# MA 109 Week 3

Mayukh Mukherjee

Department of Mathematics

December 13, 2021

# Problems centered around Rolle's Theorem

Exercise 3.3: Let $f : [a, b] \to \mathbb{R}$ be continuous and suppose $f$ is differentiable on $(a, b)$. If $f(a)$ and $f(b)$ are of opposite signs and $f'(x) \neq 0$ for all $x \in (a, b)$, then there is a unique point $x_0$ in $(a, b)$ such that $f(x_0) = 0$.

Solution: Since the Intermediate Value Theorem guarantees the existence of a point $x_0$ such that $f(x_0) = 0$, the real point of this exercise is the uniqueness.

Suppose there were two points $x_1, x_2 \in (a, b)$ such that $f(x_1) = f(x_2) = 0$. Applying Rolle's Theorem, we see that there would exist $c \in (x_1, x_2)$ such that $f'(c) = 0$ contradicting our hypothesis, This proves the exercise.

Let us look at Exercise 2.8(i): Find a function f which satisfies all the given conditions, or else show that no such function exists: $f''(x) > 0$ for all $x \in \mathbb{R}$ and $f'(0) = 1, f'(1) = 1$.

Solution: Apply Rolle's Theorem to $f'(x)$ to conclude that such a function cannot exist.

# The Mean Value Theorem

Rolle's theorem is a special case of the Mean Value Theorem (MVT).

**Theorem**

*Suppose that $f : [a, b] \to \mathbb{R}$ is a continuous function and that $f$ is differentiable in $(a, b)$. Then there is a point $x_0$ in $(a, b)$ such that*

$$\frac{f(b) - f(a)}{b - a} = f'(x_0).$$

Proof: Apply Rolle's Theorem to the function

$$g(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a} \cdot (x - a).$$

□

(Why does one think of the function $g(x)$?)

# Applications of the MVT

Here is an application of the MVT which you have probably always taken for granted:

**Theorem**

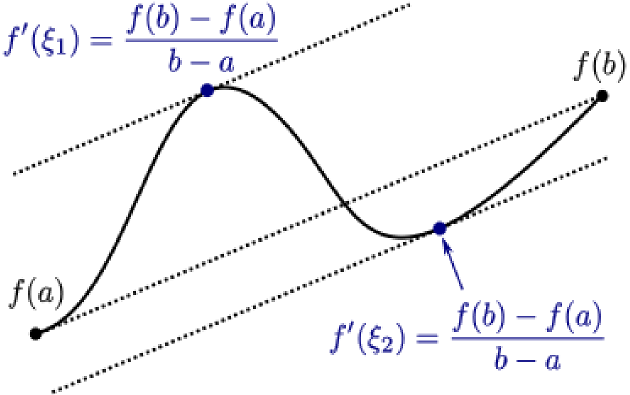
*If $f$ satisfies the hypotheses of the MVT, and further $f'(x) = 0$ for every $x \in (a, b)$, $f$ is a constant function.*

Indeed, if $f(c) \neq f(d)$ for some two points $c < d$ in $[a, b]$,

$$0 \neq \frac{f(d) - f(c)}{d - c} = f'(x_0),$$

for some $x_0 \in (c, d)$, by the MVT. This contradicts the hypothesis. $\square$

# Pictorial representation of MVT



# Applications of the MVT continued

Consider Exercise 2.6.:

Let $f$ be continuous on $[a, b]$ and differentiable on (a,b). If $f(a) = a$ and $f(b) = b$, show that there exist distinct $c_1, c_2 \in (a, b)$ such that $f'(c_1) + f'(c_2) = 2$.

Solution: The idea is that the function clearly has an average rate of growth equal to 1 on the interval $[a, b]$. If the derivative at some point is less than 1, there must be another point where it is greater than 1 so that the sum adds up to 2. How to use this idea?

Split the interval into two pieces - $[a, \frac{a+b}{2}]$ and $[\frac{a+b}{2}, b]$ - and apply the MVT to each interval.

# Darboux's Theorem

Another interesting property of differentiable functions is that their derivatives have the IVP (intermediate value property). This fact is sometimes called Darboux's Theorem.

**Theorem**

*Let $f : (a, b) \to \mathbb{R}$ be a differentiable function. If $c, d$, $c < d$ are points in $(a, b)$, then for every $u$ between $f'(c)$ and $f'(d)$, there exists an $x$ in $[c, d]$ such that $f'(x) = u$.*

Proof: We can assume, without loss of generality, that $f'(c) < u < f'(d)$, otherwise we can take $x = c$ or $x = d$. Define $g(t) = ut - f(t)$. This is a continuous function on $[c, d]$, and hence must attain its extreme values. The maximum value cannot occur at $c$ or $d$ since $g'(c) = u - f'(c) > 0$ and $g'(d) = u - f'(d) < 0$.

Suppose $g$ takes a maximum at $c$. Since $g'(c) > 0$, for $h > 0$ small enough, we must have $g(c+h) - g(c) > 0$, contradiction.

Suppose $g$ takes a maximum at $d$. Since $g'(d) < 0$, for $h < 0$ small enough, we must have $g(d+h) - g(d) > 0$, contradiction.

It follows that there exists $x \in (c, d)$ where $g$ takes a maximum. By Fermat's Theorem $g'(x) = 0$ which yields $f'(x) = u$. $\square$

# Continuity of the first derivative

We have just seen that the derivative satisfies the IVP. Can we find a function which is differentiable but for which the derivative is not continuous?

Here is the standard example:

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & \text{if } x \neq 0 \\ = 0 & \text{if } x = 0. \end{cases}$$

This function will be differentiable at 0 but its derivative will not be continuous at that point. In order to see this you will need to study the function in Exercise 2.3(ii). This will show that $f'(0) = 0$. On the other hand, if we use the product rule when $x \neq 0$ we get

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x},$$

which does not go to 0 as $x \to 0$.

# Back to maxima and minima

We will assume that $f : [a, b] \to \mathbb{R}$ is a continuous function and that $f$ is differentiable on $(a, b)$. A point $x_0$ in $(a, b)$ such that $f'(x_0) = 0$ often called a stationary point. We will assume further that $f'(x)$ is differentiable at $x_0$, that is, that the second derivative $f''(x_0)$ exists. We formulate the Second Derivative Test below.

**Theorem**

*With the assumptions above:*

1. *If $f''(x_0) > 0$, the function has a local minimum at $x_0$.*
2. *If $f''(x_0) < 0$, the function has a local maximum at $x_0$.*
3. *If $f''(x_0) = 0$, no conclusion can be drawn.*

# The proof of the Second Derivative Test

Proof: The proofs are straightforward. For instance, to prove the first part we observe that

$$0 < f''(x_0) = \lim_{h \to 0} \frac{f'(x_0 + h)}{h}.$$

It follows that for $|h|$ small enough, $f'(x_0 + h) < 0$, if $h < 0$ and $f'(x_0 + h) > 0$ if $h > 0$. It follows that $f(x_0)$ is decreasing to the left of $x_0$ and increasing to the right of $x_0$. Hence, $x_0$ must be a local minimum. A similar argument yields the second case. $\square$

If the third case of the theorem above occurs, the function may be changing from concave to convex. In this case $x_0$ is called a point of inflection. An example of this phenomenon is given by $f(x) = x^3$ at $x = 0$.

# Concavity and convexity

Let $I$ denote an interval (open or closed or half-open).
Definition: A function $f : I \to \mathbb{R}$ is said to be concave (or sometimes concave downwards) if

$$f(tx_1 + (1-t)x_2) \geq tf(x_1) + (1-t)f(x_2)$$

for all $x_1$ and $x_2$ in $I$ and $t \in [0,1]$. Similarly, a function is said to be convex (or concave upwards if

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

By replacing the $\geq$ and $\leq$ signs above by strict inequalities we can define strictly concave and strictly convex functions.

Note that if $f(x)$ is a concave function, $-f(x)$ is a convex function, so it is really enough to study one class or the other. By some convention which has cemented over time, in Mathematics one talks much more about convex functions than concave.

# Examples of concave and convex functions

Here are some examples of convex functions.

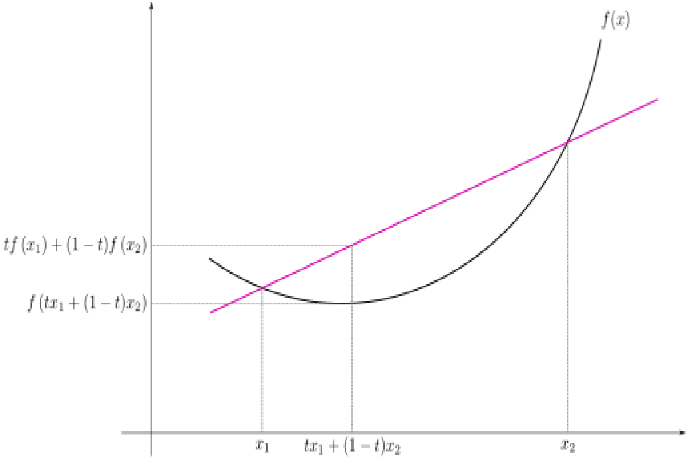
1. $f(x) = x^2$ on $\mathbb{R}$.
2. $f(x) = x^3$ on $[0, \infty)$.
3. $f(x) = e^x$ on $\mathbb{R}$.

Examples of concave functions include

1. $f(x) = -x^2$
2. $f(x) = x^3$ on $(-\infty, 0]$
3. $f(x) = \log x$ on $(0, \infty)$.

For a convex function $f$ and point $c \in (x_1, x_2)$, the point $(c, f(c))$ always lies below the line joining $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

# Convexity illustrated graphically





http://en.wikipedia.org/wiki/File:ConvexFunction.svg

# Propeties of Convex functions

Convex functions have many nice properties. For instance, it is easy to show that convex functions are continuous (do this!). More is true.

Exercise 1. Every convex function is Lipschitz continuous (a function is Lipschitz continuous if it satisfies the inequality given in Exercise 2.6 but with $\alpha = 1$). In fact, much more is true. A convex function is actually differentiable at all but at most countably many points.

A differentiable function is convex if and only if its derivative is monotonically increasing. Moreover, if a function is both differentiable and convex, it is continuously differentiable, that is, its derivative is continuous (feel free to try proving these facts).

# Convexity and the second derivative

A twice differentiable function on an interval will be convex if its second derivative is everywhere non-negative. If the second derivative is positive, the function will be strictly convex.

However, the converse of the second statement above is not true. Can you give a counter-example to the converse of the second statement?

How about $f(x) = x^4$?

Definition: A point of inflection $x_0$ for a function $f$ is a point where the function changes its behavior from concave to convex (or vice-versa). At such a point $f''(x_0) = 0$, but this is only a necessary, not a sufficient condition.(Why?) If further, we also assume that the lowest order ($\geq 2$) non-zero derivative is odd, then we get a sufficient condition.

# Concavity and convexity

Let $I$ denote an interval (open or closed or half-open).
Definition: A function $f : I \to \mathbb{R}$ is said to be concave (or sometimes concave downwards) if

$$f(tx_1 + (1-t)x_2) \geq tf(x_1) + (1-t)f(x_2)$$

for all $x_1$ and $x_2$ in $I$ and $t \in [0,1]$. Similarly, a function is said to be convex (or concave upwards) if

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

By replacing the $\geq$ and $\leq$ signs above by strict inequalities we can define strictly concave and strictly convex functions.

For various reasons, convex functions are more important in mathematics than concave functions and for this reason we will concentrate on the former rather than the latter. On the other hand, note that if $f(x)$ is a concave function, $-f(x)$ is a convex function, so it is really enough to study one class or the other.

# Examples of concave and convex functions

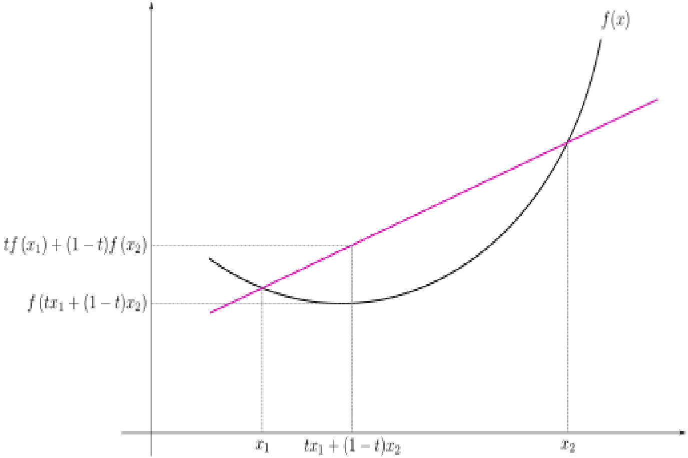
Here are some examples of convex functions.

1. $f(x) = x^2$ on $\mathbb{R}$.
2. $f(x) = x^3$ on $[0, \infty)$.
3. $f(x) = e^x$ on $\mathbb{R}$.

Examples of concave functions include

1. $f(x) = -x^2$
2. $f(x) = x^3$ on $(-\infty, 0]$
3. $f(x) = \log x$ on $(0, \infty)$.

For a convex function $f$ and point $c \in (x_1, x_2)$, the point $(c, f(c))$ always lies below the line joining $(x_1, f(x_1))$ and $(x_2, f(x_2))$.


$f(x)$
$tf(x_1) + (1-t)f(x_2)$
$f(tx_1 + (1-t)x_2)$
$x_1$
$tx_1 + (1-t)x_2$
$x_2$

# Propeties of Convex functions

Convex functions have many nice properties. For instance, it is easy to show that convex functions are continuous (do this!). More is true.

Exercise 1. Every convex function is Lipschitz continuous (a function is Lipschitz continuous if it satisfies the inequality given in Exercise 1.16 but with $\alpha = 1$). In fact, much more is true. A convex function is actually differentiable at all but at most countably many points.

A differentiable function is convex if and only if its derivative is monotonically increasing. Moreover, if a function is both differentiable and convex, it is continuously differentiable, that is, its derivative is continuous (feel free to try proving these facts).

# Convexity and the second derivative

A twice differentiable function on an interval will be convex if its second derivative is everywhere non-negative. If the second derivative is positive, the function will be strictly convex.

However, the converse of the second statement above is not true. Can you give a counter-example to the converse of the second statement?

How about $f(x) = x^4$?

Definition: A point of inflection $x_0$ for a function $f$ is a point where the function changes its behavior from concave to convex (or vice-versa). At such a point, if the function is twice differentiable, $f''(x_0) = 0$, but this is only a necessary, not a sufficient condition.(Why?) If further, we also assume that the lowest order ($\geq 2$) non-zero derivative is odd, then we get a sufficient condition.

We will now introduce some notation. The space $\mathcal{C}^k(I)$, will denote the space of $k$ times continuously differentiable functions on an (open) interval $I$, for some fixed $k \in \mathbb{N}$, that is, the space of functions for which $k$ derivatives exist and such that the $k$-th derivative is a continuous functions.

The space $\mathcal{C}^\infty(I)$ will consist of functions that lie in $\mathcal{C}^k(I)$ for every $k \in \mathbb{N}$. Such functions are called smooth or infinitely differentiable functions.

From now on we will denote the $k$-th derivative of a function $f(x)$ by $f^{(k)}(x)$.

Our aim will be to enlarge the class of functions we understand using the polynomials as stepping stones.

# The Taylor polynomials

Given a function $f(x)$ which is $n$ times differentiable at some point $x_0$ in an interval $I$, we can associate to it a family of polynomials $P_0(x), P_1(x), \dots P_n(x)$ called the Taylor polynomials of degrees $0, 1, \dots n$ at $x_0$ as follows.

We let $P_0(x) = f(x_0)$,

$$P_1(x) = f(x_0) + f^{(1)}(x_0)(x - x_0),$$

$$P_2(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{f^{(2)}(x_0)}{2!}(x - x_0)^2$$

We can continue in this way to define

$$P_n(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{f^{(2)}(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n.$$

# Taylor's Theorem

The Taylor polynomials are rigged exactly so that the degree $n$ Talyor polynomial has the same first $n$ derivatives at the point $x_0$ as the function $f(x)$ has, that is, $P^{(k)}(x_0) = f^{(k)}(x_0)$ for all $0 \leq k \leq n$, where $f^{(0)} = f(x)$ by convention.

Taylor's Theorem says that we can recover a lot of information about the function from the Taylor polynomials.

## Theorem (Taylor expansion)

*Let $I$ be an open interval and suppose that $[a, b] \subset I$. Suppose that $f \in \mathcal{C}^n(I)$ $(n \geq 0)$ and suppose that $f^{(n)}$ is differentiable on $I$. Then there exists $c \in (a, b)$ such that*

$$f(b) = P_n(b) + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1},$$

*where $P_n(x)$ denotes the Taylor polynomial of degree $n$ at $a$.*

# The proof of Taylor's theorem

Proof: From the definition, we see that

$$P_n(b) = f(a) + f^{(1)}(a)(b-a) + \cdots \frac{f^{(n)}(a)}{n!}(b-a)^n$$

Consider the function

$$F(x) = f(b) - f(x) - f^{(1)}(x)(b-x) - \frac{f^{(2)}(x)}{2!}(b-x)^2 - \ldots - \frac{f^{(n)}(x)}{n!}(b-x)^n.$$

Clearly $F(b) = 0$, and

$$F^{(1)}(x) = -\frac{f^{(n+1)}(x)(b-x)^n}{n!}. \tag{1}$$

We would like to apply Rolle's Theorem here, but $F(a) \neq 0$. So consider

$$g(x) = F(x) - \left(\frac{b-x}{b-a}\right)^{n+1} F(a)$$

(this is similar to the method by which we reduced the MVT to Rolle's Theorem), and we see that $g(a) = 0$. Applying Rolle's Theorem we see that there is a $c \in (a, b)$ such that $g'(c) = 0$.

This yields

$$F^{(1)}(c) = -(n+1)\left[\frac{(b-c)^n}{(b-a)^{n+1}}\right] F(a). \tag{2}$$

We can eliminate $F^{(1)}(c)$ using (1). This gives

$$-(n+1)\left[\frac{(b-c)^n}{(b-a)^{n+1}}\right] F(a) = -\frac{f^{(n+1)}(c)(b-c)^n}{n!},$$

from which we obtain

$$F(a) = \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c).$$

This proves what we want. $\square$

# Remarks on Taylor's Theorem and some examples

Remark 1: When $n = 0$ in Taylor's Theorem we get the MVT. When $n = 1$, Taylor's Theorem is called the Extended Mean Value Theorem.

Remark 2: The Taylor polynomials are nothing but the partial sums of the Taylor Series associated to a $\mathcal{C}^\infty$ function about (or at) the point $a$:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(b-a)^k.$$

We can show that this series converges provided we know that the difference $f(x) - P_n(x) = R_n(x)$ can be made less than any $\epsilon > 0$ when $n$ is sufficiently large. We will see how to do this for certain simple functions like $e^x$ or $\sin x$.

# The Taylor series for $e^x$

Let us show that the Taylor series for the function $e^x$ about the point 0 is a convergent series for any value of $x = b \geq 0$ and that it converges to the value $e^b$ (a similar proof works for $b < 0$).

In this case, at any point $a$, $f^{(n)}(a) = e^a$, so at $a = 0$ we obtain $f^{(n)}(0) = 1$. Hence the series about 0 is

$$\sum_{k=0}^{\infty} \frac{b^k}{k!}.$$

If we look at $R_n(b) = e^b - s_n(b)$ we obtain

$$|R_n(b)| = \frac{e^c b^{n+1}}{(n+1)!} \leq \frac{e^b b^{n+1}}{(n+1)!},$$

since $c \leq b$. As $n \to \infty$ this clearly goes to 0. This shows that the Taylor series of $e^b$ converges to the value of the function at each real number $b$.

# Defining functions using Taylor series

Instead of finding the Taylor series of a given function we can reverse the process and define functions using convergent series. Thus, one can define the function $e^x$ as

$$e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

In this case, we have to first show that the series on the right hand side converges for a given value of $x$, in which case the definition above makes sense.

We show the convergence of such series by showing that they are Cauchy series. This means that we do not have to guess at a value of the limit.

# Power series

As we have explained in the previous slide the "correct" (both from the point of view of proofs and of computation) way to define a function like $e^x$ is via convergent series involving non-negative integer powers of $x$. Such series are called power series and such functions should be viewed as the natural generalizations of polynomials.

The nice thing about power series is that once we know that they converge in some interval $(a-r, a+r)$ around $a$, it is not hard to show that the functions that they define are continuous functions. In fact, it is not too hard to show that they are smooth functions (that is, that all their derivatives exist). Thus when functions are given by convergent power series, we can automatically conclude they are smooth. This is the advantage of defining functions in this way.

# Calculating the values of functions

As we have also mentioned several times, calculators and computers calculate the values of various common functions like trigonometric polynomials and expressions in $\log x$ and $e^x$ by using Taylor series.

The great advantage of Taylor series is that one can estimate the error since we have a simple formula for the error which can be easily estimated. For instance, for the function $\sin x$, the $n$-th derivative is either $\pm \sin x$ or $\pm \cos x$, so in either case $|f^{(n}(x)| \leq 1$. Hence,

$$|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}$$

If we take $x = 1$, and we want to compute $\sin 1$ to an error of less than $10^{-16}$, we need only make sure that $(n+1)! > 10^{16}$, which is achieved when $n \geq 21$. (Can you find a value of $n$ which works for any value of $x$?)

# Computing the vaues of $\sin x$

Let us answer the question asked above.

First, remember that $\sin x$ is periodic, so we only have to look at the values of $x$ between $-\pi$ and $\pi$.

But we can do better, because $\sin(-x) = -\sin x$. So we only have to bother about the interval $[0, \pi]$.

We can do still better! Once we know $\sin x$ in $[0, \pi/2]$, we can easily figure out what it is in $[\pi/2, \pi]$.

So finally, it is enough to find the desired value of $n$ for $x \in [0, \pi/2]$.

# Computing the values of $\sin x$

We know that the remainder term satisfies

$$|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Hence, we need

$$\frac{|x|^{n+1}}{(n+1)!} < 10^{-16}.$$

We know that it is enough to look at $x \leq \pi/2$. Let us be a little careless and allow $x \leq 2$ (so we won't get the best possible $n$, maybe).

We already know that $1/(n+1)! < 10^{-16}$ if $n \geq 21$. Now $|x|^{22} \leq 2^{22}$. If we take $n = 31$, we see that $|x|^{32} \leq 2^{22} \cdot 2^{10}$,

$$1/(n+1)! = 1/32! < 10^{-16} \cdot 10^{-10} \cdot 2^{-10}.$$

# Questions for next time

If $f : \mathbb{R} \to \mathbb{R}$ is a smooth function, is it true that the Taylor series of $f$ converges?

to $f$?