University of Vienna
Faculty of Mathematics

Nonlinear Optimization Problems

Milutin Popovic

Contents

1 Sheet 2

1.1 Exercise 7

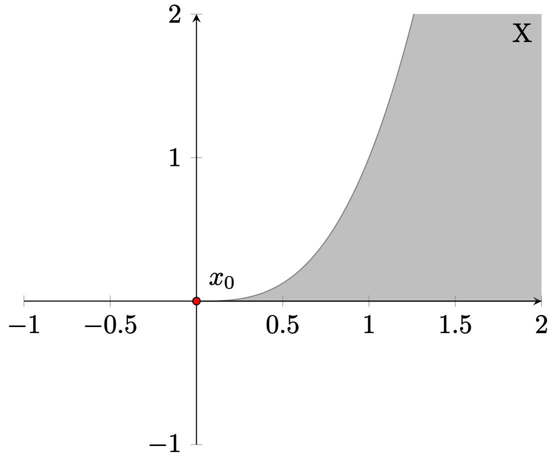
For the functions $g : \mathbb{R}^2 \to \mathbb{R}^2$, find $X = \{(x, y) \in \mathbb{R}^2 : g(x, y) \leqq 0\}$, the tangent cone and the linearized tangent cone at $x_0 \in X$ and find out if x_0 fulfills (ADABIDE-CQ), i.e. $T_{\text{lin}}(x_0) = T_X(x_0)$.

1. $g(x, y) = (y - x^3, -y)^T, \quad x_0 = (0, 0)^T$
2. $g(x, y) = (y^2 - x + 1, 1 - x - y)^T, \quad x_0 = (1, 0)^T$

For 1. we have that $g(x, y) \leqq 0$ means that

$$y - x^3 \leq_0 \tag{1}$$
$$-y \leq 0 \tag{2}$$
$$\Rightarrow 0 \leq y \leq x^3. \tag{3}$$

So is defined as $X = \{(x, y) \in \mathbb{R}^2 : 0 \leq y, y \leq x^3\}$. Graphically represented X looks like the following



Then the tangent cone is

$$T_X(x_0) = \left\{ \begin{pmatrix} \lambda \\ 0 \end{pmatrix} : \lambda \geq 0 \right\}. \tag{4}$$

Now for the linearized tangent cone we calculate, $g_1(x_0) = 0$ and $g_2(x_0) = 0$ meaning that $\mathcal{A}(x_0) = \{1, 2\}$ thereby

$$\begin{aligned} T_{\text{lin}}(x_0) &= \{d \in \mathbb{R}^2 : \nabla g_1(x_0)^T d \leq 0,\ \nabla g_2(x_0)^T d \leq 0\} && (5) \\ &= \left\{ d \in \mathbb{R}^2 : \begin{pmatrix} 0 \\ 1 \end{pmatrix}^T d \leq 0,\ \begin{pmatrix} 0 \\ -1 \end{pmatrix}^T d \leq 0 \right\} && (6) \\ &= \left\{ \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \begin{pmatrix} -\lambda \\ 0 \end{pmatrix} :\ \lambda \geq 0 \right\}. && (7) \end{aligned}$$

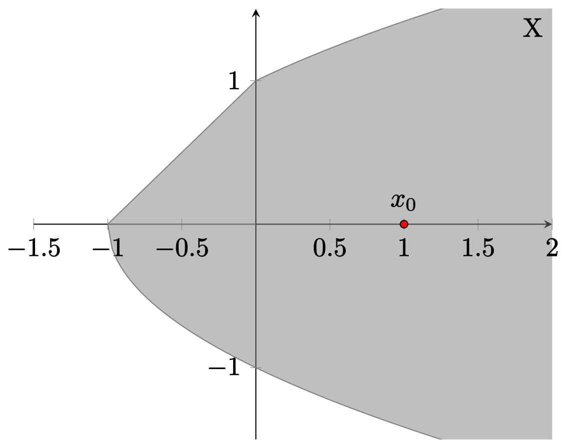
We conclude that $x_0 = (0, 0)^T$ does not satisfy the ADABIE-CQ condition for this optimization problem.
For number 2. first the domain X, $g(x, y) \leq 0$

$$\begin{aligned} y^2 - x + 1 \leq 0 \quad &\text{and} \quad 1 - x - y \leq 0 && (8) \\ y^2 - 1 \leq x \quad &\text{and} \quad y - 1 \leq x && (9) \end{aligned}$$

so X has the following form

$$X = \left\{ (x, y) \in \mathbb{R}^2 : \begin{cases} -\sqrt{x+1} \leq y \leq x + 1 & \text{for } x \in (-1, 0] \\ -\sqrt{x+1} \leq y \leq \sqrt{x+1} & \text{for } x > 0 \end{cases} \right\} \tag{10}$$

and graphically



Then the tangent cone is obviously

$$T_X(x_0) = \{d : d \in \mathbb{R}^2\} \tag{11}$$

For the linearized tangent cone we calculate $g_1(x_0) = 0$ and $g_2(x_0) = 0$, thereby $\mathcal{A}(x_0) = \{1, 2\}$ and

$$T_{\text{lin}}(x_0) = \left\{ d \in \mathbb{R}^2 : \begin{pmatrix} -1 \\ 0 \end{pmatrix}^T d \le 0,\ \begin{pmatrix} -1 \\ -1 \end{pmatrix}^T d \le 0 \right\} \tag{12}$$

$$= \left\{ \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} \lambda \\ -\lambda \end{pmatrix} : \lambda \ge 0 \right\}. \tag{13}$$

In this case x_0 also does not satisfy the ABADIE-CQ.

1.2 Exercise 8

Let (x^*, λ^*, μ^*) be a KKT point of the optimization problem

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \le 0, i = 1, \ldots, m \\ & h_j(x) = 0, j = 1, \ldots, p \\ & x \in \mathbb{R}^n \end{aligned} \tag{14}$$

for $f, g_i, h_i : \mathbb{R}^n \to \mathbb{R}$ continuously differentiable functions. Prove that x^* is a critical point of the optimization point, namely that it holds

$$\nabla f(x^*)^T \ge 0 \ \ \forall d \in T_X(x^*), \tag{15}$$

where $X = \{x \in R^n : g_i(x) \le 0, i = 1, \ldots, m, h_j(x) = 0, j = 1, \ldots, p\}$. Given a critical point x^* when do Lagrange multipliers λ^*, μ^* exist such that (x^*, λ^*, μ^*) is a KKT point?
Firs of all if (x^*, λ^*, μ^*) is a KKT point then

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0 \tag{16}$$

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla h_j(x^*) = 0 \tag{17}$$

is satisfied for the Lagrangian. Then we can take the scalar product with $d \in T_X(x^*)$. We know that $\nabla g_i(x^*)^T d \le 0$ and $\nabla h_j(x^*)^T d = 0$ for all $i = 1, \ldots, m$ and $j = 1, \ldots, p$ and $\lambda_i^* \ge 0$ which means

$$0 = \nabla f(x^*)^T d + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*)^T d + \sum_{j=1}^{p} \mu_j^* \nabla h_j(x^*)^T d \tag{18}$$

$$= \nabla f(x^*)^T d + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*)^T d \tag{19}$$

$$\le \nabla f(x^*)^T d. \tag{20}$$

This concludes

$$\nabla f(x^*)^T d \ge 0. \tag{21}$$

Now if x^* is a critical point then it is a local minimum. If it fulfills the ABADIE-CQ condition then there exist $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^p$ such that (x^*, λ^*, μ^*) is a KKT point. We know that X is convex and x^* fulfills the ABADIE-CQ then $\nabla f(x^*) \in (T_X(x^*)^*)$ and $\left(T_X(x^{(*)})\right)^* = (T_{\text{lin}}(x^*))^*$. This means that $\nabla f(x^*) \in (T_{\text{lin}}(x^*))^*$. By Farkas Lemma there exist $\lambda_i^* \ge 0$ and μ_j^*, $i = 1, \ldots, m$, $j = 1, \ldots, p$ such that $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$, then (x^*, λ^*, μ^*) is a KKT point.

1.3 Exercise 9

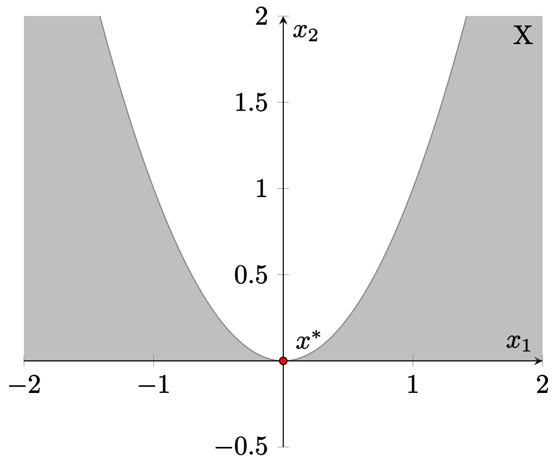
Consider the optimization problem

$$\begin{aligned} \min \quad & x_1^2 (x_2+1)^2 , && (22)\\ \text{s.t.} \quad & x_1^3 - x_2 \le 0 && (23)\\ & x_2 \le 0. && (24) \end{aligned}$$

Show that $x^* = (0,0)^T$ fulfills ABADIE-CQ but not MFCQ.
The domain X is defined by $x_1^2 \ge x^2$ and $x_2 \ge 0$,

$$X = \left\{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 \ge x_2 \ge 0\right\}, \tag{25}$$

graphically



meaning that

$$T_X(x^*) = \left\{\begin{pmatrix}-\lambda\\0\end{pmatrix}, \begin{pmatrix}\lambda\\0\end{pmatrix} : \lambda \ge 0\right\}, \tag{26}$$

Then

$$\begin{aligned} T_{\text{lin}}(x^*) &= \left\{ d \in \mathbb{R}^2 : \begin{pmatrix}0\\1\end{pmatrix}^T d \le 0,\ \begin{pmatrix}0\\-1\end{pmatrix}^T d \le 0 \right\} && (27)\\ &= \left\{\begin{pmatrix}-\lambda\\0\end{pmatrix}, (\lambda, 0) : \lambda \ge 0\right\}. && (28) \end{aligned}$$

This means that x^* fulfills the ABADIE-CQ condition. On the other hand MFCQ is fulfilled only if there exists $d \in \mathbb{R}^2$ such that $\nabla g_i(x^*)^T d < 0$, for all $i \in \mathcal{A}(x^*)$ but the problem is the strict constraint

$$\nabla g_1(x^*) = \begin{pmatrix}0\\1\end{pmatrix}, \quad \nabla g_2(x^*) = \begin{pmatrix}0\\-1\end{pmatrix}. \tag{29}$$

Any feasible solutions are of the form $(\pm\lambda, 0)^T$, $\lambda \ge 0$. Both cases always equal to 0.

1.4 Exercise 10

Consider the optimization problem

$$\begin{aligned} \min \quad & x_1^2 (x_2+1)^2 , && (30)\\ \text{s.t.} \quad & -x_1^3 - x_2 \le 0 && (31)\\ & -x_2 \le 0. && (32) \end{aligned}$$

Show that $x^* = (0,0)^T$ fulfills MFCQ but not LICQ.
The domain X is defined by $x_1^2 \geq -x^2$ and $x_2 \geq 0$,

$$X = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 0\}, \tag{33}$$

and $g_1(x^*) = 0$ and $g_2(x^*) = 0$ so $\mathcal{A}(x^*) = \{1, 2\}$.

$$\nabla g_1(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \nabla g_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{34}$$

For strict inequality $\nabla g_i(x^*)^T d <=$ for all $i \in \mathcal{A}(x^*)$ we have that $d = (0, \lambda)$ with $\lambda > 0$. This means x_0 fulfills MFCQ. On the other hand LICQ is fulfilled if

$$\{\nabla g_i(x^*)\}_{i \in \mathcal{A}(x^*)} \tag{35}$$

are linearly independent. But in our case $\nabla g_1(x^*) = \nabla g_2(x^*)$, meaning that x_0 does not fulfill LICQ.

1.5 Exercise 11

Let $U \subseteq \mathbb{R}^n$ be a nonempty, open convex set and $f \in U \to \mathbb{R}$ a differentiable function on U. Prove that the following statements are equivalent.

1. f is convex on U
2. $\langle \nabla f(x), y - x \rangle \leq f(y) - f(x) \quad \forall x, y \in U$
3. $\langle \nabla f(x) - \nabla f(y), y - x \rangle \leq 0 \quad \forall x, y \in U$
4. if f is twice differentiable on U, then $\nabla^2 f(x)$ is positively semi definite for every $x \in U$.

We start with (1) $\Leftrightarrow$ (2).
Ad $\Rightarrow$: f is convex, then for all $x, y \in U$, $\lambda \in [0,1]$ we have

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda) f(x) + \lambda(y) \tag{36}$$
$$= f(x) + \lambda (f(y) - f(x)) \tag{37}$$
$$\frac{f((1-\lambda)x + \lambda y) - f(x)}{\lambda} \leq f(y) - f(x). \tag{38}$$

Letting $\lambda \downarrow 0$ we get

$$\nabla f(x)^T (y - x) \leq f(x) - f(y) \tag{39}$$

Ad $\Leftarrow$: we have that $\forall x, y \in U$:

$$\nabla f(x)^T (y - x) \leq f(x) - f(y). \tag{40}$$

Since U is convex then the above also holds for $z \in U$ where $z = (1-\lambda)x + \lambda y$, then

$$f(x) \geq f(z) + \nabla f(z)^T (x - z) \quad | \cdot (1 - \lambda) \tag{41}$$
$$f(y) \geq f(z) + \nabla f(z)^T (y - z) \quad | \cdot \lambda \tag{42}$$

adding both of them together we get

$$(1-\lambda) f(x) + \lambda f(y) \geq f(z) + \nabla f(z)^T ((1-\lambda)x + \lambda y - z) \tag{43}$$
$$= f(z) \tag{44}$$
$$= f((1-\lambda)x + \lambda y). \tag{45}$$

This shows that f is convex on U.
Next we show (2) $\Leftrightarrow$ (3).
Ad $\Rightarrow$: We start with

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \tag{46}$$
$$f(x) \geq f(y) + \nabla f(y)^T (x - y). \tag{47}$$

Adding them together we get

$$\nabla f(y)^T(y-x) - \nabla f(x)^T(y-x) \geq 0 \tag{48}$$
$$\left(\nabla f(y)^T - \nabla f(x)^T\right)(y-x) \geq 0. \tag{49}$$

Ad $\Leftarrow$: We can just do the same operations as in $\Rightarrow$ in reverse.
Now we prove (2) $\Leftrightarrow$ (4).First we consider in one dimension and then generalize
Ad $\Rightarrow$: . In $U \subseteq \mathbb{R}$ we have that $\forall x, y \in U$

$$f(y) \geq f(x) + f(x)'(y-x) \tag{50}$$
$$f(x) \geq f(y) + f(y)'(x-y). \tag{51}$$

Let $x < y$, then

$$f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x) \quad |\frac{1}{(y-x)^2} \tag{52}$$
$$\frac{f'(y)-f'(x)}{y-x} \geq 0 \quad |y \to x \tag{53}$$
$$f''(x) \geq 0 \quad \forall x \in U. \tag{54}$$

Ad $\Leftarrow$: We use Taylors expansion formula for $f(y)$ in $x \in U$

$$f(y) = f(x) + f'(x)(y-x) + \frac{1}{2}f''(\xi)(y-x) \quad \xi \in [x, y] \tag{55}$$
$$f(y) \geq f(x) + f'(x)(y-x). \tag{56}$$

In general dimensions convexivity means convexivity along all directions, i.e. $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex if

$$g(\alpha) = f(x + \alpha d) \tag{57}$$

is convex $\forall x \in U$ and $\forall d \in \mathbb{R}^n$. This is exactly the case if

$$g''(\alpha) = d^T \nabla^2 f(x + \alpha d) d \geq 0 \quad \forall x \in U \; \forall d \in \mathbb{R}^n \; \forall \alpha \in \mathbb{R} \tag{58}$$

such that $x + \alpha d \in U$ so f is convex if and only if

$$\nabla f(x) \geq 0 \quad \forall x \in U \quad \square \tag{59}$$

1.6 Exercise 12

Let $c : \mathbb{R} \to \mathbb{R}$ be defined as

$$c(y) = \begin{cases} (y+1)^2 & y < -1 \\ 0 & -1 \leq y \leq 1 \\ (y-1)^2 & y > 1 \end{cases} \tag{60}$$

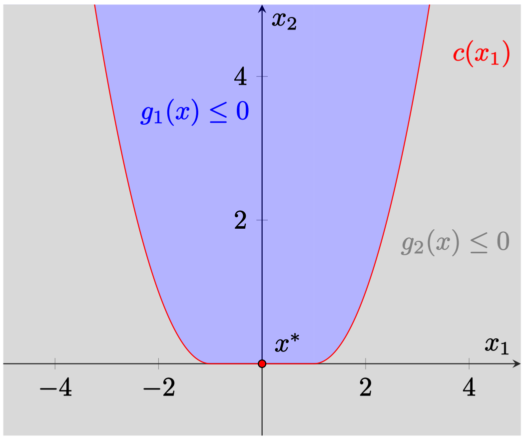
Let $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$

$$g_1(x_1, x_2) = c(x_1) - x_2 \tag{61}$$
$$g_2(x_1, x_2) = c(x_1) + x_2 \tag{62}$$
$$\tag{63}$$

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a convex function and continuously differentiable. Show that for the convex optimization problem

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \leq 0, i = 1, 2 \\ & x \in \mathbb{R}^2 \end{aligned} \tag{64}$$

ABADIE-CQ holds at $x^* = (0,0)^T$ SLATER-CQ is not satisfied.
Bellow is a graphical representation of, $c(x_1)$, $g_1(x) \leq 0$ and $g_2(x)$



So X has only elements on the curve $c(x)$, i.e. $X = \{x \in \mathbb{R}^2 : g_1(x) \leq 0, g_2(x) \leq 0\} = \{(x_1, c(x_1))^T : x_1 \in \mathbb{R}\}$ and thereby the tangent cone of X at x^* consists of tangent vectors of $c(x)$ at x^*

$$T_X(x^*) = \left\{ \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \begin{pmatrix} -\lambda \\ 0 \end{pmatrix} : \lambda \geq 0 \right\}. \tag{65}$$

For the linearized tangent cone we have that $g_1(x^*) = c(0) = 0$ and $g_2(x^*) = c(0) = 0$, then the gradients at x^* are

$$\nabla g_1(x^*) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \nabla g_2(x^*) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}. \tag{66}$$

Thereby

$$T_{\text{lin}}(x^*) = \left\{ d \in \mathbb{R}^2 : \nabla g_1(x)^T d \leq 0, \nabla g_2(x)^T d \leq 0 \right\} \tag{67}$$
$$= \left\{ \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \begin{pmatrix} -\lambda \\ 0 \end{pmatrix} : \lambda \geq 0 \right\}. \tag{68}$$

We have that x^* satisfies ABADIE-CQ.
In our case SLATER-CQ is fulfilled if there exists an $x' \in \mathbb{R}^2$ such that $g_i(x') < 0$ for all $i = 1, 2$. The problem arises because in case of strict inequality the domains defined by $g_1(x) < 0$ and $g_2(x) < 0$ do not match for any x as seen the figure above. In the relaxed case they match exactly at the line $c(x_1)$. But $c(x_1) \geq 0$. Meaning that there exists no x' such that SLATER-CQ is satisfied (in our case).