

# Econ 703 TA Note 6

Chia-Min Wei \*

August 25, 2025

## 1 Differentiable Functions on $\mathbb{R}$

**Definition 1.1 (Differentiable at a Point):** A function  $f : (a, b) \rightarrow \mathbb{R}$  is differentiable at a point  $c$  if for all  $\{x_n\} \subset [a, b]$  that converges to  $c$ ,

$$\lim_{n \rightarrow \infty} \frac{f(x_n) - f(c)}{x_n - c}$$

exists and equals the same value. This value is called the derivative of  $f$  at  $c$ , denoted  $f'(c)$ .

From the definition we can immediately see that if  $f$  is differentiable at  $c$ , then it is also continuous at  $c$ .

**Definition 1.2 (Derivative of a Function):** If a function  $f : (a, b) \rightarrow \mathbb{R}$  is differentiable everywhere in  $(a, b)$ , we say that the function is differentiable.  $f'(x)$  is well-defined for all  $x \in (a, b)$ , and thus is also a function on  $(a, b)$ . An alternative notation is  $\frac{d}{dx}f(x)$ .

Here are some common rules for taking derivatives:

- **Power Rule:**  $\frac{d}{dx}[x^n] = nx^{n-1}$ ,  $n \in \mathbb{R}$ .
- **Constant Multiple Rule:**  $\frac{d}{dx}[c \cdot f(x)] = c \cdot f'(x)$ .
- **Sum/Difference Rule:**  $\frac{d}{dx}[f(x) \pm g(x)] = f'(x) \pm g'(x)$ .
- **Product Rule:**  $\frac{d}{dx}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x)$ .
- **Quotient Rule:**  $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$ .
- **Chain Rule:**  $\frac{d}{dx}[f(g(x))] = f'(g(x)) \cdot g'(x)$ .

## 2 Big O, Little o

If a function is differentiable at a point  $c$ , then the linear function

$$h(x) = f(c) + f'(c)(x - c)$$

provides a “good” approximation of  $f(x)$  near  $x = c$ . But what exactly do we mean by good? To make this precise, let us introduce the concepts of *big O* and *little o* notation.

---

\*This TA note was prepared for the Econ PhD math camp taught by Prof. John Kennan at UW-Madison in 2025. All errors are mine.

**Definition 2.1 (Big O):** We say that a function  $f(x)$  is of  $O(g(x))$  as  $x \rightarrow c$ , written as  $f(x) = O(g(x))$  as  $x \rightarrow c$ , if there exists  $\delta, M > 0$  such that for all  $|x - c| < \delta$ ,

$$|f(x)| \leq M|g(x)|.$$

$f(x) = O(g(x))$  as  $x \rightarrow c$  means  $f(x)$  is bounded by a constant multiple of  $g(x)$  near  $x = c$ .

**Example 2.1:** If  $f$  is differentiable at  $c$ , then  $f(x) - f(c) = O(|x - c|)$ . This is because

$$\lim_{x \rightarrow c} \left| \frac{f(x) - f(c)}{x - c} \right| = |f'(c)| < \infty.$$

**Example 2.2:**  $f$  being continuous at  $c$  does not necessarily mean that  $f(x) - f(c) = O(|x - c|)$ . For example,  $f(x) = \sqrt{|x|}$  is continuous at  $x = 0$ , however,

$$\lim_{x \rightarrow 0} \left| \frac{\sqrt{|x|}}{|x|} \right| = \lim_{x \rightarrow 0} |x|^{-1/2} \rightarrow \infty.$$

**Definition 2.2 (Little o):** We say that a function  $f(x)$  is of  $o(g(x))$  as  $x \rightarrow c$ , written as  $f(x) = o(g(x))$  as  $x \rightarrow c$ , if

$$\lim_{x \rightarrow c} \left| \frac{f(x)}{g(x)} \right| = 0.$$

$f(x) = o(g(x))$  as  $x \rightarrow c$  means  $f(x)$  becomes negligible compared to  $g(x)$  as  $x$  approaches  $c$ . Here are some rules regarding big  $O$  and little  $o$ .

1.  $f(x) = o(g(x)) \implies f(x) = O(g(x))$ .
2.  $f(x) = O(g(x)) \implies \alpha f(x) = O(g(x))$ ,  $f(x) = o(g(x)) \implies \alpha f(x) = o(g(x))$ .
3.  $f(x) = O(g(x))$ ,  $h(x) = O(g(x)) \implies f(x) + h(x) = O(g(x))$ .  
 $f(x) = o(g(x))$ ,  $h(x) = o(g(x)) \implies f(x) + h(x) = o(g(x))$ .
4.  $f(x) = O(g(x))$ ,  $h(x) = O(k(x)) \implies f(x)h(x) = O(g(x)k(x))$ .  
 $f(x) = O(g(x))$ ,  $h(x) = o(k(x)) \implies f(x)h(x) = o(g(x)k(x))$ .
5.  $f(x) = O(g(x))$ ,  $g(x) = O(h(x)) \implies f(x) = O(h(x))$ .  
 $f(x) = O(g(x))$ ,  $g(x) = o(h(x)) \implies f(x) = o(h(x))$ .  
 $f(x) = o(g(x))$ ,  $g(x) = O(h(x)) \implies f(x) = o(h(x))$ .

**Theorem 2.1:** A function  $f : [a, b] \rightarrow \mathbb{R}$  is differentiable at  $c$  with derivative  $f'(c)$  if and only if

$$f(x) - (f(c) + f'(c)(x - c)) = o(|x - c|),$$

or alternatively,

$$f(x) = f(c) + f'(c)(x - c) + o(|x - c|).$$

Thus, saying a function is differentiable at  $c$  means that the error — the difference between  $f(x)$  and its linear approximation — goes to 0 faster than the distance between  $x$  and  $c$ .

Let us use the big  $O$  and little  $o$  notation to prove the chain rule.

**Theorem 2.2 (Chain Rule):** Let  $g : (a_1, b_1) \rightarrow \mathbb{R}$ ,  $f : (a_2, b_2) \rightarrow \mathbb{R}$  where  $(a_2, b_2) \supset g(a_1, b_1)$ . Suppose  $g$  is differentiable at  $c \in (a_1, b_1)$  and  $f$  is differentiable at  $g(c)$ , then  $h(x) = f(g(x))$  is differentiable at  $c$  with derivative  $f'(g(c))g'(c)$ .

*Proof.* As  $x$  approaches  $c$ ,

$$\begin{aligned} f(g(x)) &= f(g(c)) + f'(g(c))(g(x) - g(c)) + o(|g(x) - g(c)|) \\ &= f(g(c)) + f'(g(c))g'(c)(x - c) + f'(g(c))g'(c)o(|x - c|) + o(|g(x) - g(c)|) \\ &= f(g(c)) + f'(g(c))g'(c)(x - c) + o(|x - c|). \end{aligned}$$

Note that  $|g(x) - g(c)| = O(|x - c|)$ , and thus  $o(|g(x) - g(c)|) = o(|x - c|)$  by rule 5,  $f'(g(c))g'(c)o(|x - c|) = o(|x - c|)$  by rule 2, and the sum of them is still  $o(|x - c|)$  by rule 3.  $\square$

### 3 First Order Condition, Rolle's Theorem, Mean Value Theorem

**Theorem 3.1 (First Order Condition):** Let  $f : (a, b) \rightarrow \mathbb{R}$  be differentiable. If  $f$  attains a local maximum or minimum at some point  $c \in (a, b)$ , then  $f'(c) = 0$ .

*Proof.* We prove the case when  $c$  is a maximal point. Since  $f$  is differentiable at  $c$ ,

$$f(x) = f(c) + f'(c)(x - c) + r(x),$$

where  $r(x) = o(|x - c|)$  as  $x \rightarrow c$ . If  $f'(c) > 0$ , then there exists a small enough  $\delta$  such that for all  $|x - c| < \delta$ ,  $|r(x)| < \left| \frac{f'(c)}{2}(x - c) \right|$ . Hence,

$$\begin{aligned} f(x) &= f(c) + f'(c)(x - c) + r(x) \\ &\geq f(c) + f'(c)(x - c) - \left| \frac{f'(c)}{2} \right| |x - c| \\ &= f(c) + \frac{f'(c)}{2}(x - c) \end{aligned}$$

for all  $c < x < c + \delta$ . Therefore, for all  $c < x < c + \delta$ , we have  $f(x) > f(c)$ . This is a contradiction. The proof that  $f'(c)$  cannot be smaller than 0 is the same.  $\square$

Note that the First Order Condition is a **necessary** condition for a point to be a local maximal/minimal point. A function  $f$  can have a derivative equal to 0 at some point  $x$ , even if  $x$  is neither a maximum nor a minimum.

**Example 3.1:** Consider  $f(x) = x^3$ .  $f'(0) = 0$  but  $x = 0$  is neither a local maximal nor a local minimal point.

Suppose  $f : [a, b] \rightarrow \mathbb{R}$  and  $f(a) = f(b)$ , then  $f$  must attain a global maximum or minimum in  $(a, b)$ . This leads to Rolle's Theorem:

**Theorem 3.2 (Rolle's Theorem):** Let  $f : [a, b] \rightarrow \mathbb{R}$  be differentiable on  $(a, b)$ . If  $f(a) = f(b)$ , then there exists  $c \in (a, b)$  such that  $f'(c) = 0$ .

The generalization of Rolle's theorem is the Mean Value Theorem:

**Theorem 3.3 (Mean Value Theorem):** Let  $f : [a, b] \rightarrow \mathbb{R}$  be differentiable on  $(a, b)$ . Then there

exists  $c \in (a, b)$  such that

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

*Proof.* Write  $\theta = \frac{f(b)-f(a)}{b-a}$ . Let us consider the function:

$$h(x) := f(x) - \theta(x - a).$$

Note that  $h(a) = h(b) = f(a)$ . By Rolle's Theorem, there exists  $c \in (a, b)$  such that  $h'(c) = 0$ . But note that

$$h'(x) = f'(x) - \theta.$$

Therefore,  $h'(c) = 0 \implies f'(c) = \theta$ . □