Analysing Change Continuously

CMEE Maths Week

Michael Mustri

January 20, 2023

Table of Contents

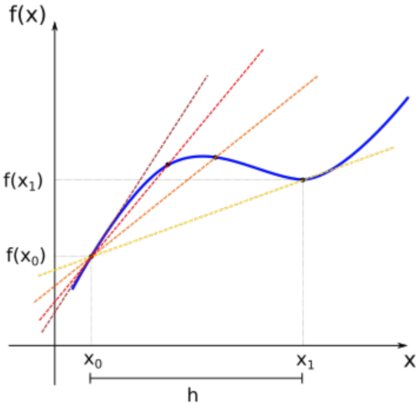
1 The Derivative

The Average Rate of Change

Let's imagine a **continuous** function $f(x)$, and define its average rate of change between two points x_0 and x_1 as:

$$\frac{\Delta f}{\Delta x} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

This quantity is the slope of the line that passes through $(x_0, f(x_0))$ and $(x_1, f(x_1))$, known as the **secant line**.

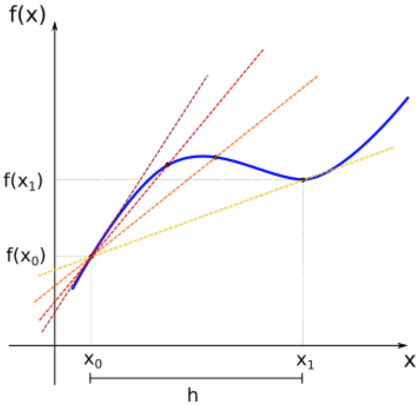


Instantaneous Rate of Change

Now, what happens when x_1 inches closer to x_0. How does the **secant line** change and what can be said about its slope?

$$\lim_{x_1 \to x_0} \frac{f(x_1) - f(x_0)}{x_1 - x_0} = L$$

As $x_1 \to x_0$ the secant becomes the **tangent** of $f(x)$ at x_0, and its slope, L is better known as $f'(x_0)$.



The Derivative

$f'(x_0)$ is denominated the **derivative** of $f(x)$ at x_0. The derivative can be more generally defined for any x in f's domain as:

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

This is mostly true when $f(x)$ is **continuous** at x, a condition we have yet to explain. However, continuous functions are not necessarily **differentiable**, which is something to keep in mind.

Table of Contents

2 Continuity

Formally, a function is **continuous** on an interval $[x_0, x_1]$ if, for every $a \in [x_0, x_1]$:

$$\lim_{x \to a} f(x) = f(a)$$

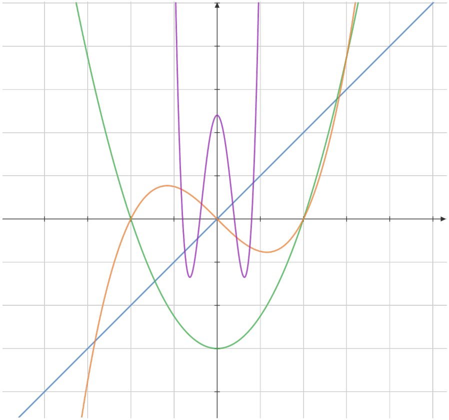
That is to say, that the limit exists and is equal to the function evaluated at a. When this limit either does not exist or is different to $f(a)$, we say that a is a **discontinuity** of f.

$$\lim_{x \to a} f(x) \neq f(a) \quad \text{or} \quad \lim_{x \to a} f(x) \to \pm\infty$$

Let's look at a few examples to illustrate what a continuous function looks like.

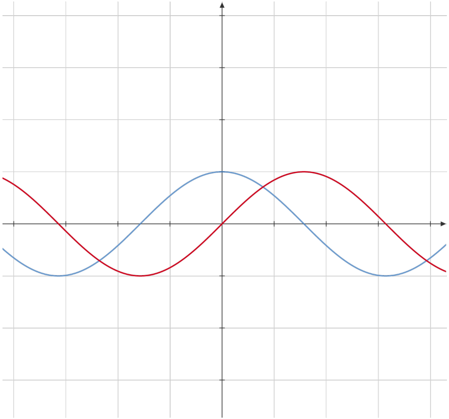
Continuous Example: Polynomials

All polynomials are continuous in $\mathbb{R}$, and as we'll later see, they are infinitely differentiable everywhere in their domain, to abbreviate we say they are class C^{∞}.



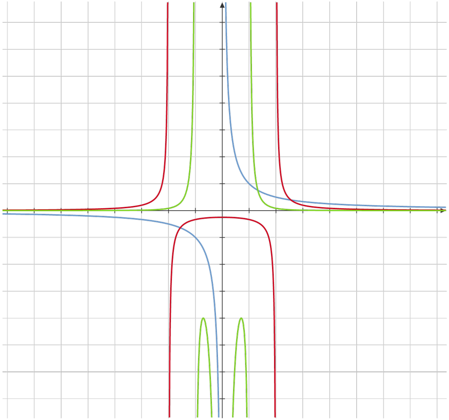
Continuous Example: $\sin$ and $\cos$

Both $\sin(x)$ and $\cos(x)$ are continuous in $\mathbb{R}$ and class C^{∞}. We'll see why this is the case later.



Discontinuous Example: Rational Functions

Rational functions are a good example of functions with discontinuities. For the most part they are continuous everywhere except at their vertical asymptotes.



Discontinuous Example: The Logarithm

Another function which is clearly not continuous is $\log(x)$, which diverges as $x \to 0$. Additionally, the logarithm isn't defined for $x < 0$.

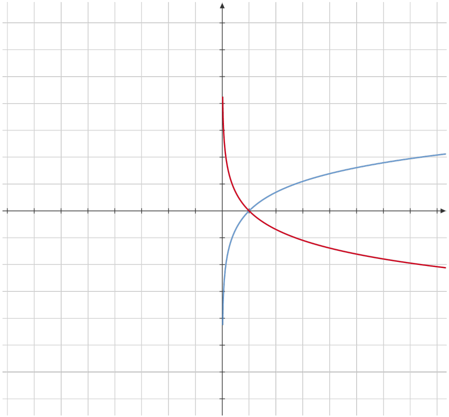


Table of Contents

3 Calculating Derivatives

Properties of Limits

While we won't go into the formal definition of limits, it is worth going through their properties before moving forward. Given two continuous functions $f(x)$ and $g(x)$, it can be proven that, if $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$, the following statements are true:

- $\lim_{x\to a} \left[f(x) + g(x)\right] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) = L + M$
- $\lim_{x\to a} \left[Cf(x)\right] = C \lim_{x\to a} f(x) = CL$, for any $C \in \mathbb{R}$
- $\lim_{x\to a} \left[f(x)g(x)\right] = \left[\lim_{x\to a} f(x)\right] \left[\lim_{x\to a} g(x)\right] = LM$
- $\lim_{x\to a} \left[\frac{f(x)}{g(x)}\right] = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)} = \frac{L}{M}$, so long as $M \neq 0$
- $\lim_{x\to a} g(f(x)) = g(\lim_{x\to a} f(x)) = g(L)$, if $g(x)$ is continuous at $x = L$

Observe that the limit is a **linear operator**.

We saw the definition of the derivative, now let's look at a simple example: the derivative of - $f(x) = x^2$.

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h}$$

By the properties seen previously, this can be decomposed into:

$$\lim_{h \to 0} \left(\frac{2xh}{h} + \frac{h^2}{h} \right) = \lim_{h \to 0} (2x) + \lim_{h \to 0} (h) = 2x + 0 = 2x$$

Thus $f'(x) = 2x$.

Calculating Derivatives: Power Functions

We can generalize our previous example - $f(x) = x^2$ - to any power function - $f(x) = x^n$ - with $n \neq 0$:

$$f(x) = x^n \Rightarrow f'(x) = nx^{n-1}$$

This follows from the binomial expansion of $(x+h)^n$, though it requires a few more steps for rational $n = \frac{p}{q}$.

The Derivative as a Linear Operator

From the properties of limits, their linearity in particular, we can make a few observations. Given two differentiable functions $f(x)$ and $g(x)$:

- $\frac{d}{dx}\left[f(x)+g(x)\right]=\frac{df}{dx}+\frac{dg}{dx}$ - The derivative preserves function addition.
- For any $C\in\mathbb{R}$ - $\frac{d}{dx}\left[Cf(x)\right]=C\frac{df}{dx}$ - The derivative preserves scalar multiplication.

That is to say, the derivative is a **linear operator**.

The Product Rule

Say that we have two differentiable functions $f(x)$ and $g(x)$, what can we conclude about the derivative of their product?

$$\frac{d}{dx}\left[f(x)g(x)\right] = \lim_{h\to 0}\frac{f(x+h)g(x+h) - f(x)g(x)}{h}$$

Notice that we can add as many zeros to the above equation without changing its result.

$$\frac{d}{dx}\left[f(x)g(x)\right] = \lim_{h\to 0}\frac{f(x+h)g(x+h) - f(x)g(x) + \left[f(x+h)g(x) - f(x+h)g(x)\right]}{h}$$

The Product Rule

Rearranging terms from the previous equation we find that:

$$\frac{d}{dx}\left[f(x)g(x)\right] = \lim_{h\to 0}\frac{f(x+h)\left[g(x+h)-g(x)\right]+g(x)\left[f(x+h)-f(x)\right]}{h}$$

From the properties of limits, we can are left with:

$$\frac{d}{dx}\left[f(x)g(x)\right] = \lim_{h\to 0}\frac{f(x+h)\left[g(x+h)-g(x)\right]}{h}+\lim_{h\to 0}\frac{g(x)\left[f(x+h)-f(x)\right]}{h}$$

Since $f(x)$ is differentiable $\lim_{h\to 0} f(x+h) = f(x)$ and we can safely factor it out from the limit.

The Product Rule

Finally, from the last slide it should be obvious that:

$$\frac{d}{dx}\left[f(x)g(x)\right] = f(x)\frac{dg}{dx} + g(x)\frac{df}{dx}$$

This last relationship is called the **Product Rule**, and apart from being an important result in differential calculus, it is particularly useful for calculating tricky derivatives.

The Chain Rule

Now that we know what the derivative of the product of functions looks like, we may ask what happens with the composition of functions. With a fair bit of work, it can be shown that, given two differentiable functions $f(x)$ and $g(x)$:

$$\frac{d}{dx}(g \circ f)(x) = f'(x)\left[g' \circ f(x)\right] = \frac{dg}{df}\frac{df}{dx}$$

This is known as the **Chain Rule** and it can be generalized to however many compositions. Along with the product rule, the chain rule greatly simplifies more complex derivatives.

The Quotient Rule

An application of both the product and the chain rule is the derivative of the quotient of two functions. Given two differentiable functions $f(x)$ and $g(x)$, with $g(x) \neq 0$:

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x) - g'(x)f(x)}{g^2(x)}$$

This is the **Quotient Rule** and while it is a direct consequence of the product and chain rules, it serves as a shortcut for differentiating rational functions.

The Inverse Function Rule

3 Calculating Derivatives

Assuming we know the derivative of a function $f(x)$, we can use what we've learned so far to find the derivative of it's inverse function. So given $f(x)$ and it's inverse $g(x)$ such that $g \circ f(x) = x$, we can see that if f and g are both differentiable:

$$\frac{d}{dx}[g \circ f(x)] = f'(x)[g' \circ f(x)] = 1 \quad \Rightarrow \quad g' \circ f(x) = \frac{1}{f'(x)}$$

Making a substitution in both sides of the equation $x = g(x)$ we find that the derivative of the inverse is given by:

$$g'(x) = \frac{1}{f'(g(x))} \quad \text{or} \quad \frac{d}{dx} f^{-1}(x) = \frac{1}{f' \circ f^{-1}(x)}$$

The Derivative of the Exponential

The function $f(x) = e^x$, has an interesting derivative worth working through explicitly.

$$\frac{d}{dx}e^x = \lim_{h\to 0}\frac{e^{x+h} - e^x}{h} = \lim_{h\to 0}\frac{e^x e^h - e^x}{h} = \lim_{h\to 0}\frac{e^x(e^h - 1)}{h}$$

Since e^x doesn't depend on h, we can factor to out from the limit:

$$\frac{d}{dx}e^x = e^x \lim_{h\to 0}\frac{e^h - 1}{h}$$

Using the definition of $e^h = \lim_{n\to\infty}(1 + \frac{h}{n})^n$ it can be proven that the above limit is equals to one:

$$\lim_{h\to 0}\frac{\lim_{n\to\infty}(1 + \frac{h}{n})^n - 1}{h} = 1 \quad \Rightarrow \quad \frac{d}{dx}e^x = e^x$$

The Derivative of the Logarithm

Now that we've seen that $\frac{d}{dx}e^x = e^x$, we can use the inverse function rule to find the derivative of the logarithm:

$$\frac{d}{dx}\ln(x) = \frac{1}{e^{\ln(x)}} = \frac{1}{x}$$

Using the rules and properties we've learned so far, we can find the derivatives of the elementary functions:

- $f(x) = a^x$ then $f'(x) = \ln(a)a^x$, for $a > 0$.
- $f(x) = \log_a(x)$ then $f'(x) = \frac{1}{\ln(a)x}$.
- $f(x) = \sin(x)$ then $f'(x) = \cos(x)$.
- $f(x) = \cos(x)$ then $f'(x) = -\sin(x)$.

These are only a handful of derivatives, but along with the product and chain rules, they pave the road to most of derivatives you'll find.

Table of Contents

4 Higher Order Derivatives

Another thing worth asking is whether you can differentiate a function more than once and, unsurprisingly, the answer depends on the function. In general, if the derivative of a function $f'(x)$ is well behaved or **differentiable**, we can safely assume its derivative exists:

$$\frac{d}{dx}f'(x) = \frac{d^2f}{dx^2} = f''(x)$$

We call this the second derivative of $f(x)$ with respect to x, and we say that $f(x)$ is **at least** twice differentiable.

Differentiability Classes

Now, we can take the differentiation process a bit further and define the n^{th} derivative of $f(x)$ with respect to x:

$$f^{(n)} = \frac{d^n f}{dx^n}$$

This is only valid as long as the lower order derivatives $\left[f^{(0)}(x), f^{(1)}(x), \ldots, f^{(n-1)}\right]$ are also differentiable. If a function can be differentiated up to n times, and $f^{(n)}$ is continuous, we say it is of **differentiability class** C^n. If a function can be differentiated infinitely many times, we call it class C^∞ or **smooth**.

Table of Contents

5 Multivariable Functions

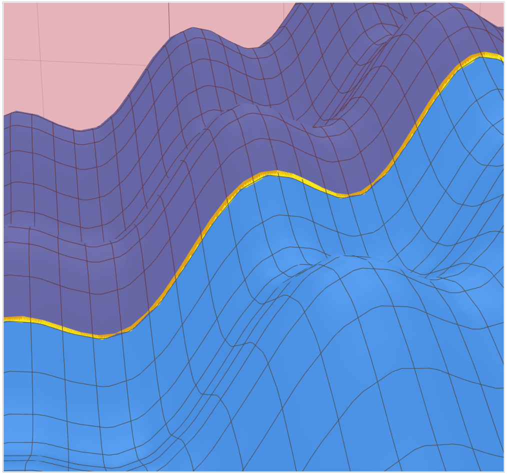
Let's consider a function $z = f(x, y)$ and note that it defines a surface on the three-dimensional xyz-coordinate space. If $f(x, y)$ is differentiable with respect to both x and y, we can define its derivatives with respect to each variable:

$$\frac{\partial}{\partial x} f(x, y) = \frac{\partial f}{\partial x} \quad \text{and} \quad \frac{\partial}{\partial y} f(x, y) = \frac{\partial f}{\partial y}$$

These are the **partial derivatives** of $f(x, y)$ with respect to x and y respectively. Partial derivatives tell us how a function varies along either of the directions defined by x or y.

A Geometric View

An intuitive way of viewing partial derivatives is by imagining a cross section of the surface defined by $z = f(x, y)$. For example, the partial derivative $\frac{\partial f}{\partial x}$ at $y = 0$, describes the rate of change of the curve at the intersection with $y = 0$.



Calculating partial derivatives is, for the most part, identical to normal derivatives, you simply set the variable not under consideration as a constant. So for example, if - $f(x, y) = x^2y$:

$$\frac{\partial f}{\partial x} = 2xy \quad \text{and} \quad \frac{\partial f}{\partial y} = x^2$$

This is a very simple example, but the principle holds for arbitrarily complex f, so long as it is differentiable in each variable.

If you've been wondering whether higher order derivatives are possible for multivariate functions, the answer is yes, but it's a bit different. If $f(x, y)$ has continuous and differentiable partial derivatives, the following derivatives exist:

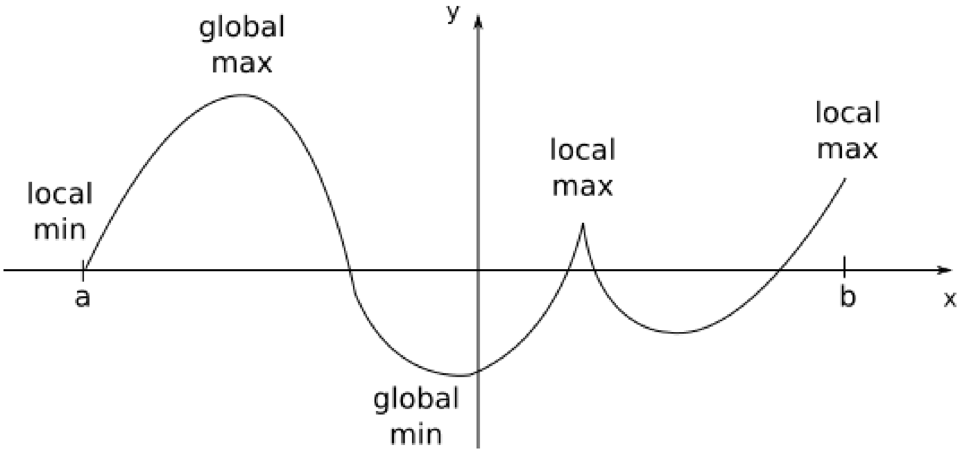
$$\frac{\partial^2 f}{\partial x^2} = f_{xx}(x, y) \qquad \frac{\partial^2 f}{\partial y^2} = f_{yy}(x, y)$$
$$\frac{\partial^2 f}{\partial x \partial y} = \left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right) = f_{xy}(x, y) \qquad \frac{\partial^2 f}{\partial y \partial x} = \left(\frac{\partial f}{\partial y}\right)\left(\frac{\partial f}{\partial x}\right) = f_{yx}(x, y)$$

Moreover, since both first order partials are differentiable, it can be proven that $f_{xy} = f_{yx}$. This result is known as **Clairaut's Theorem**.

Table of Contents

6 Optimization

Finding the Optima



A key result in calculus is the **Extreme Value Theorem**: given a function $f(x)$, that is continuous on the interval $x \in [a, b]$, we can assure that $f(x)$ attains a **maximum** and a **minimum** value at least once. In other words, there exist c and d in $[a, b]$ such that:

$$f(d) \leq f(x) \leq f(c) \quad \forall x \in [a, b]$$

We call $f(c)$ the **global maximum** of $f(x)$ on the interval $[a, b]$ and $f(d)$ its **global minimum**.

Now, it is clear that the **global extrema** will largely depend on the interval chosen. However, this allows us to develop a more general concept: **local extrema**. Assuming $f(x)$ is continuous, we say that x_0 is a **local maximum or minimum** if, for an arbitrarily small c:

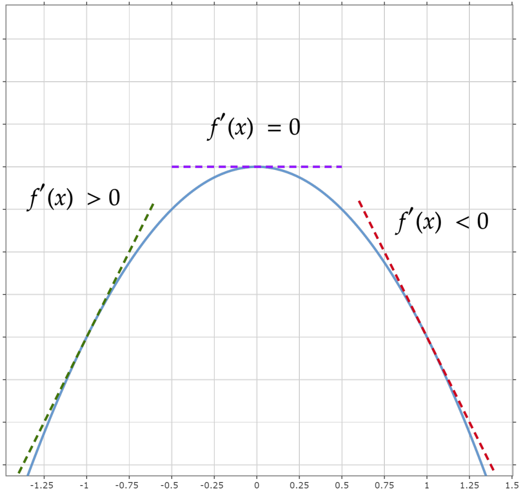
$$f(x_0) \geq f(x) \quad \text{or} \quad f(x_0) \leq f(x), \quad \forall x \in (x_0 - c, x_0 + c)$$

So as we shrink the interval - $(x_0 - c, x_0 + c)$ - around x_0, $f(x_0)$ remains the largest or smallest value on the interval.

An important result of local extrema is that if x_0 is a local extremum of $f(x)$ and $f(x)$ is differentiable at x_0, then

$$f'(x_0) = 0$$

we call x_0 a **critical point** of $f(x)$.



Not all **critical points** are local extrema, they are only potential candidates. For example: $f(x) = x^3$.

$$f'(x) = 3x^2 \quad \Rightarrow \quad f'(0) = 0$$

$x_0 = 0$ is neither a local max nor a local min. Instead x_0 is known as an **inflection point**.

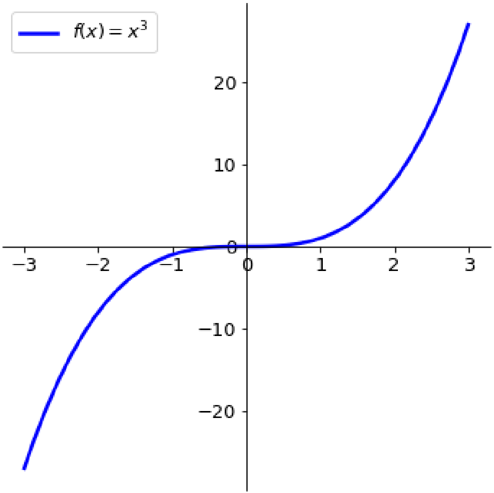


Table of Contents

7 Monotonicity and Concavity

We'll step back a bit and imagine a function $f(x)$ defined on the interval $[a, b]$. Given any two points x_1 and x_2 on the interval $[a, b]$ such that $x_2 > x_1$, we can make the following definitions:

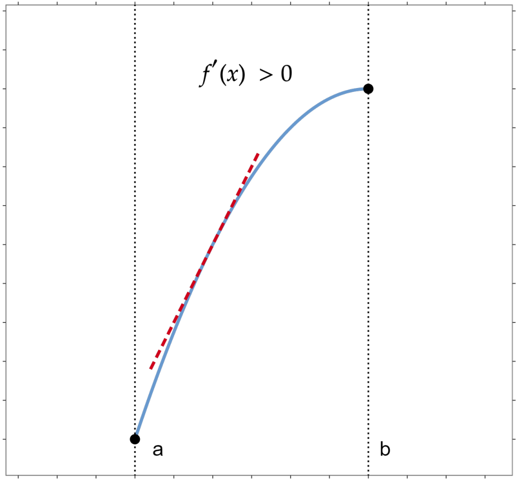
- If $f(x_2) > f(x_1) \Rightarrow f(x)$ is **increasing** on the interval $[a, b]$.
- If $f(x_2) < f(x_1) \Rightarrow f(x)$ is **decreasing** on the interval $[a, b]$.

Since we have a strict inequality, we can add that $f(x)$ is **monotonous**. Otherwise if $f(x_2) \geq f(x_1)$ or $f(x_2) \leq f(x_1)$, we would say $f(x)$ was **non-decreasing** or **non-increasing**. Moreover this generalization of increasing/decreasing functions can be stated in terms of the derivatives.

Given function $f(x)$ that is differentiable on the interval $[a, b]$ if:

$$f'(x) > 0 \quad \forall x \in [a, b]$$

Then $f(x)$ is **increasing** on the interval $[a, b]$.

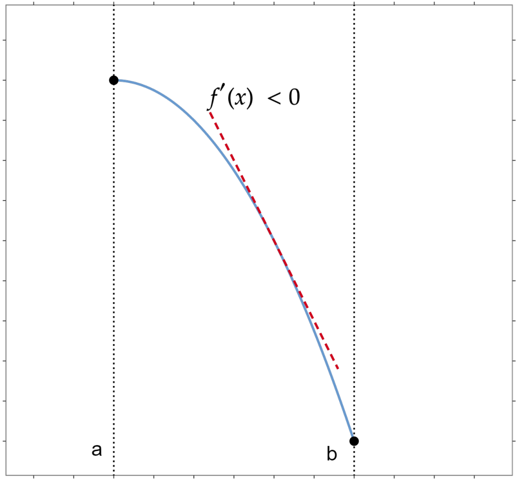


The Derivative of Decreasing Functions

Given function $f(x)$ that is differentiable on the interval $[a, b]$ if:

$$f'(x) < 0 \quad \forall x \in [a, b]$$

Then $f(x)$ is **decreasing** on the interval $[a, b]$.



Concavity

Using derivatives, can generalize what we had seen on concavity to arbitrarily defined functions. Again, let's imagine $f(x)$ differentiable on $[a, b]$ and add that it's derivative, $f'(x)$, is also differentiable on the same interval. Then, we find the following statements are true:

- If $f'(x)$ is **increasing** for $[a, b] \Rightarrow f(x)$ is **concave up**.
- If $f'(x)$ is **decreasing** for $[a, b] \Rightarrow f(x)$ is **concave down**.

This is another way of saying that the slope of the tangent lines of $f(x)$ either increase or decrease with x.

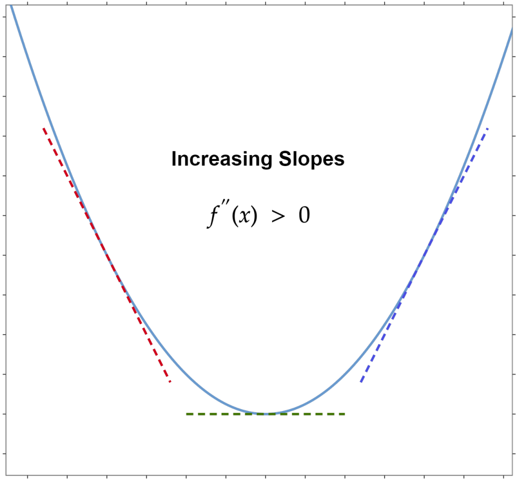
Concave Up

7 Monotonicity and Concavity

Given function $f(x)$ that is twice differentiable on the interval $[a, b]$ if:

$$f''(x) > 0 \quad \forall x \in [a, b]$$

Then $f(x)$ is **concave up** on the interval $[a, b]$.



Concave Down

Given function $f(x)$ that is twice differentiable on the interval $[a, b]$ if:

$$f''(x) < 0 \quad \forall x \in [a, b]$$

Then $f(x)$ is **concave down** on the interval $[a, b]$.

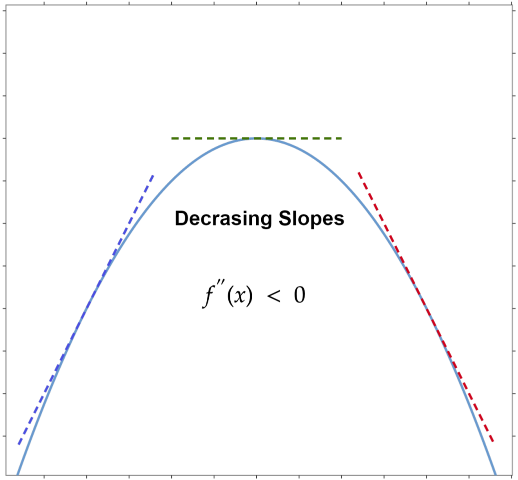


Table of Contents

8 Finding Extrema

Finding Extrema

These last few results provide a simple framework for finding local maxima and minima of a function. Given a twice differentiable function $f(x)$, defined on the interval $[a, b]$:

1. Find all critical points of $f(x)$ - $c \in [a, b]$ such that $f'(c) = 0$.
2. Compute $f''(x)$ for each critical point c.
3. Classify c according to local concavity.

The last step requires a bit of intuition on what concavity implies for c.

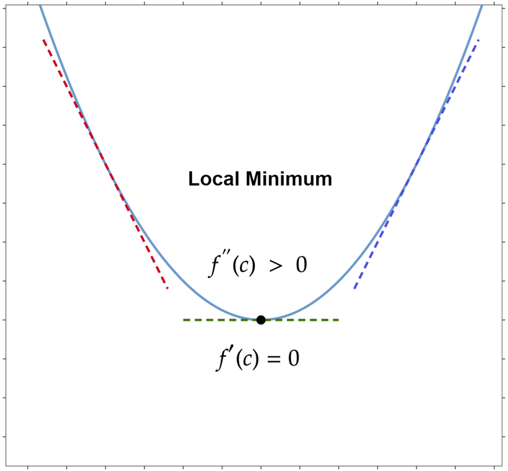
Concavity of Local Minima

8 Finding Extrema

Given function $f(x)$ that is twice differentiable on the interval $[a, b]$, if for $c \in [a, b]$:

$$f'(c) = 0 \quad \text{and} \quad f''(c) > 0$$

Then c is a **local minimum** of $f(x)$.

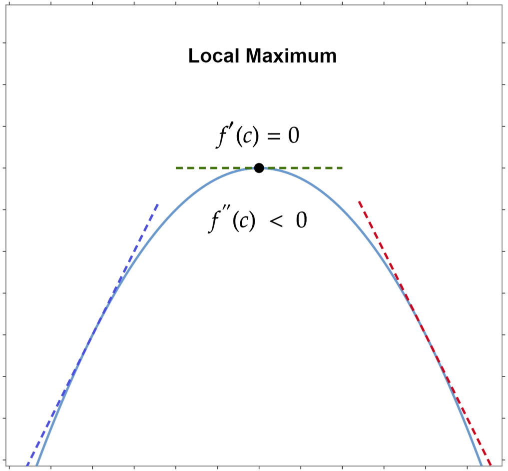


Concavity of Local Maxima

Given function $f(x)$ that is twice differentiable on the interval $[a, b]$, if for $c \in [a, b]$:

$$f'(c) = 0 \quad \text{and} \quad f''(c) < 0$$

Then c is a **local maximum** of $f(x)$.



Concavity of Inflection Points

As you may have already noted, there exists a third class of critical point. This pertains to the case:

$$f'(c) = 0 \quad \text{and} \quad f''(c) = 0$$

Here c is an **inflection point** of $f(x)$. We further observe that inflection points indicate where functions **change concavity**.

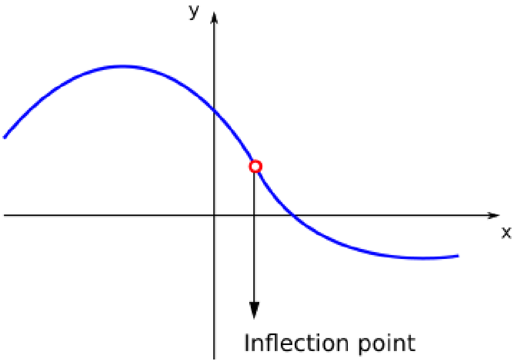


Table of Contents

9 Taylor Series

Now, let's take another look at the definition of the derivative. Given a differentiable function $f(x)$:

$$f'(a) = \lim_{x \to a} \frac{f(a) - f(x)}{a - x}$$

If x is in a small enough neighborhood around a, we can make the following approximation.

$$f'(a) \approx \frac{f(a) - f(x)}{a - x} \qquad \Rightarrow \qquad f(x) \approx f(a) + f'(a)(x - a)$$

So around a very small neighborhood of a, we can approximate $f(x)$ as a **linear function**

Assuming $f(x)$ can be differentiated n times and using the **mean value theorem**, it can be proven that:

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n$$

The above statement is known as **Taylor's Theorem**. Taylor's Theorem implies that any arbitrary C^n function can be approximated as a polynomial of its derivatives around a point a, this approximation is called the **Taylor polynomial** of $f(x)$.

Building the Taylor Polynomial

Although we won't be proving Taylor's theorem, we'll look at the motivation behind it. Assume you have an n-differentiable function $f(x)$ that can be approximated around a point a by an n^{th} degree polynomial $P(x)$:

$$P(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n$$

Since we want this to be the best approximation possible, we expect the derivatives of $f(x)$ and $P(x)$, and the functions themselves to be the same at $x = a$:

$$\begin{aligned} f(a) &= P(a) \\ f'(a) &= P'(a) \\ &\vdots \\ f^{(n)}(a) &= P^{(n)}(a) \end{aligned}$$

Building the Taylor Polynomial

Differentiating up to the n^{th} order derivative of $P(x)$, we find that:

$$
\begin{gathered}
P(a) = f(a) = a_0 \\
P'(a) = f'(a) = a_1 \\
P''(a) = f''(a) = 2 \cdot 1 \cdot a_2 \\
\vdots \\
P^{(n)}(a) = f^{(n)}(a) = n \cdot (n-1) \cdots 2 \cdot 1 \cdot a_n
\end{gathered}
$$

Defining $n! = n \cdot (n-1) \cdots 2 \cdot 1$ as the **factorial** of n, we can solve for the coefficients a_i.

Solving for the coefficients in our last set of equations, we have:

$$f^{(n)}(a) = n!a_n \quad \Rightarrow \quad a_n = \frac{f^{(n)}(a)}{n!}$$

Hence the polynomial $P(x)$ is defined by:

$$P(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{n!}f^{(n)}(a)(x-a)^n$$

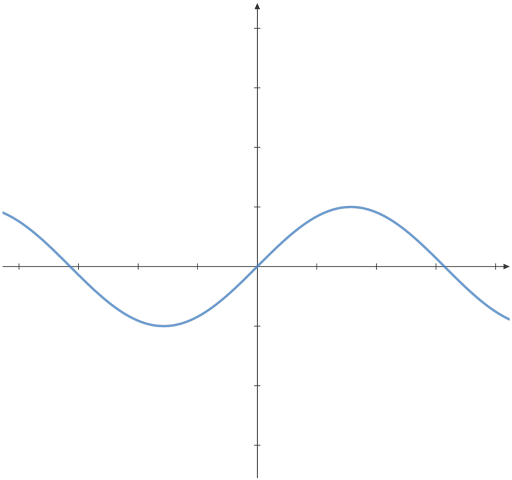
$P(x)$ is the n^{th} degree **Taylor Polynomial** of $f(x)$ defined at the point $x = a$.

Example: Taylor Expansion of $\sin(x)$

So let's go though an example of a Taylor expansion - $f(x) = \sin(x)$ - around the point $x = 0$.

$$\sin(x) = P(x)$$

First let's note that all of the even order terms $n = 2k$ of $P(x)$ are going to be zero since $\sin(0) = 0$.

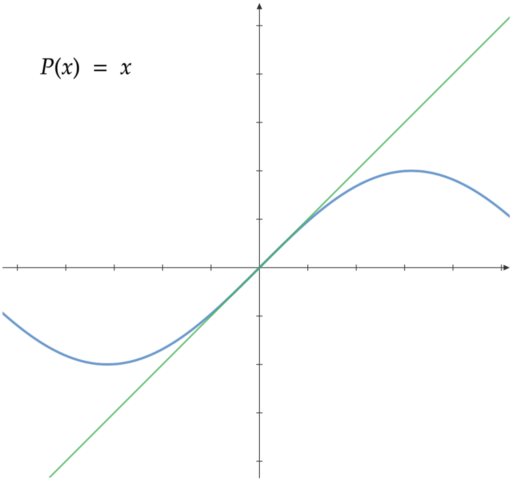


Example: Taylor Expansion of $\sin(x)$

We can jump straight into the 1^{st} order expansion.

$$\sin(x) = \sin(0) + \cos(0)(x)^1 = x$$

Turns out $\sin(x) \approx x$ is a decent approximation for $x << 1$, this is usually referred to the small angle approximation.

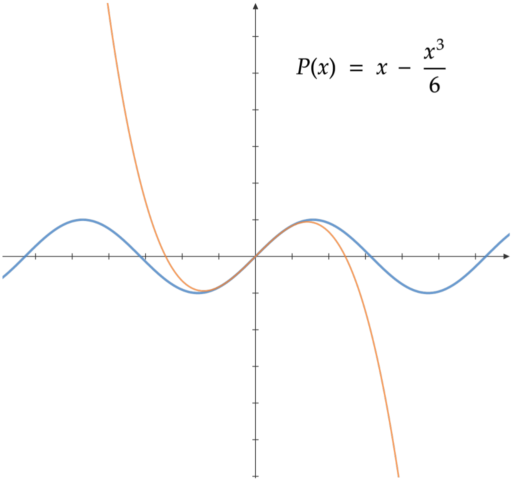


Example: Taylor Expansion of $\sin(x)$

Since even order terms don't change the polynomial, we can skip to the third order polynomial.

$$\sin(x) = x - \frac{\cos(0)}{3!}(x)^3 = x - \frac{x^3}{3!}$$

Here the negative sign appears from $\frac{d}{dx}\cos(x) = -sin(x)$. This will create an alternating $\pm$ pattern.

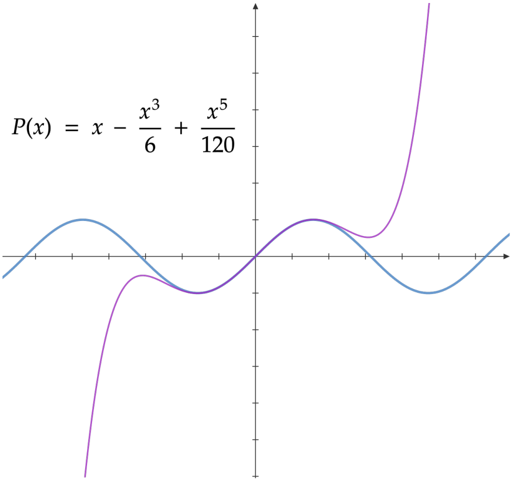


Example: Taylor Expansion of $\sin(x)$

We go on and add the fifth order term:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$$

Notice how adding terms improves the approximation.



Example: Taylor Expansion of $\sin(x)$

9 Taylor Series

And since $\sin(x)$ is C^∞ we can build an arbitrarily large polynomial

$$\sin(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{(2k+1)}}{(2k+1)!}$$

Eventually (as $n \to \infty$) $P(x)$ will be almost indistinguishable from $\sin(x)$. Though it is usually unnecessary to resort to such high order terms.