OPERATIONS RESEARCH

Degree in Management

3rd year

Academic Year 2023/2024

Linear Programming

Ana Catarina Nunes
Anabela Costa
Maria João Cortinhal
Maria João Lopes

1. Optimization

The meaning of **to optimize** depends on the nature of the problem being studied. For instance, to optimize can mean:

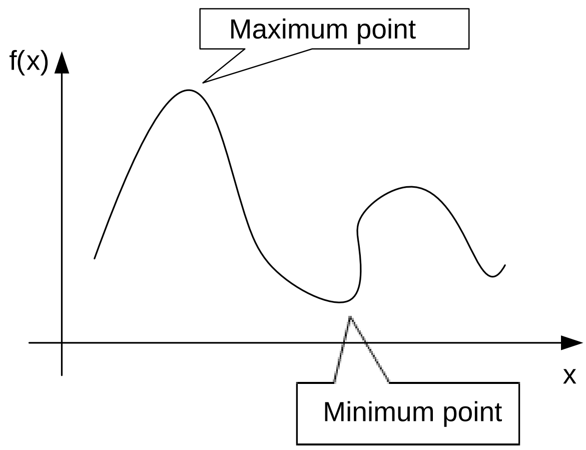
- To find the solution that maximizes the profit;
  or
- To find the solution that minimizes the use of scarce resources

Optimization can be divided into two main categories:

- Unconstrained optimization;
- Constrained optimization.

U**nconstrained optimization** deals with problems in which no constraints are imposed to the variables. The objective is to determine the solution that optimizes (minimizes or maximizes) a given function.

Example:



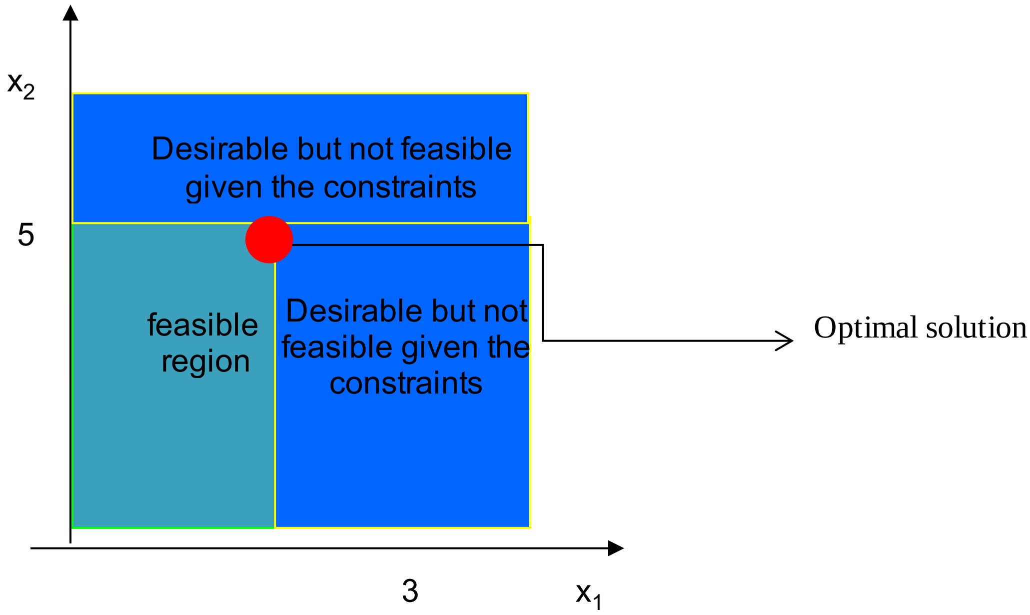
By the other hand, with the **constrained optimization** we still want to find the solution that optimizes a given function, but now the variables are subject to a set of constraints. Despite the set of constraints being more restrict, constrained problems are usually harder to solve to optimality than the unconstrained ones.

Example

Let us consider the following problem:

$$\begin{aligned} \max\ & x_1 + x_2 \\ \text{s.to.}\ \ & x_1 \leq 3 \\ & x_2 \leq 5 \\ & x_1, x_2 \geq 0 \end{aligned}$$

As it can be seen in the following figure, the imposed conditions limit the set of feasible solutions (feasible region).

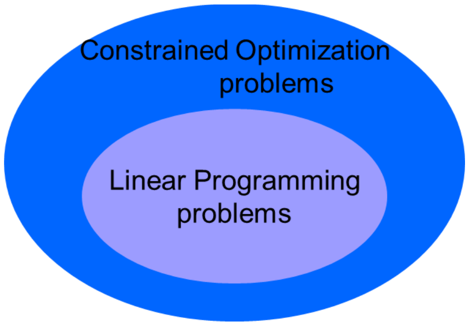


To mathematically define a constrained optimization problem, the following entities must be defined:

- **Variables** (also called decision variables): their values are not known when you start to solve the problem, and usually represent things that you adjust or control;
- **Objective function**: a mathematical expression that combines the variables to express the goal. It will required to maximize or to minimize the objective function;
- **Constraints**: mathematical expressions that combine variables to express limits on the possible solutions, e.g. the number of workers available to operate a machine is limited;
- **Variable bounds** (also called sign constraints): values that variables can take on are bounded.

2. Linear Programming

Linear Programming (LP) stands for **planning with linear functions**, which means that **linear mathematical expressions are required**. Therefore, Linear Programming problems are a subset of Constrained Optimization problems.



The mathematical formulation of a LP problem consists on the definition of a LP model.

The **general linear programming model** is as follows:

$$
\begin{array}{ll}
(\text{MIN})\ Z = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n & \rightarrow \text{OBJECTIVE FUNCTION} \\
\left.\begin{array}{l}
a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \ (\leq , = , \geq)\ b_1 \\
a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n \ (\leq , = , \geq)\ b_2 \\
\ldots \\
a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \ (\leq , = , \geq)\ b_m
\end{array}\right\} & \rightarrow \text{CONSTRAINTS} \\
x_1, x_2, \ldots, x_n \ (\geq 0, \leq 0, \text{unrestricted in sign}) & \rightarrow \text{SIGN RESTRICTIONS}
\end{array}
$$

where:

- **x_j, j = 1, …, n**, are the decision variables (i.e., are the variables that completely describe the decisions to be made);

- **c_j, j = 1, …, n**, are the objective function coefficients of the variables x_j – they may represent costs, profits, expenses, etc.;

- **a_{ij}, i = 1, …, m, j = 1, …, n**, are the technological coefficients;

- **b_i, i = 1, …, m**, are the constraints' right-hand sides;

- **Z**, is the objective function value – it may, for instance, represent the total profit or the total cost;

If all the variables are non-negative ($x_j \geq 0$, j=1,…,n), the sign restrictions are frequently mentioned as **non-negativity constraints**.

Note that, only linear expressions can be considered. Thus, variables can only be multiplied by constants. Thus, products of variables, or other functions of variables such as logarithms, roots, etc., are not allowed.

Besides the general form, it also exist the standard forms:

Standard form for a maximization problem

$$
\begin{aligned}
\text{MAX } Z = {} & c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \\
\text{s.to.} \quad & a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \leq b_1 \\
& a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n \leq b_2 \\
& \ldots \\
& a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \leq b_m \\
& x_1, x_2, \ldots, x_n \geq 0
\end{aligned}
$$

Standard form for a minimization problem

$$
\begin{aligned}
\text{MIN } Z = {} & c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \\
\text{s.to.} \quad & a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \geq b_1 \\
& a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n \geq b_2 \\
& \ldots \\
& a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \geq b_m \\
& x_1, x_2, \ldots, x_n \geq 0
\end{aligned}
$$

A LP model requires the following **assumptions**:

- **Proportionality**: The contribution of each activity is proportional to its level. If one unit of a certain activity j yields a profit of k (or consumes k units of resource i) x_j units yield a profit of $k\, x_j$ (or consume $k\, x_j$ units of resource i). This hypothesis excludes all start-up/setup costs situations, or costs/profits which depend on the activities' level;

- **Additivity**: The objective function value (or the spending of a certain resource) is the sum of the contributions of each activity (or each spending). This hypothesis excludes all situations where there is an interaction between activities or situations where there may exist gain or loss due to considering more than one activity;

- **Divisibility**: Each decision variables can take on fractional values - for example, in a production problem it may be possible to produce 0.3 of an object;

- **Certainty**: Each parameter (o. f. coefficient, technological coefficient or right-hand side) is known with certainty, and so each parameter is a constant.

Note that **proportionality** and **additivity** are guaranteed by the linear expressions, whereas **divisibility** is ensured by the signal constraints.

In real world problems, it is possible to find problems for which the LP formulation requires thousands of variables and constraints. However, due to the advances on the OR methods and on computers, some of these problems can be solved in a reasonable amount of time.

To write down a LP formulation is not a trivial task: it requires a detailed analysis of all the given information. Moreover, there are no standard procedures to follow. However, here we have some guidelines:

i) Identify the objective of the problem;
ii) Identify the decision variables and the restrictions;
iii) Write down the objective function and the functional constraints: use the given information about the problem to identify the coefficients of the decision variables;
iv) Add other relevant constraints, such as the sign restrictions.

Example

Give a LP formulation to the following problem:

A retailer wants to obtain no more than five tonnes of a certain product. The product can be ordered from two factories A and B. Each ton supplied by factory A contributes 4 thousand Euros to profit whereas each ton supplied by factory B contributes 3.5 thousand Euros to profit. Moreover, factory A can supply at most 3 tonnes. To maximize the profit, how many tonnes should be supplied by each factory?

Let us consider the **decision variables**:

x_i – tonnes of product to be ordered to factory i, i = A, B;

The problem can be formulated as a LP model as follows:

$$
\begin{array}{lll}
\text{Max } Z = 4\,x_A + 3.5\,x_B & & \text{(maximize the profit)} \\
\text{s.to.} & x_A + x_B \le 5 & \text{(no more than 5 ton can be ordered)} \\
 & x_A \le 3 & \text{(factory A cannot supply more than 3 ton)} \\
 & x_A,\, x_B \ge 0 & \text{(non-negative orders)}
\end{array}
$$

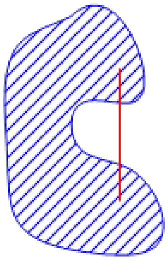
3. Definitions and properties

Let P be a LP problem, with n decision variables and m functional constraints:

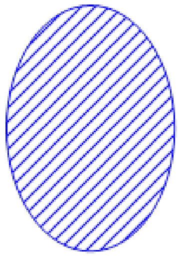
- **Solution**: is any vector of values for the decision variables;
- **Feasible solution**: is any solution that satisfies all the problem's constraints and sign restrictions;
- **Feasible region (S)**: is the set of all feasible solutions;
- **Optimal solution**: is the feasible solution with the highest (lowest) objective function value;

In LP, the following properties hold:

i) The **feasible region S** of an LP problem is a closed convex set with a finite number of extreme points.

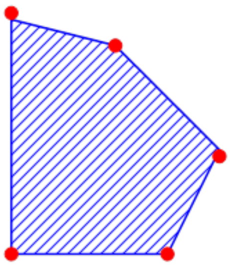


Non-convex set



Convex set

ii) If a finite optimum exists (max or min) the o. f. finds it on an extreme point in S; if the optimum is obtained in more than one extreme point, all convex combinations of the 2 points will correspond to an optimum.



4. Solving linear programming problems

The resolution of a linear programming problem is usually done using the Simplex Method.

The Simplex method has underlying that all the functional inequalities are transformed into equalities. This transformation is done with deviation variables (slack and surplus) as follows:

- For each constraints $\mathbf{a_{i1}\, x_1 + a_{i2}\, x_2 + \ldots + a_{in}\, x_n \le b_i}$ (of type "≤"):
  Consider the **slack variable s_i** and do:
  $$a_{i1}\, x_1 + a_{i2}\, x_2 + \ldots + a_{in}\, x_n + \mathbf{s_i} = b_i, \text{ with } s_i \ge 0$$

- For each constraints $\mathbf{a_{i1}\, x_1 + a_{i2}\, x_2 + \ldots + a_{in}\, x_n \ge b_i}$ (of type "≥"):
  Consider the **surplus variable t_i** and do:
  $$a_{i1}\, x_1 + a_{i2}\, x_2 + \ldots + a_{in}\, x_n - \mathbf{t_i} = b_i, \text{ with } t_i \ge 0$$

The Simplex method is an iterative process that solves linear programming problems. It departs from an extreme point and visits other adjacent extreme points, but always ensuring that the solutions that are successively obtained never worsen the value of the objective function.

The Excel Solver is one of the software's that can be used to solve linear programming problems by the Simplex Method.

To use the Excel Solver it is not necessary to transform inequalities into equalities, that is, the original constraints should be included. Nevertheless, the outputs analysis is made assuming that transformation of the constraints and thus considering the existence of the deviation variables (slack and surplus).

Section 7 discusses the use of Excel Solver for linear programming problems.

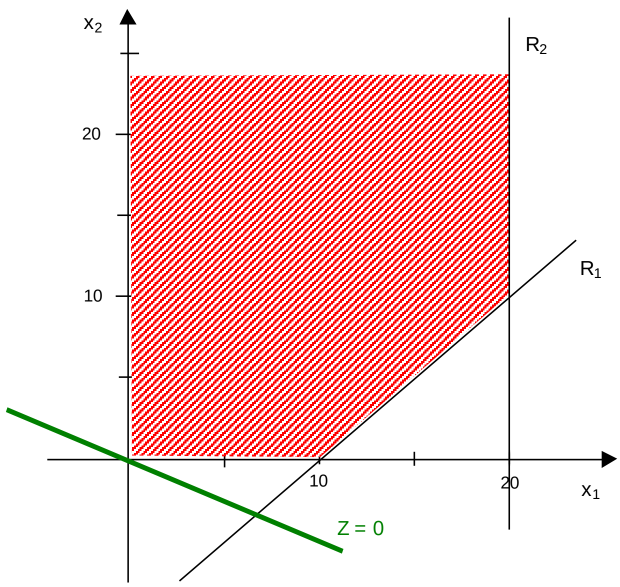
5. Special cases of solutions

Unbounded problem

In some LP models, the problem is unbounded, meaning that the objective value can be driven to infinity.

Example

$$\text{Max } Z = x_1 + 2x_2$$
$$\text{s.t. } x_1 - x_2 \leq 10 \quad (R_1)$$
$$2x_1 \leq 40 \quad (R_2)$$
$$x_1, x_2 \geq 0$$



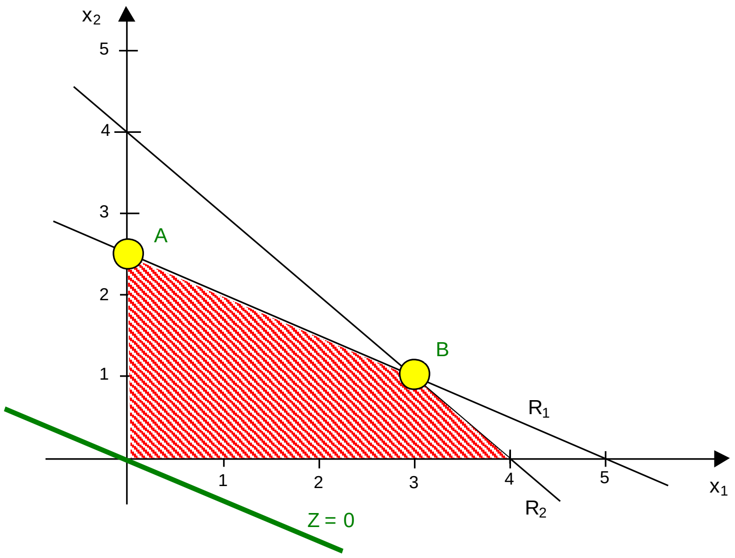
The infeasible region is unbounded in the direction of growth of x_2, which is also the direction of growth of Z (x_2 +∞ and $Z \rightarrow +\infty$). Thus, the problem is unbounded

Multiple optimal solutions

In some LP models, instead of a unique optimal solution, there are infinite optimal solutions (multiple optimal solutions), for which the objective function assumes the same optimal value.

Example

Max $Z = 2x_1 + 4x_2$
s.to $x_1 + 2x_2 \leq 5$ (R_1)
$x_1 + x_2 \leq 4$ (R_2)
$x_1, x_2 \geq 0$



There are two optimal extreme points (points A and B), thus all the points of the line segment between A and B are also optimal.

The existence of multiple optimal solutions is detected (graphically, for problems with two variables) when the line of the objective function is parallel to the line of a constraint that defines the optimal solution. The o.f. assumes the same value for any point on the line associated with that constraint.

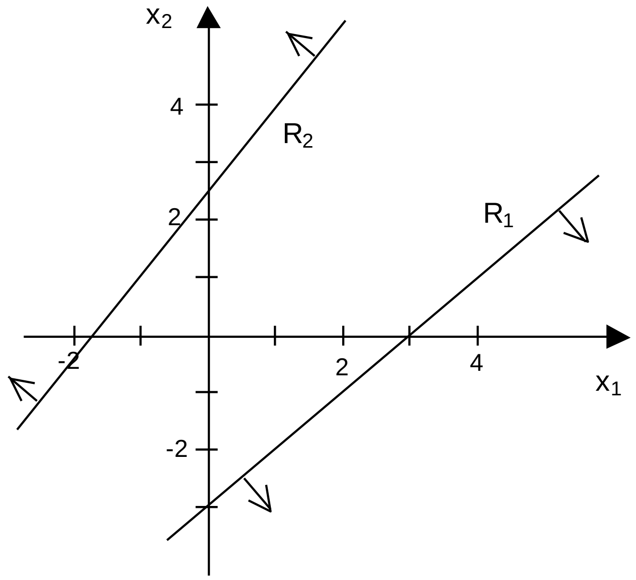
Infeasible problem

In some LP models the feasible region is empty, i.e., there are no points that satisfy all the constraints of the problem (including the sign constraints). The problem is named unfeasible.

Example

Max $Z = 3x_1 + 7x_2$
s.to $x_1 - x_2 \geq 3$ (R_1)
$-3x_1 + 2x_2 \geq 5$ (R_2)
$x_1, x_2 \geq 0$

On the first quadrant, the regions $x_1 - x_2 \geq 3$ and $-3x_1 + 2x_2 \geq 5$ do not intersect. Thus, the feasible region is empty, and the problem is infeasible.



6. Sensitivity analysis

Sensitivity analysis is helpful to determine "how" sensitive a model is to changes in the value of its parameters.

The parameters of a LP model, a_{ij}, b_i, c_j, i = 1, ..., m, j = 1, ..., n, are <u>known constants</u> but, in general, represent only estimates of values, as they represent quantities that sometimes are very difficult or even impossible to measure to a great deal of accuracy in the real world. Therefore, they should not be given an excessive degree of confidence.

By the other hand, parameters such as b_i (which can represent for instance raw-materials) are frequently subjected to external conditions, and so their values change overtime.

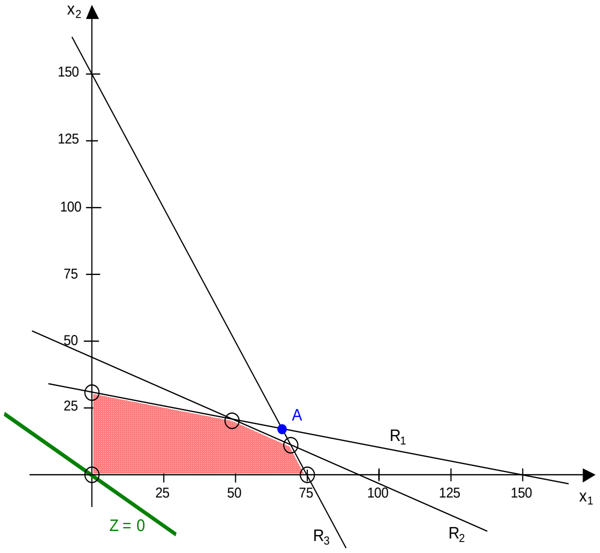
By showing how the model responds to changes in parameter values, sensitivity analysis is a useful tool in model building as well as in model evaluation. Thus, after obtaining the optimal solution, sensitivity analysis must be carried out in order to investigate what is the effect, in the optimal solution and in the optimal value, caused by changing the value of one parameter, assuming that all other parameters remain unchanged.

Through sensitivity analysis it is possible to analyse the model with a small computational effort, when compared to the alternative of solving the problem from the beginning for the new parameter.

Consider the following example:

<u>Example</u>

$$\begin{aligned} &\text{Max } Z = 30\,x_1 + 40\,x_2 \\ &\text{s.to} \quad x_1 + 5\,x_2 \leq 150 \quad (R_1) \\ &\qquad\;\; x_1 + 2\,x_2 \leq 90 \quad (R_2) \\ &\qquad\;\; 2\,x_1 + x_2 \leq 150 \quad (R_3) \\ &\qquad\;\; x_1,\, x_2 \geq 0 \end{aligned}$$



The optimal point has coordinates $(x_1,x_2) = (70,10)$, and the optimal value is
$Z^* = 30 \times 70 + 40 \times 10 = 2500$

The optimal solution is in lines of constraints R2 and R3, meaning that they are verified in their equality (i.e., $x_1^*+2\,x_2^*= 90$ and $2\,x_1^*+ x_2^*= 0$). For that reason, constraints R1 and R2 are called active constraints, while R1 is called not active.

Change in an objective function coefficient

What happens if one changes the value of c_1, which represents the objective coefficient of x_1?

The objective function $Z = 30\,x_1 + 40\,x_2$ can be rewritten as an equation of a line (y=mx+b) as follows: $x_2 = (-\ 30/40)\ x_1 + (Z/40)$.

Thus, if c_1 changes from 30 to another value, the slope will change, and the optimal point may change. Active restrictions may no longer be R2 and R3, that is, the set of variables that can take a value other than zero may be changed.

The allowable range for an objective coefficient c_j gives the range of values for c_j for which the active constraints are guaranteed to remain optimal, that is, the set of variables that can take a value other than zero remains the same. This result is obtained assuming that all other parameters remain unchanged.

For c_j within this range, the values of the decision variables remain unchanged, but the optimal value Z may or may not change.

Let AR_{cj} be the allowable range for c_j, c_j^N be the new value for parameter c_j, and Z^N be the new optimal value.

- If $c_j^N \in AR_{cj}$, then:
    - The solution remains optimal (i.e., all the variables will have the same values);
    - The optimal value may change: $Z^N = Z + \Delta Z = Z + (c_j^N - c_j) \times x_j^*$, where x_j^* is the value of x_j in the optimal solution;

- If $c_j^N \notin AR_{cj}$:
    - Active constraints are changed and thus the optimal solution and the optimal value also change. It is necessary to resolve the problem (for instance, with the Excel Solver);

Change in a right-hand side

Let us consider the constraint $x_1 + 2\,x_2 \leq 90$ and its associated line equation $x_1 + 2\,x_2 = 90$. We easily see that changing the value of parameter b_2 will move the y-intercept, which graphically corresponds to moving the line of the constraint R2 parallel to itself.

Thus, changing the right-hand side b_2 from 90 to another value will change the feasible region. Active restrictions may no longer be R2 and R3, that is, the set of variables that can take a value other than zero may be changed.

The allowable range for a right hand side b_i gives the range of values for b_i for which the active constraints are guaranteed to remain optimal, that is, the set of variables that can take a value other than zero remains the same. This result is obtained assuming that all other parameters remain unchanged.

Even if b_i remains within this range, the values of the decision variables will probably change, and the optimal value Z may or may not change.

Let AR_{bi} be the allowable range for b_i, b_i^N be the new value for parameter b_i, and Z^N be the new optimal value.

Is it possible to know by how much the optimal value will change if parameter b_i changes? The answer is yes, it is possible, from the shadow price.

The **shadow price** for the *i*-th constraint is the amount by which the optimal Z value is changed for each unit added to the right-hand side b_i. This conclusion is only valid if the new value of b_i is within its respective allowable range.

Let AR_{bi} be the allowable range for b_i, b_i^N be the new value for parameter b_i, and Z^N be the new optimal value.

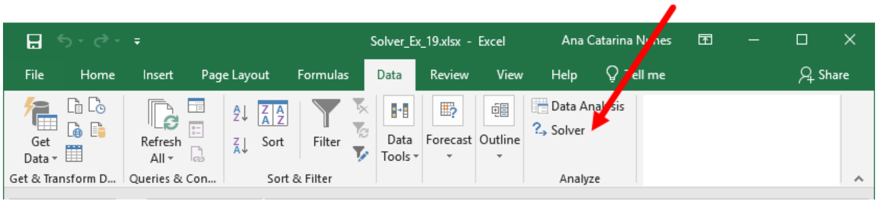
- If $b_i^N \in AR_{bi}$:
  - The variables that can take a value other than zero are the same (i.e., the active constraints are the same), although the values of the variables may change (new values can be determined using Excel Solver);
  - The shadow price of the *i*-th constraint, y_i^*, is valid;
  - The optimal value may change: $Z^N = Z + \Delta Z = Z + (b_i^N - b_i) \times y_i^*$;

- If $b_i^N \notin AR_{bi}$:
  - Active constraints are changed and thus the optimal solution and the optimal value also change. It is necessary to resolve the problem (for instance, with the Excel Solver).

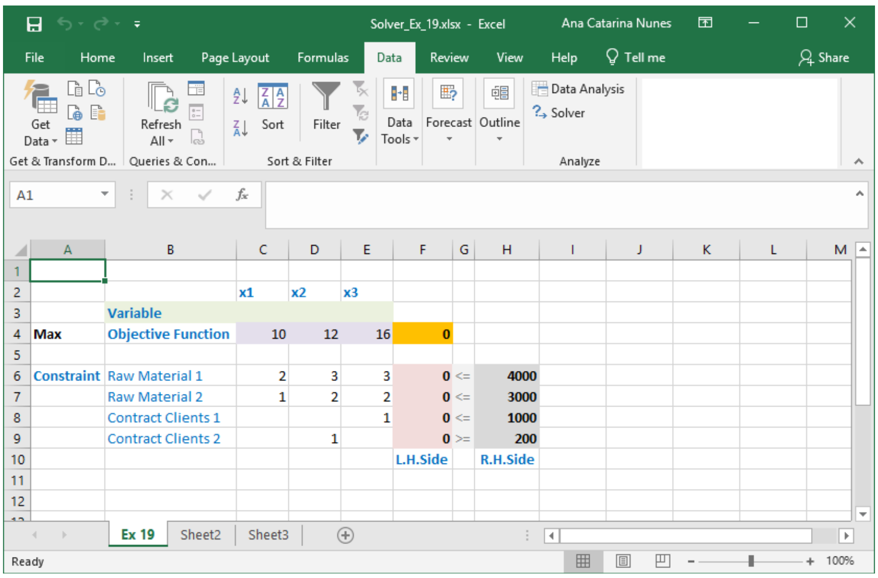
7. Using Excel Solver for LP models

This section concerns using Excel Solver to solve linear programming problems. As previously mentioned, the original constraints should be included (see the end of Section 4).

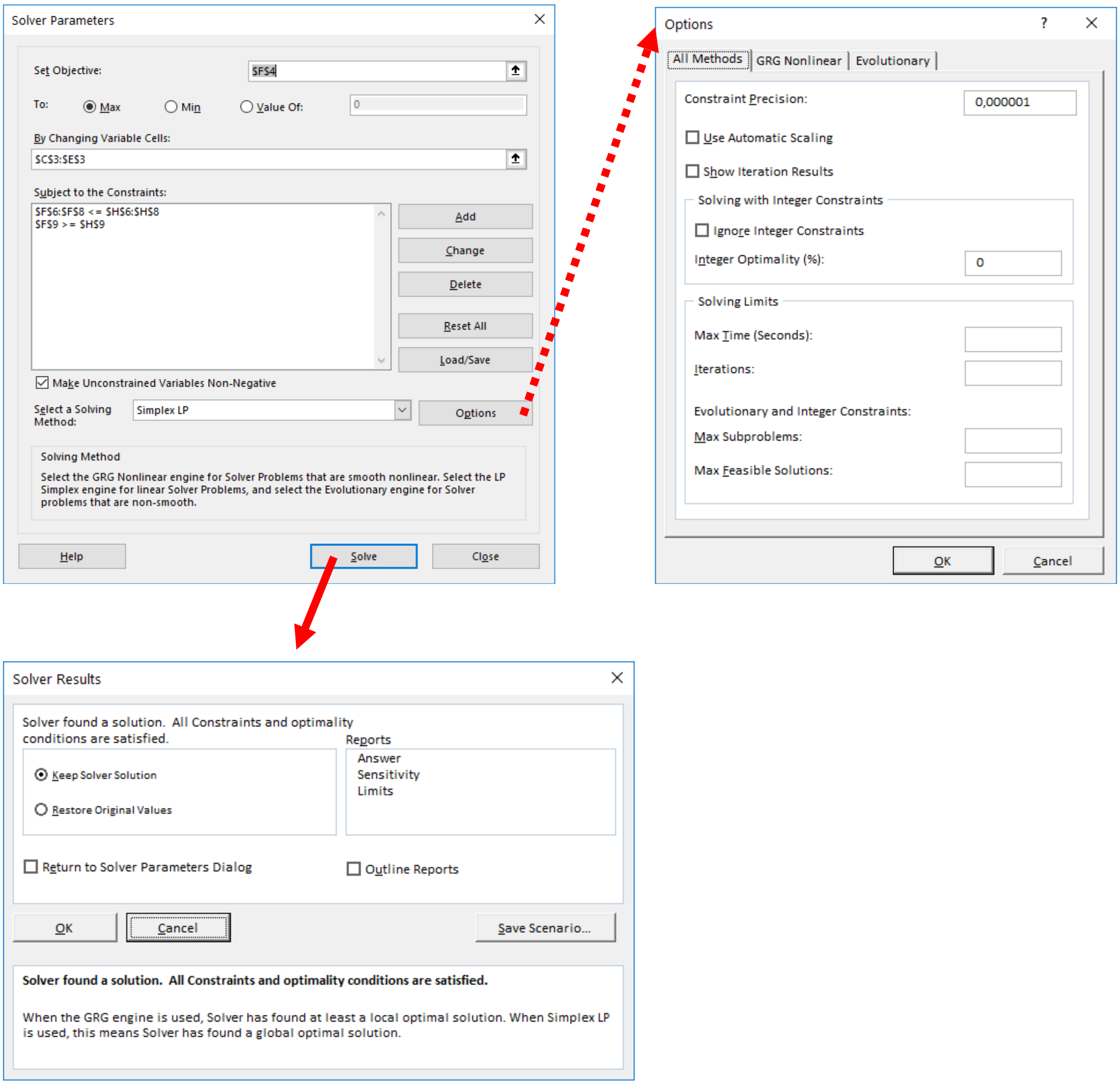
In Excel, the **Solver** command is available in the **Analysis** group on the **Data** tab. If it is not available, it must be loaded, following the instructions given at: https://support.microsoft.com/en-us/office/load-the-solver-add-in-in-excel-612926fc-d53b-46b4-872c-e24772f078ca



For an example on entering data and formulas, see the file *Solver_Ex_19.xlsx* available in the *Content Repository* of Fénix.



See the same file as an example on how to define and solve a problem using Solver.

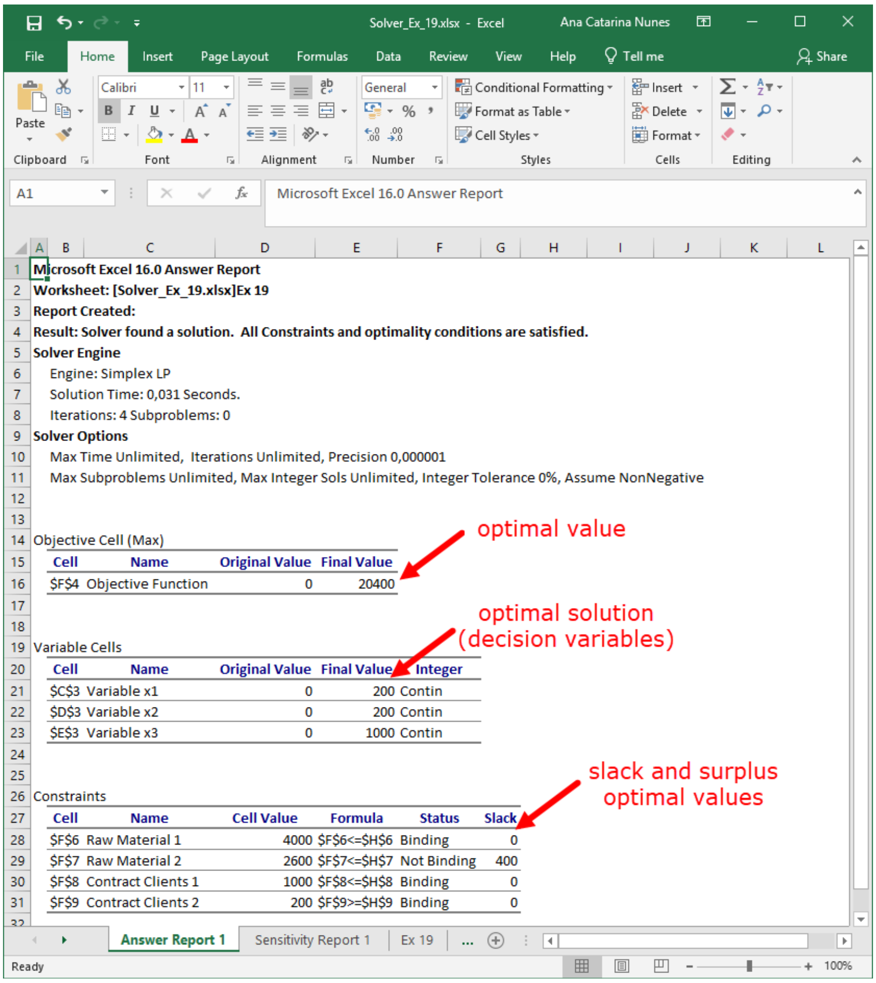


For more details:
https://support.microsoft.com/en-us/office/define-and-solve-a-problem-by-using-solver-5d1a388f-079d-43ac-a7eb-f63e45925040

Excel Solver generates several reports (Answer, Sensitivity and Limits), containing information about:

- The optimal value of the objective function;
- The optimal value of the decision and deviation variables (slack and surplus);
- The shadow price of each constraint;
- The values from which the allowable ranges are calculated (*Allowable Increase* and *Allowable Decrease*).

Answer Report



Microsoft Excel 16.0 Answer Report
Worksheet: [Solver_Ex_19.xlsx]Ex 19
Report Created:
Result: Solver found a solution. All Constraints and optimality conditions are satisfied.
Solver Engine
Engine: Simplex LP
Solution Time: 0,031 Seconds.
Iterations: 4 Subproblems: 0
Solver Options
Max Time Unlimited, Iterations Unlimited, Precision 0,000001
Max Subproblems Unlimited, Max Integer Sols Unlimited, Integer Tolerance 0%, Assume NonNegative

Objective Cell (Max)

Cell	Name	Original Value	Final Value
$F$4	Objective Function	0	20400

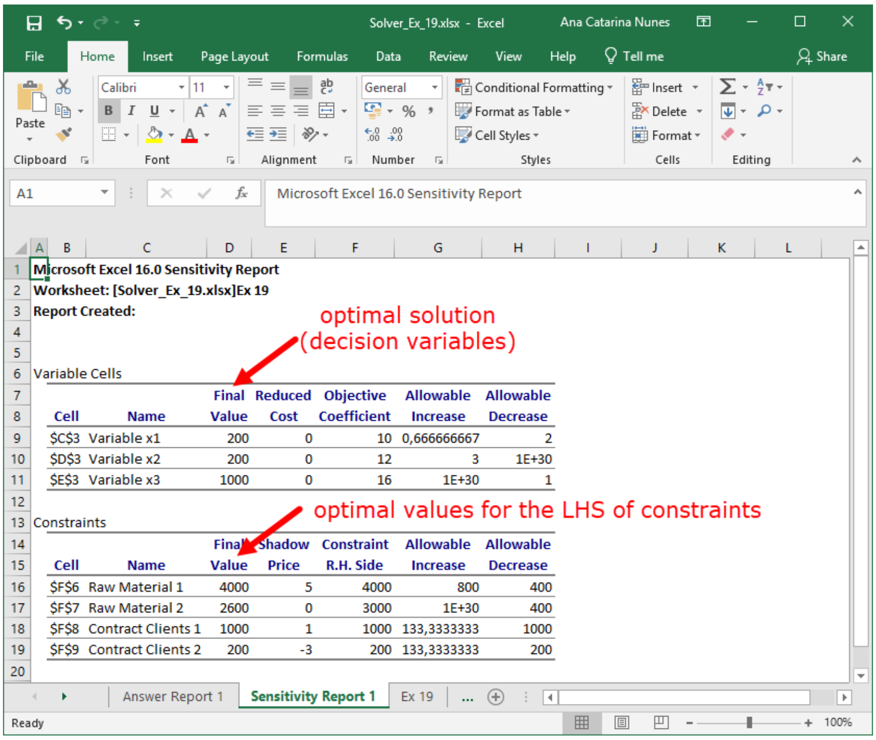
Variable Cells

Cell	Name	Original Value	Final Value	Integer
$C$3	Variable x1	0	200	Contin
$D$3	Variable x2	0	200	Contin
$E$3	Variable x3	0	1000	Contin

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
$F$6	Raw Material 1	4000	$F$6<=$H$6	Binding	0
$F$7	Raw Material 2	2600	$F$7<=$H$7	Not Binding	400
$F$8	Contract Clients 1	1000	$F$8<=$H$8	Binding	0
$F$9	Contract Clients 2	200	$F$9>=$H$9	Binding	0

Sensitivity Report



Microsoft Excel 16.0 Sensitivity Report
Worksheet: [Solver_Ex_19.xlsx]Ex 19
Report Created:

Variable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$C$3	Variable x1	200	0	10	0,666666667	2
$D$3	Variable x2	200	0	12	3	1E+30
$E$3	Variable x3	1000	0	16	1E+30	1

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$F$6	Raw Material 1	4000	5	4000	800	400
$F$7	Raw Material 2	2600	0	3000	1E+30	400
$F$8	Contract Clients 1	1000	1	1000	133,3333333	1000
$F$9	Contract Clients 2	200	-3	200	133,3333333	200

8. Transportation and assignment models

The Transportation Problem and the Assignment Problem are special linear programming models. Due to their mathematical properties, there is a guarantee that all variables will have an integer value in the optimal solution (therefore, without the need to include it explicitly in the model), as long as the right-hand sides of the constraints are integer values.

The Transportation Model (or Transportation Problem)

Suppose the situation in which it is intended to determine the minimum transportation cost plan of a certain product, from a set of m sources and a set of n destinations. A supply S_i is associated with each source i (i=1,…,m), and a demand D_j is associated to each destination j (j=1,…,n). Let c_{ij} denote the cost per unit shipped from source i to destination j. The values S_i, D_j e c_{ij} are known constants.

Considering the decision variables $\mathbf{x_{ij}}$ that represent the quantity shipped from source i to destination j (i=1,…,m, j=1,…,n), this problem may be formulated as follows:

$$\min \quad Z = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$

$$\text{s. to:} \quad \sum_{j=1}^{n} x_{ij} = S_i \quad i=1,\ldots,m$$

$$\sum_{i=1}^{m} x_{ij} = D_j \quad j=1,\ldots,n$$

$$x_{ij} \geq 0 \quad i=1,\ldots,m\ ,\ j=1,\ldots,n$$

This formulation assumes a balanced system, with $\Sigma S_i = \Sigma D_j$. If not, then it will be needed to add a dummy source (if $\Sigma Si < \Sigma Dj$) or a dummy destination (if $\Sigma Si > \Sigma Dj$).

Transportation Model (or Problem) is the name given to the structure of this mathematical model. This type of model has a wide range of real applications, some of which may not even be an effective transportation between sources and destinations.

The Assignment Model (or Assignment Problem)

Suppose the situation in which it is intended to assign, with the minimum cost, a certain number of assignees (n) to the same number of tasks. Each assignee is to be assigned to exactly one task, and each task is to be performed by exactly one assignee. Let c_{ij} be the cost of assigning assignee i to task j. The values c_{ij} are known constants.

Considering the decision variables:

$$x_{ij} = \begin{cases} 1 & \text{if task j is performed by assignee i} \\ 0 & \text{if not} \end{cases}, \text{ with } i=1,\ldots,n,\ j=1,\ldots,n,$$

this problem can be formulated as follows:

$$\min \quad Z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}\, x_{ij}$$

$$\text{s. to:} \quad \sum_{j=1}^{n} x_{ij} = 1 \qquad i=1,\dots,n$$

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

$$x_{ij} \in \{0,1\} \quad i=1,\dots,n\ ,\ j=1,\dots,n$$

This formulation assumes a balanced system: the number of tasks is exactly the same as the number of agents. If it is not the case, then it will be necessary to create dummy assignees or dummy tasks, until they are equal.

Assignment Model (or Problem) is the name given to the structure of this mathematical model. This type of model has a wide range of real applications, some of which may not even deal with assigning tasks to agents.