Methods and Models for Combinatorial Optimization

Modeling by Linear Programming

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**. It provides a formulation and a basis for standard optimization algorithms.

- **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!

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.

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

One possible modeling schema: optimal production mix

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

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

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.

$$
\begin{array}{lrcrcrcrl}
\min & 4\,x_V & + & 10\,x_M & + & 7\,x_F & & & \text{cost} \\
s.t. & 5\,x_V & + & 15\,x_M & + & 4\,x_F & \geq & 20 & \text{proteins} \\
 & 6\,x_V & + & 10\,x_M & + & 5\,x_F & \geq & 30 & \text{iron} \\
 & 5\,x_V & + & 3\,x_M & + & 12\,x_F & \geq & 10 & \text{calcium} \\
 & x_V & , & x_M & , & x_F & \geq & 0 & \text{domains of the variables}
\end{array}
$$

One possible modeling schema: minimum cost covering

- set I: available resources $I = \{V, M, F\}$
- set J: request set $J = \{proteins, iron, calcium\}$
- 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 $i \in I$

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

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 30 units, respectively. Stores 1, 2, 3 and 4 require 20, 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:	4,	2,	1,	3
from C:	4,	2,	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 $j \in J$

$$
\begin{array}{lll}
\min & \displaystyle\sum_{i \in I} \sum_{j \in J} C_{ij} x_{ij} & \\
s.t. & & \\
& \displaystyle\sum_{i \in I} x_{ij} \geq D_j & \forall\, j \in J \\
& \displaystyle\sum_{j \in J} x_{ij} \leq O_i & \forall\, i \in I \\
& x_{ij} \in \mathbb{R}_+ \; [\, \mathbb{Z}_+ \mid \{0, 1\} \,] & \forall\, i \in I\; j \in J
\end{array}
$$

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 m^2 of store. Once opened, the store in i guarantees a revenue of R_i per 100 m^2. 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.

Fixed costs: notes

Modeling fixed costs: binary/boolean variables, non linear

- set I: potential locations
- parameters W, F_i, C_i, R_i
- variables x_i: size (in 100 m^2) 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

NON LINEAR formulation (correct, but we avoid it when possible [see next slide])

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

$s.t.$

$$\sum_{i \in I} C_i \, x_i \, y_i + F_i \, y_i \leq W \qquad \text{budget}$$

$$\sum_{i \in I} y_i \leq K \qquad \text{max number of stores}$$

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

Modeling fixed costs: binary/boolean variables (linear)

- set I: potential locations
- parameters W, F_i, C_i, R_i, "large-enough" M (e.g. $M = \arg\max_{i \in I}\{W/C_i\}$)
- variables x_i: size (in 100 m^2) 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: notes

Moving scaffolds between construction yards: MILP model

[Suggestion: compose transportation and fixed cost schemas]

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

$$\begin{array}{llcll}
\text{s.t.} & \displaystyle\sum_{i\in I} x_{ij} & \geq & R_j & \forall \;\; j \in J \\
& \displaystyle\sum_{j\in J} x_{ij} & \leq & D_i & \forall \;\; i \in I \\
& x_{ij} & \leq & K\, y_{ij} & \forall \;\; i \in I, j \in J \\
& \displaystyle\sum_{i\in I, j\in J} y_{ij} & \leq & N + z & \\
& y_{A2} + y_{B2} & \leq & 1 & \\
& x_{ij} & \in & \mathbb{Z}_+ & \forall \;\; i \in I, j \in J \\
& y_{ij} & \in & \{0,1\} & \forall \;\; i \in I, j \in J \\
& z & \in & \{0,1\} &
\end{array}$$

Moving scaffolds between construction yards: variant 1

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

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

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

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

Remark: try to preserve linearity!

Exercise: complete with the "logical" constraints $y_{A2} + y_{B2} \leq 1$ and try to generalize it.

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 an 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 od potential locations ($I = \{1, 2, ..., 6\}$).

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

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

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 two signals 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, $N = 2$);
- M_j: parameter, large enough, e.g., $M_j = card(\{i \in I : \sigma_{ij} \geq T\})$.
- 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;

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

Four Italian friends [from *La Settimana Enigmistica*]

Andrea, Bruno, Carlo and Dario share an apartment and read four newspapers: "La Repubblica", "Il Messaggero", "La Stampa" and "La Gazzetta dello Sport" before going out. Each of them wants to read all newspapers in a specific order. Andrea starts with "La Repubblica" for one hour, then he reads "La Stampa" for 30 minutes, "Il Messaggero" for two minutes and then "La Gazzetta dello Sport" for 5 minutes. Bruno prefers to start with "La Stampa" for 75 minutes; he then has a look at "Il Messaggero" for three minutes, then he reads "La Repubblica" for 25 minutes and finally "La Gazzetta dello Sport" for 10 minutes. Carlo starts with "Il Messaggero" for 5 minutes, then he reads "La Stampa" for 15 minutes, "La Repubblica" for 10 minutes and "La Gazzetta dello Sport" for 30 minutes. Finally, Dario starts with "La Gazzetta dello Sport" for 90 minutes and then he dedicates just one minute to each of "La Repubblica", "La Stampa" and "Il Messaggero" in this order. The preferred order is so important that each is willing to wait and read nothing until the newspaper that he wants becomes available. Moreover, none of them would stop reading a newspaper and resume later. By taking into account that Andrea gets up at 8:30, Bruno and Carlo at 8:45 and Dario at 9:30, and that they can wash, get dressed and have breakfast while reading the newspapers, what is the earliest time they can leave home together?

Four Italian friends: a Job-Shop Scheduling Problem (JSP)

- **Jobs**: **A**ndrea, **B**runo, **C**arlo, **D**ario [set I]
- **Machines**: "La **R**epubblica", "Il **M**essaggero", "La **S**tampa" and "La **G**azzetta dello Sport" [set K]
- **Processing times and order**:

  A: R (60) $\rightarrow$ S (30) $\rightarrow$ M (2) $\rightarrow$ G (5);

  B: S (75) $\rightarrow$ M (3) $\rightarrow$ R (25) $\rightarrow$ G (10);

  C: M (5) $\rightarrow$ S (15) $\rightarrow$ R (10) $\rightarrow$ G (30);

  D: G (90) $\rightarrow$ R (1) $\rightarrow$ S (1) $\rightarrow$ M (1);

  [param: D_{ik}, processing times]
  [param: $\sigma[i, \ell] \in K$, newspaper read by i in position ℓ)]
- **Release time**: A 8:30 – B 8:45 – C 8:45 – D 9:30. [param R_i]
- Objective: Minimize the **Makespan** (job-completion time)
- **No pre-emption**

Four Italian friends: a JSP, example

- **Processing times and order**:

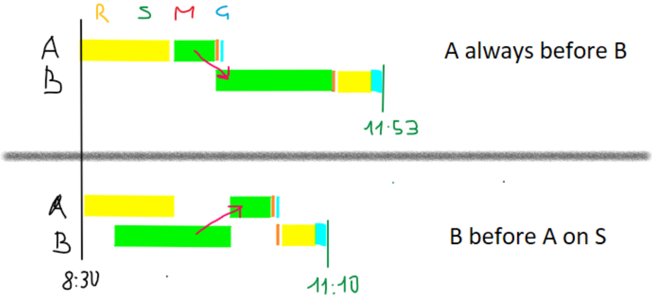
  A: R (60) → S (30) → M (2) → G (5);

  B: S (75) → M (3) → R (25) → G (10);

  C: M (5) → S (15) → R (10) → G (30);

  D: G (90) → R (1) → S (1) → M (1);

- **Release time**: A 8:30 – B 8:45 – C 8:45 – D 9:30.



LP model for JSP

- y: completion time (in minutes after 8:30);
- h_{ik}: start time (in minutes after 8:30) of $i \in I$ on $k \in K$;
- x_{ijk}: binary, 1 if $i \in I$ precedes $j \in I$ on $k \in K$, 0 otherwise.

$$
\begin{array}{lrll}
\min & & y & \\
\text{s.t.} & y & \geq h_{i\,\sigma[i,|K|]} + D_{i\,\sigma[i,|K|]} & \forall\, i \in I \\
& h_{i\,\sigma[i,1]} & \geq R_i & \forall\, i \in I \\
& h_{i\,\sigma[i,\ell]} & \geq h_{i\,\sigma[i,\ell-1]} + D_{i\,\sigma[i,\ell-1]} & \forall\, i \in I, \ell = 2 ... |K| \\
& h_{ik} & \geq h_{jk} + D_{jk} - M\, x_{ijk} & \forall\, k \in K, i \in I, j \in I : i \neq j \\
& h_{jk} & \geq h_{ik} + D_{ik} - M\,(1 - x_{ijk}) & \forall\, k \in K, i \in I, j \in I : i \neq j \\
& y & \in \mathbb{R}_+ & \\
& h_{ik} & \in \mathbb{R}_+ & \forall\, k \in K, i \in I \\
& x_{ijk} & \in \{0, 1\} & \forall\, k \in K, i \in I, j \in I : i \neq j
\end{array}
$$

Project scheduling in the boatyard industry

Constructing a boat requires the completion of the following operations :

Operations	**Duration**	**Precedences**
A	2	none
B	4	A
C	2	A
D	5	A
E	3	B,C
F	3	E
G	2	E
H	7	D,E,G
I	4	F,G

Some of the operations are alternative to each other. In particular, only one of B and C needs to be executed, and only one of F and G needs to be executed. Furthermore, if both C and G are executed, the duration of I increases by 2 days. The table also shows the precedences for each operation (i.e., operations that must be completed before the beginning of the new operation). For instance, H can start only after the completion of E, D and G (if G will be executed). Write a linear programming model that can be used to decide which operations should be executed in order to minimize the total duration of the construction of the boat.

Project scheduling in the boatyard industry: hints

$$
\begin{array}{lll}
\min & z & \\
s.t. & z \geq t_i & \forall i \in A...I \\
& t_A \geq d_A & \\
& t_B \geq t_A + d_B - M(1 - y_B) & \\
& t_C \geq t_A + d_C - M(1 - y_C) & \\
& t_D \geq t_A + d_D & \\
& t_E \geq t_B + d_E & \\
& t_E \geq t_C + d_E & \\
& t_F \geq t_E + d_F - M(1 - y_F) & \\
& t_G \geq t_A + d_G - M(1 - y_G) &
\end{array}
\qquad
\begin{array}{ll}
t_H \geq t_D + d_H & \\
t_H \geq t_E + d_H & \\
t_H \geq t_G + d_H & \\
t_I \geq t_F + d_I + 2y_{CG} & \\
t_I \geq t_G + d_I + 2y_{CG} & \\
y_B + y_C = 1 & \\
y_F + y_G = 1 & \\
y_C + y_G <= 1 + y_{CG} & \\
z, t_i \geq 0 & \forall i \in \{A...I\} \\
y_\cdot \in \{0, 1\} &
\end{array}
$$

where

- t_i completion time of operation $i \in \{A, B, C, D, E, F, G, H, I\}$;
- y_i 1 if operation $i \in \{B, C, F, G\}$ is executed, 0 otherwise;
- y_{CG} 1 if both C and G are executed, 0 otherwise;
- z completion time of the last operation;
- d_i parameter indicating the duration of operation i;
- M sufficiently large constant.

Exercise: write a more general model for generic sets of operations and precedence.

A (shift) covering problem

The pharmacy federation wants to organize the opening shifts on public holidays all over the region. The number of shifts is already decided, and the number of pharmacies open on the same day has to be as balanced as possible. Furthermore, every pharmacy is part of one shift only. For instance, if there are 12 pharmacies and the number of shifts is 3, every shift will consist of 4 pharmacies. Pharmacies and users are thought as concentrated in centroids (for instance, villages). For every centroid, we know the number of users and the set of pharmacies located in it. The distance between every ordered pair of centroids is also known. For the sake of simplicity, we ignore congestion problems and we assume that every user will go to the closest open pharmacy. The target is to determine the shifts so that the total distance covered by the users is minimized.

A (shift) covering problem: model 1

- y_{ik}: 1 if pharmacy $j \in P$ takes part in shift $k = 1 \ldots K$, 0 otherwise;
- z_{ijk}: 1 if centroid $i \in C$ uses pharmacy $j \in P$ during shift $k = 1 \ldots K$, 0 otherwise (notice: by optimality, z selects the nearest open pharmacy)

$$
\begin{array}{lll}
\min \displaystyle\sum_{k=1}^{K} \sum_{i \in C} \sum_{j \in P} D_{ij} z_{ijk} & & \text{(parameter } D_{ij}\text{: distance from } i \text{ to } j\text{)} \\
\text{s.t.} \displaystyle\sum_{k=1}^{K} y_{jk} = 1 & & \forall\, j \in P \\
\displaystyle\sum_{j \in P} z_{ijk} = 1 & & \forall\, i \in C, k = 1 \ldots K \\
x_{ijk} \leq y_{jk} & & \forall\, i \in C, j \in P, k = 1 \ldots K \\
\Big(\lfloor |P|/K \rfloor \leq \displaystyle\sum_{j \in P} y_{jk} \leq \lceil |P|/K \rceil & & \forall\, k = 1 \ldots K \Big) \\
z_{ijk}, y_{jk} \in \{0, 1\} & & \forall\, i \in C, j \in P, k = 1 \ldots K
\end{array}
$$

Notice: the model has a polynomial number of variables and constraints but suffers from symmetries, that is, the same "real" solution can be represented in many different ways, by giving different names (i.e. value of k) to the same shifts.

A (shift) covering problem: model 2

$\mathcal{P}$: set of all possible subsets of P (with balanced cardinality for balancing constraint)
$D(J)$: total distance covered by all users in C to reach the nearest pharmacy in $J \in \mathcal{P}$

- x_J: 1 if the set $J \in \mathcal{P}$ is selected as a shift, 0 otherwise;

$$
\begin{aligned}
\min \sum_{J \in \mathcal{P}} D_J x_J & \\
\text{s.t.} \sum_{J \in \mathcal{P}} x_J = K & \\
\sum_{J \in \mathcal{P} : j \in J} x_J = 1 & \qquad \forall\, j \in P \\
x_J \in \{0, 1\} & \qquad \forall J \in \mathcal{P}
\end{aligned}
$$

Notice: the model does not suffer from symmetries (a shift is directly determined by the defining subset), but has an exponential number of variables [we will see how to face this issue].

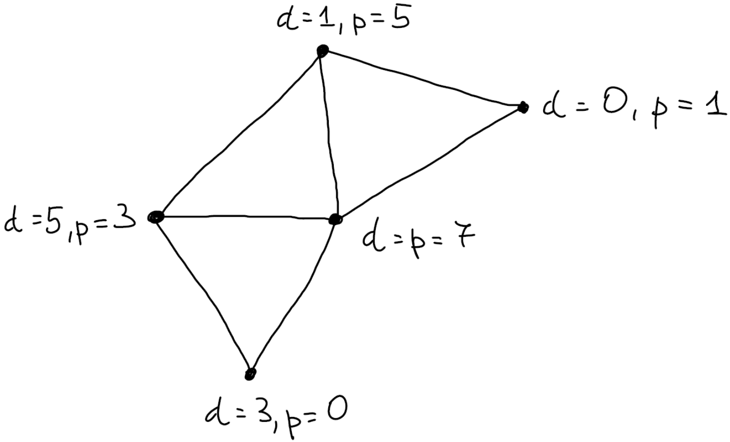
An energy flow problem

A company distributing electric energy has several power plants and distributing stations connected by wires. Each station i can:

- produce p_i kW of energy ($p_i = 0$ if the station cannot produce energy);
- distribute energy on a sub-network whose users have a total demand of d_i kW ($d_i = 0$ if the station serves no users);
- carry energy from/to different stations.

The wires connecting station i to station j have a maximum capacity of u_{ij} kW and a cost of c_{ij} euros for each kW carried by the wires. The company wants to determine the minimum cost distribution plan, under the assumption that the total amount of energy produced equals the total amount of energy required by all sub-networks.

Notes



Network flows models: single commodity

Parameters: u_{ij}, c_{ij} and

G $= (N, A)$, $N =$ power/distribution stations, $A =$ connections between stations

b_v $= d_v - p_v$, $v \in N$ [demand ($b_v > 0$)/supply (< 0)/transshipment ($= 0$) node]

Variables:

x_{ij} amount of energy to flow on arc $(i, j) \in A$

$$
\begin{array}{llcll}
\min & \displaystyle\sum_{(i,j)\in A} c_{ij} x_{ij} & & & \\
s.t. & \displaystyle\sum_{(i,v)\in A} x_{iv} - \sum_{(v,j)\in A} x_{vj} & = & b_v & \forall\, v \in N \\
 & x_{ij} & \leq & u_{ij} & \forall\, (i,j) \in A \\
 & x_{ij} & \in & \mathbb{R}_+ &
\end{array}
$$

Minimum Cost Network Flow Problem

An multi-type energy flow problem

A company distributing electric energy has several power and distributing stations connected by wires. Each station produces/distributes different kinds of energy. Each station i can:

- produce p_i^k kW of energy of type k (it may be $p_i^k = 0$);
- distribute energy of type k on a sub-network whose users have a total demand of d_i^k kW (it may be $d_i^k = 0$);
- carry energy from/to different stations.

Note that every station can produce and/or distribute different types of energy. The wires connecting station i to station j have a maximum capacity of u_{ij} kW, independently of the type of energy carried. The transportation cost depends both on the pair of stations (i, j) and the type of energy k, and is equal to c_{ij}^k euros for each kW. The company wants to determine the minimum cost distribution plan, under the assumption that, for each type of energy, the total amount produced equals the total amount of energy of the same type required by all sub-networks.

Network flows models: multi-commodity

Parameters: u_{ij}, c_{ij}^k, K (set of energy types or **commodities**) and

$G = (N, A)$, $N =$ power/distribution stations, $A =$ connections between stations

$b_v^k = d_v^k - p_v^k$, $v \in N$ [demand ($b_v^k > 0$)/supply (< 0)/transshipment ($= 0$) node]

Variables:

x_{ij}^k amount of energy of type k to flow on arc $(i, j) \in A$

$$\begin{array}{llcll}
\min & \sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k x_{ij}^k & & & \\
s.t. & \sum_{(i,v) \in A} x_{iv}^k - \sum_{(v,j) \in A} x_{vj}^k & = & b_v^k & \forall\, v \in N,\ \forall\, k \in K \\
 & \sum_{k \in K} x_{ij}^k & \leq & u_{ij} & \forall\, (i,j) \in A \\
 & x_{ij}^k & \in & \mathbb{R}_+ & \forall\, (i,j) \in A,\ \forall\, k \in K
\end{array}$$

Minimum Cost Network Multi-commodity Flow Problem