# Mid-term Exam for *Introductory Lectures on Optimization*

3220102535

Dec. 8, 2024

**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:** Since $f_i(\boldsymbol{x})$ is convex, for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, we have

$$f_i(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \le \alpha f_i(\boldsymbol{x}) + (1-\alpha) f_i(\boldsymbol{y}).$$

Since $g(\boldsymbol{x})$ is the maximum of $f_i(\boldsymbol{x})$ for $i \in I$, we have

$$\begin{aligned}
g(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &= \max_{i \in I} f_i(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \\
&\le \max_{i \in I} (\alpha f_i(\boldsymbol{x}) + (1-\alpha) f_i(\boldsymbol{y})) \\
&\le \alpha \max_{i \in I} f_i(\boldsymbol{x}) + (1-\alpha) \max_{i \in I} f_i(\boldsymbol{y}) \\
&= \alpha g(\boldsymbol{x}) + (1-\alpha) g(\boldsymbol{y}).
\end{aligned}$$

Therefore, $g(\boldsymbol{x})$ is convex. $\square$

**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. Since $f$ is convex, for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, we have

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

Since $F$ is non-decreasing, we have

$$F(f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y})) \leq F(\alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y})).$$

Since $F$ is convex, we have

$$F(\alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y})) \leq \alpha F(f(\boldsymbol{x})) + (1-\alpha) F(f(\boldsymbol{y})).$$

Therefore,

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

Therefore, $g(\boldsymbol{x})$ is convex.

2. Since $f_i(\boldsymbol{x})$ is convex, for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\alpha \in [0, 1]$, we have

$$f_i(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f_i(\boldsymbol{x}) + (1-\alpha) f_i(\boldsymbol{y}).$$

For any $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m$, and $\alpha \in [0, 1]$, we have

$$\begin{aligned} g(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) &= F(f_1(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}), \ldots, f_m(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y})) \\ &\leq F(\alpha f_1(\boldsymbol{x}) + (1-\alpha) f_1(\boldsymbol{y}), \ldots, \alpha f_m(\boldsymbol{x}) + (1-\alpha) f_m(\boldsymbol{y})) \\ &= F(\alpha(f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})) + (1-\alpha)(f_1(\boldsymbol{y}), \ldots, f_m(\boldsymbol{y}))) \\ &\leq \alpha F(f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})) + (1-\alpha) F(f_1(\boldsymbol{y}), \ldots, f_m(\boldsymbol{y})) \\ &= \alpha g(\boldsymbol{x}) + (1-\alpha) g(\boldsymbol{y}). \end{aligned}$$

Therefore, $g(\boldsymbol{x})$ is convex.

$\square$

**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:** Since $f(\boldsymbol{x}, \boldsymbol{y})$ is convex in $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n$, for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$, $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^n$, and $\alpha \in [0, 1]$, we have

$$f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2, \alpha \boldsymbol{y}_1 + (1-\alpha)\boldsymbol{y}_2) \leq \alpha f(\boldsymbol{x}_1, \boldsymbol{y}_1) + (1-\alpha) f(\boldsymbol{x}_2, \boldsymbol{y}_2).$$

Therefore,

$$\begin{aligned}
g(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) &= \inf_{\boldsymbol{y} \in Y} f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2, \boldsymbol{y}) \\
&= \inf_{\boldsymbol{y}_1 \in Y, \boldsymbol{y}_2 \in Y} f(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2, \alpha \boldsymbol{y}_1 + (1-\alpha)\boldsymbol{y}_2) \\
&\leq \inf_{\boldsymbol{y}_1 \in Y, \boldsymbol{y}_2 \in Y} (\alpha f(\boldsymbol{x}_1, \boldsymbol{y}_1) + (1-\alpha) f(\boldsymbol{x}_2, \boldsymbol{y}_2)) \\
&= \alpha \inf_{\boldsymbol{y}_1 \in Y} f(\boldsymbol{x}_1, \boldsymbol{y}_1) + (1-\alpha) \inf_{\boldsymbol{y}_2 \in Y} f(\boldsymbol{x}_2, \boldsymbol{y}_2) \\
&= \alpha g(\boldsymbol{x}_1) + (1-\alpha) g(\boldsymbol{x}_2).
\end{aligned}$$

Therefore, $g(\boldsymbol{x})$ is 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:**

1. $f(x)$ is a continuous differentiable function, and for any $x_1, x_2 \in \mathbb{R}$, $\alpha \in [0, 1]$,we have:

$$\begin{aligned}
e^{x_2 - x_1} &\geq x_2 - x_1 + 1 \\
e^{x_2} &\geq e^{x_1}(x_2 - x_1 + 1) \\
e^{x_2} &\geq e^{x_1} + e^{x_1}(x_2 - x_1)
\end{aligned}$$

   This implies that:

$$f(x_2) \geq f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle$$

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

2. For any $x \in \mathbb{R}$ with $x > 0$:

$$\nabla f(x) = p x^{p-1}$$

   For any $x \in \mathbb{R}$ with $x < 0$:

$$\nabla f(x) = -p(-x)^{p-1}$$

When $x = 0$, Since

$$\lim_{x \to 0^-} \frac{|x|^p - 0}{x} = \lim_{x \to 0^-} \left((-1)^p x^{p-1}\right) = 0$$
$$\lim_{x \to 0^+} \frac{|x|^p - 0}{x} = \lim_{x \to 0^+} x^{p-1} = 0$$

Then $\nabla f(x) = 0$. Hence, $f(x)$ is a continuous differentiable function.

Then for any $x_1, x_2 \in \mathbb{R}$, assume $x_1 \geq x_2$, we have:

- If $x_2 > 0$:
$$\langle \nabla f\left(x_1\right) - \nabla f\left(x_2\right), x_1 - x_2 \rangle = \left(p x_1^{p-1} - p x_2^{p-1}\right) \cdot \left(x_1 - x_2\right) \geq 0$$
- If $x_2 = 0$:
$$\langle \nabla f\left(x_1\right) - \nabla f\left(x_2\right), x_1 - x_2 \rangle = p x_1^p \geq 0$$
- If $x_1 > 0, x_2 < 0$:
$$\langle \nabla f\left(x_1\right) - \nabla f\left(x_2\right), x_1 - x_2 \rangle = \left(p x_1^{p-1} + p\left(-x_2\right)^{p-1}\right) \cdot \left(x_1 - x_2\right) \geq 0$$
- If $x_1 = 0, x_2 < 0$:
$$\langle \nabla f\left(x_1\right) - \nabla f\left(x_2\right), x_1 - x_2 \rangle = p\left(-x_2\right)^p \geq 0$$
- If $x_1 < 0, x_2 < 0$:
$$\langle \nabla f\left(x_1\right) - \nabla f\left(x_2\right), x_1 - x_2 \rangle = \left(-p\left(-x_1\right)^{p-1} + p\left(-x_2\right)^{p-1}\right) \cdot \left(x_1 - x_2\right) \geq 0$$

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

3. For any $x \in \mathbb{R}, x > 0$:

$$\nabla f\left(x\right) = \frac{x^2 + 2x}{\left(1 + x\right)^2}$$

For any $x \in \mathbb{R}, x < 0$:

$$\nabla f\left(x\right) = \frac{-x^2 + 2x}{\left(1 - x\right)^2}$$

When $x = 0$: Since

$$\lim_{x \to 0^-} \frac{\frac{x^2}{1+|x|} - 0}{x} = \lim_{x \to 0^+} \frac{\frac{x^2}{1+|x|} - 0}{x} = 0$$

Hence, $\nabla f(x) = 0$. Hence $f(x)$ is a continuous differentiable function.

Then for any $x_1, x_2 \in \mathbb{R}$, assume $x_1 \geq x_2$, we have:

- If $x_2 > 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \left( \frac{x_1^2 + 2x_1}{(1+x_1)^2} - \frac{x_2^2 + 2x_2}{(1+x_2)^2} \right) \cdot (x_1 - x_2)$$

$$= \frac{x_1^2 - x_2^2 + 2x_1 - 2x_2}{(1+x_1)^2 (1+x_2)^2} \cdot (x_1 - x_2) \geq 0$$

- If $x_2 = 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \frac{x_1^3 + 2x_1^2}{(1+x_1)^2} \geq 0$$

- If $x_1 > 0, x_2 < 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \frac{2x_1^2 x_2^2 - 4x_1^2 x_2 + 4x_1 x_2^2 + x_1^2 + x_2^2 - 8x_1 x_2 + 2x_1 - 2x_2}{(1+x_1)^2 (1-x_2)^2} \cdot (x_1 - x_2) \geq 0$$

- If $x_1 = 0, x_2 < 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \frac{-x_2^3 + 2x_2^2}{(1-x_2)^2} \geq 0$$

- If $x_1 < 0, x_2 < 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \left( \frac{-x_1^2 + 2x_1}{(1-x_1)^2} - \frac{-x_2^2 + 2x_2}{(1-x_2)^2} \right) \cdot (x_1 - x_2)$$

$$= \frac{-x_1^2 + 2x_1 + x_2^2 - 2x_2}{(1-x_1)^2 (1-x_2)^2} \cdot (x_1 - x_2) \geq 0$$

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

4. For any $x \in \mathbb{R}, x > 0$:

$$\nabla f(x) = \frac{x}{1+x}$$

For any $x \in \mathbb{R}, x < 0$:

$$\nabla f(x) = \frac{x}{1-x}$$

When $x = 0$, Since

$$\lim_{x \to 0^-} (|x| - \ln(1+|x|)) = \lim_{x \to 0^+} (|x| - \ln(1+|x|)) = 0$$

Then $\nabla f(x) = 0$. Hence, $f(x)$ is a continuous differentiable function.

Then for any $x_1, x_2 \in \mathbb{R}$, assume $x_1 \geq x_2$, we have:

- If $x_2 > 0$:

$$\begin{aligned}\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle &= \left( \frac{x_1}{1+x_1} - \frac{x_2}{1+x_2} \right) \cdot (x_1 - x_2) \\ &= \frac{(x_1 - x_2)^2}{(1+x_1)(1+x_2)} \geq 0\end{aligned}$$

- If $x_2 = 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \frac{x_1^2}{1+x_1} \geq 0$$

- If $x_1 > 0, x_2 < 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \frac{x_1 - x_2 - 2x_1x_2}{(1+x_1)(1-x_2)} \cdot (x_1 - x_2) \geq 0$$

- If $x_1 = 0, x_2 < 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \frac{x_2^2}{1-x_2} \geq 0$$

- If $x_1 < 0, x_2 < 0$:

$$\langle \nabla f(x_1) - \nabla f(x_2), x_1 - x_2 \rangle = \frac{(x_1 - x_2)^2}{(1-x_1)(1-x_2)} \geq 0$$

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

$\square$

**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:** Since $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$, $f$ is convex and Lipschitz continuous, which means there exists a constant $L$ such that for any $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^n$:

$$\|\nabla f(\boldsymbol{y}_1) - \nabla f(\boldsymbol{y}_2)\| \leq L\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|$$

We prove that $\phi(\boldsymbol{y}) = f(\boldsymbol{y}) - \langle \nabla f(\boldsymbol{x}_0), \boldsymbol{y} \rangle$ belongs to $\mathcal{F}_L^{1,1}(\mathbb{R}^n)$ as follows:

$$\begin{aligned}
&\|\nabla \phi(\boldsymbol{y}_1) - \nabla \phi(\boldsymbol{y}_2)\| \\
&= \|\nabla f(\boldsymbol{y}_1) - \nabla(\langle \nabla f(x_0), \boldsymbol{y}_1 \rangle) - (\nabla f(\boldsymbol{y}_2) - \nabla(\langle \nabla f(x_0), \boldsymbol{y}_2 \rangle))\| \\
&= \|\nabla f(\boldsymbol{y}_1) - \nabla f(\boldsymbol{y}_2)\| \\
&\leq L\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|
\end{aligned}$$

Since $f(\boldsymbol{y})$ is convex and $\langle \nabla f(\boldsymbol{x}_0), \boldsymbol{y} \rangle$ is linear, $\phi(\boldsymbol{y})$ is convex.

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

Let $\boldsymbol{y}^*$ be the optimal point of $\phi$. From the properties of $\mathcal{F}_L^{1,1}(\mathbb{R}^n)$, we have:

$$\nabla \phi(\boldsymbol{y}^*) = \mathbf{0}$$
$$\nabla f(\boldsymbol{y}^*) = \nabla f(\boldsymbol{x}_0)$$

Since the gradient is monotonic on convex functions, it follows that:

$$\boldsymbol{y}^* = \boldsymbol{x}_0$$

Hence, the optimal point of $\phi$ is $\boldsymbol{x}_0$. $\square$

**Excercise 6.** Proof that, for $f : \mathbb{R}^n \to \mathbb{R}$ and $\alpha$ 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:** First, we have

$$\alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) - f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq \frac{\alpha(1-\alpha)}{2L} \left\| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \right\|^2 \geq 0$$

So, $f \in \mathcal{F}^{1,1}(\mathbb{R}^n)$. Now, we have.

$$\begin{aligned} f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), (1-\alpha)(\boldsymbol{y}-\boldsymbol{x}) \rangle \\ f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq f(\boldsymbol{y}) + \langle \nabla f(\boldsymbol{y}), \alpha(\boldsymbol{x}-\boldsymbol{y}) \rangle \end{aligned}$$

Then, we have

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

Now, we let $\left\| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \right\| = k \left\| \boldsymbol{x}-\boldsymbol{y} \right\|$. We have known that $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, k \leq 2L$. So, $k$ must have upward boundary. We assume that $m = \sup_{\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n} k$. Then, we have $f \in \mathcal{F}_m^{1,1}(\mathbb{R}^n)$

$$\begin{aligned} & \frac{\alpha(1-\alpha)}{2L} \left\| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \right\|^2 \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) - f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha(1-\alpha) \frac{m}{2} \left\| \boldsymbol{x}-\boldsymbol{y} \right\|^2 \\ & \Rightarrow \left\| \nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y}) \right\| \leq \sqrt{mL} \left\| \boldsymbol{x}-\boldsymbol{y} \right\| \\ & \Rightarrow \sqrt{mL} \geq m \\ & \Rightarrow m \leq L \end{aligned}$$

Therefore, $f \in \mathcal{F}_m^{1,1}(\mathbb{R}^n) \subseteq \mathcal{F}_L^{1,1}(\mathbb{R}^n)$. $\square$

**Excercise 7.** Proof that, for $f : \mathbb{R}^n \to \mathbb{R}$ and $\alpha$ 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:**

1. Since

$$\alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) - f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq 0$$

then $f$ is convex

2. Given the inequality:

$$0 \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) - f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha(1-\alpha)\frac{L}{2}\|\boldsymbol{x}-\boldsymbol{y}\|^2$$

we can rearrange it to:

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

Applying L'Hospital's rule, we find:

$$\lim_{\alpha \to 1} \frac{f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) - \alpha f(\boldsymbol{x})}{1-\alpha} = f(\boldsymbol{x}) + \langle \nabla f(\boldsymbol{x}), \boldsymbol{y}-\boldsymbol{x} \rangle$$

Thus, we obtain:

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

3. For any $\boldsymbol{x} \in \mathbb{R}^n$, fix $\boldsymbol{x}$, and consider the function $\phi(\boldsymbol{y}) = f(\boldsymbol{y}) + g(\boldsymbol{y})$, where $g(\boldsymbol{y}) = -\langle \nabla f(\boldsymbol{x}), \boldsymbol{y} \rangle$.

For any $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^n$, we have:

$$f(\boldsymbol{y}_2) \leq f(\boldsymbol{y}_1) + \langle \nabla f(\boldsymbol{y}_1), \boldsymbol{y}_2 - \boldsymbol{y}_1 \rangle + \frac{L}{2}\|\boldsymbol{y}_2 - \boldsymbol{y}_1\|^2$$

$$g(\boldsymbol{y}_2) = g(\boldsymbol{y}_1) + \langle \nabla g(\boldsymbol{y}_1), \boldsymbol{y}_2 - \boldsymbol{y}_1 \rangle$$

then:

$$\phi(\boldsymbol{y}_2) \leq \phi(\boldsymbol{y}_1) + \langle \nabla \phi(\boldsymbol{y}_1), \boldsymbol{y}_2 - \boldsymbol{y}_1 \rangle + \frac{L}{2}\|\boldsymbol{y}_2 - \boldsymbol{y}_1\|^2$$

Since $f$ and $g$ are both convex, $\phi$ is convex.

Let the optimal point of $\phi$ be $\boldsymbol{y}^*$, then:

$$\nabla \phi(\boldsymbol{y}^*) = 0$$

$$\nabla f(\boldsymbol{y}^*) = \nabla f(\boldsymbol{x})$$

Since the gradient is monotonic on convex function, then:

$$\boldsymbol{y}^* = \boldsymbol{x}$$

Hence, we have:

$$\phi(\boldsymbol{y}^*) \leq \phi\left(\boldsymbol{y} - \frac{1}{L}\nabla\phi(\boldsymbol{y})\right)$$

$$\phi\left(\boldsymbol{y} - \frac{1}{L}\nabla\phi(\boldsymbol{y})\right) \leq \phi(\boldsymbol{y}) + \left\langle \nabla\phi(\boldsymbol{y}), \boldsymbol{y} - \frac{1}{L}\nabla\phi(\boldsymbol{y}) - \boldsymbol{y}\right\rangle + \frac{1}{2L}\|\nabla\phi(\boldsymbol{y})\|^2$$

$$\phi(\boldsymbol{y}^*) \leq \phi(\boldsymbol{y}) - \frac{1}{2L}\|\nabla\phi(\boldsymbol{y})\|^2$$

which means:

$$\phi(\boldsymbol{x}) \leq \phi(\boldsymbol{y}) - \frac{1}{2L}\|\nabla\phi(\boldsymbol{y})\|^2$$

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

Since:

$$\nabla\phi(\boldsymbol{y}) = \nabla f(\boldsymbol{y}) - \nabla f(\boldsymbol{x})$$

then:

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

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

which means:

$$\frac{1}{2L}\|\nabla f(\boldsymbol{y}) - \nabla f(\boldsymbol{x})\|^2 \leq \frac{L}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2$$

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

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

$\square$