An introduction to mathematical logic, notation, quantifiers, definitions, and proofs
Alfonso Gracia-Saz, MAT 137

In MAT137 we emphasize rigorous definitions and proofs. Many students struggle with this because they have never seen anything similar in high-school math. These notes and the references in it are designed to smooth that transition.

1 The basics

First we need some basic objects to work with. Read sections 1.2 and 1.3 from the textbook.

2 Mathematical logic

As a first introduction to mathematical logic, read section 3 (titled “Mathematical logic”) on http://uoft.me/precalc

Note: Actually, all nine sections of this website are an excellent preparation for MAT137. The other eight sections are review, things we count on you having already learned in high school. (If this is not the case, you should study them as well!) The section on Logic is probably new, which is why we emphasize it here.

3 Some pieces of notation

On page 4 of the textbook you were introduced to the standard sets of numbers. Here are their usual mathematical names:

- the naturals or positive integers: \( \mathbb{N} \)
- the integers: \( \mathbb{Z} \)
- the rationals: \( \mathbb{Q} \)
- the reals: \( \mathbb{R} \)

On section 3 of http://uoft.me/precalc you were introduced to mathematical logic. Here are some standard pieces of notation:

- \( A \implies B \) means “If A, then B”.
- \( A \iff B \) or “A if and only if B” (read “A if and only if B”) means “The statements A and B are equivalent”.
- \( \forall \) = “for all” or “for every”.
- \( \exists \) = “there exists (at least one)”

We will try not to go too heavy with notation, but you will encounter these symbols either in this course or in future courses, so it is good to be aware of them.
4 Definitions

We are going to show with an example how to write down a mathematically rigorous definition. We want to define what it means for a function to be *increasing*. We assume that you already understand intuitively what this means – something like the larger $x$ is, the larger $f(x)$ is – but that is not yet a rigorous definition.

We need to compare the values of $f$ at two points. An idea is to say that increasing means the following:

$$x_1 < x_2, \quad f(x_1) < f(x_2)$$

This is not a correct or incorrect definition. It simply is not a definition. What is $f$? What are $x_1$ and $x_2$?

In a definition, you have to introduce all the symbols you will use. Otherwise, it is nonsensical.

Okay then: $f$ is a function and $x_1$ and $x_2$ are arbitrary real numbers. How arbitrary? When we talk about a function being increasing, we have to refer to a specific interval. A function could be increasing in some intervals and decreasing in some other intervals. For example, the function whose graph is below is increasing on $[-1, 1]$ but decreasing on $[1, 3]$:

![Graph of a function](image)

This suggests that our definition might look like this:

**Definition 4.1 (Bad).** Let $f$ be a function defined on the real interval $I$. We say that $f$ is *increasing* on $I$ when

$$x_1, x_2 \in I, \quad x_1 < x_2, \quad f(x_1) < f(x_2)$$

Unfortunately, this definition is still ambiguous (or wrong, or nonsensical, but either way not good enough). Do we mean that the last two equations are true for every pair of numbers $x_1, x_2 \in I$? Do we mean that there is a pair of numbers $x_1, x_2 \in I$ for which the two equations are true? Do we mean something else?

So, let’s try to make it precise. Consider the following attempts to a more precise definition:
**Definition 4.2.** (Bad) Let \( f \) be a function defined on the real interval \( I \). We say that \( f \) is *increasing* on \( I \) when

\[
\text{For every } x_1, x_2 \in I, \quad x_1 < x_2, \quad f(x_1) < f(x_2)
\]

**Definition 4.3.** (Bad) Let \( f \) be a function defined on the real interval \( I \). We say that \( f \) is *increasing* on \( I \) when

\[
\text{There exist } x_1, x_2 \in I \text{ such that } x_1 < x_2 \text{ and } f(x_1) < f(x_2)
\]

**Definition 4.4.** Let \( f \) be a function defined on the real interval \( I \). We say that \( f \) is *increasing* on \( I \) when

\[
\text{For every } x_1, x_2 \in I, \quad \text{if } x_1 < x_2 \text{ then } f(x_1) < f(x_2)
\]

Definition 4.2 is wrong. It is unambiguous and it means something precise, but it does not mean “\( f \) is increasing”. Specifically, there isn’t any function \( f \) that satisfies the property in Definition 4.2. Do you see why?

Definition 4.3 is wrong. It is unambiguous and it means something precise, but it does not mean “\( f \) is increasing”. Specifically, there are lots of functions \( f \) that satisfy the property in Definition 4.3, but which are not increasing. Do you see why?

Definition 4.4 is the correct one! It is unambiguous, it means something precise, and it means exactly what we want “\( f \) is increasing” to mean. Do you see why?

Incidentally, there is often more than one equivalent, correct way to define a concept. In Definition 4.4, instead of

\[
\text{For every } x_1, x_2 \in I, \quad \text{if } x_1 < x_2 \text{ then } f(x_1) < f(x_2)
\]

we could have written:

\[
\text{If } x_1, x_2 \in I \text{ and } x_1 < x_2, \quad \text{then } f(x_1) < f(x_2)
\]

or

\[
\text{For every } x_1, x_2 \in I, \quad x_1 < x_2 \implies f(x_1) < f(x_2)
\]

or

\[
\forall x_1, x_2 \in I, \quad x_1 < x_2 \implies f(x_1) < f(x_2)
\]

and they all would be equally correct.

## 5 Proofs

### 5.1 Basic structure of a proof

Most mathematical theorems are of the form “If P, then Q”. For example, a theorem we will learn early in MAT 137 is:

**Theorem.** Every differentiable function is continuous.

(It is okay if you do not know yet what these concepts mean.) It may not look of the form “If P, then Q”, but what this theorem is actually saying is

**Theorem.** If a function is differentiable, then it is continuous.
Or, to be even more precise, the theorem should probably be restated as

**Theorem.** Let $f$ be a function and let $a$ be a real number inside the domain of $f$. If $f$ is differentiable at $a$, then $f$ is continuous at $a$.

The basic structure of a proof for an “If P, then Q” will always be the same. We start by assuming P to be true, then we use other results, properties, and algebraic manipulations that we know are true, and we conclude that Q has to be true. Schematically:

**Theorem 5.1.** If P, then Q.

**Proof.** Assume P is true.

Then we have concluded that Q is true.

The little symbol □ at the end of the proof means “end of the proof”. Let us illustrate this with an example.

First, here is an example of a bad theorem and a bad proof:

**Theorem 5.2 (Bad).**

$$\sqrt{xy} \leq \frac{x + y}{2}$$

**Bad proof.**

$$\sqrt{xy} \leq \frac{x + y}{2} \leq \frac{(x + y)^2}{4}$$

$$4xy \leq (x + y)^2$$

$$4xy \leq x^2 + 2xy + y^2$$

$$0 \leq (x - y)^2$$

\[\square\]

Theorem 5.2 is bad because it is not a theorem; instead, it is an equation. An equation by itself can never be a theorem. A theorem could be an equation together with an explanation of what the variables mean and when the equation is true. So, when is this equation true? For all real numbers $x$ and $y$? No, it isn’t! For example, when $x = -1$ and $y = -4$ the equation is false. This needs to be fixed.

As for the proof of Theorem 5.2, it is terrible! It has no words; it is just a bunch of disconnected statements; there is no explanation of what we are assuming and what we are trying to proof; we do not know what follows from what. Worse, it appears to start with the conclusion of the theorem, instead of ending with it. If you ever write a proof like this in a test, it will get you 0 points, and it will make your grader cry.

Here is how to rewrite this into a correct theorem and a correct proof.
Theorem 5.3. If \( x, y \) are positive real numbers, then:

\[
\sqrt{xy} \leq \frac{x + y}{2}
\]

Proof. Let \( x, y \) be any two positive real numbers. We know that \( c^2 \geq 0 \) for every real number \( c \). In particular, we know that \((x - y)^2 \geq 0\). Let’s expand this:

\[
0 \leq (x - y)^2 \\
\implies 0 \leq x^2 - 2xy + y^2 \\
\implies 4xy \leq x^2 + 2xy + y^2 \\
\implies xy \leq \frac{(x + y)^2}{4} = \left(\frac{x + y}{2}\right)^2
\]

Since both \( x \) and \( y \) are positive, so is \( xy \), and hence we can take the square root of both sides.

\[
\sqrt{xy} \leq \sqrt{\left(\frac{x + y}{2}\right)^2} = \left|\frac{x + y}{2}\right|
\]

where we have used that for any real number \( c \), \( \sqrt{c^2} = |c| \). However, in this case, \( x \) and \( y \) are positive, and so is \( x + y \), so \( |x + y| = x + y \) and we conclude that

\[
\sqrt{xy} \leq \frac{x + y}{2}
\]

which is what we wanted to prove.

What is different about this proof? We start with the hypothesis \((x \text{ and } y \text{ are positive real numbers})\) and we end with the conclusion. In between, there are words that explain how to justify every step (except for very simple algebraic manipulations) and everything follows in logical order.

The general convention is that whenever you write something in the middle of a proof, you are claiming that you have already established that it is true. If you want to write something else in the middle of the proof that you have not proven yet (for example, you want to write the statement that you are aiming for), you need to explain it explicitly (for example, you could write \(\text{“I want to prove that..”}\)).

With this in mind, here is a way to rewrite the proof of Theorem 5.3 into something that looks like the bad proof we wrote for Theorem 5.2, but which is correct!

Theorem 5.3. If \( x, y \) are positive real numbers, then:

\[
\sqrt{xy} \leq \frac{x + y}{2}
\]

Proof. Let \( x \) and \( y \) be positive real numbers. We want to prove that

\[
\sqrt{xy} \leq \frac{x + y}{2} \tag{1}
\]

Instead, we are going to prove that

\[
xy \leq \frac{(x + y)^2}{4} \tag{2}
\]
Notice that (2) is obtained from (1) by squaring both sides. Since both sides are positive in Equation (1), the two equations are equivalent. Hence, we only need to prove that (2) is true, and it will follow that (1) is true.

Now we do some algebraic manipulations, all of which are equivalences:

\[ xy \leq \frac{(x + y)^2}{4} \]  
\[ 4xy \leq (x + y)^2 \]  
\[ 4xy \leq x^2 + 2xy + y^2 \]  
\[ 0 \leq (x - y)^2 \]

All of Equations (3) through (6) are equivalent. Equation (3) is the one we want to prove. Equation (6) is clearly true. Hence, we have completed the proof.

5.2 Proof by contrapositive and by contradiction

Imagine that you need to prove a theorem of the form \( P \implies Q \). There are three standard ways to do it.

- **Direct proof.** This is what we discussed in section 5.1
  1. Assume \( P \) is true.
  2. Do stuff.
  3. Conclude that \( Q \) is true

- **Proof by contrapositive.** Instead of proving that \( [P \implies Q] \) directly, we prove the equivalent statement \( [(\not Q) \implies (\not P)] \).
  1. Assume \( Q \) is false.
  2. Do stuff.
  3. Conclude that \( P \) is false

If you are wondering why this works, review the section on Mathematical Logic at http://uoft.me/precalc once more.

- **Proof by contradiction.**
  1. Assume \( P \) is true and \( Q \) is false.
  2. Do stuff.
  3. Obtain a contradiction.

The contradiction means that it is impossible for \( P \) to be true and for \( Q \) to be false at the same time. Hence if \( P \) is true, \( Q \) must be true as well, and this is a proof of \( [P \implies Q] \).

Proofs by contrapositive and by contradiction will rarely appear in MAT137. However, they are very common in mathematics, and you will encounter them often in MAT237 and in Linear Algebra.

5.3 Proof by induction

Read section 1.8 of the textbook.