

Definition 3.2 Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuous function on \mathbb{R}^n . This function is also a (Fréchet) differentiable function on \mathbb{R}^n if for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + o(\|\mathbf{y} - \mathbf{x}\|_2),$$

where $o(r)$ is some function of $r > 0$ such that

$$\lim_{r \rightarrow 0} \frac{1}{r} o(r) = 0, \quad o(0) = 0.$$

In particular, if a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable on \mathbb{R}^n and $\mathbf{0} \neq \mathbf{s} \in \mathbb{R}^n$, then,

$$f'(\mathbf{x}; \mathbf{s}) = \langle \nabla f(\mathbf{x}), \mathbf{s} \rangle, \quad \text{for } \mathbf{x} \in \mathbb{R}^n.$$

We say that the function is continuously differentiable if the function $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuous.

Hereafter, we define for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, the standard inner product $\langle \mathbf{a}, \mathbf{b} \rangle := \sum_{i=1}^n a_i b_i$, and the associated norm $\|\mathbf{a}\|_2 := \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle}$ to it.

Definition 3.3 Let Q be a subset of \mathbb{R}^n . We denote by $\mathcal{C}_L^{k,p}(Q)$ the class of functions with the following properties:

- Any $f \in \mathcal{C}_L^{k,p}(Q)$ is k times continuously differentiable on Q ;
- Its p th derivative is Lipschitz continuous on Q with the constant $L \geq 0$:

$$\|\mathbf{f}^{(p)}(\mathbf{x}) - \mathbf{f}^{(p)}(\mathbf{y})\|_2 \leq L\|\mathbf{x} - \mathbf{y}\|_2, \quad \forall \mathbf{x}, \mathbf{y} \in Q.$$

In particular, $\mathbf{f}^{(1)}(\mathbf{x}) = \nabla f(\mathbf{x})$ and $\mathbf{f}^{(2)}(\mathbf{x}) = \nabla^2 f(\mathbf{x})$. Observe that if $f_1 \in \mathcal{C}_{L_1}^{k,p}(Q)$, $f_2 \in \mathcal{C}_{L_2}^{k,p}(Q)$, and $\alpha, \beta \in \mathbb{R}$, then for $L_3 = |\alpha|L_1 + |\beta|L_2$ we have $\alpha f_1 + \beta f_2 \in \mathcal{C}_{L_3}^{k,p}(Q)$.

Lemma 3.4 Let $f \in \mathcal{C}^2(\mathbb{R}^n)$. Then $f \in \mathcal{C}_L^{2,1}(\mathbb{R}^n)$ if and only if $\|\nabla^2 f(\mathbf{x})\|_2 \leq L$, $\forall \mathbf{x} \in \mathbb{R}^n$.

Proof:

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\begin{aligned} \nabla f(\mathbf{y}) &= \nabla f(\mathbf{x}) + \int_0^1 \nabla^2 f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x})d\tau \\ &= \nabla f(\mathbf{x}) + \left(\int_0^1 \nabla^2 f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x}))d\tau \right) (\mathbf{y} - \mathbf{x}). \end{aligned}$$

Since $\|\nabla^2 f(\mathbf{x})\|_2 \leq L$,

$$\begin{aligned} \|\nabla f(\mathbf{y}) - \nabla f(\mathbf{x})\|_2 &\leq \left\| \int_0^1 \nabla^2 f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x}))d\tau \right\|_2 \|\mathbf{y} - \mathbf{x}\|_2 \\ &\leq \int_0^1 \|\nabla^2 f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x}))\|_2 d\tau \|\mathbf{y} - \mathbf{x}\|_2 \\ &\leq L\|\mathbf{y} - \mathbf{x}\|_2. \end{aligned}$$

On the other hand, for $\mathbf{s} \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$, $\alpha \neq 0$,

$$\|\nabla f(\mathbf{x} + \alpha \mathbf{s}) - \nabla f(\mathbf{x})\|_2 \leq |\alpha|L\|\mathbf{s}\|_2.$$

Dividing both sides by $|\alpha|$ and taking the limit to zero,

$$\|\nabla^2 f(\mathbf{x})\mathbf{s}\|_2 \leq L\|\mathbf{s}\|_2, \quad \mathbf{s} \in \mathbb{R}^n.$$

Therefore, $\|\nabla^2 f(\mathbf{x})\|_2 \leq L$. ■

Example 3.5

1. The linear function $f(\mathbf{x}) = \alpha + \langle \mathbf{a}, \mathbf{x} \rangle \in \mathcal{C}_0^{2,1}(\mathbb{R}^n)$ since

$$\nabla f(\mathbf{x}) = \mathbf{a}, \quad \nabla^2 f(\mathbf{x}) = \mathbf{O}.$$

2. The quadratic function $f(\mathbf{x}) = \alpha + \langle \mathbf{a}, \mathbf{x} \rangle + 1/2\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle$ with $\mathbf{A} = \mathbf{A}^T$ belongs to $\mathcal{C}_L^{2,1}(\mathbb{R}^n)$ where

$$\nabla f(\mathbf{x}) = \mathbf{a} + \mathbf{A}\mathbf{x}, \quad \nabla^2 f(\mathbf{x}) = \mathbf{A}, \quad L = \|\mathbf{A}\|_2.$$

3. The function $f(x) = \sqrt{1+x^2} \in \mathcal{C}_1^{2,1}(\mathbb{R})$ since

$$f'(x) = \frac{x}{\sqrt{1+x^2}}, \quad f''(x) = \frac{1}{(1+x^2)^{3/2}} \leq 1.$$

Lemma 3.6 Let $f \in \mathcal{C}_L^{1,1}(\mathbb{R}^n)$. Then for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have

$$|f(\mathbf{y}) - f(\mathbf{x}) - \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| \leq \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2.$$

Proof:

For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have

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

Therefore,

$$\begin{aligned} |f(\mathbf{y}) - f(\mathbf{x}) - \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| &= \left| \int_0^1 \langle \nabla f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x})) - \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle d\tau \right| \\ &\leq \int_0^1 |\langle \nabla f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x})) - \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle| d\tau \\ &\leq \int_0^1 \|\nabla f(\mathbf{x} + \tau(\mathbf{y} - \mathbf{x})) - \nabla f(\mathbf{x})\|_2 \|\mathbf{y} - \mathbf{x}\|_2 d\tau \\ &\leq \int_0^1 \tau L \|\mathbf{y} - \mathbf{x}\|_2^2 d\tau = \frac{L}{2} \|\mathbf{y} - \mathbf{x}\|_2^2. \end{aligned}$$

Consider a function $f \in \mathcal{C}_L^{1,1}(\mathbb{R}^n)$. Let us fix $\mathbf{x}_0 \in \mathbb{R}^n$, and define two quadratic functions:

$$\begin{aligned} \phi_1(\mathbf{x}) &= f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle - \frac{L}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2, \\ \phi_2(\mathbf{x}) &= f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \frac{L}{2} \|\mathbf{x} - \mathbf{x}_0\|_2^2. \end{aligned}$$

Then the graph of the function f is located between the graphs of ϕ_1 and ϕ_2 :

$$\phi_1(\mathbf{x}) \leq f(\mathbf{x}) \leq \phi_2(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n.$$