

Methods and Models for Combinatorial Optimization

Introduction

Luigi De Giovanni, Marco Di Summa

Dipartimento di Matematica, Università di Padova

Mathematical Programming Models

A **mathematical programming model** describes the characteristics of the optimal solution of an optimization problem by means of **mathematical relations**.

- **Sets:** they group the elements of the system
- **Parameters:** the data of the problem, which represent the known quantities depending on the elements of the system.
- **Decision (or control) variables:** the unknown quantities, on which we can act in order to find different possible solutions to the problem.
- **Constraints:** mathematical relations that describe solution feasibility conditions (they distinguish acceptable combinations of values of the variables).
- **Objective function:** quantity to maximize or minimize, as a function of the decision variables.

Linear Programming models

Mathematical programming models where:

- objective function is a *linear* expression of the decision variables;
- constraints are a system of *linear* equations and/or inequalities.

Classification of linear programming models:

- **Linear Programming models (LP)**: all the variables can take real (\mathbb{R}) values;
- **Integer Linear Programming models (ILP)**: all the variables can take integer (\mathbb{Z}) values only;
- **Mixed Integer Linear Programming models (MILP)**: some variables can take real values and others can take integer values only.

Linearity limits expressiveness but allows faster solution techniques!

Linear Programming models

Mathematical programming models where:

- objective function is a *linear* expression of the decision variables;
- constraints are a system of *linear* equations and/or inequalities.

Classification of linear programming models:

- **Linear Programming models (LP)**: all the variables can take real (\mathbb{R}) values;
- **Integer Linear Programming models (ILP)**: all the variables can take integer (\mathbb{Z}) values only;
- **Mixed Integer Linear Programming models (MILP)**: some variables can take real values and others can take integer values only.

Linearity limits expressiveness but allows faster solution techniques!

Linear Programming models

Mathematical programming models where:

- objective function is a *linear* expression of the decision variables;
- constraints are a system of *linear* equations and/or inequalities.

Classification of linear programming models:

- **Linear Programming models (LP)**: all the variables can take real (\mathbb{R}) values;
- **Integer Linear Programming models (ILP)**: all the variables can take integer (\mathbb{Z}) values only;
- **Mixed Integer Linear Programming models (MILP)**: some variables can take real values and others can take integer values only.

Linearity limits expressiveness but allows faster solution techniques!

An LP model for a simple CO problem

Example

A perfume firm produces two new items by mixing three essences: rose, lily and violet. For each decaliter of perfume *one*, it is necessary to use 1.5 liters of rose, 1 liter of lily and 0.3 liters of violet. For each decaliter of perfume *two*, it is necessary to use 1 liter of rose, 1 liter of lily and 0.5 liters of violet. 27, 21 and 9 liters of rose, lily and violet (respectively) are available in stock. The company makes a profit of 130 euros for each decaliter of perfume *one* sold, and a profit of 100 euros for each decaliter of perfume *two* sold. The problem is to determine the optimal amount of the two perfumes that should be produced.

$$\max \quad 130 x_{one} + 100 x_{two}$$

$$s.t. \quad 1.5 x_{one} + x_{two} \leq 27$$

$$x_{one} + x_{two} \leq 21$$

$$0.3 x_{one} + 0.5 x_{two} \leq 9$$

$$x_{one}, x_{two} \geq 0$$

objective function

availability of rose

availability of lily

availability of violet

domains of the variables

An LP model for a simple CO problem

Example

A perfume firm produces two new items by mixing three essences: rose, lily and violet. For each decaliter of perfume *one*, it is necessary to use 1.5 liters of rose, 1 liter of lily and 0.3 liters of violet. For each decaliter of perfume *two*, it is necessary to use 1 liter of rose, 1 liter of lily and 0.5 liters of violet. 27, 21 and 9 liters of rose, lily and violet (respectively) are available in stock. The company makes a profit of 130 euros for each decaliter of perfume *one* sold, and a profit of 100 euros for each decaliter of perfume *two* sold. The problem is to determine the optimal amount of the two perfumes that should be produced.

max	130 x_{one}	+	100 x_{two}		objective function
s.t.	1.5 x_{one}	+	x_{two}	\leq 27	availability of rose
	x_{one}	+	x_{two}	\leq 21	availability of lily
	0.3 x_{one}	+	0.5 x_{two}	\leq 9	availability of violet
	x_{one}	,	x_{two}	\geq 0	domains of the variables

One possible modeling schema: optimal production mix

- **set** I : resource set $I = \{\text{rose}, \text{lily}, \text{violet}\}$
- **set** J : product set $J = \{\text{one}, \text{two}\}$
- **parameters** D_i : availability of resource $i \in I$ e.g. $D_{\text{rose}} = 27$
- **parameters** P_j : unit profit for product $j \in J$ e.g. $P_{\text{one}} = 130$
- **parameters** Q_{ij} : amount of resource $i \in I$ required for each unit of product $j \in J$ e.g. $Q_{\text{rose one}} = 1.5$, $Q_{\text{lily two}} = 1$
- **variables** x_j : amount of product $j \in J$ $x_{\text{one}}, x_{\text{two}}$

$$\begin{aligned} \max \quad & \sum_{j \in J} P_j x_j \\ \text{s.t.} \quad & \sum_{j \in J} Q_{ij} x_j \leq D_i \quad \forall i \in I \\ & x_j \in \mathbb{R}_+ \quad [\mathbb{Z}_+ \quad | \quad \{0, 1\}] \quad \forall j \in J \end{aligned}$$

The diet problem

Example

We need to prepare a diet that supplies at least 20 mg of proteins. 30 mg of iron and 10 mg of calcium. We have the opportunity of buying vegetables (containing 5 mg/kg of proteins, 6 mg/Kg of iron e 5 mg/Kg of calcium, cost 4 E/Kg), meat (15 mg/kg of proteins, 10 mg/Kg of iron e 3 mg/Kg of calcium, cost 10 E/Kg) and fruits (4 mg/kg of proteins, 5 mg/Kg of iron e 12 mg/Kg of calcium, cost 7 E/Kg). We want to determine the minimum cost diet.

min	$4x_V$	+	$10x_M$	+	$7x_F$		cost
s.t.	$5x_V$	+	$15x_M$	+	$4x_F$	≥ 20	proteins
	$6x_V$	+	$10x_M$	+	$5x_F$	≥ 30	iron
	$5x_V$	+	$3x_M$	+	$12x_F$	≥ 10	calcium
	x_V	,	x_M	,	x_F	≥ 0	domains of the variables

The diet problem

Example

We need to prepare a diet that supplies at least 20 mg of proteins. 30 mg of iron and 10 mg of calcium. We have the opportunity of buying vegetables (containing 5 mg/kg of proteins, 6 mg/Kg of iron e 5 mg/Kg of calcium, cost 4 E/Kg), meat (15 mg/kg of proteins, 10 mg/Kg of iron e 3 mg/Kg of calcium, cost 10 E/Kg) and fruits (4 mg/kg of proteins, 5 mg/Kg of iron e 12 mg/Kg of calcium, cost 7 E/Kg). We want to determine the minimum cost diet.

min	$4x_V$	+	$10x_M$	+	$7x_F$		cost
s.t.	$5x_V$	+	$15x_M$	+	$4x_F$	≥ 20	proteins
	$6x_V$	+	$10x_M$	+	$5x_F$	≥ 30	iron
	$5x_V$	+	$3x_M$	+	$12x_F$	≥ 10	calcium
	x_V	,	x_M	,	x_F	≥ 0	domains of the variables

One possible modeling schema: minimum cost covering

- **set** I : available resources $I = \{V, M, F\}$
- **set** J : request set $J = \{\textit{proteins}, \textit{iron}, \textit{calcium}\}$
- **parameters** C_i : unit cost of resource $i \in I$
- **parameters** R_j : requested amount of $j \in J$
- **parameters** A_{ij} : amount of request $j \in J$ satisfied by one unit of resource $i \in I$
- **variables** x_i : amount of resource $i \in I$

$$\begin{array}{ll}\min & \sum_{i \in I} C_i x_i \\ \text{s.t.} & \\ & \sum_{i \in I} A_{ij} x_i \geq D_j \quad \forall j \in J \\ & x_i \in \mathbb{R}_+ [\mathbb{Z}_+ \mid \{0, 1\}] \quad \forall i \in I\end{array}$$

The transportation problem

Example

A company produces refrigerators in three different factories (A, B and C) and need to move them to four stores (1, 2, 3, 4). The production of factories A, B and C is 50, 70 and 20 units, respectively. Stores 1, 2, 3 and 4 require 10, 60, 30 e 40 units, respectively. The costs in Euros to move one refrigerator from a factory to stores 1, 2, 3 and 4 are the following:

from A: 6, 8, 3, 4

from B: 2, 3, 1, 3

from C: 2, 4, 6, 5

The company asks us to formulate a minimum cost transportation plan.

One possible modeling schema: transportation

- **set** I : origins **factories** $I = \{A, B, C\}$
- **set** J : destinations **stores** $J = \{1, 2, 3, 4\}$
- **parameters** O_i : capacity of origin $i \in I$ **factory production**
- **parameters** D_j : request of destination $j \in J$ **store request**
- **parameters** C_{ij} : unit transp. cost from origin $i \in I$ to destination $j \in J$
- **variables** x_{ij} : amount to be transported from $i \in I$ to $j \in J$

$$\min \sum_{i \in I} \sum_{j \in J} C_{ij} x_{ij}$$

s.t.

$$\sum_{i \in I} x_{ij} \geq D_j \quad \forall j \in J$$

$$\sum_{j \in J} x_{ij} \leq O_i \quad \forall i \in I$$

$$x_{ij} \in \mathbb{R}_+ [\mathbb{Z}_+ \mid \{0, 1\}] \quad \forall i \in I \ j \in J$$

Fixed costs

Example

A supermarket chain has a budget W available for opening new stores. Preliminary analyses identified a set I of possible locations. Opening a store in $i \in I$ has a fixed cost F_i (land acquisition, other administrative costs etc.) and a variable cost C_i per 100 m² of store. Once opened, the store in i guarantees a revenue of R_i per 100 m². Determine the subset of location where a store has to be opened and the related size in order to maximize the total revenue, taking into account that at most K stores can be opened.

Modeling fixed costs: binary/boolean variables

- **set** I : potential locations
- **parameters** W , F_i , C_i , R_i , “large-enough” M
- **variables** x_i : size (in 100 m²) of the store in $i \in I$
- variables y_i : taking value 1 if a store is opened in $i \in I$ ($x_i > 0$), 0 otherwise

$$\max \sum_{i \in I} R_i x_i$$

s.t.

$$\sum_{i \in I} C_i x_i + F_i y_i \leq W$$

budget

$$x_i \leq M y_i \quad \forall i \in I$$

BigM constraint / relate x_i to y_i

$$\sum_{i \in I} y_i \leq K$$

max number of stores

$$x_i \in \mathbb{R}_+, y_i \in \{0, 1\} \quad \forall i \in I$$

Moving scaffolds between construction yards

A construction company has to move the scaffolds from three closing building sites (A, B, C) to three new building sites (1, 2, 3). The scaffolds consist of iron rods: in the sites A, B, C there are respectively 7000, 6000 and 4000 iron rods, while the new sites 1, 2, 3 need 8000, 5000 and 4000 rods respectively. The following table provide the cost of moving one iron rod from a closing site to a new site:

Costs (euro cents)	1	2	3
A	9	6	5
B	7	4	9
C	4	6	3

Trucks can be used to move the iron rods from one site to another site. Each truck can carry up to 10000 rods. Find a linear programming model that determine the minimum cost transportation plan, taking into account that:

- using a truck causes an additional cost of 50 euros;
- only 4 trucks are available (and each of them can be used only for a single pair of closing site and new site);
- the rods arriving in site 2 cannot come from both sites A and B;
- it is possible to rent a fifth truck for 65 euros (i.e., 15 euros more than the other trucks).

Moving scaffolds between construction yards: elements

Sets:

- I : closing sites (*origins*);
- J : news sites (*destinations*).

Parameters:

- C_{ij} : unit cost (per rod) for transportation from $i \in I$ to $j \in J$;
- D_i : number of rods available at origin $i \in I$;
- R_j : number of rods required at destination $j \in J$;
- F : fixed cost for each truck;
- N : number of trucks;
- L : fixed cost for the rent of an additional truck;
- K : truck capacity.

Decision variables:

- x_{ij} : number of rods moved from $i \in I$ to $j \in J$;
- y_{ij} : binary, values 1 if a truck from $i \in I$ to $j \in J$ is used, 0 otherwise.
- z : binary, values 1 if the additional truck is used, 0 otherwise.

Moving scaffolds between construction yards: MILP model

[Suggestion: compose transportation and fixed cost schemas]

$$\min \sum_{i \in I, j \in J} C_{ij} x_{ij} + F \sum_{i \in I, j \in J} y_{ij} + (L - F) z$$

$$\text{s.t.} \quad \sum_{i \in I} x_{ij} \geq R_j \quad \forall j \in J$$

$$\sum_{j \in J} x_{ij} \leq D_i \quad \forall i \in I$$

$$x_{ij} \leq K y_{ij} \quad \forall i \in I, j \in J$$

$$\sum_{i \in I, j \in J} y_{ij} \leq N + z$$

$$x_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J$$

$$y_{ij} \in \{0, 1\} \quad \forall i \in I, j \in J$$

$$z \in \{0, 1\}$$

Moving scaffolds between construction yards: MILP model



suggestion: compose transportation and fixed cost schemas]

$$\min \sum_{i \in I, j \in J} C_{ij} x_{ij} + F \sum_{i \in I, j \in J} y_{ij} + (L - F) z$$

$$\text{s.t.} \quad \sum_{i \in I} x_{ij} \geq R_j \quad \forall j \in J$$

$$\sum_{j \in J} x_{ij} \leq D_i \quad \forall i \in I$$

$$x_{ij} \leq K y_{ij} \quad \forall i \in I, j \in J$$

$$\sum_{i \in I, j \in J} y_{ij} \leq N + z$$

$$x_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J$$

$$y_{ij} \in \{0, 1\} \quad \forall i \in I, j \in J$$

$$z \in \{0, 1\}$$

Moving scaffolds between construction yards: variant 1

- truck capacity K does not guarantee that one truck is enough
 - ▶ how many trucks per (i,j) ? \Rightarrow variable $w_{ij} \in \mathbb{Z}_+$ instead of $y_{ij} \in \{0,1\}$

$$\min \quad \sum_{i \in I, j \in J} C_{ij} x_{ij} + F \sum_{i \in I, j \in J} w_{ij} + (L - F) z$$

$$\text{s.t.} \quad \sum_{i \in I} x_{ij} \geq R_j \quad \forall j \in J$$

$$\sum_{j \in J} x_{ij} \leq D_i \quad \forall i \in I$$

$$x_{ij} \leq K w_{ij} \quad \forall i \in I, j \in J$$

$$\sum_{i \in I, j \in J} y_{ij} \leq N + z$$

$$x_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J$$

$$w_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J$$

$$z \in \{0,1\}$$

Moving scaffolds between construction yards: variant 1

- truck capacity K does not guarantee that one truck is enough
 - how many trucks per (i,j) ? \Rightarrow variable $w_{ij} \in \mathbb{Z}_+$ instead of $y_{ij} \in \{0,1\}$

$$\min \sum_{i \in I, j \in J} C_{ij} x_{ij} + F \sum_{i \in I, j \in J} w_{ij} + (L - F) z$$



$$\text{s.t.} \quad \sum_{i \in I} x_{ij} \geq R_j \quad \forall j \in J$$

$$\sum_{j \in J} x_{ij} \leq D_i \quad \forall i \in I$$

$$x_{ij} \leq K w_{ij} \quad \forall i \in I, j \in J$$

$$\sum_{i \in I, j \in J} y_{ij} \leq N + z$$

$$x_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J$$

$$w_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J$$

$$z \in \{0,1\}$$

Moving scaffolds between construction yards: variant 2

- additional fixed cost A_i for loading operations in $i \in I$
 - ▶ does loading take place in i ? \Rightarrow variable $v_i \in \{0, 1\}$

$$\min \sum_{i \in I, j \in J} C_{ij} x_{ij} + F \sum_{i \in I, j \in J} w_{ij} + (L +$$

Control that if we have some

$$\begin{aligned} \text{s.t.} \quad & \sum_{i \in I} x_{ij} \geq R_j && \forall j \in J \\ & \sum_{j \in J} x_{ij} \leq D_i v_i && \forall i \in I \\ & x_{ij} \leq K w_{ij} && \forall i \in I, j \in J \\ & \sum_{i \in I, j \in J} y_{ij} \leq N + z \\ & x_{ij} \in \mathbb{Z}_+ && \forall i \in I, j \in J \\ & w_{ij} \in \mathbb{Z}_+ && \forall i \in I, j \in J \\ & v_i \in \{0, 1\} && \forall i \in I \\ & z \in \{0, 1\} \end{aligned}$$

Attention: try to preserve linearity!

Moving scaffolds between construction yards: variant 2

- additional fixed cost A_i for loading operations in $i \in I$
 - ▶ does loading take place in i ? \Rightarrow variable $v_i \in \{0, 1\}$

$$\min \sum_{i \in I, j \in J} C_{ij} x_{ij} + F \sum_{i \in I, j \in J} w_{ij} + (L - F)z + \sum_{i \in I} A_i v_i$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{i \in I} x_{ij} \geq R_j & \forall j \in J \\ & \sum_{j \in J} x_{ij} \leq D_i v_i & \forall i \in I \\ & x_{ij} \leq K w_{ij} & \forall i \in I, j \in J \\ & \sum_{i \in I, j \in J} y_{ij} \leq N + z \\ & x_{ij} \in \mathbb{Z}_+ & \forall i \in I, j \in J \\ & w_{ij} \in \mathbb{Z}_+ & \forall i \in I, j \in J \\ & v_i \in \{0, 1\} & \forall i \in I \\ & z \in \{0, 1\} \end{aligned}$$



Attention: try to preserve linearity!

Moving scaffolds between construction yards: variant 2



- additional fixed cost A_i for loading operations in $i \in I$
 - ▶ does loading take place in i ? \Rightarrow variable $v_i \in \{0, 1\}$

$$\begin{aligned} \min \quad & \sum_{i \in I, j \in J} C_{ij} x_{ij} + F \sum_{i \in I, j \in J} w_{ij} + (L - F)z + \sum_{i \in I} A_i v_i \\ \text{s.t.} \quad & \sum_{i \in I} x_{ij} \geq R_j \quad \forall j \in J \\ & \sum_{j \in J} x_{ij} \leq D_i v_i \quad \forall i \in I \\ & x_{ij} \leq K w_{ij} \quad \forall i \in I, j \in J \\ & \sum_{i \in I, j \in J} y_{ij} \leq N + z \\ & x_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J \\ & w_{ij} \in \mathbb{Z}_+ \quad \forall i \in I, j \in J \\ & v_i \in \{0, 1\} \quad \forall i \in I \\ & z \in \{0, 1\} \end{aligned}$$

Attention: try to preserve linearity!

Emergency location

A network of hospitals has to cover an area with the emergency service. The area has been divided into 6 zones and, for each zone, a possible location for the service has been identified. The average distance, in minutes, from every zone to each potential service location is shown in the following table.



	Loc. 1	Loc. 2	Loc. 3	Loc. 4	Loc. 5	Loc. 6
Zone 1	5	10	20	30	30	20
Zone 2	10	5	25	35	20	10
Zone 3	20	25	5	15	30	20
Zone 4	30	35	15	5	15	25
Zone 5	30	20	30	15	5	14
Zone 6	20	10	20	25	14	5

It is required each zone has an average distance from a emergency service of at most 15 minutes. The hospitals ask us for a service opening scheme that minimizes the number of emergency services in the area.

Emergency location: MILP model from covering schema

I set of potential locations ($I = \{1, 2, \dots, 6\}$).

x_i variables, values 1 if service is opened at location $i \in I$, 0 otherwise.

$$\begin{array}{llllllllll} \min & x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & + & x_6 \\ \text{s.t.} & & & & & & & & & & & \\ & x_1 & + & x_2 & & & & & & & & \geq 1 & \text{(cover zone 1)} \\ & x_1 & + & x_2 & & & & & & & + & x_6 & \geq 1 & \text{(cover zone 2)} \\ & & & & x_3 & + & x_4 & & & & & & \geq 1 & \text{(cover zone 3)} \\ & & & & x_3 & + & x_4 & + & x_5 & & & & \geq 1 & \text{(cover zone 4)} \\ & & & & & & x_4 & + & x_5 & + & x_6 & \geq 1 & \text{(cover zone 5)} \\ & & & x_2 & & & & + & x_5 & + & x_6 & \geq 1 & \text{(cover zone 6)} \\ & x_1 & , & x_2 & , & x_3 & , & x_4 & , & x_5 & , & x_6 & \in \{0, 1\} & \text{(domain)} \end{array}$$

TLC antennas location

A telephone company wants to install antennas in some sites in order to cover six areas. Five possible sites for the antennas have been detected. After some simulations, the intensity of the signal coming from an antenna placed in each site has been established for each area. The following table summarized these intensity levels:

	area 1	area 2	area 3	area 4	area 5	area 6
site A	10	20	16	25	0	10
site B	0	12	18	23	11	6
site C	21	8	5	6	23	19
site D	16	15	15	8	14	18
site E	21	13	13	17	18	22

Receivers recognize only signals whose level is at least 18. Furthermore, it is not possible to have more than one signal reaching level 18 in the same area, otherwise this would cause an interference. Finally, an antenna can be placed in site E only if an antenna is installed also in site D (this antenna would act as a bridge). The company wants to determine where antennas should be placed in order to cover the maximum number of areas.

TLC antennas location: MILP [from covering schema]

- I : **set** of sites for possible locations; J : set of areas;
- σ_{ij} : **parameter**, signal level of antenna in $i \in I$ received in $j \in J$;
- T : parameter, minimum signal level required;
- N : parameter, maximum number of non-interfering signals (here, $N = 1$);
- M_j : parameter, large enough, e.g., $M_j = \text{card}(\{i \in I : \sigma_{ij} \geq T\})$.
- x_i : binary **variable**, values 1 if an antenna is placed in $i \in I$, 0 otherwise;
- z_j : binary variable, values 1 if area $j \in J$ will be covered, 0 otherwise;

$$\begin{array}{ll}\max & \sum_{j \in J} 1 \\ \text{s.t.} & \sum_{i \in I: \sigma_{ij} \geq T} x_i \geq 1 \quad \forall j \in J \\ & \sum_{i \in I: \sigma_{ij} \geq T} x_i \leq N \quad \forall j \in J \\ & x_i \in \{0, 1\} \quad \forall i \in I \\ & z_j \in \{0, 1\} \quad \forall j \in J\end{array}$$

TLC antennas location: MILP [from covering schema]

- I : **set** of sites for possible locations; J : set of areas;
- σ_{ij} : **parameter**, signal level of antenna in $i \in I$ received in $j \in J$;
- T : parameter, minimum signal level required;
- N : parameter, maximum number of non-interfering signals (here, $N = 1$);
- M_j : parameter, large enough, e.g., $M_j = \text{card}(\{i \in I : \sigma_{ij} \geq T\})$.
- x_i : binary **variable**, values 1 if an antenna is placed in $i \in I$, 0 otherwise;
- z_j : binary variable, values 1 if area $j \in J$ will be covered, 0 otherwise;

$$\begin{array}{ll}\max & \sum_{j \in J} z_j \\ \text{s.t.} & \sum_{i \in I: \sigma_{ij} \geq T} x_i \geq z_j \quad \forall j \in J \\ & \sum_{i \in I: \sigma_{ij} \geq T} x_i \leq N + M_j(1 - z_j) \quad \forall j \in J \\ & x_i \in \{0, 1\} \quad \forall i \in I \\ & z_j \in \{0, 1\} \quad \forall j \in J\end{array}$$