# Real Analysis 5: Differentiation

Brad Baxter
Birkbeck College, University of London

June 8, 2023

You can download these slides and the lecture videos from my office server

`http://econ109.econ.bbk.ac.uk/brad/analysis/`

**Recommended books:** Lara Alcock (2014), "How to Think about Analysis", Oxford University Press.
J. C. Burkill (1978), "A First Course in Mathematical Analysis", Cambridge University Press.

## Definition

*Let* $f : [a, b] \to \mathbb{R}$. *We say that* $f$ *is* differentiable *at* $c \in (a, b)$ *if the limit*

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h}$$

*exists, and if so we write*

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

Useful alternative form: $f$ is differentiable at $c$ if there exists a number $f'(c) \in \mathbb{R}$ for which

$$f(c+h) = f(c) + f'(c)h + e(h),$$

where the **error** $e(h) \to 0$ as $h \to 0$.
A differentiable function is one that is locally almost linear.

## Example (A differentiable function is continuous)

Given any $\epsilon > 0$, choose $\delta > 0$ so small that $|e(h)| < \epsilon/2$ and $\delta < \epsilon/(2|f'(c)|)$. Then, for $|h| < \delta$,

$$\begin{aligned} |f(c+h) - f(c)| &= \left|f'(c)h + e(h)\right| \\ &\le \left|f'(c)\right| |h| + |e(h)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

## Theorem (Vector space of differentiable functions)

*Let $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ be differentiable functions. Any linear combination $\alpha f + \beta g$ is also differentiable.*

## Theorem (Product Rule)

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

## Proof.

$$\frac{f(x+h)g(x+h) - f(x)g(x)}{h}$$
$$= \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h}$$
$$= \left(\frac{f(x+h) - f(x)}{h}\right) g(x+h) + f(x)\left(\frac{g(x+h) - g(x)}{h}\right)$$
$$\to f'(x)g(x) + f(x)g'(x),$$

as $h \to 0$. □

## Theorem (Chain rule)

$$\frac{d}{dx} f\left(g(x)\right) = f'(g(x))g'(x).$$

Proof: We know that

$$g(x+h) = g(x) + hg'(x) + e_g(h),$$

where $e_g(h) \to 0$ as $h \to 0$, and

$$f(y+k) = f(y) + f'(y)k + e_f(k),$$

where $e_f(k) \to 0$ as $k \to 0$. Hence, setting $y = g(x)$ and

$$k = hg'(x) + e_g(h)$$

we obtain

$$\begin{aligned} f(g(x+h)) &= f(y+k) \\ &= f(y) + kf'(y) + e_f(k) \\ &= f(g(x)) + \left(hg'(x) + e_g(h)\right) f'(g(x)) + e_f(k) \\ &= f(g(x)) + hg'(x)f'(g(x)) + E(h,k). \end{aligned}$$

Here

$$E(h,k) = e_g(h)f'(g(x)) + e_f(k).$$

Now, as $h \to 0$, we also find that $k \to 0$, and hence $E(h,k) \to 0$. Hence $f(g(x))$ is differentiable, with derivative

$$\frac{d}{dx} f\left(g(x)\right) = f'(g(x))g'(x).$$

Traditional calculus notation: If $z = f(y)$ and $y = g(x)$, then

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

## Example (Chain Rule 1)

If $f(y) = 1/y$, then $f'(y) = -1/y^2$ and

$$\frac{d}{dx}\frac{1}{g(x)} = -\frac{g'(x)}{g(x)^2}.$$

## Example (Chain Rule 2)

If $f(y) = \exp(y)$ and $g(x) = -x^2/2$, then

$$f(g(x)) = e^{-x^2/2}$$

and its derivative is given by

$$f'(g(x))g'(x) = -e^{-x^2/2} \cdot x.$$

## Example (Chain Rule 3)

Let $y = f(x)$, for $x \in \mathbb{R}$, and suppose that the inverse function $g(y) = x$ exists. Thus $g(f(x)) = x$ and the chain rule implies

$$g'(y)f'(x) = 1,$$

i.e.

$$g'(y) = \frac{1}{f'(g(y))},$$

for $y \in \{f(x) : x \in \mathbb{R}\}$.

## Exercise

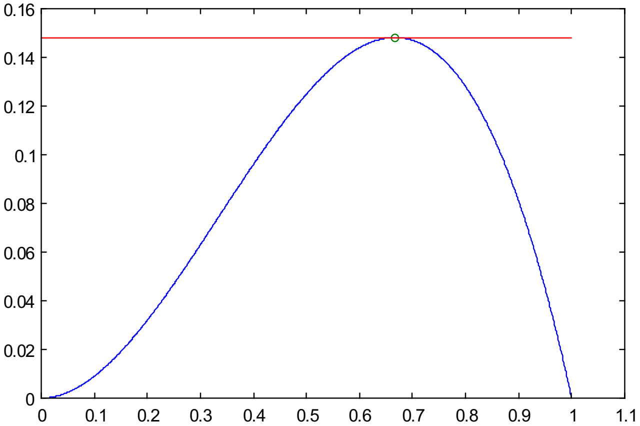
*Show that, if $f'(x) = f(x)$, for all $x \in \mathbb{R}$, then the inverse function $g(y)$ satisfies*

$$g'(y) = \frac{1}{y},$$

*for $y \in \{f(x) : x \in \mathbb{R}\}$.*

## Theorem (Rolle's Theorem)

*Let $f : [a, b] \to \mathbb{R}$ be a continuous function which is differentiable on $(a, b)$. If $f(a) = f(b) = 0$, and if $f$ is not identically zero, then there exists $c \in (a, b)$ such that $f'(c) = 0$.*



Proof:

$f$ is continuous, so it's bounded and attains its bounds. Suppose $f(c) = \sup_{[a,b]} f$ and $c \in (a, b)$. Hence

$$f(c) \geq f(x), \qquad \forall x \in [a, b].$$

If $f'(c) > 0$, then

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

so $f(c+h) - f(c) > 0$ for all sufficiently small $h > 0$, contradicting $f(c) = \sup_{[a,b]} f$.

Similarly, if $f'(c) < 0$, then

$$0 > f'(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}.$$

Thus, if $h < 0$ and $|h|$ is sufficiently small, then

$$\frac{f(c+h) - f(c)}{h} < 0,$$

i.e.

$$f(c+h) - f(c) > 0,$$

on multiplying the inequality by the negative number $h$. This again contradicts $f(c) = \sup_{[a,b]} f$. The only remaining case is $f'(c) = 0$. $\square$

## Theorem (Mean Value Theorem)

*Let* $f : [a, b] \to \mathbb{R}$ *be a continuous function which is differentiable on* $(a, b)$*. Then there exists* $c \in (a, b)$ *such that*

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

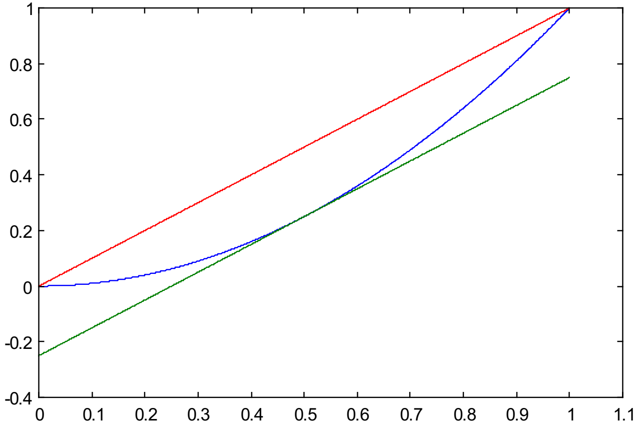



Figure: Mean Value Theorem for $f(x) = x^2$ on $[0, 1]$

Proof:

Let

$$L(x) = f(a) + \left(\frac{f(b) - f(a)}{b - a}\right)(x - a),$$

for $x \in [a, b]$, $L(x)$ is the linear function agreeing with $f$ at the end-points $a$ and $b$ of the interval. Further

$$L'(x) \equiv \frac{f(b) - f(a)}{b - a}, \quad \text{for all } x \in [a, b].$$

Define the **error**

$$E(x) = f(x) - L(x).$$

Then $E(a) = E(b) = 0$, $E\colon [a, b] \to \mathbb{R}$ is continuous, and $E$ is differentiable on $(a, b)$. Applying Rolle's theorem, we deduce the existence of at least one point $c \in (a, b)$ for which $E'(c) = 0$, that is

$$f'(c) - L'(c) = 0,$$

or

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

**Example (Strictly increasing functions never take the same value twice)**

Suppose $f'(x) > 0$ for all $x \in (a, b)$. If $a < x < y < b$, then the MVT tells us that, for some $z \in [x, y]$, we have

$$\frac{f(y) - f(x)}{y - x} = f'(z) > 0.$$

Hence $f(y) > f(x)$.

Another way to restate the Mean Value Theorem is as follows:
$\forall h > 0$, $\exists \theta \in (0,1)$ such that

$$f(a+h) = f(a) + hf'(a+\theta h).$$

**Example (Important MVT example)**

Apply the Mean Value Theorem to $f(x) = \sin x$:

$$\frac{f(x) - f(0)}{x - 0} = f'(\alpha),$$

for some $\alpha \in [0, x]$ depending on $x$, i.e.

$$\frac{\sin x - \sin 0}{x - 0} = \frac{\sin x}{x} = \cos \alpha.$$

Now $x \to 0$ implies $\alpha \to 0$. Hence

$$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{\alpha \to 0} \cos \alpha = 1.$$

## Definition (Taylor's series)

*Let $f, f', f^{(2)}, \ldots, f^{(n)}$ all exist and be continuous in $[a, b]$. The polynomial*

$$p_n(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (x - a)^k, \qquad x \in \mathbb{R},$$

*is called the n-th* Taylor partial sum based at $x = a$*, and is used to approximate the function $f(x)$. Further,*

$$p_n^{(k)}(a) = f^{(k)}(a), \qquad \text{for } 0 \le k \le n - 1.$$

## Example (Taylor series good locally but not globally)

Let $f(x) = \cos x$ and $a = 0$. Then

$$p_5(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!}.$$

$p_5(x)$ is an excellent approximation to $\cos x$ for small $|x|$:

$$\cos 0.1 = 0.995004165278026 \text{ and } p_5(0.1) = 0.995004166666667.$$

It's not good for larger $|x|$:

$$\cos 1 = 0.540302305868140 \text{ but } p_5(1) = 0.541666666666667.$$

It's useless for big $|x|$:

$$\cos 10 = -0.839071529076452 \text{ but } p_5(10) = 367.666666666667.$$

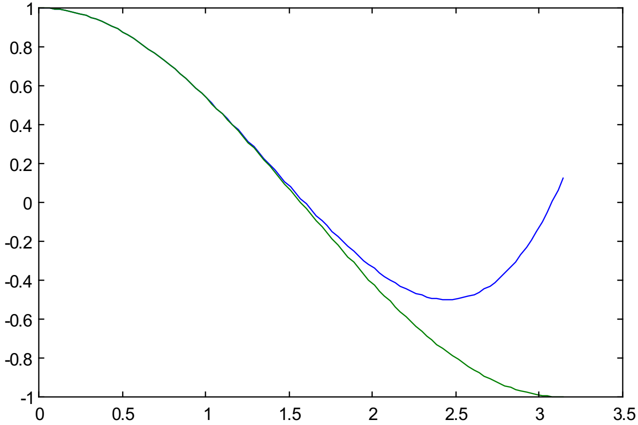



Figure: $p_5(x) = 1 - x^2/2! + x^4/4!$ and $\cos x$ on $[0, \pi]$

## Theorem (Taylor's theorem)

*Let $f, f', f^{(2)}, \ldots, f^{(n)}$ be continuous on $[a, b]$. Then, for any $x \in [a, b]$, we have*

$$\begin{aligned} f(x) &= \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{(x-a)^n}{n!} f^{(n)}(a+\theta h) \\ &= p_n(x) + e_n(x). \end{aligned} \tag{0.1}$$

*for some $\theta \in (0,1)$. Further, if $|f^{(n)}(x)| \le M$ for all $x \in [a, b]$, then*

$$|e_n(x)| \le \frac{(b-a)^n}{n!} M, \qquad \textit{for all } x \in [a, b].$$

## Example (Taylor error bounds for cosine)

Returning to $f(x) = \cos x$, its higher derivatives are either $\pm \cos x$ or $\pm \sin x$. Thus $|f^{(n)}(x)| \le 1$, for any positive integer $n$ and any $x \in \mathbb{R}$, and the error in Taylor's theorem satisfies

$$|e_n(x)| \le \frac{x^n}{n!},$$

for any positive $x$. In particular,

$$|e_5(0.1)| \le 10^{-5}/5! = 8.33333\cdots \times 10^{-8}.$$

On the other hand,

$$|e_5(1)| \le 1/5! = 0.0083333\cdots,$$

and

$$|e_5(10)| \le 10^5/5! = 833.333\cdots.$$

## Proof of Taylor's theorem:

The $n$-th Taylor partial sum satisfies

$$p_n^{(j)}(a) = f^{(j)}(a), \text{ for } j = 0, 1, \ldots, n-1.$$

Key trick: Define

$$\phi(x) = f(x) - p_n(x) - A(x-a)^n/n!, \qquad \text{for } x \in [a, a+h],$$

where $0 < h < b - a$, and choose $A$ so that $\phi(a+h) = 0$, i.e. solve

$$0 = f(a+h) - p_n(a+h) - Ah^n/n!.$$

Thus $\phi(a) = \phi(a+h) = 0$, so (Rolle) there exists $h_1 \in (0, h)$ for which $\phi'(a+h_1) = 0$. Now

$$\frac{d^j}{dx^j}(x-a)^n = n(n-1)(n-2)\cdots(n-j+1)(x-a)^{n-j},$$

and therefore vanishes when $x = a$ and $j = 0, 1, \ldots, n-1$. Hence we have

$$\phi'(a) = \phi'(a+h_1) = 0.$$

Applying Rolle to

$$\phi'(a) = \phi'(a + h_1) = 0,$$

there exists $h_2 \in (0, h_1)$ for which

$$\phi^{(2)}(a) = \phi^{(2)}(a + h_2) = 0.$$

Repeatedly applying Rolle's theorem, we find

$$\phi^{(n-1)}(a) = \phi^{(n-1)}(a + h_{n-1}) = 0,$$

where $0 < h_{n-1} < h_{n-2} < \cdots < h_2 < h_1 < h$. Applying Rolle one last time yields $h_n \in (0, h_{n-1})$ such that

$$\phi^{(n)}(a + h_n) = 0.$$

But (Exercise)

$$\phi^{(n)}(x) = f^{(n)}(x) - A.$$

Hence

$$A = f^{(n)}(a + h_n), \quad \text{where } 0 < h_n < h.$$

Then

$$e_n(a + h) = f^{(n)}(a + \theta h) h^n / n!, \quad \text{where } h_n = \theta h. \quad \square$$

Differentiable functions are extremely useful but they're a set of measure zero in continuous functions.

**Theorem (Nonexaminable:)**

*The function*

$$f(x) = \sum_{n=0}^{\infty} \frac{\sin\left[(n!)^2 x\right]}{n!}, \qquad x \in \mathbb{R},$$

*is continuous but nowhere differentiable.*