Numerical Linear Algebra Homework Project 5: Constrained Optimization

Catalano Giuseppe, Cerrato Nunzia

Exercise 1

We want to minimize the function

$$f_a(\mathbf{x}) = (x_1 - 4)^2 + x_2^2 \tag{1}$$

subject to the constraints

$$x_1 + x_2 \leq 2, \quad x_1 \geq 0, \quad x_2 \geq 0, \tag{2}$$

which identify the set $\mathcal{C}_a = \{\mathbf{x} \in \mathbb{R}^2 : x_1 + x_2 \leq 2, \ x_1 \geq 0, \ x_2 \geq 0\}$. In other words, we want to find

$$f_a(\mathbf{x}^*) = \min_{\mathbf{x} \in \mathcal{C}_a} f_a(\mathbf{x}). \tag{3}$$

Since the function $f_a(\mathbf{x})$, reported in Eq. (1), is convex and continuous and the constraint set $\mathcal{C}_a$ is convex, there exists a unique minimum. In particular, it can be seen that, since there are no mixed terms in $x_1 x_2$ in the expression of the function, its minimum can be found by minimizing $(x_1 - 4)^2$ and x_2^2 separately. x_2^2 is minimized when $x_2 = 0$, which is compatible with the constraint set, and $(x_1 - 4)^2$ is minimized when $x_1 = 4$, which is incompatible with the constraint set. Since it has to be $x_1 + x_2 \leq 2$, the optimal choice of x_1 is $x_1 = 2$. Therefore we have:

$$\mathbf{x}^* = (2, 0), \qquad f_a(\mathbf{x}^*) = 4. \tag{4}$$

We would like to find the minimum value both analytically, by solving the KKT system, and numerically, by using the interior point method.

The Lagrangian of the problem is the following:

$$\mathcal{L}_a(\mathbf{x}, \boldsymbol{\lambda}) = f_a(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{c}_a(\mathbf{x}), \quad \text{with } \mathbf{c}_a(\mathbf{x}) = \begin{bmatrix} 2 - x_1 - x_2 \\ x_1 \\ x_2 \end{bmatrix}, \tag{5}$$

where the constraints of the problem are satisfied when $c_i(\mathbf{x}) \geq 0 \ \forall i = 1, 2, 3$. From this definition, it is possible to write the KKT conditions, namely:

$$\begin{cases} \text{stationarity} & (2(x_1 - 4) + \lambda_1 - \lambda_2, 2x_2 + \lambda_1 - \lambda_3) = (0, 0) \\ \text{primal feasibility} & (2 - x_1 - x_2, x_1, x_2) \geq (0, 0, 0) \\ \text{dual feasibility} & (\lambda_1, \lambda_2, \lambda_3) \geq (0, 0, 0) \\ \text{complementarity} & (\lambda_1(2 - x_1 - x_2), \lambda_2 x_1, \lambda_3 x_2) = (0, 0, 0) \end{cases} \tag{6}$$

where the inequalities must be intended element-wise.

The complementarity conditions can be satisfied in eight different cases, depending on the fact that each constraint can be active or not. We recall that a constraint $c_i(\mathbf{x})$ is active if

and only if $c_i(\mathbf{x}^*) = 0$. In the following, we identify each case with an array of m entries, where m is the number of constraints, and we assign to the i-th entry the label A or N depending on whether the i-th constraint is active or not. Associated to each case we report the possible solutions of the constrained minimization problem which are compatible with the corresponding complementarity condition.

	A/N constraints	Possible solutions $\mathbf{x}^*$
1	(A, N, N)	$\mathbf{x}^* = (x_1^*, -x_1^* + 2)$, with $x_1^* \in (0, 2)$
2	(A, A, A)	$\nexists \, \mathbf{x}^* \in \mathbb{R}^2$
3	(N, N, N)	$\mathbf{x}^* \in \mathring{\mathcal{C}}_a$
4	(N, A, A)	$\mathbf{x}^* = (0, 0)$
5	(A, A, N)	$\mathbf{x}^* = (0, 2)$
6	(A, N, A)	$\mathbf{x}^* = (2, 0)$
7	(N, A, N)	$\mathbf{x}^* = (0, x_2^*)$, with $x_2^* \in (0, 2)$
8	(N, N, A)	$\mathbf{x}^* = (x_1^*, 0)$ with $x_1^* \in (0, 2)$

Table 1: Table of the cases satisfying the complementarity condition. The i-th row of the table contains the array of the active/non-active states of the constraints (second column) and the possible solutions $\mathbf{x}^*$ of the constrained minimization problem which are compatible with that case (third column).

We want to find the solutions to the KKT system (6) by investigating if the conditions on Lagrange multipliers imposed by the constraint states and the range of possible solutions $\mathbf{x}^*$ reported in Tab. 1, which came from the complementarity condition, are compatible with the other equations of the system. The results that we find are reported below.

1. From the case (A, N, N) we have $\lambda_2^* = \lambda_3^* = 0$. This leads to $\lambda_1^* = 2$ and $\mathbf{x}^* = (3, -1)$, which is not compatible with the constraint set.

2. For the case (A, A, A) it can be seen from the second row of the Table 1 that $\nexists \, \mathbf{x}^* \in \mathbb{R}^2$ which represents a solution to the constrained minimization problem.

3. From the case (N, N, N) we have $\lambda_1^* = \lambda_2^* = \lambda_3^* = 0$. In this case one finds $\mathbf{x}^* = (4, 0)$, which is outside the feasible set.

4. From the case (N, A, A) we have $\lambda_1^* = 0$. From the complementarity condition we already have $\mathbf{x}^* = (0, 0)$. From the stationarity condition, we find $\lambda_3^* = 0$, and $\lambda_2^* = -8$, which is not allowed.

5. From the case (A, A, N) we have $\lambda_3^* = 0$. From the complementarity condition we already have $\mathbf{x}^* = (0, 2)$. By inserting this value in the first equation of the KKT system, we find $\lambda_1^* = -4$, $\lambda_2^* = -12$, which are both not allowed values.

6. From the case (A, N, A) we have $\lambda_2^* = 0$. In this case, one has $\mathbf{x}^* = (2, 0)$ from the complementarity condition and $\lambda_1^* = 4$, $\lambda_3^* = 4$, which are both acceptable solutions. Note that this is the solution that we have already found heuristically.

7. From the case (N, A, N) we have $\lambda_1^* = 0$, $\lambda_3^* = 0$. In this case, we find $\lambda_2 = -8$, which is not an acceptable solution. Note that one can also find $x_2^* = 0$, which is not compatible with the state of the constraint.

8. From the case (N, N, A) we have $\lambda_1^* = 0$, $\lambda_2^* = 0$. In this case, we find $\lambda_3 = 0$ and $\mathbf{x}^* = (4, 0)$, which is outside the feasible set.

As emerges from the study of the KKT system, $\boldsymbol{\lambda}^* = (4, 0, 4)$ and the unique solution to the constrained minimization problem is $\mathbf{x}^* = (2, 0)$, where $f_a(\mathbf{x}^*) = 4$.

Exercise 2

We want to minimize the function

$$f_b(\mathbf{x}) = 2x_1 - x_2^2, \tag{7}$$

under the constraints

$$x_1^2 + x_2^2 \leq 1, \quad x_1 \geq 0, \quad x_2 \geq 0, \tag{8}$$

which identify the set $\mathcal{C}_b = \{\mathbf{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, \ x_1 \geq 0, \ x_2 \geq 0\}$. In other words, we want to find

$$f_b(\mathbf{x}^*) = \min_{\mathbf{x} \in \mathcal{C}_b} f_b(\mathbf{x}). \tag{9}$$

In this case, the function $f_b(\mathbf{x}^*)$ is concave, therefore, in principle, nothing can be said about its minimum. However, by exploiting the constraints, we can obtain the following chain of inequalities

$$2x_1 - x_2^2 \geq -x_2^2 \geq -1, \tag{10}$$

where in the first inequality we use that $x_1 \geq 0$, which is saturated when $x_1 = 0$, while in the second one we use that $x_2^2 \leq 1 - x_1^2 \leq 1$, which is saturated when $x_2 = \pm 1$. Finally, knowing that x_2 must be positive, we can conclude that the unique solution of the problem in Eq. (9) is $\mathbf{x}^* = (0, 1)$ and $f_b(\mathbf{x}^*) = -1$.

We would like to obtain now the same minimum point by solving the KKT system analytically and, subsequently, by exploiting the interior point method numerically.
The Lagrangian associated with the problem is the following:

$$\mathcal{L}_b(\mathbf{x}, \boldsymbol{\lambda}) = f_b(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{c}_b(\mathbf{x}), \quad \text{with } \mathbf{c}_b(\mathbf{x}) = \begin{bmatrix} 1 - x_1^2 - x_2^2 \\ x_1 \\ x_2 \end{bmatrix}, \tag{11}$$

where the constraints of the problem are satisfied when $c_i(\mathbf{x}) \geq 0 \ \forall i = 1, 2, 3$. The KKT system is:

$$\begin{cases} \text{stationarity} & (2 + 2\lambda_1 x_1 - \lambda_2, -2x_2 + 2\lambda_1 x_2 - \lambda_3) = (0, 0) \\ \text{primal feasibility} & (1 - x_1^2 - x_2^2, x_1, x_2) \geq (0, 0, 0) \\ \text{dual feasibility} & (\lambda_1, \lambda_2, \lambda_3) \geq (0, 0, 0) \\ \text{complementarity} & (\lambda_1(1 - x_1^2 - x_2^2), \lambda_2 x_1, \lambda_3 x_2) = (0, 0, 0) \end{cases} \tag{12}$$

where the inequalities must be intended element-wise.

Analogously to the previous case, we study the eight cases that can be derived from the complementarity condition. We report in the following table the possible solutions to the constrained minimization problem which are compatible with corresponding complementarity condition.

	A/N constraints	Possible solutions $\mathbf{x}^*$
1	(A, A, A)	$\nexists\, \mathbf{x}^* \in \mathbb{R}^2$
2	(A, A, N)	$\mathbf{x}^* = (0, 1)$
3	(A, N, A)	$\mathbf{x}^* = (1, 0)$
4	(A, N, N)	$x_2^* = \sqrt{(1 - x_1^{*2})}$, with $x_1^* \in (0, 1)$
5	(N, A, A)	$\mathbf{x}^* = (0, 0)$
6	(N, A, N)	$\mathbf{x}^* = (0, x_2^*)$, with $x_2^* \in (0, 1)$
7	(N, N, A)	$\mathbf{x}^* = (x_1^*, 0)$ with $x_1^* \in (0, 1)$
8	(N, N, N)	$\mathbf{x}^* \in \mathring{\mathcal{C}}_b$

Table 2: Table of the cases satisfying the complementarity condition. The i-th row of the table contains the array of the active/non-active states of the constraints (second column) and the possible solutions $\mathbf{x}^*$ of the constrained minimization problem which are compatible with that case (third column).

1. For the case (A, A, A) it can be seen from the second row of the Table 2 that $\nexists\, \mathbf{x}^* \in \mathbb{R}^2$ which represents a solution to the constrained minimization problem.

2. For the case (A, A, N) we have $\lambda_3^* = 0$. From the complementarity condition we already know that $\mathbf{x}^* = (0, 1)$. By inserting this value in the first equation of the KKT system, we find $\lambda_1^* = 1$ and $\lambda_2^* = 0$, which are allowed values.

3. For the case (A, N, A) we have $\lambda_2^* = 0$. In this case, we have already found that $\mathbf{x}^* = (1, 0)$, which leads to $\lambda_1^* = -1$, which is not allowed, and $\lambda_3^* = 0$.

4. For the case (A, N, N) we have $\lambda_2^* = \lambda_3^* = 0$. From the first equation of the KKT system, we arrive at the condition $\lambda_1 x_1 = -1$, which cannot be satisfied since both λ_1 and x_1 must be positive.

5. For the case (N, A, A) we have $\lambda_1^* = 0$. Here we already have $\mathbf{x}^* = (0, 0)$, which leads to $\lambda_2^* = 2$ and $\lambda_3^* = 0$.

6. For the case (N, A, N) we have $\lambda_1^* = \lambda_3^* = 0$. In this case, we find $\lambda_2^* = 2$ and $\mathbf{x}^* = (0, 0)$.

7. For the case (N, N, A) we have $\lambda_1^* = \lambda_2^* = 0$. By inserting these values in the stationarity condition we arrive at the expression $2 = 0$, which is absurd, therefore no solutions can be found in this case.

8. For the case (N, N, N) we have $\lambda_1^* = \lambda_2^* = \lambda_3^* = 0$. This case is analogous to the previous one, since we arrive at the same expression $2 = 0$ when considering $\lambda_1^* = 0$, therefore no solutions can be found.

From the analysis of the KKT conditions, we can conclude that there exist two solutions of the KKT system, namely $\mathbf{x}_1^* = (0, 1)$ with $\boldsymbol{\lambda}_1^* = (1, 0, 0)$ and $\mathbf{x}_2^* = (0, 0)$ with $\boldsymbol{\lambda}_2^* = (0, 2, 0)$. By evaluating the function at these points we have $f_b(\mathbf{x}_1^*) = -1$ and $f_b(\mathbf{x}_2^*) = 0$, therefore we conclude that $\mathbf{x}_1^*$ is the solution of the constrained minimization problem. However, since there are two solutions of the KKT system, we expect there could be problems when searching numerically the minimum of this function by using the interior point method.

Exercise 3

Consider the problem

$$\min f(\mathbf{x}) - \mu \mathbf{e}^T \log(\mathbf{z}) \quad \text{s.t.} \quad \mathbf{c}(\mathbf{x}) - \mathbf{z} = \mathbf{0}, \tag{13}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, is the function one wants to find the minimum, $\mathbf{c} : \mathbb{R}^n \to \mathbb{R}^m$ is a vector defining the constraints, which are assumed to be $c_i(\mathbf{x}) \geq 0$, $\forall i = 1, \ldots, m$, $\mathbf{z} \in \mathbb{R}^m$ is called slack variable and, finally, $\mathbf{e}^T = (1, \ldots, 1) \in \mathbb{R}^m$ and $\mu > 0$. Due to the presence of the logarithmic term, for sufficiently small values of the parameter μ this represents a slightly perturbed version of a constrained minimization problem.

In this case, from the KKT theorem it is possible to find the following necessary conditions:

$$\mathbf{r}_1 := \nabla f(\mathbf{x}) - \boldsymbol{\lambda}^T \nabla_{\mathbf{x}} \mathbf{c}(\mathbf{x}) = \mathbf{0} \tag{14}$$

$$\mathbf{r}_2 := \mathbf{c}(\mathbf{x}) - \mathbf{z} = \mathbf{0} \tag{15}$$

$$\mathbf{r}_2 := \mathbf{z} \odot \boldsymbol{\lambda} - \mu \mathbf{e} = \mathbf{0}. \tag{16}$$

In order to solve this system, it is necessary to update the variables $\mathbf{x}$, $\boldsymbol{\lambda}$, $\mathbf{z}$, which can define a new variable $\mathbf{y} = [\mathbf{x}, \boldsymbol{\lambda}, \mathbf{z}]^T$. By denoting with $R(\mathbf{x})$ the matrix with rows $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, the updating step is

$$\mathbf{y}_{k+1} = \mathbf{y}_k - \nabla R(\mathbf{y}_k)^{-1} R(\mathbf{y}_k), \tag{17}$$

where $\nabla R(\mathbf{y})$ represents the Jacobian of the system, given by:

$$\nabla R(\mathbf{y}) = J(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{z}) = \begin{pmatrix} H(\mathbf{x}) + K(\mathbf{x}, \boldsymbol{\lambda}) & -\nabla \mathbf{c}(\mathbf{x})^T & \mathbf{0}_{n \times m} \\ \nabla \mathbf{c}(\mathbf{x}) & \mathbf{0}_{m \times m} & -\mathbb{1}_m \\ \mathbf{0}_{m \times n} & \mathrm{diag}(\mathbf{z}) & \mathrm{diag}(\boldsymbol{\lambda}) \end{pmatrix}, \tag{18}$$

where $H(\mathbf{x}) = \nabla^2 f(\mathbf{x})$, $K(\mathbf{x}, \boldsymbol{\lambda}) = -\boldsymbol{\lambda}^T \nabla^2 \mathbf{c}(\mathbf{x})$.

We want to write down the Jacobian associated with the problem (13) considering as functions and constraints the ones of the two previous exercises.

Function $f_a(\mathbf{x})$

We recall the expression of the function $f_a(\mathbf{x}) = (x_1 - 4)^2 + x_2^2$ and that of the constraints, which are written in the form $c_i(\mathbf{x}) \geq 0$, and can be recast in the vector

$$\mathbf{c}_a(\mathbf{x}) = \begin{bmatrix} 2 - x_1 - x_2 \\ x_1 \\ x_2 \end{bmatrix}. \tag{19}$$

In order to obtain the Jacobian, we have to compute the gradient and the Hessian of the function and the constraints. We report in the following their expressions.

$$\nabla f_a(\mathbf{x}) = [2(x_1 - 4), 2x_2]^T \qquad H_{f_a}(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \tag{20}$$

$$\nabla \mathbf{c}_a(\mathbf{x}) = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \nabla^2 \mathbf{c}_a(\mathbf{x}) = [\mathbf{0}_{2 \times 2}, \mathbf{0}_{2 \times 2}, \mathbf{0}_{2 \times 2}]^T. \tag{21}$$

Therefore, $K_a(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{0}_{2 \times 2}$ as expected, since all the constraints are linear.

Function $f_b(\mathbf{x})$

As before, we recall the expression of the function $f_b(\mathbf{x}) = 2x_1 - x_2^2$ and that of the constraints, which are written also in this case in the form $c_i(\mathbf{x}) \geq 0$, and can be recast in the vector

$$\mathbf{c}_b(\mathbf{x}) = \begin{bmatrix} 1 - x_1^2 - x_2^2 \\ x_1 \\ x_2 \end{bmatrix}. \tag{22}$$

We report in the following the gradient and the Hessian of the function and the constraints:

$$\nabla f_b(\mathbf{x}) = [2, -2x_2]^T \qquad\qquad H_{f_b}(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}, \qquad\qquad (23)$$

$$\nabla \mathbf{c}_b(\mathbf{x}) = \begin{bmatrix} -2x_1 & -2x_2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad\qquad \nabla^2 \mathbf{c}_b(\mathbf{x}) = \left[\nabla^2 [\mathbf{c}_b]_1, \mathbf{0}_{2\times 2}, \mathbf{0}_{2\times 2} \right]^T, \qquad (24)$$

where $\nabla^2 [\mathbf{c}_b]_1 = -\mathbb{1}_2$ and, therefore, $K_b(\mathbf{x}, \boldsymbol{\lambda}) = \lambda_1 \mathbb{1}_2$.