Elementary Calculus Manual

Alfredo Sánchez Alberca (asalber@ceu.es)

Sep 2016

Department of Applied Math and Statistics
CEU San Pablo

Contents

1 Analytic geometry

Analytic geometry

Contents

1.1 Vectors

Scalars

Some phenomena of Nature can be described by a number and a unit of measurement.

Definition 1 (Scalar)**.** A *scalar* is a number that expresses a magnitude without direction.

Examples The height or weight of a person, the temperature of a gas or the time it takes a vehicle to travel a distance.

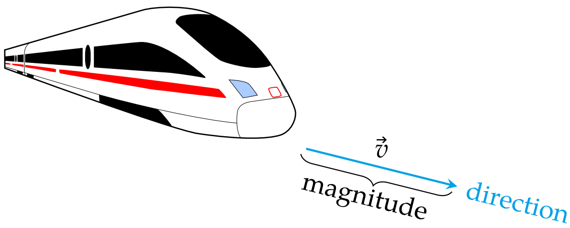
However, there are other phenomena that cannot be described adequately by an scalar. If, for instance, a sailor wants to head for seaport and only knows the intensity of wind, he won't know what direction to take. The description of wind requires two elements: intensity and direction.

Vectors

Definition 2 (Vector)**.** A *vector* is a number that expresses a magnitude and has associated an orientation and a sense.

Examples The velocity of a vehicle or the force applied to an object.

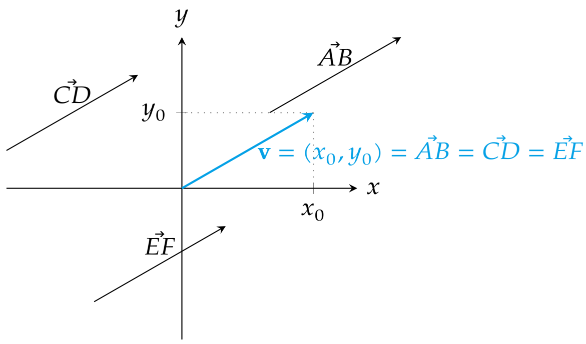
Geometrically, a vector is represented by an directed line segment, that is, an arrow.



Vector representation

An oriented segment can be located in different places in a Cartesian space. However, regardless of where it is located, if the length and the direction of the segment doesn't change, the segment represents always the same vector.

This allows to represent all vector with the same origin, the origin of the Cartesian coordinate system. Thus, a vector can be represented by the Cartesian *coordinates* of its final end in any Euclidean space.

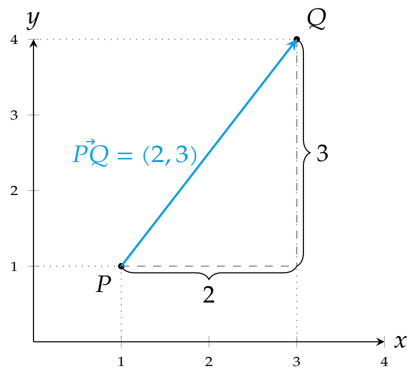


Vector from two points

Given two points P and Q of a Cartesian space, the vector that starts at P and ends at Q has coordinates $\vec{PQ} = Q - P$.

Example Given the points $P = (1, 1)$ and $Q = (3, 4)$ in the real plane $\mathbb{R}^2$, the coordinates of the vector that start at P and ends at Q are

$$\vec{PQ} = Q - P = (3, 4) - (1, 1) = (3 - 2, 4 - 1) = (2, 3).$$



Module of a vector

Definition 3 (Module of a vector)**.** Given a vector $\mathbf{v} = (v_1, \cdots, v_n)$ in $\mathbb{R}^n$, the *module* of $\mathbf{v}$ is

$$|\mathbf{v}| = \sqrt{v_1^2 + \cdots + v_n^2}.$$

The module of a vector coincides with the length of the segment that represents the vector.

Examples Let $\mathbf{u} = (3, 4)$ be a vector in $\mathbb{R}^2$, then its module is

$$|\mathbf{u}| = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

Let $\mathbf{v} = (4, 7, 4)$ be a vector in $\mathbb{R}^3$, then its module is

$$|\mathbf{v}| = \sqrt{4^2 + 7^2 + 4^2} = \sqrt{81} = 9$$

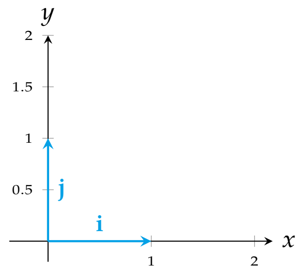
Unit vectors

Definition 4 (Unit vector)**.** A vector $\mathbf{v}$ in $\mathbb{R}^n$ is a *unit vector* if its module is one, that is, $|\mathbf{v}| = 1$.

The unit vectors with the direction of the coordinate axes are of special importance and they form *standard basis*.

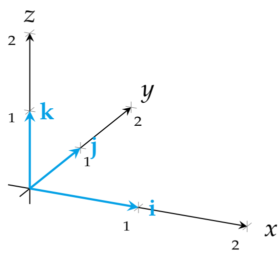
In $\mathbb{R}^2$ the standard basis is formed by two vectors

$$\mathbf{i} = (1,0) \text{ and } \mathbf{j} = (0,1)$$



In $\mathbb{R}^3$ the standard basis is formed by three vectors

$$\mathbf{i} = (1,0,0), \mathbf{j} = (0,1,0) \text{ and } \mathbf{k} = (0,0,1)$$



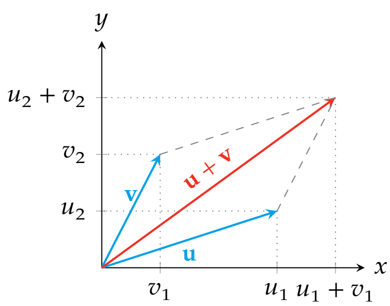
Sum of two vectors

Definition 5 (Sum of two vectors)**.** Given two vectors $\mathbf{u} = (u_1, \cdots, u_n)$ y $\mathbf{v} = (v_1, \cdots, v_n)$ de $\mathbb{R}^n$, the *sum* of $\mathbf{u}$ and $\mathbf{v}$ is

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, \ldots, u_n + v_n).$$

Example Let $\mathbf{u} = (3,1)$ and $\mathbf{v} = (2,3)$ two vectors in $\mathbb{R}^2$, then the sum of them is

$$\mathbf{u} + \mathbf{v} = (3+2, 1+3) = (5,4).$$



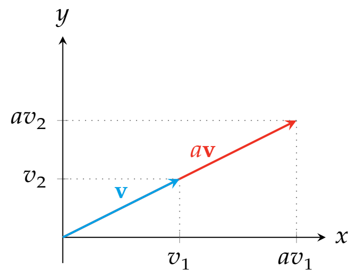
Product of a vector by an scalar

Definition 6 (Product of a vector by an scalar)**.** Given a vector $\mathbf{v} = (v_1, \cdots, v_n)$ in $\mathbb{R}^n$, and a scalar $a \in \mathbb{R}$, the *product* of $\mathbf{v}$ by a is

$$a\mathbf{v} = (av_1, \ldots, av_n).$$

Example Let $\mathbf{v} = (2,1)$ a vector in $\mathbb{R}^2$ and $a = 2$ a scalar, then the product of a by $\mathbf{v}$ is

$$a\mathbf{v} = 2(2,1) = (4,2).$$

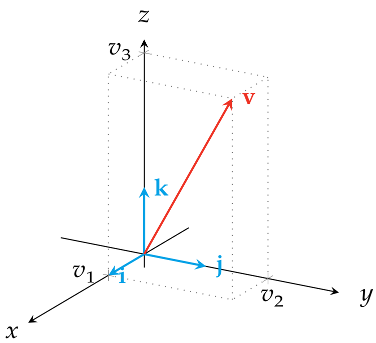


Expressing a vector as a linear combination of the standard basis

The sum of vectors and the product of vector by a scalar allow as to express any vector as a linear combination of the standard basis.

In $\mathbb{R}^3$, for instance, a vector with coordinates $\mathbf{v} = (v_1, v_2, v_3)$ can be expressed as the linear combination

$$\mathbf{v} = (v_1, v_2, v_3) = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}.$$



Dot product of two vectors

Definition 7 (Dot product of two vectors). Given the vectors $\mathbf{u} = (u_1, \cdots, u_n)$ and $\mathbf{v} = (v_1, \cdots, v_n)$ in $\mathbb{R}^n$, the *dot product* of $\mathbf{u}$ and $\mathbf{v}$ is

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + \cdots + u_n v_n.$$

Example Let $\mathbf{u} = (3,1)$ and $\mathbf{v} = (2,3)$ two vectors in $\mathbb{R}^2$, the the dot product of them is

$$\mathbf{u} \cdot \mathbf{v} = 3 \cdot 2 + 1 \cdot 3 = 9.$$

It holds that

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos \alpha$$

where α is the angle between the vectors.

Parallel vectors

Definition 8 (Parallel vectors). Two vectors $\mathbf{u}$ and $\mathbf{v}$ are *parallel* if there is a scalar $a \in \mathbb{R}$ such that

$$\mathbf{u} = a\mathbf{v}.$$

Example The vectors $\mathbf{u} = (-4,2)$ and $\mathbf{v} = (2,-1)$ in $\mathbb{R}^2$ are parallel, as there is a scalar -2 such that

$$\mathbf{u} = (-4,2) = -2(2,-1) = -2\mathbf{v}.$$

Orthogonal and orthonormal vectors

Definition 9 (Orthogonal and orthonormal vectors)**.** Two vectors $\mathbf{u}$ and $\mathbf{v}$ are *orthogonal* if their dot product is zero,

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

If in addition both vectors are unit vectors, $|\mathbf{u}| = |\mathbf{v}| = 1$, then the vectors are *orthonormal*.

Orthogonal vectors are perpendicular, that is the angle between them is right.

Example The vectors $\mathbf{u} = (2, 1)$ and $\mathbf{v} = (-2, 4)$ in $\mathbb{R}^2$ are orthogonal, as

$$\mathbf{uv} = 2 \cdot -2 + 1 \cdot 4 = 0,$$

but they are not orthonormal as $|\mathbf{u}| = \sqrt{2^2 + 1^2} \neq 1$ and $|\mathbf{v}| = \sqrt{-2^2 + 4^2} \neq 1$.

The vectors $\mathbf{i} = (1, 0)$ and $\mathbf{j} = (0, 1)$ in $\mathbb{R}^2$ are orthonormal, as

$$\mathbf{ij} = 1 \cdot 0 + 0 \cdot 1 = 0, \quad |\mathbf{i}| = \sqrt{1^2 + 0^2} = 1, \quad |\mathbf{j}| = \sqrt{0^2 + 1^2} = 1.$$

1.2 Lines

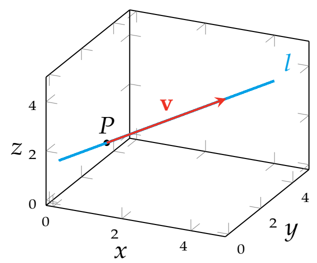
Vectorial equation of a straight line

Definition 10 (Vectorial equation of a straight line)**.** Given a point $P = (p_1, \ldots, p_n)$ and a vector $\mathbf{v} = (v_1, \ldots, v_n)$ of $\mathbb{R}^n$, the *vectorial equation of the line l* that passes through the point P with the direction of $\mathbf{v}$ is

$$l : X = P + t\mathbf{v} = (p_1, \ldots, p_n) + t(v_1, \ldots, v_n) = (p_1 + tv_1, \ldots, p_n + tv_n), \quad t \in \mathbb{R}.$$

Example Let l the line of $\mathbb{R}^3$ that goes through $P = (1, 1, 2)$ with the direction of $\mathbf{v} = (3, 1, 2)$, then the vectorial equation of l is

$$\begin{aligned} l : X = P + t\mathbf{v} = (1, 1, 2) + t(3, 1, 2) = \\ = (1 + 3t, 1 + t, 2 + 2t) \quad t \in \mathbb{R}. \end{aligned}$$



Parametric and Cartesian equations of a line

From the vectorial equation of a line $l : X = P + t\mathbf{v} = (p_1 + tv_1, \ldots, p_n + tv_n)$ is easy to obtain the coordinates of the the points of the line with n *parametric equations*

$$x_1(t) = p_1 + tv_1, \ldots, x_n(t) = p_n + tv_n$$

from where, if $\mathbf{v}$ is a vector with non-null coordinates ($v_i \neq 0\ \forall i$), we can solve for t and equal the equations getting the *Cartesian equations*

$$\frac{x_1 - p_1}{v_1} = \cdots = \frac{x_n - p_n}{v_n}$$

Example Given a line with vectorial equation $l : X = (1,1,2) + t(3,1,2) = (1+3t, 1+t, 2+2t)$ in $\mathbb{R}^3$, its parametric equations are

$$x(t) = 1 + 3t, \quad y(t) = 1 + t, \quad z(t) = 2 + 2t,$$

and the Cartesian equations are

$$\frac{x-1}{3} = \frac{y-1}{1} = \frac{z-2}{2}$$

Point-slope equation of a line in the plane

In the particular case of the real plane $\mathbb{R}^2$, if we have a line with vectorial equation $l : X = P + t\mathbf{v} = (x_0, y_0) + t(a,b) = (x_0 + ta, y_0 + tb)$, its parametric equations are

$$x(t) = x_0 + ta, \quad y(t) = y_0 + tb$$

and its Cartesian equation is

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}.$$

From this, passing b to the other side of the equation, we get

$$y - y_0 = \frac{b}{a}(x - x_0),$$

or renaming $m = b/a$,

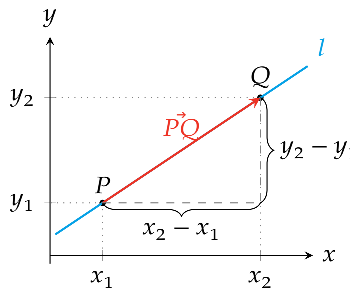
$$y - y_0 = m(x - x_0).$$

This equation is known as the *point-slope equation* of the line.

Slope of a line in the plane

Definition 11 (Slope of a line in the plane)**.** Given a line $l : X = P + t\mathbf{v}$ in the real plane $\mathbb{R}^2$, with direction vector $\mathbf{v} = (a,b)$, the *slope* of l is b/a.

Recall that given two points $Q = (x_1, y_1)$ y $Q = (x_2, y_2)$ of the line l, we can take as a direction vector the vector from P to Q, with coordinates $\vec{PQ} = Q - P = (x_2 - x_1, y_2 - y_1)$. Thus, the slope of l is $\dfrac{y_2 - y_1}{x_2 - x_1}$, that is, the ratio between the changes in the vertical and horizontal axes respectively.



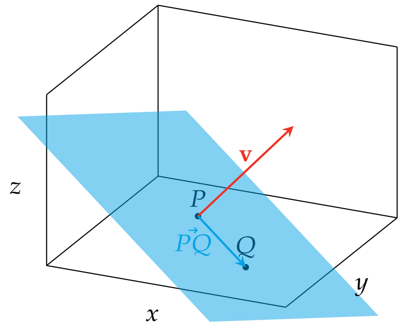
1.3 Planes

Vector equation of a plane in space

To get the equation of a plane in the real space $\mathbb{R}^3$ we can take a point of the plane $P = (x_0, y_0, z_0)$ and an orthogonal vector to the plane $\mathbf{v} = (a,b,c)$. Then, any point $Q = (x,y,z)$ of the plane meets that the vector $\vec{PQ} = (x - x_0, y - y_0, z - z_0)$ is orthogonal to $\mathbf{v}$, and therefore their dot product is zero.

$$\vec{PQ} \cdot \mathbf{v} = (x - x_0, y - y_0, z - z_0)(a,b,c) = a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

This equation is known as the *vector equation of the plane*.



Scalar equation of a plane in space

From the vector equation of the plane we can get

$$a(x-x_0)+b(y-y_0)+c(z-z_0)=0 \Leftrightarrow ax+by+cz=ax_0+by_0+cz_0,$$

that, renaming $d=ax_0+by_0+cz_0$, can be written

$$ax+by+cz=d,$$

and is known as the *scalar equation of the plane*.

Example Given the point $P=(2,1,1)$ and the vector $\mathbf{v}=(2,1,2)$, the vector equation of the plane that passes through P and is orthogonal to $\mathbf{v}$ is

$$(x-2,y-1,z-1)(2,1,2)=2(x-2)+(y-1)+2(z-1)=0,$$

and its scalar equation is

$$2x+y+2z=7.$$

2 Differential calculus with one real variable

2.1 Concept of derivative

Increment

Definition 12 (Increment of a variable)**.** An *increment* of a variable x is a change in the value of the variable and is denoted Δx. The increment of a variable x along an interval $[a,b]$ is

$$\Delta x = b - a.$$

Definition 13 (Increment of a function)**.** The *increment* of a function $y = f(x)$ along an interval $[a,b] \subseteq Dom(f)$ is

$$\Delta y = f(b) - f(a).$$

Example The increment of x along the interval $[2,5]$ is $\Delta x = 5 - 2 = 3$ and the increment of the function $y = x^2$ along the same interval is $\Delta y = 5^2 - 2^2 = 21$.

Average rate of change

The study of a function $y = f(x)$ requires to understand how the function changes, that is, how changes the dependent variable y when we change the independent variable x.

Definition 14 (Average rate of change)**.** The *average rate of change* of a function $y = f(x)$ in an interval $[a, a + \Delta x] \subseteq Dom(f)$, is the quotient between the increment of y and the increment of x in that interval, and is denoted

$$\text{ARC}\, f[a, a + \Delta x] = \frac{\Delta y}{\Delta x} = \frac{f(a + \Delta x) - f(a)}{\Delta x}.$$

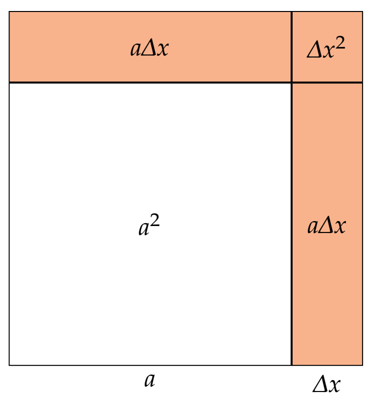
Average rate of change

Example of the area of a square

Let $y = x^2$ be the function that measures the area of a metallic square of side x.

If at any given time the side of the square is a, and we heat the square uniformly increasing the side by dilatation a quantity Δx, how much will increase the area of the square?

$$\begin{aligned}\Delta y &= f(a + \Delta x) - f(a) = (a + \Delta x)^2 - a^2 = \\ &= a^2 + 2a\Delta x + \Delta x^2 - a^2 = 2a\Delta x + \Delta x^2.\end{aligned}$$

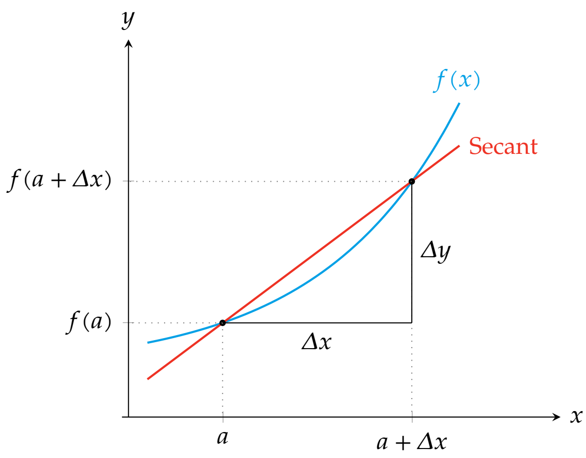


What is the average rate of change in the interval $[a, a + \Delta x]$?

$$\text{ARC}\, f[a, a + \Delta x] = \frac{\Delta y}{\Delta x} = \frac{2a\Delta x + \Delta x^2}{\Delta x} = 2a + \Delta x.$$

Geometric interpretation of the average rate of change

The average rate of change of a function $y = f(x)$ in an interval $[a, a + \Delta x]$ is the slope of the *secant* line of f through the points $(a, f(a))$ and $(a + \Delta x, f(a + \Delta x))$.



Instantaneous rate of change

Often is interesting to study the rate of change of a function, not in an interval, but in a point.

Knowing the tendency of change of a function in an instant can be used to predict the value of the function in next instants.

Definition 15 (Instantaneous rate of change and derivative)**.** The *instantaneous rate of change* of a function f in a point a, is the limit of the average rate of change of f in the interval $[a, a + \Delta x]$, when Δx tends to 0, and is denoted

$$\text{IRC}\, f(a) = \lim_{\Delta x \to 0} \text{ARC}\, f[a, a + \Delta x] = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}$$

When this limit exists, the function f is said to be *derivable* or *differentiable* at the point a, and its value is called *derivative* of f at a, and denoted $f'(a)$ (Lagrange's notation) or $\frac{df}{dx}(a)$ (Leibniz's notation).

Instantaneous rate of change

Example of the area of a square

Let's take again the function $y = x^2$ that measures the area of a metallic square of side x.

If at any given time the side of the square is a, and we heat the square uniformly increasing the side, what is the tendency of change of the area in that moment?

$$\begin{aligned} \text{IRC}\, f(a) &= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x} = \\ &= \lim_{\Delta x \to 0} \frac{2a\Delta x + \Delta x^2}{\Delta x} = \lim_{\Delta x \to 0} 2a + \Delta x = 2a. \end{aligned}$$

Thus,

$$f'(a) = \frac{df}{dx}(a) = 2a,$$

indicating that the area of the square tend to increase the double of the side.

Interpretation of the derivative

The derivative of a function $f'(a)$ shows the growth rate of f at point a:

- $f'(a) > 0$ indicates an increasing tendency (y increases as x increases).
- $f'(a) < 0$ indicates a decreasing tendency (y decreases as x increases).

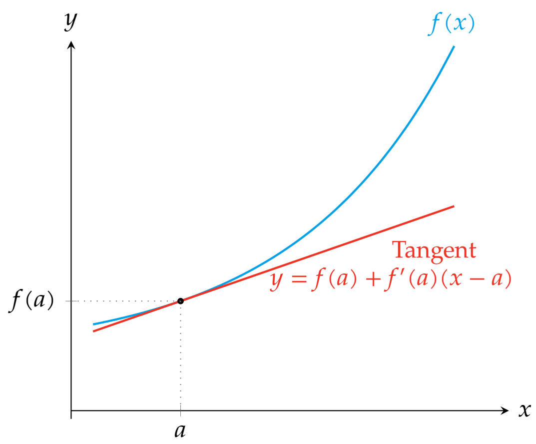
Example A derivative $f'(a) = 3$ indicates that y tends to increase triple of x at point a. A derivative $f'(a) = -0.5$ indicates that y tends to decrease half of x at point a.

Geometric interpretation of the derivative

We have seen that the average rate of change of a function $y = f(x)$ in an interval $[a, a + \Delta x]$ is the slope of the *secant* line, but when Δx tends to 0, the secant line becomes the tangent line.

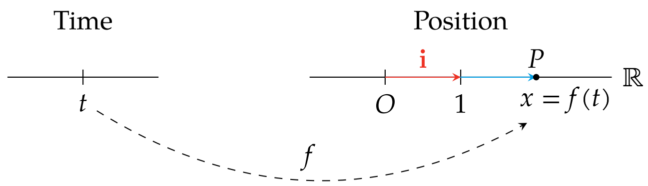
The instantaneous rate of change or derivative of a function $y = f(x)$ at $x = a$ is the slope of the *tangent line* to f at point $(a, f(a))$. Thus, the equation of the tangent line to f at the point $(a, f(a))$ is

$$y - f(a) = f'(a)(x - a) \Leftrightarrow y = f(a) + f'(a)(x - a)$$



Kinematic applications: Linear motion

Assume that the function $y = f(t)$ describes the position of an object moving in the real line at every time t. Taking as reference the coordinates origin O and the unitary vector $\mathbf{i} = (1)$, we can represent the position of the moving object P at every moment t with a vector $\vec{OP} = x\mathbf{i}$ where $x = f(t)$.



Observation It also makes sense when f measures other magnitudes as the temperature of a body, the concentration of a gas, or the quantity of substance in a chemical reaction at every moment t.

Kinematic interpretation of the average rate of change

In this context, if we take the instants $t = t_0$ and $t = t_0 + \Delta t$, both in $\mathrm{Dom}(f)$, the vector

$$\mathbf{v}_m = \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}$$

is known as the *average velocity* of the trajectory f in the interval $[t_0, t_0 + \Delta t]$.

Example A vehicle makes a trip from Madrid to Barcelona. Let $f(t)$ be the function that determine the position of the vehicle at every moment t. If the vehicle departs from Madrid (km 0) at 8:00 and arrive to Barcelona (km 600) at 14:00, then the average velocity of the vehicle in the path is

$$\mathbf{v}_m = \frac{f(14) - f(8)}{14 - 8} = \frac{600 - 0}{6} = 100 km/h.$$

Kinematic interpretation of the derivative

In the same context of the linear motion, the derivative of the function $f(t)$ at the moment t_0 is the vector

$$\mathbf{v} = f'(t_0) = \lim_{\Delta t \to 0} \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t},$$

that is known, as long as the limit exists, as the *instantaneous velocity* or simply *velocity* of the trajectory f at moment t_0.

That is, the derivative of the object position with respect to time is a vector field that is called *velocity along the trajectory f*.

Example Following with the previous example, what indicates the speedometer at any instant is the modulus of the instantaneous velocity vector at that moment.

2.2 Algebra of derivatives

Properties of the derivative

If $y = c$, is a constant function, then $y' = 0$ at any point.

If $y = x$, is the identity function, then $y' = 1$ at any point.

If $u = f(x)$ and $v = g(x)$ are two differentiable functions, then

- $(u + v)' = u' + v'$
- $(u - v)' = u' - v'$
- $(u \cdot v)' = u' \cdot v + u \cdot v'$
- $\left(\dfrac{u}{v}\right)' = \dfrac{u' \cdot v - u \cdot v'}{v^2}$

Derivative of a composite function

The chain rule

Theorem 16 (Chain rule). *If the function $y = f \circ g$ is the composition of two functions $y = f(z)$ and $z = g(x)$, then*

$$(f \circ g)'(x) = f'(g(x))g'(x).$$

It's easy to proof this using the Leibniz notation

$$\frac{dy}{dx} = \frac{dy}{dz}\frac{dz}{dx} = f'(z)g'(x) = f'(g(x))g'(x).$$

Example If $f(z) = \sin z$ and $g(x) = x^2$, then $f \circ g(x) = \sin(x^2)$. Applying the chain rule the derivative of the composite function is

$$(f \circ g)'(x) = f'(g(x))g'(x) = \cos(g(x))2x = \cos(x^2)2x.$$

On the other hand, $g \circ f(z) = (\sin z)^2$, and applying the chain rule again, its derivative is

$$(g \circ f)'(z) = g'(f(z))f'(z) = 2f(z)\cos z = 2\sin z \cos z.$$

Derivative of the inverse of a function

Theorem 17 (Derivative of the inverse function)**.** *Given a function $y = f(x)$ with inverse $x = f^{-1}(y)$, then*

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))},$$

provided that f is differentiable at $f^{-1}(y)$ and $f'(f^{-1}(y)) \neq 0$.

It's easy to proof this using the Leibniz notation

$$\frac{dx}{dy} = \frac{1}{dy/dx} = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}$$

Derivative of the inverse of a function

Example

The inverse of the exponential function $y = f(x) = e^x$ is the natural logarithm $x = f^{-1}(y) = \ln y$, so that we can compute the derivative of the natural logarithm using the previous theorem and we get

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{e^x} = \frac{1}{e^{\ln y}} = \frac{1}{y}.$$

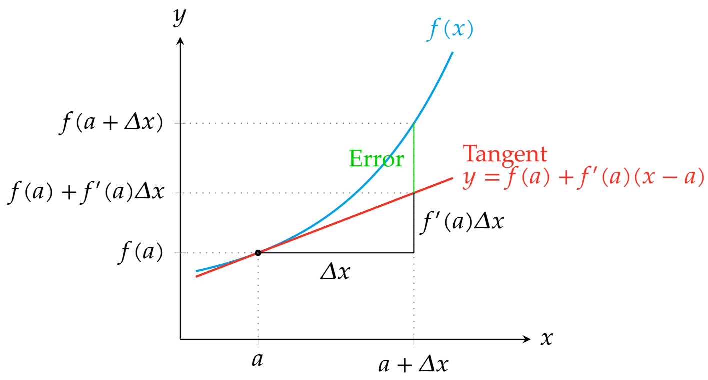
Example Sometimes is easier to apply the chain rule to compute the derivative of the inverse of a function. In this example, as $\ln x$ is the inverse of e^x, we know that $e^{\ln x} = x$, so that differentiating both sides and applying the chain rule to the left side we get

$$(e^{\ln x})' = x' \Leftrightarrow e^{\ln x}(\ln(x))' = 1 \Leftrightarrow (\ln(x))' = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$

2.3 Function approximation

Approximating a function with the derivative

The tangent line to a function $f(x)$ at $x = a$ can be used to approximate f in a neighbourhood of a.



Thus, the increment of a function $f(x)$ in an interval $[a, a+\Delta x]$ can be approximated multiplying the derivative of f at a by the increment of x

$$\Delta y \approx f'(a)\Delta x$$

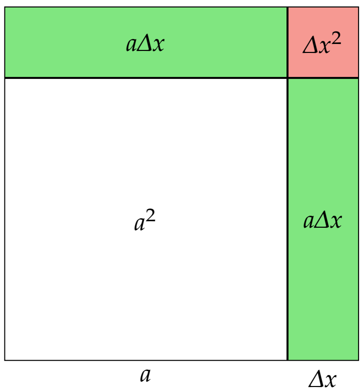
Approximating a function with the derivative

Example of the area of a square

In the previous example of the function $y = x^2$ that measures the area of a metallic square of side x, if the side of the square is a and we increment it a quantity Δx, then the increment on the area will be approximately

$$\Delta y \approx f'(a)\Delta x = 2a\Delta x.$$

In the figure below we can see that the error of this approximation is Δx^2, and is smaller than Δx when Δx tends to 0.



Approximating a function by a polynomial

Another useful application of the derivative is the approximation of functions by polynomials.

Polynomials are functions easy to calculate (sums and products) with very good properties:

- Defined in all the real numbers.
- Continuous.
- Differentiable of all orders with continuous derivatives.

Goal

Approximate a function $f(x)$ by a polynomial $p(x)$ near a value $x = x_0$.

Approximating a function by a polynomial of order 0

A polynomial of grade 0 has equation

$$p(x) = c_0,$$

where c_0 is a constant.

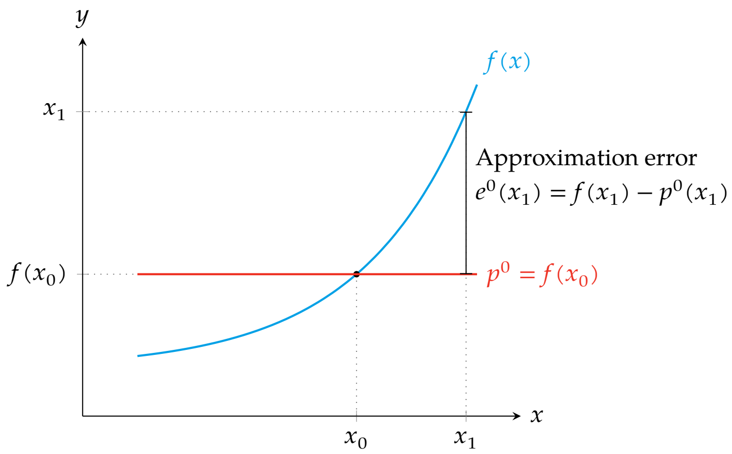
As the polynomial should coincide with the function at x_0, it must satisfy

$$p(x_0) = c_0 = f(x_0).$$

Therefore, the polynomial of grade 0 that best approximate f near x_0 is

$$p(x) = f(x_0).$$

Approximating a function by a polynomial of order 0



Approximating a function by a polynomial of order 1

A polynomial of grade 1 has equation

$$p(x) = c_0 + c_1 x,$$

but it can also be expressed

$$p(x) = c_0 + c_1(x - x_0).$$

Among all the polynomials of grade 1, the one that best approximate $f(x)$ near x_0 is that that meets the following conditions

1. p and f coincide at x_0: $p(x_0) = f(x_0)$,
2. p and f have the same rate of change at x_0: $p'(x_0) = f'(x_0)$.

The last condition guarantee that p and f have approximately the same tendency, but it requires the function f to be differentiable at x_0.

The tangent line: Best approximating polynomial of order 1

Imposing the previous conditions we have

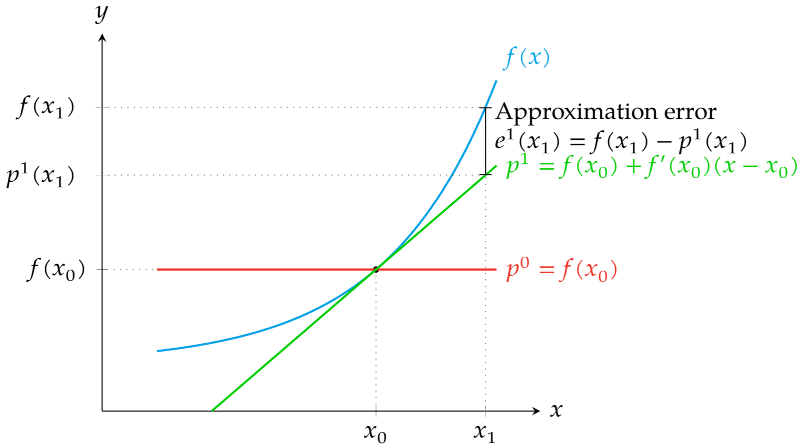
1. $p(x) = c_0 + c_1(x - x_0) \Rightarrow p(x_0) = c_0 + c_1(x_0 - x_0) = c_0 = f(x_0)$,
2. $p'(x) = c_1 \Rightarrow p'(x_0) = c_1 = f'(x_0)$.

Therefore, the polynomial of grade 1 that best approximates f near x_0 is

$$p(x) = f(x_0) + f'(x_0)(x - x_0),$$

which turns out to be the tangent line to f at $(x_0, f(x_0))$.

Approximating a function by a polynomial of order 1



Approximating a function by a polynomial of order 2

A polynomial of grade 2 is a parable with equation

$$p(x) = c_0 + c_1x + c_2x^2,$$

but it can also be expressed

$$p(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)^2.$$

Among all the polynomials of grade 2, the one that best approximate $f(x)$ near x_0 is that that meets the following conditions

1. p and f coincide at x_0: $p(x_0) = f(x_0)$,
2. p and f have the same rate of change at x_0: $p'(x_0) = f'(x_0)$.
3. p and f have the same concavity at x_0: $p''(x_0) = f''(x_0)$.

The last condition requires the function f to be differentiable twice at x_0.

Best approximating polynomial of order 2

Imposing the previous conditions we have

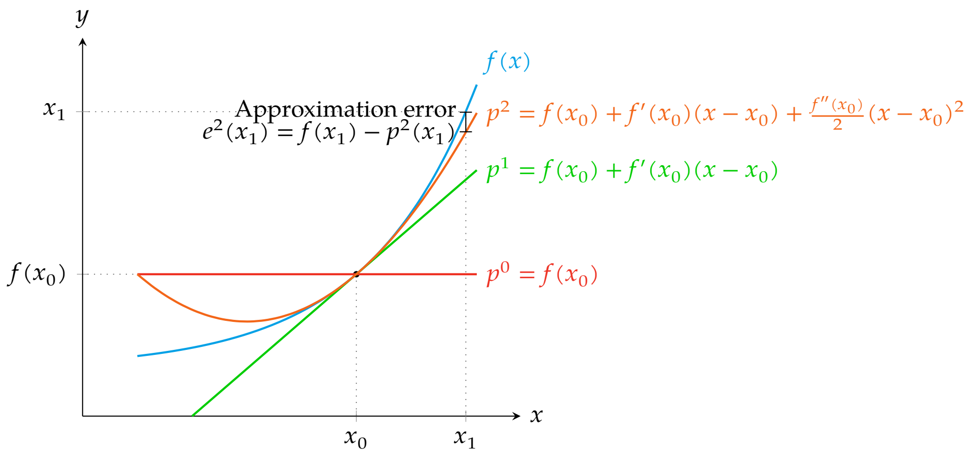
1. $p(x) = c_0 + c_1(x - x_0) \Rightarrow p(x_0) = c_0 + c_1(x_0 - x_0) = c_0 = f(x_0)$,

2. $p'(x) = c_1 \Rightarrow p'(x_0) = c_1 = f'(x_0)$.
3. $p''(x) = 2c_2 \Rightarrow p''(x_0) = 2c_2 = f''(x_0) \Rightarrow c_2 = \frac{f''(x_0)}{2}$.

Therefore, the polynomial of grade 2 that best approximates f near x_0 is

$$p(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

Approximating a function by a polynomial of order 2



Approximating a function by a polynomial of order n

A polynomial of grade n has equation

$$p(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n,$$

but it can also be expressed

$$p(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + \cdots + c_n(x - x_0)^n.$$

Among all the polynomials of grade 2, the one that best approximate $f(x)$ near x_0 is that that meets the following $n + 1$ conditions

1. $p(x_0) = f(x_0)$,
2. $p'(x_0) = f'(x_0)$,
3. $p''(x_0) = f''(x_0)$,

   ...

n+1. $p^{(n)}(x_0) = f^{(n)}(x_0)$.

Observe that these conditions require the function f to be differentiable n times at x_0.

Coefficients calculation for the best approximating polynomial of order n

The successive derivatives of p are

$$\begin{aligned} p(x) &= c_0 + c_1(x-x_0) + c_2(x-x_0)^2 + \cdots + c_n(x-x_0)^n, \\ p'(x) &= c_1 + 2c_2(x-x_0) + \cdots + nc_n(x-x_0)^{n-1}, \\ p''(x) &= 2c_2 + \cdots + n(n-1)c_n(x-x_0)^{n-2}, \\ &\vdots \\ p^{(n)}(x) &= n(n-1)(n-2)\cdots 1c_n = n!\,c_n. \end{aligned}$$

Imposing the previous conditions we have

1. $p(x_0) = c_0 + c_1(x_0-x_0) + c_2(x_0-x_0)^2 + \cdots + c_n(x_0-x_0)^n = c_0 = f(x_0)$,
2. $p'(x_0) = c_1 + 2c_2(x_0-x_0) + \cdots + nc_n(x_0-x_0)^{n-1} = c_1 = f'(x_0)$,
3. $p''(x_0) = 2c_2 + \cdots + n(n-1)c_n(x_0-x_0)^{n-2} = 2c_2 = f''(x_0) \Rightarrow c_2 = f''(x_0)/2$,

   ...

n+1. $p^{(n)}(x_0) = n!\,c_n = f^{(n)}(x_0) = c_n = \frac{f^{(n)}(x_0)}{n!}$.

Taylor polynomial of order n

Definition 18 (Taylor polynomial)**.** Given a function $f(x)$ differentiable n times at x_0, the *Taylor polynomial* of order n of f at x_0 is the polynomial with equation

$$\begin{aligned} p_{f,x_0}^n(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n = \\ &= \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i. \end{aligned}$$

The Taylor polynomial of order n of f at x_0 is the nth degree polynomial that best approximates f near x_0, as is the only one that meets the previous conditions.

Taylor polynomial calculation

Example

Let's approximate the function $f(x) = \log x$ near the value 1 by a polynomial of order 3.

The equation of the Taylor polynomial of order 3 of f at $x_0 = 1$ is

$$p_{f,1}^3(x) = f(1) + f'(1)(x-1) + \frac{f''(1)}{2}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3.$$

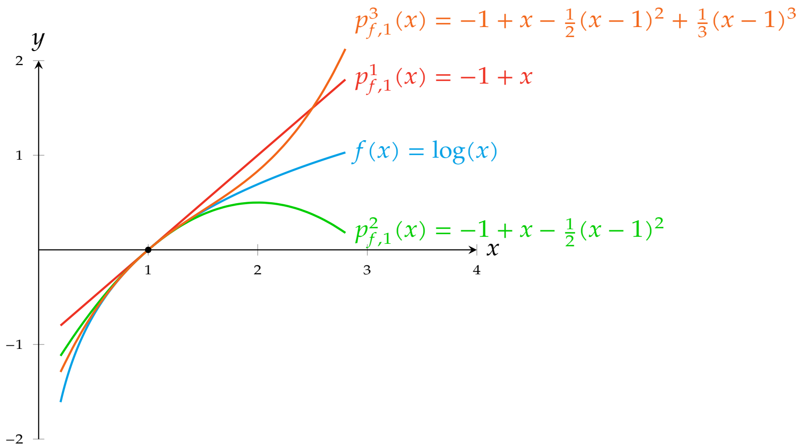
The derivatives of f at 1 up to order 3 are

$$\begin{array}{ll} f(x) = \log x & f(1) = \log 1 = 0, \\ f'(x) = 1/x & f'(1) = 1/1 = 1, \\ f''(x) = -1/x^2 & f''(1) = -1/1^2 = -1, \\ f'''(x) = 2/x^3 & f'''(1) = 2/1^3 = 2. \end{array}$$

And substituting into the polynomial equation we get

$$p_{f,1}^3(x) = 0 + 1(x-1) + \frac{-1}{2}(x-1)^2 + \frac{2}{3!}(x-1)^3 = \frac{2}{3}x^3 - \frac{3}{2}x^2 + 3x - \frac{11}{6}.$$

Taylor polynomials of the logarithmic function



Maclaurin polynomial of order n

The Taylor polynomial equation simplifies when the polynomial is calculated at 0. This special case of Taylor polynomial at 0 is known as the *Maclaurin polynomial*.

Definition 19 (Maclaurin polynomial)**.** Given a function $f(x)$ differentiable n times at 0, the *Maclaurin polynomial* of order n of f is the polynomial with equation

$$\begin{aligned} p^n_{f,0}(x) &= f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n = \\ &= \sum_{i=0}^{n} \frac{f^{(i)}(0)}{i!}x^i. \end{aligned}$$

Maclaurin polynomial calculation

Example

Let's approximate the function $f(x) = \sin x$ near the value 0 by a polynomial of order 3.

The Maclaurin polynomial equation of order 3 of f is

$$p^3_{f,0}(x) = f(0) + f'(0)x + \frac{f''(0)}{2}x^2 + \frac{f'''(0)}{3!}x^3.$$

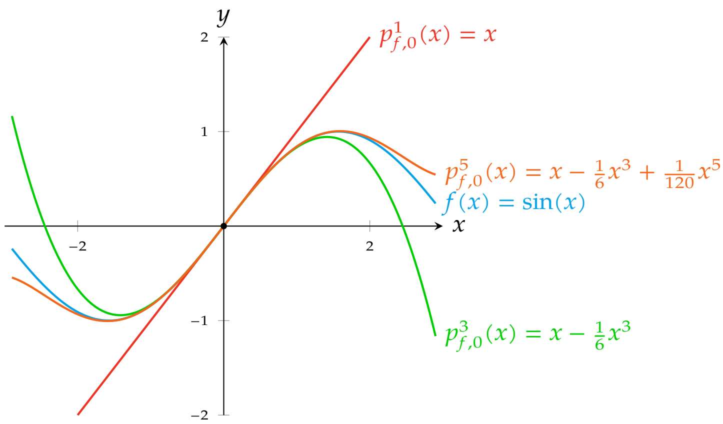
The derivatives of f at 0 up to order 3 are

$$\begin{array}{ll} f(x) = \sin x & f(0) = \sin 0 = 0, \\ f'(x) = \cos x & f'(0) = \cos 0 = 1, \\ f''(x) = -\sin x & f''(0) = -\sin 0 = 0, \\ f'''(x) = -\cos x & f'''(0) = -\cos 0 = -1. \end{array}$$

And substituting into the polynomial equation we get

$$p^3_{f,0}(x) = 0 + 1 \cdot x + \frac{0}{2}x^2 + \frac{-1}{3!}x^3 = x - \frac{x^3}{6}.$$

Maclaurin polynomial of the sine function



Maclaurin polynomials of elementary functions

$f(x)$	$p^n_{f,0}(x)$
$\sin x$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^k \frac{x^{2k-1}}{(2k-1)!}$ if $n = 2k$ or $n = 2k-1$
$\cos x$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^k \frac{x^{2k}}{(2k)!}$ if $n = 2k$ or $n = 2k+1$
$\arctan x$	$x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^k \frac{x^{2k-1}}{(2k-1)}$ if $n = 2k$ or $n = 2k-1$
e^x	$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$
$\log(1+x)$	$x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n}$

Taylor remainder and Taylor formula

Taylor polynomials allow to approximate a function in a neighborhood of a value x_0, but there is always an error in the approximation.

Definition 20 (Taylor remainder)**.** Given a function $f(x)$ and its Taylor polynomial of order n at x_0, $p^n_{f,x_0}(x)$, the *Taylor remainder* of order n of f at x_0 is de difference between the function and the polynomial,

$$r^n_{f,x_0}(x) = f(x) - p^n_{f,x_0}(x).$$

The Taylor remainder measures the error in the approximation of $f(x)$ by the Taylor polynomial and allow us to express the function as the Taylor polynomial plus the Taylor remainder

$$f(x) = p^n_{f,x_0}(x) + r^n_{f,x_0}(x).$$

This expression is known as *Taylor formula* of order n or f at x_0.

It can be proved that

$$\lim_{h\to 0} \frac{r^n_{f,x_0}(x_0+h)}{h^n} = 0,$$

which means that the remainder $r^n_{f,x_0}(x_0+h)$ is much less than h^n.

2.4 Analysis of functions

Analysis of functions: increase and decrease

The main application of derivatives is to determine the variation (increase or decrease) of functions. For that we use the sign of the first derivative.

Theorem 21. *Let $f(x)$ be a function with first derivative in an interval $I \subseteq \mathbb{R}$.*

- *If $\forall x \in I\ f'(x) \geq 0$ then f is increasing on I.*
- *If $\forall x \in I\ f'(x) \leq 0$ then f is decreasing on I.*

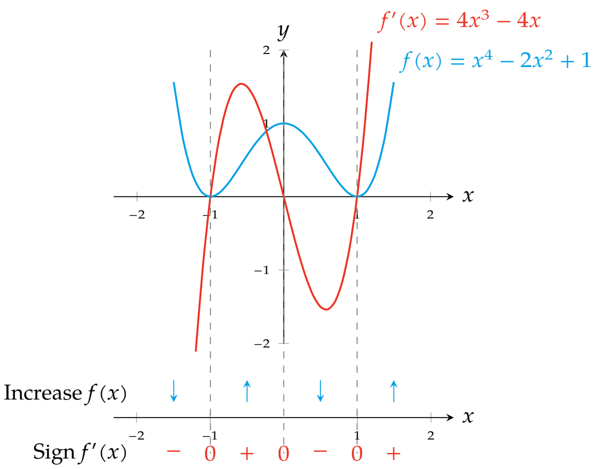
If $f'(x_0) = 0$ then x_0 is known as a *stationary point* and the function in non-increasing and non-decreasing at that point. **Example** The function $f(x) = x^3$ is increasing on $\mathbb{R}$ as $\forall x \in \mathbb{R}\ f'(x) \geq 0$.

Observation A function can be increasing or decreasing on an interval and not have first derivative.

Analysis of functions: increase and decrease

Example

Let's analyze the increase and decrease of the function $f(x) = x^4 - 2x^2 + 1$. Its first derivative is $f'(x) = 4x^3 - 4x$.



Analysis of functions: relative extrema

As a consequence of the previous result we can also use the first derivative to determine the relative extrema of a function.

Theorem 22 (First derivative test)**.** *Let $f(x)$ be a function with first derivative in an interval $I \subseteq \mathbb{R}$ and let $x_0 \in I$ be a stationary point of f ($f'(x_0) = 0$).*

- *If $f'(x) > 0$ on an open interval extending left from x_0 and $f'(x) < 0$ on an open interval extending right from x_0, then f has a relative maximum at x_0.*
- *If $f'(x) < 0$ on an open interval extending left from x_0 and $f'(x) > 0$ on an open interval extending right from x_0, then f has a relative minimum at x_0.*
- *If $f'(x)$ has the same sign on both an open interval extending left from x_0 and an open interval extending right from x_0, then f has an inflection point at x_0.*

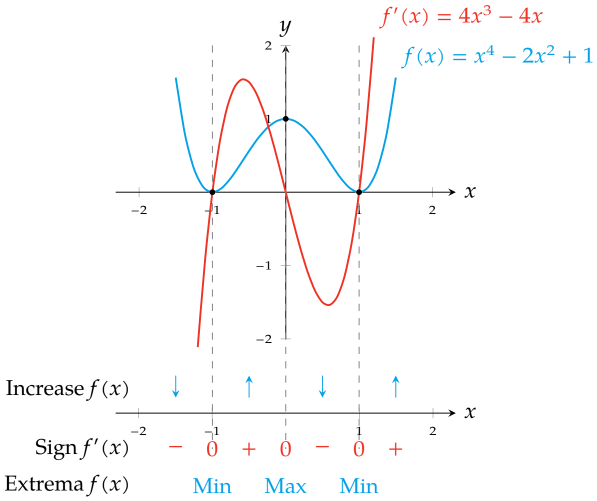
Observation *A vanishing derivative is a necessary but not sufficient condition for the function to have a relative extrema at a point.*

Example The function $f(x) = x^3$ have derivative $f'(x) = 3x^2$ and have a stationary point at $x = 0$. However it doesn't have a relative extrema at that point, but an inflection point.

Analysis of functions: relative extrema

Example

Consider again the function $f(x) = x^4 - 2x^2 + 1$ and let's analyze its relative extrema now. Its first derivative is $f'(x) = 4x^3 - 4x$.



Analysis of functions: concavity

The concavity of a function can be determined by de second derivative.

Theorem 23. *Let $f(x)$ be a function with second derivative in an interval $I \subseteq \mathbb{R}$.*

- *If $\forall x \in I\ f''(x) \geq 0$ then f is concave up (convex) on I.*
- *If $\forall x \in I\ f''(x) \leq 0$ then f is concave down (concave) on I.*

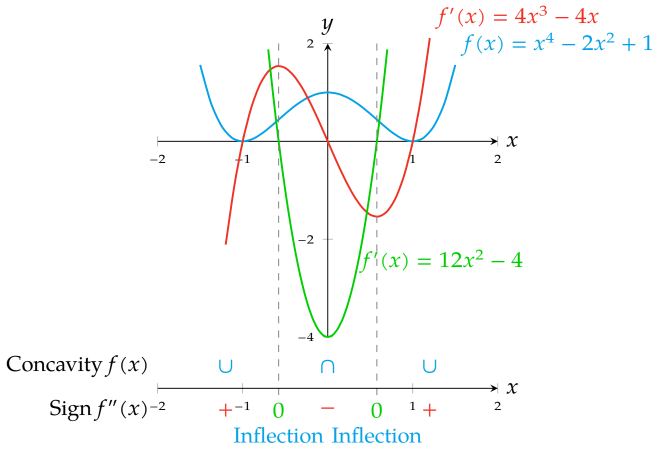
Example The function $f(x) = x^2$ has second derivative $f''(x) = 2 > 0\ \forall x \in \mathbb{R}$, so it is concave up in all $\mathbb{R}$.

Observation *A function can be concave up or down and not have second derivative.*

Analysis of functions: concavity

Example

Let's analyze the concavity of the same function of previous examples $f(x) = x^4 - 2x^2 + 1$. Its second derivative is $f''(x) = 12x^2 - 4$.


$f'(x) = 4x^3 - 4x$
$f(x) = x^4 - 2x^2 + 1$
$f'(x) = 12x^2 - 4$
y
x
2
−2
−4
−2
−1
1
2
Concavity $f(x)$
∪
∩
∪
Sign $f''(x)$
+
0
−
0
+
Inflection Inflection