HIGHER ORDER DIFFERENTIABILITY

FREDERICK TSZ-HO FONG

Definition 1. $f : \mathbb{R}^n \to \mathbb{R}$ is said to be k-th differentiable at $\vec{a} \in \mathbb{R}^n$ if there exists a k-th degree polynomial $P_k : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(\vec{x}) = P_k(\vec{x}) + o\big(\|\vec{x} - \vec{a}\|^k \big) \quad \text{as } \vec{x} \to \vec{a}.$$

Theorem 2. *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has first order partial derivatives $\frac{\partial f}{\partial x_i}$'s near $\vec{a} \in \mathbb{R}^n$, and that all $\frac{\partial f}{\partial x_i}$'s are differentiable at $\vec{a}$, then f is twice differentiable at $\vec{a}$.*

Proof. It is given that $\frac{\partial f}{\partial x_i}$ is differentiable at $\vec{a}$ for any i, so its partial derivatives $\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}$ at $\vec{a}$ exist, and we have

$$\frac{\partial f}{\partial x_i}(\vec{x}) = \frac{\partial f}{\partial x_i}(\vec{a}) + \sum_{j=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}\bigg|_{\vec{a}} (x_j - a_j) + o\big(\|\vec{x} - \vec{a}\| \big) \quad \text{as } \vec{x} \to \vec{a}. \tag{0.1}$$

Now we claim that f can be approximated by the following quadratic polynomial near $\vec{a}$:

$$P_2(\vec{x}) := f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\bigg|_{\vec{a}} (x_i - a_i) + \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}\bigg|_{\vec{a}} (x_i - a_i)(x_j - a_j).$$

For each $\vec{x}$ near $\vec{a}$, we connect $\vec{a}$ and $\vec{x}$ a straight path:

$$\gamma(t) := \vec{a} + t(\vec{x} - \vec{a}),\ 0 \le t \le 1.$$

In particular, $\gamma(0) = \vec{a}$ and $\gamma(1) = \vec{x}$. Consider the composition $g(t) := f(\gamma(t))$. One can easily verify (left as an exercise) by the chain rule that

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\bigg|_{\gamma(t)} (x_i - a_i)$$
$$g''(t) = \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}\bigg|_{\gamma(t)} (x_j - a_j)(x_i - a_i)$$

Inspired by the one-variable case, we consider the error term between g and its second-order approximation near $t = 0$:

$$E_2(t) := g(t) - g(0) - g'(0)t - \frac{g''(0)}{2}t^2.$$

Note that $E_2(1) = f(\vec{x}) - P_2(\vec{x})$ which we want to show it is of order $o\big(\|\vec{x} - \vec{a}\|^2 \big)$.

By Cauchy's mean value theorem applied to E_2 and t^2, there exists $c \in (0, 1)$ such that

$$E_2(1) = \frac{E_2(1) - E_2(0)}{1^2 - 0^2} = \frac{E_2'(c)}{2c} = \frac{g'(c) - g'(0) - g''(0)c}{2c}.$$

By the results of the above chain rule exercise and (0.1), we have:

$$\begin{aligned} &E_2(1) \\ &= \frac{1}{2c}\left[\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\bigg|_{\gamma(c)} - \frac{\partial f}{\partial x_i}\bigg|_{\vec{a}} \right)(x_i - a_i) - c\sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}\bigg|_{\vec{a}} (x_i - a_i)(x_j - a_j) \right] \\ &= \frac{1}{2c}\left\{ \sum_{i=1}^{n} \left[\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}\bigg|_{\vec{a}} (\gamma_j(c) - a_j) + o\big(\|\gamma(c) - \vec{a}\| \big) \right] (x_i - a_i) - c\sum_{i,j=1}^{n} \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}\bigg|_{\vec{a}} (x_i - a_i)(x_j - a_j) \right\}. \end{aligned}$$

Note that $\gamma(c) - \vec{a} = c(\vec{x} - \vec{a})$ and so $\gamma_j(c) - a_j = c(x_j - a_j)$, so after simplification we get:

$$|E_2(1)| = \left| o\big(\|\vec{x} - \vec{a}\|\big) \cdot \sum_i (x_i - a_i) \right| \leq 0\big(\|\vec{x} - \vec{a}\|^2 \big).$$

As a result, we have

$$|f(\vec{x}) - P_2(\vec{x})| \leq o\big(\|\vec{x} - \vec{a}\|^2 \big),$$

as desired. □

Corollary 3. *If $f : \mathbb{R}^n \to \mathbb{R}$ is C^2 on an open set $U \subset \mathbb{R}^n$, i.e. its first partial derivatives are C^1 on U, then f is twice differentiable on U.*

Proof. By the fact that C^1 implies differentiability. □

For higher-order differentiability, we have a similar theorem and corollary:

Theorem 4. *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ has $(k-1)$-th order partial derivatives near $\vec{a}$, and that all of them are differentiable at $\vec{a}$, then f is k-th differentiable at $\vec{a}$.*

Proof. Similar to the twice differentiable proof. The only thing that needs to be changed is that we should consider:

$$E_k(t) := g(t) - g(0) - g'(0)t - \frac{g''(0)}{2!}t^2 - \cdots - \frac{g^{(k)}(0)}{k!}t^k.$$

Apply Cauchy's mean value theorem on $E_k(t)$ and t^k to estimate $E_k(1)$. □

Corollary 5. *If $f : \mathbb{R}^n \to \mathbb{R}$ is C^k on an open set U in $\mathbb{R}^n$, then it is k-th differentiable on U.*