The analytical approach: an introduction to optimisation for false beginners

Useful math symbols

(For the sake of completeness an exhaustive list can be found HERE)

- $\forall$ for all (universal quantifier)
- $\exists$ there exists (existential quantifier)
- $\exists!$ there exists and is unique.
- $|$ such that.
- $:$ it follows.
- $\Rightarrow$ implicates.
- $\Leftrightarrow$ if and only if.
- $\in$ belong to/is in.
- $\{\dots\}$ set.
- $\emptyset$ empty set.
- $|\dots|$ cardinality/module.
- $\subseteq$ subset/contained.
- $\cup$ set union.
- $\cap$ set intersection.
- $\setminus$ set difference.
- $\times$ Cartesian product.
- $\wedge$ and.
- $\vee$ or.
- $\equiv$ is equivalent.
- $\approx$ is approximately.
- ∞ infinity.
- $\notin, \nexists, \neq \dots$ negations.
- $\circ$ function composition.

Set

- A set is a *primitive* concept that we can picture as a collection of objects, e.g. the set of the white shirts.
- These objects are the elements of the set. Let us indicate with A a generic set and with x its generic element, then we can state $x \in A$.
- The cardinality of a set A is the number of elements contained in A, e.g. $|A| = \{\text{\LaTeX}, 1, C\} = 3$.
- A set A is said empty when it does not contain any element $(|A| = 0 \Rightarrow A \equiv \emptyset)$.
- A set is said finite if its cardinality is a finite number $(|A| = n \in \mathbb{N})$. Conversely, it is called infinite set $(|A| = \infty)$.

Basic operations between sets

Union, intersection, difference and Cartesian product.

Given two sets $A = \{a_1, a_2, a_3, \dots\}$ and $B = \{b_1, b_2, b_3, \dots\}$:

$$A \cup B = \{x | x \in A \vee x \in B\} \qquad A \cap B = \{x | x \in A \wedge x \in B\}$$

$$A \setminus B = \{x | x \in A \wedge x \notin B\}$$

$$A \times B = \{(a, b) | a \in A \wedge b \in B\} =$$
$$\{(a_1, b_1), (a_1, b_2), (a_1, b_3), \dots, (a_2, b_1), (a_2, b_2), (a_2, b_3), \dots\}$$

The Cartesian product is a new set containing all the possible ordered pairs!

This is of particular interest for us, as the search space for a box-constrained optimisation problem can be seen as the product between the admissible ranges of the design variables.

Order relation

- Let $= A \times B$ be a Cartesian product. A relation on C is an arbitrary subset $R \subseteq C$. This subset means that some elements of A relates to B according to a certain criterion R $(aRb, a \in A, b \in B)$.

- A set A together with an order relation "$\leq$" is said to be **ordered** with respect to $\leq$. A relation "$\leq$" is an order relation if the following properties are verified:

  Reflexivity $\forall a \in A : a \leq a$.

  Transitivity $\forall a_1, a_2, a_3 \in A$: if $a_1 \leq a_2 \wedge a_2 \leq a_3 \Rightarrow a_1 \leq a_3$.

  Antisymmetry $\forall a_1, a_2 \in A$: if $a_1 \leq a_2$ then $a_2 \not\leq a_1$.

Dense and discrete sets

Let A be a ordinate set.

- if $\forall a_0 \in A : \exists a \in A || a - a_0 | < \mathcal{E}$ regardless how small $\mathcal{E}$ is taken, A is said dense set.
- Conversely, A is said to be discrete.

Alternatively, we can state that in a discrete set $\exists$ a radius $\mathcal{E}|$ the neighborhood N centred in the point $a_0 \in A$, namely $(a_0 - \mathcal{E}, a_0 + \mathcal{E})$, does not contain further elements apart from a_0: $A \cap N = \{a_0\}$. Conversely that value of the radius does not exist in a dense set.

Number sets

- Natural numbers $\mathbb{N}$ is the discrete set $\mathbb{N} = \{0, 1, 2, \dots\}$
- Relative numbers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.
- Rational numbers $\mathbb{Q}$ contains all the possible fractions $\frac{x}{y}$, $(x, y \in \mathbb{Z}, y \neq 0)$.
- Real numbers $\mathbb{R}$: the set containing Q and all the other decimal numbers that cannot be expressed as fractions of relative numbers.
- Complex numbers $\mathbb{C}$: the set of numbers that can be expressed as $a + ib$ where $a, b \in \mathbb{R}$ and the imaginary unit $i = \sqrt{-1}$.

Interval

. . . in $\mathbb{R}$

An interval is a special dense set, as it is a subset of $\mathbb{R}$:

$$(a, b) =]a, b[= \{x \in \mathbb{R} | a < x < b\}$$
$$[a, b) = [a, b[= \{x \in \mathbb{R} | a \leq x < b\}$$
$$(a, b] =]a, b] = \{x \in \mathbb{R} | a < x \leq b\}$$
$$[a, b] = \{x \in \mathbb{R} | a \leq x \leq b\}$$

Function

A relation is said to be a mapping or function (f) when it relates to any element of a set a unique element of another. Let A (domain) and B (co-domain) be two sets, a mapping $f : A \to B$ is a relation $f \subseteq A \times B | \forall a \in A : \exists! b \in B | (a, b) \in f$

$f : A \to B$ is said to be injective if:

$\forall a_1, a_2 \in A, f(a_1) = f(a_2) \Rightarrow a_1 = a_2$

$f : A \to B$ is said to be surjective if:

$\forall a \in A \exists b \in B \; | f(a) = b$

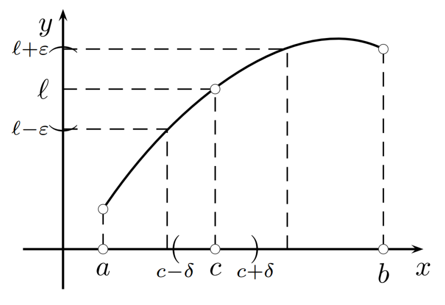
f is said to be bijective if both injective and surjective

Limit

(ϵ, δ) definition.

Given an interval $(a, b) \in \mathbb{R}$, a point $c \in (a, b)$, and a function $f : (a, b) \setminus \{c\} :\rightarrow \mathbb{R}$, then

$$\lim_{x \rightarrow c} f(x) = l \in \mathbb{R}$$



means that:

$$\forall \epsilon > 0 \exists \delta > 0 | \forall x \neq c |x - c| < \delta \Rightarrow |f(x) - l| < \epsilon$$

The limit studies the local behaviour of f

This definition can be extended for $l = \pm\infty$ and $c = \pm\infty$

(examples and exercises with solutions HERE and HERE)

Continuous and discrete functions

- A function defined over a discrete set is a discrete function (Figure 1), i.e. a discrete set itself.
- Conversely, a function $f : X \to Y$ defined on a dense set is not necessarily continuous! (Figure 2).
  - If $\lim_{x \to c} f(x) = f(c)$ then f is continuous in the point c.
  - If this condition is satisfied $\forall x \in X$ then f is said continuous (Figure 3).

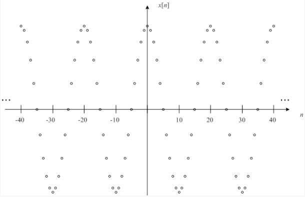


Figure 1: Discrete

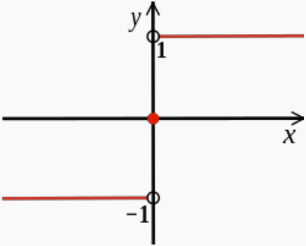


Figure 2: Discrete

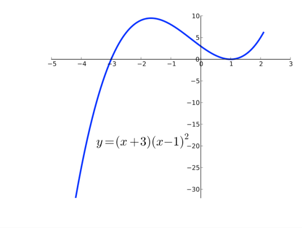


Figure 3: Continuous

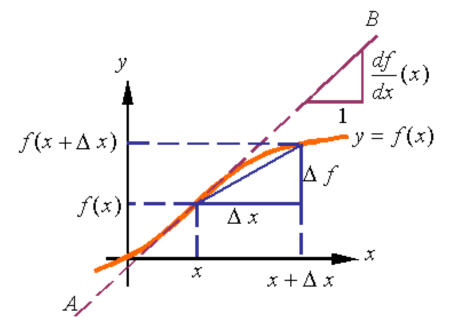
Derivative

The derivative of a function $y = f(x)$ in a point x_0 is the limit value of the ratio of the differences $\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x)}{\Delta x} = \frac{f(x_0)-f(x)}{x_0-x}$ as Δx becomes infinitely small:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x)}{\Delta x}$$

or equivalently:

$$\lim_{x \to x_0} \frac{f(x_0) - f(x)}{x_0 - x}$$



The derivative measure the sensitivity to changes in a point x_0 (i.e. angular coefficient of the tangent line in x_0)

If calculated in all the points it is itself a function $\frac{df(x)}{dx} = f'(x)$.

Basic derivatives

$f(x)$	$f'(x)$
$c \in \mathbb{R}$	0
x	1
$x^n,\ n \in \mathbb{Q}$	nx^{n-1}
$sin(x)$	$cos(x)$
$cos(x)$	$-sin(x)$
$log(x)$	$\frac{1}{x},\ x > 0$
e^x	e^x
$\lvert x \rvert$	$\frac{x}{\lvert x \rvert}$

(see more HERE)

Fundamental theorems of derivatives

Let f and g be two functions defined over $\mathbb{R}$ $(a, b, \in \mathbb{R})$.

- Differentiation is linear: $\frac{d\left(af(x)+bg(x)\right)}{dx} = a\frac{df(x)}{dx} + b\frac{dg(x)}{dx}$.
- Product rule: $\frac{d\left(f(x)g(x)\right)}{dx} = \frac{df(x)}{dx}g(x) + f(x)\frac{dg(x)}{dx}$.
- Quotient rule: $\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{\frac{df(x)}{dx}g(x)-f(x)\frac{dg(x)}{dx}}{\left(g(x)\right)^2}$.
- Chain rule: $\frac{df\circ g(x)}{dx} = \frac{f\left(g(x)\right)}{dx} = \frac{df\left(g(x)\right)}{dg(x)}\frac{dg(x)}{dx}$.

(more details HERE)

Examples

$f(x)$	$f'(x)$
$x^3 + 5x^2 + 7e^x + 8$	$3x^2 + 10x + 7e^x$
$x^2 sin(x)$	$2xsin(x) + x^2 cos(x)$
$\frac{cos(x)}{x^2}$	$\frac{-x^2 sin(x) - 2xcos(x)}{x^4}$
$log(x^2 - 2)$	$\frac{2x}{\lvert x^2-2 \rvert}$
$5^{sin(x)}$	$ln(5)5^{sin(x)} cos(x)$

(see more HERE)

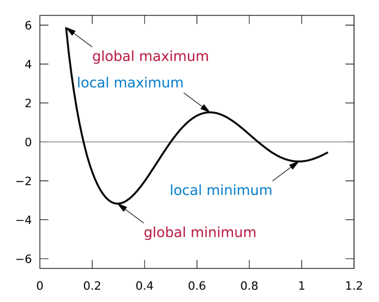
Minima and maxima
(in one variable)

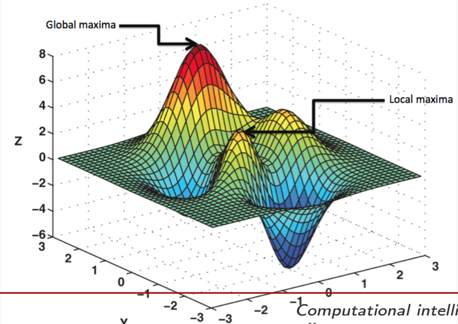
Given a function $f : X \to Y$, a point $x_0 \in X$ is local minimum if:

$$\exists\epsilon \in \mathbb{R} | \forall x \in (x_0 - \epsilon, x_0 + \epsilon) : f(x_0) < f(x)$$

and global minimum if:

$$\forall x \in X : f(x_0) < f(x)$$





N.B. change $<$ into $>$ for local and global maximum

These definitions can be easily extended for a multivariate case.

Optimisation problem:
minimisation or maximisation?

- An optimisation problem consists in finding the global minimum, that is that x_0 such that the objective-function value is the lowest/highest within the decision space (also called search space, i.e. the domain).
- It can be easily verified that $max(f(x)) = min(-f(x))$. Since the problems are equivalent, in this module (as in most papers), we will conventionally refer all the time to minimisation.

Analytical approach

- A continuous (differentiable) function having an explicit analytical expression can be tackled analytically.
- The main theoretical piece of information is that the gradient in a local/global minimum is a null vector (thus in one variable $f'(x_0) = 0$).

Exact solution!

No need to use CI!

Most real-world problem are not differentiable!

If the number of variable is high optimisation can be hard and time consuming anyhow!

Optimisation in one variable

1. All the points having null derivative (stationary/critical points) must be found.
2. Second order derivative. If positive, the external point is a minimum, if negative is a maximum.
3. The local minimum with the lowest function value is the global minimum (the bounds of the search space and critical points where $f^{'}$ is not defined must be checked as well).

N.B. if $\frac{d^2}{dx^2} = 0$ the stationary point is a flex!

A flex is neither a minimum nor a maximum, but the point where the concavity of the function changes!

Example 1

Find the global minimum of:

$f(x) = x^2 + 1$ in the interval $[-4, 4]$

This is parabola $\Rightarrow$ we know there is only a global minimum!
Let us follow the procedure:

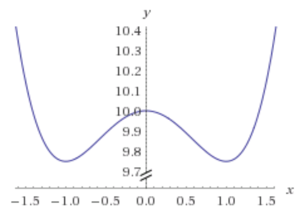
- $\frac{df(x)}{dx} = 2x \Rightarrow x = 0$ (1 stationary point).
- $\frac{d^2f(x)}{dx^2} = \frac{df^{'}(x)}{dx} = 2 > 0 \Rightarrow$ *minimum*!
- Bounds chek: $f(-4) = f(4) > f(0) = 1 \Rightarrow x = 0$ is the solution!

Example 2

Find the global minimum of:

$f(x) = \frac{1}{4}x^4 - \frac{1}{2}x^2 + 10$ in the range [0,0.5].

- $\frac{df(x)}{dx} = x^3 - x = x(x-1)(x+1)$
  $\Rightarrow$ stationary points $x = 0, \pm 1$.
- $\frac{df^2(x)}{dx} = 3x^2 - 1$ positive for:
  $x < \sqrt{\frac{1}{3}}$ and $x > \sqrt{\frac{1}{3}}$.



$\Rightarrow x = 0$ is a local maximum, $x = \pm 1$ local minima.

Since we are interested in $[0, 0.5]$, we don't consider the actual minima but the solution $x = 0.5$, where the functional value is 9.8754.

Example 3

Find the global minimum of:

$f(x) = xe^{-x^2}$ in the interval $[-100, 100]$

- $\frac{df(x)}{dx} = e^{-x^2} - 2x^2e^{-x^2} = e^{-x^2}(1-2x^2) \Rightarrow x_{stationary} = \pm\sqrt{\frac{1}{2}}$.
- $\frac{d^2f(x)}{dx^2} = -4xe^{-x^2} - 2x(1-2x^2)e^{-x^2}$.
- $\left.\frac{d^2f(x)}{dx^2}\right|_{x=-\sqrt{\frac{1}{2}}} \approx 1.71553 \Rightarrow$ (global) minimum$\in \mathbb{R}$.
  $\left.\frac{d^2f(x)}{dx^2}\right|_{x=\sqrt{\frac{1}{2}}} \approx -1.71553 \Rightarrow$ (global) maximum over $\mathbb{R}$.

$-\sqrt{\frac{1}{2}} \in [-100, 100] \Rightarrow$ is our solution with a corresponding functional value of 0.429.

Laboratory and participation work

find the global minimum of the following functions:

1. $f(x) = \sqrt{x}$ in the interval $[0, 10]$.
2. $f(x) = x^3$ in $\mathbb{R}$ and in the interval $[-1, 1]$.
3. $f(x) = x^6 - e^4x^4 + ln(5)$ in the interval $[-10, 0]$.
4. $f(x) = \frac{x}{ln(x)}$ in the interval $[0, \infty]$.
5. $f(x) = sin(x) - x$ in $\mathbb{R}$ and in the interval $[0, \pi]$.
6. $f(x) = cos(2x) - x$ in the interval $[0, \frac{2\pi}{3}]$.

Show all the calculations and comment on them.