M2PM1: Real Analysis

Dr. Claude Warnick

December 19, 2016

Abstract

In first year analysis courses, you learned about the real numbers and were introduced to important concepts such as completeness; convergence of sequences and series; continuity; differentiability of functions of one real variable. In this course we will extend these ideas to higher dimensions, as well as introducing new fundamental tools for analysis, including uniform convergence and integration.

Contents

Notes on Exercises

A crucial feature of mathematics is that it is a *practical* subject. One learns to a large extent by *doing*: solving problems, checking details of proofs, finding counterexamples.

Throughout the notes are incorporated exercises to test your understanding, develop intuition and establish results that are required later on. You should not treat these exercises as optional[1]. **You won't understand the material in this course unless you spend the time trying to solve the problems!** You will have to think hard, and you may have to come back to some of the problems over more than one session to crack them. Persist.

Notes on the Appendix

Appendix A consists of an introduction to differentiation in one dimension, and will be covered in the JMC problem classes at the start of term. Students on the mathematics course who sat the course M1P1 Analysis 1 should have seen this material before, and can treat it as optional (but they may wish to study it for revision). Those students on the joint computing and mathematics course will not have seen this material before, so should attend the problem classes and study the appendix.

Notes on non-examinable material

Material marked (∗) is not examinable, but will be valuable in future courses, especially if you take (applied or pure) analysis courses in your third year and beyond. You should certainly at least read through the notes on these sections, even if you choose not to attempt the questions. I will try to indicate in lectures when I'm covering this material.

[1] The exception here are those questions marked (∗), which go beyond the scope of the course, but which are hopefully nevertheless interesting.

Introduction

In this chapter we're going to start off by introducing a few definitions so that we've agreed on notation. We're also going to extend some of the ideas that you saw last year (such as limits and continuity) to higher dimensions. The definitions are almost identical, so this should mostly feel like a review chapter to begin with, although some of the ideas we're going to approach from a different point of view.

Notation for $\mathbb{R}^n$

Numbers

First, we quickly recall the various sets of numbers that you were introduced to last year. I'm going to define the natural numbers to start from 0, and denote by $\mathbb{N}^*$ the set of positive natural numbers:

$$\begin{aligned} \mathbb{N} &:= \{0, 1, 2, \ldots\}, \\ \mathbb{N}^* &:= \{1, 2, 3, \ldots\}. \end{aligned}$$

Be careful: not all authors use this convention, for some $\mathbb{N}$ denotes the positive natural numbers. The set of integers is denoted $\mathbb{Z}$:

$$\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

If needed, the set $\mathbb{Z}^*$ is the non-zero integers $\mathbb{Z}^* := \mathbb{Z} \setminus \{0\}$. We know how to add, subtract and multiply elements of $\mathbb{Z}$ (and hence $\mathbb{N}$, $\mathbb{N}^*$). The rational numbers are denoted $\mathbb{Q}$:

$$\mathbb{Q} := \left\{ \frac{p}{q} \;\middle|\; p \in \mathbb{Z}, \quad q \in \mathbb{Z}^* \right\},$$

where we make the identification:

$$\frac{p}{q} \sim \frac{mp}{mq}, \qquad \forall\, m \in \mathbb{Z}^*.$$

Again, if required $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$ is the set of non-zero rational numbers. We can add multiply and subtract elements of $\mathbb{Q}$, and we also can divide by elements of $\mathbb{Q}^*$. On $\mathbb{Q}$

we have a notion of ordering $\leq$, so that we can say whether a rational number is greater than, less than or equal to another.

The set $\mathbb{R}$ of real numbers is obtained as the *completion* of $\mathbb{Q}$, and inherits the operations of arithmetic and the order relation from $\mathbb{Q}$. We have:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Moreover, $\mathbb{R}$ satisfies the *completeness axiom*:

> If $A \subset \mathbb{R}$ is non-empty and bounded above, then A has a least upper bound.

An important function defined on all real numbers is the *modulus function*, given by:

$$|x| := \begin{cases} x & x \geq 0, \\ -x & x < 0. \end{cases}$$

This has the properties:

i) For all $x \in \mathbb{R}$, we have $|x| \geq 0$, with $|x| = 0$ if and only if $x = 0$.

ii) For all $x \in \mathbb{R}$, if $a \in \mathbb{R}$, then $|ax| = |a|\,|x|$

iii) The modulus function satisfies the triangle inequality:

$$|x + y| \leq |x| + |y|, \qquad \forall\, x, y, \in \mathbb{R}$$

The set $\mathbb{R}^n$

We can now introduce the space $\mathbb{R}^n$, which consists of column vectors with n real components[2]:

$$\mathbb{R}^n = \left\{ \left(x^1, \ldots, x^n\right)^t \middle|\ x^i \in \mathbb{R}, i = 1, \ldots, n \right\}$$

If $x, y \in \mathbb{R}^n$ are two vectors in $\mathbb{R}^n$ with

$$x = \left(x^1, \ldots, x^n\right)^t, \qquad y = \left(y^1, \ldots, y^n\right)^t,$$

we can define the sum of the two by:

$$x + y := \left(x^1 + y^1, \ldots, x^n + y^n\right)^t.$$

If $\lambda \in \mathbb{R}$, we define:

$$\lambda x := \left(\lambda x^1, \ldots, \lambda x^n\right)^t.$$

With these definitions, $\mathbb{R}^n$ has the structure of a *vector space* over $\mathbb{R}$.

[2] I will try to stick to the convention of using superscripts to label components of vectors, and subscripts to label different vectors, so that $x_1, x_2 \in \mathbb{R}^n$ are two different vectors, while $x^1, x^2 \in \mathbb{R}$ are the component of one vector.

We can define the inner product of two vectors by:

$$\langle x, y \rangle = \sum_{i=1}^{n} x^i y^i,$$

and we define the length, or norm of a vector by:

$$||x|| = \sqrt{\langle x, x \rangle}.$$

This has the properties:

i) For all $x \in \mathbb{R}^n$, we have $||x|| \geq 0$, with $||x|| = 0$ if and only if $x = 0$.

ii) For all $x \in \mathbb{R}^n$, if $a \in \mathbb{R}$, then $||ax|| = |a| \, ||x||$

iii) The norm satisfies the triangle inequality:

$$||x + y|| \leq ||x|| + ||y|| \,, \qquad \forall \; x, y, \in \mathbb{R}^n. \tag{I.1}$$

These properties are often abstracted away to define more general *normed spaces* [See the Functional Analysis course M3P7]. The norm gives us a useful notion of *distance* between two points. The distance from x to y is given by $||x - y||$. Notice that if $n = 1$ we have $|\cdot| = ||\cdot||$, and we will use either interchangeably in this case.

A very important class of subsets of $\mathbb{R}^n$ are the *open balls*. For $x \in \mathbb{R}^n$, $r > 0$, these are defined as follows:

$$B_r(x) = \{ y \in \mathbb{R}^n : ||x - y|| < r \} \,.$$

That is, $B_r(x)$ consists of all points in $\mathbb{R}^n$ which are a distance less than r from x.

Exercise 1.1. a) Show that the inner product has the following properties:

$$\langle x, y \rangle = \langle y, x \rangle \,, \qquad \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \,, \qquad \langle ax, y \rangle = a \langle x, y \rangle \,.$$

for all $x, y, z \in \mathbb{R}^n$ and $a \in \mathbb{R}$.

b) For $t \in \mathbb{R}$ and $x, y \in \mathbb{R}^n$, show that:

$$||x + ty||^2 = ||x||^2 + 2t \langle x, y \rangle + t^2 ||y||^2 \geq 0 \tag{I.2}$$

c) By thinking of (I.2) as a quadratic in t, and considering its possible roots, deduce the *Cauchy-Schwartz* inequality:

$$|\langle x, y \rangle| \leq ||x|| \, ||y|| \,. \tag{I.3}$$

When does equality hold?

d) Deduce the triangle inequality (I.1).

e) Show the reverse triangle inequality:

$$\big| \, ||x|| - ||y|| \, \big| \leq ||x - y||$$

f) Suppose $x = (x^1, \ldots, x^n)^t \in \mathbb{R}^n$.

i) Show that:

$$\max_{k=1,\ldots,n} \left| x^k \right| \leq ||x|| \, .$$

ii) Show that:

$$||x|| \leq \sqrt{n} \max_{k=1,\ldots,n} \left| x^k \right| .$$

Sets and convergence

Now that we have a few definitions relating to $\mathbb{R}^n$, we're ready to revisit some concepts from first year analysis and see how they can be extended to higher dimensions.

Convergence of sequences in $\mathbb{R}^n$

A sequence in $\mathbb{R}^n$ is an ordered list

$$x_0, x_1, \ldots, x_i, \ldots$$

with each $x_i \in \mathbb{R}^n$ for $i = 0, 1, 2, \ldots$. This is often written $(x_i)_{i=0}^{\infty}$, or $(x_i)_{i \in \mathbb{N}}$. A very important concept relating to sequences is that of *convergence*.

Definition I.1. A sequence $(x_i)_{i=0}^{\infty}$ with $x_i \in \mathbb{R}^n$ converges to the vector $x \in \mathbb{R}^n$ if the following holds: given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $i \geq N$ we have

$$||x_i - x|| < \epsilon.$$

We then write:

$$x_i \to x, \qquad \text{as } i \to \infty,$$

or

$$\lim_{i \to \infty} x_i = x.$$

You should go back and compare this definition to the definition of convergence for a sequence of real numbers that you saw in your first year.

Lemma I.1. *The sequence of vectors $(x_i)_{i=0}^{\infty}$ with $x_i \in \mathbb{R}^n$ converges to the vector $x \in \mathbb{R}^n$ if and only if each component of x_i converges to the corresponding component of x. That is, if we write:*

$$x_i = (x_i^1, \ldots, x_i^n)^t, \quad \textit{and} \quad x = (x^1, \ldots x^n)^t,$$

then

$$x_i^k \to x^k, \quad \textit{as} \;\; i \to \infty, \quad \forall \; k = 1, \ldots n.$$

Proof. First suppose that:

$$x_i^k \to x^k, \quad \text{as} \quad i \to \infty, \quad \forall\ k = 1, \dots n.$$

Let $\epsilon > 0$. Then for each $k = 1, \dots n$, there exists N_k such that if $i \geq N_k$ we have:

$$\left| x_i^k - x^k \right| < \frac{\epsilon}{\sqrt{n}}.$$

Let $N = \max\{N_1 \dots, N_n\}$. Then for $i \geq N$, we have:

$$\max_{k=1,\dots,n} \left| x_i^k - x^k \right| < \frac{\epsilon}{\sqrt{n}}.$$

Now, recall from Exercise 1.1 f) ii) that:

$$||x|| \leq \sqrt{n} \max_{k=1,\dots,n} \left| x^k \right|,$$

so we deduce:

$$||x_i - x|| \leq \sqrt{n} \max_{k=1,\dots,n} \left| x_i^k - x^k \right| < \epsilon.$$

This establishes the result in one direction.

Now suppose that:

$$x_i \to x, \quad \text{as} \quad i \to \infty.$$

Let $\epsilon > 0$. Then there exists N such that for $i \geq N$ we have:

$$||x_i - x|| < \epsilon.$$

Recall from Exercise 1.1 f) i) that:

$$\max_{k=1,\dots,n} \left| x^k \right| \leq ||x|| .$$

Thus for any $k = 1, \dots n$ we have that if $i \geq N$:

$$\left| x_i^k - x^k \right| \leq \max_{k=1,\dots,n} \left| x_i^k - x^k \right| \leq ||x_i - x|| < \epsilon,$$

which implies $x_i^k \to x^k$, which completes the proof. □

Exercise 1.2. Suppose that $(x_i)_{i=0}^\infty$ and $(y_i)_{i=0}^\infty$ with $x_i, y_i \in \mathbb{R}^n$ are two sequences of vectors with

$$x_i \to x, \quad y_i \to y, \qquad \text{as } i \to \infty.$$

a) Show that

$$x_i + y_i \to x + y \qquad \text{as } i \to \infty.$$

b) Show that

$$\langle x_i, y_i \rangle \to \langle x, y \rangle \qquad \text{as } i \to \infty,$$

deduce that

$$||x_i|| \to ||x|| \qquad \text{as } i \to \infty.$$

c) Suppose that $(a_i)_{i=0}^\infty$ with $a_i \in \mathbb{R}$ is a sequence of real numbers with $a_i \to a$ as $i \to \infty$. Show that:

$$a_i x_i \to a x \qquad \text{as } i \to \infty.$$

Open, closed and compact sets

In the study of the real line we are familiar with objects such as the open interval $(0,1)$, the closed interval $[0,1]$ and other sets which are neither open nor closed such as $(0,1]$. In higher dimensions we will need to consider sets which have properties such as "openness" and "closedness". In fact, the study of open and closed sets on different spaces is an entire branch of mathematics in itself: *topology*. We will take a quick and dirty approach to the area, and defer more careful treatment to the course M2PM5 Metric Spaces and Topology.

Definition I.2. A subset $U \subset \mathbb{R}^n$ is *open* if for each $x \in U$, there exists $r > 0$ such that $B_r(x) \subset U$.

In other words, about any point in an open set we can find a small ball which is entirely contained in the set. This means that an open set has no edge or boundary points.

Example 1. The ball $B_1(0)$ is open. To see this, suppose $x \in B_1(0)$, so that $||x|| < 1$. Let $r = (1 - ||x||)/2$ and suppose $y \in B_r(x)$. Then:

$$||y|| = ||y - x + x|| \leq ||y - x|| + ||x|| < \frac{1 - ||x||}{2} + ||x|| < \frac{1 + ||x||}{2} < 1$$

so that $y \in B_1(0)$. Thus $B_r(x) \subset B_1(0)$.

Example 2. The set $A = \{x \in \mathbb{R}^n : ||x|| \leq 1\}$ is not open. Clearly $y := (1, 0, \ldots, 0)^t$ belongs to A. On the other hand, if $r > 0$ then $z := (1 + r/2, 0, \ldots, 0)^t$ belongs to $B_r(y)$ but not to A, so there is no $r > 0$ such that $B_r(y) \subset A$.

Exercise 1.3. Which of the following subsets of $\mathbb{R}^n$ is open:

a) $\mathbb{R}^n$?

b) $\emptyset$?

c) $\{x = (x^1, \ldots, x^n)^t \in \mathbb{R}^n : x^1 > 0\}$?

d) $\{x = (x^1, \ldots, x^n)^t \in \mathbb{R}^n : x^i \in [0,1)\}$?

e) $\mathbb{Q}^n := \{x = (x^1, \ldots, x^n)^t \in \mathbb{R}^n : x^i \in \mathbb{Q}\}$?

Now we move on to discuss closed sets:

Definition I.3. A subset $E \subset \mathbb{R}^n$ is *closed* if for every convergent sequence $(x_i)_{i=0}^{\infty}$ with $x_i \in E$, we have $x = \lim_{i\to\infty} x_i \in E$.

In other words, E contains all of its *limit points*, that is points which can be approached by a sequence from inside E.

Example 3. The single point set $E = \{y\} \subset \mathbb{R}^n$ is closed. Any convergent sequence $(x_i)_{i=0}^{\infty}$ with $x_i \in E$ must in fact be the constant sequence $x_i = y$, for which we certainly have $y = \lim_{i\to\infty} x_i \in E$.

Example 4. The set $E = \{x \in \mathbb{R}^n : ||x|| \leq 1\}$ is closed. Any convergent sequence $(x_i)_{i=0}^{\infty}$ with $x_i \in E$ has $||x_i|| \leq 1$. Now, we know from Exercise 1.2 that if $x_i \to x$, then $||x_i|| \to ||x||$. It's also a result from last year that if we have a convergent sequence of real numbers $(a_i)_{i=0}^{\infty}$ with $a_i \in \mathbb{R}$ satisfying $a_i \leq M$, then $\lim_{i\to\infty} a_i \leq M$. Combining these two facts, we conclude that:

$$||x|| = \left|\left|\lim_{i\to\infty} x_i\right|\right| = \lim_{i\to\infty} ||x_i|| \leq 1,$$

so that $x \in E$. More generally, for $r > 0$ and $y \in \mathbb{R}^n$, the *closed ball*:

$$\overline{B_r(y)} := \{x \in \mathbb{R}^n : ||x - y|| \leq r\}$$

is closed.

Theorem I.2. *A set E is closed if and only if the set $E^c := \{x \in \mathbb{R}^n : x \notin E\}$ is open.*

Proof. First suppose that E is closed and let $y \in E^c$. We have to prove that there exists $r > 0$ such that $B_r(y) \subset E^c$. We proceed by contradiction. Suppose there is no such r, then for each $i \in \mathbb{N}$ we can find $x_i \in B_{2^{-i}}(y) \cap E$. We have:

$$||x_i - y|| < 2^{-i}$$

so $x_i \to y$ as $i \to \infty$. Since E is closed we must have $y \in E$, but this is a contradiction since $y \in E^c$. Thus if E is closed, E^c must be open.

Now consider the converse. Suppose that E^c is open and that $(x_i)_{i=0}^{\infty}$ is a convergent sequence with $x_i \in E$. Let $x = \lim_{i\to\infty} x_i$. We wish to show that $x \in E$. Again we proceed by contradiction. Suppose $x \notin E$, i.e. $x \in E^c$. Since E^c is open, there exists $r > 0$ such that $B_r(x) \subset E^c$. In particular, this means that $||y - x|| \geq r$ for all $y \in E$. However, we know that since $x_i \to x$ there exists N such that if $i \geq N$ we have:

$$||x_i - x|| < \frac{r}{2}.$$

Since $x_i \in E$ this is a contradiction. Thus we have shown that if E^c is open, then E is closed. □

This result enables us to equivalently define closed sets as being those whose complement is open. The fact that these two definitions of closedness are equivalent relies on certain properties of $\mathbb{R}^n$ which do not hold for more general topological spaces, and it is the definition in terms of open complements that is usually taken to be the correct notion to generalise.

Exercise 1.4. Which of the following subsets of $\mathbb{R}^n$ is closed:

a) $\mathbb{R}^n$?

b) $\emptyset$?

c) $\left\{x = (x^1, \ldots, x^n)^t \in \mathbb{R}^n : x^1 \geq 0\right\}$?

d) $\{x = (x^1, \dots, x^n)^t \in \mathbb{R}^n : x^i \in [0,1)\}$?

e) $\mathbb{Q}^n := \{x = (x^1, \dots, x^n)^t \in \mathbb{R}^n : x^i \in \mathbb{Q}\}$?

Definition I.4. For any set $A \subset \mathbb{R}^n$, we define:

i) the *interior* of A, denoted A° is the set of points $x \in A$ such that there exists $r > 0$ with $B_r(x) \subset A$.

ii) the *closure* of A, denoted $\overline{A}$, is the set of points $x \in \mathbb{R}^n$ such that there exists a sequence $(x_i)_{i=0}^\infty$ with $x_i \in A$ and $x_i \to x$.

iii) the *boundary* of A, denoted ∂A is the set $\overline{A} \setminus A^\circ$.

From the definition we have that if U is open then $U = U^\circ$, while if E is closed $E = \overline{E}$. We also have that for any set the inclusions:

$$A^\circ \subset A \subset \overline{A},$$

hold.

Exercise 1.5. a) Let $(x_i)_{i=0}^\infty$ be a sequence of vectors $x_i \in \mathbb{R}^n$ with $x_i \to x$. Suppose that the x_i satisfy $||x_i|| < r$ for all i and some $r > 0$. Show that:

$$||x|| \leq r.$$

b) Show that the closure of the open ball $B_r(y) := \{x \in \mathbb{R}^n : ||x - y|| < r\}$ is the closed ball $\overline{B_r(y)} := \{x \in \mathbb{R}^n : ||x - y|| \leq r\}$.

Exercise 1.6. Find A°, $\overline{A}$ and ∂A for the following sets:

a) $A = \mathbb{R}^n$.

b) $A = \{0\}$.

c) $A = \{x = (x^1, \dots, x^n)^t \in \mathbb{R}^n : x^1 \geq 0\}$?

d) $A = \{x = (x^1, \dots, x^n)^t \in \mathbb{R}^n : x^i \in [0,1)\ \forall i = 1, \dots, n\}$?

e) $A = \{x = (x^1, \dots, x^n)^t \in \mathbb{R}^n : x^i \in \mathbb{Q}\ \forall i = 1, \dots, n\}$?

Exercise 1.7. a) Show that if U_1, U_2 are open, then so are:

i) $U_1 \cup U_2$ *ii)* $U_1 \cap U_2$

b) Show that if E_1, E_2 are closed, then so are:

i) $E_1 \cup E_2$ *ii)* $E_1 \cap E_2$

*c) Suppose $\mathcal{U}$ is any collection of open sets.

i) Show that $\bigcup \mathcal{U}$ is open.

ii) Give an example showing that $\bigcap \mathcal{U}$ need not be open.

iii) What are the analogous statements for closed sets?

It is useful to also introduce the notion of *boundedness* for a subset of $\mathbb{R}^n$:

Definition I.5. A subset $A \subset \mathbb{R}^n$ is *bounded* if there exists $M \geq 0$ such that for all $x \in A$ we have:

$$||x|| \leq M.$$

A set which is both closed and bounded is called *compact*.

The Bolzano–Weierstrass theorem

The Bolzano–Weierstrass theorem is the the most basic of a family of important results in analysis that allow us to extract convergence from some sequence, provided that we have some compactness. These type of results can be very powerful, but often seem a bit mystical to begin with.

Theorem I.3 (Bolzano–Weierstrass)**.** *Suppose $E \subset \mathbb{R}^n$ is compact and that $(x_i)_{i=0}^{\infty}$ is a sequence with $x_i \in E$. Then $(x_i)_{i=0}^{\infty}$ has a subsequence which converges in E. That is, there exist integers:*

$$0 \leq i_1 < i_2 < \ldots < i_j < \ldots$$

such that the sequence $(x_{i_j})_{j=0}^{\infty}$ converges to some $x \in E$.

Before we attack the theorem in the generality it's stated, we'll first establish a simpler result, which you should have seen last year.

Lemma I.4. *Suppose $(a_i)_{i=0}^{\infty}$ is a bounded sequence of real numbers. Then$(a_i)_{i=0}^{\infty}$ has a convergent subsequence.*

Proof. We proceed by constructing a nested sequence of intervals, each containing an infinite number of elements of the sequence. First, since $(a_i)_{i=0}^{\infty}$ is bounded there exists some K such that $a_i \in [-K, K]$ for all $i = 0, 1, \ldots$. We call this interval $I_0 := [-K, K]$, and this certainly contains infinitely many elements of our sequence. Now, we can split the interval I_0 into two intervals $[-K, 0]$ and $[0, K]$. At least one of these intervals contains infinitely many elements of the sequence, so pick the leftmost and call it I_1. We proceed inductively. After j steps we have a closed interval I_j. We split this in half and choose the leftmost half which contains infinitely many elements of the sequence to call I_{j+1}.

We end up with a nested sequence:

$$I_0 \supset I_1 \supset I_2 \ldots \supset I_j \ldots$$

of closed intervals I_j of length $2^{1-j}K$, each containing infinitely many elements of the sequence (see Figure 1). We set $i_0 := 0$ and define i_j inductively by:

$$i_j := \min\{i : a_i \in I_j, \ i > i_{j-1}\}$$

for $j = 1, 2, \ldots$.

Let us denote by c_j the lower endpoint of I_j and d_j the upper endpoint, so that $I_j = [c_j, d_j]$. The sequence $(c_j)_{j=0}^{\infty}$ is monotone increasing and bounded above, so has a limit by the completeness axiom, say $c_j \to c$. Similarly, $(d_j)_{j=0}^{\infty}$ is monotone decreasing and bounded below so $d_j \to d$ for some d. Since $d_j - c_j = 2^{1-j}K$, we have that $c = d$. We also have $c_j \leq a_{i_j} \leq d_j$, so we conclude that

$$a_{i_j} \to c$$

as $j \to \infty$. Thus the subsequence $(a_{i_j})_{j=0}^{\infty}$ which we have constructed is convergent. □

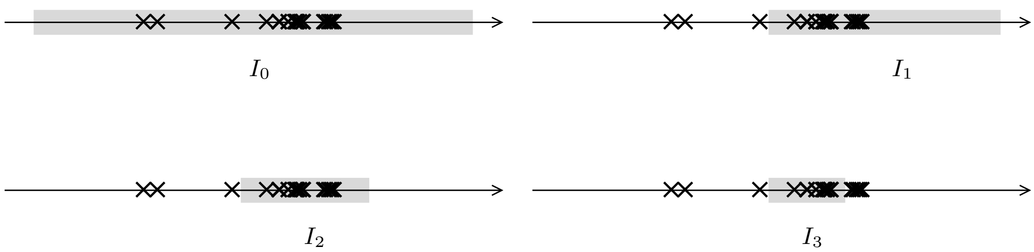

Figure 1 The nested intervals chosen in the proof of Lemma I.4

Now we can extend this result to $\mathbb{R}^n$:

Lemma I.5. *Suppose $(x_i)_{i=0}^{\infty}$, with $x_i \in \mathbb{R}^n$ is a bounded sequence. Then $(x_i)_{i=0}^{\infty}$ has a convergent subsequence.*

Proof. Since $(x_i)_{i=0}^{\infty}$ is bounded, there exists $K > 0$ such that $||x_i|| \leq K$. Recall from Exercise 1.1 f) i) that:

$$\max_{k=1,\ldots,n} \left| x^k \right| \leq ||x|| \, .$$

We conclude that in particular we have $x_i^1 \in [-K, K]$ for all $i \in \mathbb{N}$. Now, we can apply Lemma I.4 to the sequence of real numbers $(x_i^1)_{i=0}^{\infty}$ to extract a subsequence $(x_{i_{j_1}}^1)_{j_1=0}^{\infty}$ which converges to x^1. Consider the sequence $(x_{i_{j_1}})_{j_1=0}^{\infty}$. This is a subsequence of our original sequence, remains bounded and moreover $x_{i_{j_1}}^1 \to x^1$. Now we know that $x_{i_{j_1}}^2 \in [-K, K]$ for all $j_1 \in \mathbb{N}$. Again applying Lemma I.4, we can extract a subsequence $(x_{i_{j_1 j_2}}^2)_{j_2=0}^{\infty}$ which converges to x^2. The sequence $(x_{i_{j_1 j_2}})_{j_2=0}^{\infty}$ is a subsequence of our original sequence, remains bounded and moreover $x_{i_{j_1 j_2}}^1 \to x^2$ and $x_{i_{j_1 j_2}}^2 \to x^1$. Continuing in this way we extract a subsequence of $(x_i)_{i=0}^{\infty}$ such that each component converges. By Lemma I.1 we conclude that the subsequence converges. □

This previous Lemma is sometimes also known as the Bolzano–Weierstrass theorem. With this result, we can finally conclude our proof of Theorem I.3.

Proof of Theorem I.3. Since $(x_i)_{i=0}^{\infty}$ is a sequence with $x_i \in E$ and E is bounded, the sequence is bounded. By Lemma I.5 we have that $(x_i)_{i=0}^{\infty}$ admits a convergent subsequence, $(x_{i_j})_{j=0}^{\infty}$. Since $x_{i_j} \in E$ for all j, and E is closed, we have that $\lim_{j\to\infty} x_{i_j} \in E$. $\square$

Exercise 2.1. We say that a sequence $(x_i)_{i=0}^{\infty}$ with $x_i \in \mathbb{R}^n$ is a *Cauchy sequence* if the following holds: given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $i, j \geq N$ we have:

$$||x_i - x_j|| < \epsilon.$$

a) Suppose the sequence $(x_i)_{i=0}^{\infty}$ with $x_i \in \mathbb{R}^n$ converges to x. Show that $(x_i)_{i=0}^{\infty}$ is a Cauchy sequence.

b) Suppose $(x_i)_{i=0}^{\infty}$ with $x_i \in \mathbb{R}^n$ is a Cauchy sequence. Show that $(x_i)_{i=0}^{\infty}$ is bounded.

c) Show that every Cauchy sequence in $\mathbb{R}^n$ converges.

Continuity

Last year, you learned about the notion of continuity for functions from $\mathbb{R}$ (or subsets thereof) to $\mathbb{R}$. In this section we'll revisit those definitions and upgrade them for functions from (subsets of) $\mathbb{R}^n$ to $\mathbb{R}^m$. In fact, the definitions we shall give are almost identical: the only thing that changes is that we use the appropriate norm for the spaces.

Continuity at a point

We will start with the simple definition:

Definition I.6. Let $A \subset \mathbb{R}^n$ and suppose $f : A \to \mathbb{R}^m$. We say that f is continuous at $p \in A$ if the following holds: given $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in A$ with $||x - p|| < \delta$ we have:

$$||f(x) - f(p)|| < \epsilon.$$

If f is continuous at p for all $p \in A$, we say f is continuous on A.

We can think of this as saying "f maps points in A close to p to points in $\mathbb{R}^m$ close to $f(p)$". Notice that in the definition above, the symbol $||\cdot||$ is playing two slightly different roles: as the norm on $\mathbb{R}^n$ and the norm on $\mathbb{R}^m$.

Example 5. The map:

$$\begin{array}{cccc} f & : & \mathbb{R}^n & \to & \mathbb{R} \\ & & x & \mapsto & ||x|| \end{array}$$

is continuous on $\mathbb{R}^n$. To see this, pick $p \in \mathbb{R}^n$. Suppose $||x - p|| < \delta$, then by the reverse triangle inequality (see Exercise 1.1) we have:

$$|f(x) - f(p)| = \big|\, ||x|| - ||p|| \,\big| \leq ||x - p|| < \delta.$$

Thus we can take $\delta = \epsilon$ and we have satisfied the condition for continuity.

Example 6. Suppose $\Lambda : \mathbb{R}^n \to \mathbb{R}^m$ is *linear*. Then Λ is continuous. To see this, we let $\{e_j\}_{j=1}^n$ be the canonical basis for $\mathbb{R}^n$ and we calculate:

$$\begin{aligned} ||\Lambda x - \Lambda p|| = ||\Lambda(x-p)|| &= \left|\left|\Lambda\left(\sum_{j=1}^n e_j (x-p)^j\right)\right|\right| \\ &= \left|\left|\sum_{j=1}^n \Lambda e_j (x-p)^j\right|\right| \leq \sum_{j=1}^n \left|\left|\Lambda e_j (x-p)^j\right|\right| \\ &\leq \sum_{j=1}^n \left|(x-p)^j\right| ||\Lambda e_j|| \end{aligned}$$

Setting $M = \max_{j=1,\ldots,n} ||\Lambda e_j||$, we have:

$$||\Lambda x - \Lambda p|| \leq Mn \max_{j=1,\ldots,n} \left|(x-p)^j\right| \leq Mn\, ||x-p||\,,$$

using Exercise 1.1 f) i). Thus, if we take $\delta = \frac{\epsilon}{1+Mn}$, then for any x with $0 < ||x-p|| < \delta$, we have

$$||\Lambda x - \Lambda p|| < \frac{\epsilon}{1+Mn} Mn < \epsilon,$$

so Λ is continuous.

Example 7. The map

$$\begin{array}{rccl} f & : & \mathbb{R}^n & \to & \mathbb{R} \\ & & (x^1, \ldots, x^n)^t & \mapsto & x^1 \end{array}$$

is continuous on $\mathbb{R}^n$. To see this, pick $p \in \mathbb{R}^n$. Suppose $||x-p|| < \delta$, then by Exercise 1.1 f) i) we have:

$$|f(x) - f(p)| = \left|x^1 - p^1\right| \leq \max_{k=1,\ldots,n} \left|x^k - p^k\right| \leq ||x-p|| < \delta,$$

so we may take $\delta = \epsilon$ and we have satisfied the condition for continuity. Obviously the same argument shows that all of the coordinate maps (i.e. the map taking x to x^k) are continuous.

Theorem I.6. *Let $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$. Suppose $f : A \to B$ is continuous at p and $g : B \to \mathbb{R}^l$ is continuous at $f(p)$. Then $g \circ f : A \to \mathbb{R}^l$ is continuous at p.*

Proof. Let $\epsilon > 0$. Since g is continuous at $f(p)$, we know that there exists δ_1 such that for any $y \in B$ with $||y - f(p)|| < \delta_1$, we have $||g(y) - g(f(p))|| < \epsilon$. Similarly, Since f is continuous at p, we know that there exists δ such that for any $x \in A$ with $||x-p|| < \delta$, we have $||f(x) - f(p)|| < \delta_1$. Combining these two statements and taking $y = f(x)$, we deduce that if $x \in A$ with $||x-p|| < \delta$, we have $||g(f(x)) - g(f(p))|| < \epsilon$. □

It's sometimes useful to express the continuity of a function in a slightly different way, for which we need the following definition:

Definition I.7. Let $A \subset \mathbb{R}^n$ and suppose $f : A \to \mathbb{R}^m$. For $p \in A$ a limit point[3] of A, we say that $F \in \mathbb{R}^m$ is the limit of f as x tends to p if the following holds: given $\epsilon > 0$, there exists $\delta > 0$ such that for all $x \in A$ with $0 < ||x - p|| < \delta$ we have:

$$||f(x) - F|| < \epsilon.$$

In this case, we write

$$\lim_{x \to p} f(x) := F.$$

With this notion of a limit in hand, we can give the definition of continuity more compactly as:

"f is continuous at p if $\lim_{x \to p} f(x) = f(p)$."

Theorem I.7. *Suppose $A \subset \mathbb{R}^n$ and suppose $f, g : A \to \mathbb{R}$ and*

$$\lim_{x \to p} f(x) = F, \qquad \lim_{x \to p} g(x) = G,$$

then:

i) $\lim_{x \to p} f(x) + g(x) = F + G,$

ii) $\lim_{x \to p} f(x)g(x) = FG,$

iii) If, furthermore $G \neq 0$, then:

$$\lim_{x \to p} \frac{f(x)}{g(x)} = \frac{F}{G}.$$

Proof. i) Fix $\epsilon > 0$. Since $\lim_{x \to p} f(x) = F$, we know that there exists δ_1 such that if $x \in A$ with $0 < ||x - p|| < \delta_1$ then:

$$|f(x) - F| < \frac{\epsilon}{2}.$$

Similarly, there exists δ_2 such that if $x \in A$ with $0 < ||x - p|| < \delta_2$ then:

$$|g(x) - G| < \frac{\epsilon}{2}.$$

Take $\delta = \min\{\delta_1, \delta_2\}$. Then if $x \in A$ with $0 < ||x - p|| < \delta$, we have by the triangle inequality:

$$|f(x) + g(x) - (F + G)| \leq |f(x) - F| + |g(x) - G| < \epsilon.$$

[3] p is a limit point of A if $B_r(p) \cap A \neq \{p\}$ for any $r > 0$. That is, each open ball around p contains at least one other point of A. Note that every point in A° is a limit point.

ii) Fix $\epsilon > 0$, and assume without loss of generality that $\epsilon < 1$. Since $\lim_{x\to p} f(x) = F$, we know that there exists δ_1 such that if $x \in A$ with $0 < ||x-p|| < \delta_1$ then:

$$|f(x) - F| < \frac{\epsilon}{3(1+|G|)}.$$

Similarly, there exists δ_2 such that if $x \in A$ with $0 < ||x-p|| < \delta_2$ then:

$$|g(x) - G| < \frac{\epsilon}{3(1+|F|)}.$$

We note the following identity:

$$f(x)g(x) - FG = (f(x)-F)(g(x)-G) + (f(x)-F)G + (g(x)-G)F$$

Now, take $\delta = \min\{\delta_1, \delta_2\}$. If $x \in A$ with $0 < ||x-p|| < \delta$, we have by the triangle inequality:

$$\begin{aligned}|f(x)g(x) - FG| &\leq |f(x)-F|\,|g(x)-G| + |G|\,|f(x)-F| + |F|\,|g(x)-G| \\ &< \frac{\epsilon^2}{9(1+|F|)(1+|G|)} + \frac{\epsilon\,|G|}{3(1+|G|)} + \frac{\epsilon\,|F|}{3(1+|F|)} < \epsilon.\end{aligned}$$

iii) Given the previous part, it suffices to show that if $\lim_{x\to p} g(x) = G$ with $G \neq 0$, then

$$\lim_{x\to p} \frac{1}{g(x)} = \frac{1}{G}.$$

Fix $\epsilon > 0$. Since $\lim_{x\to p} g(x) = G$, we know that there exist δ such that if $x \in A$ with $0 < ||x-p|| < \delta$ then:

$$|g(x) - G| < \frac{\epsilon\,|G|^2}{2}.$$

Without loss of generality, we can assume that ϵ is sufficiently small that $\epsilon\,|G| < 1$, which implies $|g(x) - G| < \frac{|G|}{2}$ and so $|g(x)| > \frac{|G|}{2}$.
If $x \in A$ with $0 < ||x-p|| < \delta$, we can estimate:

$$\left|\frac{1}{g(x)} - \frac{1}{G}\right| = |G - g(x)| \cdot \frac{1}{|G|} \cdot \frac{1}{|g(x)|} < \frac{\epsilon\,|G|^2}{2} \cdot \frac{1}{|G|} \cdot \frac{2}{|G|} = \epsilon.$$

□

Corollary I.8. *Suppose $A \subset \mathbb{R}^n$ and $f, g : A \to \mathbb{R}$ are continuous at $p \in A$. Then:*

i) $f + g$ is continuous at p.

ii) fg is continuous at p.

iii) If, furthermore $g(p) \neq 0$, then $\frac{f}{g}$ is continuous at p.

Exercise 2.2 (*). Suppose $A \subset \mathbb{R}^n$ and $f : A \to \mathbb{R}^m$. Show that $\lim_{x\to p} f(x) = F$ if and only if for any sequence $(x_i)_{i=0}^{\infty}$ with $x_i \in A, x_i \neq p$ and $x_i \to p$ we have:

$$f(x_i) \to F, \qquad \text{as } i \to \infty.$$

Exercise 2.3. a) Show that the map

$$\begin{array}{rccc} f & : & \mathbb{R} & \to & \mathbb{R}^n \\ & & x & \mapsto & (x, 0, \ldots, 0)^t \end{array}$$

is continuous on $\mathbb{R}$.

b) Let $A \subset \mathbb{R}^n$ and suppose we are given a map:

$$\begin{array}{rcccc} f & : & A & \to & \mathbb{R}^m \\ & & (x^1, \ldots, x^n)^t & \mapsto & \left(f^1(x^1, \ldots, x^n), \ldots, f^m(x^1, \ldots, x^n)\right)^t. \end{array}$$

Show that f is continuous at $p \in A$ if and only if each map $f^k : A \to \mathbb{R}$ is continuous at p, for $k = 1, \ldots, m$.

c) Show that a map $f : \mathbb{R}^n \to \mathbb{R}$ which is polynomial in the coordinates is continuous on $\mathbb{R}^n$.

Extreme value theorem

The final result we shall establish in this chapter is the extreme value theorem. This is a valuable result which allows us to conclude something about the range of a continuous function by knowing something about the domain. More precisely, if the domain is compact, then the range is bounded.

Theorem I.9. *Let $E \subset \mathbb{R}^n$ be compact, and suppose $f : E \to \mathbb{R}$ is continuous on E. Then there exist $p_+, p_- \in E$ such that:*

$$f(p_-) \leq f(x) \leq f(p_+)$$

for all $x \in E$. That is to say that f is bounded and achieves its bounds.

Proof. We first argue that $f(x)$ is bounded above for $x \in E$. Suppose not, then for each $i \in \mathbb{N}$ there exists $x_i \in E$ such that $f(x_i) > i$. Since E is compact, by Bolzano-Weierstrass the sequence $(x_i)_{i=0}^{\infty}$ has a convergent subsequence $(x_{i_j})_{j=0}^{\infty}$ such that $x_{i_j} \to x \in E$. By the continuity of f we have:

$$\lim_{j\to\infty} f(x_{i_j}) = f(x),$$

but on the other hand $f(x_{i_j}) > i_j \to \infty$ as $j \to \infty$, a contradiction. Therefore, there exists M such that $f(x) \leq M$ for all $x \in E$. In other words, the set

$$f(E) := \{f(x) : x \in E\}$$

is bounded above. By the completeness axiom, $f(E)$ has a least upper bound, so without loss of generality we will assume $M = \sup_{x\in E} f(x)$.

Now since M is the least upper bound, for $i \geq 1$, we have that $M - 2^{-i}$ is not an upper bound. Therefore there exists $y_i \in E$ with $M - 2^{-i} < f(y_i) \leq M$. Clearly, we have that $f(y_i) \to M$ as $i \to \infty$. Now the sequence $(y_i)_{i=0}^{\infty}$ with $y_i \in E$ has a convergent subsequence $(y_{i_j})_{j=0}^{\infty}$ with $y_{i_j} \to p_+$ for some $p_+ \in E$ by the Bolzano-Weierstrass theorem. Since $f(y_{i_j}) \to M$, by the continuity of f we conclude that $f(p_+) = M$, Thus we have established that there exists $p_+ \in E$ such that $f(x) \leq f(p_+)$ for all $x \in E$. An identical argument deals with the lower bound. □

Exercise 2.4. Suppose that $E \subset \mathbb{R}^n$ is compact and $f : E \to \mathbb{R}^m$ is continuous on E.

a) Show that $f(E) := \{f(x) : x \in E\}$ is bounded.

b) Show that $f(E)$ is closed, hence compact.

c) Is it true that $f(E)$ is closed if we just assume E is closed (not necessarily compact)?

Exercise 2.5 (*)**.** a) Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous on $\mathbb{R}^n$, and suppose $U \subset \mathbb{R}^m$ is open. Show that:

$$f^{-1}(U) := \{x \in \mathbb{R}^n : f(x) \in U\}$$

is open.

b) Suppose that $f : \mathbb{R}^n \to \mathbb{R}^m$ has the property that $f^{-1}(U) \subset \mathbb{R}^n$ is open for every open $U \subset \mathbb{R}^m$. Show that f is continuous on $\mathbb{R}^n$.