51.59s$	$51.17s$	$51.14s$	$51.83s$
Freestyle	$47.85s$	$48.93s$	$48.01s$	$49.31s$

Question: Who swims which leg of the relay to minimize total time?

	S_1	S_2	S_3	S_4
L_1	51.77	51.99	52.33	51.85
L_2	58.86	58.87	58.91	58.95
L_3	51.59	51.17	51.14	51.83
L_4	47.85	48.93	48.01	49.31

Formulation

Step 1: identify decision variables

For $i = 1,2,3,4, j = 1,2,3,4$, let $x_{ij} = \begin{cases} 1 \text{ if leg i is assigned to swimmer j} \\ 0 \text{ otherwise} \end{cases}$

Step 2: determine the objective function

$$\begin{aligned} \min Z = {} & 51.77\, x_{11} + 51.99 x_{12} + 52.33\, x_{13} + 51.85\, x_{14} \\ & +58.86\, x_{21} + 58.87\, x_{22} + 58.91\, x_{23} + 58.95\, x_{24} \\ & +51.59\, x_{31} + 51.17\, x_{32} + 51.14\, x_{33} + 51.83 x_{34} \\ & +47.85\, x_{41} + 48.93\, x_{42} + 48.01 x_{43} + 49.31\, x_{44} \end{aligned}$$

Step 3: identify constraints

$$\left.\begin{aligned} x_{11} + x_{12} + x_{13} + x_{14} &= 1 \\ x_{21} + x_{22} + x_{23} + x_{24} &= 1 \\ x_{31} + x_{32} + x_{33} + x_{34} &= 1 \\ x_{41} + x_{42} + x_{43} + x_{44} &= 1 \end{aligned}\right\}$$

Leg constraints: Each leg is assigned exactly one swimmer

$$\left.\begin{aligned} x_{11} + x_{21} + x_{31} + x_{41} &= 1 \\ x_{12} + x_{22} + x_{32} + x_{42} &= 1 \\ x_{13} + x_{23} + x_{33} + x_{43} &= 1 \\ x_{14} + x_{24} + x_{34} + x_{44} &= 1 \end{aligned}\right\}$$

Swimmer constraints: Each swimmer is assigned to exactly one leg

$$x_{ij} \in \{0,1\}, i = 1,2,3,4, j = 1,2,3,4$$

Formulation

	S_1	S_2	S_3	S_4
L_1	51.77	51.99	52.33	51.85
L_2	58.86	58.87	58.91	58.95
L_3	51.59	51.17	51.14	51.83
L_4	47.85	48.93	48.01	49.31

$$\min Z = 51.77\,x_{11} + 51.99x_{12} + 52.33\,x_{13} + 51.85\,x_{14} + 58.86\,x_{21} + 58.87\,x_{22} + 58.91\,x_{23} + 58.95\,x_{24} + 51.59\,x_{31} + 51.17\,x_{32} + 51.14\,x_{33} + 51.83x_{34} + 47.85\,x_{41} + 48.93\,x_{42} + 48.01x_{43} + 49.31\,x_{44}$$

$$x_{11} + x_{12} + x_{13} + x_{14} = 1$$
$$x_{21} + x_{22} + x_{23} + x_{24} = 1$$
$$x_{31} + x_{32} + x_{33} + x_{34} = 1$$
$$x_{41} + x_{42} + x_{43} + x_{44} = 1$$
$$x_{11} + x_{21} + x_{31} + x_{41} = 1$$
$$x_{12} + x_{22} + x_{32} + x_{42} = 1$$
$$x_{13} + x_{23} + x_{33} + x_{43} = 1$$
$$x_{14} + x_{24} + x_{34} + x_{44} = 1$$

$$x_{ij} \in \{0,1\}, i = 1,2,3,4, j = 1,2,3,4$$

← BIP formulation

FACILITY LOCATION PROBLEM

Facility Location Problem

- Valmart has 5 shops in the vicinity of C-U
  - The warehouse serving these 5 shops is currently in Chicago and its operating costs have been high
  - Valmart wants to decide where to open two new warehouses that will serve these 5 shops
- Cost of opening a warehouse in each of the locations is known
- Operation costs for the next 10 years is also known:

Cost for row location warehouse to serve col location shop

Location	Champaign	Urbana	Savoy	Danville	Rantoul
Champaign	0	18	15	15	20
Urbana	18	0	21	25	19
Savoy	16	14	0	21	11
Danville	12	19	19	0	16
Rantoul	23	20	16	9	0

Questions: 1. Where should the warehouses be opened?
2. Which shop should be served by which warehouse?
(Each shop should be served by a warehouse)
… so that the total cost is minimized

Facility Location Problem

- Valmart has 5 shops in the vicinity of C-U
  - The warehouse serving these 5 shops is currently in Chicago and its operating costs have been high
  - Valmart wants to decide where to open <u>two</u> new warehouses that will serve these 5 shops
- Cost of opening a warehouse in each of the locations is known
- Operation costs for the next 10 years is also known:

Cost for row location warehouse to serve col location shop

Location
Champaign
Urbana
Savoy
Danville
Rantoul

Questions: 1. Where should the warehouses be opened?
2. Which shop should be served by which warehouse?
(Each shop should be served by a warehouse)
… so that the total cost is minimized

Formulation

Step 1: identify decision variables

$$y_i = \begin{cases} 1 \text{ if warehouse is opened at location } i \\ 0 \text{ otherwise} \end{cases}$$ for $i = 1, \ldots, 5$

$$x_{ij} = \begin{cases} 1 \text{ if warehouse at location } i \text{ serves shop at location } j \\ 0 \text{ otherwise} \end{cases}$$ for $i = 1, \ldots, 5,\ j = 1, \ldots, 5$

Step 2: determine the objective function

$$\min 25y_1 + 40y_2 + 31y_3 + 16y_4 + 15y_5 + \sum_{i=1}^{5} \sum_{j=1}^{5} c_{ij} x_{ij}$$

where c_{ij} is the cost of location i to serve location j

Step 3: identify constraints

$x_{ij} \in \{0,1\}$ for every $i = 1, \ldots, 5,\ j = 1, \ldots, 5$

$y_i \in \{0,1\}$ for every $i = 1, \ldots, 5$

$\sum_{i=1}^{5} y_i = 2$

$\sum_{i=1}^{5} x_{ij} = 1$ for every $j = 1, \ldots, 5$ (Each location is served by exactly one warehouse)

$x_{ij} \le y_i$ for every $i = 1, \ldots, 5,\ j = 1, \ldots, 5$ (Shop at location j can be served by warehouse at location i only if the warehouse at location i is open)

Formulation

BILP formulation

$$\min 25y_1 + 40y_2 + 31y_3 + 16y_4 + 15y_5 + \sum_{i=1}^{5}\sum_{j=1}^{5} c_{ij}x_{ij}$$

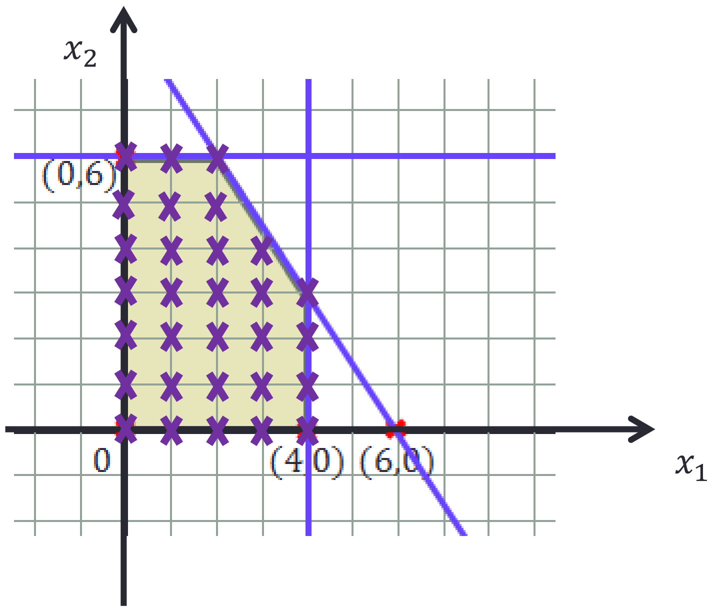
$\sum_{i=1}^{5} x_{ij} = 1$ for every $j = 1, \dots, 5$
$x_{ij} \leq y_i$ for every $i = 1, \dots, 5$, $j = 1, \dots, 5$
$\sum_{i=1}^{5} y_i = 2$
$x_{ij} \in \{0,1\}$ for every $i = 1, \dots, 5$, $j = 1, \dots, 5$
$y_i \in \{0,1\}$ for every $i = 1, \dots, 5$

IP: HOW TO SOLVE?

Caution: IPs are difficult to solve

Solving IP: LP solving techniques no longer work

IP has a smaller number of feasible solutions than the corresponding LP



If the feasible region is bounded, an IP has a finite number of feasible solutions, but that number can still be too large to check

By imposing the additional integer constraint, the feasible region has been reduced, but the problem structure has also been destroyed

Solving IP: rounding is usually not a good idea

IP:

$$\max Z = 2x_1 + x_2$$
$$3x_1 + 2x_2 \leq \frac{11}{2}$$
$$x_1, x_2 \geq 0$$
$$x_1, x_2 \in \text{integers}$$

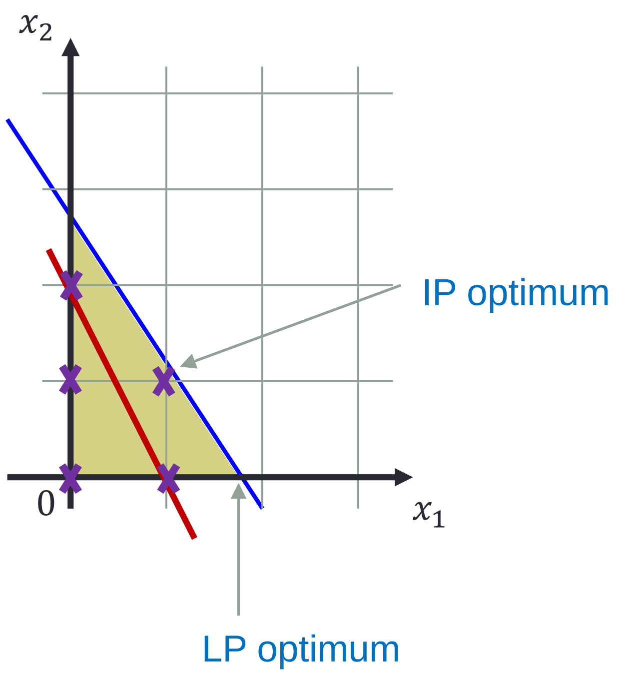
$$x_1^* = 1, x_2^* = 1, Z^* = 3$$

By dropping the integer constraints, we obtain an LP:
since the constraints are relaxed, we call it the **LP-relaxation**

LP relaxation:

$$\max Z = 2x_1 + x_2$$
$$3x_1 + 2x_2 \leq \frac{11}{2}$$
$$x_1, x_2 \geq 0$$

$$x_1^* = \frac{11}{6}, x_2^* = 0, Z^* = \frac{11}{3}$$



If we round the LP optimum solution

Either $x_1 = 1, x_2 = 0$ whose obj value is only $\frac{1}{3}$ of the IP-optimum obj value

Or $x_1 = 2, x_2 = 0$ which is infeasible

Solving IP: small problems can be extremely easy

Example:

Suppose that we have a budget of $20 that can be invested in some of the following 3 projects

project	cost ($c)	return ($r)
1	7	8
2	10	11
3	6	6

Question: To maximize the total return, which projects should we invest in?

IP formulation

$$\max 8y_1 + 11y_2 + 6y_3$$
$$7y_1 + 10y_2 + 6y_3 \leq 20$$
$$y_1, y_2, y_3 \in \{0,1\}$$

y_1	y_2	y_3	Obj
0	0	0	0
0	0	1	6
0	1	0	11
0	1	1	17
1	0	0	8
1	0	1	14
1	1	0	19
1	1	1	Infeasible

Solving IP: large problems can be very difficult

- The same capital budget problem in more generality:
  Given: Returns $r_1, \dots, r_n$ on n contracts, budget B
  Question: pick a subset to maximize the profit

IP formulation

$$\begin{aligned} &\max r_1 y_1 + r_2 y_2 + \cdots + r_n y_n \\ &c_1 y_1 + \cdots + c_n y_n \leq B \\ &\quad y_1, \dots, y_n \in \{0,1\} \end{aligned}$$

- This is also called the "knapsack problem":

  i.e., given the size (budget) of a knapsack, select the most valuable set of items to carry

- Potential number of solutions to inspect is 2^n

Combinatorial Explosion!

$$n = 100 \Rightarrow 2^n = 1.3 \times 10^{30}$$

Fastest computer in the world: 16.32×10^{15} calculations per second

Assume (extremely generous): one calculation evaluates obj value of one solution

$$t = \frac{1.3 \times 10^{30}}{16.32 \times 10^{15}} = 7.9 \times 10^{13} \text{secs} = 2.19 \times 10^{10} \text{ hours} = 2.5 \text{ million years}$$

Solving IP is difficult

- There is no SIMPLEX-like theory to dramatically reduce the search space for the optimal solution

- If we ignore integer constraints, solve the LP first, and then round solutions to nearest integers, the resulting point could be
    - Infeasible or
    - Far from the optimum

- We cannot enumerate all possible solutions, which can take millions of years to complete

How to deal with IPs?

- Settle for an "approximately" optimum solution
  - Design algorithms to find a solution <u>with a guaranteed bound</u>
    - "algorithm's objective value will differ from the optimum by a factor of at most 2"

  Area: Approximation Algorithms

  - Design algorithms to find an approximate optimum <u>with high probability</u>
    - "algorithm's objective value will differ from the optimum by a factor of at most 2 with probability at least 0.9"

  Area: Randomized Algorithms

- To obtain exact optimum solutions
  - Find intelligent ways to prune the search space, so that we do not explore all 2^n possible solutions

  1. Pre-processing: Variable Fixing
  2. Branch and Bound

SOLVING IPS

1. Pre-processing: Variable Fixing

Some simple pre-processing can go a long way in reducing the search space

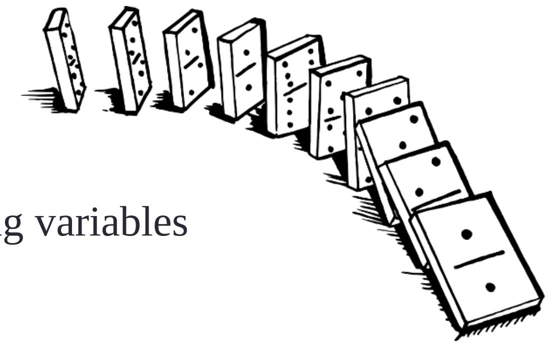
Fixing variables from constraints

- By fixing some variables, we can reduce the problem size
- For instance, if a problem has 20 binary variables (e.g., which of the 20 projects to invest in), then there are **2^{20} (1,048,576)** possible choices to check

  If we know one of these variables must be 1, then the number of possible choices reduces to **2^{19} (524, 288)**

  which is a tremendous reduction in the number of possibilities to check
- Some variables are easy to fix by just inspecting constraints
- Examples:
  1. if $4x_1 + 3x_2 + 3x_3 \leq 3$ and $x_1, x_2, x_3 \in \{0,1\}$, then $x_1 = 0$
  2. if $4x_1 - 2x_2 + x_3 \geq 2$ and $x_1, x_2, x_3 \in \{0,1\}$, then $x_1 = 1$
  3. if $4x_1 - 2x_2 + x_3 \leq -1$ and $x_1, x_2, x_3 \in \{0,1\}$, then $x_1 = 0$, $x_2 = 1$
  4. if $2x_1 - 2x_2 + x_3 \geq 3$ and $x_1, x_2, x_3 \in \{0,1\}$, then $x_1 = x_3 = 1$

“chain-reaction”



In some lucky cases, the entire problem can be solved by progressively fixing variables

Example:

$$\max Z = 3x_1 + 5x_2 - 2x_3 + x_4$$
$$x_1 + x_2 + 2x_4 \leq 1$$
$$x_1 - 2x_2 + x_4 \geq 0$$
$$x_2 + x_3 \geq 1$$
$$x_1 + 2x_2 - x_3 \geq 0$$
$$x_1, x_2, x_3, x_4 \in \{0,1\}$$

$x_1 + x_2 + 2x_4 \leq 1 \Rightarrow x_4^* = 0$

$x_1 - 2x_2 \geq 0 \Rightarrow x_2^* = 0$

$x_3 \geq 1 \Rightarrow x_3^* = 1$

$x_1 - 1 \geq 0 \Rightarrow x_1^* = 1$

$\Rightarrow \mathrm{Z}^* = 1$