Università della Svizzera italiana | Institute of Computing CI

Optimization Methods **Spring 2024**

Student: Birindelli Leonardo

Fourth assignment

Due date: Monday, 3 June 2024, 12:00 AM

Exercise 1 (20/100)

Consider the quadratic function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as:

$$f(x) = 7x^2 + 4xy + y^2 \tag{1}$$

where $\mathbf{x} = (x, y)^T$.

1. Write this function in canonical form, i.e. $f(x) = \frac{1}{2}x^T Ax - b^T x + c$, where A is a symmetric matrix.

   The function $f(x)$ can be written according to the canonical form definition previously described as follows:

$$f(x) = \frac{1}{2}\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 14 & 4 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

   It can be easily noticed that the matrix A is symmetric, the vector b is equal to $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$ because there are not any 1° grade polynomials in the expression and the constant c is equal to 0.
   We only need to prove that matrix A is positive definite, mathematically speaking it is required to prove that all its eigenvalues are stricly positive definite $\lambda > 0$, it follows that :

$$\begin{aligned}
\det(A - \lambda I) &= \det\left(\begin{bmatrix} 14-\lambda & 4 \\ 4 & 2-\lambda \end{bmatrix}\right) \\
&= (14-\lambda)(2-\lambda) - 4 \cdot 4 \\
&= \lambda^2 - 16\lambda + 8 \\
\det(A - \lambda I) &= 0 \Longrightarrow \lambda^2 - 16\lambda + 8 = 0 \\
\lambda &= \frac{16 \pm \sqrt{16^2 - 4\cdot 8}}{2} = \frac{16 \pm \sqrt{256-32}}{2} = \frac{16 \pm \sqrt{224}}{2} = \frac{16 \pm 4\sqrt{14}}{2} = 8 \pm 2\sqrt{14} \\
\lambda_1 &= 8 + 2\sqrt{14} \approx 15.48 > 0 \\
\lambda_2 &= 8 - 2\sqrt{14} \approx 0.52 > 0
\end{aligned}$$

Now, we have proved that the matrix A is symmetric and positive definite. Then, It is possible to prove that the canonical form is correct for the function $f(x)$ by computing the product of the matrices and the vectors as follows:

$$
\begin{aligned}
f(x) &= \frac{1}{2}\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 14 & 4 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + 0 \\
&= \frac{1}{2}(14x^2 + \underline{4xy} + \underline{4xy} + 2y^2) \\
&= 7x^2 + 4xy + y^2
\end{aligned}
$$

2. Describe briefly how the Conjugate Gradient (CG) Method works and discuss whether it is suitable to minimize f from equation (1). Explain your reasoning in detail **(max. 30 lines)**.

The Conjugate Gradient method is an optimization algorithm for solving linear systems of equations in the form $Ax = b$ where A is a symmetric positive definite matrix. The linear system can be interpreted as a minimization problem for a quadratic function with the matrix A s.p.d. which it can be described as :

$$\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2}x^T Ax - b^T x \tag{2}$$

The CG method aims to find the minimum of the function $f(x)$ by generating a set of vectors that guarante conjugacy with respect to the symmetric positive definite matrix A, allowing the method to converge in at most n steps. In this case, the iterative method is suitable to minimize the function $f(x)$ from equation (1) because, as I previously proved, the function can be written in the canonical form and it is possible to notice that expression corresponds exactly to the quadratic form expression as described in the minimization problem (2). Thus, we know the matrix A is symmetric positive definite, which is the necessary conditions for the method to work properly. It is worth noting that the CG is used generally for problems in which we have a large sparse matrix due to efficiency in terms of memory usage and computational complexity at each iteration (which requires a matrix-vector multiplication) and its negative point is the fact that it is affected by rounding errors. In this case, where the matrix A is very small, it would have been better to use the Gauss elimination or a factorization methods instead of the CG.
To conclude with the Conjugate Gradient method, it is required to talk about its convergence rate which is a property that directly depends on the eigenvalues distribution of the matrix: in our case, the eigenvalues distribution of matrix A is only composed by two eigenvalues which their difference is pronounced but this does not necessarily compromise the convergence of the method. Generally, the method converges slowly in the case in which the eigenvalues are distributed uniformely because it relies on the distances between of them. On the other hand, a faster convergence is reached when the eigenvalues are clustered together, this is because the method exploits the marginal differences between the eigenvalues to accelerate the descent direction towards the minimum.

Exercise 2

Consider the following constrained minimization problem for $\mathbf{x} = (x, y, z)^T$

$$\begin{aligned} \min f(x) &:= -3x^2 + y^2 + 2z^2 + 2(x + y + z) \\ \text{s.t. } c(\mathbf{x}) &= x^2 + y^2 + z^2 - 1 = 0 \end{aligned} \tag{3}$$

Write down the Lagrangian function and derive the KKT conditions for 3

To begin with, the Lagrangian function is a mathematical tool used to solve minimization or maximization optimization problems in which the objective function is subjected to a set of constraints that describe which is the feasible area of solutions. The Lagrangian function is defined as follows:

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i=1}^{n} \lambda_i c_i(x)$$

where $x \in \mathbb{R}^n$, $f(x)$ is the objective function, $c_i(x)$ are the constraints which must be satisfied and λ_i are the Lagrange multipliers which are scalar values used as coefficients in the linear combination of the constraint gradients. Moreover, those scalar values are required to guarantee that the given solution is both a stationary point of the initial function and stationary point for the Lagrangian function.
As it is possible to notice, the optimization problem is subjected to only one equality constraint in this problem, thus, it can be possible to simplify the Lagrangian function (3) as follow:

$$\begin{aligned} \mathcal{L}(x, \lambda) &= f(x) - \lambda c(x) \\ \nabla_x \mathcal{L}(x, \lambda) &= \nabla f(x) - \lambda \nabla c(x) \end{aligned} \tag{4}$$

As we proceed, it becomes important to also introduce the *Karush-Kuhn-Tucker conditions*: those requisites, also known as KKT conditions, are a set of first-order necessary conditions used to define if a solution is optimal for a constrained optimization problem or not. The KKT conditions are a generalization of the Lagrange multipliers method to equality and inequality constraints.
The theorem behind those necessary conditions, as introduced at the section ***"Theorem 12.1 (First-Order Necessary Conditions)"*** *page 321 of "Jorge Nocedal and Stephen J Wright. Numerical optimization. Springer, 1999"(1)* describes that:

Suppose that x^* is a local solution of (12.1), that the functions f and c_i in (12.1) are continuously differentiable, and that the LICQ holds at x^*. Then there is a Lagrange multiplier vector λ^*, with components λ_i^*, $i \in \mathcal{E} \cup \mathcal{I}$, such that the following conditions are satisfied at (x^*, λ^*)

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*) &= 0, \\ c_i(x^*) &= 0, \quad \text{for all } i \in \mathcal{E}, \\ c_i(x^*) &\geq 0, \quad \text{for all } i \in \mathcal{I}, \\ \lambda_i^* &\geq 0, \quad \text{for all } i \in \mathcal{I}, \\ \lambda_i^* c_i(x^*) &= 0, \quad \text{for all } i \in \mathcal{E} \cup \mathcal{I}. \quad (\textit{complementary conditions}) \end{aligned}$$

where $\mathcal{E}$ and $\mathcal{I}$ are respectively the sets of equality and inequality constraints, while the *complementary conditions* imply whether c_i is an equality or an active inequality constraint, its value is $c_i(x^*) = 0$; in the opposite case, we have that $\lambda_i^* = 0$.

Furthermore, the LICQ refers to the *Linear Independence Constraint Qualification property* that must be verified by the active constraint gradients at the point x^*. This attribute is important to guarantee that the point x^* is a local solution for the problem by verifying that the gradients of active constraints are not linearly dependent between one another at that specific point; a deeper explanation of the LICQ property and the active set can be found respectively in the sections **"Definition 12.4 (LICQ)"** *at page 320* and **"Definition 12.1"** *at page 308 of "Jorge Nocedal and Stephen J Wright. Numerical optimization. Springer, 1999"*(1).

In the case of the problem 3, the KKT conditions can be derivated in the following conditions:

$$\begin{aligned} \nabla_x \mathcal{L}(x^*, \lambda^*) &= 0, \\ c_0(x^*) &= 0, \quad \text{for all } i \in \mathcal{E}, \\ \lambda_0^* c_0(x^*) &= 0, \quad \text{for all } i \in \mathcal{E} \cup \mathcal{I} \end{aligned}$$

where $\mathcal{I} = \emptyset$, $\mathcal{E} = \{0\}$ in which $i = 0$ correspond to the unique equality constraint of the problem.

Exercise 3

1. Read the chapter on Simplex method, in particular the section 13.3 The Simplex Method, in Numerical Optimization, Nocedal and Wright. Explain how the method works, with a particular attention to the search direction.

   To begin with, linear programming is an optimization technique used to minimize or maximize a linear objective function subject to a set of linear constraints, which can be both equalities and inequalities. Generally, linear programs are described and analyzed in the *standard form*, which is an important formulation necessary for optimization algorithms in order to resolve the problems. For minimization problems, it is described as follows:

$$\begin{aligned} &\min\ c^T x \\ &\text{s.t. } Ax \geq b \\ &\qquad x \geq 0 \quad (\textit{non-negativity constraints}) \end{aligned}$$

   where $c, x \in \mathbb{R}^n$ are respectively the vector of coefficients and the one of unknowns, $A \in \mathbb{R}^{m \times n}$ is a full rank matrix containing the coefficients of constraints. Notice that the procedure to find the optimal value in this kind of problem is described by the **Fundamental Theorem of Linear Programming**(3) which explains that :

   If a linear program admits a solution (i.e., if it is neither unfeasible nor unbounded), this solution will lie on a vertex of the polytope defined by the feasible region. In the case in which two vertices are both maximisers (or minimisers) of the objective function, then also all the points lying on the line segment between them will represent optimal solutions of the linear programming problem.

   The Simplex method is the most widely used algorithm to solve linear programming problems, in particular, it is used when there are many variables to be optimized with several constraints. The idea behind its computations is to move from one vertex of set of basic feasible points to another in order to find the optimal solution.

Geometrically, the method, starts from an initial vertex, moves along the edges of the polytope until it reaches the optimal vertex that minimizes or maximizes the initial problem. The Simplex method is based on the concept of a *basis* and *non-basis* variables, indeed the matrix A is divided in two sub matrixes $B \in \mathbb{R}^{m\times m}$, which is a squared and invertible matrix with m linearly independent columns that contains the coefficients of the actual basic solution, and $N \in \mathbb{R}^{m\times n}$, which contains the remaining columns. Moreover, it is also required to split the vectors x and c in two sub-vectors $x_B, c_B \in \mathbb{R}^{m\times 1}$ and $x_N, c_N \in \mathbb{R}^{n\times 1}$: x_B and x_N are respectively the basic and non-basic variables. Through this splitting, it is possible to rewrite the system $Ax = b$ as follows:

$$\begin{aligned} Ax &= b \\ Bx_B + Nx_N &= b \\ Bx_B &= b - Nx_N \\ x_B &= B^{-1}b - B^{-1}Nx_N \end{aligned}$$

The previous formula represents the general solution for computing the optimal basis solution; notice that if $x_N = 0$, then $x_B = B^{-1}b$ is a basic solution if the non-negativity constraint is verified. The iterative procedure of finding the solution around the vertices of the polytope continues until the optimality condition or pricing rule is satisfied. The pricing rule is a method used to determine the entering variable in the basis, which is the variable that will be added to the basis in order to reach the optimal solution. The optimality condition is a rule that determines when the algorithm has reached the optimal solution, which is when the reduced costs of the non-basic variables are all non-negative. That rule can be described as follows:

$$r_N = c_N^T - c_B^T B^{-1} N$$

For minimization problems the pricing is reached when $r_N \geq 0$. If the condition is not reached the Simplex method aims to interchange the basic and non-basic variables in order to get closer to the optimal solution.
Referring to the optimal solution, it is required to notice that the KKT conditions can be applied also for linear programming problems given that we are working on a constrained problem thus it is possible to extend them to this kind of optimization problem. It can be derived by the (***First-Order Necessary Conditions***) that:
The Lagragian function is defined as follows:

$$\begin{aligned} \mathcal{L}(x,\lambda,s) &= c^T x + \lambda^T (Ax - b) - s^T x \\ \nabla\mathcal{L}(x,\lambda,s) &= c^T - \lambda^T A - s^T \end{aligned}$$

Hence, the KKT conditions can be written as follows:

$$\begin{aligned} \nabla\mathcal{L}(x,\lambda,s) &= 0 \\ Ax &= b \\ x &\geq 0 \\ s &\geq 0 \\ x_i s_i &= 0, \quad i = 1, 2, ..., n \end{aligned}$$

where $\lambda \in \mathbb{R}^m$ represents the vector of Lagrange multipliers, while $s \in \mathbb{R}^n$ is the vector of Lagrangian multipliers for inequalities constraints $x \geq 0$. In practice, the simplex method verifies these conditions implicitly through iterations and pivoting operations as I will introduce soon.

The application of the iterative rule is crucial because if the optimality condition is not reached, we need to find a new direction to move towards the next optimal solution among the vertices of the polytope that could verify the pricing rule. For minimization problems, we firstly need to identify the variable with the lowest value of reduced cost coefficients r_N; Secondly, it is necessary to identify the departing variable which will be extracted from the basis and replaced by the new entering variable. The departing variable will be set to zero in the next iteration, thus, requiring us to find the smallest positive ratio between the right-hand side of the constraints and the column of the basis matrix corresponding to the entering variable. The ratio is calculated as follows:

$$\frac{B^{-1}b}{B^{-1}N}$$

Thirdly the interchange operation (also named *pivoting*) between the basic and non-basic variables is performed and the sub-matexes B and N are updated at each iteration of the process until the optimality condition is satisfied .

2. Consider the following contrained minimization problem, $\mathbf{x} = (x_1, x_2)^T$;

$$\min_x f(x) = 4x_1 + 3x_2 \tag{5}$$

subject to:

$$\begin{aligned} 6 - 2x_1 - 3x_2 &\geq 0 \\ 3 + 3x_1 - 2x_2 &\geq 0 \\ 5 - 2x_2 &\geq 0 \\ 4 - 2x_1 - x_2 &\geq 0 \\ x_2 &\geq 0 \\ x_1 &\geq 0 \end{aligned}$$

a) Sketch the feasible region for this problem.

Firstly, let us analyze the inequality constraints to determine the boundaries of the feasible region. The optimization problem is restricted by the following constraints:

$$\begin{aligned} 6 - 2x_1 - 3x_2 &\geq 0 \\ 3 + 3x_1 - 2x_2 &\geq 0 \\ 5 - 2x_2 &\geq 0 \\ 4 - 2x_1 - x_2 &\geq 0 \\ x_2 &\geq 0 \\ x_1 &\geq 0 \end{aligned}$$

Now, by solving the system of inequality constraints in terms of x_2, we can determine the boundaries of the feasible region of the problem (5), what I obtained is the following:

$$\begin{cases} 6 - 2x_1 - 3x_2 \geq 0 & \rightarrow x_2 \leq -\frac{2x_1}{3} + 2 \\ 3 + 3x_1 - 2x_2 \geq 0 & \rightarrow x_2 \leq \frac{3}{2}x_1 + \frac{3}{2} \\ 5 - 2x_2 \geq 0 & \rightarrow x_2 \leq \frac{5}{2} \\ 4 - 2x_1 - x_2 \geq 0 & \rightarrow x_2 \leq -2x_1 + 4 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

Secondly , I represent the feasible area marked by the constraints on the Cartesian plane and I obtain:

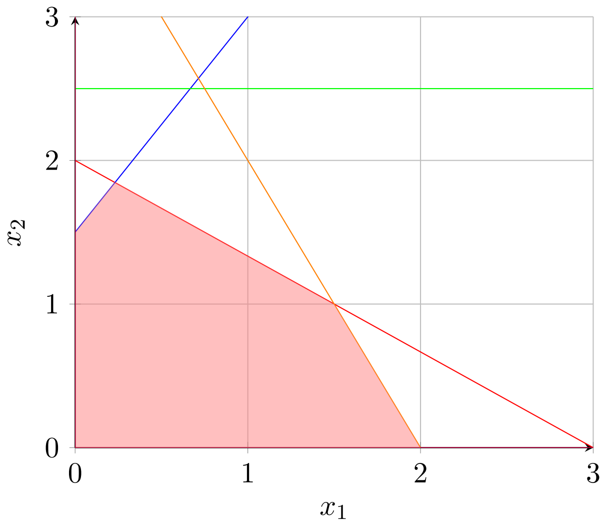

Figure 1: Feasible region for the problem (5)

Note that I decide to color each straight line with different dyes in order to easily understand the plane: The red line represents the line equation of the constraint $x_2 \leq -\frac{2x_1}{3} + 2$, the green line is the equation corresponding to the constraint $x_2 \leq \frac{5}{2}$, the blue one is referring to the line equation of the constraint $x_2 \leq \frac{3}{2}x_1 + \frac{3}{2}$, the orange one is referring to the line equation of the constraint $x_2 \leq -2x_1 + 4$ and ,in conclusion, the barely visible purple lines correspond to the non-negativity constraints $x, y \geq 0$

b) Which are the basic feasible points of the problem (5)? Compute them by hand using the geometrical interpretation and find the optimal point $\mathbf{x}^*$ that minimizes f subject to the constraints.

First of all, I need to find the intersection points between the constraints in order to find the basic feasible points. As the theorem described in the section **"Theorem 13.3 "** *page 365 of "Jorge Nocedal and Stephen J Wright. Numerical optimization. Springer, 1999"*(1) affirms that all the basic feasible points are always located at vertexes of the feasible region. Therefore, it is required to look for the point coordinates of the ones that delimitate the feasible area through the computation of the intersections between the different inequalities.

To begin with, it is really easy to notice that one point of the basic feasible points is the origin $A(0,0)$ because it is the intersection between the non-negativity constraints $x_1 \geq 0$ and $x_2 \geq 0$.
Subsequently, I aimed to find the intersection point between y-axis and $x_2 \leq \frac{3}{2}x_1 + \frac{3}{2}$, thus, what I obtained is the following result:

$$\begin{cases} x_1 = 0 \\ x_2 = \frac{3}{2}x_1 + \frac{3}{2} \end{cases} \rightarrow \begin{cases} x_1 = 0 \\ x_2 = \frac{3}{2} \cdot 0 + \frac{3}{2} = \frac{3}{2} \end{cases}$$

The intersection point between those two lines is $B(0, \frac{3}{2})$.

Secondly, I found the intersection point between $x_2 \leq \frac{3}{2}x_1 + \frac{3}{2}$ and $x_2 \leq -\frac{2x_1}{3} + 2$, thus, the computations are the following:

$$\begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ x_2 = \frac{3}{2}x_1 + \frac{3}{2} \end{cases} \rightarrow \begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ -\frac{2x_1}{3} + 2 = \frac{3}{2}x_1 + \frac{3}{2} \end{cases}$$

$$\begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ \frac{-4x_1+12}{\cancel{6}} \cdot \cancel{6} = \frac{9x_1+9}{\cancel{6}} \cdot \cancel{6} \end{cases} \rightarrow \begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ -4x_1 - 9x_1 = -12 + 9 \end{cases}$$

$$\begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ -13x_1 = -3 \end{cases} \rightarrow \begin{cases} x_2 = -\frac{2 \cdot \frac{3}{13}}{3} + 2 \\ x_1 = \frac{3}{13} \end{cases}$$

$$\begin{cases} x_2 = -\frac{2}{13} + 2 \\ x_1 = -\frac{3}{13} \end{cases} \rightarrow \begin{cases} x_2 = \frac{24}{13} \\ x_1 = \frac{3}{13} \end{cases}$$

The intersection point between those two lines is $C(\frac{3}{13}, \frac{24}{13})$.

Thirdly, I found the intersection point between $x_2 \leq -\frac{2x_1}{3} + 2$ and $x_2 \leq -2x_1 + 4$, thus, the computations are the following:

$$\begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ x_2 = -2x_1 + 4 \end{cases} \rightarrow \begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ -\frac{2x_1}{3} + 2 = -2x_1 + 4 \end{cases}$$

$$\begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ \frac{-2x_1+6}{\cancel{3}} \cdot \cancel{3} = -\frac{-6x_1+12}{\cancel{3}} \cdot \cancel{3} \end{cases} \rightarrow \begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ -2x_1 + 6 = -6x_1 + 12 \end{cases}$$

$$\begin{cases} x_2 = -\frac{2x_1}{3} + 2 \\ 4x_1 = 6 \end{cases} \rightarrow \begin{cases} x_2 = -\frac{2 \cdot \frac{3}{2}}{3} + 2 \\ x_1 = \frac{\cancel{6}^3}{\cancel{4}^2} \end{cases}$$

$$\begin{cases} x_2 = -1 + 2 \\ x_1 = \frac{3}{2} \end{cases} \rightarrow \begin{cases} x_2 = 1 \\ x_1 = \frac{3}{2} \end{cases}$$

The intersection point between those two lines is $D(\frac{3}{2}, 1)$.

To conclude with the intersections computations, I calculated the final intersection point between $x_2 \leq -2x_1 + 4$ and the x-axis , thus, the computations are the following:

$$
\begin{cases} x_2 = -2x_1 + 4 \\ x_2 = 0 \end{cases} \rightarrow \begin{cases} x_2 = -2x_1 + 4 \\ 0 = -2x_1 + 4 \end{cases}
$$
$$
\begin{cases} x_2 = -2x_1 + 4 \\ 2x_1 = 4 \end{cases} \rightarrow \begin{cases} x_2 = -2 \cdot 2 + 4 \\ x_1 = 2 \end{cases}
$$
$$
\begin{cases} x_2 = 0 \\ x_1 = 2 \end{cases}
$$

The intersection point between those two lines is $E(2, 0)$.

To sum up, the basic feasible points are $A(0, 0)$, $B(0, \frac{3}{2})$, $C(\frac{3}{13}, \frac{24}{13})$, $D(\frac{3}{2}, 1)$ and $E(2, 0)$; Now, we need to identify which point from the set is the one that minimized the function in order to resolve the optimization problem, thus, I evaluate the objective function at those vertices and pick the minimum value:

$$
\begin{aligned}
&\text{at vertex } A(0,0) : f(A) = 4 \cdot 0 + 3 \cdot 0 = 0 \\
&\text{at vertex } B(0, \frac{3}{2}) : f(B) = 4 \cdot 0 + 3 \cdot \frac{3}{2} = \frac{9}{2} = 4.5 \\
&\text{at vertex } C(\frac{3}{13}, \frac{24}{13}) : f(C) = 4 \cdot \frac{3}{13} + 3 \cdot \frac{24}{13} = \frac{12}{13} + \frac{72}{13} = \frac{84}{13} \approx 6.46 \\
&\text{at vertex } D(\frac{3}{2}, 1) : f(D) = \cancel{4}^{2} \cdot \frac{3}{\cancel{2}^{1}} + 3 \cdot 1 = 6 + 3 = 9 \\
&\text{at vertex } E(2,0) : f(E) = 4 \cdot 2 + 3 \cdot 0 = 8
\end{aligned}
$$

In conclusion, it can be noticed that the point $A(0, 0)$ is the coordinate that minimize the objective function (5), then , $x_1 = 0$ and $x_2 = 0$ is the optimal solution for the linear programming problem. Next, I added the graph of the feasible region with the set of basic optimal solutions:

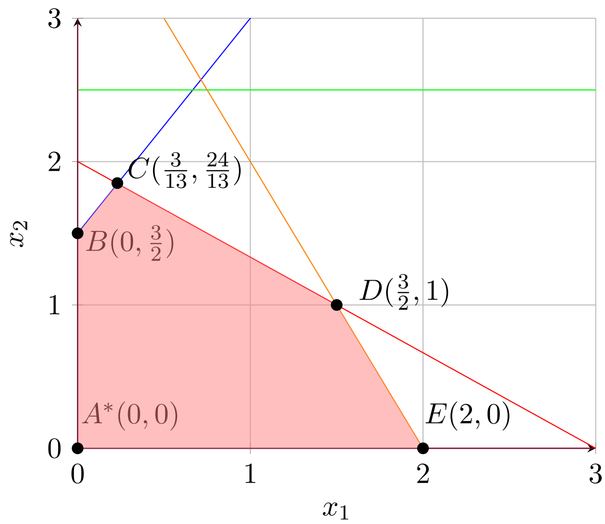

Figure 2: Feasible region for the problem (5)

c) Prove that the first order necessary conditions holds for the optimal point.

To begin with the verification of the first order necessary conditions, as first task, we need to remark the KKT conditions: the general case for these conditions has been explained in the ***Theorem 12.1*** in the previous exercise; in particular, KKT conditions for this constrained optimization can be described as following:

$$\begin{aligned}
\nabla_x \mathcal{L}(x^*, \lambda^*) &= 0, \\
c_i(x^*) &\geq 0, \quad \text{for all } i \in \mathcal{I}, \\
\lambda_i^* &\geq 0, \quad \text{for all } i \in \mathcal{I}, \\
\lambda_i^* c_i(x^*) &= 0, \quad \text{for all } i \in \mathcal{E} \cup \mathcal{I}.
\end{aligned}$$

where $\mathcal{E} = \emptyset$, $\mathcal{I} = \{0, 1, 2, 3, 4, 5\}$ in which $i = 0, 1, 2, 3, 4, 5$ correspond to the inequality constraints in the problem.
After this remark, it is necessary to compute the gradient of the objective function and the active constraints in order to evaluate the KKT conditions at the optimal point $A(0, 0)$. Remark: the active set of constraints in which given the feasible point x^* the conditions are $c_i(x^*) = 0$, mathematically $\mathcal{A} = \mathcal{E} \cup \{i \in \mathcal{I} : c_i(x^*) = 0\}$, therefore, the gradient of the objective function and the constraints are calculated as follows:

$$\nabla f(x) = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

Now, let us find the active constraints at the point $A(0, 0)$, hence:

$$\begin{cases}
6 - 2 \cdot 0 - 3 \cdot 0 & \rightarrow 6 \geq 0 \\
3 + 3 \cdot 0 - 2 \cdot 0 & \rightarrow 3 \geq 0 \\
5 - 2 \cdot 0 & \rightarrow 5 \geq 0 \\
4 - 2 \cdot 0 - 0 & \rightarrow 4 \geq 0 \\
0 \geq 0 & \rightarrow 0 = 0 \\
0 \geq 0 & \rightarrow 0 = 0
\end{cases}$$

Then, the active constraints at the point $A(0, 0)$ are the constraints c_4 and c_5,thus, $\mathcal{A} = \{4, 5\}$.
Subsequently, we can compute and substitute the optimal point in the active constraint gradients as follow:

$$\nabla c_4(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\nabla c_5(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Notice that, the LICQ property is satisfied at the point $A(0, 0)$ because the gradients of the active constraints are not linearly dependent between each other at the specific point. Next, we need to check that $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$, thus it is possible to write as follows:

$$\begin{aligned}
\nabla_x \mathcal{L}(x^*, \lambda^*) &= \nabla f(x) - \lambda_4 \nabla c_4(x) - \lambda_5 \nabla c_5(x) \\
&= \begin{bmatrix} 4 \\ 3 \end{bmatrix} - \lambda_4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \lambda_5 \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\end{aligned}$$

It can easily be noticed that the gradient of the Lagrange function is equal to zero at the optimal point A when Lagrangian multipliers associated to the active constraints are $\lambda_4 = 4$ and $\lambda_5 = 3$. In conclusion, it is possible to affirm that the first order necessary conditions hold for the optimal point $A(0, 0)$.

References

1. Jorge Nocedal and Stephen J Wright. Numerical optimization. Springer, 1999
2. "Notes on the optimization" available on iCorsi platform
3. Ron Larson,"Elementary Linear Algebra" chapter 9 about linear programming and the simplex method