MA2201: Analysis II

Differentiation

Spring 2021

Satvik Saha
19MS154

Indian Institute of Science Education and Research, Kolkata, Mohanpur, West Bengal, 741246, India.

The origins of differential calculus lie in our attempts to approximate various functions using linear ones. Suppose that we have been given a curve described by the function f, and we want to *locally* approximate the function around a point x using a straight line. In other words, for a small shift h, we want to write

$$f(x+h) \approx f(x) + kh.$$

Here, k is the slope of the straight line. In order to obtain k, we can rearrange the above to get

$$k \approx \frac{f(x+h) - f(x)}{h}.$$

As we pick smaller and smaller neighbourhoods of x, we want our approximation to get better and better. Thus, if such an approximation is possible, then the value of k must stabilize. This means that the limit

$$k = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

must exist. Note that this immediately forces the continuity of f, since

$$\lim_{h \to 0} f(x+h) - f(x) = \lim_{h \to 0} h \cdot \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = 0k = 0,$$

whereby $\lim_{x \to a} f(x) = f(a)$. Splitting the limit is justified because the individual limits exist. If such a limit k exists, we call it the derivative of f at x, denoted $f'(x)$. We are now able to write

$$f(x+h) \approx f(x) + f'(x)h.$$

Definition 2.1 (Derivative)**.** The derivative of a function $f\colon [a,b] \to \mathbb{R}$ at a point $x \in [a,b]$ is defined as

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

if such a limit exists. Note that we only demand a one-sided limit if x is an endpoint. If the derivative of f exists at every point in $[a,b]$, we say that f is differentiable on $[a,b]$.

Example. Consider the map $x \mapsto x^n$, where $n \in \mathbb{N}$. Using the binomial theorem, we can write

$$(x+h)^n = x^n + nx^{n-1}h + \cdots + h^n,$$

which means that

$$\frac{d}{dx}x^n = \lim_{h\to 0}\frac{1}{h}\left[(x+h)^n - x^n\right] = \lim_{h\to 0}\left[nx^{n-1} + \binom{n}{2}x^{n-2}h + \cdots + h^{n-1}\right] = nx^{n-1}.$$

Note that the process of differentiation we described can be generalised to multivariable functions. The idea is to locally approximate a function with an affine function.

Theorem 2.1. *If $f\colon (a,b) \to \mathbb{R}$ is differentiable on (a,b), then it is also continuous on (a,b).*

Theorem 2.2. *Let $f\colon I \to \mathbb{R}$ be a continuous function. Then,*

1. *f maps compact sets to compact sets.*
2. *f maps connected sets to connected sets.*

Corollary 2.2.1. *A continuous function $f\colon I \to \mathbb{R}$ maps intervals to intervals.*

Corollary 2.2.2. *A continuous function $f\colon [a,b] \to \mathbb{R}$ attains its minimum and maximum on $[a,b]$.*

Definition 2.2. Given $f\colon (a,b) \to \mathbb{R}$, a point $c \in (a,b)$ is said to be a point of local maximum if there exists a neighbourhood I_c of c such that

$$f(c) > f(x),$$

for all $x \in I_c \setminus \{c\}$. There is an analogous definition for a local minimum.

Theorem 2.3. *If $f\colon (a,b) \to \mathbb{R}$ is differentiable and $c \in (a,b)$ is a point of local minimum or maximum, then $f'(c) = 0$.*

Remark. The converse is not true. Note that the derivative of $x \mapsto x^3$ vanishes at $x = 0$, but that is not a local minimum or maximum.

Proof. Let c be a local minimum or maximum of f, but suppose that $f'(c) \neq 0$. Define the function

$$g\colon (a,b) \to \mathbb{R}, \qquad g(x) = \begin{cases} (f(x) - f(c))/(x-c), & \text{if } x \neq c \\ f'(c), & \text{if } x = c \end{cases}$$

We note that g is continuous. Also, $f'(c) = g(c) \neq 0$. If $g(c) > 0$, there exists a neighbourhood $I_\delta = (c-\delta, c+\delta)$ such that for all $x \in I_\delta$, $g(x) > 0$, from the continuity of g. This means that on I_c,

$$\frac{f(x) - f(c)}{x - c} > 0,$$

which gives $f(x) > f(c)$ on $(c, c+\delta)$ and $f(x) < f(c)$ on $(c-\delta, c)$. This means that c cannot be a local minimum, nor a local maximum. There is an analogous case assuming $g(c) < 0$, which leads to the same contradiction. Thus, we must have $f'(c) = g(c) = 0$. □

Theorem 2.4. *If $f\colon (a,b) \to \mathbb{R}$ is twice differentiable, and $c \in (a,b)$ is such that $f'(c) = 0$ and $f''(c) < 0$, then c is a point of local maximum. If $f'(c) = 0$ and $f''(c) > 0$, then c is a point of local minimum.*

Theorem 2.5 (Rolle's Theorem)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous, and differentiable on (a,b), with $f(a) = f(b)$. Then, there exists $c \in (a,b)$ such that $f'(c) = 0$.*

Proof. Set $f(a) = f(b) = \kappa$. From the continuity of f, note that the image of the closed interval $[a,b]$ is another closed interval $[\alpha, \beta]$. This means that $\alpha \leq \kappa \leq \beta$. Note that if $\alpha = \beta = \kappa$, then the function f is identically equal to the constant κ, hence $f'(x) = 0$ everywhere on $[a,b]$. By the continuity of f, it must attain its maximum and minimum on $[a,b]$. If $\beta > \kappa$, then the maximum is al least β and is hence not attained at the endpoints, which means that the point of maximum lies in (a,b). If $\alpha < \kappa$, then the same argument shows that f attains a minimum in (a,b). Thus, in either case, we have found $c \in (a,b)$ which is either a maximum or minimum of f, i.e. $f'(c) = 0$. □

Theorem 2.6 (Mean Value Theorem)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be continuous, and differentiable on (a,b). Then, there exists $c \in (a,b)$ such that*

$$f(b) - f(a) = f'(c)\,(b-a).$$

Proof. Apply Rolle's Theorem on the function defined as

$$g\colon [a,b] \to \mathbb{R}, \qquad g(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a} \cdot (x - a).$$

Note that g is continuous on $[a,b]$, differentiable on (a,b), and $g(a) = g(b) = 0$. □

Theorem 2.7. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable, and $f'(x) > 0$ for all $x \in \mathbb{R}$. Then, f is strictly increasing on $\mathbb{R}$.*

Proof. Let $x_2 > x_1$. By the mean value theorem, we pick $c \in (x_1, x_2)$ such that

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) > 0.$$

□

Remark. The converse is not true. The map $x \mapsto x^3$ is strictly increasing, but its derivative vanishes at 0.

Theorem 2.8 (Chain rule)**.** *Let f and g be differentiable on $\mathbb{R}$. Then, $f \circ g$ is also differentiable, with*
$$(f \circ g)' = (f' \circ g) \cdot g'.$$

Proof. Fix $a \in \mathbb{R}$. Define the functions

$$\varphi \colon \mathbb{R} \to \mathbb{R}, \qquad \varphi(x) = \begin{cases} (g(x) - g(a))/(x-a) & \text{if } x \neq a \\ g'(a), & \text{if } x = a \end{cases},$$

$$\psi \colon \mathbb{R} \to \mathbb{R}, \qquad \psi(y) = \begin{cases} (f(y) - f(b))/(y-b) & \text{if } y \neq b \\ f'(b), & \text{if } y = b \end{cases}.$$

Note that φ and ψ are continuous. Also, when $x \neq a$, we have

$$g(x) - g(a) = \varphi(x)(x-a).$$

Set $b = g(a)$, and write

$$f(g(x)) - f(g(a)) = \psi(g(x))(g(x) - g(a)) = \psi(g(x))\,\varphi(x)(x-a).$$

Setting $h = f \circ g$, we have
$$\frac{h(x) - h(a)}{x - a} = \psi(g(x))\varphi(x).$$

Taking limits $x \to a$, we use the continuity of φ, ψ and g to conclude that the derivative of h is indeed defined at a, and

$$h'(a) = \psi(g(a))\,\varphi(a) = f'(g(a))\,g'(a).$$ □

Definition 2.3 (Intermediate Value Property)**.** Let $f \colon (a,b) \to \mathbb{R}$ be such that for all $c, d \in (a,b)$ such that $f(c) < f(d)$ and $\lambda \in (f(c), f(d))$, there exists $x_0 \in (a,b)$ such that $f(x_0) = \lambda$. Then, we say that f has the intermediate value property.

Theorem 2.9 (Intermediate Value Theorem)**.** *All continuous functions $f \colon (a,b) \to \mathbb{R}$ have the intermediate value property.*

Theorem 2.10. *Let $f \colon (a,b) \to \mathbb{R}$ be differentiable. Then, f' satisfies the intermediate value property.*

Proof. Let $c, d \in (a,b)$ and let $\lambda \in \mathbb{R}$ such that $\lambda \in (f'(c), f'(d))$. We wish to find $x_0 \in (a,b)$ such that $f'(x_0) = \lambda$. Define

$$g \colon (a,b) \to \mathbb{R}, \qquad g(x) = f(x) - \lambda x.$$

Note that $g'(x) = f'(x) - \lambda$, so $g'(c) < 0$ and $g'(d) > 0$. Thus, g is decreasing near c and increasing near d, so we can find $t_1, t_2 \in (c,d)$ such that $g(t_1) < g(c)$ and $g(t_2) < g(d)$. This means that g has no local minimum at c nor d. From the continuity of g, there exists $x_0 \in [c,d]$ such that $g(x_0) = \inf_{[c,d]} g(x)$. We have already shown that x_0 is neither c, nor d, so $x_0 \in (c,d)$. Hence, $g'(x_0) = 0$, which gives $f'(x_0) = \lambda$. □

Lemma 2.11. *If $f\colon (a,b) \to (c,d)$ is surjective, continuous and strictly increasing, then f is invertible with a continuous inverse.*

Theorem 2.12 (Inverse function theorem)**.** *Let $f\colon (a,b) \to (c,d)$ be surjective and differentiable, with $f'(x) \neq 0$ everywhere. Then, f is invertible, with a differentiable inverse whose derivative is given by*

$$(f^{-1})'(f(x)) = \frac{1}{f'(x)}.$$

Proof. Given $f'(x) \neq 0$ on (a,b). Then intermediate value property gives either $f'(x) > 0$ for all $x \in (a,b)$, or $f'(x) < 0$. Without loss of generality, assume the former. This means that f is strictly increasing on (a,b), continuous, and surjective. Our lemma gives the existence of a continuous inverse f^{-1}.

Let $y \in (c,d)$, and let $x = f^{-1}(y)$. From the continuity of f^{-1}, we can always write $f^{-1}(y+\kappa) = x + h$. Thus,

$$\lim_{\kappa\to 0} \frac{f^{-1}(y+\kappa) - f^{-1}(y)}{\kappa} = \lim_{\kappa\to 0} \frac{x+h-x}{\kappa} = \lim_{\kappa\to 0} \frac{h}{\kappa}.$$

Note that $h \to 0$ as $\kappa \to 0$. Thus, this limit can be written as

$$(f^{-1})'(y) = \lim_{h\to 0} \frac{h}{f(x+h) - f(x)} = \frac{1}{f'(x)}.$$

□

Corollary 2.12.1. *Let f be continuously differentiable on $\mathbb{R}$, with $f'(x_0) \neq 0$ for some $x_0 \in \mathbb{R}$. Then, there exists some neighbourhood of x_0 on which f is invertible, with a continuously differentiable inverse.*

Theorem 2.13. *Let $f_n \to f$ pointwise and $\{f_n'\}$ converge uniformly on some interval $[a,b]$. Then, $f_n \to f$ uniformly.*

Proof. We show that $\{f_n\}$ is uniformly Cauchy on E. Note that for some fixed t, we can write

$$|f_n(x) - f_m(x)| \leq |(f_n(x) - f_m(x)) - (f_n(t) - f_m(t))| + |f_n(t) - f_m(t)|.$$

Using the Mean Value Theorem, the first term can be bounded as

$$|(f_n(x) - f_m(x)) - (f_n(t) - f_m(t))| = (f_n' - f_m')(x_0)|x - t|,$$

where x_0 is between x and t. From the pointwise convergence of $f_n \to f$, we have

$$|f_n(t) - f_m(t)| < \frac{\epsilon}{2}$$

for all $n, m \geq N_t$. The uniform convergence of $\{f_n'\}$ means that we can find N_0 such that

$$|f_n'(x_0) - f_m'(x_0)| < \frac{\epsilon}{2(b-a)}$$

for all $n, m > N_0$. Thus, for all $x \in [a,b]$, and $n, m \geq N_t + N_0$, we have

$$|f_n(x) - f_m(x)| < \frac{\epsilon}{2(b-a)} \cdot (b-a) + \frac{\epsilon}{2} = \epsilon.$$

This means that $\{f_n\}$ is uniformly Cauchy on $[a,b]$, which gives the uniform convergence of $\{f_n\}$.

□

Theorem 2.14. *Let $\{f_n\}$ be a sequence of differentiable functions on some bounded interval (a, b) such that $f_n \to f$ pointwise and f_n' converges uniformly. Then, f is differentiable and $f_n' \to f'$.*