CHAPTER 4
SOLVING LINEAR PROGRAMMING MODELS

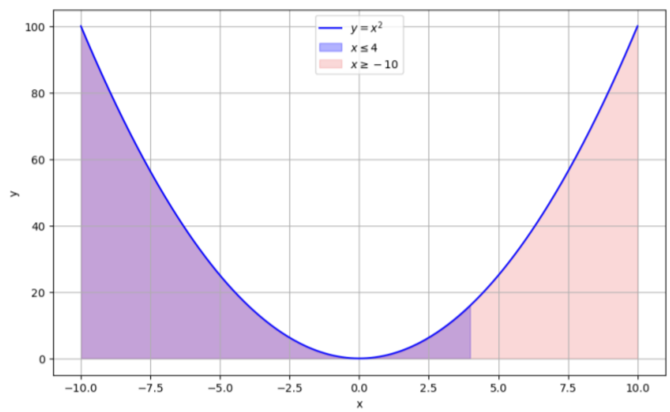
- We will learn some basic terminology relevant to discuss all types of optimization problems
- We will explore the fundamentals of the simplex method, the most popular algorithm for the solution of linear programming problems.

Basic terminologies

- *Active/inactive constraints:* consider the following optimization problem

$$\min(x^2)$$
$$st. \begin{array}{l} x \leq 4 \\ x \geq -10 \end{array}$$

This problem is illustrated in the following figure.

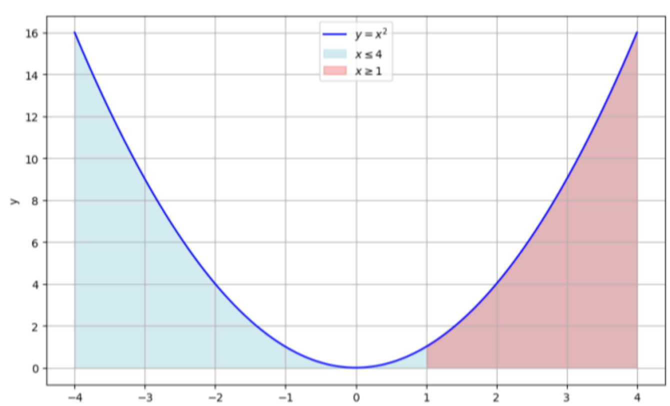


The optimal solution for this optimization problem is $x^* = 0$, note that in this case we can completely ignore the constraints and still get the same answer. If this is the case, we say that the *constraints are inactive* at the solution.

Now let's consider the following problem:

$$\min(x^2)$$
$$st. \begin{array}{l} x \leq 4 \\ x \geq 1 \end{array}$$

The graphical representation of the problem is shown below:



Note that the optimal value of this function occurs at $x = 1$, thus we see that the first inequality can be ignored without causing a change in the solution, but the second inequality, which is verified exactly at the solution (i.e., $x^* = 1$), cannot be ignored. We say that this constraint is active.

Definition: Consider $g: R^n \to R^p$ and $h: R^n \to R^m$. For $1 \leq i \leq p$ an inequality constraint

$$g_i(x) \leq 0$$

Is said to be active in x^* if

$$g_i(x^*) = 0$$

and inactive in x^* if

$$g_i(x) < 0$$

By extension, for $1 \leq i \leq m$ an equality constraint

$$h_i(x) = 0$$

is said to be active if it is satisfied in x^*, i.e.,

$$h_i(x^*) = 0$$

The set of indices of the active constraints in x^* is denoted by $\mathcal{A}(x^*)$.

- *Feasible directions*: consider a general optimization problem and a feasible point $x \in R^n$. A direction d is said to be feasible in x if there exists $\eta > 0$ such that $x + \alpha d$ is feasible for any $0 < \alpha \leq \eta$.

  When the set X of the constraints is convex, the identification of a feasible direction in $x \in X$depends on the identification of a feasible point $y \in X$, other than x.

  Theorem: Let X be a feasible convex set and consider $x, y \in X, y \neq x$. The direction $d = y - x$ is a feasible direction in x and $x + \alpha d = x + \alpha(y - x)$ is feasible for any $0 \leq \alpha \leq 1$

    Corollary: Let $X \subseteq R^n$ and $x \in R^n$ be an interior point of X. Here, any direction $d \in R^n$ is feasible in x.

  Theorem (linear case): consider the optimization problem $\min f(x)\ subject\ to\ Ax = b, x \geq 0$, and let x^+ be a feasible point. A direction d is feasible in x^+ if and only if

  $$Ad = 0$$

  This first point is easy to understand, note that if we are on a feasible direction we are moving from one feasible point to another, thus $Ax = b$ should also hold true for this new point such that:

  $$b = A(x^+ + \alpha d) = Ax^+ + \alpha Ad = b + \alpha Ad$$

  $$d_i \geq 0 \text{ when } x_i^+ = 0$$

- *Polyhedron*: *a polyhedron is a set of points of* R^n delimited by hyperplanes.

  $$\{x \in R^n | Ax \leq b\}$$

$$A \in R^{m \times n} \text{ and } b \in R^m.$$

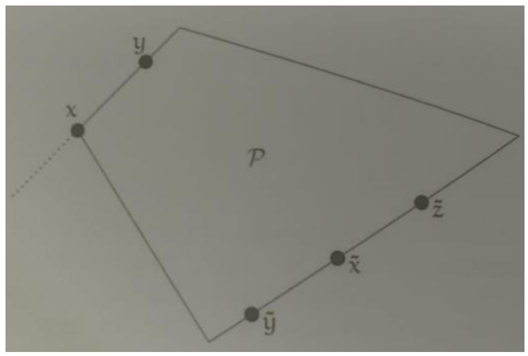
There exists an alternative representation for a polyhedron, which is known as the standard form, in this case a polyhedron is defined in the following way:

$$\{x \in R^{n'} | Ax' = b, x' \geq 0\}$$

(Can you spot the connection with a LP problem? Note that the set of constraints of a linear programming problem constitute a polyhedron)

- *Vertex*: let $\mathcal{P}$ be a polyhedron. A vector $x \in \mathcal{P}$ is a vertex of $\mathcal{P}$ if it is imposible to find two vectors $y, z \in \mathcal{P}$, different from x such that x is a convex combination of y and z. That is, such that there exist a real number $0 < \lambda < 1$ such that:

$$x = \lambda y + (1 - \lambda) z$$



Note that for the vertex point x there is not a second vector z that can be used in a convex combination to produce x.

One critical question is how we can find/identify the vertexes of a polyhedron, toward this end, we can use the following theorem:

Theorem: Let $\mathcal{P} = \{x \in R^n | Ax = b, x \geq 0\}$ be a polyhedron represented in standard form, with $A \in R^{m \times n}$ of full rank and $b \in R^m$ and $n \geq m$. Consider m linearly independent columns of A, and call B the matrix containing these m columns, and N the matrix containing the remaining $n - m$ columns, such that

$$AP = (B|N)$$

Where P is the appropriate permutation matrix. Consider the vector

$$x = P \begin{pmatrix} B^{-1}b \\ 0_{R^{n-m}} \end{pmatrix}$$

If $B^{-1}b \geq 0$, then x is a vertex of $\mathcal{P}$. We will not prove the theorem, but note that this ensures that the condition that $x \geq 0$ for a standard polyhedral is satisfied.

- Basic solution: Let $P = \{Ax = b, x \geq 0\}$ be a polyhedron represented in standard form, with $A \in R^{m \times n}$ and $n \geq m$. A vector $x \in R^n$ such that $Ax = b$ along with a set of indices $j_1, \dots, j_m$ is said to be a basic solution of $\mathcal{P}$ if:
  - The matrix $B = (A_{j1}, \dots, A_{jm})$ composed of columns $j_1, \dots, j_m$ of the matrix A is nonsingular and
  - $x_i = 0$ if $i \neq j_1, \dots, j_m$

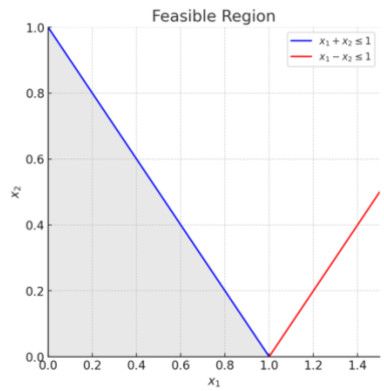
If moreover, $x = B^{-1}b \geq 0$, the vector x is called a feasible basic solution.

Example: Consider the following optimization problem.

$$\min c^T x$$

$$s.t. \begin{array}{c} x_1 + x_2 \leq 1 \\ x_1 - x_2 \leq 1 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{array}$$

This feasible region is presented below.



This problem is not in standard form, if we are to represent this in standard form, then the resulting problem is shown below:

$$\min c^T x$$

$$s.t. \begin{array}{c} x_1 + x_2 + s_1 = 1 \\ x_1 - x_2 + s_2 = 1 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ s_1 \geq 0 \\ s_2 \geq 0 \end{array}$$

The feasible region for the problem in standard form can be represented by the following matrices:

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Let's explore the different basic and feasible basic solutions, they can be found by selecting different columns (columns: 1-2, 1-3,1-4,2-1,2-3,3-4). In total there are six.

Basic solution 1: x_1 and x_2 in the basis (columns 1 and 2 $j_1 = 1, and\ j_2 = 2$).

$$B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}; B^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}; x_B = B^{-1}b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

If we plot this solution on the same figure as before (note that we can ignore the slack variables here), we will find that this is a corner point of the feasible region. This means that this is a basic feasible solution.

Basic solution 2: x_1 and x_3

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}; B^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}; x_B = B^{-1}b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Basic solution 3: x_1 and x_4

$$B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}; B^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}; x_B = B^{-1}b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; x = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Basic solution 4: x_2 and x_3

$$B = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}; B^{-1} = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}; x_B = B^{-1}b = \begin{bmatrix} -1 \\ 2 \end{bmatrix}; x = \begin{bmatrix} 0 \\ -1 \\ 2 \\ 0 \end{bmatrix}$$

Basic solution 5: x_2 and x_4

$$B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}; B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}; x_B = B^{-1}b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; x = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 2 \end{bmatrix}$$

Basic solution 6: x_3 and x_4

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; B^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; x_B = B^{-1}b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}; x = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

As a conclusion to this example, we should note that the basic variables are obtained by solving $x_B = B^{-1}b$, and the non-basic variables are arbitrarily selected equal to zero, such that we ensure that the second condition that determines that something is a basic solution is satisfied.

The notion of a basic solution enables us to analyze the polyhedron in terms of active constraints of the optimization problem. Let x be a feasible basic solution such that $x_B = B^{-1}b > 0$ (note that to be feasible the variables x_B need to be positive because we have imposed this constraint in the standard polyhedron formulation), we say that this is a non-degenerate solution. In this case there are exactly n active constraints (the m equality constraints and the $n - m$ non basic variables which are zero. The constraint $x_i \geq 0$ corresponding to the basic variables are all inactive because $x_B > 0$.

In some cases, some components of $x_B = B^{-1}b$ are zero, in these cases there are more than *n* active constraints. And the feasible basic solution is defined by more than *n* equations in a space of n dimensions. If this is the case, we say that we are dealing with a degenerate feasible basic solution.

Theorem: the point x^* is a vertex of P if and only if it is a feasible basic solution.

- Basic direction: If x is a feasible basic solution, **then a basic direction, is a feasible direction along the edges of the polyhedron,** where the edges are given by the constraints adjacent to the vertex corresponding to the feasible basic solution x. To define these directions, we consider a feasible basic solution:

$$\boldsymbol{x} = \binom{x_B}{x_N} = \binom{B^{-1}b}{0}$$

Consider a non-basic variable, for instance the variable with index p, and define a direction that gives positive values to this non basic variable, all the while maintaining the other non-basic variables at zero. Then:

$$d = \binom{d_B}{d_N} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \\ d_{m+1} \\ \vdots \\ d_{p-1} \\ d_p \\ d_{p+1} \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Since the $d_N$part is defined, we now need only to define d_B. If that is the case, then we can make use of a previously presented theorem (see feasible directions subsection) and write:

$$Ad = 0 = [BN]\binom{d_B}{d_N} = Bd_B + Nd_N = Bd_B + \sum_{j=m+1}^{n} A_j d_j = Bd_B + A_p = 0$$

That is,

$$d_p = -B^{-1}A_p$$

Now we can formally define a basic direction: Let $P = \{x \in R^n | Ax = b, x \geq 0\}$ be a polyhedron represented in standard form, with $A \in R^{m\times n}$ and $b \in R^m$ and $n \geq m$ and let $x \in R^n$ be a feasible basic solution of P. A direction d is called the pth basic direction if p is the index of a non-basic variable and

$$d_p = \boldsymbol{P}\binom{d_{B_p}}{d_{N_p}}$$

where $\boldsymbol{P}$ is the permutation matrix corresponding to the basic solution, $d_{B_p} = -B^{-1}A_p$ (note that A_p is the pth column), and d_{N_p} is such that

$$d_{N_p} = \boldsymbol{P^T} e_p = \begin{pmatrix} 0 \\ d_{N_p} \end{pmatrix}$$

That is all the elements of d_{N_p} are zero, except the one corresponding to the variable p which is 1.

Example: Consider a polyhedron represented in standard form

$$P = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} | Ax = b, x \geq 0 \right\}$$

With:

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix}, b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Consider the basic solution where x_2 and x_4 are in the basis, in this case we have that:

$$x = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}$$

(this numbers can be found following the approach described in the basic solution section). The permutation matrix is such that we reorganize the elements of the solution, in this case the solution was found such that x_2 and x_4 are in the basis and x_1 and x_3 are not, this leads to a vector as follows:

$$d_p = \boldsymbol{P} \begin{pmatrix} x_2 \\ x_4 \\ x_1 \\ x_3 \end{pmatrix}$$

The permutation matrix required to reorganize this is:

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

An easy way to remember how to understand the effect of this matrix on vector x is as follows: when we are pre-multiplying the matrix we read it row by row, from left to right, the element that contains the number one will be the one that will be placed in the corresponding row. For example, for row 1, we see a number one in element 3, this means that after applying the permutation the third element of the column vector (i.e., x_1)will be placed in the first row.

In this case matrixes B and B^{-1} are given by:

$$B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}; B^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

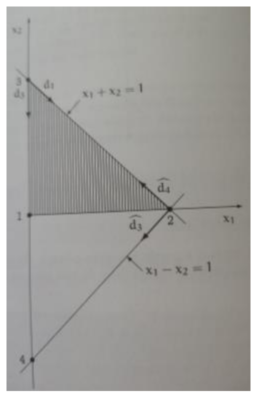
In this case the basic direction corresponding to the non-basic variable x_1 is as follows:

$$d_1 = \boldsymbol{P}\begin{pmatrix} d_{B_1} \\ d_{N_1} \end{pmatrix} = \begin{pmatrix} -B^{-1}A_p \\ 1 \\ 0 \end{pmatrix} = \boldsymbol{P}\begin{bmatrix} -\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ 1 \\ 0 \end{bmatrix} = \boldsymbol{P}\begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ -2 \end{pmatrix}$$

In this case the basic direction corresponding to the non-basic variable x_3 is as follows:

$$d_3 = \boldsymbol{P}\begin{pmatrix} d_{B_1} \\ d_{N_1} \end{pmatrix} = \begin{pmatrix} -B^{-1}A_p \\ 0 \\ 1 \end{pmatrix} = \boldsymbol{P}\begin{bmatrix} -\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ 1 \\ 0 \end{bmatrix} = \boldsymbol{P}\begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

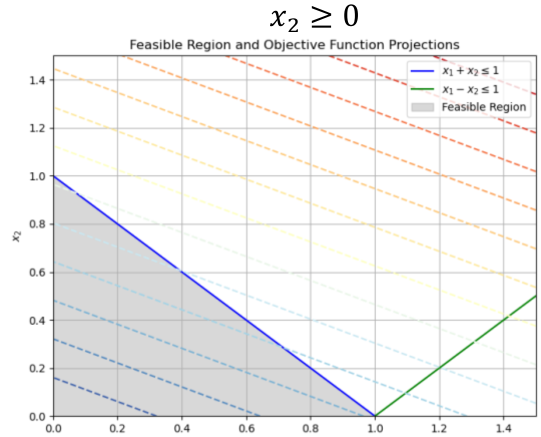
A graphical representation is shown below, note that we are at point three, and that we found two basic directions d_1 and d_3 that move along the axes (to avoid confusion, note that the arrow depicted is such that you are moving along $x_3 + \alpha d_i$ where $i = 1{,}3$.



The simplex algorithm

- *Introduction*. The first thing that we are going to do, is to try to understand the solution of a linear programming problem from a geometrical point of view. With that aim we will use a 2D problem, which can be easily visualized. Let's consider

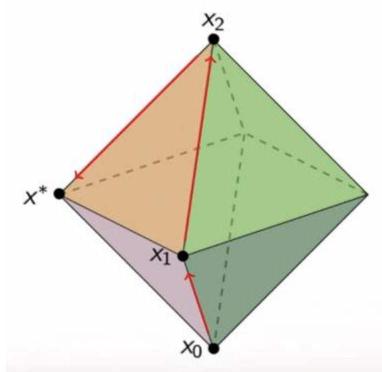
$$\begin{array}{rr} & \min(-x_1 - 2x_2) \\ & x_1 + x_2 \leq 1 \\ s.t. & x_1 - x_2 \leq 1 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}$$



The feasible region is shown in grey, and the dashed lines represent contours for the objective function. The key observation from the simplex method is that the optimal value (if it exists) will lie on a vertex of the polytope defined by the intersection of the constraints.

The main question is how to survey the vertexes of the polytope systematically, since visiting them all may take too long. The simplex method uses a smart approach, at each iteration, it identifies a descent direction, and then it moves a bit in that direction. In general, the geometric interpretation can be summarized as follows:

- Start from a known vertex (note that we are starting from a vertex this in itself needs to be discussed)
- Identify an edge of the polyhedron along which the objective function decreases. If no such edge exists, the current vertex is an optimal solution. In the figure, we start at x_0, and the idea of the algorithm as it iterates is to *move along the edges* until a new vertex is found.



- The identified edge is followed until the next vertex is reached.

- *The mathematical foundation of the algorithm.* Remember that we are trying to solve a problem in standard form.

$$\min(c^T x)$$
$$Ax = b$$
$$x \geq 0$$

The problem has *n* variables and *m* constraints. Thus:

$$c \in R^n$$
$$b \in R^m$$
$$A^{m \times n}$$

The algorithm starts at a vertex, note that as we discussed before, this vertex algebraically can be described as equivalent to a basic feasible solution, which in turn has a basic matrix associated with it. This basic matrix (B) can be constructed by selecting some of the columns of matrix A. The indexes of the columns used to construct the basic matrix B are given by the set:

$$J^k = \left(J_1^k, J_2^k, \dots J_m^k\right)$$

In this set the super-index k is used to denote the iteration. Note that matrix B is squared and invertible. This matrix can be written in terms of the selected columns as follows:

$$B = [A_{j_1^k}, A_{j_2^k}, \dots, A_{j_m^k}]$$

Where $A_{j_i^k}$ denotes column j_i^k of matrix A.

Using the basic matrix, we can calculate the basic variables, as we have seen

$$x_B = B^{-1}b$$

With the understanding that all non-basic variables are equal to zero (remember that this means that these constraints are active).

$$x_N = 0$$

We will consider that we have a basic feasible direction p, this direction is represented as follows:

$$d_p = \boldsymbol{P}\begin{pmatrix} d_{B_p} \\ d_{N_p} \end{pmatrix}$$

This basic direction will be constructed by taking one variable p from the non-basic variables (current value is zero), and bringing it into the basis (i.e., making its value different than zero), remember that we just did this when we studied basic directions! We have seen that the basic part of this direction can be calculated as follows:

$$d_p = -B^{-1}A_p$$

Where A_p is the pth column of A.

Now we can calculate the directional derivative along this basic direction:

$$\nabla f(x)^T d_p = c^T d_j = \mathrm{c}_B^T\, d_{B_p} + \mathrm{c}_N^T d_{N_p} = -c_B^T B^{-1} A_p + c_p$$

This directional derivative in the context of linear optimization is called the ***reduced cost***.

- *Algorithm*
  - The first thing that we need to define is the input, the linear optimization problem can be presented as follows:

$$\begin{gathered} \min(c^T x) \\ Ax = b \\ x \geq 0 \end{gathered}$$

The input for this algorithm is

  - $c \in R^n, b \in R^m$ and $A^{m \times n}$
  - And a set of indexes $J^0 = (j_1^0, \dots, j_m^0)$ of the basic variables corresponding to a basic feasible solution (this means that we are starting at a vertex of the polyhedron).

Steps of the algorithm

  - Let $B = (A_{j_1^k}, \dots, A_{j_m^k})$ be the matrix formed by the columns of A corresponding to the indices of J_k.
  - Calculate the reduced cost in all basic directions

- If $\bar{c}_j = c_j - c_B^T B^{-1} A_j \geq 0, \forall j \notin J^k$ then optimal solution

  In this step, we are finding a descent directions and selecting one of them to follow, we calculate the reduced cost. If there is no direction in which the reduced cost is negative, i.e., if there is no descent direction, it means that we are at the optimum

- Else p=smallest indext such that $\bar{c}_p < 0$.

  If there is more than one descent direction, we will select one of them, the rule is that we choose the one with the smallest index. Note that this is a bit counterintuitive as one may think that choosing the one with the smallest reduced cost is the better choice, but in the implementation of the algorithm, especially for large problems, using this heuristic has proven more efficient. The reason is that one need to calculate only the first p reduced derivatives and not all of them. This saves space and, in many cases, reduce time.

- Calculate the basic variables $x_B = B^{-1}b$ (here we are finding the coordinates of the point were we are)
- Calculate the basic components of the p^th basic direction $d_b = -B^{-1}A_p$ (here we are finding the direction that we will follow)

  Now that we have chosen a descent direction that we will follow, we proceed to calculate the components (the basic ones) of the vector that defines this descent direction, the non-basic components are obviously equal to zero. Note that we do not need to do this to estimate the reduced cost (i.e., the directional derivative), which we calculated previously.

- For each $i = 1, \dots, m$, calculate the distance to the non-negativity constraint (here we are finding for how long we can move along direction d_b)

$$\alpha_i = \begin{cases} -\dfrac{(x_b)_i}{(d_b)_i} & if\ (d_B)_i < 0 \\ +\infty & otherwise \end{cases}$$

  Once we have identified a descent direction that we are going to follow, then we need to decide for how long we are going to follow this direction (note that we are moving following an edge) to find another corner point.

  The idea is that we can follow this direction until we violate one of the constraints (importantly, note that the polyhedron is in standard form, so the constraints are of the form $x_i \leq 0$). Therefore, we need to identify which constraint is violated first, toward this end, we calculate the distance toward each constraint from the current point. We can do this using the following approach:

$$\begin{pmatrix} x_{B_1} \\ \vdots \\ x_{B_i} \\ \vdots \\ x_{B_N} \end{pmatrix} + \alpha_i \begin{pmatrix} d_1 \\ \vdots \\ d_{B_i} \\ \vdots \\ d_{B_N} \end{pmatrix} = \begin{pmatrix} x_{B_1} + \alpha_i d_1 \\ \vdots \\ 0 \\ \vdots \\ x_{B_N} + \alpha_i d_{B_N} \end{pmatrix}$$

This is repeated for each variable, note that once we hit the constraint, it becomes active (inequality is satisfied with an equality), so the variable that was hit first when following the selected descent direction leaves the basis).

An important point is that if $\alpha_i < 0$, in other words $d_{B_i} > 0$, this means that to hit the constraint, we need to move opposite to the descent direction, this means that if we move in the direction d, we can continue forever, we denote this as $+\infty$.

- Let q be the smallest index such that $\alpha_q = \min_i \alpha_i$
- If $\alpha_q = \infty$ then the problem is unbounded and has no optimal solution
- $J^{k+1} = J^k \cup \{p\} \backslash j_q^k$

  In general, we say that p is entering the base and q is leaving it.
- $k = k + 1$

  These steps are repeated until convergence is achieved.

Example: *consider the following optimization problem (same as before!)*

$$\min(-x_1 - 2x_2)$$
$$s.t. \quad \begin{aligned} x_1 + x_2 &\leq 1 \\ x_1 - x_2 &\leq 1 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix}, b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, c = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 0 \end{pmatrix}$$

Iteration 1

Consider $J^0 = \{3,4\}$, thus

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

If we solve for $x_b = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = B^{-1}b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, remember that non basic variables are zero, then the value of the solution vector for this point is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Note that this is a basic feasible solution, thus we are at a corner point, and we can proceed with the algorithm. The value of the objective function at this point is given by

$$c^T x = 0$$

Now we can calculate the reduced cost in all directions.

$$\bar{c}_j = c_j - c_B^T B^{-1} A_j \geq 0, \forall j \notin J^k$$

Since we need to do it for all variables not in the base, we will calculate the reduced cost for indexes 1 and 2.

$$\bar{c}_1 = c_1 - c_B^T B^{-1} A_1 = -1 - \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1$$

$$\bar{c}_2 = c_2 - c_B^T B^{-1} A_2 = -2 - \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2$$

We choose index $p = 1$ as the variable that is going to leave the basis. Note that in practical implementations of the algorithm it is not necessary to calculate $\bar{c}_2$ since we have already identified the variable with the lower index that gives a descent direction.

Once we identify x_1 as the variable to leave the basis, we proceed to identify the descent direction that we will follow.

$$d_p = \boldsymbol{P} \begin{pmatrix} d_{B_p} \\ d_{N_p} \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} x_3 \\ x_4 \\ x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} ? \\ ? \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ ? \\ ? \end{pmatrix}$$

Note that we have a 1 at the position were the variable x_1 (that will enter the basis) is located. Now we calculate the value of the basic variables

$$d_p = -B^{-1} A_p = -B^{-1} A_1 = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

Then:

$$d_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ -1 \end{bmatrix}$$

Now we calculate the distance to the non-negativity constraints to select the variable that is going to leave the basis, so in this case we do it for x_3 and x_4.

$$\alpha_3 = -\frac{(x_b)_1}{(d_b)_1} = -\frac{1}{-1} = 1$$
$$\alpha_4 = -\frac{(x_b)_2}{(d_b)_2} = -\frac{1}{-1} = 1$$

Note that the 1 and 2 subindexes in the equation represent the first and second element of the basis. Based on these results we can say that we will remove x_3 from the basis. $q = 3$

Then in the original set $J^0 = \{3,4\}$ we replace index $q = 3$ by index $p = 1$, such that the set for the next iteration is $J^1 = \{1,4\}$

Iteration 2

$$J^1 = \{1,4\}$$

$$B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\bar{c}_2 = -1$$
$$\bar{c}_3 = 1$$

$$p = 2$$
$$d_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}$$
$$\alpha_3 = 1$$
$$\alpha_4 = \infty$$
$$q = 1$$
$$J^2 = \{2,4\}$$

Iteration 3

$$J^2 = \{2,4\}$$
$$B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}$$
$$\bar{c}_1 = 1$$
$$\bar{c}_3 = 2$$

Since there is no descent direction, we have found the optimal value, and the value of the objective function is:

$$c^T x = -2$$

- *The simplex tableau.* As we saw in the previous discussion, the simplex algorithm requires a significant computational effort. This computational effort is centered in the calculation of the following quantities

$$c^T - c_B^T B^{-1} A \text{ (Reduced cost)}$$
$$B^{-1} b \text{ (Calculate current value of basic variables)}$$
$$-B^{-1} A_p \text{ (Calculate the basic direction)}$$

The simplex tableau is a computational strategy in which we store the value of $B^{-1}A$ and $B^{-1}b$ instead of the values of the original variables. Let's consider a linear optimization problem in standard form and let us take a basic matrix B corresponding to a basic feasible solution x^*. The table

$B^{-1}A$	$B^{-1}b$
$c^T - c_B^T B^{-1} A$	$-c_B^T B^{-1} b$

is called the simplex tableau, and can be represented in more detail as follows:

n			1	
$B^{-1}A_1$	$\cdots$	$B^{-1}A_n$	x_{j1} $\vdots$ x_{jm}	*m*

$c_1 \quad \cdots \quad c_n$	$-c^T x$	*1*

- Each column of the left side of the tableau corresponds to a variable of the problem.
- Each row of the upper part of the tableau corresponds to a basic variable.
- The *m* first rows of the last column contain the values of the basic variables, the other variables, that is, the non-basic ones, are always zero.
- The last element of the last column contains the value of the objective function with opposite sign.
- It is important to note that if variable I is a basic variable, then its reduced cost is zero, and the column associated with is the same as in an identity matrix.

i	1	
$\vert \;\cdots\; 0 \;\cdots\; \vert$	x_{j1}	
$\vert \;\cdots\; 1 \;\cdots\; \vert$	$\vdots$	*m*
$\vert \;\cdots\; 0 \;\cdots\; \vert$	x_{jm}	
$\vert \;\cdots\; 0 \;\cdots\; \vert$	$-c^T x$	*1*

As an example, let's consider a system as follows:

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix}, b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, c = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 0 \end{pmatrix}$$

If we select variables 2 and 4 as the basic variables then:

$$B = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, B^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Based on this information we can construct the following tableau:

x_1	x_2	x_3	x_4		
1	1	1	0	1	x_2
2	0	1	1	2	x_4
1	0	2	0	2	$-c^T x$

Note that based on these results this tableau is optimal, can you identify the optimal solution?

If a tableau is available, an iteration of the simplex algorithm greatly simplified, since many of the required quantities in the algorithm are available in the tableau can be directly read it. There are two questions that remain:

- How do we update the tableau from one iteration to the next one?
- How do we generate an initial tableau?

- *Pivoting*. This step is aimed at finding theujjujj new tableau using as input the tableau from a previous step. Conceptually we can think of this process as follows:

$$[A_{j1} \quad \cdots \quad A_{jq} \quad \cdots \quad A_{jm}] \rightarrow [A_{j1} \quad \cdots \quad A_p \quad \cdots \quad A_{jm}]$$

$B^{-1}A$	$B^{-1}b$
$c^T - c_B^T B^{-1}A$	$-c_B^T B^{-1}b$

That gets transformed into:

$\bar{B}^{-1}A$	$\bar{B}^{-1}b$
$c^T - c_B^T \bar{B}^{-1}A$	$-c_B^T \bar{B}^{-1}b$

This means that between iterations column q gets replaced with column p. The problem is that if we do this naively, we will need to repeat all the calculations. The good news is that we can circumvent these calculations by using the information already contained in the tableau. Formally, we are interested in finding a transformation Q such that

$$QB^{-1} = \bar{B}^{-1} \rightarrow QB^{-1}\bar{B} = I$$

This means that matrix Q transforms $B^{-1}\bar{B}$ into an identity matrix, the key to move forward is to understand that B and $\bar{B}$ are very similar and differ only by one column. Consequently, their product looks something like this (i.e., almost the identity matrix except for one column):

$$\begin{bmatrix} 1 & 0 & | & u_1 & | & 0 \\ 0 & 1 & | & u_2 & | & 0 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ \vdots & \vdots & | & u_q & | & \vdots \\ \vdots & \vdots & | & \vdots & \ddots & \vdots \\ 0 & 0 & | & u_m & | & 1 \end{bmatrix}$$

Where we have that $u = -d_p = B^{-1}A_p$ (note that this is the value of the column with index p that we are trying to bring into the basis). We must transform this equation into the identity matrix, this can be easily achieved by using elementary row operations (same as those used in Gauss elimination). These elementary row operations are of two types:

- If we are at row q, we will divide by u_q such that we obtain the unity.
- If we are at any other row, then we add the row q multiplied by $-u_i/u_q$.

These elementary row operations can be represented as a matrix, such that, if we define an elementary row operation on A as consisting in multipliying row j by a constant β and adding the result to row i, then

$$a_i = a_i + \beta a_j$$

Where a_i denotes the ith row of A. This operation consists in multiplying A by the matrix Q_{ij}, which is the identity matrix, of which element (I,j) is replaced by β. When several of these operations are applied sequentially, one can create an overall Q resulting from the composition. This intuition can be used to transform the simplex tableau, in which elementary row operations are used to obtain a new tableau in which one of the non-basic variables is transformed into a basic one (rember that basic variables have zero reduced cost and they have only one non zero element on the upper right part of the tableau).

$QB^{-1}A$	$QB^{-1}b$	=	$\bar{B}^{-1}A$	$\bar{B}^{-1}b$

Example: Let's consider a simplex tableau given below, let's try to identify the variable leaving the basis and the variable entering the basis, the pivot, and then let's try to calculate the new tableau. In the example the number highlighted in red corresponds to the pivot.

x_1	x_2	x_3	x_4	x_5	x_6			
0	1.5	1	1	-0.5	0	10	x_4	$\alpha = \frac{10}{1.5}$
1	0.5	1	0	0.5	0	10	x_1	$\alpha = \frac{10}{0.5}$
0	1	-1	0	-1	1	0	x_6	$\alpha = \frac{0}{1}$
0	-7	-2	0	5	0	100		

x_1	x_2	x_3	x_4	x_5	x_6			
0	0	2.5	1	1	-1.5	10	x_4	$\alpha = \frac{10}{2.5}$
1	0	1.5	0	1	-0.5	10	x_1	$\alpha = \frac{10}{1.5}$
0	1	-1	0	-1	1	0	x_6	$\alpha = \frac{0}{-1}$
0	0	-9	0	-2	7	100		

- *The initial tableau*. The last problem that we need to deal with is that of calculating the first tableau. To do this we will use the simplex algorithm, but instead of searching for the optimal solution, we will look for a feasible solution. Let's consider the following problem:

$$\min c^T x$$

$$s.t. \begin{array}{c} Ax = b \\ x \geq 0 \\ A \in R^{m\times n}, b \in R^m, c \in R^n \\ b \geq 0 \end{array}$$

To find the feasible solution, we will use two tricks (1) we will define a positive auxiliary variable per each equality constraint (2) we will forget about the objective function and replace it by the sum of the auxiliary variables as follows:

$$\min \sum_{i=1}^{m} x_i^{aux}$$

$$s.t. \quad \begin{gathered} Ax + Ix^{aux} = b \\ x \geq 0 \\ x^{aux} \geq 0 \\ A \in R^{m\times n}, b \in R^m, c \in R^n \\ b \geq 0 \end{gathered}$$

Note that this problem can be solved using the simplex algorithm, but now it is extremely simple to find the initial tableau, because you just say that the basis variables are all the auxiliary variables, then $x = 0$ and $Ix^{aux} = b$, thus your system is feasible.

Once the problem is solved, it is possible to reuse some of the information contained in the tableau to generate the initial tableau of the original problem, to that end we need to do couple of things:

- We need to check which variables ended up in the basis for the final solution, if any of the auxiliary variables is in the final solution, then we need to remove it from the basis by pivoting it out (exchanging it with one variable in the original problem)
- Once we have pivoted the required variables, and all auxiliary variables are out of the basis, then we remove the corresponding columns.
- Finally, before we solve the original problem we recalculate the last row of the tableau.

Example: Now we will develop an example in which we explore the complete simplex algorithm for a problem, this algorithm is called the two step simplex, in the first part we will find the initial tableau, and in the second part we will look for the optimal solution. The problem can be presented as follows:

$$\begin{gathered} \min(2x_1 + 3x_2 + 3x_3 + x_4 - 2x_5) \\ x_1 + 3x_2 + 4x_4 + x_5 = 2 \\ x_1 + 2x_2 - 3x_4 + x_5 = 2 \\ -x_1 - 4x_2 + 3x_3 = 1 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{gathered}$$

Note that in this case we have that:

$$A = \begin{bmatrix} 1 & 3 & 0 & 4 & 1 \\ 1 & 2 & 0 & -3 & 1 \\ -1 & -4 & 3 & 0 & 0 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, c = \begin{bmatrix} 2 \\ 3 \\ 3 \\ 1 \\ -2 \end{bmatrix}$$

Phase I: First, we will find the initial tableau, we define auxiliary variables for each of the equality constraints such that the auxiliary problem can be written as follows:

$$\min(x_1^{aux} + x_2^{aux} + x_3^{aux})$$

$$x_1 + 3x_2 + 4x_4 + x_5 + x_1^{aux} = 2$$
$$x_1 + 2x_2 - 3x_4 + x_5 + x_2^{aux} = 2$$
$$-x_1 - 4x_2 + 3x_3 + x_3^{aux} = 3$$
$$x_1, x_2, x_3, x_4, x_5, x_1^{aux}, x_2^{aux}, x_3^{aux} \geq 0$$

Note that for this auxiliary problem we can also define the matrixes and vectors characteristic of an LP:

$$A^{aux} = \begin{bmatrix} 1 & 3 & 0 & 4 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & -3 & 1 & 0 & 1 & 0 \\ -1 & -4 & 3 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, b^{aux} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, c^{aux} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Based on this information we can construct the initial tableau, note that if we choose the auxiliary variable to be in the basis, then we get the following matrix B:

$$B = B^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

x_1	x_2	x_3	x_4	x_5	x_1^{aux}	x_2^{aux}			α
1	3	0	4	1	1	0	0	2=x_1^{aux}	2
1	2	0	-3	1	0	1	0	2=x_2^{aux}	2
-1	-4	3	0	0	0	0	1	1=x_3^{aux}	inf
-1	-1	-3	-1	-2	0	0	0		

Remember that to calculate the last row, we have used the definition of the reduced cost: $\bar{c}_j = c_j - c_B^T B^{-1} A_j$.

x_1	x_2	x_3	x_4	x_5	x_1^{aux}	x_2^{aux}	x_3^{aux}		α
1	3	0	4	1	1	0	0	2=x_1	Inf
0	-1	0	-7	0	-1	1	0	0=x_2^{aux}	inf
0	-1	3	0	4	1	1	0	1=x_3^{aux}	1/3
0	2	-3	3	-1	1	0	0		

x_1	x_2	x_3	x_4	x_5	x_1^{aux}	x_2^{aux}	x_3^{aux}		α
1	3	0	4	1	1	0	0	2=x_1	
0	-1	0	-7	0	-1	1	0	0=x_2^{aux}	
0	-1/3	1	4/3	1/3	1/3	0	1/3	1=x_3	
0	1	0	7	0	2	0	1		

This last tableau is optimal! but this is just for the auxiliary problem, now we need to transform this tableau into the initial tableau for the full problem. The problem with this tableau is that it has one auxiliary variable as part of the basis, to be able to use the tableau to generate the one for the original problem we will pivot the auxiliary variable out of the basis. We can do this by selecting the element at the intersection between a column with a non-basic variable in the original problem (in this case we will use x_2) and the row associated with x_2^{aux}. Note that you need a good pivot here, that is x_5 would not be a good choice because the coefficient is zero, and we cannot pivot.

x_1	x_2	x_3	x_4	x_5	x_1^{aux}	x_2^{aux}	x_3^{aux}		α
1	3	0	4	1	1	0	0	2=x_1	
0	-1	0	-7	0	-1	1	0	0=x_2^{aux}	
0	-1/3	1	4/3	1/3	1/3	0	1/3	1=x_3	
0	1	0	7	0	2	0	1		

The resulting tableau after pivoting is shown below:

x_1	x_2	x_3	x_4	x_5	x_1^{aux}	x_2^{aux}	x_3^{aux}		α
1	0	0	-17	1	-2	3	0	2=x_1	
0	1	0	7	0	1	-1	0	0=x_2	
0	0	1	3.67	1/3	2/3	-1/3	1/3	1=x_3	
0	0	0	0	0	1	1	1		

Now the only thing left is to remove the auxiliary variables. And we get the following tableau. Note that I have intentionally left blank the reduced cost, as the data that we have is for the auxiliary problem, and we need to calculate again for the original one.

x_1	x_2	x_3	x_4	x_5		α
1	0	0	-17	1	2=x_1	
0	1	0	7	0	0=x_2	
0	0	1	3.67	1/3	1=x_3	

Phase II: Finally, we use the initial tableau estimated as the starting point for the second part of the algorithm.

Thankfully the reduced cost for the basic variable is zero, for the non-basic variables we do the following estimation (note that the value of $B^{-1}A_p$ can be directly read from the previous simplex tableau.

$$\bar{c}_4 = c_4 - c_B^T B^{-1} A_4 = 1 - \begin{bmatrix} 2 & 3 & 3 \end{bmatrix} \begin{bmatrix} -17 \\ 7 \\ 3.67 \end{bmatrix} = 3$$

$$\bar{c}_5 = c_5 - c_B^T B^{-1} A_5 = -2 - [2 \quad 3 \quad 3]\begin{bmatrix}1\\0\\\frac{1}{3}\end{bmatrix} = -5$$

x_1	x_2	x_3	x_4	x_5		α
1	0	0	-17	1	2=x_1	2
0	1	0	7	0	0=x_2	Inf
0	0	1	3.67	1/3	1=x_3	3
0	0	0	3	-5	-7	

x_1	x_2	x_3	x_4	x_5		α
1	0	0	-17	1	2=x_5	
0	1	0	7	0	0=x_2	0
-1/3	0	1	9.33	0	1/3=x_3	0.04
5	0	0	-82	0	3	

x_1	x_2	x_3	x_4	x_5		α
1	2.43	0	0	1	2=x_5	
0	0.14	0	1	0	0=x_4	0
-1/3	-1.33	1	0	0	1/3=x_3	0.04
5	11.71	0	0	0	3	

- **Some notes on the interpretation of optimization results obtained from the optimal simplex tableau:**
  - The first thing that we need to consider is the reduced cost at the optimal point, these values are informative in a number of ways
    - For basic variables its value is zero.
    - For non-basic variables its value is greater or equal to zero
      - If the value is greater than zero, that value has two interpretations
        - The reduced cost is the directional derivative, a value greater than zero tells us by how much the objective function changes if the value of the non-basic variable increases by one
        - The reduced cost can also be read as the change in the objective function coefficient that is required for the non-basic variable to be able to enter the basis.
      - In some instances, non-basic variables have a value that is equal to zero, if that is the case it means that we have a degenerate solution. This implies that there is more than one solution. This can be good news, because if for some reason we do not like or find useful the first solution that we find, we can always find an alternative solution. v