Coursework (3) for *Introductory Lectures on Optimization*

Hanxuan Li
3220106039

Nov. 21, 2023

Excercise 1. Proof that if $f_i(\boldsymbol{x})$, $i \in I$, are convex, then

$$g(\boldsymbol{x}) = \max_{i \in I} f_i(\boldsymbol{x})$$

is also convex.

Proof of Excercise 1: Because $\forall i \in I$, $f_i(\boldsymbol{x})$ are convex, we have:

$$\forall i \in I, f_i(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f_i(\boldsymbol{x}) + (1-\alpha) f_i(\boldsymbol{y}) \leq \alpha g(\boldsymbol{x}) + (1-\alpha) g(\boldsymbol{y}) \tag{1}$$

where $\alpha \in [0,1]$. So we have:

$$g(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) = \underset{i \in I}{max} \{f_i(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y})\} \leq max \{\alpha g(\boldsymbol{x}) + (1-\alpha) g(\boldsymbol{y})\} = \alpha g(\boldsymbol{x}) + (1-\alpha) g(\boldsymbol{y}) \tag{2}$$

where $\alpha \in [0,1]$. So $g(\boldsymbol{x})$ is also convex. □

Excercise 2. Proof that

1. if f is a convex function on $\mathbb{R}^n$ and $F(\cdot)$ is a convex and non-decreasing function on $\mathbb{R}$, then $g(\boldsymbol{x}) = F(f(\boldsymbol{x}))$ is convex.
2. If $f_i, i = 1, \ldots, m$ are convex functions on $\mathbb{R}^n$ and $F(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m)$ is convex and non-decreasing (component-wise) in each argument, then

   $$g(\boldsymbol{x}) = F(f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x}))$$

   is convex.

Proof of Excercise 2:

1. Because f is convex, so $\forall \alpha \in [0,1]$, we have

   $$f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha)(f(\boldsymbol{y})) \tag{3}$$

   So we have $\forall \alpha \in [0,1]$:

   $$\begin{aligned} g(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &= F(f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y})) & (4) \\ &\leq F(\alpha f(\boldsymbol{x}) + (1-\alpha)(f(\boldsymbol{y}))) & (5) \\ &\leq \alpha F(f(\boldsymbol{x})) + (1-\alpha) F(f(\boldsymbol{y})) & (6) \\ &= \alpha g(\boldsymbol{x}) + (1-\alpha) g(\boldsymbol{y}) & (7) \end{aligned}$$

   We obtain (5) because F is non-decreasing and (3), and we obatin (6) because F is convex.Thus from (7) we obtain that $g(\boldsymbol{x})$ is convex.

2. Because f is convex, so $\forall \alpha \in [0,1]$, we have

$$f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha)(f(\boldsymbol{y})) \tag{8}$$

So we have $\forall \alpha \in [0,1]$:

$$\begin{aligned}
g(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &= F(f_1(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}), \cdots, f_m(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y})) &(9)\\
&\leq F(\alpha f_1(\boldsymbol{x}) + (1-\alpha)(f_1(\boldsymbol{y})), \cdots, \alpha f_m(\boldsymbol{x}) + (1-\alpha)(f_m(\boldsymbol{y}))) &(10)\\
&\leq \alpha F(f_1(x), \cdots, f_m(x)) + (1-\alpha)F(f_1(x), \cdots, f_m(x)) &(11)\\
&= \alpha g(\boldsymbol{x}) + (1-\alpha)g(\boldsymbol{y}) &(12)
\end{aligned}$$

We obtain (10) because F is non-decreasing and (8), and we obatin (11) because F is convex.Thus from (12) we obtain that $g(\boldsymbol{x})$ is convex.

□

Excercise 3. Proof that if $f(\boldsymbol{x}, \boldsymbol{y})$ is convex in $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n$ and Y is a convex set, then

$$g(\boldsymbol{x}) = \inf_{\boldsymbol{y} \in Y} f(\boldsymbol{x}, \boldsymbol{y})$$

is convex.

Proof of Excercise 3: For given $\boldsymbol{x_1}, \boldsymbol{x_2}$, we define $\left\{\boldsymbol{y_n}^{(1)}\right\}, \left\{\boldsymbol{y_n}^{(2)}\right\}$ as follows:

$$\inf_{\boldsymbol{y} \in Y} f(\boldsymbol{x_1}, \boldsymbol{y}) = \lim_{n\to\infty} f(\boldsymbol{x_1}, \boldsymbol{y_n}^{(1)}) \tag{13}$$

$$\inf_{\boldsymbol{y} \in Y} f(\boldsymbol{x_2}, \boldsymbol{y}) = \lim_{n\to\infty} f(\boldsymbol{x_2}, \boldsymbol{y_n}^{(2)}) \tag{14}$$

So we have:

$$\begin{aligned}
\alpha g(\boldsymbol{x_1}) + (1-\alpha)g(\boldsymbol{x_2}) &= \alpha \inf_{\boldsymbol{y} \in Y} f(\boldsymbol{x_1}, \boldsymbol{y}) + (1-\alpha) \inf_{\boldsymbol{y} \in Y} f(\boldsymbol{x_2}, \boldsymbol{y}) &(15)\\
&= \alpha \lim_{n\to\infty} f(\boldsymbol{x_1}, \boldsymbol{y_n}^{(1)}) + (1-\alpha) \lim_{n\to\infty} f(\boldsymbol{x_2}, \boldsymbol{y_n}^{(2)}) &(16)\\
&= \lim_{n\to\infty} (\alpha f(\boldsymbol{x_1}, \boldsymbol{y_n}^{(1)}) + (1-\alpha) f(\boldsymbol{x_2}, \boldsymbol{y_n}^{(2)})) &(17)\\
&\geq \lim_{n\to\infty} f(\alpha \boldsymbol{x_1} + (1-\alpha)\boldsymbol{x_2}, \alpha \boldsymbol{y_n}^{(1)} + (1-\alpha)\boldsymbol{y_n}^{(2)}) &(18)\\
&\geq \inf_{\boldsymbol{y} \in Y} f(\alpha \boldsymbol{x_1} + (1-\alpha)\boldsymbol{x_2}, \boldsymbol{y}) &(19)\\
&= g(\alpha \boldsymbol{x_1} + (1-\alpha)\boldsymbol{x_2}) &(20)
\end{aligned}$$

We obtain (18) because $f(\boldsymbol{x}, \boldsymbol{y})$ is convex, and we obtain (19) since Y is a convex set. So from (20) we have proved that $g(\boldsymbol{x})$ is also convex. □

Excercise 4. Proof that the following univariate functions are in the set of $\mathcal{F}^1(\mathbb{R})$

$$\begin{aligned}
&f(x) = e^x,\\
&f(x) = |x|^p, \ p > 1,\\
&f(x) = \frac{x^2}{1+|x|},\\
&f(x) = |x| - \ln(1+|x|).
\end{aligned}$$

Proof of Excercise 4: From lecture, we have a continuously differentiable function f belongs to $\mathcal{F}^1(\mathbb{R}^n)$ iff for any $x, y \in \mathbb{R}^n$ we have:

$$\langle \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle \geq 0 \tag{21}$$

1. $f(x) = e^x$ Obviously, we have:

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle = (e^x - e^y)(x - y) \geq 0 \tag{22}$$

   So $f(x) = e^x$ is in the set of $\mathcal{F}^1(\mathbb{R})$

2. $f(x) = |x|^p, p > 1$ we have:

$$f'(x) = \begin{cases} p|x|^{(p-1)} & x > 0 \\ 0 & x = 0 \\ -p|x|^{(p-1)} & x < 0 \end{cases} \tag{23}$$

   Now divide the problem into four cases:

   (a) $x \geq 0, y \geq 0$: we have $\langle \nabla f(x) - \nabla f(y), x - y \rangle = p(x^{p-1} - y^{p-1})(x - y) \geq 0$

   (b) $x \geq 0, y < 0$ we have $\langle \nabla f(x) - \nabla f(y), x - y \rangle = p(x^{p-1} + |y|^{p-1})(x - y) \geq 0$

   (c) $x < 0, y \geq 0$ we have $\langle \nabla f(x) - \nabla f(y), x - y \rangle = p(-|x|^{p-1} - y^{p-1})(x - y)$

   since $x - y < 0$ and $-|x|^{p-1} - y^{p-1} < 0$, we have $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$

   (d) $x < 0, y < 0$ we have $\langle \nabla f(x) - \nabla f(y), x - y \rangle = p(-|x|^{p-1} + |y|^{p-1})(x - y) \geq 0$

   Above all, we have $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$, So end of the proof.

3. $f(x) = \frac{x^2}{1+|x|}$ Firstly, we have

$$f^{'}(x) = \begin{cases} 1 - \frac{1}{(1+|x|)^2} & x > 0 \\ 0 & x = 0 \\ \frac{1}{(1+|x|)^2} - 1 & x < 0 \end{cases} \tag{24}$$

   Then:

   (a) $x \geq 0, y \geq 0$ $\langle \nabla f(x) - \nabla f(y), x - y \rangle = (\frac{1}{(1+|x|)^2} - \frac{1}{(1+|y|)^2})(x - y) \geq 0$

   (b) $x \geq 0, y < 0$ $\langle \nabla f(x) - \nabla f(y), x - y \rangle = (2 - \frac{1}{(1+|x|)^2} - \frac{1}{(1+|y|)^2})(x - y)$ Since $x > y$ and $f(x) = \frac{1}{(1+|x|)^2} \leq 1$,so $2 - \frac{1}{(1+|x|)^2} - \frac{1}{(1+|y|)^2} \geq 0$, thus $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$

   (c) $y \geq 0, x < 0$ simliar as case (b), just switching x, y.

   (d) $x < 0, y < 0$ $\langle \nabla f(x) - \nabla f(y), x - y \rangle = (\frac{1}{(1+|y|)^2} - \frac{1}{(1+|x|)^2})(x - y) \geq 0$

   Above all, $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$, so end of the proof.

4. $f(x) = |x| - ln(1 + |x|)$ Firstly, we have:

$$f^{'}(x) = \begin{cases} 1 - \frac{1}{1+|x|} & x > 0 \\ 0 & x = 0 \\ \frac{1}{1+|x|} - 1 & x < 0 \end{cases} \tag{25}$$

   Then:

(a) $x \geq 0, y \geq 0$ $\langle \nabla f(x) - \nabla f(y), x - y \rangle = (\frac{1}{1+|x|} - \frac{1}{1+|y|})(x - y) \geq 0$

(b) $x \geq 0, y < 0$ $\langle \nabla f(x) - \nabla f(y), x - y \rangle = (2 - \frac{1}{1+|y|} - \frac{1}{1+|x|})(x - y)$ Also, we have $f(x) = \frac{1}{1+|x|} \leq 1$,So $2 - \frac{1}{1+|y|} - \frac{1}{1+|x|} \geq 0$, thus $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$

(c) $x < 0, y \geq 0$: simliar as case (b), just switching x, y.

(d) $x < 0, y < 0$: $\langle \nabla f(x) - \nabla f(y), x - y \rangle = (\frac{1}{1+|y|} - \frac{1}{1+|x|})(x - y) \geq 0$

Above all, we have $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$, so end of the proof.

□

Excercise 5. For $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$ and function $\phi(\boldsymbol{y}) = f(\boldsymbol{y}) - \langle \nabla f(\boldsymbol{x}_0),\ \boldsymbol{y} \rangle$, prove that $\phi \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$, and its optimal point is $\boldsymbol{y}^* = \boldsymbol{x}_0$.

Proof of Excercise 5:

1. $\phi(y)$ is continuously differentiable: Obviously we have:

$$\nabla \phi(y) = \nabla f(y) - \nabla f(x_0) \tag{26}$$

   Since $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$, so $f(x)$ is continuously differentiable. So $\phi(y)$ is continuously differentiable.

2. $\phi(y)$ is convex: Since $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$, so $f(x)$ is convex, which means $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0$ So

$$\langle \nabla \phi(x) - \nabla \phi(y), x - y \rangle = \langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0 \tag{27}$$

   which means $\phi(x)$ is convex.

3. $\phi(y)$ satisfies the Lipschitz continuous with constant L. Since $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$, so $|\nabla f(x) - \nabla f(y)| \leq L\|x - y\|^2$ we have

$$|\nabla \phi(x) - \nabla \phi(y)| = |\nabla f(x) - \nabla f(y)| \leq L\|x - y\|^2 \tag{28}$$

   Above all, $\phi(y) \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$ We can easily find that $\nabla \phi(\boldsymbol{x_0}) = 0$.

   Because of the properties that the convex function satisfies, $y^* = \boldsymbol{x_0}$ must be the global minimum, i.e, optimal point.

□

Excercise 6. Proof that, for $f : \mathbb{R}^n \to \mathbb{R}$ and α from $[0, 1]$, if

$$\begin{aligned} \alpha f(\boldsymbol{x}) + (1 - \alpha) f(\boldsymbol{y}) \geq f(\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) \\ + \frac{\alpha(1 - \alpha)}{2L} \left\| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \right\|^2, \end{aligned}$$

then $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$.

Proof of Excercise 6:

$$\begin{aligned} \alpha f(\boldsymbol{x}) + (1 - \alpha) f(\boldsymbol{y}) &\geq f(\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) + \frac{\alpha(1 - \alpha)}{2L} \left\| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \right\|^2 \\ &\geq f(\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) \end{aligned} \tag{29}$$

So $f(x)$ is convex.

Now we should prove that f satisfies the Lipschitz condition.We have:

$$\alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) \geq f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) + \frac{\alpha(1-\alpha)}{2L} \left\| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \right\|^2$$

$$f(\boldsymbol{y}) \geq \frac{f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) - \alpha f(\boldsymbol{x})}{1-\alpha} + \frac{\alpha}{2L} \| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \|^2 \tag{30}$$

When $\alpha \to 1$,

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{1}{2L} \| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \|^2 \tag{31}$$

switching $\boldsymbol{x}, \boldsymbol{y}$ in (27),and add them together, we obtain:

$$\frac{1}{L} \| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \|^2 \leq \langle \nabla f(\boldsymbol{y}) - \nabla f(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle \tag{32}$$

Applying Cauchy-Schwarz inequailty,

$$\frac{1}{L} \| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \|^2 \leq \langle \nabla f(\boldsymbol{y}) - \nabla f(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle \tag{33}$$

$$\leq \| \nabla f(\boldsymbol{y}) - \nabla f(\boldsymbol{x}) \| \| \boldsymbol{y} - \boldsymbol{x} \| \tag{34}$$

That is: $|\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})| \leq L \| \boldsymbol{y} - \boldsymbol{x} \|$.

So, $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$

□

Excercise 7. Proof that, for $f : \mathbb{R}^n \to \mathbb{R}$ and α from $[0,1]$, if

$$\begin{aligned} 0 &\leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) - f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \\ &\leq \alpha(1-\alpha) \frac{L}{2} \left\| \boldsymbol{x} - \boldsymbol{y} \right\|^2, \end{aligned}$$

then $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$.

Proof of Excercise 7: From the inequailty given above, we can obtain:

$$f(\boldsymbol{y}) \leq \frac{f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) - \alpha f(\boldsymbol{x})}{1-\alpha} + \frac{\alpha L}{2} \| \boldsymbol{y} - \boldsymbol{x} \|^2 \tag{35}$$

When $\alpha \to 1$, we have

$$f(\boldsymbol{y}) \leq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{L}{2} \| \boldsymbol{y} - \boldsymbol{x} \|^2 \tag{36}$$

And it is equivalent to the conclusion that $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$. So end of the proof. □