Coursework (3) for *Introductory Lectures on Optimization*

Your name
Your ID

Nov. 17, 2022

Excercise 1. Prove the following resuls. Define

$$\psi_Q(\boldsymbol{g}) = \sup\{\langle \boldsymbol{g},\ \boldsymbol{x}\rangle \mid \boldsymbol{x} \in Q\}.$$

Let Q_1 and Q_2 be two closed convex sets.

1. If for any $\boldsymbol{g} \in \mathrm{dom}\ \psi_{Q_2}$ we have $\psi_{Q_1}(\boldsymbol{g}) \leq \psi_{Q_2}(\boldsymbol{g})$, then $Q_1 \subseteq Q_2$.
2. Let $\mathrm{dom}\ \psi_{Q_1} = \mathrm{dom}\ \psi_{Q_2}$ and for any $\boldsymbol{g} \in \mathrm{dom}\ \psi_{Q_1}$, we have $\psi_{Q_1}(\boldsymbol{g}) = \psi_{Q_2}(\boldsymbol{g})$. Then $Q_1 \equiv Q_2$.

Proof of Excercise 1:

1. Assume that there exists $\boldsymbol{x}_0 \in Q_1$ and $\boldsymbol{x}_0 \notin Q_2$. Since Q_2 is a closed convex set, $\boldsymbol{x}_0$ is strongly separable from Q_2, which means

   $$\langle \boldsymbol{g},\ \boldsymbol{x}\rangle < \gamma < \langle \boldsymbol{g},\ \boldsymbol{x}_0\rangle, \quad \forall \boldsymbol{x} \in Q_2,\ \text{and } \boldsymbol{x} \in Q_1.$$

   Obviously it is contradict with $\psi_{Q_1}(\boldsymbol{g}) \leq \psi_{Q_2}(\boldsymbol{g})$. Thus, in this case we have $Q_1 \subseteq Q_2$.

2. Based on the first statement:

   (a) $\forall g \in \mathrm{dom}\ \psi_{Q_1},\ \psi_{Q_1}(\boldsymbol{g}) = \psi_{Q_2}(\boldsymbol{g}) \quad \Longrightarrow \quad Q_1 \subseteq Q_2$;

   (b) $\forall g \in \mathrm{dom}\ \psi_{Q_2},\ \psi_{Q_2}(\boldsymbol{g}) = \psi_{Q_1}(\boldsymbol{g}) \quad \Longrightarrow \quad Q_2 \subseteq Q_1$.

   Above all, we can get $Q_1 \equiv Q_2$.

□

Excercise 2. Prove the following result. Let f be a closed convex function. For any $\boldsymbol{x}_0 \in \mathrm{int}(\mathrm{dom}\ f)$ and $\boldsymbol{p} \in \mathbb{R}^n$ we have

$$f'(\boldsymbol{x}_0; \boldsymbol{p}) = \max\{\langle \boldsymbol{g},\ \boldsymbol{p}\rangle \mid \boldsymbol{g} \in \partial f(\boldsymbol{x}_0)\}.$$

Proof of Excercise 2: According to related definitions:

$$f'(\boldsymbol{x}_0; \boldsymbol{p}) = \lim_{\alpha \to 0} \frac{1}{\alpha}[f(\boldsymbol{x}_0 + \alpha \boldsymbol{p}) - f(\boldsymbol{x}_0)].$$

and

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}_0) + \langle \boldsymbol{g},\ \boldsymbol{y} - \boldsymbol{x}_0\rangle, \quad \forall \boldsymbol{g} \in \partial f(\boldsymbol{x}_0).$$

Obviously, we have

$$\begin{aligned} f'(\boldsymbol{x}_0, \boldsymbol{p}) &= \lim_{\alpha \to 0} \frac{1}{\alpha}[f(\boldsymbol{x}_0 + \alpha \boldsymbol{p}) - f(\boldsymbol{x}_0)] \\ &\geq \lim_{\alpha \to 0} \frac{1}{\alpha}\langle \boldsymbol{g},\ \alpha \boldsymbol{p}\rangle = \langle \boldsymbol{g},\ \boldsymbol{p}\rangle. \end{aligned}$$

Here $\boldsymbol{g}$ is from $\partial f(\boldsymbol{x}_0)$. Therefore, the subdifferential of the function $f'(\boldsymbol{x};\boldsymbol{p})$ at $\boldsymbol{p}=0$ is not empty and $\partial f(\boldsymbol{x}_0) \subseteq \partial_2 f'(\boldsymbol{x}_0;0)$.

Since $f'(\boldsymbol{x};\boldsymbol{p})$ is convex in $\boldsymbol{p}$, we have

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}_0) + f'(\boldsymbol{x}_0;\boldsymbol{y}-\boldsymbol{x}_0) \geq f(\boldsymbol{x}_0) + \langle \boldsymbol{g},\ \boldsymbol{y}-\boldsymbol{x}_0\rangle.$$

where $\boldsymbol{g} \in \partial_2 f'(\boldsymbol{x}_0;0) \subseteq \partial f(\boldsymbol{x}_0)$ and we can get $\partial f(\boldsymbol{x}_0) = \partial_2 f(\boldsymbol{x};0)$.

Consider $\boldsymbol{g} \in \partial_2 f'(\boldsymbol{x}_0;0)$. Thus for $\tau > 0$

$$\tau f'(\boldsymbol{x}_0;\boldsymbol{v}) = f'(\boldsymbol{x};\tau\boldsymbol{v}) \geq f'(\boldsymbol{x}_0;\boldsymbol{p}) + \langle \boldsymbol{g},\ \tau\boldsymbol{v}-\boldsymbol{p}\rangle.$$

Considering $\tau \to \infty$ we get $f'(\boldsymbol{x}_0;\boldsymbol{p}) - \langle \boldsymbol{g},\ \boldsymbol{p}\rangle \leq 0$. Thus we conclude that $\langle \boldsymbol{g},\ \boldsymbol{p}\rangle = f'(\boldsymbol{x};\boldsymbol{p})$. $\square$

Excercise 3. Let f be closed and convex. Assume that it is differentiable on its domain. Then $\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$ for any $\boldsymbol{x} \in \operatorname{int}(\operatorname{dom} f)$.

Proof of Excercise 3: For any direction p, we have

$$f'(\boldsymbol{x};\boldsymbol{p}) = \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{p}\rangle.$$

Since $f(\boldsymbol{x}+\boldsymbol{p}) \geq f(\boldsymbol{x}) + f'(\boldsymbol{x};\boldsymbol{p}) \geq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{p}\rangle$, we can get $\nabla f(\boldsymbol{x}) \in \partial f(\boldsymbol{x})$. In the meanwhile

$$f'(\boldsymbol{x};\boldsymbol{p}) = \max\{\langle \boldsymbol{g},\ \boldsymbol{p}\rangle \mid \boldsymbol{g} \in \partial f(\boldsymbol{x}_0)\} = \langle \nabla f(\boldsymbol{x}),\ \boldsymbol{p}\rangle$$

Similarly, according to the statement in Excercise. 1, we can get $\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$. $\square$

Excercise 4. Let Δ be a set and $f(\boldsymbol{x}) = \sup\{\phi(\boldsymbol{y},\boldsymbol{x}) \mid \boldsymbol{y} \in \Delta\}$. Suppose that for any fixed $\boldsymbol{y} \in \Delta$ the function $\phi(\boldsymbol{y},\boldsymbol{x})$ is closed and convex in $\boldsymbol{x}$. Then $f(\boldsymbol{x})$ is closed convex.

Moreover, for any $\boldsymbol{x}$ from

$$\operatorname{dom} f = \{\boldsymbol{x} \in \mathbb{R}^n \mid \exists\gamma : \phi(\boldsymbol{y},\boldsymbol{x}) \leq \gamma,\ \ \forall \boldsymbol{y} \in \Delta\}$$

we have

$$\partial f(\boldsymbol{x}) \supseteq \operatorname{Conv}\{\partial\phi_{\boldsymbol{x}}(\boldsymbol{y},\boldsymbol{x}) \mid \boldsymbol{y} \in I(\boldsymbol{x})\},$$

where $I(\boldsymbol{x}) = \{\boldsymbol{y} \mid \phi(\boldsymbol{y},\boldsymbol{x}) = f(\boldsymbol{x})\}$.

Proof of Excercise 4: Here we define

$$\hat{Q} = \left\{\boldsymbol{x} \in Q \mid \sup_{\boldsymbol{y}\in\Delta} \phi(\boldsymbol{x},\boldsymbol{y}) < +\infty\right\}.$$

According to last equation, it is without any doubt that $f(\boldsymbol{x}) < +\infty\ \forall \boldsymbol{x} \in \hat{Q}$ and we can conclude that $Q \in \operatorname{dom} f$. In addition, it is obvious that $\boldsymbol{x}, t \in \operatorname{epi}_Q(f)$ if and only if

$$\boldsymbol{x} \in Q, \quad t \geq \phi(\boldsymbol{x},\boldsymbol{y}), \quad \forall \boldsymbol{y} \in \Delta.$$

This means that

$$\operatorname{epi}_Q(f) = \bigcap_{\boldsymbol{y}\in\Delta} \operatorname{epi}_Q(\phi(\cdot,\boldsymbol{y})).$$

Since each set $\operatorname{epi}_Q(\phi(\cdot,\boldsymbol{y}))$ is closed and convex, $\operatorname{epi}_Q(f)$ is also closed and convex. Thus, f is closed and convex on $\hat{Q}$.

In the meanwhile, for all $\boldsymbol{x} \in \hat{Q}$, $\boldsymbol{y}_0 \in I(\boldsymbol{x}_0)$, and $\boldsymbol{g}_0 \in \partial_{Q,\boldsymbol{x}}\phi(\boldsymbol{x}_0,\boldsymbol{y}_0)$, we have

$$f(\boldsymbol{x}) \geq \phi(\boldsymbol{x},\boldsymbol{y}_0) \geq \phi(\boldsymbol{x}_0,\boldsymbol{y}_0) + \langle \boldsymbol{g}_0,\ \boldsymbol{x}-\boldsymbol{x}_0\rangle = f(\boldsymbol{x}_0) + \langle \boldsymbol{g}_0,\ \boldsymbol{x}-\boldsymbol{x}_0\rangle$$

Therefore, we can prove the second statement. $\square$

Excercise 5. Caculate the subdifferentials of the following functions.

1. $f(\boldsymbol{x}) = |\boldsymbol{x}|, \boldsymbol{x} \in \mathbb{R}^1$.
2. $f(\boldsymbol{x}) = \sum_{i=1}^m |\langle \boldsymbol{a}_i,\ \boldsymbol{x}\rangle - \boldsymbol{b}_i|$.
3. $f(\boldsymbol{x}) = \max_{1\le i\le n} \boldsymbol{x}^{(i)}$.
4. $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$.
5. $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_1 = \sum_{i=1}^n |\boldsymbol{x}^{(i)}|$.

Solution of Excercise 4:

1. $f(\boldsymbol{x}) = |\boldsymbol{x}| = \max\{-\boldsymbol{x}, \boldsymbol{x}\} \quad \Rightarrow \quad \partial f(\boldsymbol{x}) = [-1, 1]$.
2. Here we define
$$\begin{aligned} I_+(\boldsymbol{x}) &= \{i \mid \langle \boldsymbol{a}_i,\ \boldsymbol{x}_i\rangle - \boldsymbol{b}_i > 0\}, \\ I_-(\boldsymbol{x}) &= \{i \mid \langle \boldsymbol{a}_i,\ \boldsymbol{x}_i\rangle - \boldsymbol{b}_i < 0\}, \\ I_0(\boldsymbol{x}) &= \{i \mid \langle \boldsymbol{a}_i,\ \boldsymbol{x}_i\rangle - \boldsymbol{b}_i = 0\}. \end{aligned}$$
Then we have
$$\partial f(\boldsymbol{x}) = \sum_{i\in I_+(\boldsymbol{x})} \boldsymbol{a}_i - \sum_{i\in I_-(\boldsymbol{x})} \boldsymbol{a}_i + \sum_{i\in I_0(\boldsymbol{x})} [-\boldsymbol{a}_i, \boldsymbol{a}_i].$$
3. Here we define $I(\boldsymbol{x}) = \{i \mid \boldsymbol{x}^{(i)} = f(\boldsymbol{x})\}$. Then
$$\partial f(\boldsymbol{x}) = \begin{cases} \mathrm{Conv}\{\boldsymbol{e}_i \mid 1 \le i \le n\}, & \boldsymbol{x} = 0, \\ \mathrm{Conv}\{\boldsymbol{e}_i \mid i \in I(\boldsymbol{x})\}, & \boldsymbol{x} \ne 0. \end{cases}$$
4. $$\partial f(\boldsymbol{x}) = \begin{cases} B_2(0,1) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| \le 1\}, & \boldsymbol{x} = 0, \\ \{\boldsymbol{x}/\|\boldsymbol{x}\|\}, & \boldsymbol{x} \ne 0. \end{cases}$$
5. Here we define
$$\begin{aligned} I_+(\boldsymbol{x}) &= \{i \mid \boldsymbol{x}^{(i)} > 0\}, \\ I_-(\boldsymbol{x}) &= \{i \mid \boldsymbol{x}^{(i)} < 0\}, \\ I_0(\boldsymbol{x}) &= \{i \mid \boldsymbol{x}^{(i)} = 0\}. \end{aligned}$$
Then we have
$$\partial f(\boldsymbol{x}) = \begin{cases} B_\infty(0,1) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \max_{1\le i\le n} |\boldsymbol{x}^{(i)}| \le 1\}, & \boldsymbol{x} = 0, \\ \sum_{i\in I_+(\boldsymbol{x})} \boldsymbol{e}_i - \sum_{i\in I_-(\boldsymbol{x})} \boldsymbol{e}_i + \sum_{i\in I_0(\boldsymbol{x})} [-\boldsymbol{e}_i, \boldsymbol{e}_i], & \boldsymbol{x} \ne 0. \end{cases}$$

□