IE 411: Introduction to Nonlinear Optimization

Fall 2022 - Homework Assignment 5

Due: December 20 2022

Question 1. Consider the optimization problem:

$$\begin{aligned} \text{minimize} \quad & x_1 - 4x_2 + x_3 \\ \text{subject to} \quad & x_1 + 2x_2 + 2x_3 = -2 \\ & x_1^2 + x_2^2 + x_3^2 \leq 1. \end{aligned}$$

a) Given a KKT point of this problem, must it be an optimal solution? Explain/show your reasoning.

b) Solve the problem using KKT conditions.

Sol:

a)
- $x_1 - 4x_2 + x_3, x_1^2 + x_2^2 + x_3^2 - 1$ are continuously differentiable convex functions over $\mathbb{R}^3$, (first one is an affine function and the second one is $\|\mathbf{x}\|^2 - 1$),
- $x_1 + 2x_2 + 2x_3 + 2$ is an affine function.

Hence KKT points are sufficient for optimality.

b) We write KKT conditions,

$$\begin{bmatrix} 1 \\ -4 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$
$$\lambda(x_1^2 + x_2^2 + x_3^2 - 1) = 0,$$
$$\lambda \geq 0.$$

$\lambda = 0$ is not possible, so we must have $x_1^2 + x_2^2 + x_3^2 = 1$. Then we have the following system of equations,

$$x_1 = \frac{-1-\mu}{2\lambda}, \quad x_2 = \frac{4-2\mu}{2\lambda}, \quad x_3 = \frac{-1-2\mu}{2\lambda}.$$

Using this system we obtain the following system,

$$\frac{-1-\mu}{2\lambda} + 2\left(\frac{4-2\mu}{2\lambda}\right) + 2\left(\frac{-1-2\mu}{2\lambda}\right) = -2$$
$$\left(\frac{-1-\mu}{2\lambda}\right)^2 + \left(\frac{4-2\mu}{2\lambda}\right)^2 + \left(\frac{-1-2\mu}{2\lambda}\right)^2 = 1$$

Then we have only one solution satisfying $\lambda \geq 0$, that is

$$\mu = \frac{5+2\sqrt{\frac{137}{5}}}{9} \approx 1.7188, \quad \lambda = \frac{\sqrt{\frac{137}{5}}}{2} \approx 2.6173.$$

x_1, x_2, x_3 can be found accordingly. Due to the choice of λ, μ it will satisfy $x_1 + 2x_2 + 2x_3 = -2$ and $x_1^2 + x_2^2 + x_3^2 = 1$.

Correction: Previously I wrote that there are no feasible KKT points. I made a mistake while calculating the quadratic roots. Nevermind, it is deleted now.

Question 2. Consider the optimization problem:

$$\begin{aligned} \text{minimize} \quad & x_1^2 - x_2^2 - x_3^2 \\ \text{subject to} \quad & x_1^4 + x_2^4 + x_3^4 \leq 1. \end{aligned}$$

a) Is this a convex programming problem? Explain/show your reasoning.

b) Find all the KKT points of the problem.

c) Find the optimal solution of the problem.

Sol:

a) Objective is non-convex since Hessian is $\begin{bmatrix} 2 & & \\ & -2 & \\ & & -2 \end{bmatrix}$. So the problem is not a convex problem.

b) We have the KKT conditions

$$\begin{bmatrix} 2x_1 \\ -2x_2 \\ -2x_3 \end{bmatrix} + \lambda \begin{bmatrix} 4x_1^3 \\ 4x_2^3 \\ 4x_3^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$
$$\lambda(x_1^4 + x_2^4 + x_3^4 - 1) = 0$$
$$x_1^4 + x_2^4 + x_3^4 - 1 \leq 0$$
$$\lambda \geq 0.$$

c) If $\lambda = 0$, then we should have $(0,0,0)$ as the KKT point. If $\lambda \neq 0$, then $x_1^4 + x_2^4 + x_3^4 = 1$ and

$$2x_1(2\lambda x_1^2 + 1) = 0$$
$$2x_2(2\lambda x_2^2 - 1) = 0$$
$$2x_3(2\lambda x_3^2 - 1) = 0$$

Then $x_1 = 0$, and $x_2, x_3 = \{0, \pm\sqrt{\frac{1}{2\lambda}}\}$, which yields,

λ	x_1	x_2	x_3	Obj
0	0	0	0	0
$\frac{1}{2}$	0	± 1	0	-1
$\frac{1}{2}$	0	0	± 1	-1
$\frac{1}{\sqrt{2}}$	0	$\pm\sqrt[4]{\frac{1}{2}}$	$\pm\sqrt[4]{\frac{1}{2}}$	$-\sqrt{2}$

Therefore the solution is $(0, \pm\sqrt[4]{\frac{1}{2}}, \pm\sqrt[4]{\frac{1}{2}})$ with objective $-\sqrt{2}$.

Question 3. Use KKT conditions to solve the following problem. Explain/show your reasoning in detail.

$$\begin{aligned} \text{minimize} \quad & x_1^4 - x_2^2 \\ \text{subject to} \quad & x_1^2 + x_2^2 \leq 1 \\ & 2x_2 + 1 \leq 0. \end{aligned}$$

Sol: KKT conditions are

$$\begin{bmatrix} 4x_1^3 \\ -2x_2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\lambda_1(x_1^2 + x_2^2 - 1) = 0$$
$$\lambda_2(2x_2 + 1) = 0$$
$$x_1^2 + x_2^2 \leq 1$$
$$2x_2 + 1 \leq 0$$
$$\lambda_1, \lambda_2 \geq 0$$

Using the fact that $x_1^2 + x_2^2 \leq 1$ we know that

$$x_1^4 - x_2^2 \geq x_1^4 + x_1^2 - x_1^2 - x_2^2 \geq 0 + 0 - 1 \geq -1.$$

So if we show that $(0, -1)$ is a KKT point, then we are done.

$$\begin{bmatrix} 0 \\ 2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 0 \\ -2 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$\lambda_2 = 0$$
$$\lambda_1, \lambda_2 \geq 0$$

For $\lambda_1 = 1, \lambda_2 = 0$ our point is a KKT point and it attains the lower bound for the objective.

Question 4. Consider the optimization problem:

$$\begin{aligned} \text{minimize} \quad & (x_1 - 3)^2 + (x_2 - 2)^2 \\ \text{subject to} \quad & x_1 + x_2 = 1 \\ & x_1, x_2 \geq 0. \end{aligned}$$

a) Solve the problem using KKT conditions. Explain each step clearly.

b) Derive the Lagrange dual problem. What can you say about strong duality without solnving the dual problem.

c) Solve the dual problem.

Sol:

a) We write the KKT conditions

$$
\begin{aligned}
\begin{bmatrix} 2x_1 - 6 \\ 2x_2 - 4 \end{bmatrix} + \lambda_1 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ 1 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\lambda_1 x_1 &= 0 \\
\lambda_2 x_2 &= 0 \\
x_1 + x_2 &= 1 \\
x_1, x_2 &\geq 0 \\
\lambda_1, \lambda_2 &\geq 0.
\end{aligned}
$$

Then we have three cases:

Case 1: $x_1 = 0, x_2 > 0$.

Then we have $x_1 = 0, x_2 = 1, \lambda_1 = -4, \lambda_2 = 0, \mu = 2$ by inspection. INFEASIBLE.

Case 2: $x_1 > 0, x_2 = 0$.

Then we have $x_1 = 1, x_2 = 0, \lambda_1 = 0, \lambda_2 = 0, \mu = 4$ by inspection. KKT POINT.

Case 3: $x_1, x_2 > 0$.

Then we have $x_1 = 1, x_2 = 0, \lambda_1 = 0, \lambda_2 = 0, \mu = 4$ by inspection. INFEASIBLE.

Hence $(1, 0)$ is the optimal solution for the problem with objective 8.

b) Let

$$
\begin{aligned}
L(x_1, x_2, \lambda_1, \lambda_2, \mu) &= (x_1 - 3)^2 + (x_2 - 2)^2 - \lambda_1 x_1 - \lambda_2 x_2 + \mu(x_1 + x_2 - 1) \\
&= x_1^2 - (6 + \lambda_1 - \mu)x_1 + x_2^2 - (4 + \lambda_2 - \mu)x_2 + 13 - \mu \\
&= \left(x_1 - \frac{6 + \lambda_1 - \mu}{2}\right)^2 + \left(x_2 - \frac{4 + \lambda_2 - \mu}{2}\right)^2 + 13 \\
&\quad - \mu - \left(\frac{6 + \lambda_1 - \mu}{2}\right)^2 - \left(\frac{4 + \lambda_2 - \mu}{2}\right)^2 \\
&\geq 13 - \mu - \left(\frac{6 + \lambda_1 - \mu}{2}\right)^2 - \left(\frac{4 + \lambda_2 - \mu}{2}\right)^2.
\end{aligned}
$$

$L(x_1, x_2, \lambda_1, \lambda_2, \mu)$ cannot be $-\infty$ for fixed $\lambda_1, \lambda_2, \mu$ so the dual problem turns out to be

$$\max\ 13 - \mu - \left(\frac{6 + \lambda_1 - \mu}{2}\right)^2 - \left(\frac{4 + \lambda_2 - \mu}{2}\right)^2 \text{ subject to } \lambda_1, \lambda_2 \geq 0, \mu \in \mathbb{R}.$$

- $\{x_1, x_2 : x_1 + x_2 = 1\}$ is a convex set.
- $(x_1 - 3)^2 + (x_2 - 2)^2, -x_1, -x_2$ are convex functions.
- Problem has a finite optimal value (by part a)
- $(x_1, x_2) = (0.5, 0.5)$ satisfies $-x_1 < 0$ and $-x_2 < 0$.

Then by Theorem 12.8 we have strong duality.

c) We substitute $\max g(\lambda_1, \lambda_2, \mu) = -\min -g(\lambda_1, \lambda_2, \mu)$. Then $-g(\lambda_1, \lambda_2, \mu)$ is a convex function with a convex set so KKT is sufficient. Let $x_1, x_2 \geq 0$ such that KKT conditions are

$$\begin{bmatrix} 1 - (\frac{6+\lambda_1-\mu}{2}) - (\frac{4+\lambda_2-\mu}{2}) \\ \frac{6+\lambda_1-\mu}{2} \\ \frac{4+\lambda_2-\mu}{2} \end{bmatrix} - x_1 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - x_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$x_1\lambda_1 = 0$$

$$x_2\lambda_2 = 0$$

$$\lambda_1, \lambda_2, x_1, x_2 \geq 0.$$

Case 1: $\lambda_1, \lambda_2 = 0$.

Then $\mu = 4$ which makes $x_1 = 1, x_2 = 0$ a KKT point. It is sufficient and we stop. It returns -8 as the objective. But the original objective will be 8, as expected due to strong duality.

Question 5. Consider the optimization problem:

$$\begin{aligned} \text{minimize} \quad & x_1^2 + 2x_2^2 + 2x_1x_2 + x_1 - x_2 - x_3 \\ \text{subject to} \quad & x_1 + x_2 + x_3 \leq 1 \\ & x_3 \leq 3. \end{aligned}$$

a) Is the problem convex?

b) Find an optimal solution to this problem. Explain each step clearly.

c) Derive the Lagrange dual problem. What can you say about strong duality without solnving the dual problem.

d) Solve the dual problem.

Sol:

a) Problem is a minimization problem, it has a convex objective with affine constraints, so it is a convex problem.

b) KKT conditions are sufficient, so

$$\begin{bmatrix} 2x_1 + 2x_2 + 1 \\ 2x_1 + 4x_2 - 1 \\ -1 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \lambda_1(x_1 + x_2 + x_3 - 1) &= 0 \\ \lambda_2(x_3 - 3) &= 0 \\ x_1 + x_2 + x_3 &\leq 1 \\ x_3 &\leq 3 \\ \lambda_1, \lambda_2 &\geq 0 \end{aligned}$$

We have 4 cases:

Case 1: $\lambda_1 = \lambda_2 = 0$.

INFEASIBLE, since first equality returns $-1 = 0$.

Case 2: $\lambda_1 = 0, \lambda_2 > 0$.

Then $\lambda_2 = 1, x_1 = -\frac{3}{2}, x_2 = 1, x_3 = 3$. INFEASIBLE.

Case 3: $\lambda_1 > 0, \lambda_2 = 0$.

Then $\lambda_1 = 1, x_1 = -2, x_2 = 1, x_3 = 2$. KKT POINT.

Case 4: $\lambda_1, \lambda_2 > 0$.

Then $\lambda_2 = -2$. INFEASIBLE.

Hence the optimal solution is $(-2, 1, 2)$ with objective value -3.

c) We write the Lagrangian,

$$
\begin{aligned}
L(x_1, x_2, x_3, \lambda_1, \lambda_2) &= x_1^2 + 2x_2^2 + 2x_1x_2 + x_1 - x_2 - x_3 \\
&\quad + \lambda_1(x_1 + x_2 + x_3 - 1) + \lambda_2(x_3 - 3) \\
g(\lambda_1, \lambda_2) &= \inf_{x_1, x_2} \{x_1^2 + 2x_1x_2 + 2x_2^2 + x_1 - x_2 + \lambda_1 x_1 + \lambda_1 x_2\} \\
&\quad + \inf_{x_3} \{(\lambda_1 + \lambda_2 - 1)x_3\} - \lambda_1 - 3\lambda_2 \\
&= \begin{cases} \inf_{x_1,x_2}\{\mathbf{x}^\mathsf{T} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \lambda_1 + 1 \\ \lambda_1 - 1 \end{bmatrix}^\mathsf{T} \mathbf{x}\} - \lambda_1 - 3\lambda_2, & \lambda_1 + \lambda_2 = 1, \\ -\infty, & \lambda_1 + \lambda_2 \neq 1. \end{cases}
\end{aligned}
$$

Hessian of $\mathbf{x}^\mathsf{T} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \lambda_1 + 1 \\ \lambda_1 - 1 \end{bmatrix}^\mathsf{T} \mathbf{x}$ is a p.d. matrix so the function is convex and attains its minimum at stationary points, that is

$$
\left.\begin{aligned} 2x_1 + 2x_2 &= -\lambda_1 - 1 \\ 2x_1 + 4x_2 &= -\lambda_1 + 1 \end{aligned}\right\} \Rightarrow x_1 = \frac{-\lambda_1 - 3}{2}, x_2 = 1
$$

Then

$$
g(\lambda_1, \lambda_2) = \begin{cases} \left(-\frac{\lambda_1^2 + 2\lambda_1 + 5}{4}\right) - \lambda_1 - 3\lambda_2, & \lambda_1 + \lambda_2 = 1, \\ -\infty, & \lambda_1 + \lambda_2 \neq 1. \end{cases}
$$

Then we have

$$
\max -\frac{1}{4}\lambda_1^2 - \frac{3}{2}\lambda_1 - 3\lambda_2 - \frac{5}{4} \text{ subject to } \lambda_1 + \lambda_2 = 1, \lambda_1, \lambda_2 \geq 0.
$$

Substitute $\lambda_2 = 1 - \lambda_1$. Then

$$
\max -\frac{1}{4}\lambda_1^2 + \frac{3}{2}\lambda_1 - \frac{17}{4} \text{ subject to } 1 \geq \lambda_1 \geq 0.
$$

- $x_1^2 + 2x_2^2 + 2x_1x_2 + x_1 - x_2 - x_3, x_1 + x_2 + x_3 - 1, x_3 - 3$ are convex functions,
- $\mathbb{R}^3$ is a convex set,
- Problem has a finite optimal value by part b).
- $(x_1, x_2, x_3) = (0.3, 0.3, 0.3)$ satisfies $x_1 + x_2 + x_3 < 1$ and $x_3 < 3$.

Hence Theorem 12.8 provides strong duality.

d) We may solve

$$-\min \frac{1}{4}\lambda_1^2 - \frac{3}{2}\lambda_1 + \frac{17}{4} \text{ subject to } 1 \geq \lambda_1 \geq 0.$$

It is a convex quadratic function, stationary point is at $\lambda_1 = 3$ which is out of feasible region. Hence optimal solution must be on one of the endpoints which is $\lambda_1 = 1$ with the original objective -3, as expected due to strong duality.