MATH3611 / MATH5705

Chapter 0: Introduction

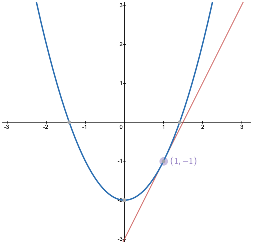
Roughly speaking, *analysis* (in the mathematical sense) is a generalization of calculus. But the topics, questions, and methods that come up often look quite different than what you may have seen in first and second year calculus courses! Rather than precisely defining analysis, in this introduction we will consider a few questions and concepts to motivate the material that we will study in this course.

Newton's method to find the square root of 2

We can use calculus to help solve various types of equations. For example, suppose we want to find the square root of 2, or equivalently a (positive) root of the differentiable function

$$f(x) = x^2 - 2.$$

Newton's method
is an iterative procedure, where we start with an initial value x_0, and then from each x_n we choose the next value x_{n+1} to be the root of the linear function which passes through the point $(x_n, f(x_n))$ and has slope $f'(x_n)$:



(from *desmos.com*)

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

The square root of 2 is clearly between 1 and 2, so we might try starting at $x_0 = 1$. It turns out that Newton's method then actually gives a sequence which converges to $\sqrt{2}$.

Now let's consider a differential equation. Differential equations are ubiquitous across all areas of science, and are often used to model the dynamics of systems (for example an epidemic spread). Consider the differential equation

$$y' = \cos(xy).$$

Question: Does this equation have a solution? And if so can we find one, or at least approximate one?

Here it's understood that x is a real variable, and $y = y(x)$ is function. To specify a solution uniquely, we would presumably need an initial condition such as

$$y(0) = 1.$$

In this case, the variable y that we are trying to solve for is not a *number*, but is rather a *function*. If we are to have any hope of using an iterative method to approximate a solution to a differential equation, we'll need to develop a theory of convergence for functions.

For a function $f : \mathbb{R} \to \mathbb{R}$,

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

means
"For any $\epsilon > 0$, there is a $\delta(\epsilon)$ such that $|f(x) - L| < \epsilon$ when $|x - a| < \delta$."

The absolute value of a difference corresponds to *distance* on the real line:

$$d(x, y) = |x - y|.$$

Or:

"For any $\epsilon > 0$, there is a $\delta(\epsilon)$ such that $d(f(x), L) < \epsilon$ whenever $d(x, a) < \delta$,"

where by definition $d(x, y) = |x - y|$. This generalizes nicely to two dimensions. For $f : \mathbb{R}^2 \to \mathbb{R}$, what does it mean to say

$$\lim_{(x,y) \to (a,b)} f(x, y) = L?$$

Question: What does $d((x_1, y_1),(x_2, y_2))$ mean in $\mathbb{R}^2$?

Some possibilities:

- $\sqrt{|x_1 - x_2|^2 + |y_1 - y_2|^2}$ ("Euclidean distance" or "2-distance ")
- $|x_1 - x_2| + |y_1 - y_2|$ ("1-distance")
- $(|x_1 - x_2|^p + |y_1 - y_2|^p)^{\frac{1}{p}}$ for some fixed number $p > 1$ ("p-distance")
- $\max\{|x_1 - x_2|, |y_1 - y_2|\}$ ("∞-distance")

Question: What does the set

$$\{(x, y) : d_p((x, y), (0, 0)) \leq 1\}$$

look like for $p = 1, 2, 3, \infty$?

Question: In the definition of a limit for $f : \mathbb{R}^2 \to \mathbb{R}$, does it matter which p we use? Why or why not?

The answer is that for the definition of limits it doesn't matter which p we use. Try to think about why the two definitions of limit (using $p = 1$ or $p = 2$) are equivalent. Later in the course, we will say that d_1 and d_2 are different "metrics" but give the same "topology". The *metric* is the notion of distance, while the *topology* is the notion of convergence (we will give precise definitions later). As we have seen for $\mathbb{R}$ and $\mathbb{R}^2$, one way to define limits/convergence is in terms of distance between points. Two main abstractions in this course:

- **Metric** - an abstract notion of distance in a space (not necessarily $\mathbb{R}$ or $\mathbb{R}^n$)
- **Topology** - an abstract notion of convergence (even in spaces when there is no underlying notion of distance).

The notion of p-distance makes sense in $\mathbb{R}^n$. For fixed $p > 1$ we can define for the p-distance of two vectors

$$\mathbf{x} = (x_1, \dots, x_n) \text{ and } \mathbf{y} = (y_1, \dots, y_n)$$

as

$$d_p(\mathbf{x}, \mathbf{y}) = \left(\sum_{k=1}^{n} |x_k - y_k|^p \right)^{\frac{1}{p}};$$

and

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, \dots, |x_n - y_n|\}.$$

Question: What about for infinite sequences $\mathbf{x} = (x_1, x_2, \dots)$ and $\mathbf{y} = (y_1, y_2, \dots)$?

Answer: Not in general. The formula above would now require an infinite sum, and for arbitrary sequences this doesn't necessarily converge. As we'll see later in the course, we can sometimes work around this by restricting our attention to only those sequences for which the appropriate sum converges (this condition depends on p).

What about functions? How do we measure the "distance" between two functions? For simplicity let's restrict our attention to the continuous real-valued functions on the interval $[0, 1]$, which we call $C[0, 1]$.

Question: How can we define the distance between two functions in $C[0, 1]$?

Some possibilities:

- The "maximum vertical gap" between the graphs of the functions - formally:

$$d_\infty(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|.$$

(Question: does this always exist?)

- The area between the graphs of the functions - formally:

$$d_1(f,g) = \int_0^1 |f(x) - g(x)| \, dx.$$

(Question: does this always exist?)

- For any $p > 1$, we can define d_p in a similar way:

$$d_p(f,g) = \left(\int_0^1 |f(x) - g(x)|^p \, dx \right)^{\frac{1}{p}}.$$

Let's look at a couple of examples in $C[0,1]$.

Example

Let

$$f_n(x) = \begin{cases} n - n^2 x & 0 \le x \le \frac{1}{n} \\ 0 & \frac{1}{n} < x \le 1 \end{cases}.$$

(Draw it!) What is the distance between f_n and the constant 0 function? (For d_1 and d_∞). (Depends on the metric!) Try d_1 and d_∞.

Example

What is the distance between $g(x) = x^n$ and the constant 0 function on $[0, 1]$?
(Again, answer for both d_1 and d_∞.)

Analysis on vector spaces

As we can see from these examples, deciding what is meant by distance or convergence is not so straightforward once we are dealing with complicated things like functions, and the "correct" definition will depend on what you are trying to do. This will be a major theme in this course. In many examples of interest, there is also an underlying vector space structure, and a large part of modern analysis deals with studying convergence in (usually infinite-dimensional) vector spaces.

Exercise

Consider $C[0,1]$ as a vector space over $\mathbb{R}$ (with pointwise operations). Show that this vector space is not finite-dimensional.

Before getting on to our main business of analysis, we will take a brief detour into the foundations of mathematics in *Chapter 1: Sets and Cardinality*.