Integer Linear Programming

(Wolsey, Chap. 1)

Formulations

Let A : $m \times n$ matrix

c : n-dimensional vector

b : m-dimensional vector

x : n-dimensional vector of variables

- If some but not all variables are restricted to be integer:

$$
\begin{aligned}
\max\ & cx + hy \\
& Ax + Gy \leq b \\
(\text{MILP}) \quad & x \geq 0 \ \text{integer} \\
& y \geq 0
\end{aligned}
$$

where y are continuos variables

Mixed Integer Linear Program (Problem)

- If all variables are integer:

$$\max \; cx$$

$$(\underline{ILP}) \qquad Ax \leq b$$

$$x \geq 0 \text{ integer}$$

<u>Integer Linear Program</u>

- If all variables are restricted in $\{0,1\}$:

$$\max \; cx$$

$$(\underline{BIP}) \qquad Ax \leq b$$

$$x \in \{0,1\}^n$$

<u>0-1 or Binar Integer Program</u>

- <u>Combinatorial optimization problems:</u>

$$(\underline{COP}) \qquad \min \sum_{j \in S} c_j$$

$$S \in F$$

where $N = \{1, \ldots n\}$ basic "objects"

c_j : cost of j, $j = 1, \ldots n$

F : feasible subsets of N

Example 1

0-1 Knapsack problem

$N = \{1, \dots n\}$ objects that can be put into a knapsack

c_j : profit of j, $j = 1, \dots n$

a_j : weight of j, $j = 1 \dots n$

b : capacity of the knapsack

COP:
$$\max \sum_{j \in S} c_j$$
$$\underbrace{S \subseteq N : \sum_{j \in S} a_j \leq b}_{\text{feasible set } F}$$

- Often a COP can be formulated as an ILP or a BIP:

e.g.
$$x_j = \begin{cases} 1 & \text{if } j \text{ is selected} \\ 0 & \text{otherwise} \end{cases} \qquad j = 1, \dots n$$

Knapsack formulation

$$\max \sum_{j=1}^{n} c_j \cdot x_j \qquad (KP_1)$$
$$F \begin{cases} \sum_{j=1}^{n} a_j \cdot x_j \leq b \\ x_j \in \{0, 1\} \qquad j = 1, \dots n \end{cases}$$

Example 2

Uncapacitated Facility Location (UFL)

Let $N = \{1, .. n\}$ set of potential depots

$M = \{1, .. m\}$ clients

f_j : fixed cost to open depot j, $j = 1, .. n$

c_{ij} : transportation cost if client i is served by depot j,

$i = 1, . m$, $j = 1, .. n$

* Which depots to open, and how to serve the clients in order to minimize the total (fixed + transportation) cost? *

Variables:

$$y_j = \begin{cases} 1 & \text{if depot } j \text{ is open} \\ 0 & \text{otherwise} \end{cases} \qquad j = 1, .. n$$

x_{ij} : fraction of the demand of i served by depot j, $i = 1, .. m$, $j = 1 .. n$

$$\min \sum_{i \in M} \sum_{j \in N} c_{ij} \cdot x_{ij} + \sum_{j \in N} f_j \cdot y_j$$

(UFL1)

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad i = 1, \dots m$$

* satisfaction of client demands *

$$\sum_{i \in M} x_{ij} \leq m \cdot y_j \qquad j = 1, \dots n$$

* linking constraints *

$$x_{ij} \geq 0 \qquad i \in M \;\; j \in N$$

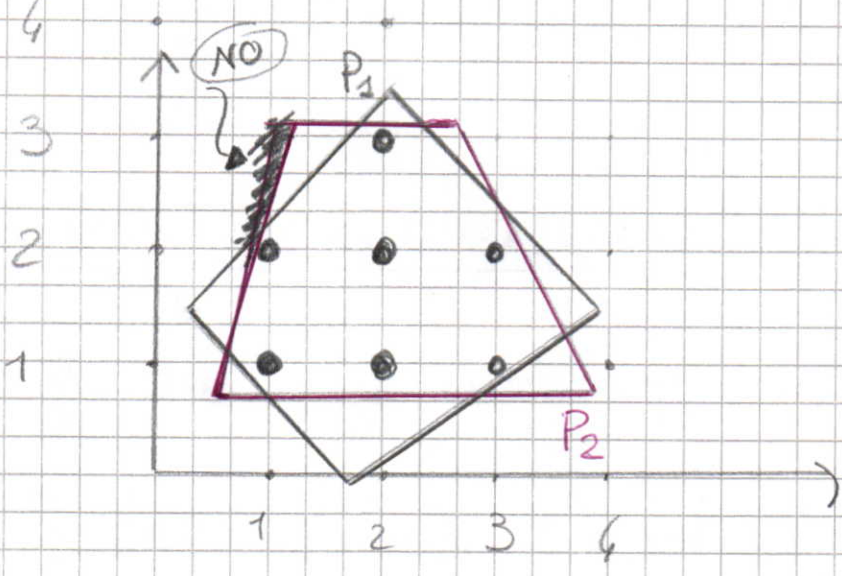
$$y_{ij} \in \{0, 1\} \qquad j \in N$$

Often there are <u>alternative formulations</u> for a given problem : some of them can be "better" than others ?

<u>Def</u> : A subset of $\mathbb{R}^n$ described by a finite set of linear constraints, e.g. $P = \{ x \in \mathbb{R}^n : Ax \leq b \}$ is a <u>polyhedron</u>

<u>Def</u> : A polyhedron $P \subseteq \mathbb{R}^{n+p}$ is a <u>formulation for a set</u> $X \subseteq \mathbb{Z}^n \times \mathbb{R}^p$ if and only if $X = P \cap (\mathbb{Z}^n \times \mathbb{R}^p)$

<u>e.g.</u> if $X = \{(1,1), (2,1), (3,1), (1,2), (2,2), (3,2), (2,3)\}$



P_1 and P_2 are two alternative formulations for X

An alternative formulation for UFL

$$\min \sum_{i \in M} \sum_{j \in N} c_{ij} x_{ij} + \sum_{j \in N} f_j \cdot y_j$$

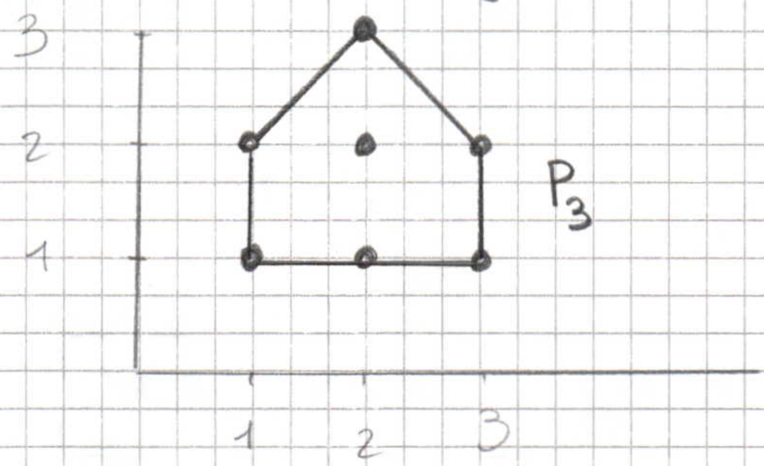
$$(UFL2) \quad \begin{aligned} &\sum_{j \in N} x_{ij} = 1 && i \in M \\ &\boxed{x_{ij} \leq y_j \qquad i \in M, j \in N} \\ &x_{ij} \geq 0 && i \in M, j \in N \\ &y_j \in \{0, 1\} && j \in N \end{aligned}$$

<u>In general</u>: there can be many different formulations of the same problem (e.g. UFL). How can we choose?

The geometry can help to find an answer.

Consider the same set X as before:

$$X = \{(1,1),(2,1),(3,1),(1,2),(2,2),(3,2),(2,3)\}$$



Formulation P_3 is "ideal". Why?

Because if we would solve the linear program $\max cx$

$$x \in P_3 \leftarrow \text{polyhedron}$$

then an optimal solution would be at an extreme point. But in this "ideal" polyhedron all the extreme points are integer! So, the optimal solution of the linear program ("easy"!) would be an optimal solution to the ILP!!

In other words: ILP would be solved via LP

Let us formalize this:

<u>Def</u>: given a set X in $\mathbb{R}^n$, the <u>convex hull of X</u> is:

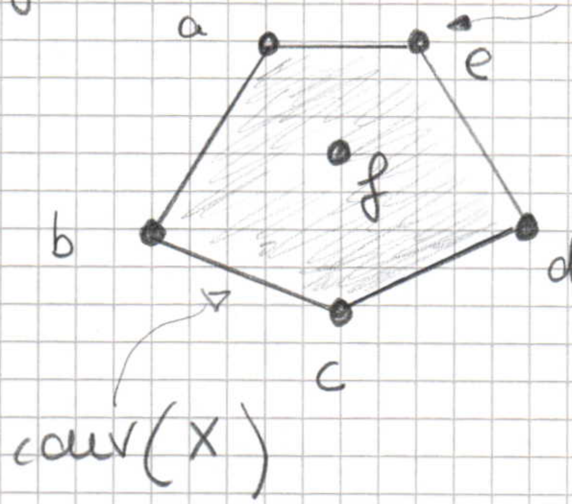
$$\mathrm{conv}(X) = \{ x : x = \sum_{i=1}^{t} \lambda_i \cdot x^i, \ \sum_{i=1}^{t} \lambda_i = 1,$$

$$\lambda_i \geq 0 \quad i = 1 \ldots t, \text{ over all finite subsets } \{x^1, \ldots x^t\} \text{ of } X \}$$

<u>Prop</u>: $\mathrm{conv}(X)$ is a polyhedron

<u>Prop</u>: the extreme points of $\mathrm{conv}(X)$ belong to X

<u>e.g.</u>



$\{a, b, c, d, e\} \subseteq X$ are the extreme points of $\mathrm{conv}(X)$

$X = \{a, b, c, d, e, f\}$

So, in order to solve

$$(ILP) \quad \max\ cx \quad \boxed{x \in X}$$

we could solve the equivalent (LP):

$$(LP) \quad \max\ cx \quad \boxed{x \in conv(X)}$$

(this ideal reduction also holds when X is unbounded and/or mixed integer)

<u>However</u>: this is only a theoretical approach, since we do not know $conv(X)$ (generally, there is no simple characterization)

Since $\boxed{X \subseteq conv(X) \subseteq P}$ for all formulations P for X ($conv(X)$ is the "best"), in practice we can compare two alternative formulations as follows:

Def : given a set X and two formulations P_1 and P_2 for X, P_1 is better than P_2 if $P_1 \subset P_2$

* This concept will be useful from an algorithmic point of view *

Examples

1) Knapsack set

Let $\mathbf{X} = \{(0,0,0,0),(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1),(0,1,0,1),(0,0,1,1)\}$

$$P_1 = \{ x \in \mathbb{R}^4 : 0 \leq x \leq 1, \; 83x_1 + 61x_2 + 49x_3 + 20x_4 \leq 100$$

is a formulation for $\mathbf{X}$.

(12)

$P_2 = \{ x \in \mathbb{R}^4 : 0 \leq x \leq 1,$

$$4x_1 + 3x_2 + 2x_3 + x_4 \leq 4$$

is also formulation for X.

Since

$$4x_1 + 3x_2 + 2x_3 + x_4 \leq 4$$

can be obtained from

$$83x_1 + 61x_2 + 49x_3 + 20x_4 \leq 100$$

as

$$:25 \quad \frac{83}{25}x_1 + \frac{61}{25}x_2 + \frac{49}{25}x_3 + \frac{20}{25}x_4 \leq 4$$

by rounding up the fractional coefficients,

then $P_2 \subset P_1$ (P_2 is better than P_1)

$$P_3 = \left\{ x \in \mathbb{R}^4 : \begin{array}{l} 4x_1 + 3x_2 + 2x_3 + x_4 \leq 4 \\ x_1 + x_2 + x_3 \leq 1 \\ x_1 + x_4 \leq 1 \\ 0 \leq x \leq 1 \end{array} \right\}$$

Also P_3 is a formulation for X.

Since $P_3 \subset P_2$, P_3 is better than P_2

It is possible to show that $P_3 = \text{conv}(X)$, and so it is ideal for X.

2) UFL

Let P_1 be the formulation ($\equiv$ polyhedron) corresponding to (UFL_1) and P_2 the one corresponding to (UFL_2).

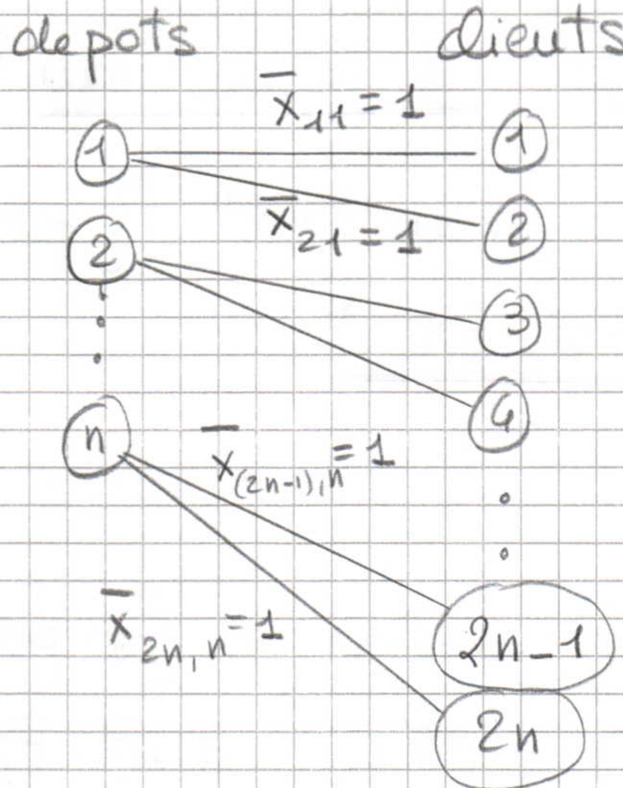
Since $x_{ij} \leq y_j$ for $i \in M, j \in N$, (in P_2)

imply $\sum_{i \in M} x_{ij} \leq m \cdot y_j$, for each j, (in P_1)

then $P_2 \subseteq P_1$

To show $P_2 \subset P_1$ we must find $\{\overline{x_{ij}}\}$ and $\{\overline{y_j}\}$ which belongs to $P_1 \setminus P_2$

e.g. assume $m = 2n$



$\bar{x}_{ij} = 0$ otherwise

$$\bar{y}_j = \frac{2}{m} \qquad j = 1, \dots n$$

This solution belongs to P_1, but lies in $P_1 \setminus P_2$:

P_2 is better than P_1

< skip the examples on >

ULS