# IE 411: Introduction to Nonlinear Optimization

Fall 2022 - Homework Assignment 4
Due: December 5, 2022

**Question 1.** Let $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$. Show that exactly one of the following two systems is feasible:

a) $Ax \geq 0, x \geq 0, c^\mathsf{T} x > 0.$

b) $A^\mathsf{T} y \geq c, y \leq 0.$

**Hint:** Rewrite system given by (a) in the form of $\tilde{A}\tilde{x} \geq 0, \tilde{c}^\mathsf{T}\tilde{x} > 0$ and apply Farkas' Lemma. How would you define $\tilde{A}, \tilde{x}, \tilde{c}$ in terms of $A, x, c$?

**Sol:** Let $\tilde{A} = \begin{bmatrix} A \\ I_n \end{bmatrix}$, $\tilde{x} = -x$ and $\tilde{c} = -c$. Then system given in a) can be written as $\tilde{A}\tilde{x} \leq 0, \tilde{c}^\mathsf{T}\tilde{x} > 0$. Then we use Farkas Lemma in order to claim that exactly one of the following systems has a solution

a*) $\tilde{A}\tilde{x} \leq 0, \tilde{c}^\mathsf{T}\tilde{x} > 0,$

b*) $\tilde{A}^T w = \tilde{c}, w \geq 0.$

Let us separate $w$ in to parts, $w = \begin{bmatrix} -y \\ z \end{bmatrix}$, where $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$. Then b*) is written as

$$-A^\mathsf{T} y + z = -c, y \leq 0, z \geq 0.$$

Now we may eliminate $z$ by combining two constraints and obtain $A^\mathsf{T} y^- c = z \geq 0$ and we have $A^\mathsf{T} y \geq c, y \leq 0$ as desired.

**Question 2.** Consider the maximization problem

$$\begin{array}{rl} \text{maximize} & x_1^2 + 2x_1x_2 + 2x_2^2 - 3x_1 + x_2 \\ \text{subject to} & x_1 + x_2 = 1 \\ & x_1, x_2 \geq 0 \end{array}$$

a) Is the problem convex?

b) Find all the KKT points of the problem.

c) Find the optimal solution of the problem.

**Sol:**

a) "$\max f(x) = -\min -f(x)$". Hessian of the minimization objective $-f(x)$, returns as a constant matrix $\begin{bmatrix} -2 & -2 \\ -2 & -4 \end{bmatrix}$ which is a negative definite matrix. Hence problem fails to be convex.

b) "Equality constraints can be handled by using Lagrangian".

Also, you may describe an equality constraint using two inequalities. We use Theorem 10.5 to find a necessary condition since we have a linearly constrained nonconvex problem. Our problem can be written as

$$\begin{aligned} -\text{minimize} \quad & -x_1^2 - 2x_1x_2 - 2x_2^2 + 3x_1 - x_2 \\ \text{subject to} \quad & x_1 + x_2 \leq 1 \\ & -x_1 - x_2 \leq -1 \\ & -x_1 \leq 0 \\ & -x_2 \leq 0 \end{aligned}$$

We define $f(x) = -x_1^2 - 2x_1x_2 - 2x_2^2 + 3x_1 - x_2$, $a_1 = (1,1)^\mathsf{T}$, $a_2 = (-1,-1)^\mathsf{T}$, $a_3(-1,0)^\mathsf{T}$, $a_4 = (0,-1)^\mathsf{T}$, $b_1 = 1$, $b_2 = -1$ and $b_3 = b_4 = 0$. Then for any local minimizer $x^*$ we have the KKT conditions:

$$\begin{bmatrix} -2x_1^* - 2x_2^* + 3 \\ -2x_1^* - 4x_2^* - 1 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} -1 \\ -1 \end{bmatrix} + \lambda_3 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \lambda_1(x_1^* + x_2^* - 1) &= 0 \\ \lambda_2(-x_1^* - x_2^* + 1) &= 0 \\ \lambda_3(-x_1^*) &= 0 \\ \lambda_4(-x_2^*) &= 0 \\ \lambda_i \geq 0, \text{i=1,}&\ldots\text{,4} \end{aligned}$$

c) We try to find solutions of this system by cases:

**Case 1:** $x_1^* = 0$. Then,

$$\begin{aligned} -2x_2^* + 3 + \lambda_1 - \lambda_2 - \lambda_3 &= 0 \\ -4x_2^* - 1 + \lambda_1 - \lambda_2 - \lambda_4 &= 0 \\ \lambda_1(x_2^* - 1) &= 0 \\ \lambda_2(-x_2^* + 1) &= 0 \\ \lambda_4 x_2^* &= 0 \end{aligned}$$

**Case 1a:** $x_1^* = 0, x_2^* = 1$. Then,

$$\begin{aligned} \lambda_1 - \lambda_2 - \lambda_3 &= -1 \\ \lambda_1 - \lambda_2 - \lambda_4 &= 5 \\ \lambda_4 &= 0 \end{aligned}$$

Then $(x_1^*, x_2^*, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0, 1, \alpha, \alpha - 5, 6, 0)$ for any $\alpha \geq 5$. $f(x_1^*, x_2^*) = -3$.

**Case 1b:** $x_1^* = 0, x_2^* \neq 1$, Then

$$\begin{aligned} -2x_2^* + 3 + \lambda_1 - \lambda_2 - \lambda_3 &= 0 \\ -4x_2^* - 1 + \lambda_1 - \lambda_2 - \lambda_4 &= 0 \\ \lambda_1 = \lambda_2 &= 0 \\ \lambda_4 x_2^* &= 0 \end{aligned}$$

Then $(x_1^*, x_2^*, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = (0, 0, 0, 0, 3, -1)$ and $(0, \frac{1}{4}, 0, 0, \frac{5}{2}, 0)$. $f(x_1^*, x_2^*) = 0$ and $\frac{-3}{8}$ respectively. But both are infeasible.

**Case 2:** $x_1^* \neq 0$. Then,

$$\begin{aligned} -2x_1^* - 2x_2^* + 3 + \lambda_1 - \lambda_2 &= 0 \\ -2x_1^* - 4x_2^* - 1 + \lambda_1 - \lambda_2 - \lambda_4 &= 0 \\ \lambda_1(x_1^* + x_2^* - 1) &= 0 \\ \lambda_2(-x_1^* - x_2^* + 1) &= 0 \\ \lambda_4 x_2^* &= 0 \end{aligned}$$

**Case 2a:** $x_1^* \neq 0, x_2^* = 0$. Then,

$$\begin{aligned} -2x_1^* + \lambda_1 - \lambda_2 &= -3 \\ -2x_1^* + \lambda_1 - \lambda_2 - \lambda_4 &= 1 \\ \lambda_1(x_1^* - 1) &= 0 \\ \lambda_2(-x_1^* + 1) &= 0 \end{aligned}$$

Then $(x_1^*, x_2^*, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = (\frac{3}{2}, 0, 0, 0, 0, -4)$ and $(1, 0, \alpha, \alpha + 1, 0, -4)$ for $\alpha \geq 0$. $f(x_1^*, x_2^*) = \frac{9}{2}$ and 2 respectively. But first point is infeasible.

**Case 2b:** $x_1^* \neq 0, x_2 \neq 0$. Then,

$$\begin{aligned} -2x_1^* - 2x_2^* + 3 + \lambda_1 - \lambda_2 &= 0 \\ -2x_1^* - 4x_2^* - 1 + \lambda_1 - \lambda_2 &= 0 \\ \lambda_1(x_1^* + x_2^* - 1) &= 0 \\ \lambda_2(-x_1^* - x_2^* + 1) &= 0 \end{aligned}$$

Then $(x_1^*, x_2^*, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = (\frac{7}{2}, -2, 0, 0, 0, 0)$ and $(3, -2, \alpha, \alpha + 1, 0, 0)$ for $\alpha \geq 0$. $f(x_1^*, x_2^*) = \frac{25}{4}$ and 6 respectively. But both points are infeasible.

Then we are minimizing a continuous function over a compact feasible region, so Weierstrass theorem grants a global optimal solution, which must be one of the KKT points. Then $(x_1^*, x_2^*) = (0, 1)$ maximizes the original problem with objective equals to 3.

**Question 3.** Consider the problem

$$\begin{aligned} \text{minimize} \quad & -x_1 x_2 x_3 \\ \text{subject to} \quad & x_1 + 3x_2 + 6x_3 \leq 48 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

a) Write the KKT conditions for the problem.

b) Find the optimal solution of the problem.

**Sol:**

a) Hessian of $f(x_1, x_2, x_3) = -x_1x_2x_3$ is negative semidefinite on $x_1, x_2, x_3 \geq 0$, so the objective is nonconvex. We use Theorem 10.5 to find a necessary condition since we have a linearly constrained nonconvex problem. We define $a_1 = (1, 3, 6)^\mathsf{T}$, $a_2 = (-1, 0, 0)^\mathsf{T}$, $a_3 = (0, -1, 0)^\mathsf{T}$, $a_4 = (0, 0, -1)^\mathsf{T}$, $b_1 = 48$ and $b_2 = b_3 = b_4 = 0$. Then for any local minimizer $x^*$ we have the KKT conditions:

$$\begin{bmatrix} -x_2^*x_3^* \\ -x_1^*x_3^* \\ -x_1^*x_2^* \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix} + \lambda_2 \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} + \lambda_3 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + \lambda_4 \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\lambda_1(x_1^* + 3x_2^* + 6x_3^* - 48) = 0$$
$$\lambda_2(-x_1^*) = 0$$
$$\lambda_3(-x_2^*) = 0$$
$$\lambda_4(-x_3^*) = 0$$
$$\lambda_i \geq 0, i = 1, \ldots, 4$$

b) For $x_1, x_2, x_3 \geq 0$ we have $f(x_1, x_2, x_3) = -x_1x_2x_3 \leq 0$. Then any solution with at least one 0 entry has objective 0. Then we check the following system with $x_1^*, x_2^*, x_3^* \neq 0$.

$$\begin{aligned} -x_2^*x_3^* + \lambda_1 &= 0 \\ -x_1^*x_3^* + 3\lambda_1 &= 0 \\ -x_1^*x_2^* + 6\lambda_1 &= 0 \\ \lambda_1(x_1^* + 3x_2^* + 6x_3^* - 48) &= 0 \end{aligned}$$

Using $x_1^*, x_2^*, x_3^* \neq 0$ we obtain $\lambda_1 \neq 0$ and the following set of equalities:

$$\begin{aligned} x_1^*x_3^* = 3\lambda_1 = 3x_2^*x_3 &\Rightarrow \frac{1}{3}x_1^* = x_2^* \\ x_1^*x_2^* = 6\lambda_1 = 6x_2^*x_3^* &\Rightarrow \frac{1}{6}x_1^* = x_3^* \\ x_1^* + 3x_2^* + 6x_3^* &= 48 \end{aligned}$$

Then $(x_1^*, x_2^*, x_3^*, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = (16, \frac{16}{3}, \frac{8}{3}, \frac{128}{3}, 0, 0, 0)$ with $f(x_1^*, x_2^*, x_3^*) = -\frac{2048}{9}$. All other KKT points have at least one entry equal to 0. Then we are minimizing a continuous function over a compact feasible region, so Weierstrass theorem grants a global optimal solution, which must be one of the KKT points. Then $(x_1^*, x_2^*, x_3^*) = (16, \frac{16}{3}, \frac{8}{3})$ minimizes the original problem with objective equals to $-\frac{2048}{9}$.

**Question 4.** Consider the problem

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 + x_1 \\ \text{subject to} \quad & x_1 + x_2 \leq a, \end{aligned}$$

where $a \in \mathbb{R}$ is a parameter.

a) Solve the problem using KKT conditions. (The solution will be in terms of the parameter $a$. You may need to consider different cases for $a$.)

b) Let $h(a)$ be the optimal value of the problem with parameter $a$. Write an explicit expression for $h$.

c) Show that $h : \mathbb{R} \to \mathbb{R}$ is a convex function.

**Sol:**

a) Hessian is positive definite, so the objective function is convex. For any choice of $\alpha$ feasible region is determined by a half-space so it is convex. Hence we have a convex problem. We use Theorem 10.6 to find necessary and sufficient conditions since we have a linearly constrained convex problem. We define $f(x_1, x_2) = x_1^2 + x_2^2 + x_1$, $a = (1,1)^\mathsf{T}$ and $b = \alpha$. Then $x^*$ is an optimal solution if and only if we have the KKT conditions:

$$\begin{aligned} \begin{bmatrix} 2x_1^* + 1 \\ 2x_2^* \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \lambda(x_1 + x_2 - \alpha) &= 0 \\ \lambda &\geq 0 \end{aligned}$$

Then we have to cases:
**Case 1:** $\alpha < -\frac{1}{2}$. In this case $\lambda = \frac{-2\alpha-1}{2} > 0$ and $x_1^* + x_2^* = \alpha$. Then $(x_1^*, x_2^*) = (\frac{2\alpha-1}{4}, \frac{2\alpha+1}{4})$ and $f(x_1^*, x_2^*) = \frac{8\alpha^2+8\alpha-2}{16}$.
**Case 2:** $\alpha \geq -\frac{1}{2}$. In this case one possible solutions is $\lambda = 0$, $x_1^* = -\frac{1}{2}$ and $x_2^* = 0$ with $f(x_1^*, x_2^*) = -\frac{1}{4}$. These values can be obtained by using the results in **Case 1**, plugging in $\alpha = -\frac{1}{2}$. If one picks $\lambda \neq 0$, then the system above forces $\lambda \geq 0$ and $\lambda \leq 0$, a contradiction.

Hence solutions are determined by $(x_1^*, x_2^*) = \begin{cases} (\frac{2\alpha-1}{4}, \frac{2\alpha+1}{4}), & \alpha < -\frac{1}{2}, \\ (-\frac{1}{2}, 0), & \alpha \geq -\frac{1}{2}. \end{cases}$

b) $h(\alpha)$ can be easily obtained by plugging in the values we found in part a), that is $h(\alpha) = \begin{cases} \frac{8\alpha^2+8\alpha-2}{16}, & \alpha < -\frac{1}{2}, \\ -\frac{1}{4}, & \alpha \geq -\frac{1}{2}. \end{cases}$.

c) Both pieces are convex functions on the whole line. So it suffices to check convexity for points $\alpha < -\frac{1}{2} < \beta$ and $\lambda \in (0,1)$. There are two possible cases:
**Case 1:** "$\lambda\alpha + (1-\lambda)\beta < -\frac{1}{2}$". Then, observe that $\frac{8\alpha^2+8\alpha-2}{16}$ is convex on $\mathbb{R}$ and check that (draw these points on the real line and observe) for our specific choice of $\alpha, \beta$

$$\lambda\alpha + (1-\lambda)\beta = \Big(\frac{\lambda(\beta-\alpha) - \beta - \frac{1}{2}}{-\alpha - \frac{1}{2}}\Big)\alpha + \Big(\frac{(1-\lambda)(\beta-\alpha)}{-\alpha-\frac{1}{2}}\Big)\Big(-\frac{1}{2}\Big).$$

Also, check that $\Big(\frac{\lambda(\beta-\alpha)-\beta-\frac{1}{2}}{-\alpha-\frac{1}{2}}\Big) \in (0,1)$. Then, using the fact that $h(\alpha) > h\Big(-\frac{1}{2}\Big) = h(\beta)$ we have .

$$\begin{aligned} h(\lambda\alpha + (1-\lambda)\beta) &= h\Bigg(\Big(\frac{\lambda(\beta-\alpha) - \beta - \frac{1}{2}}{-\alpha - \frac{1}{2}}\Big)\alpha + \Big(\frac{(1-\lambda)(\beta-\alpha)}{-\alpha-\frac{1}{2}}\Big)\Big(-\frac{1}{2}\Big)\Bigg) \\ &\leq \Big(\frac{\lambda(\beta-\alpha) - \beta - \frac{1}{2}}{-\alpha - \frac{1}{2}}\Big)h(\alpha) + \Big(\frac{(1-\lambda)(\beta-\alpha)}{-\alpha-\frac{1}{2}}\Big)h\Big(-\frac{1}{2}\Big) \\ &= \Big(\frac{\lambda(\beta-\alpha) - \beta - \frac{1}{2}}{-\alpha - \frac{1}{2}}\Big)h(\alpha) + \Big(\frac{(1-\lambda)(\beta+\frac{1}{2})}{-\alpha-\frac{1}{2}}\Big)h\Big(-\frac{1}{2}) \\ &+ (1-\lambda)h\Big(-\frac{1}{2}\Big) \\ &\leq \Big(\frac{\lambda(\beta-\alpha) - \beta - \frac{1}{2}}{-\alpha - \frac{1}{2}}\Big)h(\alpha) + \Big(\frac{(1-\lambda)(\beta+\frac{1}{2})}{-\alpha-\frac{1}{2}}\Big)h(\alpha) \\ &+ (1-\lambda)h\Big(-\frac{1}{2}\Big) \\ &\leq \lambda h(\alpha) + (1-\lambda)h\Big(-\frac{1}{2}\Big) = \lambda h(\alpha) + (1-\lambda)h(\beta). \end{aligned}$$

**Case 2:** "$\lambda\alpha + (1-\lambda)\beta \geq -\frac{1}{2}$". Then

$$\begin{aligned} h(\lambda\alpha + (1-\lambda)\beta) = h\Big(-\frac{1}{2}\Big) &= \lambda h\Big(-\frac{1}{2}\Big) + (1-\lambda)h\Big(-\frac{1}{2}\Big) \\ &\leq \lambda h(\alpha) + (1-\lambda)h(\beta). \end{aligned}$$

**Question 5.** Use the KKT conditions to solve the problem

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & -2x_1 - x_2 + 10 \le 0 \\ & x_2 \ge 0 \end{aligned}$$

**Sol:** Hessian is positive definite, so the objective function is convex. Feasible region is determined by a half-space so it is convex. Hence we have a convex problem. We use Theorem 10.6 to find necessary and sufficient conditions since we have a linearly constrained convex problem. We define $f(x_1, x_2) = x_1^2 + x_2^2$, $a_1 = (-2, -1)^\mathsf{T}$, $a_2 = (0, -1)^\mathsf{T}$, $b_1 = -10$ and $b_2 = 0$. Then $x^*$ is an optimal solution if and only if we have the KKT conditions:

$$\begin{aligned} \begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} + \lambda_1 \begin{bmatrix} -2 \\ -1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \lambda_1(-2x_1^* - x_2^* + 10) &= 0 \\ \lambda_2(-x_2) &= 0 \end{aligned}$$

We try to find solutions of this system by cases:
**Case 1:** $x_2 = 0$. Then,

$$\begin{aligned} 2x_1^* - 2\lambda_1 &= 0 \\ -\lambda_1 - \lambda_2 &= 0 \\ \lambda_1(-2x_1^* + 10) &= 0 \end{aligned}$$

Then $(x_1^*, x_2^*, \lambda_1, \lambda_2) = (0, 0, 0, 0)$ and $(5, 0, 5, -5)$. $f(x_1^*, x_2^*) = 0$ and $25$ respectively. But both points are infeasible.
**Case 2:** $x_2 \neq 0$. Then,

$$\begin{aligned} 2x_1^* - 2\lambda_1 &= 0 \\ 2x_2^* - \lambda_1 - \lambda_2 &= 0 \\ \lambda_1(-2x_1^* - x_2^* + 10) &= 0 \\ \lambda_2 &= 0 \end{aligned}$$

Then $(x_1^*, x_2^*, \lambda_1, \lambda_2) = (4, 2, 4, 0)$ with $f(x_1^*, x_2^*) = 20$, which is the optimal value of the problem.