Nonlinear Programming Assignment 001

Nguyen Quan Ba Hong*
Students at Faculty of Math and Computer Science,
Ho Chi Minh University of Science, Vietnam
email. nguyenquanbahong@gmail.com
blog. www.nguyenquanbahong.com †

January 1, 2018

*Student ID: 1411103.

Contents

1 Problems

Problem 1. *Let $f:\mathbb{R}\to\mathbb{R}^2$ be a mapping defined by*

$$f(x)=\begin{cases}(x,x^2) & \text{if } x\neq 0,\\ (0,0) & \text{if } x=0.\end{cases} \tag{1.1}$$

1. *Is f directional differentiable at $x_0=0$?*
2. *Is f Gâteaux differentiable at $x_0=0$?*
3. *Is f Fréchet differentiable at $x_0=0$?*

SOLUTION.

1. Let $d\in\mathbb{R}$, at $x_0=0$, we have

$$\lim_{t\to 0}\frac{f(x_0+td)-f(x_0)}{t}=\lim_{t\to 0}\frac{f(td)-f(0)}{t} \tag{1.2}$$
$$=\lim_{t\to 0}\left(d,td^2\right)^T \tag{1.3}$$
$$=(d,0)^T, \tag{1.4}$$

i.e., f is directional differentiable at $x_0=0$ and its directional derivative is given by $f'(0;d)=(d,0)^T$.

2. We write the vector-valued function $f:\mathbb{R}\to\mathbb{R}^2$ defined by (1.1) as

$$f(x)=\begin{pmatrix} f_1(x)\\ f_2(x)\end{pmatrix}, \tag{1.5}$$

where the coordinate functions of f are given by $f_1(x)=x, f_2(x)=x^2$. From the above result, we have

$$f'(0;d)=(d,0)=\begin{pmatrix}1\\0\end{pmatrix}d,\quad \forall d\in\mathbb{R}, \tag{1.6}$$

which is linear in d. Hence, f is Gâteaux differentiable at $x_0=0$.

3. Since f is Fréchet differentiable at x if and only if each coordinate function f_i is. So it suffices to prove that f_1,f_2 are Fréchet continuous at $x_0=0$. This is obvious since $f_1,f_2\in C^\infty(\mathbb{R})$. Hence, f is Fréchet differentiable at $x_0=0$. Moreover, $Df(0)=\left({f_1}'(0),{f_2}'(0)\right)=(1,0)$.

This completes our solution. □

Problem 2. *Let $f:\mathbb{R}^2\to\mathbb{R}$ be a mapping defined by*

$$f(x,y)=\begin{cases}\dfrac{x^2y^4}{x^4+y^8} & \text{if } (x,y)\neq(0,0),\\ 0 & \text{if } (x,y)=(0,0).\end{cases} \tag{1.7}$$

1. *Is f directional differentiable at $x_0=(0,0)$?*
2. *Is f Gâteaux differentiable at $x_0=(0,0)$?*

3. *Is f Fréchet differentiable at $x_0 = (0,0)$?*

SOLUTION.

1. Let $d \in \mathbb{R}^2$ be a vector $d = (d_1, d_2)^T$, and $x_0 = (0,0)$. If $d = (0,0)$, we have $f'(x_0; (0,0)) = 0$ by the definition of directional derivative. If $d \neq (0,0)$, we compute

$$\lim_{t \to 0} \frac{f(x_0 + td) - f(x_0)}{t} = \lim_{t \to 0} \frac{f(td_1, td_2) - f(0,0)}{t} \tag{1.8}$$

$$= \lim_{t \to 0} \frac{td_1^2 d_2^4}{d_1^4 + t^4 d_2^8}. \tag{1.9}$$

We consider the following two cases depending on d_1. If $d_1 = 0$, then the limit in (1.9) equals 0. If $d_1 \neq 0$, we estimate this limit as follows.

$$\lim_{t \to 0} \left| \frac{td_1^2 d_2^4}{d_1^4 + t^4 d_2^8} \right| \leq \lim_{t \to 0} \frac{|t| \, d_1^2 d_2^4}{d_1^4} = 0, \tag{1.10}$$

i.e., the limit in (1.9) also equals 0 in this case. Combining both cases, we deduce that f is directional differentiable at x_0 and its directional derivative is given by $f'(x_0; d) = 0$ for all $d \in \mathbb{R}^2$.

2. From the above result, we have

$$f'(x_0; d) = 0 = \begin{pmatrix} 0 & 0 \end{pmatrix} d, \quad \forall d \in \mathbb{R}^2, \tag{1.11}$$

which is linear in d. Hence, f is Gâteaux differentiable at $x_0 = (0,0)$.

3. We claim that f is not Fréchet differentiable at $x_0 = (0,0)$. To this end, we suppose for the contrary that f is Fréchet differentiable at $x_0 = (0,0)$, by definition, there exists a linear function $l : \mathbb{R}^2 \to \mathbb{R}$, $l(x) = \langle l, x \rangle = l_1 x_1 + l_2 x_2$ such that

$$\lim_{\|h\| \to 0} \frac{f(x_0 + h) - f(x_0) - \langle l, h \rangle}{\|h\|} = 0. \tag{1.12}$$

Denote $h = (h_1, h_2)^T \in \mathbb{R}^2$, then (1.12) becomes

$$\lim_{\|h\| \to 0} \frac{1}{\|h\|} \left(\frac{h_1^2 h_2^4}{h_1^4 + h_2^8} - l_1 h_1 - l_2 h_2 \right) = 0. \tag{1.13}$$

In particular, if we take $h = (h_1, 0)$ for which $h_1 \neq 0$ and $h_1 \to 0$, then (1.13) gives $\lim\limits_{h_1 \to 0} \frac{l_1 h_1}{|h_1|} = 0$, i.e., $l_1 = 0$. Similarly, taking $h = (0, h_2)$ for which $h_2 \neq 0$ and $h_2 \to 0$ gives $l_2 = 0$. Substituting $l_1 = l_2 = 0$ back to (1.13) gives

$$\lim_{\|h\| \to 0} \frac{h_1^2 h_2^4}{(h_1^4 + h_2^8) \sqrt{h_1^2 + h_2^2}} = 0. \tag{1.14}$$

But (1.14) is not true since, for instance, taking $h_2^2 = h_1$ in (1.14), i.e., $h = \left(h_1, \sqrt{h_1}\right)$, gives

$$\lim_{h_1 \to 0} \frac{1}{2\sqrt{h_1^2 + h_1}} = 0, \tag{1.15}$$

which is absurd, since the limit in the left-hand side of (1.15) is $+\infty$.

This contradiction ends our proof. $\square$

Problem 3. *Let X be a normed space, $M \subset X$ and $x_0 \in X$. The contingent cone (or tangent cone, Bouligand cone) of M at x_0 is defined by the following formula*

$$T(M, x_0) = \left\{u \in X | \exists t_n \to 0^+, u_n \to u, x_0 + t_n u_n \in M, \ \ \forall n \in \mathbb{N}\right\}. \tag{1.16}$$

By this definition, compute the following contingent cones of

1. $M = \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1^3 + x_2^2 = 0\right\}$ and $x_0 = (0, 0)$.
2. $M = \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1 + x_2 \geq 2, x_2 \leq x_1^3\right\}$ and $x_0 = (1, 1)$.

Solution. An alternative definition of *tangent cone* can be found in [1], Def. 2.28, p.47.

1. Setting $X = \mathbb{R}^2$, we notice $x_0 = (0, 0) \in M$. We claim that

$$T(M, x_0) = \widehat{T}(M, x_0) := \left\{(x, 0) \in \mathbb{R}^2 | x \leq 0\right\}. \tag{1.17}$$

To prove (1.17), we prove the following inclusions.

(a) *Prove $T(M, x_0) \subset \widehat{T}(M, x_0)$.* Taking $u = (x, y) \in T(M, x_0)$, by definition (1.16), there exist a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and a sequence $\{u_n\}_{n=1}^{\infty} \subset \mathbb{R}^2$ such that $u_n \to u$ as $n \to \infty$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. Set $u_n := (x_n, y_n)$, the fact $u_n \to u$ implies that $x_n \to x$ and $y_n \to y$, and the fact $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$ gives

$$t_n^3 x_n^3 + t_n^2 y_n^2 = 0, \ \ \forall n \in \mathbb{N}. \tag{1.18}$$

Since $t_n > 0$ for all $n \in \mathbb{N}$, (1.18) then implies

$$t_n x_n^3 + y_n^2 = 0, \ \ \forall n \in \mathbb{N}. \tag{1.19}$$

We see at a glance from (1.19) that $x_n \leq 0$ for all $n \in \mathbb{N}$. Hence, $x \leq 0$ (since $x_n \to x$ as $n \to \infty$). Now let $n \to \infty$ in (1.19) and use the given limits $x_n \to x, y_n \to y$ and $t_n \to 0^+$, we obtain $y = 0$. Hence, $u \in \widehat{T}(M, x_0)$ and our first inclusion is proved.

(b) *Prove $\widehat{T}(M, x_0) \subset T(M, x_0)$.* Taking $u = (x, 0)$ satisfying $x \leq 0$, we claim that $u \in T(M, x_0)$. To this end, we now choose $x_n = x - \frac{1}{n} < 0, y_n = \frac{1}{n^2}$[1]. This choice ensures that $u_n := (x_n, y_n) \to u := (x, 0)$ as $n \to \infty$. It then suffices to prove that there exists a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. The latter gives, using (1.19) again,

$$t_n \left(x - \frac{1}{n}\right)^3 + \frac{1}{n^4} = 0, \ \ \forall n \in \mathbb{N}. \tag{1.20}$$

[1] If $x < 0$, we can choose $x_n = x, y_n = \frac{1}{n}$ and then (1.19) gives $t_n = -\frac{1}{n^2 x^3} \to 0^+$ as $n \to \infty$, as desired. Unfortunately, this choice does not work for $x = 0$, so we used the above choice.

i.e.,

$$t_n = -\frac{1}{n^4\left(x-\frac{1}{n}\right)^3}, \quad \forall n \in \mathbb{N}. \tag{1.21}$$

It is easy to check that $t_n > 0$ (since $x \le 0$) and $t_n \to 0^+$ as $n \to \infty$.[2] Hence, $u \in T(M, x_0)$, the second inclusion is also proved.

Combining these, we conclude that (1.17) holds, i.e.,

$$T(M, x_0) = \left\{(x, 0) \in \mathbb{R}^2 | x \le 0\right\}. \tag{1.22}$$

2. Let $X = \mathbb{R}^2$ again, note that $x_0 = (1, 1) \in M$, we claim that

$$T(M, x_0) = \widehat{T}(M, x_0) := \left\{(x_1, x_2) \in \mathbb{R}^2 | x_1 + x_2 \ge 0, x_2 \le 3x_1\right\}. \tag{1.23}$$

We also prove the following two inclusions as before.

(a) *Prove* $T(M, x_0) \subset \widehat{T}(M, x_0)$. Taking $u = (x, y) \in T(M, x_0)$, by (1.16), there exist a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and a sequence $\{u_n\}_{n=1}^{\infty} \subset \mathbb{R}^2$ such that $u_n \to u$ as $n \to \infty$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. Set $u_n := (x_n, y_n)$, the fact $u_n \to u$ implies that $x_n \to x$ and $y_n \to y$, and the fact $x_0 + t_n u_n = (1 + t_n x_n, 1 + t_n y_n) \in M$ for all $n \in \mathbb{N}$ gives

$$(1 + t_n x_n) + (1 + t_n y_n) \ge 2 \text{ and } 1 + t_n y_n \le (1 + t_n x_n)^3, \tag{1.24}$$

for all $n \in \mathbb{N}$, equivalently,[3]

$$x_n + y_n \ge 0 \text{ and } y_n \le 3x_n + 3t_n x_n^2 + t_n^2 x_n^3, \quad \forall n \in \mathbb{N}. \tag{1.25}$$

Now let $n \to \infty$ in (1.25) and use the given limits $x_n \to x, y_n \to y$ and $t_n \to 0^+$, we obtain $x + y \ge 0$ and $y \le 3x$. Hence, $u \in \widehat{T}(M, x_0)$ and our first inclusion is proved.

(b) *Prove* $\widehat{T}(M, x_0) \subset T(M, x_0)$. Taking $u = (x, y) \in \widehat{T}(M, x_0)$, i.e., x, y satisfy $x + y \ge 0$ and $y \le 3x$, we claim that $u \in T(M, x_0)$. To this end, first notice $4x \ge x + y \ge 0$, so $x \ge 0$. We then choose $u_n = (x_n, y_n)$ where $x_n = x + \frac{1}{n} \ge \frac{1}{n}, y_n = y$ and $t_n \to 0^+$ arbitrarily, so that $u_n \to u$ as $n \to \infty$. It is easy to check that (1.25) holds for chosen x_n, y_n and t_n:

$$x_n + y_n = x + y + \frac{1}{n} \ge \frac{1}{n} > 0, \tag{1.26}$$

$$3x_n + 3t_n x_n^2 + t_n^2 x_n^3 = 3x + \frac{3}{n} + 3t_n x_n^2 + t_n^2 x_n^3 > 3x \ge y = y_n. \tag{1.27}$$

Hence, $u \in T(M, x_0)$, the second inclusion is also proved.

[2] If $x = 0$, then $t_n = \frac{1}{n} \to 0$ as $n \to \infty$. If $x < 0$, then $t_n \to -\frac{1}{x^3} \lim_{n\to\infty} \frac{1}{n^4} = 0$.

[3] It can be deduced from (1.24) that $t_n(x_n + y_n) \ge 0$. Since $t_n \to 0^+$, $t_n > 0$ for all $n \in \mathbb{N}$ and the factor t_n can be dropped as in (1.25).

Combining these cases, we conclude that (1.23) holds, i.e.,

$$T(M,x_0)=\left\{(x_1,x_2)\in\mathbb{R}^2|x_1+x_2\geq 0,x_2\leq 3x_1\right\}. \tag{1.28}$$

This completes our proof. □

The following exercise will show us some properties of contingent cone of first order.

Problem 4. *Let X be a normed space, $M\subset X$ and $x_0\in X$.*

1. *If $T(M,x_0)\neq\emptyset$ then $x_0\in\overline{M}$ (where $\overline{M}$ is the closure of the set M).*
2. *$T(M,x_0)$ is a closed cone.*
3. *$T(M,x_0)\subset\overline{cone\,(M-x_0)}$.*

   Moreover, if M is a convex set then
4. *$T(M,x_0)=\overline{\mathrm{cone}\,(M-x_0)}$, and hence, $T(M,x_0)$ is a convex set.*
5. *$T(M,x_0)=\{v\in X|\forall t_n\to 0^+,\forall v_n\to v,x_0+t_nv_n\in M\}$.*

Solution.

1. Suppose that $T(M,x_0)\neq\emptyset$, we can take, for instance, $u\in T(M,x_0)$. Then there exist a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n\to 0^+$ and a sequence $\{u_n\}_{n=1}^{\infty}\subset X$ such that $u_n\to u$ and $x_0+t_nu_n\in M$ for all $n\in\mathbb{N}$. Set $x_n:=x_0+t_nu_n\in M$. Since $u_n\to u$, there exists $N\in\mathbb{N}$ such that

   $$n\geq N\Rightarrow\|u_n-u\|\leq 1. \tag{1.29}$$

   We now prove that $x_n\to x_0$ as $n\to\infty$. Indeed, for $n\geq N$,

   $$\|x_n-x_0\|=\|t_nu_n\|=t_n\|u_n\|\leq t_n(\|u\|+1). \tag{1.30}$$

   Since $t_n\to 0^+$, (1.30) implies that $x_n\to x_0$ as $n\to\infty$, i.e., $x_0\in\overline{M}$.

2. We first prove that $T(M,x_0)$ is a cone. Let $u\in T(M,x_0)$ arbitrarily, we need to prove that $tu\in T(M,x_0)$ for all $t>0$. By (1.16), there exist a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n\to 0^+$ and a sequence $\{u_n\}_{n=1}^{\infty}\subset X$ such that $u_n\to u$ and $x_0+t_nu_n\in M$ for all $n\in\mathbb{N}$. Fix $t>0$ arbitrarily, if we set $v_n:=tu_n$ and $s_n=\frac{t_n}{t}$ for all $n\in\mathbb{N}$, then $s_n\to 0^+$, $v_n\to tu$ and $x_0+s_nv_n=x_0+t_nu_n\in M$ for all $n\in\mathbb{N}$, i.e., $tu\in T(M,x_0)$. Since $t>0$ and $u\in T(M,x_0)$ are chosen arbitrarily, this implies that $T(M,x_0)$ is a cone.

   To prove that $T(M,x_0)$ is closed, let $\{u_n\}_{m=0}^{\infty}\subset T(M,x_0)$ such that $u_m\to u$ as $m\to\infty$. We need to prove that $u\in T(M,x_0)$. To this end, by definition (1.16), for each $m\in\mathbb{N}$, there exist a sequence $t_{m,n}\to 0^+$ as $n\to\infty$ and a sequence $\{u_{m,n}\}_{n=0}^{\infty}\subset X$ such that $u_{m,n}\to u_m$ as $n\to\infty$

and $x_0 + t_{m,n}u_{m,n} \in M$ for all $n \in \mathbb{N}$, in addition, $\|u_{m,m} - u_m\| \leq \frac{1}{m}$ for all $m \in \mathbb{N}$[4]. We claim that

$$u_{m,m} \to u \text{ and } x_0 + t_{m,m}u_{m,m} \in M, \quad \forall m \in \mathbb{N}. \tag{1.33}$$

The latter is obvious since $x_0 + t_{m,n}u_{m,n} \in M$ for all $m, n \in \mathbb{N}$. We now prove the former in (1.33). With the help of triangle inequality for the norm of X,

$$\begin{aligned} \|u_{m,m} - u\| &\leq \|u_{m,m} - u_m\| + \|u_m - u\| && (1.34)\\ &\leq \frac{1}{m} + \|u_m - u\| \to 0 \text{ as } m \to \infty, && (1.35) \end{aligned}$$

i.e., $u \in T(M, x_0)$. Hence, $T(M, x_0)$ is a closed cone.

3. The convex conical hull of $M - x_0$ is given by (see, e.g., [1], Def. 4.19, p.94)

$$\operatorname{cone}(M - x_0) := \left\{ \sum_{i=1}^{k} \lambda_i x_i : x_i \in M - x_0, \lambda_i > 0, k \geq 1 \right\}. \tag{1.36}$$

Take $u \in T(M, x_0)$ arbitrarily, we need to prove that $u \in \overline{\operatorname{cone}(M - x_0)}$. By (1.16) again, there exist a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and a sequence $\{u_n\}_{n=1}^{\infty} \subset X$ such that $u_n \to u$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. The fact that $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$ gives us $t_n u_n \in M - x_0$ for all $n \in \mathbb{N}$. Choosing $k = 1, \lambda_1 = \frac{1}{t_n} > 0, x_1 = t_n u_n \in M - x_0$ in (1.36) gives $u_n \in \operatorname{cone}(M - x_0)$ for all $n \in \mathbb{N}$. Combining this with the fact that $u_n \to u$, we conclude that $u \in \overline{\operatorname{cone}(M - x_0)}$. Therefore,

$$T(M, x_0) \subset \overline{\operatorname{cone}(M - x_0)}. \tag{1.37}$$

4. First Proof. We now assume (until the end of the proof of this problem) that M is a convex set and $x_0 \in M$[5]. To prove $T(M, x_0) = \overline{\operatorname{cone}(M - x_0)}$, due to (1.37), it suffices to prove that $T(M, x_0) \supset \overline{\operatorname{cone}(M - x_0)}$. First, we need the following lemma (see, e.g., [2], Lemma 2.4.11, p.41).

Lemma 4.1. *Let M be a nonempty convex set and $x_0 \in M$. Then*

$$M - x_0 \subset T(M, x_0). \tag{1.38}$$

[4]This is possible, since for each $m \in \mathbb{N}$, there exists a sequence $\{u_{m,n}\}_{n=0}^{\infty} \subset X$ such that $u_{m,n} \to u_m$ as $n \to \infty$. By definition of limits, there exists $N \in \mathbb{N}$ such that

$$n \geq N \Rightarrow \|u_{m,n} - u_m\| \leq \frac{1}{m}. \tag{1.31}$$

Hence, we can drop all the terms $u_{m,1}, \ldots, u_{m,n-1}$ from the sequence. Re-indexing $\widehat{u}_{m,n} := u_{m,N+n-1}$ for all $n \in \mathbb{N}$, we have, in particular,

$$\|\widehat{u}_{m,m} - u_m\| = \left\|\widehat{u}_{m,N+m-1} - u_m\right\| \leq \frac{1}{m}. \tag{1.32}$$

We now ignore the old sequence $\{u_{m,n}\}_{n=0}^{\infty}$ and use the new sequence, by abuse notation, $\{u_{m,n}\}_{n=0}^{\infty}$ which is exactly $\{\widehat{u}_{m,n}\}_{n=0}^{\infty}$ just defined.

[5]The definition of tangent cone in [1] also requires this.

Proof of Lemma 4.1. Let $u \in M$. We need to show that $u - x_0 \in T(M, x_0)$. To this end, choose $\{t_n\}_{n=1}^{\infty} \subset [0,1]$ such that $t_n \to 0^+$, and put $u_n := u - x_0$ (hence $u_n \to u - x_0$ obviously) and put

$$x_n := x_0 + t_n (u - x_0) \tag{1.39}$$
$$= (1 - t_n) x_0 + t_n u \in M, \quad \forall n \in \mathbb{N}, \tag{1.40}$$

as M is convex. By (1.16), $u - x_0 \in T(M, x_0)$. ☐

Return to our proof, since we have proved that $T(M, x_0)$ is a closed cone, we only need to prove prove that $T(M, x_0) \supset \operatorname{cone}(M - x_0)$. Using the fact that the convex conical hull of an arbitrary nonempty set is the intersection of all closed convex cones that contain that sets, it suffices to prove that $T(M, x_0)$ is convex (and thus is a closed convex cone). Take $u, v \in T(M, x_0)$, we need to prove that $\lambda u + (1 - \lambda) v \in T(M, x_0)$ for all $\lambda \in [0,1]$. But since $T(M, x_0)$ is a cone, we deduce that $\lambda u \in T(M, x_0)$ and $(1 - \lambda) v \in T(M, x_0)$. Hence, it suffices to prove the following stronger statement[6]

$$u + v \in T(M, x_0), \quad \forall u, v \in T(M, x_0). \tag{1.41}$$

By (1.16), there exists sequences of positive reals $\{t_n\}_{n=1}^{\infty}, \{s_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and $s_n \to 0^+$ and sequences $\{u_n\}_{n=1}^{\infty}, \{v_n\}_{n=1}^{\infty}$ such that $u_n \to u, v_n \to v$ and

$$x_0 + t_n u_n \in M, \quad x_0 + s_n v_n \in M, \quad \forall n \in \mathbb{N}. \tag{1.42}$$

Since M is convex, it is deduced from (1.42) that

$$\alpha (x_0 + t_n u_n) + (1 - \alpha)(x_0 + s_n v_n) \in M, \quad \forall \alpha \in [0,1], n \in \mathbb{N}. \tag{1.43}$$

In particular, choosing $\alpha = \frac{s_n}{t_n + s_n}$ in (1.43) gives

$$x_0 + \frac{t_n s_n}{t_n + s_n} (u_n + v_n) \in M, \quad \forall n \in \mathbb{N}. \tag{1.44}$$

Hence, if we choose $w_n := u_n + v_n \to u + v$ and $r_n := \frac{t_n s_n}{t_n + s_n} \to 0^{+}$[7]. By (1.16), $u + v \in T(M, x_0)$. This completes our proof. ☐

Second Proof. We have the following result (see, e.g., [2], Proposition 2.4.8, p.40)

$$\operatorname{cone} S = \mathbb{R}_+ (\operatorname{conv} S) = \operatorname{conv}(\mathbb{R}_+ S), \tag{1.45}$$

for an arbitrary nonempty set S. Since M is convex, $M - x_0$ is also convex (as a Minkowski sum of convex sets), hence $\operatorname{conv}(M - x_0) = M - x_0$ (see [1], Corollary 4.12, p.91) and

$$\overline{\operatorname{cone}(M - x_0)} = \overline{\mathbb{R}_+ (\operatorname{conv}(M - x_0))} = \overline{\mathbb{R}_+ (M - x_0)}. \tag{1.46}$$

[6] *A cone K is convex if and only if $K + K \subset K$.* (see, e.g., [2], Proposition 2.4.2, p.38.)

[7] Indeed, $0 < r_n = t_n \underbrace{\frac{s_n}{t_n + s_n}}_{<1} < t_n \to 0^+$ as $n \to \infty$.

It suffices to prove $\overline{\mathbb{R}_+ (M - x_0)} \subset T(M, x_0)$. By Lemma 4.1, we have $M - x_0 \subset T(M, x_0)$. Since $T(M, x_0)$ is a closed cone, this yields $\overline{\mathbb{R}_+ (M - x_0)} \subset T(M, x_0)$. A direct consequence of this fact is that $T(M, x_0)$ is a closed convex cone.

5. *(Need correcting)* Suppose the set in the right-hand side is nonempty, i.e., there exists $v \in X$ such that

$$\forall t_n \to 0^+, \forall v_n \to v, x_0 + t_n v_n \in M, \quad \forall n \in \mathbb{N}. \tag{1.47}$$

If we take t_1 and v_1 arbitrarily, then $x_0 + t_1 v_1$ still belongs to M. Hence, $M = X$? Should (1.47) be corrected as "$\forall t_n \to 0^+, \forall v_n \to v, x_0 + t_n v_n \in M$ *for n large enough*"? This problem needs correcting.

We end our proof. □

Problem 5 (Formula for computing contingent cone of a system of constrained inequalities). *Suppose $g_i : \mathbb{R}^n \to \mathbb{R}$ are Fréchet differentiable functions for all $i = 1, \ldots, m$. The set M is defined by*

$$M = \{x \in \mathbb{R}^n | g_i(x) \le 0, \quad \forall i = 1, \ldots, m\}. \tag{1.48}$$

Take $x_0 \in M$, set the index set

$$I(x_0) = \{i \in \{1, \ldots, m\} | g_i(x_0) = 0\}. \tag{1.49}$$

Then, we have

1. *If $I(x_0) = \emptyset$ then $T(M, x_0) = \mathbb{R}^n$.*
2. *If $I(x_0) \ne \emptyset$ then*

$$T(M, x_0) \subset \{v \in \mathbb{R}^n | \nabla g_i(x_0)(v) \le 0, \quad \forall i \in I(x_0)\}. \tag{1.50}$$

3. *Moreover, if the following condition is satisfied*

$$\exists \bar{v} \in \mathbb{R}^n \text{ s.t. } \nabla g_i(x_0)(\bar{v}) < 0, \quad \forall i \in I(x_0), \tag{1.51}$$

then we have

$$T(M, x_0) = \{v \in \mathbb{R}^n | \nabla g_i(x_0)(v) \le 0, \quad \forall i \in I(x_0)\}, \tag{1.52}$$

where $\nabla g_i(x_0)(v)$ is Fréchet derivative of g_i at x_0 applying to vector v.

Solution.

1. We assume $I(x_0) = \emptyset$, i.e., $g_i(x_0) < 0$ for all $i = 1, \ldots, m$. For each $i \in \{1, \ldots, m\}$, since g_i is Fréchet differentiable and thus continuous, there exists $\delta_i > 0$ such that

$$x \in B_{\delta_i}(x_0) \Rightarrow g_i(x) < 0. \tag{1.53}$$

Choosing $\delta := \min\{\delta_i | i = 1, \ldots, m\} > 0$, we have

$$g_i(x) < 0, \quad \forall i = 1, \ldots, m, \quad \forall x \in B_\delta(x_0). \tag{1.54}$$

Taking $u \in \mathbb{R}^n$ arbitrarily, we prove that $u \in T(M, x_0)$. The case $u = \mathbf{0} \in \mathbb{R}^n$ is obvious (take $u_n := 0$ and $t_n \to 0^+$ arbitrarily). If $u \neq 0$, we choose $u_n := u$ and a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that

$$t_n < \frac{\delta}{\|u\|}, \;\; \forall n \in \mathbb{N} \text{ and } t_n \to 0^+, \tag{1.55}$$

for instance, we can choose a monotone decreasing sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_1 < \frac{\delta}{\|u\|}$. The choice (1.55) ensures that $x_0 + t_n u_n \in B_\delta(x_0)$ for all $n \in \mathbb{N}$. Combining this with (1.54) gives $g_i(x_0 + t_n u_n) < 0$ for all $i = 1, \ldots, m$, i.e., $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. Hence, by (1.16), $u \in T(M, x_0)$. Since u is chosen arbitrarily, we conclude that $T(M, x_0) = \mathbb{R}^n$.

2. Suppose that $I(x_0) \neq \emptyset$, we take $u \in T(M, x_0)$ and try to prove that

$$\langle \nabla g_i(x_0), u \rangle \leq 0, \;\; \forall i \in I(x_0). \tag{1.56}$$

By (1.16), there exist a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and a sequence $\{u_n\}_{n=1}^{\infty} \subset \mathbb{R}^n$ such that $u_n \to u$ as $n \to \infty$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$, i.e.,

$$g_i(x_0 + t_n u_n) \leq 0, \;\; \forall i = 1, \ldots, m, \;\; \forall n \in \mathbb{N}. \tag{1.57}$$

To prove (1.56), we have $g_i(x_0) = 0$ for all $i \in I(x_0)$. Combining this with (1.57) and the following first order multivariate Taylor's formula (see, e.g., [1], Theorem 1.23, p.15)

$$g_i(x_0 + t_n u_n) = g_i(x_0) + t_n \langle \nabla g_i(x_0 + \alpha_n t_n u_n), u_n \rangle, \tag{1.58}$$

for some $\alpha_n \in (0, 1)$, for all $i \in I(x_0)$ and for all $n \in \mathbb{N}$, we deduce that

$$\langle \nabla g_i(x_0 + \alpha_n t_n u_n), u_n \rangle \leq 0, \;\; \forall i \in I(x_0), \;\; \forall n \in \mathbb{N}. \tag{1.59}$$

Letting $n \to \infty$ in (1.59) gives (1.56) as desired. Therefore, (1.50) holds.

3. First of all, the existence of $\overline{v}$ satisfying (1.51) implies that the set in the right-hand side of (1.52) is nonempty (at least $\overline{v}$ belongs to that set). Since we have proved (1.50), it suffices to prove the reverse inclusion

$$T(M, x_0) \supset \{v \in \mathbb{R}^n | \nabla g_i(x_0)(v) \leq 0, \forall i \in I(x_0)\}. \tag{1.60}$$

Taking u belonging to the right-hand side of (1.60), i.e.,

$$\nabla g_i(x_0)(u) \leq 0, \;\; \forall i \in I(x_0), \tag{1.61}$$

we need to prove that $u \in T(M, x_0)$. We choose

$$u_n := \frac{1}{n}\overline{v} + \frac{n-1}{n} u \to u \text{ as } n \to \infty. \tag{1.62}$$

It suffices to prove that there exists a sequence of positive reals $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$, i.e.,

$$g_i(x_0 + t_n u_n) \leq 0, \;\; \forall i = 1, \ldots, m, \;\; \forall n \in \mathbb{N}. \tag{1.63}$$

We consider the following two cases depending on the index i.

Case $i \in I(x_0)$**.** For each $i \in I(x_0)$, $g_i(x_0) = 0$ and (1.58) then gives

$$g_i(x_0 + t_n u_n) = t_n \langle \nabla g_i(x_0 + \alpha_n t_n u_n), u_n \rangle \tag{1.64}$$
$$= \frac{t_n}{n} \langle \nabla g_i(x_0 + \alpha_n t_n u_n), \overline{v} \rangle \tag{1.65}$$
$$+ \frac{n-1}{n} t_n \langle \nabla g_i(x_0 + \alpha_n t_n u_n), u \rangle . \tag{1.66}$$

Combining the fact that g_i is Fréchet differentiable, (1.51) and (1.61) yields that there exists $\delta_i > 0$ such that

$$x \in B_{\delta_i}(x_0) \Rightarrow \langle \nabla g_i(x), \overline{v} \rangle < 0, \langle \nabla g_i(x), u \rangle \leq 0. \tag{1.67}$$

Take $\delta := \min\{\delta_i | i \in I(x_0)\}$, then

$$x \in B_\delta(x_0) \Rightarrow \langle \nabla g_i(x), \overline{v} \rangle < 0, \langle \nabla g_i(x), u \rangle \leq 0, \quad \forall i \in I(x_0). \tag{1.68}$$

We then take $\{t_n\}_{n=1}^\infty$ such that $t_n \to 0^+$ and

$$t_n < \frac{\delta}{\frac{\|\overline{v}\|}{n} + \frac{n-1}{n}\|u\|}, \quad \forall n \in \mathbb{N}. \tag{1.69}$$

Then

$$\|\alpha_n t_n u_n\| \leq t_n \|u_n\| = t_n \left\| \frac{1}{n}\overline{v} + \frac{n-1}{n} u \right\| \tag{1.70}$$
$$\leq t_n \left(\frac{\|\overline{v}\|}{n} + \frac{n-1}{n}\|u\| \right) < \delta, \quad \forall n \in \mathbb{N}. \tag{1.71}$$

i.e., $x_0 + \alpha_n t_n u_n \in B_\delta(x_0)$ for all $n \in \mathbb{N}$. Combining this with (1.64)-(1.66) and (1.68) yields

$$g_i(x_0 + t_n u_n) \leq 0, \quad \forall i \in I(x_0), \quad \forall n \in \mathbb{N}. \tag{1.72}$$

Case $i \notin I(x_0)$**.** For each $i \notin I(x_0)$, we have $g_i(x_0) \neq 0$. In addition, $x_0 \in M$, i.e., $g_i(x_0) \leq 0$, then we must have $g_i(x_0) < 0$. By (1.59) again, we have

$$g_i(x_0 + t_n u_n) = g_i(x_0) + \frac{t_n}{n} \langle \nabla g_i(x_0 + \alpha_n t_n u_n), \overline{v} \rangle \tag{1.73}$$
$$+ \frac{n-1}{n} t_n \langle \nabla g_i(x_0 + \alpha_n t_n u_n), u \rangle . \tag{1.74}$$

In order that $g_i(x_0 + t_n u_n) \leq 0$ for all $n \in \mathbb{N}$, we choose $\{t_n\}_{n=1}^\infty$ such that

$$- \frac{t_n}{n} \langle \nabla g_i(x_0 + \alpha_n t_n u_n), \overline{v} \rangle - g_i(x_0) \tag{1.75}$$
$$\geq \frac{n-1}{n} t_n \left| \langle \nabla g_i(x_0 + \alpha_n t_n u_n), u \rangle \right|, \quad \forall n \in \mathbb{N}. \tag{1.76}$$

To (1.75)-(1.76) holds, we can make a stronger assumption on t_n's, that is

$$t_n \|u_n\| \leq 1, \ \ \forall n \in \mathbb{N} \text{ and} \tag{1.77}$$

$$-g_i(x_0) \geq \frac{n-1}{n} t_n \|u\| \sup_{x \in B_1(x_0)} \|\nabla g_i(x)\|, \ \ \forall n \in \mathbb{N}. \tag{1.78}$$

i.e.,

$$t_n \leq \min \left\{ \frac{1}{\|u_n\|}, -\frac{n}{n-1} \cdot \frac{g_i(x_0)}{\|u\| \sup\limits_{x \in B_1(x_0)} \|\nabla g_i(x)\|} \right\}, \ \ \forall n \in \mathbb{N}. \tag{1.79}$$

If we choose t_n's satisfying (1.79) then $g_i(x_0 + t_n u_n) \leq 0$ for all $n \in \mathbb{N}$. Hence, if we choose t_n's such that

$$t_n \leq \min \left\{ \frac{1}{\|u_n\|}, -\frac{n}{n-1} \cdot \frac{g_i(x_0)}{\|u\| \sup\limits_{x \in B_1(x_0)} \|\nabla g_i(x)\|} : i \in \{1, \ldots, m\} \setminus I\{x_0\} \right\}, \tag{1.80}$$

then

$$g_i(x_0 + t_n u_n) \leq 0, \ \ \forall i \in \{1, \ldots, m\} \setminus I\{x_0\}, \ \ \forall n \in \mathbb{N}. \tag{1.81}$$

Combining two discussed cases, we now choose t_n's such that $t_n \to 0^+$ and satisfy both (1.69) and (1.80), i.e.,

$$t_n \leq \min \left\{ \begin{array}{l} \frac{1}{\|u_n\|}, \dfrac{\delta}{\frac{\|\overline{v}\|}{n} + \frac{n-1}{n}\|u\|}, \\ -\frac{n}{n-1} \cdot \dfrac{g_i(x_0)}{\|u\| \sup\limits_{x \in B_1(x_0)} \|\nabla g_i(x)\|} : i \in \{1, \ldots, m\} \setminus I\{x_0\} \end{array} \right\}. \tag{1.82}$$

Then (1.72) and (1.81) gives

$$g_i(x_0 + t_n u_n) \leq 0, \ \ \forall i = 1, \ldots, m, \ \ \forall n \in \mathbb{N}, \tag{1.83}$$

i.e., $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. By definition of tangent cone, we deduce that $u \in T(M, x_0)$.

This completes our proof. □

Remark 5.1. The assumption of existence of $\overline{v}$ (1.51) can be removed. Indeed, suppose for the contrary that there does not exist $\overline{v}$ satisfying (1.51), we then choose

$$u_n := \frac{1}{n} u_0 + \frac{n-1}{n} u \to u \text{ as } n \to \infty, \tag{1.84}$$

for arbitrarily chosen point u_0. Since we have assumed that (1.51) fails, then there exists an index $i_0 \in I(x_0)$ depending on u_0 such that

$$\nabla g_{i_0}(x_0)(u_0) \geq 0. \tag{1.85}$$

We now let $n = 1$, with replacing $\overline{v}$ by u_0, in (1.64)-(1.66) gives

$$g_{i_0}(x_0 + t_1 u_1) = t_1 \langle \nabla g_{i_0}(x_0 + \alpha_1 t_1 u_1), u_0 \rangle. \tag{1.86}$$

Since $\nabla g_{i_0}(x_0)(u_0) \geq 0$, we can not make any assumption on t_1 in order that $g_{i_0}(x_0 + t_1 u_1) \leq 0$ (it is possible that $t_1 \langle \nabla g_{i_0}(x_0 + \alpha_1 t_1 u_1), u_0 \rangle \geq 0$ for all $t_1 \in \mathbb{R}$, and our entire argument collapses). This is the reason why the assumption (1.51) cannot be excluded.

Problem 6. *Use the results of Problem 5 to compute contingent cones in Problem 3.2.*

Solution. Applying the result in Problem 5.3 to

$$M = \left\{(x_1, x_2) \in \mathbb{R}^2 | g_i(x) \leq 0, \quad i = 1, 2\right\}, \tag{1.87}$$

where $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ are Fréchet differentiable functions defined by

$$g_1(x) = 2 - x_1 - x_2, \tag{1.88}$$

$$g_2(x) = x_2 - x_1^3, \tag{1.89}$$

for all $x = (x_1, x_2) \in \mathbb{R}^2$. With $x_0 = (1, 1) \in M$, the index set $I(x_0)$ is given by

$$I(x_0) = \{i \in \{1, 2\} | g_i(1, 1) = 0\} = \{1, 2\}. \tag{1.90}$$

Next, we have

$$\nabla g_1(x) = (-1, -1)^T, \nabla g_2(x) = \left(-3x_1^2, 1\right)^T, \quad \forall x = (x_1, x_2) \in \mathbb{R}^2. \tag{1.91}$$

Take $\overline{v} = (1, 0)^T \in \mathbb{R}^2$, we have

$$\nabla g_1(x_0)(\overline{v}) = \left\langle (-1, -1)^T, (1, 0)^T \right\rangle = -1, \tag{1.92}$$

$$\nabla g_2(x_0)(\overline{v}) = \left\langle (-3, 1)^T, (1, 0)^T \right\rangle = -3, \tag{1.93}$$

i.e., (1.51) holds. Now we apply the result in Problem 5.3 to our setting to to obtain

$$\begin{aligned} T(M, x_0) &= \left\{v \in \mathbb{R}^2 | \nabla g_i(x_0)(v) \leq 0, \quad i = 1, 2\right\} && (1.94) \\ &= \left\{(v_1, v_2)^T \in \mathbb{R}^2 | - v_1 - v_2 \leq 0, -3v_1 + v_2 \leq 0\right\} && (1.95) \\ &= \left\{(v_1, v_2)^T \in \mathbb{R}^2 | v_1 + v_2 \geq 0, v_2 \leq 3v_1\right\}, && (1.96) \end{aligned}$$

which is exactly (1.28). □

Problem 7 (Geometric form of first order optimality condition for unconstrained problem). *Consider the following problem* (P)

$$(P) \qquad \min f(x) \text{ s.t. } x \in \Omega. \tag{1.97}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ *and* $\Omega \subseteq \mathbb{R}^n$. *Prove the following geometric form of the first order optimality condition*

$$\text{If } x_0 \text{ is a local minimum of } (P) \text{ then } \forall u \in T(\Omega, x_0) : \langle \nabla f(x_0), u \rangle \geq 0. \tag{1.98}$$

Solution. Assume that x_0 is a local minimizer of f in Ω, (see, e.g., [1], Def. 2.1, p.32) there exists $r > 0$ such that $B_r(x_0) \subset \Omega$ and

$$x \in B_r(x_0) \Rightarrow f(x_0) \leq f(x). \tag{1.99}$$

Take $u \in T(\Omega, x_0)$ arbitrarily, by (1.16), there exist a sequence of real positive $\{t_n\}_{n=1}^{\infty}$ such that $t_n \to 0^+$ and a sequence $\{u_n\}_{n=1}^{\infty} \subset \mathbb{R}^n$ such that $u_n \to u$ as $n \to \infty$ and $x_0 + t_n u_n \in \Omega$ for all $n \in \mathbb{N}$. Then there exists $N \in \mathbb{N}$ such that

$$x_0 + t_n u_n \in B_r(x_0), \quad \forall n \geq N. \tag{1.100}$$

Combining (1.99) with (1.100) yields

$$f(x_0) \leq f(x_0 + t_n u_n), \quad \forall n \geq N. \tag{1.101}$$

By the first order multivariate Taylor formula, we have

$$f(x_0 + t_n u_n) = f(x_0) + t_n \langle \nabla f(x_0 + \alpha_n t_n u_n), u_n \rangle, \quad \forall n \in \mathbb{N}, \tag{1.102}$$

for some $\alpha_n \in (0,1)$. Combining (1.101) with (1.102) yields

$$\langle \nabla f(x_0 + \alpha_n t_n u_n), u_n \rangle \geq 0, \quad \forall n \geq N. \tag{1.103}$$

Letting $n \to \infty$ in (1.103), we obtain

$$\langle \nabla f(x_0), u \rangle \geq 0. \tag{1.104}$$

Since u is taken from $T(\Omega, x_0)$ arbitrarily, (1.98) holds and we complete our proof. □

Problem 8. *Consider the following problem*

$$(P) \qquad \min x^2 + y \text{ s.t. } (x,y) \in \Omega := \left\{(x,y) \in \mathbb{R}^2 | x^2 + y^3 = 0\right\}. \tag{1.105}$$

1. *Compute the tangent cone of* Ω *at* $x_0 = (0,0)$.
2. *Applying the result of Problem 7, prove that* $x_0 = (0,0)$ *is not a local minimizer of* (P).

Solution.

1. We have computed the tangent cone of Ω at $x_0 = (0,0)$ in Problem 3.1 with the interchange of x_1 and x_2

$$T(\Omega, x_0) = \left\{(0,y) \in \mathbb{R}^2 | y \leq 0\right\}. \tag{1.106}$$

2. Suppose for the contrary[8] that $x_0 = (0,0)$ is a local minimizer of (P). Set

$$f(x,y) = x^2 + y, \quad \forall (x,y) \in \mathbb{R}^2. \tag{1.108}$$

which has $\nabla f(x,y) = (2x, 1)$ for all $(x,y) \in \mathbb{R}^2$. Then (1.98) gives

$$\forall u \in T(\Omega, x_0) : \langle \nabla f(x_0), u \rangle \geq 0, \tag{1.109}$$

equivalently,

$$\forall x \leq 0 : 0 \leq \langle \nabla f(x_0), (0,x) \rangle = \langle (0,1), (0,x) \rangle = x, \tag{1.110}$$

which is absurd. Therefore, x_0 is not a local minimizer of (P).

This completes our proof. □

Problem 9. *Consider the following problem*

$$(P) \qquad \min x + 2y \text{ s.t. } x^2 + y^2 \leq 1, x + y \leq 1. \tag{1.111}$$

Applying the result in Problem 7, check whether $x_0 = (0,1)$ *is a local minimizer of* (P).

Solution. Setting

$$f(x,y) = x + 2y, \forall (x,y) \in \mathbb{R}^2, \tag{1.112}$$
$$M = \left\{ (x,y) \in \mathbb{R}^2 | g_i(x,y) \leq 0, \quad i = 1,2 \right\}, \tag{1.113}$$

where $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ are Fréchet differentiable functions defined by

$$g_1(x,y) = x^2 + y^2 - 1, \tag{1.114}$$
$$g_2(x,y) = x + y - 1, \tag{1.115}$$

for all $(x,y) \in \mathbb{R}^2$, we first compute the tangent cone of M at $x_0 = (0,1) \in M$ by using the result of Problem 5. The index set $I(x_0)$ is given by

$$I(x_0) = \{ i \in \{1,2\} | g_i(0,1) = 0 \} = \{1,2\}. \tag{1.116}$$

Next, we have

$$\nabla g_1(x,y) = (2x, 2y), \nabla g_2(x,y) = (1,1), \quad \forall (x,y) \in \mathbb{R}^2. \tag{1.117}$$

Take $\overline{v} = (0,-1)^T \in \mathbb{R}^2$, we have

$$\nabla g_1(x_0)(\overline{v}) = \langle (0,2), (0,-1) \rangle = -2, \tag{1.118}$$
$$\nabla g_2(x_0)(\overline{v}) = \langle (1,1), (0,-1) \rangle = -1, \tag{1.119}$$

i.e., (1.51) holds. Now we apply the result in Problem 5.3 to our setting to obtain

$$T(M, x_0) = \left\{ v \in \mathbb{R}^2 | \nabla g_i(x_0)(v) \leq 0, \quad i = 1,2 \right\} \tag{1.120}$$

[8] A shorter proof is as follows. Take $u = (0,-1) \in T(\Omega, x_0)$, then

$$\langle \nabla f(x_0), u \rangle = \langle (0,1), (0,-1) \rangle = -1 < 0. \tag{1.107}$$

then (1.98) implies that x_0 is not a local minimizer of (P). □

$$= \left\{ (v_1, v_2) \in \mathbb{R}^2 | v_2 \le 0, v_1 + v_2 \le 0 \right\}. \tag{1.121}$$

Take $u = (-1, 0) \in T(M, x_0)$, we have

$$\langle \nabla f(x_0), u \rangle = \langle (1, 2), (-1, 0) \rangle = -1 < 0. \tag{1.122}$$

Hence, by (1.98), we deduce that x_0 is not a local minimizer of (P). □

Problem 10. *Consider the following problem*

$$(P) \qquad -xy \text{ s.t. } x + y = 8, \ \ x \ge 0, y \ge 0. \tag{1.123}$$

Applying the result in Problem 7, check whether $x_0 = (4, 4)$ is a local minimizer of (P).

Solution. We apply the well-known Cauchy-Schwarz inequality

$$xy \le \left(\frac{x+y}{2} \right)^2 = \frac{8^2}{4} = 16. \tag{1.124}$$

Hence, $-xy \ge -16$ for all x, y such that $x + y = 8, x \ge 0, y \ge 0$. The equality happens if and only if $x = y = 4$. Thus, $x_0 = (4, 4)$ is a local minimizer. □

Remark 10.1. The result in Problem 7 is only an necessary but not sufficient condition. Hence, we can only use it to disprove the statement "x_0 is a local minimizer of (P)" (i.e., prove that x_0 is not a local minimizer as we did in Problem 8, 9) but can not use it to check whether x_0 is a local minimizer of (P).

Problem 11. *Consider the following problem*

$$(P) \qquad \min x^2 + y^2 \text{ s.t. } x^2 - (y-1)^3 = 0. \tag{1.125}$$

1. *Use algebraic or geometrical methods to solve (P).*
2. *Examine the necessary condition in Problem 7.*

Solution.

1. We deduce from (1.125) that

$$x^2 = (y-1)^3 \ge 0. \tag{1.126}$$

The last inequality implies $y \ge 1$. Substituting x^2 giving by (1.126) into our problem yields

$$x^2 + y^2 = f(y) := (y-1)^3 + y^2, \ \ \forall y \ge 1. \tag{1.127}$$

The first order derivative of f is

$$f'(y) = 3(y-1)^2 + 2y > 0, \ \ \forall y \ge 1. \tag{1.128}$$

Hence,

$$\min_{x^2-(y-1)^3=0} \left(x^2 + y^2 \right) = \min_{y \ge 1} f(y) = f(1) = 1, \tag{1.129}$$

which holds if and only if $x = 0, y = 1$.

2. Setting $x_0 = (0, 1)$ and

$$f(x, y) = x^2 + y^2, \ \forall (x, y) \in \mathbb{R}^2, \tag{1.130}$$

$$M = \left\{ (x, y) \in \mathbb{R}^2 | x^2 - (y - 1)^3 = 0 \right\}, \tag{1.131}$$

we have $\nabla f(x, y) = (2x, 2y)$ for all $(x, y) \in \mathbb{R}^2$. Now we compute the tangent cone $T(M, x_0)$ as in Problem 3. Notice that $x_0 \in M$, we claim that

$$T(M, x_0) = \widehat{T}(M, x_0) := \left\{ (0, y) \in \mathbb{R}^2 | y \geq 0 \right\}. \tag{1.132}$$

To this end, we prove the following inclusions.

(a) *Prove* $T(M, x_0) \subset \widehat{T}(M, x_0)$. Taking $u = (x, y) \in T(M, x_0)$, by (1.16), there exist a sequence of positive reals $\{t_n\}_{n=1}^\infty$ such that $t_n \to 0^+$ and a sequence $\{u_n\}_{n=1}^\infty \subset \mathbb{R}^2$ such that $u_n \to u$ as $n \to \infty$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. Set $u_n := (x_n, y_n)$, the fact $u_n \to u$ implies that $x_n \to x$ and $y_n \to y$, and the fact $x_0 + t_n u_n$ for all $n \in \mathbb{N}$ gives

$$t_n^2 x_n^2 = t_n^3 y_n^3, \ \forall n \in \mathbb{N}. \tag{1.133}$$

Since $t_n > 0$ for all $n \in \mathbb{N}$, (1.133) then implies

$$x_n^2 = t_n y_n^3, \ \forall n \in \mathbb{N}. \tag{1.134}$$

We see at a glance from (1.134) that $y_n \geq 0$ for all $n \in \mathbb{N}$. Hence, $y \geq 0$ (since $y_n \to y$ as $n \to \infty$). Now let $n \to \infty$ in (1.134) and use the given limits $x_n \to x, y_n \to y$ and $t_n \to 0^+$, we obtain $x = 0$. Hence, $u \in \widehat{T}(M, x_0)$ and our first inclusion is proved.

(b) *Prove* $\widehat{T}(M, x_0) \subset T(M, x_0)$. Taking $(0, y)$ satisfying $y \geq 0$, we claim that $u \in T(M, x_0)$. To this end, we choose $x_n = \frac{1}{n^2}, y_n = y + \frac{1}{n} > 0$. This choice ensures that $u_n := (x_n, y_n) \to u := (0, y)$ as $n \to \infty$. It then suffices to prove that there exists a sequence of positive reals $\{t_n\}_{n=1}^\infty$ such that $t_n \to 0^+$ and $x_0 + t_n u_n \in M$ for all $n \in \mathbb{N}$. The latter gives, using (1.134) again,

$$\frac{1}{n^4} = t_n \left(y + \frac{1}{n} \right)^3, \ \forall n \in \mathbb{N}, \tag{1.135}$$

i.e.,

$$t_n = \frac{1}{n^4 \left(y + \frac{1}{n} \right)^3}, \ \forall n \in \mathbb{N}. \tag{1.136}$$

It is easy to check that $t_n > 0$ (since $y \geq 0$) and $t_n \to 0^+$ as $n \to \infty$. Hence, $u \in T(M, x_0)$, the second inclusion is also proved.

Combining these cases, we conclude that (1.132) holds, i.e.,

$$T(M, x_0) = \left\{ (0, y) \in \mathbb{R}^2 | y \geq 0 \right\}. \tag{1.137}$$

We now prove the necessary condition stated in Problem 7 for our setting, i.e.,

$$\forall u \in T(M, x_0) : \langle \nabla f(x_0), u \rangle \geq 0, \tag{1.138}$$

equivalently,

$$\forall y \geq 0 : \langle (0,2), (0,y) \rangle = 2y \geq 0, \tag{1.139}$$

which is obvious. Hence, the necessary condition in Problem 7 holds in our setting.

This completes our proof. □

Problem 12. *Let $X = \mathbb{R}^n$ (a finite-dimensional space), consider the norm function*

$$f(x) = \|x\|. \tag{1.140}$$

1. *Prove that $\nabla f(a) = \|a\|^{-1} a$ for all $a \neq 0$.*
2. *Prove that f is not Fréchet differentiable at $x = \mathbf{0}$.*

Solution.

1. For $x \neq 0$, write $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we have

$$f(x) = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}}, \tag{1.141}$$

$$\frac{\partial f}{\partial x_i}(x) = \frac{x_i}{\left(\sum\limits_{i=1}^n x_i^2 \right)^{\frac{1}{2}}} \tag{1.142}$$

$$= \frac{x_i}{\|x\|}, \quad \forall i = 1, \ldots, n. \tag{1.143}$$

Hence,

$$\nabla f(a) = \frac{a}{\|a\|}, \quad \forall a \neq 0. \tag{1.144}$$

2. Suppose for the contrary that f is Fréchet differentiable at $x = \mathbf{0}$, by definition, there exists a linear function $l : \mathbb{R}^n \to \mathbb{R}$, $l(x) = \langle l, x \rangle$, such that

$$\lim_{\|h\| \to 0} \frac{f(h) - f(0) - \langle l, h \rangle}{\|h\|} = 0. \tag{1.145}$$

Write $h = (h_1, \ldots, h_n), \langle l, h \rangle = \sum\limits_{i=1}^n l_i h_i$, (1.145) becomes

$$\lim_{\|h\| \to 0} \frac{1}{\|h\|} \sum_{i=1}^n l_i h_i = 1. \tag{1.146}$$

In particular, for an arbitrary index i, if we choose $h_i = \frac{1}{n}, h_j = 0$ for all $j \neq i$, (1.146) gives $l_i = 1$. Otherwise, if we choose $h_i = -\frac{1}{n}, h_j = 0$ for all $j \neq i$, (1.146) gives $l_i = -1$, which is absurd. Hence, f is not Fréchet differentiable at $x = \mathbf{0}$. This completes our proof. □

References

[1] O. Güler, *Foundations of Optimization*, Graduate Texts in Mathematics 258, Springer.

[2] Tim Hoheisel, *Convex Analysis*, University of Würzburg, Germany, Lecture Notes, Summer term 2016.