Proof:
Left for exercise. ∎

Example 6.10 The following functions are differentiable and convex:

1. $f(x) = e^x$
2. $f(x) = |x|^p, \quad p > 1$
3. $f(x) = \frac{x^2}{1+|x|}$
4. $f(x) = |x| - \ln(1 + |x|)$
5. $f(\boldsymbol{x}) = \sum_{i=1}^m e^{\alpha_i + \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle}$
6. $f(\boldsymbol{x}) = \sum_{i=1}^m |\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle - b_i|^p, \quad p > 1$

Theorem 6.11 Let f be a twice continuously differentiable function. Then $f \in \mathcal{F}^2(\mathbb{R}^n)$ if and only if

$$f''(\boldsymbol{x}) \succeq \boldsymbol{O}, \quad \forall \boldsymbol{x} \in \mathbb{R}^n.$$

Proof:
Let $f \in \mathcal{F}^2(\mathbb{R}^n)$, and denote $\boldsymbol{x}_\tau = \boldsymbol{x} + \tau \boldsymbol{s}$, $\tau > 0$. Then, from the previous result

$$\begin{aligned} 0 &\leq \frac{1}{\tau^2}\langle f'(\boldsymbol{x}_\tau) - f'(\boldsymbol{x}), \boldsymbol{x}_\tau - \boldsymbol{x} \rangle = \frac{1}{\tau}\langle f'(\boldsymbol{x}_\tau) - f'(\boldsymbol{x}), \boldsymbol{s} \rangle \\ &= \frac{1}{\tau}\int_0^\tau \langle f''(\boldsymbol{x} + \lambda \boldsymbol{s})\boldsymbol{s}, \boldsymbol{s} \rangle d\lambda \\ &= \frac{F(\tau) - F(0)}{\tau} \end{aligned}$$

where $F(\tau) = \int_0^\tau \langle f''(\boldsymbol{x} + \lambda \boldsymbol{s})\boldsymbol{s}, \boldsymbol{s} \rangle d\lambda$. Therefore, tending τ to 0, we get $0 \leq F'(0) = \langle f''(\boldsymbol{x})\boldsymbol{s}, \boldsymbol{s} \rangle$, and we have the result.

Conversely, $\forall \boldsymbol{x} \in \mathbb{R}^n$,

$$\begin{aligned} f(\boldsymbol{y}) &= f(\boldsymbol{x}) + \langle f'(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle + \int_0^1 \int_0^\tau \langle f''(\boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x}))(\boldsymbol{y} - \boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle d\lambda d\tau \\ &\geq f(\boldsymbol{x}) + \langle f'(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle. \end{aligned}$$

∎

6.3 Differentiable Convex Functions with Lipschitz Continuous Gradients

Corollary 6.12 Let f be a two times continuously differentiable function. $f \in \mathcal{F}_L^{2,1}(\mathbb{R}^n)$ if and only if $\boldsymbol{O} \preceq f''(\boldsymbol{x}) \preceq L\boldsymbol{I}, \quad \forall \boldsymbol{x} \in \mathbb{R}^n$.

Theorem 6.13 Let f be a continuously differentiable function in $\mathbb{R}^n$, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, and $\alpha \in [0, 1]$. Then the following conditions are equivalent:

1. $f \in \mathcal{F}_L^{1,1}(\mathbb{R}^n)$.
2. $0 \leq f(\boldsymbol{y}) - f(\boldsymbol{x}) - \langle f'(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle \leq \frac{L}{2}\|\boldsymbol{x} - \boldsymbol{y}\|_2^2$.
3. $f(\boldsymbol{x}) + \langle f'(\boldsymbol{x}), \boldsymbol{y} - \boldsymbol{x} \rangle + \frac{1}{2L}\|f'(\boldsymbol{x}) - f'(\boldsymbol{y})\|_2^2 \leq f(\boldsymbol{y})$.
4. $0 \leq \frac{1}{L}\|f'(\boldsymbol{x}) - f'(\boldsymbol{y})\|_2^2 \leq \langle f'(\boldsymbol{x}) - f'(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle$.