Introduction to Optimization: Homework #5

Due on December 28, 2020

Professor Andre Milzarek

Peng Deng

Assignment A5.1

Consider the nonlinear optimization problem

$$\min_{x\in\mathbb{R}^2} f(x) := (x_1+1)^3 + 2\,(x_2 - x_1) + (x_2+1)^2 \qquad \text{s.t.} \qquad g(x) \le 0$$

where the constraint function $g : \mathbb{R}^2 \to \mathbb{R}^3$ is given by

$$g_1(x) := (x_1+1)^2 + x_2^2 - 4, \quad g_2(x) := x_1 + 1, \quad g_3(x) := x_2$$

Let us further set $\bar{x} := (-1,-2)^\top$.

a) Sketch the feasible set $X := \left\{x \in \mathbb{R}^2 : g(x) \le 0\right\}$.

b) Determine the active set $\mathcal{A}(\bar{x})$ and show that $\bar{x}$ is a regular point.

c) Investigate whether $\bar{x}$ is a KKT point of problem (1) and calculate a corresponding Lagrange multiplier $\bar{\lambda} \in \mathbb{R}^3$

d) Compute the associated critical cone $\mathcal{C}(\bar{x})$ and simplify as far as possible.

e) Investigate whether $\bar{x}$ is a local solution of problem (1) and explain your answer.

Solution

Subproblem (a)

As we can see from Figure 1, the yellow area is corresponded to $g_1(x) \le 0$, the purple area is corresponded to $g_2(x) \le 0$, the green area is corresponded to $g_3(x) \le 0$. Thus, the intersection of the three areas is corresponded to $g(x) \le 0$.

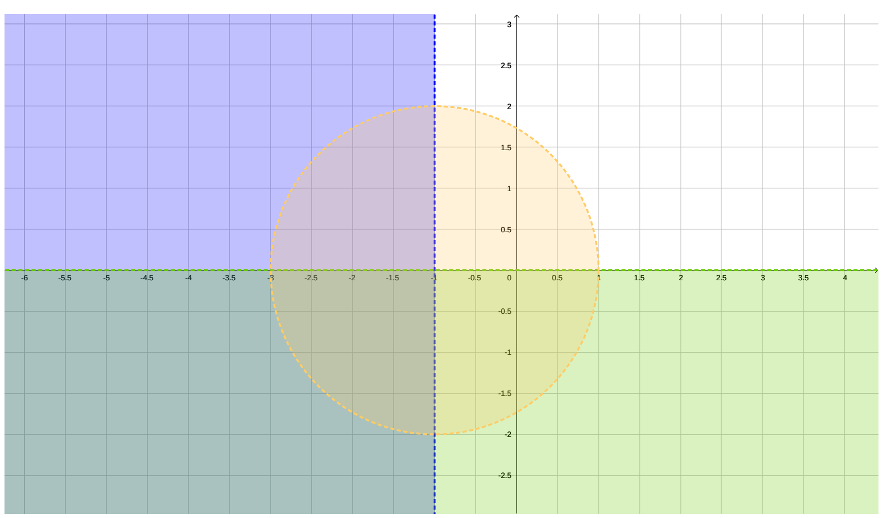

Figure 1 The feasible set X

Subproblem (b)

○ $\bar{x} := (-1,-2)^\top$, so we have

$$\begin{aligned} g_1(\bar{x}) &= (-1+1)^2 + (-2)^2 - 4 = 0 \\ g_2(\bar{x}) &= x_1 + 1 = -1 + 1 = 0 \\ g_3(\bar{x}) &= x_2 = -2 \end{aligned} \tag{1}$$

Assignment A5.1 continued on next page...

Thus, we can get $\mathcal{A}(\bar{x}) = \{1, 2\}$.
∘ In order to show that $\bar{x}$ is a regular point, we have to show that $\bar{x}$ is a feasible set and satisfy the LICQ. Firstly, it is easy to find that $\bar{x}$ is a feasible point. Then, we have to prove that it satisfies the LICQ. We have

$$\begin{aligned} \begin{pmatrix} \nabla g_1(\bar{x}) & \nabla g_2(\bar{x}) \end{pmatrix} &= \begin{pmatrix} \nabla g_1(\bar{x}) & \nabla g_2(\bar{x}) \end{pmatrix} \\ &= \begin{pmatrix} 2(x_1+1) & 1 \\ 2x_2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \end{aligned} \tag{2}$$

We can see the rank of this matrix is 2, which means the matrix has full rank, so $\nabla g_1(\bar{x})$ and $\nabla g_2(\bar{x})$ are linear independent. Thus, we have proved that $\bar{x}$ satisfies the LICQ. In all, we have showed that $\bar{x}$ is a regular point.

Subproblem (c)
Firstly, we derive the gradients and Hessian as follow:

$$\begin{aligned} &\nabla f(x) = \begin{pmatrix} 3\left(x_1+1\right)^2 - 2 \\ 2x_2 + 4 \end{pmatrix} \qquad \nabla^2 f(x) = \begin{pmatrix} 6x_1 + 6 & 0 \\ 0 & 2 \end{pmatrix} \\ &\nabla g_1(x) = \begin{pmatrix} 2\left(x_1+1\right) \\ 2x_2 \end{pmatrix} \qquad \nabla^2 g_1(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \\ &\nabla g_2(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \nabla^2 g_2(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ &\nabla g_3(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \nabla^2 g_3(x) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned} \tag{3}$$

Let's check the KKT-conditions. Due to $\mathcal{I}(\bar{x}) = \{3\}$, we know $\bar{\lambda}_3 = 0$.
∘ Main Cond:

$$\begin{aligned} &\nabla f(\bar{x}) + \bar{\lambda}_1 \nabla g_1(\bar{x}) + \bar{\lambda}_2 \nabla g_2(\bar{x}) = 0 \\ \Longrightarrow &\begin{pmatrix} -2 \\ 0 \end{pmatrix} + \bar{\lambda}_1 \begin{pmatrix} 0 \\ -4 \end{pmatrix} + \bar{\lambda}_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \\ \Longrightarrow &\bar{\lambda}_1 = 0, \quad \bar{\lambda}_2 = 2 \end{aligned} \tag{4}$$

∘ Dual Feasibility

$$\begin{aligned} \bar{\lambda}_1 &= 0 \geq 0 \\ \bar{\lambda}_2 &= 2 \geq 0 \\ \bar{\lambda}_3 &= 0 \geq 0 \end{aligned} \tag{5}$$

∘ Complementarity

$$\begin{aligned} &\bar{\lambda}_1 \cdot g_1(\bar{x}) = 0 \\ &\bar{\lambda}_2 \cdot g_2(\bar{x}) = 2 \cdot (-1+1) = 0 \\ &\bar{\lambda}_3 \cdot g_3(\bar{x}) = 0 \end{aligned} \tag{6}$$

∘ Primal Feasibility

$$\begin{aligned} g_1(\bar{x}) &= 0 \leq 0 \\ g_2(\bar{x}) &= 0 \leq 0 \\ g_3(\bar{x}) &= -2 \leq 0 \end{aligned} \tag{7}$$

Thus, $\bar{x}$ is a KKT point. The corresponding Lagrange multiplier $\bar{\lambda}=\left(\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3\right)^\top = (0,2,0)^\top$.

Subproblem (d)

$$\begin{aligned}
\mathcal{C}(\bar{x}) &= \left\{d \in \mathbb{R}^2 : \nabla f(\bar{x})^\top d = 0, \nabla g_i(\bar{x})^\top d \leq 0, \forall i \in \mathcal{A}(\bar{x})\right\} \\
&= \left\{d \in \mathbb{R}^2 : (-2,0)\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = 0;\ (0,-4)\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \leq 0;\ (1,0)\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} \leq 0\right\} \\
&= \left\{d \in \mathbb{R}^2 : d_1 = 0, d_2 \geq 0\right\} \\
&= \left\{(0,t)^\top, t \geq 0\right\}
\end{aligned} \tag{8}$$

Subproblem (e)

In order to investigate whether $\bar{x}$ is a local solution, we need to check whether $\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right)$ is positive definite. We have

$$\begin{aligned}
\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) &= \nabla^2 f(\bar{x}) + \sum_{i=1}^{3} \lambda_i \nabla^2 g_i(\bar{x}) \\
&= \nabla^2 f(\bar{x}) + \bar{\lambda}_2 \nabla^2 g_2(\bar{x}) \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}
\end{aligned} \tag{9}$$

Then we can get

$$d^\top \nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) d = (0,t)\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ t \end{pmatrix} = 2t^2 \tag{10}$$

So we have $d^\top \nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) d > 0$ for all $d \in \mathcal{C}(\bar{x}) \backslash \{0\}$. $\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right)$ is positive definite on $\mathcal{C}(\bar{x})$. Thus, we can conclude that $\bar{x}$ is a local solution.

Assignment A5.2

Let $n \in \mathbb{N}$ be given. Let us consider the problem

$$\min_{x \in \mathbb{R}^n} f(x) = \sum_{j=1}^{n} x_j^j \quad \text{s.t.} \quad g(x) = 1 - \|x\|_2^2 \leq 0$$

and let $X := \{x \in \mathbb{R}^n : g(x) \leq 0\}$ denote the corresponding feasible set.

a) Show that the LICQ is satisfied at every feasible point of (2) .

b) Verify that the point $\bar{x} = (1,0,0,\ldots,0)^\top \in \mathbb{R}^n$ and the multiplier $\bar{\lambda} = \frac{1}{2}$ form a KKT pair of problem (2)

c) Compute the critical cone $\mathcal{C}(\bar{x})$ and simplify as far as possible.

d) Apply the second-order necessary and sufficient conditions and show that $\bar{x}$ is a local solution of problem (2) if and only if $n \leq 2$.

Subproblem (a)

In order to show that LICQ satisfied at every point, we have to show that the vector set $\{\nabla g_i(\bar{x}) : i \in \mathcal{A}(\bar{x})\}$ is linearly independent. We have

$$\nabla g(\bar{x}) = \left(-2\bar{x}_1, -2\bar{x}_2, \ldots, -2\bar{x}_n\right)^\top \tag{11}$$

When $g(\bar{x}) < 0$, the active set $\mathcal{A}(\bar{x})$ is empty, so the LICQ is satisfied automatically at this point. When $g(\bar{x}) = 0$, the active set $\mathcal{A}(\bar{x}) = \{1\}$. So we have to prove $\{\nabla g_1(\bar{x})\}$ is linearly independent. We have

$$\begin{aligned} g(\bar{x}) &= 0 \\ \Longrightarrow \|\bar{x}\|_2^2 &= 1 \end{aligned} \tag{12}$$

Thus, $\bar{x} \neq 0$, which means $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ can not equal 0 simultaneously. So we can derive that $\nabla g_1(\bar{x}) \neq 0$, so the vector set $\{\nabla g_1(\bar{x})\}$ is linearly independent.
Above all, we have showed thta LICQ is satisfied at every feasible point.

Subproblem (b)
Firstly, we derive the gradients and Hessian as follow:

$$\begin{aligned} \nabla f(x) = \begin{pmatrix} 1 \\ 2x_2 \\ 3x_3^2 \\ \vdots \\ nx_n^{n-1} \end{pmatrix} \qquad \nabla^2 f(x) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 6x_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n(n-1)x_n^{n-2} \end{pmatrix} \\ \nabla g(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \\ -2x_3 \\ \vdots \\ -2x_n \end{pmatrix} \qquad \nabla^2 g(x) = \begin{pmatrix} -2 & 0 & 0 & \cdots & 0 \\ 0 & -2 & 0 & \cdots & 0 \\ 0 & 0 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -2 \end{pmatrix} \end{aligned} \tag{13}$$

Let's check the KKT-conditions. We have $g(\bar{x}) = 0$, so $\mathcal{A}(x) = \{1\}$.
○ Main Cond:

$$\nabla f(\bar{x}) + \bar{\lambda} \nabla g(\bar{x}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \bar{\lambda} \cdot \begin{pmatrix} -2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} -2 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0 \tag{14}$$

○ Dual Feasibility

$$\bar{\lambda} = \frac{1}{2} \geq 0 \tag{15}$$

○ Complemenrarity

$$\bar{\lambda} \cdot g(\bar{x}) = \frac{1}{2} \cdot 0 = 0 \tag{16}$$

○ Primal Feasibility

$$g(\bar{x}) = 0 \leq 0 \tag{17}$$

Thus, the point $\bar{x}$ satisfied the KKT conditions with $\bar{\lambda} = \frac{1}{2}$, and it is easy to find that $\bar{x}$ is a feasible point.
Above all, we have verified that the point $\bar{x}$ and the multiploer $\bar{\lambda}$ from a KKT pair.

Subproblem (c)

$$\begin{aligned} \mathcal{C}(\bar{x}) &= \left\{ d \in \mathbb{R}^n : \nabla f(\bar{x})^\top d = 0, \nabla g_i(\bar{x})^\top d \leq 0, \forall i \in \mathcal{A}(\bar{x}) \right\} \\ &= \left\{ d \in \mathbb{R}^n : (1, 0, 0, \cdots, 0) \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{pmatrix} = 0;\ (-2, 0, 0, \cdots, 0) \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{pmatrix} \leq 0 \right\} \\ &= \{ d \in \mathbb{R}^n : d_1 = 0 \} \end{aligned} \tag{18}$$

Assignment A5.2 continued on next page...

Subproblem (d)

In order to investigate whether $\bar{x}$ is a local solution, we need to check whether $\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right)$ is positive definite. We have

$$\begin{aligned}\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) &= \nabla^2 f(\bar{x}) + \bar{\lambda}\nabla^2 g(\bar{x}) \\ &= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} + \frac{1}{2} \cdot \begin{pmatrix} -2 & 0 & 0 & \cdots & 0 \\ 0 & -2 & 0 & \cdots & 0 \\ 0 & 0 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -2 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix}\end{aligned} \tag{19}$$

◦ When $n = 1$, we have

$$d^\top \nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) d = (0)\left(-1\right)\left(0\right) = 0 \tag{20}$$

We can see $\mathcal{C}(\bar{x}) = \{(0)\}$ in this situation. Thus, $\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right)$ is positive definite on $\mathcal{C}(\bar{x})$. Thus, $\bar{x}$ is a local solution.

◦ When $n = 2$, we have $\mathcal{C}(\bar{x}) = \left\{(0, d_2)^\top\right\}$

$$d^\top \nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) d = (0, d_2)\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ d_2 \end{pmatrix} = d_2^2 \tag{21}$$

So we have $d^\top \nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) d > 0$ for all $d \in \mathcal{C}(\bar{x})\backslash\{0\}$. $\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right)$ is positive definite on $\mathcal{C}(\bar{x})$. Thus, $\bar{x}$ is a local solution.

◦ When $n \geq 3$, we have $\mathcal{C}(\bar{x}) = \left\{(0, d_2, d_3, \cdots, d_n)^\top\right\}$

$$d^\top \nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) d = (0, d_2, d_3, \cdots, d_n)\begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix}\begin{pmatrix} 0 \\ d_2 \\ d_3 \\ \cdots \\ d_n \end{pmatrix} = d_2^2 - \sum_{i=3}^{n} d_i^2 \tag{22}$$

Because d_i is arbitrary for all $i \geq 2$, so we do not have $d^\top \nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right) d > 0$ for all $d \in \mathcal{A}(\bar{x})\backslash\{0\}$, which means $\nabla^2_{xx} L\left(\bar{x}, \bar{\lambda}\right)$ is not positive definite on $\mathcal{C}(\bar{x})$. Thus, $\bar{x}$ is not local solution.

Assignment A5.3

Let $a \in \mathbb{R}^n\backslash\{0\}$ and $b \in \mathbb{R}$ be given and define the set $C := \left\{x \in \mathbb{R}^n : a^\top x \leq b\right\}$. Compute the projection $\mathcal{P}_C(x)$ for $x \in \mathbb{R}^n$, i.e., solve the optimization problem

$$\min_{y \in \mathbb{R}^n} \frac{1}{2}\|y - x\|^2 \quad \text{s.t.} \quad a^\top y \leq b$$

We have

$$\begin{aligned} f(y) &= \frac{1}{2}\|y - x\|^2 & \nabla f(y) &= y - x \\ g(y) &= a^\top y - b \leq 0 & \nabla g(y) &= a \end{aligned} \tag{23}$$

Assignment A5.3 continued on next page...

We apply KKT condition as follow, suppose y^* is the solution.
○ Main Cond

$$\nabla f(y^*) + \lambda \nabla g(y^*) = y^* - x + \lambda \cdot a = 0 \tag{24}$$

○ Dual Feasibility

$$\lambda \geq 0 \tag{25}$$

○ Complementarity

$$\lambda \cdot g(y^*) = 0 \tag{26}$$

○ Primal Feasibility

$$g(y^*) \leq 0 \tag{27}$$

• Case 1: $a^\top y^* < b$, which means $g(y^*) < 0$
Due to equation 26, we can have $\lambda = 0$. Due to equation 24, we can have $y^* = x$. Due to $a^\top y^* < b$ and $y^* = x$, we can get $a^\top x < b$. Thus, when $a^\top x < b$, we have $y^* = x$.
• Case 2: $a^\top y^* = b$, which means $g(y^*) = 0$
In this situation, Complementarity is satisfied. And we have

$$\begin{aligned} & g(y^*) = 0 \\ \Longrightarrow & a^\top y^* = b \end{aligned} \tag{28}$$

Combine the equation 24 and equation 28, we can get

$$\begin{aligned} \lambda &= \frac{a^\top x - b}{a^\top a} \\ y^* &= x - \frac{a^\top x - b}{a^\top a} \cdot a \end{aligned} \tag{29}$$

In this situation, the λ should satisfy $\lambda \geq 0$. So

$$\begin{aligned} \lambda &= \frac{a^\top x - b}{a^\top a} \geq 0 \\ \Longrightarrow & a^\top x \geq b \end{aligned} \tag{30}$$

Thus, when $a^\top x \geq b$, we have $y^* = x - \left(a^\top x - b\right) \cdot \left(a^\top a\right)^{-1} \cdot a$.
It is easy to show that the function $f(y)$ is convex. Because $g(x) = x^2, (x \in \mathbb{R})$ is convex and non-decreasing. Besides, $h(x) = \|x\|, (x \in \mathbb{R}^n)$ is convex. So the composition $g(h(x))$ is convex, so $f(y)$ is convex. On the other hand, the set $\left\{y \in \mathbb{R}^n : a^\top y \leq b\right\}$ is a half space, so it is convex. Overall, the problem is a convex optimization problem. So the KKT-point y^* is a global solution.
Thus we have derived the projection

$$\mathcal{P}_C(x) = \begin{cases} x & \left(a^\top x < b\right) \\ x - \left(a^\top x - b\right) \cdot \left(a^\top a\right)^{-1} \cdot a & \left(a^\top x \geq b\right) \end{cases} \tag{31}$$