CHAPTER 1

Real Analysis Foundations for this Book

The beginner should not be discouraged if he finds he does not have the prerequisites for reading the prerequisites.

Paul Halmos

1.1 INTRODUCTION AND OBJECTIVES

In this chapter we introduce a number of mathematical concepts and methods that underlie many of the topics in this book. The most urgent attention points revolve around functions of real variables, their properties and the ways they are used in applications. We discuss the most important topics from *real analysis* to help us in our understanding of partial differential equations (PDEs). A definition of real analysis is:

> In mathematics, real analysis is the branch of mathematical analysis that studies the behavior of real numbers, sequences and series of real numbers, and real functions. Some particular properties of real-valued sequences and functions that real analysis studies include convergence, limits, continuity, smoothness, differentiability and integrability.
>
> Real analysis is distinguished from complex analysis, which deals with the study of complex numbers and their functions.
>
> *(Wikipedia)*

A related branch of mathematics is *calculus,* which we learn at school:

> Calculus, originally called infinitesimal calculus or 'the calculus of infinitesimals', is the mathematical study of continuous change, in the same way that geometry is the study of shape and algebra is the study of generalizations of arithmetic operations.

> It has two major branches, differential calculus and integral calculus; the former concerns instantaneous rates of change, and the slopes of curves, while integral calculus concerns accumulation of quantities, and areas under or between curves. These two branches are related to each other by the fundamental theorem of calculus, and they make use of the fundamental notions of convergence of infinite sequences and infinite series to a well-defined limit.
>
> *(Wikipedia)*

In practice, there is a distinction between calculus and real analysis. Calculus entails techniques (and tricks) to differentiate and integrate functions. It does not discuss the conditions under which a function is continuous or differentiable. It assumes that it is allowed to carry out these operations on functions. Real analysis, on the other hand, does discuss these issues and more; for example:

- Continuous functions: How do we recognise them and prove that a function is continuous?
- The different kinds of discontinuous functions.
- Differential calculus from a real-analysis viewpoint.
- Taylor's theorem.
- An introduction to metric spaces and Cauchy sequences.

In our opinion, these topics are necessary prerequisites for the rest of this book.

Knowledge of vector (linear) analysis and numerical linear algebra is also a prerequisite for computational finance. To this end, we devote Chapters 4 and 5 to these topics. Finally, complex variables and complex functions (which are at the heart of complex analysis) are introduced in Chapter 16. We use the notation $\forall$ to mean 'for all' and $\exists$ to mean 'there exists'.

1.2 CONTINUOUS FUNCTIONS

In this section we are mainly concerned with real-valued functions of a real variable, that is $f : \mathbb{R} \to \mathbb{R}$. In rough terms, a continuous function is one that can be drawn by hand without taking the pen from paper. In other words, a continuous function does not have jumps or breaks, but it is allowed to have sharp bends and kinks. Examples of continuous functions are:

$$
\begin{aligned}
&f : \mathbb{R} \to \mathbb{R},\ \ f(x) = x^2 \\
&f : \mathbb{R} \to \mathbb{R},\ \ f(x) = \max(0, x).
\end{aligned} \tag{1.1}
$$

We can see that these functions are continuous just by drawing them. The first function is 'smoother' than the second function, the latter being similar to a one-factor call or put payoff on the one hand and a *Rectified Linear Unit* (ReLU) activation function on the other hand (Goodfellow, Bengio and Courville (2016)). Intuitively, a function f is continuous if $f(x) \to f(p)$ when $x \to p$, no matter how x approaches p. Alternatively, small changes in x lead to small changes in $f(x)$.

If we formally differentiate the above ReLU function (1.1), we get the famous discontinuous *Heaviside function*:

$$H(x) = \begin{cases} 0, x < 0 \\ 1, x \geq 0. \end{cases} \tag{1.2}$$

A *discontinuous function* is one that is not continuous. Another discontinuous function is:

$x \in \mathbb{R},\ [x] \equiv$ largest integer n *s.t.* $n \leq x \leq n + 1$.

Define $f(x) = [x]$; let $p \in \mathbb{Z}$ (integer).
Then taking left and right limits gives different answers, showing that the function is not continuous.

1. $x < p \Rightarrow f(x) = p - 1$
2. $x > p \Rightarrow f(x) = p$

Thus $\lim\limits_{x \to p-} f(x) = p - 1,\ \lim\limits_{x \to p+} f(x) = p$.

1.2.1 Formal Definition of Continuity

The following definition is based on the fact that small changes in x lead to small changes in $f(x)$.

Definition 1.1

$$\lim_{x \to p} f(x) = A \text{ means } \forall \varepsilon > 0\ \ \exists \delta > 0 \ s.t.$$

$$| f(x) - A | < \varepsilon \text{ when } 0 < | x - p | < \delta.$$

Some properties of continuous functions f(x) and g(x) are:

$$\lim_{x \to p} (f(x) \pm g(x)) = \lim_{x \to p} f(x) \pm \lim_{x \to p} g(x)$$

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

$$\lim_{x \to p} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x \to p} f(x)}{\lim\limits_{x \to p} g(x)}, \quad g(x) \neq 0.$$

1.2.2 An Example

It can be a mathematical challenge to prove that a function is continuous using the above '*epsilon-delta*' approach in Definition 1.1. One approach is to use the well-known technique of splitting the problem into several mutually exclusive cases, solving each case separately and then merging the corresponding partial solutions to form the desired solution. To this end, let us examine the square root function:

$$f : \mathbb{R}^+ \to \mathbb{R}^+, f(x) = \sqrt{x}. \tag{1.3}$$

We show that there exists $\delta > 0$ such that for $x \geq 0$:

$$|x - y| < \delta \Rightarrow |\sqrt{x} - \sqrt{y}| < \epsilon \;\; \forall y \in \mathbb{R}^+.$$

Then:

$$\sqrt{x} - \sqrt{y} = \frac{x - y}{\sqrt{x} + \sqrt{y}}.$$

We now consider two cases:

Case 1 : $x > 0$. Then:

$$|x - y| < \delta \Rightarrow |\sqrt{x} - \sqrt{y}| \leq \frac{|x - y|}{\sqrt{x}} = \frac{\delta}{\sqrt{x}} = \epsilon.$$

Choose $\delta = \epsilon\sqrt{x}$.

Case 2: $x = 0$. Then:

$$|x - y| < \delta \Rightarrow |\sqrt{x} - \sqrt{y}| = \frac{|x - y|}{\sqrt{x} + \sqrt{y}} = \frac{y}{\sqrt{y}} = \sqrt{y} = \epsilon.$$

Hence:

$$|-y| = |y| < \delta \Rightarrow \sqrt{y} = \epsilon \Rightarrow \sqrt{\delta} < \epsilon \Rightarrow \delta < \epsilon^2.$$

Choose $\delta = \epsilon^2$.

We have thus proved that the square root function is continuous.

1.2.3 Uniform Continuity

In general terms, *uniform continuity* guarantees that $f(x)$ and $f(y)$ can be made as close to each other as we please by requiring that x and y be sufficiently close to each other. This is in contrast to ordinary continuity, where the distance between $f(x)$ and $f(y)$ may depend on x and y themselves. In other words, in Definition 1.1 δ depends only on ϵ and not on the points in the domain. Continuity itself is a *local property* because a function f is or is not continuous at a particular point and continuity can be determined by looking at the values of the function in an arbitrary small neighbourhood of that point. Uniform continuity, on the other hand, is a *global property* of f because the definition refers to pairs of points rather than individual points. The new definition in this case for a function *f* defined in an interval *I* is:

$$\forall \epsilon > 0 \; \exists \delta > 0 \; s.t. \; \forall x, y \in I : |x - y| < \delta \Rightarrow \left| f(x) - f(y) \right| < \epsilon.$$

Let us take an example of a uniformly continuous function:

$$f : \mathbb{R} \to \mathbb{R}, \;\; f(x) = 3x + 7. \tag{1.4}$$

Then $\left| f(x) - f(y) \right| = |3x + 7 - (3y + 7)| = 3|x - y| < 3\delta < \epsilon, \;\; (x, y \in \mathbb{R})$.

Choose $\delta = \epsilon/3$.

In general, a continuous function on a closed interval is uniformly continuous. An example is:

$$f(x) = x^2 \text{ on } I = [0, 2]. \tag{1.5}$$

Let $x, y \in I$. Then:

$$|f(x) - f(y)| = (x + y)|x - y| < (2 + 2)\delta = \epsilon.$$

Choose $\delta = \epsilon/4$.

An example of a function that is continuous and nowhere differentiable is the *Weierstrass function* that we can write as a *Fourier series*:

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x), \;\; 0 < a < 1, \tag{1.6}$$

b is a positive odd integer and $ab > 1 + \frac{3}{2}\pi$.

This is a jagged function that appears in models of Brownian motion. Each partial sum is continuous, and hence by the *uniform limit theorem* (which states that the uniform limit of any sequence of continuous functions is continuous), the series (1.6) is continuous.

1.2.4 Classes of Discontinuous Functions

A function that is not continuous at some point is said to be *discontinuous* at that point. For example, the Heaviside function (1.2) is not continuous at $x = 0$. In order to determine if a function is continuous at a point x in an interval (a, b) we apply the test:

$$f(x+) = q \text{ if } f(t_n) \to q, n \to \infty \text{ for all sequences } \{t_n\} \text{ in } (x, b) \; s.t. \;\; t_n \to x$$

$$f(x-) = q \text{ if } f(t_n) \to q, n \to \infty \text{ for all sequences } \{t_n\} \text{ in } (a, x) \; s.t. \;\; t_n \to x$$

$$\exists \lim_{t \to x} f(t) \Longleftrightarrow f(x+) = f(x-) = \lim_{t \to x} f(t) = f(x).$$

There are two (simple discontinuity) main categories of discontinuous functions:

- *First kind*: $f(x+) = \lim_{t \to x+} f(t)$ and $f(x-) = \lim_{t \to x-} f(t)$ exists. Then either we have $f(x+) \neq f(x-)$ or $f(x+) = f(x-) \neq f(x)$.
- *Second kind*: a discontinuity that is not of the first kind.

Examples are:

$$f(x) = \begin{cases} 1, x \ \text{ rational } (x \in \mathbb{Q}) \\ 0, x \ \text{ not rational}, x \notin \mathbb{Q} \\ \text{2nd kind: Neither } f(x+) \text{ nor } f(x-) \text{ exists.} \end{cases}$$

$$f(x) = \begin{cases} x+2, \quad -3 < x < -2 \\ -x-2, \quad -2 \leq x < 0 \\ x+2, \quad 0 \leq x < 1 \\ \text{Simple discontinuity at } x = 0. \end{cases}$$

You can check that this latter function has a discontinuity of the first kind at $x = 0$.

1.3 DIFFERENTIAL CALCULUS

The *derivative* of a function is one of its fundamental properties. It represents the rate of change of the slope of the function: in other words, how fast the function changes with respect to changes in the independent variable. We focus on real-valued functions of a real variable.

Let $f : \mathbb{R} \to \mathbb{R}$. Then the derivative of f at x (if it exists) is defined by the limit for $x \in [a, b]$:

$$\begin{aligned} &\varphi(t) = \frac{f(t) - f(x)}{t - x} \ (t \neq x), \\ &f'(x) = \lim_{t \to x} \varphi(t) \\ &\text{or} \\ &\frac{df(x)}{dx} = f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}. \end{aligned} \tag{1.7}$$

This limit may not exist at certain points, and it is possible to define right-hand and left-hand limits, that is, one-sided derivatives.

Some results that we learn in high school are:

$$\begin{aligned} &(f+g)'(x) = f'(x) + g'(x) \\ &(fg)'(x) = f'(x)g(x) + f(x)g'(x) \\ &\left(\frac{f}{g}\right)'(x) = \frac{g(x)f'(x) - g'(x)f(x)}{g^2(x)} \ (g(x) \neq 0). \end{aligned} \tag{1.8}$$

A *composite function* is a function that we can differentiate using the *chain rule* that we state as follows:

$$
\begin{aligned}
& x \in [a,b], \quad \exists f'(x) \text{ with } g \text{ differentiable at } f(x). \\
& \text{Then:} \\
& h(t) \equiv g(f(t)), \quad a \le t \le b \text{ has derivative} \\
& h'(x) = g'(f(x))f'(x).
\end{aligned} \tag{1.9}
$$

A simple example of use is:

$$
\begin{aligned}
f(x) &= x^2, \quad g(y) = 2y + 1 \\
h(x) &= g(f(x)) = g(x^2) = 2x^2 + 1 \\
h'(x) &= g'(f(x))f'(x) = 4x \; (= 2 * 2x).
\end{aligned}
$$

More challenging examples of composite functions are:

$$
\begin{aligned}
& f(x) = \begin{cases} x \sin \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases} \\
& f'(x) = \sin \frac{1}{x} - \frac{1}{x} \cos \frac{1}{x}, \quad x \neq 0 \\
& f'(0) \text{ does not exist.} \\
& f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases} \\
& f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}, \quad x \neq 0 \\
& f'(0) = \lim_{t \to 0} \frac{f(t) - f(0)}{t - 0} = 0.
\end{aligned}
$$

1.3.1 Taylor's Theorem

Taylor's theorem allows us to expand a function as a series involving higher-order derivatives of a function. We take the *Cauchy form* (with exact remainder):

$$
\begin{aligned}
& f \text{ is } n \text{ times differentiable} \\
& f(b) = \sum_{k=0}^{n-1} \frac{(b-a)^k}{k!} f^{(k)}(a) + R_n \\
& \text{where:} \\
& R_n = \frac{(b-\xi)^n f^{(n)}(\xi)}{n!}, \quad a < \xi < b
\end{aligned} \tag{1.10}
$$

and:

$$f' = f^{(1)} = \frac{df}{dx}, \; f^{(2)} = \frac{d^2 f}{dx^2}$$
$$f^{(n)}(x) = \left(f^{(n-1)}(x)\right)' = \frac{d}{dx}\left(f^{(n-1)}(x)\right).$$

We conclude with a discussion of the *exponential function*. It is the only function that is the same as its derivative. To see this, we use the formal definition (1.7) of a derivative (and noting that $e^x e^y = e^{x+y}, x, y \in \mathbb{R}$):

$$\frac{d}{dx}e^x = \lim_{h\to 0}\left(\frac{e^{x+h} - e^x}{h}\right) = e^x \lim_{h\to 0}\frac{e^h - 1}{h} = e^x, x \in \mathbb{R}.$$

We summarise some useful properties of the exponential function:

$$e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$$
$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$$
$$\frac{d}{dx}e^x = e^x$$
$$e^{x+y} = e^x e^y$$
$$y = \log x \Longleftrightarrow x = e^y$$
$$\log(ab) = \log a + \log b.$$
$$\frac{d^n}{dx^n}e^x = e^x \;\; \forall n \geq 1$$
$$e^x = \sum_{k=0}^{n-1}\frac{x^k}{k!} + \mathbb{R}_n \text{ where } \mathbb{R}_n = \frac{x^n}{n!}e^{\xi}, \;\; \xi < x.$$

1.3.2 Big O and Little o Notation

For many applications we need a definition of the *asymptotic behaviour* of quantities such as functions and series; in particular we wish to find bounds on mathematical expressions and applications in computer science. To this end, we introduce the *Landau symbols* O and o.

Definition 1.2 ***(O-Notation).***

$$f(x) = O(g(x)) \text{ as } x \to \infty \text{ if } \exists M > 0, \;\; \exists x_0 \;\; s.t.$$
$$|f(x)| \leq M|g(x)| \text{ for } x > x_0$$
$$f(x) = O(g(x)) \text{ as } x \to a \text{ if } |f(x)| \leq M|g(x)| \text{ for } |x - a| < \delta$$
$$\text{Unified definition: } \lim_{x\to a}\frac{f(x)}{g(x)} < \infty.$$

An example is:

$$f(x) \equiv 6x^4 - 7x^2 + 2$$
$$g(x) \equiv x^4$$
$$f(x) = O(g(x)) \text{ as } x \to \infty$$
$$f_n \equiv 2n^3 + 6n^2 + 5(\log n)^3$$
$$f_n = O(n^3) \text{ as } n \to \infty.$$

Definition 1.3 ***(o-Notation).***

$$f(x) = o(g(x)) \text{ as } x \to \infty$$
$$\text{if } \lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

An example is:

$$2x = o(x^2)$$
$$2x^2 \neq o(x^2)$$
$$1/x = o(1).$$

We note that *complexity analysis* applies to both continuous and discrete functions.

1.4 PARTIAL DERIVATIVES

In general, we are interested in functions of two (or more) variables. We consider a function of the form:

$$z = f(x, y).$$

The variables x and y can take values in a given bounded or unbounded interval. First, we say that $f(x, y)$ is continuous at (a, b) if the limit:

$$\lim_{\substack{x\to a\\ y\to b}} f(x, y)$$

exists and is equal to $f(a, b)$. We now need definitions for the derivatives of f in the x and y directions.

In general, we calculate the *partial derivatives* by keeping one variable fixed and differentiating with respect to the other variable; for example:

$$z = f(x, y) = e^{kx}\, cosmy$$
$$\frac{\partial z}{\partial x} = ke^{kx}\, cosmy$$
$$\frac{\partial z}{\partial y} = -me^{kx}\, sinmy.$$

We now discuss the situation when we introduce a change of variables into some problem and then wish to calculate the new partial derivatives. To this end, we start with the variables (x, y), and we define new variables (u, v). We can think of these as 'original' and 'transformed' coordinate axes, respectively. Now define the function $z(u, v)$ as follows:

$$z = z(u,v),\ u = u(x,y),\ v = v(x,y).$$

This can be seen as *a function of a function*. The result that we are interested in is the following: if z is a differentiable function of (u, v) and u, v are themselves continuous functions of x, y, with partial derivatives, then the following rule holds:

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x}\\ \frac{\partial z}{\partial y} &= \frac{\partial z}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial y}.\end{aligned} \tag{1.11}$$

This is a fundamental result that we shall apply in this chapter. We take a simple example of Equation (1.11) to show how things work. To this end, consider the *Laplace equation* in Cartesian geometry:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

We now wish to transform this equation into an equation in a circular region defined by the polar coordinates:

$$x = r\,cos\theta,\ y = r\,sin\theta.$$

The derivative in r is given by:

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} = cos\theta\frac{\partial u}{\partial x} + sin\theta\frac{\partial u}{\partial y}$$

and you can check that the derivative with respect to θ is:

$$\frac{\partial u}{\partial \theta} = -rsin\theta\frac{\partial u}{\partial x} + rcos\theta\frac{\partial u}{\partial y}$$

hence:

$$\begin{aligned}\frac{\partial u}{\partial x} &= cos\theta\frac{\partial u}{\partial r} - \frac{1}{r}sin\theta\frac{\partial u}{\partial \theta}\\ \frac{\partial u}{\partial y} &= sin\theta\frac{\partial u}{\partial r} + \frac{1}{r}cos\theta\frac{\partial u}{\partial \theta}\end{aligned}$$

and:

$$\frac{\partial^2 u}{\partial x^2} = cos\theta \frac{\partial}{\partial r}\left(\frac{\partial u}{\partial x}\right) - \frac{1}{r} sin\theta \frac{\partial}{\partial \theta}\left(\frac{\partial u}{\partial x}\right)$$
$$\frac{\partial^2 u}{\partial y^2} = sin\theta \frac{\partial}{\partial r}\left(\frac{\partial u}{\partial y}\right) + \frac{1}{r} cos\theta \frac{\partial}{\partial \theta}\left(\frac{\partial u}{\partial y}\right).$$

Combining these results allows us to write Laplace's equation in polar coordinates as follows:

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = 0.$$

Thus, the original heat equation in Cartesian coordinates is transformed to a PDE of convection-diffusion type in polar coordinates.

We can find a solution to this problem using the *Separation of Variables method*, for example.

1.5 FUNCTIONS AND IMPLICIT FORMS

Some problems use functions of two variables that are written in the *implicit form*:

$$f(x,y) = 0.$$

In this case we have an implicit relationship between the variables x and y. We assume that y is a function of x. The basic result for the differentiation of this *implicit function* is:

$$df \equiv \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0 \tag{1.12a}$$

or:

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y} \tag{1.12b}$$

We now use this result by posing the following problem. Consider the transformation:

$$\left.\begin{aligned} u &= u(x,y) \\ v &= v(x,y) \end{aligned}\right\} \text{original equations}$$

and suppose we wish to transform back:

$$\left.\begin{aligned} x &= x(u,v) \\ y &= y(u,v) \end{aligned}\right\} \text{find } x, y \text{ (inverse functions).}$$

To this end, we examine the following *differentials*:

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy$$
$$dv = \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy. \tag{1.13}$$

Let us assume that we wish to find dx and dy, given that all other quantities are known. Some arithmetic applied to Equation (1.13) (two equations in two unknowns!) results in:

$$dx = \left(\frac{\partial v}{\partial y}du - \frac{\partial u}{\partial y}dv\right)/J$$

$$dy = \left(-\frac{\partial v}{\partial x}du + \frac{\partial u}{\partial x}dv\right)/J$$

where J is the *Jacobian determinant* defined by:

$$J = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial(u,v)}{\partial(x,y)}.$$

We can thus conclude the following result.

Theorem 1.1 The functions $x = F(u,v)$ and $y = G(u,v)$ exist if:

$$\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}$$

are continuous at (a, b) and if the Jacobian determinant is non-zero at (a, b).

Let us take the example:

$$u = x^2/y, \ v = y^2/x.$$

You can check that the Jacobian is given by:

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} 2x/y & -x^2/y^2 \\ -y^2/x^2 & 2y/x \end{vmatrix} = 3 \neq 0$$

Solving for x and y gives:

$$x = u^{2/3}v^{1/3}, y = u^{1/3}v^{2/3}.$$

You need to be comfortable with partial derivatives. A good reference is Widder (1989).

1.6 METRIC SPACES AND CAUCHY SEQUENCES

Section 1.6 may be skipped on a first reading without loss of continuity.

1.6.1 Metric Spaces

We work with sets and other mathematical structures in which it is possible to assign a so-called *distance function* or *metric* between any two of their elements. Let us suppose that X is a set, and let x, y and z be elements of X. Then a *metric* d on X is a non-negative real-valued function of two variables having the following properties:

$$D1: d(x,y) \geq 0; \quad d(x,y) = 0 \text{ if and only if } x = y$$

$$D2: d(x,y) = d(y,x)$$

$$D3: d(x,y) \leq d(x,z) + d(z,y) \text{ where } x,y,z \in X.$$

The concept of distance is a generalisation of the difference between two real numbers or the distance between two points in n-dimensional Euclidean space, for example.

Having defined a metric d on a set X, we then say that the pair (X, d) is a *metric space*. We give some examples of metrics and metric spaces:

1. We define the set X of all continuous real-valued functions of one variable on the interval $[a, b]$ (we denote this space by $C[a, b]$)), and we define the metric:

$$d(f,g) = \max\left\{|f(t) - g(t)|;\ t \in [a,b]\right\}.$$

Then (X, d) is a metric space.

2. n-dimensional Euclidean space, consisting of vectors of real or complex numbers of the form:

$$x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n)$$

with metric:

$$d(x,y) = \max\{|x_j - y_j|; j = 1, \ldots, n\} \text{ or using the notation for a norm } d(x,y) = \|x - y\|_\infty.$$

3. Let $L^2[a,b]$ be the space of all square-integrable functions on the interval $[a, b]$:

$$\int_a^b |f(x)|^2 dx < \infty.$$

We can then define the distance between two functions f and g in this space by the metric:

$$d(f,g) = \|f - g\|_2 \equiv \left\{\int_a^b |f(x) - g(x)|^2\right\}^{1/2}.$$

This metric space is important in many branches of mathematics, including probability theory and stochastic calculus.

4. Let X be a non-empty set and let the metric d be defined by:

$$d(x,y) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{if } x \neq y. \end{cases}$$

Then (X, d) is a metric space.

Many of the results and theorems in mathematics are valid for metric spaces, and this fact means that the same results are valid for all specialisations of these spaces.

1.6.2 Cauchy Sequences

We define the concept of convergence of a sequence of elements of a metric space X to some element that may or may not be in X. We introduce some definitions that we state for the set of real numbers, but they are valid for any *ordered field*, which is basically a set of numbers for which every non-zero element has a multiplicative inverse and there is a certain ordering between the numbers in the field.

Definition 1.4 A sequence (a_n) of elements on the real line $\mathbb{R}$ is said to be *convergent* if there exists an element $a \in \mathbb{R}$ such that for each positive element ε in $\mathbb{R}$ there exists a positive integer n_0 such that:

$$|a_n - a| < \varepsilon \text{ whenever } n \geq n_0.$$

A simple example is to show that the sequence $\left\{\frac{1}{n}\right\}, n \geq 1$ converges to 0. To this end, let ε be a positive real number. Then there exists a positive integer $n_0 > 1/\varepsilon$ such that $\mid \frac{1}{n} - 0 \mid = \frac{1}{n} < \varepsilon$ whenever $n \geq n_0$.

Definition 1.5 A sequence (a_n) of elements of an ordered field F is called a *Cauchy sequence* if for each $\varepsilon > 0$ in F there exists a positive integer n_0 such that:

$$|a_n - a_m| < \varepsilon \text{ whenever } m, n \geq n_0.$$

In other words, the terms in a Cauchy sequence get close to each other while the terms of a convergent sequence get close to some fixed element. A convergent sequence is always a Cauchy sequence, but a Cauchy sequence whose elements belong to a field F does not necessarily converge to an element in F. To give an example, let us suppose that F is the set of rational numbers; consider the sequence of integers defined by the *Fibonacci recurrence relation*:

$$\begin{aligned} F_0 &= 0 \\ F_1 &= 1 \\ F_n &= F_{n-1} + F_{n-2}, \quad n \geq 2. \end{aligned}$$

It can be shown that:

$$F_n = \frac{1}{\sqrt{5}}[\alpha^n - \beta^n]$$
$$\text{where } \alpha = \frac{1+\sqrt{5}}{2}\beta = \frac{1-\sqrt{5}}{2}. \tag{1.14}$$

Now define the sequence of rational numbers by:

$$x_n = F_n/F_{n-1}, \quad n \geq 1.$$

We can show that:

$$\lim_{n\to\infty} x_n = \alpha = \frac{1+\sqrt{5}}{2} \text{ (the \textit{Golden Ratio})}$$

and this limit is not a rational number. The Fibonacci numbers are useful in many kinds of applications, such as optimisation (finding the minimum or maximum of a function) and random number generation.

We define a *complete metric space* X as one in which every Cauchy sequence converges to an element in X. Examples of complete metric spaces are:

- Euclidean space $\mathbb{R}^n$.
- The metric space $C[a, b]$ of continuous functions on the interval $[a, b]$.
- By definition, *Banach spaces* are complete normed linear spaces. A *normed linear space* has a norm based on a metric, as follows $d(x, y) = ||x - y||$.
- $L^p(0, 1)$ is the Banach space of functions $f : [0, 1] \to \mathbb{R}$ defined by the norm $||f||_p = \left(\int_0^1 |f(x)|^p dx\right)^{1/p} < \infty$ for $1 \leq p < \infty$.

Definition 1.6 An *open cover* of a set E in a metric space X is a collection $\{G_j\}$ of open subsets of X such that $E \subset \cup_j G_j$.

Finally, we say that a subset K of a metric space X is *compact* if every open cover of K contains a *finite subcover*, that is $K \subset \cup_{j=1}^N G_j$ for some finite N.

1.6.3 Lipschitz Continuous Functions

We now examine functions that map one metric space into another one. In particular, we discuss the *concepts of continuity* and *Lipschitz continuity*.

It is convenient to discuss these concepts in the context of metric spaces.

Definition 1.7 Let (X, d_1) and (Y, d_2) be two metric spaces. A function f from X into Y is said to be *continuous at the point* a $\in X$ if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that:

$$d_2(f(x), f(a)) < \varepsilon \text{ whenever } d_1(x, a) < \delta.$$

This is a generalisation of the concept of continuity in Section 1.2 (Definition 1.1). We should note that this definition refers to the continuity of a function at a single point. Thus, a function can be continuous at some points and discontinuous at other points.

Definition 1.8 A function f from a metric space (X, d_1) into a metric space (Y, d_2) is said to be a *uniformly continuous* on a set $E \subset X$ if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that:

$$d_2(f(x), f(y)) < \varepsilon \text{ whenever } x, y \in E \text{ and } d_1(x, y) < \delta.$$

If the function f is uniformly continuous, then it is continuous, but the converse is not necessarily true. Uniform continuity holds for all points in the set E, whereas normal continuity is only defined at a single point.

Definition 1.9 Let $f\colon [a, b] \to \mathbb{R}$ be a real-valued function and suppose that we can find two constants M and α such that $|f(x) - f(y)| \leq M|x - y|^{\alpha}, \forall x, y \in [a, b]$. Then we say that f satisfies a *Lipschitz condition* of order α, and we write $f \in Lip(\alpha)$.

We take an example. Let $f(x) = x^2$ on the interval $[a, b]$.
Then:

$$|f(x) - f(y)| = |x^2 - y^2| = |(x + y)(x - y)| \leq (|x| + |y|)|x - y|$$

$$\leq M|x - y|, \text{where } M = 2\max(|a|, |b|).$$

Hence $f \in Lip(1)$.
A concept related to Lipschitz continuity is called a *contraction*.

Definition 1.10 Let (X, d_1) and (Y, d_2) be metric spaces. A transformation T from X into Y is called a *contraction* if there exists a number $\lambda \in (0, 1)$ such that:

$$d_2(T(x), T(y)) \leq \lambda d_1(x, y) \text{ for all } x, y \in X.$$

In general, a contraction maps a pair of points into another pair of points that are closer together. A contraction is always continuous.

The ability to discover and apply contraction mappings has considerable theoretical and numerical value. For example, it is possible to prove that stochastic differential equations (SDEs) have unique solutions by the application of *fixed point theorems*:

- Brouwer's fixed point theorem
- Kakutani's fixed point theorem
- Banach's fixed point theorem
- Schauder's fixed point theorem

Our interest here lies in the following fixed point theorem.

Theorem 1.2 (Banach Fixed Point Theorem) Let T be a contraction of a complete metric space (X, d) into itself:

$$d(T(x), T(y)) \leq \lambda d(x, y), \quad \lambda \in (0, 1).$$

Then T has a unique fixed-point $\bar{x}$. Moreover, if x_0 is any point in X and the sequence (x_n) is defined recursively by the formula $x_n = T(x_{n-1}), \; n = 1, 2, \ldots$, then $\lim x_n = \bar{x}$ and:

$$d(\bar{x}, x_n) \leq \frac{\lambda}{1-\lambda} d(x_{n-1}, x_n) \leq \frac{\lambda^n}{1-\lambda} d(x_0, x_1).$$

In general, we assume that X is a Banach space and that T is a linear or non-linear mapping of X into itself. We then say that x is a *fixed point* of T if $Tx = x$.

1.7 SUMMARY AND CONCLUSIONS

In this chapter we gave an introduction to a number of relevant mathematical concepts from real analysis that are used throughout this book, directly or indirectly. We also have introduced other relevant topics in other chapters. To summarise:

- Chapter 1: Real analysis
- Chapter 4: Finite dimensional vector spaces
- Chapter 5: Numerical linear algebra
- Chapter 16: Complex analysis

In this way we hope that this book becomes more self-contained than otherwise.